Automated On-line Fault Prognosis for Wind Turbine Monitoring using SCADA data

CHEN, BINDI

Automated On-line Fault Prognosis for Wind Turbine Monitoring using SCADA data

Bindi Chen

Supervised by Dr. Peter Matthews and Prof. Peter Tavner

A thesis submitted for the degree of Doctor of Philosophy.

School of Engineering and Computing Sciences

Durham University

United Kingdom

Acknowledgement

In the past three years from Jan 2011, many people have helped me completing the work describes in this thesis and I would like to take this opportunity to thank all of them.

In particular:

- My PhD supervisors, Dr. Peter Matthews and Prof. Peter Tavner, for giving me the opportunity to carry out this research and the thoughtful guidance & positive words throughout the duration of my research.
- The EU FP7 RELIAWIND and UK EPSRC SuperGen Wind projects for funding the work.
- Dr. William Wei Song and Dr. Yanhui Feng for their initial proposal of this research.
- Dr. Yanhui Feng, Dr. Yingning Qiu, Dr. Chris Crabtree, Donatella Zappala and Jamie Godwin for sharing and discussing data and knowledge.
- The IT officer Tony McFarlane and Martin Edney for their helps to setup the ReliaWind Server.
- The master students Chan Xie and Espen Norevik for their helps to analyse Mitsubishi data.
- My friends Hongbo Shao, Peter Wyllie, Terry Ho, Wenjuan Wang, Chris Watson, Nick Cresswell, Sean Norris and Mustafa for the joys in the department.
- Finally, I would like to thank to my parents and my brother for your unconditional support throughout this period of my life.

Bindi Chen

December 2013

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.

Abstract

Current wind turbine (WT) studies focus on improving their reliability and reducing the cost of energy, particularly when WTs are operated offshore. A Supervisory Control and Data Acquisition (SCADA) system is a standard installation on larger WTs, monitoring all major WT sub-assemblies and providing important information. Ideally, a WT's health condition or state of the components can be deduced through rigorous analysis of SCADA data. Several programmes have been made for that purpose; however, the resulting cost savings are limited because of the data complexity and relatively low number of failures that can be easily detected in early stages.

This thesis develops an automated on-line fault prognosis system for WT monitoring using SCADA data, concentrating particularly on WT pitch system, which is known to be fault significant. A number of preliminary activities were carried out in this research. They included building a dedicated server, developing a data visualisation tool, reviewing the existing WT monitoring techniques and investigating the possible AI techniques along with some examples detailing applications of how they can be utilised in this research.

The a-priori knowledge-based Adaptive Neuro-Fuzzy Inference System (APK-ANFIS) was selected to research in further because it has been shown to be interpretable and allows domain knowledge to be incorporated. A fault prognosis system using APK-ANFIS based on four critical WT pitch system features is proposed. The proposed approach has been applied to the pitch data of two different designs of 26 Alstom and 22 Mitsubishi WTs, with two different types of SCADA system, demonstrating the adaptability of APK-ANFIS for application to variety of technologies. After that, the Alstom results were compared to a prior general alarm approach to show the advantage of prognostic horizon. In addition, both results are evaluated using Confusion Matrix analysis

and a comparison study of the two tests to draw conclusions, demonstrating that the proposed approach is effective.

Publications

Journal Papers

1. **Chen, Bindi**, Matthews, P.C. & Tavner, P.J (2014). Automated On-line Fault Prognosis for Wind Turbine Pitch Systems using SCADA data, IET Renewable Power Generation, Awaiting peer review.
2. **Chen, Bindi**, Matthews, P.C. & Tavner, P.J (2013). Wind turbine pitch faults prognosis using a-priori knowledge-based ANFIS. Expert Systems with Applications 40(17): 6863.
3. Qiu, Y N, Feng, Y H, Tavner, P J, Richardson, P, Erdos, G, & **Chen, Bindi** (2012). Wind turbine SCADA alarm analysis for improving reliability. Wind Energy 15(8): 951-966.

Conference Papers

1. **Chen, Bindi**, Matthews, PC & Tavner, P J (2013), Automated Wind Turbine Pitch Fault Prognosis using ANFIS, EWEA 2013. Vienna, Austria, European Wind Energy Association.
2. **Chen, Bindi**, Tavner, P.J. , Feng, Y. Song, W.W. & Qiu, Y.N. (2012), Bayesian Network for Wind Turbine Fault Diagnosis, EWEA 2012. Copenhagen, Denmark, European Wind Energy Association.
3. **Chen, Bindi**, Song, W.W. Qiu, Y.N. Feng, Y. & Tavner, P.J. (2011), Knowledge-based Information Systems – A Wind Farm Case Study, 20th International Conference on Information Systems Development. Edinburgh, Scotland, August 24 -26, 2011, Springer.
4. **Chen Bindi**, Qiu Y.N., Feng Y., Tavner, P.J. & Song, W.W. (2011), Wind turbine SCADA alarm pattern recognition, Renewable Power Generation (RPG 2011), IET Conference on. Edinburgh, UK, IET, pp. 1-6.

Last updated in December 2013

List of Abbreviation

ACC	Accuracy
AI	Artificial Intelligence
ANFIS	Adaptive Neuro Fuzzy Inference System
APK-ANFIS	A-priori Knowledge-based ANFIS
ANN	Artificial Neural Network
BN	Bayesian Network
BPN	Back Propagation Network
BTA	Boosting Tree Algorithm
CART	Classification and Regression Tree
CCF	Critical Characteristic Feature
CMS	Condition Monitoring System
CONMOW	Condition Monitoring for Offshore Wind Farms
CWT	Continuous Wavelet Transform
DBN	Dynamic Bayesian Network
DFIG	Doubly Fed Induction Generator
ER	Error Rate
F	F-measure
FCM	Fuzzy c-Mean
FFT	Fast Fourier Transform
FDD	Fault Detection and Diagnosis
FIS	Fuzzy Inference System

FN	False Negative
FP	False Positive
GUI	Graphical User Interface
GP	Genetic Programming
IDFT	Iterative Localised Discrete Fourier Transform
KPI	Key Performance Indicator
k-NN	k-nearest Neighbours
KMS	Knowledge Management System
MF	Membership Function
MSE	Mean Square Error
NPT	Node Probability Table
NN Ensemble	Neural Network Ensemble
O&M	Operation and Maintenance
P	Precision
RC	Recall
RMSE	Root Mean Square Error
SBPF	Sideband Power Factor
SCADA	Supervision Control and Data Acquisition
SIMAP	Intelligent System for Predictive Maintenance
SOM	Self-Organising Feature Mapping
SVM	Support Vector Machine
TP	True Positive

TN	True Negative
UPS	Uninterruptable Power Supply
WF	Wind Farm
WS	Window Size
WT	Wind Turbine
WTCMTR	Wind Turbine Condition Monitoring Test Rig
UPS	Uninterruptable Power Supply

Nomenclature

A_i	The *i*th linguistic variables of the membership function (Chapter 4)
ACC	Accuracy rate (Chapter 5)
a_i	Premise parameter of the *i*th rule (Chapter 4)
a_i	The *i*th weight (Section 4.2.2)
a^j	Parameter of Gaussian basis that capture the domain knowledge (Chapter 4)
B_i	The *i*th linguistic variables of the membership function (Chapter 4)
b_i	Premise parameter of the *i*th rule (Chapter 4)
C_j	The jth centroid of the cluster in Fuzzy c-Means (Section 3.2.10)
C_n^i	The *n*th rule centre (Chapter 4)
C_i^B	The *i*th rule centre (Chapter 4)
C_i^F	The *i*th favourable rule centre (Chapter 4)
c_i	Premise parameter of the *i*th rule (Chapter 4)
D_n	Desired output of the *n*th input pattern (Section 3.2.6)
D_{ij}	The Euclidian distance of the *i*th rule centre to *j*th rule centre (Chapter 4)
E	The mean square error (Section 3.2.6)
ER	Error rate

F	Feature variable (Section 3.2.2)
F	F-measure (Chapter 5)
F	False value (Section 3.2.8)
f	ANFIS output (Chapter 4)
f_i	The *i*th consequent function (Section 3.2.4 & 4.1.1)
f_i	The activation function at *i*th layer (Section 3.2.6)
H	A m dimensional column vector representing ANN hidden layer (Section 3.2.6)
I	A k dimensional column vector representing ANN input (Section 3.2.6)
j_m	Fuzzy c-Means objective function (Section 3.2.10)
K	The number of cluster (Section 3.2.9 & 3.2.10)
$K_m(x, y, \dots, z)$	The *m*th centroid (Section 3.2.9)
k	The number of the nearest neighbours (Section 3.2.11)
k_n^i	The *n*th coefficient of the ANFIS consequent linear model for the *i*th rule (Chapter 4)
M_m , M_k	The corrective maintenance at time m and k (Chapter 5)
O	A column vector representing ANN output (Section 3.2.6)
O_n	Actual output of the *n*th input pattern (Section 3.2.6)
$O_{A,i}$	The *i*th membership grade of the input to membership function A (Chapter 4)

$O_{B,i}$	The *i*th membership grade of the input to membership function B (Chapter 4)
P	Precision (Chapter 5)
P	Probability (Section 3.2.8)
P_i	A vector represents fault behaviours (Chapter 4)
p_i	The *i*th parameter of the consequent function (Chapter 4)
$p(C_i)$	Probability for the event C_i (Section 3.2.2)
q_i	The *i*th parameter of the consequent function (Chapter 4)
R	ANFIS rule base (Chapter 4)
r_i	The *i*th parameter of the consequent function (Chapter 4)
RC	Recall
T	Time (Chapter 5)
T	Threshold
T	True value (Section 3.2.8)
V	Wind speed (m/s) (Section 3.3)
W_i	The *i*th weight matrix in ANN (Section 3.2.6)
w_i	Firing strength of the *i*th rule (Chapter 4)
$\overline{w}_i$	Normalised firing strength of the *i*th rule (Chapter 4)
WS	Window Size (Chapter 5)
X	A set of feature variables (Section 3.2.2)
x	ANFIS input (Chapter 4)
x_i	The *i*th ANFIS input (Chapter 4)

x_i^{new}	The normalised value of x_i (Chapter 5)
x_{max}	The maximum of variable x (Chapter 5)
x_{min}	The minimum of variable x (Chapter 5)
y	ANFIS input (Chapter 4)
μ	The mean (Section 3.2.2)
σ	The standard deviation (Section 3.2.2)
$\sigma(x)$	The sigmoid function (Section 3.2.6)
$\|\ast\|$	Similarity between any measure data (Section 3.2.10)
μ_{ij}	The degree of membership of x_i in the cluster j (Section 3.2.10)
ε	Exponent in Fuzzy c-Means (Section 3.2.10)
β	Blade angle (°) (Section 3.3)
Ω	Rotor speed (r/s) (Section 3.3)
$\mu_{A_i}(x)$	The *i*th membership grade of the input to membership function A (Chapter 4)
θ_n^i	The *n*th coefficient in Taylor series (Chapter 4)
Φ_r^j	Gaussian basis function to model from *j*th favourable rule to *r*th rule (Chapter 4)
σ_i^j	Parameter of Gaussian basis used to govern the model output decay (Chapter 4)
γ_i^j	Weight of the closeness of the *i*th rule centre to *j*th favourable (Chapter 4)

Table of Contents

List of Figures

List of Tables

Chapter 1.

Introduction

Wind energy is currently the fastest growing renewable source used for electrical generation around the world. It is expected that a large number of wind turbines (WTs), especially offshore, will be employed in the near future with the aim of achieving the desired carbon emission targets and providing alternative energy sources for customers (Krohn et al. 2009). WTs are designed to be operated around 20 years and their life-time reliability is the viable factor for the success of any wind farm (WF) project.

Following a rapid acceleration of wind energy development in the late 20th & early 21st century, current studies of WTs are beginning to focus on improving the cost of energy. The main reason is to ensure that wind generated electricity is competitive with other generation sources. Costs for wind generated electricity can be higher because O&M costs constitute a significant share of the annual cost of a WF and WT downtime. With the rapid growth of wind energy and more offshore WTs to be employed in the near future, there is a commercial interest in ensuring reduced O&M costs by increasing reliability and having more economical operations. The essence of improving WT reliability is to reduce the downtime and increase the availability by optimising both the WT design and its maintenance schedule (Tavner et al. 2007). Both these strategies require a full understanding of the WT system and a detailed analysis of its failure mechanisms. Most modern large WTs are now manufactured with some types of SCADA and CMS systems that monitor the main components and it is possible for WF operators to analyse these data to identify WT's systematic performance.

Ideally, a WT's health condition or state of the turbine's component can be deduced through rigorous analysis of SCADA and CMS data. This information would also be very useful to plan power outages and schedule effective maintenance schemes. However, many WF operators have been unable to make full use of these available, due to large unmanageable volumes of data and lack of domain knowledge impeding its analysis and interpretation.

1.1. Wind Energy Development & Cost of Energy

Figure 1.1 shows the European Commission's forecasts on annual wind power investments in EU-27 from 2000 to 2030 (Krohn et al. 2009). The market is growing, with a gradually increasing share of investments going to offshore. By 2020, the annual market for wind power capacity will have grown to €17 billion annually with approximately half of investments going to offshore. By 2030, annual wind energy investment will reach almost €20 billion with 60% of investments offshore.

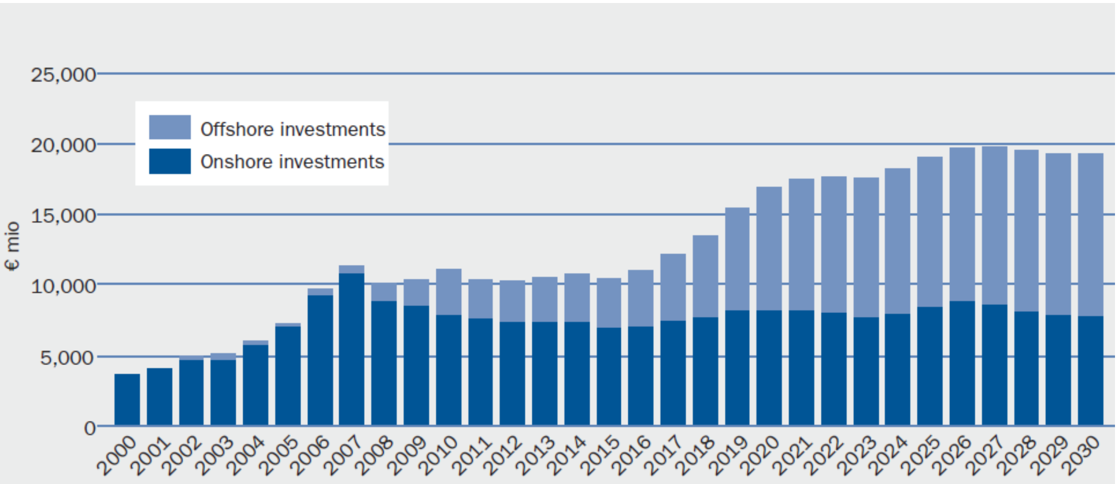


Figure 1.1:Wind energy investments 2000-2030 (€ million) (Krohn et al. 2009)

In UK, the government has agreed to a legally binding target for 15% of energy production from renewable source by 2020, increasing from 1.5% in 2006 (GOV.UK 2009). According to a consultative document GOV.UK (2011) published by the Department for Business, Enterprise and Regulatory Reform in

June 2008, offshore wind power could contribute up to 19% of UK renewable energy target by 2020. As an island UK has large potential for offshore WTs to be installed to achieve this greater wind energy harvest. However, due to the lack of operating experience on large-scale offshore WFs, this could also increase the risk to energy capture from low reliability and availability, in view of the difficulties of accessing offshore WTs for maintenance (Feng et al. 2010).

1.1.1. Trends in Wind Turbine Size

For the development of WT machine itself, with a focus on increasing MW ratings over the last decade (Krohn et al. 2009), turbine sizes have become larger. This is due to the fact that more energy can be captured by a greater swept area, as well as a cost benefit due to the scale, which means that the cost of energy will be more competitive for larger WTs. The size evolution of modern WTs since the 1980s is depicted in Figure 1.2 below.

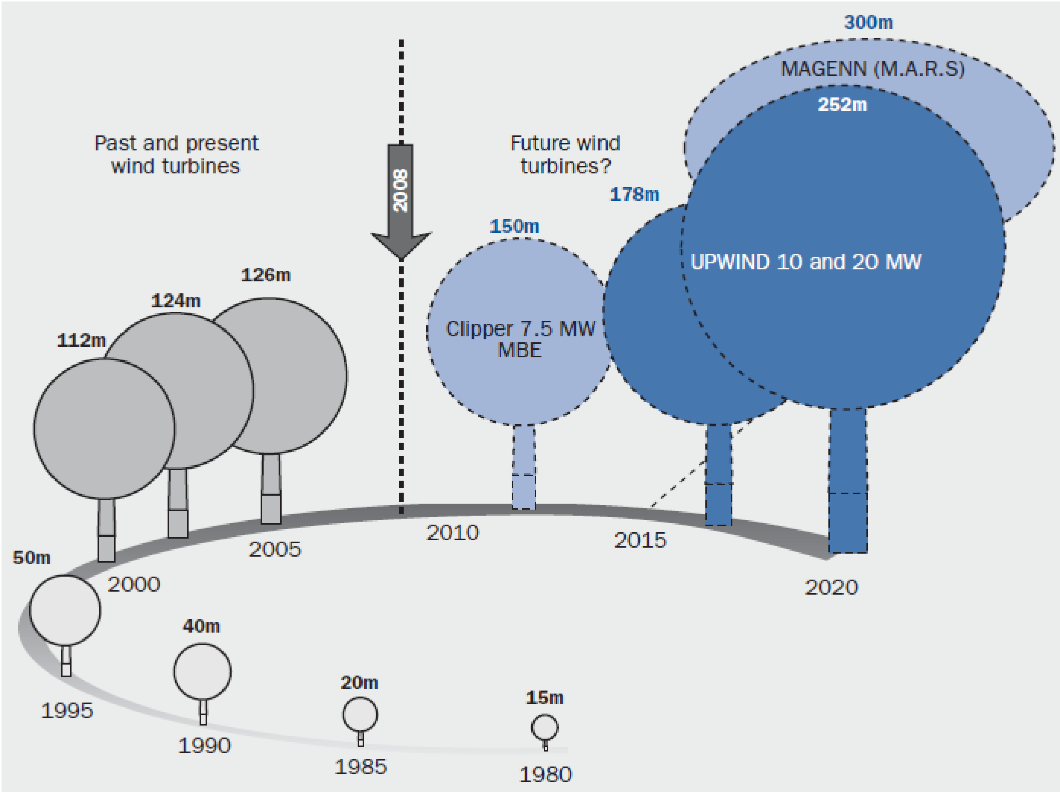


Figure 1.2: Growth in size of wind turbine design (Krohn et al. 2009).

1.1.2. Moving Offshore

Currently, most WFs have been sited onshore, but there is growing interest in installing them offshore to take advantage of the stronger winds and lower environmental impact of offshore locations, as shown in the investment trend Figure 1.1.

1.1.3. Cost of Energy

The key elements that determine the basic costs of wind energy are listed below (Blanco 2009):

- **Capital costs**, including WTs, foundation, road construction and grid connection, which can be as much as 80% of the total project cost over lifetime. The capital costs are the initial fixed costs and determined by market, location and WT design at the time of installation;

- **Variable costs**, the most significant are the O&M costs, but also including other categories such as land rental, insurance, taxes and administration. The variable costs constitute around 20% of the total cost, with variations between countries, regions and sites;
- **The electricity production**, which depends on local climate, site characteristics, WT technical specifications and power generation reduction. The indicator that best reflect this element is capacity factor, which expresses the ratio of actual energy produced in a year;
- **The discount rate and economic lifetime of the investment**. These reflect the perceived risk of the project.

For wind energy to be competitive with other energy sources, it is essential to make every effort to reduce the cost of energy from wind. At present, one of the priorities of WT research is to lower the variable costs, mainly related to O&M. This is because O&M costs constitute a sizable share of the total cost, 12% for onshore and 23% for offshore in some EU WFs (Musial et al. 2006; Feng et al. 2010). In addition, a German study shows that O&M costs are likely to increase over WT lifetime (Report 2002). The detection of incipient WT failures in their early stages and the identification of their root causes would improve O&M, leading to better WT availability and decreased cost of energy.

1.2. Current Challenges affecting the Wind Turbine O&M

The first challenge that affects WT O&M is difficulty of access to the WT in a WF. A WF is a group of WTs in the same location used to produce electric power. A large WF may contain more than 100 individual WTs and covers an extended area of tens of square km. The sizes of WTs have become physically larger and the WF locations are usually built on remote plains, hills or inshore sea regions. The growing interest in installing WF offshore increases the difficulty of access and results in potentially high O&M costs. The main factors affecting offshore WT O&M costs are:

- Difficulties of site access, corresponding repair and maintenance techniques will require reaching the WT by vessel or helicopter;
- Delays in the first opportunity to carry out a visual inspection of suspected failure, caused by adverse weather conditions, may be several days or even weeks, resulting lost revenue;
- Extreme weather conditions may reduce the ability to perform certain maintenance procedures to effect a repair;

The second challenge arises from the data collected by the SCADA & CMS systems. They quickly accumulate large, unmanageable volumes of data, making data analysis difficult. It would be impractical and maybe impossible to carry out WT SCADA & CMS data analysis manually.

The third challenge is the lack of WT domain knowledge impeding incipient fault detection and interpretation. The lack of defect knowledge and expertise on how faults manifest themselves in the data is ultimately the main aspect of this challenge, which may allow an undetected component fault to develop into major failure and could even damage the whole WT.

With the increase of wind energy development, especially the increasingly numerous offshore WFs, questions regarding O&M are gaining more importance. This heightens the need for comprehensive automated on-line fault detection that can use of existing WT SCADA & CMS data to provide accurate

WT fault prediction, so that impending component failures can be detected and reasonable maintenance actions are scheduled to minimise WT downtime, improve reliability, maximise availability and improve profitability.

1.3. Research Aim

In order to overcome the challenges to WT O&M leading to better WT availability and decreased cost of energy, a system that can automatically analyse and interpret the large volumes of data from SCADA & CMS systems is required. This would reduce WF operators' workload as only high level information about WT health or component state would be presented to them, dramatically reducing the complexity of having to carry out data analysis manually and help operators to make timely maintenance decisions.

The aim of this research is to provide a suitable solution that can be used as a framework for automated on-line WT fault prognosis based upon a specific, important fault area, the WT pitch system. This research speculates that an AI algorithm is available to build a robust and effective WT fault prognosis system. This potential AI system would be able to work on different designs and locations of WTs, and produce better fault prognosis result than the existing solutions.

1.4. Structure of the Thesis

This thesis is organised into a number of chapters in order to reflect the progress and results of the research since January 2011.

Chapter 1 briefly introduces wind energy development and the research background. The challenges to WT O&M are discussed and the aims of the research are introduced.

Chapter 2 begins with the brief introduction of different types of WT and the "Danish Concept" is discussed in detail because of its current wide application. WT SCADA & CMS systems are introduced and current WT reliability studies are discussed. Then, the review summarises and discusses the recent research about the WT fault detection and diagnosis (FDD). In the end, the focus in this research is presented and followed by the introduction of the available data and research facilities in Durham University.

Chapter 3 reviews a number of AI techniques, commonly used in the field of data classification, which can be used for WT FDD. In the end of this chapter a-priori knowledge-based Adaptive Neuro-Fuzzy Inference System (APK-ANFIS) was selected to analyse WT SCADA data in further. In addition, six known pitch faults were found and used as the knowledge base for the proposed diagnosis approach in Chapter 4.

Chapter 4 introduces the APK-ANFIS in detail and proposes an automated on-line fault prognosis system. With the a-priori knowledge incorporation, the proposed system should have improved ability to interpret previously unseen conditions. The data of the six known pitch faults were labelled and used to train the proposed system with a-priori knowledge incorporated.

Chapter 5 shows the trained system was applied to data from two different designs of WTs, manufactured by Alstom & Mitsubishi, with two different types of SCADA systems, demonstrating the adaptability of the proposed approach to variety of technologies. The results were further evaluated by Confusion Matrix analysis to check the validity of the results.

Chapter 6 discusses how the proposed system meets the aim of this research and lists the advantages of the proposed approach. This chapter also provides conclusions from this research and proposes the further work.

Chapter 2.

Wind Turbine Monitoring

As has been noted in Chapter 1, the detection of WT faults, particularly offshore, is gaining greater importance because of their remote location and inaccessibility, the difficulties of maintenance and the significance of machine failures on availability. The need for successfully detecting incipient faults before they develop into serious failures, to increase availability and lower cost of energy, has led to the development of a large number of WT SCADA & CMS systems (Crabtree 2010; Chen and Zappala 2011).

This chapter begins with the brief introduction of different types of WT and then the "Danish Concept" is discussed in detail because of its wide applications. After that, WT monitoring systems are introduced and followed by a summary of current WT reliability studies. Then, the review summarises and discusses the recent researches about the WT fault detection and diagnosis (FDD). In the end, the focus in this research is presented and followed by the introduction of the available data and research facilities.

2.1. Modern Wind Turbines

WTs are mechanical devices that convert kinetic energy from the wind into rotational mechanical energy at the shaft, then conversion into electrical energy in a rotating generator. Several WT designs have been devised over the last 100 years. However, in general, WTs can be classified into vertical-axis and horizontal-axis ones depending on the position of the WT's rotor axis (Bianchi et al. 2006).

The most successful vertical-axis WT is the Darrieus (Hau and Platz 2006), as illustrated in Figure 2.1(a). The advantage of this type of WT is that any gearbox and conversion systems are placed at ground level and the WT is independent of the wind direction. However, maintenance of these WTs is not straightforward as rotor removal is often required. In addition, the captured energy is not efficient and large WT plan areas are required for the guy-wires necessary to support the vertical axis structure. For these reasons only a few large vertical-axis WTs have gone into operation and none have been installed offshore (Ackermann and Söder 2002).

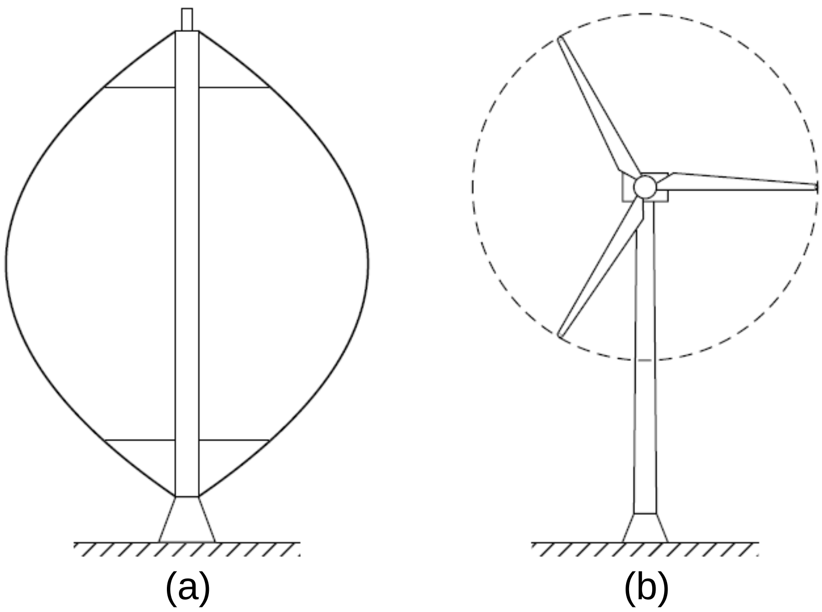


Figure 2.1: (a) Vertical-axis & (b) Horizon-axis WTs (Bianchi et al. 2006).

Today, the majority of modern WTs used horizontal-axis, two or three blades designs with the three-blade option being by far the majority (Hau and Platz 2006), as illustrated in Figure 2.1(b).

Three-blade option is the preferred solution for modern horizontal-axis WT because of the integrated considerations of power coefficient, noise emission and visual effect (Hau and Platz 2006). In addition, the horizontal-axis WTs also can be classified into up-wind and down-wind designs depending on the position of the WT's rotor. Down-wind WTs have the rotor on the back of the turbine and doesn't require a separate yaw system to yaw the turbine into the wind. However, the advantages of down-wind design can also be a disadvantage as it responds to wind directional changes more slowly and this is not practicable for modern large WT.

The vast majority of current horizontal-axis WTs have up-wind rotors to create a more practicable system for wind energy capture. The tower holds up the nacelle containing the generator, which in the case of a direct drive WT the generator is a low speed synchronous machine but the majority of current machines installed are indirect drive and high speed induction or asynchronous generator is assembled in the nacelle with the gearbox. There is a yaw mechanism that turns the rotor and nacelle into the wind, in order to capture as much energy as possible. The power electronics converter changes the generator frequency to the grid frequency and the transformer changes the generator voltage to the grid voltage. The converter and transformer are usually arranged at the base of the tower, on the ground for onshore WTs but for offshore WTs they are usually contained within the tower. Only three-blade, horizontal-axis, up-wind WTs are considered in this research, since they represent the vast majority of large WTs currently installed.

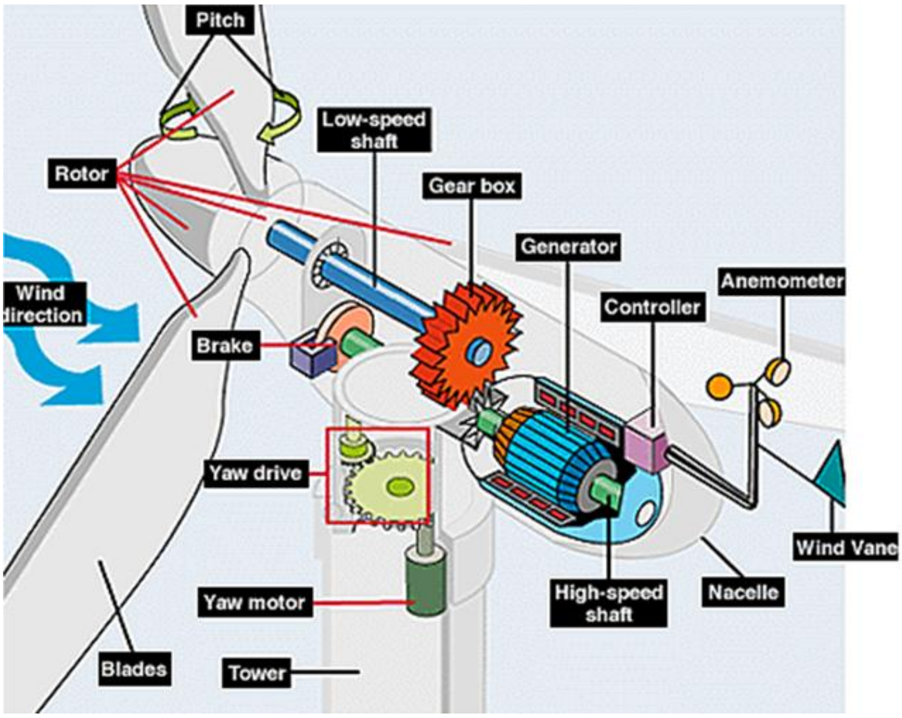


Figure 2.2: Wind turbine nomenclature (Darling 2011).

There are two types of three-blade, horizontal-axis, up-wind WTs currently in operation and considered in this research.

- The fixed speed, variable pitch WT using the blade pitch-to-stall control to adjust the WT power;
- The variable speed, variable pitch WT using generator and blade pitch-to-feather control to adjust the WT power.

Figure 2.2 shows the components involved in both types of three-blade horizontal-axis, up-wind WTs. The following are brief descriptions of the main WT sub-assemblies (Darling 2011), important for the gathering of knowledge from the WT:

- **Rotor:** The WT blades and hub together.

- **Blades:** Extract kinetic energy from the wind and converts it into rotational mechanical shaft energy as a driving torque and WT speed at a certain wind speed.
- **Pitch System:** Controls the angle of attack of the blades to the wind to control the extraction of kinetic energy and avoid rotor over-speed at high winds speed.
- **Brake:** A disc brake to slow down and stop the rotor at cut-out wind speed or in over-speed emergencies.
- **Low-speed shaft:** Turned by the WT rotor.
- **Gearbox:** Used to transfer rotational mechanical energy from the low speed shaft to the high speed shaft.
- **High speed shaft:** Driven by the gearbox output, coupled to the generator and drives the generator.
- **Generator:** Converts the rotational mechanical shaft energy from the high speed shaft into electrical energy, developing a reaction torque to the high speed shaft.
- **Converter:** Controls the flow of electrical energy from the generator by adjusting its frequency.
- **Controller:** Starts up and shuts down the WT at the cut-in and cut-out wind speeds, controls the pitch and converter to extract the maximum energy, and the yaw system to point the WT into the wind and develop the appropriate reaction torque to the WT at a given wind speed.
- **Anemometer:** Measures the wind speed and sends the data to the controller to assist in the development of the reaction torque.
- **Wind Vane:** Measures the wind direction and sends the data to the controller to control the yaw system.
- **Nacelle:** Housing on the top the tower to yaw into the wind and protect the drive-train assemblies, shafts, gearbox and generator.

- **Yaw Drive:** Used to control the nacelle to face the wind as wind direction changes.
- **Yaw Motors:** Power the yaw drive.
- **Tower:** Supports the nacelle at an appropriate height, as wind speed increases with height, taller towers enable WTs to capture more energy and generate more electricity.

There are more than 50 large horizontal-axis WT manufacturers in the world (Krohn et al. 2009) and their WT designs vary from manufacturer to manufacturer. However, the sub-assemblies described above are common to almost all manufacturers and this gives confidence in taking a common approach to all WTs FDD.

2.2. Wind Turbine Monitoring System

WTs are monitored for a variety of reasons, including measuring the meteorological data to forecast their prospective power generation. Figure 2.3 shows a number of different monitoring systems typically installed on a large, > 1.5MW, WT.

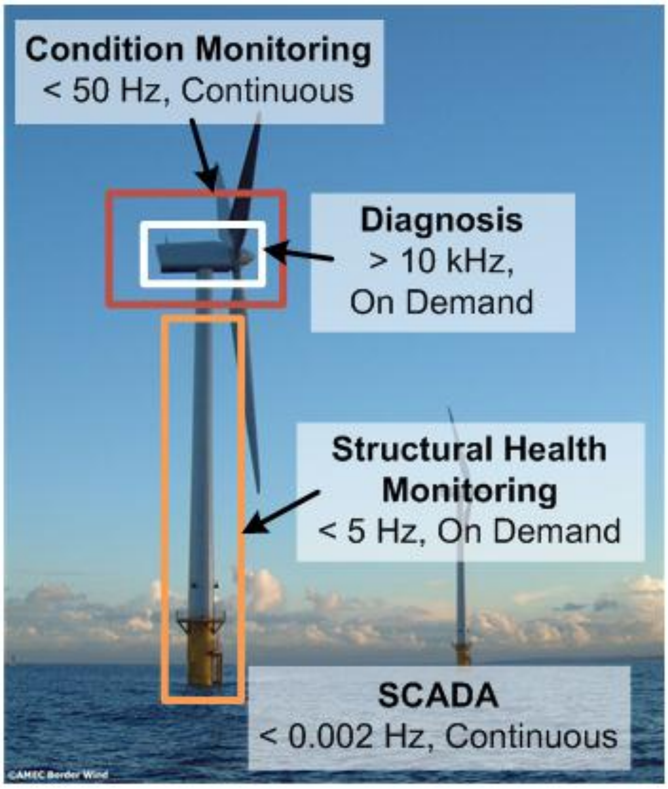


Figure 2.3: Monitoring systems on WT (Crabtree 2010).

2.2.1. Supervisory Control and Data Acquisition

The Supervisory Control and Data Acquisition (SCADA) system is a standard installation on large WTs; its data is collected from individual WT controllers. According to (Zaher et al. 2009), the SCADA system assesses the status of the WT and its sub-assemblies using sensors fitted to the WT, such as anemometers, thermocouples, accelerometers and switches. The meteorological information, e.g. wind speed, direction and turbulence, and turbine operating information, e.g. rotor speed, blade angle, power output, and lubrication oil temperature, are measured by these instruments at a low data rate, usually at 5-10 minutes interval. The SCADA system data are transmitted to a central database for WF operators to monitor the WT & WF performance.

The initial design of the SCADA system is not necessary to give an indication of the WT health or provide warning of impending malfunctions. However, some studies (Feng et al. 2011; Yang and Jiang 2011) have pointed out that the SCADA data is a rich resource potentially useful for maintenance optimisation. This is largely because SCADA system covers almost all major WT sub-assemblies and archives comprehensive signal information, historical alarms and detailed fault logs, as well as environmental and operational condition. It is believed that a WT's health condition can be monitored through rigorous analysis of the information collected by SCADA system (Chen et al. 2013).

2.2.2. Condition Monitoring & Diagnosis

Many large WTs are now fitted with Condition Monitoring system (CMS), initially encouraged by insurance companies on early large WTs to reduce drive train outages due to the prominent gearbox failures, which monitor sensors associated with the rotating drive train, such as accelerometers, proximeters and oil particle counters. As a CMS is installed on the WT's drive train, it may be considered as a method for determining whether a WT is operating correctly or whether a fault is present or developing. Once a fault has been detected, CMS can diagnose automatically or via a monitoring engineer to determine the exact nature and the location of the fault. For this analysis the data must be recorded at higher sampling frequencies, however CMS diagnosis will only be required on an intermittent basis (Crabtree 2011), when incipient faults are indicated by the SCADA. The CMS is separated from SCADA system and as explained collects data at much higher data rates.

According to (Yang and Jiang 2011), the cost of CMS is much higher than SCADA system due to expensive CMS transducers, data processing equipment and a higher installation cost. In addition, the WF operators usually have a great challenge to transmit, store and analyse data with such high rates.

2.2.3. Structural Health Monitoring

On larger WTs, > 2MW, there is frequently a Structural Health Monitoring (SHM) system installed below the nacelle. The SHM system uses low frequency sampling, < 5Hz, of accelerometers or strain gauges, to determine the structural integrity of the WT tower and foundation for faults, driven by blade-passing frequencies, wind gusts and wave slam.

Structural faults are slow to develop and do not need continuous monitoring. They are better for consideration during perhaps an annual structural survey.

2.2.4. Survey of the Commercially Available Monitoring Systems

In order to understand the scope of current systems, two surveys were made by C.J. Crabtree and this author to investigate all commercially available CMS and SCADA systems respectively (Crabtree 2010; Chen and Zappala 2011) for the UK EPSRC Supergen Wind Research Consortium. The information was collected from WT manufacturers, Renewable Energy Consultancies, Industrial Software Companies, WT Operating Company, Electrical Equipment Provider, papers and the internet. The survey observations were made concern with the nature of systems currently available and apparent future development of WT monitoring systems rather than their detailed effectiveness. The survey conclusions have shown that it would be beneficial, from the perspective of WF operators, if the WT monitoring data could be integrated, analysed and interpreted automatically to support the operators identifying WT defects.

2.3. Reliability of Wind Turbines

WT reliability is largely dependent on the design of the machine, along with the quality of its sub-assemblies and the manufactured quality of its components (Hau and Platz 2006). WT technology has developed and matured in recent years, the design of large WTs has become fairly standardised, centring around the three-blade, horizontal-axis, up-wind design of the original "Danish Concept". However, there may be some variations within this general design. For example, the more popular concept is a WT with a gearbox linked to a high-speed induction/asynchronous generator and this concept has lower capital cost, more recent WTs have been developed with a direct drive system connected to a low-speed synchronous generator and fully-rated converter and this new concept may have the potential to be more reliable and suffer lower losses in low wind (Hau and Platz 2006). However, regardless of which concept is considered there are a number of sub-assemblies within a WT which could be the potential source of failure.

Some studies have analysed publicly available data in an attempt to gain knowledge of overall WT reliability, whilst also ascertaining the reliability of particular sub-assemblies in relation to the whole system. Existing research has taken many different approaches to analyse the public available data. Some have looked at reliability based on WT rating (Spinato et al. 2009) or weather & location (Tavner et al. 2013). Some studies have investigated the reliability of different WT sub-assemblies (Spinato et al. 2009; Wilkinson et al. 2010).

