Variable Selection for Wind Turbine Condition Monitoring and Fault Detection System

Yifei Wang

This thesis is submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

January 2016

Lancaster University
Faculty of Science and Technology
Engineering Department

Declaration

I declare that this dissertation consists of original work undertaken solely by myself at Lancaster University and not been submitted in substantially the same form towards the award of a degree or other qualificatory work, and affirm that acknowledgement has been made to assistance given and that all major sources have been appropriately referenced.

Yifei Wang

January 2016

Abstract

With the fast growth in wind energy, the performance and reliability of the wind power generation system has become a major issue in order to achieve cost-effective generation. Integration of condition monitoring system (CMS) in the wind turbine has been considered as the most viable solution, which enhances maintenance scheduling and achieving a more reliable system. However, for an effective CMS, large number of sensors and high sampling frequency are required, resulting in a large amount of data to be generated. This has become a burden for the CMS and the fault detection system.

This thesis focuses on the development of variable selection algorithm, such that the dimensionality of the monitoring data can be reduced, while useful information in relation to the later fault diagnosis and prognosis is preserved.

The research started with a background and review of the current status of CMS in wind energy. Then, simulation of the wind turbine systems is carried out in order to generate useful monitoring data, including both healthy and faulty conditions.

Variable selection algorithms based on multivariate principal component analysis are proposed at the system level. The proposed method is then further extended by introducing additional criterion during the selection process, where the retained variables are targeted to a specific fault. Further analyses of the retained variables are carried out, and it has shown that fault features are present in the dataset with reduced dimensionality.

Two detection algorithms are then proposed utilising the datasets obtained from the selection algorithm. The algorithms allow accurate detection, identification and severity estimation of anomalies from simulation data and supervisory control and data acquisition data from an operational wind farm.

Finally an experimental wind turbine test rig is designed and constructed. Experimental monitoring data under healthy and faulty conditions is obtained to further validate the proposed detection algorithms.

Acknowledgement

I am extremely grateful to my supervisor Dr Xiandong Ma for believing in me and offering his support throughout the project. His guidance, feedback and patience were essential not only for the progress of the research work, but also for the personal development during the PhD study.

I would like to thank my second supervisor Prof. Malcolm Joyce for his important support throughout this work.

I wish to express my special and sincere thanks to Prof. Roger Kamp and Robert Mackin for their support in the design of the experimental test rig.

Thank you for the technicians from the engineering department, particularly Andy Baker, Dr. James Dickinson, Jonathan Gates, Barry Noble, Mark Salisbury and Andrew Verden, who have helped with the experimental test rig.

I also appreciate the supports from Peter Caley, Jan Henschke and Kathryn Rucastle in the development of the test rig.

The permission of use SCADA data from Wind Prospect Ltd. is also gratefully acknowledged.

Finally, I would like to express my gratitude to my parents. Their continuous encouragement and support have been invaluable for me throughout the PhD study. This thesis is dedicated to them.

Table of Content

Nomenclature

Abbreviations

ANN	Artificial Neural Network
CMS	Condition Monitoring System
DFIG	Doubly-Fed Induction Generator
EMD	Empirical Mode Decomposition
FFT	Fast Fourier Transform
FMEA	Failure Mode and Effect Analysis
FMECA	Failure Mode Effects and Criticality Analysis
HAWT	Horizontal Axis Wind Turbine
HHT	Hilbert-Huang Transform
IGBT	Insulated Gate Bipolar Transistor
IMF	Intrinsic Mode Function
LCOE	Levelised Cost of Energy
MPPT	Maximum Power Point Tracking
O&M	Operation and Maintenance
PCA	Principal Component Analysis
PC	Principal components
PI	Proportional Integral
PLL	Phase Locked Loop
PMSG	Permanent Magnet Synchronous Generator
PRESS	Predicted Residual Sum of Squares
PWM	Pulse Width Modulation
RMSE	Root Mean Squared Error
SCADA	Supervisory Control and Data Acquisition
SCIG	Squirrel-Cage Induction Generator
SHM	Structural Health Monitoring
SSE	Sum Squared Error
WT	Wind Turbine
WRIG	Wound Rotor Induction Generator

Roman letters

a	Axial induction factor
a'	Angular induction factor
$a_i(t)$	Instantaneous amplitude at level i
A	Turbine rotor swap area
$\mathbf{b}$	Bias matrix
c	Blade chord
C_{dc}	DC-link capacitance
$c_i(t)$	ith intrinsic mode function
C_p	Turbine power coefficient
cpm	Combination performance measure
$cppv$	Cumulative partial covariance matrix
$\mathcal{D}$	Discarded dataset
D_{eq}	Equivalent damping factors
e	Error between the actual and predicted signal
$E(X)$	Information entropy of variable X
F_D	Drag force
F_L	Lift force
F_N	Normal force
F_T	Tangential force
h	Sum of the squared correlation
h_{ik}	ith temporary IMF at k iteration
$H(X)$	Normalised information entropy of variable X
i	Current
I_{dc}	DC bus current
$\mathbf{I}$	Generator current vector
$\mathbf{J}$	Jacobian matrix
$J_{b,gb,g}$	Moment of inertias of the blades, gearbox and generator
K_{eq}	Equivalent elasticity stiffness
$\mathbf{L}$	Characteristic roots/Eigenvalues
$L_{\sigma s,\, \sigma r,\, ms}$	Magnetising and leakage inductances of the stator and rotor windings in DFIG model
$L_{D,Q,d,q,f}$	Inductances of stator, rotor and field windings in PMSG model
m	Mass of the air
$m_{ik}(t)$	Mean value of the upper and lower envelope of ith IMF

M	Pitching moment
$M_{f,fd,sd,sq}$	Mutual inductances between stator, rotor and field winds in PMSG model
P	Active power of the generator
P_w	Power extracted from the wind
P_{ref}	Reference power (Rated power of the turbine)
Q	Reactive power of the generator
$\mathbf{Q}$	Residual Q statistics
r	Pear's correlation coefficient
$\mathbf{r}$	Generator winding resistance matrix
r_z	Fisher's correlation coefficient
$r_i(t)$	ith residual signal in EMD
R	Turbine rotor radius
$\boldsymbol{\mathcal{R}}$	Retained dataset
R^2	Squared correlation coefficient
$\mathbf{S}$	Covariance matrix
$S_{rr.d}$	Partial covariance matrix for retained variable set
t_m	Total cumulative percentage
T^2	Hotelling's T2 statistic
T_e	Electromagnetic torque
$T_{b,t,gb,g}$	Torques in the blades, turbine, gearbox and generator
TC_i	Relative contribution index
u_{rel}	Incident air stream flows over the blade (wind speed)
$\mathbf{U}$	Characteristic vectors/Eigenvectors
v	Voltage
V_{dc}	DC bus voltage
$\mathbf{V}$	Generator voltage vector
$\mathbf{W}$	Weight matrix
$\mathbf{W_q}$	Scaled eigenvectors of first q column
$x_{up}(t)$	Upper envelope of the signal
$x_{low}(t)$	Lower envelope of the signal
$\mathbf{X}$	Original dataset
X_{ms}	Mutual reactance in DFIG model
$X_{mD,mQ}$	Main reactance for PMSG
$X_{\sigma s,\, \sigma r}$	Leakage reactance of stator and rotor windings in DFIG model

$X_{\sigma D,\, \sigma Q,\, \sigma d,\, \sigma q}$	Leakage reactance of stator and rotor windings in PMSG model
$\mathbf{Y}$	Output signals
$\mathbf{Y_q}$	Scaled PC scores of first q column
$\mathbf{Z}$	Principal component scores

Greek letters

α	Angle of attack
β	Pitch angle
η_e	Percentage information entropy
θ	Angular displacement of the reference frame
θ_t	Twist angle
θ_r	Angular displacement of generator rotor
$\theta_i(t)$	Instantaneous phase angle at level i
$\theta_{b,t,gb,g}$	Angular displacements of the blades, turbine, gearbox and generator
λ	Tip speed ratio
$\boldsymbol{\mu}$	Arithmetic mean of the dataset
ρ	Air density
$\boldsymbol{\sigma}$	Standard deviations of the dataset
φ	Angle of relative wind
ψ	Flux linkage of each winding
Ψ	Flux linkage vector
ω	Angular velocity of the reference frame
ω_r	Angular velocity of generator rotor
ω_e	Angular velocity of the grid voltage
ω_s	Synchronous speed
$\omega_i(t)$	Instantaneous frequency
$\omega_{b,t,gb,g}$	Angular speeds of the blades, turbine, gearbox and generator
Ω	Angular speed of turbine rotor

Subscripts

abc	Three phase normal reference frame
cond	Conduction
d	Freewheeling diode
d,q	dq0 reference frame

f	Field winding
igbt	IGBT
s	Stator windings
sw	Switching
r	Rotor windings
tar	Targeting signal of interest

Superscript

‘	Equivalent rotor parameters in stator winding
h	Number of hidden layers
i	Input layer of the neural network
o	Output layer of the neural network
pc	Principal components
ref	Reference value for the controller
T	Transpose of matrix
var	Variables

Accent marks

^	Estimated value of the signal
-	Mean value of the signal

List of Figures

List of Tables

Publications

Wang, Y. Ma, X. and Joyce, M. 2016, Reducing sensor complexity for monitoring wind turbine performance using principal component analysis, *Renewable Energy*. (Under Review)

Ma, X., Wang, Y., & Qin, J. (2013). Generic model of a community-based microgrid integrating wind turbines, photovoltaics and CHP generations. *Applied Energy*, 112, 1475–1482. 10.1016/j.apenergy.2012.12.035

Wang, Y., Ma, X., & Joyce, M. (2013). Enhancing condition monitoring of distributed generation systems through optimal sensor selection. *In Industrial Electronics Society, IECON 2013 - 39th Annual Conference of the IEEE*. pp. 7610-7616. Piscataway, N.J.: IEEE. 10.1109/IECON.2013.6700401

Wang, Y., Ma, X., & Joyce, M. (2012). Optimal sensor selection for wind turbine condition monitoring using multivariate principal component analysis approach. *Paper presented at Proceedings of the 18th International Conference on Automation & Computing (ICAC'12)*, Loughborough University, Leicestershire, UK, 8 September 2012, Loughborough, United Kingdom.

Cross, P., Ma, X., & Wang, Y. (2013). Feature selection for artificial neural network model-based condition monitoring of wind turbines. *In Proceedings of the 10th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies (CM 2013 & MFPT 2013)*. BINDT.

Qian, P., Ma, X., & Wang, Y. (2015). Condition monitoring of wind turbines based on extreme learning machine. *In ICAC2015: Proceedings of the 21st International Conference on Automation and Computing*. pp. 37-42. IEEE. 10.1109/IConAC.2015.7313974

Chapter 1. Introduction

This chapter begins with a brief overview of the current state of wind power. It then focuses on the importance of wind energy research and the need of a condition monitoring system for a wind turbine system. Thereafter, the motivation behind research in developing variable selection techniques for a cost-effective condition monitoring system is stated. The aim and objectives involved in the research are also listed. Finally, the layout of the dissertation is described.

1.1 The importance of wind energy research

In recent decades, due to the fast pace of technology development, the demand for electrical energy consumption has rapidly increased. Additionally, due to being faced with the reality of diminishing fossil fuel resources and the obligation to reduce greenhouse gases due to climate change, the urge to implement renewable energy resources is imperative. Many countries have issued policies and targets in favour of the establishment of commercial renewable power generation projects. The European Union (EU) has also published official directives that require 20% of the energy consumption in the EU to be renewable by 2020, of which a 15% share of energy from renewable generation was set aside for the United Kingdom [1]. In 2014, additional objectives have been proposed by the EU targeting an increasing percentage of renewable power generation to 27% and a reduction of greenhouse gas emissions of 40%, compared to the levels in 1990, by 2030 [2]. As in the United Kingdom, according to the roadmap published by the Department of Energy and Climate Change, the targets of renewable electricity by 2020 are 100% for Scotland, 40% for Wales and Northern Ireland is aimed to reach twice the amount of the current status, respectively [3].

With the recent advances in technology and commercial development in wind power, of all the renewable energy sources, wind power has shown the highest net capacity growth in Europe of 116.8 GW since 2000. By the end of 2014, a cumulative power of 128.8 GW had been installed, and it was predicted to produce, on average, 284 TWh yearly, which covers 10.2% of the total electricity consumption in Europe [4]. Moreover, Figure 1.1 shows the annual onshore and offshore wind power installed in Europe between 2001 and 2014. As can be seen, there is a steady growth rate of the annual total wind turbines that have been installed since 2001. The annual installation capacity of onshore turbines has become relatively stable since 2007. In contrast, there is a significant increase in offshore wind turbines installed yearly. In addition, according to the 'Wind energy scenarios for 2030' published by the European Wind Energy Association (EWEA), the estimated installed capacity by 2030, using a central scenario for onshore and offshore wind turbines, are 253.6 GW and 66.5 GW, respectively [5].

This corresponds to an increased rate of 23.6 and 43.84 times for onshore and offshore turbines compared to the 2014 data. These statistics show the great potential of wind power in energy generation, notably for the offshore system.

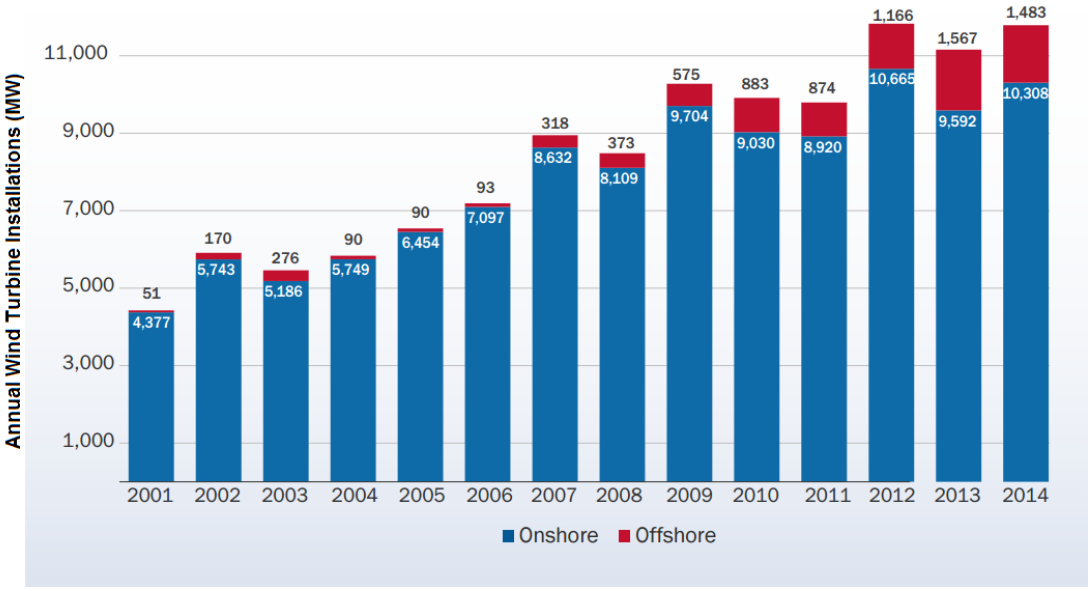


Figure 1.1 - Annual onshore and offshore installations (MW) in Europe [4]

Despite the significant energy generation potentials and benefits toward environmental issues of wind power, one of the most critical concerns is the high levelised cost of energy (LCOE) compared to the traditional generation technologies. This difference is primarily due to the high capital cost and the operation and maintenance cost. The estimated LCOE varies significantly according to the location, weather conditions and project. However, in the report published by the Department of Energy and Climate Change, the estimated LCOE for onshore and offshore wind turbines developed in 2010 are approximately 75 - 127 £/MWh and 149 - 191 £/MWh, respectively [3]. Although the onshore wind turbines are showing high competitiveness to conventional power generation techniques, offshore wind turbines still require further development and improvement in terms of lowering the LCOE, hence establishing stable, cost effective and environmental friendly energy sources. It is further mentioned in the report that, by 2020, the plan is to reduce the LCOE for onshore and offshore wind turbines to 71 - 122 £/MWh and 102 - 176 £/MWh, respectively [3]. In order to achieve these ambitious targets, further research and development in wind energy is essential.

Operation and maintenance (O&M) costs for wind turbines mainly involve the costs associated with routine maintenance, engineering and technician staff salaries, replacement parts and unexpected damage or maintenance. The O&M cost can take up to 10-15% of the total LCOE for onshore wind turbines, and the percentage goes up to approximately 25% for off-shore turbines [18, 49] primarily because of the extra cost from routine servicing and maintenance due to the poor accessibility of offshore sites. Weather conditions that result in a harsh sea or poor visibility also have a significant impact on operation and maintenance tasks, hence causing an increase in O&M expenditure. Consequently, any reduction of the cost in operation and maintenance will result in a lowering of the LCOE, hence enhancing the competitiveness and cost effectiveness of other energy sources.

When comparing the operation and maintenance aspects, the expenses that correspond with maintenance have a much higher percentage than that of operation, particularly for turbines that have been in operation for a lengthy period of time. After all, apart from the routine maintenance and exchange of parts, there is a considerable uncertainty of unexpected service or maintenance required under extreme conditions, which could cost a tremendous amount. Therefore, an effective maintenance strategy is necessary to provide reliable operation and to avoid catastrophic damages.

There are two main types of maintenance strategies [30]: corrective maintenance and preventive maintenance. The difference between the two types of maintenance primarily depends on the actions carried out before or after a failure has occurred. For corrective maintenance, action is performed after a fault is identified, and it is mainly aimed to restore the wind turbine to a functioning state. Preventive maintenance, on the other hand, is performed prior to the identification of a failure, hence reducing the probability of failure for the wind turbine and maintaining its healthy operation. This strategy can be further divided into scheduled maintenance and condition monitoring based maintenance.

Scheduled maintenance is executed at a regular time interval or quantitative measure which can be age related (i.e. the percentage of wear) [31]. This periodic

interval is system dependent, and it should be designed to minimise the probability of a failure rate. As such, the wind turbine is maintained in a healthy condition to minimise the cost and down time. As an example, for an onshore wind turbine, a scheduled maintenance can be performed every three months. As for the condition monitoring based maintenance, the service is carried out on the basis of the evaluation of a wind turbine's condition, where the turbine's operational performance is continuously monitored online through the condition monitoring system, hence achieving the optimal maintenance actions.

Figure 1.2 shows the comparison of the three types of maintenance strategies. As can be seen, the maintenance is performed posterior to the failure (breakdown point) for the corrective maintenance; the obvious disadvantage of such a strategy concerns the higher downtime for the wind turbine and the potential for severe consequential damage due to the failure. On the contrary, the scheduled maintenance is performed on a regular basis. However, this method is only suitable for age related faults where the fault can be quantified using probability distribution. The periodic inspection and replacement of the components might result in more frequent maintenance activities than actually required, thus limiting the life span for certain components. By incorporating the CMS, based on the performance of the wind turbine, maintenance is able to be performed optimally prior to the failure so that downtime can be minimised and for the full life span of the components are utilised. Furthermore, planning ahead for the maintenance becomes possible. This would be particular beneficial for offshore wind turbines, where uncertainties that stem from harsh weather conditions and deep sea water installations must be considered, hence resulting in an overall reduction in the O&M cost, which in turn reduces the LCOE. Consequently, an accurate and reliable CMS serves as the fundamental basis for the condition based maintenance.

1.2 Project motivation and aim

The purpose of the condition monitoring system is to monitor the performance and status of the wind turbine through sensor measurements and to identify the imminent or incipient failures of the wind turbine through the diagnosis and prognosis process by applying advanced signal processing and feature extraction

algorithms and techniques to the monitoring data. Therefore, it will allow an appropriate and logical scheduling of maintenance to be arranged as described earlier.

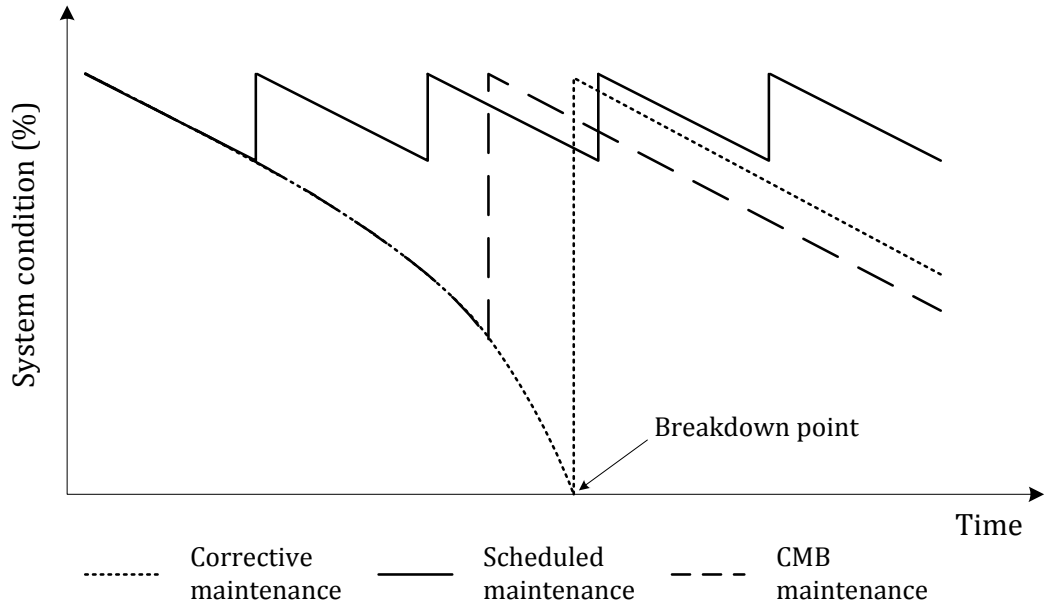


Figure 1.2 - Type of maintenance strategies: corrective, scheduled and condition monitoring based (CMB) maintenance

The CMS process is composed of data acquisition, signal processing and diagnosis and prognosis [3]. Extensive researches have been carried out in the development of effective and accurate signal processing techniques or diagnosis and prognosis algorithms, as will be shown in Chapter 2. As previous studies have found, accurate and reliable measurement data form the fundamental basis for CMS. Currently, the information maximisation principle is adopted for CMS data acquisition, which means sensors are installed to obtain as much information as possible. Typically, a wind turbine has approximately 150-250 monitoring points installed [6], and the number increases significantly for wind farms. The large amount of data (e.g. with a sampling rate of 10 minutes, for a wind farm with 20 wind turbines, there will be approximately 1.6×10^8 measurements per year) will be stored and processed. This particular application will present a realistic challenge for condition monitoring techniques optimised for fault detection and diagnosis. Moreover, with a higher number of monitoring points required, this might reduce the overall reliability and accountability of the data obtained [77, 78].

The physical and geometrical placement of the monitoring sensors has the potential to cause redundancy within the measurement data, and repeated information might be acquired. An appropriate variable selection technique is desirable in order to identify and remove the unnecessary repetitiveness due to there being too many sensors carrying out similar functions. In the meantime, the method should be able to retain the provision of vital information, which is critical for fault diagnosis, prognosis and maintenance scheduling. The aim of this research is to develop variable selection algorithms that are able to identify an optimal set of monitoring variables from a large amount of data to establish more effective condition monitoring of the wind turbines. In this case, the reliability of data measurement improves, and the amount of data to be stored and processed is reduced, thus improving the overall performance, accuracy and cost of the development of the wind turbine CMS.

1.3 Objectives

In relation to the proposed research, the following objectives will be addressed:

Objective 1 – To conduct a theoretical review of the fundamental working principles, the parameters involved and the common failure modes of a wind turbine system.

Objective 2 – To review current researches in wind turbine condition monitoring algorithms and techniques.

Objective 3 – To model and simulate a wind turbine system with different configurations and to further investigate the dynamic behaviour of a wind turbine with different fault scenarios.

