Advanced data-driven modelling approaches to alarm-related fault detection and condition monitoring of wind turbines

Yueqi Wu

This dissertation is submitted for the degree of Doctor of Philosophy

January 2022

Faculty of Science and Technology

Engineering Department

Declaration

This thesis has not been submitted in support of an application for another degree at this or any other university. It is the result of my own work and includes nothing that is the outcome of work done in collaboration except where specifically indicated. Many of the ideas in this thesis were the product of discussion with my supervisor Dr. Xiandong Ma.

Yueqi Wu

January 2022

Lancaster University, UK

Abstract

With the increasing demand of renewable energy, the installation of wind power capacity has been rising exponentially in the past decades. With growing size of modern variable speed wind turbines, the operation and maintenance (O&M) cost can be very high. In order to effectively reduce the O&M cost and maximise the reliability of wind turbines, the condition monitoring has been considered as a most viable solution.

However, to improve the reliability of the condition monitoring system, large number of sensors are required and thus numerous data are produced. This can increase the complexity of the monitoring system and bring heavy burden to the computation process. Besides, limited researches has been conducted to study the relationships between alarms and faults.

The thesis starts with overview of current condition monitoring technologies and systems. Then the monitoring data used in the research are explained, which include both supervisory control and data acquisition (SCADA) data and data produced from simulation models. A statistical tool based on Kullback-Leibler divergence (KLD) is proposed for feature extraction, considering normal and abnormal behaviour presented in the monitoring data. The proposed method is improved with kernel support vector machine (KSVM), thus capable of classifying the normal, alarm and fault condition of the operational wind turbines. Furthermore, an approach is proposed based on long short-term memory (LSTM) incorporating a KLD for fault detection and identification of representative faults. This method can effectively distinguish the alarms from the faults, from which the distinguished alarms can be considered as an early warning of the fault occurrence. The true positive rate of the proposed LSTM-KLD is 94% whereas the true negative rates for alarm and fault are 96% and 90.9%, respectively. In the end, a wind turbine test rig is designed and developed, from which experimental data are obtained to validate the proposed fault detection algorithms and models.

The contributions of the research mainly have three aspects. For fault classification, the kernel function is adapted by both principal component analysis (PCA) and support vector machine (SVM) in order to transfer the linearly inseparable problems into linearly separable problems. Besides, variable selection with kernel PCA is proposed for effective condition monitoring, which can reduce the computation load while retaining the most useful information in the monitoring data. For alarm detection, a statistical tool based on KLD is employed to discover the behaviours of

specific components in different operation conditions. By incorporating LSTM with KLD, the fault can be localised by correlating the alarms in the data. With this hybrid method, the fault severity can be estimated based on the alarm signal since it can provide sufficient information as required to indicate early warning of the fault. For experimental validation, a PMSG (permanent magnet synchronous generator) based wind turbine test rig is designed and constructed to emulate the operational behaviours of the turbine under various type of faults in order to collect sufficient experimental data to validate the proposed algorithms for fault detection and severity estimation.

Acknowledgements

First of all, I would like to give my great gratitude to my supervisor Dr. Xiandong Ma for offering me the opportunity in research and giving me his support during the whole time. With his support, feedback and patience, I have gained a lot in the progress of my research work. His caring has also improved my personal development during the study.

I also would like to thank my second supervisor Prof. Allan Rennie and my panel member Dr. Denes Csala for supervising my research progress and giving me suggestions base on my progress appraisal reports.

My sincere thanks also go to visiting scholar Prof. Jun Song, who gave me lots of technical support in design the experimental test rig, especially for circuit design and testing.

I wish to express my thanks to the technicians from the engineering department, particularly Andy Baker, Andrew Verden, Thomas Clayton, James Dickinson and Andrew Gavriluk, for their technical support during my study.

I also appreciate administration support from Deborah Lund, Janet wood, Laura Gracie and Kathryn Rucastle.

The permission of using operational wind farm SCADA data from Wind Prospect Ltd. is also gratefully appreciated.

At last, I would like to express my extreme thankfulness to my parents who believed me the whole time and gave me continuous support during the study.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Acronyms

Abbreviations

AC	Alternating Current
AE	Acoustic Emission
ANN	Artificial Neural Network
ApEn	Approximate Entropy
APSO	Adaptive Particle Swarm Optimisation
BDFIG	Brushless Doubly-Fed Induction Generator
BDS	Blind Deconvolution Separation
BMS	Blade Monitoring System
BPNN	Back Propagation Neural Network
CCEB	Current Condition Evaluation-Based
CDBN	Convolutional Deep Belief Network
CFF	Current Form Factor
CMS	Condition Monitoring System
CNN	Convolutional Neural Network
CPSO	Cross-Validated Particle Swarm Optimised
DBN	Deep Belief Network
DCS	Distributed Control Systems
DDS	Direct Digital Synthesis
DDT	Digital Drive Technology
DFIG	Doubly-Fed Induction Generator
DNN	Deep Neural Network
dq0	direct-quadrature zero
DVAR	Dynamic Voltage-Ampere Reactive
EMA	Exponential Moving Average
FBG	Fibre Bragg Gratings
FCPB	Future Condition Prediction-Based
FFT	Fast Fourier Transform
FIS	Fuzzy Inference System
GBM	Gradient Boosting Machine
GRU	Gated Recurrent Unit
GSC	Grid Side Converter
HAWT	Horizontal Axis Wind Turbine
H-ELM	Hierarchical Extreme Learning Machine
HHT	Hilbert-Huang transform
HMI	Human Machine Interface
I/O	Input and Output
IADT	Iterative Atomic Decomposition Thresholding
ICS	Industrial Control System
IEDS	Intelligent Electronic Devices
IGBT	Insulated Gate Bipolar Transistor

IPCA	Independent Principal Components Analysis
KELM	Kernel Extreme Learning Machine
KICA	Kernel Independent Component Analysis
KKT	Karush–Kuhn–Tucker
KLD	Kullback-Leibler Divergence
KLOF	Kernel based Local Outlier Factor
KPCA	Kernel Principal Components Analysis
KSVM	Kernel Support Vector Machine
LCOE	Levelised Cost of Energy
LDA	Linear Discriminant Analysis
LS-SVM	Least Square Support Vector Machine
LSTM	Long Short Term Memory
MDUE	Minimisation of Dissimilarity-and-Uncertainty-based Energy
ML	Machine Learning
MTU	Master Terminal Unit
NI Crio	National Instrument Compact Rio
O&M	Operation and Maintenance
PCA	Principal Components Analysis
PDD	Probability Density Distribution
PHM	Prognostics and Health Management
PLC	Programmable Logic Circuit
PLP	Power Law Process
PMSG	Permanent Magnetic Synchronous Generator
RAMS	Reliability, Availability, Maintainability, and Safety
RBF	Radical Basis Function
RMS	Root Mean Square
RNN	Recurrent Neural Network
RPM	Revolutions Per Minute
RTU	Remote Terminal Units
RUL	Remaining Useful Life
SCADA	Supervisory Control and Data Acquisition
SCIG	Squirrel Cage Induction Generator
SHM	Structural Health Monitoring
SIMAP	Intelligent System for Predictive Maintenance
SNR	Signal to Noise Ratio
SPWM	Sinusoidal Pulse Width Modulation
SR	Stochastic Resonance
STFT	Short-Time Fourier Transform
SVD	Singular Value Decomposition
SVM	Support Vector Machine
SVPWM	Space Vector Pulse Width Modulation
SVR	Support Vector Regression
VAWT	Vertical Axis Wind Turbine

WAN	Wide Area Network
WPD	Wavelet Packet Decomposition
WPT	Wavelet Packet Transform
WRIG	Wound Rotor Induction Generator
WT	Wind Turbine
WTCM	Wind Turbine Condition Monitoring

Roman letters

a_i	Variance contribution
A_r	Rotor blade surface area
b	Bias
$COST$	Cost function
C_p	Power coefficient
C_t	Cell state
Eq''_{cmd}	Voltage command
erf	Error function
$f(t)$	Failure density function
$fa(t)$	Failure intensity function
f_c	Carrier frequency
$f_{cut\text{-}off}$	Cut-off frequency
f_{ref}	Reference frequency
f_t	Forget gate
$h(x)$	Classification function
H_0	Fault-free condition
H_1	Fault condition
$h(t)$	Hidden state
$I(p\|\|g)$	Kullback-Leibler information between two probability density distributions
I_{pcmd}	Current command
$Iphase$	Phase current
$i(t)$	Input gate
$K(x,xi)$	Kernel function of x
K_b	Fixed constant
$KLD(p,g)$	Divergence of the two distributions
KLD_M	Mean value of Kullback-Leibler divergence
M	Covariance matrix
m_a	Frequency modulation index
m_f	Identical components
N_0	Number failed at time t
$N_{f(t)}$	Number surviving at time t
$N_{s(t)}$	Alarm value probability

P_{AV}	Generator active power
P_{gen}	Mechanical power
P_{mech}	Normal value probability
P_{NV}	Active power to the system based on the inputs from the turbine model
P_{ord}	Unreliability function
$Q(t)$	Reactive power command
Q_{cmd}	Generator reactive power
Q_{gen}	Reactive power to the system based on the inputs from the turbine model
Q_{ord}	Reliability function
$R(t)$	i-th principal component' variance
s_i	Sign function
sgn	Standard deviation
std	Time
t	Training dataset
$T(x,y)$	Hyperbolic tangent
$tanh$	Left singular vectors
U	Carrier wave
V_c	DC voltage
V_{dc}	Fundamental frequency
V_o	Output voltage of LC filter
V_{out}	Reference wave
V_{ref}	Terminal voltage
V_{term}	Wind speed
v_w	Weight
w	Right singular vectors
W^T	Standardised dataset
z_j	

Greek letters

α_i	Lagrange multipliers
$\alpha_{i,j}$	Pre-defined coefficients used for wind power C_p curve
β	Dimensionless parameter determining the shape of the curve
γ	Decision boundary
λ	Ratio of rotor blade tip speed and wind speed
Λ	Width of the kernel
μ_i	Mean value of i th row
ρ	Scale parameter
σ	Sigmoid function
$\tau(w)$	Optimal hyperplane with maximal margin
$\phi(x)$	Mapping of x in hyperplane
ω_{ref}	Reference angular speed

List of Appendices

Publications

Y. Wu and X. Ma, “A hybrid LSTM-KLD approach to condition monitoring of operational wind turbines,” *Renewable Energy*, vol. 181, pp. 554–566, Jan. 2022, doi: 10.1016/j.renene.2021.09.067.

M. Benbouzid, T. Berghout, N. Sarma, S. Djurović, Y. Wu, and X. Ma, “Intelligent Condition Monitoring of Wind Power Systems: State of the Art Review,” *Energies*, vol. 14, no. 18, p. 5967, Sep. 2021, doi: 10.3390/en14185967.

Y. Wu and X. Ma, “Alarms-related wind turbine fault detection based on kernel support vector machines,” *The Journal of Engineering*, vol. 2019, no. 18, pp. 4980–4985, Jul. 2019, doi: 10.1049/joe.2018.9283.

Y. Wu and X. Ma, “Kullback-Leibler divergence based wind turbine fault feature extraction,” in *2018 24th International Conference on Automation and Computing (ICAC)*, Sep. 2018, pp. 1–6. doi: 10.23919/IConAC.2018.8749103.

Chapter 1. Introduction

This chapter starts with a brief overview of the current state of wind energy. Then the importance of wind energy research especially in wind turbine condition monitoring is introduced. The motivation in research and development of fault detection and fault severity estimation techniques towards intelligent wind turbine condition monitoring is described. The aim and objectives for the research are also demonstrated. Furthermore, the novelty of this research is summarised, followed by the structure of the thesis at the end.

1.1 Current state of wind energy

Currently, the traditional fossil fuels are dominating the global energy supply. The energy related carbon emissions have reached the highest level in history recently [1], which has resulted in the serious global environmental problems [2]. However, the demand of electrical energy consumption is still rapidly increasing. The use of renewable energy has become more imperative in order to reduce the carbon emissions. The United Nations General Assembly (UNGA) set up the Sustainable Development Goals (SDGs) in 2015 to provide an international cooperation framework in order to protect the planet environment and thus bring a sustainable future for the planet. In "Agenda 2030", sustainable energy is the core to the success [3]. There are three key targets set to achieve the SDG, including to i) increase substantially the share of renewable energy in the global energy market, ii) ensure affordable, reliable and universal access to modern energy services, and iii) double the global rate of improvement in energy efficiency. This has been a focus of a number of renewable energy studies [4-10]. The European Union (EU) is taking an important role in global energy transition process. It is estimated that in 2030, the savings due to use of renewable energy rather than the import of fossil fuels should reach EUR 58 billion [11]. Besides, by 2020, 10% energy consumption in transportation came from renewable energy [12]. In addition, in 2014, the European Council promised to reduce at least 40% of greenhouse gas emissions and increase 32% share of renewable energy in all energy resources as compared with 1990 by the year 2030 [13]. The COP25 climate summit in Madrid in 2019 agreed the European Green Deal, by which the EU will reach climate-neutral economy by 2050 [14]. As a common type of renewable energy, wind power has been widely used and continually growing over the past decades. Fig. 1.1 shows the historic development of total wind energy installations. In the begging of the new millennium, the total wind energy installations were only 6.5 GW, including both onshore and offshore turbines. It can be observed from the Fig. 1.1 that the total installation has been rising exponentially over the past two decade. The total wind turbine (WT) installation in 2020 was 743GW with 93 GW newly installed capacity, as compared with 60.8 GW newly installed capacity in 2019 [15].

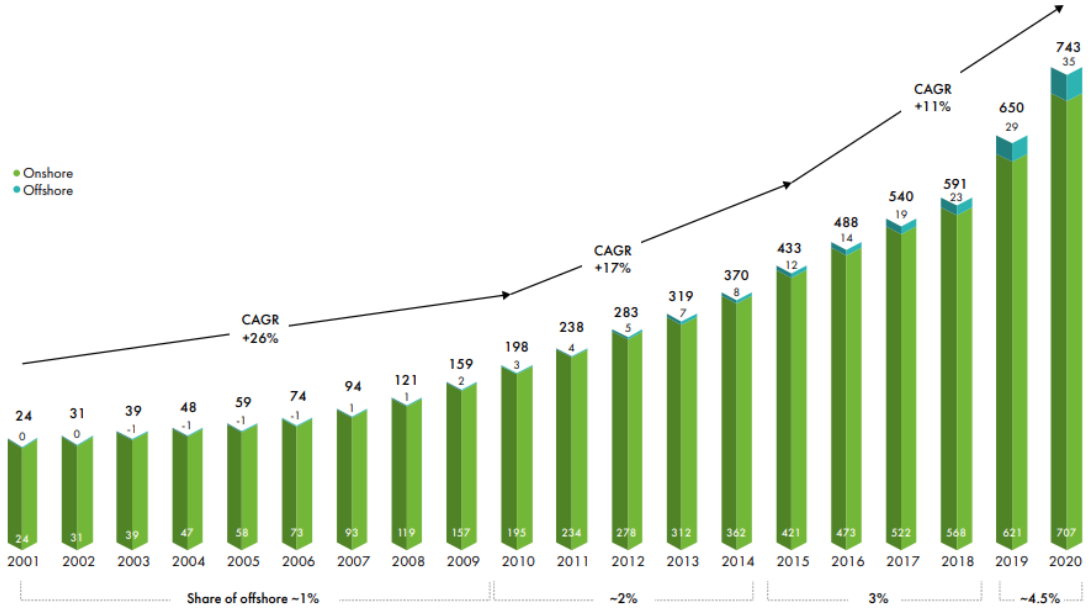


Figure 1.1 Historic development of total WT installations (GW) [15]

With the exponentially installed capacity of WTs, one of the most concerned problems is the levelised cost of energy (LCOE). The LCOE, also known as the levelised cost of electricity, is an economic index used to calculate the average total cost required to build and operate a power-generating system over its lifetime divided by the total power generated of the system over that lifetime [16]. The LCOE of an onshore wind farm can vary from $50/MWh to $60/MWh. However, it can increase from $700/MWh to $1,200/MWh under unfavourable wind conditions [17]. The LCOE for typical onshore and distributed WT installations was observed to be $37/MWh. The full range of single-variable sensitivity based on turbine distance for the onshore turbines is estimated to be $29–$78/MWh. The LCOE for the offshore turbines tend to be higher, which is estimated to be $85/MWh for fixed-bottom substructures and $132/MWh for floating substructures, with a single-variable sensitivity range of $63–$122/MWh for fixed bottom structure and $83–$174/MWh for floating structure, respectively [18].

The O&M cost mainly involves the costs associated with operating and maintaining process such as insurance, regular maintenance, repairs, spare parts and administrations. The annual O&M cost is on average 3% of the original cost of the turbines. For offshore wind farms, the O&M costs occupy 30% of the total income of the turbines over 20 years of their operating lifetime [19], [20]. Since WTs are usually installed in remote areas and moving further offshore, more efforts and costs

have been required to ensure their reliable operation. Therefore, it is important to detect the WT faults at an early stage, thus improve their reliability. In order to reduce the O&M cost and minimise the economic loss caused by downtime, appropriate condition monitoring techniques that are able to detect fault in the early stage before being developed into catastrophic ones would be crucial.

1.2 Research motivations

There are various types of signals that are be used for wind turbine condition monitoring (WTCM). The signals mainly include acoustic emission, torque, temperature, lubrication oil, vibration, strain and electrical powers. Monitoring data are usually acquired from the SCADA system for modern WTs [21]. In terms of the condition monitoring models, the data-driven model-based methods, as compared to physical-based models, are focusing on constructing the relationship between inputs and outputs of the system. It does not require building a mathematical model of the physical system [22]. Numerous statistical, machine learning (ML) and hybrid data-driven approaches have been studied. In statistical approaches, the factors such as variance, kurtosis, skewness, mean value and root mean square (RMS) value are acquired from time-series data to monitor the rotor performance, blade surface roughness and gearbox failures of the WTs [20], [23-25]. Usually, the deviations of those statistical values are monitored during the operation. The statistical methods have been proven as a mature and flexible solution to implement; however, inaccurate decisions might be made because of their high sensitivity to noise since the noise tolerance of the statistical methods is lower.

ML based methods such as artificial neural network (ANN) and SVM have been developed rapidly in the past decade. These human neurons-inspired methods have been widely applied in different WT subsystems such as gearbox, generator, and power transmission systems [26-28]. In [29], the author proposed a condition monitoring method based on fusion of spatial-temporal features of SCADA data by convolutional neural network (CNN) and gated recurrent unit (GRU) to predict the output for recognising the WT condition. Wang et al [30] developed a deep neural network (DNN) based framework to model lubricant pressure. Their results showed that the DNN model is more accurate when compared with k-nearest neighbour, Lasso, Ridge, SVM, and ANN. Hu et al. [31] proposed a deep belief network (DBN)

based prognostic model to improve the prediction accuracy and facilitate prognostic uncertainty for bearing degradation detection. In order to improve the accuracy and robustness of the condition monitoring methods, hybrid methods have also been developed. A hybrid statistical-ML method based on fast spectral kurtosis and multi-branch CNN was proposed for identifying the complex fault in WT gearbox [32]. It has been proved that this method can diagnose the gearbox fault with over 97% accuracy. Pan et al. [33] combined DBN, self-organising map and particle filtering to evaluate the degradation process and predict the remaining useful life (RUL) of WT gearbox effectively. Several ML algorithms, such as CNN, recurrent neural network (RNN) and LSTM, can also be combined together to provide a more accurate condition monitoring method at the cost of model complexity [34, 35].

Efforts have also be made to develop condition monitoring methods for early fault detection. Bangalore et al. proposed an ANN based condition monitoring method that used SCADA data to produce the early fault warning on WT gearbox in order to arrange maintenance resources [36]. Wang et al. employed operational condition clustering and optimised DBN modelling for early fault detection of the main bearing [37]. However, most condition monitoring approaches have not yet utilised multiple variables being monitored, e.g., temperatures and pressures, from the same WT subsystem to improve the condition monitoring capability. Besides, only few researches have taken SCADA alarms into consideration. In addition to the operational and environmental parameters, a SCADA system includes a detailed record of alarm logs, revealing the malfunction of particular parameters of subsystems and components of the turbines. There are two types of alarm signals. The first type is usually triggered when certain measurement value of the components exceeds the pre-set threshold, which is considered as true alarms [38]. Another type is triggered when the system has experienced transient changes due to such as acute changes and disturbances of wind speed, and is thus considered as false alarms [39]. Due to inherent correlation between alarms and fault, it would be sensible to investigate if the WT alarm signals can be verified as an early warning for performance monitoring. By analysing the alarms and using time-sequence probability-based optimisation, the maintenance schedule can be well-organised and thus the WT reliability is improved [40]. Alarm signals can also be used to find the root cause of a fault or stoppage because the early warning raised by the alarms can be related to the fault occurrence [41]. Hence, alarm signals can be used to

crosscheck potential faults identified from the data against what was actually happening and thus can play a significant role in WTCM.

1.3 Aims and objectives

To address the issues mentioned above, intelligent WTCM technologies need to be developed. The aim of the research is to develop cost-effective, accurate, and robust condition monitoring technologies to improve the reliability of WTs. Therefore, the O&M costs can be reduced. The data used in the research contain three types, namely SCADA data, simulation dare and experimental data. The historical SCADA data are sampled at 10 mins/sample that show almost every detail of the operational WT condition. However, such a low sampling rate could only reflect the operation trend of mechanical components. The electrical signals are changing fast which cannot be captured instantaneously. Besides, the historical data do not contain electrical faults. Hence, simulation models will be built to simulate the dedicated electrical faults. As a key component of the turbine, the power conversion unit is playing a vital role to connect the turbine and the grid. Thus, a power conversion test rig will also be built to emulate those typical electrical faults including capacitor aging. To fulfil the research aim, the following objectives are listed:

- Conduct critical literature reviews of current WT technologies, including the fundamental working principle, common failure modes, condition monitoring technologies, and maintenance strategies.
- Model and simulate WT systems with different configurations under different operation scenarios such as normal condition and fault condition.
- Understand the physical relationships among monitoring variables not only in the simulation data but also in the real measurement data.
- Develop an appropriate methodology that can distinguish and separate the faults from the healthy conditions effectively.
- Develop effective models to distinguish the normal, alarm, and fault conditions of the WTs whilst reducing the computation load by optimising the variable selection for dedicated condition monitoring.
- Develop novel methods to estimate fault severity index that can measure the severity of abnormal conditions of the WT.

- Design and develop an operational WT power conversion test rig to further experimentally validate the proposed condition monitoring algorithms and models.

1.4 Novelty of the research

As described, the aim of the research is to design cost-effective, accurate, and robust intelligent WTCM technologies. The novelty of the research are summarised below:

A statistical tool based on KLD is proposed for monitoring variable selection for fault classification. The divergence value can be used to indicate the differences between normal and abnormal behaviour of the monitoring data, which helps localise the fault.

A kernel principal components analysis (KPCA) is further proposed for variable selection to classify operation conditions. The method can significantly reduce the computation load while maintaining the most information presented in the original datasets. Compared with the PCA, the KPCA with an adaptive kernel function can solve the non-linear problems.

The proposed kernel function is extended to be used in SVM. The kernel support vector machine (KSVM) can project the original dataset into a higher dimension that is then divided with a hyperplane. This can transfer the linearly inseparable problems into linearly separable problems.

A novel data-driven model-based condition monitoring method based on LSTM with KLD is developed. The proposed method is able to implement a condition monitoring that can analyse the WT operating conditions automatically and detect both alarms and faults simultaneously.

LSTM is developed to capture relation features in temporal dependencies among monitoring data in an iterative manner, thus improving the prediction capability. The KLD value is used as fault indicator, which measures the severity of the fault by comparing the probability distributions between normal data and test data. By adopting a cost function based on normal value probability and alarm value probability of the calculated KLD values, the optimised thresholds are determined to distinguish the normal, alarm, and fault conditions.

1.5 Layout of the thesis

The first chapter starts with description of the importance of wind power and research motivation. The research aim and related objectives are listed. Furthermore, the layout of the thesis is presented at end of this chapter.

In the second chapter, the overview of WT technologies including structure, configuration and common failure modes is firstly presented. The state-of-the-art condition monitoring technologies are then reviewed, along with the WT maintenance strategies.

In the third chapter, the monitoring data used in this research is interpreted, which is organised by two major sections. The first section discusses the SCADA data obtained from an operational wind farm. The second section explains the process required to develop simulation models of doubly-fed induction generator (DFIG) and PMSG based on MATLAB/ Simulink, respectively. The models are simulated under different operation conditions to generate sufficient data for model validation.

In the fourth chapter, a method based on KLD and KSVM is proposed for fault feature selection. This method is developed to distinguish particular faults and extract the fault feature of different faults.

In the fifth chapter, a cost-effective condition monitoring method that can distinguish the normal, alarm, and fault is proposed through a combination of variable selection based on KPCA and operation condition classification based on KSVM. The computation load reduction is also considered and assessed during the development of the models.

In the sixth chapter, a novel data-driven model-based condition monitoring method based on LSTM with KLD is proposed. The proposed method aims to implement a WTCM that can analyse the WT operating conditions automatically and detect both alarms and faults simultaneously. The fault severity is also estimated by the proposed hybrid approach.

In the seventh chapter, the details of the design and construction of WT test rig are presented, which is to emulate the operation of PMSG WT, focusing on its power conversion unit. The condition monitoring algorithms and models proposed in the

previous chapters are also validated by the experimental data acquired from this test rig.

In the eight chapter, conclusions of the research are summarised, followed by contributions to the knowledge arising and achievements obtained from this PhD research. The limitations of the research and recommendations for the further improvement are also discussed.

Chapter 2. Overview of wind turbine condition monitoring technologies

In this chapter, the state-of-the-art wind turbine condition monitoring system is reviewed. First of all, the background of current wind turbine technologies including different designs and configurations are introduced. Then the common failure modes of wind turbine subsystems are described, followed with the wind turbine reliability analysis. The modern wind turbine condition monitoring technologies are also reviewed with respect to conventional and more recent ML technologies. The statistical methods, signal processing methods and ML methods all are reviewed. Moreover, the maintenance strategies based on wind turbine fault diagnosis and prognosis are addressed. At the end of this chapter, the future challenges and improvements of wind turbine condition monitoring are discussed.

2.1 Overview of WT technologies

Wind power is one of the most popular, sustainable, and renewable energy sources. Compared with conventional fossil-fuel energy sources, wind power has far less environmental impact. Wind energy provides mechanics through the blades of WTs to turn electric generators for producing electrical power.

Human use of wind energy can be traced back thousands of years. The wind wheel of Hero of Alexandria is considered the first recorded instance of a wind-powered machine in history (10 AD–70 AD) [42]. In the medieval period, the first recorded use of wind power occurred in England in the 11^{th} or 12^{th} century [43]. A modern horizontal-axis wind turbine (HAWT) with a 100-kW generator on a 30-m tower connected to a local 6.3 kV distribution system was installed at the Soviet Union in 1931 [44].

Various modern WTs have been developed. There are two main types of WTs, which differ with regard to the direction of the main rotor shaft: vertical-axis wind turbines (VAWTs) and HAWTs. These are the two most commonly used designs. VAWTs can be divided into two main categories: Savonius and Darrieus, as shown in Figures 2.1(a) and (b), respectively [45]. A Savonius turbine is composed of two or more semi-cylindrical buckets, which is the simplest design. Savonius turbines have lower efficiencies than that for the other designs [46]; hence, they can only be used in small scale applications. The second type of VAWT is the Darrieus turbine. Usually, three aerofoil-shaped blades are installed to generate a lift force for the main shaft rotation. The three common Darrieus turbines are shown in Figures 2.1(a)–(c): helical blades, Darrieus, and H-shaped blades, respectively [47]. Compared with Savonius turbines, Darrieus turbines have lower starting torques and higher efficiencies and rotation speeds.

Another type of WT is the HAWT. There are two types of HAWT designs: the three-bladed upwind HAWT and the two-bladed downwind HAWT. Upwind HAWTs are more popular than downwind HAWTs because their active yaw mechanism is not needed in the early development of wind power, and there is no risk of blades hitting the tower. Additionally, downwind turbines have the fatal problem that the turbulence induced by the tower causes periodic loads on the blades, leading to power fluctuation, which is also known as ‘tower shadow’ [48]. In contrast, the

blades and rotors of upwind turbines are placed along the wind direction. Hence, there is no need to consider the side effects of tower shadows. The mainstream design of a commercial HAWT typically comprises three blades. This is because further increasing the number of blades increases the wind resistance, reducing the power generation, leading to an efficiency lower than that of a three-bladed turbine. For turbines with two blades, gyroscopic precession occurs, leading to turbulence.

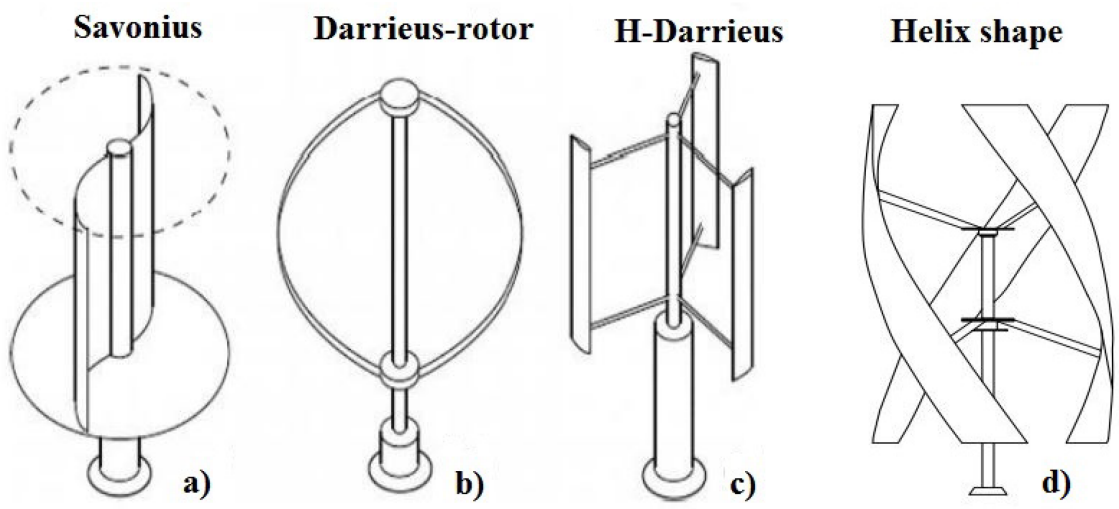


Figure 2.1 Different types of VAWTs: (a) Savonius; (b) Darrieus; (c) H-shaped blades; (d) Helix [45]

For power generation under a steady wind speed, HAWTs are better than VAWTs. Under the condition of a steady and high-speed wind stream, HAWTs can generate significantly more energy than VAWTs [49]. Compared with HAWTs, VAWTs are more difficult to start and stop and have lower efficiencies. This dissertation focuses on the condition monitoring of HAWTs, as they are the most commonly used WTs.

The structure and subsystems of the HAWT are illustrated in Figure 2.2 [50]. The turbine tower is designed to capture high-speed winds and to avoid turbulent air layers close to the ground. The rotor and nacelle of the HAWT are elevated and installed on top of the tower. To optimise the balance between energy and cost, the height of the tower is typically between two and three times the length of the blades. Depending on the installation area, the foundation beneath the tower can be either solid (for onshore WTs) or floating (for offshore WTs).

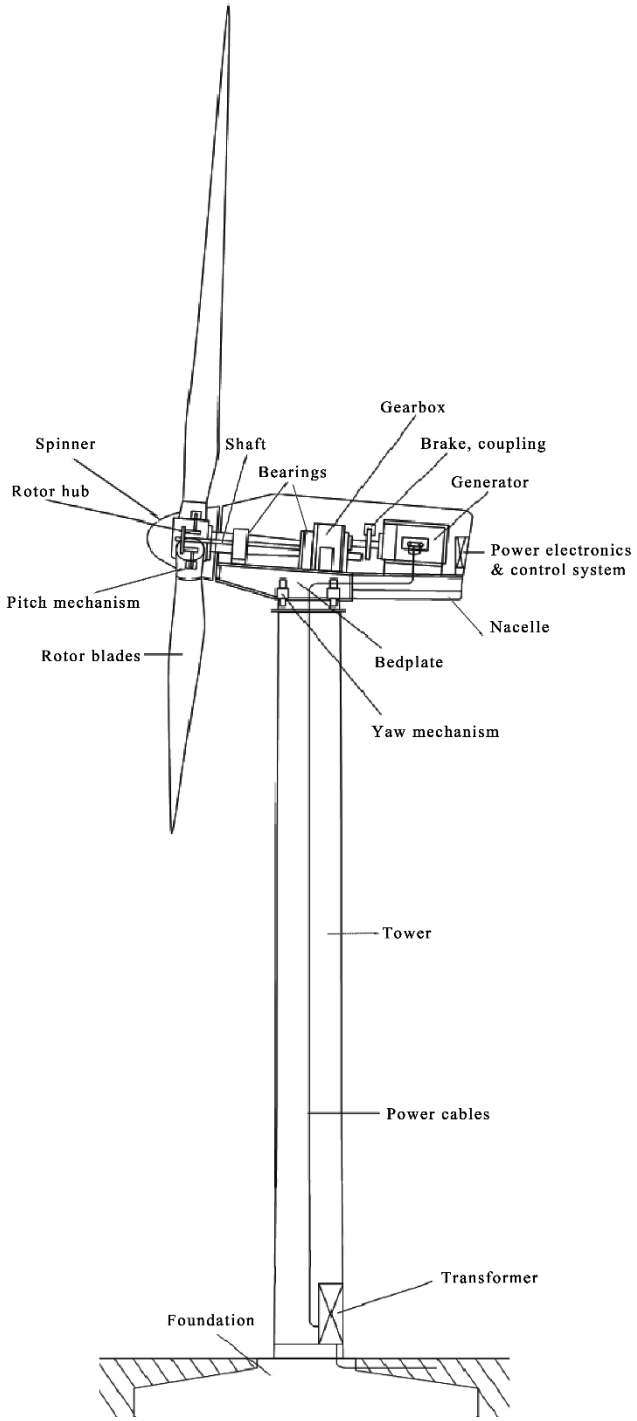


Figure 2.2 Typical HAWT structure and subsystems [50]

A typical rotor system of an HAWT usually has three blades connected to a horizontal shaft. The blades are designed as aerofoils instead of being completely flat to satisfy the aerodynamic requirements. A pitch system is used to control the pitch angle when the HAWT faces the wind. The blades receive an uplifting force from the wind, and a driving torque is generated, causing rotation. The diameters of typical modern HAWTs are between 40 and 90 m.

The nacelle is directly connected to the rotor. The gearbox, generator, shaft, bearing, brake, and controller are installed inside the nacelle house. The low-speed shaft is connected to the rotor and gearbox. It rotates at approximately 10–20 revolutions per minute (rpm), typically transferring mechanical power to the gearbox [51]. The gearbox consists of a set of gears that increase the output speed of the shaft to approximately 1500 rpm at a frequency of 50 Hz (depending on the requirement of the local grid) [51]. The high-speed shaft connects the gearbox and the generator. The output speed of the shaft is vital if it is directly connected to a generator. This is because the frequency of the alternating current (AC) produced is directly related to the rotation speed. The generator converts mechanical power into electrical power.

On the opposite side of the rotor, an external anemometer is installed on the nacelle, which is responsible for measuring the wind speed and wind direction. According to the received data, the controller determines the turbine behaviour. In commercial WTs, a programmable logic controller (PLC) is embedded in the anemometer to determine the direction of the WT for maximising energy harvesting. The turbine can also be shut down to prevent overspeeding.

The yaw system is located at the top of the tower and connected to the nacelle. The turbine is rotated to ensure that the blades face the incoming wind.

For all types of HAWTs, variable-speed WTs have more advantages than fixed-speed WTs, e.g. lower component stresses, higher energy yields, and fewer grid connection power peaks. To maximise the power production, the WTs operate in two primary regimes: below rated wind speed and above rated wind speed. When the wind speed is below the rated wind speed, the turbine operates at a variable rotor speed to capture the maximum amount of wind energy. The rotor speed is controlled by changing the generator torque. The blade pitch angle is kept constant. When the wind speed is above the rated wind speed, the generator torque is kept constant, and the blade pitch angle varies depending on the wind speed. Thus, a constant power output is maintained [51].

The aforementioned basic structures are similar to those of the modern WT design. However, there are differences in the electrical configurations of WT drivetrains. Figure 2.3 shows the four main drivetrain configurations currently used in industry.

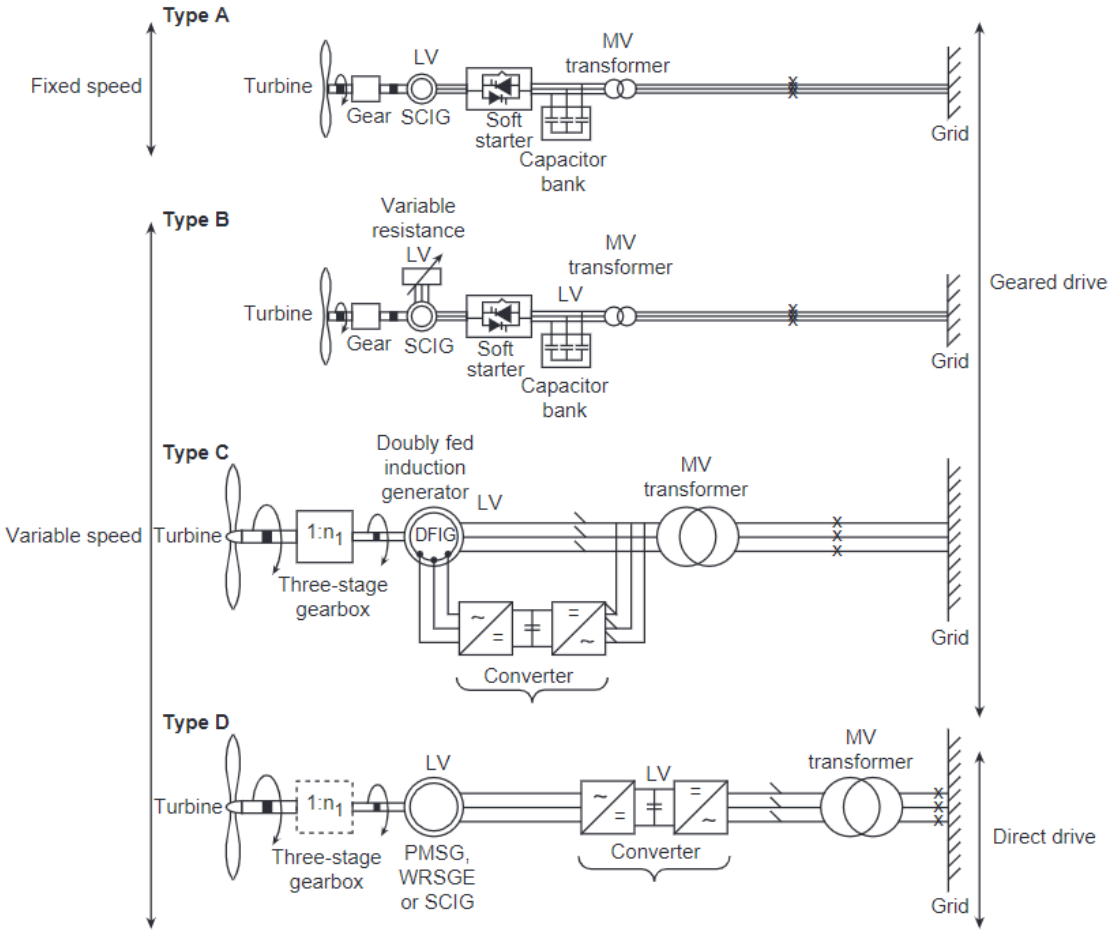


Figure 2.3 Main electrical configurations for the current WT drivetrains [52]

The type A configuration is used for fixed- or dual-speed and stall-regulated WTs. A geared-drive low-voltage squirrel-cage induction generator (SCIG) is installed. This generator is connected to a medium-voltage grid via a transformer. The synchronisation inrush current can be reduced via power-factor correction and a soft starter. This type of drivetrain configuration has a simple structure, a low initial cost, and high reliability. However, sudden and substantial changes in the wind speed require higher mechanical robustness. Additionally, a robust network is needed because of the fluctuations in the electrical output.

In contrast to type A, the type B and C configurations are used for variable-speed turbines. Type B is for fixed- or dual-speed, stall-regulated, or variable-speed, controlled-stall-regulated WTs. A geared-drive low-voltage wound rotor induction generator (WRIG) with a variable rotor resistance is connected via a transformer directly to the medium-voltage grid. Power-factor correction and a soft starter are used to reduce the synchronisation inrush current.

As shown in Figure 2.3, types C and D use a four-quadrant converter instead of a soft starter connected to a medium-voltage grid through a transformer. For type C, a geared-drive low-voltage WRIG is used. The rotor is connected to the converter, and the stator is connected to the transformer. This connection is called the DFIG and is one of the commonly used for industry WTs [52].

The type D configuration is designed for variable-speed variable-pitch WTs with a direct-drive PMSG or SCIG, where the stator is connected to the converter and also connected to the medium-voltage grid via a transformer.

In addition to the aforementioned commonly used WT drivetrain configurations, new designs have been researched, as shown in Figure 2.4. They offer potential reliability benefits for both electrical and hydraulic options.

Types C' and C'' are derivatives of type C. Type C' uses a low-voltage brushless doubly fed induction generator (BDFIG) instead of a DFIG. The BDFIG feeds through the tow stator windings with different pairs of poles to avoid direct coupling. However, it also provides a lower-speed generator. Thus, two-stage gearboxes can be adopted instead of the three-stage gearboxes. Similar to type C', type C'' removes the brush gear and slip rings. A WRIG with a self-driven three-phase AC brushless exciter is adopted to replace the DFIG.

Type D' is based on type D but uses a single- or two-stage gearbox with a low-speed generator and a fully rated converter. Additionally, the brush gear and slip rings are removed in this configuration. It has higher gearbox reliability and better power quality than type D.

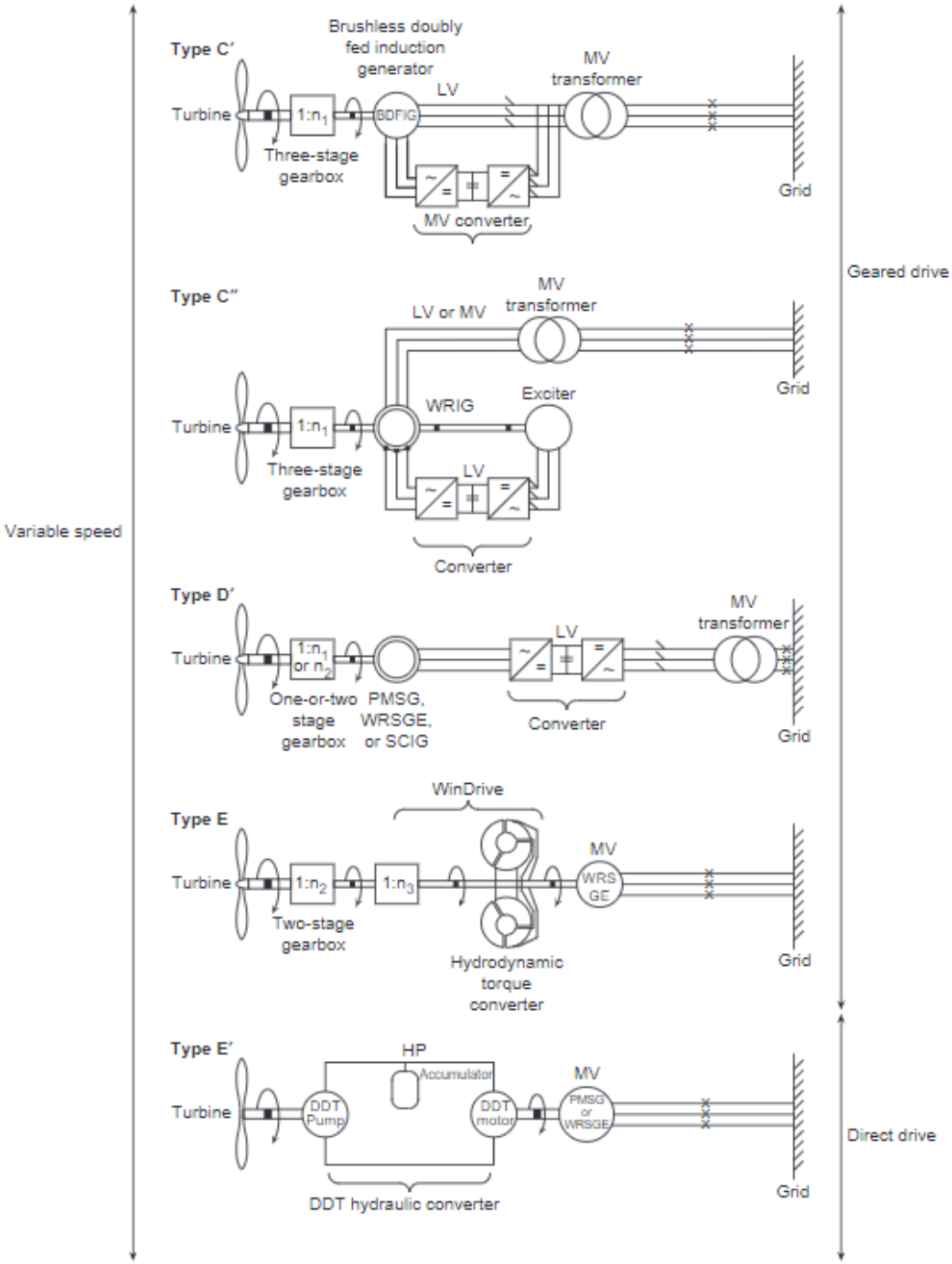


Figure 2.4 Main electrical configurations for innovative WTs [52]

Type E is a hydraulic arrangement based on the conventional geared drivetrain; however, a limited-speed range hydraulic torque converter is used to drive a medium-voltage wound rotor synchronous generator. Synchronous generation at a medium voltage is used to replace the converter and transformer, which reduces maintenance costs.

Type E' is an advanced hydraulic solution based on Type E. Digital drive technology (DDT) from Artemis Innovative Power is adopted to drive a low-speed pump. The

low-speed hydraulic pump then provides a high-pressure hydraulic fluid to the high-speed hydraulic motor. The advantage of this design is that the gearbox, converter, and transformer are eliminated. However, it has not yet been implemented in industry.

2.2 Common failure modes of WT subsystems

To capture the maximum amount of wind energy and minimise environmental side effects, both onshore and offshore WTs are installed in remote areas. However, this has harsh drawbacks, such as unstable wind profiles and low temperatures, which present challenges when failures occur. In addition to the unpredictable environmental conditions, the development of WTs size tends to be larger and heavier, which increases the failure rate. Figures 2.5 and 2.6 show the failure rate and downtime, respectively, for critical subassemblies and components of WTs; the data were collated from 18 publicly available databases (including >18000 WTs, corresponding to >90000 turbine years). Here, the inner and outer circles correspond to onshore and offshore turbines, respectively [53].

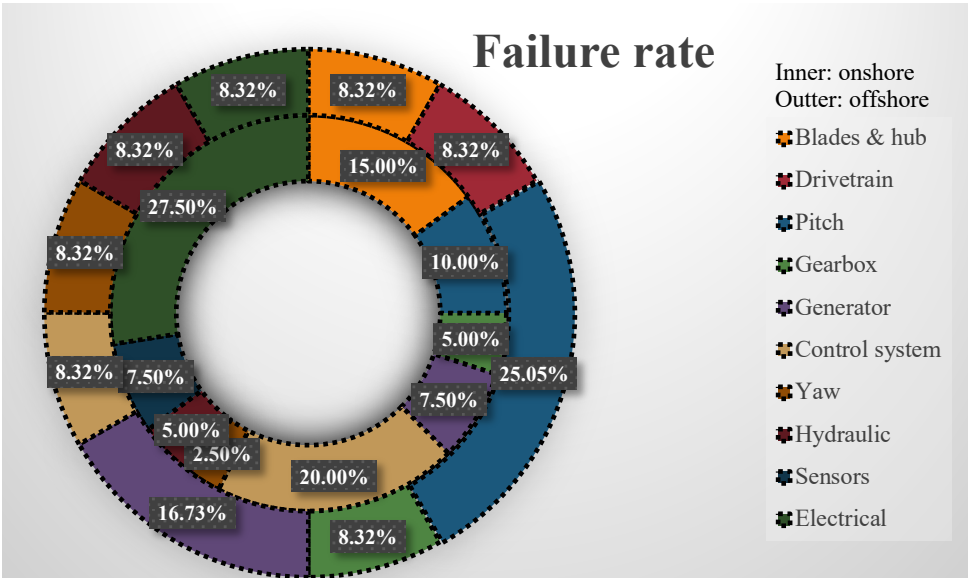


Figure 2.5 Critical subassemblies in terms of failure rate

As shown in Figure 2.5, the critical subassemblies of onshore WTs are as follows (top five failure rates, in descending order): electrical, control system, blades and hub, pitch, and generator. These subassemblies are also critical for offshore turbines; however, the order of criticality is slightly different.

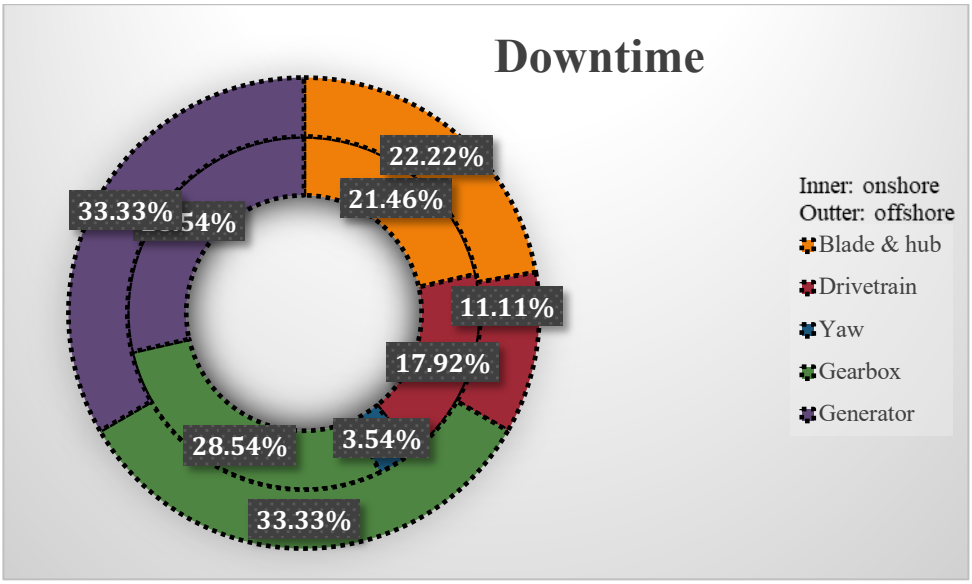


Figure 2.6 Critical components in term of downtime

The rotor blades of WTs convert wind energy into mechanical energy. Because of the complexity of the environment and the fluctuating stress, the blades suffer a high failure rate. The most common failure modes of blades are fracture, fatigue, freezing, sensor failure, cracking, and wear. It is difficult to repair and maintain the blades because of their high installation locations. Thus, it is important to study the failure modes of the blades and improve their reliability.

Owing to the complex structure of the gearbox, the failure of any key component can lead to high maintenance costs and a long repair time (particularly for offshore WTs), as shown in Figure 2.6. Fatigue is the most common failure mode [54]. The remaining failure modes are insufficient lubrication, fracture, wear, and bearing failure. The strong and unbalanced axial forces on the bolts can lead to the failure of the bolts that connect the front box, ring, and middle box.

The generator converts mechanical power into electrical power through the rotation of a high-speed shaft connected to the gearbox. The three most common failure modes are bearing failure, magnetic-wedge loss, and contamination.

Table 2.1 Summary of failure modes of components [55]

Objects	Functions	Failure modes	Causes
Blades	Capture wind	Fracture, edge crack, stuck, motor failure, pitch bearing failure	Fatigue loads underestimated; extreme loads; environmental effects; imbalance
Main shaft	Transmit high torque	Fracture	Fatigue loads underestimated; operation of generator under off-design conditions; material properties below specs
Yaw system	Enable the nacelle to rotate on the tower	Increased bearing friction	Cracked roller; galled surface; lack of lubrication
High-speed shaft	Stop and hold the shaft during shutdown and operation	Low or high brake torque	Environment effects
Gearbox	Transmit torque with speed increase	Internal gear tooth failure	Fatigue loads underestimated; exceeding design load; improper materials; loss of lubricating oil
Hub assembly	Transmit torque from blades	Structure failure, bolt failure	Excessive design loads; excessive preload; stress corrosion
Oil seals	Retain oil in main bearing housing; exclude foreign matter	Cut or wear in lip	Installation damage; wear
Filters	Extract all particulate contaminants from hydraulic fluid and hold them	Case leakage	Damage to case or seals
Generator	Generate electrical power	Overheating, fault, jammed bearing, bearing seizure, overspeed	Overload; no excitation; environmental effects; misalignment; fatigue; mechanical failure; loss of drivetrain control
Blades	Capture wind	Fracture, edge crack, stuck, motor failure, pitch bearing failure	Fatigue loads underestimated; extreme loads; environmental effects; imbalance

As part of the drivetrain, the main shaft and high-speed shaft contribute to the drivetrain failure. The most common failure modes of the two shafts are fractures and low or high brake torques [56]. The fatigue of the components and high torque can devastate the bearing, leading to failure of the shafts.

Besides the mechanical failures occurring on the critical subassemblies, the failures that occur on the electrical components cannot be ignored. Power electronics failure is the major failure mode of electrical components. Sudden strong electrical pulses can be fatal to power electronics and lead to converter failures. Numerous failure

modes of the main subassemblies have been explained and details regarding the failure modes of the critical components are presented in Table 2.1 [55].

2.3 Monitoring WTs

Several systems have been developed for monitoring the conditions of modern WTs. The monitoring systems can be categorised as condition monitoring systems (CMSs), SCADA systems, and structural health monitoring (SHM) systems. The CMS provides the highest resolution among the three types of monitoring systems. It also allows monitoring of high-risk subassemblies of the WT for diagnosis and prognosis of faults. The blade monitoring system was designed as a part of the CMS to detect early defects in blades. The SCADA system provides a low-resolution system that allows WTs to supervise operating conditions. In addition, alarms are generated from the WTs. The SHM system was designed to monitor key items of the WT structure by providing low-resolution signals.

2.3.1 Condition monitoring systems

The CMS monitors the conditions of the machinery parameters. By analysing the significant changes in the target parameter, developing faults can be discovered. This is also a major process in predictive maintenance. By adopting CMS, major faults that shorten the useful life of a WT can be addressed in advance. CMSs have been widely used in auxiliary systems, rotating equipment, and other machinery, such as pumps, electric motors, compressors, and internal combustion engines. In addition, non-destructive testing and fit-for-service evaluation methods used for steam boilers and heat exchangers are adopted [57].

The CMS was first applied in the WT industry in the 1900s. It was initially developed from reputable condition monitoring original equipment manufacturers, such as National Instruments, Bently Nevada, Bruel, and Kajaer. Initially, the systems were mostly based on previous rotating machine vibration condition monitoring experience [58]. However, the mechanism of modern WTs is not the same as that of conventional rotating machines. It has been proven that CMSs are successful in onshore WTs when used by experienced operators. Currently, CMSs are installed in most new WTs (≥1.5 MW) and have been fitted to almost all offshore turbines. A typical wind farm CMS is shown in Figure 2.7.

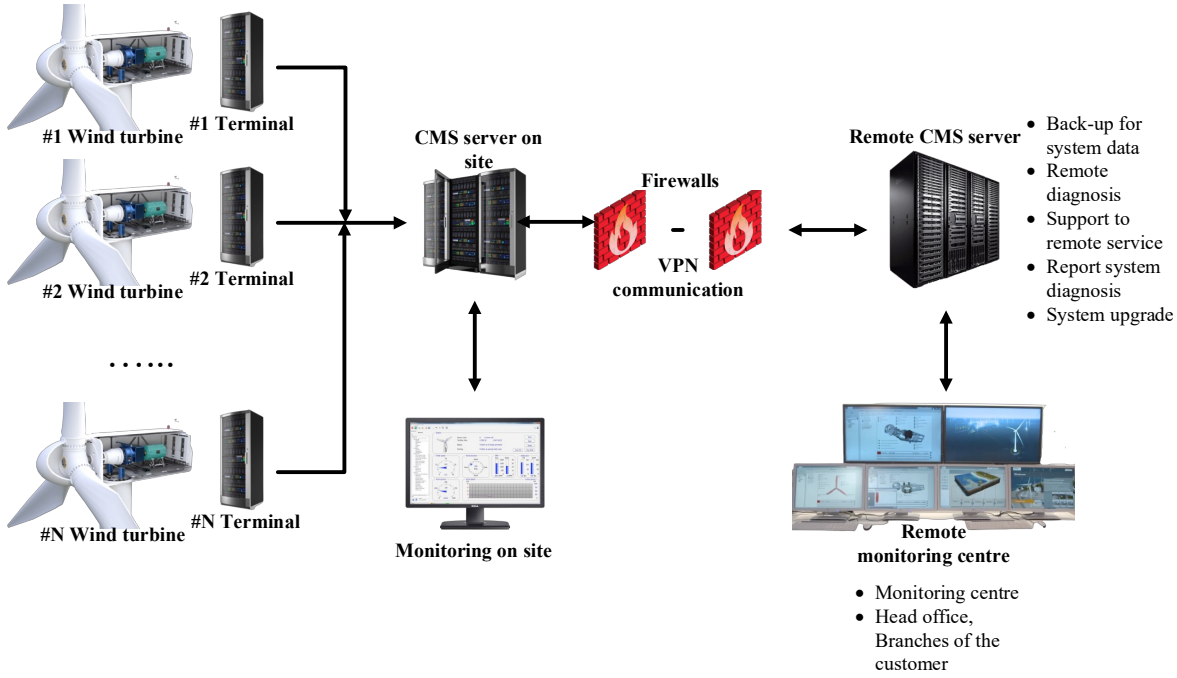


Figure 2.7 Wind farm condition monitoring system

The four main techniques used for the WT CMS are vibration, oil debris, strain, and electrical analyses. Vibration analysis is first used to monitor the main bearing, the gearbox, and the generator. To use vibration analysis, a set of low-frequency accelerometers is installed on the main bearings, while higher-frequency accelerometers are installed on the gearbox and generator. In some cases, proximeters are installed on the generator for more accurate measurements. Because of the rapid change in the drivetrain torque, the vibration periods and amplitudes change with time. When a fault occurs, the vibration signals contain a high impulsive content. It is recommended to use a wavelet transform to analyse the vibration signals [59]. However, this can be computationally expensive. The use of the Fourier transform to analyse a period of limited, predefined speed and power range vibration signals has become mainstream in industry.