2.3.1. Reliability Study based on Wind Turbine Rating

Spinato et al. (2009) carried out a failure rate study based on WT rating as specified in the LWK data for onshore WTs, considering 158-643 WTs and age up to 15 years, the results of which are shown in Figure 2.4. It is apparent that there is a general trend of increasing failure rate with WT rating. Based on this study, it may be more difficult to decrease failure rates as WTs continue to grow in rating. This will be more significant as they move to offshore location, where larger WTs are needed. However, experience also shows that as larger WTs are introduced there is a learning curve and failure rates can reduce with time as appropriate O&M procedures are learnt on the larger machine (Blanco 2009).

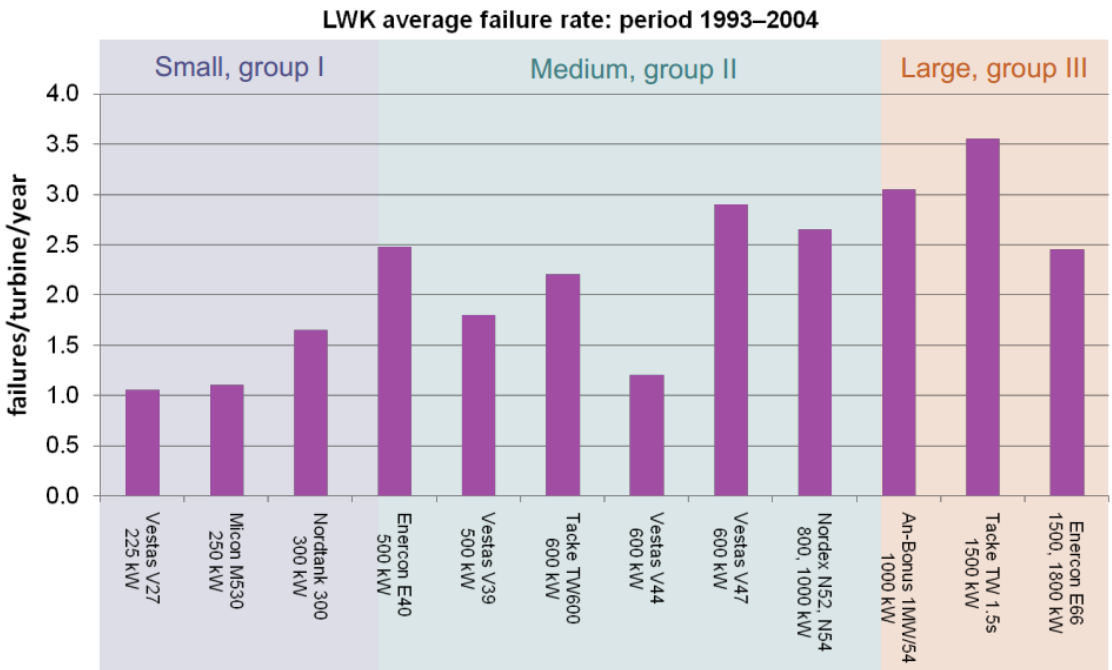


Figure 2.4: Variation of failure rates with different wind turbine rating (Spinato et al. 2009).

2.3.2. Reliability Study based on Weather and Location

Tavner et al. (2006) investigated the influence of wind speed on WT reliability, using about 16000 WT-years of Windstats data from 1994 to 2004 of WTs located in Denmark. The study has some analytical difficulties; the population changed size and consisted of various WT models located all over the country in varying climatic conditions, and Danish national information was used to characterise the weather over the large geographical area. Nevertheless, it did clearly show that higher mean wind speeds resulted in an increased number of WT failures. In Tavner et al. (2013), a more precise analysis using WMEP data was made at three specific locations on one WT type, with 201 WT-years of data, using weather data from local meteorological masts. Cross-correlation was shown to be effective for relating weather conditions to failure rate and this research has shown that there was significant cross-correlation between the failure rate, weather and turbulence for all three sites rather than just wind speed.

2.3.3. Reliability Study based on Wind Turbine Sub-assemblies

Quantitative studies of WT faults from existing public databases have been carried out by Tavner et al. (2007) and Spinato et al. (2009) on 25,322 WT-years of data. Figure 2.5 shows the comparison between failure rate and downtimes of different WT sub-assemblies, such as those described in Section 2.1, from three large EU surveys of onshore WTs.

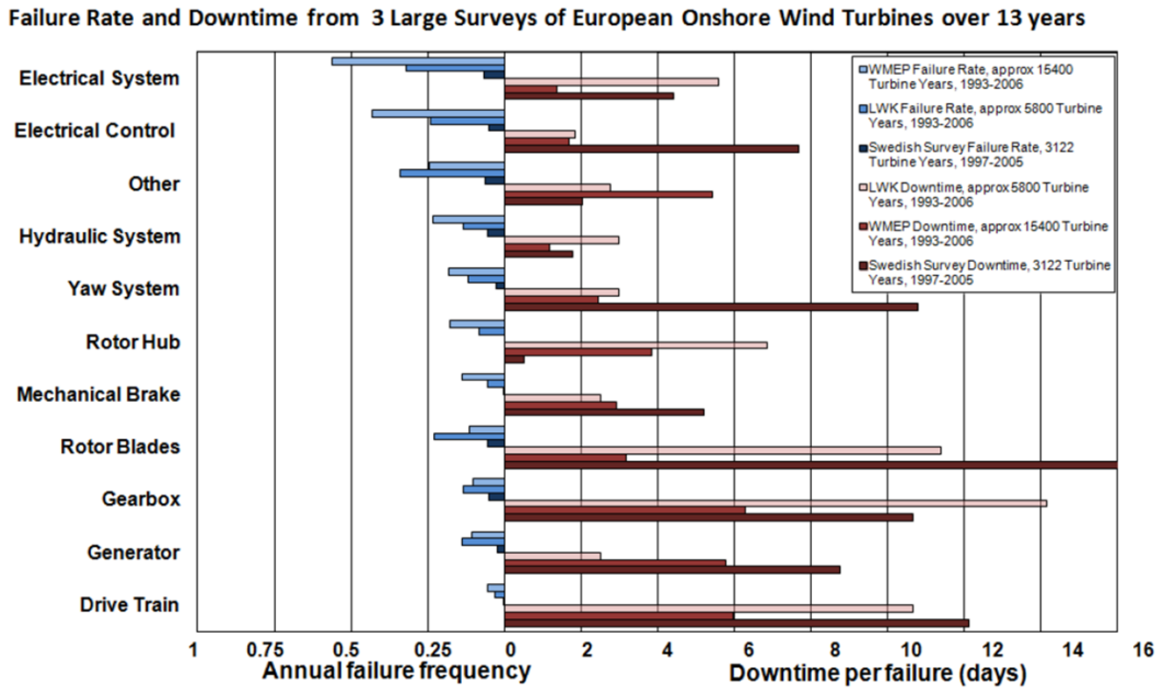


Figure 2.5: Failure/WT/year and downtime results, 25,322 WT-year LWK, WMEP and Swedish surveys, 1993-2006 (Spinato et al. 2009).

It can be seen that Electrical System and Electrical Control have the highest failure rate, but the corresponding WT downtimes are not high. The major sources of downtimes have their root causes centred on the drive train, which refers to the large rotating components including the rotor, main bearing, main shaft, coupling, gearbox and generator. Although their failure rates are not high, their downtimes are the highest of all sub-assemblies as shown in Figure 2.5. This is because the repair procedures in drive train are complex and this will be aggravated particularly offshore, requiring not only special lifting equipment such as crane, but also vessels and the weather conditions will have to be considered.

Another study of WT sub-assembly reliability was carried out by more recent ReliaWind project (Wilkinson et al. 2010). Figure 2.6 shows more detailed breakdown results of WT sub-assemblies with data covering 1400 WT-year. The failure rate lessons from ReliaWind project are similar to the last study, but the downtime lessons are different showing greater emphasis on the power and rotor modules because it is believed these newer variable speed WTs have not yet experienced any major gearbox, generator or blade failure to date in service (Tavner 2012).

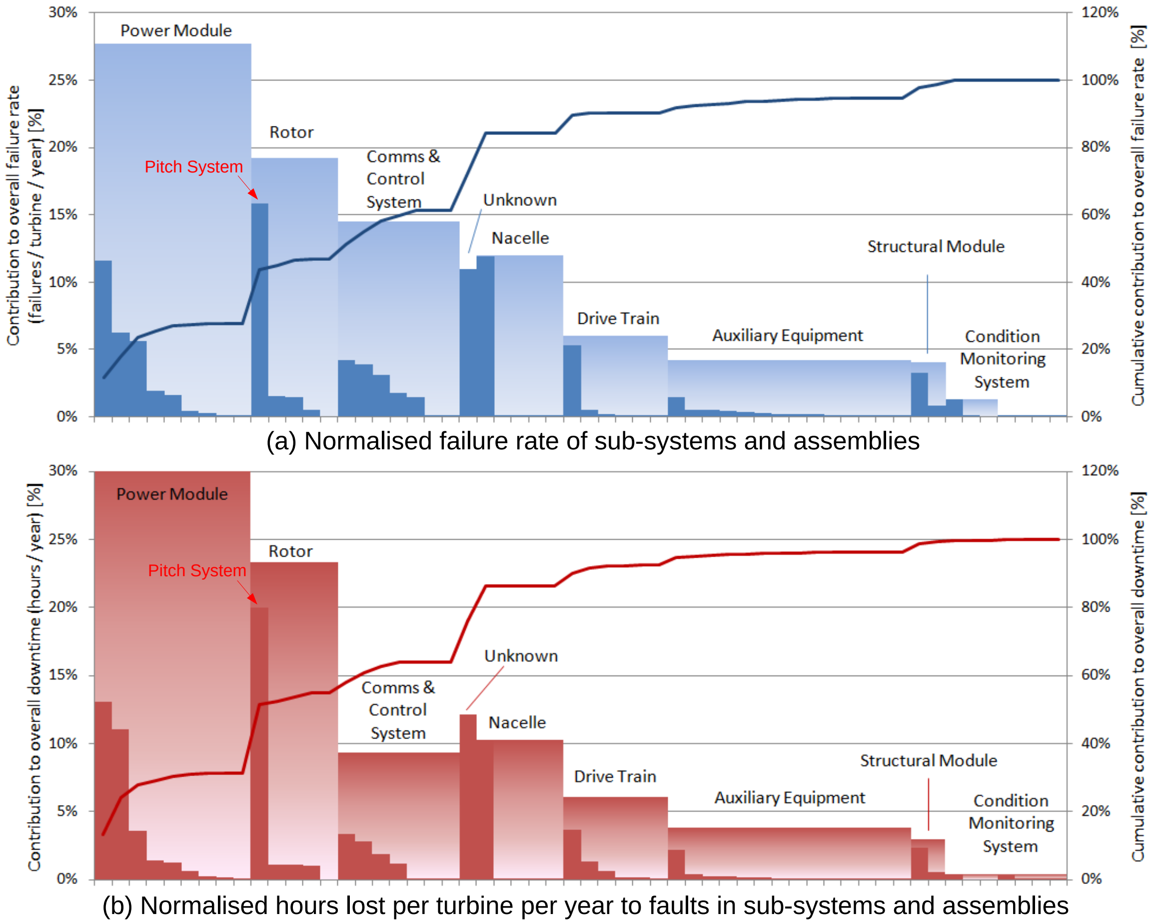


Figure 2.6: WT sub-assembly reliability analysis, the 1,400 WT-year, 2004-2010 (Wilkinson et al. 2010).

2.3.4. Current Reliability Knowledge

On the basis of above studies, the current knowledge of WT reliability can be summarised as follows:

- Firstly, WT failure rates tend to increase as WT rating grows. This will be of more concern as WTs move to offshore locations.
- Secondly, higher mean wind speeds result in increased WT failures.
- Thirdly, there is a significant cross-correlation between the failure rate, weather and turbulence rather than just wind speed.
- Finally, WT sub-assemblies with the highest failure rate and downtime from public domain survey are shown in Table 2.1, in descending order of significance:

	Failure Rate	**Downtime**
High	Pitch system	Gearbox
	Converter	Generator
	Electrical system	Rotor blades
	Rotor blades	Pitch system
	Generator	Converter
	Hydraulics	Electrical system
Low	Gearbox	Hydraulics

Table 2.1: Wind turbine sub-assemblies failure rate and downtime

In addition, a recent work by (Tavner et al. 2007) has shown that for onshore WTs have failure rates of around 1-3 failures per WT per year, for failures of >24 hours, are common. In a later paper, Spinato et al. (2009) suggested that a maximum failure rate of 0.5 failures per WT per year is likely to be necessary on offshore WT, where planned maintenance visits occurring no more than once per year would be desirable.

Moreover, according to (Faulstich et al. 2011), 75% of onshore WT faults cause 5% of the downtime, whereas 25% of the faults cause 95% of the downtime. Most of the downtime on onshore WT is dominated by a few large faults, many associated with gearboxes, generators and blades, where the corresponding replacement procedure is complex and costly. Work by Tavner et al. (2011) notes that the 75% of failures with low downtimes onshore will have a significant effect in the move to offshore as quick repairs will not be possible due to access issues.

The recent ReliaWind project, shown in Figure 2.6, has shown that in that survey the pitch system was responsible for 15.5% of failures and 20% of total downtime and is the largest contributing assembly in both cases.

2.4. Review of Wind Turbine Fault Detection & Diagnosis Methods

Fault detection & diagnosis (FDD) schemes in industrial processes are becoming increasingly important, because of growing demands for higher product quality, safety and operational reliability (Braun and Herrick 2003). Fault diagnosis refers to the ability to identify the nature and cause of a specific fault (Isermann 2006). Fault prognosis refers to a reliable and sufficiently accurate prediction of the remaining useful life of equipment in service (Muller et al. 2008). An FDD scheme for a WT would allow a predictive maintenance scheme to be introduced, reducing WT downtime and increasing the annual energy production.

As has been noted in Chapter 1, that the detection of WT faults, particularly offshore, is gaining importance because of the need to reduce the Cost of Energy. This idea of successfully detecting incipient faults before they develop into failures has led to the development of a number of WT FDD methods. According to Isermann and Balle (1997), FDD methods can be classified by the way process knowledge is incorporated into either model or signal-based methods. When a process is too complex to be modelled analytically and signal analysis does not yield an unambiguous diagnosis, a fault detection approach based on AI can be used. This section covers the research found in the literature that is more closely related to WT FDD.

2.4.1. Model-based Approaches

In model-based FDD techniques some models of the system are used to decide about the occurrence of fault (Venkatasubramanian et al. 2003). In this approach, the system models can be mathematical or knowledge-based. Some methods, e.g. parameter estimation, parity equations or state observers are often used to indicate abnormal status. After that, following the different symptoms' fault diagnosis procedure, faults are determined by applying inference or classification methods. A typical example is the WT Condition Monitoring Test Rig (WTCMTR) Matlab model developed at Durham University by Zaggout (2013).

The work by Zaggout (2013) developed a diagnostic technique for detecting rotor and stator electrical asymmetries in a WT doubly fed induction generator (DFIG) using generator control signals. A mathematical model of the WTCMTR was developed and built in the Matlab representing the electrical and mechanical parts for each component, grid and losses. The effectiveness of the proposed method has been evaluated by using this WTCMTR Matlab model and control the DFIG with a stator flux oriented vector control.

2.4.2. Signal-based Approaches

A signal or feature-based method for fault detection is based on the analysis of measured output signals. Suitable features of the measured signals are used to evaluate the operating conditions. These features can be studied in either the time or frequency domains, some typical examples used for WT are signal analysis method for detecting incipient WT gearbox failure developed by

Crabtree (2011) and variance analysis in WT gearbox developed by Feng et al. (2011).

The work by Crabtree (2011) developed a signal processing algorithm based on experience with analysis techniques and their relationship with the variable speed characteristic of a large WT. The algorithm is based on iterative localised discrete Fourier transform ($IDFT_{local}$) analysis and allows the analysis of fault-related speed-dependent frequencies within non-stationary signals such as those encountered on a WT. The verification of the algorithm has been tested in three possible fault-like conditions on the Durham WTCMTR:

- Rotor electrical asymmetry;
- High speed shaft mass unbalance;
- High speed gear tooth damage;

The work concludes with a comparison between $IDFT_{local}$ and localised continuous wavelet transform (CWT_{local}) analysis in terms of clarity of result and computational intensity. The result shows that $IDFT_{local}$ analysis has less computing time and better clarity of the results, which is highly important in the field of FDD.

A signal-based approach for detecting WT gearbox and generator faults using SCADA and CMS signals was proposed by Feng et al. (2011). The authors summarised the typical principal failure modes of WT gearbox and the relevant SCADA & CMS measurements for health monitoring. Starting with the basic physics of heat generation and temperature rise due to a fault based on the first law of thermodynamics, they then derive the relationship between temperature, efficiency, and power output or rotational speed. This lead to a new algorithm using oil and bearing temperature SCADA data to detect gearbox failures. In the this work, another case study extracted diagnostic information from both enveloped amplitude and oil debris particle count CMS data and plotted it against WT energy generation. This research showed that SCADA signal analyses using simple algorithms can give early warning of gearbox failures and the analysis of CMS signals can locate and diagnose failures with detailed information. The research suggested that in future WT monitoring systems both SCADA and CMS signals should be used to detect faults and schedule maintenance.

Zappalà et al. (2013) proposed an approach based on Sideband Power Factor (SBPF) algorithm for incorporation into a commercial CMS for automatic gear fault detection and diagnosis. The algorithm has been successfully tested on the Durham WTCMTR from which a gear condition indicator has been proposed to evaluate the gear damage during non-stationary load and speed operating conditions. The performance of the proposed approach has also been successfully tested on signals from a field test of full-size WT gearbox which has sustained gear damage. The result shows that the proposed technique proves efficient and reliable for detecting gear damage. The author concluded that once implemented into WT CMSs, this algorithm can automate data interpretation reducing the quantity of information that WT operators must handle.

2.4.3. Expert System & AI approaches

Sometimes, a process is too complex to be modelled analytically and a signal analysis approach will not yield a reliable FDD scheme, for example certain fault combinations have different effects on the system behaviour. It is then possible to classify faulty behaviour by using qualitative process knowledge to evaluate relations between measured signals and the current operating condition. These methods include approaches like probabilistic methods, fuzzy logic techniques, artificial neural network (ANN), and Bayesian network (BN).

All of these methods can be used for the purpose of extracting or inferring knowledge from large volumes of sensor data. The following section provides an outline of the named methodologies along with some examples from the literature detailing applications of how they have been utilised for the application of early fault detection for WTs.

The Condition Monitoring for Offshore Wind Farms (CONMOW) (Wiggelinkhuize et al. 2007) was a collaborative project carried out by a number of large and well-established institutes in WT research. Its purpose was to investigate the notion of a cost-effective integral CM system for WT monitoring with a specific focus on the development of data analysis algorithms. These algorithms were to be integrated into the SCADA systems to produce accurate information to aid O&M whilst attempting to lower the cost of CM systems. At the time of drafting the state of the art of CM techniques report, the authors stated that there were no successful WT CM example applications to be found in the literature.

Research by Yang and Jiang (2011) pointed out that the SCADA data are the simplest resource for developing a WT condition monitoring system. A basic idea of how a WF SCADA system could contributes to establishing a Reliability Centred Maintenance strategy is described in this paper. Some examples have been given for providing a clear explanation of the opinions. The paper concludes with the opinion that SCADA data are of importance for carrying out reliability analysis and is the cheapest resource for existing WTs.

In 2007, Singh et al. (2007) utilised the ANN approach for WT power generation forecasting. The paper mentioned that the various factors aside from the obvious wind speed and direction which can affect WT power output, such as air density, topography of the site e.g. hills & mountains, which can cause the wind profile to deviate from the ideal case. The paper concluded with a comparison between traditional methods of power estimation, using the manufacturer's power curve, and the ANN approach. The results show that the ANN offered over a monthly period a much more accurate estimation, closer to actual generated power, than the traditional method. Although this research was not intended for WT FDD, power estimation is potentially useful for FDD because abnormal power output without human interaction can be regarded as a possible fault.

A system called SIMAP was developed by Garcia et al. (2006) based on ANN for detecting and diagnosing gearbox faults. The system was split into a number of components:

- A fault detection module based on normal behaviour modelling utilising ANN's tailored towards the gearbox of a WT.
- A diagnosis module based on a simple fuzzy expert system consisting of three main rules.
- An automated maintenance scheduling calendar.

The result of the ANN normal behaviour model shows its capability of detecting a gearbox fault 2 days before the actual failure which is a positive and interesting result, although 2 days warning is rather too short for offshore application. The aim of the system was to aid the operator in their decision-making process by informing them of events that are important to them. In this way, the WT operator can make the decision based on the evidence supplied by the system, rather than being detached from the decision making process completely.

Zaher et al. (2009) proposed an automated analysis system also based on ANN. The study described a set of techniques that can be used for early fault identification for the main components of a WT. The results from 52 WT-year of data have shown that those techniques can automatically interpret the large volumes of SCADA data presented to an operator and highlight only the important aspects of interest to them. In this way, the system dramatically reduces the information presented to the operator, therefore allowing them to make more informed decisions regarding the WT maintenance. In addition, the proposed multi-agent platform allows the techniques to be brought together to corroborate their output for more robust fault detection. It also allows the development of a system that can be used to apply the techniques across a complete WF, therefore offering only one point of contact for an operator that provides all of this information in a clear and concise manner.

A recent study from Qiu et al. (2012) has investigated the Key Performance Indicators (KPI) of SCADA alarms from 4 onshore WFs, with considering 153 1.67MW variable speed, pitch controlled WTs over 2 years and 366 2.5MW variable speed, variable pitch WTs over 1 year, that is about 672 WT-years of data. The results show an average alarm rate varying from 4-20 per 10 minutes and maximum alarm rate varying from 390-1,500 per 10 minutes. These are very high figures from relatively small onshore WFs and the alarm rate would need to be reduced to be interpretable by operators or maintainers. Qiu et al. then introduced a time-sequence and a probability-based analysis method to analyse SCADA alarm data. These two methods have proved to be potential for rationalising and reducing alarm data providing fault detection, diagnosis and prognosis from the conditions generating the alarms.

A pattern recognition approach for identifying WT pitch fault was proposed by Chen et al. (2011) based on a feasibility study of SCADA alarm processing and diagnosis on 10 WT-year of data using ANN. At the beginning, 3 criteria were defined to identify pitch system faults and used to generate the training data. The back-propagation network (BPN) algorithm was used to supervise a three layers network with different numbers of neuron in the hidden layer to identify WT pitch fault. The trained ANN was then applied to another 4 WTs to find similar pitch faults. In the end, the highest accuracy rate 47% was gained from one tested WT with 50 neurons in the hidden layer. This study found that the general mapping capability of ANN can help to identify those most likely WT faults from SCADA alarm signals, but a wide range of representative alarm patterns are necessary for supervisory training.

A data-driven approach for monitoring WT pitch faults was proposed by Kusiak and Verma (2011) based on 9 WT-year of data. At the beginning, two pitch faults, blade angle asymmetry and blade angle implausibility were analysed to determine the associations between them and the components/sub-assemblies of the WT. After that, five different data-mining algorithms were studied to evaluate the quality of the models for prediction of WT pitch faults. Genetic programming (GP) was found to have the best accuracy and was selected to perform prediction at different time stamps. In the end, the solutions obtained by GP provide an easy-to-understand relationship between the input parameters that classify an output as a fault/non-fault. However, in this research, due to the limited volume of the data, only the blade angle implausibility was predicted.

Gray and Watson (2010) present a methodology for damage calculation applied to a typical 3 stages WT gearbox design from 400 WT-year of data based on the concept of physics of failure. The authors state that damage is generally accumulated due to an "irreversible change that takes place in the microstructure of a component subjected to certain loading or environmental conditions". The methodology is illustrated using a case study of a large wind farm where a significant number of gearbox failures occurred within a short space of time. The proposed methodology is extremely positive and clear, however it requires an in depth understanding of the dynamics of the gearbox under all kinds of conditions and loads.

Schlechtingen et al. (2012) proposed a system for WT FDD using Adaptive Neuro-Fuzzy Inference System (ANIFS). ANFIS normal behaviour models were developed in this research in order to detect abnormal behaviour of the captured signals and indicate component malfunctions or faults using the prediction error. In addition, the Fuzzy Inference Systems were also used to capture the existing expert knowledge linking observed prediction error patterns to specific faults. The final outputs are the condition of the component and a possible root cause for the anomaly. This work is based on continuously measured SCADA data from 18 2MW WTs covering a period of 30 months, that is 45 WT-year of data. The proposed method in this paper shows a novelty which regards the usage of ANFIS and the application of the proposed procedure to a wide range of SCADA signals. However, the number of the membership function in ANFIS models and the probability 0.01% for identifying the prediction error are not clear; both of them need to be classified.

Kusiak and Li (2011) presented a methodology to predict WT faults using SCADA data. The methodology involves three consecutive levels:

- The existence of a status/fault was identified;
- The category/severity of the fault was predicted;
- A specific fault was predicted;

Four data sets, each of them was collected at period of three months, were used in this study. Several data-mining algorithms, the ANN, the Neural Network Ensemble (NN Ensemble), the Boosting Tree Algorithm (BTA), the Support Vector Machine (SVM), the Standard Classification and Regression Tree

(CART) have been applied at each level to extract the model. The research demonstrated that faults can be predicted with a reasonable accuracy 60 min before they occur in most cases. The author concluded that prediction accuracy of the fault category is somewhat lower but acceptable and this may due to data limitation which cause less accurate in identifying a specific fault.

A summary of all these AI method is given in Table 2.2.

Author	Type of WT	Specific WT problem	Method	Data	Estimated Prognostic Horizon
Singh et al. (2007)	Zond 500 kW	Power curve	ANN	1,500 sets of 10 minutes data	Not applicable
Garcia et al. (2006)	Not known, (Owned by Molinos del Ebro S.A., Spain)	Gearbox	ANN, Fuzzy Inference System	Not known	2 days
Zaher et al. (2009)	Bonus 600 kW, stall-regulated indirect drive WT	Gearbox Generator	ANN	52 WT-year	6 months for gearbox, more than a year for generator
Qiu et al. (2012)	Variable speed variable pitch indirect drive WT	Pitch System Converter	KPIs & Venn Diagrams	306 & 366 WT-year	Not known
Chen et al. (2011)	Variable speed variable pitch indirect drive WT	Pitch System	ANN	10 WT-year	Fault Identification only
Gray and Watson (2010)	Variable speed variable pitch indirect drive WT	Gearbox	Statistics and Physics of Failure Model	400 WT-year	Fault Identification only
Schlechtingen et al. (2012)	Not known	45 normal behaviour models	ANFIS	45 WT-year	Fault Identification only
Kusiak and Li (2011)	Not known	7 most frequent faults	ANN, NN Ensemble, BTA, SVM, CART	1 WT-year	1 hour
Kusiak and Verma (2011)	Not known	Pitch System	k-NN, GP, PART, ANN, bagging	9 WT-year	Fault Identification only

Table 2.2: Summary of previous application of AI methods to WT monitoring.

2.4.4. Review Summary

From the above literature, it can be seen that:

- Most of existing WT FDD researches focus on a specific problem or component of WT in isolation, due to the complexity of modern WT and the complex nature of each individual problem.
- Fault prognostic research has been made in the gearbox and generator. Among them, Zaher's results show 6 months fault prognostic horizon for the gearbox and more than 1 year for the generator, which give enough time for WT operator to investigate the problem and schedule maintenance. However, these results were based on historical data and the final time of failure was known from maintenance logs. Therefore, it was unclear whether a prediction about time to failure could be made in a real application.
- No fault prognostics research for pitch system was found in the literature. However, the fault diagnostics for pitch systems have been studied by Kusiak and Verma (2011) and Chen et al. (2011). The importance of the pitch system has been discussed. However, both of the studies used small data sizes. It would be beneficial from FDD point of view to have more data for use in research.

None of these previous studies have tested their proposed approaches on different designs and locations of WTs, to demonstrate the adaptability of the proposed approach. More testing on different designs and locations of WTs are necessary.

2.5. Focus in this Research

The purpose of this section, based upon this earlier work, is to identify the main focus in this research and determine the appropriate data and facilities to be used.

2.5.1. Pitch System

As mentioned in Section 2.1, the pitch system controls the angle of attack of the blades to the wind to control the extraction of kinetic energy and avoid rotor over-speed at high wind speeds. The pitch system is a vital part of the modern fixed or variable speed WTs, whether it is for they use pitch-to-stall or pitch-to-feather control. This is because of the pitch system is not only responsible for regulating the WT's power output by controlling the blade angle to enhance wind energy conversion efficiency, but it also provides security braking in the case of emergency situations and high wind speeds. It requires that, the WT can be brought to a stop, with the rotor blades driven into their feathered positions, using power from a back-up system, even in the event of grid power failure (Bianchi et al. 2006; Hau and Platz 2006).

In today's wind industry, there are primarily two types of pitch system: hydraulic pitch system and electrical pitch system.

Hydraulic Pitch System

Most of earlier WTs use hydraulic pitch system (Hau and Platz 2006; Clarkson 2010), which has hydraulic actuators in the rotor hubs, rotating the blades either directly or via mechanical linkages, as shown in Figure 2.7. The most significant advantages of a hydraulic pitch system are:

- Its high driving force or torque, available from hydraulic rams power to the blade;
- Its simplicity and a robust back-up power supply, in the form of a pressurised hydraulic accumulator.

Because of these advantages, hydraulic pitch systems have historically dominated WT pitch control in Europe, North America & China. Its disadvantages include:

- Poor reliability, particularly with ram oil leakages and difficulty of detecting faults remotely;
- Power hungry as a hydraulic pump is required to operate constantly;
- In colder climates the hydraulic oil viscosity rises as temperature drops;

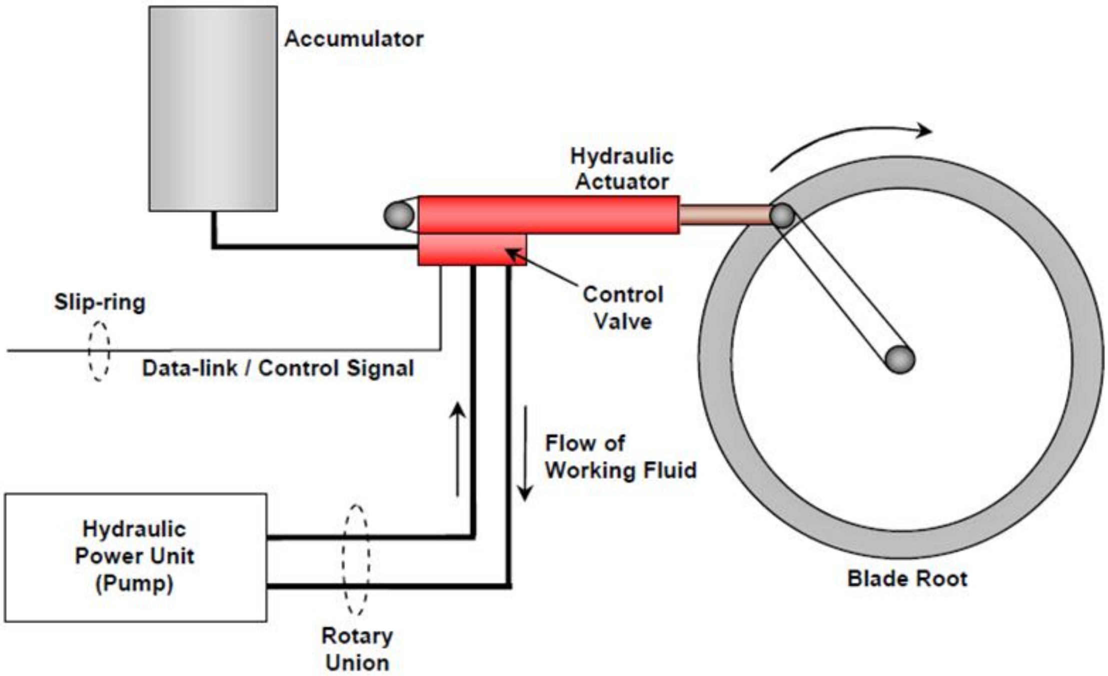


Figure 2.7: Hydraulic pitch system (Clarkson 2010).

Electrical Pitch System

On the other hand, there has been considerable progress in the development of electrical pitch systems. In an electrical pitch system (Hau and Platz 2006; Clarkson 2010), as shown in Figure 2.8, each blade is controlled by an electric servo motor connected to a gearbox which reduces the motor speed to a level to apply a high torque to the blade. The gearbox output then drives a pinion gear which engages with an internal ring gear that is rigidly attached to the rotor blade root. This kind of pitch system offers a number of benefits over hydraulic pitch system:

- It has a higher efficiency than hydraulic system;
- There is no risk of oil leakage;
- It doesn't require constantly running pump for actuation, it is therefore more power efficient;
- The major advantage of the electrical pitch system is its extended control possibilities and greater precision.

For these reasons, the electrical pitch system has become more and more popular in WTs in recent years. However, there is no doubt that the back-up power supply in electrical pitch system is the main weakness, which is usually provided by batteries with a lifetime of two or possible three years, although new battery types and ultra-capacitors are now being considered (Schneuwl, 2013, Hau and Platz 2006).

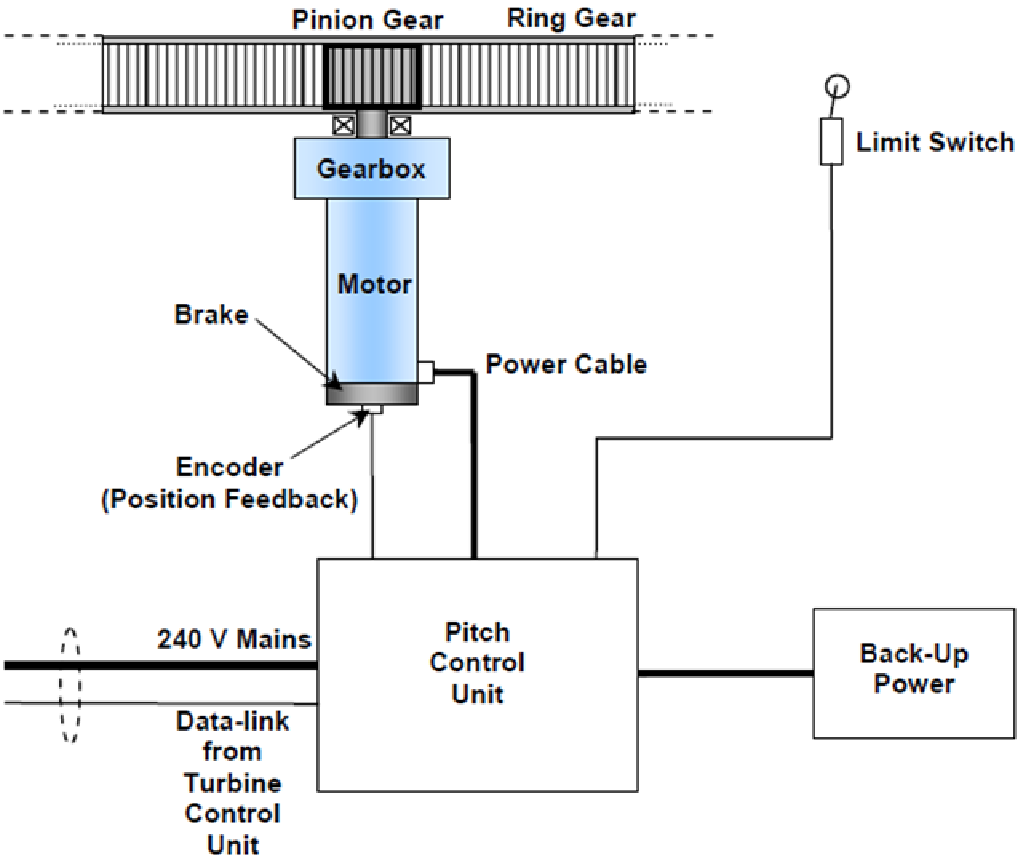


Figure 2.8: Electrical Pitch System (Clarkson 2010)

2.5.2. Why Pitch System & Aim of this Research

Nowadays, both hydraulic and electrical pitch systems are widely used in wind industry and their market share is about 55% and 45% respectively in 2009 (Dvorak 2009).

As mentioned in Section 2.3, Tavner et al. (2007) investigated WT subassembly reliability in two WT national populations during the period 1994-2004 showed that the pitch system generally has the highest failure rate.

Another recent study (Wilkinson et al. 2010), as shown in Figure 2.6, has shown that the pitch system is responsible for 15.5% of failures and 20% of the total downtime and is the largest contributing sub-assembly in both cases.

In addition, as has been noted in last Section, the pitch system is a vital part of the modern variable speed WT. Moreover, at the time of starting this research in Jan 2011, no successful WT pitch fault detection systems were reported in the literature.

Thus, this research had decided to focus on analysing WT pitch faults with the objective of developing a fault detection approach, which will extend the prognostic horizon. This research will involve the following tasks:

- Development of server and software platform to process and display existing WT data held at Durham University;
- Using the data stored in the above server to develop algorithms for automatically diagnosing and processing incipient WT pitch faults to plan maintenance intervention;
- The research will investigate the use of AI approaches to detect WT faults and manage the large-scale WF.

This thesis only considers AI approaches to develop the pitch fault detection for WT, since the other approaches have already been considered by other Durham University researchers. The main potential benefits of using AI are as follows:

- Human knowledge can be translated into a computer language to improve FDDs;
- Machine learning methods can be used to generalize system behaviour from examples;
- High potential for automation, essential for offshore WFs;
- Easy to construct as a whole system to provide WF management.

2.5.3. Research Data

There are about 2 Terabytes of real WT data available to us including SCADA and CMS data. CMS data is not considered in this research as it only monitors the WT drive train, which doesn't have any pitch data for the author to analyse.

For the SCADA data, 49GB from 5 different companies are available. However, by considering the number of WT and data availability, only two of them are suitable for this research. The information of these two suitable SCADA data is listed in Table 2.3:

Location	**Company**	**Data Size**	**Data Description**
Various locations in Spain	Alstom, Eco 80	35.2 GB	• 1.67 MW variable pitch, variable speed indirect drive machine; • Onshore; • Electrical Pitch System, pitch-to-feather; • 6 WFs; • 153 WTs; • 10 minutes data contain alarms, maintenance log; • Available from Jun 2006 to Oct 2008;
Brazos, Texas, USA	Mitsubishi M1000	13 GB	• 1 MW class variable pitch, fixed speed indirect drive machine; • Onshore; • Hydraulic Pitch System, pitch-to-stall; • 2 WFs; • 160 WTs; • 10 minutes data contain monthly report; • Available from Jun 2004 to Nov 2006;

Table 2.3: Two suitable SCADA data sources

The original SCADA data is in Office Access format. In order to conduct an efficient research, both of them have been imported into the Microsoft SQL

Server database and connected to a Data Visualisation tool developed by the author.

2.5.4. Research Facility

This research aims at improving the efficiency of investigations into the reliability of WT. This is part of large activity to research improved Knowledge Management for WFs and WTs therein. In order to conduct efficient laboratory research in this area, the New & Renewable Energy Research Group at Durham, School of Engineering and Computing Sciences has proposed the building of a dedicated Server and development of customised application to analyse the existing WT data, obtained under confidentiality agreements from various operators and manufacturers.

The Server environment includes a HP Server, Backup Devices, Uninterruptable Power Supply (UPS), and software. Their specifications are listed in Appendix A.

2.5.5. WT Data Visualisation Tool

The WT Data Visualisation Tool is a Client/Server-based application. It is a unified platform developed by the author to assist experts to conduct efficient laboratory research on WT reliability data. The Server side of this application is placed on the ReliaWind Server; it is used to handle data and process Client requests. The Client side is a graphical user interface (GUI) - data visualisation interface, which allows users to request the Server's content or services, for example raw data, data plot and data aggregation. The screenshots of the Data Visualisation Tool are shown in Appendix B.

Compared to the Excel, Access and Matlab Plot, the developed WT Data Visualisation Tool has the following advantages which could assist experts to conduct efficient laboratory research on WT reliability data:

- Fast data loading;
- Fast plotting including Line Plot and Scatter Plot;
- Fast zoom in/out with the caption to show the corresponding period in viewport;
- Be able to plot SCADA alarm data;
- Order data in different format e.g. datetime, number and character;
- Allow to see the data information by clicking the data point in viewport;

2.6. Chapter Summary

Only three-blade, horizontal-axis, up-wind WTs are considered in this research, since they represent the vast majority of WTs currently installed. The current WT reliability knowledge have shown that WT failure rates increase as the size of WT grows and this will be of more concern when WTs move to offshore locations. The existing researches about WT FDD were reviewed in this chapter. The review found most of the existing researches focus on a specific problem or component of the WT in isolation. This is mainly due to the complexity of modern WT and the complex nature of each individual problem. Fault diagnostics for the pitch system and fault prognostics for gearbox & generator systems have been reported. However, the existing fault prognostics researches are still unclear to make a prediction about the time to failure in real application. Small sizes of data were used in the fault diagnostics researches for pitch system and it would be beneficial to use more data. In addition, the review also found that none of the studies have tested their approaches on different designs and locations of WTs.