Objective 4 – To propose appropriate variable selection methodologies in order to minimise redundancies in the measurement data, whilst still maintaining sufficient information to detect faults.

Objective 5 – To validate and analyse the proposed algorithms for the particular target signal with simulation data and SCADA data from the operational wind farm.

Objective 6 – To further evaluate if vital information relating to the particular fault feature is maintained in the reduced variable set obtained from the proposed variable selection algorithms.

Objective 7 – To locate the fault and identify the severity of the fault through an analysis of a data anomaly utilising variables retained from the selection algorithm.

Objective 8 – To design and construct a wind turbine experiment test rig to further validate the proposed algorithm for the detection and severity estimation of the fault.

1.4 Layout of the thesis

The layout of this thesis is as follows: In the first chapter, the importance of wind energy research and motivation of the projects is discussed. Additionally, the aims and objectives associated with the research and the layout of the dissertation are presented.

In the second chapter, a review is performed on the working principle of modern wind power generation systems with different configurations, the common failure modes for the key components of wind turbines and the current researches in the wind turbine CMS techniques and algorithms.

In the third chapter, the mathematical modelling and simulation of wind turbines with a permanent magnetic synchronised generator (PMSG) and a doubly fed induction generator (DFIG) are carried out in a PSCAD/EMTDC package, respectively. The simulation of a DC-link capacitor ageing fault and a grid line-ground short circuit fault are performed, where the dynamic behaviour of the system during fault is studied. Lastly, simulation data is collected under different operation conditions, which will be used in later chapters for the validation of the proposed algorithms. .

In the fourth chapter, a variable selection method based on the multivariate feature extraction through the principal component analysis is proposed, aiming to identify a set of variables with maximal information retained and minimal information redundancy at the system level. Furthermore, a modified method to

select a set of variables related specifically to a particular fault signal of interest is also proposed. Performance measures are offered and combined in order to evaluate the proposed selection algorithms. The proposed methods are then analysed and validated against simulation data obtained in Chapter 3 and SCADA data from a wind farm under different operational conditions.

The fifth chapter focuses on the further evaluation of the retained variable sets obtained from the selection algorithms proposed in Chapter 4 and demonstrates that vital information relating to the particular fault is still maintained. Two methods are adopted: a feature based technique is proposed to examine if the feature of a known fault is present in the retained variable set through a correlation test, and the ANN prediction model is developed to further evaluate the accuracy of the retained variable set.

In the sixth chapter, two fault detection algorithms are proposed by utilising the retained variable set obtained in Chapter 4. The former method incorporates the Hotelling's T^2 statistic to detect the anomaly in the measurement data. Thereafter, the identification of this anomaly is conducted by decomposing the T^2 statistic (during the period with abnormality), from which the variables contributing the most to this anomaly are identified. In the latter case, an empirical model for the particular fault is established based on the features obtained from the principal component analysis. This model then allows for the determination of the type of fault and for the estimation of the fault severity.

In the seventh chapter, the development of an experimental PMSG wind turbine test rig is presented, where the selection of components is discussed. An overview of the experiments carried out on the test rig is also given. Further validation of the fault detection algorithm proposed in Chapter 6 by means of the experimental data is given in the last section.

In the eighth chapter, a conclusion for the current research is given. Achievements corresponding to the objectives set out in this chapter and the contributions to knowledge arising from this research are presented. Finally, the limitations of the algorithms proposed in this research and recommendations for future research are discussed.

Chapter 2. Overview of the Wind Turbine Condition Monitoring Technologies

In this chapter, a review of the condition monitoring system and fault detection system (FD) for a wind turbine (WT) is given. Firstly, the background of current wind turbine technologies is described. Common failure modes of wind turbines are then discussed, followed by the techniques for evaluating turbine reliability. Subsequently, the review focuses on the condition monitoring system and fault detection, where different algorithms and methods are considered. These techniques range from statistical, time domain, frequency domain and time-frequency domain-based methods. Moreover, the integration of artificial intelligence in condition monitoring and fault detection is also addressed. Finally, this chapter covers the instrumentation required for effective data acquisition for wind turbine condition monitoring and a fault detection system, where challenges involved in the instrumentation are described.

2.1 Overview of WT technologies [7, 85, 86]

This section provides a general background of wind turbines. A detailed theoretical analysis and mathematical modelling of the wind turbine system and its subsystem is given in Chapter 3.

Wind energy is one of the most promising renewable power generation technologies. A wind turbine extracts energy stored in the airflow of natural wind and converts the kinematic energy into mechanical energy stored at the rotating turbine blades. This mechanical energy is then transformed into electrical power through an alternator or electrical generator.

The concept of capturing power from wind has existed for the last thousand years, where initially the wind energy was converted to mechanical power with the windmill [7]. The development of wind power began to decline after the appearance of thermal power, such as steam and oil engines. The attention to wind power then returned with the popularisation of electricity and the rapid development in aerodynamics through aeronautic research in the nineteenth century [7].

For modern wind turbines, different designs for capturing the wind power have been invented. The most popular concept is the horizontal axis turbine with two or three blades (Fig 2.1 a and c); However, vertical axis turbines, such as the Darrieus WT (Fig 2.1 b) and the Savonius WT (Fig 2.1 d) are also commonly used. Currently, the horizontal axis WT (HAWT) with three blades is the most mainstream of commercial wind turbines. Its customary use is mainly because the HAWT is considerably superior in regard to long-term investment due to higher generation efficiency and the possibility of developing larger sized turbines as a result of the less complicated design and configuration of HAWT [8].

A typical modern HAWT system consists of basic components/subsystems as shown in Figure 2.2 (a) and (b). The foundation is aimed to establish a reinforced and stable ground on which the wind turbine is stationed. It can be both a solid foundation for on-shore wind turbines and a floating foundation for off-shore wind turbines. The tower is installed to elevate the altitude of the turbine rotor. It has two main purposes: to permit a minimal ground clearance between the

lowest point of rotor swap area and the ground, where the swap area is an imaginary vertical surface created by the blade during rotation, and to increase energy capture, as the wind speed increases with elevation.

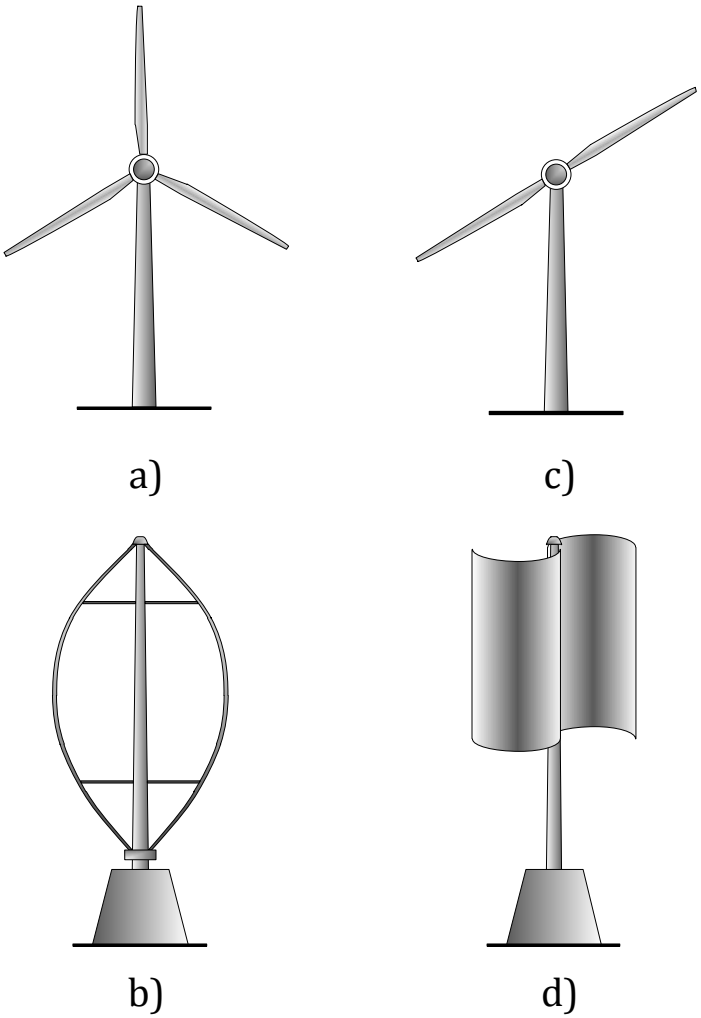


Figure 2.1 - Different types of wind turbine design concepts

The actual energy extraction from wind into mechanical rotation occurs through the turbine rotor. The rotor consists of a rotor hub and blades. The rotor hub acts as a central connection point between the blades and the rest of the system; the aerodynamic behaviour of the blades provides the capture of wind energy, and it can be either rigidly mounted to the rotor hub or allowed some degree of movement to optimise energy capture through a mechanism known as pitching. A detailed discussion of the blade and pitch is given in later sections.

The mechanical rotation from the rotor is transmitted to the electrical generator through the drive train. The drive train should withstand weight, reaction forces

and torque exerted from other subsystems. Moreover, a gearbox can be installed to step-up the rotational speed depending on the types of generator used. The generator then converts the mechanical rotational energy into electrical energy. The most commonly used generators in the wind industry are the AC induction generator and the permanent magnetic synchronous AC generator.

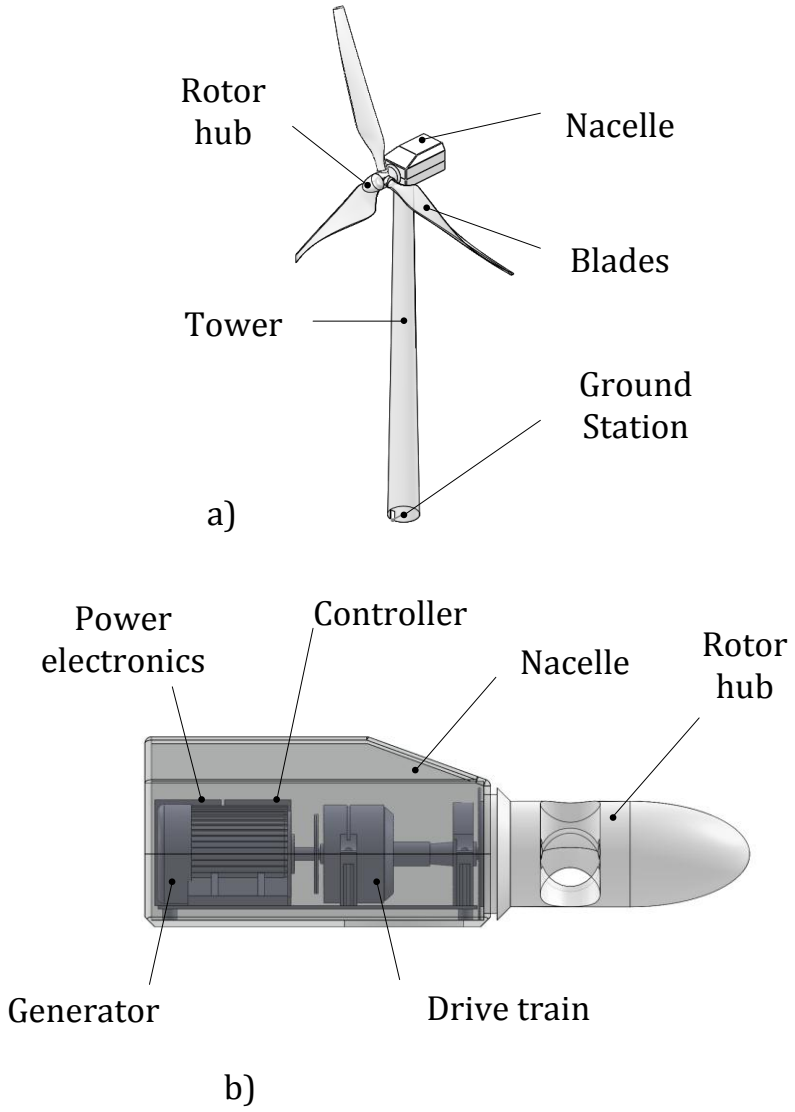


Figure 2.2 - Basic structure and subsystems of HAWT with three blades

The power generation has a direct relationship to the mechanical rotation, which is associated with the incoming wind. Because of the variable and stochastic behaviour of wind speed, the generated electrical power will have a random frequency and voltage. In order to deliver the power to the grid network, the regulation of the generated electrical power is acquired. Various techniques have been proposed; however, power electronics inverters have been proven to be the

best solution. It is adopted in order to realise the voltage stabilisation, filtering, grid synchronisation and other aspects. Moreover, a controller is installed to adjust the operating states of wind turbines for the achievement of certain functions, such as an effective power generation and optimal performance. This also includes systems such as the supervisory control and data acquisition (SCADA), which is responsible for the data acquisition, remote monitoring and control of the wind power generation system.

A wind turbine is generally installed in a harsh environment in order to increase its generation capacity factor. All the generation components are enclosed in the nacelle, where it acts as protection for the components inside. It also serves as the primary coupling between the tower and rest of the system, where a yaw drive mechanism is fitted and permits the turbine to adjust its area of swap facing the incoming wind at all times. Moreover, it provides a sturdy platform on which the drive train, generators and other components can be mounted. Finally, a ground station is generally constructed at the bottom of the tower. It serves the purpose of interfacing between the wind turbine and the grid network or load.

Wind turbines are usually installed either onshore or offshore. The main advantages of offshore wind turbines over onshore system are the significantly higher generation capacity factor due to a stronger wind profile and fewer environmental effects. These advantages are more obvious as the turbine is deployed farther away from the shore. However, positioning the wind turbine at a great distance from the shore will introduce several technical challenges and will also increase the capital and O&M cost.

All modern HAWT WTS have a similar structure, as described earlier. However, there are numerous configurations available that are categorised according to the properties or types of generators incorporated. Among these varieties, the four most commonly applied system architectures are shown in Figure 2.3 [9].

Figure 2.3 (a) shows the wind turbine architecture with constant speed. This means the turbine is optimised at one particular speed; any fluctuations in the wind speed will be transferred directly to the power output. A squirrel-cage induction generator (SCIG) is generally equipped and designed to operate at this specific speed. The gearbox is responsible for the increasing of the low rotor

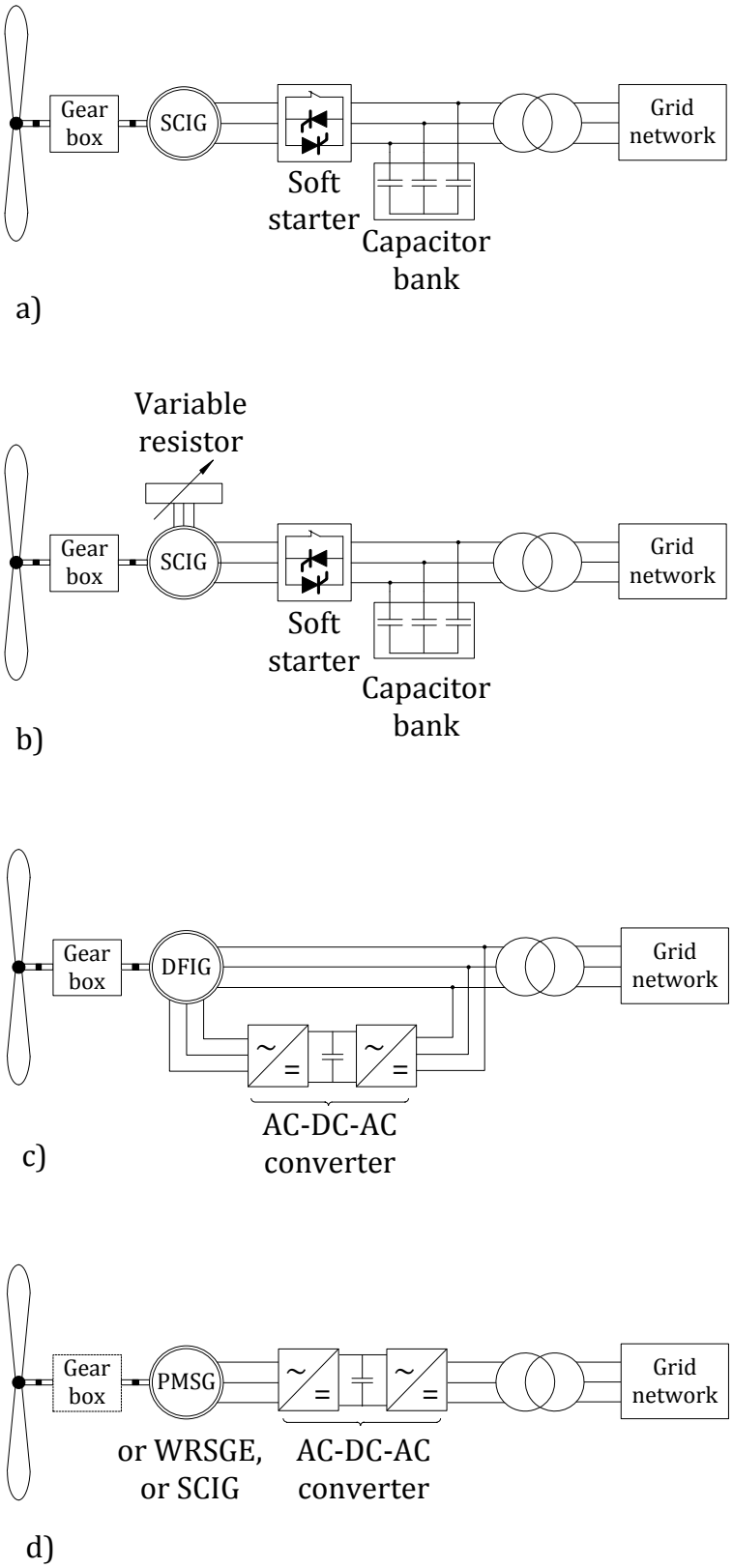


Figure 2.3 - Four commonly applied WT configurations

speed to match the generator rating. There are occasions where two sets of windings are installed on the generator, such that at different wind speeds,

different windings are used to generate power efficiently. Moreover, the soft starter, the capacitor bank and the transformer are used to optimise performance and regulate the power generated to meet grid connection standards. Although, this type of system includes the advantages of a simple structure, a lower initial cost, higher robustness and reliability, it suffers from the following major drawbacks: 1) higher mechanical requirements of the components to withstand sudden and substantial changes of the wind speed; 2) the need for a stiff grid network due to the fluctuation of the electrical output; 3) the turbine power is optimised at a single wind speed.

The rest of the configurations in Fig 2.3 (b) - (d) all incorporate variable speed control, where the turbine is designed to maximise power over a range of wind speeds. Figure 2.3 (b) achieves the variable speed by using a wound rotor induction generator (WRIG) and a variable resistor connected in a series with the rotor winding of the generator. The speed of the generator is controlled by the slip, which in turn is the result of varying the total rotor resistance. The range of the variable speed is determined by the size of the variable resistance, and it is typically around 0-10% of the synchronous speed [9]. Again, the gearbox, the soft starter and the capacitor banks serve the same purpose as previous turbine architecture. This configuration has the advantage of being more efficient due to the variable speed. Also, it reduces the high requirement of mechanical components as there is less fluctuation. However, the system clearly becomes more complex, and the initial cost and reliability have been recognisable issues.

With the growth in power electronic industry, new wind turbine configurations have recently been developed in order to achieve variable speed operations through power converters. Fig 2.3 (c) shows the configuration utilising power converters; such a system is known as the doubly-fed induction generator (DFIG). The system uses a WRIG where the stator is directly connected to the grid and a power converter is installed in the rotor circuit of the generator. The variable speed is realised by controlling the voltage and frequency fed into the rotor winding, such that the slip can be controlled. Typically, the converter takes up to approximately 30% of the nominal power of the generator, which will result in a variable speed range of +/- 30% in relation to synchronous speed

[109]. The reason for positive and negative speed ranges is that the system is capable of a bi-directional power flow depending on the slip, hence the term 'doubly fed'. When comparing the concept with variable resistance, where the energy produced during super synchronous speed is dissipated through the resistor, this power can also be captured and delivered to the network by integrating an AC-DC-AC power converter in the rotor circuit. Moreover, with the appropriate control of the power converter, it is capable of regulating the power output to meet grid connection standards. As such, the soft starter and the capacitor bank are eliminated. A detailed overview of the DFIG system will be given in the next chapter.

Finally, the last configuration shown in Fig 2.3 (d) uses a permanent magnetic synchronous generator (PMSG) and a full-scaled power converter. For the PMSG, it operates at a low rotational speed with a high number of poles; thus, the gearbox can be removed from the system. Power fluctuations due to variable wind speed are handled by the power converter, as well as other power regulations such as reactive power compensation and smooth grid connection. Because the gearbox has a high failure rate, the elimination of the gearbox increases the reliability of the system [10]. Thus, such a system is regarded as mainstream in off-shore wind turbine application. The utilisation of the power converter enables the system to have control over active and reactive power, to have faster response to transient behaviour and to increase the quality of the power output.

These different turbine architectures all require power control. This component serves two main purposes: to optimise power capture efficiency and to limit the power output at a very strong wind speed and correspondingly protect the turbine from getting damaged. This power control is usually carried out through the designing of the turbine blades' aerodynamics or through the incorporation of a pitching mechanism. There are three main alternatives for power control: a) Passive control; b) Pitch (Active) control and c) Active stall control.

Passive control (stall control) operates purely based on the design of the blades' aerodynamic, such that the laminar flow terminates when the wind speed reaches a limit (normally it is a turbine's rated power) and the replaced

turbulence flow will increase the drag force on the turbine blades and cause stalling, thus preventing the turbine speed to increase further.

A pitching mechanism is installed which allows the blade to rotate out of the wind using hydraulic or electrical motors. As a result, it changes its aerodynamic property, which is known as active control (pitch control). A threshold is predefined (at a specific wind speed); therefore, if the power output exceeds this limit, the pitching mechanism is activated in order to keep the power output at the rated level.

Active stall control is the combination of the above two control methods. During low wind speed, the blades are rotated at an angle to obtain a larger torque, thus achieving higher efficiency. However, when the limit is reached, different from the active control, the blades rotate in the opposite direction to increase the angle of attack, causing the blades to stall and stop the rotation. This method is commonly adopted for large powered turbines.

2.2 Common failure modes of the wind turbine system

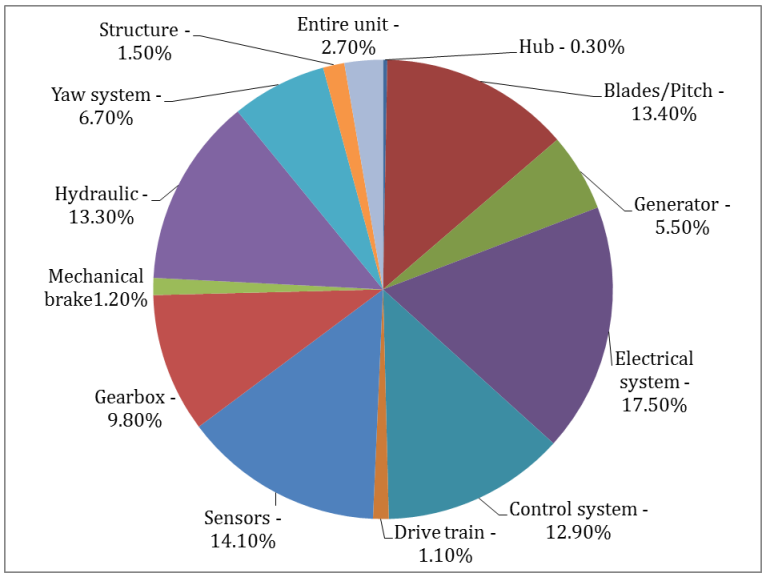
Due to the harsh environments where wind turbines are installed, such as a strong turbulent wind profile, extremely low temperatures and an icing effect on the blades, the reliability of the system has a great influence on the efficiency, performance and cost of electricity, especially for off-shore WTs. In order to determine the reliability of a wind turbine, knowledge of the failure rate, λ of the wind turbine and subsystems is need, such that it is the reciprocal of the mean time before failure. With the information on failure rates, a comparison of WTs' reliability and a further investigation and identification of unreliable sub-assemblies can be achieved. Failure rates of the components vary according to the different designs, configurations and manufacturers. Unfortunately, for WT, there is not yet a formal standard on the information gathering on failure rates as there is in other industries [11]. However, studies have been carried out in developing standard approaches for data collection of reliability [12, 13]. To collect the data, the development of a standard structure or taxonomy of WT is necessary, and it should provide a detailed definition of the data that should be collected in terms of both the structure and functionality of WT. Tavner *et al*. [18]

have adopted the standardised taxonomy developed by the ReliaWind project in their study on WT's reliability and detailed information can be found in reference.