Owing to the basic properties of lubrication oil, it is added to gear pairs and bearings. The oil has three main functions: cooling the gearbox, lubrication for the rolling element bearings, and lubrication for the meshing gears. However, it is inevitable for gears and bearings to produce ferrous and non-ferrous debris, because they are all wearing components. Lubricating oil is pumped from the gearbox sump. It then passes through the cooler and the in-line filter. Finally, it reaches the top of the gearbox and is sprayed using numerous nozzles. Consequently, the oil stream gains

heat and debris from all parts of the gearbox. This is attractive for condition monitoring because it has affects the entire gearbox. The analysis results of oil debris detection can provide a warning of upcoming failure, providing time for inspection and maintenance.

Strain-gauge measurement is used for blade pitch control at the beginning because of the process of independently pitching the three blades of the turbine. This control method reduces the torque and lateral force loads on the WT. Thus, the useful life of the WT is prolonged. In addition, the independent blade root bending moment can be measured by adopting circular fibre optics incorporating fibre Bragg grating strain gauges [60]. The measurement can also be used for the CMS.

The electrical signals also need to be mentioned as a type of CMS signals. The voltage, current, and power signals used for generator speed control and excitation have also been used for condition monitoring and their coupled drivetrain for many years [61]. One difficulty in analysing these signals is that rich harmonic electrical information is mixed with the required signals. One method uses fast Fourier transform (FFT) spectral cursors for vibration interpretation to facilitate electrical interpretation [62].

2.3.2 Supervisory control and data acquisition

In addition to processing with condition monitoring signals, the SCADA system is extensively used in condition monitoring. The SCADA system was designed for the oil, gas, and process industries for accurate measurement of the statuses of pumps, valves, and storage vessels. It has been applied in many areas and has evolved into an industrial control system and further into a distributed control system. The UK power industry has been using the SCADA system for modern power station control for over 36 years (since 1985) [52].

The SCADA system includes both input/output (I/O) signals and alarms. A SCADA system usually takes samples at 10-min intervals, which has lower sampling rate than those for the CMS. The system samples valuable monitoring variables such as the wind speed, power output, lubrication oil temperature, and pressure. The maximum, mean, minimum, and standard deviation are recorded for these variables. Most SCADA data are output as they flow through the WT to the control room. The

rest of the SCADA data are input as they are fed from the control room to the WT to manipulate the behaviour of the WT.

Most SCADA signals and alarms are generated from inside the WT controller, and an industrial PLC is commonly used to ensure that the WT operates normally and stops action in an emergency event.

One of the advantages of the SCADA system is that it can provide a universal view of the entire turbine and generate alarms in the case of fatal damage. However, the low data sampling rate limits the analysis accuracy since the transient changes are ignored. The SCADA system is a low-cost monitoring system that integrates measurement, information, and communication technologies. Details regarding the system are presented in Chapter 3.

2.3.3 Structural health monitoring

As the size of WTs increases and a large amount of WTs are installed in remote areas for richer wind sources, the demand for SHM is increasing. SHM has become a useful method because of its ability to enhance operational management and optimise maintenance activities [63]. It aims to reduce unwanted downtime and prevent unnecessary replacements, thereby increasing availability. Because SHM allows monitoring of the structural integrity, improvements can be made, such as the selection of lighter blades. Thus, the performance was enhanced, and the conservative margins of safety were reduced [64]. In this section, common SHM techniques are introduced.

Acoustic emission (AE) is an effective technique for detecting microscale failures. Failures, such as crushing, excessive deformation, debonding, cracking, impacts, and delamination, can cause transient changes and release elastic energy at specific points. An AE signal is defined by its amplitude and energy [65]. The energy released at the failure points can be captured by AE signals. The cracks in the blade can be located, and fatigue tests can be monitored [66]. The sound produced by the stress-released waves and energy dissipation can be measured by piezoelectric sensors [67].

The thermal imaging method aims to detect anomalies beneath the material surface using the subsurface temperature gradients. Changes in thermal diffusivity can be detected according to irregularities in material damage [68]. According to the thermal excitation methods, thermal imaging methods can be classified as active or passive. Passive thermal imaging methods are not commonly used, because they measure materials at temperatures outside the range for typical environments. The active methods use an external stimulus and measure the material temperature.

Ultrasound is widely used for the detection of the inner structures of solid objects [69]. In this technique, ultrasonic waves are affected by the tested materials. The method can detect cracks within the millimetre range.

Since damage to turbine structures can be catastrophic, fatigue and modal-property monitoring is one of the most important SHM techniques. Accelerometers are installed to analyse the dynamic response of the structure. The changes in the damping coefficient, resonance frequency, and modal curvatures can be detected according to the changes in different physical properties, such as the reductions in the stiffness and mass [70].

2.4 Existing condition monitoring techniques

WTCM systems consist of a combination of sensors and signal processing equipment [52]. Condition monitoring techniques include statistical analysis, signal processing, and ML techniques. These techniques are used to monitor the statuses of major WT subsystems, such as blades, gearboxes, generators, and nacelles. The monitoring process can be online or offline. Online monitoring provides real-time data that reflect the instantaneous feedback of the operation conditions. Offline monitoring involves the collection of data at regular time intervals using different data-acquisition systems [53]. With appropriate condition monitoring techniques, maintenance actions can be planned in time to prevent further damage to the turbine while the components are still operational. Thus, the downtime, maintenance, and operational costs are reduced [54]. As mentioned previously, both the condition monitoring and SCADA systems can be used to monitor the condition of WTs. In this subsection, modern WTCM technologies are reviewed and summarised according to the different condition monitoring methods and WT subassemblies.

2.4.1 Conventional methods

Statistical measures such as the root mean square error (RMS) and peak–peak amplitude are used for fault detection. Jinhyuk et al. used vibration signals collected from strain gauges on a gearbox and generator to detect faults by calculating the RMS value [71]. The change in the signal amplitude can be detected as a short fault by setting an appropriate threshold [72]. Other statistical parameters, such as the standard deviation and kurtosis, are also used for WTCM [73, 74]. However, these condition monitoring methods cannot process real-time data. In addition, these techniques focus on a single feature of the variables; thus, extra time is needed to process different types of data.

For WTCM, trend analysis involves the collection of data from various sensors and the prediction of their trends. It is often incorporated with other algorithms, such as co-integration residual analysis, semi-Markov models, and maximum likelihood [75-77]. The applications of trend analysis include monitoring the bearing of the turbine and extracting feature patterns from the generator.

Time-domain analysis includes all monitoring methods based on time-domain signals. It is used to examine physical and transient changes over time. Statistical methods and trend analysis are mostly based on time-domain analysis because statistical values, such as the mean, maximum, minimum, and kurtosis, are extracted from the data for a period of time. Variations in the time-domain signal are easy to observe and are typically used for vibration analysis [78], oil analysis [79], and AE analysis [80]. Cepstrum analysis is a time-domain analysis approach; the cepstrum is defined as ‘the inverse Fourier transform of the logarithmic power spectrum’ [81]. A novel local cepstrum method was proposed for reducing the background noise and non-harmonic components in the vibration signal and power spectrum. Thus, the effectiveness of fault detection can be improved [82]. Another time-domain analysis method is time-synchronous averaging, which is also called time-domain averaging. Time-synchronous averaging aims to increase the signal-to-noise ratio by averaging a set of replicate measurements. This method can be used to identify the vibration sources of a WT gearbox [83].

By transferring time-domain data to the frequency domain, various frequency-domain analysis methods can be applied. The relationship between pre-twist blades

and cracked blades was investigated using the FFT to analyse the natural frequencies of the turbine blades [84]. The wavelet transform can also be adopted for feature extraction in WTCM. An enhanced empirical wavelet transform was obtained by developing a feasible and efficient spectrum-segmentation method to extract gearbox fault features [85]. Compared with time-domain analysis, the main advantage of frequency-domain analysis is its ability to identify and isolate target frequencies.

Time-frequency-domain analysis is a signal-processing method that allows the simultaneous observation of signals in the time and frequency domains. Compared with one-dimensional time/frequency signals, time–frequency-domain analysis can use the two-dimensional signal for WTCM to achieve a higher accuracy. The limitation of frequency-domain analysis is that the Fourier transform assumes an infinite length of periodic signals. Besides, localising the time domain from the frequency domain is impossible. In contrast, Time–frequency-domain analysis can locate the time domain from the frequency domain, and time–frequency-domain analysis methods, including the wavelet transform, short-time Fourier transform, and Hilbert–Huang transform, are commonly used in this area [85-87].

2.4.2 Machine learning techniques

Owing to the increasing number of sensors and increasing complexity of data structures, ML techniques have attracted attention in the past decade. Conventional WTCM systems either focus on a single parameter or offline data, and they have difficulties dealing with massive heterogeneous data. ML techniques are designed to solve heterogeneous massive data problems, such as variable selection, dimension reduction, feature extraction, and online processing. ML algorithms are commonly used for data-driven condition monitoring and do not require any prior knowledge of the turbine [22].

Owing to the large amount of data and untraceable data sources, the raw data may contain considerable noise. Incorrect data may lead to misjudgements during condition monitoring. Cleaning is necessary before processing the data. A kernel-based local outlier factor (KLOF) was proposed for data cleaning. The data were divided into several segments. Then, the features extracted from these segments, such as the mean, maximum, and peak-to-peak values, were used to evaluate the

degree of each segment by adopting the KLOF. A proper threshold was set to distinguish incorrect data from correct data. The results indicated that the proposed method could effectively identify incorrect data and abnormal segments [88]. Another method based on the minimisation of dissimilarity-and-uncertainty-based energy was proposed for data cleaning [89]. It transforms scattered data into a digital image on a grey scale. Then, the optimum threshold is determined according to intensity-based class uncertainty and shape dissimilarity. Subsequently, the abnormal data are marked by applying an imaging threshold.

Dimension reduction techniques are widely used to reduce the complexity of the original dataset and thus reduce the computational load while processing a large amount of data. PCA is able to extract principal components (PCs) from different variables. They are often used for dimension reduction and feature extraction. By applying PCA to a massive dataset, the computational load can be significantly reduced. Wang et al. proposed a PCA based method for selecting certain variables among all variables related to the target fault. The proposed method reduced the dimensions of two different datasets to 51.7% (15 out of 29 variables) for the simulation data and 45.4% (35 out of 77 variables) for SCADA data. The average correlation and information entropy after dimension reduction were 99.81%, 0.0082, and 81.32%, respectively, for the simulation data and 99%, 0.162, and 88.88%, respectively, for the SCADA data. It has been proven that this method can detect faults efficiently and accurately, while reducing the set of variables [90]. Other ML techniques, such as k-means clustering, auto-encoders, and deep belief networks, have also proven to be capable of dimension reduction and feature extraction [88-93].

ML techniques specialise in dealing with heterogeneous massive data in condition monitoring, particularly in reducing the computational load. There is still a problem with condition monitoring: online processing. In the context of streaming/online data, most of ML algorithms may not capable for solving those data, because they are trained using training set based on historical data [94]. In this scenario, incremental learning was considered to prevent the re-training of the previous model. An incremental learning method based on support vector regression and the Karush–Kuhn–Tucker conditions was developed [95]. The dimensions of the training dataset change if a new sample is introduced. The weights are updated

automatically without retraining the data. Therefore, online monitoring can be achieved without building a new model for training. In addition to the incoming data-side problems, data uploading-side problems are considered. To solve these problems, a hierarchical extreme learning machine embedded with a cloud-computing platform is proposed to reduce the quantity of data uploaded [96]. Results indicated that the uploaded data volume was reduced to 12.5% of the original data size before compression, and the data transmission security was also improved, because the parameters of the model and original input data were compressed in the first hidden layer. Thus, the data transmission efficiency and security were improved.

In the past decades, researchers have developed various fault-detection algorithms for WTCM. Algorithms developed in recent years usually combine multiple basic algorithms or extend existence algorithms. In this subsection, state-of-the-art modern WTCM algorithms are reviewed.

2.4.3 Applications of condition monitoring techniques to WTs

Rotor hub and blade

The rotor of the WT consists of a hub and a blade. Errors in the blade pitch angle or rotor mass imbalance can cause rotor asymmetry [97]. The impact of wind on the blades and material aging can result in fatigue. Long-term fatigue can lead to delamination of a blade's glass or carbon fibre-reinforced plastic structure, reducing the stiffness of the blade. It also causes cracks on the blade surface or the internal structure.

Fatigue, cracking, and stiffness reduction result from structural changes in the blade material. Hence, AE sensors installed on the blade can detect these failures. Zhou et al. introduced a blind deconvolution separation (BDS) approach for the detection of WT blade fatigue cracks [98]. Simple values were extracted to compensate for conventional AE measures. To estimate the fatigue crack conditions, BDS algorithms were used to analyse the frequency components. Although the BDS method can enhance the separation quality of specific frequency components, the noise spectra of different wind farms are not the same. Hence, the parameters must be modified when applied to different WT systems.

Xu et al. identified the WT blade damage mode using wavelet packet decomposition (WPD) [66]. A 59.5-m-long composite blade under a fatigue load was used to verify the proposed method. The raw AE waveform was obtained after wave attenuation calibration and sensor array arrangement. WPD was used to capture the features—particularly the information contained in the original AE signals. A clustering analysis based on the extracted features was performed to identify the damage mode and detect a singular signal.

The aforementioned WTCM approaches were applied to WT blades in previous studies. However, if the defects develop to certain levels and lead to vibration on the blades, vibration signal-based WTCM can be adopted. In addition, vibration on the main shaft can result in abnormal behaviours of the rotor, such as rotor asymmetry. Hence, electrical signal-based WTCM methods are also considered.

Awadallah et al. proposed a WTCM method based on a special kriging analysis to detect faults in WT blades [99]. Cracks and defects in the blades cause spectral shifts in the resonant peaks. According to this phenomenon, a special kriging model was proposed. FFT was employed to extract the spectral information of the vibration. The spectral information was compared between healthy and cracked structures to produce the kriging model. It was verified that the kriging model could identify microcrack locations with an accuracy of 95%.

The magnitudes of particular fault signature components generated by WT faults based on current signals have been widely used for detecting early WT abnormalities. However, owing to the constant change in the generator speed, the current signals vary with time, which causes the fault signature components to be close to the supply frequency and its harmonics. To address this problem, Ibrahim et al. used electrical signals to detect rotor asymmetry faults [100]. First, analytical expressions defining the rotor electrical asymmetry fault signature in the DFIG stator current were used to ensure that the fault signature components could be recalculated over time. Then, the adaptive extended Kalman filter tracker was used to extract the instantaneous fault signature, which is based on the machine speed signals and estimated error covariance. In each time step, the calculated fault signature components from the current signals were processed using an extended Kalman filter to predict the future states of the fault signature components. Thus,

continuously updating the future state makes real-time monitoring possible. This method was compared with the continuous wavelet transform and iterative localised discrete Fourier transform and proved to be superior with regard to both computational efficiency and accuracy of fault diagnosis.

With the improvement of computing elements, ML algorithms have been increasingly applied to the condition monitoring of WT rotor hubs and blades. Reddy et al. proposed a convolutional neural network (CNN)-based image analysis method for the detection of WT blade cracks [101]. They used drone-captured images to train the CNN model to identify the fault and fault-free conditions of blades. First, a drone was used to capture images of the WT blades. Manual annotations were then added to distinguish between binary classification data and multiple classification data. These two types of data were later trained via a CNN. The trained model was saved and used in the sliding window for damage detection in binary classification and multiple classification models. In the last step, the trained model is deployed using a flask micro framework. Compared with other WTCM methods, the use of drones reduces the safety risk of manual inspections. Moreover, the CNN can reach accuracies of 94.94% in binary fault classification and 91% in multiple fault classification. In addition to the CNN, other ML methods, such as long short-term memory (LSTM), deep neural networks (DNNs), and support vector machines, are used for WT blade condition monitoring [102-104].

Generator

As described previously in Section 2.2, the failure modes of a generator can be classified into two categories: electrical and mechanical faults. The electrical faults of the generator mainly include the open/short circuit of the stator phase winding and a broken rotor bar. Mechanical faults primarily involve failed bearings, unbalanced rotors, misaligned rotors, and bowed rotors. These faults result in changes in the vibration signals, torque, temperature, and current signals [105]. The contributions of the components in induction machines to the failure rates are as follows: bearing, 40%; stator, 38%; rotor, 10%; others, 12% [106, 107].

The difficulty of vibration analysis of generator bearings is the extraction of fault features from nonstationary and noisy raw signals. To solve this problem, time-domain analysis, frequency-domain analysis, and time–frequency-domain analysis

are widely adopted. Chen et al. proposed a method based on an empirical wavelet transform for diagnosing faults in a WT generator bearing [108]. First, the abnormal vibration data were extracted from the WT drivetrain. Then, the wavelet spatial neighbouring coefficient was used for denoising with a data-driven threshold. Subsequently, an empirical wavelet transform was deployed to extract the fault features. Finally, a diagnosis decision based on the fault features was made.

In addition to the wavelet transform, stochastic resonance (SR)-based methods have been proposed for denoising. The first step of the SR methods is to shift the original signal to the target frequency with the carrier signal. Then, parameter tuning methods, such as discrete wavelet transform and quantum particle swarm optimisation, are used to detect weak fault signals [109, 110]. Li et al. proposed a method based on coupled bistable systems for improving the SR filtering performance [111]. It was effective for suppressing low- and high-frequency interferences.

In addition to the aforementioned methods, empirical mode decomposition, Hilbert transform, variational mode decomposition, and integral extension local mean decomposition have proven to be capable of denoising raw signals and localising faults [112-114].

When a fault occurs in the generator bearing, the fluctuation in the WT shaft can result in a stator current at the bearing defect frequencies. Hence, a fault in the WT generator can be detected from the electrical signals. Shahriar et al. acquired the current signal from a generator to estimate the bearing rotation speed and resampled the nonstationary current signals. Subsequently, the fault type can be determined in the sequence according to the resampled signals [115].

Apart from the conventional methods, ML methods have attracted increasing attention in the past decade. Ziani et al. [116] compared an ANN and a genetic algorithm (GA) for the detection of the early stage of bearing degradation through vibration analysis. Both the normal and defective vibration signals were processed for feature extraction. One normal and four different fault class levels were set for the ANN classifier. The GA was used for feature selection, and the trace criterion from linear discriminant analysis was used as the evaluation function. The ANN was used to train a subset of the experimental data for the known machine conditions.

The remainder of the dataset was used for testing. Unala et al. [117] developed a fault-estimation method based on an ANN that aimed to extract features for bearing faults. In this method, a GA is used to optimise the model. Ma et al. [118] combined kernel independent component analysis (KICA) and the least-squares support vector machine (LS-SVM) to classify bearing faults. The KICA was used to map the vibration signal, and the LS-SVM was used for feature vector classification.

In addition to ML algorithms, deep-learning algorithms have been adopted. Shao et al. [93] improved a convolutional deep belief network (CDBN) with compressed sensing (CS) for feature learning and fault diagnosis of generator bearings. First, CS was employed to reduce the vibration dataset for increasing the analysis efficiency. Then, Gaussian visible units were used to enhance the feature-learning ability of the CDBN model for the compressed data. Finally, the exponential moving average technique was adopted to improve the generalisation performance. Experimental results indicated the effectiveness of the proposed method. Sun et al. [92] proposed a method in which a DNN is used for feature extraction and classification. A stacked sparse auto-encoder-based deep neural network was designed with an unsupervised learning procedure followed by a supervised fine-tuning process. Their methods are more effective than the conventional method for handling large amounts of data.

Gearbox

According to Figure 2.6, the gearbox has the longest downtime among all subassemblies in both onshore and offshore turbines owing to the complexity of its structure. Both condition monitoring and SCADA data were used to develop the WTCM methods.

Vibration analysis is also an effective method for monitoring the WT gearbox conditions. However, the noise in the vibration signal makes fault diagnosis difficult. Researchers have conducted a series of studies in recent years to solve these problems. Feng and Liang proposed the iterative atomic decomposition thresholding (IADT) method for extracting the true constituent components of complex signals [119]. Thus, the background noise interference is suppressed. This method was designed for planetary gearboxes under constant running conditions. In the IADT method, the Fourier dictionary is used to match harmonic waves in the frequency domain and localise the gear fault characteristic frequency. Numerically

simulated, laboratory (experimental), and onsite collected signals have proven that the proposed method can localise faults accurately in the WT gearbox. Feng also proposed a time–frequency-domain analysis method based on the Vold–Kalman filter and higher-order energy separation for WT planetary gearbox fault diagnosis under non-stationary conditions [120]. Owing to the high adaptability of local signal changes, higher-order energy separation is used to estimate the transient frequency. The Vold–Kalman filter is used to decompose an arbitrarily complex signal into mono-components.

He et al. proposed a novel order-tracking method based on a discrete spectrum correction technique for WT gearbox vibration analysis [121]. The shaft rotating speed is identified from the vibration signal with the amplitudes of the gear meshing components using the proposed method. Additionally, the inherent shaft misalignment in the fixed-shaft gear is revealed.

Li et al. proposed a method for the detection of gear cracks in a WT gearbox through a supervised order-tracking bounded component analysis of vibration signals [122]. In the proposed method, order tracking is incorporated into a bounded component analysis framework to eliminate the noise and disturbance in the signal. An autoregressive model was built to supervise the reconstruction of the crack-vibration source signature. Owing to the dependence tolerance ability of the BCA framework, the supervised order tracking bounded component analysis can recognise the interfering vibration sources that are dependent on or correlated with the crack. Thus, the vital information that related to the fault that can be preserved in the reconstructed signal.

Kong et al. used spectral kurtosis and time wavelet energy spectra to extract the fault features of a WT gearbox planetary gear [123]. First, the spectral kurtosis and kurtogram of the vibration signals were calculated. The computed kurtosis and kurtogram were used to optimise the filtering parameters for the following bandpass filter. The optimal frequency band with the highest spectral kurtosis was applied to extrude periodic transient impulses. In the final step, the Morlet wavelet was selected as the mother wavelet to implement the time-wavelet energy spectrum analysis. Hence, the fault features of the gearbox planetary gear were extracted.

In addition to fault feature extraction methods and denoising methods, multi-fault detection methods have been studied. A complex wavelet transform-based method was proposed by Teng et al. for WT gearbox multi-fault detection and failure analysis [124]. Distinct harmonics representing the gearbox broken teeth were detected via the Hilbert transform and narrow-band filtering. The cepstrum method was employed to distinguish the different frequency components. A multiscale enveloping spectrogram provided by a complex wavelet transform was used to demodulate and decompose the signals. By analysing multi-scale enveloping spectrograms at different scales, weak fault features can be discovered from intensive energies spectrums.

With the development of ML techniques in recent years, many neural network-based, fuzzy logic-based, and information fusion-based WT gearbox condition monitoring methods have been proposed. A light gradient boosting machine (LightGBM)-based method was proposed for analysing SCADA data and selecting the features of gearbox faults [125]. First, the maximum information coefficient is utilised to analyse the correlation of the WT SCADA data for selecting the fault features. Subsequently, a performance evaluation criterion is used to improve the LightGBM. A confusion matrix was used to evaluate the performance of the proposed method. The proposed method was confirmed to have a lower false-alarm rate than previous by employing a three-year SCADA dataset from a wind farm located in Southern China.

Owing to the characteristics of LSTM, it lacks the ability to handle discrimination tasks. To solve this problem, Yin et al. used LSTM incorporating cosine loss to diagnose WT gearbox faults [126]. By adopting cosine loss, the loss can be converted from the Euclidean space to the angular space. Thus, the effect of the signal strength can be eliminated, and the diagnostic accuracy can be improved. Subsequently, the energy sequence features and wavelet energy entropy of the vibration signals were used to evaluate the Cos-LSTM network.

In addition to the application of CMS, various combined methods have been proposed. Hu et al. developed a method that uses the time-domain sequence approximate entropy (ApEn) adaptive strategy incorporating a wavelet packet transform filter and a cross-validated particle swarm optimised kernel extreme learning machine to improve the accuracy of WT gearbox fault diagnosis [127].

Chang et al. proposed a concurrent convolutional neural network for adaptive learning to resolve the decline in the discrimination accuracy [128].

In addition to ML algorithms, fuzzy logic-based methods have been employed to diagnose gearbox faults. Qu et al. used non-singleton fuzzy logic and expanded the linguistic terms to detect gearbox faults [129]. By generating fuzzy inputs, a non-singleton fuzzy inference system can be constructed for WT gearbox fault detection. The expanded terms and rules can provide information that can help detect faults in the early stages. The result is defuzzified and defined as a fault factor, which is used as an indicator to measure the fault severities.

Power converter

Since the 1980s, the penetration of power electronics in WTs has continuously increased. First, only thyristor-based soft starters were used to interconnect WTs. Soon thereafter, they were bypassed, and the generator was connected directly to the grid [130]. In the 1990s, rotor resistance control with a diode bridge and a power electronic switch was primarily used. Eventually, the back-to-back configuration emerged. In the best-seller range of 1.5–3 MW WTs, the use of two two-level voltage source converters in a back-to-back configuration is the most widely adopted solution [131]. However, the power electronics converter is a subassembly with one of the highest failure rates. In addition to downtime-related revenue loss, the failure of the converter can cause substantial secondary damage.

Qiu et al. proposed a method that can accurately detect and locate a WT insulated gate bipolar transistor fault in circuit arms [132]. A wind speed-based normalised current trajectory method was proposed for handling the uncertainty due to the stochastic wind speed. The proposed method used a wind speed-based normalised current trajectory method to detect the open circuit fault of the generator side converter. The phase current's park vectors are firstly normalised. Then the current vectors are used for fault diagnosis by comparing different wind conditions such as step wind and turbulent wind. By comparing the fault current vector pattern to the fault-free pattern, the fault can be discovered and localised.

Jlassi et al. developed a multi-fault detection method for back-to-back PMSG drives [133]. The proposed method is based on a Luenberger observer with an adaptive

threshold to improve the reliability of the diagnosis results. First, the current form factors (CFFs) of the measured current and observed current are calculated, where the CFF equals the phase RMS value divided by the phase average value. When an open-circuit fault occurs, the CFF values exhibit different behaviours. Thus, the CFF is useful for diagnostic purposes. A Luenberger observer is then used to estimate the three-phase generator and grid current. Subsequently, two adaptive thresholds are calculated for the generator-side converter and grid-side converter, which should be higher than the CFF value under healthy operating conditions, because the thresholds must have strong immunity against current and speed transients. Finally, the residuals between the adaptive thresholds and CFF values are calculated. Using this method, open-circuit faults on different phases and different sides (generator side and grid side) can be detected and localised.

The method proposed by Zhao is based on the phase current of the converter [134]. This method was developed based on the absolute normalised Park's vector approach and can be used for multi-open-circuit switch fault detection. In addition to phase current monitoring, methods based on the emitter voltage, humidity, junction temperature, and thermal coupling have been developed for monitoring the condition of power-conversion units [135-137].

In addition to conventional condition monitoring methods for power-conversion units, AI-based WTCM methods have been developed. Zhang et al. proposed a method based on a backpropagation neural network (BPNN) for detecting WT power converter faults [138]. First, the grid-side converter three-phase bridge leg voltages under the fault-free operating condition and different fault operating conditions are acquired. Then, the acquired voltage signals are decomposed and reconstructed using a wavelet transform. Third, the divergence of the reconstructed wavelet signals is amplified to obtain the feature. Finally, the processed signals are used as inputs of the BPNN to produce a decision and classification. This method can classify single and double open-circuit faults.

Xiao et al. developed a deep-learning method for WT converter fault detection [139]. First, the fault indicator variables extracted from the WT SCADA system are selected. Radar charts are then generated from the fault indicator variables. A CNN model was used to extract features from radar charts. The ResNet50 backbone network was combined with an attention octave convolution structure to overcome

the issue of information asymmetry caused by the asymmetry of the sampling method. In addition, the damage to the original features in the high- and low-frequency domains was resolved. The proposed method achieved a fault-detection accuracy of 98.0%, confirming its effectiveness.

Xue et al. used LSTM to diagnose multiple open-circuit faults in a WT power-conversion unit [140]. Before the LSTM network was constructed, the raw current and voltage signals were normalised, and a fault dataset was created. The training set was processed using LSTM, and a loss function was produced. Finally, the trained network was compared with the test set to detect faults. Compared with the LS-SVM, BPNN, and RNN, the LSTM-based method had the highest accuracy and robustness, with a short time delay.

In the Section 2.4, various conventional and machine learning-based condition-monitoring techniques are reviewed with regards to their capabilities. The condition monitoring methods for WT components are divided into two categories: intrusive (including vibration analysis, oil debris, and shock pulses) and non-intrusive (including ultrasonic testing, visual inspection, AE, thermography, and power signal analysis). The key subsystems of the WT include the nacelle, rotor, gearbox, generator, bearing, and blades. Table 2.2 presents the WTCM methods and their applicable areas.

Table 2.2 Condition monitoring techniques for WTs

		Nacelle	Rotor	Gearbox	Generator	Bearing	Blades
Intrusive	Vibration Analysis	✓	✓	✓	✓	✓	✓
	Oil debris			✓	✓	✓	
	Shock Pulse					✓	
Non-intrusive	Ultrasonic testing						✓
	Acoustic emission			✓		✓	✓
	Thermography			✓	✓		✓
	Power signals		✓	✓	✓		

2.5 WT maintenance strategies

2.5.1 Reliability of WT systems

The reliability of WT systems is a significant factor in wind energy projects. The reliability of a WT system can significantly affect a project's revenue stream by influencing the operation and maintenance costs and downtime. For evaluating the

reliability of the WT system, a mathematical description of the relationships between the various reliability functions is presented below [141, 142]. We consider N_0 identical WT components to be tested.

$$N_s(t) = number\ surviving\ at\ time\ t \tag{2-1}$$

$$N_f(t) = number\ failed\ at\ time\ t \tag{2-2}$$

Therefore,

$$N_s(t) + N_s(t) = N_0 \tag{2-3}$$

At any time t, the survivor or reliability function $R(t)$ is given as

$$R(t) = \frac{N_s(t)}{N_0} \tag{2-4}$$

The unreliability function $Q(t)$ is given as

$$Q(t) = \frac{N_f(t)}{N_0} \tag{2-5}$$

where $R(t) = 1 - Q(t)$.

The failure density function $f(t)$ is given as

$$f(t) = \frac{1}{N_0}\left(\frac{dN_f(t)}{dt}\right) \tag{2-6}$$

The failure intensity function $fa(t)$ is given as follows:

$$fa(t) = \frac{1}{N_s(t)}\left(\frac{dN_f(t)}{dt}\right) \tag{2-7}$$

$$fa(t) = \frac{1}{R(t)}\left(\frac{dR(t)}{dt}\right) \tag{2-8}$$

The failure density function normalised by the number of survivors is shown in Figure 2.8.

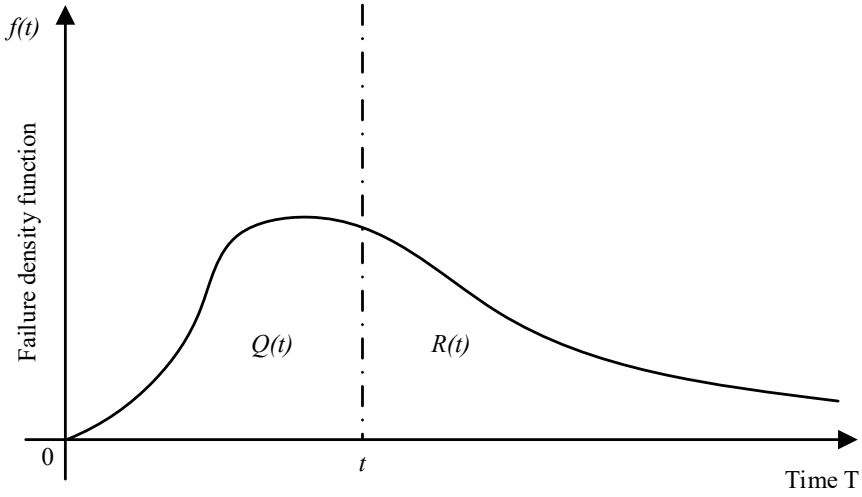


Figure 2.8 Failure density function with respect to time, indicating the reliability $R(t)$ and $Q(t)$

For any reparable machinery, the intensity function indicating the reliability varies throughout the life ('bathtub curve'), as shown in Figure 2.9, where β is a dimensionless parameter determining the shape of the curve.

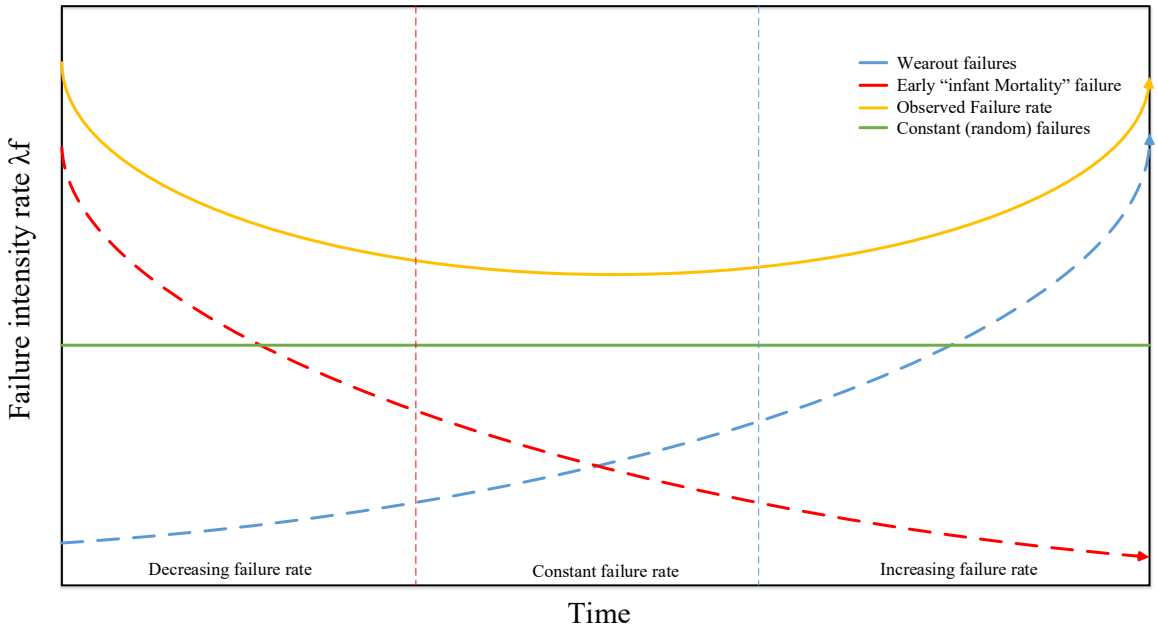


Figure 2.9 Bathtub curve

The bathtub curve is derived from the power law process (PLP), which is a special case of the Poisson process with a failure intensity function:

$$fa(\tau) = \rho\beta t^{\beta-1} \tag{2-9}$$

where ρ is a scale parameter with unit year^{-1}. For $\beta < 1$ or $\beta > 1$, the bathtub curve exhibits an upward or downward trend. When $\beta = 1$, the intensity function of the PLP is ρ. This can be observed at the bottom of the bathtub curve, which is called the intrinsic failure phase. *fa* represents the average failure rate.

To enhance the reliability of the WT system, actions can be planned or performed at different time periods, including the design, commissioning, and maintenance stages.

2.5.2 Maintenance strategies

The conventional maintenance theory is either corrective or scheduled. The corrective maintenance is performed when the fault is detected of the WT's components, which can be caused by many reasons, such as components fatigue, unreliability design and environmental operational factors. The corrective maintenance is implemented during an inspection of a WT or when the turbine shuts down due to a fault. The unscheduled downtime of WT will happen due to the maintenance being scheduled after the faults. Thus, the O&M cost of corrective is the highest among all maintenance strategies. The scheduled maintenance is also known as the time-based maintenance strategy was carried out by repairing at fixed time intervals which is usually recommended by the supplier. The fatigue components can be replaced before failure [143]. The processes of the two maintenance strategies are shown in Figure 2.10.

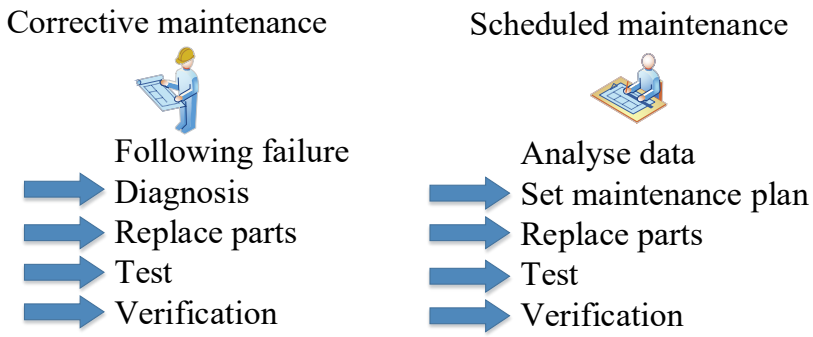


Figure 2.10 Processes of corrective and scheduled maintenance

The scheduled maintenance strategy is also known as periodic-based maintenance. It is based on failure time analyses [144]. The failure rate as shown in Figure 2.9 is estimated based on failure time data or used-based data.

Figure 2.9 also shows the three-phase of the failure rate in the WT life span. In the early stage of a WT, the failures tend to be the combination of inadequate quality and incorrect installation and maintenance. After the early useful stage, the failure rate will be reduced to a minimum as the faults are often caused by the unreliably inherent in the design. By the end of useful age, the failure rate will rise because of the fatigue of the components.

Scheduled maintenance strategy can indeed reduce un-scheduled downtime. However, setting maintenance tasks more frequently than usual would increase the O&M cost since the replaced components have not reached their full useful life.

Thus, the condition-based predictive maintenance strategy is taken into consideration to mitigate major component failures. This strategy includes a whole set of data acquisition, data processing and data analysis system in order to produce optimal maintenance actions [145]. The maintenance tasks can be scheduled in advance of faults [146]. By adapting the condition-based preventive maintenance strategy, un-scheduled time and extra maintenance tasks are prevented. Hence, the O&M cost is reduced.

2.5.3 Condition-based predictive maintenance

The condition-based predictive maintenance has drawn more attention in recent years because of the widely use of ML algorithms in condition monitoring. Compared with corrective maintenance and scheduled maintenance, the condition-based predictive maintenance has the lowest O&M cost. Based on the Energy Roadmap 2050, the European electricity would be supplied by wind energy from 31.6% to 48.7% [147]. Owing to the large energy market, it is necessary to reduce the O&M cost. The location of the offshore wind farms is usually deployed in remote areas for richer wind resources, which caused more difficulties in maintenance [148]. Therefore, offshore wind farms are constantly under cost optimisation pressure. Hence, it is vital for operating companies to carry out suitable predictive maintenance strategies in order to increase the useful lifetime for WTs [149]. Using technologies such as real-time processing, storage and analytic queries in the WTCM would carry out a more reliable predictive maintenance strategy to the energy markets. The time-series data acquired from the WTs are multi-dimensional that

need a precise modelling method to predict faults [150]. The condition-based predictive maintenance is able to gather necessary information from CMS to analyse the status of the components in order to prevent major failures from happening [151]. Condition-based maintenance strategies minimise the O&M cost and increase WTs reliability, while the monitoring devices require extra costs. Many condition monitoring techniques are applied to monitor and inspect the components in a WT [152].

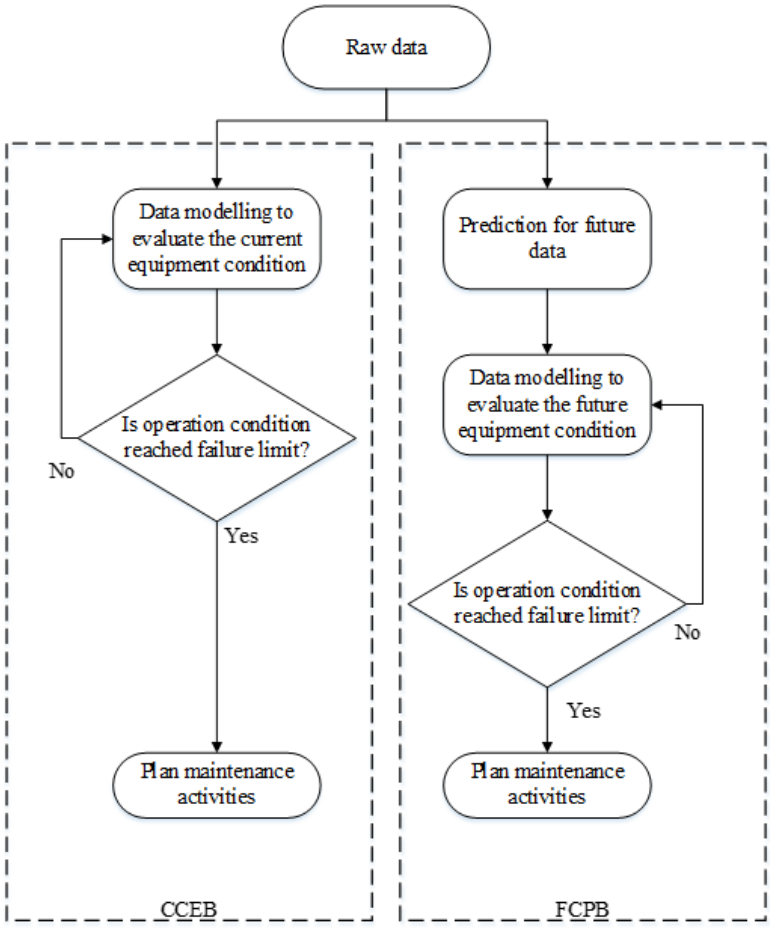


Figure 2.11 Typical decision framework of CCEB and FCPB

The maintenance decision can be classified into two: diagnosis and prognosis. The main aim of diagnosis is to provide early warning signals to the user when the turbine has abnormal behaviours, while the prognosis is to provide further information that the turbine might fail. Compared with the diagnostic process, the prognostic process is superior because it can prevent failures in advance thus saving the O&M costs. Maintenance decision making via condition-based predictive maintenance is implemented by two methods: current condition evaluation-based (CCEB) and

future condition prediction-based (FCPB) [153]. The major difference between the two decision-making methods is that CCEB more focuses on the current state while FCPB focuses on the future state. The decision framework of those two methods is shown in Figure 2.11. Both of the decision methods highly rely on the condition monitoring systems. Maintenance activities will be scheduled as long as the data exceed a certain threshold.

2.5.4 Condition-based predictive maintenance applications

The state-of-arts condition monitoring methods have been introduced in the previous chapters. Various types of fault diagnosis methods are implemented via condition monitoring methods. Another type of condition-based predictive maintenance application is focusing on the process of prognoses and remaining useful life (RUL) estimation. Cheng et al. has proposed a fault prognosis and RUL prediction method for WT gearbox [154]. In the proposed method, the adaptive neuro-fuzzy inference system is used to learn the state transition function of the fault feature. Then the particle filtering algorithm can predict the RUL of the gearbox via the learned state transition function. The effectiveness of this method has been approved by their run-to-failure test. Another case study also showed that a power purchase agreement managed wind farm, the estimation of WT RUL enables predictive maintenance for wind, thus avoiding corrective maintenance that reduces the cost and downtime [155]. Zhang et al. has proposed a fatigue prediction model of the blade to reproduce the fatigue damage evolution in composite blades subjected to aerodynamic loadings by cyclical winds. Then the lifetime probability of fatigue failure of the blades was carried out by stochastic deterioration modelling. At last, a cost-benefit model was built to optimise the maintenance cost [156]. To estimate the RUL of a system, the Prognostics and Health Management (PHM) techniques can be adapted. A turbine with a PHM system is used with a stochastic jump-diffusion model to model the random evolution of the deterioration process as well as production output. The Monte-Carlo simulation is used to find the optimal maintenance data as well as the lowest maintenance cost [156]. Not only is the mechanical part of WT used to estimate their RUL, but the RUL estimation of the electrical system is also necessary. The Gaussian Process regression technique has proposed the estimate the RUL for degraded power insulated gate bipolar transistors (IGBTs). This method is

also compatible with real device accelerated ageing database set under thermal overstress utilising a DC at the gate [157].

As the literature shows that both diagnostic and prognostic/ RUL estimation strategies provide valuable information for condition-based preventive maintenance. On the other hand, scheduling optimization has also been presented by some researchers. Garcia et al. proposed a maintenance system called Intelligent System for Predictive Maintenance (SIMAP) for the WT gearbox. It showed that the SIMAP can adapt the maintenance calendar according to its needs and operating time [158]. Zhong et al. proposed a maintenance scheduling optimization model as a 2-phase solution framework integrating the operational law of fuzzy arithmetic and the non-dominated sorting genetic algorithm. The schedules are derived from the trade-offs between the maximum reliability and minimum cost [159]. Except for the condition monitoring methods for WT components, the labour cost and production loss being used as objective functions are also taken into consideration for maintenance scheduling decision. By analysing historical weather data and producing a statistical model for weather description, the maintenance problem is formulated compactly as a mixed-integer linear programming model. Compared with the periodic preventive maintenance plan, the expected labour cost and production loss are reduced by approximately 30% and 20%, respectively [160]. Other parameters such as maintenance vessel allocation, electrical price and dynamic safe access pre-requisites for WT and crane are also playing a role in maintenance scheduling optimization [161, 162].

In condition-based predictive maintenance, the main idea of decision making is to perform the fault prediction of the components. The faults can be detected if they exceed predetermined limits. Then the maintenance action is decided by an appropriate maintenance scheme. CCEB and FCPB are the two main condition-based predictive maintenance strategies. However, both strategies have their shortcomings in real industrial practice. As a matter of fact, when implementing CCEB, it may not have enough time for maintenance planning if the condition shows the components has already reached the fault limit. FCPB can indeed solve this problem since it predicts the future condition of the components. However, the reliability of short-term predictions is higher than that of long-term ones. When dealing with long-term prediction problems, FCPB might not precise enough. To

produce a reliable maintenance decision, either CCEB or FCPB needs to be chosen carefully for optimal decision.

The increasing number of sensors have improved the measurement accuracy, in the meantime, the redundancy increased, which brought increasing system cost and complexity. The sensor failures and misreporting jeopardise the risk of system reliability. An unnecessary maintenance plan can be scheduled for this. Employing low-resolution SCADA data with high-resolution condition monitoring data is considered as a solution to improve system accuracy. However, it is a challenge to distinguish whether a fault is true or false. Besides, the complexity of the data brought problems to collecting and analysing. New condition monitoring methods such as ML based methods are studied in recent years to improve the robustness and accuracy of the system.

Lacking detail in the existing data collection system is another shortcoming for condition-based predictive maintenance. The reliability, availability, maintainability, and safety (RAMS) databases are carried out to provide more detailed information. These databases provide information on maintenance planning, scheduling optimisation and life-cycle cost minimisation [163].

The last limitation to condition-based predictive maintenance is the data reliability. The data can be lost, noised and hacked during the transmission process. In order to improve the accuracy and reliability of the condition monitoring system, data cleaning and data encryption were also taken into consideration.

Table 2.3 Difference between three maintenance strategies

	Corrective maintenance	Scheduled maintenance	Condition-based predictive maintenance
Initial cost	Low	Medium	High
Operating cost	High	Medium	Low
Unplanned maintenance	High	Low	Medium
Downtime	High	Medium	Low
Level of automation	Low	Medium	High

Compared with corrective maintenance and scheduled maintenance, condition-based predictive maintenance can schedule time for maintenance based on reliability requirements. The differences among the three maintenance strategies are shown in

Table 2.2. The total maintenance cost is evaluated for arranging maintenance actions. By adopting proper condition monitoring methods, the failed components would be replaced. The components which tend to have faults would be also replaced to prevent further damage to the turbine. Thus, the downtime of the turbine is reduced and the O&M cost is minimised.

2.6 Summary

This chapter reviewed the state-of-art condition monitoring technologies of WT. The common failure modes of major sub-assemblies are also discussed. The reliability of the WT system and modern condition-based maintenance strategies are also covered. This chapter is focusing on the WTCM methods. Those methods have been discussed in terms of conventional approaches and machine-learning approaches. Various methods have been discussed based on either in analysing perspective or in sub-assemblies perspective. With the development of ML algorithms technologies, ML based methods are also emphasised. However, those methods were mostly focused on single fault and without sensor fusion. Besides, the use of alarm signals was also omitted. It is necessary to develop a WTCM system that can diagnose the fault with a global view of the turbine. Additionally, the alarm signals can also be used wisely for early fault warnings.

Chapter 3. Wind turbine data interpretation

This chapter presents interpretations of the data used in the research. The first part describes the SCADA data collected from an operational wind farm. The SCADA system is introduced first; then the data used for condition monitoring purpose are explained. The second part describes the simulation models of DFIG and PMSG WTs, respectively. It begins with system descriptions and simulation model constrictions based on MATLAB/ Simulink, where both healthy conditions and faulty conditions are simulated. In faulty condition simulations, three-phase short circuit fault and capacitor breakdown are simulated. The results give a comprehensive understanding of transient changes when the WT experiences different faults. Further analysis of the simulated data will be shown in the next chapters.

3.1 SCADA data exploration

3.1.1 Introduction to SCADA systems

In modern industrial and manufacturing processes, long-distance equipment and systems are necessary for public and private utilities, security, and mining industries. The communication distance can range from a few meters to thousands of kilometres, and telemetry is required for this communication. SCADA involves a combination of data acquisition and telemetry. The required information is collected and transferred via SCADA. This information is then displayed on screens for further analysis, and SCADA systems convey the required actions back to the process [164].

In the early days of data acquisition, relay logic was widely adopted for industrial control. With the development of electronic devices and CPUs, digital electronics have been incorporated into relay-logic equipment. Currently, programmable logic controllers (PLCs) are the most commonly used control systems in industries. The figure below shows a diagram of PLCs and distributed control systems (DCSs).

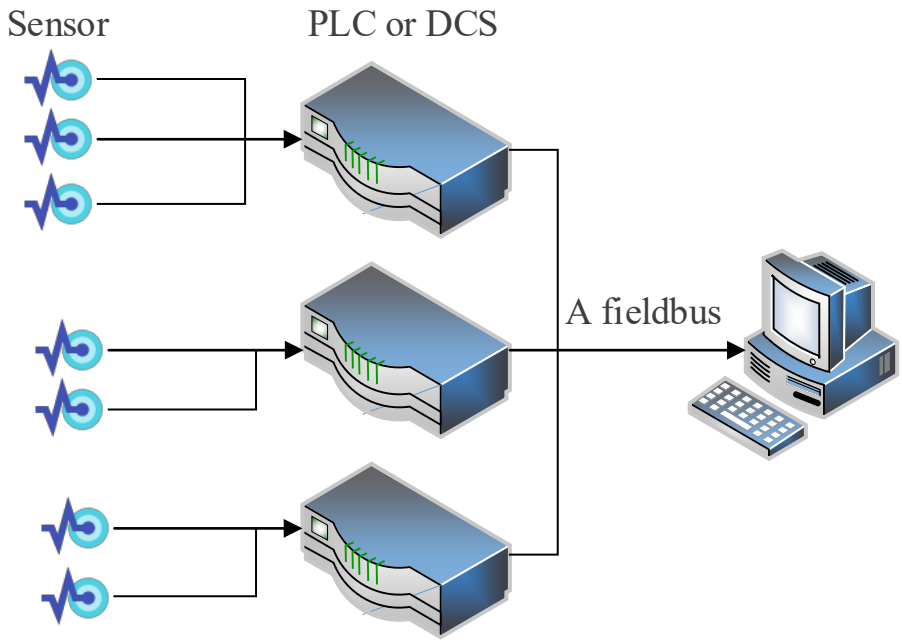


Figure 3.1 Computer to PLC or DCS with a fieldbus and sensor

For PLC- and DCS-based SCADA systems, the computer can record and store a very large amount of data, and the data can be viewed anywhere based on user

requirements. However, to do this, users are required to have multiple complex operating skills. In addition, the resulting large amount of wires are difficult to deal with.

With the development of SCADA systems, smaller and smarter devices have become increasingly popular. Intelligent electronic devices (IEDs) are connected via a fieldbus to the computer. A typical IED combines an analogue input sensor, an analogue output, the memory, the control, and a communication system in a single device [165]. However, there remain problems with IED-based SCADA systems, such as the need for more sophisticated systems, more expensive sensors, and a greater reliance on communication systems.

For a typical modern SCADA system, there are five essential hierarchies in hardware: field-level instrumentation and control devices, marshalling terminals and remote terminal units (RTUs), communication systems, master stations, and data-processing department computer systems. Meanwhile, there are two types of SCADA software: proprietary and open software, and corporations develop their own software to communicate with their own hardware. SCADA software must include features such as a user interface, alarms, scalability, database, fault tolerance, client distributed processing, networking, access to data, RTU interfaces, trends, and graphical display [164].

Figure 3.2 shows a typical SCADA system used to control infrastructure for utilities such as gas, oil, or wind industries. It usually consists of a control centre and a number of field sites. These field sites may be located in distributed areas and connected to the control centre by different communication technologies, such as wide area networks (WANs), radio, cellular networks, or satellites [166]. PLCs and RTUs are installed on the on-site machines to communicate with the control centre.

The control centre consists of a human–machine interface (HMI), server or master terminal unit, and a database management system (historian). The HMI interacts with the operator. The information is then transferred between the field sites, and data is sent to the field devices. The data can then be sent to the historian for archiving purposes.

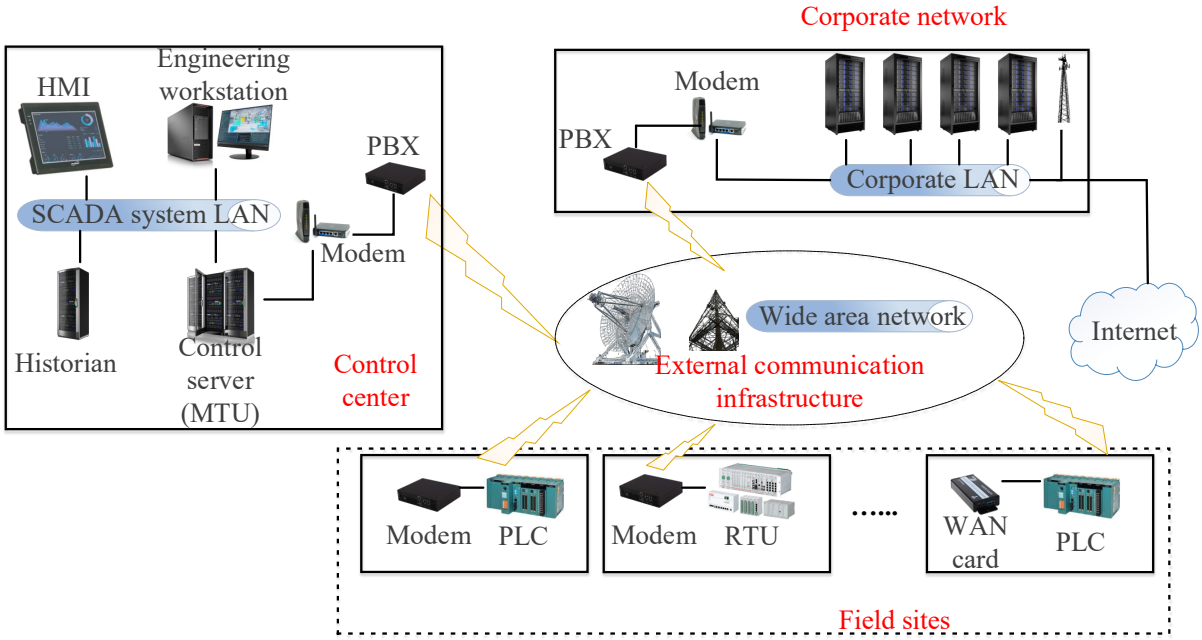


Figure 3.2 Simplified logical view of a typical SCADA architecture

3.1.2 Success of SCADA systems in WTs

SCADA systems have been widely used in industries for accurate fault detection, diagnosis, and prognosis [167]. For instance, with appropriate algorithms, failures such as compressor bleed band failure, fuel supply system faults, and turbine degradation can be automatically detected in turbine engines. Because SCADA data can provide continuous-time measurements, the overall turbine performance can be exploited. The isolation of individual component fault schemes is carried out with performance monitoring. The WT converts wind kinetic energy into useful electrical energy. As the WT components deteriorate, the electrical energy conversion efficiency decreases. This performance degradation indicates the presence of potential problems in the turbine, such as generator winding faults, drivetrain misalignment, pitch control system degradation, or even blade aerodynamic degradation.