Pitch system is a vital part of the modern variable speed WT and it was found to have high failure rate. There were not many successful WT pitch fault detection found in the literature, this research then decided to focus on analysing WT pitch fault with the objective of developing the fault prognosis approach. The author only considers AI approaches since the other approaches are already being considered by other Durham University researchers.

After that, two different SCADA data sets, each of them has more than 380 WT-year data, were found to be suitable in this research. In order to conduct efficient laboratory research on WT reliability, a dedicated server called ReliaWind and a WT Data Visualisation Tool were built by author in early 2011.

In conclusion, this research speculates that there is a robust and effective AI solution which allows us to build an automated on-line fault prognosis system for WT pitch system. The system would be able to work on different designs and locations of WTs, and produce better fault prognosis results. The key questions need to be answered in this research are:

- Which AI algorithm makes possible the identification of an incipient WT pitch faults?
- Once the AI algorithm is decided, how to construct an automated on-line fault prognosis system?
- Whether the proposed system can be applied on other types of WT?
- Does this AI approach produce a better result? What is evaluation approach?

Chapter 3.

Possible Fault Detection & Diagnosis using AI Methods

AI techniques have increasingly been used in the field of FDD of industrial systems because of increasing system complexity. According to Angeli and Chatzinikolaou (2004), in very complicated non-linear systems, where valid mathematical models do not exist, the AI techniques would allow the development of new approaches for fault detection. The reason behind these uses is that AI provides the necessary association, reasoning and decision-making processes that mimic human thought processes when solving diagnostic problems. In the field of WT reliability research a large amount of data is collected from WT through various monitoring systems, e.g. SCADA and CMS. The dynamic nature of the WT operational environment makes the development of mathematical models, that capture and interpret WT operational behaviours, extremely difficult. In contrast, AI techniques can be used in a variety of ways that transform data into knowledge for the purpose of achieving a suitable mechanism for FDD. In addition, the model-based and signal-based approaches are already being considered by other Durham University researchers. These reasons formulated the basis for the decision to investigate the use of AI techniques to detect WT faults.

The FDD task consists of determining the fault types, with as many details as possible, such as the fault size, location and time of detection (Braun and Herrick 2003; Isermann 2006). The diagnostics procedure is based upon the observed analytical and heuristic symptoms and the heuristic knowledge of the process. In order to accomplish FDD, many AI approaches have been proposed and that procedure can be considered as a classification task.

3.1. Supervised & Unsupervised Learning Modes

In general, there are two different classification modes: supervised and unsupervised learning modes. It is important to understand their differences so that the potential techniques reviewed in this chapter can be put into perspective. The supervised and unsupervised learning modes refer to the training procedure and how their requirements can affect the construction of the model. Typically supervised learning requires the training data to be fully labelled, for example in the field of FDD each data instance is assigned with either a normal or abnormal class. Any unseen data is compared against the trained model to determine which class it belongs to. Example supervised learning algorithms are Naïve Bayes, ANN, SVM and ANFIS. These algorithms are reviewed in detail within this chapter.

By contrast, unsupervised learning doesn't have any classes assigned to training data. The algorithm itself needs to determine what those classes are and how to separate them. The most well-known unsupervised learning algorithms are k-Means Clustering, Fuzzy c-means (FCM) and Self-organising feature mapping (SOM). These are also reviewed in detail within this chapter.

In addition, semi-supervised learning is a class of supervised learning tasks. It uses training data that consist of labelled as well as unlabelled samples. The challenge in semi-supervised learning is that the labelled samples should be fairly accurate. Its applicability is dependent on whether the accurate labelling can be made regarding the training data.

The following section provides an outline of the possible techniques, including supervised and unsupervised learning, along with some examples detailing applications of how they can be utilised in the field of WT FDD.

3.2. Possible AI Techniques

As has been noted before, the aim of this research is to examine the use of AI techniques for WT FDD with the objective of developing an automated, on-line fault prognosis using SCADA data. This section will describe the findings of some possible AI classification techniques that could be potentially useful for early WT fault detection. This section concludes with a summary of the reviewed techniques and gives a most feasible data analysis solution. The criteria used to evaluate potential techniques in this research are: interpretability of output, accuracy of diagnosis, and availability of necessary data.

3.2.1. Acquiring Data for Investigation Purposes

The WT power curve, plotting wind speed vs. power output, is the most well-known 2D plot widely used to indicate WT performance. It also been applied for WT fault detection because any large deviation from the factory supplied power curve could be regarded as a possible fault.

For this reason, two sets of Alstom WT power data, named Data A and B, both for a six months period, are prepared from the ReliaWind server for the investigation of possible AI techniques. The two data sets are presented in Figure 3.1.

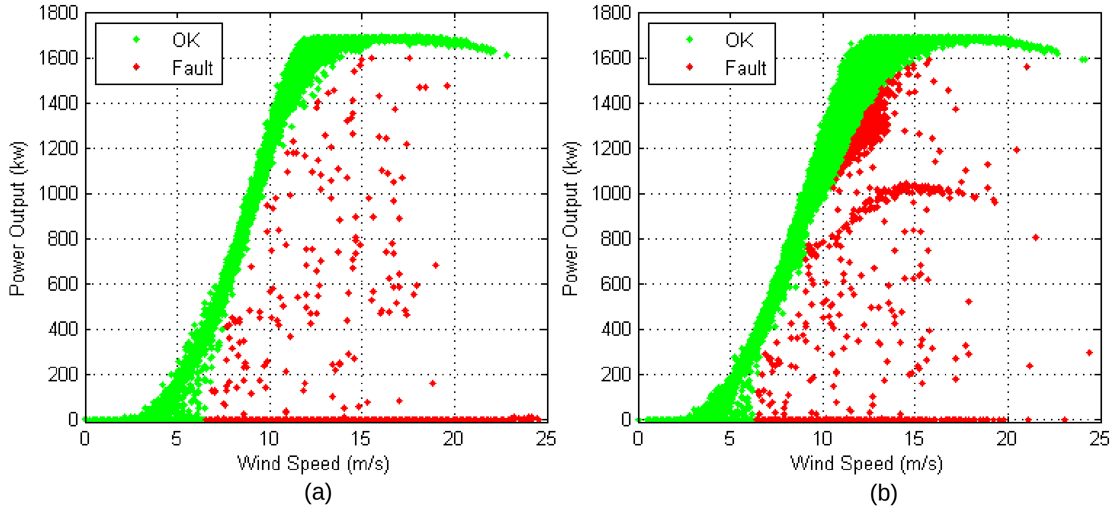


Figure 3.1: WT power data for investigation purpose (a) Data A; (b) Data B.

In addition, two 3D data sets, representing variable-speed pitch-to-feather control strategy (Bianchi et al. 2006) of a WT, are prepared and shown in Figure 3.2. Both of them are six months period and named Data C and D. The variable-speed pitch-to-feather control strategy plot on the wind speed, rotor speed and pitch angle space.

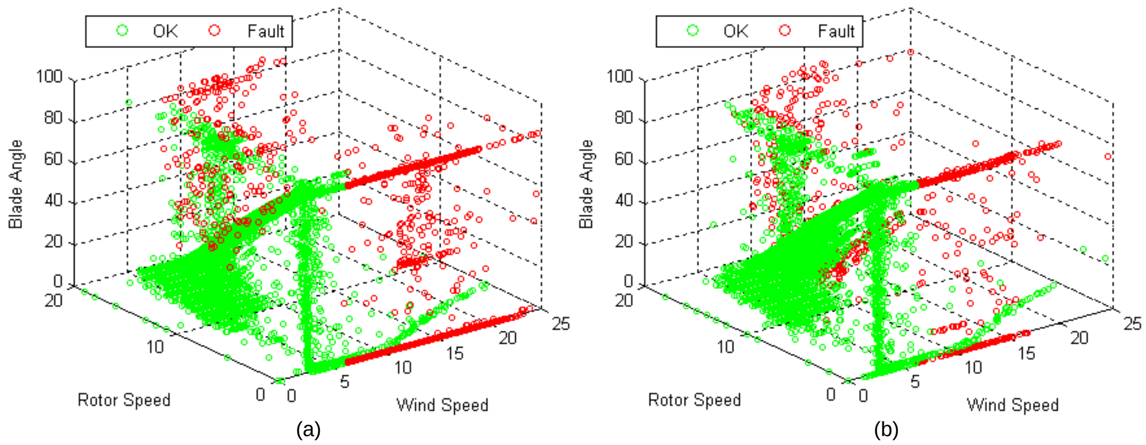


Figure 3.2: WT 3D data sets (a) Data C; (b) Data D.

3.2.2. Naïve Bayes Classifier

One classification scheme is given by the so-called Naïve Bayes Classifier. It is a statistically-based supervised learning technique. The approach, based on Bayes' theorem, is particularly suited when inputs dimensionality is high. Despite its simplicity, Naïve Bayes can often outperform more sophisticated classification methods (Wu et al. 2008).

Naïve Bayes classifier can handle an arbitrary number of independent variables whether continuous or categorical. Given a set of feature variables, $X = \{ F_1, F_2, \dots, F_n \}$, and suppose the task is to construct the posterior probability for the event $C_j, 1 \leq j \leq i$ among a set of possible outcomes $C = \{C_1, C_2, \dots C_i\}$. According to Bayes' theorem, the task can be represented as:

$$p(C_j | F_1, \dots F_n) = \frac{p(C_j) * p(F_1, \dots, F_n | C_j)}{p(F_1, \dots, F_n)} \tag{3.1}$$

In plain English the above equation can be written as:

$$posterior = \frac{prior \times likelihood}{evidence} \tag{3.2}$$

In practice we are only interested in the numerator of that fraction, since the denominator does not depend on C and the values of the features F_i are given, so that the denominator is effectively constant. According to Bayes' theorem, the numerator is equivalent to the joint probability model:

$$p(C_j) * p(F_1, \dots, F_n | C_j) = p(C_j, F_1, \dots, F_n) \tag{3.3}$$

By introducing the "naive" conditional independence assumption, which assume that each feature F_i is conditional independent of every other feature F_j for $i \neq j$. Then, the joint model can be expressed as:

$$p(C_j, F_1, \dots, F_n) = p(C_j)p(F_1|C_j)p(F_2|C_j)p(F_3|C_j) \cdots p(F_n|C_j)$$

$$= p(C)\prod_{i=1}^{n} p(F_i|C) \qquad (3.4)$$

Therefore, under the above independence assumptions, the conditional distribution over the class variable C_j can be expressed like this:

$$p(C_j|F_1, \dots, F_n) = \frac{1}{Z}p(C_j)\prod_{i=1}^{n} p(F_i|C_j) \qquad (3.5)$$

where Z, the evidence, is a scaling factor dependent only on $F_1, \dots, F_n$.

Naïve Bayes can be modelled in several different ways including Normal, Lognormal, Gamma and Poisson density function. However, a common procedure is to assume Normal density function like below:

$$p(F_i|C_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}{}^2}} e^{-\frac{(F_i-\mu_{ij})^2}{2\sigma_{ij}{}^2}} \qquad (3.6)$$

where μ_{ij} is the mean and σ_{ij} is standard deviation.

In the implementation of WT fault diagnosis, supposed to have the following training data for a specific design of 2MW WT.

WT Status	*Power Output (kW)*	*Wind Speed (m/s)*
OK	1800	12
OK	……	
OK	1650	10
Fault	800	6
Fault	……	
Fault	1300	9

Table 3.1: Labelled training data

The classifier created from above training data using a **Normal Distribution** assumption would be:

WT Status	*Mean (Power Output, kW)*	*Std_Dev (Power Output, kW)*	*Mean (Wind Speed, m/s)*	*Std_Dev (Wind Speed, m/s)*
OK	1716.7	76.4	11	1
Fault	1066.7	251.7	7.3	1.5

Table 3.2: Mean and standard deviation of the training data

Then, there is a sample to be classified as "OK" or "Fault":

WT Status	*Power Output (KW)*	*Wind Speed (m/s)*
To be classified	1750	12

Table 3.3: A sample to be classified

We wish to determine which posterior is greater, "OK" or "Fault". For the classification as "OK" the posterior is given by:

$$posterior(ok) = \frac{p(ok) * p(PowOutput|ok) * p(WindSpeed|ok)}{evidence} \quad (3.7)$$

For the classification as "Fault" the posterior is given by:

$$posterior(Fault) = \frac{p(Fault) * p(PowOutput|Fault) * p(WindSpeed|Fault)}{evidence} \quad (3.8)$$

The evidence may be ignored as it is the same in $posterior(ok)$ and $posterior(Fault)$. Finally, based on the Normal Distribution assumption, we can get the posterior numerator of $p(ok) = 0.062$ and the posterior numerator of $p(Fault) = 0.0057$. Since posterior numerator is greater in the "OK" case, this sample is identified as "OK".

Naïve Bayes classifier is designed for use when features are independent of one another within each class. One test, using it to learn the knowledge from WT power data A, applied the classifier to data A and B, Figure 3.1, received the error rates 4.22% and 4.37% respectively. The other test, using it to learn the knowledge from WT 3D data C and applied the classifier to data C and D have received the error rate 4.1% and 13.25% respectively. The two tests show that Naïve Bayes is possible to detect WT fault in practice even if the independence assumption is not valid, for example in the first test the independence assumption between wind speed and power output is invalid, in the second test the independence assumption between wind speed, rotor speed and blade angle are invalid either. The accuracy result has shown that Naïve Bayes has the potential to detect WT fault, however, the independence assumption has made it difficult to interpret the classification result.

3.2.3. Rule-based Expert System

A rule-based expert system employing logic programming can also be an effective classifier (Stuart and Peter 2010). IF-THEN rule-based algorithms are attractive when the patterns representative of a particular class of operation can be easily identified. The main benefit of using them is because it clearly demonstrates the state of the art in building knowledge representation and illustrates the cause and effect of the problem. In a rule-based expert system, much of the knowledge is represented as conditional sentences relating statements of facts with one another. Modus ponens is the primary rule of inference by which a system adds new facts to growing knowledge database. The basic framework of rule-based system is showed below:

IF **A** IS **TRUE**	OR	**IF A**
AND A IMPLIES **B**		**AND A → B**
THEN **B** IS **TRUE**		--------------
		THEN B

In a real WT gearbox fault case, rules can be defined as:

Rule 1 **IF** The cooler oil temperature is NORMAL **AND** Gearbox main bearing temperature is HIGH **THEN** Failure in the gearbox main bearing is CERTAIN ***Note:*** NORMAL and HIGH can be defined using fuzzy logic

In real system development, the Propositional Logic is widely used for the representation of information and knowledge in Computer System (Stuart and Peter 2010). Propositional Logic is fairly restrictive, which allows us to write sentences about propositions – statements about the world – which can either be true or false. The symbols in this logic are:

- Capital letters represent proposition such as P represents "The cooler oil temperature is Normal";
- Connectives which are : and (∧), or (∨), implies (⟶) and not (¬);
- Brackets;
- T stands for the proposition "true" and F stands for the proposition "false".

Therefore, above rule can be represented as:

Rule 1 $P \wedge Q \longrightarrow \mathrm{R}$

Moreover, if we program an intelligent agent with the semantics of the above propositional logic, for example in our WT gearbox example we could tell it that if "Cooler oil temperature is Normal" and "Gearbox main bearing temperature is High", it can infer "Failure in the gearbox main bearing". In the real implementation, known facts about a domain is converted into a Truth Table and used to deduce new facts (Stuart and Peter 2010). This enables an intelligent agent to prove things, for example to start with a set of statements, Axioms in a Truth Table, we believe to be true and deduce whether another statement, or Theorem, is true or not. This technique is known as Making Deductive Inferences and it has been widely used for developing rule-based expert system. For WT reliability analysis, the rule-based expert system has been applied in SIMAP (Garcia et al. 2006), an intelligent system for predictive WT gearbox maintenance. Another application, employing thresholds as shown in Figure 3.3, was made by Moorse (2010) in Durham University to investigate the healthy and faulty behaviour of WT pitch system, where faults had been observed in the data held on the ReliaWind Server. Due to the lack of diagnostic knowledge available to codify as rules, the author decided not to take this technique any further.

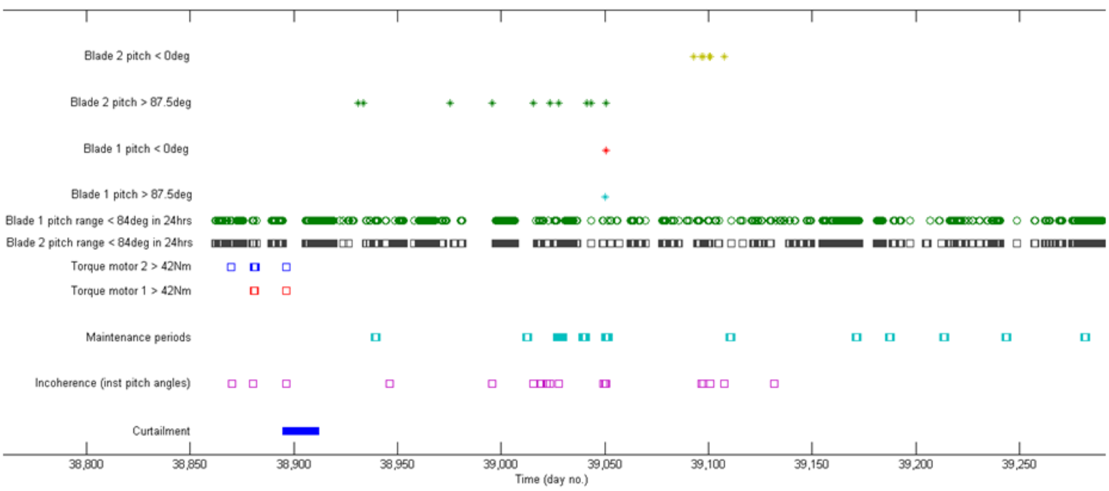


Figure 3.3: A rule-based algorithm employing thresholds to investigate the healthy and faulty behaviour of WT pitch system.

3.2.4. Fuzzy Inference System

The Fuzzy Inference System (FIS) is a computing framework based on the concepts of fuzzy set theory, fuzzy if-then rules, and fuzzy reasoning. It has found successful applications in a wide variety of field, such as automatic control, data classification, decision analysis and expert systems (Sturart and Peter 2010, Jang et al. 1997). The basic structure of a FIS consists of three conceptual components:

- A rule base, which contains a selection of fuzzy rules;
- A database or dictionary, which defines the membership functions used in the fuzzy rules;
- A reasoning mechanism, which performs the inference procedure upon the rules and gives the reasonable output.

The most widely used FISs are Mamdani, Takagi-Sugeno and Tsukamoto FISs (Jang et al. 1997). All of them take crisp inputs and are able to generate crisp outputs by applying the defuzzification method. The differences between them lie in the consequents of their fuzzy rules and the aggregation & defuzzification procedures. The detail description of these three FISs can be found from (Jang et al. 1997).

The initial test of using Takagi-Sugeno FIS to estimate the WT power output was implemented by the author. Three fuzzy rules, as listed below, were obtained based on the knowledge about how the WT power output varies with wind speed, as shown in Figure 3.3 (a). The generalised bell function was chose for the membership function because of its smoothness and concise notation (Jang et al. 1997). In addition, linear regression was utilised on the normal power data A to obtain the corresponding consequent functions, as shown in Figure 3.3 (b).

$$\begin{cases} If\ X\ is\ low, then\ Y\ is\ f_1(x) \\ If\ X\ is\ medium, then\ Y\ is\ f_2(x) \\ If\ X\ is\ high, then\ Y\ is\ f_3(x) \end{cases} \quad (3.9)$$

where X represents wind speed and Y represents the WT power output.

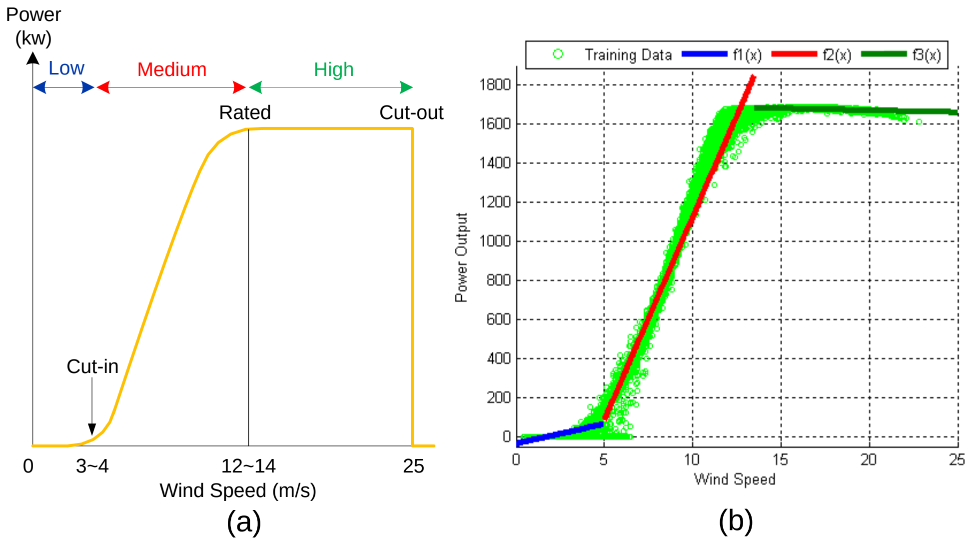


Figure 3.4: (a) WT power curve; (b) Linear regression applied to the data A.

By using Matlab Fuzzy toolbox, the Takagi-Sugeno FIS model for WT power output estimation was built and the results shown in Figure 3.4. The fault then can be identified when observation deviate from this obtained power curve. However, the real WT power output data cannot match the obtained power curve precisely, for reasons of the dynamic variations in energy in the wind and the dynamic state of the WT, as shown in Figure 3.3(b). Therefore, an acceptable upper and lower bound of the power curve is needed to avoid false identifications. This approach can also be improved by introducing more fuzzy rules, for example rule to describe the cut-in, rated speed and cut-out conditions.

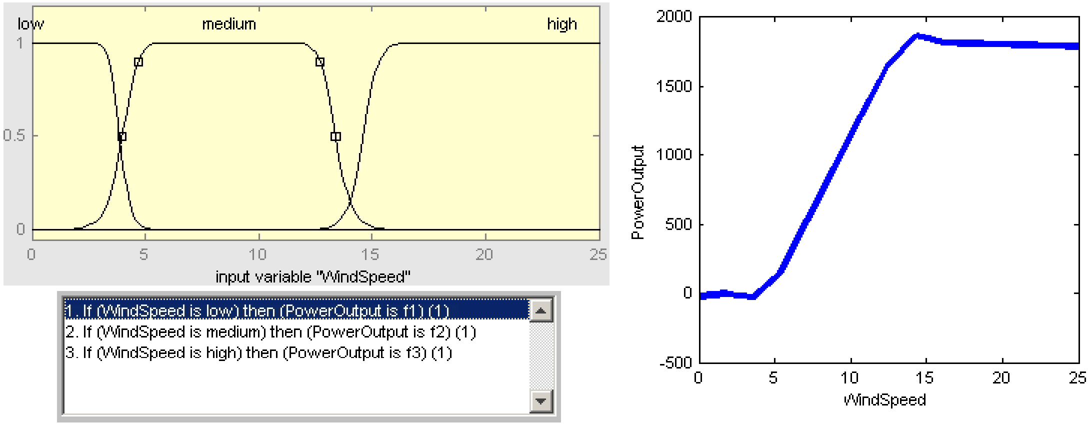


Figure 3.5: The Takagi-Sugeno FIS for WT power output estimation

The advantages of this approach are simple, interpretable and of low computational cost. However, the FIS's disadvantages or problems are seen to be difficult for non-experts for the following reasons:

- No standard methods for transforming human knowledge into a FIS. The if-then rules, the type of Membership Functions (MF) and parameters for them are normally prepared by an expert;
- The need for expert fine-tuning for finding optimal fuzzy rules, MF type and parameter value of MF. There is a need for effective methods to do this;
- The upper and lower bounds of the obtained power curve are needed to avoid false identifications.

The aforementioned problems of the FIS can be solved by the integration of FIS and ANN. This approach was developed by Jang (Jang 1993; Jang et al. 1997) and named the Adaptive Neuro-Fuzzy Inference System (ANFIS).

3.2.5. Decision Tree

A decision tree is established as a graphical tool for the visualisation of relationships in decision analysis, to help identify a strategy most likely to reach a goal (Wu et al. 2008). Decision trees are non-parametric supervised learning which do not require any assumptions about the distribution of variables in each class.

Decision trees are useful and intuitive graphical tools for displaying relationships that lead to faults with a hierarchical structure that aids human comprehension. They are common for the analysis and diagnosis of safety-critical applications. Quantitative failure probabilities can also be derived from a decision tree. They require information about the failure probabilities of individual elements. By combining these with the relationships from the decision tree it is possible to calculate the probability of a system or component fault. An algorithm for learning a Decision Tree is a matter of choosing which attribute to test at each node in the tree. The widely used algorithms are ID3 and C4.5 (Wu et al. 2008). Both of them define a measure called information gain, also known as Entropy, to decide which attribute to test at each node.

It is interesting to see how the Decision Tree method can be used for WT FDD. A simple test was made using Matlab statistics toolbox applying the Decision Tree to the WT power data A. The result is shown in Figure 3.6 with two different views, the traditional tree and the region views. The region view, Figure 3.6(b), clearly shows that the Decision Tree method has successfully learned the fault pattern. However, some sub-regions, as encircled, are mislabelled as OK because the WT cannot produce power when wind speed is too low. The reason for this is because these sub-regions have sparse or no training data and cause the Decision Tree result to become inconsistent with domain knowledge.

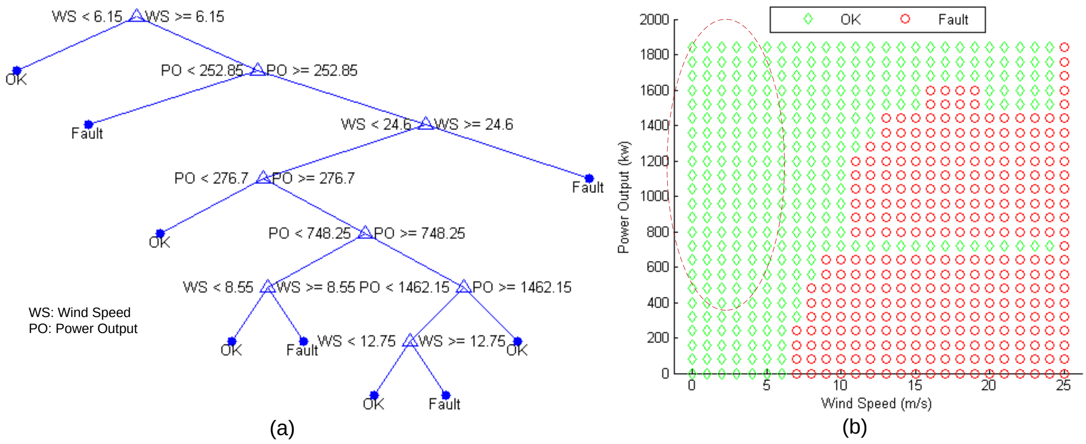


Figure 3.6: Decision tree result, (a) Tree view; (b) Region View.

3.2.6. Artificial Neural Network

An Artificial Neural Network (ANN) is a computational supervised learning model made up of many processing neurons that have a natural propensity for storing experiential knowledge and making it available for use (Wu et al. 2008; Haykin et al. 2009). Typically, the structure of an ANN consists of three layers, as shown in Figure 3.7. The first layer of inputs nodes, are connected to neurons of a hidden layer, which are connected to the neurons of the third or output layer.

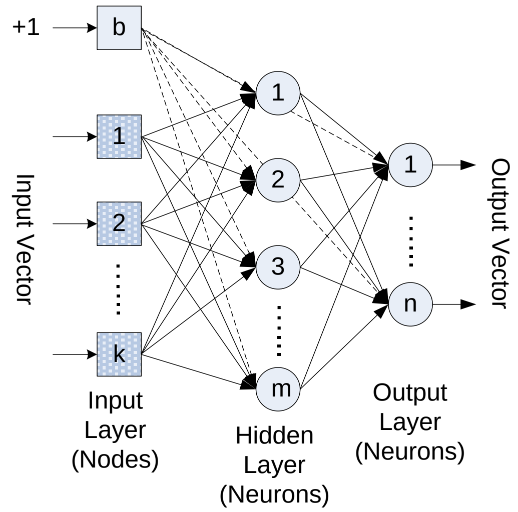


Figure 3.7: A k×m×n feed-forward ANN

The training data knowledge is stored in the connection strengths, known as weights, between the layers acquired during the training process. The network is trained in accordance with a training algorithm, which governs how connection weights are modified and adjusted in response to the training data feed into the inputs nodes and what is desired at the output neurons. In the recall process, the trained ANN accepts data fed into the input nodes and then produces a response at the output neurons as a consequence of calculated weights.

Generally, the relation of a three layers ($k \times m \times n$) feed-forward ANN can be represented in vector notation (Haykin et al. 2009). If I is a k dimensional column vector presenting inputs and H is a m dimensional column vector representing the results of hidden layer, then:

$$H = f_1(W_1 I) \qquad (3.10)$$

where W_1 is a $m \times k$ weight matrix assigned to the connection between the input and hidden layers and f_1 is an activation function using a sigmoid function, as follows:

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{3.11}$$

This function has the ability to produce a continuous non-linear threshold function and transform the input between $-\infty$ and $+\infty$ into real numbers between 0 and 1. Similarly, the n dimensional column vector for the output layer can be represented as follow:

$$O = f_2(W_2 H) = f_2\big(W_2 f_1(W_1 I)\big) \tag{3.12}$$

where f_2 is another activation function using a sigmoid function as shown in Eq. 3.11 and $W_2$is a $n \times m$ weight matrix for the connection between the hidden and output layers. As we can see from Eq. 3.12, the calculations of the output of the ANN involve two weight matrix multiplications and two applications of the activation function. Therefore, the use of this ANN will require some computational overhead. ANN training algorithms follow an iterative gradient descent principle designed to minimize the overall mean square error E, defined as:

$$\mathrm{E} = \frac{1}{N}\sum_{\mathrm{n}=1}^{\mathrm{N}} \|D_n - O_n\| \tag{3.13}$$

where:

- N denotes the number of training pattern presented to the input layer;
- D_n represents the desired output of the nth input pattern;
- O_n is the actual output of the same input pattern;
- Both D_n and O_n are vectors.

Currently, the widely used training algorithms are Back-propagation, Delta Rule, Evolutionary and Perception learning (Haykin et al. 2009).

Two ANN models were built using Matlab ANN Pattern Recognition Tool (nprtool) and data A & C respectively. Both training processes applied the default setting with 10 neurons in hidden layer and 1 neuron in output layer. The trained ANN models were tested on training data and new data (Data B & D, respectively). The testing result is shown in Table 3.4 and it demonstrates that ANN has strong potential to be applied in the field of WT FDD. However, the biggest problem for ANN is the difficulty to interpret the result because of its "black-box" nature.

ANN Model	**Testing Result**	
	Training Data (Error Rate)	**New Data (Error Rate)**
Power Curve (Trained using Data A)	0.3%	1.7%
Pitch Mechanism (Trained using Data C)	0.5%	1.9%

Table 3.4: ANN testing result.

In addition, the ANN has shown to be potential useful in the field of WT FDD on a number of occasions. SIMAP was developed by using ANN to detect gearbox faults (Garcia et al. 2006). In 2009, A. Zaher et al. introduced Gearbox and Generator Normal Behaviour Modelling (Zaher et al. 2009) also using ANN and the results provided an early warning of gearbox and generator problems. R.F. Mesquita et al. introduced a similar approach for WT gearbox condition monitoring (Brandão et al. 2010). In 2011, B. Chen et al. investigated ANN to analyse SCADA alarms for the automatic detection of WT pitch system faults (Chen et al. 2011), an ANN was constructed to learn pitch system faults from one WT and then applied on the other four other WTs to detect similar symptoms. The results show that ANN is a feasible method for on-line WT fault diagnosis and has the potential to rapidly identify WT failures and reduce WT alarm rates.

3.2.7. Self-organising Feature Mapping

Self-organizing feature mapping (SOM) was developed by Teuvo Kohonen to provide a data visualisation technique, which helps to understand high-dimensional data by reducing the data dimensions to a map. SOM also represents unsupervised learning by grouping similar data together. It can be said that SOM reduces data dimensions and displays data similarities (Kohonen 2001).

With SOM, clustering is performed by having several units compete for the current object. Once the data have been entered into the system, the network of artificial neurons is trained by providing information about inputs. The unit with the weight vector closest to the current object becomes the winning or active unit. During the training stage, the values for the input variables are gradually adjusted in an attempt to preserve neighbourhood relationships that exist within the input data set. As it gets closer to the input object, the weights of the winning unit are adjusted as well as its neighbours.

SOM clustering was tested on data A and C, representing the power curve and pitch mechanism of a WT respectively. The SOM results are shown in Figure 3.8 & 3.9.

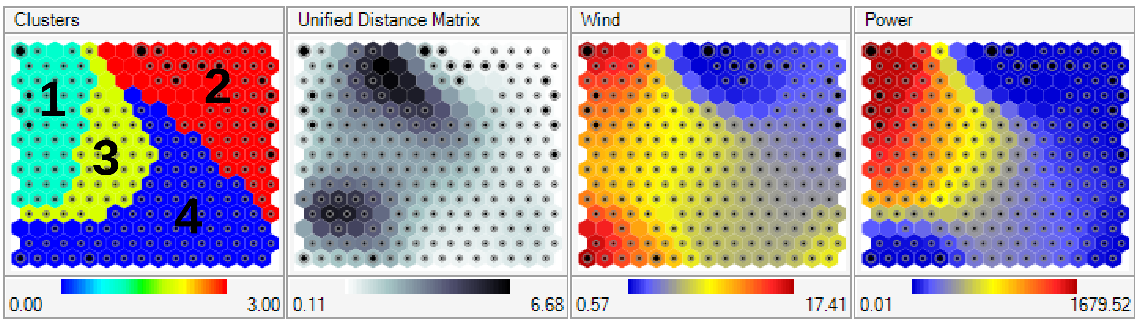


Figure 3.8: SOM test on data A.

Figure 3.8 is the SOM clustering result of training data set A. Based on the training result, different clusters revealed different operational stages of the WT.

- The Cluster 1 represents WT operation between rated and cut-out wind speeds, in which the WT produces max power output.
- The Cluster 2 represents WT operation below the cut-in wind speed. In this period, the WT is disconnected with grid and does not generate electricity.
- The Cluster 3 represents the operational period above the cut-in but below rated wind speeds. As wind speed increases, the generator produces more power.
- In Cluster 4 the wind speed is high but there is no WT power output. This may illustrate a possible WT fault.

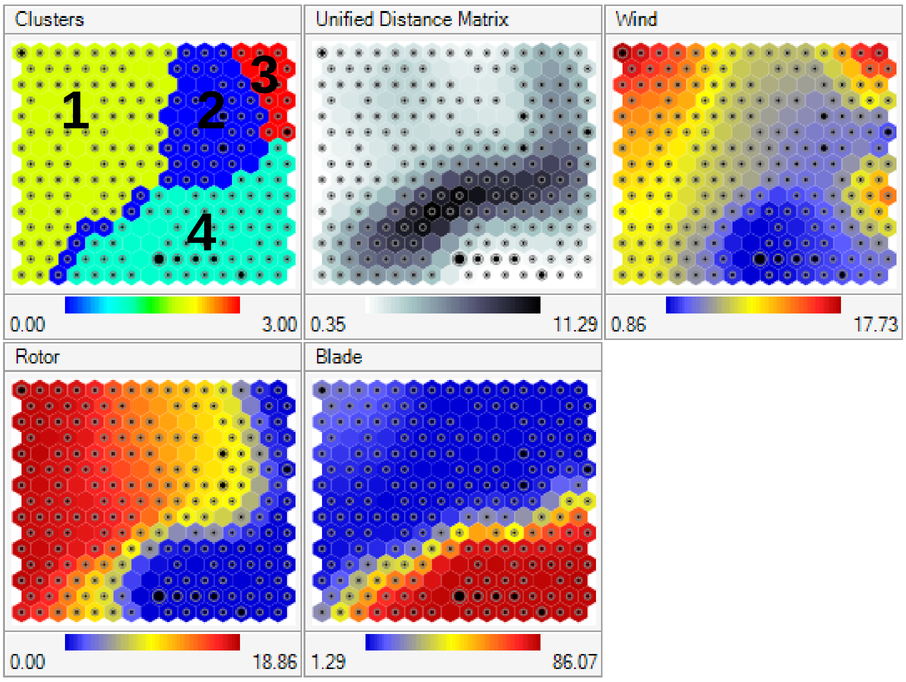


Figure 3.9: SOM test on Data C.

Figure 3.9 is clustering result of the variable-speed pitch-to-feather control strategy. Based on the training result, different clusters revealed different operational stages of the WT pitch mechanism.

- The Cluster 1 represents WT operation above rated wind speed, in which WT rotor speed remains constant. In this period, the blade angle is regulated to produce constant power output.
- The Cluster 2 represents WT operational period represents above the cut-in but below rated wind speeds. As wind speed increases, the generator produces more power.
- In Cluster 3 the wind speed is high but there is no rotor speed and blade angle is almost at optimal position. This may illustrate a possible WT pitch fault.
- The Cluster 4 represents WT operation below the cut-in wind speed, in which the WT is disconnected with grid and does not generate electricity.

SOM results have shown the capability to identify abnormal event, however, its results have to be interpreted by expert.

3.2.8. Bayesian Network

A Bayesian network (BN) is a probabilistic graphical model that represents a set of variables and their conditional dependencies via a directed acyclic graph (Korb and Nicholson 2003). For example, a BN could represent the probabilistic relationship between faults and alarms. Given alarms, the network can be used to compute the probabilities the presence of various faults.

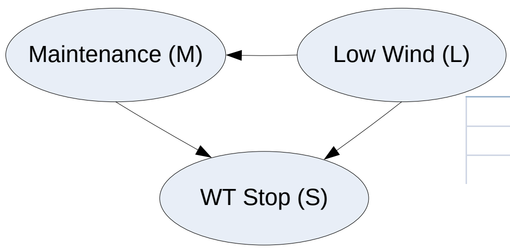
Formally, BN represented variables as nodes linked in a directed graph, as in a cause & effect model. Conditional probabilities are specified for every node with a Node Probability Table (NPT). Root causes just have an "a priori" probability. Run-time calculation generates probability estimates for every node and changes when any node receives a new observed value. Thus, BNs could perform on-line diagnosis or even prediction.

The equation of BN can be written as a product of the individual density function, conditional on their parent variables:

$$p(x) = \prod_{v \in V} p(x_v | x_{pa(v)}) \qquad (3.14)$$

where $pa(v)$ is the set of parents of v, i.e. those vertices pointing directly to v via a single edge. For example, suppose that there were two events which could cause a WT to stop: either Low Wind or Maintenance. Suppose also that the Low Wind has a direct effect on Maintenance. Then the situation can be modelled with a BN in Figure 3.10. All three variables have two possible values, T for true and F for false.

	Maintenance	
Low Wind	T	F
F	0.4	0.6
T	0.01	0.99



	Low Wind	
	T	F
	0.2	0.8

		WT Stop	
Maintenance	Low Wind	T	F
F	F	0	1
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

Figure 3.10: BN case study.

The joint probability function is $P(S, M, L) = P(S|M, L)P(M|L)P(L)$. The model can answer question like "What is the probability that it is low wind that has caused the WT to stop?". By using the conditional probability formula and summing over all nuisance variables (In probability theory, the nuisance variable is a variable that value does not affect the result of a probabilistic model. In this example, the nuisance variables are M in the numerator, M & L in the denominator, as shown in the equation below):

$$P(L = T|S = T) = \frac{P(S = T, L = T)}{P(S = T)} = \frac{\sum_{M\epsilon\{T,F\}} P(S = T, M, L = T)}{\sum_{M,L\epsilon\{T,F\}} (S = T, M, L)}$$

$$= \frac{(0.99 * 0.01 * 0.2) + (0.8 * 0.99 * 0.2)}{0.00198 + 0.288 + 0.1584 + 0} \approx 35.77\%$$

As in the previous example the numerator is highlighted explicitly, the joint probability function is used to calculate each iteration of the summation function, in the numerator marginalising over M and in the denominator marginalising over M and L.