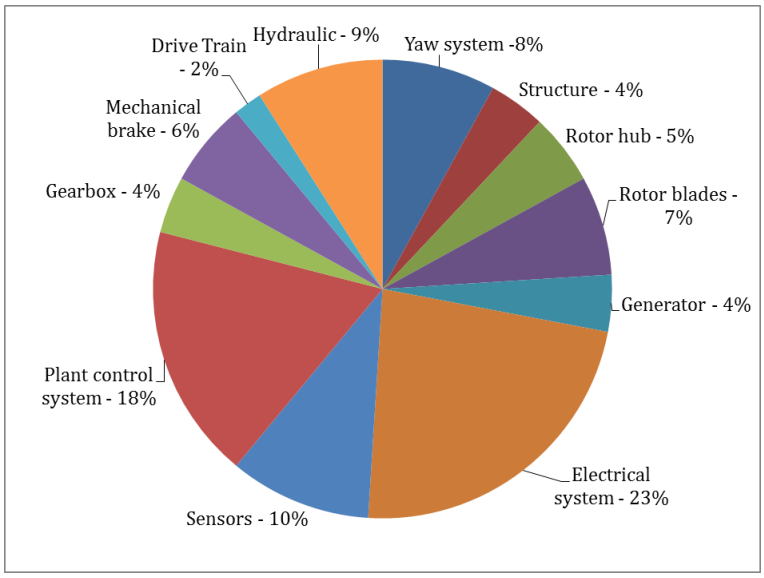
A number of researchers have conducted studies on the failure modes of turbines' sub-assemblies and downtime from various sources. For instance, Ribrant *et al*. [14] investigated Swedish wind power plants from 1997-2005 and determined the percentage number of failures between 2000 and 2004 (Fig 2.4 a). Results show that the electrical system, sensors, blades/pitch and hydraulics have the highest failure rates. The downtime for the individual failure mode is also calculated in the study, where gears, the control system and the electrical system have the highest amount of downtime. Moreover, the institut fur solare energieversornungstechnik (ISET) conducted a survey on the reliability of 1500 WTs in operation in Germany over a period of 15 years [15, 16]. In this study, the sub-assemblies that had the highest failure rates proved to be the electrical system, the control system, the hydraulics and the sensors, as shown in Fig 2.4 b. It can be seen that the sub-assemblies with the highest failure rates from the two individual studies are similar. Furthermore, for each failure mode in the ISET study, the generator, gears and drive train have the longest downtime. Again, this is very similar to the findings from Ribrant *et al*.

Furthermore, Tavner *et al*. [18] have studied turbines' reliability, availability and maintenance using data from different sources including the Windstats surveys in Denmark and Germany (WSDK and WSD) on wind turbines over a period of 25 years, the Landwirtschaftskammer (LWK) survey and the scientific measurement and evaluation programme (WMEP) survey on turbines in over 15 years of operation in Germany. All these surveys contain failure data from turbines with fixed and variable speeds as well as geared or direct drive configurations. Figure 2.5 shows an example of the failure rates and downtime of turbine sub-assembly using LWK and WMEP data [17, 18]. On the left-hand side of the plot, from point zero, are the failure rates per turbine per year for the two datasets, and the downtimes per failure (in days) for individual sub-system are shown on the right-hand side from zero. It can be seen that the electrical system has the

highest failure rate, while the gearbox has the highest downtime even though its failure rate is relatively low.



a)



b)

Figure 2.4 - Percentage number of failures of WT sub-assemblies. a) Swedish wind power plants 2000-2004 [8] b) ISET (Institut fur Solare Energieversornungstechnik) report [15, 16]

The number of failure rates and the downtime differ depending on the turbine type and configuration. The study shows that for a turbine, there is an overall failure rate of 1-3 failures per year with a downtime above 24 hours is common for onshore turbine. This number increases with turbine size. Moreover, a general trend of reduction in failure rates with time is found, indicating

advancement in WT technologies and the production of more reliable wind turbines [18]. It is concluded that the sub-assemblies of WT with the highest failure rates, in descending order, are the rotor pitch system, the power electronics converter, the electrical system, the rotor blades, the generator, the hydraulics and the gearbox. Apart from the yearly failure rate, the average downtimes of failure also have a tremendous impact on maintenance. A different ranking of failure downtime is obtained for these sub-assemblies, which is listed here in descending order: gearbox, generator, rotor blades, pitch system, power electronics converter, electrical system and hydraulics.

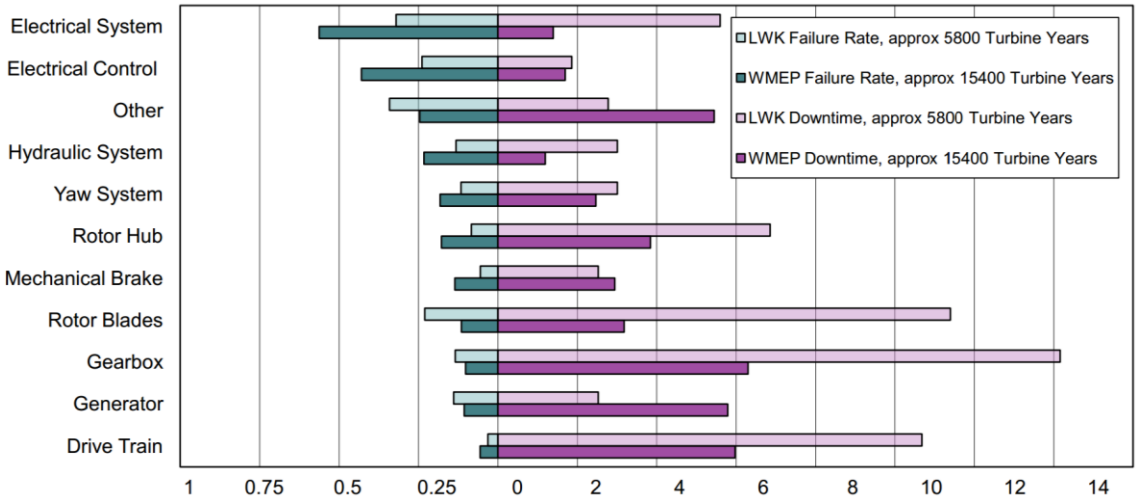


Figure 2.5 - WT sub-assembly reliability from the LWK and WMEP surveys [18]

It can be noticed that the sub-assemblies with highest downtime are the principal component of the WTS, such as the gearbox and generator. Any replacement or maintenance of these components will require a much complex procedure and longer time. Furthermore, a recent study carried out by Faulstich *et al*. [20] has shown that for onshore WTs, 5% of the total downtime is caused by 75% of the faults; the remaining 25% of the faults are responsible for 95% of the downtime. Typical faults and their consequences for major components from the above mentioned list are discussed in detail in the following sections.

2.2.1 Turbine rotor failures

The common failures of a pitch system include mechanical issues, such as pitch bearing failure, oil leakage or pump failure for a hydraulic pitch system; while motor failure, motor converter failure and battery failure are categorised as an

electrical pitch system. Any failure of the pitch system will affect the performance of the WT, causing asymmetrical aerodynamics forces. At most, it can damage principal components of the turbine as the turbine will operate above the rated value due to a high input wind profile [21]. The rotor blade is the fundamental component in the WTS, and common failure modes include mechanical imbalance, aerodynamic imbalance, lamination de-bonding, a crack due to lightning and icing [69]. These faults have direct effects on the blades' aerodynamics and can cause excessive stresses on other components, thus significantly reducing the profitability of WTs. Fatigue is considered to be the major contributor for the blade failure, as unevenly distributed forces are applied on the blade and will result in excessive stress, which in turn causes microscopic cracks in the material. A crack will then propagate with time and lead to further deterioration of the composite material used for the blade. Moreover, lamination de-bonding and broken fibres are all caused by excessive tension or a compression load. Finally, extreme environmental conditions such as low temperatures, lightning and debris present in the wind are also factors that are responsible for certain failure modes. Such extreme conditions have the potential to alter the aerodynamic behaviour of the blade, increase wear or cause exposure to conditions beyond the designed conditions [22].

2.2.2 Drive train failures

The main bearing and shaft are responsible for the main power transmission between the turbine rotor and generator. A failure of the bearing will induce radial movement, torque variation and vibration. Common faults include bearing lubrication failure, bearing roller wear or corrosion, shaft misalignment, shaft bending as well as other effects. Notably, manufacturing errors, installation defects, improper maintenance, fatigue and extreme loads are the dominant causes for the failure. A gearbox is used for most of the WTS configurations and, as described earlier, it is the major contributor to the turbine's downtime. Common failure modes are the gearbox bearing fault and gear teeth fault. Surveys have shown that the root cause of gearbox failure is due to the rapid change of torque from the stochastic wind profile, which will create an uneven load for the bearing and cause misalignment of the gear teeth. These phenomena

will accelerate the wear of the components. Moreover, other causes of bearing and gear teeth failure are the elevated operating temperature and excessive contamination of the cooling lubricant due to failure of the gearbox cooling system. Any fault from the gearbox can result in an abnormal input to the generator, reducing efficiency or damaging the generator in extreme cases [18, 23].

2.2.3 Electrical generation failures

Previous studies have shown that the failure of the generator due to bearing, stator and rotor are 40%, 38% and 10%, respectively [49]. For the stator and rotor failures, the inter-turn short circuit fault is the most widespread fault. It can be caused by the degradation of winding insulation due to electrical and thermal stresses. When this happens, the motor winding can be partially or completely shorted; as a result, the output of the generator will be reduced due to winding asymmetry. In the case of rotor windings, deformation of the winding due to the exposure of high centrifugal forces is also a factor for the failure. This can have significant effects on the generator such as increased winding temperature and vibration and deformed voltage waveform. Moreover, any failure in the generator bearing can result in excessive vibration and rotor misalignment, which in turn cause eccentricity. The eccentricity of a motor occurs when the air-gap between the stator and rotor are non-uniform, thus leading to an unbalanced magnetic flux and effects on the output current. As stated earlier, the operating condition of the WT is non-stationary; hence, the bearing is subjected to a non-uniform load. Additionally, misalignment has a significant impact on the performance of the bearing. Under this condition, the bearing can suffer from abnormal stresses which wear the bearing faster and in-turn increase the mechanical vibration, consequently worsening the degradation to a greater extent. As a generator is responsible for the power generation, any failures of the components depending on the location and the severity can reduce the quality of power generated, damage the generator or even cause catastrophic damage to the turbine [24].

The failures of the power electronics not only have an effect on the turbine itself but can also affect the grid due to the poor quality of electrical outputs. Common failure modes are over-temperature, dc-link capacitor ageing, switching

components failure due to bond wire lift off caused by temperature swinging (e.g. insulated-gate bipolar transistor (IGBT) and thyristor module). Soldering fatigue is also a main contributor to the converter failure. Because the converter suffers from thermal cycles, a crack can occur at the edge area of the soldering. As the degradation propagates, there is less area for the heat to dissipate, thus increasing the thermal resistance between the chip and the heat sink and overheating the component. Moreover, electrical overstress, electrostatistic discharge and latching-up of the switching components are also possible causes of converter failure. These may cause a short circuit or thermally damage the module. Lastly, the DC chopper and crowbar can also suffer from fuse failure [24, 62].

A number of failure modes and the effects of the major subsystems of a WT are listed. The severity of these failures modes varies, where with critical failure, it can result in WT malfunction, loss of power generation or broken sub-assemblies; On the other hand, for minor failures such as oil leakage and a loose connection, it can be less significant. However, these failures can develop into critical failures if proper maintenance is not performed. Studies on the root causes of these failure modes can be found in several literatures. Table 2.1 lists all the possible root causes for the failure modes of a WT, and it is categorised into environmental, structural, electrical and wear [18, 26]. Any of the failure modes as described earlier can be traced to these root causes.

Environmental	Structural	Electrical	Wear
Strong turbulent wind	Material defects	Overload	Aging
Icing	Installation defects	Insulation failure	Corrosion
Lightning	Manufacturing defects	Connection fault	Fatigue
Debris in the wind	Mechanical overload	Software failure	Thermal overload
-	Design defect	Calibration error	-
-	Maintenance error	Maintenance error	-

Table 2.1 - List of root causes for wind turbine system failure modes [18, 26]

2.3 Reliability of WT system

To effectively carry out a reliability analysis for WTs, a quantitative measure needs to be developed. Generally, the failure intensity function is used. Assume N_0 number of identical components to be examined. At any time t, there are Ns number of fault free components and N_f number of components with faults. Their respective reliability function/cumulative distribution function $R(t)$ and $Q(t)$ are [18, 19]:

$$R(t) = \frac{N_s(t)}{N_0} \tag{2-1}$$

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

$$R(t) = 1 - Q(t) \tag{2-3}$$

and the failure intensity function $f(t)$ and the normalised failure intensity function to the number of fault free components $\lambda(t)$ are:

$$f(t) = \frac{1}{N_0}\left(\frac{dN_f(t)}{dt}\right) \tag{2-4}$$

$$\lambda(t) = \frac{1}{R(t)}\left(\frac{dR(t)}{dt}\right) \tag{2-5}$$

For any repairable system, its failure intensity can be modelled with the bathtub curve as shown in Figure 2.6 a. The bathtub curve for WTs' failure intensity represents a discrete stochastic process, which can be described mathematically using the PLP (Power Law Process) [19]:

$$\lambda(t) = \frac{\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1} \tag{2-6}$$

where, β is a dimensionless parameter determining the shape of the curve, θ is the scaling parameter with a unit of time.

During each stage of the bathtub with respect to time, the failure intensity function can be modelled with equation (2-6) using different β as shown in Figure 2.6 b. During the early failure, it represents the infant mortality failure when new components are introduced and the failure rate decreases with time, β <1. Then it moves to the intrinsic failure stage, where random faults occurs, but it

is relatively constant during this period and represents the product's useful life with β=1. Finally, at deterioration or the wear out stage, the component has reached the end of its life cycle, with a fast growth of failure rate and β >1.

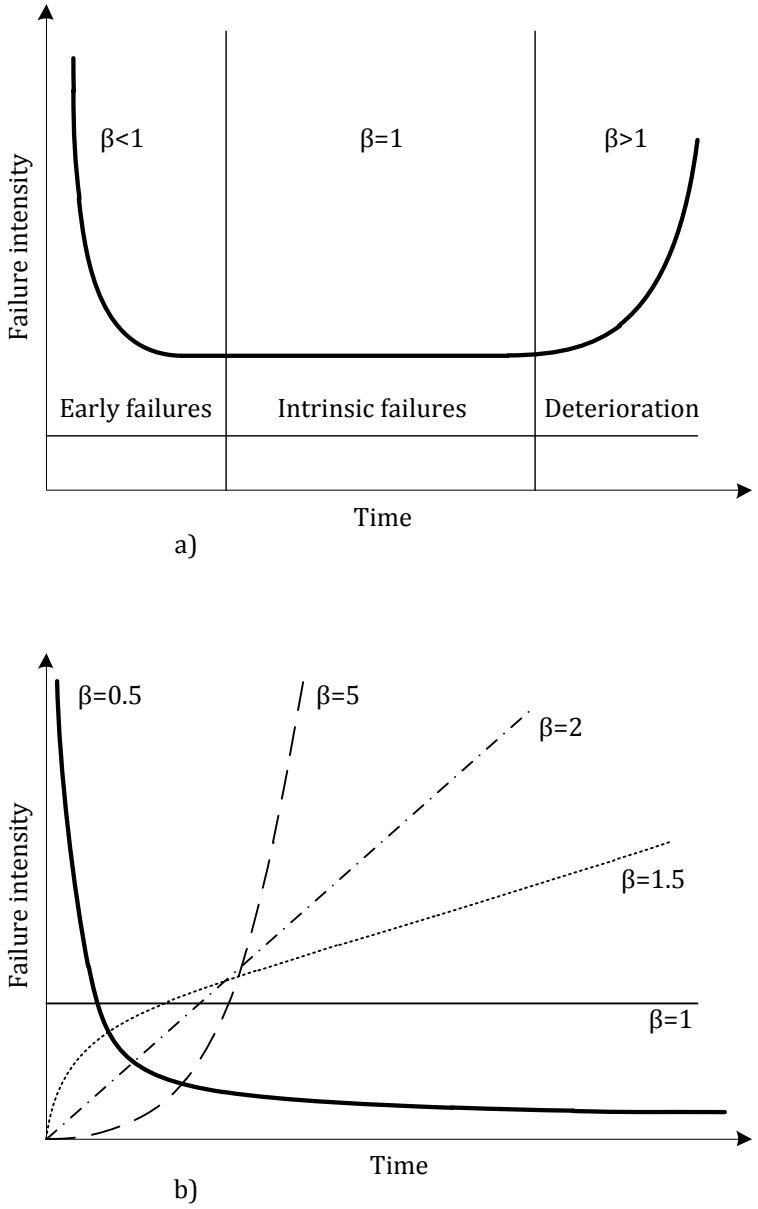


Figure 2.6 - Bathtub curve of repairable components' failure intensity with time. a) bathtub curve at different stages in time; b) modelled bathtub curve with different shape parameter β

The model can be implemented to sub-systems. The reliability block diagram method is then adopted to connect individual sub-systems (both in parallel and in series) to represent the whole system's functionality.

To actually enhance the reliability of WTS, actions can be carried out at different time periods: during the designing of the wind turbine, at the commissioning stage of the turbine and throughout the operation and maintenance of the turbine.

2.3.1 Failure mode and effective analysis

In his book on WT reliability, Tavner discussed several ways to improve reliability during the designing stage, such as adopting new design concepts and optimising wind farm configurations utilising FMEA (Failure Mode and Effects Analysis). Both Spinato *et al*. [10] and Polinder *et al*. [25] have compared turbines with the direct drives and geared generators from different sources. Spinato has shown that the reliability of the direct drive concept is not necessarily higher than the geared concept. This is mainly due to the fact that by eliminating the gearbox, the aggregated failure rate of the generator and converter from the direct drive concept has increased dramatically, as the generator used in this configuration is much more complex and the converters are fully scaled. However, the direct drive WT has better availability than the geared drive WT due to the fact that downtime relating the gearbox failures are substantially lower than electrical [25].

FMEA and FMECA (Failure Mode Effects and Criticality Analysis) are common bottom-up proactive failure analysis tools used during the designing of a product. Though the FMEA was designed to identify risks within a system, it can also be useful in determining components with higher failure rate, thus improve reliability. Generally, it evaluates a system by calculating the RPN (Risk Priority Number) for each component. The RPN is the product of occurrence, which is the frequency of root causes of a particular failure mode to occur; severity, which is the magnitude of the consequence of the failure; and detectability, which is the possibility for the fault to be detected before failure. FMECA is the extension of FMEA which also considers the relationship between the probability of failure mode and severity.

The application of FMEA and FMECA on WT can be found in several literatures. In [26], the researcher managed to show there is a relationship between the occurrence and mean time between failure, and between severity and mean time

to repair. Moreover, both Shafiee *et al*. [28] and Kahrobaee *et al*. [27] have developed FMEA for WTs, where the failure severity is dependent on both the consequences of failure and the cost of damage to the equipment.

2.3.2 Enhancement of reliability through testing and commissioning

By carrying out sufficient testing and commissioning, the reliability of the system can be improved; the technique such as accelerated life testing is commonly adopted. System and sub-assemblies are tested under a controlled environment to simulate ageing, and the failure rates for individual components are collected. These data can provide valuable information in revising the design of WTs through FMEA/FMECA analysis. The accelerated life testing can be time consuming and costly for certain sub-assemblies of the WT. In the absence of such a technique, components or sub-assemblies of the WT are constantly tested throughout each stage from prototype to production to commissioning in order to raise the reliability of WT [18].

2.3.3 Maintenance of wind turbine

The improvement of reliability through designing and testing are carried out before commissioning of the WT. Methods such as FMEA require knowledge of the failure mode of the system and its sub-systems, as well as the root causes. Root cause analysis is a top-down process, where it tries to identify the main reason for specific failure mode, such as the main shaft failure as shown in Figure 2.7. However, the reliability can also be enhanced through appropriate monitoring of the turbine and effective maintenance. This process is regarded as a bottom-up process, as the monitoring system identifies anomalies through data collected and alerts the operator. This will allow the operator to carry out maintenance with provisional spare parts to reduce downtime and avoid catastrophic failure. This procedure leads to the discussion of condition monitoring system as detailed in the later sections.

2.4 Wind turbine condition monitoring system

A typical monitoring system setup of modern wind turbines is shown in Figure 2.8. It comprises of a SCADA system and a CMS, which serve different purposes. Both systems are reviewed in the following sections.

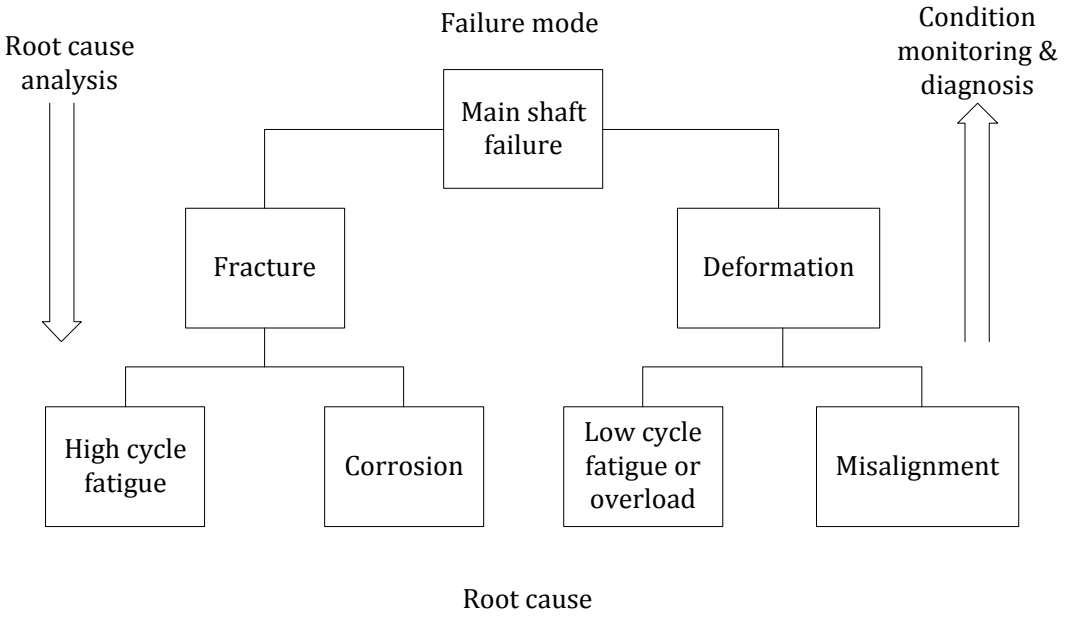


Figure 2.7 - Example of root cause analysis and condition monitoring & diagnosis for WT main shaft failure

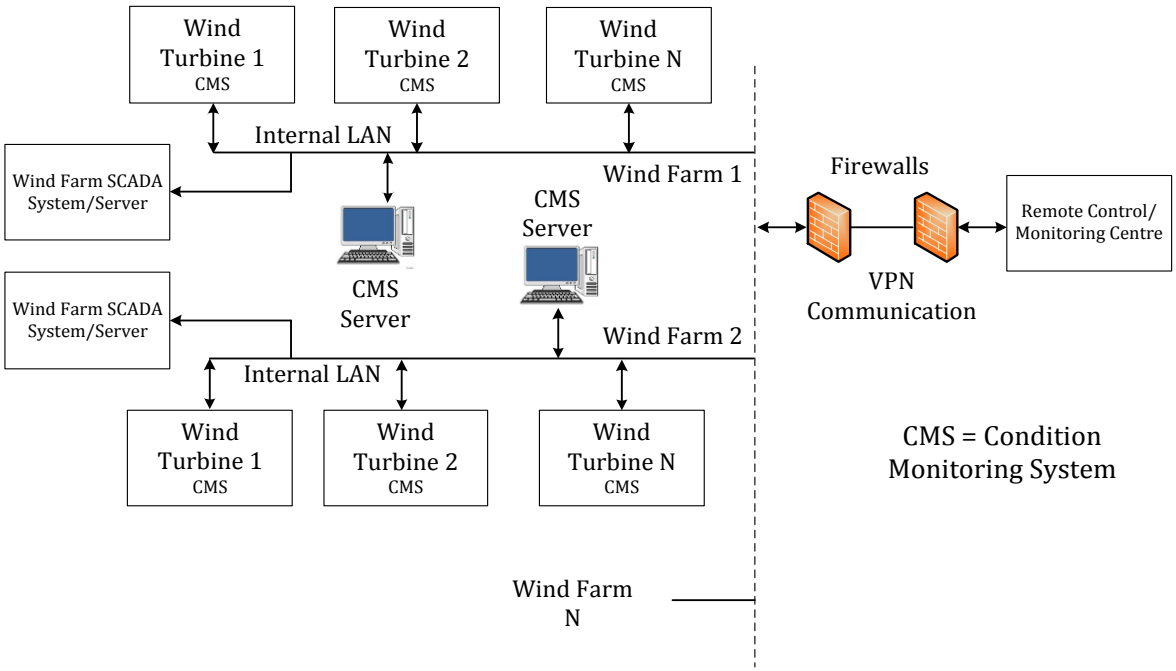


Figure 2.8 - Wind farm monitoring system with SCADA and CMS

2.4.1 SCADA system

The SCADA system was first introduced in the oil and gas industry. It was then adopted by the power generation industry in 1985 [18]. As its name implies, it

allows communication between the turbine and the remote control and monitoring centre. The channels used permit dual way communication, where the remote centre has direct control of the wind turbines and it can receive signals acquired from sensors. However, for wind turbine applications, the system is used more for monitoring than control. The turbine control is mainly performed through an on-board controller; however, operators have some degree of override authority.