Figure 3.3 shows the functional elements of performance monitoring. First, the performance parameter is computed based on sensor measurements. This parameter can be raw data, pre-processed data with calibration from prior knowledge, or residuals based on the WT model. Anomaly detection can be implemented using single or multiple performance parameters. Furthermore, the SCADA data are also

involved in trending prediction wherein parameters or fault indicators are trended, and failure time is predicted.

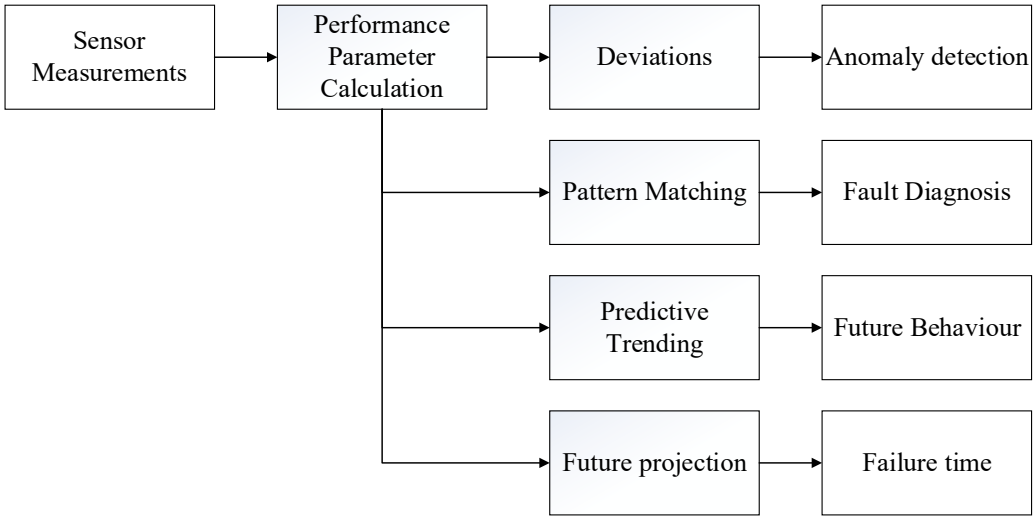


Figure 3.3 SCADA data based monitoring system

3.1.3 Description of SCADA data used in the research

The SCADA data used in the research were acquired from an operational wind farm with 26 turbines. Unlike high-frequency condition monitoring data, SCADA data have a low sampling rate, usually at 10 min/sample, to reduce data storage while still maintaining vital information about the operation and performance of the WTs [52]. The monitoring variables for each turbine consisted of 128 readings from various physical and electrical signals.

Based on the components, the monitoring variables can be classified as the generator, nacelle, yaw, wind, pitch, gearbox, rotor, hydraulic system, main bearing, and electrical system. Monitoring variables such as pressure, temperature, vibration, speed, angle, and electrical signals were collected from the system and sent to the user. The digital constant is sent back to control the turbine. In addition to the 128 monitoring variables, 480 types of alarms and 480 types of events were recorded for further diagnosis. The cut-in wind speed, rated speed, and cut-out speed for the 26 turbines were 3.5 m/s, 14 m/s, and 25 m/s, respectively. The rated power output was 2,500 kW. Figure 3.4 shows a typical power curve of the turbine. Owing to the shape of the power curve, it is also called an ‘S curve’. The turbine started to rotate at a wind speed of 3.5 m/s. When the wind speed was 14 m/s, the turbine reached its

rated output power of 2,500 kW. Finally, to prevent any damage to the turbine, it was stopped at a wind speed of 25 m/s.

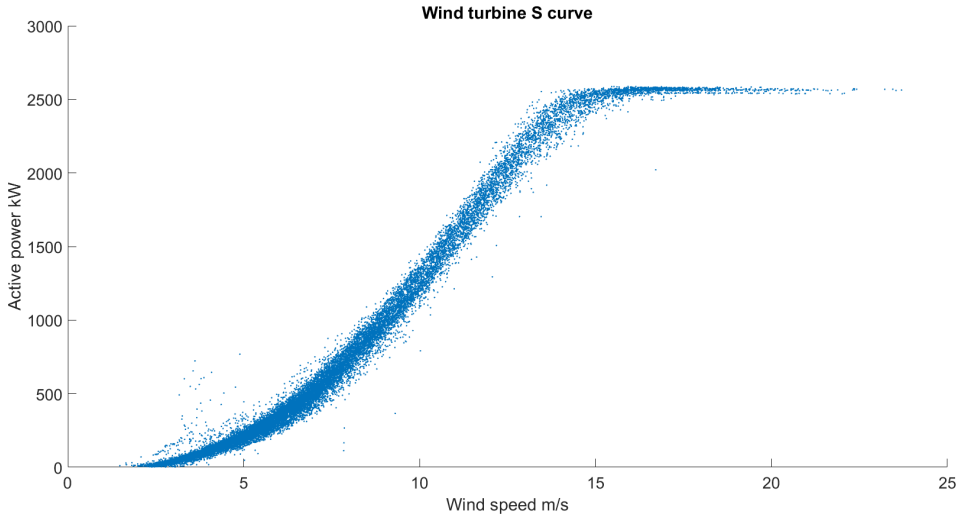


Figure 3.4 WT power curve

As described in the previous subsection, an anomaly can be detected by monitoring the turbine performance. The output power is an indicator that clearly demonstrates the performance of the turbine. Figure 3.5 shows the output power measurements obtained during both the healthy and faulty conditions of the two turbines with faults.

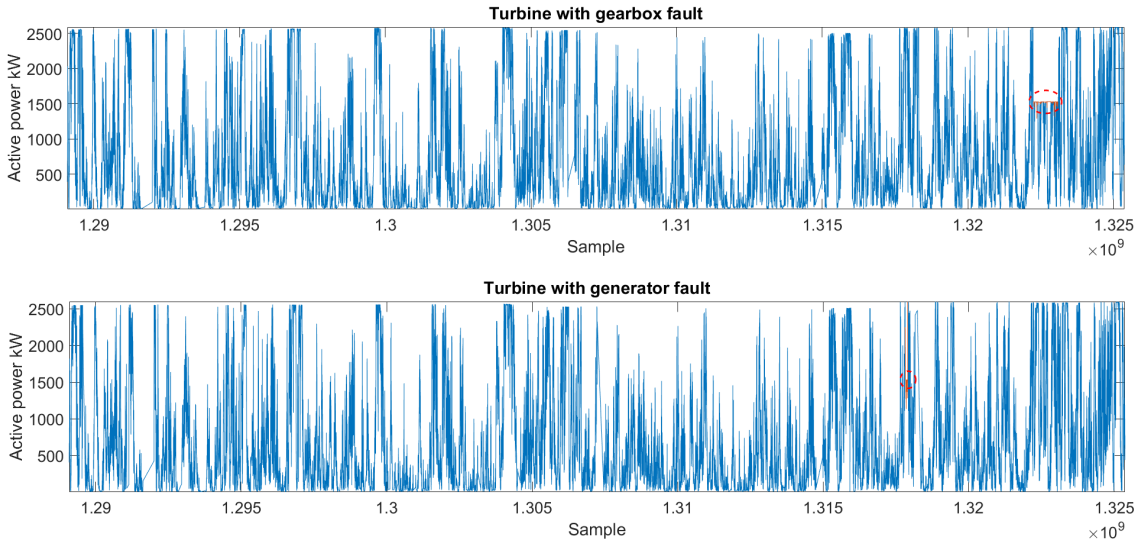


Figure 3.5 Active power measurements of two faulty turbines (Upper: WT1, Lower: WT2)

The upper figure shows the output power within one year of the turbine operation with a gearbox bearing fault, and the lower figure shows the output power within

one year of the turbine operation with a generator winding fault. As can be observed from the figure, the active power changes constantly for most of the time. However, during the faulty time period, the active power is reduced to 1,500 kW to avoid further damage, which can be observed in the red circles in Fig 3.5. The faulty time periods are circled in red for both turbines.

As an example, Figure 3.6 shows the behaviours of the six monitoring variables under both normal and faulty conditions. These six monitoring variables are related to the WT gearbox: gearbox bearing temperature 1 at the main speed shaft bearing connected to the rotor, gearbox oil pressure, gearbox oil heat exchanger output temperature, gearbox oil sump temperature, gearbox oil pressure behind the pump, and gearbox bearing temperature 2 at the high-speed shaft connected to the generator. In the figure, the normal operation time is labelled with a blue line, and the faulty time is labelled with red crosses. As can be observed from the figure, the temperatures tended to be higher than the normal time, and the pressures tended to be lower than the normal time.

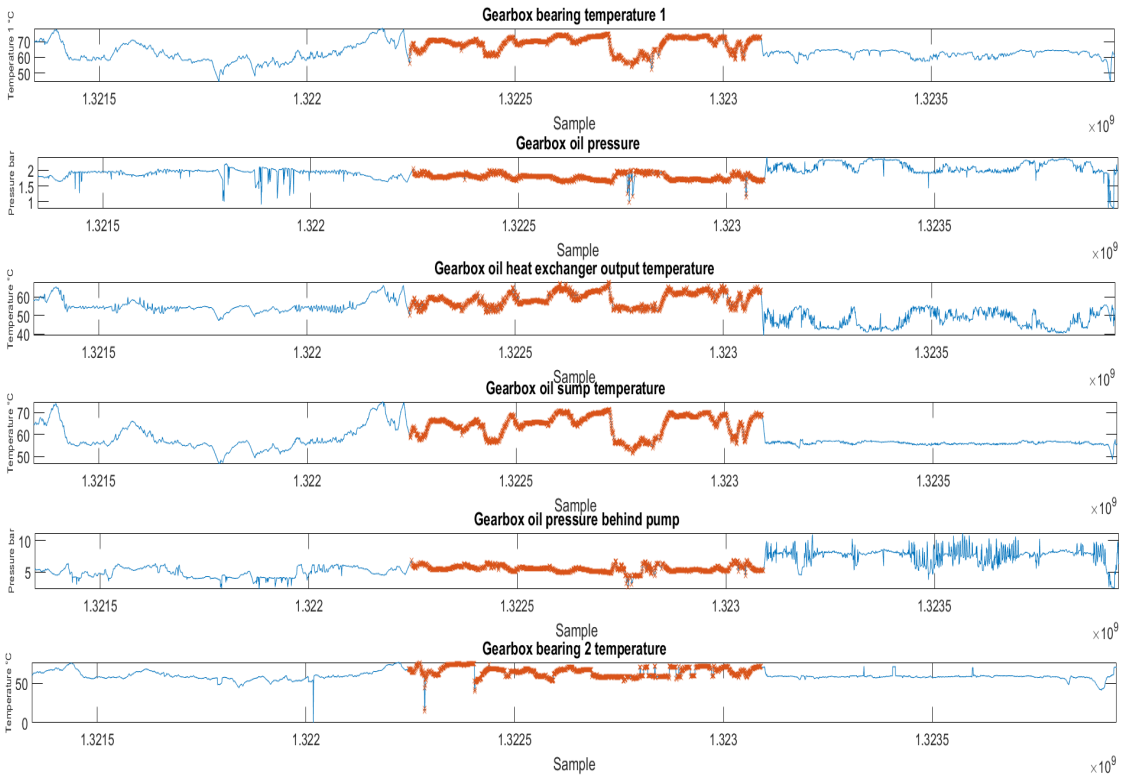


Figure 3.6 Example of gearbox subsystem monitoring variables

Among the 26 turbines, two have an obvious performance anomaly that is easy to observe. The research was primarily based on historical SCADA data. Fault diagnosis models have mostly been developed based on faulty turbines in operational wind farms. The details of the turbines and the fault diagnosis methods are introduced in the following sections.

3.2 Modelling of WT simulation

The main aim of building a WT simulation model is to simulate the different operating conditions of WTs and acquire relevant data. Compared with existing SCADA data, the simulated data have a higher sampling rate, which provides a deeper understanding of the dynamic behaviour of WTs. In addition, the simulation model can simulate different types of faults that SCADA data do not have. Two types of WTs are commonly used: PMSG based variable-speed WTs and DFIG-based variable-speed WTs. Owing to the PMSG power conversion unit test rig being built in practice, for which the details can be found in Chapter 7, in the simulation process, the focus is on the DFIG-based WT.

3.2.1 DFIG WT model

The DFIG is a type of induction generator, which allows the generator to run slightly higher or lower than the natural synchronous speed. Owing to sudden changes in the wind speed, it is considered useful for large variable-speed WTs. The speed of the synchronous generator is locked in synchronous speed with the frequency of the grid with which they are connected, which resists acceleration when a gust of wind hits the turbine. This causes damage to the mechanical systems; thus, an asynchronous generator was used to prevent this problem. As a result, the stress on the mechanical components was lower, and the power from the wind gust could also be converted to electrical power.

In this DFIG simulation model, a WRIG was adopted. Figure 3.7 shows the structure and key components of the wound-rotor induction generator. The wound-rotor induction generator is a type of asynchronous generator in which the rotor windings are connected to an external resistance via slip rings to adjust the speed/torque of the generator. Compared with the squirrel-cage generator, the wound-rotor induction generator has more winding turns, which results in a higher induced voltage and a

lower current. While the stator of the generator is connected to the grid directly, the rotor is connected to the grid via a power converter. Figure 3.8 shows the equivalent circuit of the wound-rotor induction generator. If the generator speed exceeds a threshold, the converter converts the excess energy to the grid, whereas the stator feeds a constant 60 Hz to the grid directly in this study. When the rotor is running under speed, the converter transfers power from the grid to the rotor.

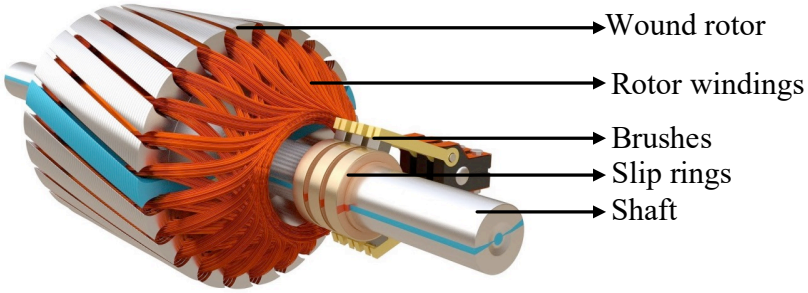


Figure 3.7 Structure of WRIG [168]

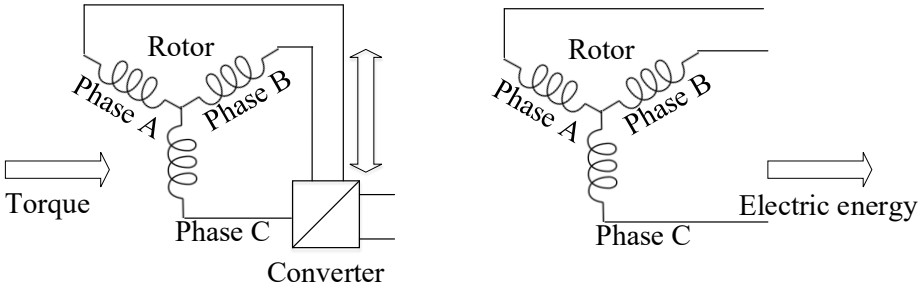


Figure 3.8 Equivalent circuit of WRIG

For the purpose of simulation, only the interface between the WT generator electrical controller and the network was provided in the generator model; user settable and control functions were not included. Additionally, mechanical state-related variables were not included. Compared to conventional generator models, the electrical/flux state variables are simplified as algebraic equivalents. The generator model is thus considered an algebraic, controlled current source that produces active and reactive power in the network.

3.2.1.1 Fundamentals

The DFIG is a popular alternative solution for variable-speed WTs. Unlike PMSGs, DFIGs use wound-rotor induction machines. The stator winding of the DFIG is connected directly to the grid, while the rotor is connected to the AC-DC-AC

converter. Compared with PMSG, the manufacturing and assembling costs can be high because of the complicated structure and enormous size of the generator. However, the DFIG does not require permanent magnets, which reduce the manufacturing and assembling cost. Besides, the AC-DC-AC converter of the DFIG only requires approximately 30% of the nominal generated power [169].

A DFIG-based WT was simulated to produce the dynamic behaviour of the system. Figure 3.9 demonstrates the structure of the simulation model, which was designed based on the dynamic modelling of the GE 1.5 and 3.6 MW WT generators [170]. To construct a fully functional WT model, models of the generator, electrical control, turbine, wind power, and power converter control were constructed.

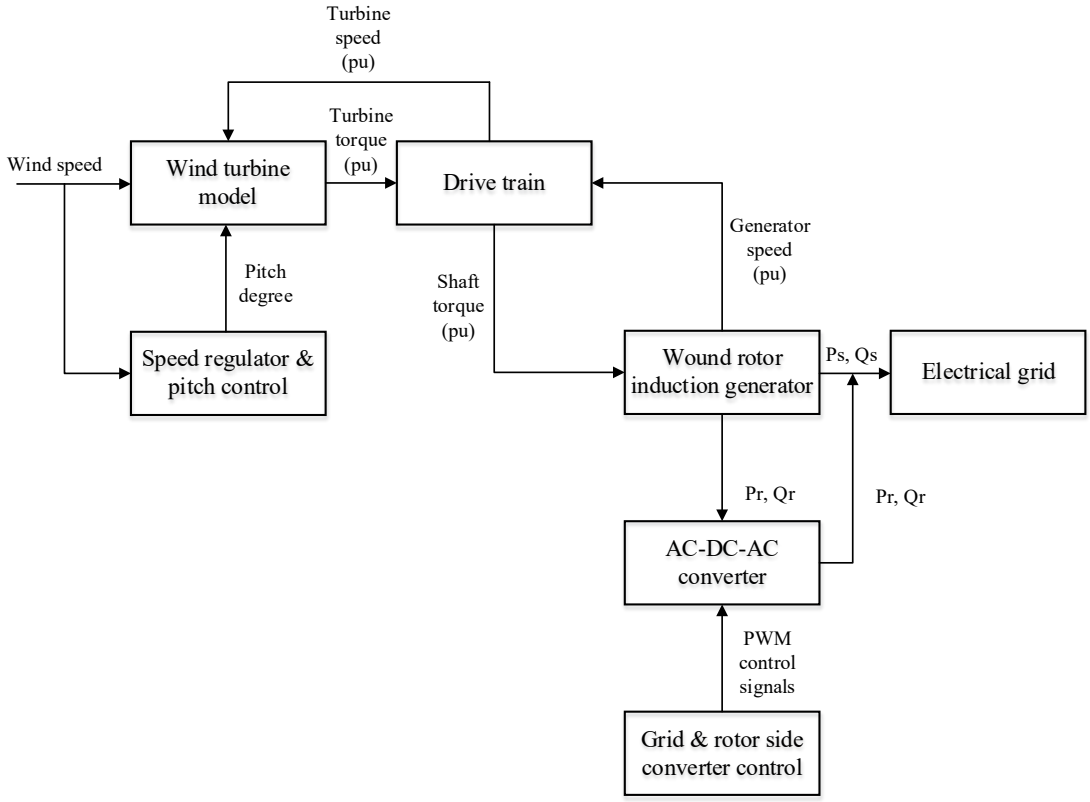


Figure 3.9 Overall structure of DFIG simulation

Besides the WT modelling, the load flow analysis is also simulated. The generator is connected to a 575 V PV bus which is connected to the power system through a step-up transformer. The bus on the generator terminal side is connected to the collector system bus via a rated transformer. The generator has rated power of 1.5 MW with reactive power capability of each individual machine is 0.9 pf under-excited and 0.95 pf over excited. The supervisory control is also applied to regulate collector bus voltage to a specific level.

3.2.1.2 Generator model

3.2.1.3 WT electrical controller model

The electrical controller model delivers active and reactive power to the system based on the inputs from the turbine model (P_{ord}) and supervisory voltage-ampere reactive (VAR) controller (Q_{ord}). Q_{ord} can also be obtained from a dynamic voltage-ampere reactive (DVAR) function. The overall controller model consists of three control functions: wind VAR emulation, open-loop control logic, and an electrical controller. Figure 3.10 shows the overall structure of the electrical controller model.

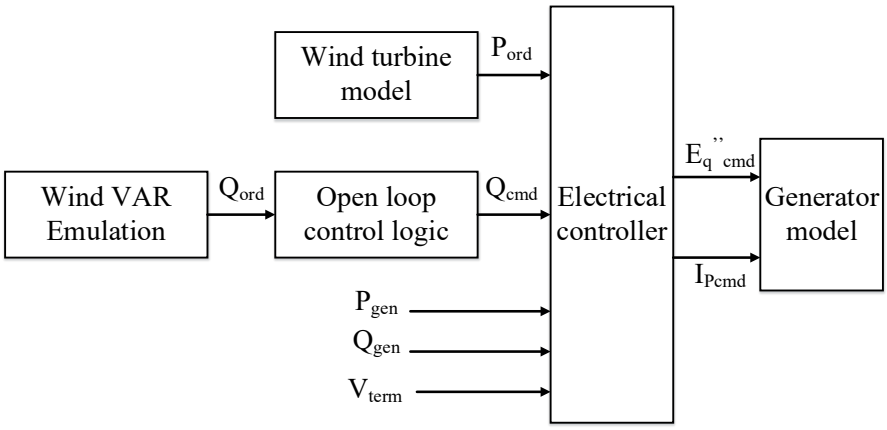


Figure 3.10 Overall electrical controller model

The wind VAR emulation is a simplified equivalent of the supervisory VAR controller that is used for the entire turbine. It can also be considered as a closed-loop controller. The bus voltages were monitored and compared with the reference voltage, which is the initial voltage of the regulated bus. The regulator can be modified by compensating for the impedance, which can be omitted or selected by the user. The regulator then produces a reactive power modulation to achieve power factor control. The initial power factor was set based on the initial conditions. However, as the reactive power output changes, it becomes necessary to maintain a specific power factor. The reactive power order was then transmitted to the open-loop control for further processing.

As shown in Fig. 3.10, the open-loop control logic block in the model is aimed at improving system performance when system events trigger large voltage deviations, and it is usually disabled when the terminal voltage is within the normal range. Open-loop control forces the reactive power to be determined at pre-specified levels

while maintaining the voltage deviations. Similar to other open-loop controllers of this type, hunting is prevented by hysteresis. When the voltage threshold is exceeded and the open-loop reactive power command is issued, the threshold voltage is levelled up by a certain amount. The reactive power command (Q_{cmd}) is further processed in an electrical controller alongside the generator active power (P_{gen}), generator reactive power (Q_{gen}), and terminal voltage (V_{term}) from the WT model.

The electrical controller is designed as a simplified model of the excitation system to produce the voltage and current commands Eq''_{cmd} and I_{pcmd}, respectively, by monitoring the generator reactive power Q_{gen} and terminal voltage $V_{term.}$ Q_{gen} is regulated by the voltage command $Eq''_{cmd,}$ which is computed by the integral of the error between Q_{gen} and $Q_{cmd.}$ The power order P_{ord} from the WT model is used to compute the current command $I_{pcmd.}$

3.2.1.4 WT model

The WT model was used to simulate the very complicated electromechanical system in a simple manner. All related mechanical dynamics and relevant control strategies were simulated in the WT model. The rotor speed was obtained by determining the electrical power from the generator model and the mechanical power from the wind power model. A two-mass rotor model was adopted to transfer the shaft torque from the turbine torque, as shown in Figure 3.11. All shaft parameters are referred to high-speed shaft (generator shaft). The wind turbine torque and generator speed are used as inputs to produce the transmitted torque and speed of the shaft.

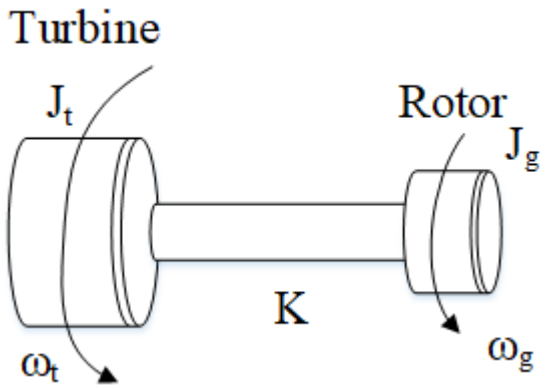


Figure 3.11 Two-mass rotor model

The equations regarding the dynamics of the employed two mass rotor model are given as below:

$$2J_t \frac{d\omega_t}{dt} = T_m - K\theta - D\omega_b(\omega_t - \omega_g) \tag{3-1}$$

$$-2J_g \frac{d\omega_g}{dt} = K\theta + D\omega_b(\omega_t - \omega_g) - T_e \tag{3-2}$$

$$\frac{d\theta}{dt} = \omega_b(\omega_t - \omega_g) \tag{3-3}$$

where the Tm and Te are the mechanical and electromagnetic torques. Jt and Jg are the inertia constants of the wind turbine and generator rotor. K is the shaft stiffness. θ represents the rotating angle. D is the shaft damping coefficient. ω_t, ω_g and ω_b are the turbine shaft, rotor, and electrical base speed, respectively.

Practical WT control aims to maximise power delivery over a range of wind conditions. When the input wind power exceeds the power rating of the turbine, the mechanical power (P_{mech}) is reduced to the equipment rating by pitch control. When the input wind power is lower than the rated power, the blades are fixed to maximise the mechanical power, and torque control is employed by enforcing the turbine operating speed as the specified speed reference. The pitch compensation block compares the current pitch status with reference angle to produce pitch angle compensation. A block diagram of the control strategy is shown in Figure 3.12.

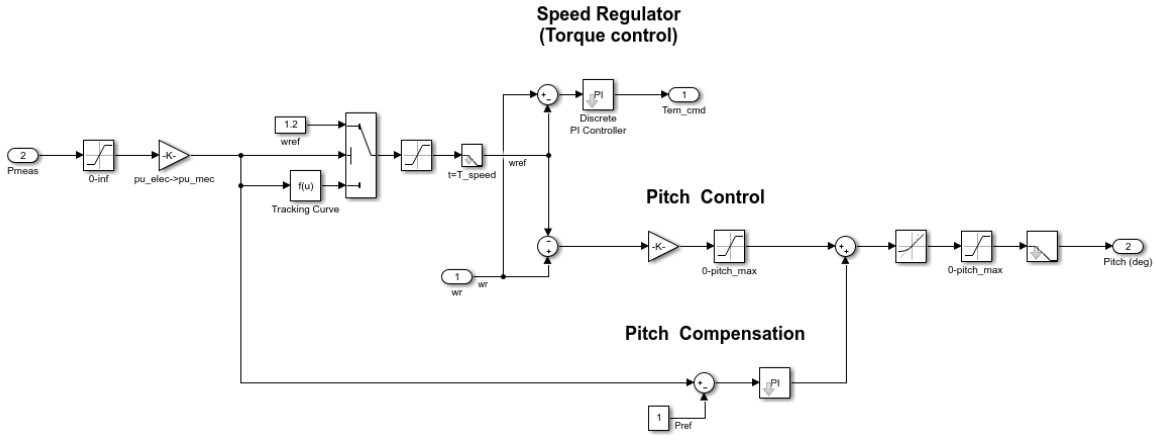


Figure 3.12 Turbine control strategy model

The electrical control system receives a power order from the turbine control model and then requests the converter to deliver power to the grid. Electric power is delivered to the grid and is used to determine the rotor speed set point. In dynamic control, blade pitch control and electric power order are combined into two different operating mechanisms. The turbine speed is primarily controlled by the electric

power order of the specified speed reference. The rotor speed is primarily determined by pitch control when the power levels are above the rated level.

The reference speed is normally 1.2 pu. However, it is reduced when the power level is below 75%. The reference angular speed for power below 0.75 pu is regulated by the following equation:

$$\omega_{ref} = -0.0.67P^2 + 1.42P + 0.51 \tag{3-4}$$

Because the machine speed is allowed to change around the reference speed, the inertia of the machine is considered a buffer to the mechanical power variation.

3.2.1.5 Wind power model

The wind power produced by the WT's mechanical power extracts energy from the wind using the following equation:

$$P_{mech} = \frac{\rho}{2} A_r v_w^3 C_p(\lambda, \theta) \tag{3-5}$$

where P_{mech} is the mechanical power generated, ρ is the air density in kg/m^3, A_r is the rotor blade surface area, v_w is the wind speed in m/s, and C_p is the power coefficient, which is a function related to the ratio λ of the rotor blade tip speed to the wind speed (v_{tip}/v_w) and the blade pitch angle θ. The calculation of λ is as follows:

$$\lambda = K_b(\omega / v_w) \tag{3-6}$$

where the ω is the generator rotor speed and K_b is a fixed constant.

The C_p curve of the designed turbine simulation model can be described as the following mathematical representation:

$$C_p(\theta, \lambda) = \sum_{i=0}^{4} \sum_{j=0}^{4} \alpha_{i,j} \theta^i \lambda^j \tag{3-7}$$

where $\alpha_{i,j}$ is pre-defined coefficients that use for wind power C_p curve.

3.2.1.6 WT power converter control model

A schematic of the DFIG WT power converter is shown in Figure 3.13. Unlike PMSGs, power converters are installed on the rotor side instead of handling the full power rating of the generator, which typically accounts for 30% of the rated power

of the generator. In this model, the mechanical torque Tm is used to produce electric power.

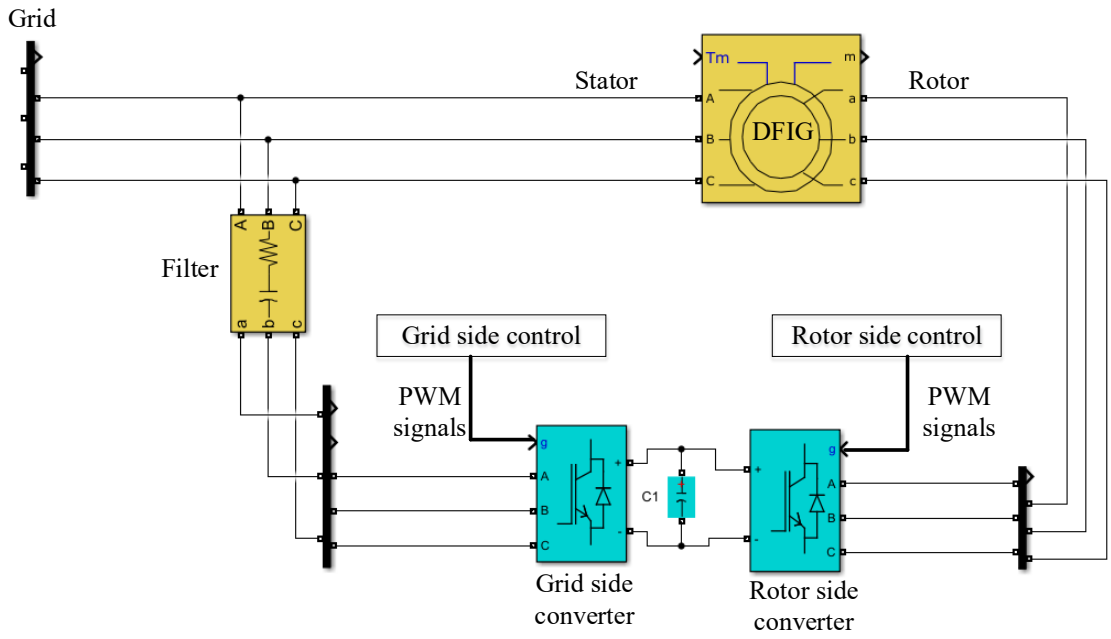


Figure 3.13 Schematic of the DFIG WT power converter

A rotor-side converter controller is used to excite the rotor to achieve active and reactive power control. A block diagram of the rotor-sider converter is shown in Figure 3.14.

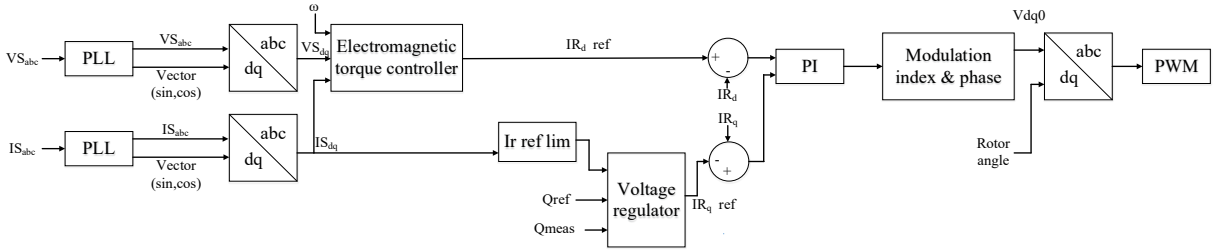


Figure 3.14 Rotor side converter control

Stator voltages and currents were first passed through the phase-locked loops to track their frequency and phase. Then, the three-phase stator voltages and currents are transformed in the direct-quadrature zero (dq0) axis to simplify calculations in the electrical system. Subsequently, the modulation index and phase were calculated to produce the correct pulse-width modulation (PWM) signals to rectify the rotor-side voltages.

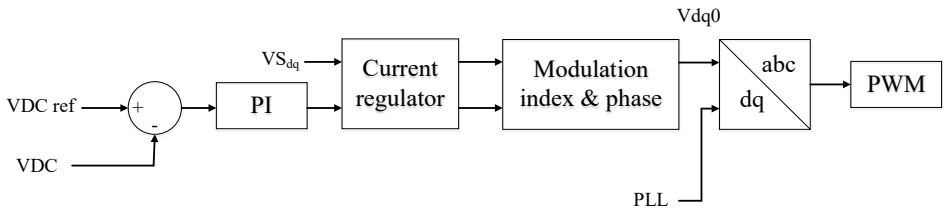


Figure 3.15 Grid side converter control

Figure 3.15 shows a schematic of the grid-side converter controller. The grid-side converter control enables the power electronics to perform the function of the inverter. Space vector pulse width modulation (SVPWM) technology incorporating dq0 transformation was adopted to produce PWM control signals to the IGBTs.

3.2.2 PMSG WT model

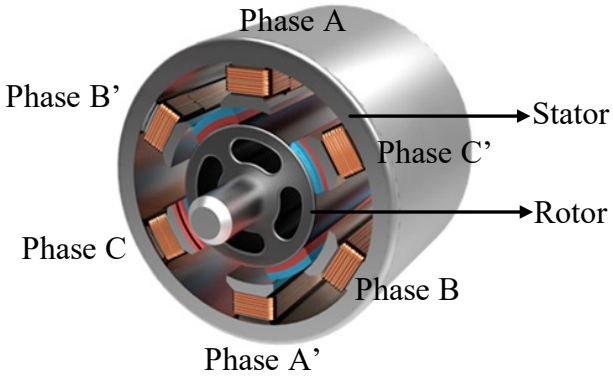


Figure 3.16 Schematic of two-pole three-phase PMSG [171]

A PMSG WT simulation model was built based on the DFIG WT simulation model. A major modification was the generator structure, while other control system models remained the same. A PMSG is a type of synchronous generator that uses a permanent magnet instead of a coil to provide an excitation field.

Figure 3.16 shows the basic structure of the PMSG [171]. When the WT starts to operate, the mechanical force drives the generator shaft to create a rotating magnetic field in the rotor poles with an angular speed ω_r. An EMF is induced in each stator coil when the magnetic flux cuts the individual stator coil. The three-phase stator and rotors A-B-C and A'-B'-C' are electrically 120° apart.

Mathematical modelling of PMSGs has been discussed in many studies. The Park transformation also known as dq0 transformation is vital for modelling, and the details can be found in the literature [172].

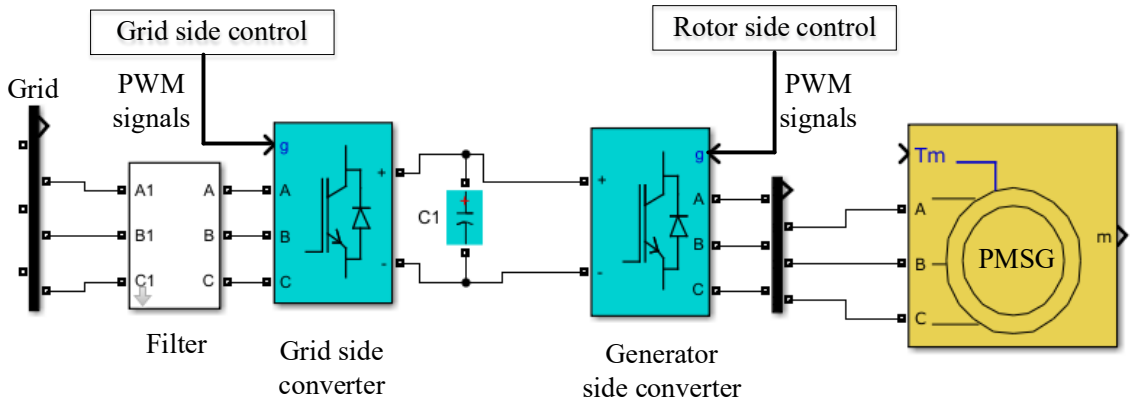


Figure 3.17 Schematic of PMSG power electronic converter topology

Figure 3.17 shows the simulation schematic of the PMSG and its power converter. A power converter is also required for a variable-speed PMSG. Similar to the DFIG, it consists of three major components: a generator-side converter, a DC-link capacitor, and a grid-side converter. The generator-side converter converts the variable AC power output from the generator to DC power and extracts the maximum power, which depends on the generator speed. A DC-link capacitor is used for power storage and transfers energy to the grid-side converter. The grid-side converter is then converted into a three-phase AC that can satisfy grid requirements. Conventionally, a diode rectifier is used as a generator-side converter, and a thyristor inverter is used as a grid-side converter. In this simulation, IGBTs are used to perform both generator-side and grid-side conversions because of the increasing size of WTs and thus satisfy the high-reliability requirements of power switching components. The other settings of the PMSG based WT simulation remain similar to those of the DFIG-based WT simulation, which will not be described here again.

3.2.3 WT simulation model performance

The previous subsection describes the simulation model based on MATLAB/Simulink. The model was built using an industrial turbine with a graphical interface to facilitate the study of transient behaviours of complex electrical networks. The designed model can simulate healthy operation conditions and various types of faults, such as three-phase short circuits and capacitor bank

breakdown. In this subsection, the performance of the DFIG-based WT and PMSG based WT models is presented.

Both turbines have a rated power of 1.5 MW at a rated wind speed of 12 m/s with cut-in and cut-out speeds of 3.5 m/s and 25 m/s, respectively. Dynamic pitch and torque controls were employed, and the pitch angle was set to 0° when the simulation was initiated. The generator produced a peak-to-peak voltage of 575 V at 60 Hz. The nominal voltage of the DC link was 1,200 V.

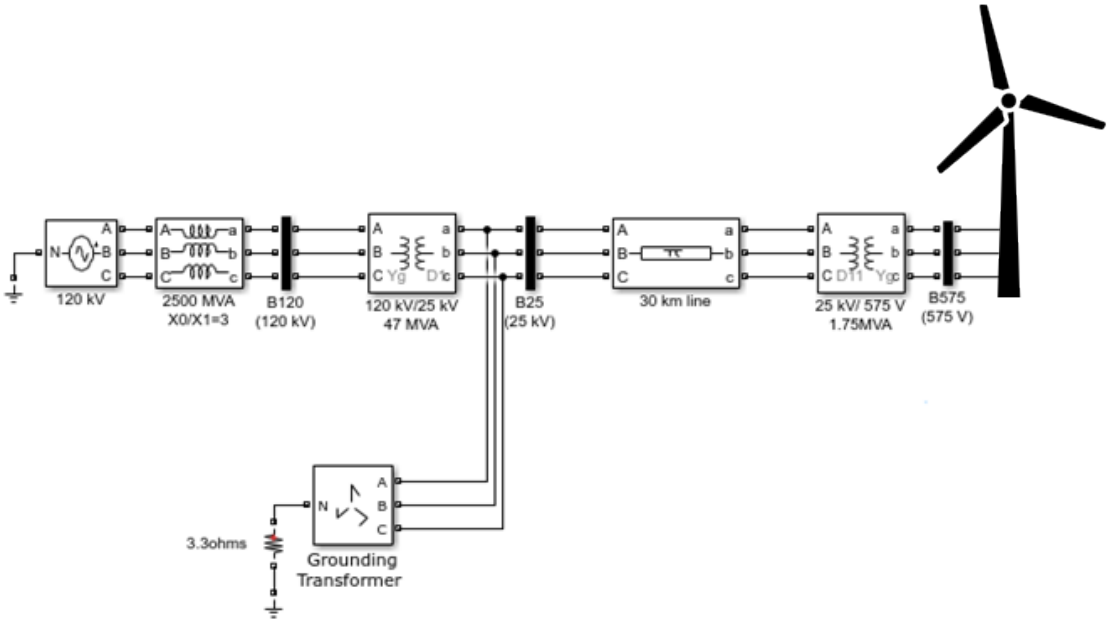


Figure 3.18 WT simulation of electrical network connection

A simple distributed model was used to simulate the electrical network, as shown in Figure 3.18. The simulated WT had a phase-ground voltage of 575 V at 60 Hz. The output of the WT was connected to a 1.75 MVA 575 V/25 kV transformer. A 30 km transmission line was connected between the WT power output and an electrical network to simulate transmission loss.

3.2.3.1 Healthy condition simulation

First, the DFIG WT was simulated. As an example, Figure 3.19 shows the simulation results for some of the outputs. From top to bottom, it includes the input wind speed, pitch angle of the blade, torque of the turbine, generator speed, voltage of the DC link, current of the DC link, active and reactive power of the turbine, RMS voltages at two different locations, and RMS currents at two different locations. A constant

wind speed of 15 m/s was used as the input. The first 15 s shows the WT start-up behaviours.

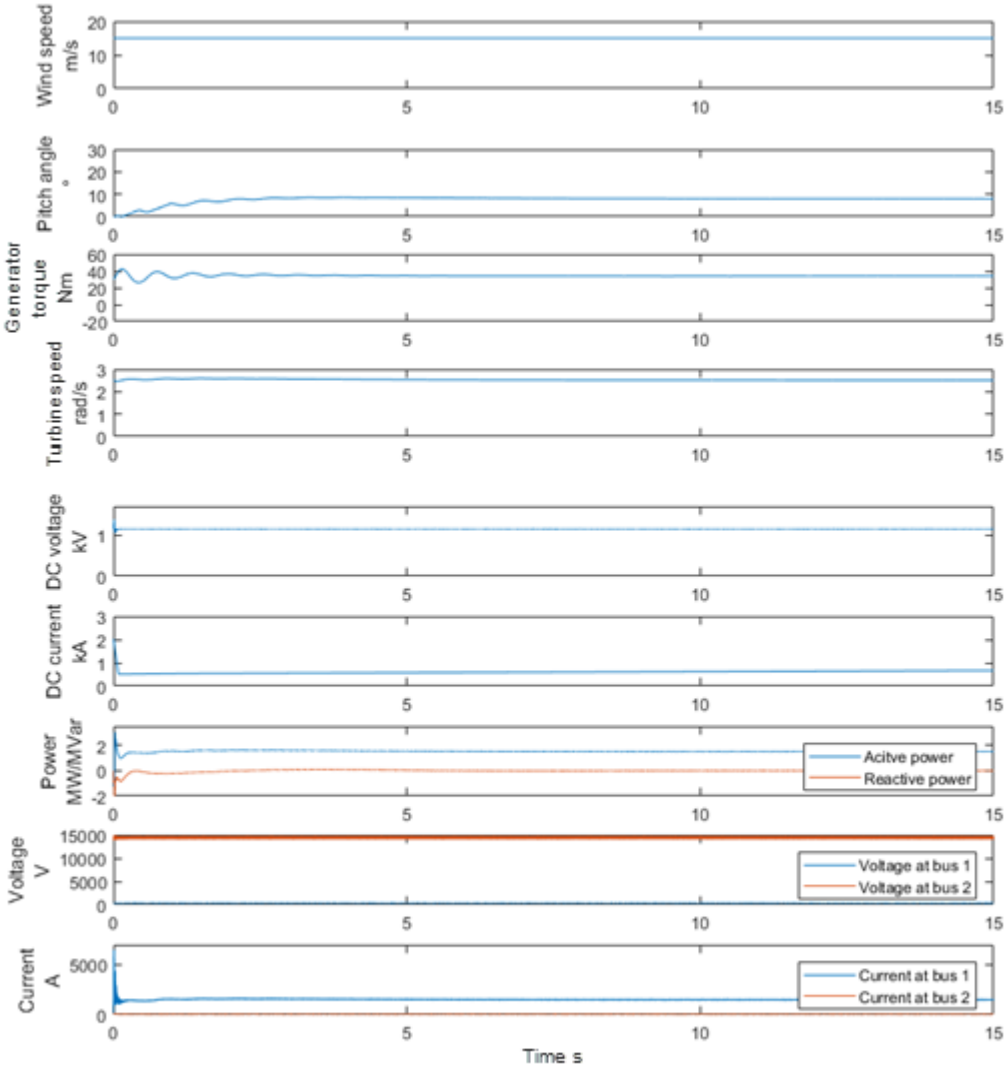


Figure 3.19 Simulation results of the fault-free DFIG WT under a constant wind speed

It can be observed from the figure that the system took approximately 3 s to reach the steady state. At the beginning of the simulation, the DC link had a pre-charged voltage of 1,175 V. Hence, pre-charging was not required. Then, it increased to 1,300 V when the simulation began, and it fell back to 1,175 V when the system reached a steady state. The system took approximately 4 s to reach a steady state and deliver constant power. In the steady state, the pitch angle was 9°, and the torque was 34 Nm. The output active and reactive powers in the steady state are approximately 1.5 MW and 0 MVar, respectively. The RMS voltages and currents

at the 575 V bus and 120 kV bus are approximately 340 V, 15 kV, 1.5 kA, and 34 A, respectively.

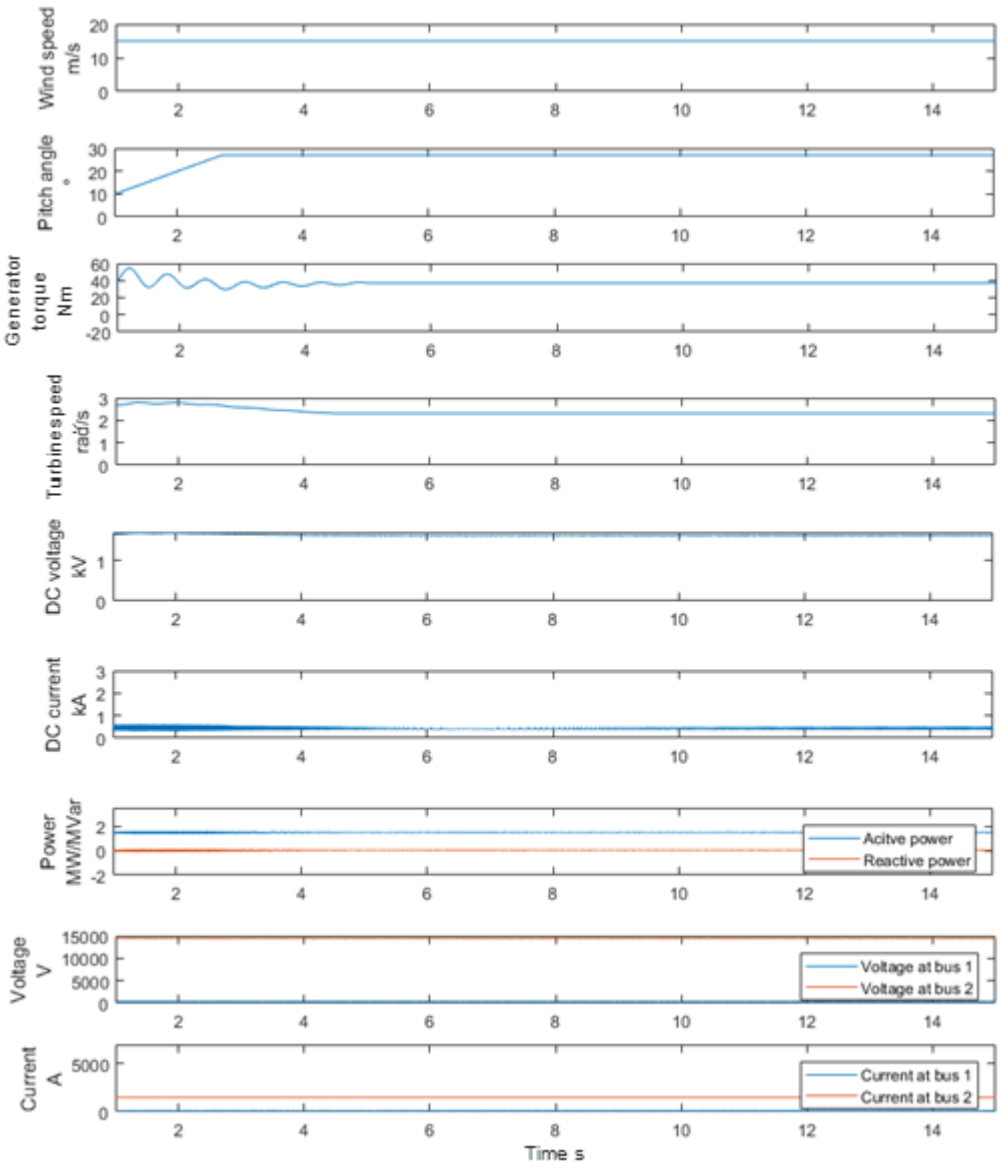


Figure 3.20 Simulation results of the fault-free PMSG WT under constant wind speed

Similar to the DFIG WT simulation, a complete PMSG WT model was built and tested for its performance. As an example, Figure 3.20 shows the simulation results of the PMSG WT under fault-free conditions with a constant wind speed of 15 m/s. Similar to the DFIG WT, the results include the input wind speed, pitch angle of the blade, torque of the turbine, generator speed, voltage of the DC link, current of the DC link, active and reactive power of the turbine, and RMS voltages and currents at two different locations. The first 15 s show the WT start-up behaviours.

Compared with the DFIG, it can be observed from the figure that the system took more than 4 s to stabilise. The DC-link voltage stabilised at 1,300 V at a steady state. In the steady state, the pitch angle was 27°, and the torque was approximately 37 Nm. In addition, the generator torque took longer to reach the steady state, as can be observed from the figure. The output active and reactive powers in the steady state are approximately 1.5 MW and 0 MVAr, respectively. The RMS voltages and currents at the 575 V and 120 kV buses are approximately 330 V, 15 kV, 1.5 kA, and 33 A, respectively.

3.2.3.2 Faulty condition simulation

Because the PMSG model was built based on the DFIG model, the performance of the monitoring variables was similar. Hence, only the DFIG faulty condition simulation is described in this subsection. Besides, the experimental power conversion test rig described in Chapter 7 was built based on the PMSG. The details of the PMSG based power conversion unit simulation are presented in Chapter 7. The previous section discussed the fault-free condition of the DFIG WT, and the faulty conditions were simulated. To improve the simulation performance, the input signal wind speed was simulated using historical SCADA data. In this section, the generation of a three-phase short-circuit fault at the grid-side converter between 10 s and 13 s is also described. Figure 3.21 presents the simulation results of the faulty DFIG WT under a constant wind speed.

It can be observed from the figure that a three-phase short-circuit fault is triggered, and the DC voltage has a sharp pulse in a very short time period. During the faulty period, the active power, reactive power, and voltage at the 575 V bus dropped to zero. The currents at both buses reached very high levels during the faulty period. Unlike electrical parameters that exhibit transient changes, mechanical components such as the pitch angle, torque, and generator speed require a longer time to return to the steady state.

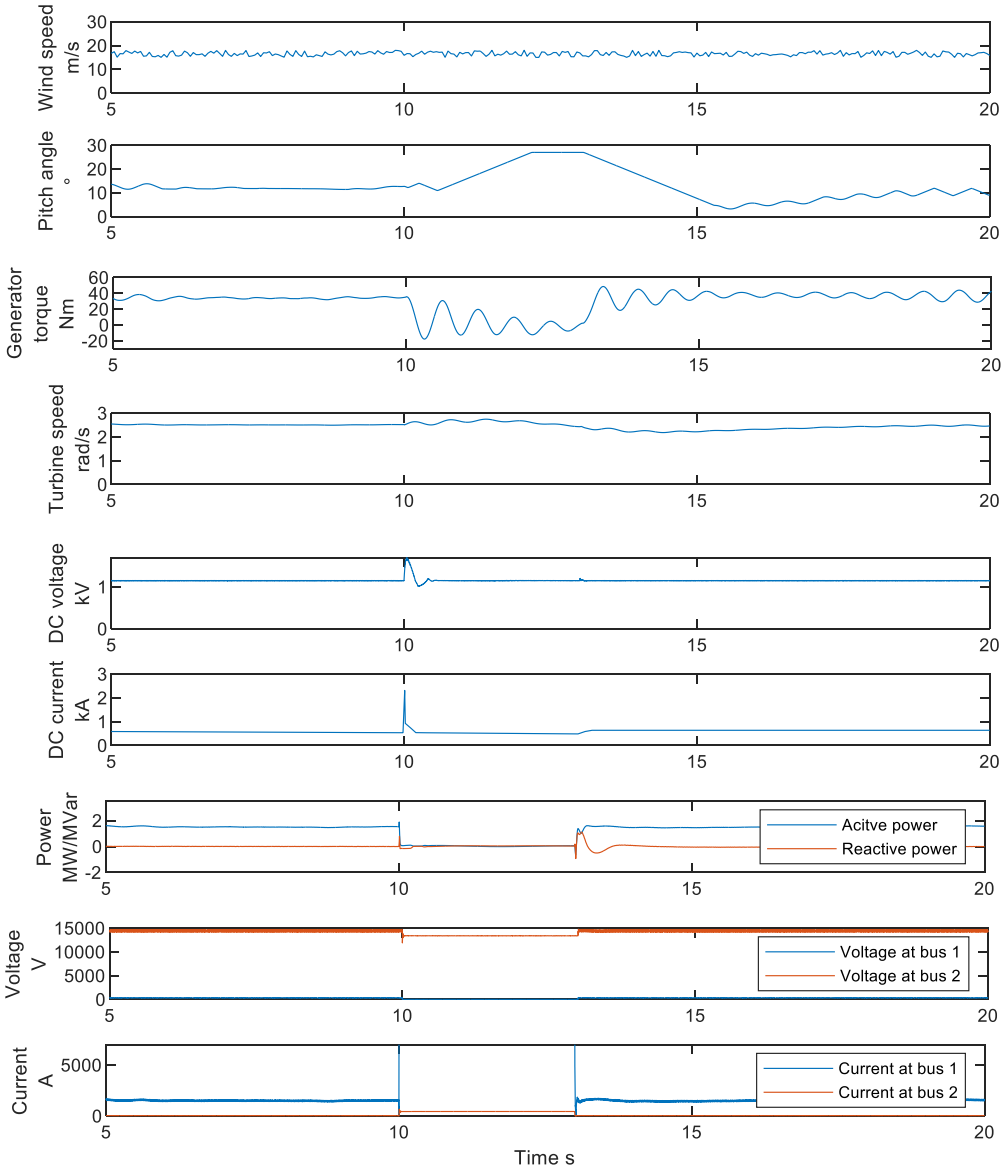


Figure 3.21 Simulation results of the DFIG WT three-phase short circuit fault

In addition to the three-phase short-circuit fault, capacitor breakdown was also simulated. For modern megawatt WTs, power converters play an important role. DC-link capacitors are considered to be the weakest components in power converters [171]. The constant high ripple current flowing in the DC-link capacitor can result in the degradation of the electrolyte material, especially at high operating temperatures. In addition, the useful life of the DC-link capacitor depends on its operating current and voltage. Once the DC-link capacitor is broken down, the DC voltage is derated. Hence, DC-link capacitor failure can result in unreliable wind power generation.