Chen et al. (2012) investigated the possibility of using BN to analyse WT SCADA data with the objective of on-line fault diagnosis. The BN model was derived from an existing probability-based analysis method, the Venn diagram (Qiu et al. 2012), and based upon 26 months of historical SCADA data. The research results have shown that the BN is a valuable tool for WT fault diagnosis and has great potential to rationalise failure root causes. Compared to the Venn diagram approach the BN could rationalise the data better and was more feasible for on-line fault diagnosis. A drawback of using the BN, pointed out by authors, was that BN complexity grows exponentially with the increase of parent node numbers. Due to the complexity of BN and lack of diagnostic knowledge to build a BN for WT FDD, we decided not to take this technique any further.

3.2.9. k-Means Clustering

k-Means clustering algorithm was developed by MacQueen (1967). Simply speaking k-means clustering is an unsupervised learning algorithm to classify or to group data based on features into K number of group. K is positive integer number. The grouping is done by minimising the sum of squares of distances between data and the corresponding cluster centroid, for example the most widely used Euclidean distance.

The k-Means clustering algorithm is composed of the following steps:

- Classify the number of the cluster K.
- Randomly place K points into the space and each of them represent the centroid of the corresponding cluster.
- Assign each object to the closest centroid.
    - When all objects have been assigned, recalculate the positions of the K centroids.
    - $K_m(x, y, \ldots, z) = \left(\frac{\sum_{i=1}^{n} x_i}{n}, \frac{\sum_{i=1}^{n} y_i}{n}, \ldots, \frac{\sum_{i=1}^{n} z_i}{n}\right)$
- where n is the number of objects that has the closest centroid to cluster K_m.
- Repeat Step 3 and 4 until the centroids no longer move. This produces a separation of the objects into k groups.

The k-Means clustering with $k = 4$ was tested on the WT data A & C and their results are shown in Figure 3.11 & 3.12. Both results have failed to identify the abnormal data, therefore k-Means clustering is unlikely to give good WT fault detection.

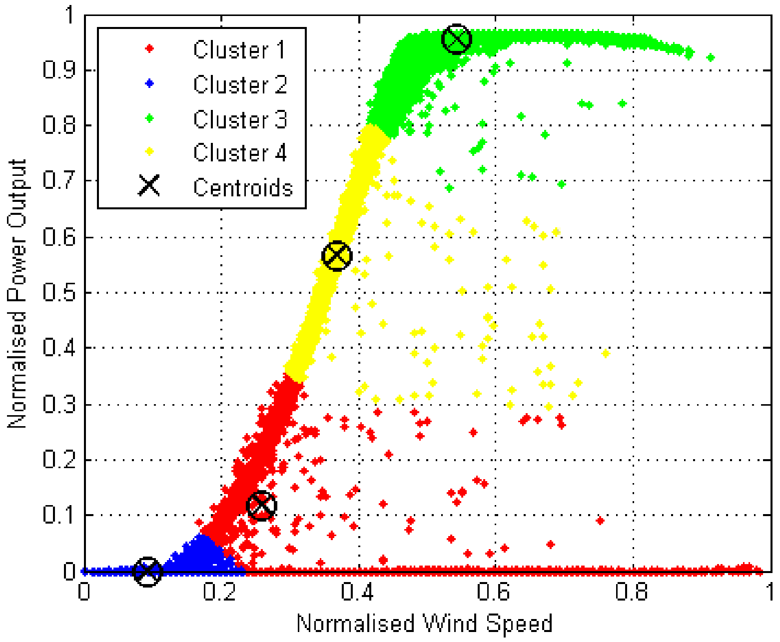


Figure 3.11: k-Means clustering result with k=4.

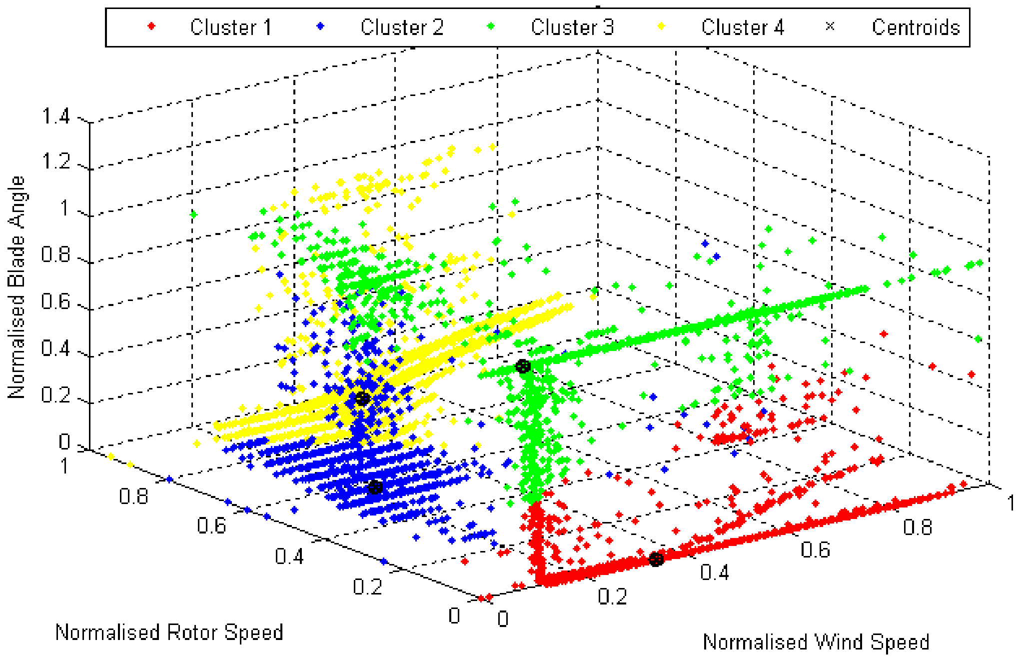


Figure 3.12: Data C, k-Means clustering result with k=4.

According to Stuart and Peter (2010) and Witten et al. (2011) , the weaknesses of k-Means cluster are:

- The number of K clusters must be determined at the start of the analysis;
- It is sensitive to initial position of cluster centroid, different initial position may produce different cluster results;
- It is difficult to know which feature contributes more to the clustering process, general scaling can only be applied if we assume each feature has the same weight.

3.2.10. Fuzzy c-Means Clustering

Fuzzy c-Means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. This method was developed by Dunn (1973) and improved by Bezdek (1981), it was frequently used in pattern recognition. It is based on minimisation of the following objective function:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m}\left\|x_i - c_j\right\|^2 , \quad 1 \leq m < \infty \qquad (3.15)$$

where m is any real number greater than 1, u_{ij} is the degree of membership of x_i in the cluster j, x_i is the ith of d-dimensional measured data, c_j is the d-dimension centre of the cluster, and $\|*\|$ is any norm expressing the similarity between any measured data and the centre.

Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership u_{ij} and the cluster centres c_j by:

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\left\|x_i - c_j\right\|}{\left\|x_i - c_k\right\|}\right)^{\frac{2}{m-1}}} \qquad (3.16)$$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m * x_i}{\sum_{i=1}^{N} u_{ij}^m} \tag{3.17}$$

where $\|x_i - c_j\|$ is the distance from point i to current cluster j, $\|x_i - c_k\|$ is the distance from point i to the other clusters k. This iteration will stop when $\max\left\{\left|u_{ij}^{(k+1)} - u_{ij}^{(k)}\right|\right\} < \varepsilon$, where ε is a termination criterion between 0 and 1, whereas k is the iteration step.

The advantages of the FCM are:

- It gives best result for overlapped data sets, comparatively better than k-means algorithm;
- Unlike k-means, where a data point must belong exclusively to one cluster centre, an FCM data point can be assigned membership to each cluster centre.

The disadvantages are:

- The number of clusters have to be classified at the start of the analysis;
- With lower value of ε we get the better results, but it takes more number of iteration;
- Euclidean distance measures can unequally weight underlying factors.

The Fuzzy c-means with using 4 clusters and exponent 2 was tested on the WT power data A and the result is shown in Figure 3.13. The results have failed to identify the abnormal data, therefore the FCM clustering is unlikely to give good WT fault detection.

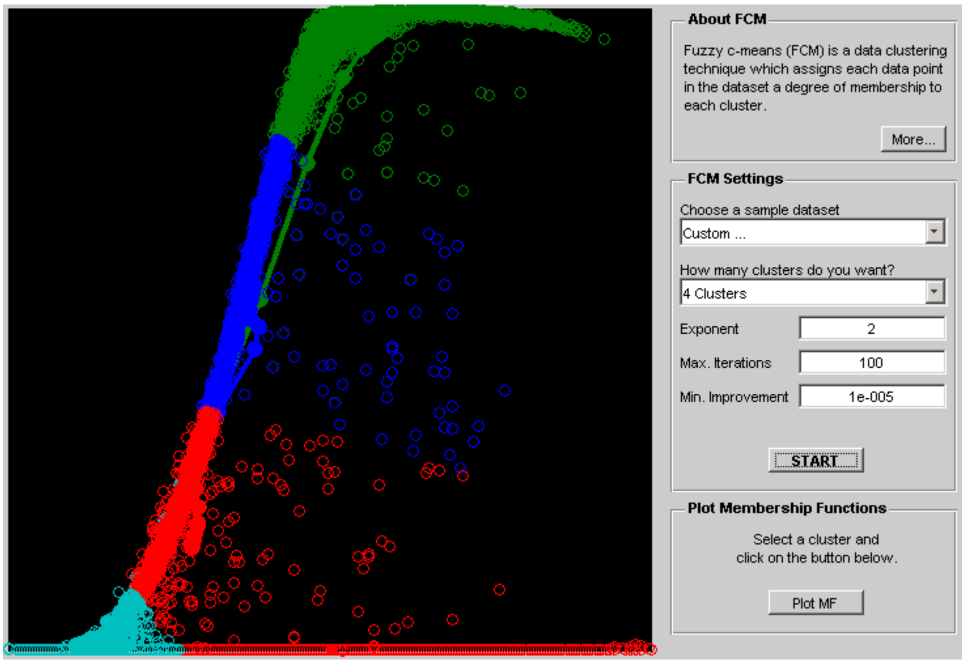


Figure 3.13: Fuzzy c-means result

3.2.11. k-Nearest Neighbours

k-Nearest Neighbours (k-NN) is supervised learning that has been used in many applications in the field of data mining and pattern recognition (Witten et al. 2011). It classifies objects based on closest training examples in the feature space. k-NN is a type of instance-based learning where the function is only approximated locally and all computation is deferred until classification. An object is classified by a majority vote of the closest k neighbours or the distance-weighted average of the closest k neighbours if the class is numeric. If k=1, then the object is simply assigned to the class or the value of that single nearest neighbour.

In general, the k-NN algorithm is composed of the following steps:

- Determine parameter k, which is the number of nearest neighbours;

- Calculate the distance between the query-object and all the training samples;
- Sort the distance and determine the nearest k neighbours based on minimum distance;
- Gather the category Y of the k nearest neighbours;
- Use simple majority of the category of nearest neighbours (or the distance-weighted average if the class is numeric) as the prediction value of the query-instance.

According to (Witten et al. 2011), the advantages of k-NN are:

- Robust to noisy training data, especially if the inverse square of weighted distance is used as the distance metric;
- Effective in training procedure compare to the other algorithms;

However, the disadvantages are:

- Need to determine the value of k, which is the number of nearest neighbours;
- k-NN is a type of distance based learning and which type of distance can produce the best result is not clear;
- Computation cost is high because the algorithm needs to calculate the distance of each query-object to all training samples.

The k-NN was trained using the WT power data A and the classification was tested using a test data with k=10 as shown in Figure 3.14. We found that this test data is misidentified as OK because majority of the 10 nearest data to the test data are OK.

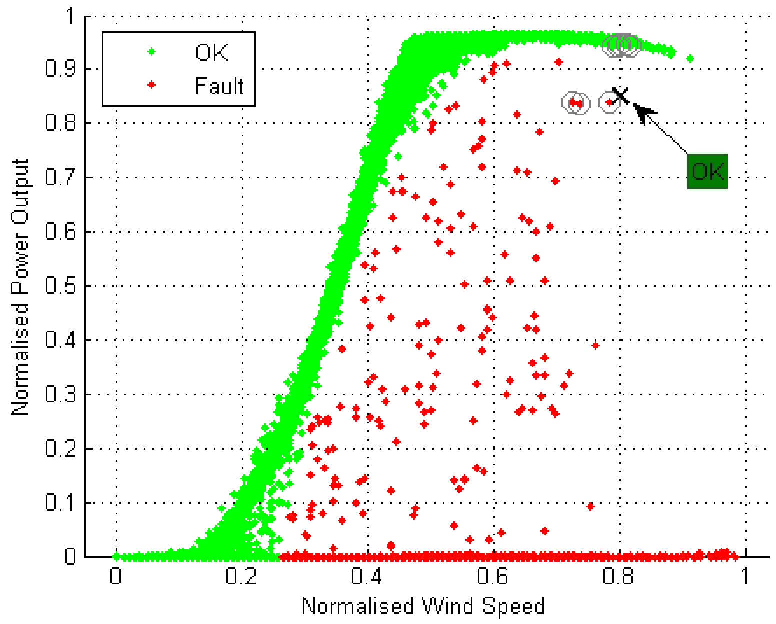


Figure 3.14: k-NN result with a data A

Due to the uncertainty of k and the high computation cost in every classification instance, this algorithm is unlikely to meet the requirement of the on-line FDD.

3.2.12. Support Vector Machine

The Support Vector Machine (SVM) is a supervised learning algorithm that classifies both linear and non-linear data based on maximising margin between support points and a non-linear mapping to transform the original training data into a higher dimension. It was originally developed by Vapnik (Wu et al. 2008) based on the groundwork of statistical learning theory. Similar to ANN, SVM has successfully demonstrated the capability to perform classification, pattern recognition and prediction (Witten et al. 2011). The main difference between SVM and ANN is the method used for minimising the error during training procedure. Unlike ANN, which uses empirical risk minimisation, the SVM makes use of structural risk minimisation, which is alleged to provide better generalisation abilities (Bennett and Campbell 2000).

In a two-class learning task, the aim of SVM is to find the best classification function to distinguish the members of two classes in training data. The metric of SVM can be realised geometrically. For a linearly separable dataset, a linear classification function is a boundary that passes through the middle of the classes separating them into two. However, there are many possible boundaries, as shown in Figure 3.15(a). The optimal boundary is the classification function with the maximum margin to several data points from both classes in perpendicular way, as shown in Figure 3.15(b). The data points that are touched by the margin are special because they determine the margin and the classification function. Therefore, a special name, is given to them, Support Vector.

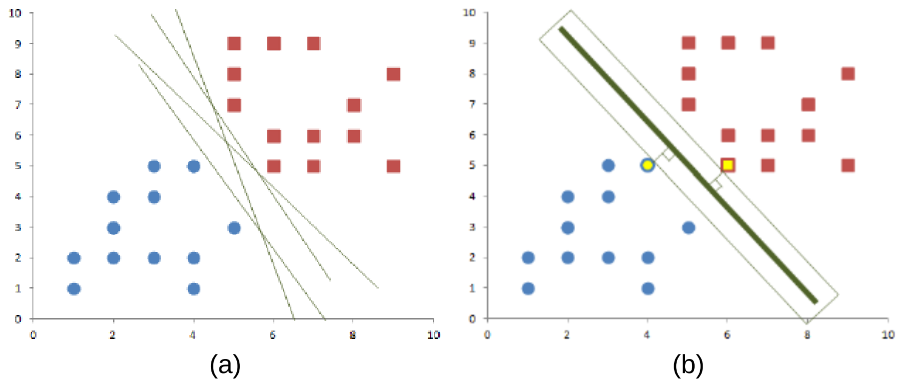


Figure 3.15: Linearly separable dataset

There are a number of mathematical methods that can be used to determine the optimal classification function. One method is to find these support vectors by solving a quadratic problem as shown in (Bennett and Campbell 2000), another method is to maximise the margin between two parallel supporting planes also discussed in (Bennett and Campbell 2000). Both methods result in the same solution; however quadratic programming method is currently better founded.

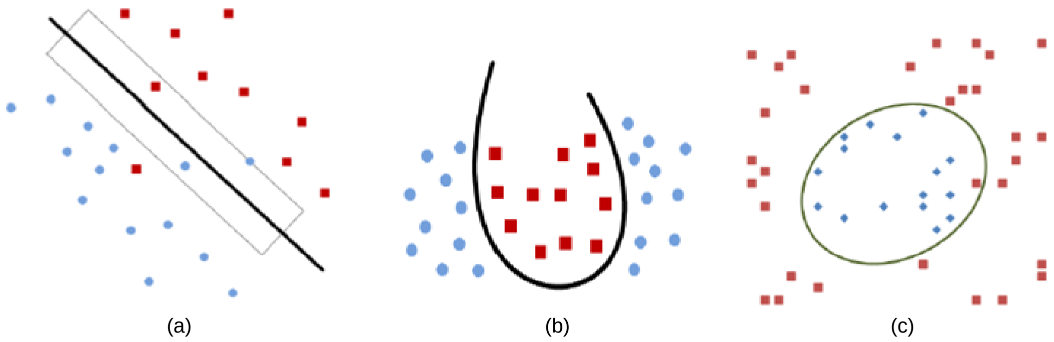


Figure 3.16: Linear non-separable dataset

For a linearly non-separable dataset, SVM utilises slack variables and a kernel trick. Figure 3.16 shows some examples of a linearly non-separable case. The Figure 3.16(a) is almost linearly separable but there are a few data points that can go wrongly to the other side, this type of dataset could be transformed into a linearly separable case using slack variables (Witten et al. 2011). The Figure 3.16(b)&(c) are clearly linearly non-separable as the decision boundary is non-linear. For these types, a kernel trick is applied to transform the non-linear decision boundary into linear decision boundary in a higher dimension. The Figure 3.16(c) also contains misclassification on non-linear decision boundary. In this case, both slack variables and the kernel trick are needed. A mathematically rigorous explanation of the SVM can be found from Wu et al. (2008) and Witten et al. (2011).

According to Fletcher (2009), an SVM framework is currently the most popular approach for "off-the-shelf" supervised learning which means that the SVM is an excellent method to try first if no specialized prior domain knowledge is available. There are three advantages that make SVMs attractive:

- SVMs construct a maximum margin separator, or decision boundary, with the largest possible support vector margin;

- SVMs have the ability to embed the data into higher-dimensional space, using the kernel trick, data that was not linearly separable in the original input space can easily be separable in higher-dimensional space.
- SVMs are a non-parametric method; they retain training examples and potentially need to store them all. On the other hand, in practice they often end up retaining only a small fraction of the number of examples, sometimes as few as a small constant time the number of dimensions. Thus SVMs combine the advantages of non-parametric and parametric models, they have the flexibility to represent complex functions, but they are resistant to over-fitting.

SVM have been applied successfully to a number of applications, ranging from speech recognition, signal prediction and industrial machinery FDD (Witten et al 2011, Widodo and Yang 2007). An interesting application of SVM was made by Assunção et al. (2006) to estimate the transformer top oil temperature and the result was compared to an ANN model. The comparison showed that the SVM estimation performance was slightly better than the ANN. In addition, a survey published by Widodo and Yang (2007) reviewed the application of SVM to the diagnosis of rolling element bearings, induction motors, diagnosis of machine tools and a number of other industrial condition monitoring based scenarios. The survey result stated that the SVM provides good performance for classification.

Initial research using SVM with the Radial Basis kernel and the scaling factor sigma set to 1 to detect WT fault was applied to data A & C (MathWorks 2013). The results are shown in Figure 3.17 & 3.18, which demonstrate that the SVM almost successfully separated the data into two classes, OK & Faulty. SVM has shown potential to detect WT fault, however, the author found that the SVM might have the difficulty interpreting results in high dimension and the model currently doesn't support domain knowledge incorporation.

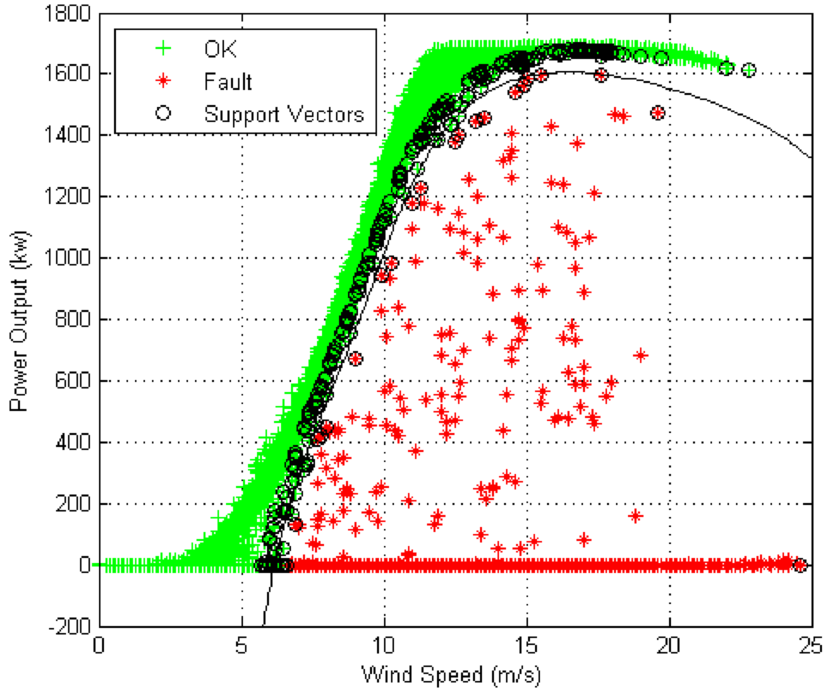


Figure 3.17: SVM result on WT power data A.

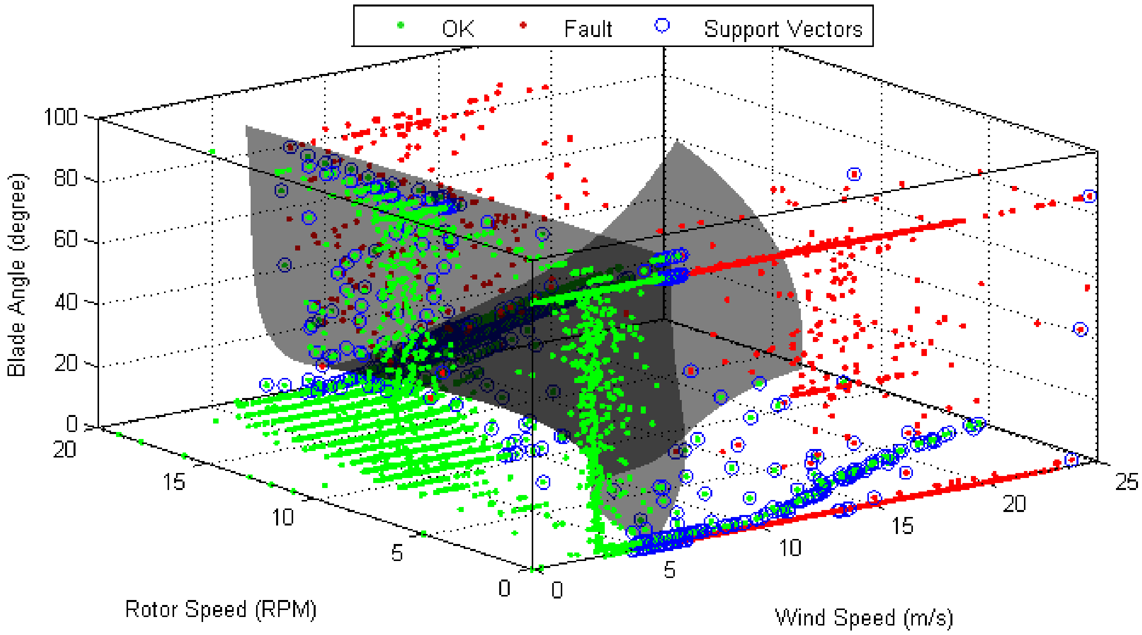


Figure 3.18: SVM result on WT pitch mechanism data C. The decision hyper plane is shown in grey colour.

3.2.13. Adaptive Neuro-Fuzzy Inference System

A neuro-fuzzy system is a fusion of two different systems that has a combination of advantages from ANN, such as robustness, learning & training, and FIS, such as interpretability. Strictly speaking, an ANFIS is a specific kind of the most widely used neuro-fuzzy system and is functionally equivalent to a Takagi-Sugeno FIS (Jang 1993). ANFIS is a multilayered feed-forward network consisting of a number of nodes connected through directional links. ANFIS is adaptive because some or all nodes contain modifiable parameters which can be updated by the learning algorithm. ANFIS is a powerful approach for building complex non-linear relationships between sets of input and output data. An ANFIS system can be trained without the expert knowledge usually required by FIS. Both numerical and linguistic knowledge can be combined into a rule base by employing the fuzzy method. Fuzzy Membership Functions (MFs) can be optimally tuned by using optimisation algorithms. Another advantage of the

ANFIS is its capacity for fast learning and adaptation. Because of these attractive features, ANFIS has been employed directly in a variety of modelling, diagnosis, decision making, signal processing and control applications (Korbicz and Kowal 2007, Mote and Lokhande 2012, Tran et al. 2009, Zio and Gola 2009).

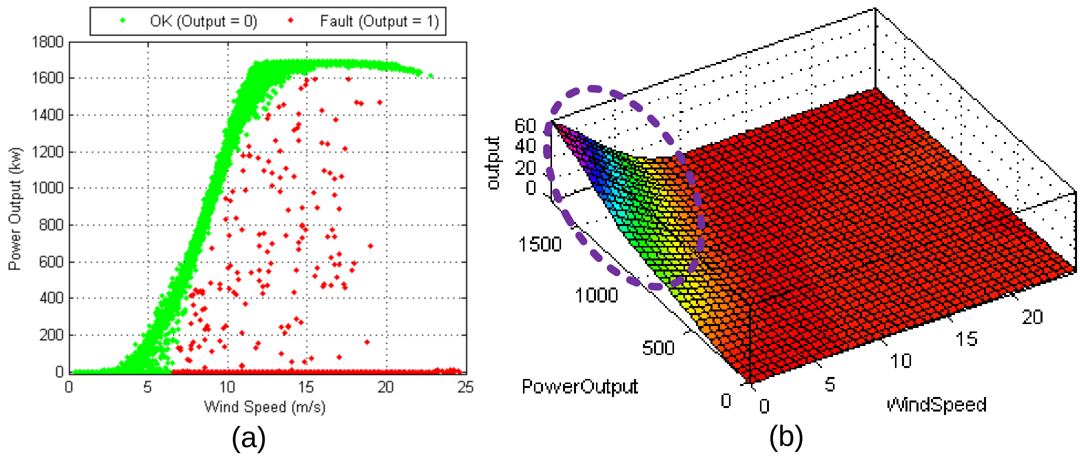


Figure 3.19: (a) WT power data A; (b) ANFIS result, the output z-axis is used to indicate the OK and Fault state of the pitch as defined in training data and shown in Fig 3.19(a);

In order to illustrate the ANFIS, the aforementioned fault detection using WT power curve was studied. The WT power data A, as shown in Figure 3.19(a), with output value 0 and 1 to indicate the OK and Fault state of the pitch was used in this study. The ANFIS was built with 3 MFs in each input, maximum training iteration 50 and minimum error 0.01. Figure 3.19(b) shows the output generated by the trained ANFIS. Clearly, the ANFIS output range is much larger than the defined output value in training data, as encircled in Figure 3.19(b). This is largely because of the insufficient training data in the corresponding areas and result in the trained model behaves erratically in never-seen input conditions.

Recent research has proposed that A-Priori Knowledge (APK) can be incorporated into an ANFIS model, for example a technique called APK-ANFIS from Tewari (2009) allowed domain knowledge to be introduced into the ANFIS model, for example if the WT power output was high at low wind speed, this could be regarded as a possible sensor fault. Therefore, another test using APK-ANFIS with the introduction of two favourable rules, as specified in Figure 3.20(a), using the same number of MFs, maximum training iteration and minimum error was trained on the same data and the result is shown in Figure 3.20 (b).

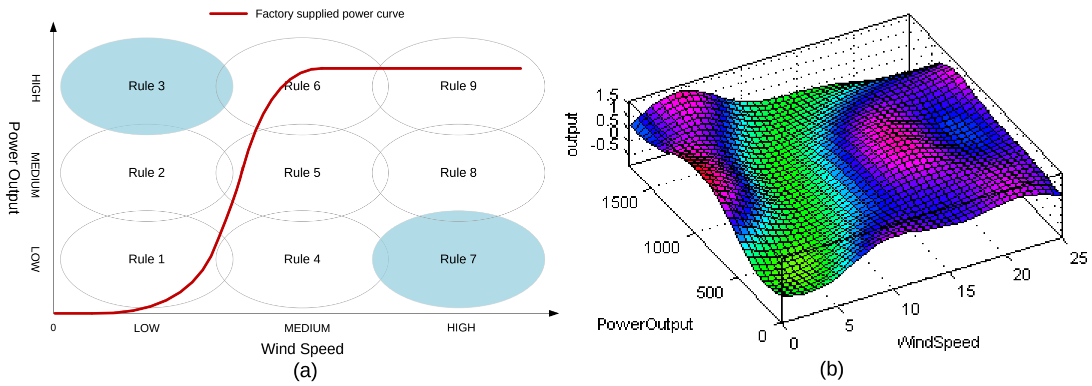


Figure 3.20: APK-ANFIS result (a) 2 favourable rule corresponds to the a-priori domain knowledge. (b) The APK-ANFIS result.

With the incorporation of domain knowledge, the APK-ANFIS showed relatively better interpretability and is able to maintain the consistency of the model even in the regions with few data points.

3.2.14. Findings & Conclusion

The findings of the possible AI approaches are shown in Table 3.2. In the view of these techniques, unsupervised learning algorithms are relatively difficult to be used in this research. The main reason for this is because the nature disadvantages of the algorithms themselves and fault pattern in this study are unlikely to have a strong geometrical difference from normal pattern.

For the supervised learning algorithms, SVM has shown strong potential to do the fault detection, however, with more practices and we found the SVM is unlike to incorporate any domain knowledge which can be used to solve the problem when training dataset is inadequate. But, this problem may be solved through adding manual created data to the training dataset. For ANN, it is not transparent and incapable of explaining a particular decision to the user in comprehensible form. FIS requires fine-tuning to obtain an acceptable rule base and optimal parameters for available data. Although individual ANN and FIS problems can be solved by the integration of both methods, ANFIS, there is also a significant disadvantage of hybrid models, as the learning becomes entirely data driven, it imposes stringent requirements on the quality of training dataset. If the training dataset is inadequate, then the trained model can behave erratically in unseen input conditions and becomes uninterpretable. Fortunately, recent research has proposed a-priori knowledge technique can be incorporated into ANFIS model and this have shown relatively better interpretability and is able to maintain the consistency of the model even in the regions with few data points, as shown in Figure 3.16.

Therefore, this study has decided focus on using APK-ANFIS to analyse WT SCADA data and proves its feasibility for WT fault prognosis, concentrating particularly on WT electric pitch system faults, which are known to be significant.

AI Approach	Advantages	Disadvantages	Findings	Reason not to take further
Naïve Bayes Classifier	• Simple; • Good for high dimensional inputs;	• Independence assumption;	Naïve Bayes classifier is designed for use when feature are independent of one another within each class. However, the two tests showed it can detect WT fault in practice even if the independence assumption is not valid.	Difficult to interpret the result because of the independence assumption.
Rule-based Expert System	• Model is clear and ; • Natural Knowledge representation;	• Expertise needed; • Difficult to build; • Inability to learn itself;	Rule-based expert system employing logic programming has been widely used for knowledge representation. However, substantial amount of expert knowledge is needed and it doesn't have ability to learn by itself.	Lack of diagnostic knowledge available to codify as rules.
Takagi Sugano FIS	• Simple and interpretable; • Low computational cost;	• Difficult to transform human knowledge; • The need for expert to fine-tuning the model;	Takagi Sugano FIS is simple, interpretable and low computational cost. However, no standard methods for transforming human knowledge into a FIS and it needs expert for fine-tuning to find the optimal MF and parameters.	Lack of diagnostic knowledge. In addition, no standard methods for transforming knowledge into a FIS.
Decision Tree	• Visualisation of the relationship; • Non-parametric supervised learning;	• Uncertainty of obtaining optimal decision tree; • Difficult to learn some problems e.g. XOR;	Decision tree is an intuitive graphical tool for displaying relationships that lead to faults. The test demonstrated decision tree has successfully learned the fault pattern. However, some sub-regions with sparse training data cause the decision tree to become inconsistent with domain knowledge.	The decision tree model has the difficulty to deal with near-seen data.
ANN	• Implicitly detect complex nonlinear relationships; • Great fault tolerance; • Adaptive learning;	• "Black-box" nature; • High computational cost; • Empirical nature of model development;	The two tests demonstrated the general mapping capability of the ANN help to identify those most likely faults very well.	Difficult to interpret result because of the "Black-box" nature.
SOM	• High dimensional data visualisation; • Reduce data dimensions;	• Need sufficient data; • Similar groups may appear in different areas;	The two tests showed that the SOM has the capability to identify abnormal data. However, the result needs to be interpreted by expert.	Domain knowledge is required in order to interpret the result.
Bayesian	• Represent the	• Domain knowledge	The research results have shown that the BN is a valuable	Lack of diagnostic

Network	conditional dependences; • Represent the causal relationship;	required; • Difficult to build BN structure; • Difficult to define NPT;	tool for WT fault diagnosis and has great potential to rationalise failure root causes. However, the BN complexity grows exponentially with the increase of parent node numbers.	knowledge to build a robust BN. In addition, its complexity grows exponentially with the increase of node.
k-Means Clustering	• Simple and easy to understand; • Computationally faster than other clustering algorithm, e.g. SOM; • Give best result when data set are distinct;	• The number of k cluster has to be determined; • Sensitive to initial position of cluster centroid; • Difficult to know which feature contribute more to the clustering process.	Two tests results have shown that the k-Means clustering has failed to identify the abnormal data.	Uncertainty of the cluster number and test has failed to identify the abnormal data.
Fuzzy c-means	• Give best result for overlapped datasets; • Data point can be assigned membership to each cluster;	• The number of cluster has to be determined; • Lower value of ε would get better result, but it takes more times;	A test has shown that the FCM clustering has failed to identify the abnormal data.	Uncertainty of the cluster number and test has failed to identify the abnormal data.
k-Nearest Neighbours	• Robust to noisy data; • Effective in training procedure;	• k has to be determined; • High computation cost; • Distance based learning and which	Due to the uncertainty of k and the high computation cost in every classification instance, this algorithm is unlikely to meet the requirement of the on-line FDD.	A misclassification problem was shown in the example and the technique has high computation cost in every classification instance.

		type of distance has best result is not clear;		
SVM	• Non-parametric approach; • Structural risk minimisation approach; • Effective in high dimensional spaces;	• Give poor performance if the number of features is much greater than the number of samples; • Do not provide probability estimates;	Initial research has demonstrated that the SVM almost successfully separated the normal and abnormal data.	It might have the difficulty to interpret the result in high dimension. In addition, SVM is unlike to incorporate any domain knowledge.
ANFIS	• One of the best trade-off between ANN and FIS; • Fast learning and adaptation;	• Only allow one output; • Strong computational complexity;	ANFIS has a combination of advantages from ANN and FIS. However, if the training dataset is inadequate then the trained model can behave erratically in unseen input conditions and becomes un-interpretable.	Trained model can behave erratically in unseen input conditions and becomes un-interpretable.
APK-ANFIS	• Inherited the advantages of ANFIS; • A-prior domain knowledge incorporation;	• High computational cost; • Domain knowledge is required to build the model;	It allows a-priori domain knowledge to be incorporated into the ANFIS training procedure and restrict the output of sub-regions. The built model showed relatively better interpretability.	N/A

Table 3.5: Findings of the possible AI approaches.

3.3. Available SCADA data from WT Pitch Systems

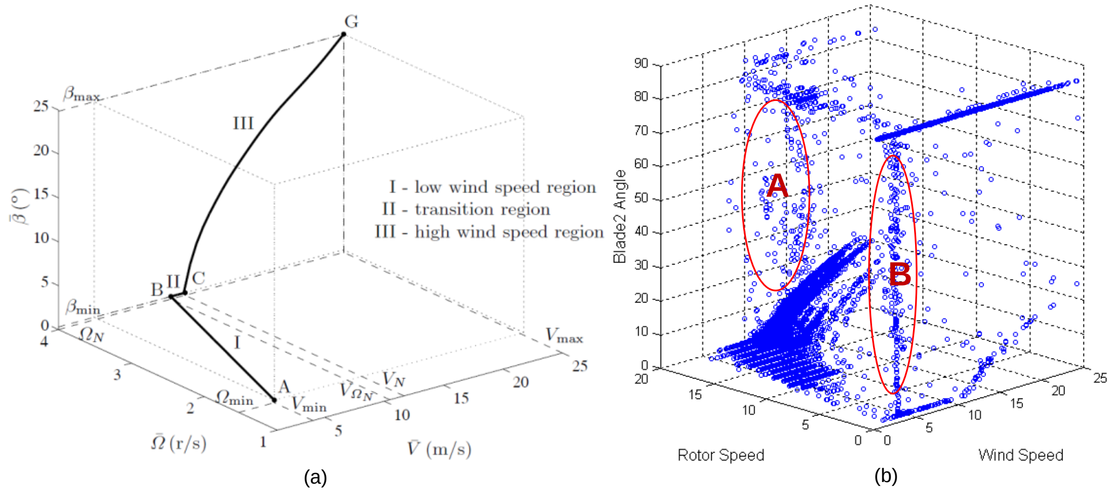


Figure 3.21: (a) Variable-speed pitch-to-feather control strategy; (b) Real Alstom WT data plot;

Figure 3.21(a) shows a variable-speed pitch-to-feather control strategy (Bianchi et al. 2006) plotted on the pitch angle, rotor speed and wind speed space.

- Region I shows the rotor speed is increased in proportion to the wind speed from its minimum Ω_{min} at cut-in (Point A) to its rated value Ω_N at point B.
- Region II is the transition region, where wind speed is between $V_{\Omega N}$ and V_N, the pitch angle remains constant at β_{min}.
- Region II for wind speeds larger than V_N the pitch angle is increased to β_{max} to avoid the rotor over speed and limit the power output to its rated value.

Figure 3.21(b) displays the actual SCADA plots for comparison of the characteristics seen in Figure 3.21(a). In general, there are clear similarities to Figure 3.21(a), with additional transition regions A and B for blade adopt feather. Regions A & B are noisy because the blade movement is slow and it can occupy a number of angular positions in those conditions.

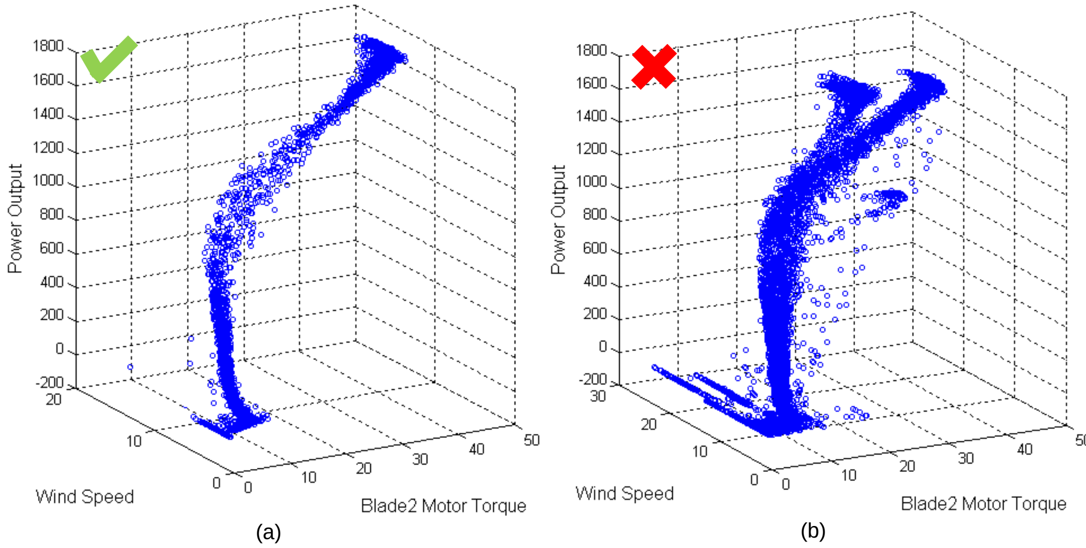


Figure 3.22: (a) Pitch torque power curve characteristic; (b) An abnormal WT;

Figure 3.22(a) shows WT pitch torque power curve characteristic and these figures are less noisy than Figure 3.21. According to Moorse (2010), a normal WT should have the above characteristic trends during different periods of operation. Faults can be identified when a number of observations deviate from the characteristic curve, as shown in Figure 3.22(b).