The SCADA system usually has a sampling rate of <0.002 Hz, i.e. data is sampled every 5-10 minutes. However, the transient maximum, minimum, mean and standard deviation for some signals are also included, such as wind speed and power output. The measured signals vary as a result of different manufacturers; nonetheless, critical parameters such as vibration, temperature and electrical are always required. Moreover, systems also maintain an alarm and event log, registering any failures and vital operations. The monitoring data is usually stored at the central database or is transmitted to a centralised remote centre, where management and control of the WTs are performed. Although fast transient behaviour of the monitoring signals is not captured by the SCADA system due to the low sampling rates, the fact that data is already collected can be beneficial, which leads to the great potential of long-term fault diagnosis and prognosis utilising SCADA data [29].

2.4.2 Structural health monitoring

The structural health monitoring system is dedicated to ensuring the wind turbine's structural components, such as the tower and blades, are in healthy condition. The importance of SHM is obvious, since the structure of the wind turbine provides the integrity of the system. As structural damages are usually a long-term process, the data sampling rate for the SHM system is typically low <5Hz [18]. Methods adopted in SHM are non-destructive, and many manufacturers have built-in sensor networks in the turbine structure for data collection. Commonly used techniques include thermal imaging, acoustic emission and ultrasonic methods [69].

2.4.3 Process of condition monitoring system

The condition monitoring system is aimed to determine the health condition of the wind turbine based on the signals collected continuously from sensors and predict incipient fault, together with fault identification. It has a much higher sampling frequency than SCADA data, which ranges from 1 Hz to 20 kHz [18]. This is because these data are reserved for diagnostic purposes; the transient behaviour of a certain fault can occur in milliseconds, i.e. a generator wind short circuit fault. A low sampling rate will not be able to capture these features, thus not being able to detect the faults accurately. Moreover, Hameed *et al*. [30, 50] have discussed the benefits of implementing condition monitoring-based maintenance, which has advantages over traditional corrective and preventive maintenance strategies.

The process of a CMS is shown in Figure 2.9. It mainly consists of three steps: data collection, data processing and decision making (diagnosis, prognosis and maintenance scheduling) [31]. To achieve an effective CMS, accurate data collection from sensors and precise diagnostic algorithms are needed. The former requires knowledge in sensors' characteristics, performance, reliability, selection and optimal placement, while the latter focuses on the actual detection algorithms for individual failure mode, especially for sub-assemblies having higher failure intensity as discussed earlier.

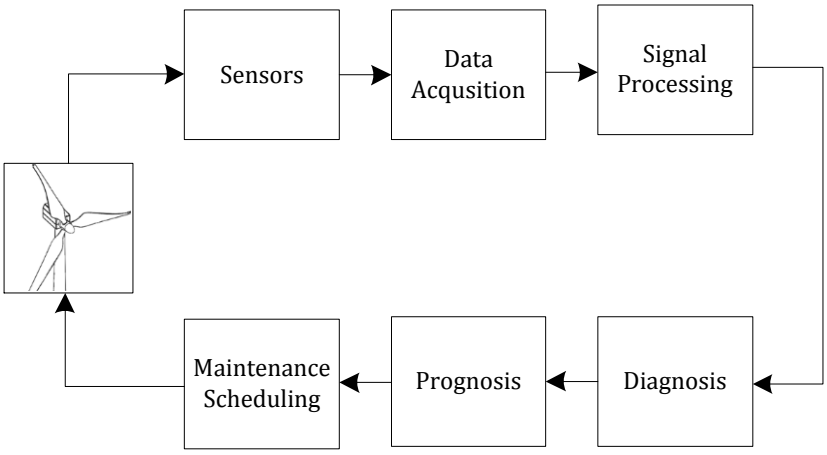


Figure 2.9 - Process chart of condition monitoring and fault detection system for a wind turbine

2.5 Existing techniques for wind turbine condition monitoring

Essentially, condition monitoring is the action of processing information collected from sensors in a suitable format through the transformation or methods of some sort, such that a fault signature can be identified from a normal operation [48]. Typical processing techniques are given in different modes and can be categorised into time, frequency, (TF) Time-Frequency domain analysis and statistical and artificial intelligence techniques.

2.5.1 Statistical methods

Due to the high number of monitoring points and the size of the signals, it is highly common to use a statistical tool in WT fault detection. Based on the historical data of the WT, fault features or patterns are extracted from data to establish models for future FD. Several techniques have been adopted by researchers, including probabilistic models such as the Bayesian network [42] and estimation models such as maximum likelihood [43]. Clustering and classification based on statistical variances [33] and information entropy [44] have all been adopted by researchers. Moreover, the statistical process control method from quality control is also commonly used in wind turbine applications. Detailed applications of these techniques are discussed in later sections.

2.5.2 Time domain analysis

It might be advantageous to see the physical variation and transient change in time, where fault features may be obvious. Most of the descriptive statistic methods are carried out in time domain, such as the mean, median, maximum, standard deviation and skewness and kurtosis, which are high order statistics. Another technique commonly used in time domain is trend analysis [32]; it includes a synchronous average where the mean values of the subset of original data at a predefined window are calculated.

$$\bar{S}(t) = \frac{1}{N}\sum_{n=0}^{N-1} s(t+nT), 0 \le t \le T \tag{2-7}$$

Where $\bar{s}(t)$ is the result signal, N is the number of samples within the selected averaging period T, and s is the original time series data. The method is aimed to reduce noise level and effectively disclose a signal of interest. An extension of this

approach includes an autoregressive moving average. The method fits the original time series data to a parametric model where features can be extracted. Moreover, other methods such as correlation analysis [33] and principal component analysis (the focus of this thesis) are all based on time domain data.

2.5.3 Frequency domain analysis

In contrast, spectral analysis reveals the variation of signals with respect to frequency. It is helpful in the identification and separation of frequency components that are relevant. However, it usually requires an additional transformation from time domain data. The most common and conventional transformation is the fast Fourier transform (FFT) [34-36]. It is not unusual for information in a frequency domain signal to be contained more saliently than in a time domain. Moreover, high order spectral analysis not only considers frequency variation of the signal but also includes the phase information of the harmonic components and therefore improves the signal to noise ratio [37, 38]. Methods such as bispectrum analysis are also often found in the application of wind turbine condition monitoring.

2.5.4 Time-frequency domain analysis

Signals from wind turbines are considered as highly dynamic and non-stationary due to the variable load and speed. This limits the application of frequency domain techniques such as Fourier transform due to the fact that it is assumed for infinite length or periodic signals. Also, localisation in the time domain is not possible. Time-frequency domain analysis resolves this issue. Instead of representing the signal in one-dimension, signals are visualised in a two-dimensional plane with respect to both time and frequency. Common time-frequency transforms include STFT (short-time Fourier transform), wavelet analysis and HHT (Hilbert-Huang transform) [39-41].

2.5.5 Artificial intelligence methods

Implementations of artificial intelligence-methods in WT FD have also shown increasing trend. It is aimed to solve the problem by introducing human reasoning or mimicking the functionality of the human brain. Thus, solving problems which are non-parametric and non-linear is possible, as in ANN

(artificial neural network) [45]. Also, it is possible to realise continuous monitoring and make accurate decisions utilising the expert system [46]. Moreover, fuzzy logic allows the system to make judgements beyond the binary right or wrong [47].

2.5.6 Application of condition monitoring techniques in wind turbine

Based on the techniques described earlier, researchers have developed many fault detection algorithms during the recent decades. These methodologies usually adopt multiple techniques or a combination of these techniques to achieve their target for different failure modes. In later sections, common methodologies employed in developing CMS for wind turbine applications are described. The review is carried out based on the turbine's taxonomy described earlier, and it is focused on components that have high failure rates.

2.5.6.1 Generator

A thorough review on the condition monitoring of rotating machines has been given in the reference [24]. Different types of monitoring techniques are revised, including temperature, chemical, vibration, electrical and artificial intelligence. Each of these techniques is implemented to different failure modes and mechanisms, including ageing due to temperature, electricity, mechanical forces and environmental factors; insulation failures, stator and rotor windings failures; and mechanical failures concerning the bearing and cooling system. However, due to the non-stationary and stochastic behaviour of wind, only some aspects of these techniques can be adopted and modified for wind turbine application.

It has been found that major failures of wind turbine generators are related to the stator, rotor and bearing with percentages of 38%, 10% and 40%, respectively [34]. Then, the stator unbalance, rotor unbalance and turn to turn fault of a wind turbine with a DFIG configuration are studied. The CSA (current signature analysis) of time domain data was used to extract the trend change of the current amplitude. Thereafter, FFT was applied to identify components that are related to the faults under investigation. Data collected from the experimental platform were used to validate the proposed methods. Results showed the unbalance of the stator and rotor and turn-to-turn fault can be identified by the line current.

Watson [55] used a wavelet and CSA to detect a generator fault from an instantaneous power signal. It was concluded by the author that the rotor eccentricity in DFIG can be identified by the component from a wavelet at twice the slip frequency divided by pole pairs. This rotor eccentricity is mostly caused by bearing wear; hence, it can be used as a measure for bearing fault prediction. The reason for using a wavelet over traditional FFT is that the variable frequency characteristic of DFIG makes FFT infeasible in tracking the measure proposed. Moreover, literature [39] also presented the application of a wavelet to detect the short winding faults of a synchronous machine. The author proved there is a relationship between a short coil and the ratio of mechanical speed to reactance. Furthermore, short winding and rotor mass unbalance are investigated by Wilkinson [56]. The short time power spectral density method was applied to the mechanical rotational speed to detect faults. Finally, Saidi [37] proposed the application of higher order spectral techniques to detect the broken-bar faults of induction machines. Results have shown the method can effectively distinguish between normal and faulty conditions. Further analysis of the robustness of the proposed techniques was carried out through independent Monte-Carlo experiments by the author.

Apart from CSA and spectrum methods, Bennouna [57-59] has developed model dependent approaches for determining physical anomalies in generators. The data reconciliation concept is adopted to estimate the stator/rotor current and voltage in the *dq0* axes. The errors between estimation and actual measurement are calculated dynamically. Based on the error and classical threshold test, an abnormality can be identified. Both linear and nonlinear systems are considered, which allows the application of the method to variable speed wind turbines. Finally, despite the ability of fault detection, this technique was capable of determining signal signatures relating to the load as well.

The information entropy-based method was proposed for the fault detection of a generator unbalanced mass [44]. The information entropy is the measure of randomness of a signal. The frequency spectra of WT signals during the development of fault generally produce the growth of characteristic frequencies relating to fault, increase the number of harmonics and create a more substantial

noise. These changes of the frequency spectra will result in the changing of its information entropy. Yang [44] has implemented this concept to test the unbalance of a generator shaft. Results show a clear relationship between a fault and its severity with information entropy.

In addition, the application of artificial intelligence methods in fault detection is proposed by [60]. The author established a data driven fuzzy model for the wind turbine fault diagnosis system. By employing the Takagi-Sugeno model, unbiased parameters of the model are estimated based on errors in the variable identification strategy. The author then applied clusters of rules to determine the output from parameterised functions (estimators). For different rules, only the parameters are changed. The signals used to create the model are selected based on measures from FMEA of failure modes. Different strategies of using banks of fuzzy estimators are also proposed, including multiple inputs and multiple outputs, and multiple inputs and a single output. Thereafter, threshold logic is used according to the residual mean and standard deviation. Data is sampled from a simulated wind turbine benchmark at 100 Hz. Several pitch faults are considered, such as the fixed pitch angle or scaling error. The proposed method is then compared with several other methods, including the Gaussian kernel and estimation-based solution, up-down counters, the combined observer, the Kalman filter and the general fault model. The measures used to compare methods are the false alarm rate, the missed fault rate, the true detection/isolation rate and the mean detection isolation delay. Results show that the technique is adequate for non-linear, non-stationary FD of a wind turbine. Moreover, the simplified fuzzy model is good for an online monitoring application. However, only simulation data is used in the validation. Moreover, Cross [61] has also proposed the state dependent and neural network model for the generator fault detection of wind turbines using SCADA data. The fuzzy inference threshold method is then proposed for identifying faults at three different warning levels in ascending order. Results have shown that an incipient level 1 warning is successfully detected 20 hours ahead [61].

2.5.6.2 Power electronics converter

Bin Lu carried out an in-depth review of fault detection and protection methods for general use of the IGBT power inverters [63]. He investigated a gate drive open circuit fault, a transistor short circuit fault and a gate misfiring. Out of the three faults, the misfiring fault had the fewest immediate consequences, but it was determined that it could damage the inverter in the long term, hence causing catastrophic failure. In regard to the transistor short circuit fault, it will lead to IGBT rupture, which is caused by an inaccurate gate voltage or component intrinsic failure. The gate drive open circuit fault is induced by wire de-bonding due to fast thermal change. A study that compared the multiple detection techniques was also carried out, which included the use of Park's Vector in identifying a gate drive open circuit, where the fault is related to the magnitude and shape of the space vector trajectory calculated from the current in the *dq0* axes. Notably, there is a change of the virtual capacitance in IGBT under a faulty condition, which in turn affects the gate voltage, and the transistor short circuit fault can be detected by monitoring the gate voltage. However, complex protection circuitry is needed, as is the use of a pattern recognition method for identifying a misfiring fault through a set of fuzzy rules. Again, some of the methods reviewed by Lu may not be applicable for wind turbine power converters due to the non-stationary and stochastic behaviour of the wind.

Moreover, the application of fault detection in wind turbine power electronics can be found in [64-68]. Ko *et al*. [64, 65] carried out a faulty switching detection of three-parallel voltage-source converter for wind turbine. The authors utilised three phase current vector patterns to capture features in a *dq*-reference frame. A fuzzy logic-based threshold is set for identifying an abnormality. Open fault switching and short fault switching are studied. The former can induce noise and vibration in the system, and the latter can cause an over-current and the demagnetisation of a synchronous generator. The method is beneficial as no additional sensors or excessive computation is needed. Further studies of the detection of simultaneous faulty switching were carried out. A neural network was adopted to recognise and classify these patterns. The authors also proposed fault tolerant procedures by introducing additional triacs to isolate the faulty

switch device, hence reducing the system structure and corresponding power rating until the fault is fixed.

As described in earlier sections, power converters for a PMSG wind turbine with a synchronous generator generally have a full power rating. However, in the case of a DFIG system, the converter only handles a portion of the rated power. Additionally, because the converter is bi-directional, the fault behaviour and detection may be different from PMSG. You *et al*. [66] have investigated short circuits due to high voltage drops and heat dissipation when operating at a full rating for converters in DFIG. The possible operation conditions from the generator side of the converter proved to be normal, mono-tube short circuit, mono-tube turnoff, crossover short circuit, crossover turnoff, paracentric short circuit, ipsilateral full short circuit and ipsilateral full turn off. The actual fault can be any combination or permutation of these intrinsic conditions. The authors found that each of these faults distorted the phase voltage differently. With the use of a self-organising feature map neural network, which is a type of NN model with unsupervised learning, all the fault patterns are learnt and then applied to recognise and classify the faults. The method needs no supervision and is capable of self-learning. Even with the presence of noise, results showed a 98% correct fault recognition rate. The method is adaptive and able to expand its knowledge by adding a new fault pattern at any time. Similarly, the detection of a converter fault based on a neural network using the converter output voltage waveform is conducted in reference [67]. Authors carried out detection using spectrum magnitude, harmonics of the voltage signal and the phase angle harmonics in the direct-axis reference frame. Finally, Duan *et al*. [68] proposed a sample-based detection method for an open circuit fault based on voltage and current signals. He utilised the phenomenon that occurs during the open-circuit, when the free-wheeling diodes will supress the phase current and force it to zero. With an appropriate threshold setting, the number of samples of the measurements at normal and faulty conditions is different.

2.5.6.3 Gearbox and bearing

Turbine's gearbox and bearing has the highest downtime due to the complexity of repair work needed. Researchers have proposed FD algorithms to predict an

early warning for it, with different types of measurement and processing techniques having been adopted. The spectrum analysis of vibrational data is most commonly used. In literature [33], Zhang *et al*. analysed gearbox faults using jerk data from multiple accelerometers installed at different locations on an experimental testing platform. Correlation analysis is used to establish a baseline comparing data from normal and faulty operations at different speeds and torque settings in order to demonstrate the existence of faults. The authors further studied this phenomena using clustering, where data is categorised into groups with respect to speed. The Euclidian distance between these clusters can indicate any presence of failure within a gearbox. The same method is applied to the power spectrum of the vibration data, and similar results are obtained. In both cases, damages at the high-speed stage of the gearbox were identified. A data driven model was further established by using seven statistical methods, including Neural Networks (NN), Neural Network Ensemble (NNE), Boosting Tree Regression (BTR), Random Forest Regression, Classification and Regression Tree, Support Vector Machine and K Nearest Neighbor. Statistical control charts are used to detect the faults. The model is used to monitor vibration excitement on the high speed side of a gearbox. BTR yields the best accuracy and performance. The severity of the vibration is quantified, and control charts are implemented to identify abnormality. However, there was some inaccuracy because environmental factors were not considered.

The gear eccentricity of a gearbox with a synchronous generator was studied in [39]. Yang *et al*. used a wavelet-based technique; a thresholding wavelet coefficient was developed in order to reduce noise. Two thresholding approaches are proposed in the literature. Based on the data collected from an experimental test rig, the low-frequency amplitude of the modulated vibration signal was obtained. The modulated signal has a period of approximately 0.3s. The author demonstrated this phenomenon is related to gear-eccentricity. Moreover, the diagnosis of mechanical faults from electrical properties is also presented. Both of the wavelet maps of the generator current and gearbox vibration demonstrated a characteristic frequency of 300 Hz. This result is due to the strong linear relationship between the drivetrain torque and generator current.

Yang *et al*. [51] presented a gearbox fault diagnosis using the Artificial Neural Network (ANN) with a vibration signal. Wavelet analysis was employed to decompose the vibration signal into levels of interest. Decomposed levels were fed into the BP NN with four identifiable outputs: normal, gentle fault, fault and bad fault. The results showed promising fault detection ability, where with the excellent property of self-learning, adaptive and non-linearity of the methodology, it can have wider applications. Moreover, Huang *et al*. [52] also investigated the application of wavelet neural networks of a gearbox fault based on vibration signals. Furthermore, a statistical tool such as the Maximum Likelihood Estimation (MLE)-based fault detection of a gearbox using vibration data was carried out in [43]. Amplitude demodulation was used to remove the white noise from the signal, where the parameters were estimated by MLE. The method was tested with real bearing signals with an outer race fault at 12 kHz. It demonstrated that the method can reveal characteristic frequency and associated harmonics of the signal by adjusting the signal to a noise ratio, thus identifying the failure.

Other measurements such as temperature and acoustic waves are also considered in a gearbox fault diagnosis. Guo *et al*. [32] proposed a method of using Auto-associative Kernel Regression (AAKR) to build a normal operational model of gearbox temperature and then compare it to an actual fault, which requires the construction of a memory matrix. The moving window statistical method was adopted to identify the residual's mean and standard deviation as a measure of the performance. Power, wind speed, ambient temperature, gearbox temperature and one time delay of the gearbox temperature from SCADA data were used to build, validate and test the model, and the testing of faulty data was obtained through simulation. A threshold is defined as acting as a limit in determining a fault. Results show the model is capable of identifying rises in temperature and confirmed the SCADA alarm log. The author further commented that the method can be sensitive to the selection of the memory matrix, window size and the implementation of this concept to other sub-systems. Moreover, Acoustic Emission (AE) sensing is also used in gearbox fault detection. It is the measure of surface stress waves due to friction of the component surface. Lekou *et al*. [53] used AE in conjunction with vibration, temperature and rotational

speed to detect gearbox fault. However, a high sampling rate is necessary for an AE sensor to be effective. Thus, it can increase the cost of a condition monitoring system.

Finally, there is also research on utilising an expert system in the gearbox fault diagnosis. The SIMAP (Intelligent System for Predictive Maintenance) expert system was developed to perform online gearbox fault diagnosis and prognosis and to achieve optimum maintenance scheduling in reference [46]. It uses the gearbox bearing temperature, the gearbox oil temperature and the digital state of a cooler fan speed to continuously monitor the health condition of the gearbox. A prediction ANN model was established using normal operating data. This particular model is used to estimate the gearbox main bearing temperature, the thermal difference and the cooling oil temperature, and compares these elements to the online measurements, hence achieving fault detection. Expert fuzzy logic rules are adopted to perform a diagnostic test of the faults, *i.e.* if the gearbox main bearing temperature and thermal difference are all high and the cooling oil temperature is normal, then failure in the gearbox main bearing is certain. A further knowledge base was developed in order to accomplish CM-based maintenance scheduling. The performance of the autonomous maintenance scheduling was assessed by the change of health conditions before and after the maintenance. Moreover, Yang *et al*. [54] also used the expert system for identifying gearbox faults, such as a shaft fault and a bearing fault, based on a fault tree analysis in accordance to the physical structure of the gearbox. Fault tree analysis is a systematic top-down technique for identifying an undesirable state of a system using Boolean logic, and it is used in the development of a root cause analysis. Based on the fault tree model of the gearbox, a qualitative analysis of the basic events is carried out to obtain the diagnostic knowledge base for the expert system.