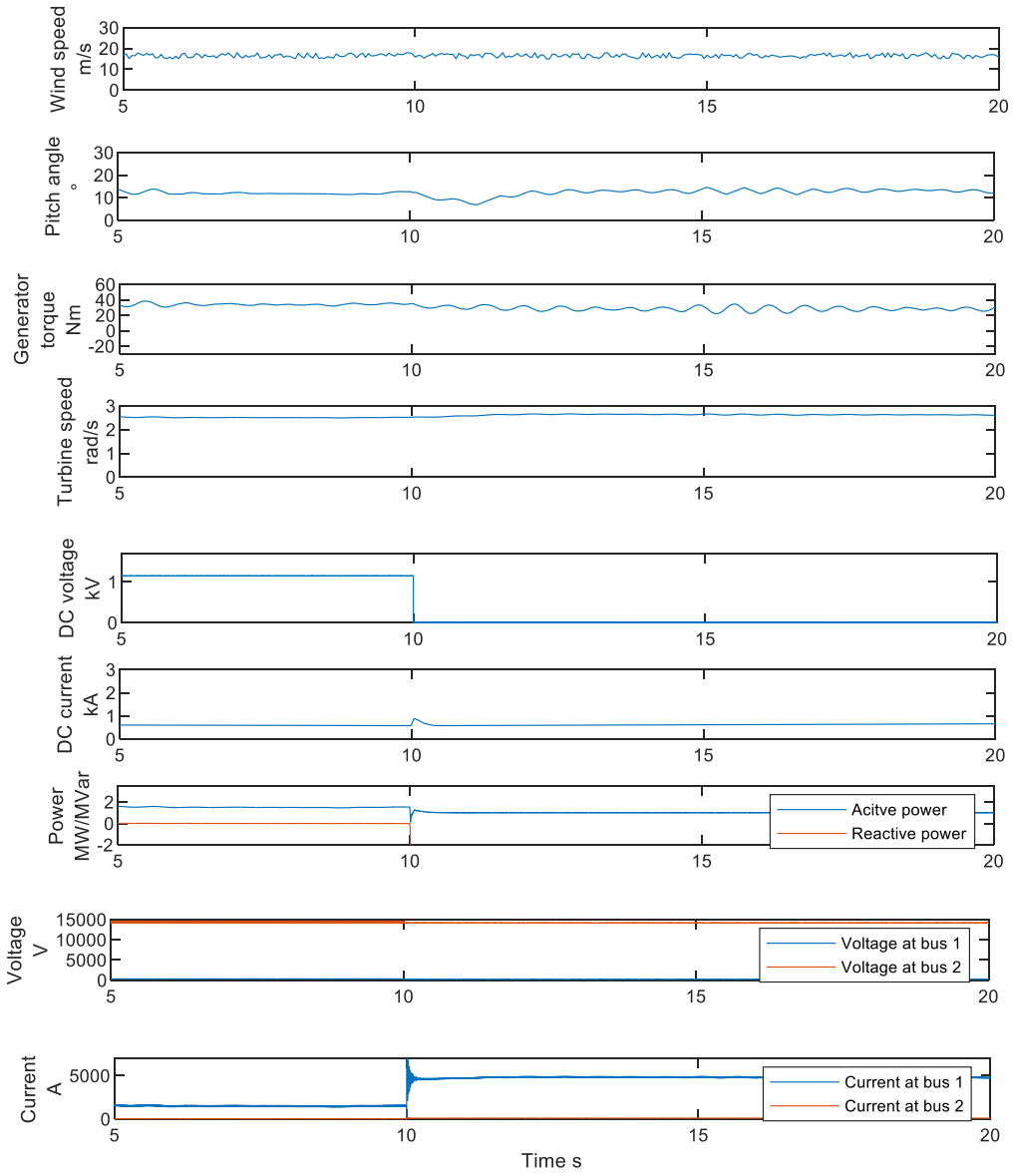


Figure 3.22 Simulation results of the DFIG WT capacitor breakdown fault

The simulation was run for 20 s, and the fault was injected for 10 s. The results are shown in Figure 3.22. The data before starting up (t < 5 s) are neglected, and the figure only shows the time period between 5 s and 20 s. As can be observed from the figure, electrical parameters such as the DC link voltage, active power, and 575 V bus voltage dropped to zero immediately when the fault was triggered. The DC current and 575 V bus current increased after the fault occurred.

3.3 Summary

In this chapter, historical data collected from an operational wind farm and modelling of the DFIG and PMSG variable-speed WTs were discussed. The

historical SCADA data could only provide existing faults which are mainly mechanical faults and cannot show the transient changes by electrical faults. The simulation models can be used to simulate electrical faults which complement the information that cannot be provided by historical SCADA data. The models were built and simulated based on MATLAB/Simulink to determine the dynamic changes under both healthy and faulty conditions. Compared with SCADA data, high-frequency sampling rate data provided more information on transient changes in the system. Two types of faults were simulated, and sufficient information was provided for further analysis. More results will be presented in the following chapters, combined with real historical data collected from operational wind farms, for the validation of the proposed algorithms

Chapter 4. Kullback-Leibler divergence based wind turbine fault feature extraction

In this chapter, a multivariate statistical technique combined with a ML algorithm is proposed to provide a fault classification and feature extraction approach for the WTs. As the probability density distributions (PDDs) of the monitoring variables can illustrate the inner correlations among variables, the dominant factors causing the failure are figured out, with the comparison of PDD of the variables under the healthy and unhealthy scenarios. Then the selected variables are used for fault feature extraction by using kernel support vector machine. The presented algorithms are implemented and assessed based on the SCADA data acquired from an operational wind farm. The results show the features relating specifically to the faults are extracted to be able to identify and analyse different faults for the WTs.

4.1 Introduction

The chapter addresses the problem of fault feature extraction and selection of monitoring variables by employing KLD and KSVM. In this paper, different faults are discussed and classified. The highly correlated variables with regards to the WT failures are extracted. In section 4.2, the basic knowledge of KLD and SVM with radical basis function kernel are introduced first. The SCADA data and the modelling procedures are introduced in Section 4.3. It also includes the analysis of the PDD of each variable under different operation scenarios. Then the PDD differences between the healthy turbine and faulty turbine for the same variables are analysed by KLD. At last, the proposed model is also validated by the simulated data. Larger KLD values indicate the variable has a larger difference between the healthy and faulty conditions. By setting an appropriate threshold for the KLD value, the variables, which have larger chances to trigger a specific fault are selected. The fault features are also extracted based on the KLD selected variables.

4.2 Methodologies

4.2.1 Kullback-Leibler divergence

The KLD, also called Shannon entropy, is a probability measurement from information theory perspective. It is used to compare the one probability distribution against with another one [173]. To evaluate the discrimination between two PDDs, *p(x)* and *g(x)*, of a random variable x, the Kullback-Leibler information $I(p \parallel g)$ is calculated:

$$I(p \parallel g) = \int p(x) log \frac{p(x)}{g(x)} dx \tag{4-1}$$

The divergence of the two distributions is a symmetric version of the information:

$$KLD(p, g) = I(p \parallel g) + I(p \parallel f) \tag{4-2}$$

It is only produced when $\sum_x p = \sum_x g = 1$ and the two distributions share the same support set. When the two distributions are same, the KLD of the two distributions is 0. When the difference between the two distributions increases, the calculated value of KLD also increases. In other words, the bigger the KLD value, the larger the difference between two distributions that can be found.

4.2.2 Kernel support vector machine

The SVM is a ML algorithm that is developed from the statistical learning with a better performance than many other methods, such as decision trees, discriminant analysis and nearest neighbour. It is a supervised learning method that has been widely applied in regression and classification [174, 175].

It is assumed that there are two classes for separation, which are denoted as {-1, 1}. For a given training data set $T(x, y)$

$$T(x, y), (x_1, y_1), (x_2, y_2), \dots, (x_n, y_n),$$

$$x_n \in R^n, \;\; y_n \in \{-1,1\}. \tag{4-3}$$

where x are the training vectors and y are the indicator vectors indicating the class of x, respectively. In this case, {-1, 1} indicates the turbine under faulty and healthy condition, respectively. The hyperplane created by SVM is used to separate the two different classes that is specified by its weight w and the bias b [175]. The hyperplane can be described as

$$(w, x) + b = 0 \qquad w \in R^n, b \in R \tag{4-4}$$

It yields a corresponding decision function:

$$f(x) = sgn((w, x) + b) \tag{4-5}$$

The sign of $f(x)$ depends on the side of the hyperplane where the sample lies. An optimal hyperplane is the one that maximises the distance between the hyperplane and the nearest points to the hyperplane of both classes. The points of having the minimal distances to the hyperplane are called support vectors. The distance between the support vectors is called margin. In order to obtain the hyperplane, the weight and bias of support vectors should satisfy the equation$|(w, x) + b| = 1$. The optimal hyperplane with maximal margin $\tau(w)$ can thus be formulated as follows

$$\min_{w \in R^n, b \in R} \tau(w) = \frac{1}{2}\|w\|^2 \tag{4-6}$$

subject to

$$y_i\big((w,x)+b\big) \geq 1 \quad \forall i = 1,\dots,n \tag{4-7}$$

The decision function $f(x_i)$ yields +1 for $y_i \in \{1\}$ while -1 for $y_i \in \{-1\}$ when the constraint (4-7) is satisfied. Furthermore, the Lagrange multipliers α_i are introduced to adapt the Karush Kuhn Tucker conditions in order to optimise the *w* and *b* [176].

$$max_{\alpha \in R^n} \; L(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j (x_i, x_j) \tag{4-8}$$

subject to

$$\alpha_i \geq 0 \quad \forall i = 1,\dots,n \tag{4-9}$$

and

$$\sum_{i=1}^{n} \alpha_i \, y_i = 0 \tag{4-10}$$

The optimal parameters w^* and b^* can be written as

$$w^* = \sum_{i=1}^{n} \alpha_i y_i x_i \tag{4-11}$$

$$b^* = -\frac{1}{2}(w^*, x_a + x_b) \tag{4-12}$$

where $x_a \in \{1\}$ and $x_b \in \{-1\}$ indicate the support vectors from two classes. Then the decision function from (4-5) can be written as

$$f(x) = sgn((w^*, x) + b^*) \tag{4-13}$$

By adapting the optimal parameters w^* and b^* from (4-11, 4-12), the (4-13) in the feature space can be updated as

$$f(x) = sgn\left(\big(w \cdot \phi(x)\big) + b\right)$$

$$= \sum_{i=1}^{n} y_i \alpha_i \phi(x)\phi(x_i) + b^* \tag{4-14}$$

The above equations are only suitable for linear separation problem. In general, most of the cases are linear non-separable. In order to solve the linear non-separable problem, the kernel function is introduced for improving the computation efficiency. In this specific condition, the Radical Basis Function (RBF) or Gaussian kernel is used.

The dot product (x, x_i) can be substituted by the kernel

$$k(x, x_i) = (\phi(x) \cdot \phi(x_i)) \tag{4-15}$$

where $\phi(x)$ and $\phi(x_i)$ are the mapping of x, x_i in hyper dimension. By adapting a RBF kernel, the kernel function can be written as

$$K(x, x_i) = e^{\left(\frac{-\|x-x_i\|^2}{2\sigma^2}\right)} = e^{(-\gamma\|x-x_i\|^2)} \tag{4-16}$$

where Λ is the width of the kernel, which controls the smoothness of the decision boundary in the feature space and $\gamma = -\frac{1}{2\sigma^2}$.

Thus, the decision function $f(x)$ for an unknown input x in the feature space can be extended to

$$f(x) = sgn\Big(\big(w \cdot \phi(x)\big) + b\Big)$$

$$= \sum_{i=1}^{n} y_i \alpha_i e^{(-\gamma\|x-x_i\|^2)} + b^* \tag{4-17}$$

In the regression cases, the SVM classification function from (4-13) can be expressed as follows to predict the value *h(x)*

$$h(x) = (w^*, x) + b^* \tag{4-18}$$

By adapting RBF kernel to the regression function (4-18) in feature space, the prediction value can be extended to

$$f(x) = (w^*, x) + b^*$$

$$= \sum_{i=1}^{n} \alpha_i e^{(-\gamma\|x-x_i\|^2)} + b^* \tag{4-19}$$

4.2.3 Feature extraction model

The data used in this paper are from an operational wind farm in one-year duration consisting of 26 turbines with each turbine having 128 monitoring variables, including temperatures, active power and various speeds. The SCADA system applied in this particular wind farm sampled signals at 10-mins intervals to cover a

range of physical and electrical variables and digital control signals. However, the digital signals cannot reveal the working condition of the WT; it is thus necessary to eliminate those variables before further processing [22].

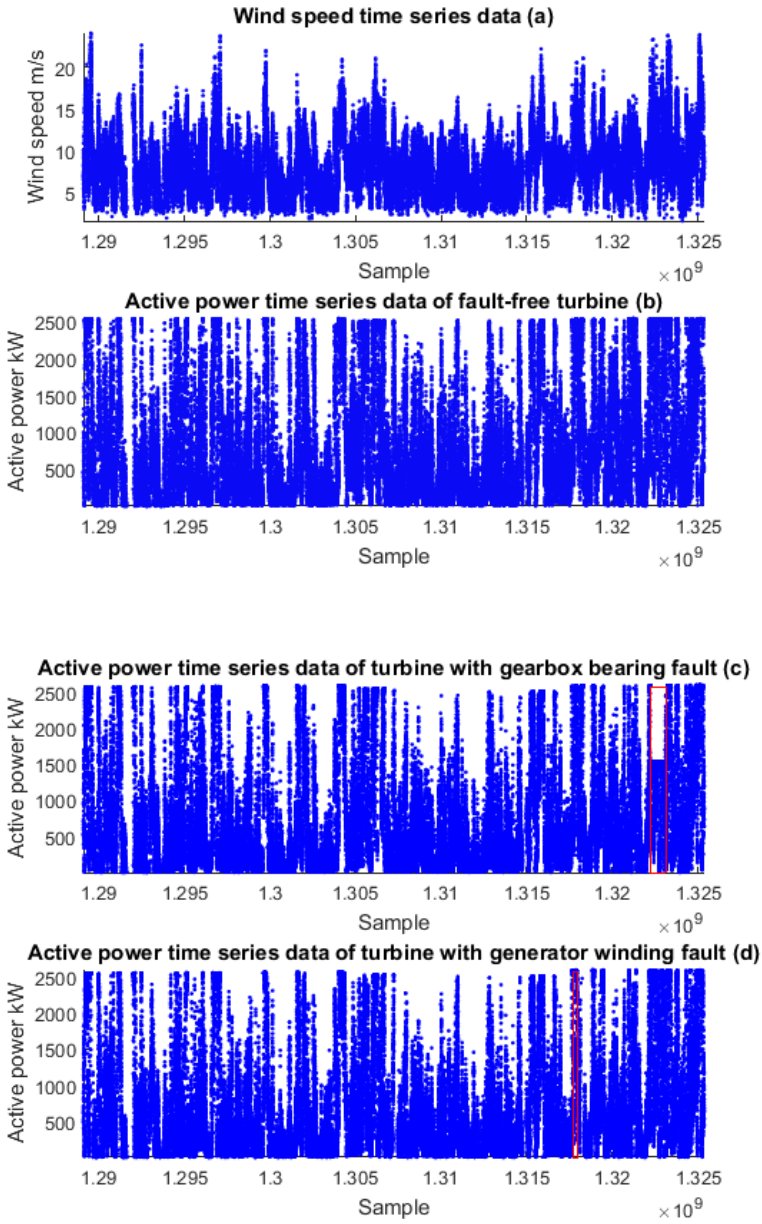


Figure 4.1 Time series data of (a) wind speed; (b) active power of a fault-free turbine; (c) active power of a turbine with gearbox bearing fault; (d) active power of a turbine with generator winding fault

Figure 4.1 shows wind speed and the output active power of the turbines working in different conditions over one year. Figure 4.1(a) shows the wind speed data measured on both healthy and faulty turbine. Figure 4.1 (b-d) show the active power outputs over one year for the fault-free turbine and turbines with gearbox bearing fault and generator winding fault, respectively. The occurrence of the faults are

labelled in red boxes. Compared with the turbine having a gearbox fault, the turbine with a generator fault clearly has a shorter faulty time duration. In order to prevent overheating to the critical components and further damage the whole turbine, the turbines were operated with power output being reduced to half.

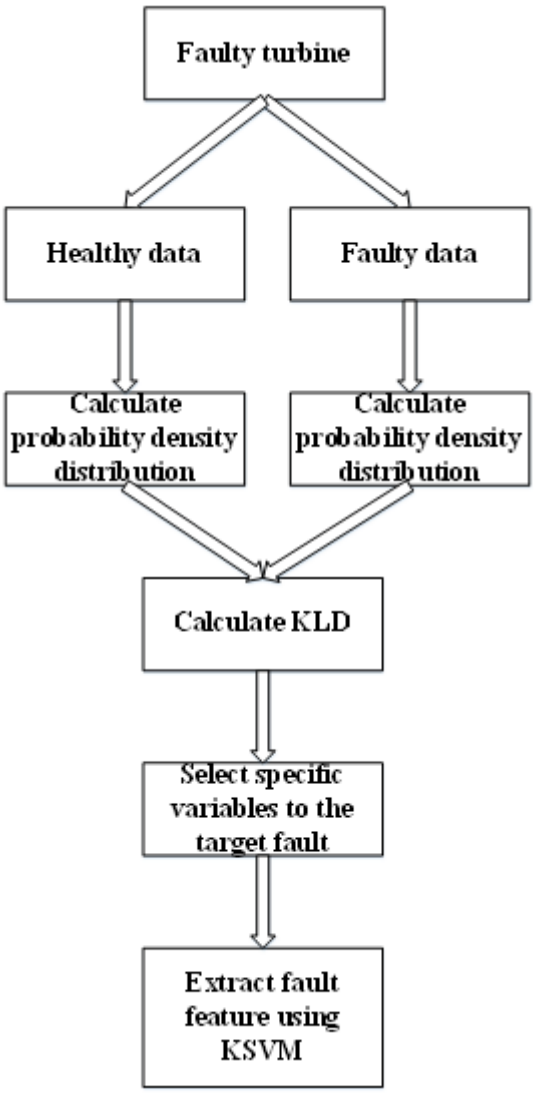


Figure 4.2 Fault classification and feature extraction

Figure 4.2 illustrates the procedures of variable selection and fault feature extraction. In order to obtain the PDD ideally, the data samples should be selected as much as possible. In the fault classification and feature extraction procedures, 50000 samples measured over one year are selected. Two turbines are selected for modelling: one has a gearbox fault while another has a generator winding fault. The faulty data can be selected based on time sequential information. First, the PDDs of all the monitoring variables are calculated with the data under healthy and faulty

conditions. Then the two types of PDD are compared by calculating the KLD values between them. By sorting the KLD values of the monitoring variables in descending order, the monitoring variables related specifically to the fault can be discovered. Finally, the features associated with the gearbox bearing fault and generator winding fault are extracted by the KSVM.

4.3 Model validation

4.3.1 Variable selection using KLD of SCADA data

As mentioned above, the digital constants and control signals are removed in the data pre-processing, totally 78 variables are then retained for further analysis. The KLD values are calculated based on same variable under healthy and faulty conditions. By sorting the KLD values in descending order and eliminate the variables that are infinite or cannot be computed, the KLD values of 22 variables related to the gearbox fault are given in Table 4.1.

Table 4.1 KLD values of gearbox bearing fault variables

Variable	Nacelle temperature	Wind direction	Nacelle position	Pitch converter temperature 2
KLD	15.89	14.95	14.51	7.47
Variable	Hydraulic fluid temperature	Pitch angle 2	Pitch Converter Temperature 1	Pitch Converter Temperature 3
KLD	6.49	5.52	4.87	4.68
Variable	Power Switchboard Temperature	Gearbox Oil Pressure Behind Pump	Temperature of generator cooling water return	Temperature of gearbox oil heat exchanger output
KLD	4.68	2.44	1.69	1.14
Variable	Hygrometer 2 temperature	Generator speed	Generator speed (to-default)	Wind speed right-hand
KLD	1.11	0.86	0.7	0.51
Variable	Wind speed	Wind speed left-hand	Minimum pitch angle	Oscillation signal Z
KLD	0.51	0.5	0.39	0.19
Variable	Maximum pitch speed	Temperature of Cooling water return		
KLD	0.16	0.16		

All 22 KLD values show the differences between healthy condition and faulty condition. The larger value indicates larger difference. It can be observed from the Table I that some variables are related to the environmental condition, such as wind speed and direction. The environmental related variables can be ignored in the variable selection because they do not have a direct impact on the fault. By

eliminating these variables, pitch system, gearbox and generator related variables are retained.

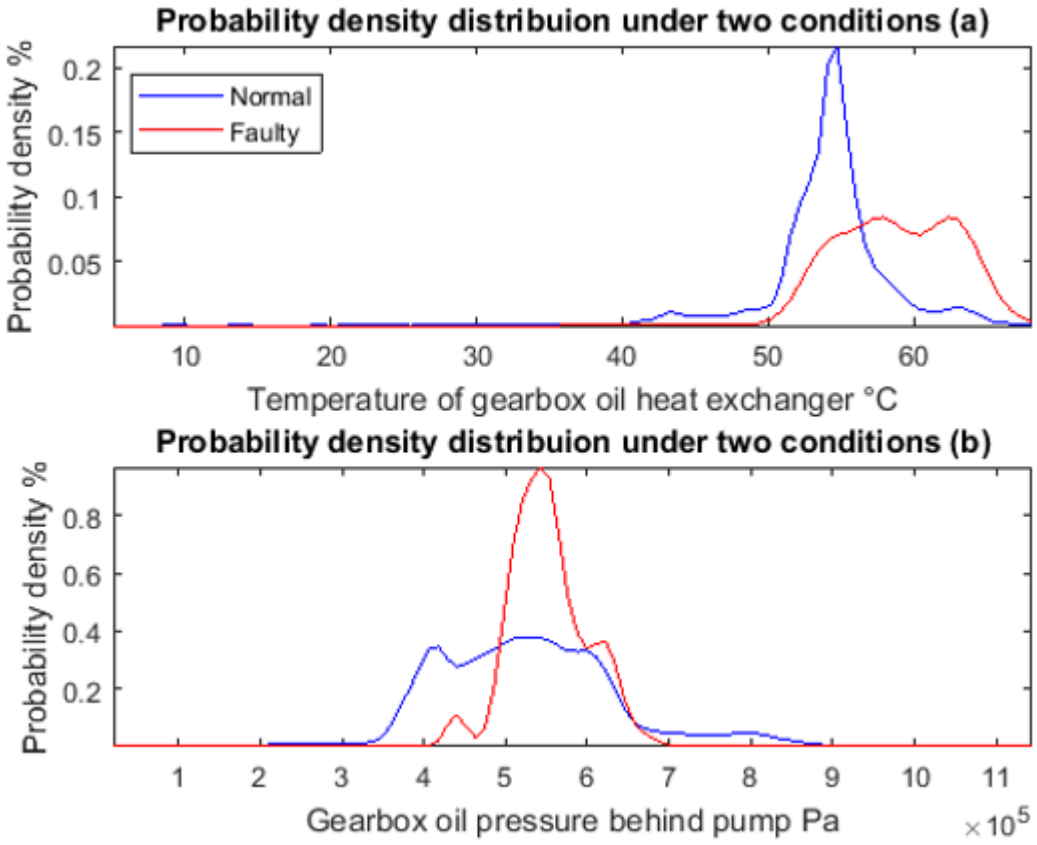


Figure 4.3 PDDs between healthy and faulty conditions of (a) gearbox oil heat exchanger temperature and (b) gearbox oil pressure behind pump

Figure 4.3 shows examples of the PDDs of the gearbox oil exchanger temperature and gearbox oil pressure behind pump under healthy and faulty conditions. It can be observed from the top plot of Figure 4.3 that when the turbine is under the healthy condition, the PDD of gearbox oil exchanger is similar to Gaussian distribution with an average temperature around 55 °C. In addition, the gearbox oil pressure behind pump also indicates the abnormal behaviour of this turbine. Consequently, the pressure tended to be higher when fault occurred, compared with the healthy condition.

Table 4.2 show KLD values of the selected variables related to generator winding fault. There are 21 variables are selected for this specific fault. It can be observed from the table that the pitch system, gearbox and generator related variables are selected.

Table 4.2 KLD values of generator winding fault variables

Variable	Temperature of cooling water return	Wind speed left-hand	Wind speed	Pitch converter temperature 2
KLD	37.89	11.73	11.7	11.62
Variable	Wind speed right-hand	Hygrometer 2 temperature	Gearbox oil pressure Behind Pump	Pitch converter temperature 1
KLD	11.25	10.26	8.26	5.5
Variable	Temperature of generator cooling water return	Oscillation signal Z	Hydraulic fluid temperature	Nacelle temperature
KLD	3.7	2.92	1.81	1.43
Variable	Temperature of gearbox oil heat exchanger output	Pitch angle 1	Minimum pitch angle	Pitch angle 3
KLD	1.06	0.45	0.45	0.37
Variable	Pitch angle average 1-3	Generator speed	Generator speed (to-default)	Pitch angle 2
KLD	0.33	0.28	0.28	0.26
Variable	Generator temperature			
KLD	0.2			

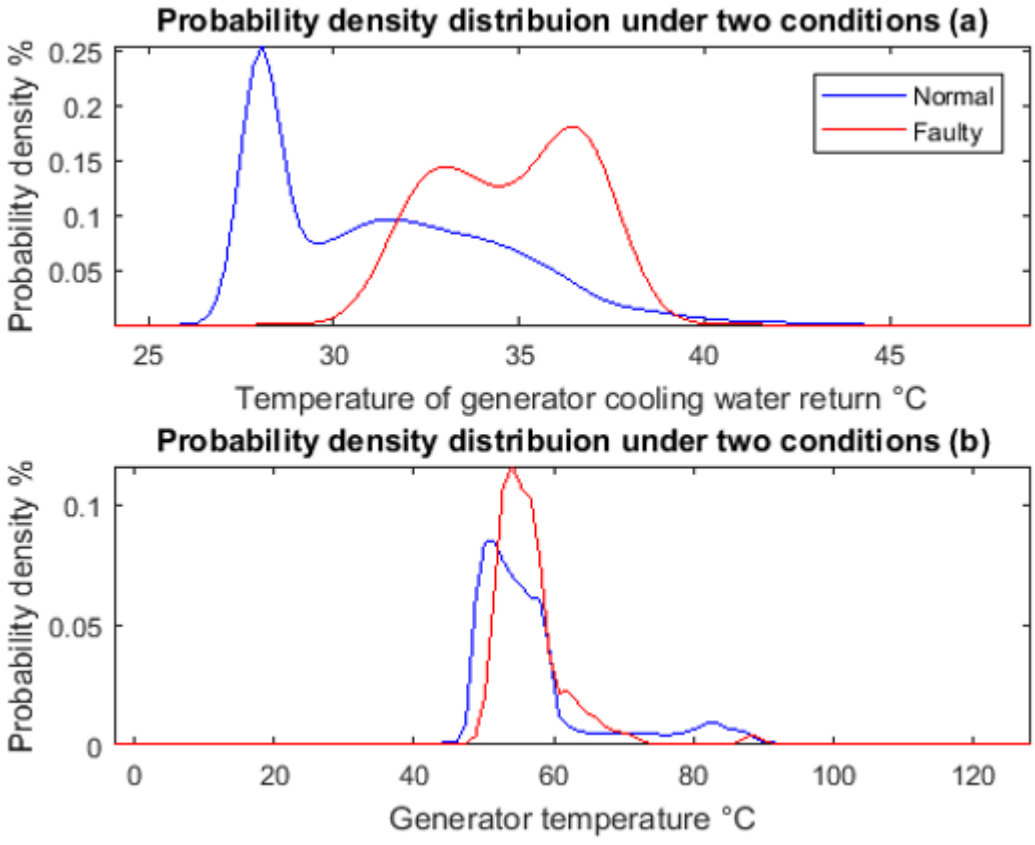


Figure 4.4 PDDs between healthy and faulty conditions of (a) generator cooling water return temperature and (b) generator winding temperature

Figure 4.4 shows two variables' PDDs of the turbine with generator winding fault. For this specific fault, both the temperatures of generator cooling water and

generator winding tended to shift towards higher temperature zone compared with the normal condition. This clearly indicates the abnormal behaviour of generator when fault appears.

4.3.2 Feature extraction using KSVM of SCADA data

It can be concluded from the both Table 4.1 and 4.2 that the pitch system, generator system and gearbox system tended to be the three most abnormal systems of the turbines when faults occur. The variables are chosen from gearbox, generator and pitch systems to separate the fault and normal conditions.

The relationship between generator speed and wind speed can be hard to discover, which is shown at top plot of Figure 4.5. In order to extract the real behaviour of the generator speed against the wind speed, the KSVM is applied as a regression model to find the relationship between generator speed and wind speed. A threshold is also added to help the data selection for this regression model. The bottom plot of Figure 4.5 shows the comparison between original generator speed curve and filtered generator speed curve, where the yellow curve indicates the filtered generator speed curve. The relationship between generator speed and wind speed can be revealed clearly now by adapting KSVM regression.

Then the KSVM is applied to divide the normal and fault data into two groups. Figure 4.6 shows the extracted feature of the gearbox bearing fault. The blue dots indicate the normal feature and the red crosses indicate the gearbox bearing fault, which is separated by the yellow plane. It can be observed from the figure in order to prevent dramatic damages to the turbine, the power output is reduced to half. This is due to the fact that the changing of pitch angle controls the output power. Any points within the yellow plane are considered as the abnormal date related to the fault. The plane created by KSVM is used to extract the fault feature can be described as $(x - 1136.12)^2 + (y - 2319.04)^2 + (z + 1916.93)^2 - 9300391.69 < 0$. When the equation is above zero, it can be considered as healthy condition. When the gearbox bearing fault occurs, the inequality is satisfied.

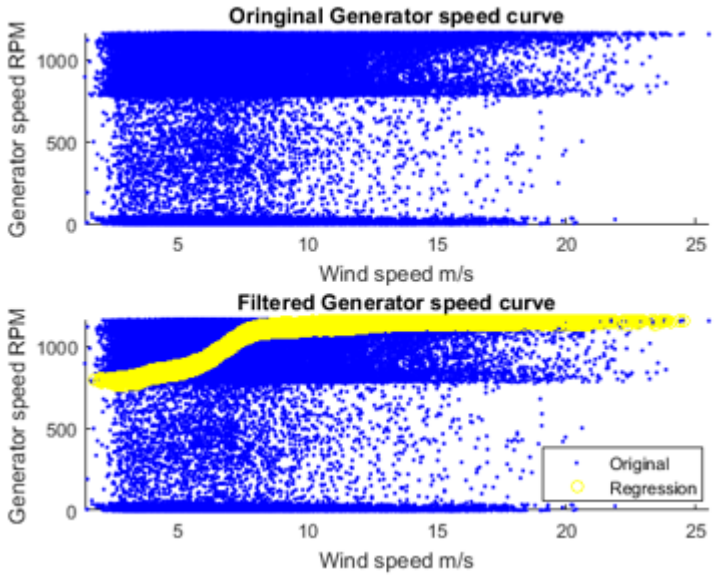


Figure 4.5 Original generator speed curve vs. filtered generator speed curve of the gearbox bearing fault turbine

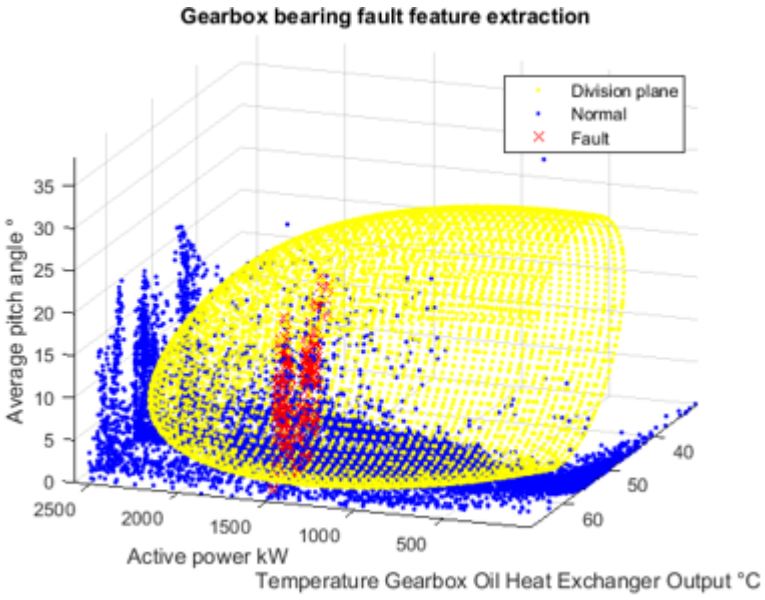


Figure 4.6 Gearbox bearing fault feature extraction

Figure 4.7 shows the generator speed curve against wind speed for the turbine with generator winding fault. In order to extract the key information, same method has been applied. As the bottom plot of Figure 4.7 shows, the feature is revealed by KSVM.

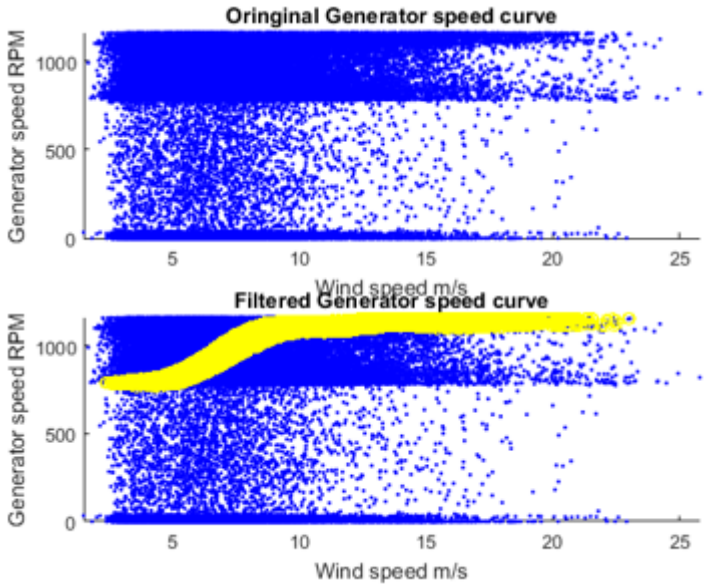


Figure 4.7 Original generator speed curve vs. filtered generator speed curve of the generator winding fault turbine

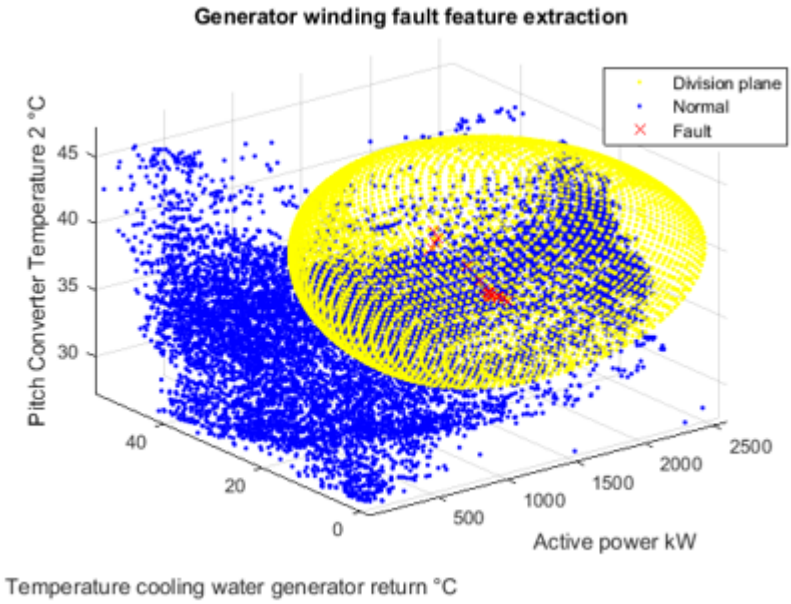


Figure 4.8 Generator winding fault feature extraction

Based on the active power, generator cooling water temperature and pitch converter temperature, the generator winding fault feature can be extracted as shown in Figure 4.8. The plane which is used to extract the fault feature is also created by KSVM. It is described as the inequality $(x - 1251.33)^2 + (y + 5895.003)^2 + (z - 6673.873)^2 - 79617582.07 < 0$. When the inequality is established, the generator winding fault can be found. It can be observed from the figure that when the fault

appears, the pitch converter temperature tends to be higher while the generator temperature is still in a low temperature zone.

4.3.3 Variable selection using KLD of simulation data

To verify the capability of the proposed model, the DFIG simulation model was tested with two different faults at variable wind speeds. First, the three-phase short-circuit fault was simulated, as shown in Figure 4.9. From top to bottom, it includes the input wind speed, pitch angle of the blade, torque of the turbine, generator speed, voltage of the DC link, current of the DC link, active and reactive power of the turbine, RMS voltages of two different locations, and RMS currents of the 575 V and 25 kV buses. The simulation was run for 20 s and the fault was injected at 13 s and ended at 16 s. The data before starting up ($t < 5$ s) are neglected, and the figure only shows the time period between 5 s and 20 s. It can be observed from the figure that the monitoring variables varied with different wind speeds. Owing to the control strategy, the torque and pitch angle were changed constantly to maintain the maximum power output. When a fault is injected, the active power output decreases to zero. The current also drops to zero because the three-phase current output is short-circuited with the ground. The DC voltage and current exhibited a sharp impulse when the fault was first injected.

Because the wind speed is the input of the system and active/reactive power are the outputs of the system, these three variables are neglected in the variable selection. Similar to the SCADA data, the KLD values were calculated based on the same variables under healthy and faulty conditions. The KLD values of the nine variables related to the three-phase short-circuit fault are listed in Table 4.3.

Table 4.3 KLD values of three-phase short circuit fault variables

Variable	Pitch degree	Torque	Generator speed	V_{dc}
KLD	25.39	20.16	2.15	43.54
Variable	I_{dc}	575V bus voltage	25kV bus voltage	575V bus current
KLD	39.42	5.68	1.32	2.51
Variable	25kV bus current			
KLD	78.32			

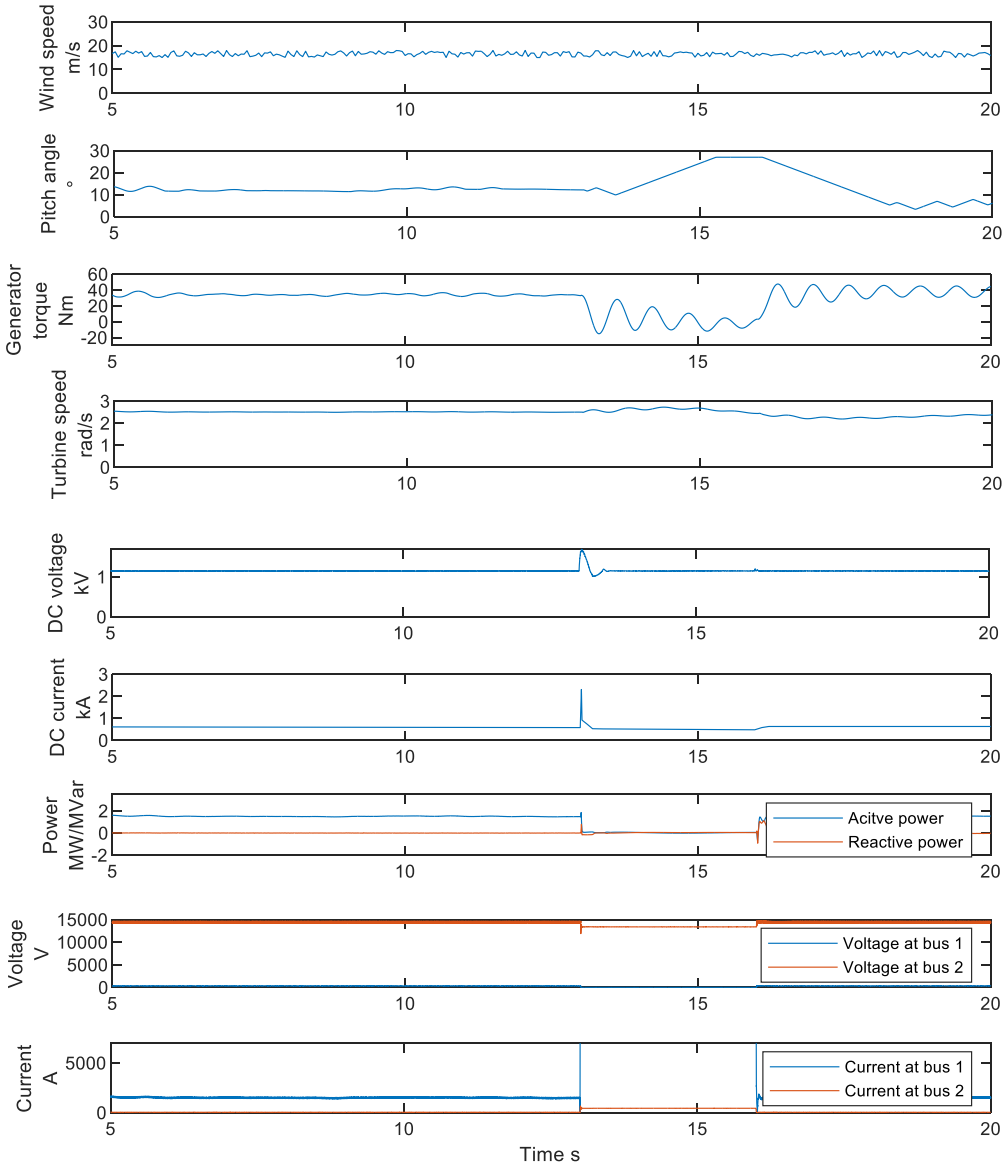


Figure 4.9 Simulated DFIG WT three-phase short-circuit fault under variable wind speed

All nine KLD values showed differences between healthy and faulty conditions. A larger value indicates a larger difference. As shown in Table 4.3, the KLD values for the pitch degree, torque, generator speed, V_{dc}, I_{dc}, 575 V and 25 kV bus voltages, 575 V and 25 kV bus currents are retained for comparison. The 25 kV bus current and V_{dc} have the largest KLD values, indicating that these variables have the highest correlation with the fault. Consequently, these two variables were selected for feature extraction.

In addition to the three-phase short-circuit fault, a capacitor breakdown fault was also simulated for model validation. The simulation results are shown in Figure 4.10.

The fault was injected within 13 s. The DC-link voltage drops to zero because the capacitor has already experienced breakdown. The other electrical variables also change abruptly when a fault occurs. Physical variables such as the pitch angle, torque, and generator have more turbulence after the fault is triggered.

The KLD values for the nine variables are listed in Table 4.4. Among all the variables, the 25 kV bus voltage and turbine torque have the largest KLD values, which indicate the highest differences between the normal and faulty conditions. Thus, these two variables were selected for further feature extraction.

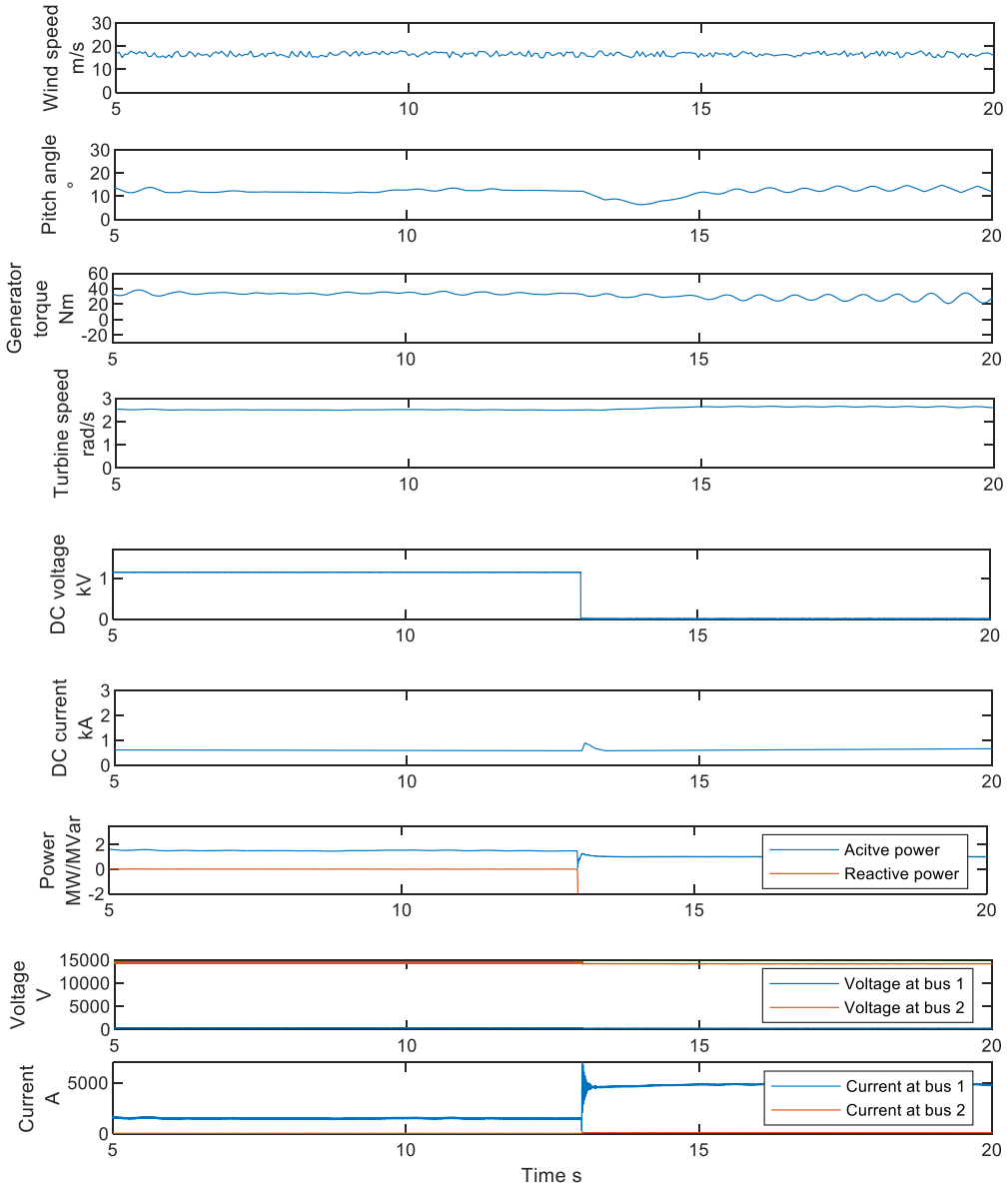


Figure 4.10 Simulated DFIG WT capacitor breakdown fault under variable wind speeds

Table 4.4 KLD values of capacitor breakdown fault variables

Variable	Pitch degree	Torque	Generator speed	V_{dc}
KLD	12.71	25.36	0.22	20.19
Variable	I_{dc}	575V bus voltage	25kV bus voltage	575V bus current
KLD	5.88	3.41	55.24	17.38
Variable	25kV bus current			
KLD	2.12			

4.3.4 Feature extraction using KSVM of simulation data

To understand the relationship between the DC-link voltage and 25 kV bus current and to extract the fault feature, KSVM was applied. It was then applied to classify the normal and fault data into two groups. Figure 4.1 shows the extracted features of the three-phase short-circuit fault. The blue dots indicate the normal feature and the red crosses indicate the three-phase short-circuit fault, which is separated by the yellow division plane.

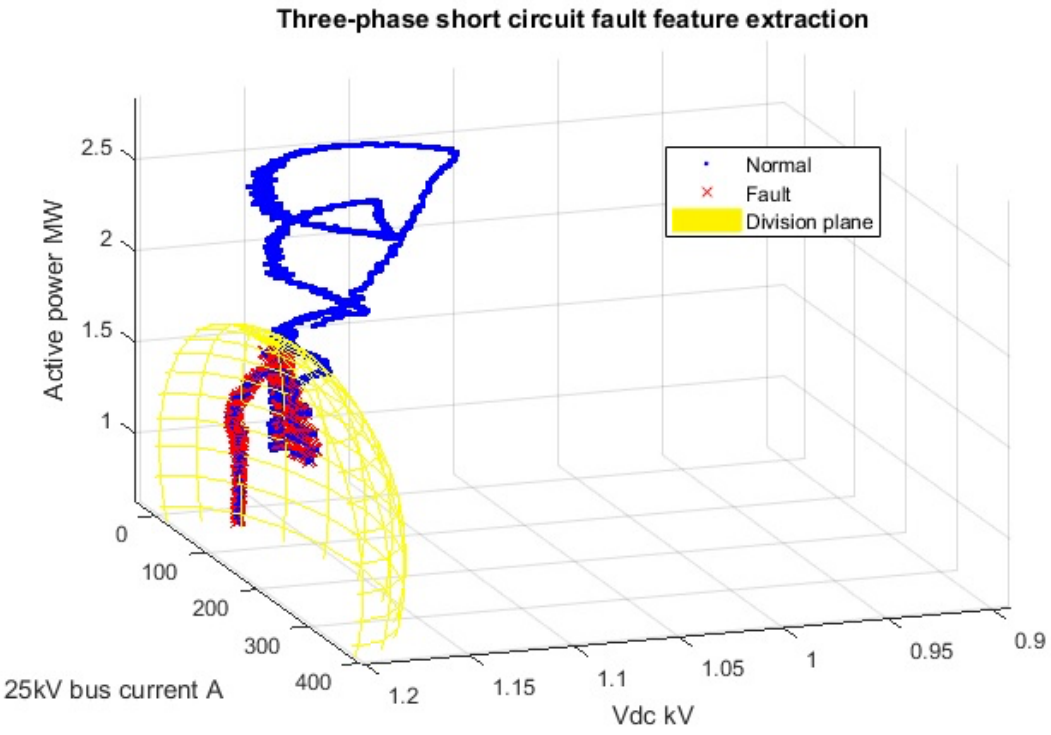


Figure 4.11 Three-phase short circuit fault feature extraction

It can be observed from the figure that when the turbine operates under normal conditions, V_{dc} and the active power remain constant. When a fault occurs, the relationship between the active power, V_{dc}, and current changes. Any points within

the yellow plane were considered abnormal data related to the fault. The plane created by KSVM is used to extract the fault feature and can be described as$(x-441.22)^2+(y-1.15)^2+(z+0.03)^2-10.02<0$. When the value is above zero, it can be considered as a healthy condition. When a three-phase short-circuit fault occurs, the inequality is satisfied.

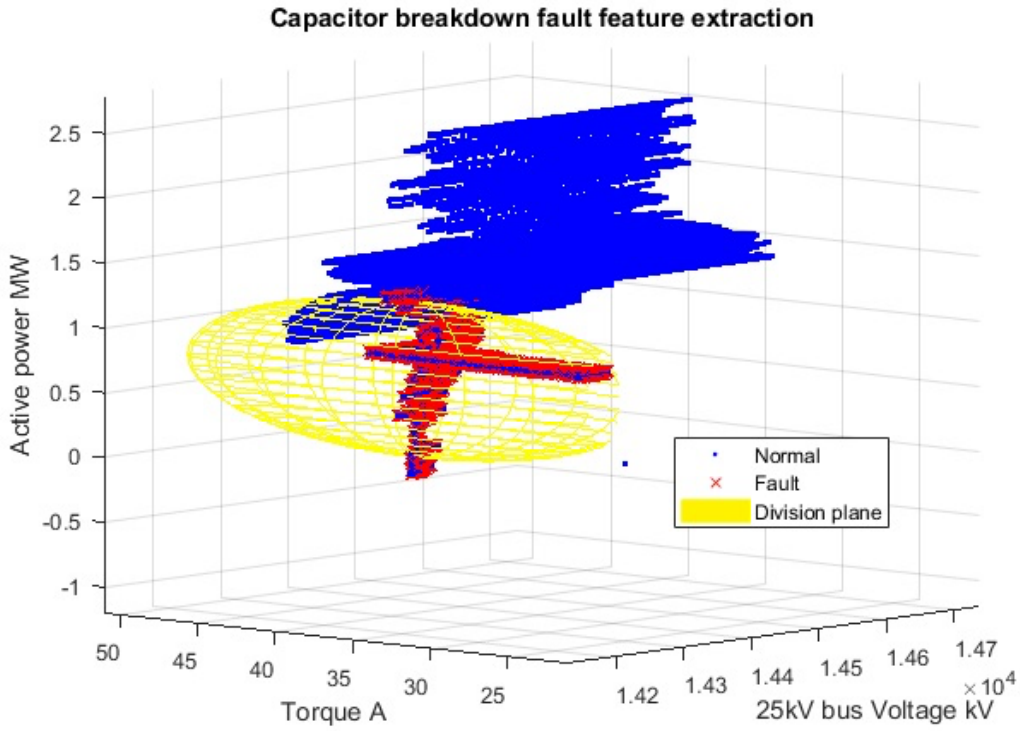


Figure 4.12 Capacitor breakdown fault feature extraction

KSVM is also applied to extract the feature of the capacitor breakdown fault, as shown in Figure 4.12. Similar to the previous case, the blue dots indicate the normal feature and the red crosses indicate the capacitor breakdown fault, which is separated by the yellow division plane. Any data points within the division plane are considered faults. The division plane used to extract the fault feature can be described as $(x-14168.84)^2+(y-29.99)^2+(z+1.01)^2-130.11<0$. When this inequality is satisfied, a capacitor breakdown fault can be observed. The figure shows that when a fault occurs, the torque tends to be lower, and thus, the power output is reduced.

4.4 Summary

This chapter presents a novel fault diagnosis and detection method based on a multivariate statistical technique combined with the ML algorithm based on KSVM.

The turbines with a gearbox bearing fault and a generator winding fault are tested by using the proposed model. The results can be concluded that the fault can be localised via calculating the KLD values between healthy and faulty conditions. The amplitudes of KLD values indicates the contributions brought by the variables to the fault. By selecting the most significant variables for classification model training, it reduces computation load and improve efficiency. The fault feature separation plane can extract the gearbox bearing fault the generator winding fault accurately.

Further work will focus on applying the fault feature model to the turbines with other faults based on deep learning algorithms. Other multivariate statistical techniques such as Wasserstein metric will be considered and incorporated for fault diagnosis.

Chapter 5. Alarm-related wind turbine fault detection based on kernel support vector machines

This chapter proposes a data driven model based fault detection method for WT systems, in which the detection process is based on the alarm signals and modelling by KPCA and KSVM. The deployment of large-scale WT systems requires better-scheduled maintenance schemes to ensure the systems are under normal, safe and cost-effective working condition. To improve the performance of WT CMS and reduce the O&M cost, and offer an efficient approach to achieve this goal, a model based data driven fault detection method based on KPCA and KSVM is developed. This method is validated with real historical SCADA data from an operational wind farm.

5.1 Introduction

This chapter presents research results of a model-based data- driven WT fault detection method, which creates a relationship to identify the false alarms and true alarms related to the faults. The model is performed using the KSVM incorporating the KPCA based on the historical SCADA data. The alarm of WT system can be triggered when key component signals exceed the pre-defined threshold limits usually due to design defects, changing of WT running states and components malfunction [40]. Since the alarms could reveal the working conditions of the turbine's components, it can be regarded as a significant index to indicate an early warning of the vital faults. Firstly, the computation load can be reduced by choosing specific PCs. Secondly, the chosen PCs are used to build normal-abnormal classification model. Finally, a classification model based on the extracted abnormal data is built to classify the alarms and faults.

5.2 Methodologies

5.2.1 Principal components analysis

The PCA transforms a set of correlated variables into a set of linearly uncorrelated variables, which are the PCs of the original dataset. It has been widely used to visualize relatedness and genetic distance between variables. The process can be achieved by calculating the eigenvalues of the covariance matrix or singular values of non-orthogonal matrix condition [7, 8]. PCA has shown its strong capability in dimension reduction and been verified by researches in different fields [177]. By selecting first few PCs, the major information can be maintained and the dimension of the original dataset is then dramatically reduced. Hence, this technique has been widely applied in feature extraction and incorporated with various ML algorithms such ANN to monitor and predict the performance of WTs [178]. To obtain the PCs from a dataset X with n-by-p dimensions, where p is the number of the variables and n is the number of the samples of each variable, Eigen-analysis for the covariance matrix M of original dataset X needs to be performed. First, the dataset X need to be standardised:

$$z_j = \frac{x_j - \overline{x_j}}{\sigma_{x_j}} \qquad (j = 1,2,\dots,p) \tag{5-1}$$

where $\overline{x_j}$ is the mean value of x_j, σ_{x_j} is the standard deviation of x_j and $\boldsymbol{Z} = [z_1, .., z_p]$ is denoted as the standardised dataset with n-by-p dimensions. The covariance matrix M of Z is defined as:

$$M_{i,j} = cov(Z_i, Z_j) = E\left[(Z_i - \mu_i)(Z_j - \mu_j)\right]$$

$$= E\left[Z_i Z_j\right] - \mu_i \mu_j \tag{5-2}$$

where $\mu_i = E(Z_i)$ is the mean value of i th row of Z. The PCs can be derived from the covariance matrix by using Singular Value Decomposition (SVD). The singular values of the matrix M can be calculated by

$$M = USW^T \tag{5-3}$$

where S is an n-by-p rectangular matrix contains the i-th singular values of M. U is an n-by-n matrix called the left singular vectors consists of the n largest eigenvalues of MM^T and W^T is a p-by-p matrix called right singular vectors associated with the orthonormalised eigenvectors of M^TM [179]. By sorting the singular values in descending order and finding their corresponding singular vectors in the same order, the i-th principal component can be obtained by following equation:

$$Y_i = U_{i1}z_1 + U_{i2}z_2 + \cdots + U_{ip}z_p \qquad (i = 1,2, \dots, p) \tag{5-4}$$

The singular values of M are the variances of their corresponding PCs. Hence, the magnitudes of each singular value represent the weighted information contained in the original dataset. To select the number of PCs, the accumulated variance contributions from each principal component need to be calculated. The variance contribution a_i of i-th principal component' variance s_i is defined as:

$$a_i = \frac{s_i}{\sum_{i=1}^{p} s_i} \tag{5-5}$$

To obtain the information from the original dataset, the selection of k PCs should be as large as possible ($k<p$). However, the number of PCs must be compromised in order to achieve the dimension reduction. In our study, the accumulated variance contribution is selected no smaller than 85%.

5.2.2 Kernel function in PCA and SVM

The SVM could only solve linear separable problems. Hence, to solve a larger dataset with a linear inseparable problem, kernel function is introduced. By using kernel, the linear operations of PCA are performed in a reproducing kernel Hilbert space. Therefore, the linear inseparable problem can be solved by using kernel function projecting to a higher dimension.

KPCA is an extension version of the PCA using the kernel function to perform the originally linear operations in a reproducing kernel Hilbert space. As introduced above, the calculation of PCA can be transferred into the Eigen-analysis. By mapping the original data points in the feature space using the RBF kernel. It is defined as:

$$K(Z) = K(Z, Z^T) = e^{\Lambda(-|Z-Z^T|^2)} \tag{5-16}$$

where the Z is the original input dataset and Z^T is its transpose [180]. $|Z - Z^T|^2$ is considered as the squared Euclidean distance between them. The Λ is the width of the kernel, which cannot be predicted precisely and has to be constrained by the model or defined by the user [181].

By replacing the original dataset with the kernel, the covariance matrix of Equation (5-2) can be rewritten as

$$C = \frac{1}{n}\sum_{i=1}^{n}(K(z_i) - \overline{K})(K(z_i) - \overline{K})^T \tag{5-17}$$

$$\overline{K} = \frac{1}{n}\sum_{i=1}^{n} K(z_i) \tag{5-18}$$

Then, following the same procedures as described from Equations (5-3)-(5-4), the singular values and vectors, the kernelised PCs can be obtained.