Initially, an inspection of six known pitch faults (Cases 1-6, as shown in Table 3.6), using both typical variable-speed pitch-to-feather control strategy and pitch torque power curve characteristics, has been made to find the common pitch fault symptom, as shown in Figure 3.23.

WT	Case	Developing Fault	Maintenance	After Maintenance
A	Case 1	05/01/2008 ~ 15/02/2008	16/02/2008 ~ 21/02/2008	22/02/2008 ~ 03/03/2008
	Case 2	20/12/2006 ~ 14/01/2007	15/01/2007 ~ 25/01/2007	26/02/2007 ~ 10/02/2007
B	Case 3	22/08/2007 ~ 04/09/2007	05/09/2007 ~ 09/09/2007	10/09/2007 ~ 18/09/2007
	Case 4	17/10/2006 ~ 28/10/2006	29/10/2006 ~ 29/10/2006	30/10/2006 ~ 04/11/2006
	Case 5	10/08/2008 ~ 27/08/2008	28/08/2008 ~ 30/08/2008	31/08/2008 ~ 10/09/2008
	Case 6	20/09/2006 ~ 13/10/2006	14/10/2006 ~ 19/10/2006	19/10/2006 ~ 22/10/2006

Table 3.6: Six pitch fault cases from the same WF. The three periods were used to inspect the WT's status at different stages.

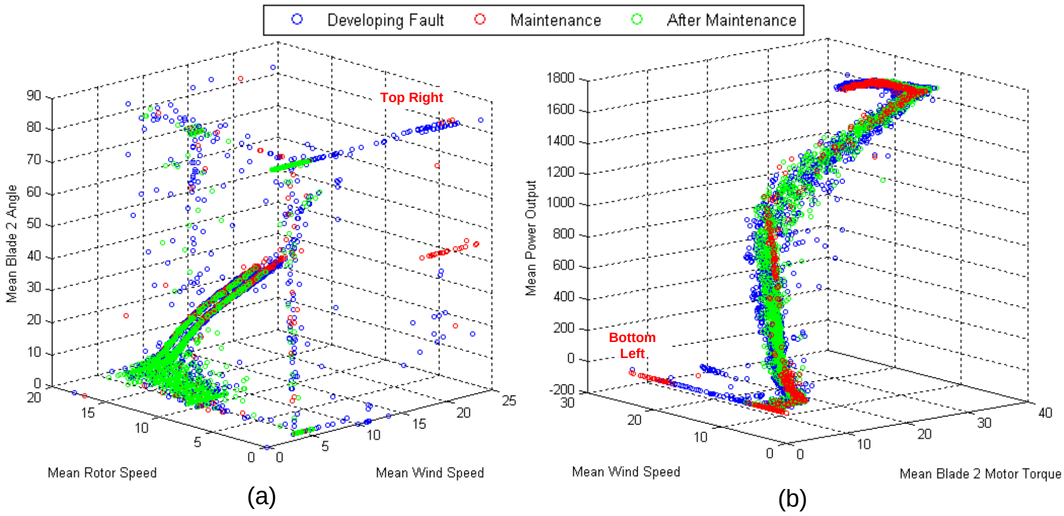


Figure 3.23: (a) Typical variable-speed pitch-to-feather control plot for Case 1; (b) Pitch torque power curve plot for Case 1;

In Figure 3.23 (a), no *After Maintenance* data can be found on the top right corner, representing high wind speeds, high blade angle and low rotor speed. A normal running turbine should not have feathered blades and zero rotor speed when the wind speed is greater than cut-in. Thus, any data appearing on top right corner of this 3D plot can be regarded as a possible pitch fault.

For Figure 3.23 (b), no *After Maintenance* data can be found on bottom left corner, representing high wind speeds, low motor torque and low power output. This is because a normal running turbine should start generating power when the wind speed is greater than cut-in. Meanwhile, blade pitch motor torque is needed to change the blade angle to prevent rotor over-speed. Thus, any data appearing in the bottom left of this 3D plot could be caused by a pitch fault.

These graphs were presented in 3D but analysis in one plane would simplify the algorithm to two variables, 2D views are shown in Figure 3.24. By comparing and analysing the difference between *Developing Fault* and *After Maintenance* periods, four 2D views, numbered 1, 2, 4 & 5, were found and they clearly showing abnormal SCADA data in the *Developing Fault* period, as circled in Figure 3.24. Therefore, these four 2D views, known as Critical Characteristic Features (CCFs), can be used to identify WT pitch faults.

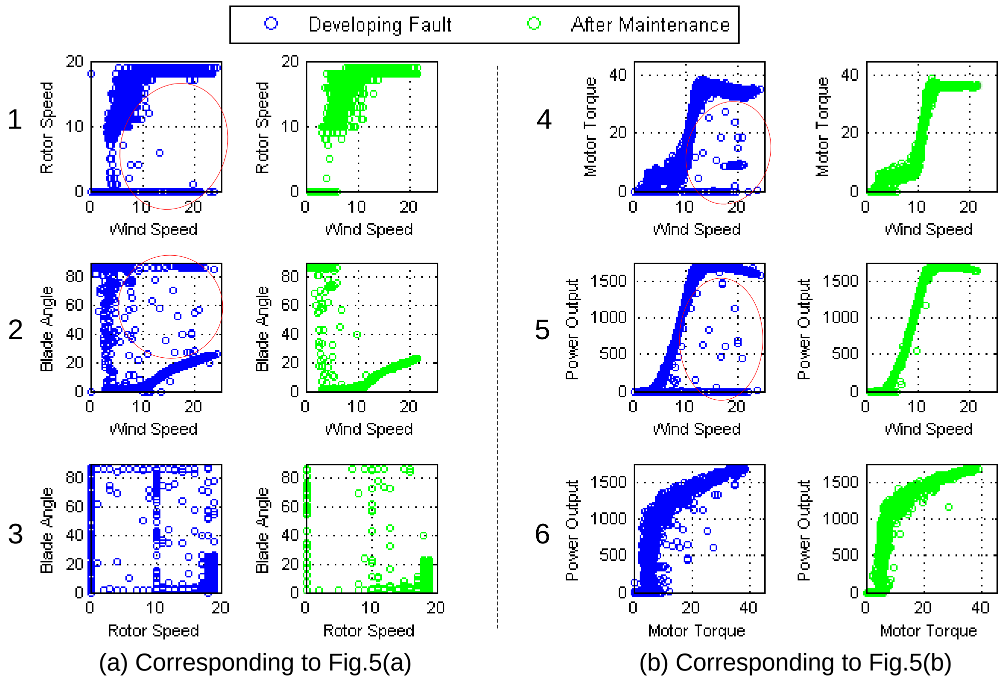


Figure 3.24: 2D views of Fig. 3.18 covering Developing Fault and After Maintenance. The subfigures 1-6 are the corresponding 2D plot.

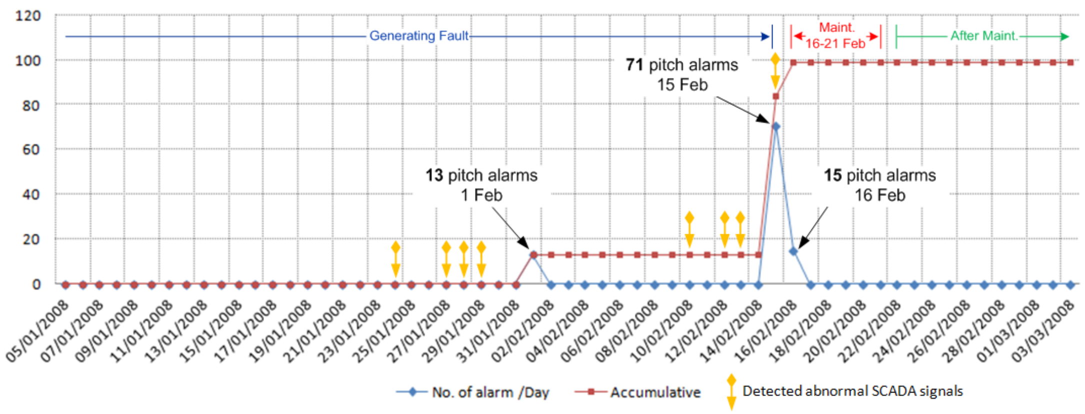


Figure 3.25: Day-by-day analysis

In addition, a day-by-day analysis through plotting the aforementioned four CCFs in the *Developing Fault* period and checking against the corresponding

SCADA pitch alarms had shown that the SCADA signals is able to provide fault detection and it is much earlier than SCADA alarms, as shown in Figure 3.25.

Considering WT O&M economic factors and based on the findings received here, this research is extended to develop an automated fault prognosis for WT pitch system. A diagnosis procedure is proposed in the next Chapter by applying an APK-ANFIS. The six known pitch faults will be used as the knowledge base for the proposed diagnosis procedure.

3.4. Anomaly Detection

After assessing the possible AI techniques and considering the available SCADA data, it quickly becomes evident that anomaly detection is the most suitable method for fault detection. This is due to the absence of domain knowledge about how incipient component faults are represented in the SCADA data. In addition, the lack of WF operator support and resource, as well as the time-consuming and complex nature of analysing large SCADA data volumes, to understand every single pitch fault symptom. Moreover, limited access to data with fault records also prohibits the development of WT fault detection system. These limitations ultimately lead towards the need to utilise anomaly detection to achieve fault detection through abnormal behaviour observation.

Anomaly detection refers to the problem of discovering patterns in data that do not conform to expected behaviour (Chandola et al. 2009). Usually, the anomalies are infrequent, but their importance is significant and critical in most application domains; this makes their detection extremely important. In the field of WT FDD using SCADA data, an anomaly detection model would receive data from various sensors mounted on the WT components and the objective is to capture how the data evolves and changes with respect to factors that may influence it under normal circumstances. This would therefore allow for the detection of the abnormal behaviour, which does not subject to the physical properties of a normal running WT. Usually, a model is built to provide an estimation of a senor output based on the relevant signals from the other sensors. The estimation can then be compared to the real value recorded by the senor, where a significant deviation from the estimated value would be viewed as an abnormality. In this way, incipient faults can be highlighted and presented to the operator, dramatically reducing the complexity of their task since only significant information of relevance to the WT health is presented to them.

As the knowledge of every individual fault is not included in the model, the developed anomaly detection cannot be used to exactly identify fault in a specific component. However, the aim of this research is to provide the initial stage of the fault identification process. Early detection of failures and problems would allow operators to schedule maintenance schemes appropriately; and an experienced operator could take this opportunity to have a look the problems in detail, for example by turning on the high frequency CMS to diagnosis the component fault.

3.5. Chapter Summary

This chapter has investigated a number of AI techniques, commonly used in the field of data classification, which can be used for WT FDD. SVM and APK-ANFIS are found to be potential useful in this research. For APK-ANIFS, a-priori knowledge technique can be incorporated into ANFIS model and this have shown relatively better interpretability and is able to maintain the consistency of the model even in the regions with few data points. Therefore, this study has decided to focus on using APK-ANFIS to analyse WT SCADA data and proves its feasibility for WT fault prognosis, concentrating particularly on WT pitch system faults, which are known to be significant.

In addition, an inspection of six known pitch faults, using the variable-speed pitch-to-feather control and pitch torque power curves, have been used to identify common pitch faults. The six known pitch faults will be used as the knowledge base for the proposed diagnosis procedure in next Chapter.

Finally, for this particular application, the anomaly detection approach was found to be the necessary approach due to the lack of domain knowledge about how incipient component faults are represented in the SCADA data and complex nature of analysing large volumes of SCADA data.

Chapter 4.
Proposed On-line Fault Diagnosis Procedure

As has been noted in Chapter 3, this study has decided to use APK-ANFIS to analyse WT SCADA data and proves its feasibility for WT fault prognosis, using SCADA data. The study concentrates particularly on WT pitch faults as it is known to be significant. This chapter aims to introduce the APK-ANFIS in detail and propose an on-line fault diagnosis system with the possibility for fault prognosis. With the a-priori knowledge incorporation, the proposed system is expected to have improved ability to interpret previously unseen conditions and thus fault diagnoses should be improved over the conventional ANFIS. The data of the six known WT pitch faults, introduced in Section 3.3, are labelled and used to train the APK-ANFIS system with a-priori knowledge incorporated.

This chapter is organised as follows: it begins with the introduction of the ANFIS and the a-priori knowledge incorporation. Based on the findings from Section 3.3, the four critical characteristic features (CCFs) for analysing WT pitch fault are discussed in detail. After that, a fault diagnosis procedure using the APK-ANFIS and the four CCFs is proposed. The data of the six known WT pitch faults are labelled according to expert knowledge and the WT's physical properties. The proposed system is trained with a-priori knowledge incorporated. Finally, a trained system is obtained and a demonstration of the diagnosis system is shown.

4.1. Development of ANFIS & A-priori Knowledge Incorporation

ANN research started in the 1940s (Haykin et al. 2009) and has been proven as a reliable technique in the field of pattern recognition, estimation and prediction, as well as fault detection. However, ANNs are not transparent and are incapable of explaining a particular decision to the user in comprehensible form. FIS research started in the 1960s (Jang et al. 1997) and has been widely used for control engineering and knowledge representation. FIS has the ability to model human knowledge in a form of an if-then rule. It also has the capability of transforming linguistic and heuristic terms into crisp numerical values for use in complex machine computation, via fuzzy rules and Membership Functions (MF). The if-then rules and the initial parameters of MFs are normally prepared by an expert. Thus, FIS requires fine-tuning to obtain an acceptable rule base and optimal parameters for available data. In the 1990s, researchers came up with an idea to combine ANN and FIS to achieve learning and readability at the same time. This idea was first proposed by Kosko and Isaka (1993) and named neuro-fuzzy. Compared to ANN and FIS, neuro-fuzzy research is relatively new, but it has gained fast development, especially with the introduction of the ANFIS by Jang et al. (1997).

In general, a neuro-fuzzy system is a fusion or hybrid of two different systems that combine the advantages of ANN, robustness, learning and training, and with the advantage of FIS, interpretability. An ANFIS is a specific kind of neuro-fuzzy system most widely used and functionally equivalent to Takagi-Sugeno FIS (Takagi and Sugeno 1985). ANFIS is a multi-layered feed forward network consisting of a number of nodes connected through directional links.

Adaptive in the title of ANFIS signifies that some nodes contain modifiable parameters which can be updated by the learning algorithm.

ANFIS is a powerful approach for building complex non-linear relationships between sets of input and output data. An ANFIS system can be trained without the expert knowledge usually required by FIS. Both numerical and linguistic knowledge can be combined into a rule base by employing the fuzzy method. Fuzzy MFs can be optimally tuned by using optimisation algorithms. According to Jang et al. (1997), another advantage of the ANFIS is its capacity for fast learning and adaptation. Because of these attractive features, ANFIS has been employed directly in a variety of modelling, diagnosis, decision making, signal processing and control applications.

4.1.1. ANFIS Architecture

A typical ANFIS architecture is functionally equivalent to a first-order Takagi-Sugeno FIS (Jang 1993), as shown in Figure 4.1, where two inputs x and y and one output f are assumed. The common rule set for two fuzzy if-then rules can be expressed as follows:

- **Rule 1:** if x is $\mathrm{A_1}$ and y is $\mathrm{B_1}$, then $\mathrm{f_1 = p_1x + q_1y + r_1}$;
- **Rule 2:** if x is $\mathrm{A_2}$ and y is $\mathrm{B_2}$, then $\mathrm{f_2 = p_2x + q_2y + r_2}$;

where A and B represent the linguistic variables of the MF, p, q and r are the parameters of the consequent first-order polynomial function to be determined during the training stage.

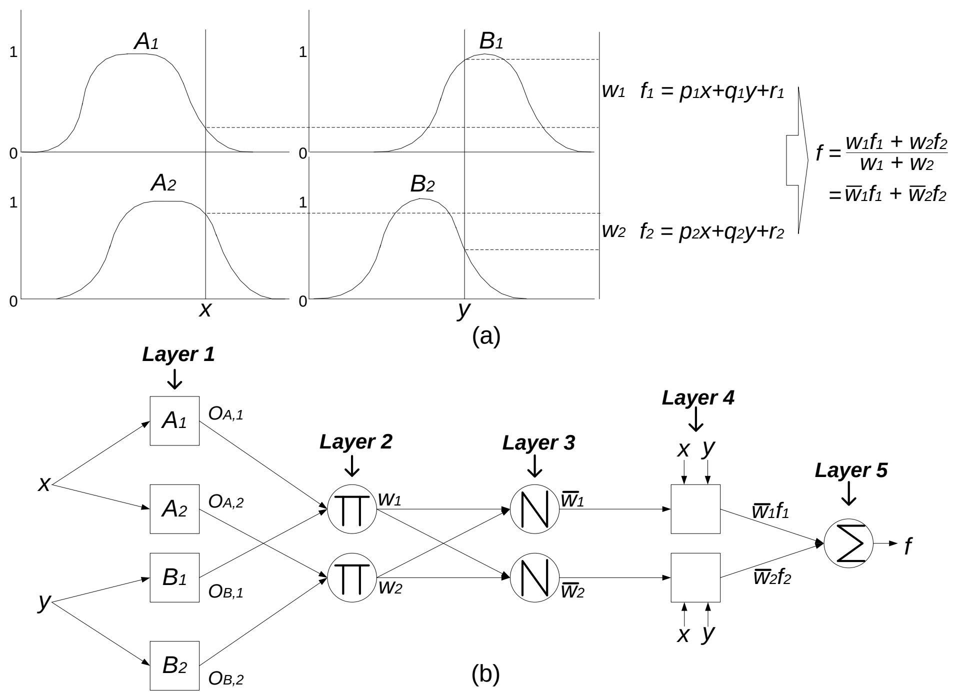


Figure 4.1: (a) A two inputs first order Takagi-Sugeno FIS model with two rules; (b) The equivalent ANFIS architecture.

Figure 4.1(a) illustrates the reasoning mechanism for this Takagi-Sugeno FIS model; the corresponding equivalent ANFIS architecture is shown in Figure 4.1(b), where only Layer 1 & 4 are adaptive and nodes of the same layer have similar functions, as described below:

Layer 1: Fuzzy Layer: all crisp input variables are assigned equivalent linguistic fuzzy labels based on the MFs in this layer. Every node in this layer is adaptive with an output $O_{A,i} = \mu_{A_i}(x)$ or $O_{B,i} = \mu_{B_i}(y)$ which represents the membership grade of the input x or y to the fuzzy MF (A_1 , A_2 , B_1 or B_2). A_i and B_i are linguistic fuzzy labels, such as "small" or "large", associated with the node. MF can be any appropriate parameterised MF introduced in (Jang et al. 1997), such as the generalised bell function:

$$\mu(x) = \frac{1}{1 + \left|\frac{x - c_i}{a_i}\right|^{2b_i}} \qquad (4.1)$$

where $\{a_i\ b_i\ c_i\}$ is the *premise parameter set* and it will be updated during the training stage. As the values of these parameters update, the bell-shaped function varies accordingly and finally adapts to the given training data.

Layer 2: Product Layer: this layer is used to combine the incoming signals and every node in the layer is fixed. They are labelled $\prod$ to indicate that they multiply the incoming signals to produce the *firing strength* of a rule, defined as:

$$w_i = \mu_{A_i}(x) * \mu_{B_i}(y), \qquad i = 1, 2 \qquad (4.2)$$

Layer 3: Normalization Layer: this layer will generate the corresponding ratio of the firing strength and every node in this layer is also fixed. They are labelled N to indicate the calculation of the ratio of the i^{th} rule's firing strength to the sum of all rules' firing strengths, defined as:

$$\bar{w}_i = \frac{w_i}{w_1 + w_2}, \qquad i = 1,2 \qquad (4.3)$$

For convenience, outputs of this layer are called *normalised firing strengths*.

Layer 4: Defuzzify Layer; every rule in this layer will obtain a crisp output and all nodes in this layer are adaptive. The output of each node in this layer is simply the product of the *normalised firing strength* and a *first-order polynomial function*. Parameters $\{p_i\ q_i\ r_i\}$ are referred to as the *consequent parameter set*.

$$\bar{w}_i f_i = \bar{w}_i(p_i x + q_i y + r_i) \qquad (4.4)$$

A special case in this model is that the consequents are expressed by constant values r_i and this model is usually called 0^{th} order ANFIS.

Layer 5: Output Layer; this layer has only one node labelled $\sum$ to indicate that it computes the overall output as the summation of all incoming signals:

$$Output = f = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} = \frac{w_1 f_1 + w_2 f_2}{w_1 + w_2}\,, \quad i = 1,2 \qquad (4.5)$$

Thus, an adaptive network that is functionally and structurally equivalent to a Takagi-Sugeno FIS has been constructed, as shown in Figure 4.1.

4.1.2. Learning Algorithm

As mentioned in Section 4.1.1, there are two adaptive layers, 1 & 4, in ANFIS architecture. Each node in layer 1 has three parameters $\{a_i\ b_i\ c_i\}$ representing the premise parameters. Every node in layer 4 also has three tuneable parameters $\{p_i\ q_i\ r_i\}$ pertaining to the first order polynomial of the consequent part of the rules. The objective of the learning algorithm is to optimise all these parameters to make the ANFIS output best match the training data.

From the ANFIS architecture shown in Figure 4.1(b) and Eq. 4.5, we found that the ANFIS output can be expressed as a linear combination of the consequent parameters when the values of the premise parameters are fixed. In symbols, the output f can be rewritten as:

$$\begin{aligned} f &= \overline{w}_1 f_1 + \overline{w}_2 f_2 \\ &= \overline{w}_1(p_1 x + q_1 y + r_1) + \overline{w}_2(p_2 x + q_2 y + r_2) \\ &= (\overline{w}_1 x)p_1 + (\overline{w}_1 y)q_1 + (\overline{w}_1)r_1 + (\overline{w}_2 x)p_2 + (\overline{w}_2 y)q_2 + (\overline{w}_2)r_2 \qquad (4.6) \end{aligned}$$

which is linear in the consequent parameters $p_1, q_1, r_1, p_2, q_2, r_2$. From this observation, it can be found that the premise parameters are non-linear and consequent parameters are linear. Therefore, a hybrid algorithm combining the least squares method and the gradient descent method is proposed. The hybrid algorithm consists of a forward pass and a backward pass, as listed in Table 4.1.

In the forward pass, the least square method is used to optimise the consequent parameters with the fixed premise parameters. Once the optimal consequent parameters are found, the backward pass commences immediately. In the backward pass, the gradient descent method is used to adjust the premise parameters corresponding to the fuzzy set in the input domain, whilst the consequent parameters remain fixed. This procedure is repeated until the overall squared error between desired output and actual output is less than a specified value or the learning has reached the maximum iteration.

	Forward Pass	**Backward Pass**
Premise parameters	Fixed	Gradient descent
Consequent parameters	Least-square method	Fixed
Signals	Node outputs	Error signals

Table 4.1: Hybrid algorithm for ANFIS training

4.1.3. Down-side of ANFIS

ANFIS has been researched for many different fault detection and diagnosis applications. Some recent applications include induction motor fault diagnosis (Tran et al. 2009), bearing fault diagnosis (Zio and Gola 2009), rotating machinery fault diagnosis (Lei et al. 2007) and dynamic system fault detection (Korbicz and Kowal 2007). Although the integration of ANN with FIS has proved useful, there is also a noticeable down-side of this hybrid model. As the learning becomes entirely data driven, it imposes stringent requirements on the quality of training dataset. If the training dataset is inadequate then the trained model can behave erratically in unseen input conditions and becomes un-interpretable. For example a WT power curve input subspace has sparse data distribution and could result in the corresponding ANFIS region's output become inconsistent with domain knowledge. In other words, the model loses its consistency with domain knowledge and behaves like a black-box. This down-side has led to the development of various parameter estimation techniques aiming to improve the interpretability of this hybrid model. Yen et al. (1998) proposed a combination of global and local identification of consequent parameters to certain constraints. In another approach, Bikdash (1999) proposed the use of spline-based fuzzy FMs and rule-centred linear models that can be treated as 1st order Taylor series expansion about the rule centres. This method improved both the interpretability and the estimation capability of fuzzy rules. Abonyi et al. (1999) proposed a technique which the a-priori knowledge is incorporated in Takagi-Sugeno FIS models by introducing a set of parameters constraints and penalty functions.

Quite often, a domain expert is capable of identifying a few sub-regions in the input space where the corresponding output is expected to be more or less than the others. Tewari (2009) proposed a technique that allows a-priori domain knowledge to be incorporated into the ANFIS training procedure and restrict the output of sub-regions. The core idea is to incorporate the expert's qualitative domain knowledge into the parameter estimation step of ANFIS training procedure. This proposed technique enables the ANFIS to remain consistent with domain knowledge even when the certain input sub-regions are sparse or training data has noise. With these advantages, this research has decided to use the Tewari's technique to analyse WT SCADA data. This new technique would allow the expert's domain knowledge such as "If WT power output is high at low wind speed, this can be regarded as a possible fault" to be incorporated.

4.1.4. A-priori Knowledge Incorporation

The latest development of the ANFIS has enabled the incorporation of domain knowledge into the ANFIS training procedure. Assuming that we are working on fault detection using the WT power curve, 3 MFs will be associated with each input signal. The 3 MFs are used to represent the linguistic labels *Low*, *Medium* and *High* for both power output and wind speed. Thus, the total number of rules R is $3 \times 3 = 9$ in this study. In addition, some a-priori domain knowledge, known as favourable rule, are allowed to be identified by domain expert. The following domain knowledge are available to us prior to the ANFIS training:

- A possible fault is detected if Wind Speed is *LOW* and Power Output is *HIGH* (Shown as Rule 3 in Figure 4.2 (a));
- A possible fault is detected if Wind Speed is *HIGH* and Power Output is *LOW* (Shown as Rule 7 in Figure 4.2 (a));

The above statements enabled us to distinguish two favourable rules, whose locations in the input space are highlighted in Figure 4.2(a). Figure 4.2(a) is the region view of the input space, which has been partitioned into 9 overlapping sub-regions to represent the aforementioned 9 rules. Amongst the 9 rules, rule 3 & 7 are the favourable rules. Figure 4.2(b) shows the real power curve and the encircled areas correspond to the favourable rules which have sparse data distribution.

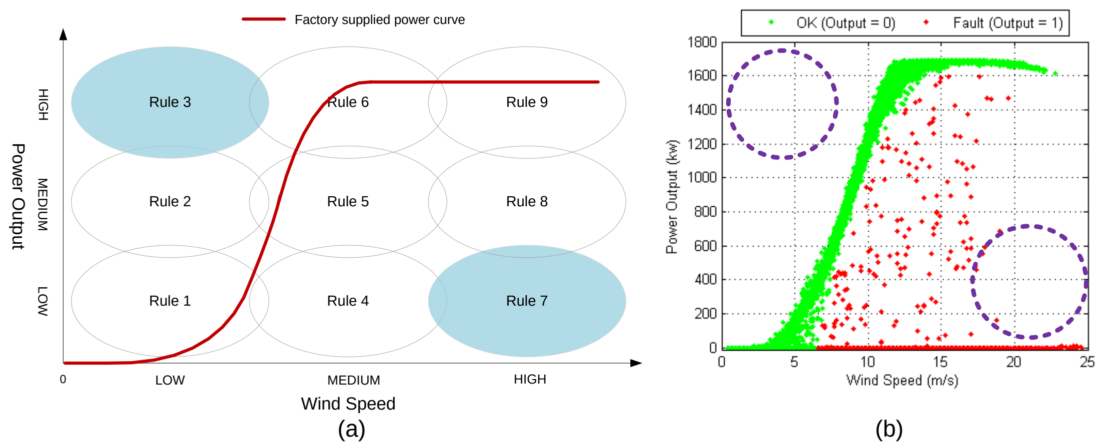


Figure 4.2: (a) 2D input space that has been partitioned by 9 overlapping fuzzy sets. The shaded subspaces correspond to the a-priori domain knowledge. (b) The corresponding real data.

The *i*th rule in conventional ANFIS, as mentioned in Section 4.1.1, can be transformed and represented as:

$$IF\ x_1\ is\ A_1^i\ AND\ x_2\ is\ A_2^i\ \dots AND\ x_n\ is\ A_n^i,$$
$$THEN\ f_i = k_0^i + k_1^i x_1 + \cdots + k_n^i x_n \tag{4.7}$$

where $k_0^i, k_1^i, \dots, k_n^i$ are the coefficients of the consequent linear model defined for the *i*th rule. $X = [x_1\ x_2\ \dots\ x_n]^T$ represents an input vector with n dimensions. A_n^i represents the linguistic labels, *Low*, *Medium* & *High*, assigned to the input variable x_n . f_i is the consequent first-order polynomial function to be determined.

According to Tewari (2009), the technique is based on the framework of rule-centred Takagi-Sugeno-Kang (TSK) fuzzy model in which the consequent first-order polynomial function, the f_i as mentioned in Eq. 4.7, can be interpreted as the 1st order Taylor series approximation, as shown in Eq. 4.8.

$$IF\ x_1\ is\ A_1^i\ AND\ x_2\ is\ A_2^i\ \dots AND\ x_n\ is\ A_n^i,$$
$$THEN\ f_i \approx \theta_0^i + \theta_1^i\left(x_1 - c_1^i\right) + \cdots + \theta_n^i\left(x_n - c_n^i\right) \tag{4.8}$$

where $c^i = [c_1^i, \dots, c_n^i]$ is a n dimension vector representing the *i*th rule centre having the same dimension as input. The centre of a rule is simply the geometric centre of the corresponding multivariate fuzzy set. θ_n is the corresponding coefficients in Taylor series. The coefficients k_n^i as shown in Eq. 4.7 have a straight forward relationship with the coefficients θ_n:

$$\begin{cases} k_0^i = \theta_0^i - \sum_{n=1}^{N} \theta_n^i c_n^i \\ k_n^i = \theta_n^i \end{cases} \tag{4.9}$$

The Eq. 4.8 can be further expanded as follows:

$$f_i \approx f_i(c^i) + \frac{df_i}{dx_1^i}(x_1 - c_1^i) + \cdots + \frac{df_i}{dx_n^i}(x_n - c_n^i) \qquad (4.10)$$

Then, we find that the parameter θ_0^i signifies the underlying function value at the ith rule centre and parameter θ_n^i represents the gradient of the function along the n dimension about the ith rule centre. The relationship is shown in Eq. 4.11.

$$\begin{cases} \theta_0^i = f_i(c^i) \\ \theta_n^i = \dfrac{df_i}{dx_n^i} \end{cases} \qquad (4.11)$$

After that, the 1st order ANFIS can be expressed as follows:

$$f \approx \sum_{i=1}^{R} \left(\theta_0^i + \sum_{n=1}^{N} \theta_n^i (x_n^i - c_n^i) \right) * \bar{w}_i \qquad (4.12)$$

where R is the total rules and N is the dimension of the input vector.

The qualitative domain knowledge is incorporated into the model in the form of Gaussian basis function Φ at the centre of each favourable rule, i.e.

$$\Phi_r^j = a^j * \exp\left(- \sum_{i=1}^{n} \left(\frac{c_i^r - c_i^j}{\sigma_i^j} \right)^2 \right) \qquad j \in J, r \in R \; and \; J \subseteq R \qquad (4.13)$$

where J are the available favourable rules out of R total rules. This Gaussian basis function, with parameters α^j and σ_i^j, is at the centre of each jth favourable rule. The Gaussian functions are used to mimic the available domain knowledge. Since there can be several favourable rules, the model output at ith rule can be represented as the weighted geometric mean of the individual Gaussians, as follows:

$$\theta_0^i = \prod_{j=1}^{J} (\Phi_i^j)^{\gamma_i^j} \qquad (4.14)$$

$$\gamma_i^j = \frac{1}{\sum_{n=1}^{J}(D_{ij}/D_{rn})} \qquad (4.15)$$

where θ_0^i is the consequent parameter of the *i*th rule of the 0th order ANFIS. The γ_i^j represents the weight that signifies the degree of closeness of the *i*th rule centre to the centre of *j*th favourable. D_{ij} is the Euclidean distance between two rule centres. Therefore, the 0th order ANFIS can be computed using the following equation:

$$f = \prod_{i=1}^{R} (\theta_0^i)^{\overline{w}_i} \qquad (4.16)$$

where $\overline{w}_i$ is the *normalised firing strength* of each rule. Eq. 4.16 is completely different from conventional way of computing the output of a Takagi-Sugeno FIS, the rationale of defining the output in this way is that it renders the model output intrinsically linear in terms of parameters α^j and σ_i^j (Jang et al. 1997).

The definition of the 1st order ANFIS, as shown in Eq. 4.12, can be transformed as:

$$f \approx \sum_{i=1}^{R} \left(\theta_0^i + \sum_{n=1}^{N} \left(\alpha_n^i \theta_n^i + (1 - \alpha_n^i) \right) \theta_n^i \left(x_n^i - c_n^i \right) \right) * \overline{w}_i \qquad (4.17)$$

where α_n^i assumes a value of either 0 or 1 depends on the input location to the corresponding rule centre as follows:

$$\begin{cases} \alpha_n^i = 1 & if \ x_n^i - c_n^i > 0 \\ \alpha_n^i = 0 & if \ x_n^i - c_n^i < 0 \end{cases} \qquad (4.18)$$

The entire procedure of the APK-ANFIS is outlined in Tewari (2009), the learning algorithm of this APK-ANFIS used the quadratic programming for solving the constraint bound given by domain knowledge in the forward pass and the gradient-decent method was used in the backward pass because of its ease of implementation. According to two examples shown in Tewari (2009), the results of this technique showed relatively better interpretability under two different conditions: data with noise or certain inputs space is sparse.

In order to illustrate the advantage of incorporating prior domain knowledge into the APK-ANFIS, fault detection using the WT power curve was studied using real data from Figure 4.2(b). The ANFIS and APK-ANFIS were built in a similar manner with 3 MFs in each input, maximum training iteration 50 and minimum error 0.01. Figure 4.3 (a) & (b) show the output surface generated by the trained ANFIS and APK-ANFIS. Clearly, the conventional ANFIS resulted in an output surface inconsistent with domain knowledge, shown as large Z scale and is encircled in Figure 4.3(a). This is because of insufficient training data in the various areas and resulted in a trained model that behaved erratically under unseen input conditions. On the contrary, with the incorporation of domain knowledge, Figure 4.3(b), APK-ANFIS result has shown relatively better interpretability and was able to maintain model consistency even in the regions with few data points. With this advantage of improved interpretability, this research is going to apply the APK-ANFIS to build a fault diagnosis procedure for a WT pitch system.

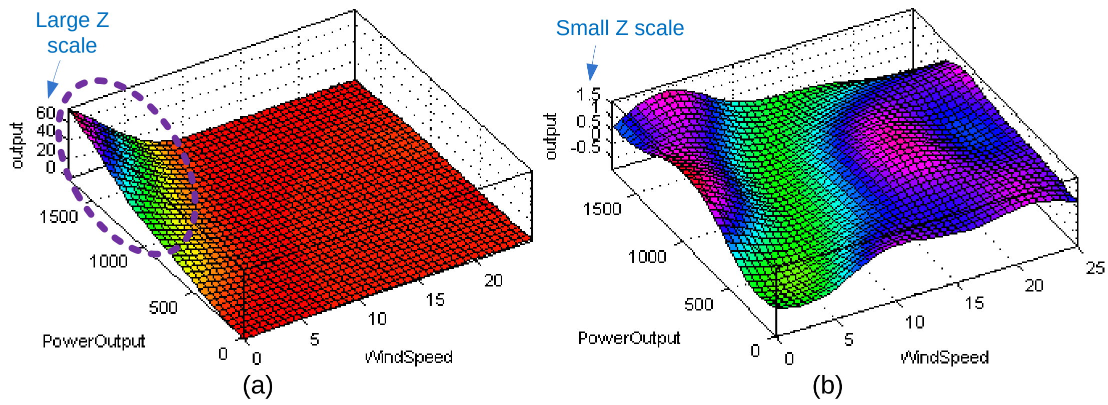


Figure 4.3: (a) Conventional ANFIS result without a-priori knowledge, (b) APK-ANFIS result;

4.2. Proposed Fault Diagnosis Procedure

The use of APK-ANFIS was explained in Section 4.1 and this section will propose a new approach for analysing WT SCADA data using the APK-ANFIS with the aim of achieving automated detection of pitch faults.

4.2.1. Four Critical Characteristic Features

As has been mentioned in Section 3.3, six known pitch faults from 2 WTs have been analysed using typical variable-speed pitch-to-feather control and pitch torque power curve characteristics. The analysis results have shown that the pitch fault symptom can be clearly identified from four features, named as CCFs, as shown in Figure 4.4 and the corresponding physical properties are described below:

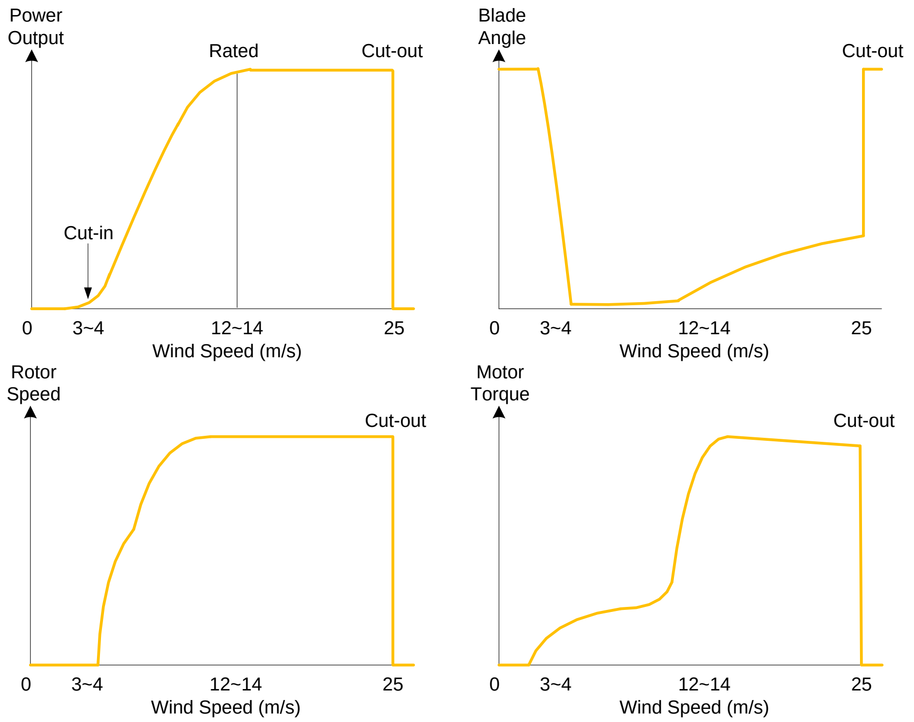


Figure 4.4: Four CCFs for pitch fault analysis

- **Wind Speed vs. Power Output:** This is the well-known power curve, as shown in Figure 4.4(a), which has been widely used for WF operator to monitor the WT's performance. In general, a WT start generating power when the wind speed rises above the cut-in speed, at 3-4 m/s. Then, the power generation is in proportion to wind speed until it rises to rated speed, at 12-14 m/s. At higher wind speed, the WT is arranged to limit the power to the rated power output and this is done by adjusting the blade angle to keep the power at the constant level.
- **Wind Speed vs. Blade Angle:** In a variable speed WT, its blade angles are regulated to enhance wind energy conversion and limit the power to the rated power output at high wind speeds. At very low wind speeds, WT blades are usually kept in feathered position for safety. They are adjusted to the optimal angle to extract optimum power between cut-in and rated wind speed. When wind speed rises above the rated speed, the WT adjust

the blade angle to keep the power at rated power output. The blade is driven into their feathered positions when wind speed is larger than cut-out. The sketch diagram is shown in Figure 4.4(b).