2.5.6.4 Rotor blades and others

Ciang *et al*. [69] conducted a review of the common techniques used in structural health monitoring of a wind turbine system. Methods in identifying failures from the turbine rotor, blade and tower were explained, and included the acoustic emission method, the thermal imaging method, the ultrasonic method, the

modal-based method, the fibre optics method, the laser Doppler vibrometer method, the strain memory alloy method, the x-radioscopy method and the eddy current method. A summary of the strength and weakness of each method was given. However, detection algorithms are required for these techniques to function, and different methodologies have been proposed by researchers as will be discussed in the following sections.

Rumsey *et al*. [71] presented a fault detection of turbine blades through experiments. A 9 metre glass-epoxy and a carbon-epoxy blade were used in a fatigue test with a cyclic load at 1.2 Hz. During 700k cycles, incipient damage started, where a fine gel coating crack was found on the high pressure side. The damage continued until a visible crack near the mid-point on the high pressure side parallel to the chord axis was found at 2.5 million cycles. During the testing, several non-destructive detection techniques were implemented and monitored the blade. These techniques included a spectrum analysis of acoustic emission signals (proposed by the author), the NASA wave propagation-based method using fibre optic sensors and a modal analysis of blade deflection measured from a tri-axial accelerometer. Results showed that the spectrum analysis of acoustic emission signals installed near a crack has a clear increase of its acoustic energy from 0.01 MeV/cycle (no fault) up to 30 MeV/cycle (peak damage), which indicates the building up of fatigue damage. Moreover, a trend representing the deterioration of the blade was also found by the two other methods. However, results also suggested that these two methods can be sensitive to several factors. First, the locations of the sensors installed have a great impact on the detection accuracy. Second, it was demonstrated that the variation in ambient temperature would have a direct impact on the blade dynamic, such as the modal analysis of blade deflection. Thirdly, the presence of high noise in the measurements might affect the results.

The identification of blade imbalance through bi-spectrum and bi-coherence can be found in reference [38]. The bi-spectrum shows the third order cumulant sequence of the original time series data, which can also reveal the phase coupling between frequency components. The author has implemented the method and successfully identified blade imbalance fault using blade tip

acceleration and power output signals. Moreover, time-frequency domain analysis is commonly adopted in blade fault detection. A wavelet-based fault detection for turbine blades can be found in [70]. The author proposed the use of the Meyer function as the base function for wavelet analysis. The symmetrical and orthogonal properties of the Meyer function enhance the computational speed of the wavelet analysis. This method is validated through several different scenarios such as normal operation, changing the sensor's location and a damaged blade. Through a 2-D grey scale of the time-frequency domain representation, the damage of the blade can be identified. However, the author did not offer a detailed discussion of the types of faults and types of signals that were used in his study. Furthermore, in [41], blade imbalance, misalignment and inner race bearing fault were identified through the Hilbert-Huang transform, which is another time-frequency transformation method. The author first incorporated a phase loop lock to estimate the angular speed of the shaft from the stator current. This is beneficial as a low number of input signals are needed and reduce the signal-to-noise ratio. He then transformed the estimated angular speed into a time-frequency domain using HHT. The sum of amplitudes that transformed signals at a specific frequency is used to predict a fault and its severity. Since HHT is designed for non-linear and non-stationary signals, it makes it the perfect solution in wind turbine application.

Apart from fault detection algorithms targeting a specific failure mode, there are many researches in the diagnosis of wind turbine from a system level. With the aid of artificial intelligence, it is possible to achieve an autonomous condition monitoring and fault detection system. In [45], the author has adopted and compared several techniques in fault detection and identification at the system level, including ANN (Artificial Neural Network), a standard classification and regression tree, a boosting tree algorithm and a support vector machine. SCADA data is used for validating the proposed algorithms, and information about turbine failures came from the alarm and event log from SCADA data and the S curve. For fault detection and identification, three levels of prediction are proposed: determination if a fault is present, then the categorisation of fault severity and fault identification and, finally, actual predictions that are made with 5-60 min ahead. Three measures are used to validate the results: accuracy

(number of correct prediction/total number of fault), sensitivity (number of correct prediction/number of fault) and specification (number of correct prediction of normal/total number of normal). The results showed that the standard classification and regression tree method has the best performance and is used as main prediction algorithm with industrial SCADA data.

A study of fault diagnosis of wind turbine gearbox lubrication and cooling system is carried out in [72]. The author has implemented the Lating nestling method (petri nets), which is a mathematical graphical modelling language that describes discrete events, characterising behaviour and interactions between sub processes. Based on this graphical model, rule bases are established to identify faults through a series of processes, sub-processes, events and locations. Measurements from sensors are used to determine the state of the system. Finally, individual and combination failure modes can be determined from the model. The author created the model and proved faults can be identified but did not discuss the validation process with data from a wind turbine in detail.

A model-based solution using qualitative physics is proposed by Echavarria *et al.* [73]. It determines the system's behaviour through qualitative characteristics of the physical system in time and quantifies this change. The behaviour is determined through physical phenomenon, system status and physical laws. Once the healthy model is established, any variation in the model indicates abnormality. Unlike other methods, historical data is needed, and this method is able to predict fault without the need of knowing it first. Moreover, sensor faults should also be able to be identified though this method. According to the author, for this method, minimal information is needed for the reasoning process, and no complex system of equations needs to be solved. The model is reusable, and robust. However, one major drawback is that it is time consuming to build the model, especially a comprehensive model representing a real wind turbine.

2.6 Instrumentation for condition monitoring system

Despite the fault detection algorithms reviewed in previous sections, data is the basis in realising a wind turbine condition monitoring system; therefore, data collection is an essential step. As shown in Figure 2.8, data collection consists of

the physical installation of sensors and digitises the analogue sensor outputs using an analogue-to-digital converter. The latter is a topic beyond this dissertation. There are three main issues regarding the physical sensors, which may have a great impact on the turbine condition monitoring and fault detection. These are sensor placement, sensor reliability and sensor /variable selection.

The physical placement of sensors has a direct effect on the data it collects, especially if the measurement is dedicated to a specific purpose. Information entropy, Fisher's information and mutual information are the common techniques adopted in the optimal sensor placement on structures [74]. Moreover, Kincaid *et al*. [75] proposed an experimental design method for solving sensor and actuator locations for complex truss structures built at the NASA-Langley Research Centre through a discrete D-optimal design method.

Sensor reliability may be different for the same measurement. Factors such as environmental conditions, manufacturing errors and meteorological conditions can all affect the performance of the sensors. Rogova and Nimier have classified the reliability of sensors into three categories, the first being the physical characteristics such as the resolution, linearity and accuracy of the sensors, which are determined by the physical design, the material used and the manufacturing process. The second category is the reliability of the time domain waveform measurement, where the data collected can be directly evaluated with the desired or expected value, e.g. temperature and pressure. The third category involves the reliability of high dimensional measurements such as thermal imaging, which is usually used for pattern recognition and classification problems [76]. Both Elouedi *et al*. [77] and Guo *et al*. [78] have carried out reliability studies of sensors in classification problems based on the belief function from evidence theory. A comparison of the discounting factor as a measure for sensor reliability was performed, where the discounting factor is a coefficient that weights the belief function between the trustworthy sensor and the unreliable sensor.

For a complex system with a large sensor network, it has theoretically been proven that multisensory data fusion should improve the performance in aspects such as classification and recognition [76]. However, this is true only when the

optimal set of measurements, which are related to the problem in hand, is obtained. There are many researches on optimal sensor selection for different applications. Commonly adopted methods have included information entropy and Fisher's information-based method. The application of the optimal selection for target localisation can be found in [79-81]; [82] has shown the application of information-based selection in sensor network managements. In addition, a model-based sensor selection strategy for health monitoring systems targeting to particular fault condition or diagnosis can be found in [83]. Finally, Hovland *et al*. [84] have suggested a stochastic dynamic programming method for solving the sensor selection of robotic systems in real time, where it is aimed to obtain a set of sensors dynamically which provides high quality and reliable measurements for a robotic controller. Unfortunately, there are limited researches on the optimal sensor selection for a wind turbine's condition monitoring system.

2.7 Summary and discussion

In this chapter, different wind turbines' architectures and their subsystems have been reviewed. The corresponding common failure modes and consequences have also been discussed. Additionally, the review covered the importance of reliability for a wind turbine system and the possible solutions for improving turbine reliability. The chapter then focused on the existing condition monitoring techniques and specifically addressed the components/subsystems that have the highest failure rate, as concluded earlier. Finally, the importance of reliable and accurate measurements for a condition monitoring system was discussed. The review has shown a substantial amount of literature on sensor selection in other applications. However, there are limited researches pertaining to variable selection for a wind turbine condition monitoring system; therefore, it is necessary to develop variable selection techniques for wind turbine condition monitoring and fault detection applications. In the next chapter, wind turbine modelling and simulation are presented.

Chapter 3. Simulation of wind turbine system under healthy and faulty conditions

This chapter describes the mathematical modelling and simulation of a wind turbine system. It begins with a mathematical analysis of a wind turbine system with a PMSG (permanent magnetic synchronous generator) and DFIG (doubly fed induction generator). Thereafter, simulation models of the respective turbine system are created in PSCAD/EMTDC and the results of the model under various normal operation conditions are discussed. Finally, simulations of the wind turbine under abnormal conditions are performed, where the AC-DC-AC converter DC-link capacitor ageing fault and grid line-ground short fault are studied.

3.1 Wind turbine modelling [85, 86]

The main objective in building a wind turbine simulation model is to acquire relevant data of the turbine under different operation conditions, which can be used to validate the proposed algorithms in later chapters. Moreover, the process provides a thorough understanding of a wind turbine's dynamic behaviour. The two most commonly used wind turbine designs are studied and simulated: a variable speed wind turbine with a PMSG (permanent magnetic synchronised generator) and pitch control, and a variable speed wind turbine with a DFIG (double fed induction generator).

3.1.1 Wind turbine with permanent magnet synchronous generator [97-100]

The subsystems of a PMSG wind turbine model are shown in Figure 3.1. The turbine mainly consists of the turbine rotor model, which includes: the wind turbine model, the speed and pitch controller, the mechanical drive train model, the permanent magnet synchronised generator model, the back-to-back power electronic converter model and controller and the electrical grid network model. Each model is discussed in detail in the following sections.

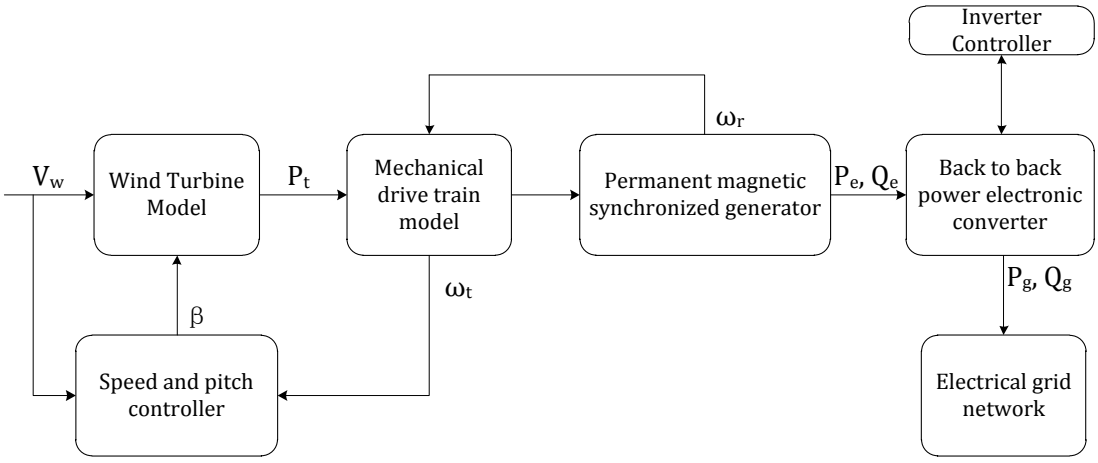


Figure 3.1 - PMSG wind turbine model block diagram

3.1.1.1 Turbine model

The kinetic energy stored in a unit of air travelling at speed u can be found by:

$$KE = \frac{1}{2}mu^2 = \frac{1}{2}\rho Axu^2 \tag{3-1}$$

where m is the mass of the air, ρ is the density of air, A is the cross-sectional area and x is the width of this unit of air. The actual power within this unit of air can then be determined by finding the derivative of the kinetic energy with respect to time:

$$P_w = \frac{1}{2}\,\rho A u^3 \tag{3-2}$$

This equation calculates the actual power within the wind that can be extracted. However, in reality, only a portion of this power can be extracted due to the physical presence of the wind turbine and the aerodynamic performance of the turbine blades. Betz has developed an actuator disk model which represents the ideal turbine rotor, as shown in Figure 3.2 [7, 85].

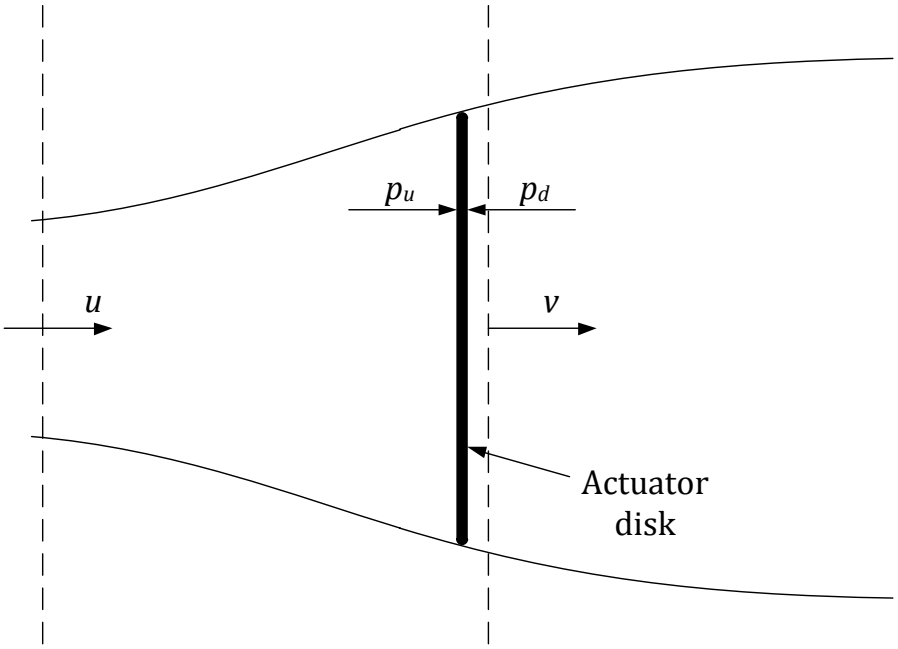


Figure 3.2 - Betz's actuator disk model of a wind turbine

In Figure 3.2, a free stream with an incoming wind speed of u is going through an actuator disk (turbine rotor) with a radius of r. The wind speed after the turbine rotor (downstream speed) is v, and the pressure of the wind at upstream and downstream is p_u and p_d, respectively. Based on the linear momentum theory and Bernoulli equation, Betz has been able to show that the power that can be extracted from the actuator disk is:

$$P_w = \frac{1}{2} C_p\,\rho A u^3 \tag{3-3}$$

$$C_{p=}4a(1-a)^2 \tag{3-4}$$

$$a = \frac{u-v}{u} \tag{3-5}$$

where C_p is known as the power coefficient, and a is the axial induction factor. Betz has also proven that the maximum value of C_p is 0.593, which means that when the aerodynamic of the rotor blades is disregarded, only 59.3% of the power in the wind can be extracted.

The previous model is a linear representation of the ideal turbine. In reality, the rotation of the rotor blades will create angular momentum, which in turn will induce a rotational flow after the rotor in the opposite direction. This process is known as the wake rotation. It has been proven by researchers that the presence of wake rotation will reduce the efficiency of energy extraction. Based on the actuator disk model and inclusive of the wake rotation, the power coefficient is:

$$C_p = \frac{8}{\lambda}\int_0^{\lambda} a'(1-a)\lambda_r^3 d\lambda_r \tag{3-6}$$

$$\lambda = \frac{\Omega R}{u} \tag{3-7}$$

$$\lambda_r^2 = \frac{\Omega^2 r^2}{u^2} = \frac{a(1-a)}{a'(1+a)} \tag{3-8}$$

where λ is the tip speed ratio, Ω is the angular speed of turbine rotor, R is the radius of the rotor and λ_r is the local speed ratio at radius r. It is also a function of axial induction factor a and angular induction factor a'. After solving the above integral, the maximum power coefficient can be found [86]:

$$C_{p,m} = \frac{8}{729\lambda^2}\left\{\frac{64}{5}x^5 + 72x^4 + 124x^3 + 38x^2 - 63x - 12[\ln x] - 4x^{-1}\right\}_{x=(1-3a_2)}^{x=0.25} \tag{3-9}$$

where, x is the substitution of $(1\text{-}3a_2)$ and a_2 is the upper limit of the axial induction factor when λ_r=λ. It can be seen in Figure 3.3 that the theoretical maximum Cp increases as the tip speed ratio λ becomes larger, though it will never exceed the Betz limit.

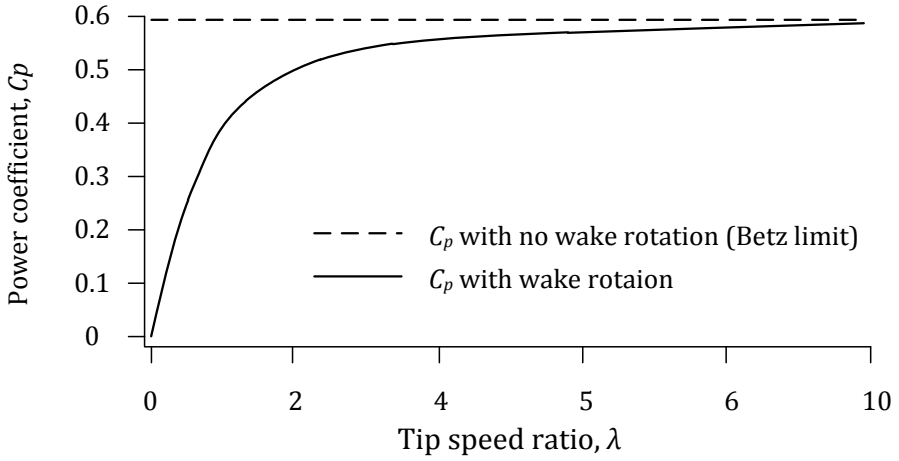


Figure 3.3 - Theoretical power coefficient curve with and without wake rotation

The above analysis covers the maximum conditions of power available in the wind stream for a turbine to extract. The actual extraction is dependent on the aerodynamic characteristic of the blades and the blades' geometry. There is an entire area of study dedicated to the design and selection of turbine blades, though it is beyond the focus of this dissertation. However, the basic theory behind it will be laid out for simulation purposes. Figure 3.4 shows the cross-sectional view of a blade aerofoil with parameters that are relevant in the analysis.

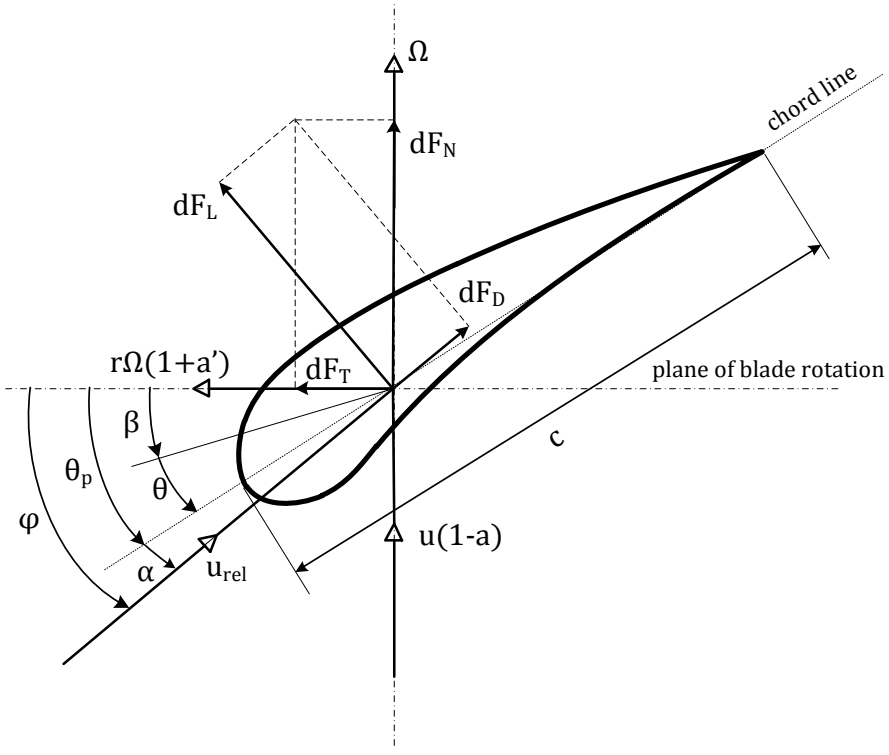


Figure 3.4 - Blade geometry and parameters of a horizontal axis wind turbine

In Figure 3.4, a stream of air flows over the aerofoil u_{rel} at an angle φ, known as the angle of relative wind, which is the sum of the section pitch angle θ_p (consists of the blade pitch angle β and the section twist angle θ) and the angle of attack α. This air flow creates a different pressure distribution across the top and bottom surfaces. These pressures and frictional force will resolve into lifting force F_L (dF_L in Figure 3.4 is the incremental lift force, which is used in a later analysis), drag force, F_D (dF_D) and pitching moment M acting along the chord line c, which is the dotted line connecting the leading edge and the trailing edge of the aerofoil. It is common in an analysis to use coefficients to represent the relationships between these forces and the fluid flow conditions. The corresponding coefficients for the two forces C_L and C_D and the moments C_M are:

$$C_L = \frac{F_l}{\frac{1}{2}\rho c l u^2} \tag{3-10}$$

$$C_D = \frac{F_d}{\frac{1}{2}\rho c l u^2} \tag{3-11}$$

$$C_M = \frac{M}{\frac{1}{2}\rho c^2 l u^2} \tag{3-12}$$

where ρ is the air density, c is the chord length, l is the aerofoil span (width) and u is the wind speed. The lifting and drag forces are always perpendicular and parallel to the relative wind speed, respectively. The relative wind speed is the vector sum of the wind speed at the rotor, $u(1-a)$ and the wind speed due to rotor rotation, $r\Omega(1+a')$, where Ω is the angular speed of the rotor. The lifting and drag forces can then be resolved into the normal F_N (dF_N) and tangential forces F_T (dF_T), which contributes to the trust and torque generated from the aerofoil.