5.2.3 KPCA and KSVM based classification model

The SCADA data used in this paper were acquired from an operational wind farm which consists of 26 turbines over a period of 12 months. To test and validate the proposed classification model, it is necessary to use historical data from an operational wind farm. Unlike the high-frequency condition monitoring data,

SCADA data have a low sampling rate usually at 10 minutes/sample in order to reduce data storage amount while still maintaining the vital information about the operation and performance of the WTs [52]. The monitoring variables for each turbine consist of 128 readings among various types of physical and electrical signals, such as temperatures, pressures, power outputs and control signals. Pre-processing to the data is essential for further analysis due to the occasions that the turbines are in inactive during the periods of low and high speeds. Besides, the digital and constant data need to be removed to prevent inferences to the processing [22, 182].

As the examples, Figures 1-3 show the wind power curve of three different turbines. For WTs, the S-curve refers to the relationship between the output power and wind speed [183]. The output power would often be reduced when the fault occurs in order to prevent the fault being developed into the detrimental one. The dashed box indicates the fault area. As can be observed from the figures, the turbine with generator winding fault has a shorter time period of fault exposure compared to the turbine with gearbox bearing fault.

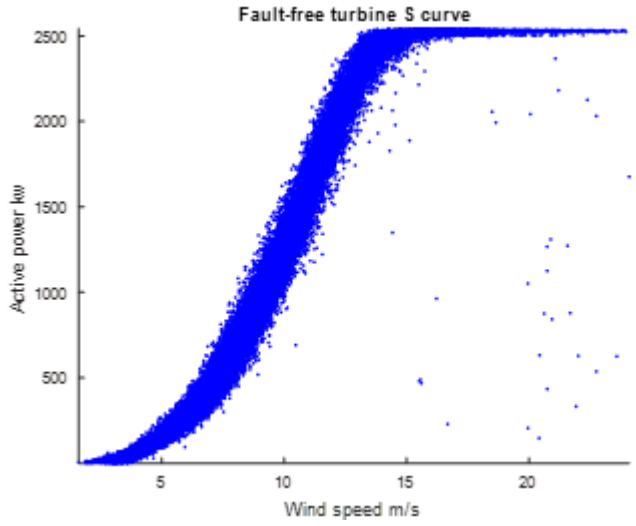


Figure 5.1 Power curve of the fault-free turbine

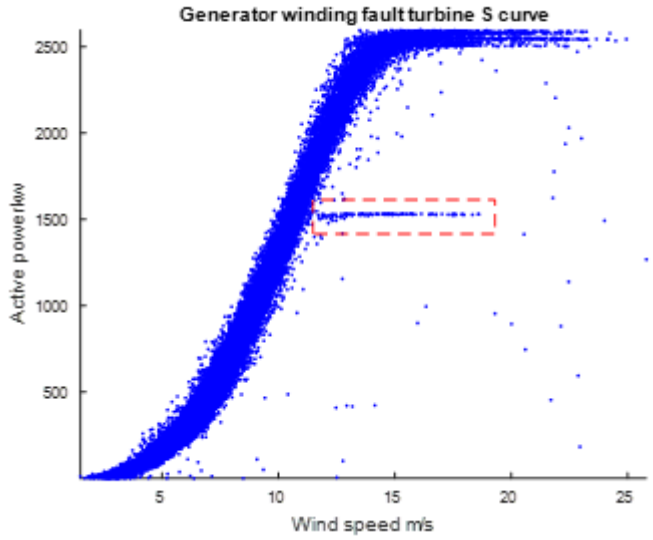


Figure 5.2 Power curve of the turbine with generator winding fault

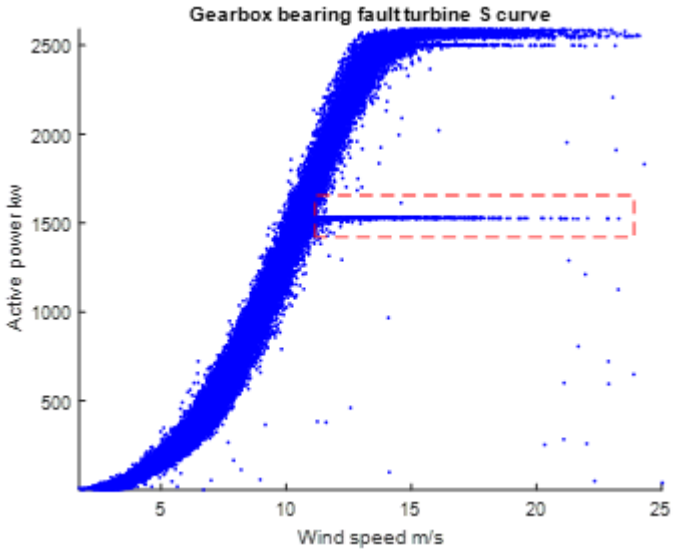


Figure 5.3 Power curve of the turbine with gearbox bearing fault

To detect the faulty condition of the WT, a two-stage classification method is proposed, as illustrated in Figure 5.4. By checking time-series data, the original dataset includes data under the normal working condition and those alarm data. The alarm data also contain the fault data related to the alarms triggered during the fault period. Then abnormal data are further classified into the true positive signals, indicating the occurrence of a real fault, and false positive signals, which can be considered as a warning.

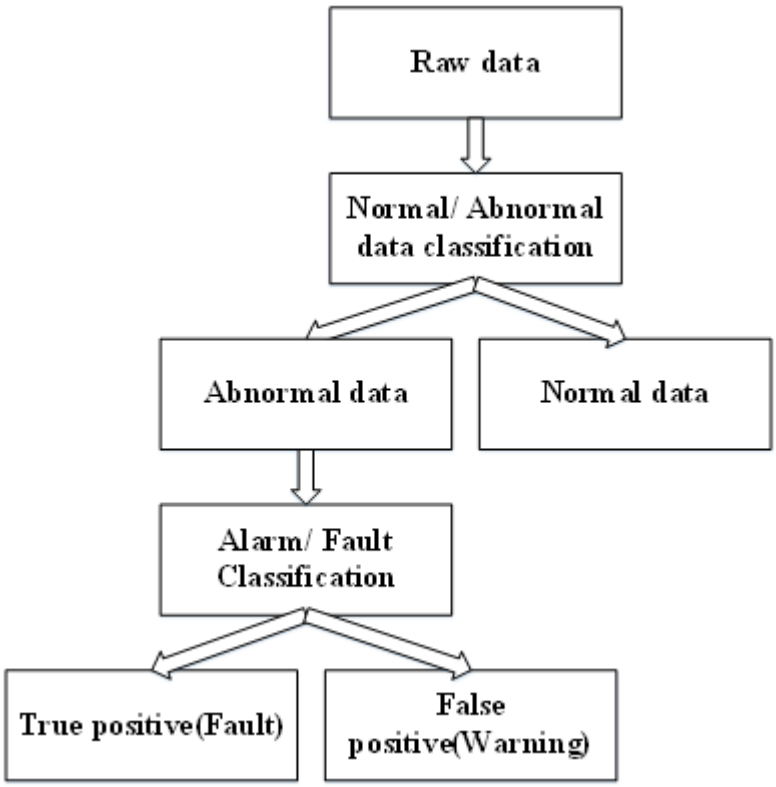


Figure 5.4 Overall modelling procedures

Three normal data selection methods are used in our study. The first one is to choose first 5000 samples in the original dataset, which is referred as to the method 1. The second method is to choose 2500 samples before and after the fault respectively, which is referred as to the method 2. The last method is to choose 5000 samples randomly among the normal data, which is referred as to the method 3. The fault detection method is then applied to both faulty turbines, as shown in Figures 5.2 and 5.3. The results given in the next section are based on the turbine with a gearbox bearing fault with the normal data being selected using method 3.

5.3 Model validation

5.3.1 Monitoring variable selection of SCADA data

After pre-processing the original data by removing those control and DC signals, there are 78 variables in total remaining for further data dimension reduction. All the data samples relating to the fault are selected and processed with KPCA. To select the appropriate PCs, the variance contribution of each principal component needs to be calculated, as given in Table 5.1. 16 PCs are therefore selected in order to meet the requirement of achieving 85% accumulated variance contribution.

Table 5.1 Variance contribution of the principal components

PCs	1	2	3	4	5	6
Contribution%	6.59	6.56	6.34	6.34	6.17	5.64
PCs	7	8	9	10	11	12
Contribution%	5.64	5.53	5.49	5.47	5.13	4.42
PCs	13	14	15	16	…	78
Contribution%	4.41	4.4	3.72	3.68	…	0.000975

5.3.2 Normal-abnormal condition classification using SCADA data

The selected PCs will be further processed by KSVM. Since KSVM is a supervised learning algorithm, the dataset needs to be divided into two groups, the data under normal conditions and the data under abnormal condition (formed by false alarms and true alarms related to the fault). Since it is impossible to plot 16 dimensional graph form the selected 16 PCs, all the results will be plotted in 2D space in relation to wind speed and active power.

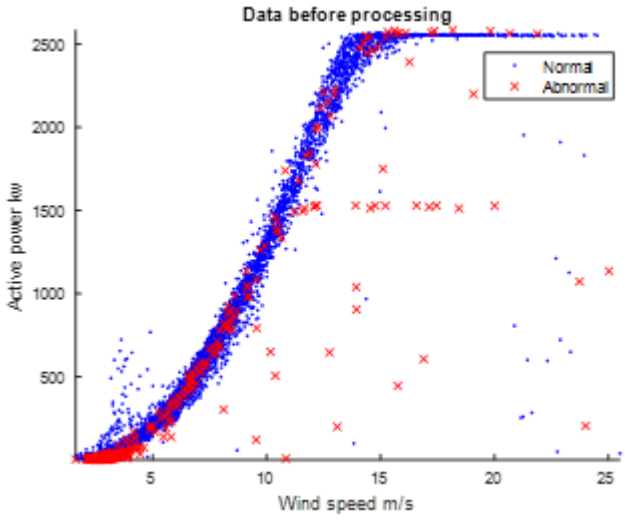


Figure 5.5 Power curve of normal-abnormal data

Figure 5.5 gives an example of the data needing to be processed for normal-abnormal classification, where the blue dots represent the normal data and red crosses represent the abnormal data.

As mentioned above, to process the data using KSVM algorithm, the linear inseparable data in a lower dimension can be projected into a higher dimension and thus differentiated by a hyperplane. As an example, Figure 5.6 shows the working principle of the KSVM, where the blue dots represent the normal data and red dots

represent the abnormal data. The support vectors are labelled by green circles while the fitted hyperplane is demonstrated in gradient colour. The function of the fitted hyperplane is expressed as

$$z = 51.68 + 1.382x - 41.55y - 4.574x^2 + 28.84xy + 17.82y^2 \qquad (5\text{-}16)$$

where x, y are the wind speed and active power respectively. The coefficient of determination r^2 is used to evaluate the accuracy of the fitting and the value of this fitted plane is 0.8605.

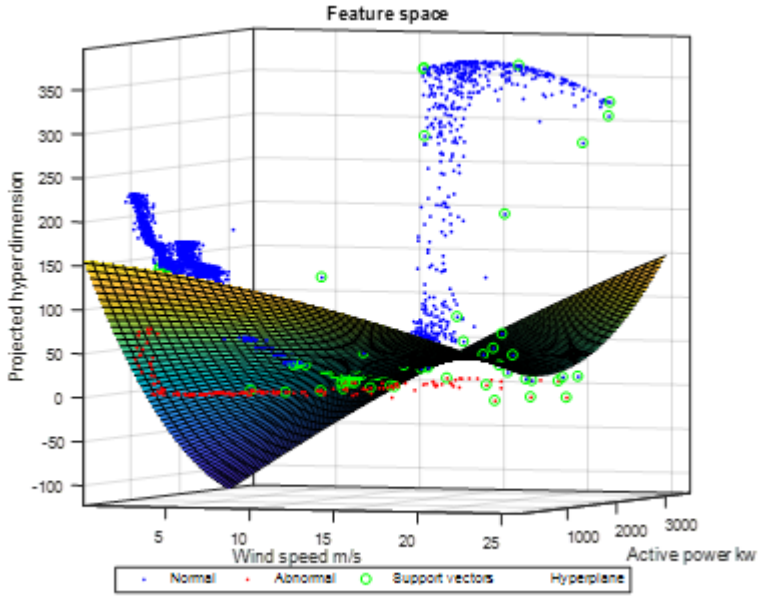


Figure 5.6 Normal-abnormal data classification using KSVM in the hyper dimension

During this process, 70% of the data were used as the training set and 30% of the data were used for the validation. The validation result is displayed in Figure 5.7 and 5.8. In Figure 5.7, the normal data classified as normal are shown in blue dots while the normal data classified as alarm are shown in blue crosses; the alarm data classified as alarm are shown in red dots and alarm data classified as normal data are shown in red crosses. Figure 5.8 shows the confusion map of the normal-alarm classification result, which is used to evaluate the performance of the algorithm. The white areas show rates of both normal and alarm data were predicted correctly and the yellow areas show the misclassified data. As can be observed from the figure, the predicted normal data has reached 99.9% true and alarm data has reached 90.9% true, leading to a total accuracy of 99.4%.

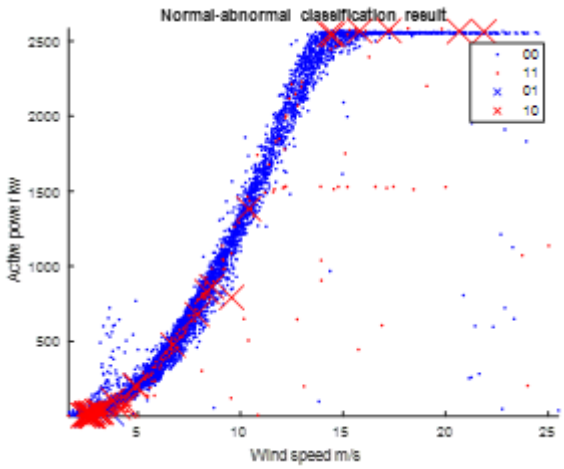


Figure 5.7 Normal-abnormal classification result

		Normal	Abnormal
Actual Class	Normal	99.9%	0.1%
	Abnormal	9.1%	90.9%
		Prediction Class	

Figure 5.8 Confusion map of normal-abnormal classification result

5.3.3 Alarm-fault classification using SCADA data

After the procedure of normal-alarm classification, the alarm-fault classification is then processed. The fault data contained in abnormal data was selected in another group and model with other non-fault alarms.

Figure. 5.9 shows the alarm-fault classification in the relationship between wind speed and active power. The blue dots represent alarm signals and red dots represent for fault signals. The support vectors are labelled in the green circle and the classification hyperplane is fitted as

$$z = 68.2 + 14.47x - 16.24y + 6.274x^2 + 8.044xy - 8.719y^2 \qquad (5\text{-}17)$$

where *x, y* are the wind speed and active power. The r^2 of the fitted plane is 0.7639.

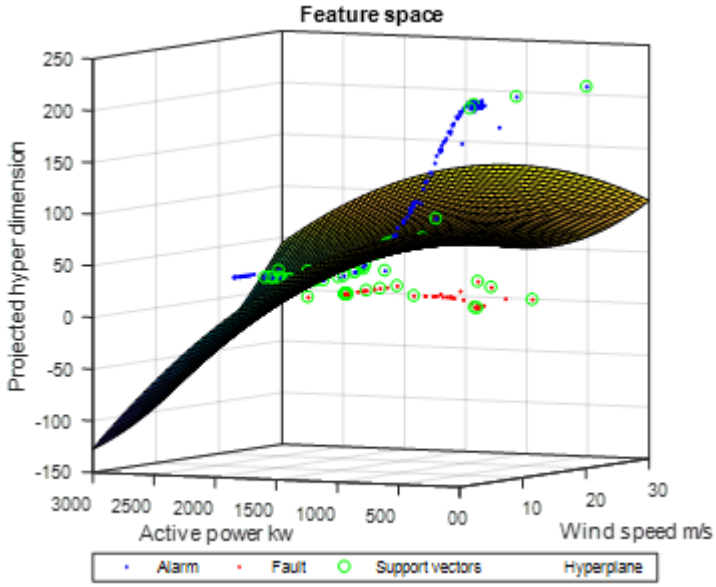


Figure 5.9 Alarm-fault classification using KSVM in hyper dimension

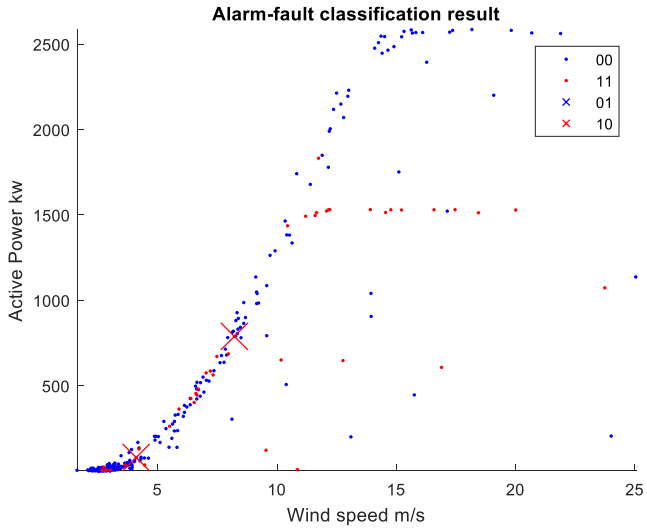


Figure 5.10 Alarm-fault classification result

Figure 5.10 and 5.11 give the validation result of the alarm-fault classification model. The accuracy of alarm-fault classification model is also evaluated by the confusion map. It can be seen from the figure that it reached 100.0% accuracy on alarm signal classification and 95.4% accuracy on fault signal classification. The total accuracy on alarm-fault classification has reached 99.3%.

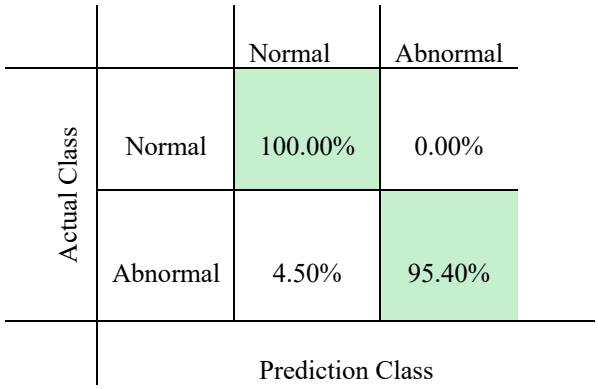


Figure 5.11 Confusion map of alarm-fault classification result

5.3.4 SCADA based classification comparison with other methods

Table 5.2 Classification results for the faulty turbines based on three different data selection methods

	Normal-abnormal classification	Normal-abnormal classification
Gearbox bearing fault turbine with method 1	0.997	0.963
Gearbox bearing fault turbine with method 2	0.998	0.986
Generator winding fault turbine with method 1	0.969	0.931
Generator winding fault turbine with method 2	0.973	0.762
Generator winding fault turbine with method 3	0.936	0.857

In order to examine the robustness of the proposed methods, more turbines are tested with different SCADA data selection methods. It can be observed from the Table 5.2 that the performances of the turbine with generator winding fault are not as good as the turbine with gearbox bearing fault. This might be due to the insufficient samples acquired from alarm and fault signals. It can be believed that if the amount of abnormal data increases, the classification model accuracy can be improved further.

5.3.5 Monitoring variable selection of simulation data

As described in chapter 3, two types of faults were simulated. A three-phase short-circuit fault was generated first.

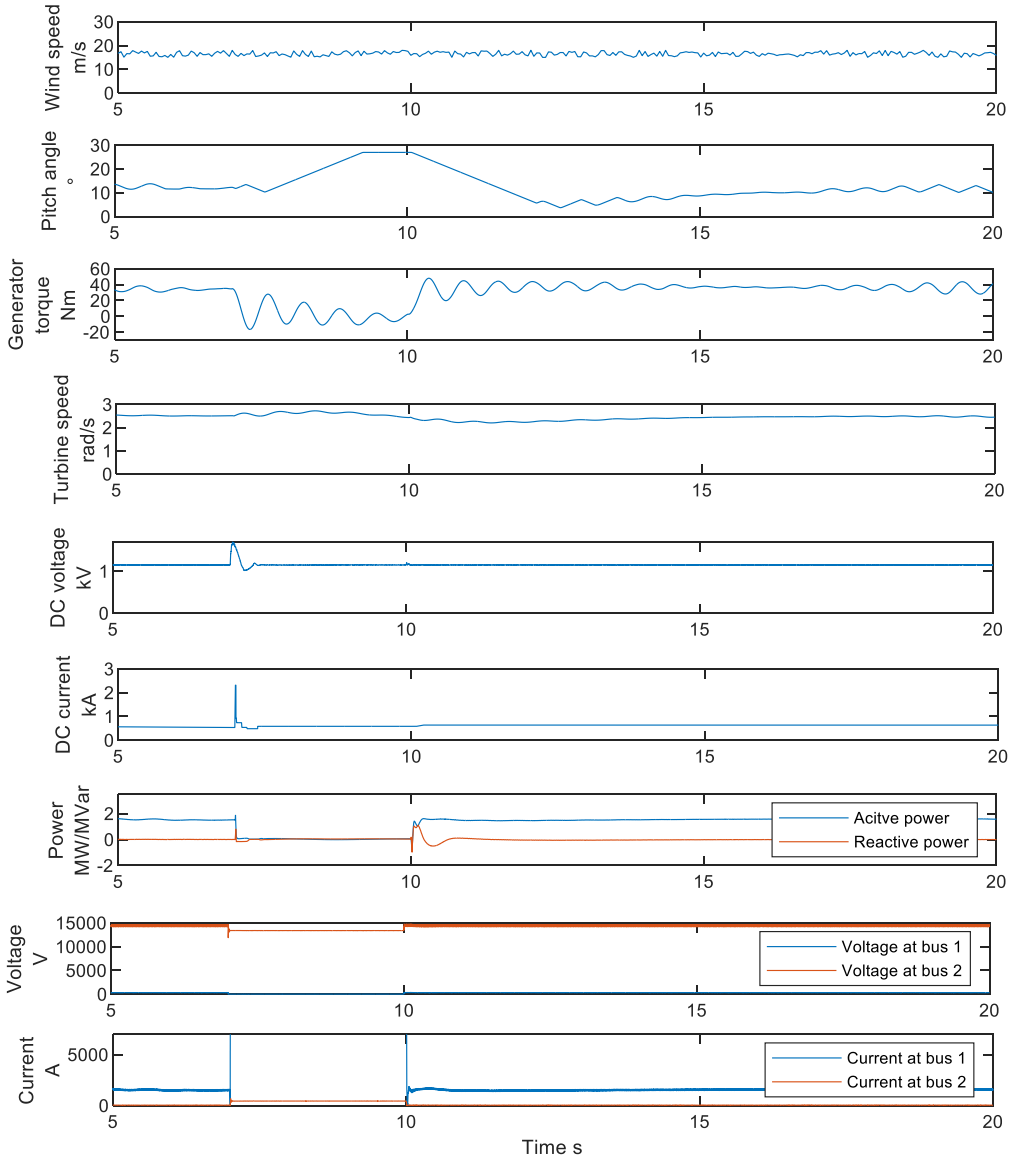


Figure 5.12 Simulated DFIG WT three-phase short circuit fault

Figure 5.12 shows the simulation results for the DFIG WT three-phase short-circuit fault. From top to bottom, it presents the input wind speed, pitch angle of the blade, torque of the turbine, generator speed, voltage of the DC link, current of the DC link, active and reactive power of the turbine, RMS voltages at two different locations, and RMS currents at two different locations. The simulation was run for 20 s. The fault was injected at 7 s and ended at 10 s. The data of WT starting up period ($t < 5$ s) were neglected, and the figure only shows the time period between 5 and 20 s.

After the original data were acquired and the power output was removed, 10 variables remained for further data dimension reduction. All data samples related to the fault were selected and processed using KPCA. To select the appropriate PCs, the

variance contributions of the PCs were calculated and sorted in descending order, as shown in Table 5.3. Therefore, 16 PCs were selected to satisfy the requirement of an accumulated variance contribution of 85%.

Table 5.3 Variance contribution of the PCs for a three-phase short-circuit fault

PCs	1	2	3	4	5
Contribution%	70.35	29.61	0.02	0.02	0.0011
PCs	6	7	8	9	10
Contribution%	0.000441	0.00000713	1.26E-07	3.48E-46	0

As indicated by the table, the first two PCs already satisfy the requirement of variance contribution in this circumstance; they are the 25-kV bus voltage and 575-V bus current. Hence, the KSVM modelling is based on these two monitoring variables.

A capacitor breakdown fault was generated to test the accuracy of the proposed model. Figure 5.13 shows the simulation results for the DFIG WT capacitor breakdown fault.

Similar to the three-phase short-circuit fault simulation, the capacitor breakdown fault was simulated for 20 s, and the fault was injected at 7 s. The figure shows only the time period between 5 and 20 s. Similarly, 10 variables and their variance contributions are presented in Table 5.4. As shown, the first two PCs already satisfy the requirement of variance contribution, with an accumulated contribution of >85%. Because the PCs were obtained from the same system, the order of the calculated PCs was similar to that reported previously. The 25-kV bus voltage and 575-V bus current were selected for further classification modelling.

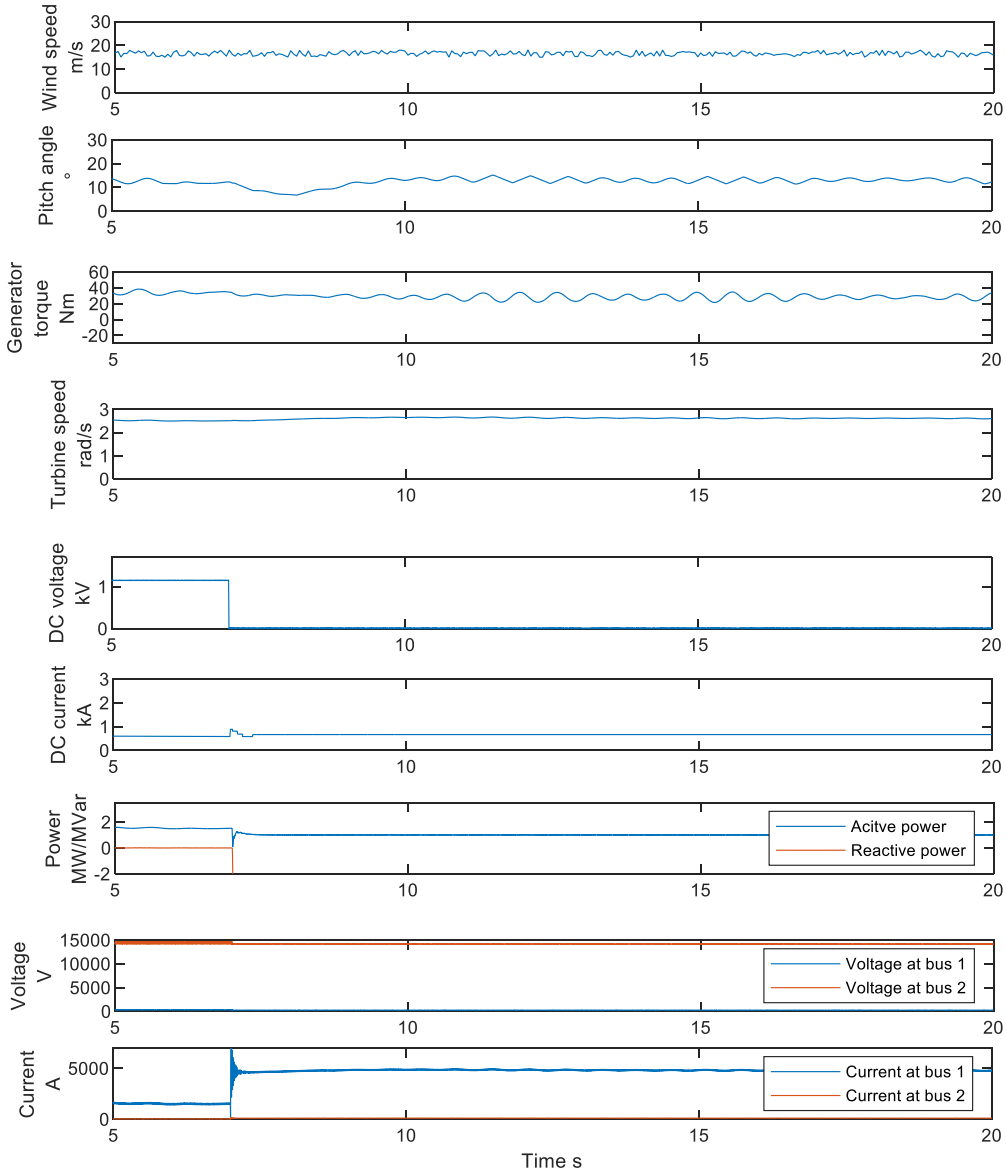


Figure 5.13 Simulated DFIG WT capacitor breakdown fault

Table 5.4 Variance contributions of the PCs for the capacitor breakdown fault

PCs	1	2	3	4	5
Contribution%	77.72	20.09	0.1338	0.0521	3.11E-05
PCs	6	7	8	9	10
Contribution%	4.66E-06	1.58E-07	7.77E-10	2.11E-11	0

5.3.6 Normal-fault condition classification using simulation data

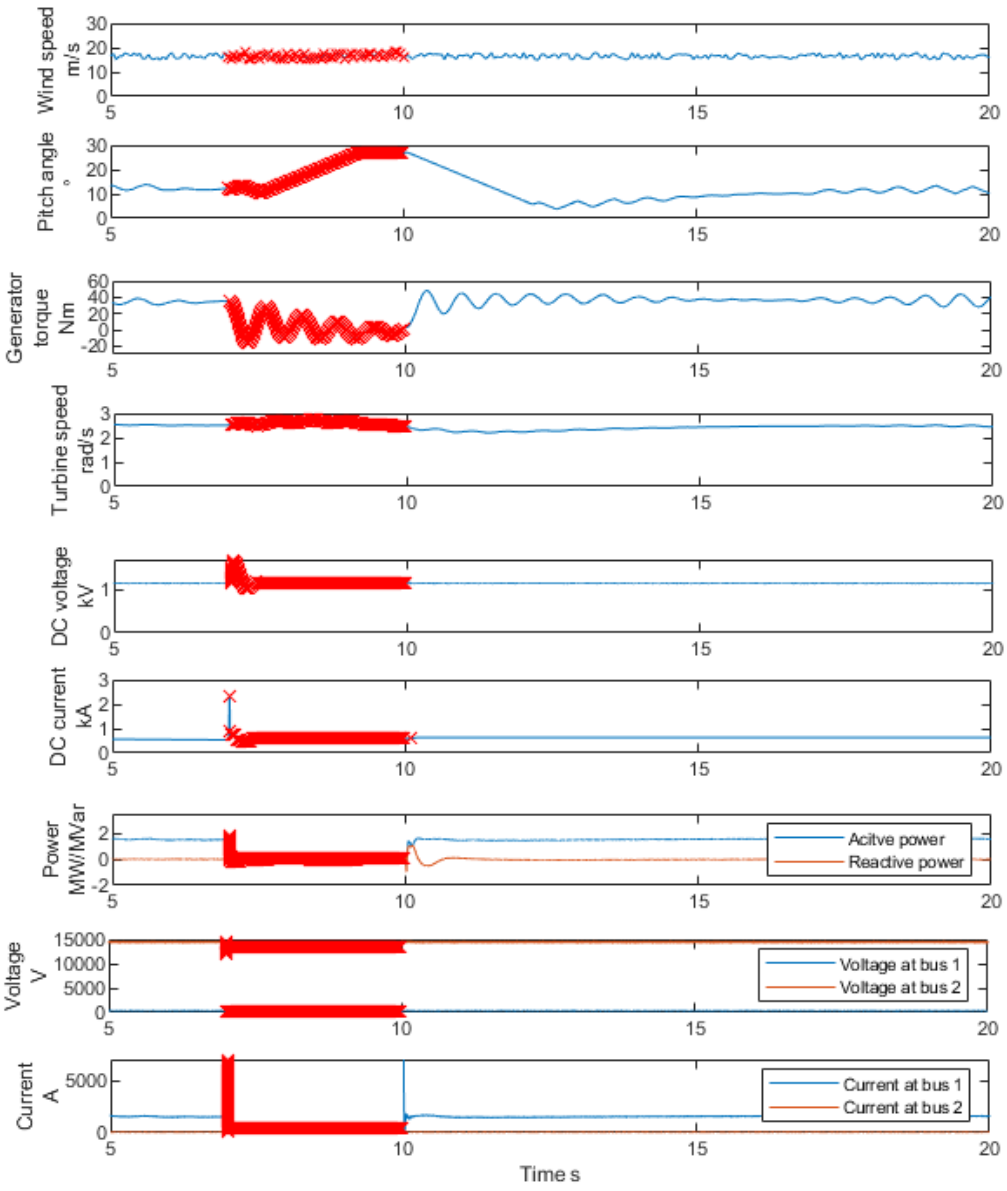


Figure 5.14 Classification results for the DFIG WT three-phase short-circuit fault

Because the simulation model cannot provide alarm signals, the classification model was only tested in normal-fault condition classification. Figures 5.14 and 5.15 show the classification results for the DFIG WT three-phase short-circuit fault and capacitor breakdown fault. Instead using hyper dimension to demonstrate results, time-series data are used in this chapter to show the capability of KSVM in solving time-series data. Because of the limitation of the computer, only limited data can be simulated. The data in the time period are indicated by the red crosses. As indicated by the figures, faulty time periods were distinguished.

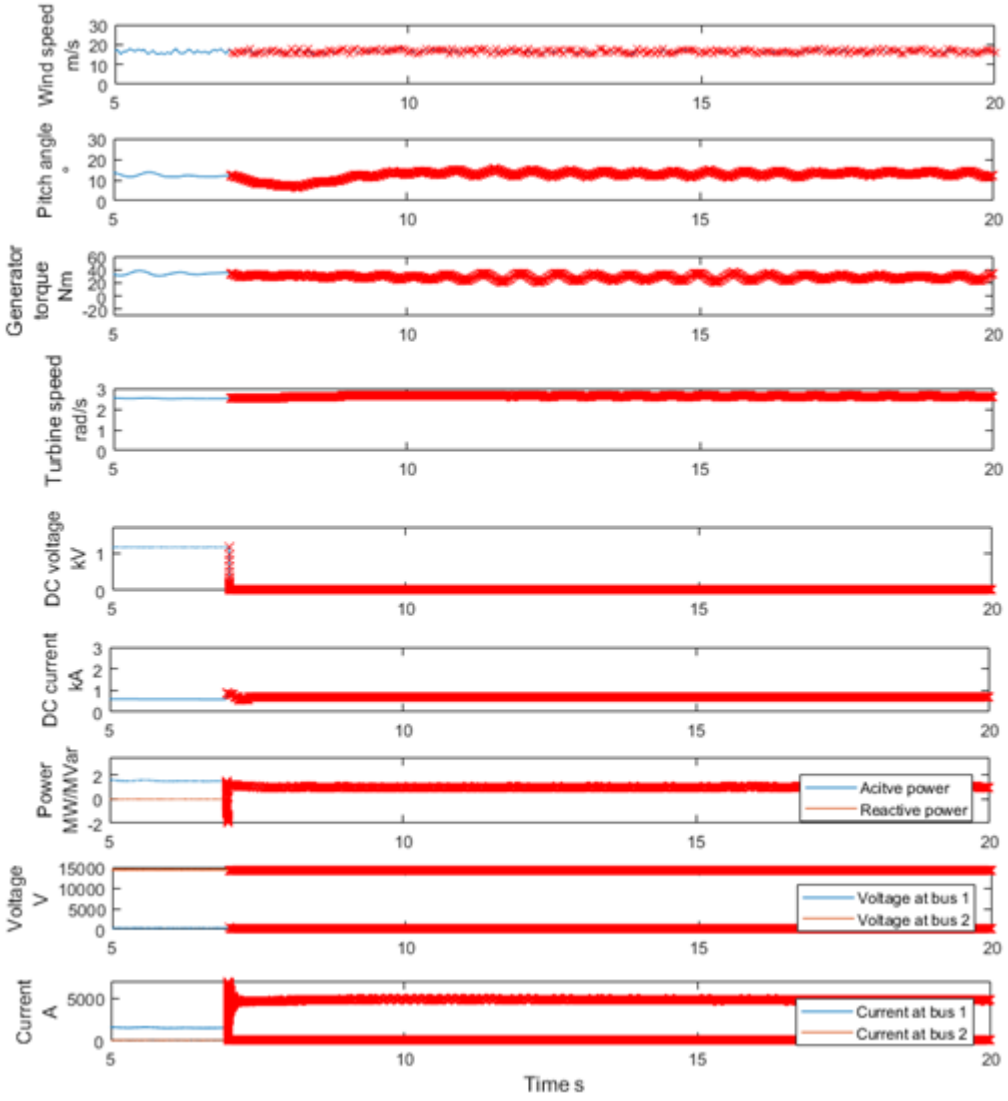


Figure 5.15 Classification results for the DFIG WT capacitor breakdown fault

To further study the accuracy of the proposed classification model, a confusion matrix was created. Figure 5.16 presents the validation results for the three-phase short-circuit fault. As shown, the fault classification accuracy reached 100.0%. However, approximately 0.2% of the normal data were misclassified as abnormal data. The overall accuracy of three-phase short-circuit fault classification was 99.9%. Figure 5.17 shows the validation results for the three-phase short-circuit fault. Similar to the three-phase short-circuit fault classification, the capacitor breakdown fault classification accuracy reached 100%. The normal classification accuracy was slightly lower (99.3%). The total accuracy of capacitor breakdown fault classification was 99.7%.

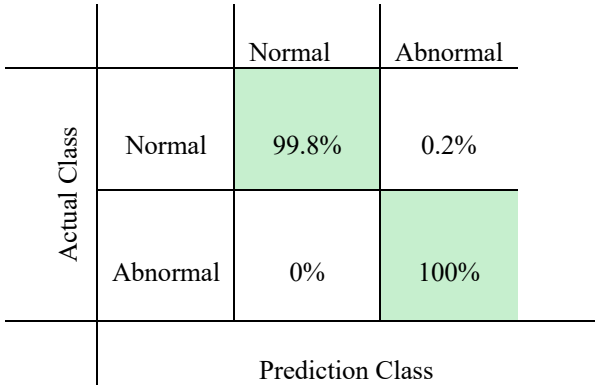


Figure 5.16 Confusion map of three-phase short circuit fault classification results

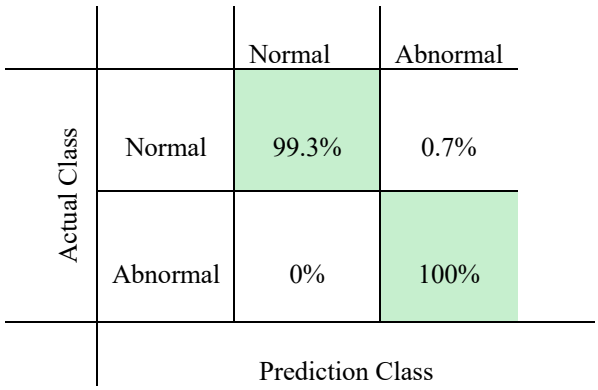


Figure 5.17 Confusion map of capacitor breakdown fault classification results

5.4 Summary

With these alarm signals being identified, the fault can be warned at an early stage, which leaves sufficient time for maintenance scheduling. According to the results, it can be concluded that the accumulated variance of the PCs can be regarded as the most significant factor by selecting PCs of the monitoring variables. Compared with other ML algorithms, the SVM has its strength on solving the two-group classification problem. Compared with the decision tree and discriminant analysis algorithms, the SVM demonstrates more accurate results. In terms of sample data selection, the turbine, which has large amount of abnormal data, shows a better classification performance, indicating the influence of the sample selection. The KPCA can reduce the dimension in an acceptable range while the KSVM demonstrates excellent results for the two-stage classification. Further work will be focused on examination of the proposed approach incorporating with deep learning algorithms and verification of the results with more data from both simulations and physical test rig.

Chapter 6. A hybrid LSTM-KLD approach to condition monitoring of operational wind turbines

This chapter presents a novel ML model-based data-driven approach to accurately evaluate the performance of the turbines and diagnose the faults. The approach is based on LSTM incorporating a statistical tool named KLD. The hybrid LSTM-KLD method has been applied to two faulty WTs with gearbox bearing fault and generator winding fault respectively for fault detection and identification. The proposed method is then compared with three other well-established machine-learning algorithms to investigate its superiority. The results show that the proposed method can produce a more effective detection with accuracy reaching 94% and 92% for the turbines, respectively. Furthermore, the proposed method can effectively distinguish the alarms from the faults, from which the distinguished alarms can be considered as an early warning of the fault occurrence.

6.1 Introduction

In this chapter, a novel condition monitoring approach is presented based on a deep learning algorithm incorporating a statistic tool to estimate the operating condition of WTs. Multiple monitoring variables that contribute to the specific subsystems are taken into consideration to improve the reliability of fault diagnosis. Besides, the alarms are also fully used as significant evidence to support diagnostic results. Specifically, the LSTM is adopted to achieve the behaviour prediction of the key subsystems and then the KLD is employed to detect the fault by comparing probability distributions of the variables over the time. In the end, the monitoring data can be classified as normal, fault, true alarm and false alarm while the severity of the fault is also evaluated.

6.2 Methodologies

The LSTM algorithm was first proposed in 1995 by Sepp Hochreiter and Jurgen Schmidhuber [184]. It is a type of deep learning algorithm whose structure is implemented based on RNN. Unlike other regression based ML algorithms, the cell memory of LSTM can preserve the hidden state through time, while, in the meantime, adding new information. The KLD used in this paper is also called relative entropy, which has been applied as a measure of data representativeness. The KLD is a statistical tool developed to measure difference between two probability density distributions. It has been widely used in neuroscience and ML due to its ability in characterising relative entropy in information systems [173].

6.2.1 Long-short term memory

As mentioned, LSTM is developed based on RNN in order to solve the vanishing gradient problem. The feedback loops are embedded in every recurrent layer of the RNN, and thus the information can be preserved. However, with the increasing feedback loops in RNN, the gradient of the loss function decays exponentially with time. Compared to RNN, the LSTM has four interacting layers (cell state, forget gate, input gate and output gate) inside a LSTM cell [185], where, except for the standard units of RNN, a set of cell states and gates are added to control which memory should be stored. With this structure, the gradient decent problem of RNN

is solved and "long-term memory" is achieved. The overall structure of the LSTM is illustrated in Figure 6.1.

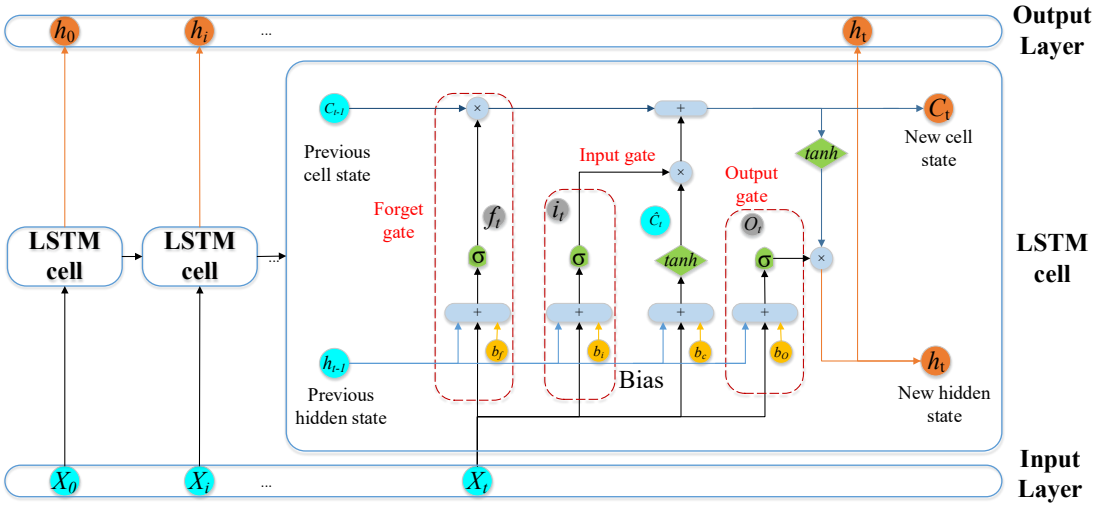


Figure 6.1 Schematic diagram of the LSTM structure

At time step t, the LSTM cell has an input vector $[h_{t-1}, X_t]$. The new cell state, denoted by C_t, and the new output, denoted by h_t, inherit information from former states (C_{t-1} and h_{t-1}) and are transmitted to the next cell at time step $t+1$. The output h_t of the cell is determined by the current input X_t and the previous outputs h_{t-1}, h_{t-2}, while the cell state C_t can preserve information to flow forward without any losses. The gates in a LSTM cell include a forget gate (f), an input gate (i), and an output gate (O). The forget gate determines what information will be added (or forgot) to the cell state C_{t-1} when new input enters the network. The input gate decides what new information from the input will be updated into the cell state. The output gate decides what information updated in the cell state is sent to the network as input for next time step $t+1$, as represented by h_t.

Mathematically, the output f_t of the forget gate is given by

$$f_t = \sigma(W_f \odot [h_{t-1}, X_t] + b_f) \tag{6-1}$$

where σ is the sigmoid activation function; W_f and b_f are the weight and bias of the forget gate, respectively, and the operator $\odot$ denotes element-wise multiplication. The output of the sigmoid function is a value between 0 and 1, with 0 representing to forget this value and 1 representing to preserve this value.

Two activation functions are then used to decide which information to be stored in the cell. The sigmoid function σ at the input gate is used to decide which information to

update while the hyperbolic tangent (*tanh*) function is used to decide what new information to add to the cell state denoted by $\hat{C}_t$.

$$i_t = \sigma(W_i \odot [h_{t-1}, X_t] + b_i) \tag{6-2}$$

$$\hat{C}_t = tanh\,(W_C \odot [h_{t-1}, X_t] + b_C) \tag{6-3}$$

where i_t is the output of the input gate; W_i and b_i are the weight and bias of the input gate whereas W_C and b_C are the weight and bias of the *tanh* function, respectively.

This leads to the cell state C_t from C_{t-1} by computing the function below.

$$C_t = C_{t-1} f_t + i_t \hat{C}_t = C_{t-1} f_t + tanh\,(W_C \odot [h_{t-1}, X_t] + b_C) i_t \tag{6-4}$$

Then the LSTM cell is to compute the output O_t of the output gate by computing the following function.

$$O_t = \sigma(W_O \odot [h_{t-1}, X_t] + b_O) \tag{6-5}$$

where W_O and b_O are the weight and bias of the output gate, respectively. The sigmoid function σ in Equation (6-5) is used to decide which part of the cell state should be outputted.

The last step is to compute the output h_t of the cell, as described below.

$$h_t = O_t \tanh\,(C_t) \tag{6-6}$$

where the *tanh* function is used to fit the output value in the range between 0 and 1.

The vanishing gradient problem in LSTM is solved by back-propagation. The gradient is calculated through back propagation along time using the chain rule. With all gradients calculated according to the corresponding error term (loss function), the weights associated with input gate, output gate, and forget gate are updated. More details about the back-propagation process of the LSTM can be refereed in the literature [186].

Compared to other ML algorithms such as CNN and DBN, one of the significant advantages of LSTM is that it can link the previous information to the current state. LSTM is a time related RNN, suitable for processing and predicting important events with relatively long time intervals and delays in time series. Many researches have

employed LSTM and proved that it has advantages in solving time-series data especially for predicting performance of the rotating machinery components [187-189]. Hence, LSTM is chosen as the data modelling method in our study.

6.2.2 Hybrid method incorporating LSTM and KLD

To implement the proposed model, four steps are required following pre-processing of the raw data, as illustrated in Figure 6.2. First, wind speed and active power output are selected as the model inputs, while temperature and pressure variables, which reflect the operation condition of the subsystem as a whole, are selected as the target output. The LSTM model is then built based on those variables. To train the LSTM model, the size of training dataset is vital, which is trade-off between the prediction performance and computation time. Small datasets can ease computation complexity but likely overfit the training data, resulting in poor performance. Larger datasets can help better learn model parameters but the dataset may be over-representative of the problem along with high computation demand. In our study, ten days' data (1440 points per variable considering ten-minute interval of SCADA measurements) are used to train the model. The data are then predicted on a daily basis, i.e., by sliding window with 24 hours, continuously for 30 days to ensure the modelling accuracy. It was found that increasing the training dataset size does not have significant improvement on the performance in our study. The probability density distributions of prediction data and original data are then calculated respectively. The third step is to evaluate the discrimination between the prediction and the real data in terms of their KLD values. Note that daily data are considered to calculate the probability density distribution. The prediction values are considered as the condition at which the turbine supposes to operate while the measurement data represent the real operation condition of the turbine. Their difference can be evaluated by KLD and a larger KLD value represents a worse condition of the turbine. Finally, the normal, alarm and fault conditions are distinguished by introducing a two-level threshold strategy of the KLD values, which are defined as fault-free condition (H_0) and fault condition (H_1).

The probability density function of the KLD can be described as:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(-x-KLD_M)^2/2\sigma^2} \tag{6-7}$$

where KLD_M and *std* are the mean value and standard deviation of the KLD, respectively.

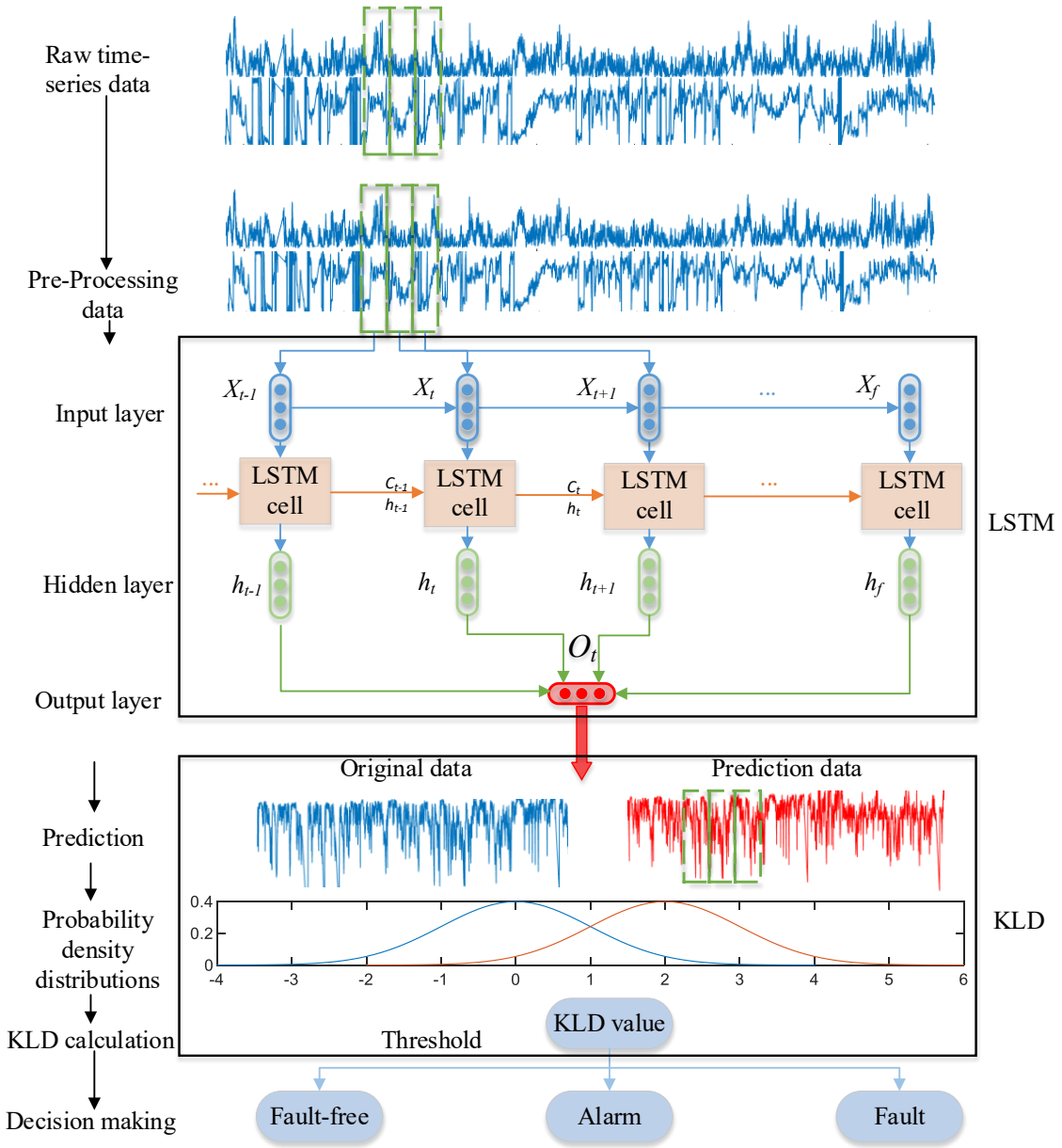


Figure 6.2 Structure of the proposed LSTM-KLD fault diagnosis method

Referring to [190], to determine the thresholds, a normal value probability (P_{NV}) and an alarm value probability (P_{AV}) are adopted here. The P_{NV} and P_{AV} can be calculated from the error function (6-8) of the KLD:

$$erf = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{6-8}$$

$$P_{NV} = 0.5(1 + \mathrm{erf}\left(\frac{h - KLD_{M0}}{\sigma_0 \sqrt{2}}\right)) \tag{6-9}$$

$$P_{AV} = 1 - 0.5(1 + \mathrm{erf}\left(\frac{h - KLD_{M1}}{\sigma_1\sqrt{2}}\right)) \quad (6\text{-}10)$$

where KLD_{M0}, std_0, KLD_{M1} and std_1 are the mean and standard deviation values at two different conditions. By combining Equations (6-9) and (6-10), a COST function can be formed to evaluate the performance of the thresholds.

$$COST = P_{NV} + P_{AV} \quad (13)$$

The optimal thresholds are chosen when the COST reaches its minimum [191] at which both P_{NV} and P_{AV} are minimum as well. The threshold H_0 is determined when KLD_M=KLD_{M0} with a corresponding standard deviation of std_0. It indicates the system is in a normal condition. When KLD_M=KLD_{M1} and $std = std_1$, the system is in an alarm condition. The data lower than H_0 are considered as normal condition while the data higher than H_1 are considered as faulty condition. The calculated KLD values between H_0 and H_1 are considered as alarm condition. Examples of these thresholds for specific subsystems of the faulty turbines will be given in the subsequent section.

6.3 Case studies

6.3.1 Case 1: Gearbox fault

The first subsystem to be studied is the gearbox that is used to transmit kinetic power to the generator from the rotor. The WT torque control is made based on the gearbox by adjusting the rotation speed and torque accordingly. The two common faults associated with the gearbox are bearing and gear teeth faults. The unpredictable wind profiles can cause rapid changes of the torque, which may lead to misalignment of the gear teeth and uneven load for the bearing. Besides, the failure of the gearbox cooling system can also result in failures of bearing and gear teeth [184]. The gearbox under our study has 6 monitoring variables, including gearbox bearing temperature 1 at the main speed shaft bearing connected to the rotor, gearbox oil pressure, gearbox oil heat exchanger output temperature, gearbox oil sump temperature, gearbox oil pressure behind pump, and gearbox bearing temperature 2 at the high-speed shaft connected to the generator. To improve the diagnosis accuracy, all these six monitoring variables are analysed.

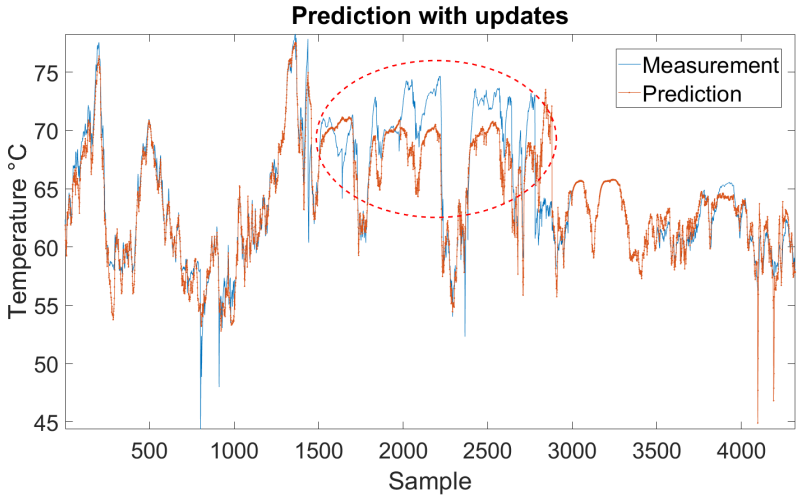


Figure 6.3 Prediction result during gearbox fault time

Figure 6.3 gives an example of the prediction result of the gearbox bearing temperature 1 as compared with the real measurements from 30 days, i.e., 4320 data samples, where the faulty time period of 10 days is circled in red. Because the prediction model is trained based on healthy data, the predicted results therefore can represent the performance of the target subsystem under fault-free condition. It can be observed from the figure that the predicted temperature is lower than the measured temperature during the fault time. The probability density distributions of both predicted and measured data are shown in Figure 6.4, where the KLD value indicating overall divergence of the two probability density distributions is 7.2893. Note that only 10-days data are shown in the Figure 6.4, focusing on the period of fault occurrence.