- **Wind Speed vs. Rotor Speed:** The rotor speed refers to the WT main shaft rotation speed. In a variable speed WT, rotor speed is changed with respect to the change of wind speed, as shown in Figure 4.4(c). In general, the main shaft doesn't move until the wind speed rises above cut-in. Then, the rotor speed is in proportion of wind speed until the wind speed reaches the rated speed. After that, the rotor speed is kept at the constant maximum level until cut-out.
- **Wind Speed vs. Blade Motor Torque:** In electrical pitch system, each blade is fitted with a pitch motor to control the blade angle, extract optimum power from the wind and avoid rotor over-speed. At the very low wind speed, blade motor torque is not required as blade is in the feathered position. Once the wind speed rises above cut-in, the blade motor should start working to regulate the blade to enhance wind energy conversion. Large torque is required from rated speed to cut-out to regulate and maintain blade angle.

It is widely believed that a normal WT's pitch system should subject to above physical properties. Any observation deviates far away from the above curves can be regarded as a possible fault. The idea of using anomaly detection, as mentioned in Section 3.4, is lay in these physical properties in a normal WT pitch system.

4.2.2. Proposed Fault Diagnosis Procedure

In this work, the aforementioned four CCFs and the SCADA alarms are utilized for detecting the incipient WT pitch faults. The proposed fault diagnosis procedure consists of 4 modules, as shown in Figure 4.5, are specifically explained as follows:

- **Data Acquisition:** This module will collect data from the SCADA system and ensure no maintenance or manual stop in the selected period. In addition, data must be not NULL and subject to factory supplied ranges:
  - Wind speed range from 0 to 25 m/s;
  - Rotor speed range from 0 to 20 rpm;
  - Blade motor torque range from 0 to 45 kN;
  - Power output range from -10 to 1750 kw;
  - Blade angle range from 0 to 90 degree;

  Note: WT use electricity from the grid is shown as negative in power output.
- **Feature Extraction:** Valid data are divided into signals and alarms. Four CCFs, as mentioned in Section 4.2.1, will be extracted from signals. Alarm distribution & showers will then be produced from alarm data to validate the final result. (Alarm distribution & showers (Qiu et al. 2012) is the number of alarms during a certain period of times, for example a day in this proposed system.)
- **Multiple Diagnosis:** The four CCFs will be passed to the corresponding trained APK-ANFIS to calculate the fault degree. The overall result is the aggregation of the 4 individual APK-ANFISs, defined as:

  $$Result = \frac{\sum_{i=1}^{4} a_i * APKANFIS_i}{\sum_{i=1}^{4} a_i} \qquad (4.19)$$

  where a_i is the corresponding weight. All a_i are set to 1 for calculating the average in this case.
- **Fault Diagnosis Result:** Finally, the overall result will be checked against SCADA alarms to provide the warning to the WF operator.

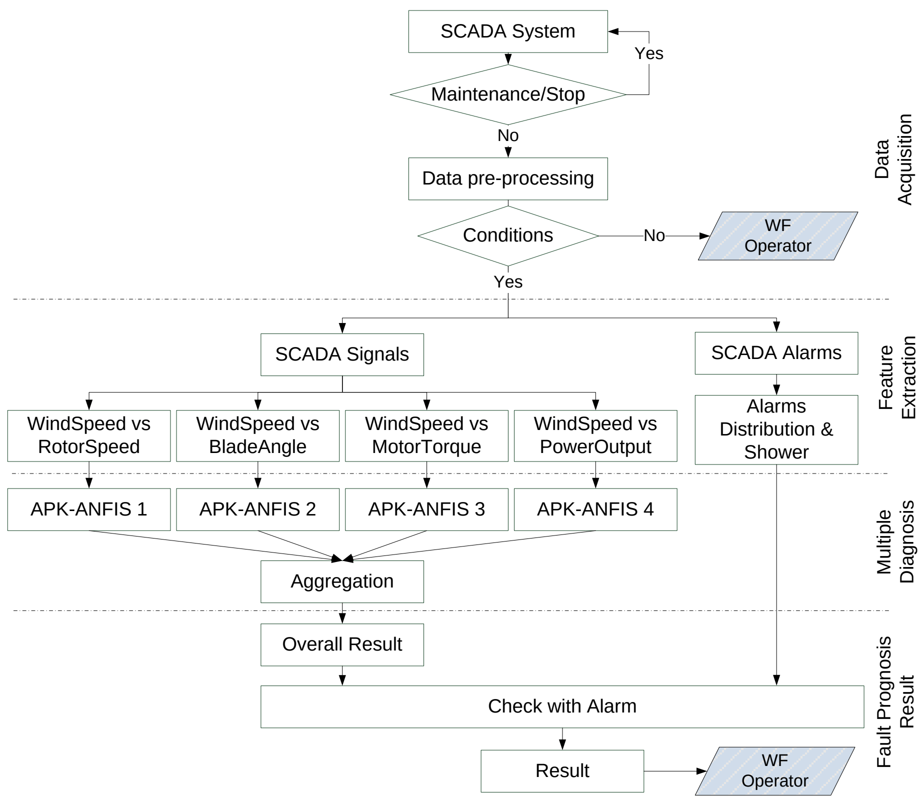


Figure 4.5: The proposed diagnosis procedure

It can be seen from above description that WT stop and maintenance are not considered in the proposed system. This is due to the nature complexity of human intervention that may cause the signals to be outputted like a fault and result in false identification. In addition, any observed CCFs' values beyond the factory supplied range are largely caused by sensor fault. For simplicity, these situations are not diagnosed by the proposed system and they will be recorded and forwarded to the WF operators for further inspection. Next, this research would have to train the proposed system using the six known pitch fault data as mentioned in Section 3.3. The first step is to collect and label the training data.

4.2.3. Acquiring Training Data

In order to construct the proposed diagnosis procedure, the data of the six known pitch faults are used as a knowledge base for training and testing the individual APK-ANFIS. The fault behaviours of the four CCFs can be represented using a vector as follows:

$$P_i = \left[I_{i,1}, I_{i,2}, O_i\right]^T, \quad i \in [1,2,3,4]$$

where P_i correspond to the four CCFs as mentioned in Figure 4.4 and the aggregation of them can be considered to characterise pitch fault. $I_{i,1}$ and $I_{i,2}$ are the inputs of the ith CCF. The O_i is the corresponding output and it takes one of the values 0 and 1, which indicate the *Absent* and *Present* state of the pitch fault. Thus, abnormal data, such as a possible pitch fault, were given value 1 and the remainders were given value 0, to represent *No* pitch fault.

A labelling procedure is needed in order to assign the correct class, *Absent* or *Present* state of the pitch fault, to each of the training data. The procedure is listed as follows:

- Giving value 0 to data in *After Maintenance* period, which indicates the *Absent* state of pitch fault;
- Based on the 4 CCFs' physical properties as discussed in Section 4.2.1 and the comparison between *Generating fault* and *After Maintenance* periods, the abnormal data in *Generating Fault* period are found and given value 1, which indicated the *Present* state of the pitch fault. The reminders represent *Absent* of pitch fault and are given value 0;
- Data in *Maintenance* period are not included in the training data because the nature complexity of human intervention may cause the signals to be outputted like a fault;

An example of labelling abnormal data in Case 1 is shown as encircled regions in Figure 4.6.

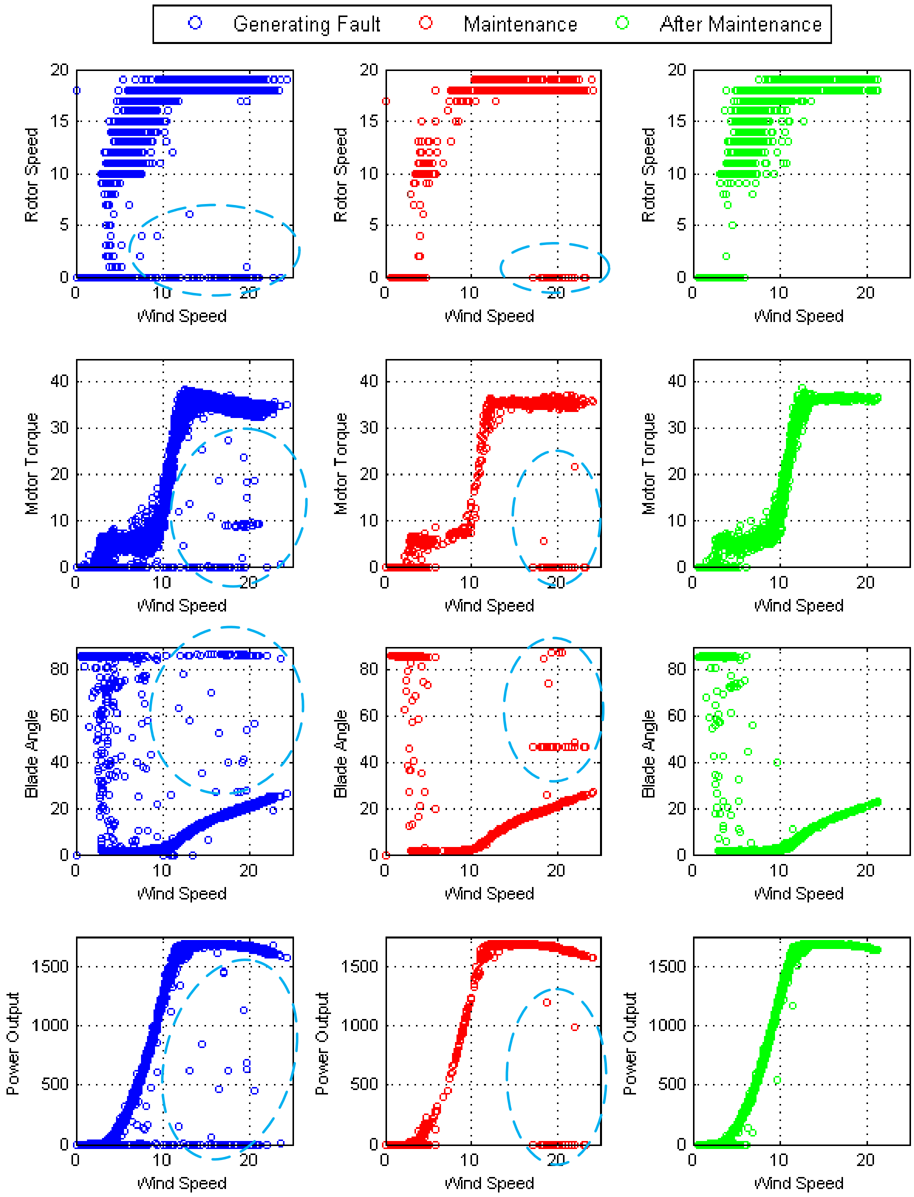


Figure 4.6: Labelling the abnormal data in Case 1

By doing above labelling procedure for all the 6 known pitch faults and putting six pitch faults' data together, 26,971 sets of data are collected, as shown in Figure 4.7.

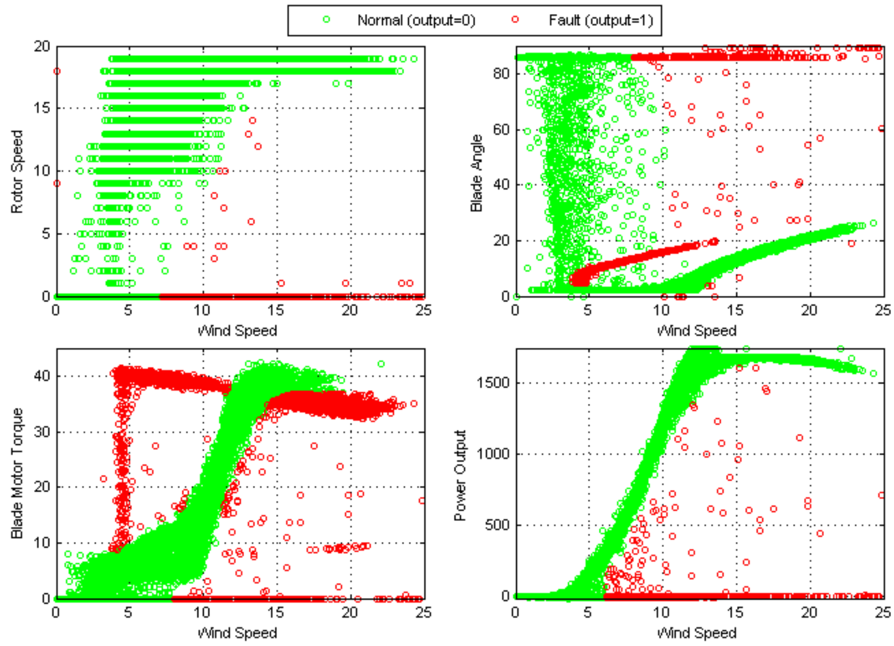


Figure 4.7: Training data from the six known pitch faults.

Figure 4.7 shows the labelled six pitch faults data and it also demonstrates that some input sub-regions have sparse data distribution, which a WT is hardly or impossible to produce signal there, for example: the sub-region where wind speed is low and power output is high. For those sub-regions, the a-priori domain knowledge is allowed to be identified by domain expert. The following sub-regions, as encircled in Figure 4.8, are expected to have 1 output to indicate the *Present* of fault. They are identified by domain expert and available to us prior to the ANFIS training.

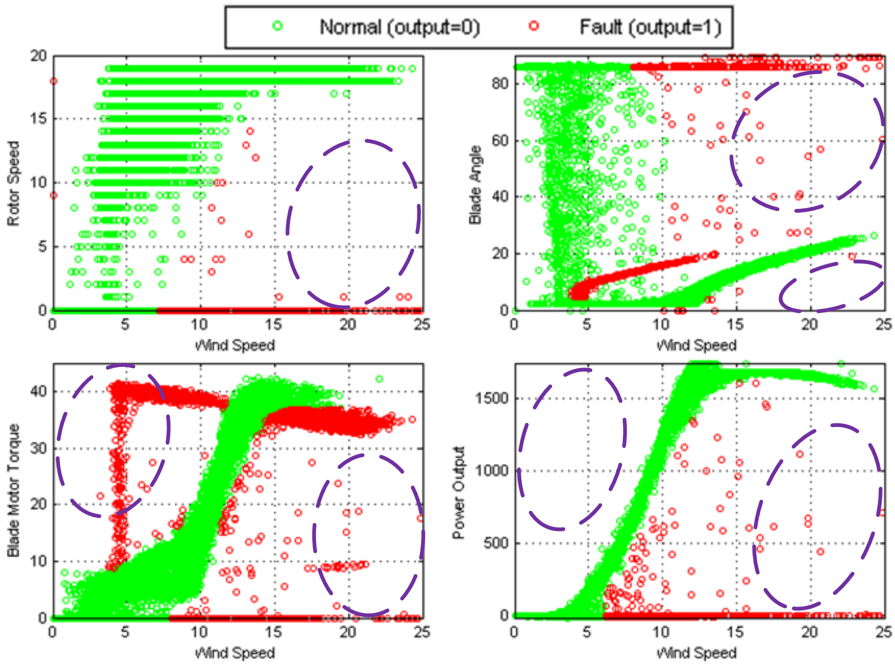


Figure 4.8: Training data from the six known pitch faults. Encircled areas have insufficient data and a-priori approach is required.

4.2.4. Training Process

The training process is shown in Figure 4.9. During the training, an input vector was fed into the input layer of the APK-ANFIS and the desired output corresponding to input vector was used to compare with the actual APK-ANFIS output. If the result of the comparison was unacceptable, the Hybrid training algorithm adjusted the APK-ANFIS parameters to be consistent with the imposed input vector and desired output. The parameters were then readjusted to accommodate new input vector with the corresponding desired output. The training process is repeated until convergence within a specified error or the learning has reached the maximum iteration. In this work, the minimum value was set to 0.01 and the maximum iteration was set to 150. The training was performed on ReliaWind Server with 2 processors, 16 cores, 48G memory and 8TB hard drive and using 64-bit Matlab.

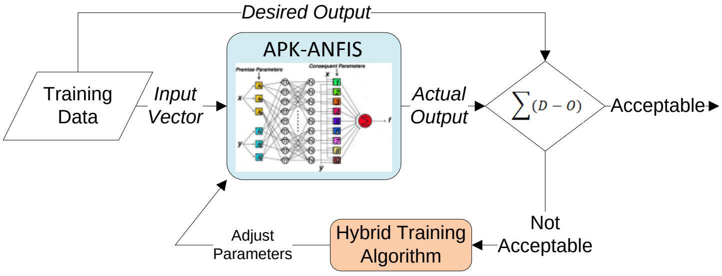


Figure 4.9: APK-ANFIS training process

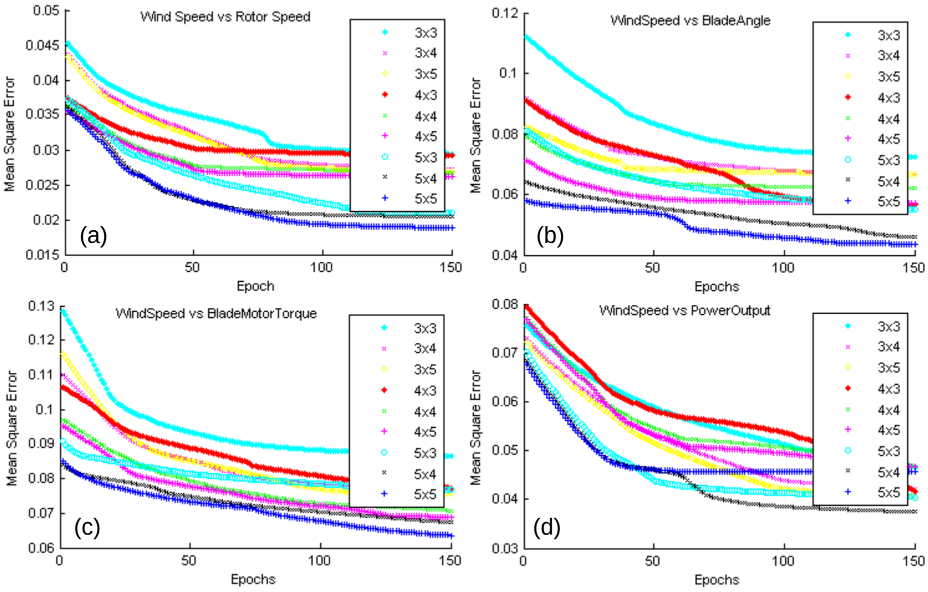


Figure 4.10: RMSE curve of different ANFIS structures; (a) WindSpeed vs RotorSpeed ANFIS; (b) WindSpeed vs BladeAngle; (c) WindSpeed vs BladeMotorTorque; (d) WindSpeed vs PowerOutput;

In order to find the optimal structure for each individual APK-ANFIS, a batch testing with different number of MFs for each input were examined, as shown in Figure 4.10. These calculate the mean square error of different structures and finally the optimal structures were chose, as shown in Table 4.2.

APK-ANFIS model	Optimal structure (The number of MFs)
Wind Speed vs Rotor Speed	5-by-5
Wind Speed vs Blade Angle	5-by-5
Wind Speed vs Blade Motor Torque	5-by-5
Wind Speed vs Power Output	5-by-4

Table 4.2: The optimal ANFIS structures

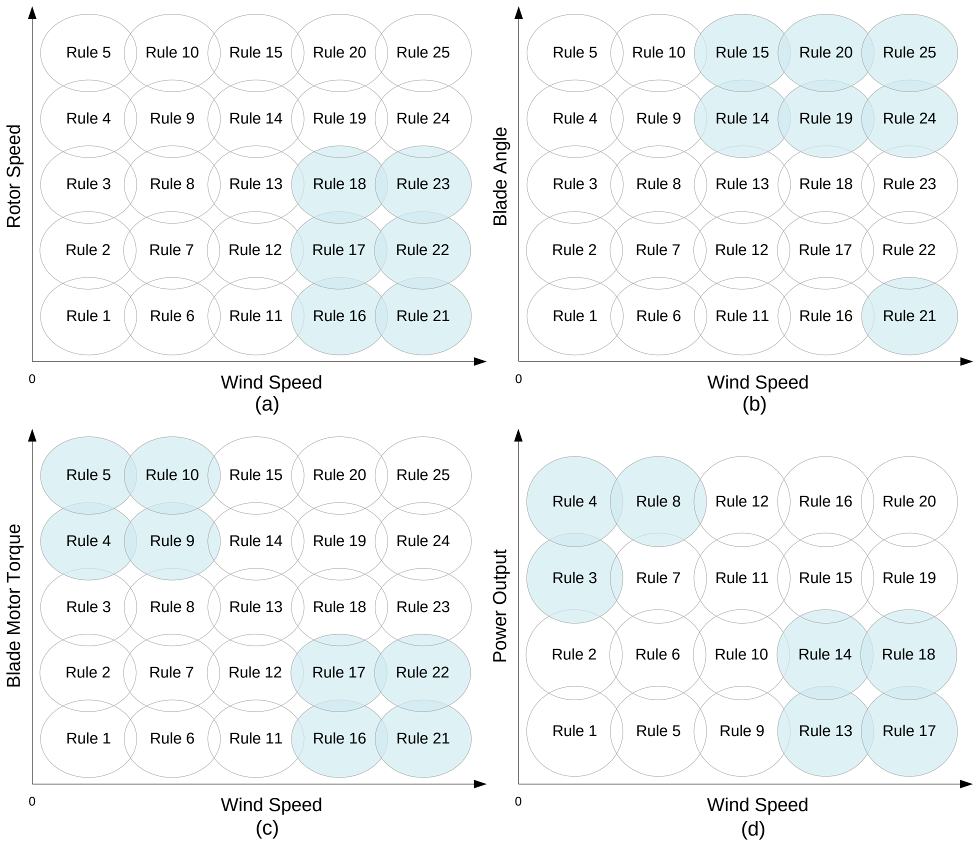


Figure 4.11: Sub-regions view of Figure 4.8. Corresponding sub-regions are expected to have 1 output to indicate the Present of fault.

Then, the corresponding sub-regions, which are expected to have 1 output to indicate the *Present* of fault, can be specified and shown in Figure 4.11. After that, the data were partitioned into training and testing data sets. Cases 1-5 provided the training data and Case 6 was used to test the trained model. Its success at actual outputs that are as close as possible to the desired outputs determines how well the network has learned or captured the relations between the inputs and outputs.

4.2.5. Trained System

Finally, the output surfaces generated by individual trained APK-ANFIS models are shown in Figure 4.12. This clearly demonstrates that abnormal data will give a large output, close to 1 as shown in the "Hill", while normal data will give a small output, close to 0 and shown as the "Valley". A demonstration of the proposed diagnosis system with an arbitrary threshold 0.5 was made and shown in Figure 4.13, where Figure 4.13(a) demonstrates a normal running WT and Figure 4.13(b) demonstrates the detection of a possible pitch fault for which an "Alarm" has been triggered. The Figure 4.14 is the demonstration of the diagnosis system with showing every individual APK-ANFIS result.

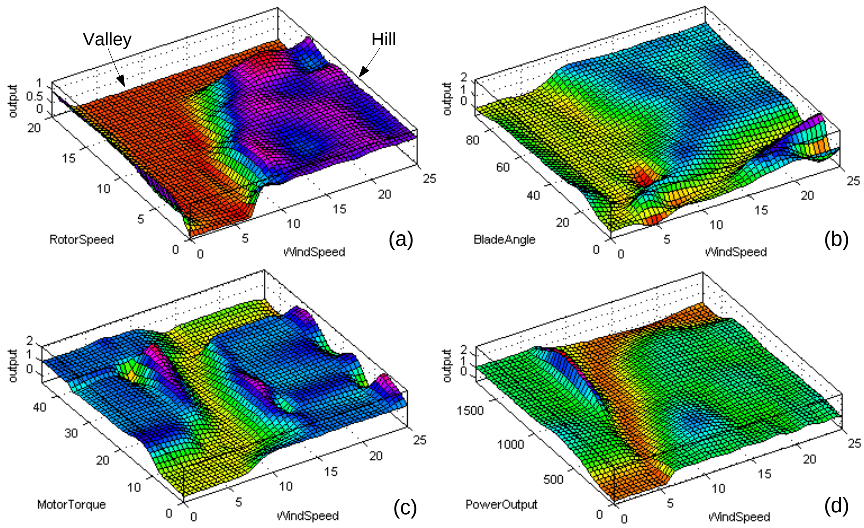
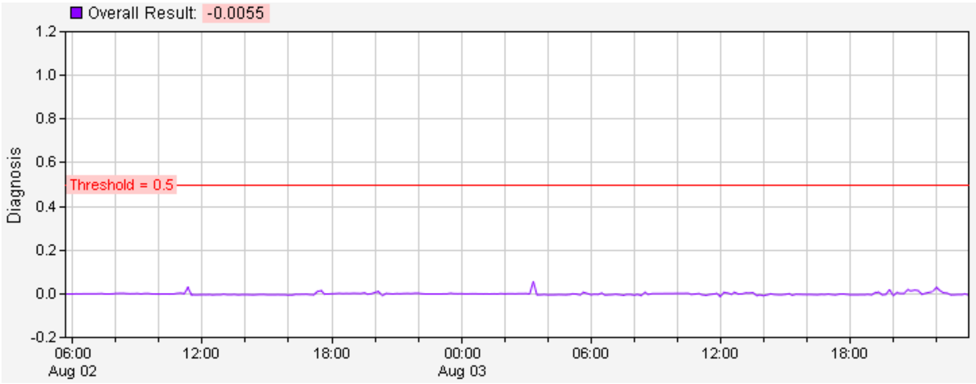
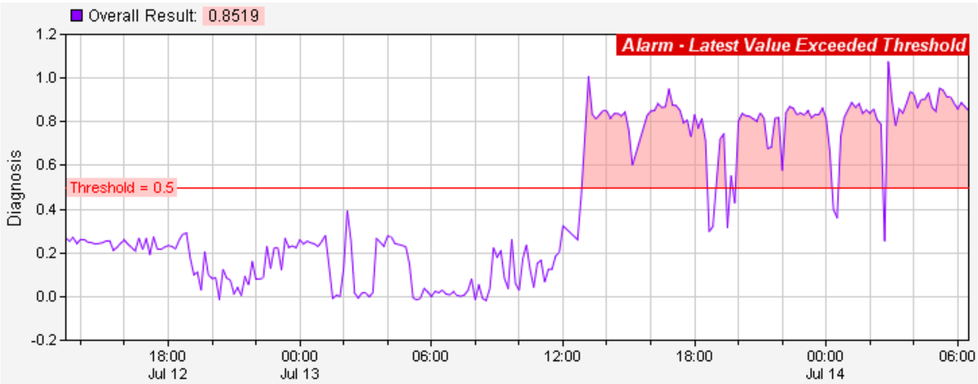


Figure 4.12: Output surfaces generated from the trained APK-ANFIS



(a) A normal running WT



(b) A possible pitch fault has been detected

Figure 4.13: Demonstration of the diagnosis system with an arbitrary threshold 0.5

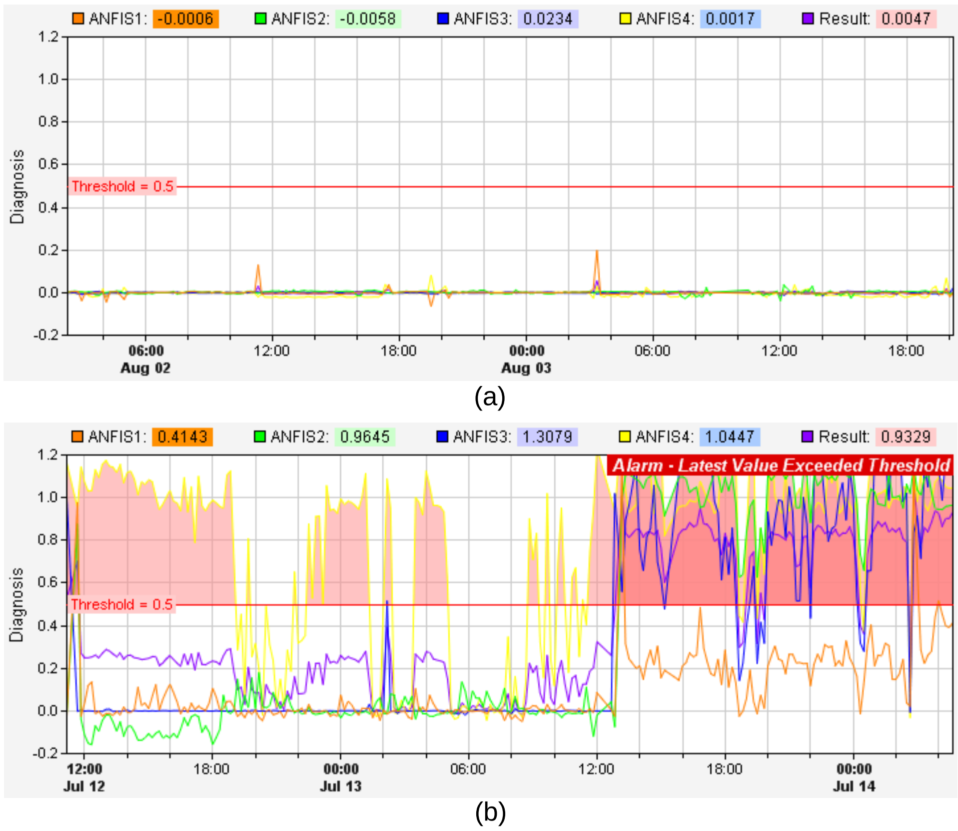


Figure 4.14: Demonstration of the diagnosis system with showing individual result

4.3. Chapter Summary

The development of ANFIS has been introduced and the rationale behind the APK-ANFIS approach selected to develop an automated on-line fault diagnosis system was discussed. A fault diagnosis system with using the APK-ANFIS and the four critical CCFs is proposed. The data of the six known WT pitch faults are labelled based on expertise. After that, the proposed system is trained using the six known pitch faults with a-priori knowledge incorporated. In the end of this chapter, a trained system is obtained and a demonstration of the diagnosis system is shown.

The next chapter will present the effectiveness of the proposed approach using five metrics: (1) the trained system will be tested in another WF containing

26 same manufactured WTs to show its prognosis ability; (2) the first test result will be compared to a general alarm approach; (3) a Confusion Matrix analysis is made to demonstrate the accuracy of the proposed approach; (4) this approach is applied to pitch data from different manufactured WTs and the result is analysed using Confusion Matrix analysis; (5) the comparison of the results of applying the proposed approach to two different manufactured WTs.

Chapter 5.

Test Results & Validation

The previous chapter has introduced the development of ANFIS and described the rationale behind the APK-ANFIS approach selected to develop an automated on-line fault diagnosis system. The data from six known pitch faults, developed from ReliaWind knowledge are labelled and used to build the proposed system with a-priori knowledge incorporated. The output surfaces generated from the trained system have shown improved ability to interpret the previously unseen conditions and the system is expected to analyse WT SCADA data to provide early pitch fault identification.

This chapter aims to show the effectiveness of the proposed approach by applying the approach to the data from two different designs of WTs, manufactured by Alstom & Mitsubishi, with two different types of SCADA, demonstrating the adaptability of APK-ANFIS for application to a variety of technologies. The details of the two tests are listed as follows:

- A WF in Spain, containing 34 Alstom WTs, out of which 26 were tested to show the method's prognosis ability. The result is compared to a general alarm approach and analysed by Confusion Matrix analysis to show the accuracy and precision.
- Another WF in Brazos, Texas, USA, containing 160 Mitsubishi WTs, out of which 22 were tested using a similar system built to show the generalised effectiveness of this approach.

A comparison study of applying this approach on two different designs of WTs is shown at the end of this chapter and this comparison of the application of the approach to two widely different designs of WTs strengthens the applicability of the proposed approach.

5.1. Comparing Alstom & Mitsubishi Wind Turbines

The proposed approach is applied to pitch data from two different types of WT with different pitch systems, described earlier in the Thesis:

- **Alstom WTs:** 26 Alstom ECO80 WTs, 1.67MW variable speed, variable pitch WTs with electric pitch-to-feather control will be used in the first test;
- **Mitsubishi WTs:** 22 Mitsubishi M1000 WTs, 1MW fixed speed, variable pitch WTs with hydraulic pitch-to-stall control will be used in the second test.

The purpose of this comparison is based upon the similarities between their respective control systems but also to identify the control signals that will be useful in each case to apply APK-ANFIS approaches to detect the prognoses faults in their respective pitch systems. These control signals will not be the same in the two cases, depending on the WT technology and the SCADA system applied to those WTs.

In order to make a comparison the results from the two WT types have been normalised using a feature scaling:

$$x_i^{new} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \qquad (5.1)$$

This will rescale the variable ranges to [0, 1].

5.1.1. Power Curves

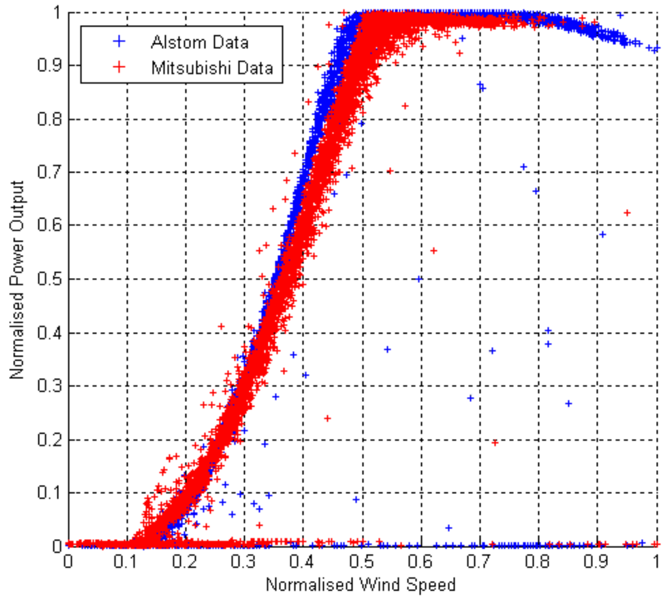
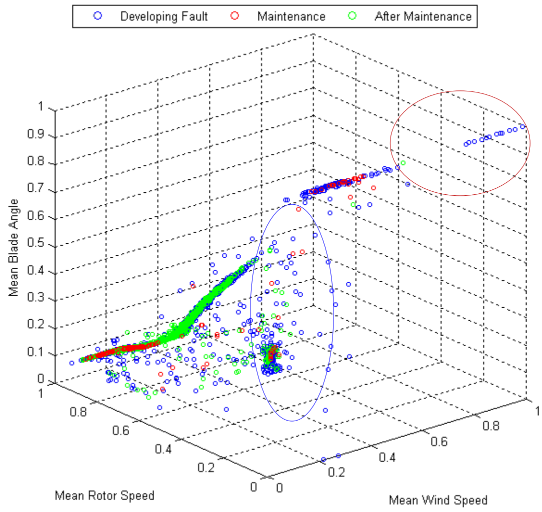


Figure 5.1: Alstom & Mitsubishi WTs, power curve comparison

The two types of WT display similar power curves, however, note the more abrupt, sharper effect of stall control in the Mitsubishi WT and the softer effect of variable speed control in the Alstom WT.

5.1.2. Pitch Control Plot Comparison

For both Alstom and Mitsubishi WTs the variable-speed pitch-to-feather/stall controls were recorded from SCADA data. The Mitsubishi WT blade angles range from -20 º to -110º because the machine is pitch-to-stall rather than pitch-to-feather. In order to compare them directly with Alstom data, Mitsubishi WT blade angle data was multiplied by -1. An example of the raw Mitsubishi SCADA blade data can be seen in Appendix C.



(a) Mitsubishi WT Case 1 A01

(b) Alstom WT Case 1

Figure 5.2: The variable-speed pitch-to-feather/stall plot. Circled in red represent the anomalies, circled in blue represent the noise data.

Figure 5.2 shows some differences as follows:

- For the Mitsubishi WTs, in Brazos, Texas, USA, the range of wind speeds was less than for the Alstom WTs in Spain, as noted from Figure 5.1;
- The mean rotor speed for the Mitsubishi WTs was fixed and high compared to the Alstom WTs, as expected for a comparison between fixed and variable speed WTs;
- For the two WT types the operational range of blade pitch angles was similar, despite the different technologies. The Mitsubishi WT blade angles are measured with an offset of about -4º with an operational range in the pitch-stall direction of about 17.5º, whereas the Alstom blades had an operational range in the pitch-to-feather direction of about 24.5º;
- Pitch anomalies are circled in red in Figure 5.2, both of them can be regarded as possible pitch faults and they demonstrated the common pitch fault symptom;

For Alstom WTs the SCADA also records blade pitch motor torque, because an electric pitch system is being used, therefore a pitch-torque power-curve plot can be recorded, as shown in Figure 5.3. Comparison between Figure 5.2 & 5.3 shows that the latter is less noisy.

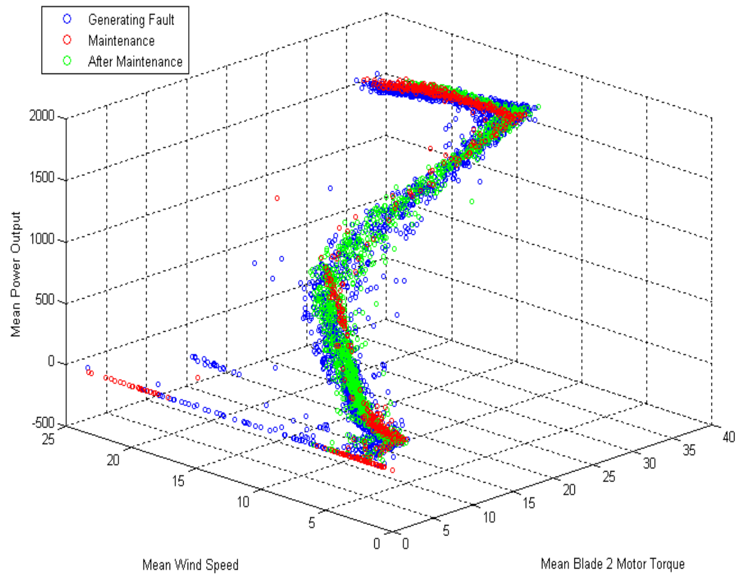


Figure 5.3: Alstom WT, Pitch Control Plot, Pitch-Torque Power-Curve plot.

The noise in the variable-speed pitch-to-feather/stall plots is primarily due to SCADA data collected when the WT is shut down with the blades either in the parked or fully pitched positions. These spurious points can clearly be seen in Figure 5.2, as encircled in blue, for both the Alstom and Mitsubishi WTs.

5.1.3. Comparison Conclusion

The comparison study has shown that:

- Pitch control data for WTs of different architecture can be analysed using the same methods;
- The SCADA-derived data pitch control analysis used the following:
    - Pitch-torque power-curve plot, Figure 5.3;

  - Variable-speed pitch-to-feather/stall Control Plot, Figure 5.2.
- It was clear from the analysis that the pitch-torque power-curve plot was less noisy than the variable-speed pitch-to-feather/stall control plot, because it was less contaminated by WT blade parked & stop positions;
- Therefore wherever possible, i.e. on electrical pitch WTs, the pitch-torque power-curve plot should be used for pitch fault analysis;
- Nonetheless the variable-speed pitch-to-feather/stall plot from hydraulic pitch WTs yielded valuable fault detection results;

5.2. Test on Alstom Wind Turbines

In this section, the trained system, described in Section 4.2.5, was tested in a WF to validate its effectiveness. After that, the results will be compared to a commonly used alarm approach to demonstrate the advantage of the prognostic horizon. The result will also be evaluated by a Confusion Matrix analysis to check the validity of the results.

5.2.1. Data Preparation & Selection

The trained system has been applied to a WF containing 26 Alstom WTs. There were 34 WTs in this WF but 8 had insufficient SCADA data for analysis, so were ignored. The data period was 28 months, from 01/Jun/2006 to 30/Sep/2008. For the selected 26 WTs 910 pitch corrective maintenance records were found in this period, these were further reduced to 487 records according to the following criteria:

- A maintenance followed by another maintenance within an interval of not more than 2 days was considered as one effective maintenance record;

Some summary statistics for this WF are:

- An average of 18.7 pitch effective corrective maintenances per WT in this 28 month period;
- That is an average of 0.67 pitch effective corrective maintenances per WT per month;

It can be seen from this that the pitch system is a significant fault for this WT.