During the analysis of the blade with the length of R, it is common to divide the blade into N elements, where each has an incremental length of dr, as shown in Figure 3.5, known as the blade element theory. Based on the parameters defined in Figure 3.4, the following relationships can be obtained:

$$\tan\varphi = \frac{u(1-a)}{\Omega r(1+a')} = \frac{(1-a)}{(1+a')\lambda_r} \tag{3-13}$$

$$u_{rel} = u(1-a)/\sin\varphi \tag{3-14}$$

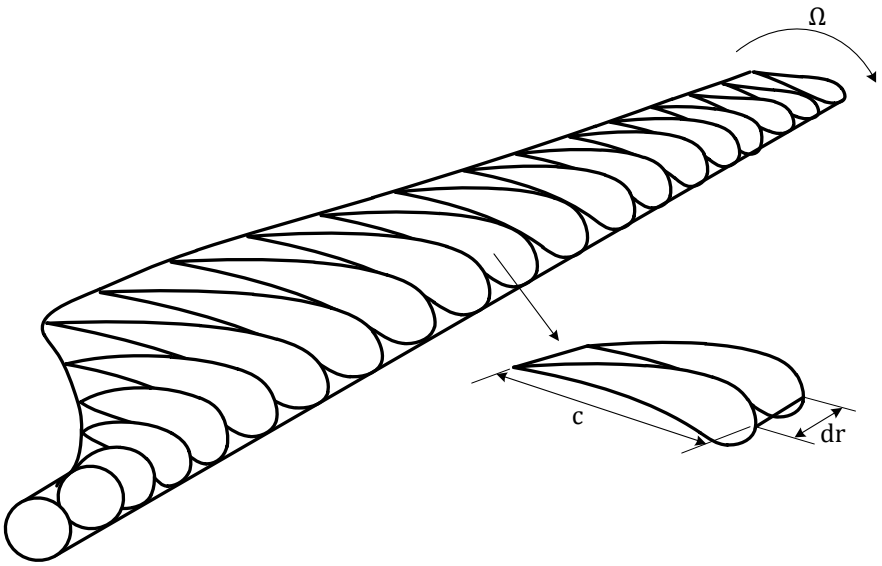


Figure 3.5 - Schematic of wind turbine blade for momentum and blade element theory analysis

$$dF_L = \frac{1}{2} C_L \rho u_{rel}^2 c dr \tag{3-15}$$

$$dF_D = \frac{1}{2} C_D \rho u_{rel}^2 c dr \tag{3-16}$$

$$dF_N = dF_L \cos\varphi + dF_D \sin\varphi \tag{3-17}$$

$$dF_T = dF_L \sin\varphi - dF_D \cos\varphi \tag{3-18}$$

For a rotor with B number of blades, the corresponding total normal force on the section and the differential torque due to tangential force with distance r from the centre are:

$$dF_N = \frac{1}{2} B \rho u_{rel}^2 (C_L \cos\varphi + C_D \sin\varphi) c dr \tag{3-19}$$

$$dQ = \frac{1}{2} B \rho u_{rel}^2 (C_L \sin\varphi - C_D \cos\varphi) c dr \tag{3-20}$$

Based on the linear momentum theory, the differential trust dT and differential torque dQ as a function of axial and angular induction factors are:

$$dT = 4a(1-a)\rho\pi r u^2 dr \tag{3-21}$$

$$dQ = 4a'(1-a)\rho\pi r^3 u\Omega dr \tag{3-22}$$

With the four equations from momentum and blade element theories, it is possible to determine the parameters of the flow conditions with some assumptions. There are two main methods used: one is a graphical method determining the lifting coefficient and axial induction factors [87]; the other is an iterative method solving for axial and angular induction factors [88]. With the flow conditions, the power coefficient can then be calculated by:

$$C_p = \frac{8}{\lambda^2}\int_{\lambda_h}^{\lambda}[1-(C_D/C_L)\cot\varphi]a'(1-a)\lambda_r^3 d\lambda_r \tag{3-23}$$

where, λ_h is the local tip speed ratio at the centre of the rotor and λ_r is the local tip speed ratio at distance r from the centre. As it can be seen, the power coefficient is now a function of both the tip speed ratio and the pitch angle β (β=φ-α-θ). In the simulation model, equation (3-3) is used to calculate the actual power extracted with the power coefficient at different tip speed ratios and pitch angles.

3.1.1.2 Speed and power control

Based on equations (3-3) and (3-23), the power extraction rate is dependent on the tip speed ratio, which is a function of wind speed, rotor rotational speed and the pitch angle. By controlling these parameters, the power extraction efficiency can be optimised.

Speed control in this context refers to the turbine's ability to operate at a range of wind speed. Initially, wind turbines are designed and operated at a fixed wind speed. This implies a disregard of the wind speed; the rotational speed of a turbine rotor is dependent on the gearbox specification, grid supply frequency and generator characteristics. Such a design has the main drawback of a low efficiency of power extraction, as the power coefficient of the turbine blades is designed optimally at a fixed speed. Moreover, the fixed speed wind turbine lacks power quality control and induces high mechanical stresses to the components.

The PMSG wind turbine is normally of variable speeds. This is realised by the presence of a full scaled power converter, such that the fast wind fluctuation is absorbed by the power converter and maintains a stable output that meets the grid requirement through appropriate control algorithm. In this case, the turbine

blade can be designed to have maximum power efficiency over a range of wind speed, where pitch angle control can be integrated. Compared to the fixed wind speed turbine, variable speed control provides a higher power capture rate and reduces the high mechanical stresses due to wind fluctuation. Moreover, the power converter allows the system to improve the quality of the electricity generated. An obvious disadvantage of variable speed turbines concerns the additional losses due to power converters. However, the improvement of the power capture rate can easily compensate for this loss and results in an overall efficiency enhancement [86]. Moreover, as more components are needed for this design, the capital cost and reliability issues increase.

The power extraction can also be optimised by controlling the aerodynamic behaviour of turbine blades. The relationship of wind speed and power output can show the power extraction efficiency with different control strategies as in Figure 3.6. According to the blade characteristics and the natural distribution of wind speed, there are three main stages that are critical for a wind turbine. First is the cut-in speed, which is the minimal wind speed required for the turbine rotor to have a sufficient torque to rotate. No power is generated below this speed, which is typically around 3-4 m/s. Second, with increasing wind speed until the rated speed, the turbine power increases up to the rated speed with designed behaviour. This is the normal operational zone for the turbine. Finally, the cut-out speed is where the turbine stops generating power due to the high forces, which may post as a risk in terms of damaging the turbine structure. As can be observed in the Figure 3.6, different control strategies will result in a different power curve.

In Figure 3.6, the first thing to notice is the difference between the fixed and variable speeds during a normal operation zone. For the fixed speed, the power extraction is optimal at a single speed; for variable speed, this is optimised for a range of wind speed. Moreover, by utilising dynamic pitch control, the power generation above the rated speed can be optimised. In contrast, with stall control, the power extraction will start to decrease due to the increase of turbulence and cause the turbine to stall, such as the curves with fixed pitch in Figure 3.6. This

behaviour is dependent on the inherent aerodynamic characteristics of the turbine blade during the design.

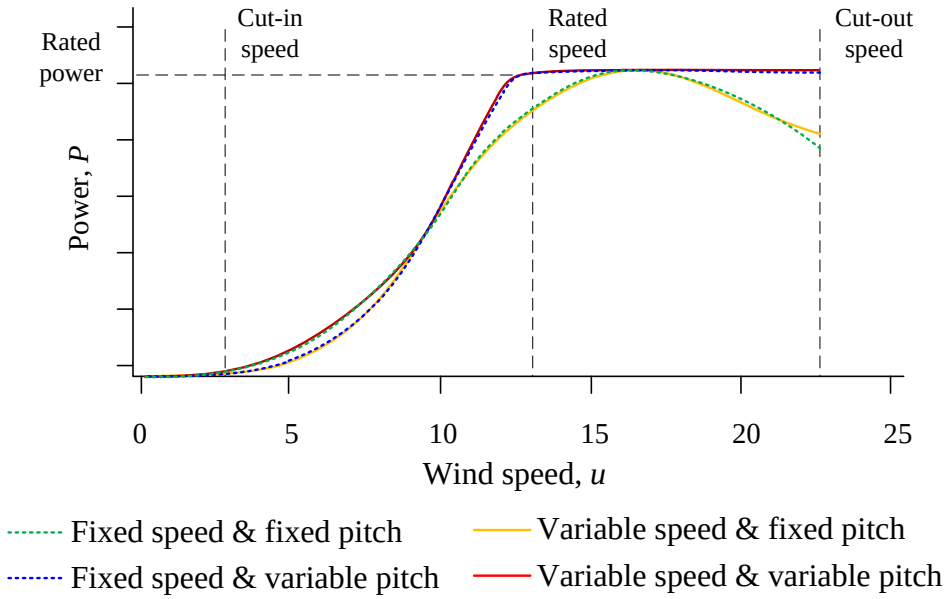


Figure 3.6 - Wind turbine power curve characteristics under different conditions

In the simulation of a PMSG wind turbine, a dynamic (active) pitch control strategy is adopted. The controller block diagram is shown in Figure 3.7. The controller consists of two parts: the generation of reference pitch angle and the pitch actuator model. There are two modes of reference pitch angle, which depend on the actual power P_g. The selector outputs the reference pitch angle β_{ref} to be 0 when the actual power is lower than the reference power (rated power) P_{ref}. On the contrary, when the actual power is greater than the reference power, the β_{ref} is obtained from the proportion-integral (PI) controller based on the difference of actual power and reference power, P_{g_err}. Because the sensitivity of the torque is minimal in relation to the pitch angle at the rated power, a gain scheduling block is used to compensate for the higher controller gain required [89, 90]. Moreover, the reference pitch angle is passed to the actuator model and limited between 0 and 25°. The final output pitch angle β is fed into the aerodynamic model as shown in Figure 3.1.

3.1.1.3 Drive train model

A wind turbine drive train model mainly consists of blades and a pitching mechanism, a rotor hub and a rotor shaft that act as the main coupling component between the rotor and the generator, and the gearbox that may be

included depends on the turbine design. The model represents the dynamical behaviour of these mechanical components in terms of the torque, rotational speed, inertia and other aspects. These mechanical behaviours have a direct influence on the output electrical power, such as the torsional mechanics of components, tower effect and vibration due to wind shear and tower shadow. In terms of the power generation point of view, the latter two behaviours have less of an impact than the torsional mechanics; thereby, the drive train model mainly analyses the torsional dynamics of the components.

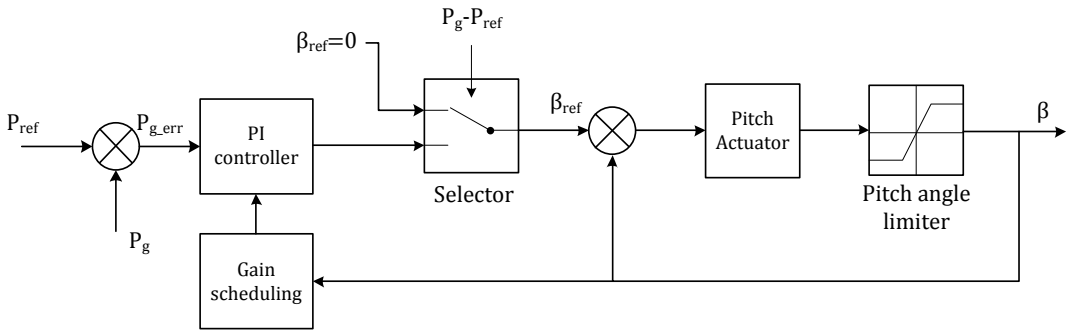


Figure 3.7 - Pitch angle controller block diagram

The drive train model can be considered as a multi-mass system and as adopting a classical rotational analysis; the dynamic equations can be formulated with the physical parameters of the components. Many different models have been developed in the recent decades. O. Wasynczuk *et al*. [91] established a 5-mass drive train model of a MOD 2 (HAWT with 2 blades) type wind turbine, where the dynamical behaviour of the turbine under random wind fluctuation with both fixed and variable pitch modes was investigated. Moreover, a 5-mass model of a fixed speed HAWT with three blades was studied by I.P. Girsang *et al*. [92] from NREL (National Renewable Energy Laboratory). In their model, a detailed dynamic analysis of the multistage gearbox, including a planetary gear set and two parallel gear sets, was carried out. On the contrary, a 3-mass model was presented in [93], where the three masses represent the bending effects of the flexible part of turbine blades, the turbine rotor including the rigid part of the turbine blades, the hub and low speed shaft and the gearbox, high speed shaft and generator. The researchers believe that this flexibility has a substantial impact on the power generation. However, the main issue with such a model is the difficulty of distinguishing between the rigid and flexible parts of the turbine

blades. It can be seen that each of these models focuses on a different aspect of component dynamics. In their literature, Muyeeen *et al*. [90, 94] determined that a 6-mass drive train model is needed to precisely represent the transient behaviour of a wind turbine system. The model consists of the inertia from each blade, the hub, the gearbox and the generator as shown in Figure 3.8 a). The torque, angular speed and displacement for each component are represented as T, ω and θ with respective subscription. The elasticity stiffness constant, K, and damping factors, D, used in the analysis are not shown in the figure. The author further discussed simplified 3-mass and 2-mass models, and it has been concluded that it is sufficient to transform the 6-mass drive train model into a 2-mass model with equivalent parameters as shown in Figure 3.8 b). The simplified 2-mass model is widely used in wind turbine simulation, where the authors from [95-97] have adopted this model to simulate gearless variable speed wind turbines and study the dynamics of a wind turbine system.

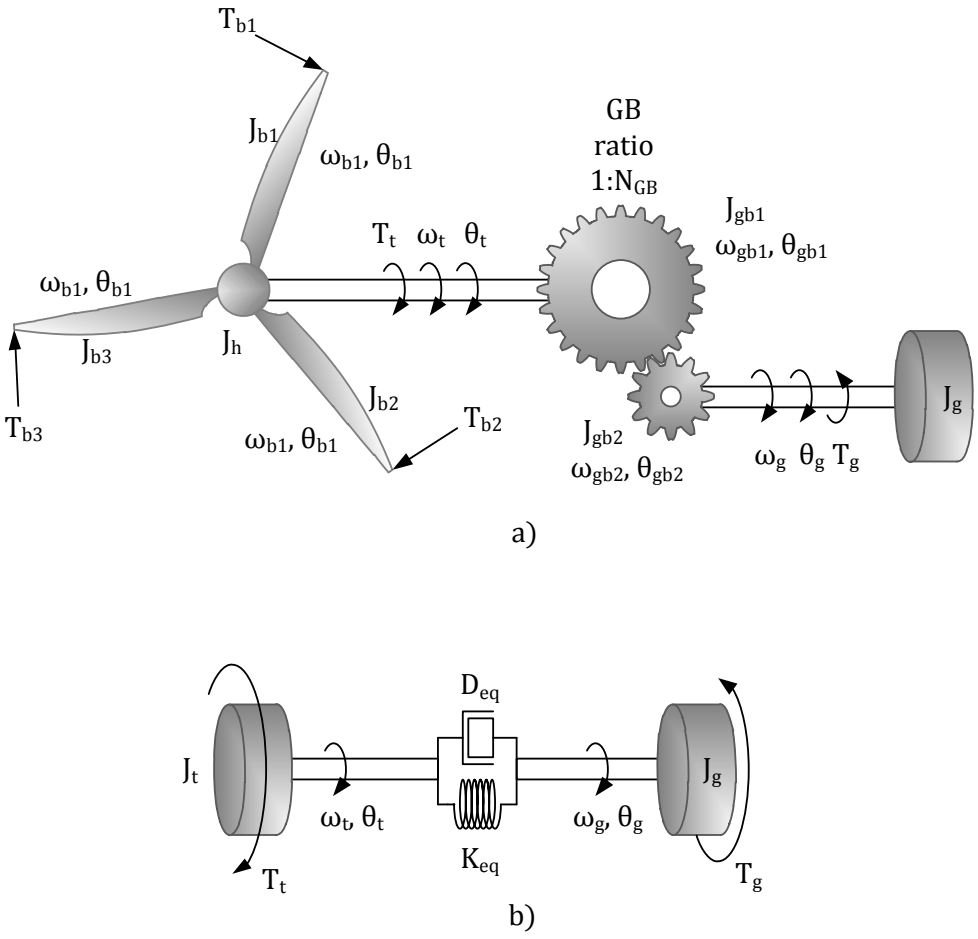


Figure 3.8 - Wind turbine drive train model: a) 6-mass drive train model; b) 2-mass drive train model.

In our simulation, the 2-mass drive train model is implemented. As shown in Figure 3.8 b), the model consists of the equivalent turbine inertia, J_t, and generator inertia, J_g. The mathematical representation of the model is:

$$J_t \frac{d\omega_t}{dt} = T_t - K_{eq}(\theta_g - \theta_t) - D_{eq}(\omega_g - \omega_t) \quad (3\text{-}24)$$

$$J_g \frac{d\omega_g}{dt} = -T_g + K_{eq}(\theta_g - \theta_t) + D_{eq}(\omega_g - \omega_t) \quad (3\text{-}25)$$

$$\frac{d\theta_t}{dt} = \omega_t \quad (3\text{-}26)$$

$$\frac{d\theta_g}{dt} = \omega_g \quad (3\text{-}27)$$

where K_{eq} and D_{eq} are the equivalent stiffness and damping factors, ω_t and ω_g are the turbine and generator angular speed and θ_t and θ_g are the angular displacement of the turbine and generator, respectively.

3.1.1.4 Permanent magnetic synchronised generator model

The modelling of a PMSG can be derived from the classical analysis of synchronous machines. The graphical representation of a typical three-phase synchronous machine with two salient poles and star connection is shown in Figure 3.9. The overall coordination system used in the analysis is defined in Figure 3.9 a), where a_s-a_s', b_s-b_s' and c_s-c_s' are the three sinusoidally distributed windings with a displacement of 120°. The rotor is rotating with an angular speed of ω_r, and the machine angular displacement is θ_r; a direct axis is defined for the rotor, and it is 90°, lagging the quadrature axis. Figure 3.9 b) and c) show the windings for the stator and rotor, respectively. Because a symmetrical system is considered, the three-phase stator windings are the same and have a resistance of r_s and a number of turns of N_s. The rotor has a field winding f_d and a damper winding k_d on the d-axis, as well as two other damper windings, k_{q1}, and k_{q2}, in the q-axis.

In developing the mathematical models for the synchronous machine, certain assumptions are made in order to simplify the analysis: the machine stator windings are considered to be symmetrical; the flux distribution across the air-gap between the stator and rotor is perfectly sinusoidal; the physical position of the rotor has no effect on the permeance of the magnetic flux; and the saturation and hysteresis losses are not considered. These assumptions have been proven to

be valid for obtaining a precise model of the machine [98]. With the predefined convention of the generator, the voltage equations of the synchronous machine in the machine parameters are:

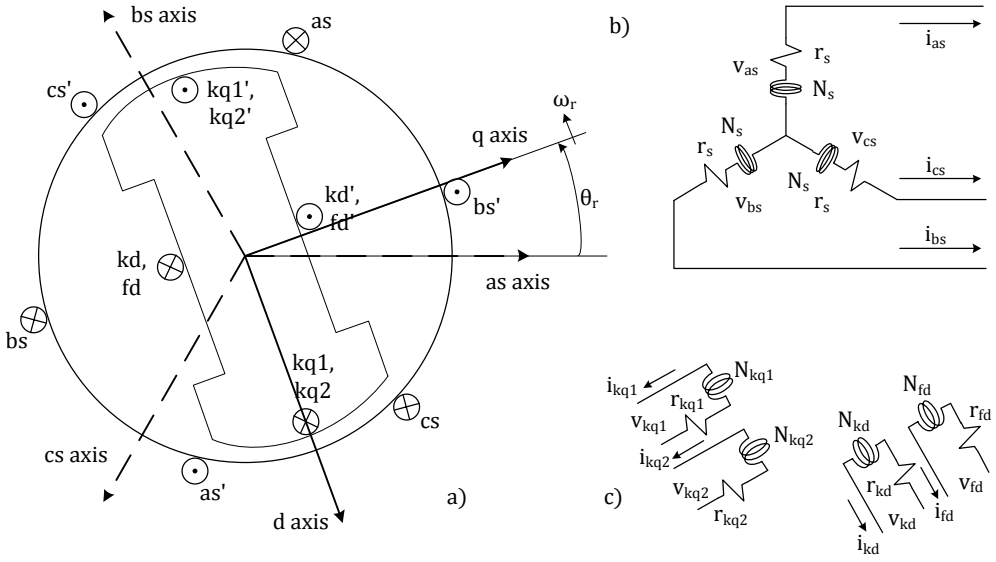


Figure 3.9 - Schematics of a two pole 3-phase synchronous machine in star connection with salient-pole: a) overall coordination system; b) stator winding; c) rotor winding [98]

$$\mathbf{V}_{abcs} = -\mathbf{r}_s\mathbf{I}_{abcs} + \frac{d}{dt}\boldsymbol{\Psi}_{abcs} \tag{3-28}$$

$$\mathbf{V}_{dqr} = \mathbf{r}_{dqr}\mathbf{I}_{dqr} + \frac{d}{dt}\boldsymbol{\Psi}_{dqr} \tag{3-29}$$

$$\begin{bmatrix}\boldsymbol{\Psi}_{abcs}\\ \boldsymbol{\Psi}_{dqr}\end{bmatrix} = \begin{bmatrix}\mathbf{L}_s & \mathbf{L}_{sr}\\ \mathbf{L}_{sr}{}^T & \mathbf{L}_r\end{bmatrix}\begin{bmatrix}-\mathbf{I}_{abcs}\\ \mathbf{I}_{dqr}\end{bmatrix} \tag{3-30}$$

where $\mathbf{V}_{abcs}$, $\mathbf{I}_{abcs}$ and $\boldsymbol{\Psi}_{abcs}$ are vectors of a three-phase stator voltage, current and flux respectively; $\mathbf{V}_{dqr}$, $\mathbf{I}_{dqr}$ and $\boldsymbol{\Psi}_{dqr}$ are vectors of rotor voltage, current and flux in the *dq*-axes; $\mathbf{r}_{abcs}$ and $\mathbf{r}_{dqr}$ are the stator and rotor winding resistance matrix; The flux vectors are a function of current and inductances as shown in equation (3-30), and the inductance is dependent on the physical parameters of each winding, such as the number of turns and the angular position of the rotor. Detailed derivation of the inductance matrix can be found in [98]. The presence of the time varying inductances has increased the complexity in solving equations (3-28) to (3-30). Therefore, an alternative approach of adopting a *dq0*

reference frame transformation in order to simplify the generator equations by eliminating the time dependent inductance as shown in Appendix A.1.

In the case of a permanent magnet synchronous machine, it has similar structure to a synchronous machine including stator windings, damper winding in direct and quadrature axes, and the field winding, which is replaced with a magnet that provides the equivalent magnetic field. The presence of the permanent magnet enables the machine to excite by itself, thus reducing the interaction between the stator and rotor windings and simplifies the mathematical model. The mathematical representation of the PMSG with one damper winding on the direct and quadrature axes in the *dq0* frame using equation (A-1) can be written as:

$$v_D = -r_s i_D - \omega_r \psi_Q + \frac{d\psi_D}{dt} \tag{3-31}$$

$$v_Q = -r_s i_Q + \omega_r \psi_D + \frac{d\psi_Q}{dt} \tag{3-32}$$

$$v_f = r_f i_f + \frac{d\psi_f}{dt} = 0 \tag{3-33}$$

$$v_d = r_d i_d + \frac{d\psi_d}{dt} \tag{3-34}$$

$$v_q = r_q i_q + \frac{d\psi_q}{dt} \tag{3-35}$$

where v_D, v_Q, ψ_D and ψ_Q are the voltage and flux linkage of the stator in the *dq0* frame; v_f, r_f, i_f and ψ_f are the field winding voltage, resistance, current and linkage flux, respectively; and the voltage, resistance, current and linkage flux of the direct and quadrature damper winding have the subscript *d* and *q*, respectively. The flux linkage equation is:

$$\psi_D = -L_D i_D + m_f i_f + m_{sd} i_d \tag{3-36}$$

$$\psi_Q = -L_Q i_Q + m_{sq} i_q \tag{3-37}$$

$$\psi_f = L_f i_f \tag{3-38}$$

$$\psi_d = L_d i_d + m_{fd} i_f - m_{sd} i_D \tag{3-39}$$

$$\psi_q = L_q i_q - m_{sq} i_Q \tag{3-40}$$

where L_D, L_Q, L_d and L_q are the inductances of stator and damper windings in the *dq0* frame; the mutual inductances between the field winding and the *d*-axis stator winding, the damper windings are m_f and $m_{fd,}$ respectively; and L_f is the main field winding inductance. The mutual inductances between stator windings and damper windings in the *dq* axes are m_{sd} and $m_{sq,}$ respectively.

It can be seen that in order to solve this equation, inductances of each winding are needed. However, in practice, the manufacturers generally do not provide inductances in the forms as described in equations (3-36) to (3-40). Instead, Table 3.1 lists all the machine parameters provided by the manufacturers in reactances and time constants [98, 100]. Each of these parameters is determined from the equivalent circuit of the machine as shown in Figure 3.10. The mathematical relationship between the parameters given by manufacturers and the parameters in equations (3-31) to (3-40) is given in the Appendix A.2.