Table 6.1 Thresholds of the gearbox components

Threshold	Gearbox bearing temperature 1 (°C)	Gearbox oil pressure (bar)	Gearbox oil heat exchanger output temperature (°C)	Gearbox oil sump temperature (°C)	Gearbox oil pressure behind pump (bar)	Gearbox bearing temperature 2 (°C)
H_0	0.3188	0.3477	0.5462	0.1197	0.1838	0.3381
H_1	1.9212	4.3833	3.0523	2.3068	4.2164	1.9404

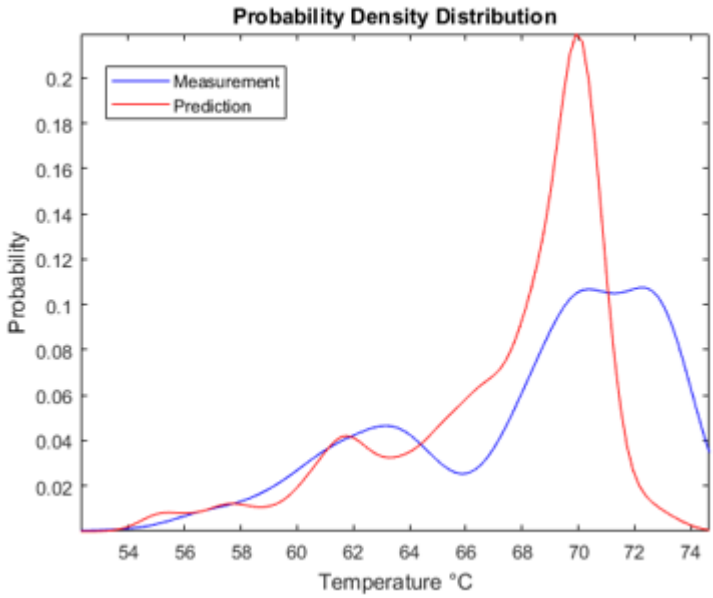


Figure 6.4 Probability density distributions of predicted and measured data of gearbox bearing temperature 1

The predicted data of all six variables are compared with measured data on a daily basis to calculate the KLD between them. The performance of the gearbox is shown in Figure 6.5, where the components 1-6 represent gearbox bearing temperature 1, gearbox oil pressure, gearbox oil heat exchanger output temperature, gearbox oil sump temperature, gearbox oil pressure behind pump, and gearbox bearing temperature 2, respectively. The fault index is measured in terms of KLD. The corresponding thresholds H_0 and H_1 for each component are also displayed in the figure to determine whether there is a fault or just alarms. Table 6.1 gives the thresholds associated with each component of the gearbox.

Figure 6.5 shows that all the components are working normally for the first 5 days and alarms occur frequently since day 6. By checking the alarm logs, it is found that both false alarms and true alarms were triggered accordingly. The alarm named "Gearbox oil sump high temperature warning" and "Gearbox oil sump stop" were triggered before the fault occurrence. The two alarms clearly indicated the gearbox malfunction and therefore considered as true alarms. Thus the proposed method can give early warning of the fault 5 days in advance. After day 10, the gearbox bearing temperature 2 exhibited the first fault behaviour which were then propagated to other gearbox components in the following days. In day 16, all the gearbox components presented temporally normal behaviour due to the decreased wind speed. With the fault severity increased, the fault index reached its maximum on day 18. Overall, the

fault lasted for ten days with a high fault index appearing on all components. Since all the faulty components related to temperature or oil pressure, it can be concluded that the fault occurred in the cooling system of the gearbox. The pressure sensors are usually installed at the end of the filter whereas the temperature sensors are installed in the oil sump to monitor the temperature of the lubricating oil. With the dysfunction of the cooling system, the heat cannot be dissipated actively, thus leading to further damage of the gearbox due to the fault.

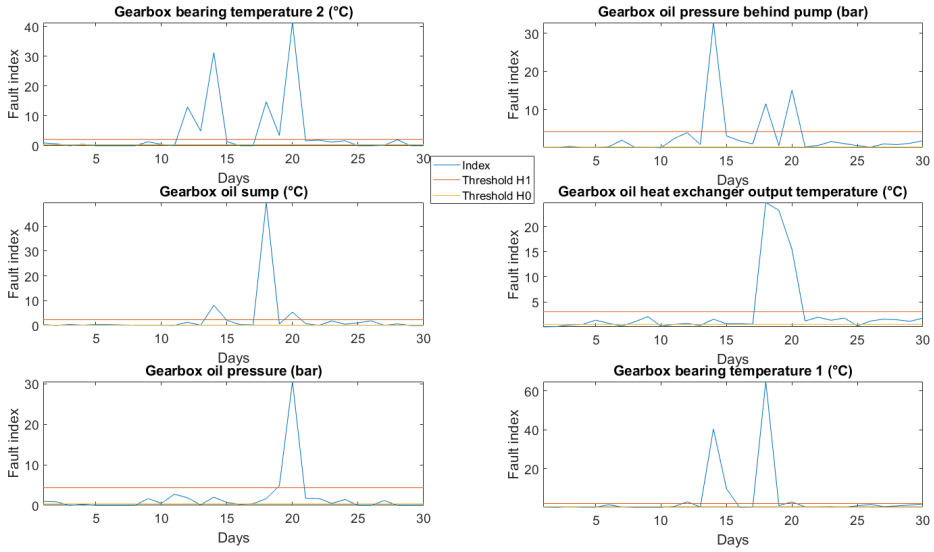


Figure 6.5 Fault index in terms of KLD of the gearbox variables

6.3.2 Case 2: Generator fault

The second subsystem to be studied is the generator that plays a key role in WT operation and produces electric power by converting the rotation mechanical power. The most common fault occurring in the generator is associated with the generator winding fault. This can be caused by phase-to-phase fault, phase to earth fault, or by additional thermal and mechanical stresses on the machine winding resulting from these electrical faults. The fault in the generator usually results in rising of generator temperatures. The generator under our study also has 6 monitoring variables, including generator stator top side temperature, generator stator bottom side temperature, generator bearing temperature phase A, generator bearing temperature phase B, generator cooling water advance temperature, and generator cooling water return temperature. Temperature sensors are embedded within the winding to monitor top and bottom side temperatures of the generator stator whereas bearing

temperature sensors are normally fitted to the bearing cartridge to monitor the temperatures of the generator bearings. Temperature sensors are also installed on the two ends of the cooling water system to monitor the temperatures of cooling water.

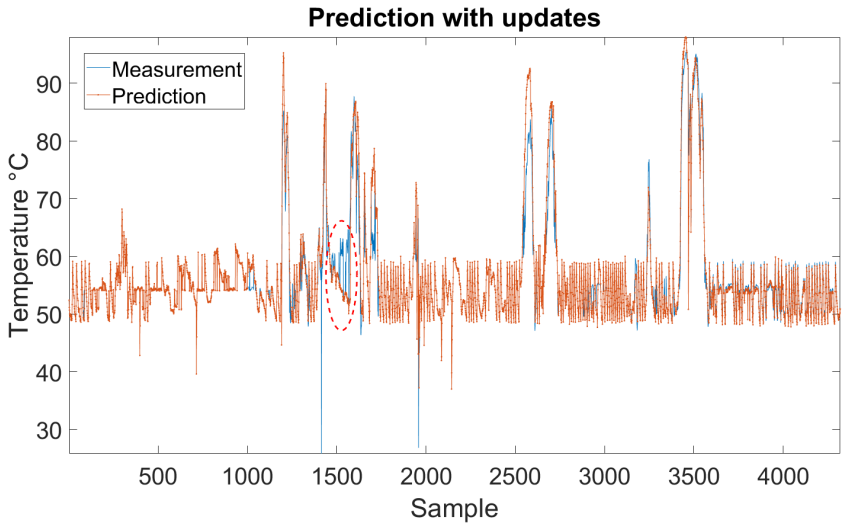


Figure 6.6 Prediction result during generator fault time

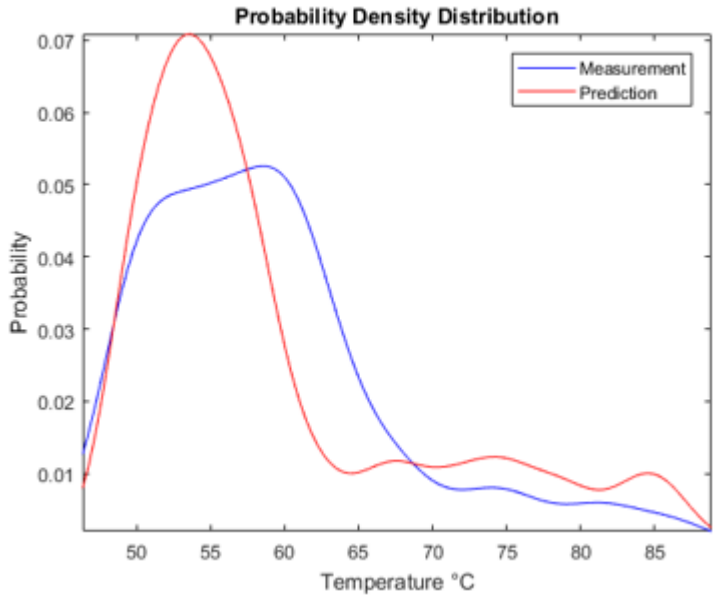


Figure 6.7 Probability density distributions of predicted and measured data for top side temperature of the generator stator

As an example, Figure 6.6 shows predicted and measured values of the generator stator top side temperature from 30 days, i.e., 4320 data samples, where the faulty time period of 3 days is circled in red. Clearly the measured temperature is different from the predicted one during this faulty time period, indicating that the component

experienced abnormal behaviour. As can be seen from the figure, the temperature is normally around 52° and there are several large spikes in the temperature. The occurrence of these temperature spikes are due to larger wind speeds and hence the higher power output during these time periods, as indicated from investigation of the SCADA data. The probability density distributions of measurement and prediction values are shown in Figure 6.7, where the KLD of these two distributions is 16.1747. Note that only 3-days data are shown in the Figure 6.7, focusing on the period of fault occurrence.

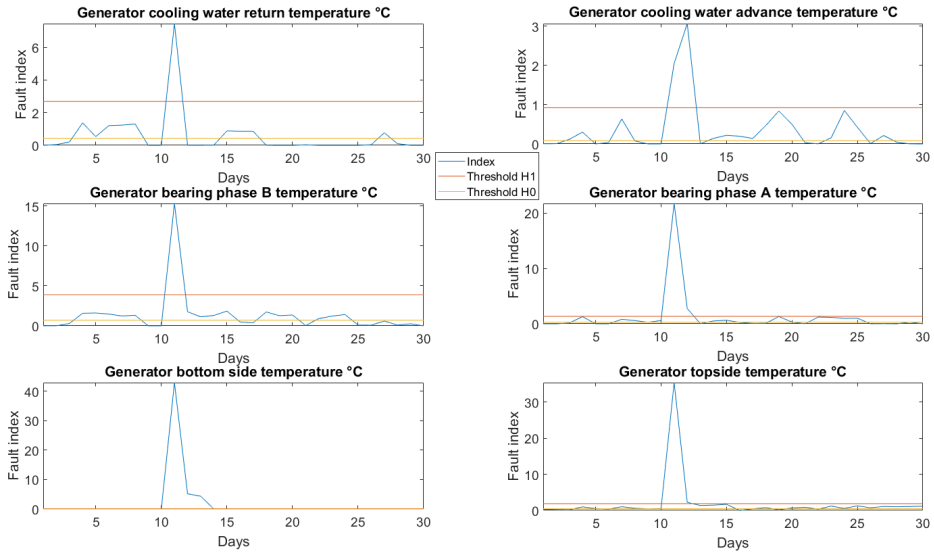


Figure 6.8 Fault index in terms of KLD of the generator variables

Table 6.2 Thresholds of generator components

Threshold	Generator topside temperature (°C)	Generator bottom side temperature (°C)	Generator bearing phase A temperature (°C)	Generator bearing phase B temperature (°C)	Generator cooling water advance temperature (°C)	Generator cooling water return temperature (°C)
H_0	0.4377	0.0088	0.2865	0.7004	0.0861	0.4322
H_1	1.3135	0.1199	1.3505	2.9033	0.9223	2.6808

Figure 6.8 demonstrates behaviour of these generator components during one month in terms of KLD values (i.e., fault index). Table.6.2 gives the thresholds associated with each component of the generator. For the first three days, the generator operated normally since all the KLD of the monitoring variables were below *H0*. Alarms were triggered after day 3. Having checked SCADA alarm logs, it is found that the "Machine interface high temperature warning" alarm was triggered several times

before the fault occurrence, which also confirm that the predicted alarms are correct. On day 10, all fault indices raised to a comprehensive high level, where the two highest indices represent generator stator top side temperature and generator stator bottom side temperature, respectively. This verifies that the fault happens on the generator winding, which lasted for three days before being fixed by maintenance and then returned to fault free condition.

6.4 Performance evaluation

It is clear that the proposed method can distinguish the normal, alarm and fault states of the operating turbines. To further test the accuracy and effectiveness of the hybrid method, dataset size will be extended to 100 days. The proposed LSTM-KLD will be compared with three other well-established machine-learning algorithms combined with KLD, which are SVM-KLD, CNN-KLD and DBN-KLD, respectively. The implementation of these hybrid methods are shown in Figure 6.9, where the steps of these hybrid methods are essentially same. Only the SVM, CNN and DBN are used to replace the LSTM to produce the prediction results, which are then incorporated with KLD to calculate the fault index.

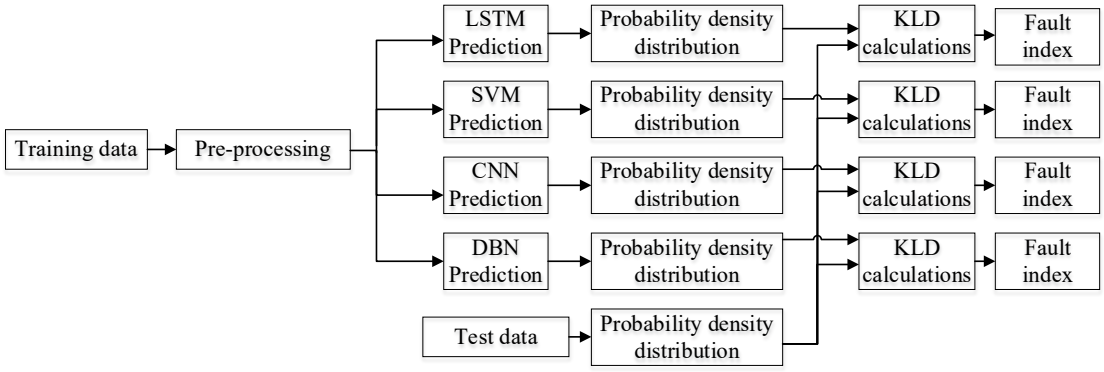


Figure 6.9 Schematic diagram of the hybrid method implementation

LSTM-KLD		Normal	Alarm	Fault
Actual Class	Normal	60 93.80%	4 6.20%	0 0.00%
	Alarm	0 0.00%	24 96.00%	1 4.00%
	Fault	1 9.10%	0 0.00%	10 90.90%
		Prediction class		

(a)

SVM-KLD		Normal	Alarm	Fault
Actual Class	Normal	58 90.60%	6 9.40%	0 0.00%
	Alarm	0 0.00%	22 88.00%	3 12.00%
	Fault	0 0.00%	1 9.10%	10 90.90%
		Prediction class		

(b)

CNN-KLD		Normal	Alarm	Fault
Actual Class	Normal	58 90.60%	6 9.40%	0 0.00%
	Alarm	2 8.00%	19 76.00%	4 16.00%
	Fault	0 0.00%	0 0.00%	11 100.00%
		Prediction class		

(c)

DBN-KLD		Normal	Alarm	Fault
Actual Class	Normal	28 43.80%	34 53.10%	2 3.10%
	Alarm	0 0.00%	16 64.00%	9 36.00%
	Fault	0 0.00%	0 0.00%	11 100.00%
		Prediction class		

(d)

Figure 6.10 Confusion matrices from four hybrid models for gearbox diagnosis (a) LSTM-KLD; (b) SVM-KLD; (c) CNN-KLD; (d) DBN-KLD

The performances of those four hybrid methods are evaluated by a confusion matrix to determine the accuracy rate when classifying normal, alarm and fault data. As is well known, the confusion matrix is also called as the error matrix, which is often used for solving classification issues [192]. It is essentially a visualised table showing the performance of a classification algorithm, where the rows of the confusion matrix show the prediction results while the columns show the actual data class [192]. For a confusion matrix, there are four main categories representing true positive, false positive, true negative and false negative. In this specific application, true positive indicates that the normal condition data are categorised correctly whereas false positive indicates that alarms and fault data are mistakenly considered as normal. True negative means the alarm and fault are classified precisely whereas the false negative shows how many normal data are mistakenly considered as alarm or fault data. By combing the true positive rate and true negative rate, the accuracy of the method is thus obtained. The decision of each class is made on a daily basis due to requirement of KLD calculation as mentioned previously. Hence, 100 decisions are made for confusion matrix evaluation.

Table 6.3 The overall accuracy of four hybrid methods for gearbox fault diagnosis

Accuracy	LSTM-KLD	SVM-KLD	CNN-KLD	DBN-KLD
TRUE	94%	90%	88%	55%
FALSE	6%	10%	12%	45%

Figure 6.10 and Table 6.3 shows confusion matrices and the overall accuracy of these hybrid methods for gearbox fault diagnosis, respectively. The true positive rate can be obtained by using the correct predictions divided by total samples available in the same row. For example, the true positive rate of the proposed LSTM-KLD can be calculated as 60/ (60+4) = 93.8% whereas the true negative rate for alarm and fault are 96% and 90.9%, respectively. Clearly, the LSTM-KLD outperforms three other hybrid methods.

Likewise, Figure 6.11 and Table 6.4 show confusion matrices and the overall accuracy of four hybrid methods for generator winding fault diagnosis, respectively. Again, the LSTM-KLD outperforms three other hybrid methods.

LSTM-KLD		Normal	Alarm	Fault
Actual Class	Normal	64 90.10%	6 8.50%	1 1.40%
	Alarm	0 0.00%	25 96.20%	1 3.80%
	Fault	0 0.00%	0 0.00%	3 100.00%
		Prediction class		

(a)

SVM-KLD		Normal	Alarm	Fault
Actual Class	Normal	62 87.30%	8 11.30%	1 1.40%
	Alarm	0 0.00%	25 96.20%	1 3.80%
	Fault	0 0.00%	0 0.00%	3 100.00%
		Prediction class		

(b)

CNN-KLD		Normal	Alarm	Fault
Actual Class	Normal	64 90.10%	7 9.90%	0 0.00%
	Alarm	0 0.00%	23 88.50%	3 11.50%
	Fault	0 0.00%	0 0.00%	3 100.00%
		Prediction class		

(c)

DBN-KLD		Normal	Alarm	Fault
Actual Class	Normal	40 56.30%	23 32.40%	8 11.30%
	Alarm	0 0.00%	18 69.20%	8 30.80%
	Fault	0 0.00%	0 0.00%	3 100.00%
		Prediction class		

(d)

Figure 6.11 Confusion matrices from four hybrid models for generator winding diagnosis (a) LSTM-KLD; (b) SVM-KLD; (c) CNN-KLD; (d) DBN-KLD

The computation time of the four hybrid methods is also compared to evaluate the efficiency of the proposed method. The programmes are run on MATLAB R2018a under the computer with Intel R Core ™ i7- 6820 HK CPU @ 2.70GHz and 16.0 GB RAM. The GPU acceleration was disabled during the training process. The computation process includes training process and judgement process. The training processing is completed by the ML algorithms whereas the judgement process is completed by KLD. The four methods are run 10 times each on both faulty turbines and the average time is listed in Table 6.5. As can be seen from the table, the SVM-KLD consumes the shortest computation time since SVM relies only on the kernel function. The other three methods are essentially deep-learning algorithms and possess a much more complex structure in comparison to the SVM. Thus, those three methods consume a longer computation time, among which the LSTM-KLD clearly consumes shortest computation time.

Table 6.4 The overall accuracy of four hybrid methods for generator winding fault diagnosis

Accuracy	LSTM-KLD	SVM-KLD	CNN-KLD	DBN-KLD
TRUE	92%	90%	90%	61%
FALSE	8%	10%	10%	39%

Table 6.5 Average computation time required from four hybrid methods for fault diagnosis

Computation time	LSTM-KLD	SVM-KLD	CNN-KLD	DBN-KLD
Training (s)	167.536	32.101	306.808	279.027
Judgement (s)	21.683	12.994	21.148	27.495

6.5 Summary

This paper presents a novel WTCM approach by a hybrid method of LSTM and KLD. The effectiveness of the proposed method has been evaluated and validated by SCADA data acquired from an operational wind farm. Two case studies are carried out to detect gearbox bearing fault and generator winding fault respectively from the turbines. In order to improve the reliability of diagnostic results, multiple monitoring variables relating to the specific WT subsystem are analysed. The probability density distributions of the measured data and the predicted data are compared and calculated in order to bring a first intuitive impression of the health condition of the turbine subsystem. The KLD values are calculated and used as a fault index to quantify the fault severity of the turbines. By cross-checking the

predicted alarms and SCADA alarm logs, an early warning of the fault can be detected 5 days in advance. In order to further evaluate the performance and effectiveness, the proposed LSTM-KLD is compared with three other ML algorithms combined with KLD and the results prove superiority of the LSTM-KLD method.

The results demonstrate that the proposed data-driven model-based approach is accurate and sensitive for fault detection. It is capable of identifying different faults occurring in the different WT subsystems. With the proposed method, normal, alarm, fault conditions can be clearly distinguished, thus enhancing detection robustness. The results also show that the LSTM-KLD requires a shortest computation time in comparison to other deep learning approaches. Furthermore, by adopting the proposed method, the false alarm rate from the SCADA system is reduced, thus improving the diagnosis confidence. Future work will focus on the datasets with higher sampling rate to reveal fault dynamics and hence the fault mechanism. This method can also be extended to detect early faults of other WT subsystems and components.

Chapter 7. Wind turbine power conversion test rig design and experiments

This chapter presents the design of WT power conversion test rig. The test rig is used to simulate various operation conditions including different kinds of faults. The experiments procedures and results are also presented in this chapter. The data obtained from the test rig with high sampling frequency at 1 kHz are then used to validate the condition monitoring algorithms and models proposed in the previous chapters.

7.1 Description of the power conversion test rig

Because the electrical system has the highest failure rate among all subsystems of the WT, it is necessary to understand the mechanism of the WT power conversion unit [193]. A power conversion test rig was built to simulate the dynamic operation of a WT under different operating conditions. Measurement data were thus acquired and used for analysis. Compared to SCADA data acquired from an operational wind farm, the measurement data had a much higher sampling frequency. In addition, the data could also simulate various types of faults. This is significant for condition monitoring because the transient behaviour that occurs during the faults may be lost when the sampling rate is low. Existing SCADA data contain only limited types of faults, which is not ideal for fault diagnosis modelling. Compared with the MATLAB simulation results, the experimental results were more realistic because the computer simulation was based on the mathematical model of a WT, which may not be suitable for practical conditions. For example, the virtual grounding in the simulation has completely different behaviour with experiment test rig. Therefore, it is necessary to design and construct a WT power conversion unit.

Owing to safety concerns, a miniaturised PMSG based power conversion unit was designed for the experiments. The overall topology and the major component list of the power conversion unit are shown in Figure 7.1 and Table 7.1, respectively. The test rig comprises six major modules: a phase conversion module, rectifier module, DC-link module, inverter module, load module, and signal conditioning module. The phase conversion module is connected to a signal generator to model the rotor-side power generation. The converted three-phase AC is then converted to DC by passing it through the rectifier module. The DC-link module is used to eliminate ripples from the converted DC and to store energy in the capacitors. The inverter module is used to convert DC into three-phase AC with the required phase and frequency. Finally, the three-phase resistive load is connected to deplete the electrical energy. All monitoring signals pass through the signal conditioning module to be converted into a certain voltage level that satisfies the data acquisition system. The control and data acquisition of the test rig were performed by the host computer. The computer communicated with a stand-alone field programmable gate array (FPGA) control system and a data acquisition card. Details of these modules are discussed in the following sections.

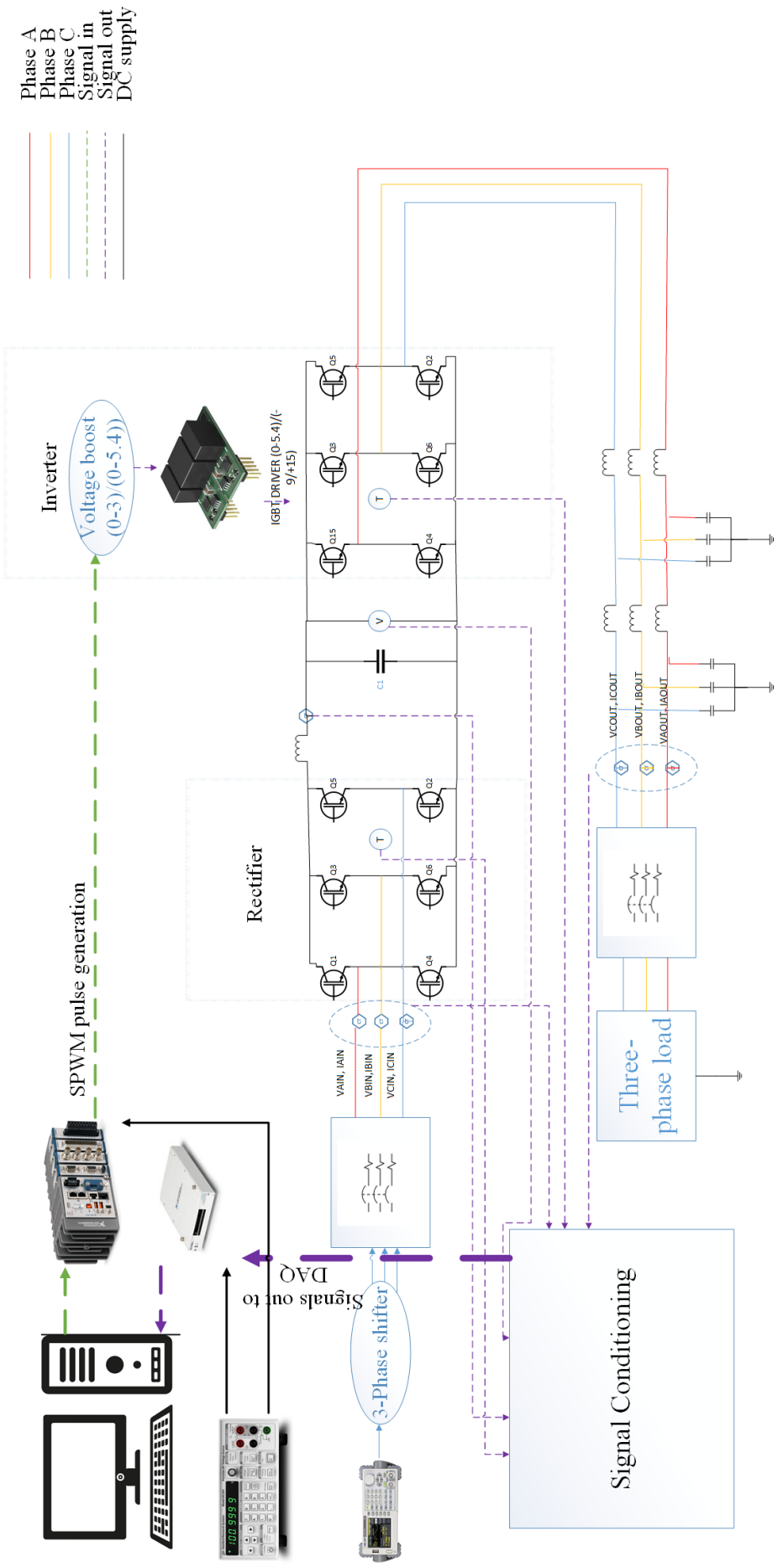


Figure 7.1 Main components of the power conversion test rig

Table 7.1 Main components of the power conversion test rig

Main functionality	Components
Rotor side simulation	Signal generator Phase converter
Rectifier module	IGBT driver IGBT module
DC-link	DC filter DC-link capacitor
Inverter module	Voltage booster IGBT driver IGBT module
Load module	LC filter Passive resistive load bank
Measurements and signal interface	Current Transducers Signal conditioning modules National instrument data acquisition card
Controller	National instrument Compact Rio system Host computer
Miscellaneous	DC power supplies, circuit breakers and fuses

7.1.1 Rotor-side modelling and power generation

Owing to the limitations of the experiment, modelling of the mechanical rotation could not be implemented. Hence, a signal generator with a phase-shifter module was adapted to model the rotation of the rotor.

Figure 7.2 Signal generator

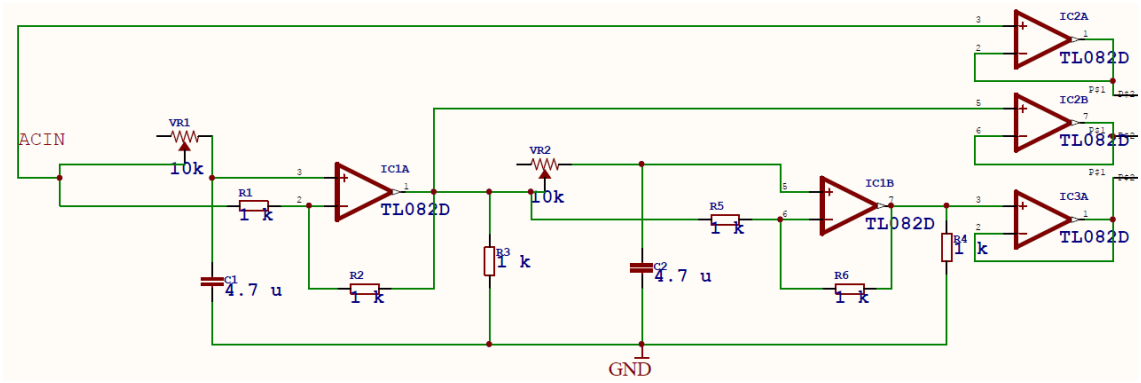


Figure 7.3 Phase shifter circuit

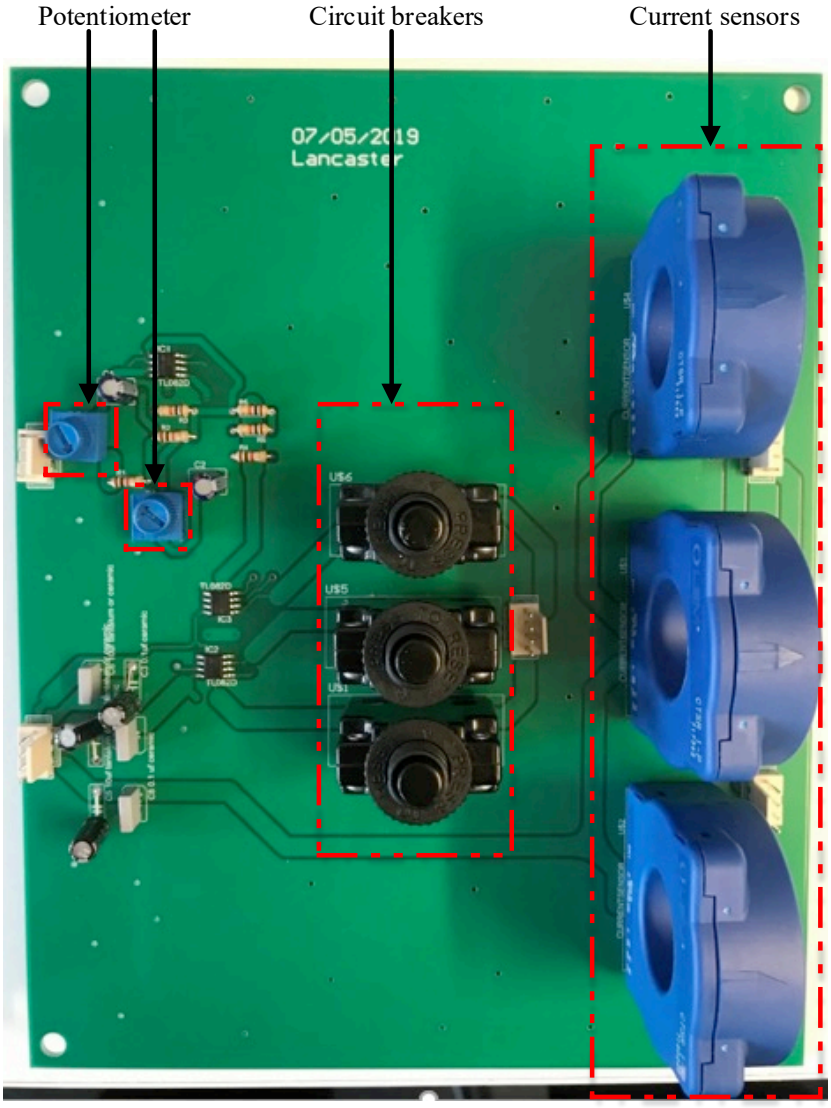


Figure 7.4 Phase shifter module

The Thandar TG501 function generator shown in Figure 7.2 can generate signals at frequencies ranging from 0.005 Hz to 5 MHz with high quality at all frequencies levels. The sine, square, triangle, ramp, pulse, and haverwave signals can be

generated with a variable start/stop phase. The generated voltage level ranges from 20 mV to 20 V peak-to-peak from 50 W, plus the transistor−transistor logic output. In the designed test bench, the function generator is used to generate a 50 Hz sinusoidal wave and is then connected to the phase shifter module, which can be seen in Figure 7.4. The phase shifter transforms a single-phase sine wave into three-phase sine waves to simulate the behaviour of the rotor. Figure 7.3 shows the mechanism of the phase shifter. The phases and amplitudes of the three-phase sine waves can be varied using two potentiometers. Three circuit breakers were installed for circuit protection against faults.

7.1.2 Rectifier module

A rectifier module was used to convert the three-phase AC into DC. It consists of two main parts: an IGBT module and IGBT drivers.

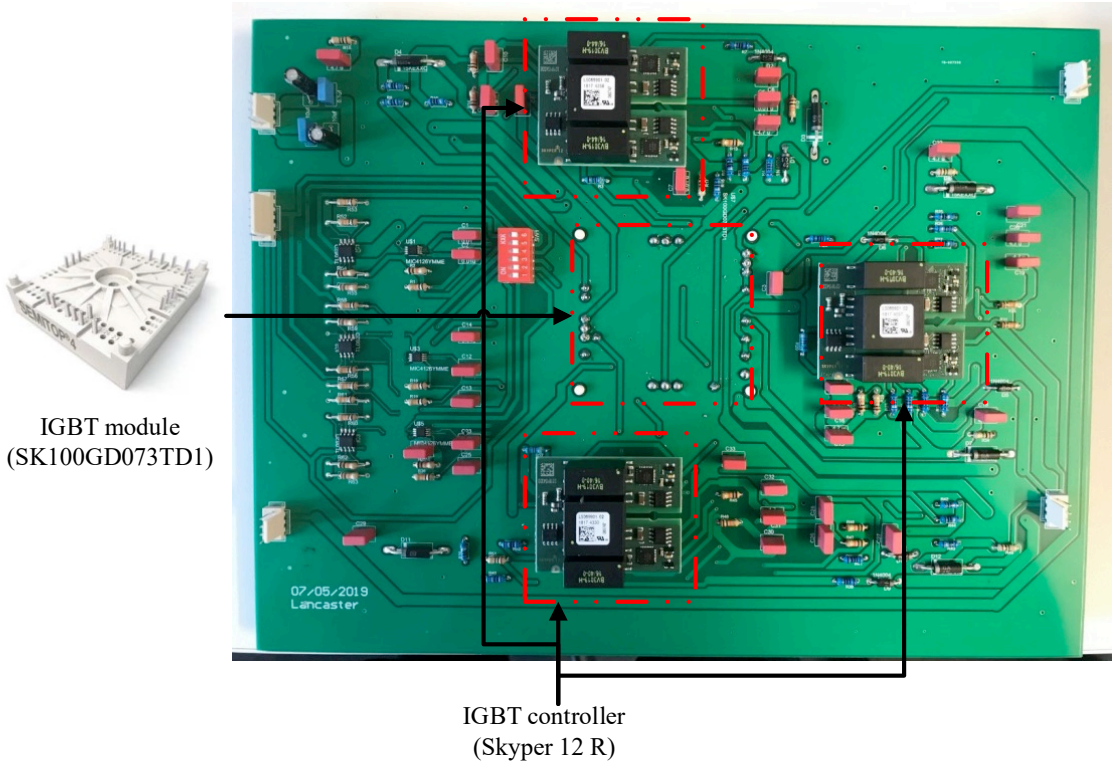


Figure 7.5 Rectifier module

The core of the rectifier module is the IGBT module and its controller, as shown in Figure 7.5. The IGBT module consists of six IGBT gates, and each controller controls the shutdown states of two IGBT gates (one phase). The IGBT controllers are used as voltage booster producing maximum 25 $V_{p\text{-}p}$ and 50 mA output current

to drive the IGBT modules. For cooling requirements, the IGBT module was installed on the back of the board to provide sufficient space for the heat sink and cooling fan. Figure 7.6 shows the topology of the IGBT module.

The IGBT module used in this study was SK100GD073TD1 from Semikron with V_{CES} 650 V and maximum power rating of 60.45 kW. As can be observed from the figure, each IGBT was connected in parallel with a diode. It has a rapid switching 650 V diode technology that can satisfy experimental requirements. It can also handle a forward current of 105 A at 25 °C and 38.9 A at 150 °C. During the component selection process, all the parameters were set much higher than the experimental requirements to prevent potential damage, such as overflow current to the components. In addition to the electrical limitation, the temperature of the IGBT could be extremely high when operating under rated conditions. Hence, a heat sink and a cooling fan were required for the IGBT module.

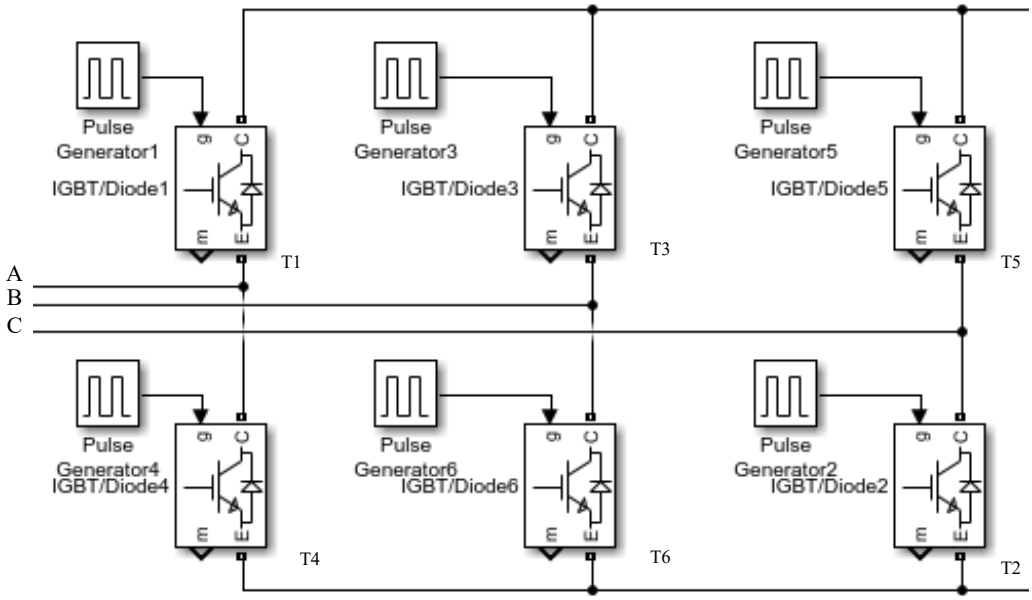


Figure 7.6 Rectifier structure

In the rectifier, all IGBTs were set as open. In other words, the IGBTs were disabled, and diodes were used to rectify the three-phase sine wave. For the further experiment improvement, the IGBTs are used instead of the diode bridges. As shown in Figure. 7.7, the conduction sequences of the diodes were D1D2, D2D3, D3D4, D4D5, D5D6, and D6D1, respectively, and the conduction angle for each diode was $2\pi/3$.

Ideally, the average value of the output can be found using:

$$V_{dc} = \frac{6}{2\pi}\int_{\pi/3}^{2\pi/3}\sqrt{3}\,V sin\theta d\theta \tag{7-1}$$

where V represents the input three phase voltage and V_{dc} represents the DC-link voltage. Or it be expressed as:

$$V_{dc} = V\frac{3\sqrt{3}}{\pi} = 1.654V \tag{7-2}$$

In addition, the current in each phase can be found using:

$$I_{phase} = I\sqrt{\frac{2}{\pi}\left(\frac{\pi}{6}+\frac{\sqrt{3}}{4}\right)} = 0.78I \tag{7-3}$$

where I is the three phase line current. The RMS current through each diode is given by:

$$I_{diode} = I\sqrt{\frac{1}{\pi}\left(\frac{\pi}{6}+\frac{\sqrt{3}}{4}\right)} = 0.552I \tag{7-4}$$

where $I = 1.73V/R$.

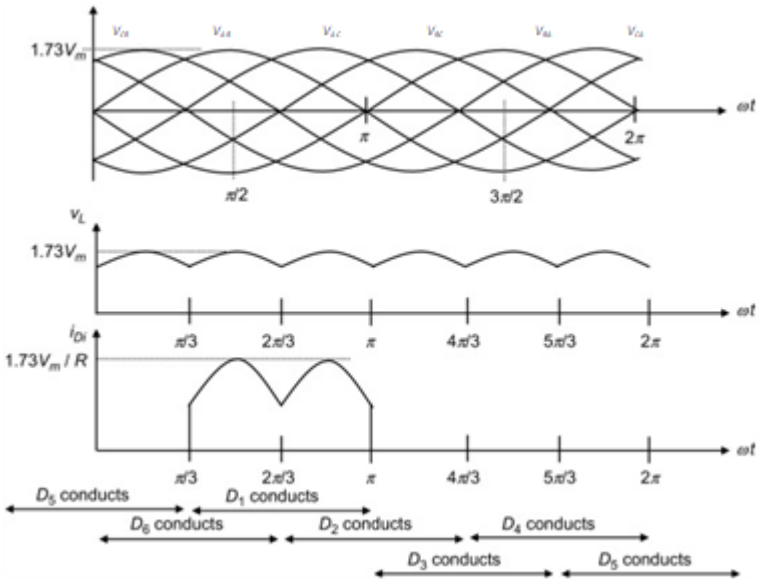


Figure 7.7 Voltage and current waveforms of the three-phase full bridge rectifier [186]

7.1.3 Capacitor bank

The DC-link capacitor bank used in the test bench was used to eliminate ripples in the rectified DC voltage, as shown in Figure 7.8. To simulate different conditions of

the DC-link module, such as open circuit, short circuit, and capacitor fatigue, a total of eight capacitors were combined to achieve different capacitance values.

Figure 7.8 Capacitor bank module

In this design, two 3,900 μF, two 2,700 μF, two 1,000 μF, and two 470 μF, with 100 V rated voltage capacitors were used. Different capacitance combinations can be achieved by connecting the capacitors in parallel. According to the design, the maximum capacitance can reach $(3{,}900 + 2{,}700 + 1{,}000 + 470) \times 2 = 16{,}140$ μF when all the capacitors are connected in the circuit. The minimum capacitance is 470 μF, and it results when only one 470 μF capacitor is connected. The capacitor bank can be disabled by connecting none of the capacitors to the circuit. A 5 W, 25 Ω resistor was also connected in parallel with the capacitors to limit the inrush current flow. The structure of the capacitor bank is shown in Figure 7.9.

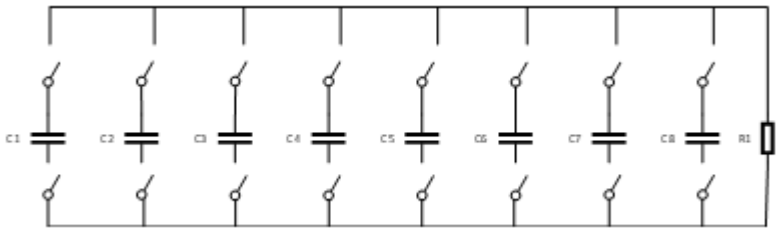


Figure 7.9 Capacitor bank structure

7.1.4 Inverter module

Similar to the rectifier module, the inverter module consists mainly of one IGBT module and three IGBT drivers, as shown in Figure 7.10. A three-phase full-bridge inverter is used to convert the DC voltage into three-phase AC voltage.

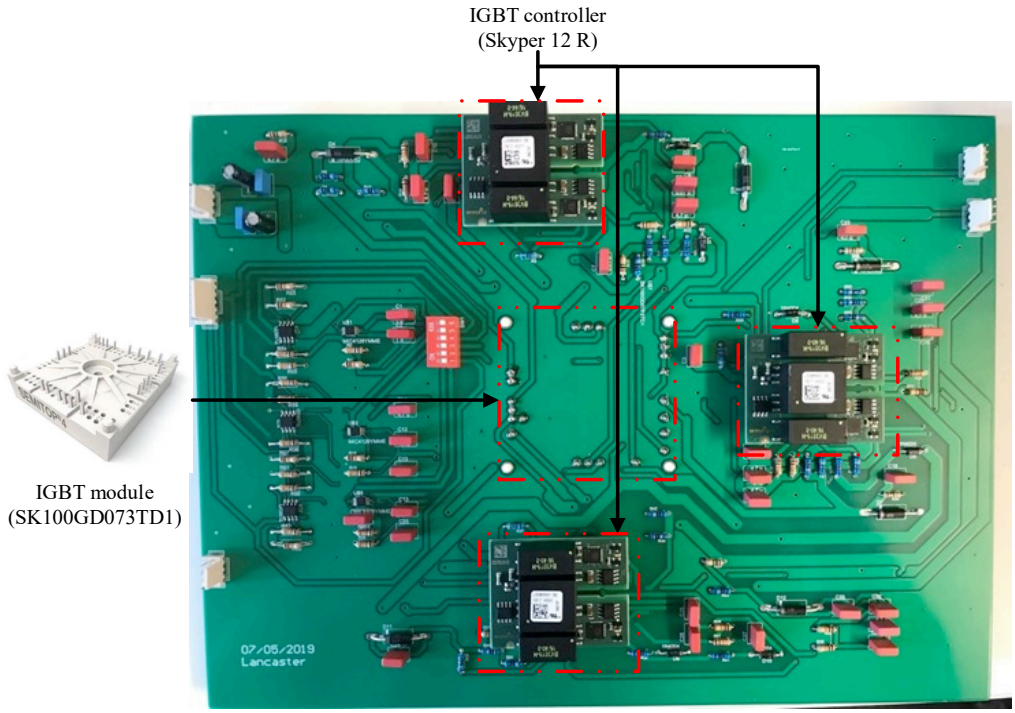


Figure 7.10 Inverter module

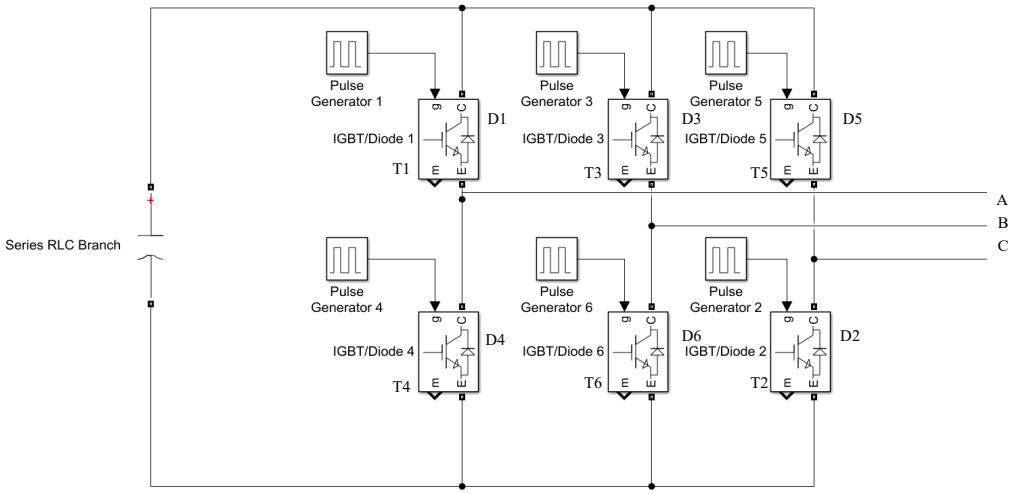


Figure 7.11 Inverter structure

Figure 7.11 demonstrates the structure of the inverter. Sinusoidal pulse-width modulation (SPWM) is employed to convert DC to three-phase AC. The sinusoidal AC reference wave V_{ref} was compared with a high-frequency triangular carrier wave V_c to determine the switching states for each IGBT. The switching states for each IGBT were determined using two rules [194]:

- $V_{ref} > V_c$, upper arm turned on. ($V_{upper}=V_{dc}/2$)
- $V_{ref} > V_c$, lower arm turned on. ($V_{lower}=-V_{dc}/2$)

Here, the peak-to-peak voltage is given as the DC-link voltage V_{dc}. The SPWM technique for one phase is illustrated in Figure 7.12.

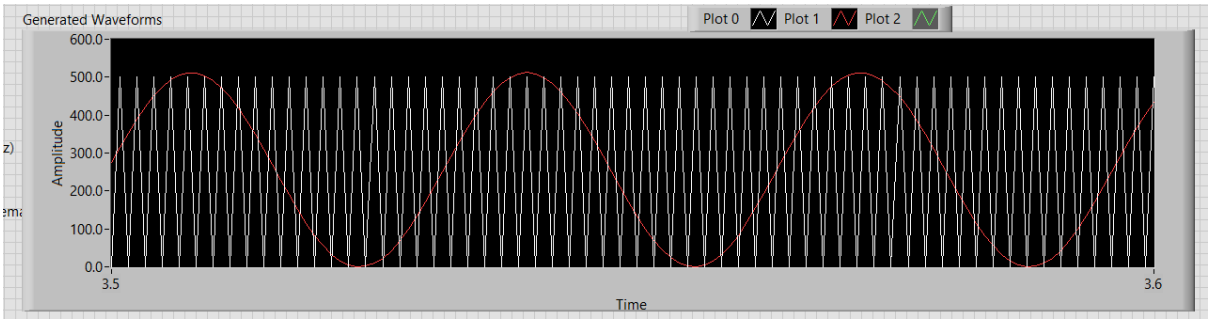


Figure 7.12 SPWM generation for one phase

The ratio of the wave amplitude to the triangular carrier wave amplitude is defined as the modulation index (m_a).

$$m_a = \frac{V_{ref}}{V_c} \tag{7-5}$$

The ratio of the wave frequency to the triangular carrier wave frequency is defined as the frequency modulation index (m_f).

$$m_f = \frac{f_{ref}}{f_c} \tag{7-6}$$

The fundamental component of the output V_o changes linearly with the reference voltage V_{ref} if the carrier frequency is sufficiently large compared to the fundamental frequency and can be expressed as:

$$V_o = V_{ref} sin\omega t \tag{7-7}$$

In terms of the modulation index, Equation (7-7) can be rewritten as:

$$V_o = \frac{V_{dc}}{2} m_a sin\omega t \tag{7-8}$$

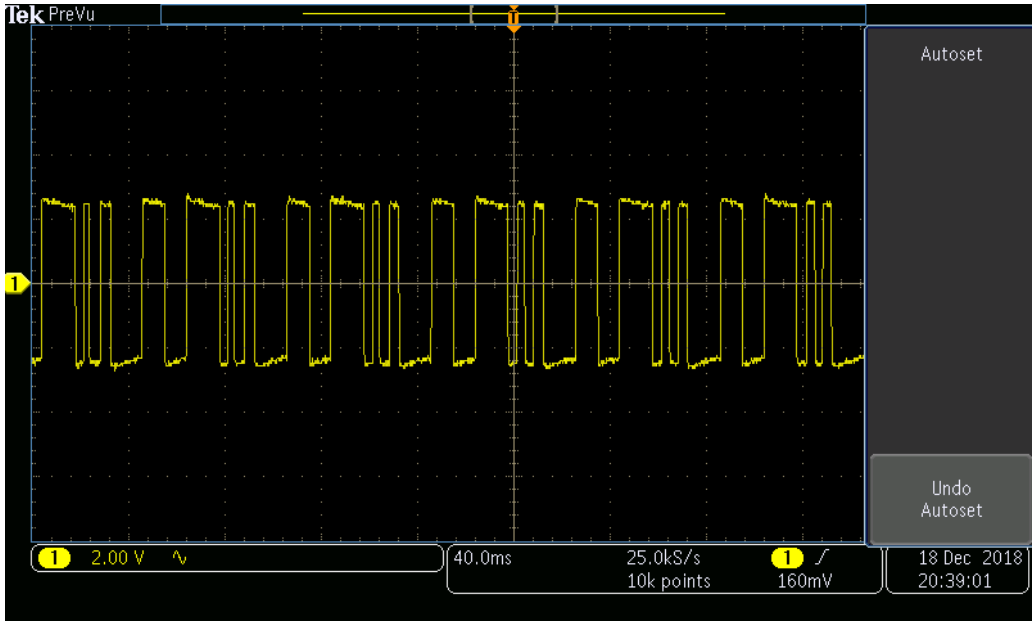


Figure 7.13 SPWM signal for one IGBT gate

Figure 7.13 shows an example of the SPWM modulation result for one IGBT gate when f_{ref}=50, f_c=3,000, and m_a=1.

7.1.5 Resistive load module

Figure 7.14 shows the resistive load module for the designed test bench, which dissipates the generated power.

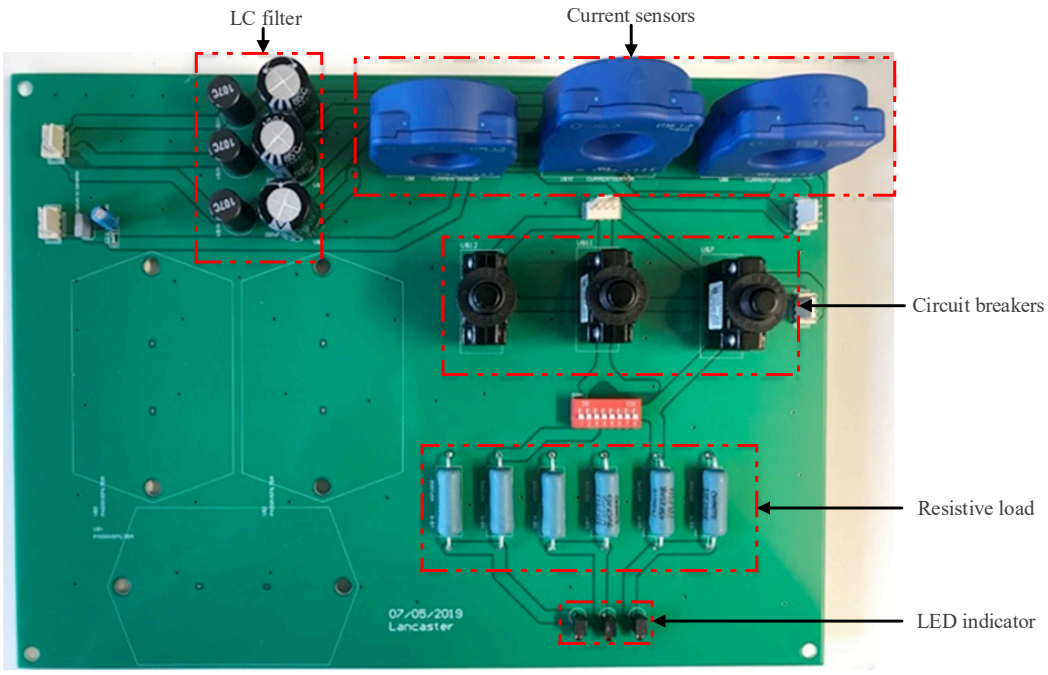


Figure 7.14 Resistive load module

On the input side of the module, a 10 mH inductor and 220 μF capacitor were used to build an LC low-pass filter to remove the harmonics and smooth the waveforms. The output of the LC filter is given as:

$$V_{out} = \frac{1}{1-\omega^2 LC} V_{in} \tag{7-9}$$

In addition, the cut-off frequency for the designed LC filter is:

$$f_{cut-off} = \frac{1}{2\pi\sqrt{LC}} \tag{7-10}$$

In the designed LC filter, the cut-off frequency is 107.3 Hz, which is 2−3 times the fundamental frequency. Thus, the harmonics can be effectively removed.

Six 5 W, 25 Ω power resistors were used to model the function as a load. Each of the two resistors was connected in parallel in one phase to perform normal, open-circuit, and short-circuit fault analysis.

7.1.6 Signal conditioning module

Various sensors were installed on the test bench to collect data for monitoring the conditions. The AC voltage, AC current, DC voltage, DC current, and temperature of the IGBT modules were monitored on the test bench. However, as the acquired

signals may be too large for a data-acquisition system, the signal conditioning module is required to adjust the amplitude of the signals to an acceptable range for data acquisition.