5.2.2. Fault Prognosis using Proposed System

In order to test the trained system with new WF data, an algorithm was written to apply the trained diagnosis procedure to calculate the prognostic horizon for every pitch corrective maintenance activity. The Pseudo-code is shown in Table 5.1. Three potential prognostic horizons of 7, 14, or 21 days, were selected to avoid the false identifications. For example a half year early warning probably has nothing to do with this corrective maintenance. In addition, to further reduce false identification, required Threshold and Window Sizes were defined as follows:

- **Threshold (T)** is the critical level for WF operator to consider investigating a possible fault. It is the aggregation of the four APK-ANFIS results and its output range is from 0 to 1, as shown in Figure 4.5.
- **Window Size (W)** is the number of the consecutive data used to identify the incipient fault. The SCADA data used in this research was measured every 10 minutes; however a single measurement is insufficient to demonstrate a possible fault, thus this work chose a Window Size of 6, 18 and 48 10 min intervals, representing 1, 3 and 8 hours respectively, to avoid false identification.

The initial test had applied the trained system against to its training data, that is the 6 pitch fault cases shown in Table 3.5. The average prognosis horizons (in days) with different window sizes and thresholds are shown in Figure 5.3. It can be seen that no prognosis horizon is longer than 20 days. In addition, the algorithm shown in Table 5.1 had been used to test the potential prognosis horizon out to 30 days. A small peak prognostic horizon occurring between 22 to 26 days was found, as encircled in Figure 5.4. According to these two pieces of evidences, the author suspected that a pitch fault is most likely to reach failure 21 days after the first indication. Therefore, the author decided to use 21 days as the maximum prognosis horizon for the application of the online fault prognosis test.

Prognosis Horizon (days)		Threshold							
Window Size (Consecutive Data)		0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.5 hour	3	19.55	19.15	19.03	16.98	15.27	13.56	10.19	10.19
1 hour	6	19.15	19.00	17.10	15.26	13.06	13.00	10.17	10.17
1.5 hours	9	19.12	17.08	17.07	13.06	12.99	12.98	10.16	10.15
2 hours	12	19.10	17.07	13.06	13.05	12.98	12.97	7.86	6.73
3 hours	18	18.97	13.05	13.04	10.74	9.56	9.34	6.51	6.50
4 hours	24	18.40	13.04	10.74	9.57	9.33	9.31	2.78	2.76
5 hours	30	18.38	10.75	9.57	9.33	9.31	5.59	2.63	2.60
6 hours	36	15.47	9.62	9.34	9.31	9.27	6.00	2.60	1.70
7 hours	42	15.46	9.58	9.50	9.27	6.01	5.97	1.70	1.65
8 hours	48	15.45	9.52	9.47	10.44	5.98	5.91	2.24	0.42
9 hours	54	15.44	9.52	9.45	5.99	5.95	7.85	0.43	0.38

Figure 4 Test against to the training data and the result of average prognosis horizons in days

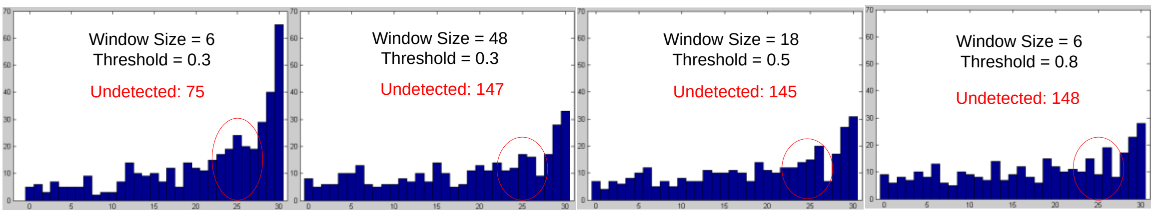


Figure 5.4 Plot of distribution of the prognosis horizon in 30 days

Step 1:
Data Cleansing – remove data when it has maintenance;
Step 2:
Define H, W, T to represent Prognostic Horizon, Window Size, Threshold respectively
Declare H = 7, 14 or 21; W = 6, 48 or 18; T = 0.3, 0.5 or 0.8;
For each WT in the WF
For each "pitch corrective maintenance record" in the selected WT
Within the given Potential_Horizon = H days
Find the earliest date when Window_Size = W and Threshold >= T
Prognosis_Day = Maintenance_date – The_Earliest_date
End

Table 5.1: Pseudo-code for calculating the fault prognosis horizon.

The prognosis results with different potential prognostic horizons are shown in Figure 5.5. The x-axis is the prognostic horizon in days, the y-axis is the number of pitch corrective maintenance activities. Each data group is for the proposed thresholds (T) and window sizes (W). The Undetected showing in graph legend is the number of undetected pitch corrective maintenance activities, out of 487. Figure 5.5 clearly shows that the proposed approach giving a significant warning of pitch faults with a long prognostic horizon up to 21 days, depending on the potential Prognostic Horizon, Window Size and Threshold.

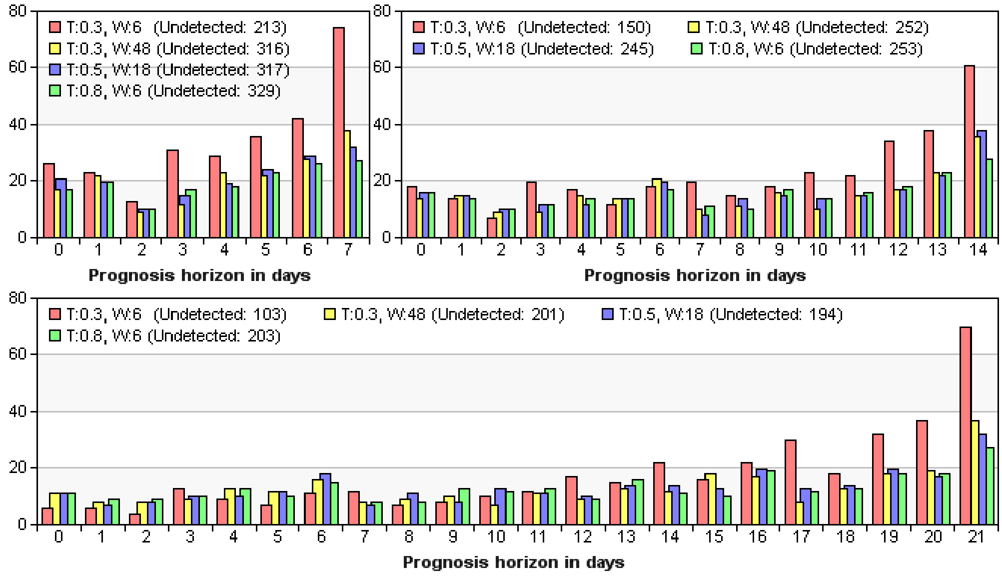


Figure 5.5: Plot of distribution of APK-ANFIS prognosis horizon in days with different potential prognostic horizons 7, 14 and 21 days. (T stands for Threshold and W stands for Window Size).

5.2.3. Fault Prognosis using SCADA Alarms

As the priority of SCADA alarm is unclear, a common approach for identifying WT pitch faults is to count the number of alarms during a certain period of time (Qiu et al. 2012). As long as the number of alarms is less than a defined threshold, the situation can be considered safe. Conversely, a possible fault is identified when the number of alarms is larger than the threshold and operators should start investigating the problem. A study using this approach to examine the efficiency of SCADA pitch alarms for fault prognosis was applied to the same WF. The threshold was taken as the average number of SCADA pitch alarms per day. At the beginning of the testing a number of thresholds were considered as follows: 2, 5, 10 and 15. Three potential prognostic horizons of 7, 14 and 21 days, were applied to avoid false identification. An algorithm, shown in Table 5.2, was also written to calculate the fault prognosis using above approach.

```
Step 1:
Data Cleansing – remove data when it has maintenance;
Step 2:
Define H, T to represent Prognostic Horizon, Threshold respectively
Declare H = 7, 14 or 21; T = 2, 5, 10 or 15
For each WT in the WF
    For each "pitch corrective maintenance record" in the selected WT
        Within the given Potential_Horizon = H days
            Find the earliest date when Total_No_of_Pitch_Alarm >= T
                Prognosis_Day = Maintenance_date – The_Earliest_date
End
```

Table 5.2: Pseudo-code for calculating the fault prognosis horizon using SCADA Alarms.

The prognosis results with different potential prognostic horizons are shown in Figure 5.6. As large numbers of detections were found close to 0 which is the day of conducting the corrective maintenance, this has demonstrated that the Alarm approach gives very little prognostic horizon.

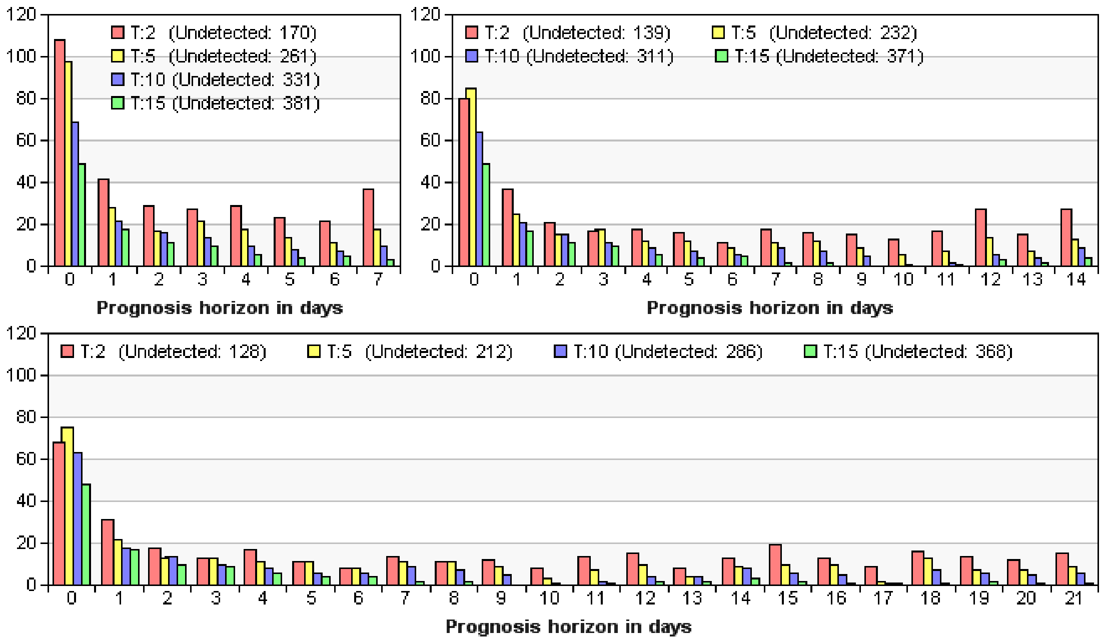


Figure 5.6: Plot of distribution of SCADA Alarms prognosis horizon in days with different potential prognostic horizons 7, 14 and 21 days.

By comparison between Figures 5.5 and 5.6 it is clear that the proposed approach with the APK-ANFIS using SCADA pitch signals gives prognostic warning of faults ahead of SCADA pitch alarms.

5.2.4. Confusion Matrix Analysis

Sections 5.2.2 & 5.2.3 have demonstrated the proposed approach gives prognostic warning of pitch fault ahead of pitch alarms. In this section, a Confusion Matrix analysis was generated to demonstrate the accuracy of the proposed approach.

The Confusion Matrix (Witten et al. 2011) contains information about actual and predicted diagnosis done by the proposed system and it is defined as follows:

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">Predicted</td></tr>
<tr><td>Needs Maintenance</td><td>No Maintenance</td></tr>
<tr><td rowspan="2">Actual</td><td>Had Maintenance</td><td>TP</td><td>FN</td></tr>
<tr><td>No Maintenance</td><td>FP</td><td>TN</td></tr>
</table>

- **True Positive (TP):** actual maintenance correctly predicted;
- **False Positive (FP):** incorrectly predicted as Needs Maintenance;
- **False Negative (FN):** incorrectly predicted as No Maintenance;
- **True Negative (TN):** correctly predicted as No Maintenance;

An algorithm was written to count TP, TN, FP and FN for every pitch corrective maintenance activity in the testing WF. The Pesudo-code is shown in Table 5.3.

Define H to represent Prognosis Horizon
Declare H = 7, 14 or 21 days;
For each WT in the WF
    For each "pitch corrective maintenance record" in the selected WT, marked as $\mathbf{M_m}$
        If the interval between $\mathbf{M_m}$ and the previous $\mathbf{M}$ is > H days, as shown in Figure 5.7(a).
            Range 1 and *Range 2* are found for the Maintenance $\mathbf{M_m}$.
                1) Within *Range 2*, results from the first positive to $\mathbf{M_m}$ are marked as TP;
                2) Remainder results in *Range 2* are marked as FN;
                3) Any positive result in *Range 1* are marked as FP;
                4) Any negative result in *Range 1* are marked as TN;
        Else if the interval between $\mathbf{M_m}$ and the previous $\mathbf{M}$ is <= H days, as shown in Figure 5.7(b).
            Range 2 is found for the Maintenance $\mathbf{M_m}$.
                1) Within *Range 2*, results from the first positive to $\mathbf{M_m}$ are marked as TP;
                2) Remainder results in *Range 2* are marked as FN;
End

Table 5.3: Pesudo-code for counting count TP, TN, FP and FN.

The two different situations are described in Figure 5.7:

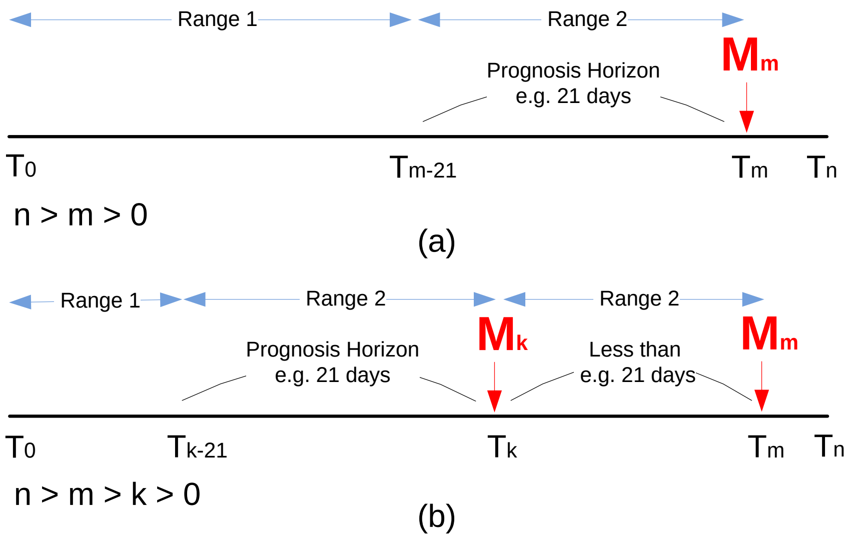


Figure 5.7: Two different situations to count TP, TN, FP and FN.

- **Figure 5.7 (a):** Maintenance M_m is far away from the previous maintenance and a Range 1, which represents the period from the end of previous maintenance to the beginning of the current potential Prognosis Horizon (Range 2), can be found;

- **Figure 5.7 (b):** Maintenance M_m is close to the previous maintenance M_k with the interval less than potential Prognosis Horizon. In this case, Range 1 is not available for the Maintenance M_m.

In addition, a further in-depth analysis of the data is presented utilising (Witten et al. 2011):

- Accuracy (ACC) is the proportion of the total number of predictions that are correct. This is one of the key aspects to determine the success of this approach.
- Error rate (ER) is the proportion of the total number of predictions that are wrong. Usually, $ER = 1 - ACC$.
- Recall (RC) is the proportion of actual maintenance cases that are predicted as positive. This value need to be high because an undetected failure might result in a catastrophic fault.
- Precision (P) is the proportion of the predicted positive cases that are truly positive. This value need to be as high as possible in order to void the additional cost caused by false maintenance request.
- F-measure (F) is a trade-off between precision and recall. It has been widely applied to identify the optimal setting of a classification system.

These are defined as follows:

$$ACC = \frac{TP + TN}{TP + FP + TN + FN}$$

$$ER = \frac{FP + FN}{TP + FP + TN + FN}$$

$$RC = \frac{TP}{TP + FN}$$

$$P = \frac{TP}{TP + FP}$$

$$F = \frac{2 * P * RC}{P + RC}$$

The Confusion Matrix analysis results of the proposed approach applied to the tested WF are shown in Table 5.4.

	ACC	*ER*	*RC*	*P*	*F*
T:0.3 WS:6	88.3%	11.7%	37.0%	76.4%	49.9%
T:0.3 WS:48	86.0%	14.0%	22.6%	66.1%	33.7%
T:0.5 WS:18	86.4%	13.6%	21.2%	72.8%	32.8%
T:0.8 WS:6	86.6%	13.4%	19.6%	79.3%	31.4%

Potential Prognostic Horizon = 7 days

	ACC	*ER*	*RC*	*P*	*F*
T:0.3 WS:6	85.1%	14.9%	48.2%	89.2%	62.6%
T:0.3 WS:48	80.6%	19.4%	30.7%	83.9%	45.0%
T:0.5 WS:18	81.0%	19.0%	30.6%	88.5%	45.5%
T:0.8 WS:6	81.0%	19.0%	29.0%	91.9%	44.1%

Potential Prognostic Horizon = 14 days

	ACC	*ER*	*RC*	*P*	*F*
T:0.3 WS:6	85.9%	14.1%	62.2%	94.4%	75.0%
T:0.3 WS:48	79.4%	20.6%	43.3%	92.1%	58.9%
T:0.5 WS:18	79.3%	20.7%	41.8%	94.4%	58.0%
T:0.8 WS:6	78.9%	21.1%	39.4%	96.2%	55.9%

Potential Prognostic Horizon = 21 days

Table 5.4: Confusion matrix analysis results with different potential prognosis horizons.

The table shows the high accuracy and precision of the proposed approach. It also can be seen that the precision is increase with the prognostic horizon out to 21 days, whilst the accuracy falls slightly. In addition, recall was improved greatly along with the increase of the potential prognostic horizon. Finally, the 21 days potential prognostic horizon is found reasonable as the error rate doesn't increase very much with the Recall, Precision and F-measure are improved greatly. The optimal Threshold and Window Size are 0.3 and 6 respectively in terms of Accuracy, Recall and F-measure. However, in terms of Precision, the optimal Threshold and Window Size are 0.8 and 6 respectively.

5.2.5. Result Conclusion

From the above results we can draw the following conclusions:

- The proposed approach using an APK-ANIFS on SCADA pitch signals gave significant warning of pitch faults with a prognostic horizon up to 21 days, depending on the window size and threshold;
- SCADA pitch alarms also detected pitch faults but counting them gave very little or no prognostic horizon of impending pitch faults;
- Confusion Matrix analysis of the SCADA signal analysis has shown that regardless of window size and threshold the precision of prediction increases the prognostic horizon out to 21 days, whilst the accuracy of detection falls slightly;
- These results all suggest that whilst SCADA alarm analysis may help to identify pitch fault root causes they cannot predict faults, whereas SCADA signal analysis using APK-ANFIS gives good prediction with a prognostic horizon up to 21 days, a valuable period for WF operators to repair notified pitch faults.

5.3. Test on Mitsubishi WTs

The proposed method has also been applied to Mitsubishi WTs of different technology to Alstom WTs also utilising a different SCADA system to collect data. Parts of this work were delegated to two Durham Master's students from January-August 2013 (Norevik 2013; Xie 2013).

5.3.1. Brazos Wind Farm and the Available Data

The Brazos WF is located in Borden and Scurry counties in Texas, US (Wikipedia 2013), as shown in Figure 5.8. It has 160 Mitsubishi 1000 WTs, each rated at 1MW fixed speed, variable pitch with hydraulic pitch-to-stall control. The WF project was completed in December 2003 supplying approximately 30,000 homes.

(a) (b)

Figure 5.8: Brazos Wind Farm (a) View of Wind Farm; (b) Location; Source from (Wikipedia 2013)

The data consists of SCADA signals, Maintenance Log (Monthly Reports), Met Mast data and WT locations. Two different SCADA systems were installed in this WF, named Type 1 and Type 2, shown in Table 5.5. All Mitsubishi WTs' historical SCADA data were available from 31/May/2004 to 24/Nov/2006.

Type	Installed at Brazos in WTs	Data Description
1	WTs A12-18 and A33	• 10 minutes interval; • 60 channels; • No alarm data;
2	Rest of WTs	• 10 minutes interval; • 60 channels; • Has alarm event data, recorded as EventDefnID and EventDesc;

Table 5.5: SCADA systems in Brazos WF.

Most of the general signals, such as wind speed, blade angle, rotor speed, power output etc., are provided with Max, Min, Mean and Standard Deviation value by Mitsubishi WT's SCADA system. However, by comparison to the Alstom SCADA system, the Mitsubishi SCADA system does not record the blade torque or ram force signal, which is one of the most valuable WT pitch system signals. In addition, the Mitsubishi SCADA system records overall blade angle, assumed aggregated from the all three blade angles, whereas the Alstom SCADA system records each blade angle signal separately. Moreover, no independent SCADA alarms are available in Mitsubishi WT SCADA Type 1, but the Mitsubishi SCADA Type 1 system has alarm information to indicate the WT state, which is called event and measured every 10 minutes. By analysing the Mitsubishi SCADA Type 1 data of 152 WTs over 2 years, 183 identical events or alarms were found. Among them, 16 were relevant to pitch system, as shown in Table 5.6.

EventDefnID	**EventDesc**
1595	F04: Blade pitch angle signal fault
1603	F12: Blade pitch control error large
1604	F13: Blade pitch control fault
1605	F14: Pitch transducer fault
1615	F24: Pitch hunt Fault
1640	F49: Blade pitch slow
1641	F50: Actual pitch error large
1673	F82: Blade pitch failure at the test mode
1705	A14: Pitch Span Data Illegal
1715	A24: Pitch hunt Alarm
1726	A35: Pitch angle diff. large (auto stop)
1759	Blade Realignment
1779	Blade Repair
1791	Blade Inspection
1825	Lightning Damage Blade
1832	A14:Pitch Span Data Illegal

Table 5.6: Pitch Alarms in Mitsubishi WT.

The maintenance log for Mitsubishi WTs was reported monthly and saved in PDF format. The monthly report has a list of the top non-manufacturer and manufacturer downtimes. However, the records are unclear, and don't show the exact start and end date for each individual maintenance activity, increasing the difficulty of researching the cause and effect of Mitsubishi pitch unreliability.

Met Mast and WT location data were available from Mitsubishi WTs. They are very useful as the power generation can be estimated from the Met Mast data and the performance of a WT can be checked against neighbouring WTs. However, these tasks are not included in this study and will be considered as further works.

5.3.2. Data Preparation & Selection

This study intended to apply the proposed APK-ANFIS approach on the pitch data from Mitsubishi WTs to confirm the effectiveness of the proposed approach on a different WT and SCADA system data. As has been noted in Section 5.3.1, the Mitsubishi SCADA System does not record blade torque or ram force signal, therefore only 3 CCFs can be considered for Mitsubishi WTs rather than the 4 CCFs for Alstom WTs. The 3 CCFs are power curve, rotor speed curve and pitch angle curve.

The data selection and testing process in this research relies on the maintenance logs, but Mitsubishi WT maintenance logs are unclear. Thus an approach needed to be developed to identify the exact maintenance period for each individual maintenance activity. By analysing monthly reports, two downtime categories were found useful, however only WT manufacturer downtime is relevant to this research, as shown in Figure 5.9. By looking in detail at the corresponding SCADA data, an approach is proposed by the author to identify pitch maintenance period, as described below and shown in Figure 5.10

- By searching the keyword "Pitch" in the EventDesc, "Manufacturer" in the DTDesc and "Maintenance" in the AvDesc, as shown in Figure 5.9;

Top 15 Manufacturer Downtime Categories – September

Fault and/or Downtime Definition	MWh	% Total	Count	% Total	Hours	% Report Time
F49: Blade pitch slow	179.57	25.18	87.00	10.81	208.83	0.18
Spindle Failure	133.47	18.71	4.00	0.50	209.67	0.18
F50: Actual pitch error large	51.33	7.20	16.00	1.99	59.50	0.05
High Speed Shaft Movement	40.07	5.62	4.00	0.50	92.50	0.08
F04: Blade pitch angle signal fault	32.90	4.61	4.00	0.50	86.17	0.07
F36: L.O. pressure low	27.40	3.84	7.00	0.87	35.00	0.03
F48: Mechanical brake pads worn or pressure lost	23.69	3.32	7.00	0.87	83.33	0.07
F22: Software over speed 115%	22.69	3.18	4.00	0.50	29.17	0.03
Lightning	17.30	2.43	7.00	0.87	77.17	0.07
F82: Blade pitch failure at the test mode	16.12	2.26	5.00	0.62	31.67	0.03
F42: Mechanical brake ON	14.26	2.00	30.00	3.73	89.50	0.08
F32: G.O. pressure low	13.37	1.87	10.00	1.24	44.67	0.04
F73: Mechanical Brake OFF Failure	12.54	1.76	9.00	1.12	25.67	0.02
Hub/Spindle Inspections	11.48	1.61	46.00	5.71	38.33	0.03
F66: MCCB-1 trip	11.09	1.55	7.00	0.87	17.83	0.02

Figure 5.9: Mitsubishi WT manufacturer downtime

SiteID	NodeID	Turbine	TStampOfMidPoint	EventDefnID	EventDesc	DownTimeID	DTDesc	AvailCatID	AVDesc
9	1	A01	2006-04-26 11:05:00.000	1	No fault code	0	No Downtime	4	On line
9	1	A01	2006-04-26 11:15:00.000	1	No fault code	0	No Downtime	4	On line
9	1	A01	2006-04-26 11:25:00.000	1	No fault code	0	No Downtime	7	Released to run - wind in limits
9	1	A01	2006-04-26 11:35:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	9	Fault - wind in limits
9	1	A01	2006-04-26 11:45:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	9	Fault - wind in limits
9	1	A01	2006-04-26 11:55:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	9	Fault - wind in limits
9	1	A01	2006-04-26 12:05:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	9	Fault - wind in limits
9	1	A01	2006-04-26 12:15:00.000	1	No fault code	0	No Downtime	7	Released to run - wind in limits
9	1	A01	2006-04-26 12:25:00.000	1	No fault code	0	No Downtime	4	On line
9	1	A01	2006-04-26 12:35:00.000	1	No fault code	0	No Downtime	4	On line
9	1	A01	2006-04-26 12:45:00.000	1	No fault code	0	No Downtime	4	On line
9	1	A01	2006-04-26 12:55:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 13:05:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 13:15:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 13:25:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 13:35:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 13:45:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 13:55:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 14:05:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 14:15:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 14:25:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 14:35:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 14:45:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 14:55:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 15:05:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 15:15:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 15:25:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 15:35:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 15:45:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 15:55:00.000	1604	F13: Blade pitch control fault	1	Manufacturer	11	Maintenance - wind in limits
9	1	A01	2006-04-26 16:05:00.000	1	No fault code	0	No Downtime	4	On line

consecutive

Maintenance

Figure 5.10: Approach used to find the exact pitch maintenance period

Using above approach, 5 typical pitch faults were found and shown in Table 5.7 and they will be used for the training dataset. In the actual data selection process, SCADA noises signals, such as "release to run" and "WT shut down" were removed to avoid false identifications. A labelling procedure, mentioned in Section 4.2.3, was applied to the 5 known pitch faults. By applying the labelling procedure for all 5 known pitch faults and aggregating them together, 4,790 sets of data were collected.

WT	Case	Generating Fault	Maintenance	After Maintenance
A01	Case 1	01/04/2006 ~ 26/04/2006	26/04/2006 ~ 27/04/2006	27/04/2006 ~ 05/05/2006
A19	Case 2	15/05/2005 ~ 05/06/2005	05/06/2005 ~ 10/07/2005	10/07/2005 ~ 15/07/2005
A21	Case 3	01/01/2005 ~ 10/01/2005	10/01/2005 ~ 10/01/2005	10/01/2005 ~ 16/01/2005
A24	Case 4	18/06/2004 ~ 01/07/2004	01/07/2004 ~ 02/07/2004	02/07/2004 ~ 08/07/2004
A26	Case 5	01/06/2005 ~ 09/06/2005	09/06/2005 ~ 09/06/2005	09/06/2005 ~ 13/06/2005

Table 5.7: Five Mitsubishi pitch fault cases.

5.3.3. Training & Training Result

A similar training procedure was applied to this Mitsubishi study as was applied to the Alstom study. The generalised bell MF was selected with giving the minimum value 0.01 and maximum iteration 200. In order to find the optimal structure for each individual APK-ANFIS, a batch test with different numbers of MFs for each input were examined. Finally, the optimal structures were chosen, as shown in Table 5.8:

APK-ANFIS model	Optimal Structure
Wind Speed vs. Rotor Speed	3-by-3
Wind Speed vs. Blade Angle	3-by-3
Wind Speed vs. Power Output	3-by-3

Table 5.8: The optimal APK-ANFIS structure.

After that, the data are partitioned into training and testing sets. Cases 1-4 provided the training data and Case 5 was used to test the trained model. Its success at actual outputs that are as close as possible to the desired outputs determines how well the APK-ANFIS has learned or captured the relations between the inputs and outputs. Finally, the output surfaces generated by individual trained APK-ANFIS models are shown in Figure 5.11. The surface chart clearly demonstrates that abnormal data will give a large output, close to 1 and shown as "Hill", while normal data will give a small output, close to 0 and shown as the "Valley".

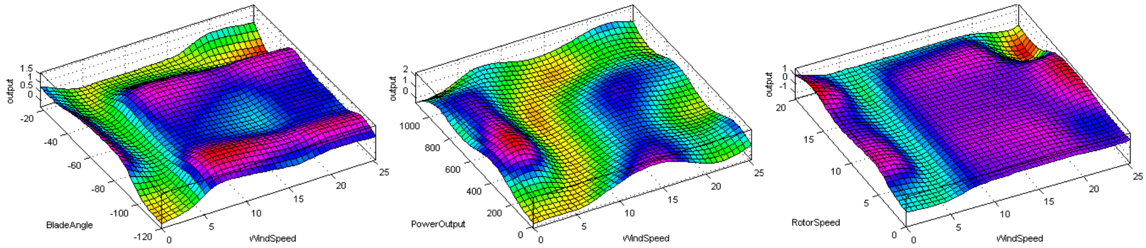


Figure 5.11: Output surfaces generated from the trained APK-ANFIS models.

5.3.4. Fault Prognosis using Proposed Approach

The trained system was tested on the pitch data from the other 22 Brazos WTs to exam its fault prognosis ability. In the data selection procedure of the 22 WTs, SCADA noises signals were removed to avoid false identifications. During the procedure, almost one sixth of the data were found to be "Release to run" and most of them with good wind speed. We conjectured that this is due to the low demand from the grid and high wind power availability so WF operators have curtailed their WTs.

For simplicity, the potential Prognosis Horizon was given 21 days as it is likely to produce a better result. Window Size 3 and 6, represent 0.5 and 1 hour intervals respectively, were chosen and the corresponding Threshold was given as follows:

Window Size	Threshold
3	0.5
6	0.3

Note that larger applied Window Size and Threshold would produce better warning but were impractical for this data.

A similar algorithm to Table 5.1 was applied to Mitsubishi pitch data from the 22 WTs. Finally, the prognosis results are shown in Figure 5.12. The result shows that the proposed method does not give significant pitch fault warning. However, the result still demonstrates that the proposed approach can be used for WT pitch fault detection, even on a WT of different technology and a different SCADA system.

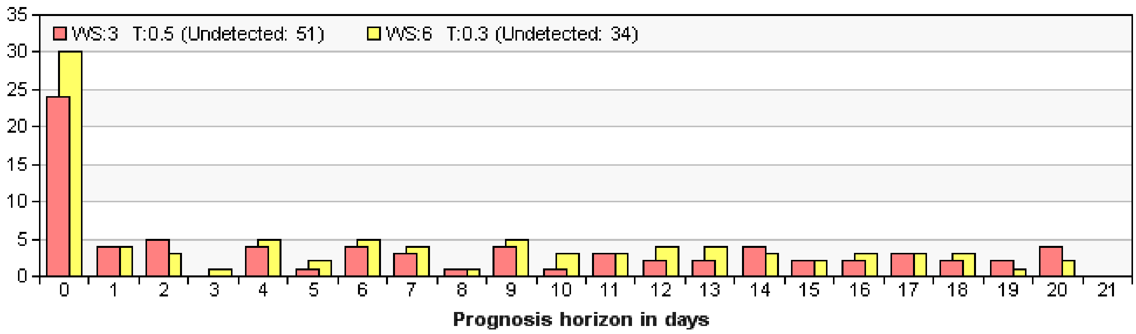


Figure 5.12: Plot of distribution of SCADA Signals prognosis horizon in days

5.3.5. Confusion Matrix Analysis

The Confusion Matrix analysis was used to evaluate the fault prognosis results. The method, as described in Section 5.2.4, was applied and the results are shown in Table 5.9. According to Table 5.9, the two results are very close, but the result from Window Size 3 and Threshold 0.5 is better as it has higher accuracy, recall and precision. In general, the evaluation using Confusion Matrix analysis has demonstrated the proposed approach gives high accuracy and precision. However, by checking against to the number of pitch faults, we found the high accuracy and precision is largely contributed by normal data as the Mitsubishi WT experiencing less pitch faults.

	ACC	*ER*	*RC*	*P*	*F*
T:0.5 WS:3	91.5%	8.5%	34.0%	96.0%	50.2%
T:0.3 WS:6	91.2%	8.8%	32.0%	91.0%	47.3%

Table 5.9: Confusion Matrix analysis results with Potential Prognostic Horizon = 21 days.

5.3.6. Result Conclusion

In this study, a similar system was built and tested on Mitsubishi WTs to show the generalised effectiveness of the proposed APK-ANFIS approach. However, the results from Mitsubishi WTs do not give significant pitch fault warnings. Although higher accuracy was obtained from the Confusion Matrix analysis, this is due to the Mitsubishi WTs experiencing less pitch faults and accuracy is being contributed mostly by normal data.

5.4. Comparison of Alstom & Mitsubishi WT Results

5.4.1. Prognostic Horizon Results

The APK-ANFIS approach has been applied to pitch data from both Alstom & Mitsubishi WTs. The results are shown in Table 5.10 and plotted in Figure 5.13, where the effective prognostic horizons of the two methods can be seen.

	Day	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Mitsubishi	WS: 3; T: 0.5	24	4	5	0	4	1	4	3	1	4	1	3	2	2	4	2	2	3	2	2	4	0
	WS: 6; T: 0.3	30	4	3	1	5	2	5	4	1	5	3	3	4	4	3	2	3	3	3	1	2	0
Alstom	WS: 6; T: 0.3	6	6	5	13	9	7	11	12	7	8	10	12	17	15	22	16	22	30	18	32	37	70
	WS: 48; T: 0.3	11	8	8	9	13	12	16	8	9	10	7	11	9	13	12	18	17	8	13	18	19	37
	WS: 18; T: 0.5	11	7	8	10	10	12	18	7	11	8	13	11	10	14	14	13	20	13	14	20	17	32
	WS: 6; T: 0.8	11	9	9	10	13	10	15	8	8	13	12	13	9	16	11	10	19	12	13	18	18	27

Table 5.10: Prognostic Horizon Results. Numbers represent the number of detected pitch maintenance activities.

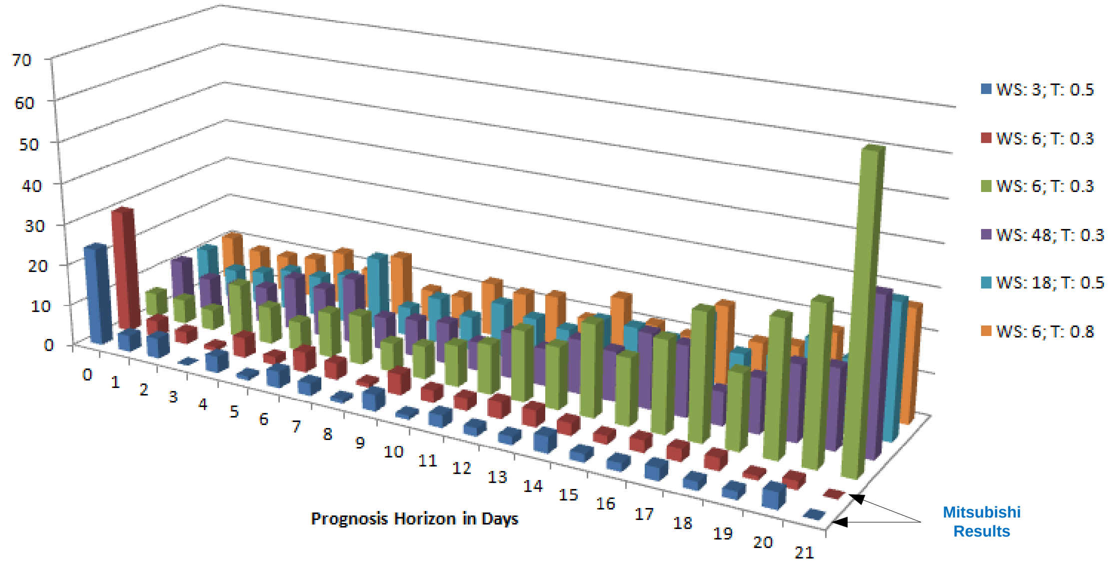


Figure 5.13: Prognostic Horizon Comparison

Figure 5.13 shows that APK-ANFIS gives a good fault prognosis horizon from the Alstom data but not from Mitsubishi data. By reviewing this study, we believe Mitsubishi data has the following difficulties, which have a big impact to the prognosis results:

- The Mitsubishi Maintenance Log is unclear having a big impact on data selection. The Monthly Reports did not give exact start and end dates for each corrective maintenance;
- The Mitsubishi SCADA System does not record blade torque or ram force signal, which is one of the most valuable WT pitch system signals.
- About one sixth of the Mitsubishi WT data used for Fault Prognosis testing shows "Release to Run" with good wind speed but WTs not operating, probably due to curtailments at times of low grid demand.
- However, the APK-ANFIS worked satisfactorily on an entirely different WT pitch technology.

5.4.2. Confusion Matrix Results

The Confusion Matrix analysis results are shown in Table 5.11 below:

		ACC (%)	ER (%)	RC (%)	P (%)	F (%)
Mitsubishi	WS: 3; T: 0.5	91.5	8.5	34.0	96.0	50.2
	WS: 6; T: 0.3	91.2	8.8	32.0	91.0	47.3
Alstom	WS: 6; T: 0.3	85.9	14.1	62.2	94.4	75.0
	WS: 48; T: 0.3	79.4	20.6	43.3	92.1	58.9
	WS: 18; T: 0.5	79.3	20.7	41.8	94.4	57.9
	WS: 6; T: 0.8	78.9	21.1	39.4	96.2	55.9

Table 5.11: Confusion Matrix analysis results.

The Mitsubishi data shows higher accuracy than Alstom because the Mitsubishi WTs experienced less pitch faults and therefore accuracy is contributed primarily by normal data.

5.5. Chapter Summary

This chapter has described the methodology used to apply the proposed approach to the pitch data from two different designs of WTs, manufactured by Alstom & Mitsubishi, with two different types of SCADA system, demonstrating the adaptability of APK-ANFIS for application to variety of technologies. The results were further evaluated by Confusion Matrix analysis to check the validity of the results. A comparison study of Alstom & Mitsubishi results was also made.

The results from Alstom WTs have shown significant warning of pitch faults with a long prognostic horizon up to 21 days, depending on window size and threshold. However, results from Mitsubishi WTs did not give good fault prognosis horizon, but still showed strong fault detection ability. We believed the results from the Alstom WTs proved more convincing than from the Mitsubishi machines. This was primarily because of poor clarity in the

Mitsubishi Maintenance records, lack of an important pitch signal and the effects of curtailment, rather than noise in the variable-speed pitch-to-feather/stall control. Both Confusion Matrix analysis results have shown high accuracy and precision. Among them, the Mitsubishi results showed higher accuracy because the Mitsubishi WTs experiencing less pitch faults and therefore accuracy is being contributed mostly by normal data.

The next chapter discusses how the proposed system meets the aim of this research and lists the advantages of the proposed approach. A summary of novel contributions from this research and the further works are also presented.

Chapter 6.

Discussion, Conclusions and Further Works

6.1. Discussion

6.1.1. Meeting the Research Aim

As has been stated in Chapter 1, the aim of this research was to develop an automated on-line fault prognosis system for WT monitoring using SCADA data. This objective has been achieved in two main areas.