Machine parameters in reactance and time constants	
$\boldsymbol{X_d}$	Unsaturated *d*-axis synchronous reactance
$\boldsymbol{X_q}$	Unsaturated *q*-axis synchronous reactance
$\boldsymbol{\dot{X}_d}$	Unsaturated *d*-axis synchronous transient reactance
$\boldsymbol{\ddot{X}_d}$	Unsaturated *d*-axis synchronous sub-transient reactance
$\boldsymbol{\ddot{X}_q}$	Unsaturated *q*-axis synchronous sub-transient reactance
$\boldsymbol{\dot{T}_d}$	Unsaturated *d*-axis short circuit transient time constant
$\boldsymbol{\dot{T}_{do}}$	Unsaturated *d*-axis open circuit transient time constant
$\boldsymbol{\ddot{T}_d}$	Unsaturated *d*-axis short circuit sub-transient time constant
$\boldsymbol{\ddot{T}_{do}}$	Unsaturated *d*-axis open circuit sub-transient time constant
$\boldsymbol{\ddot{T}_q}$	Unsaturated *q*-axis short circuit sub-transient time constant
$\boldsymbol{\ddot{T}_{qo}}$	Unsaturated *q*-axis open circuit sub-transient time constant

Table 3.1 - Synchronous machine parameters given by the manufacturer

Figure 3.10 shows the equivalent circuit of the PMSG model in a *dq0* frame. It is common practice to refer variables of the rotor side to the stator side to reduce complexity, and this can be found based on the turn ratio between each winding. Based on the equivalent circuit [98, 99], the PMSG can be modelled as:

$$v_D = -r_s i_D - \omega_r \psi_Q + \frac{d\psi_D}{dt} \tag{3-41}$$

$$v_Q = -r_s i_Q + \omega_r \psi_D + \frac{d\psi_Q}{dt} \tag{3-42}$$

$$v'_f = r_f i_f + \frac{d\psi_f}{dt} = 0 \tag{3-43}$$

$$v'_d = r'_d i'_d + \frac{d\psi'_d}{dt} \tag{3-44}$$

$$v'_q = r'_q i'_q + \frac{d\psi'_q}{dt} \tag{3-45}$$

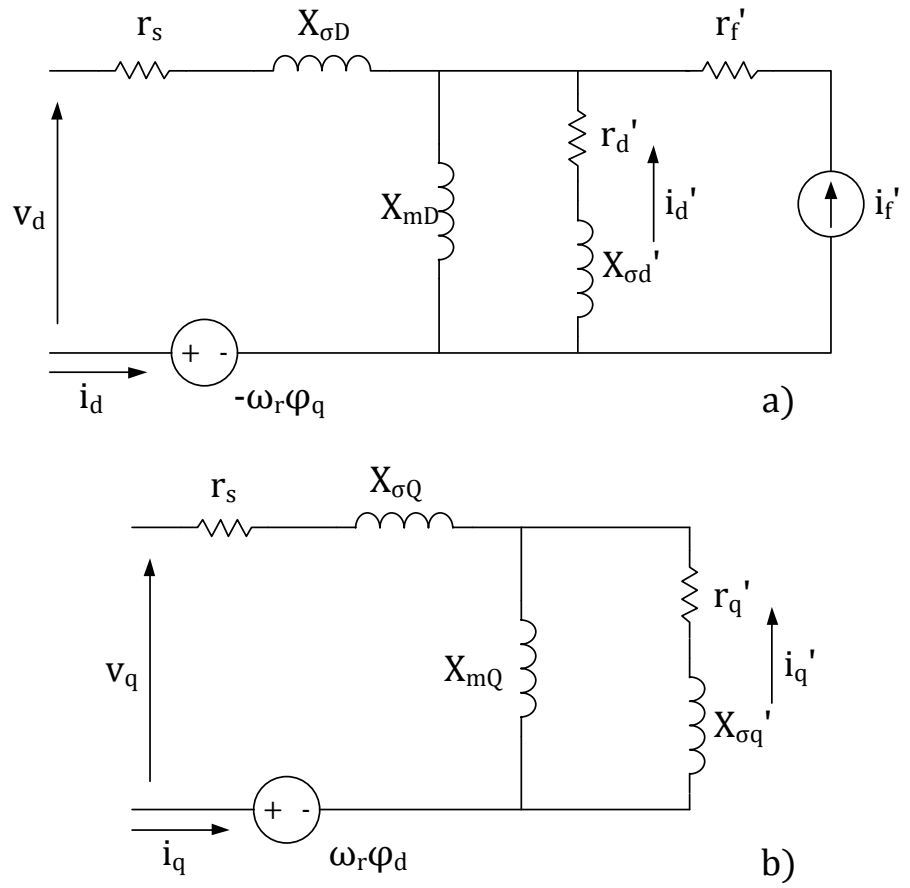


Figure 3.10 - Equivalent circuit model of PMSG in *dq0* frame: a) *d*-axis; b) *q*-axis

$$\psi_D = X_{mD}\left(-i_D + i'_d + i'_f\right) - X_{\sigma D} i_D \quad (3\text{-}46)$$

$$\psi_Q = X_{mQ}\left(-i_Q + i'_q\right) - X_{\sigma Q} i_Q \quad (3\text{-}47)$$

$$\psi'_f = constant \quad (3\text{-}48)$$

$$\psi'_d = X_{mD}\left(-i_D + i'_d + i'_f\right) + X'_{\sigma d} i'_d \quad (3\text{-}49)$$

$$\psi'_q = X_{mQ}\left(-i_Q + i'_q\right) + X'_{\sigma q} i'_q \quad (3\text{-}50)$$

where the superscript ' represents the equivalent parameters of the rotor side in the stator side. X_{mD} and X_{mQ} are the main reactances for the direct axis and the quadrature axis; $X_{\sigma D}$, $X_{\sigma Q}$, $X_{\sigma d}$, $X_{\sigma q}$ are the leakage reactance of stator and damper windings in the *dq0*-axes, and the permanent magnet is modelled as a current source with a constant magnetic flux with an equivalent resistance of r_f'.

By solving equations (3-41) to (3-50) with the parameters of a given machine, the stator voltage, current and flux can be found. These variables can then be used to calculate the electromagnetic torque T_e of the machine:

$$T_e = \frac{3}{2} p\left(\psi_D i_Q - \psi_Q i_D\right) \quad (3\text{-}51)$$

where p is the number of magnet pole pairs.

In the modelling of the PMSG for a wind turbine, it is commonly found in literature that the damper windings of the machine are neglected and still provide sufficient accuracy. This can further simplifies equations (3-41) to (3-50), and left with the voltage and flux linkage equation of the stator windings in the *dq0* frame and the constant flux produced by the magnet [101, 102]. However, in this paper, the complete model with damper windings is adopted.

3.1.1.5 Full rated power electronic converter and control

For a variable speed, direct driven PMSG wind turbine, the back-to-back power converter is an essential component. It consists of a machine side converter, which converts the variable AC power output from the generator to DC power and maximises power extraction through the regulation of the generator speed; a DC-link capacitor that acts as an energy storage; and a grid side converter which is aimed to invert the DC power back into stable AC to meet the grid standard. Traditionally, the back-to-back converter uses the diode rectifier and thyristor inverter. Because of the fast development in power electronics, with increases in the size and reliability of power switching transistors such as insulated-gate bipolar transistor (IGBT), this make the implementation of IGBT in wind power applications possible [101]. Different converter topologies and control algorithms have been proposed by researchers. A converter with an uncontrolled rectifier with a DC-DC converter known as a switch mode rectifier and IGBT inverter for stand-alone PMSG wind turbine is proposed in [103]. Moreover, a converter with controllable IGBT rectification and an inverter can be found in [104-107]. The converter output and the DC-link voltage are controlled to its rated value by the converter current in the *dq0*-axes. For the machine side converter, an independent active and reactive power scheme is used. The active power is controlled to be optimal at any given speed through maximum power point tracking (MPPT) by the *d*-axis rotor current, and the reactive power is controlled by the *q*-axis rotor current.

In our simulation, the aim is to generate wind turbine data to validate the proposed algorithm. Therefore, a benchmark model developed by PSCAD/EMTDC with a diode rectifier and thyristor converter is adopted, as

shown in Figure 3.11 [112]. For this model because the AC to DC rectification is not controllable, the converter has no control over the generator speed. An RLC circuit is then used to filter out noise and stabilise the electrical voltage input for the inverter. The inverter converts the DC voltage back to AC voltage, where it is controlled by the firing pulses generated from the controller based on the DC bus current and voltage. An additional phase locked loop (PLL) is used to synchronise the phase angle of the converted AC.

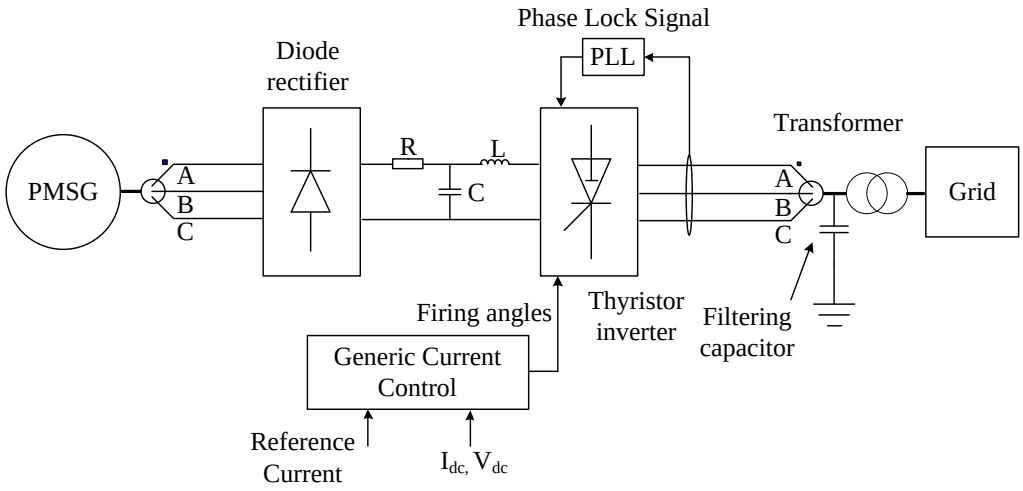


Figure 3.11 - Schematic of the PMSG wind turbine power electronic converter topology and control [112]

The structure of the controller used to generate the firing pulses is shown in Figure 3.12. The controller used a voltage dependent current limiter based on the DC-link voltage and DC-link current before the capacitor to maintain the DC-link voltage within a predefined limit. It then produces a reference DC-link current after the capacitor. The error between the reference and actual current is passed through a gain, PI controller and limiter to produce the acquired firing pulses to operate the thyristor. The controller stabilises the DC-link voltage, thus avoiding its collapse due to a high current drawn from the converter. Simulation model is developed in PSCAD and validated under various operational conditions as shown in later section.

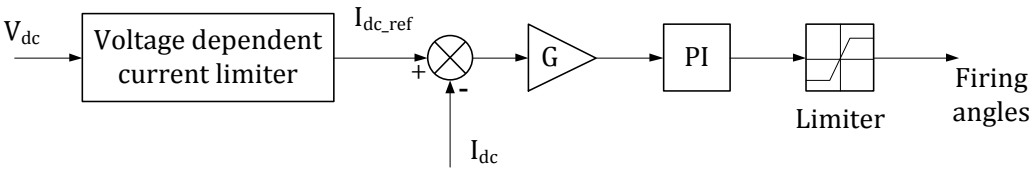


Figure 3.12 - Grid side thyristor controller block diagram

3.1.2 Wind turbine with double-induction generator [98, 108, 109]

In order to have a better understanding of wind turbine systems, the double fed induction generator (DFIG) wind turbine is also simulated. DFIG is one of the alternative solutions for variable speed wind turbines. In contrast to the PMSG, the DFIG wind turbine generally uses a wound rotor induction machine with stator windings directly connected to the grid and the rotor connected to the AC-DC-AC power converter. Because both stator and rotor windings are connected to the electrical source, the term 'doubly fed' is used. As for the PMSG, the high number of poles enables the generator to operate at a low speed that matches the turbine rotor speed, which can eliminate the use of a gearbox. However, the high number of poles also dramatically increases the size and complexity of the generator, which in turn increases the manufacturing and assembling cost. Moreover, for the DFIG, the AC-DC-AC converter is aimed to control the excitation voltage and the frequency of the rotor at a desired value, and it only handles a portion of the rated power. In contrast to a PMSG, a full rated power converter is not needed. This reduced power rating is not only reflected in cost and reliability but is also responsible for fewer harmonics induced to the grid [110].

A wind turbine with a doubly fed induction generator is simulated to study the dynamic behaviour of the system, and a comparison with a turbine with a PMSG is also carried out. Figure 3.13 shows the block diagram of a DFIG wind turbine model. In later sections, a detailed discussion of the DFIG modeling is given, including induction generator, power electronic converter and control strategies.

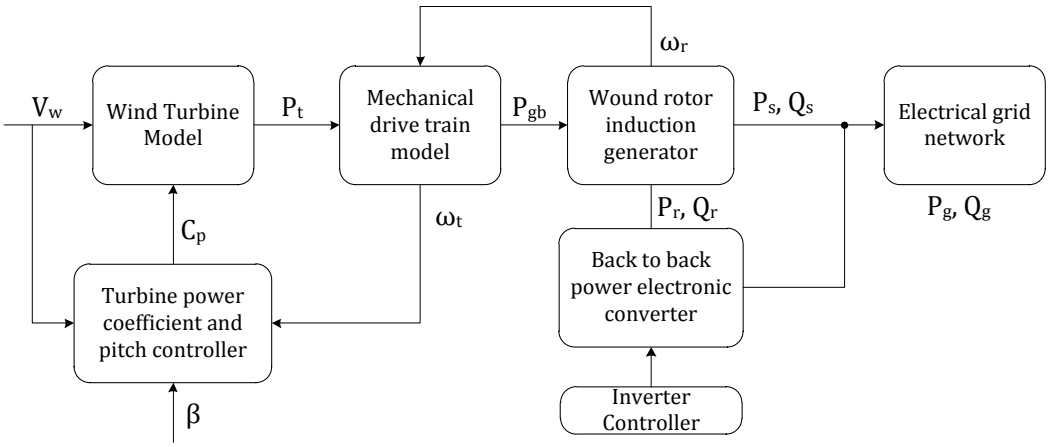


Figure 3.13 - DFIG wind turbine model block diagram

3.1.2.1 Double fed induction generator model

Wound rotor induction machines are normally used in a DFIG wind turbine, where the separated static rotor winding terminal is accomplished through a slip-ring mechanism. Figure 3.14 shows the graphical representation of a two-pole three-phase induction machine in a star connection. The coordination system of the machine is shown in Figure 3.14 a). a_s-a_s', b_s-b_s', c_s-c_s' and a_r-a_r', b_r-b_r' and c_r-c_r' are the three-phase stator and rotor windings displaced with 120°. The rotor will rotate at an angular velocity of ω_r and has a displacement with the stator θ_r. The stator and rotor windings are shown in Figure 3.14 b) and c), with each having resistances of r_s and r_r and the number of turns of N_s and N_r, respectively.

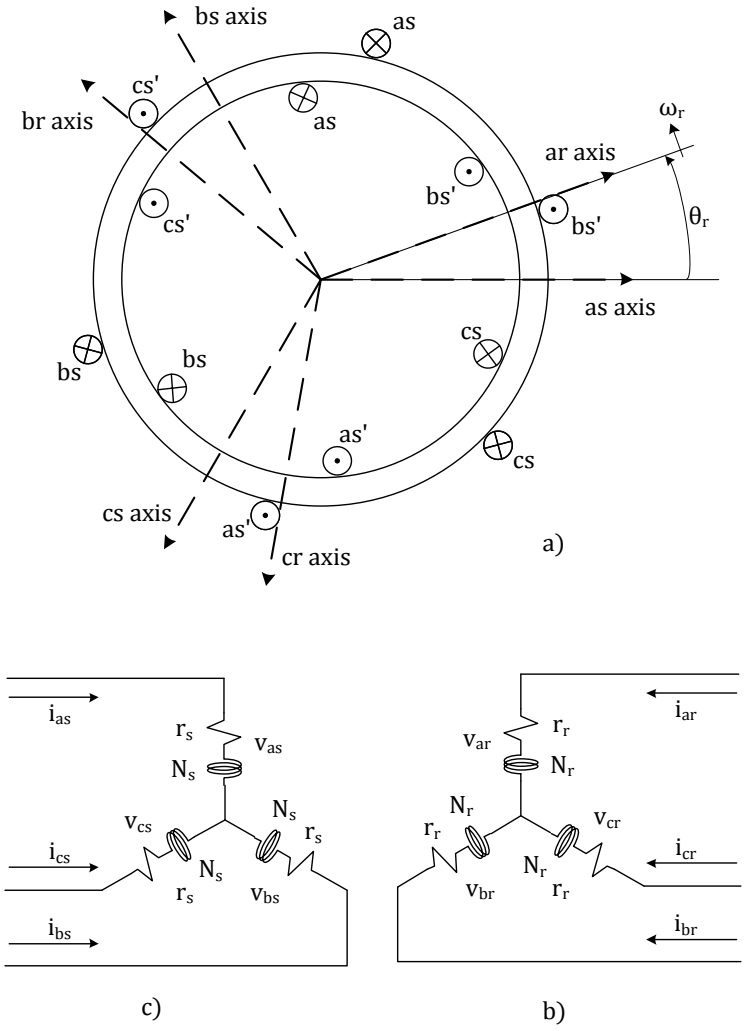


Figure 3.14 - Schematics of a two-pole three-phase induction machine in a star connection: a) coordination system; b) stator winding; c) rotor winding [98]

Again, assumptions are made during the formulation of a mathematical model for the induction machine: The machine windings are considered geometrically symmetrical and sinusoidally distributed; equivalent concentrated windings are used to replace the actual windings; stator and rotor magnetic flux harmonics are negligible; magnetic saturation, core loss and skin effect are not included in the analysis; finally, windings resistance and reactance do not vary with the temperature. Based on these assumptions and the machine convention defined in Figure 3.14, the voltage and flux equations of the machine are:

$$\mathbf{V}_{abcs} = \mathbf{r}_s \mathbf{I}_{abcs} + \frac{d}{dt}\boldsymbol{\Psi}_{abcs} \tag{3-52}$$

$$\mathbf{V}'_{abcr} = \mathbf{r}'_r \mathbf{I}'_{abcr} + \frac{d}{dt}\boldsymbol{\Psi}'_{abcr} \tag{3-53}$$

$$\begin{bmatrix}\boldsymbol{\Psi}_{abcs} \\ \boldsymbol{\Psi}'_{abcr}\end{bmatrix} = \begin{bmatrix}\mathbf{L}_s & \mathbf{L}'_{sr} \\ \mathbf{L}'^{\,T}_{sr} & \mathbf{L}'_r\end{bmatrix}\begin{bmatrix}\mathbf{I}_s \\ \mathbf{I}'_r\end{bmatrix} \tag{3-54}$$

where, the subscript *abcs* and *abcr* denotes the parameters of the three-phase stator and rotor windings, the superscript ' indicates the equivalent rotor winding variables in the stator winding based on an appropriate turn ratio; **V**, **I** and **Ψ** are the voltage, current and flux linkage vectors, **r** is the diagonal matrix of the winding resistance, **L** are the inductances matrix associated with each winding including self-inductances $\mathbf{L}_s$ and $\mathbf{L}'_r$ and mutual inductances $\mathbf{L}_{sr}$ and $\mathbf{L}'^{T}_{sr}$, respectively. Similar to the PMSG model, reference transformation is applied in order to simplify the equations by removing the time dependent inductances as shown in Appendix A.1.

However, it is worth mentioning that there are three reference frames commonly selected for particular purposes (Fig 3.15), where the selected reference frame has a displacement of θ to the stator winding *A* and rotating at a speed of ω. One is the rotor reference frame ($\omega=\omega_r$), where an analysis of any unbalance or discontinuity in the rotor winding is preferred in the rotor reference frame. On the contrary, an analysis with an unbalance or discontinuity in the stator winding is ideally carried out in stationary reference frame ($\omega=0$). Finally, a synchronous reference frame ($\omega=\omega_s$, ω_s is the synchronous speed) is normally used in a control system analysis or in a simulation of transient behaviour of the machine.

Depending on the purpose of the study, an appropriate reference frame can be selected.

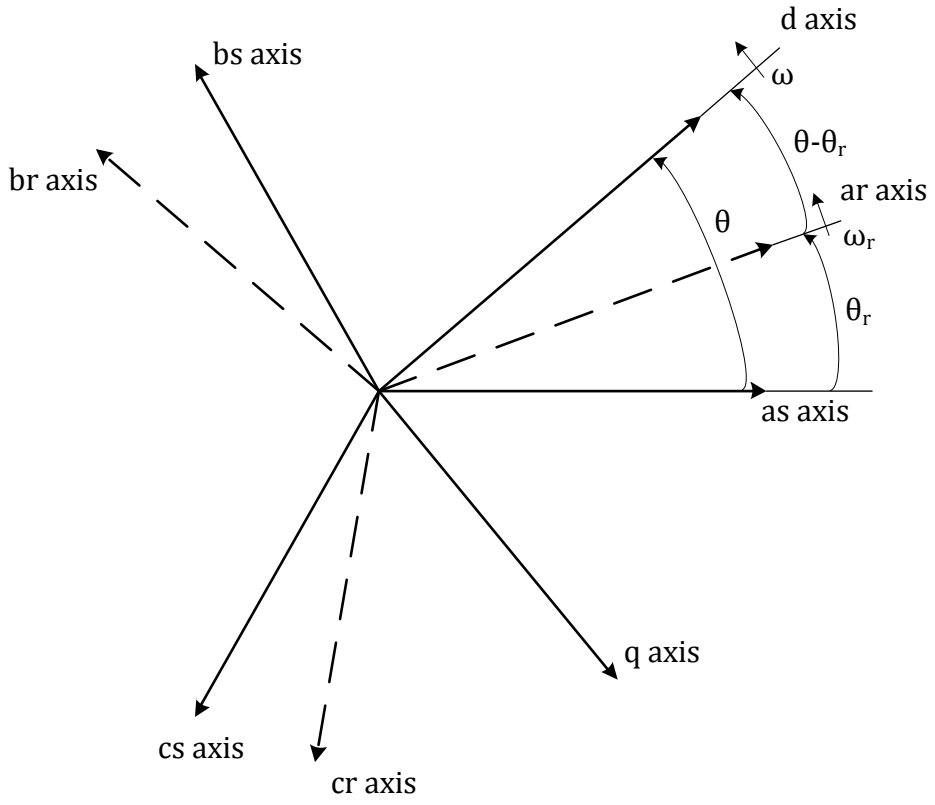


Figure 3.15 - Coordination system of the machine stator winding, rotor winding and transformed reference

Similar to the PMSG model, it is common practice to use reactance instead of inductances in the voltage equation, and the equivalent circuits of the induction machine are shown in Figure 3.16. The corresponding voltage and flux linkage equations of the machine in the arbitrary frame are given:

$$v_{ds} = r_s i_{ds} - \omega\psi_{ds} + \frac{d\psi_{ds}}{dt} \tag{3-55}$$

$$v_{qs} = r_s i_{qs} + \omega\psi_{qs} + \frac{d\psi_{qs}}{dt} \tag{3-56}$$

$$v_{0s} = r_s i_{0s} + \frac{d\psi_{0s}}{dt} \tag{3-57}$$

$$v'_{dr} = r'_r i'_{dr} - (\omega - \omega_r)\psi'_{dr} + \frac{d\psi'_{dr}}{dt} \tag{3-58}$$

$$v'_{qr} = r'_r i'_{qr} + (\omega - \omega_r)\psi'_{qr} + \frac{d\psi'_{qr}}{dt} \tag{3-59}$$

$$v'_{0s} = r'_r i'_{0r} + \frac{d\psi'_{0r}}{dt} \tag{3-60}$$

$$\psi_{ds} = X_{\sigma s} i_{ds} + \frac{3}{2} X_{ms}(i_{ds} + i'_{dr}) \quad (3\text{-}61)$$

$$\psi_{qs} = X_{\sigma s} i_{qs} + \frac{3}{2} X_{ms}(i_{qs} + i'_{qr}) \quad (3\text{-}62)$$

$$\psi_{0s} = X_{\sigma s} i_{0s} \quad (3\text{-}63)$$

$$\psi'_{dr} = X'_{\sigma r} i'_{dr} + \frac{3}{2} X_{ms}(i_{ds} + i'_{dr}) \quad (3\text{-}64)$$

$$\psi'_{qr} = X'_{\sigma r} i'_{qr} + \frac{3}{2} X_{ms}(i_{ds} + i'_{dr}) \quad (3\text{-}65)$$

$$\psi'_{0r} = X'_{\sigma r} i'_{0r} \quad (3\text{-}66)$$

where the subscript *ds*, *qs*, *0s*, *dr*, *qr* and *0r* refers the stator and rotor in the *dq0* reference frame, respectively; the superscript ' represents the equivalent parameters of the rotor side in the stator side. X_{ms} is the mutual reactance of the stator and rotor windings; $X_{\sigma s}$ and $X_{\sigma r}$ are the leakage reactance of the stator and rotor windings in *dq*-axis; r_s and r_r are the winding resistance of both stator and rotor; *v*, *i* and ψ are the voltage, current and magnetic flux of the machine.