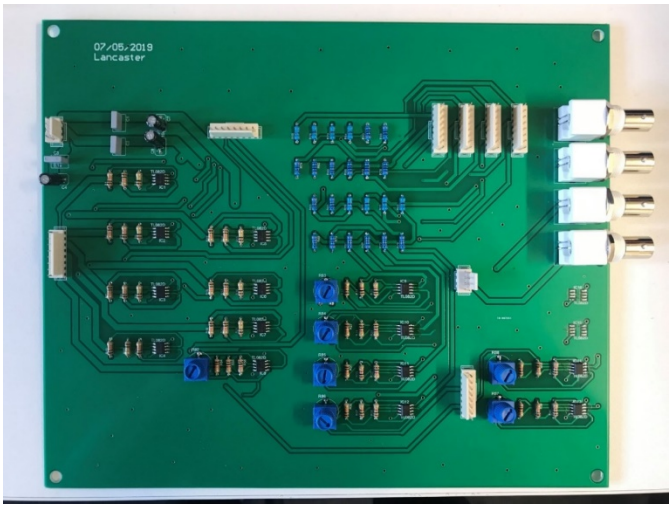


Figure 7.15 Signal conditioning module

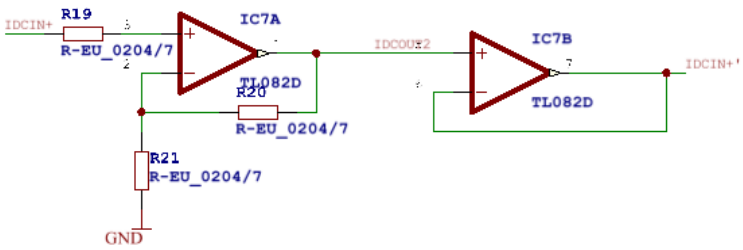


Figure 7.16 Example of signal conditioning circuit

Figure 7.15 shows the signal conditioning module, and Figure 7.16 illustrates an example of a signal conditioning circuit. Non-inverting amplifiers are connected to vary the signal amplitude without changing the phase. A voltage follower is connected between the non-inverting amplifier and the data acquisition system to eliminate loading effects while maintaining the same voltage from the former side.

7.1.7 Control and data acquisition system

To design the control system of the test bench, National Instrument CompactRIO (NI CRio) hardware, NI C series modules, and LabVIEW 2014a were chosen for real-time signal generation. The control signals to the rectifier were set as high at all times, and SPWM was employed as the gate control of the inverter. Thus,

compared with the former SPWM wave generator, a new signal generator has been designed based on direct digital synthesis (DDS).

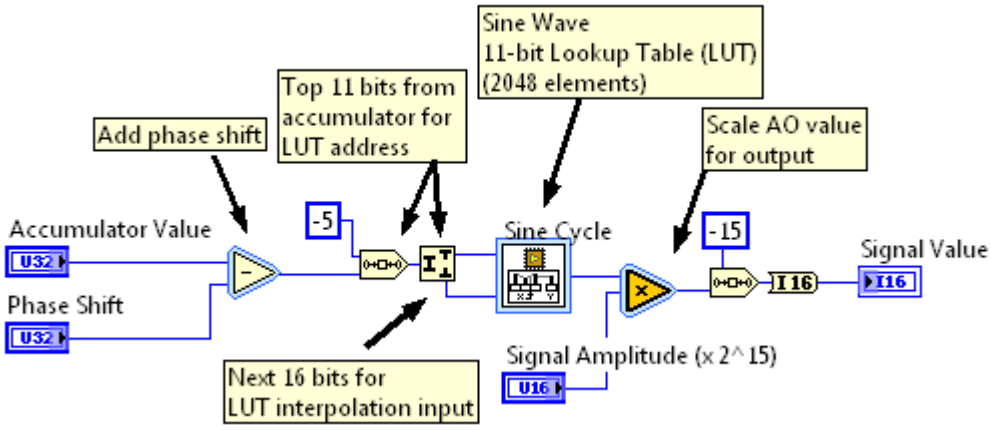


Figure 7.17 Waveform lookup function for DDS generation [195]

The control signal-producing algorithm consists of two parts: an accumulator and a waveform lookup table. The accumulator is a 32-bit counter that increases the current phase value by a specified increment. It also embeds a function that can synchronise multiple DDS generators. The waveform lookup portion of the DDS generator uses the current accumulator phase value to return the current waveform value from the reference waveform lookup table. In this design, a 2,048-sample reference waveform representing one cycle was implemented. Figure 7.17 demonstrates an example of a lookup function for DDS generation.

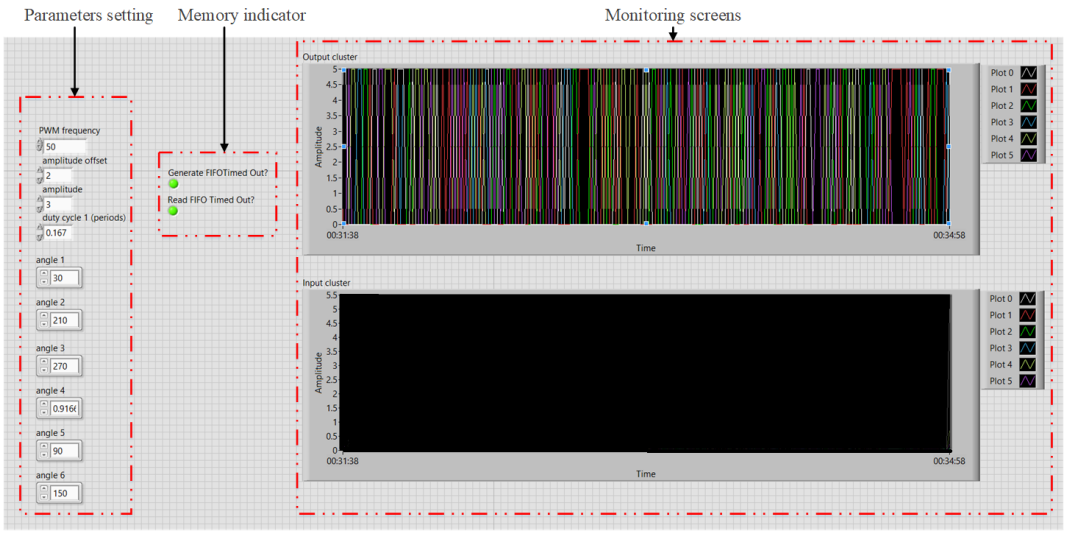


Figure 7.18 Control system interface

Figure 7.18 shows the interface of the control system. As shown in the figure, the interface has three sections: parameter setting, memory indicators, and monitoring screens. Parameter settings include the frequency, offset, amplitude, conducting angles, and duty cycle of the SPWM control gate control signals. The two memory indicators show the communication status of the first in first out method between the FPGA and the computer. The screens monitor the control signals sent from the FPGA. As an example, the output cluster on the monitoring screen shows the control signals sent to the inverter.

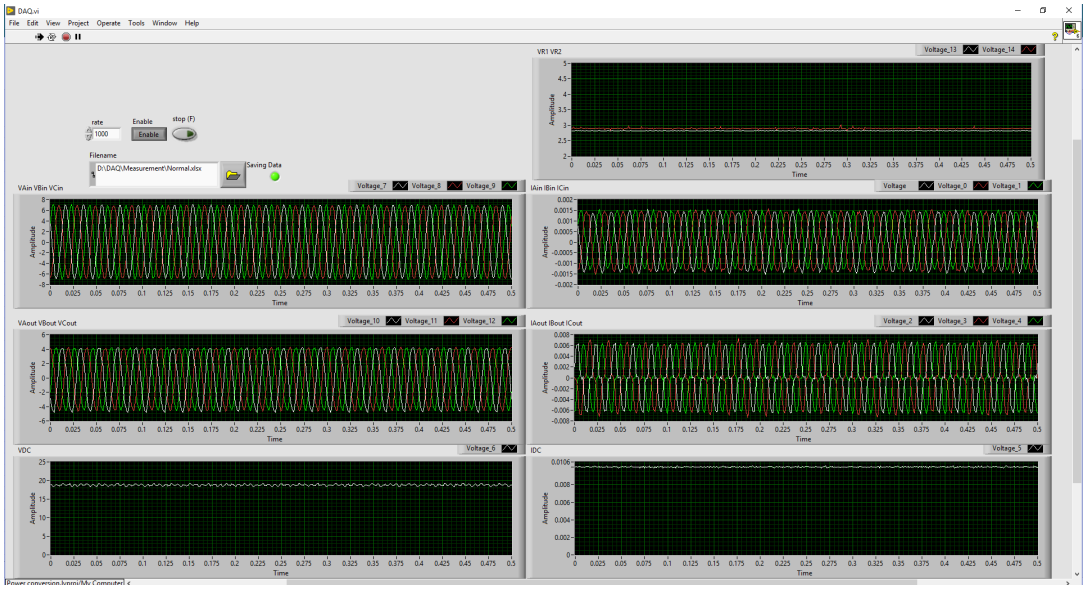


Figure 7.19 Data acquisition system interface

The NI USB-6229 data acquisition module was used for data acquisition and transmission with a graphical user interface on a desktop PC. The data acquisition system can simultaneously monitor signals and save data. In addition, the sampling frequency can be manually selected by users. A total of 16 variables were monitored, including six AC voltages, six AC currents, one DC voltage, one DC current, and two temperature signals, as shown in Figure 7.19. The overall hardware system layout of the WT power conversion test bench is shown in Figure 7.20.

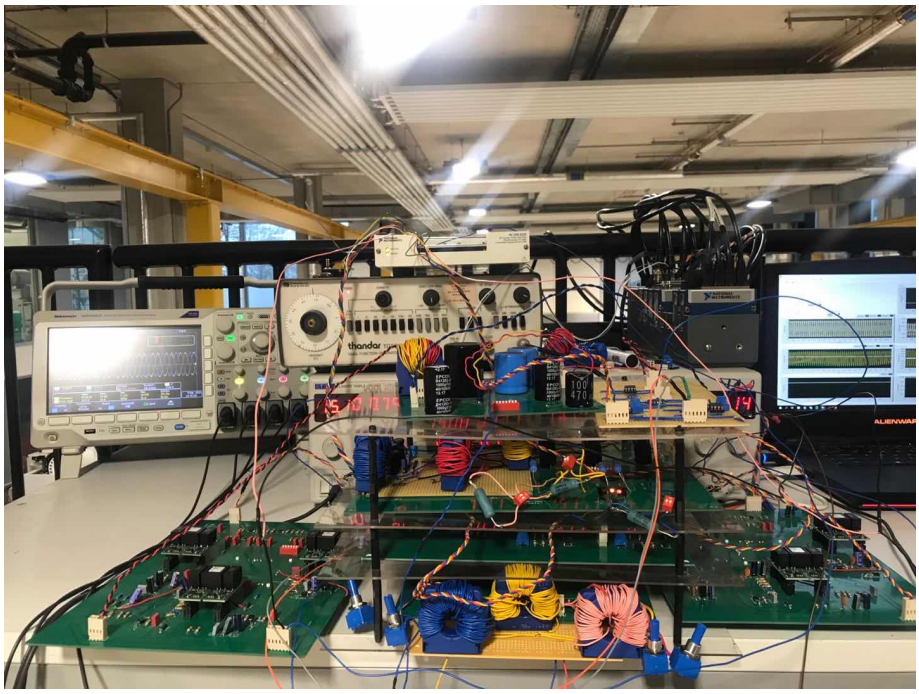


Figure 7.20 Hardware system layout of WT power conversion test rig

7.2 Test bench experiments

7.2.1 Experiment arrangement

The aim of the test rig was to simulate multiple faults in a WT power conversion unit. As mentioned in the previous section, 16 variables were monitored. The locations of the sensors are shown in Figure 7.1. For safety purposes, the voltage levels in the experiments were reduced. Constant power was generated from the waveform generator and fed to the power conversion test rig. To control the SPWM inverter, the switching frequency and modulation index were set to 2 kHz and 0.8, respectively. The inverted AC output was then dissipated by a 5 Ω with 25 W resistive load per phase. First, normal operating conditions of the test rig were simulated. Figure 7.21 shows the normal operation condition after capacitor charging of 16 variables with a sampling rate of 1,000 samples/s. As shown in the figure, the amplitude of the three-phase AC input voltage is 5.5 V_{pp}. The inverter AC output voltage is approximately 4 $V_{pp.}$ The voltage in the DC-link module is approximately 21 V. Under normal working conditions, the IGBT temperature was maintained below 50 °C. Owing to the limitation of the input power, the input and output current flows in the system are of the order of several mA.

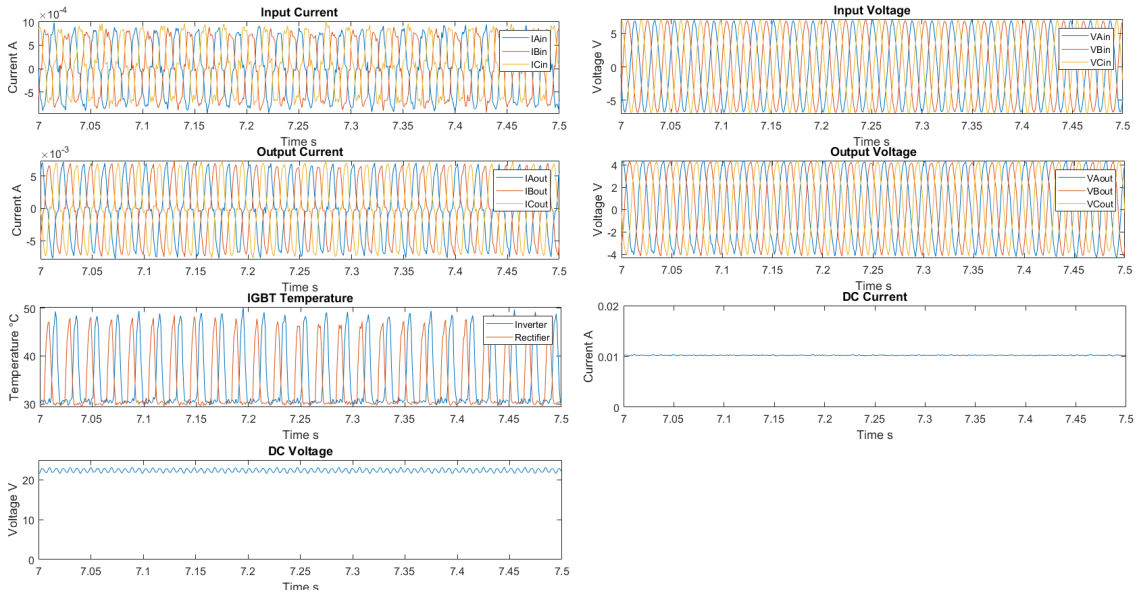


Figure 7.21 Normal operation condition of test rig

In addition to the simulation of the normal operation condition, open-circuit fault, phase-to-ground short-circuit fault, capacitor ageing, stator over/under synchronous speed, imbalanced load, inverter phase misalignment, inverter control signal update error, inverter control signal under-sampled, and inverter control signal over-sampled were also simulated. Of these nine types of faults, the open-circuit fault, short-circuit fault, capacitor aging, stator under/over synchronous speed, and imbalanced load were simulated using hardware, and the rest of the faults were simulated using software.

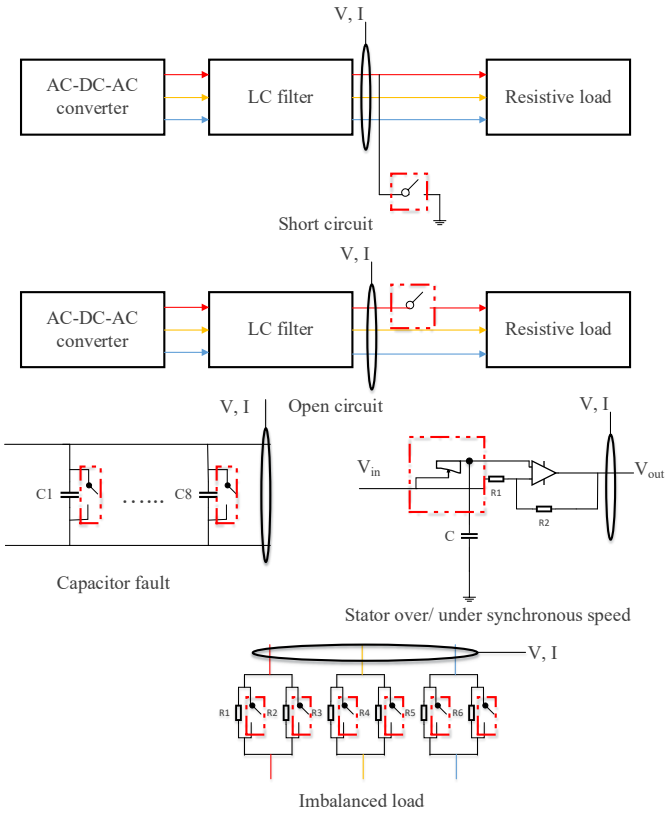


Figure 7.22 Block diagrams for simulation of five types of hardware faults on the power conversion test rig

Figure 7.22 shows the five types of hardware generating methods, where all the switching positions are circled in red and the measurement locations are also specified. A phase-to-ground short-circuit fault is generated by connecting the target phase directly to the ground to create a short circuit. The phase voltages and currents were measured after LC filtering, as shown in the figure. An open-circuit fault is generated by opening the connection between the filtered AC and resistive load. The capacitor breakdown fault can be generated by shortening all capacitors in the DC-link module. Two potentiometers installed on the phase converter module were used to model the behaviour of the stator over/under synchronous speed. By rotating these potentiometers, the input voltage amplitude and phase could be misaligned. For an imbalanced load fault generation, each phase has two power resistors connected in parallel. An imbalanced load fault is modelled by changing the connection status of the resistors. Other faults including inverter phase misalignment, inverter control signal update error, inverter control signal under-sampled, and inverter control signal over-sampled are generated by the control system of the test rig. The parameters in the control system interface were changed separately to model the different types of software faults.

7.2.2 Data acquisition

In the experiments with the WT power conversion unit test rig, nine different faults were simulated. As an example, Figure 7.23 shows the simulation results of the open-circuit fault. To demonstrate the amplitudes and observe small changes of electrical parameters clearer, the figures show the sinusoidal form instead of the RMS form.

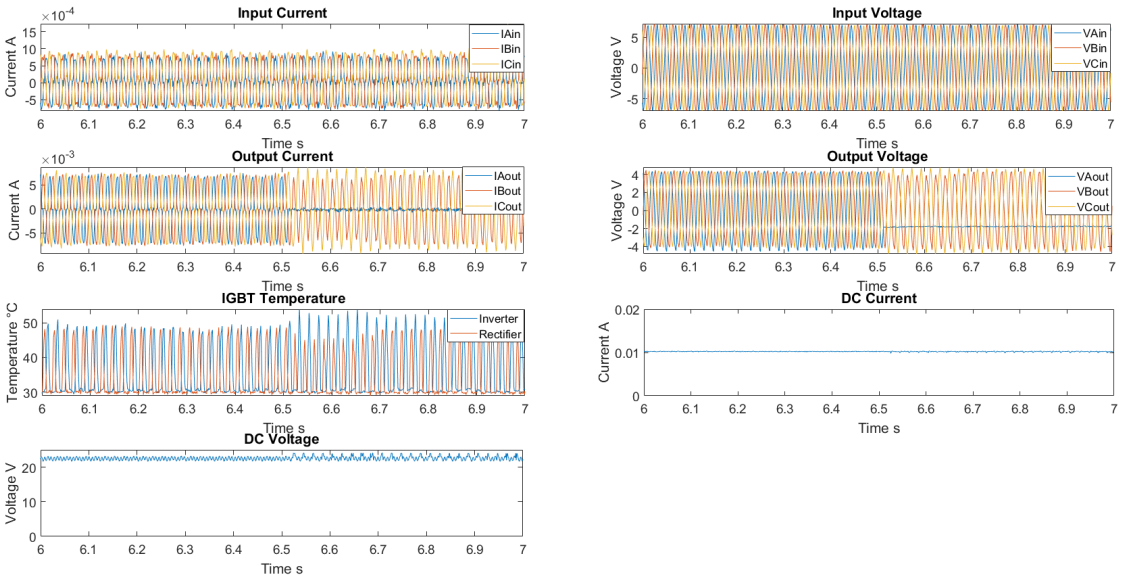


Figure 7.23 Simulation results of open-circuit fault

It can be observed from the figure that the fault is triggered at 6.5 s. Because the fault is triggered on the inverter side, there is no interference with the input side. The output current of phase A instantly reduces to zero and causes turbulence in the other two phases. The on and off voltages of the switching signals for one bridge of the IGBT are approximately +15 V and -8 V, respectively, which causes different voltage references between the IGBT and the ground. When phase A is an open-circuit, its reference voltage is no longer the same as that of the ground. This causes the output voltage of phase A to reach −2 V instead of 0 V. When the fault is triggered, the inverter-side temperature of the IGBT is higher than the temperature of the rectifier. When phase A is an open-circuit, the LC filter connected to phase A is also disabled, which results in more harmonics in the DC voltage. For the DC current, there appears to be no significant change after the fault is triggered.

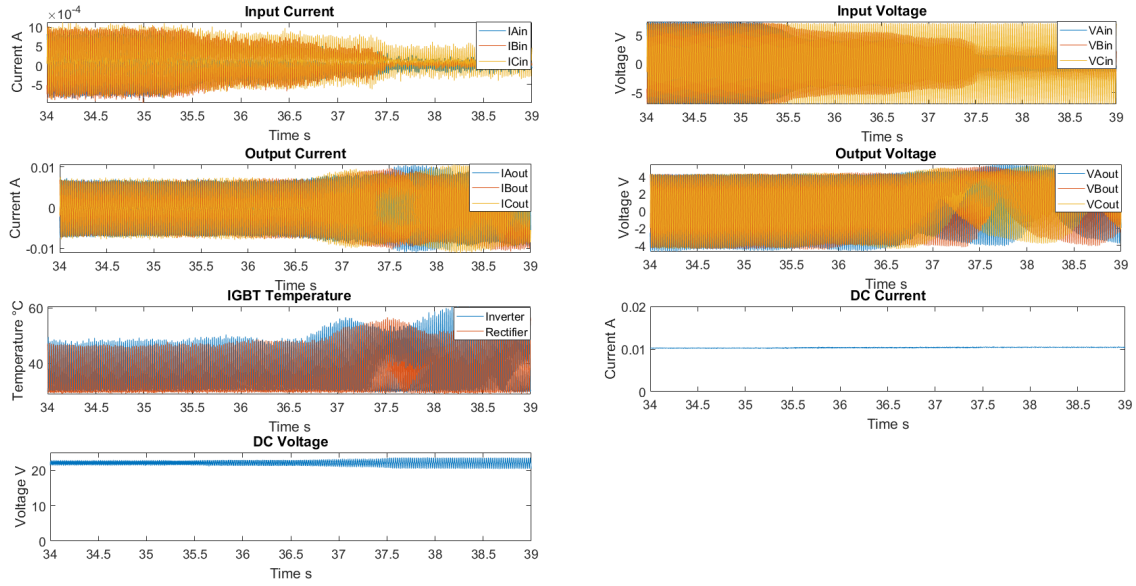


Figure 7.24 simulation results of stator over/ under synchronous speed

The different degrees of severity of the faults were also simulated. As an example, the stator over/under synchronous speeds with different degrees of severity are shown in Figure 7.24. At 34 s, all of the readings were normal. After 35.5 s, the stator over/under synchronous speed was generated manually. It can be observed from the figure that the input currents and voltages started to reduce when the fault was triggered. As the fault severity increased with time, the output-side voltages and currents were also affected. The temperatures of both the inverter and rectifier increased after 37 s. The changes in the output voltage exceeded the design limitation of the RC filter in the DC links. Hence, more ripples were added to the DC voltage.

7.3 Model validation

Several data-driven model-based WTCM methods are proposed in Chapters 4, 5, and 6. Based on the functionalities, the proposed models can be classified into two categories: fault classification and fault-severity estimation. The data acquired via the test bench were used for fault classification and fault severity estimation.

7.3.1 Normal-fault condition classification

Owing to data limitation, the previous chapter proposed methods focused on single-fault classification or extraction. The first type of data was obtained from an operational wind farm that could only provide a single type of fault on one turbine.

The other type of data is the software simulation data that can only reflect ideal operation conditions. Because the power conversion unit test rig was designed and built, multiple faults could be simulated in this way.

Based on the model proposed in Chapter 5, normal-fault classification can be realized using KSVM. In this section, imbalanced load and short-circuit faults are generated. A total of 16 monitoring variables were used for modelling, as described above. Two KSVM classification models were built to distinguish between the two types of faults. The model was first built to distinguish between imbalanced loads. It was trained using normal and imbalanced load fault data and was then validated for a longer time period. The results are shown in Figure 7.25, where the classified faulty data are labelled using red crosses. As can be observed from the figure, the fault occurs in two periods. Both faulty periods are clearly distinguishable in the figure.

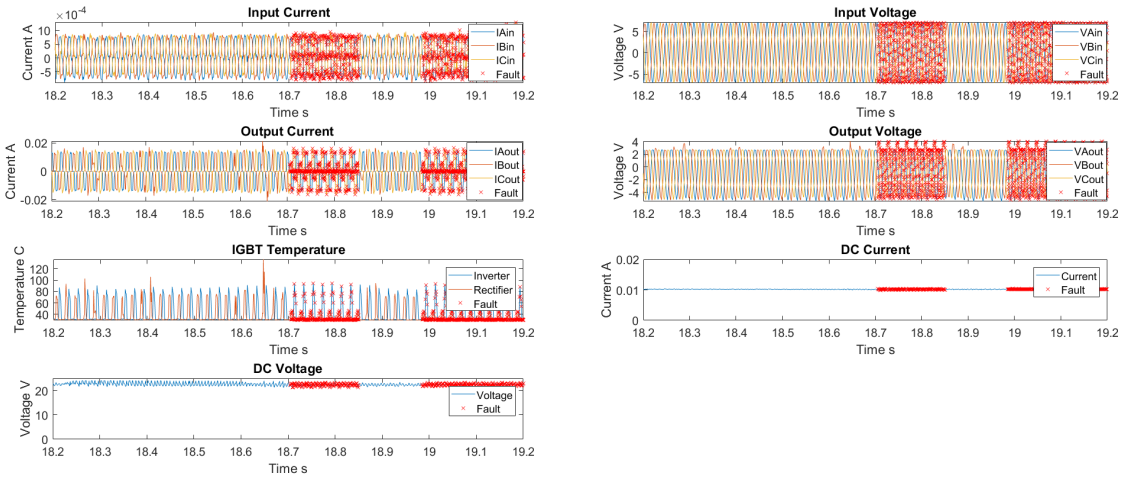


Figure 7.25 Imbalanced load fault classification results

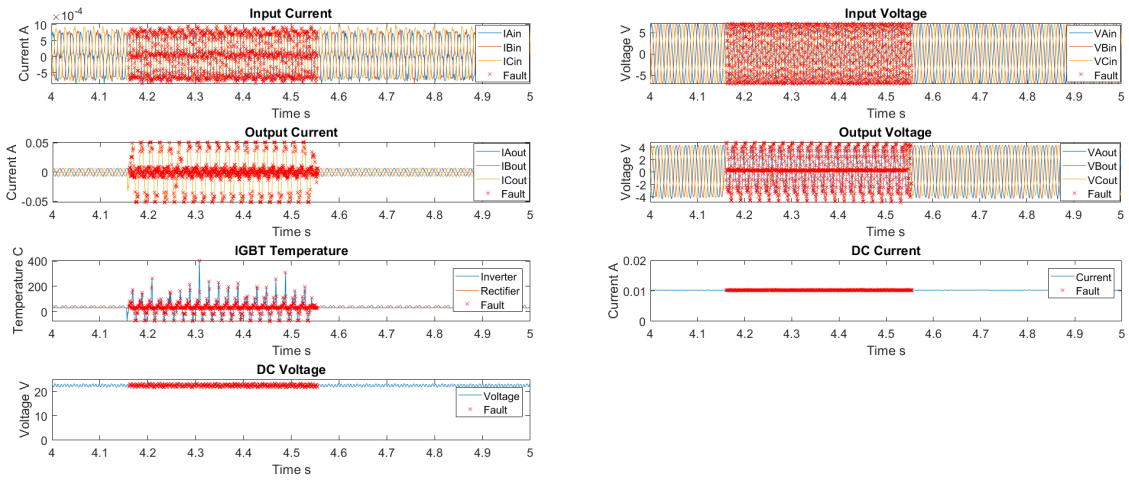
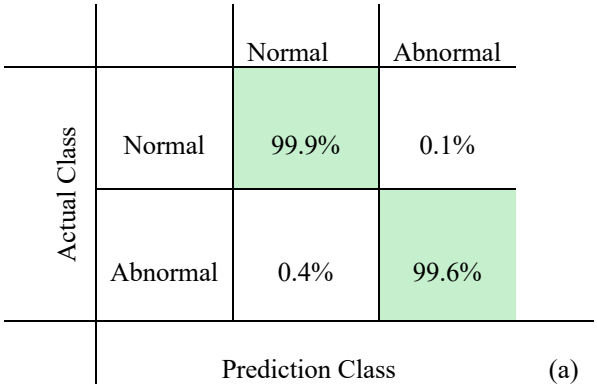


Figure 7.26 Short circuit fault classification results



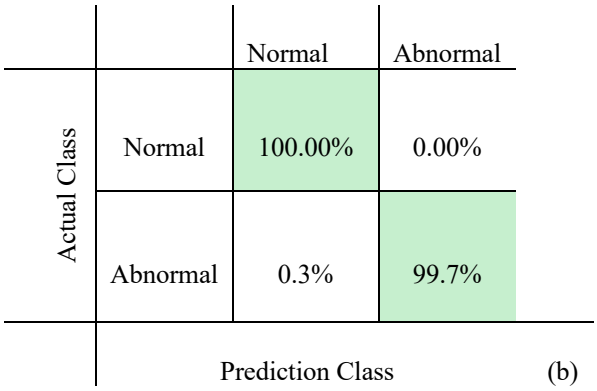


Figure 7.27 Confusion matrices of classification model for two faults: (a) imbalanced load; (b) short circuit

Similarly, the short-circuit fault classification model was trained and validated, and the results are shown in Figure 7.26. The red crosses in the figure indicate the faulty time period. It can be observed from the figure that the short-circuit fault generated

on the grid side has a significant effect on the output current, output voltage, and temperature of the inverter. The fault was clearly distinguished using the KSVM classification model. The accuracy of both the imbalanced load and short-circuit fault classification results were evaluated using confusion matrices, as demonstrated in Figure 7.27.

It can be observed from Figure 7.27 that both classification models have extremely high accuracy. The overall accuracies for the imbalanced load fault and short-circuit fault were 99.7% and 99.8 %, respectively.

7.3.2 Fault severity estimation

As described in Chapter 6, KLD can be used as a fault index to evaluate the fault severity. The fault severity estimation is based on the simulation stator over/under synchronous speed, as shown in Figure 7.24.

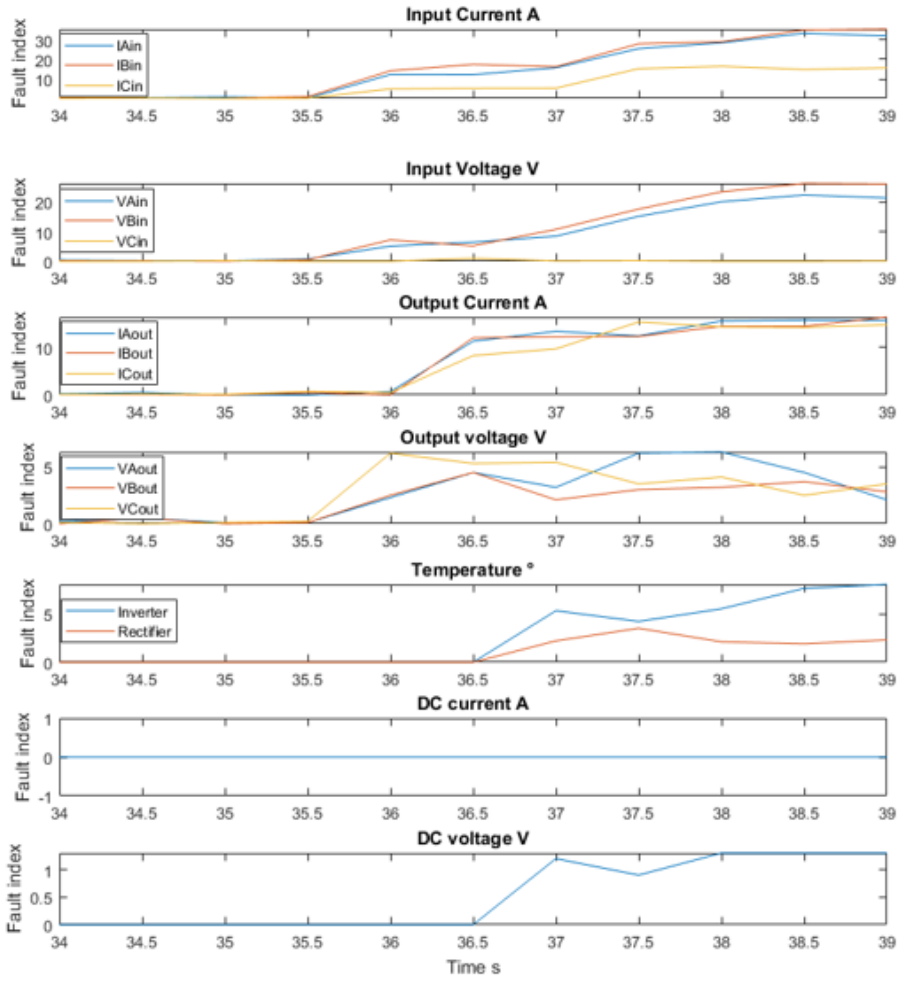


Figure 7.28 Fault index in terms of KLD of the test rig variables

The fault indices of the stator over/under synchronous speeds are shown in Figure 7.28. The KLD values are calculated every 0.5 s for all 16 variables. It can be observed from the figure that the fault was injected at 36.5 s on the input side. The input-side variables, such as the three-phase current and voltage, responded immediately when the fault occurred and became increasingly severe based on the fault index. However, it appears that the fault does not have a significant influence on the phase C input voltage. This is due to the fault generation mechanism, which is not affected by the phase C input voltage. It can also be observed from the figure that the inverter temperature suffers more than that of the rectifier.

7.4 Summary

The design and construction of the WT power conversion unit test rig were described in this chapter, and the selection of hardware based on their characteristics was discussed in detail. The design of the control and data acquisition systems was also described. The main aim of the test rig was to simulate different types of operating conditions that might occur. The open-circuit fault, short-circuit fault, capacitor ageing fault, stator over/under synchronous speed, imbalanced load, inverter phase misalignment, inverter control signal update error, inverter control signal under-sampled, and inverter control signal over-sampled were simulated. The experimental procedures and results were also provided. The next chapter provides a summary of the research achievements and contributions. Potential research developments are also discussed.

Chapter 8. Conclusions and future improvements

This chapter summarises achievements from the research. It also describes how the aims and objectives listed in chapter one are fulfilled. In addition, the contributions of knowledge in the line of research are explained, followed by the limitations of the research and prospects to the future work.

8.1 Summary of current research achievements

As a major type of renewable energy, the use of wind energy has become more and more popular in the past decades especially for offshore wind energy. The location of WTs are usually in the remote areas or offshore installations, thus increasing the O&M costs. In order to reduce the O&M costs, the condition based maintenance strategies have been carried out, which requires high reliability condition monitoring techniques. However, there are still challenges existing in current CMS. The redundant monitoring variables and large amount of monitoring data would increase not only the CMS complexity but also the computation load, thus compromising the CMS reliability. Hence, developing an effective condition monitoring method that can achieve a dedicated and rapid fault diagnosis and prognosis while maintaining the sufficient information is necessary. Besides, the information about the alarm signals contained in the SCADA system is often neglected.

In this research, the KSVM based fault feature extraction model is firstly proposed to extract particular fault features during the WT operation. The model is tested and validated against both SCADA data from an operational wind farm and simulation data. A statistical tool based on KLD is adopted for monitoring variable selection and fault localisation. Only the variables with highest KLD values are selected for further fault detection

To further discover the relationship among monitoring variables and find out how much information is contained in the variables, the KPCA incorporating with KSVM are considered. Compared with PCA, the KPCA has the capability to solve non-linear mathematic issues, which is more suitable for monitoring variables in the WTs. The KPCA is used for variable selection to reduce the information redundancy and thus improve the calculation efficiency. To select appropriate monitoring variables, the accumulated variance contribution is set to 85%. Furthermore, the use of alarm signals in the SCADA data is also taken into consideration, which can provide vital information relating to the fault. The 2-step KSVM model is used to distinguish the normal, alarm and fault condition of the turbine. The normal and abnormal data are firstly classified and the abnormal data are then further classified as alarm and fault data. Thus, the three operation states of the turbine are distinguished and the relationship between alarm and fault data are obtained. These variable selection and operation condition classification methods

are not only tested with historical SCADA data but also evaluated with simulation data with short circuit fault and capacitor breakdown fault.

In addition, the fault localisation and severity estimation are performed by incorporating LSTM with KLD. Multiple monitoring variables that contribute to the specific subsystems are taken into consideration to improve the reliability of fault diagnosis. Besides, by cross-checking alarm logs, the predicted alarms from the algorithm can be used as a significant evidence to support diagnostic results as early warning of the fault. Specifically, the LSTM is used to achieve the behaviour prediction of the key subsystems such as gearbox and generator. Then the KLD is employed to detect the fault by comparing probability distributions of the variables over the time between predicted data and test data. Consequentially, the monitoring data can be classified as normal, fault, true alarm and false alarm while the severity of the fault is also evaluated.

Furthermore, the research also involves design and construction of a PMSG based WT power conversion test rig. The representative faults emulated in the experiments include open circuit fault, short circuit fault, capacitor aging fault, stator over/ under synchronous speed, imbalanced load, inverter phase misalignment, inverter control signal update error, and inverter control signal under-sampled and over-sampled faults. A sampling rate of 1 kHz is sufficient to capture the transient change of the fault, as compared with low sampling rate SCADA data that are usually sampled in the interval of minutes. The data acquired from the test rig are also used for further evaluation of the proposed methods for fault classification and fault severity estimation. All the results demonstrate the effectiveness of the proposed condition monitoring algorithms and models.

The research objectives presented in the previous Section 1.3 and their corresponding main achievements are summarised in the table below.

Table 8.1 Objective achievements

Objectives	Achievements
Conduct critical literature reviews of current WT technologies, including the fundamental working principle, common failure modes, condition monitoring technologies and maintenance strategies.	The overview of current wind energy technologies, including WT common structure, configuration, working principle and failure modes, are described in Sections 2.1 and 2.2. Sections 2.3 and 2.4 review the state-of-the-art condition monitoring technologies. The modern maintenance strategies are also reviewed in Section 2.5.
Model and simulate WT systems with different configurations under different operation scenarios such as normal condition and fault condition.	The simulation models of both DFIG and PMSG are built based on MATLAB/Simulink. The models are built based on a real GE 1.5 MW WT. The model details including control strategies are presented in Section 3.2. Besides, different representative faults are simulated in order to study the dynamic behaviour of WTs under different fault scenarios.
Understand the physical relationships among monitoring variables not only in the simulation data but also in the real measurement data.	Section 2.1 introduces the WT structure in details, including physical relationships among different components. Section 3.1 is focused on the SCADA data, where impact of the faults on the monitoring data is also described.
Develop an appropriate methodology that can distinguish and separate the faults from the healthy conditions effectively.	A statistical tool based on KLD incorporating with KSVM is used to classify the representative faults (bearing fault and generator winding fault in this research) from the operating WTs, which is presented in Chapter 4.
Develop effective models to distinguish the normal, alarm and fault conditions of the WTs whilst reducing the computation load by optimising the variable selection for dedicated condition monitoring.	The kernel function is used in PCA to solve the nonlinear problem in the condition monitoring of WTs. The KPCA is further used to select fault related variables based on accumulative variance contribution. The alarm signals

	are introduced to provide more information of the fault. The proposed model is tested and validated with both SCADA data and simulation data, which are shown in Section 5.3.
Develop novel methods to estimate fault severity index that can measure the severity of abnormal conditions of the WT.	To locate and estimate the abnormality of the turbines, a hybrid LSTM-KLD method is proposed in Chapter 6. LSTM is adopted to capture relation features in temporal dependencies among monitoring data in an iterative manner, thus improving the prediction capability. The KLD is used as fault indicator, which measures the severity of the fault by comparing probability distributions between predicted data and test data. By adopting a COST function based on normal value probability and alarm value probability of the calculated KLD values, the optimised thresholds are determined to distinguish the normal, alarm and fault conditions.
Design and develop an operational WT power conversion test rig to further experimentally validate the proposed condition monitoring algorithms and models.	The test rig is described in Section 7.1, mainly including designs of hardware modules, control system and data acquisition system. The proposed condition monitoring algorithms and models are further validated by experimental data in Section 7.3.

8.2 Contributions in the line of research

The contributions in the line of research are summarised as follows:

1. In the study, the statistical tool based on KLD is applied to investigate the difference of probability density distribution in different operation conditions for fault localisation. Besides, the KLD can also be used for variable selection in order to reduce the dataset size for an effective condition monitoring. By

combining with the COST function, the fault severity of the WT components can be estimated.

2. The kernel function is further applied in both SVM and PCA. The employment of the kernel function can solve non-linear mathematic issues. The KSVM is able to project the original dataset into hyper-dimension and convert the linearly inseparable problems into linearly separable problems. The different operating conditions of the WTs can be separated in hyper-dimension by the hyperplanes. The confusion matrix is adopted for estimation of the classification accuracy.
3. To further quantify the effect of dimension reduction in fault classification, KPCA is applied. The KPCA is aimed to maximise dataset variability and maintain the maximum information entropy. The accumulative variation contribution is used to estimate the performance of variable selection and dimension reduction.
4. The LSTM is investigated for WT behaviour prediction. The study of alarm signals was dug further to correlate the relationship with the targeted particular fault. The distinguished alarms are then cross-referenced with SCADA alarm logs in order to provoke early warning of the fault. The proposed approach is validated by two faulty WTs with gearbox bearing fault and generator winding fault. By comparing the results with other well-established methods, the effectiveness of the proposed method is demonstrated in terms of detection accuracy and time complexity.
5. A PMSG based WT power conversion test rig is designed and constructed to emulate the operational behaviours of the turbine under various type of faults in order to collect sufficient experimental data to validate the proposed algorithms. Imbalanced load and short circuit faults are emulated, which have been clearly distinguished by the proposed algorithms. Furthermore, the fault severity of the stator over/ under synchronous speed is also estimated.

8.3 Prospects to the future work

The future improvements to study the proposed WTCMs are summarised below:

1. The KLD used for variable selection and fault localisation has been verified with different faults with both SCADA data and simulation data. In simple cases, zero of relative entropy indicates that the two distributions have

identical quantities of information. However, this is a distribution-wise asymmetric measure, which does not qualify as a statistical metric of spread. In other words, the KLD does not satisfy the triangle inequality. To solve this problem, the Wasserstein metric can be used to compare the difference between two distributions since Wasserstein metric does not neglect the geometrical features of two distributions and thus leading to more precise results.

2. Compared with ordinary PCA method, the KPCA provides a non-linear solution to the linearly inseparable issues. KPCA is a nonlinear PCA by generalising the kernel method into linear PCA. However, the high generalisation performance of KPCA is obtained at the cost of a large amount of computation time. The future work can also focus on exploring ways of reducing the large computation cost resulting from KPCA. The independent principal components analysis (IPCA) and other selection algorithms such as U matrix can be taken into consideration.

3. The KSVM has shown satisfactory results for the proposed fault feature extraction and normal-alarm-fault classification. However, the SVM does not perform very ideal when dealing with large datasets especially when the dataset contains the noise. The target classes due to ambiguous regions in the data could overlap with the noise, leading to misclassification in the results. Other deep-learning algorithms, such as deep belief network (DBN) and convolutional neural network (CNN), can be used to improve the accuracy of the model.
4. The LSTM has shown benefits on the trend predictions as compared with RNN. The LSTM is able to solve the gradient decent problems and hence has a better performance for maintaining information on inputs. However, the computation cost of LSTM can be massive. Other optimisation methods, such as adaptive particle swarm optimisation (APSO), can be introduced to improve the performance of LSTM, with the normalised mean absolute error and naive ratio being used for performance evaluation.
5. Due to the time limitation, the design and construction of WT power conversion test rig is simplified. For example, the phase locked loop was not installed on the test rig, resulting in an open loop control of the inverter.

In addition, for the rotor side converter, the IGBT is used as the function of rectifier. Further work can be focused on the hardware improvement of technologies.

6. Because of the safety issue, the test rig was designed for low-voltage implementation. The dynamic behaviours of the test rig can be different when compared with high voltage configurations. Besides, some high-voltage components like IGBT generate extra noise when running in a low-voltage condition. Therefore, enhancing safety level of the test rig can be further developed so that high-voltage tests can be implemented.

References

[1] International energy agency, "*Transforming our world: the 2030 Agenda for Sustainable Development,"* 2018 [online] Available at: <https://iea.blob.core.windows.net/assets/23f9eb39-7493-4722-aced-61433cbffe10/Global_Energy_and_CO2_Status_Report_2018.pdf> [Accessed 31 December 2021].

[2] B. Roy and A. Schaffartzik, "Talk renewables, walk coal: The paradox of India's energy transition," *Ecological Economics*, vol. 180, Feb. 2021, doi: 10.1016/j.ecolecon.2020.106871.

[3] United nations, "*Global Energy & CO2 Status Report,"* 2015 [online] Available at: < https://sdgs.un.org/2030agenda> [Accessed 10 March 2022].

[4] D. Griggs, M. Nilsson, A. Stevance and D. McCollum, "*A Guide to SDG Interactions,"* Paris: International Council for Science, 2017.

[5] C. Allen, G. Metternicht, and T. Wiedmann, "National pathways to the Sustainable Development Goals (SDGs): A comparative review of scenario modelling tools," *Environmental Science and Policy*, vol. 66, pp. 199–207, Dec. 2016, doi: 10.1016/j.envsci.2016.09.008.

[6] D. L. McCollum, L. G. Echeverri, S. Busch, S. Pachauri, S. Parkinson, J. Rogelj, V. Krey, J. C. Minx and M. Nilsson , "Connecting the sustainable development goals by their energy inter-linkages," *Environmental Research Letters*, vol. 13, no. 3. Institute of Physics Publishing, 2018. doi: 10.1088/1748-9326/aaafe3.

[7] F. F. Nerini, J. Tomei, L. S. To, I, Bisaga, P. Parikh, M. Black, A. Borrion, C. Spataru, V. C. Broto, G. Anandarajah, B. Milligan and Y. Mulugetta, "Mapping synergies and trade-offs between energy and the Sustainable Development Goals," *Nature Energy*, vol. 3, no. 1, pp. 10–15, Jan. 2018, doi: 10.1038/s41560-017-0036-5.

[8] OECD and Food and Agriculture Organization of the United Nations, "*OECD-FAO Agricultural Outlook 2015,"* OECD, 2015. doi: 10.1787/agr_outlook-2015-en.

[9] C. V. Stechow, J. C. Minx, K. Riahi, J. Jewell, D. L. McCollum, M. W. Callaghan, C. Bertram, G. Luderer and G. Baiocchi, "2°C and SDGs: United they stand, divided they fall?," *Environmental Research Letters*, vol. 11, no. 3, Mar. 2016, doi: 10.1088/1748-9326/11/3/034022.

[10] K. J. Bowen, N. A. Cradock-Henry, F. Koch, J. Patterson, T. Hayha, J. Vogt and F. Barbi., "Implementing the 'Sustainable Development Goals': towards addressing three key governance challenges—collective action, trade-offs, and accountability," *Current Opinion in Environmental Sustainability*, vol. 26–27. Elsevier B.V., pp. 90–96, Jun. 01, 2017. doi: 10.1016/j.cosust.2017.05.002.

[11] European Commission DG Energy, 2017. "*Study on Technical Assistance in Realisation of the 2016 Report on Renewable Energy, in preparation of the Renewable Energy Package for the Period 2020-2030 in the European Union 'RES-Study'*,". [online] Available at: <https://ec.europa.eu/energy/sites/ener/files/documents/res-study_final_report_170227.pdf> [Accessed 31 December 2021].

[12] European Environment Agency, 2014. "*Use of renewable energy for transport in Europe*,". [Online] Available at: <https://www.eea.europa.eu/data-and-maps/indicators/use-of-cleaner-and-alternative-fuels-2> [Accessed 11 March 2022].

[13] The guardian, 2021. "*EU leaders agree to cut greenhouse gas emissions by 40% by 2030*,". [Online] Available at: <https://www.theguardian.com/world/2014/oct/24/eu-leaders-agree-to-cut-greenhouse-gas-emissions-by-40-by-2030> [Accessed 11 March 2022].

[14] J. Brodny, M. Tutak, and P. Bindzár, "Assessing the level of renewable energy development in the european union member states. A 10-year perspective," *Energies*, vol. 14, no. 13, Jul. 2021, doi: 10.3390/en14133765.

[15] Global wind energy council, 2021. "*Global Wind Report 2021,"*. [online] Brussels: Global wind energy council. Available at: <https://gwec.net/wp-content/uploads/2021/03/GWEC-Global-Wind-Report-2021.pdf> [Accessed 31 December 2021].

[16] W. Short, D. Packey and T. Holt, "*A manual for the economic evaluation of energy efficiency and renewable energy technologies,"* Honolulu, Hawaii: University Press of the Pacific, 2005.

[17] C. Kost, S. Shammugam, V. Jülch, H.-T. Nguyen, and T. Schlegl, "*Levelised cost of electricity renewable energy*," 2018. [Online]. Available at: < https://www.ise.fraunhofer.de/content/dam/ise/en/documents/publications/studies/EN2018_Fraunhofer-ISE_LCOE_Renewable_Energy_Technologies.pdf> [Accessed 31 December 2021].

[18] T. Stehly, P. Beiter, and P. Duffy, "*2019 Cost of Wind Energy Review*," 2019. [Online]. Available: www.nrel.gov/publications. [Accessed 31 December 2021].

[19] J. Zhou, Y. Liu and Y. Chen, "Face Recognition Using Kernel PCA and Hierarchical RBF Network," *6th International Conference on Computer Information Systems and Industrial Management Applications (CISIM'07)*, 2007, pp. 239-244, doi: 10.1109/CISIM.2007.28.

[20] F. P. García Márquez, A. M. Tobias, J. M. Pinar Pérez, and M. Papaelias, "Condition monitoring of wind turbines: Techniques and methods," *Renewable Energy*, vol. 46, pp. 169–178, Oct. 2012, doi: 10.1016/j.renene.2012.03.003.

[21] W. Qiao and D. Lu, "A Survey on Wind Turbine Condition Monitoring and Fault Diagnosis - Part II: Signals and Signal Processing Methods," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10. Institute of Electrical and Electronics Engineers Inc., pp. 6546–6557, Oct. 01, 2015. doi: 10.1109/TIE.2015.2422394.

[22] P. Qian, X. Ma, and P. Cross, “Integrated data-driven model-based approach to condition monitoring of the wind turbine gearbox,” in *IET Renewable Power Generation*, Jul. 2017, vol. 11, no. 9, pp. 1177–1185. doi: 10.1049/iet-rpg.2016.0216.

[23] D. Lu and W. Qiao, "Frequency demodulation-aided condition monitoring for drivetrain gearboxes," 2013 *IEEE Transportation Electrification Conference and Expo (ITEC)*, 2013, pp. 1-6, doi: 10.1109/ITEC.2013.6574526.

[24] P. Caselitz and J. Giebhardt, "Rotor Condition Monitoring for Improved Operational Safety of Offshore Wind Energy Converters ." *ASME. J. Sol. Energy Eng*. May 2005; 127(2): 253–261. https://doi.org/10.1115/1.1850485.

[25] P. Guo and N. Bai, “Wind turbine gearbox condition monitoring with AAKR and moving window statistic methods,” *Energies*, vol. 4, no. 11, pp. 2077–2093, 2011, doi: 10.3390/en4112077.

[26] P. Bangalore and L. B. Tjernberg, “An artificial neural network approach for early fault detection of gearbox bearings,” *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 980–987, Mar. 2015, doi: 10.1109/TSG.2014.2386305.

[27] Y. Zhao, D. Li, A. Dong, D. Kang, Q. Lv, and L. Shang, “Fault prediction and diagnosis of wind turbine generators using SCADA data,” *Energies*, vol. 10, no. 8, Aug. 2017, doi: 10.3390/en10081210.

[28] H. Soliman, H. Wang, B. Gadalla, and F. Blaabjerg, “Condition monitoring for DC-link capacitors based on artificial neural network algorithm,” in *International Conference on Power Engineering, Energy and Electrical Drives*, Sep. 2015, vol. 2015-September, pp. 587–591. doi: 10.1109/PowerEng.2015.7266382.

[29] Z. Kong, B. Tang, L. Deng, W. Liu, and Y. Han, “Condition monitoring of wind turbines based on spatio-temporal fusion of SCADA data by convolutional neural networks and gated recurrent units,” *Renewable Energy*, vol. 146, pp. 760–768, Feb. 2020, doi: 10.1016/j.renene.2019.07.033.

[30] L. Wang, Z. Zhang, H. Long, J. Xu, and R. Liu, “Wind Turbine Gearbox Failure Identification with Deep Neural Networks,” *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1360–1368, Jun. 2017, doi: 10.1109/TII.2016.2607179.

[31] C. H. Hu, H. Pei, X. S. Si, D. B. Du, Z. N. Pang, and X. Wang, “A prognostic model based on DBN and diffusion process for degrading bearing,” *IEEE Transactions on Industrial Electronics*, vol. 67, no. 10, pp. 8767–8777, Oct. 2020, doi: 10.1109/TIE.2019.2947839.

[32] J. Zhang, B. Xu, Z. Wang, and J. Zhang, “An FSK-MBCNN based method for compound fault diagnosis in wind turbine gearboxes,” *Measurement: Journal of the International Measurement Confederation*, vol. 172, Feb. 2021, doi: 10.1016/j.measurement.2020.108933.

[33] Y. Pan, R. Hong, J. Chen, and W. Wu, “A hybrid DBN-SOM-PF-based prognostic approach of remaining useful life for wind turbine gearbox,” *Renewable Energy*, vol. 152, pp. 138–154, Jun. 2020, doi: 10.1016/j.renene.2020.01.042.

[34] B. Zhao, C. Cheng, Z. Peng, X. Dong, and G. Meng, “Detecting the Early Damages in Structures with Nonlinear Output Frequency Response Functions and the CNN-LSTM Model,” *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 12, pp. 9557–9567, Dec. 2020, doi: 10.1109/TIM.2020.3005113.

[35] D. Wei, B. Wang, G. Lin, D. Liu, Z. Dong, H. Liu and Y. Liu, “Research on unstructured text data mining and fault classification based on RNN-LSTM with malfunction inspection report,” *Energies*, vol. 10, no. 3, 2017, doi: 10.3390/en10030406.

[36] P. Bangalore and M. Patriksson, “Analysis of SCADA data for early fault detection, with application to the maintenance management of wind turbines,” *Renewable Energy*, vol. 115, pp. 521–532, 2018, doi: 10.1016/j.renene.2017.08.073.

[37] H. Wang, H. Wang, G. Jiang, J. Li, and Y. Wang, “Early fault detection of wind turbines based on operational condition clustering and optimized deep belief network modeling,” *Energies*, vol. 12, no. 6, 2019, doi: 10.3390/en12060984.

[38] Y. Qiu, Y. Feng, and D. Infield, “Fault diagnosis of wind turbine with SCADA alarms based multidimensional information processing method,” *Renewable Energy*, vol. 145, pp. 1923–1931, Jan. 2020, doi: 10.1016/j.renene.2019.07.110.

[39] H. Zhao, H. Liu, W. Hu, and X. Yan, “Anomaly detection and fault analysis of wind turbine components based on deep learning network,” *Renewable Energy*, vol. 127, pp. 825–834, Nov. 2018, doi: 10.1016/j.renene.2018.05.024.

[40] Y. Qiu, Y. Feng, P. Tavner, P. Richardson, G. Erdos, and B. Chen, “Wind turbine SCADA alarm analysis for improving reliability,” *Wind Energy*, vol. 15, no. 8, pp. 951–966, Nov. 2012, doi: 10.1002/we.513.

[41] K. Leahy, C. Gallagher, P. O’Donovan, and D. T. J. O’Sullivan, “Cluster analysis of wind turbine alarms for characterising and classifying stoppages,” *IET Renewable Power Generation*, vol. 12, no. 10, pp. 1146–1154, Jul. 2018, doi: 10.1049/iet-rpg.2017.0422.

[42] B. Sorensen, *Renewable Energy: Four Volume Set. Routledge*. Routledge, 2018.

[43] R. Redlinger, D. Andersen, and E. Morthorst, *Wind energy in the 21st century: Economics, policy, technology and the changing electricity industry*. Spinger, 2016.

[44] A. Wyatt, *Electric Power: Challenges and Choices*. Toranto: Book Press, 1986.

[45] F. Castellani, D. Astolfi, M. Peppoloni, F. Natili, D, Buttà adn A, Hirschl, “Experimental Vibration Analysis of a Small-Scale Vertical Wind Energy System for Residential Use,” *Machines*, vol. 7, no. 35, 2019, https://doi.org/10.3390/machines7020035

[46] B. D. Altan and M. Atilgan, “An experimental and numerical study on the improvement of the performance of Savonius wind rotor,” *Energy Conversion and Management*, vol. 49, no. 12, pp. 3425–3432, Dec. 2008, doi: 10.1016/j.enconman.2008.08.021.

[47] L. Battisti, A. Brighenti, E. Benini, and M. R. Castelli, “Analysis of Different Blade Architectures on small VAWT Performance,” in *Journal of Physics: Conference Series*, Oct. 2016, vol. 753, no. 6. doi: 10.1088/1742-6596/753/6/062009.

[48] T. Burton, N. Jenkins, D. Sharpe, and E. Bossanyi, *Wind Energy Handbook*. Chichester, UK: John Wiley & Sons, Ltd, 2011. doi: 10.1002/9781119992714.