The first area is the development of the mechanisms used to interpret the raw SCADA data. This objective is achieved by using APK-ANFIS and the four CFs. The use of APK-ANFIS has enabled the system to inherit the interpretability presented in FIS. Therefore, any observed numerical data can be transformed into linguistic and heuristic terms, which are normally expressed in a form of an if-then rule, for example "If WT power output is high and wind speed is low, then a possible fault is detected". The four CFs, as described in Section 4.2.1, reflect the physical properties of a running WT. In addition, each time, a new observation from SCADA system can also be interpreted through displaying the data on the output surface of the system, for example an observation in Table 6.1 is shown in Figure 6.1.

Wind Speed (m/s)	Rotor Speed (rpm)	Blade Angle (degree)	Motor Torque (kN)	Power Output (kW)
15	19	14	38	1680

Table 6.1: An example of an observation from the SCADA system

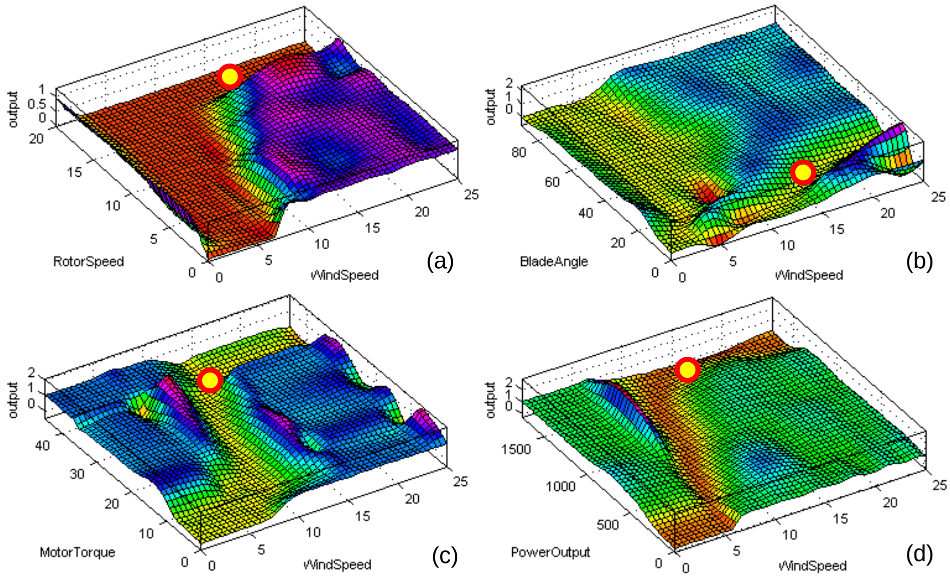


Figure 6.1: An observation, as shown in Table 6.1, is displayed on the output surface of the proposed system.

The second area is the automation of the on-line fault identification. As described in Section 4.2.2, the proposed fault diagnosis procedure consists of 4 modules: *Data Acquisition*, *Feature Extraction*, *Multiple Diagnosis* and *Fault Diagnosis Result*. Each of them is automated, so that the whole system should have the ability to work automatically. In addition, the training procedure of this proposed system takes quite a long time, but once the system is trained, each module doesn't need much computational cost. In other word, the inputs variable of the trained system can be taken in real-time and a diagnosis output will be obtained in real-time too.

6.1.2. Advantages of the Proposed Approach

SIMAP (Garcia et al. 2006), the Venn diagram (Qiu et al. 2012), Alarm Pattern Recognition (Chen et al. 2011), data-driven (Kusiak and Verma 2011) and normal behaviour models (Schlechtingen et al. 2012) approaches were found to be similar to this proposed approach. Compared to them, the proposed approach has shown the following advantages:

- **Better interpretability:** The latest development of ANFIS, the APK-ANFIS is also a hybrid system that contains the advantages of both ANN and FIS; therefore the proposed approach will inherit the interpretability present in FIS.
- **Better rationalisation of the data:** This is because the four CCFs, as mentioned in Figure 4.4, reflect the physical properties of a running WT.
- **Incorporation of domain knowledge:** The latest developments of ANFIS allow experts to introduce domain knowledge into the ANFIS training procedure and it has better interpretability for the never-seen input conditions.
- **Move convincing prognosis result:** The prognosis result is more convincing because this approach has been applied to data from two different designs of WTs. In addition, the Alstom results were also compared to an alarm approach to show its advantages.
- **More feasible online fault prognosis:** The input variables of this proposed approach are taken in real-time and a prognosis output is obtained in real-time too, as shown in Figure 4.5 & 4.13.

6.2. Conclusions

The monitoring of WT can allow prevention of downtime, a longer operational life and a reduction in the cost of energy over WT lifetime. As pitch system is a vital part of the modern variable speed WT and few successful WT pitch fault detection systems were found in the literature, this research was focused on analysing WT pitch faults with the objective of developing an automated on-line fault detection approach. The author only considered AI approaches since AI techniques have increasingly been used in the field of FDD and the other analytical CM approaches were already being considered by other Durham University researchers.

This thesis has investigated a number of AI techniques, including supervised and unsupervised learning, along with some examples detailing applications of how they can be utilised in the field of WT FDD. In the end of this investigation, the APK-ANFIS was selected to research in further as it has better interpretability and allows domain knowledge to be incorporated. After that, a fault prognosis procedure using APK-ANFIS and four CCFs was proposed. The data of the six known WT pitch faults were labelled and used to train the proposed system with a-priori knowledge incorporated.

The proposed approach using APK-ANFIS have been applied to the data from two different designs of WTs, manufactured by Alstom & Mitsubishi, with two different types of SCADA system, demonstrating the adaptability of APK-ANFIS for application to variety of technologies. The results from Alstom WTs have shown significant warning of pitch faults with a long prognostic horizon up to 21 days, depending on window size and threshold. However, results from Mitsubishi WTs did not give good fault prognosis horizon, but still showed strong fault detection ability. The author believes the results from Alstom WTs proved more convincing than from the Mitsubishi WTs. This was primarily because of poor clarity in the Mitsubishi Maintenance record, lack of an important pitch signal and the effects of operational curtailment.

In addition, the Alstom result was compared to a common alarm approach applied on the same data. The comparison result suggested that whilst SCADA alarm analysis may help to identify pitch fault root causes they cannot predict faults, whereas SCADA signal analysis using APK-ANFIS give good prediction. Moreover, both results from Alstom and Mitsubishi WTs were analysed using

Confusion Matrix analysis. The analysis result showed that the proposed approach gave high accuracy and precision. Among them, the Mitsubishi results showed higher accuracy because Mitsubishi WTs experienced less pitch faults and therefore accuracy is being contributed mostly by normal data.

In summary, the novel contributions delivered by the research are:

- This research has introduced a fault diagnosis model using AI technique and demonstrated that the proposed approach gives prognostic warning of pitch faults up to 21 days.
- The robustness and effectiveness of this approach have been demonstrated by:
  - Applying the proposed approach to pitch data from two different designs and locations of WTs.
  - Results were evaluated using Confusion Matrix analysis to show the validity.
- Considerable larger sizes of WT data were used in this research, compared to previous work. There were 26 Alstom WTs with 63 WT-year data and 22 Mitsubishi WTs with 53 WT-year data.
- Online fault diagnosis is possible if the input variables of this proposed approach are taken in real-time and a diagnosis output is obtained in real-time too.
- In addition, the robustness of the system was improved by the strong interpretability of the fault diagnosis model from two different aspects:
  - Domain knowledge incorporation: APK-ANFIS allows experts to introduce domain knowledge to the system model.
  - Rationalisation of the Data: four CCFs, as mentioned in Section 4.2.1, reflect the physical properties of the running WT.

In conclusion, this research has presented a fault diagnosis model using APK-ANFIS and demonstrated that the proposed APK-ANFIS approach gives prognostic warning of pitch faults up to 21 days. The SCADA signal analysis using APK-ANFIS has strong potential to provide automated online WT fault detection and prognosis.

6.3. Further Works

The following areas have been identified as possible works for further research.

6.3.1. Improved APK-ANFIS for Curtailed Situations

In real operation, the WF operators have to curtail the WTs' power output if there is a low grid demand. This manual intervention will result in some WT's not performing to their factory supplied specification. In this situation, the possible faults are much more difficult to be identified. For simplicity, the proposed approach currently does not consider this situation and the test on Mitsubishi WTs has shown the impact of this problem. Therefore, it would be beneficial if the proposed APK-ANFIS approach could be improved to consider its operation under the effect of WT curtailment.

6.3.2. Modular Architecture

It would be beneficial to have a completed APK-ANFIS system with tasks organised in a modular architecture, as presented in SIMAP. Four modules are suggested by the author as follows:

- A pitch Health Condition Assessment Module could be added to evaluate on-line health condition of WT pitch systems;

- A diagnosis Expert Module could be developed to identify the possible failure modes;
- A predictive Maintenance Scheduling Module could be added with the goal of scheduling WT maintenance actions optimally;
- A maintenance Effectiveness Assessment Module could be developed to measure the effectiveness of each applied maintenance action.

In addition, the WT converters should be studied in the way set out for pitch systems in this thesis and organised as a converter health condition assessment module. It was known from ReliaWind project (Wilkinson et al. 2010) that the pitch system & converters were major fault items. The converter contains many similarities with the electric pitch system since it also consists of a power electronic converter, rich in SCADA alarm and signal data. It is also known that in offshore WTs the convert is a significant source of downtime, primarily due to the logistic delay in attending many minor electronic defects.

6.3.3. More Data and Test on more Modern WTs

For the research carried out in this thesis, only four CFs (five signals in total) were applied and 5-6 pitch fault cases were studied. It would be beneficial from machine learning point of view to have more related signals and data to be introduced into the model and this is under discussion with our industrial partners DONG Energy. However, over-fitting needs to be solved if too many signals and data are used.

In addition, in order to demonstrate the adaptability of the proposed approach, more testing on different designs and locations of modern WTs are necessary. The use of data from different locations would serve the purpose of determining whether the approach can adjust to different environmental conditions. Data from different designs of WT would serve the purpose of determining how robust the proposed approach is to do the fault prognosis.

References

Abonyi, J., Babuska, R., Setnes, M., Verbruggen, H. and Szeifert, F. (1999). Constrained parameter estimation in fuzzy modeling. *Fuzzy Systems Conference Proceedings, 1999. FUZZ-IEEE'99. 1999 IEEE International*, IEEE.

Ackermann, T. and Söder, L. (2002). An overview of wind energy-status 2002. *Renewable and sustainable energy Reviews* **6**(1): 67-127.

Angeli, C. and Chatzinikolaou, A. (2004). On-Line Fault Detection Techniques for Technical Systems: A Survey. *IJCSA* **1**(1): 12-30.

Assunção, T., Silvino, J. and Resende, P. (2006). Transformer top-oil temperature modeling and simulation. *Trans. Eng. Comput. Technol* **15**: 240-245.

Bennett, K. P. and Campbell, C. (2000). Support vector machines: hype or hallelujah? *ACM SIGKDD Explorations Newsletter* **2**(2): 1-13.

Bezdek, J. C. (1981). Models for Pattern Recognition. Pattern Recognition with Fuzzy Objective Function Algorithms, *Springer*: 1-13.

Bianchi, F., De Battista, H. and Mantz, R. (2006). *Wind turbine control systems: Principles, modelling and gain-scheduling design (advances in industrial control)*.

Bikdash, M. (1999). A highly interpretable form of Sugeno inference systems. *Fuzzy Systems, IEEE Transactions on* **7**(6): 686-696.

Blanco, M. I. (2009). The economics of wind energy. *Renewable and Sustainable Energy Reviews* **13**(6): 1372-1382.

Brandão, R. M., Carvalho, J. B. and Barbosa, F. M. (2010). Neural networks for condition monitoring of wind turbines. *Modern Electric Power Systems (MEPS), 2010 Proceedings of the International Symposium*, IEEE.

Braun, J. E. and Herrick, R. W. (2003). "Automated Fault Detection and Diagnostics for Vapor Compression Cooling Equipment." 2011, from http://poet.lbl.gov/diagworkshop/proceedings/braun.htm.

Chandola, V., Banerjee, A. and Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys (CSUR)* **41**(3): 15.

Chen, B., Matthews, P. C. and Tavner, P. J. (2013). Wind turbine pitch faults prognosis using a-priori knowledge-based ANFIS. *Expert Systems with Applications* **40**(17): 6863-6876.

Chen, B., Qiu, Y., Feng, Y., Tavner, P. and Song, W. (2011). Wind turbine SCADA alarm pattern recognition. *Renewable Power Generation (RPG 2011), IET Conference on*, IET.

Chen, B., Tavner, P. J., Feng, Y., Song, W. W. and Qiu, Y. Bayesian Networks for Wind Turbine Fault Diagnosis.

Chen, B. and Zappala, D. (2011). "Survey of commercial SCADA analysis tool." Retrieved Sep, 2011, from http://www.supergen-wind.org.uk/.

Clarkson, M. (2010). A Study of Wind Turbine Structure and Data to Assist in the Development of a Reliability-Centred 10 MW Offshore Wind Turbine. School of Engineering and Computing Science, Durham University. **MEng Thesis**.

Crabtree, C. J. (2010). "Survey of commercially available condition monitoring systems for wind turbines." Retrieved Aug, 2011, from http://www.supergen-wind.org.uk/.

Crabtree, C. J. (2011). Condition Monitoring Techniques for Wind Turbine. School of Engineering and Computing Science, Durham University. **PhD Thesis**.

Darling, D. (2011). "The Encyclopaedia of Alternative Energy and Sustainable Living." Retrieved Jul, 2011, from http://www.daviddarling.info/encyclopedia/W/AE_wind_turbine.html.

Dunn, J. C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters.

Dvorak, P. (2009). "Hydraulic pitch control for wind turbine blades." Retrieved Feb, 2012, from http://www.windpowerengineering.com/design/mechanical/gearboxes/hydraulic-pitch-control-for-wind-turbine-blades/.

Faulstich, S., Hahn, B. and Tavner, P. (2011). Wind turbine downtime and its importance for offshore deployment. *Wind Energy* **14**(3): 327-337.

Feng, Y., Qiu, Y., Crabtree, C. J., Long, H. and Tavner, P. J. (2011). Use of SCADA and CMS signals for failure detection and diagnosis of a wind turbine gearbox. *EWEA*.

Feng, Y., Tavner, P. and Long, H. (2010). Early experiences with UK Round 1 offshore wind farms. *Proceedings of the Institution of Civil Engineers: energy*. **163**(4): 167-181.

Fletcher, T. (2009). Support vector machines explained. *Tutorial paper., Mar*.

Garcia, M. C., Sanz-Bobi, M. A. and del Pico, J. (2006). SIMAP: Intelligent System for Predictive Maintenance: Application to the health condition monitoring of a windturbine gearbox. *Computers in Industry* **57**(6): 552-568.

GOV.UK. (2009). "Report from Department of Energy & Climate Change, 2009a." Retrieved Jul, 2011, from https://www.gov.uk/government/organisations/department-of-energy-climate-change.

GOV.UK. (2011). "UK Renewable Energy Roadmap." Retrieved Sep, 2011, from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/48128/2167-uk-renewable-energy-roadmap.pdf.

Gray, C. S. and Watson, S. J. (2010). Physics of failure approach to wind turbine condition based maintenance. *Wind Energy* **13**(5): 395-405.

Hau, E. and Platz, H. (2006). *Wind Turbines-Fundamentals, Technologies, Application, Economics*, Springer.

Haykin, S. S., Haykin, S. S., Haykin, S. S. and Haykin, S. S. (2009). *Neural networks and learning machines*, Prentice Hall New York.

Isermann, R. (2006). *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*, Springer.

Isermann, R. and Balle, P. (1997). Trends in the application of model-based fault detection and diagnosis of technical processes. *Control engineering practice* **5**(5): 709-719.

Jang, J.-S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *Systems, Man and Cybernetics, IEEE Transactions on* **23**(3): 665-685.

Jang, J.-S. R., Sun, C.-T. and Mizutani, E. (1997). Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence [Book Review]. *Automatic Control, IEEE Transactions on* **42**(10): 1482-1484.

Kohonen, T. (2001). *Self-organizing maps*, Springer.

Korb, K. B. and Nicholson, A. E. (2003). *Bayesian artificial intelligence*, CRC press.

Korbicz, J. and Kowal, M. (2007). Neuro-fuzzy networks and their application to fault detection of dynamical systems. *Engineering Applications of Artificial Intelligence* **20**(5): 609-617.

Kosko, B. and Isaka, S. (1993). Fuzzy logic. *Scientific American* **269**(1): 62-7.

Krohn, S., Morthorst, P. E. and Awerbuch, S. (2009). *The economics of wind energy*, European Wind Energy Association.

Kusiak, A. and Li, W. (2011). The prediction and diagnosis of wind turbine faults. *Renewable Energy* **36**(1): 16-23.

Kusiak, A. and Verma, A. (2011). A data-driven approach for monitoring blade pitch faults in wind turbines. *Sustainable Energy, IEEE Transactions on* **2**(1): 87-96.

Lei, Y., He, Z., Zi, Y. and Hu, Q. (2007). Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs. *Mechanical Systems and Signal Processing* **21**(5): 2280-2294.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, California, USA.

MathWorks (2013). "Train support vector machine classifier." from http://www.mathworks.co.uk/help/stats/svmtrain.html.

Moorse, J. (2010). Analysis of SCADA Data from Large Wind Farms to Provide Incipient Fault Detection. School of Engineering and Computing Science, Durham University. **MEng Thesis**.

Mote, T. P., & Lokhande, D. S. (2012). Temperature Control System Using ANFIS. International Journal of Soft Computing and Engineering (IJSCE) ISSN, 2231-2307.

Muller, A., et al. (2008). "Formalisation of a new prognosis model for supporting proactive maintenance implementation on industrial system." Reliability Engineering & System Safety 93(2): 234-253.

Musial, W., Butterfield, S. and Ram, B. (2006). Energy from offshore wind.

Norevik, E. (2013). Fault diagnosis in 1MW Wind Turbines' hydraulic pitch system using Dynamic Bayesian Networks. School of Engineering and Computing Science, Durham University. **MSc Thesis**.

Qiu, Y., Feng, Y., Tavner, P., Richardson, P., Erdos, G. and Chen, B. (2012). Wind turbine SCADA alarm analysis for improving reliability. *Wind Energy* **15**(8): 951-966.

Report, D. (2002). "EDWI. Studie zur aktuellen kostensituation 2002 der Windenergienutzung in Deutschland." 2002, from http://www.dewi.de/.

Schlechtingen, M., Santos, I. F. and Achiche, S. (2012). Wind turbine condition monitoring based on SCADA data using normal behavior models: Part 1–system description. *Applied Soft Computing*.

Schneuwl, A. "Ultracapacitors and Wind Turbine Pitch Systems." 2013, from http://www.altenergymag.com/emagazine.php?issue_number=04.10.01&article=maxwell

Singh, S., Bhatti, T. and Kothari, D. (2007). Wind power estimation using artificial neural network. *Journal of Energy Engineering* **133**(1): 46-52.

Spinato, F., Tavner, P., Van Bussel, G. and Koutoulakos, E. (2009). Reliability of wind turbine subassemblies. *IET Renewable Power Generation* **3**(4): 387-401.

Stuart, R. and Peter, N. (2010). *Artificial intelligence: a modern approach*, Prentice Hall Upper Saddle River, New Jersey.

Takagi, T. and Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *Systems, Man and Cybernetics, IEEE Transactions on*(1): 116-132.

Tavner, P., Edwards, C., Brinkman, A. and Spinato, F. (2006). Influence of wind speed on wind turbine reliability. *Wind Engineering* **30**(1): 55-72.

Tavner, P., Greenwood, D., Whittle, M., Gindele, R., Faulstich, S. and Hahn, B. (2013). Study of weather and location effects on wind turbine failure rates. *Wind Energy* **16**: 175-187.

Tavner, P., Xiang, J. and Spinato, F. (2007). Reliability analysis for wind turbines. *Wind Energy* **10**(1): 1-18.

Tavner, P. J. (2012). *Offshore Wind Turbines: Reliability, Availability and Maintenance*, Institution of Engineering and Technology.

Tavner, P. J., Faulstich, S., Hahn, B. and van Bussel, G. J. W. (2011). Reliability & Availability of Wind Turbine Electrical & Electronic Components. *Power Electronics Journal*.

Tewari, A. (2009). Prior knowledge based identification of TSK fuzzy model for static nonlinear systems. Engineering Science and Mechanics, The Pennsylvania State University. **PhD Thesis**.

Tran, V. T., Yang, B.-S., Oh, M.-S. and Tan, A. C. C. (2009). Fault diagnosis of induction motor based on decision trees and adaptive neuro-fuzzy inference. *Expert Systems with Applications* **36**(2): 1840-1849.

Venkatasubramanian, V., Rengaswamy, R., Yin, K. and Kavuri, S. N. (2003). A review of process fault detection and diagnosis: Part I: Quantitative model-based methods. *Computers & chemical engineering* **27**(3): 293-311.

Widodo, A. and Yang, B.-S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing* **21**(6): 2560-2574.

Wiggelinkhuize, E. J., Rademakers, L. W. M. M., Verbruggen, T. W., Watson, S. J., Xiang, J., Giebel, G., Norton, E. J., Tipluica, M. C., Christensen, A. J. and Becker, E. (2007). "Conmow Final Report." Retrieved Feb, 2012, from http://www.ecn.nl/docs/library/report/2007/e07044.pdf.

Wikipedia. (2013). "Brazos Wind Farm." Retrieved Jan, 2012, from http://en.wikipedia.org/wiki/Brazos_Wind_Farm.

Wilkinson, M., Hendriks, B., Spinato, F., Gomez, E., Bulacio, H., Roca, J., Tavner, P. J., Feng, Y. and Long, H. (2010). Methodology and results of the ReliaWind Reliability Field Study. EWEC.

Witten, I. H., Frank, E. and Hall, M. A. (2011). *Data Mining: Practical Machine Learning Tools and Techniques: Practical Machine Learning Tools and Techniques*, Elsevier.

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B. and Philip, S. Y. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems* **14**(1): 1-37.

Xie, C. (2013). Wind Turbine Pitch Fault Detection with Adaptive Neuro-Fuzzy Inference System. School of Engineering and Computing Science, Durham University. **MSc Thesis**.

Yang, W. and Jiang, J. (2011). Wind turbine condition monitoring and reliability analysis by SCADA information. *Mechanic Automation and Control Engineering (MACE), 2011 Second International Conference on*, IEEE.

Yen, J., Wang, L. and Gillespie, W. (1998). A global-local learning algorithm for identifying Takagi-Sugeno-Kang fuzzy models. *Fuzzy Systems Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on*, IEEE.

Zaggout, M. N. (2013). Wind Turbine Generator Condition Monitoring via the Generator Control Loop. School of Engineering and Computing Sciences, Durham University. **PhD Thesis**.

Zaher, A., McArthur, S., Infield, D. and Patel, Y. (2009). Online wind turbine fault detection through automated SCADA data analysis. *Wind Energy* **12**(6): 574-593.

Zappalà, D., Tavner, P., Crabtree, C. and Sheng, S. (2013). Sideband Algorithm for Automatic Wind Turbine Gearbox Fault Detection and Diagnosis. EWEA. Vienna.

Zio, E. and Gola, G. (2009). A neuro-fuzzy technique for fault diagnosis and its application to rotating machinery. *Reliability Engineering & System Safety* **94**(1): 78-88.

Appendices

A. Research Facility

The ReliaWind Server environment incudes a HP Server, Backup Devices, Uninterruptable Power Supply (UPS), and software.

A.1 ReliaWind Server

The ReliaWind Server is a powerful computer that used to store data and provide powerful computing services across the department network.

Server Type:	Rack Mountable – 2U
Server Model:	HP ProLiant DL180 G6
Processor:	2 x Intel Xeon E5620 (2.4 GHz, 4 cores, 8 Threads, 12M cache)
Memory:	48GB DDR3-1333 MHz (4 x 4GB + 4 x 8GB)
Hard Drive:	4 x 2TB 3.5″ 7.2K rpm (Hot Plug) 2 x 160GB 3.5″ 7.2K rpm (Hot Plug)
Power Supply Unit:	2 x 750W PSU (Hot Plug)
Warranty:	HP 3 years next business day warranty.
Scalability:	Up to 2x 6 core Processor Up to 8 x 3.5″ HD = 16TB Up to 192GB memory Be able to connect with Storage Area Network (SAN) or Network Attached Storage (NAS)

Table A.1.1: ReliaWind Server Specification

A.2 Backup Device

In order to avoid a data loss event, backup copies of the data will be used to restore the original. For this research, two external hard drives are used to backup data.

Type:	WD My Book Edition II 4TB Dual Drive Network Storage (3.5”)
Brand:	Western Digital
Capacity:	2 x 4TB = 8TB

Table A.2.1: Backup device specification

A.3 Uninterruptable Power Supply

The UPS is an electrical apparatus that provides emergency power to the ReliaWind Server when the input power source fails.

Model:	Smart-UPS RM 1000VA USB 2U
Brand:	APC
Type:	Rack-Mountable – 2U
Power Capacity:	600 Watts / 1000VA
Voltage:	230V
Output:	4 IEC Power Sockets

Table A.3.1: UPS specification

A.4 Main Software

The Table 2.7 lists the main software installed on ReliaWind Server.

Application	Software Name	License
Operating System:	Windows Server 2008 R2 Enterprise edition – 64 bit	KMS
Database:	Microsoft SQL Server 2008 R2 Enterprise edition – 64 bit	KMS
Anti-Virus Software:	McAfee Enterprise edition	Provided by ITS
Matlab:	Matlab R2010a – 64 bit	Site license
Mathematica:	Mathematica 8 – Higher Education	2 years license from Mar/2011

Table A.4.1: Software installed on ReliaWind Server

- *Windows Server 2008 R2 Enterprise edition* improves application responsiveness for worker who access content from servers in remote locations. In addition, Millions of application available for windows OS, easy configuration & management, easy to construct the connection with our existing client OS (Windows) have led us to choose Windows Server 2008 R2 Enterprise edition.
- *SQL Server 2008 R2 Enterprise edition* support up to 524PB database, memory utilization 2TB, powerful data query ability, and no hidden cost. It is the

cheapest Enterprise level database solution among DB2, Microsoft SQL Server and Oracle.

- *McAfee* is one of the best anti-virus software. The Enterprise edition significantly increases the security confidence.
- *Matlab* is a numerical computing environment and a 4th generation programming language. It allows matrix manipulations, plotting of functions and data, implementation of algorithms, and interfacing with other programming language, e.g. C, C++. The 64 bit version of Matlab allows for much larger memory usage, up to 2^64 bytes (Unlike 32-bit which is limited to 2^32 bytes = 4GB).
- *Mathematica* is a computational software program used in scientific, engineering, and mathematical fields and other areas of technical computing.

B Data Visualisation Tool

The WT Data Visualisation Tool is a Client/Server-based application. It is a unified platform developed by the author to assist experts to conduct efficient laboratory research on WT reliability data. The Server side of this application is placed on the ReliaWind Server; it is used to handle data and process Client requests. The Client side is a graphical user interface (GUI) - data visualisation interface, which allows users to request the Server's content or services, for example raw data, data plot and data aggregation.

The tool is fairly easy to use. Simply double-click "WT_VisualizationTool.exe" to launch it. Then, choose a Data Source from the pop window like Figure B.1 and click Launch.

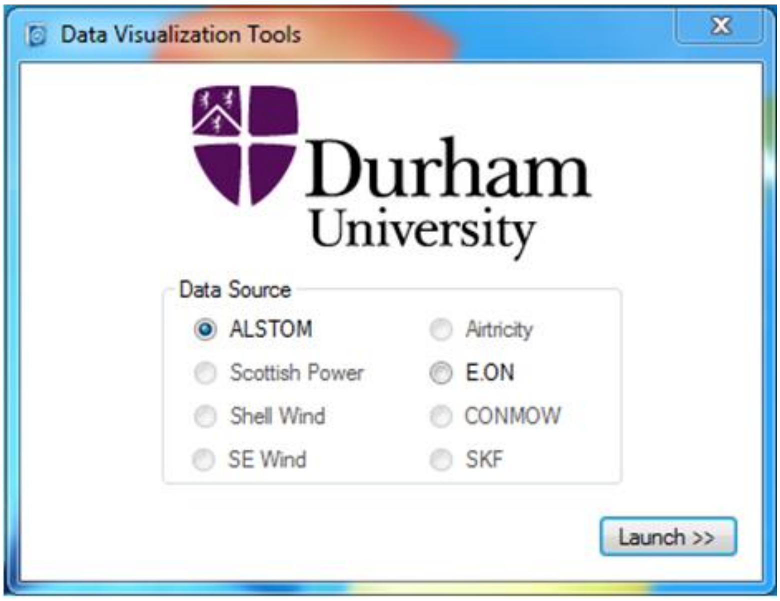


Figure B.1: Pop window.

On the Main Interface, choose any provided chart type from top left corner to display the data source. Other functionalities like Print/Export Data, Zoom in/out, Refine, and Data Aggregation are provided.

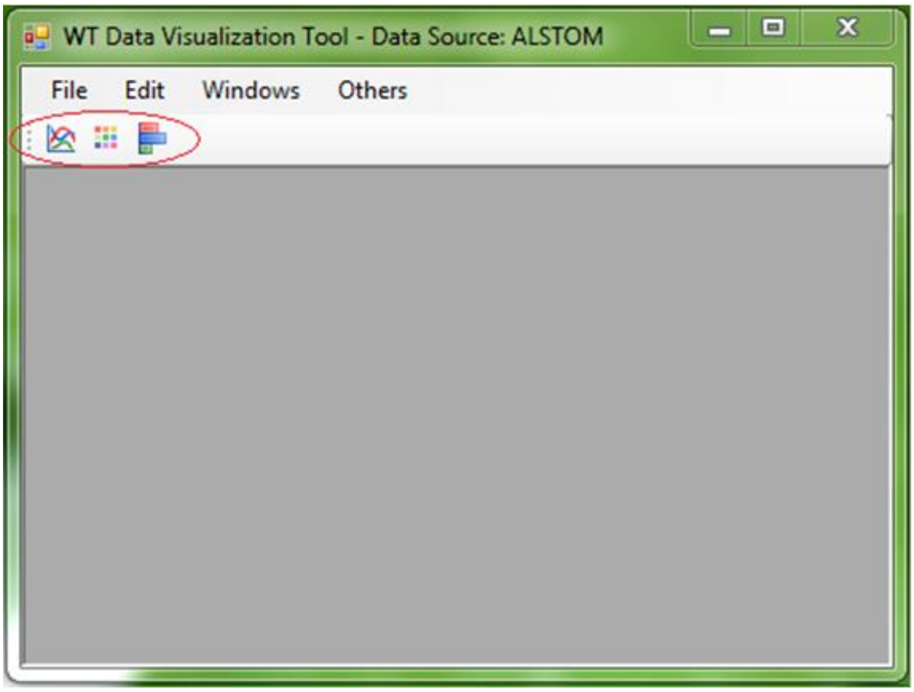


Figure B.2: WT Data Visualisation Tools – Main Interface.

Figures B.3 – B.5 show some demonstrations of the WT Data Visualisation Tool.

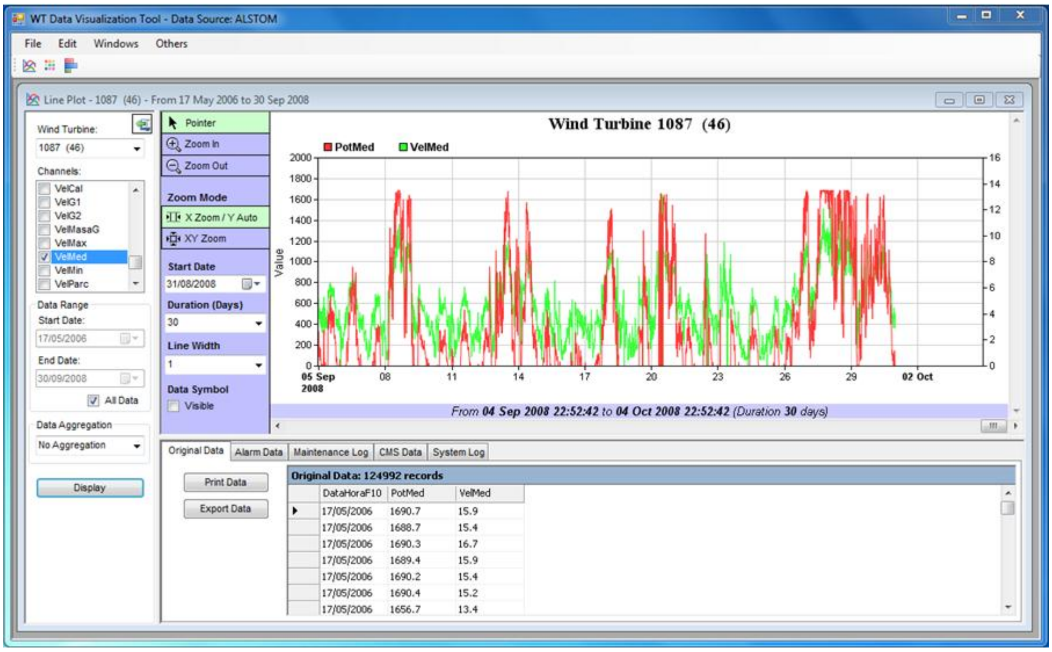


Figure B.3: Alstom Line Plot.

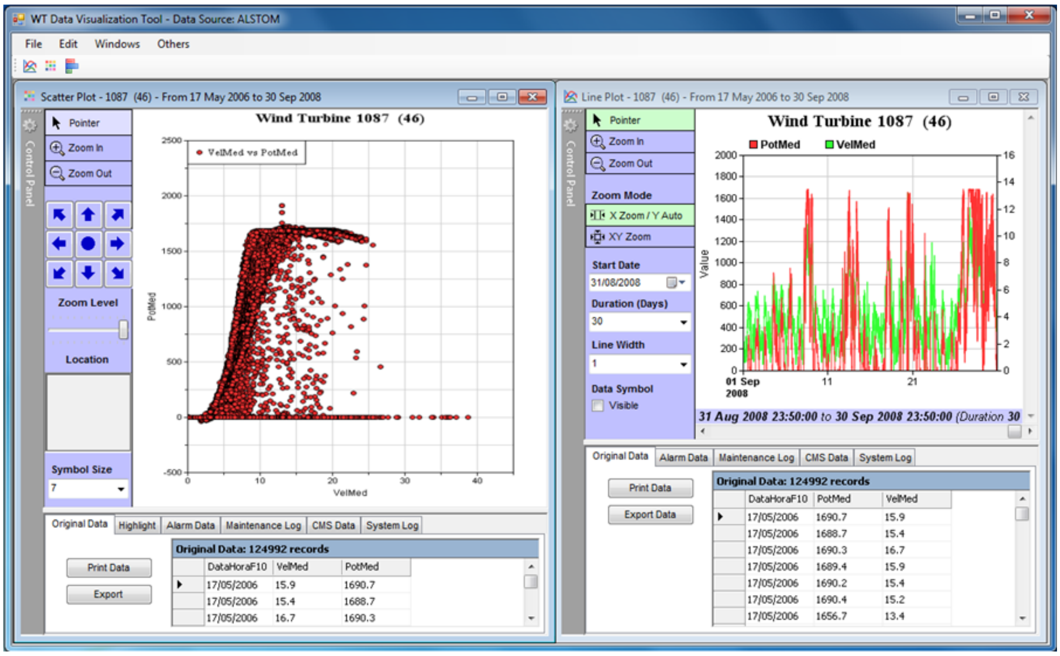


Figure B.4: Scatter Plot and Line Plot.

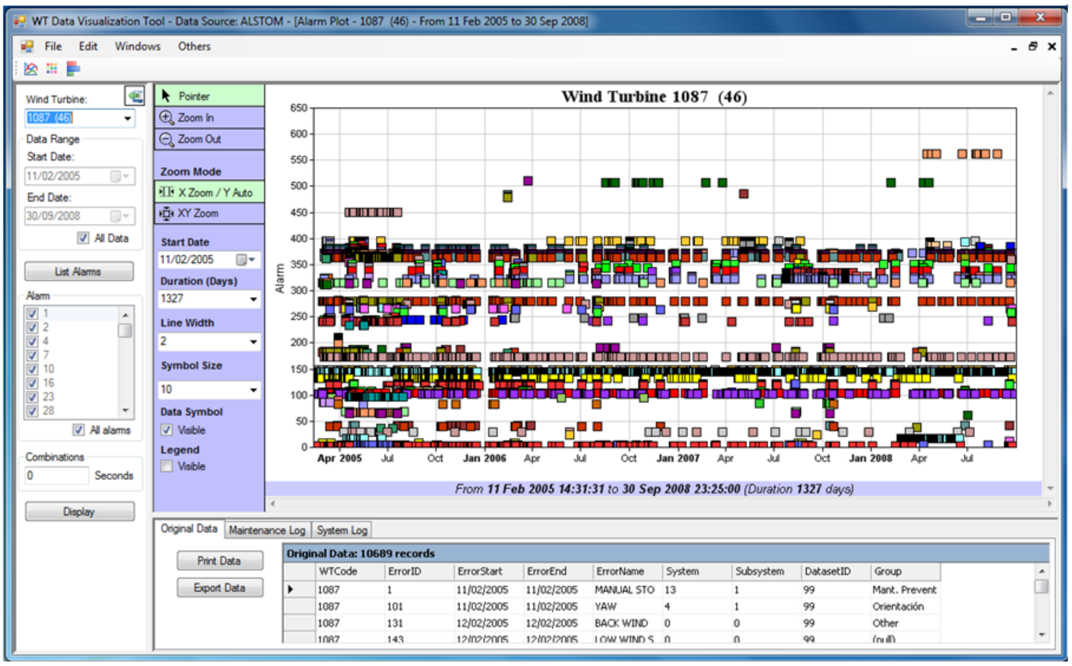


Figure B.5: Alarm Plot.

C Raw Mitsubishi SCADA blade data

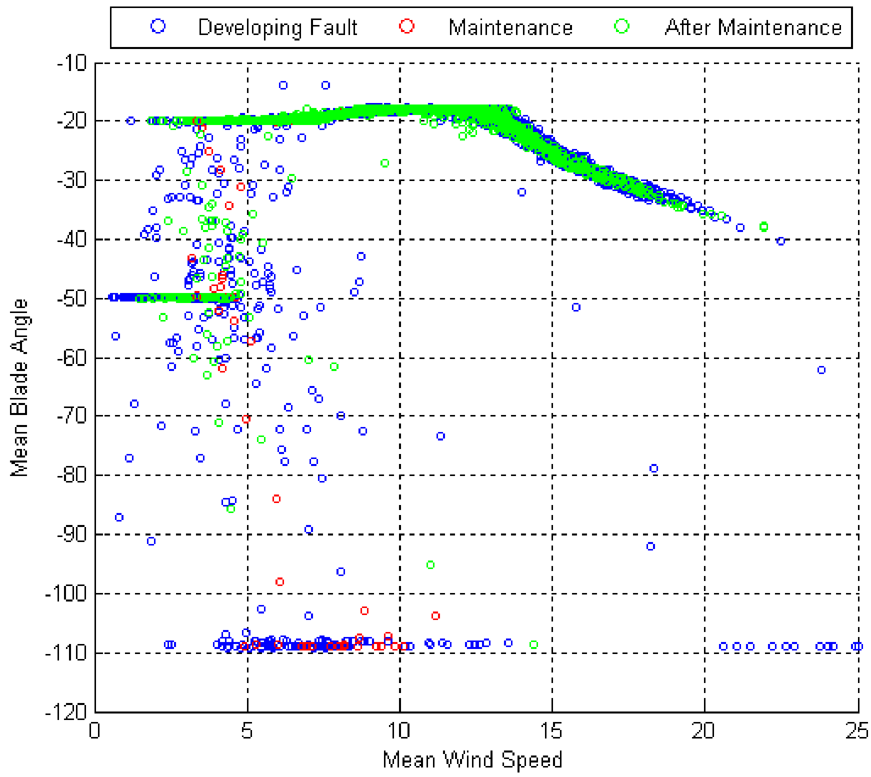


Figure C.1: Original plot of wind speed vs blade angle using Mitsubishi data, where the blade angle is negative because in fixed speed, pitch-controlled WT the blades operate in pitch-to-stall.