[49] M. K. Johari and A. F. Abdelgawad, “*Comparison of horizontal axis wind turbine (HAWT) and vertical axis wind turbine (VAWT)*,” 2018. [Online]. Available: www.sciencepubco.com/index.php/IJET [Accessed 31 December 2021].

[50] J. Huenteler, J. Ossenbrink, T. S. Schmidt, and V. H. Hoffmann, “How a product’s design hierarchy shapes the evolution of technological knowledge - Evidence from patent-citation networks in wind power,” *Research Policy*, vol. 45, no. 6, pp. 1195–1217, Jul. 2016, doi: 10.1016/j.respol.2016.03.014.

[51] A. Merabet, J. Thongam, and J. Gu, "Torque and pitch angle control for variable speed wind turbines in all operating regimes," *2011 10th International Conference on Environment and Electrical Engineering*, 2011, pp. 1-5, doi: 10.1109/EEEIC.2011.5874598.

[52] P. J. Tavner, *Offshore wind turbines : reliability, availability and maintenance*. Institution of Engineering and Technology, 2012.

[53] C. Dao, B. Kazemtabrizi, and C. Crabtree, “Wind turbine reliability data review and impacts on levelised cost of energy,” *Wind Energy*, vol. 22, no. 12. John Wiley and Sons Ltd, pp. 1848–1871, Dec. 01, 2019. doi: 10.1002/we.2404.

[54] K. Okedu, *Stability Control and Reliable Performance of Wind Turbines*. InTech, 2018. doi: 10.5772/intechopen.72160.

[55] C. Zhu, and Y. Li, "Reliability Analysis of Wind Turbines", in *Stability Control and Reliable Performance of Wind Turbines*. London, United Kingdom: IntechOpen, 2018, doi: 10.5772/intechopen.74859

[56] Y. Li, C. Zhu, X. Chen, and J. Tan, “Fatigue reliability analysis of wind turbine drivetrain considering strength degradation and load sharing using survival signature and FTA,” *Energies*, vol. 13, no. 8, Apr. 2020, doi: 10.3390/en13082108.

[57] P. Dong and J. K. Hong, “On the Residual Stress Profiles in New API 579/ASME FFS-1 Appendix E,” *Welding in the World*, vol. 51, pp. 119–127, 2007.

[58] L. Germanischer, “Guideline for the Certification of Wind Turbines,” Hamburg, 2007. [Online]. Available: www.gl-group.com/GLRenewables, [Accessed 31 December 2021].

[59] S. J. Watson, B. J. Xiang, W. Yang, P. J. Tavner, and C. J. Crabtree, “Condition monitoring of the power output of wind turbine generators using wavelets,” *IEEE Transactions on Energy Conversion*, vol. 25, no. 3, pp. 715–721, Sep. 2010, doi: 10.1109/TEC.2010.2040083.

[60] R. Janeliukstis and D. Mironovs, “Smart Composite Structures with Embedded Sensors for Load and Damage Monitoring – A Review,” *Mechanics of Composite Materials*, vol. 57, no. 2, pp. 131–152, May 2021, doi: 10.1007/s11029-021-09941-6.

[61] W. Yang, P. J. Tavner, C. J. Crabtree, and M. Wilkinson, “Cost-effective condition monitoring for wind turbines,” *IEEE Transactions on Industrial Electronics*, vol. 57, no. 1, pp. 263–271, Jan. 2010, doi: 10.1109/TIE.2009.2032202.

[62] W. Yang, P. J. Tavner, and M. Wilkinson, "Condition monitoring and fault diagnosis of a wind turbine with a synchronous generator using wavelet transforms," *2008 4th IET Conference on Power Electronics, Machines and Drives*, 2008, pp. 6-10, doi: 10.1049/cp:20080473.

[63] E. P. Carden and P. Fanning, “Vibration based condition monitoring: A review,” *Structural Health Monitoring*, vol. 3, no. 4. pp. 355–377, 2004. doi: 10.1177/1475921704047500.

[64] M. J. Schulz and M. J. Sundaresan, “*Smart Sensor System for Structural Condition Monitoring of Wind Turbines: May 30, 2002 - April 30, 2006*,” 2002. [Online]. Available: http://www.osti.gov/bridge, [Accessed 31 December 2021]

[65] A. G. Dutton, “*Thermoelastic stress measurement and acoustic emission monitoring in wind turbine blade testing*.” [Online]. Available: https://www.researchgate.net/publication/30411736, [Accessed 31 December 2021]

[66] D. Xu, P. F. Liu, and Z. P. Chen, “Damage mode identification and singular signal detection of composite wind turbine blade using acoustic emission,” *Composite Structures*, vol. 255, Jan. 2021, doi: 10.1016/j.compstruct.2020.112954.

[67] Z. Liu, X. Liu, K. Wang, Z. Liang, J. A. F. O. Correia, and A. M. P. de Jesus, “GA-BP neural network-based strain prediction in full-scale static testing of wind turbine blades,” *Energies*, vol. 12, no. 6, 2019, doi: 10.3390/en12061026.

[68] N. P. Avdelidis, D. P. Almond, C. Ibarra-Castanedo, A. Bendada, S. Kenny, and X. Maldague, “Structural integrity assessment of materials by thermography.” *Damage in Composite Materials CDCM*, Stuttgart, Germany 2006.

[69] S. Li, K. Shi, K. Yang, and J. Xu, “Research on the defect types judgment in wind turbine blades using ultrasonic NDT,” in *IOP Conference Series: Materials Science and Engineering*, Jul. 2015, vol. 87, no. 1. doi: 10.1088/1757-899X/87/1/012056.

[70] D. Wu, M. Zeng, H. Zhao, Y. Wang, and Y. Du, "Detection and localization of debonding beneath concrete pavement using transmissibility function analysis," *Mechanical Systems and Signal Processing*, vol. 159, Oct. 2021, doi: 10.1016/j.ymssp.2021.107802.

[71] J. Son, D. Kang, D. Boo, and K. Ko, "An experimental study on the fault diagnosis of wind turbines through a condition monitoring system," *Journal of Mechanical Science and Technology*, vol. 32, no. 12, pp. 5573–5582, Dec. 2018, doi: 10.1007/s12206-018-1103-y.

[72] Y. Peng, W. Qiao, L. Qu, and J. Wang, "Sensor Fault Detection and Isolation for a Wireless Sensor Network-Based Remote Wind Turbine Condition Monitoring System," in *IEEE Transactions on Industry Applications*, Mar. 2018, vol. 54, no. 2, pp. 1072–1079. doi: 10.1109/TIA.2017.2777925.

[73] A. Romero, S. Soua, T. H. Gan, and B. Wang, "Condition monitoring of a wind turbine drive train based on its power dependant vibrations," *Renewable Energy*, vol. 123, pp. 817–827, Aug. 2018, doi: 10.1016/j.renene.2017.07.086.

[74] M. Elforjani and E. Bechhoefer, "Analysis of extremely modulated faulty wind turbine data using spectral kurtosis and signal intensity estimator," *Renewable Energy*, vol. 127, pp. 258–268, Nov. 2018, doi: 10.1016/j.renene.2018.04.014.

[75] P. B. Dao, W. J. Staszewski, T. Barszcz, and T. Uhl, "Condition monitoring and fault detection in wind turbines based on cointegration analysis of SCADA data," *Renewable Energy*, vol. 116, pp. 107–122, Feb. 2018, doi: 10.1016/j.renene.2017.06.089.

[76] H. Wu, Z. Yu, and Y. Wang, "Real-time FDM machine condition monitoring and diagnosis based on acoustic emission and hidden semi-Markov model," *International Journal of Advanced Manufacturing Technology*, vol. 90, no. 5–8, pp. 2027–2036, May 2017, doi: 10.1007/s00170-016-9548-6.

[77] F. P. García Márquez, D. J. Pedregal Tercero, and F. Schmid, "Unobserved Component models applied to the assessment of wear in railway points: A case study," *European Journal of Operational Research*, vol. 176, no. 3, pp. 1703–1712, Feb. 2007, doi: 10.1016/j.ejor.2005.10.037.

[78] A. Joshuva and V. Sugumaran, "A lazy learning approach for condition monitoring of wind turbine blade using vibration signals and histogram features," *Measurement: Journal of the International Measurement Confederation*, vol. 152, Feb. 2020, doi: 10.1016/j.measurement.2019.107295.

[79] M. S. Rauscher, A. J. Tremmel, M. Schardt, and A. W. Koch, "Non-dispersive infrared sensor for online condition monitoring of gearbox oil," *Sensors (Switzerland)*, vol. 17, no. 2, Feb. 2017, doi: 10.3390/s17020399.

[80] Z. Liu, X. Wang, and L. Zhang, "Fault Diagnosis of Industrial Wind Turbine Blade Bearing Using Acoustic Emission Analysis," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6630–6639, Sep. 2020, doi: 10.1109/TIM.2020.2969062.

[81] M. J. Kabir, A. M. T. Oo and M. Rabbani, "A brief review on offshore wind turbine fault detection and recent development in condition monitoring based maintenance system," *2015 Australasian Universities Power Engineering Conference (AUPEC)*, 2015, pp. 1-7, doi: 10.1109/AUPEC.2015.7324871.

[82] X. Zhang, R. Zhou, and W. Zhang, “Improved local cepstrum and its applications for gearbox and rolling bearing fault detection,” *Measurement Science and Technology*, vol. 30, no. 7, Jun. 2019, doi: 10.1088/1361-6501/ab1235.

[83] F. Combet and L. Gelman, “An automated methodology for performing time synchronous averaging of a gearbox signal without speed sensor,” *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2590–2606, Aug. 2007, doi: 10.1016/j.ymssp.2006.12.006.

[84] R. P. B. Kocharla and R. K. Bandlamudi, “Condition Monitoring of Turbine Blades with Experimental Validation Using FFT Analyzer,” 2020, pp. 211–222. doi: 10.1007/978-981-15-1307-7_23.

[85] P. Shi, W. Yang, M. Sheng, and M. Wang, “An enhanced empiricalwavelet transform for features extraction fromwind turbine condition monitoring signals,” *Energies*, vol. 10, no. 7, 2017, doi: 10.3390/en10070972.

[86] D. Zhong, W. Guo and D. He, "An Intelligent Fault Diagnosis Method based on STFT and Convolutional Neural Network for Bearings Under Variable Working Conditions," *2019 Prognostics and System Health Management Conference (PHM-Qingdao)*, 2019, pp. 1-6, doi: 10.1109/PHM-Qingdao46334.2019.8943026.

[87] H. Malik, “Wavelet and Hilbert Huang transform based wind turbine imbalance fault classification model using k-nearest neighbour algorithm,” *International Journal of Renewable Energy Technology*, vol. 9, no. 1/2, p. 66, 2018, doi: 10.1504/IJRET.2018.090105.

[88] X. Xu, Y. Lei, and Z. Li, “An Incorrect Data Detection Method for Big Data Cleaning of Machinery Condition Monitoring,” *IEEE Transactions on Industrial Electronics*, vol. 67, no. 3, pp. 2326–2336, Mar. 2020, doi: 10.1109/TIE.2019.2903774.

[89] G. Liang, Y. Su, F. Chen, H. Long, Z. Song, and Y. Gan, “Wind Power Curve Data Cleaning by Image Thresholding Based on Class Uncertainty and Shape Dissimilarity,” *IEEE Transactions on Sustainable Energy*, vol. 12, no. 2, pp. 1383–1393, Apr. 2021, doi: 10.1109/TSTE.2020.3045782.

[90] Y. Wang, X. Ma, and M. J. Joyce, “Reducing sensor complexity for monitoring wind turbine performance using principal component analysis,” *Renewable Energy*, vol. 97, pp. 444–456, Nov. 2016, doi: 10.1016/j.renene.2016.06.006.

[91] W. Zhang and X. Ma, “Simultaneous fault detection and sensor selection for condition monitoring of wind turbines,” *Energies*, vol. 9, no. 4, Apr. 2016, doi: 10.3390/en9040280.

[92] J. Sun, C. Yan, and J. Wen, “Intelligent bearing fault diagnosis method combining compressed data acquisition and deep learning,” *IEEE Transactions on*

Instrumentation and Measurement, vol. 67, no. 1, pp. 185–195, Jan. 2018, doi: 10.1109/TIM.2017.2759418.

[93] H. Shao, H. Jiang, H. Zhang, W. Duan, T. Liang, and S. Wu, "Rolling bearing fault feature learning using improved convolutional deep belief network with compressed sensing," *Mechanical Systems and Signal Processing*, vol. 100, pp. 743–765, Feb. 2018, doi: 10.1016/j.ymssp.2017.08.002.

[94] A. L'Heureux, K. Grolinger, H. F. Elyamany, and M. A. M. Capretz, "Machine Learning with Big Data: Challenges and Approaches," *IEEE Access*, vol. 5, pp. 7776–7797, 2017, doi: 10.1109/ACCESS.2017.2696365.

[95] B. Gu, V. S. Sheng, Z. Wang, D. Ho, S. Osman, and S. Li, "Incremental learning for ν-Support Vector Regression," *Neural Networks*, vol. 67, pp. 140–150, Jul. 2015, doi: 10.1016/j.neunet.2015.03.013.

[96] P. Qian, D. Zhang, X. Tian, Y. Si, and L. Li, "A novel wind turbine condition monitoring method based on cloud computing," *Renewable Energy*, vol. 135, pp. 390–398, May 2019, doi: 10.1016/j.renene.2018.12.045.

[92] J. Zeng, D. Lu, Y. Zhao, Z. Zhang, W. Qiao, and X. Gong, "Wind turbine fault detection and isolation using support vector machine and a residual-based method," in *Proceedings of the American Control Conference*, 2013, pp. 3661–3666. doi: 10.1109/acc.2013.6580398.

[98] Z. Bo, Z. Yanan, and C. Changzheng, "Acoustic emission detection of fatigue cracks in wind turbine blades based on blind deconvolution separation," *Fatigue and Fracture of Engineering Materials and Structures*, vol. 40, no. 6, pp. 959–970, Jun. 2017, doi: 10.1111/ffe.12556.

[99] M. Awadallah and A. El-Sinawi, "Effect and detection of cracks on small wind turbine blade vibration using special Kriging analysis of spectral shifts," *Measurement: Journal of the International Measurement Confederation*, vol. 151, Feb. 2020, doi: 10.1016/j.measurement.2019.107076.

[100] R. K. Ibrahim, S. J. Watson, S. Djurović, and C. J. Crabtree, "An effective approach for rotor electrical asymmetry detection in wind turbine DFIGs," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 11, pp. 8872–8881, Nov. 2018, doi: 10.1109/TIE.2018.2811373.

[101] A. Reddy, V. Indragandhi, L. Ravi, and V. Subramaniyaswamy, "Detection of Cracks and damage in wind turbine blades using artificial intelligence-based image analytics," *Measurement: Journal of the International Measurement Confederation*, vol. 147, Dec. 2019, doi: 10.1016/j.measurement.2019.07.051.

[102] D. E. Choe, H. C. Kim, and M. H. Kim, "Sequence-based modeling of deep learning with LSTM and GRU networks for structural damage detection of floating offshore wind turbine blades," *Renewable Energy*, vol. 174, pp. 218–235, Aug. 2021, doi: 10.1016/j.renene.2021.04.025.

[103] Z. Guangfei, T. Wen, and Z. Da, “Ice Detection for Wind Turbine Blades Based on PSO-SVM Method,” in *Journal of Physics: Conference Series*, Oct. 2018, vol. 1087, no. 2. doi: 10.1088/1742-6596/1087/2/022036.

[104] H. Liu, Z. Zhang, H. Jia, Q. Li, Y. Liu, and J. Leng, “A novel method to predict the stiffness evolution of in-service wind turbine blades based on deep learning models,” *Composite Structures*, vol. 252, Nov. 2020, doi: 10.1016/j.compstruct.2020.112702.

[105] B. Yang, R. Liu, and X. Chen, “Fault Diagnosis for a Wind Turbine Generator Bearing via Sparse Representation and Shift-Invariant K-SVD,” *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1321–1331, Jun. 2017, doi: 10.1109/TII.2017.2662215.

[106] R. N. Bell, C. R. Heising, P. O’Donnell, S. J. Wells, and C. Singh, “Report of Large Motor Reliability Survey of Industrial and Commercial Installations, Part II,” *IEEE Transactions on Industry Applications*, vol. IA-21, no. 4, pp. 865–872, 1985, doi: 10.1109/TIA.1985.349533.

[107] O. E. Hassan, M. Amer, A. K. Abdelsalam, and B. W. Williams, “Induction motor broken rotor bar fault detection techniques based on fault signature analysis – A review,” *IET Electric Power Applications*, vol. 12, no. 7. Institution of Engineering and Technology, pp. 895–907, Aug. 01, 2018. doi: 10.1049/iet-epa.2018.0054.

[108] J. Chen, J. Pan, Z. Li, Y. Zi, and X. Chen, “Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals,” *Renewable Energy*, vol. 89, pp. 80–92, Apr. 2016, doi: 10.1016/j.renene.2015.12.010.

[109] Z. Liu and L. Zhang, “A review of failure modes, condition monitoring and fault diagnosis methods for large-scale wind turbine bearings,” *Measurement: Journal of the International Measurement Confederation*, vol. 149. Elsevier B.V., Jan. 01, 2020. doi: 10.1016/j.measurement.2019.107002.

[110] J. Li, M. Li, J. Zhang, and G. Jiang, “Frequency-shift multiscale noise tuning stochastic resonance method for fault diagnosis of generator bearing in wind turbine,” *Measurement: Journal of the International Measurement Confederation*, vol. 133, pp. 421–432, Feb. 2019, doi: 10.1016/j.measurement.2018.10.054.

[111] J. Li, J. Zhang, M. Li, and Y. Zhang, “A novel adaptive stochastic resonance method based on coupled bistable systems and its application in rolling bearing fault diagnosis,” *Mechanical Systems and Signal Processing*, vol. 114, pp. 128–145, Jan. 2019, doi: 10.1016/j.ymssp.2018.05.004.

[112] X. Zheng, G. Zhou, J. Wang, H. Ren, and D. Li, “Variational mode decomposition applied to offshore wind turbine rolling bearing fault diagnosis,” in *Chinese Control Conference, CCC*, Aug. 2016, vol. 2016-August, pp. 6673–6677. doi: 10.1109/ChiCC.2016.7554407.

[113] W. Y. Liu, Q. W. Gao, G. Ye, R. Ma, X. N. Lu, and J. G. Han, “A novel wind turbine bearing fault diagnosis method based on Integral Extension LMD,” *Measurement: Journal of the International Measurement Confederation*, vol. 74, pp. 70–77, Jul. 2015, doi: 10.1016/j.measurement.2015.06.005.

[114] L. Song, X. R. Zhang, L. W. Su, and Y. J. Gu, "Fault Diagnosis Approach for Incipient Bearing Fault in Wind Turbine under Variable Conditions," *Applied Mechanics and Materials*, vol. 599–601, pp. 312–320, Aug. 2014, doi: 10.4028/www.scientific.net/AMM.599-601.312.

[118] M. R. Shahriar, P. Borghesani, and A. C. C. Tan, "Electrical Signature Analysis-Based Detection of External Bearing Faults in Electromechanical Drivetrains," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5941–5950, Jul. 2018, doi: 10.1109/TIE.2017.2782240.

[116] R. Ziani, R. Zegadi, A. Felkaoui, and M. Djouada, "Bearing Fault Diagnosis Using Neural Network and Genetic Algorithms with the Trace Criterion," in *Condition Monitoring of Machinery in Non-Stationary Operations*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 89–96. doi: 10.1007/978-3-642-28768-8_10.

[117] M. Unal, M. Onat, M. Demetgul, and H. Kucuk, "Fault diagnosis of rolling bearings using a genetic algorithm optimized neural network," *Measurement: Journal of the International Measurement Confederation*, vol. 58, pp. 187–196, Dec. 2014, doi: 10.1016/j.measurement.2014.08.041.

[118] B. Ma, J. de Wu, J. Ma, X. D. Wang, and Y. G. Fan, "Fault Monitoring and Classification Method of Rolling Bearing Based on KICA and LSSVM," *Advanced Materials Research*, vol. 971–973, pp. 476–480, Jun. 2014, doi: 10.4028/www.scientific.net/AMR.971-973.476.

[119] Z. Feng and M. Liang, "Complex signal analysis for wind turbine planetary gearbox fault diagnosis via iterative atomic decomposition thresholding," *Journal of Sound and Vibration*, vol. 333, no. 20, pp. 5196–5211, Sep. 2014, doi: 10.1016/j.jsv.2014.05.029.

[120] Z. Feng, S. Qin, and M. Liang, "Time-frequency analysis based on Vold-Kalman filter and higher order energy separation for fault diagnosis of wind turbine planetary gearbox under nonstationary conditions," *Renewable Energy*, vol. 85, pp. 45–56, Jan. 2016, doi: 10.1016/j.renene.2015.06.041.

[121] G. He, K. Ding, W. Li, and X. Jiao, "A novel order tracking method for wind turbine planetary gearbox vibration analysis based on discrete spectrum correction technique," *Renewable Energy*, vol. 87, pp. 364–375, Mar. 2016, doi: 10.1016/j.renene.2015.10.036.

[122] Z. Li, X. Yan, X. Wang, and Z. Peng, "Detection of gear cracks in a complex gearbox of wind turbines using supervised bounded component analysis of vibration signals collected from multi-channel sensors," *Journal of Sound and Vibration*, vol. 371, pp. 406–433, Jun. 2016, doi: 10.1016/j.jsv.2016.02.021.

[123] Y. Kong, T. Wang, Z. Li, and F. Chu, "Fault feature extraction of planet gear in wind turbine gearbox based on spectral kurtosis and time wavelet energy spectrum," *Frontiers of Mechanical Engineering*, vol. 12, no. 3, pp. 406–419, Sep. 2017, doi: 10.1007/s11465-017-0419-0.

[124] W. Teng, X. Ding, X. Zhang, Y. Liu, and Z. Ma, "Multi-fault detection and failure analysis of wind turbine gearbox using complex wavelet transform," *Renewable Energy*, vol. 93, pp. 591–598, Aug. 2016, doi: 10.1016/j.renene.2016.03.025.

[125] M. Tang, Q. Zhao, S. X. Ding, H. Wu, L. Li, W. Long and B. Huang, "An improved lightGBM algorithm for online fault detection of wind turbine gearboxes," *Energies*, vol. 13, no. 4, 2020, doi: 10.3390/en13040807.

[126] A. Yin, Y. Yan, Z. Zhang, C. Li, and R. V. Sánchez, "Fault diagnosis of wind turbine gearbox based on the optimized LSTM neural network with cosine loss," *Sensors (Switzerland)*, vol. 20, no. 8, Apr. 2020, doi: 10.3390/s20082339.

[127] W. Hu, H. Chang, and X. Gu, "A novel fault diagnosis technique for wind turbine gearbox," *Applied Soft Computing Journal*, vol. 82, Sep. 2019, doi: 10.1016/j.asoc.2019.105556.

[128] Y. Chang, J. Chen, C. Qu, and T. Pan, "Intelligent fault diagnosis of Wind Turbines via a Deep Learning Network Using Parallel Convolution Layers with Multi-Scale Kernels," *Renewable Energy*, vol. 153, pp. 205–213, Jun. 2020, doi: 10.1016/j.renene.2020.02.004.

[129] F. Qu, J. Liu, H. Zhu, and B. Zhou, "Wind turbine fault detection based on expanded linguistic terms and rules using non-singleton fuzzy logic," *Applied Energy*, vol. 262, Mar. 2020, doi: 10.1016/j.apenergy.2019.114469.

[130] F. Blaabjerg, M. Liserre, and K. Ma, "Power electronics converters for wind turbine systems," in *IEEE Transactions on Industry Applications*, Mar. 2012, vol. 48, no. 2, pp. 708–719. doi: 10.1109/TIA.2011.2181290.

[131] M. Liserre, R. Cárdenas, M. Molinas, and J. Rodríguez, "Overview of multi-MW wind turbines and wind parks," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 4. pp. 1081–1095, Apr. 2011. doi: 10.1109/TIE.2010.2103910.

[132] Y. Qiu, H. Jiang, Y. Feng, M. Cao, Y. Zhao, and D. Li, "A New Fault Diagnosis Algorithm for PMSG Wind Turbine Power Converters under Variable Wind Speed Conditions," *Energies*, vol. 9, no. 7, p. 548, Jul. 2016, doi: 10.3390/en9070548.

[133] I. Jlassi, J. O. Estima, S. Khojet El Khil, N. Mrabet Bellaaj, and A. J. Marques Cardoso, "Multiple open-circuit faults diagnosis in back-to-back converters of PMSG drives for wind turbine systems," *IEEE Transactions on Power Electronics*, vol. 30, no. 5, pp. 2689–2702, May 2015, doi: 10.1109/TPEL.2014.2342506.

[134] H. Zhao and L. Cheng, "Open-circuit faults diagnosis in back-to-back converters of DF wind turbine," *IET Renewable Power Generation*, vol. 11, no. 4, pp. 417–424, Mar. 2017, doi: 10.1049/iet-rpg.2016.0150.

[135] A. Brunko, W. Holzke, H. Groke, B. Orlik, and N. Kaminski, "Model-Based Condition Monitoring of Power Semiconductor Devices in Wind Turbines," 2019.

[136] H. Li, X. Liao, Z. Zeng, Y. Hu, Y. Li, S. Liu, and L. Ran, "Thermal Coupling Analysis in a Multichip Paralleled IGBT Module for a DFIG Wind Turbine Power Converter," *IEEE Transactions on Energy Conversion*, vol. 32, no. 1, pp. 80–90, Mar. 2017, doi: 10.1109/TEC.2016.2614526.

[137] B. Ji, V. Pickert, W. Cao, and B. Zahawi, "In situ diagnostics and prognostics of wire bonding faults in IGBT modules for electric vehicle drives," *IEEE Transactions on Power Electronics*, vol. 28, no. 12, pp. 5568–5577, 2013, doi: 10.1109/TPEL.2013.2251358.

[138] J. Zhang, H. Sun, Z. Sun, W. Dong, and Y. Dong, "Fault Diagnosis of Wind Turbine Power Converter Considering Wavelet Transform, Feature Analysis, Judgment and BP Neural Network," *IEEE Access*, vol. 7, pp. 179799–179809, 2019, doi: 10.1109/ACCESS.2019.2958409.

[139] C. Xiao, Z. Liu, T. Zhang, and X. Zhang, "Deep learning method for fault detection of wind turbine converter," *Applied Sciences (Switzerland)*, vol. 11, no. 3, pp. 1–22, Feb. 2021, doi: 10.3390/app11031280.

[140] Z. Y. Xue, K. S. Xiahou, M. S. Li, T. Y. Ji, and Q. H. Wu, "Diagnosis of Multiple Open-Circuit Switch Faults Based on Long Short-Term Memory Network for DFIG-Based Wind Turbine Systems," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 8, no. 3, pp. 2600–2610, Sep. 2020, doi: 10.1109/JESTPE.2019.2908981.

[141] U. Bhardwaj, A. P. Teixeira, and C. G. Soares, "Reliability prediction of an offshore wind turbine gearbox," *Renewable Energy*, vol. 141, pp. 693–706, Oct. 2019, doi: 10.1016/j.renene.2019.03.136.

[142] Y. Zheng, J. Wei, K. Zhu, and B. Dong, "Reliability analysis assessment of the wind turbines system under multi-dimensions," *Advanced Composites Letters*, vol. 29, p. 2633366X2096633, Jan. 2020, doi: 10.1177/2633366X20966337.

[143] M. Ben-Daya, S. O. Duffuaa, A. Raouf, J. Knezevic, and D. Ait-Kadi, *Handbook of Maintenance Management and Engineering*. London: Springer London, 2009. doi: 10.1007/978-1-84882-472-0.

[144] J. Lee, J. Ni, D. Djurdjanovic, H. Qiu, and H. Liao, "Intelligent prognostics tools and e-maintenance," *Computers in Industry*, vol. 57, no. 6, pp. 476–489, Aug. 2006, doi: 10.1016/j.compind.2006.02.014.

[145] A. Raouf, *Maintenance Excellence: Optimizing Equipment Lifecycle Decision*, vol. 10, no. 1. 2004. doi: 10.1108/13552510410526893.

[146] D. J. Pedregal, F. P. García, and C. Roberts, "An algorithmic approach for maintenance management based on advanced state space systems and harmonic regressions," *Annals of Operations Research*, vol. 166, no. 1, pp. 109–124, Feb. 2009, doi: 10.1007/s10479-008-0403-5.

[147] R. Sahal, J. G. Breslin, and M. I. Ali, "Big data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case," *Journal of Manufacturing Systems*, vol. 54, pp. 138–151, Jan. 2020, doi: 10.1016/j.jmsy.2019.11.004.

[148] W. Yang, R. Court, and J. Jiang, "Wind turbine condition monitoring by the approach of SCADA data analysis," *Renewable Energy*, vol. 53, pp. 365–376, May 2013, doi: 10.1016/j.renene.2012.11.030.

[149] M. Dahane, M. Sahnoun, B. Bettayeb, D. Baudry, and H. Boudhar, “Impact of spare parts remanufacturing on the operation and maintenance performance of offshore wind turbines: a multi-agent approach,” *Journal of Intelligent Manufacturing*, vol. 28, no. 7, pp. 1531–1549, Oct. 2017, doi: 10.1007/s10845-015-1154-1.

[150] J. Heinermann and O. Kramer, “Machine learning ensembles for wind power prediction,” *Renewable Energy*, vol. 89, pp. 671–679, Apr. 2016, doi: 10.1016/j.renene.2015.11.073.

[151] M. Shafiee, "Maintenance logistics organization for offshore wind energy: Current progress and future perspectives", *Renewable Energy*, vol. 77, pp. 182-193, 2015. doi: 10.1016/j.renene.2014.11.045.

[152] Z. Ren, A. S. Verma, Y. Li, J. J. E. Teuwen, and Z. Jiang, “Offshore wind turbine operations and maintenance: A state-of-the-art review,” *Renewable and Sustainable Energy Reviews*, vol. 144. Elsevier Ltd, Jul. 01, 2021. doi: 10.1016/j.rser.2021.110886.

[153] R. Ahmad and S. Kamaruddin, “An overview of time-based and condition-based maintenance in industrial application,” *Computers and Industrial Engineering*, vol. 63, no. 1, pp. 135–149, 2012, doi: 10.1016/j.cie.2012.02.002.

[154] F. Cheng, L. Qu, and W. Qiao, “Fault prognosis and remaining useful life prediction of wind turbine gearboxes using current signal analysis,” *IEEE Transactions on Sustainable Energy*, vol. 9, no. 1, pp. 157–167, Jan. 2018, doi: 10.1109/TSTE.2017.2719626.

[155] X. Lei and P. A. Sandborn, “Maintenance scheduling based on remaining useful life predictions for wind farms managed using power purchase agreements,” *Renewable Energy*, vol. 116, pp. 188–198, Feb. 2018, doi: 10.1016/j.renene.2017.03.053.

[156] H. Ghamlouch, M. Fouladirad, and A. Grall, “The use of real option in condition-based maintenance scheduling for wind turbines with production and deterioration uncertainties,” *Reliability Engineering and System Safety*, vol. 188, pp. 614–623, Aug. 2019, doi: 10.1016/j.ress.2017.10.001.

[157] A. Ismail, L. Saidi, M. Sayadi, and M. Benbouzid, “Gaussian Process Regression Remaining Useful Lifetime Prediction of Thermally Aged Power IGBT,” in *Proceedings, IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society*, 2019, pp. 6004–6009.

[158] M. C. Garcia, M. A. Sanz-Bobi, and J. del Pico, “SIMAP: Intelligent System for Predictive Maintenance. Application to the health condition monitoring of a windturbine gearbox,” *Computers in Industry*, vol. 57, no. 6, pp. 552–568, Aug. 2006, doi: 10.1016/j.compind.2006.02.011.

[159] S. Zhong, A. A. Pantelous, M. Goh, and J. Zhou, “A reliability-and-cost-based fuzzy approach to optimize preventive maintenance scheduling for offshore wind farms,” *Mechanical Systems and Signal Processing*, vol. 124, pp. 643–663, Jun. 2019, doi: 10.1016/j.ymssp.2019.02.012.

[160] Y. Zhou, J. Miao, B. Yan, and Z. Zhang, “Bio-objective long-term maintenance scheduling for wind turbines in multiple wind farms,” *Renewable Energy*, vol. 160, pp. 1136–1147, Nov. 2020, doi: 10.1016/j.renene.2020.07.065.

[161] N. Y. Yürüşen, P. N. Rowley, S. J. Watson, and J. J. Melero, “Automated wind turbine maintenance scheduling,” *Reliability Engineering and System Safety*, vol. 200, Aug. 2020, doi: 10.1016/j.ress.2020.106965.

[162] D. Fan, Y. Ren, Q. Feng, B. Zhu, Y. Liu, and Z. Wang, “A hybrid heuristic optimization of maintenance routing and scheduling for offshore wind farms,” *Journal of Loss Prevention in the Process Industries*, vol. 62, Nov. 2019, doi: 10.1016/j.jlp.2019.103949.

[163] Z. Hameed, J. Vatn, and J. Heggset, “Challenges in the reliability and maintainability data collection for offshore wind turbines,” *Renewable Energy*, vol. 36, no. 8, pp. 2154–2165, Aug. 2011, doi: 10.1016/j.renene.2011.01.008.

[164] D. Bailey and E. Wright, *Practical SCADA for Industry*. Elsevier, 2003. doi: 10.1016/B978-0-7506-5805-8.X5000-4.

[165] R. J. Real-Calvo, A. Moreno-Munoz, V. Pallares-Lopez, M. J. Gonzalez-Redondo, and I. M. Moreno-Garcia, “Design of an Intelligent Electronic Device to control a private microgrid,” in *2012 IEEE Second International Conference on Consumer Electronics - Berlin (ICCE-Berlin)*, Sep. 2012, pp. 99–101. doi: 10.1109/ICCE-Berlin.2012.6336492.

[166] I. Ahmed, S. Obermeier, M. Naedele, and G. G. Richard III, “SCADA Systems: Challenges for Forensic Investigators,” *Computer*, vol. 45, no. 12, pp. 44–51, Dec. 2012, doi: 10.1109/MC.2012.325.

[167] K. Kim, G. Parthasarathy, O. Uluyol, W. Foslien, H. S. Sheng, and P. Fleming, “*Use of SCADA Data for Failure Detection in Wind Turbines*,” 2011. [Online]. Available: http://www.osti.gov/bridge, [Accessed 31 December 2021].

[168] Electrical4U, "*Wound Rotor Induction Motor: What is it? (Diagram & Speed Control)* “, 2022. [Online]. Available: https://www.electrical4u.com/wound-rotor-induction-motor/. [Accessed: 8 March 2022].

[169] J. B. Ekanayake, L. Holdsworth, XueGuang Wu, and N. Jenkins, “Dynamic modeling of doubly fed induction generator wind turbines,” *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 803–809, May 2003, doi: 10.1109/TPWRS.2003.811178.

[170] N. W. Miller, J. J. Sanchez-Gasca, W. W. Price, and R. W. Delmerico, “Dynamic modeling of GE 1.5 and 3.6 MW wind turbine-generators for stability simulations,” in *2003 IEEE Power Engineering Society General Meeting (IEEE Cat. No.03CH37491)*, pp. 1977–1983. doi: 10.1109/PES.2003.1267470.

[171] Functionbay, "*PMSM (Permanent Magnet Synchronous Machine)* “, 2022. [Online]. Available: https://functionbay.com/documentation/onlinehelp/default.htm#!Documents/pmsmpermanentmagnetsynchronousmachine.htm. [Accessed: 8 March 2022].

[172] P. Sundararajan, M. H. Sathik, F. Sasongko, C. S. Tan, J. Pou, F. Blaabjerg, A. K. Gupta, "Condition Monitoring of DC-Link Capacitors Using Goertzel Algorithm for Failure Precursor Parameter and Temperature Estimation," *IEEE Transactions on Power Electronics*, vol. 35, no. 6, pp. 6386–6396, Jun. 2020, doi: 10.1109/TPEL.2019.2951859.

[173] T. van Erven and P. Harremoes, "Rényi Divergence and Kullback-Leibler Divergence," *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 3797–3820, Jul. 2014, doi: 10.1109/TIT.2014.2320500.

[174] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY: Springer New York, 2000. doi: 10.1007/978-1-4757-3264-1.

[175] B. Scholkopf and A. Smola, *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT Press, 2018.

[176] J. Harmouche, C. Delpha, and D. Diallo, "Incipient fault detection and diagnosis based on Kullback–Leibler divergence using Principal Component Analysis: Part I," *Signal Processing*, vol. 94, pp. 278–287, Jan. 2014, doi: 10.1016/j.sigpro.2013.05.018.

[177] C. Skittides and W.-G. Früh, "Wind forecasting using Principal Component Analysis," *Renewable Energy*, vol. 69, pp. 365–374, Sep. 2014, doi: 10.1016/j.renene.2014.03.068.

[178] I. Colak, S. Sagiroglu, and M. Yesilbudak, "Data mining and wind power prediction: A literature review," *Renewable Energy*, vol. 46, pp. 241–247, Oct. 2012, doi: 10.1016/j.renene.2012.02.015.

[179] J. Dongarra, M. Gates, A. Haidar, J. Kurzak, P. Luszczek, S. Tomov and I. Ymazaki "The Singular Value Decomposition: Anatomy of Optimizing an Algorithm for Extreme Scale," *SIAM Review*, vol. 60, no. 4, pp. 808–865, Jan. 2018, doi: 10.1137/17M1117732.

[180] J. Zhou, Y. Liu, and Y. Chen, "Face Recognition Using Kernel PCA and Hierarchical RBF Network," in *6th International Conference on Computer Information Systems and Industrial Management Applications (CISIM'07)*, Jun. 2007, pp. 239–244. doi: 10.1109/CISIM.2007.28.

[181] R. Kline, *Principles and practice of structural equation modeling*. Guilford publications, 2015.

[182] Y. Wang, X. Ma, and P. Qian, "Wind Turbine Fault Detection and Identification Through PCA-Based Optimal Variable Selection," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 4, pp. 1627–1635, Oct. 2018, doi: 10.1109/TSTE.2018.2801625.

[183] M. Lydia, S. S. Kumar, A. I. Selvakumar, and G. E. Prem Kumar, "A comprehensive review on wind turbine power curve modeling techniques," *Renewable and Sustainable Energy Reviews*, vol. 30, pp. 452–460, Feb. 2014, doi: 10.1016/j.rser.2013.10.030.

[184] G. van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5929–5955, Dec. 2020, doi: 10.1007/s10462-020-09838-1.

[185] Le, Ho, Lee, and Jung, "Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting," *Water*, vol. 11, no. 7, p. 1387, Jul. 2019, doi: 10.3390/w11071387.

[186] K. Yang, Y. Liu, Y. Yao, S. Fan, and A. Mosleh, "Operational time-series data modeling via LSTM network integrating principal component analysis based on human experience," *Journal of Manufacturing Systems*, vol. 61, pp. 746–756, Oct. 2021, doi: 10.1016/j.jmsy.2020.11.020.

[187] M. Jalayer, C. Orsenigo, and C. Vercellis, "Fault detection and diagnosis for rotating machinery: A model based on convolutional LSTM, Fast Fourier and continuous wavelet transforms," *Computers in Industry*, vol. 125, Feb. 2021, doi: 10.1016/j.compind.2020.103378.

[188] Z. K. Abdul, A. K. Al-Talabani, and D. O. Ramadan, "A Hybrid Temporal Feature for Gear Fault Diagnosis Using the Long Short Term Memory," *IEEE Sensors Journal*, vol. 20, no. 23, pp. 14444–14452, Dec. 2020, doi: 10.1109/JSEN.2020.3007262.

[189] R. Zhao, J. Wang, R. Yan, and K. Mao, "Machine health monitoring with LSTM networks," in *2016 10th International Conference on Sensing Technology (ICST)*, Nov. 2016, pp. 1–6. doi: 10.1109/ICSensT.2016.7796266.

[190] A. Youssef, C. Delpha, and D. Diallo, "An optimal fault detection threshold for early detection using Kullback–Leibler Divergence for unknown distribution data," *Signal Processing*, vol. 120, pp. 266–279, Mar. 2016, doi: 10.1016/j.sigpro.2015.09.008.

[191] S. Dasgupta, "A cost function for similarity-based hierarchical clustering," in *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, Jun. 2016, pp. 118–127. doi: 10.1145/2897518.2897527.

[192] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, "The impact of class imbalance in classification performance metrics based on the binary confusion matrix," *Pattern Recognition*, vol. 91, pp. 216–231, Jul. 2019, doi: 10.1016/j.patcog.2019.02.023.

[193] Y. Lee, M. Chow. *Power electronics handbook*, Butterworth-Heinemann, 2017.

[194] S. L. Simonetti, E.A. Amorim, and D.C. Flávio, *Advances in Renewable Energies and Power Technologies*. Elsevier, 2018. doi: 10.1016/B978-0-12-812959-3.00015-0.

[195] D. (Archived), "DDS Waveform Generation Reference Design for LabVIEW FPGA (Archived)", *Forums.ni.com*, 2020. [Online]. Available: https://forums.ni.com/t5/Example-Code/DDS-Waveform-Generation-Reference-Design-for-LabVIEW-FPGA/ta-p/3996130. [Accessed: 31- Dec- 2021].

Appendices

Appendix A: Simulation parameters

Table A.1 Electrical control parameters of wind turbine generator

Parameter Name	Recommended Value
T_r	0.05
T_v	0.05
K_{pv}	20
K_{iv}	2
K_{Qi}	0.05
K_{Vi}	20
T_{vz}	1
Q_{max}	0.29
Q_{min}	-0.432
X_{IQmax}	0.07
X_{IQmin}	-0.07
V_{max}	1.05
V_{min}	0.95
V_{L1}	0.9
V_{H1}	1.1
T_{L1}	0.1
T_{L2}	0.5
T_{H1}	0.1
T_{H2}	1
Q_{L1}	0.45
Q_{L2}	0.45
Q_{H1}	0.45
Q_{H2}	-0.245
Q_{H3}	0.45
V_{hyst}	0.05
Z_c	0

Table A.2 C_p coefficients $a_{i,j}$

i	j	aij
4	4	4.97E-10
4	3	-7.15E-08
4	2	1.62E-06
4	1	-9.48E-06
4	0	1.48E-05
3	4	-8.92E-08
3	3	5.99E-06
3	2	-1.05E-04
3	1	5.41E-04
3	0	-8.60E-04
2	4	2.79E-06
2	3	-1.49E-04
2	2	2.15E-03
2	1	-1.10E-02
2	0	1.57E-02
1	4	-2.39E-05
1	3	1.07E-03
1	2	-1.39E-02
1	1	6.04E-02
1	0	-6.76E-02
0	4	1.15E-05
0	3	-1.34E-04
0	2	-1.24E-02
0	1	2.18E-01
0	0	-4.19E-01

Table A.3 Turbine control parameters

Parameter Name	Recommended Value
K_{pp}	150
K_{ip}	25
T_p (second)	0.01
Θ_{max}(degreees)	27
Θ_{min}(degreees)	0
$d\theta/dt_{max}$ (degree/second)	10
$d\theta/dt_{min}$ (degree/second)	-10
P_{max} (pu)	1
P_{min} (pu)	0.1
dP/dt_{max} (pu/second)	0.45
dP/dt_{min} (pu/second)	-0.45
K_{pc}	3
K_{ic}	30
K_{ptrq}	3
K_{itrq}	0.6
T_{pc}	0.05

Appendix B: Schematics of power converter test rig

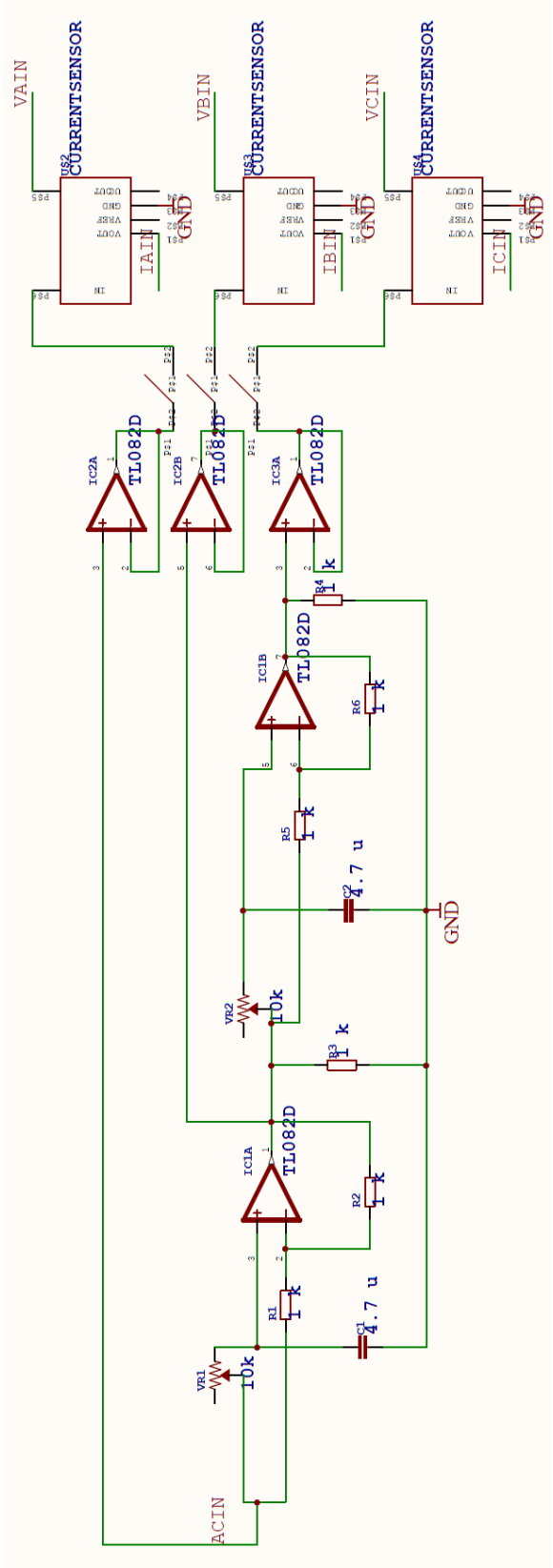


Figure B.1 Schematic of phase converter

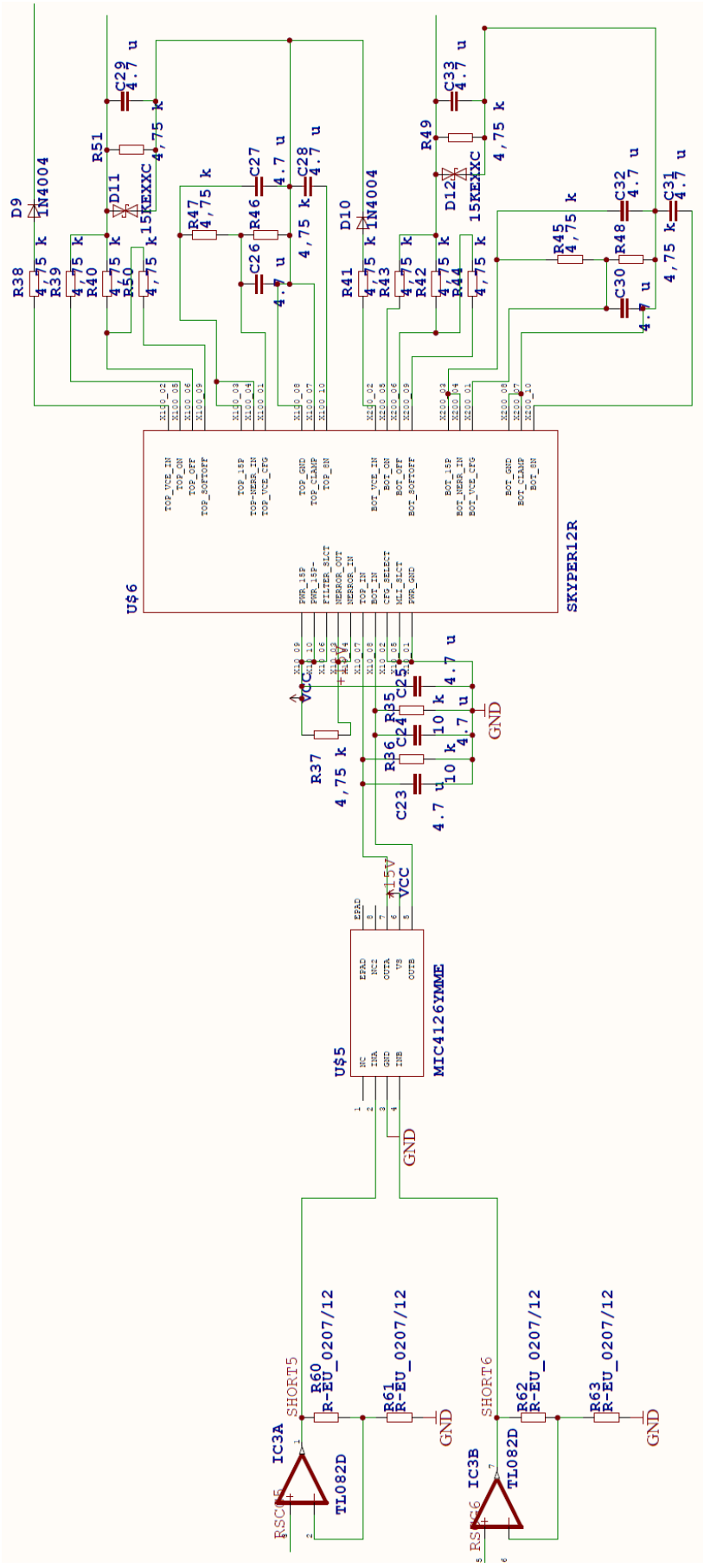


Figure B.2 Schematic of IGBT control circuit

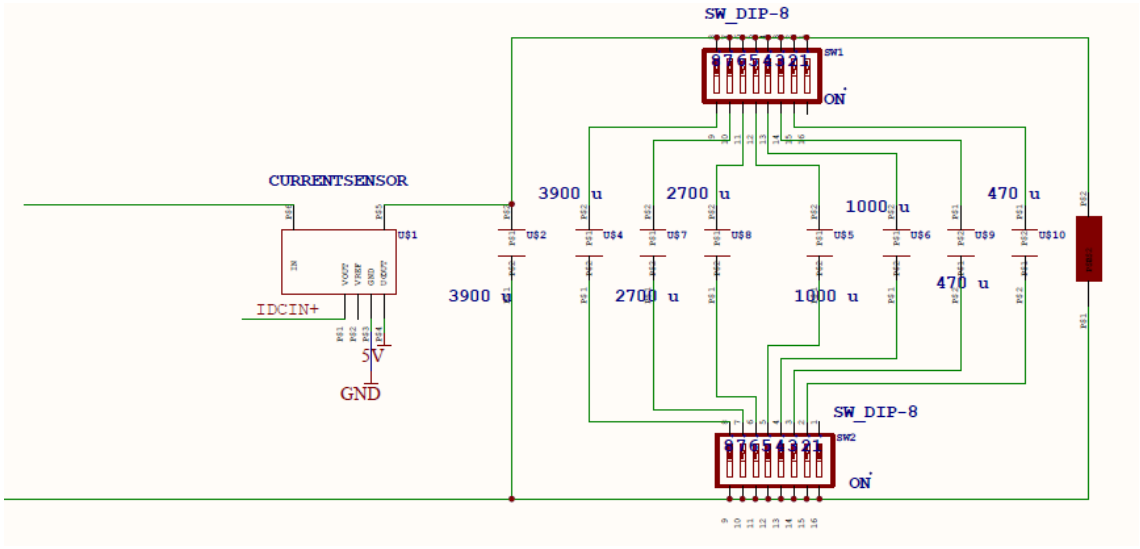


Figure B.3 Schematic of DC-link

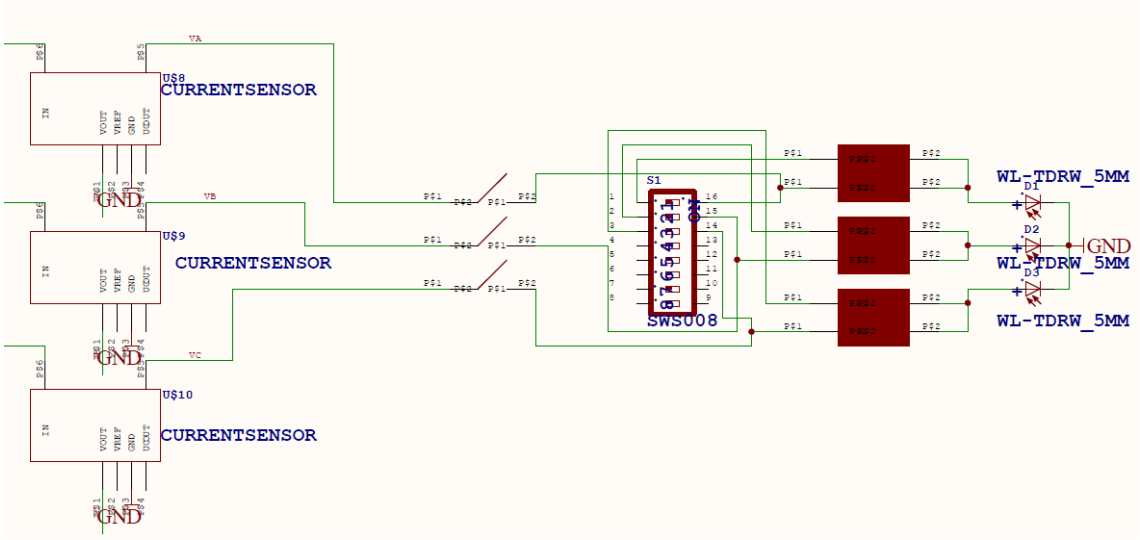


Figure B.4 Schematic of load

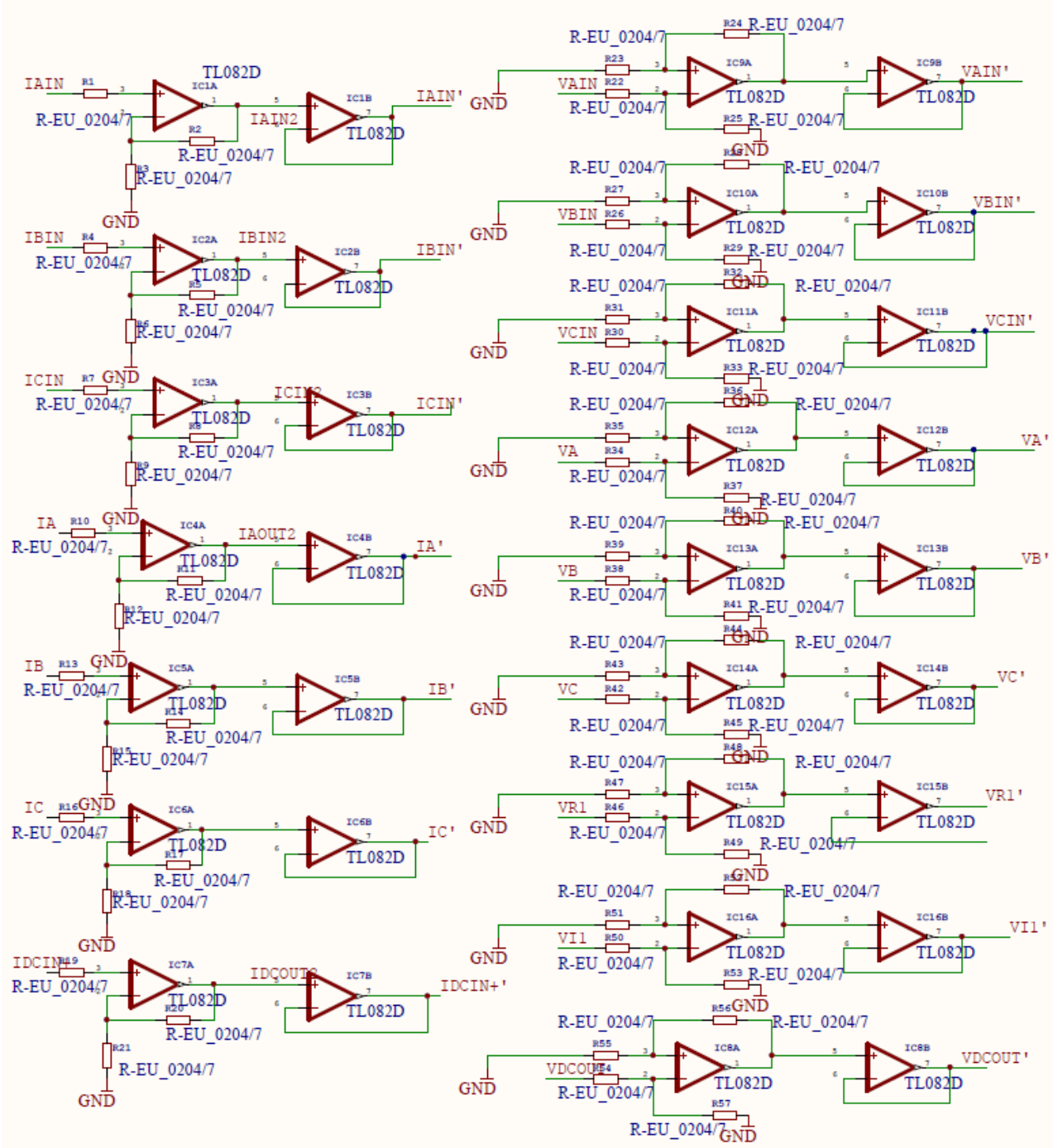


Figure B.5 Schematic of signal conditioning