With equations (3-55) to (3-66), the dynamic behaviour of the induction machine can be analysed. The electrical torque of the system is calculated using equation (3-67).

$$T_e = \frac{3}{2} p(\psi_{ds} i_{qs} - \psi_{qs} i_{ds}) \quad (3\text{-}67)$$

Based on equations (3-55) to (3-66), the active and reactive power *P* and *Q* of the machine can be found by:

$$P = \frac{3}{2}(v_d i_d + v_q i_q) \quad (3\text{-}68)$$

$$Q = \frac{3}{2}(v_q i_d - v_d i_q) \quad (3\text{-}69)$$

During steady state condition, the power balance equation can be expressed as:

$$P_{loss} = P_{elec} - P_{mech} \quad (3\text{-}70)$$

where P_{loss}, P_{elec} and P_{mech} are the power loss, active power and mechanical power, respectively. As the electrical power consists of stator power P_s and rotor power P_r, together with the assumptions given earlier (a lossless machine is considered), the power transfer equation between the stator-rotor air gap for a lossless induction machine is:

$$P_r = -sP_s \quad (3\text{-}71)$$

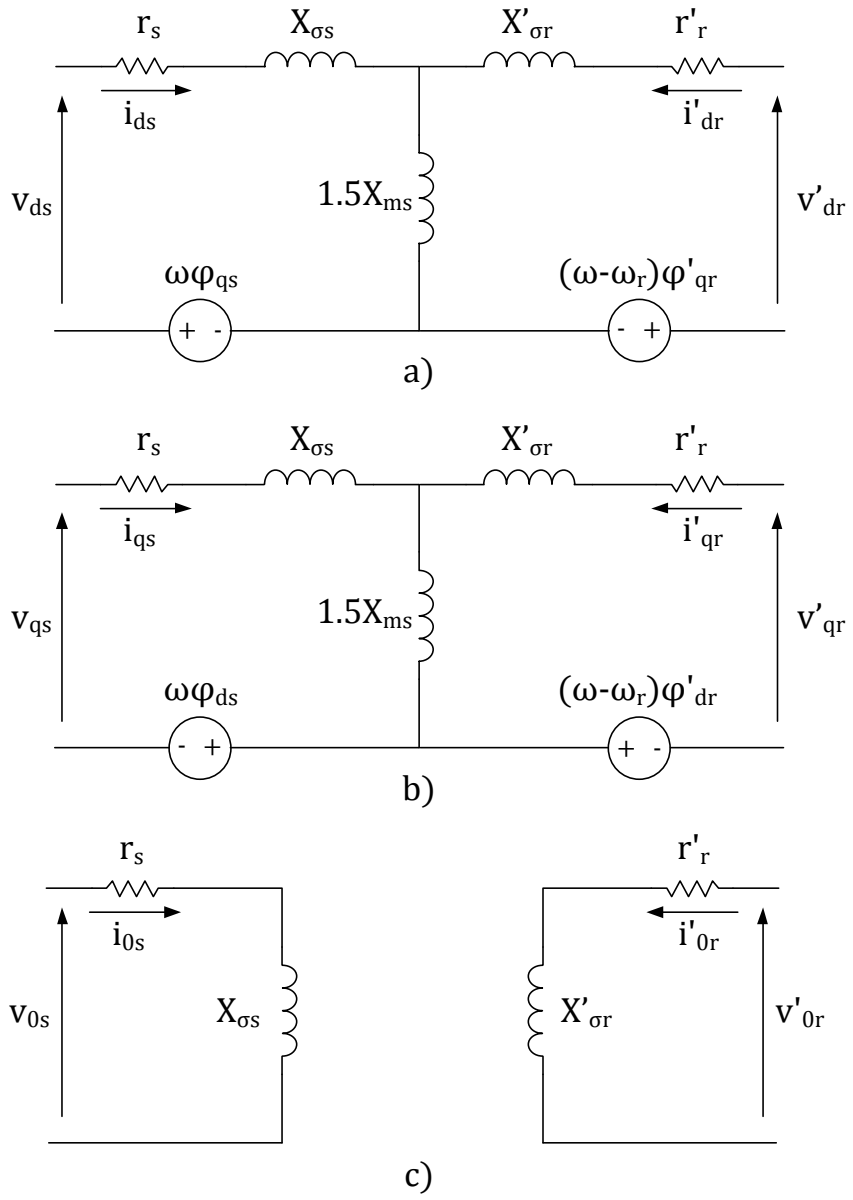


Figure 3.16 - Equivalent circuit model of wound rotor induction machine in *dq0* arbitrary frame: a) *d*-axis; b) *q*-axis; c) *0*-axis

where the negative sign indicates an opposite direction of power flow to the convention defined, and *s* is the slip of the generator and is:

$$s = \frac{\omega_s - \omega_r}{\omega_s} \tag{3-72}$$

where ω_s is the synchronous speed of the machine at 50 Hz. Based on equations (3-71) and (3-72), it can be seen that there are several operation modes for the induction machine, and it is dependent on the rotor speed of the machine, as depicted in Table 3.2.

Operation mode	Motor mode (supplied from main, $P_s > 0$)	Generator mode (fed to main, $P_s < 0$)
Super-synchronous speed ($\omega_r > \omega_s$)	Supplied from main $P_r > 0$	Fed to main, $P_r < 0$
Sub-synchronous speed ($\omega_r < \omega_s$)	Fed to main, $P_r < 0$	Supplied from main, $P_r > 0$

Table 3.2 - Power flow under different operation modes of an induction machine

During motor mode and with motor convention, as defined earlier, power is supplied from the main (P_s>0). Rotor power is supplied from main (P_r>0) when the rotor speed is greater than the synchronous speed and the rotor delivers power to the main (P_r<0) when the speed is smaller than the synchronous speed. Similarly, for the generator mode where power is fed to the main (P_s<0), during sub-synchronous speed, the rotor outputs power (P_r<0) and during super-synchronous speed, the rotor absorbs power from the main (P_r>0). Because of this unique property of an induction machine, with appropriate control strategies, the DFIG can be realised [108].

3.1.2.2 Partial rated power converter and control

A schematic diagram of the DFIG wind turbine with a power converter is shown in Figure 3.17. Unlike in the PMSG, in which the power converter has to handle the full power rating of the generator; for the DFIG, the converter is placed on the rotor circuit, which typically handles 30% of the rated power of the generator. Again, the converter consists of a rotor side converter, a DC-link and a grid side converter [111]. Each of these components and their respective control is discussed in later sections.

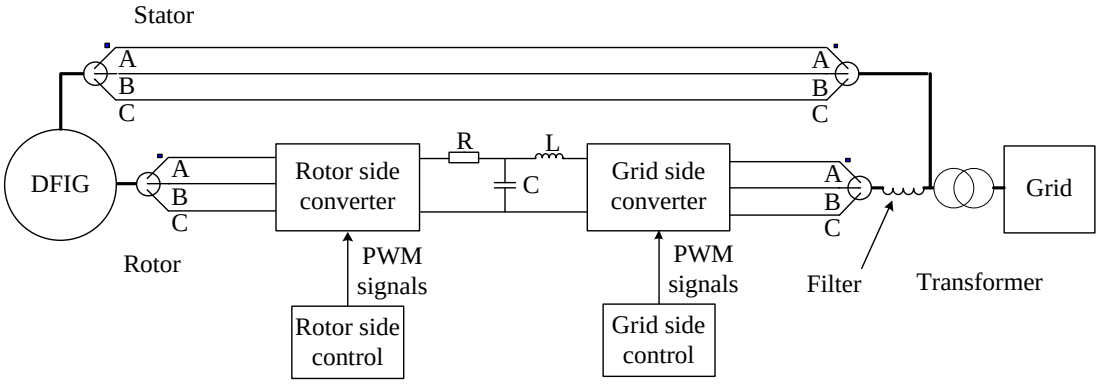


Figure 3.17 - Schematic of the DFIG wind turbine with power electronic converter

The main purposes of the rotor side converter are to provide excitation to the rotor and to achieve active and reactive power control. The control block diagram of the rotor side converter is shown in Figure 3.18 a). It consists of two consecutive PI controllers, where the outer PI loop controls the active and reactive power flow. The *d*-axis reference rotor current is forced to 0 and controls all reactive power required by the generator. This reference current i_{rd}^{ref} is obtained from the PI controller providing the error of reactive power between reference and actual value. The active power is controlled by the *q*-axis reference rotor current, where it is obtained from the error signal of the actual and reference rotor speed obtained from maximum power point tracking with the PI controller. The inner control loop is aimed to produce an optimal rotor excitation current to the induction machine by controlling the rotor current. The *dq*-axes rotor voltage v_{rd}' and v_{rq}' are obtained from the *dq*-axes current error signal in association with the PI controller. The decoupling of the *dq*-axes rotor voltage is obtained by compensating terms Δv_{rd} and Δv_{rq}, and achieving independent control of active and reactive power.

$$\Delta v_{rd} = \sigma(\omega_s - \omega_r) L_r i_{rq} \tag{3-73}$$

$$\Delta v_{rq} = (\omega_s - \omega_r)(L_o i_{ms} + \sigma L_r i_{rd}) \tag{3-74}$$

$$\sigma = 1 - \frac{L_{ms}^2}{L_s L_r} \tag{3-75}$$

$$L_o = \frac{L_{ms}^2}{L_s} \tag{3-76}$$

where σ and L_o are the leakage factor and equivalent inductance; i_{ms} is the stator flux current in the stator flux orientation. L_s and L_r are the stator and rotor self-inductances ($L_s = L_{ms}+L_{\sigma s}$ and $L_r = L_{ms}+L_{\sigma r}$). Finally, the reference *dq*-axes voltage is converted back to three-phase voltage and the switching signals for the IGBTs are obtained with the use of pulse width modulation (PWM).

The grid side converter controls the flow of active and reactive power, the power factor and the magnitude of the rotor power disregard of the direction of flow; the grid side converter also maintains the DC-link voltage. Moreover, the converter output is synchronised to match the grid standards. The control block diagram of grid side converter is shown in Figure 3.18 b). It also adopted a

double closed loops decoupling vector control strategy. The outer control loop stabilises the DC-link voltage and determines the *d*-axis reference current i_{sd}^{ref}, which is then used to control the flow of active power. The inner loop forces the power factor to unity by controlling the *q*-axis current i_{sq} to the reference value i_{sq}^{ref}. These current terms are then used to produce a reference voltage (v_{sd}^{ref}, v_{sq}^{ref}) through decoupling and compensation with equation 3-77 and 3-78.

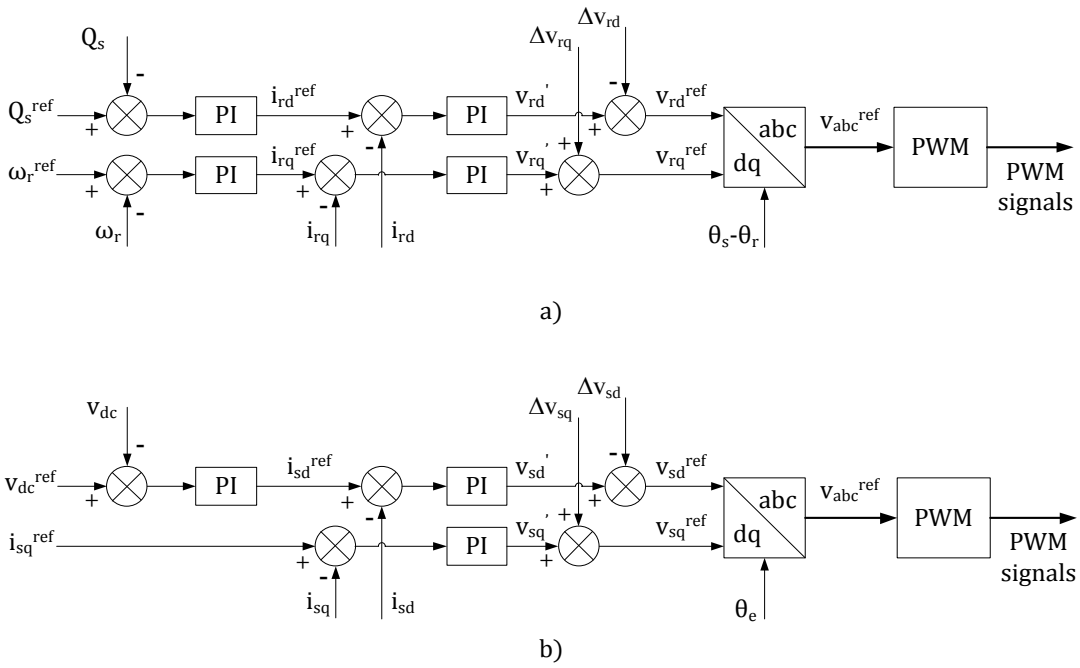


Figure 3.18 - Vector control block diagram of power converters: a) rotor side converter control; b) grid side converter control

$$\Delta v_{rd} = \omega_e L i_{sq} + v_{sd} \tag{3-77}$$

$$\Delta v_{rq} = \omega_e L i_{sd} \tag{3-78}$$

where ω_e is the angular velocity of the grid voltage and *L* is the inductance of the filter. Based on the reference voltages, the switching signals are then produced with PWM. In this model, switching dynamics of the IGBTs are not included in the model. During a steady state condition, the stator voltage is constant, and the active power and reactive power are proportional to the *dq*-axes current. Detailed description and derivation of the controller can be found in [109].

$$P = \frac{3}{2} v_s i_{sd} \tag{3-79}$$

$$Q = \frac{3}{2} v_s i_{sq} \tag{3-80}$$

3.1.3 Simulation results and discussion

The wind turbines are simulated in PSCAD/EMTDC. PSCAD/EMTDC is a general-purpose industrial standard simulation program with a graphical interface for studying the transient behaviour of complex electrical networks. The software allows a flexible time step ranging from nanoseconds to seconds to simulate electromagnetic transients in the electrical network. Also, the graphical environment enables the user to analyse the results and manage the data accordingly. Benchmark models of both the PMSG and DFIG wind turbines in PSCAD/EMTDC are adopted and the simulation results are given in the following sections [112, 113].

3.1.3.1 PMSG wind turbine

Based on the model discussed in Section 3.1.1, and using the model developed by PSCAD/EMTDC as a basis, the complete PMSG wind turbine simulated in PSCAD/EMTDC is shown in Figures 3.19-22. The turbine has a rated power of 3 MW at 13 m/s and a cut-in and cut- out speed of 4 m/s and 25 m/s, respectively. The dynamic pitch control is adopted, where the pitch angle is set to zero when the wind speed is below the rated value. The generator has a rated phase-to-phase voltage of 690 Vac at 50 Hz and parameters associated with the generator as described in equations (3-41) to (3-50) are given in the Appendix A.3. The DC-link has a rated value of 1613.9 V and upper and lower limits of 1760 V and 1440 V, respectively. The upper limit acts as over-voltage protection, as the output voltage of the generator is proportional to the speed, and there is no speed control for the generator. The lower limit is responsible for short-circuit protection in case of a grid fault to secure the DC bus.

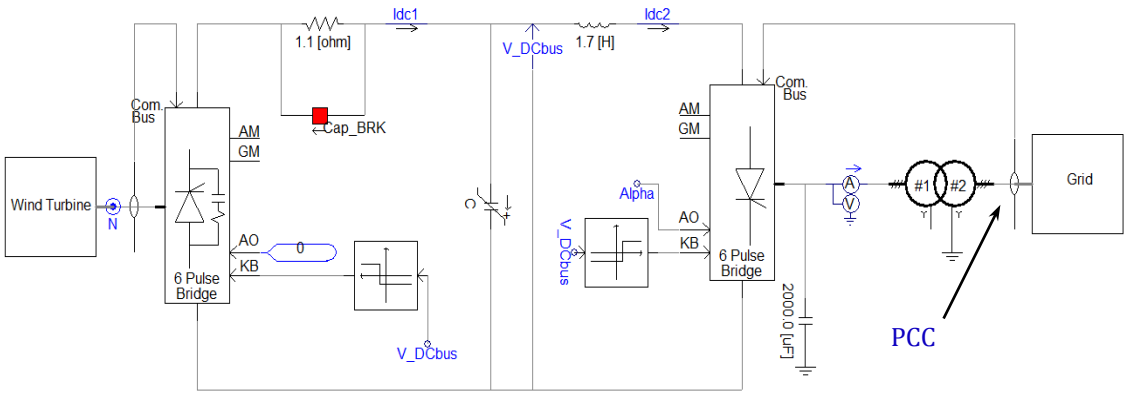


Figure 3.19 - PMSG wind turbine model with an electrical network connection

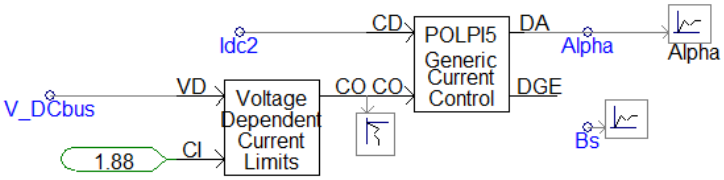


Figure 3.20 - Controller of the AC-DC-AC converter for the PMSG wind turbine model.

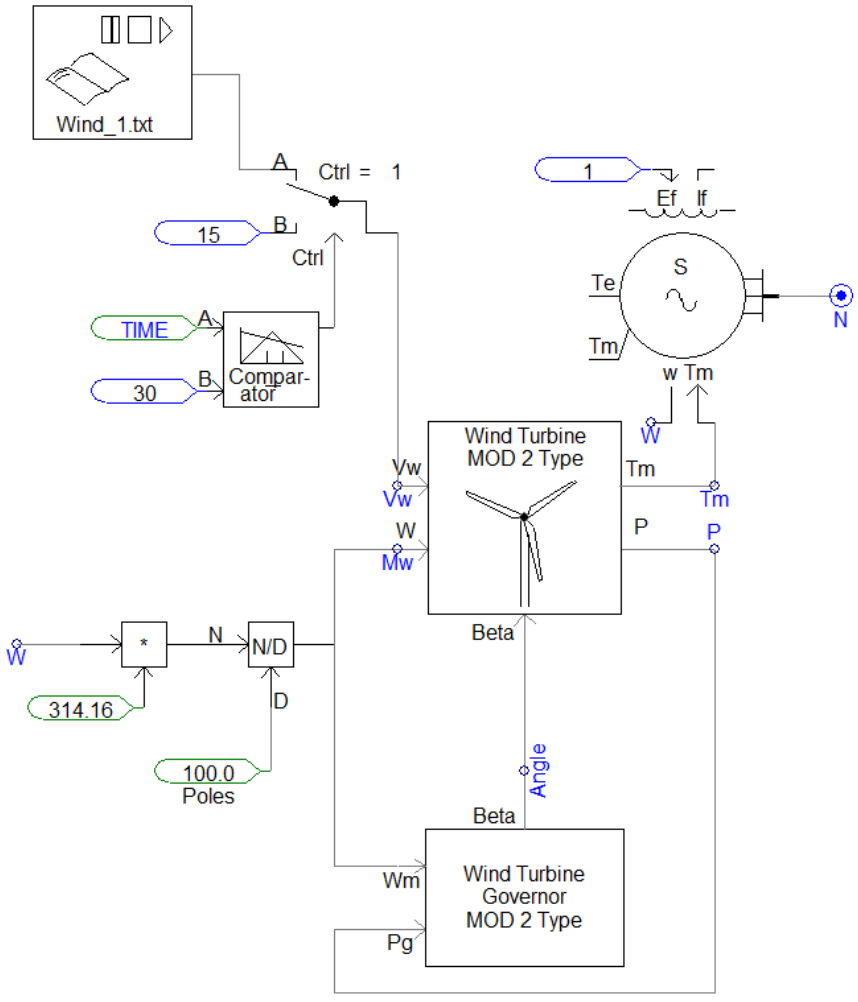


Figure 3.21 - PMSG wind turbine power generation model

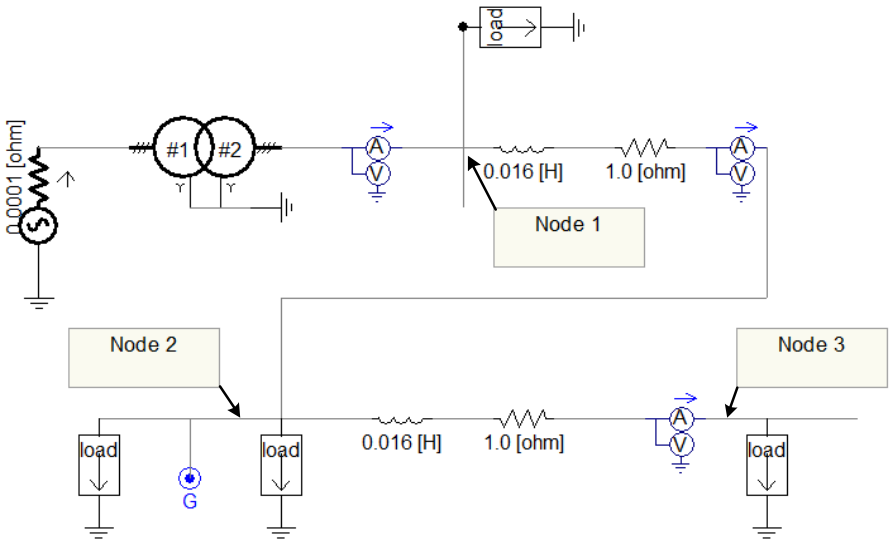


Figure 3.22 - Grid network model of the PMSG wind turbine

A simple radial distribution model with 34.5kV/300MVA is considered to simulate the grid network as shown in Figure 3.22. Several loads are connected at different locations of the grid model, where it is labelled as nodes 1, 2 and 3. The transmission losses between each node are considered, and it is modelled with equivalent resistance and inductance. The output voltage drop, due to losses of the grid, behaves the same as described by:

$$\Delta V = RI\cos\varphi + LI\omega\sin\varphi \tag{3-81}$$

where I is the grid current, L and R are the equivalent transmission line inductance and resistance, ω is the angular velocity of grid frequency and φ is a leading or lagging phase angle. Moreover, a load with a power rating of 2.133 MW/1.6 MVar is connected at node 1; loads with ratings of 0.266MW/0.2MVar are connected at nodes 2 and 3 as shown in Fig 3.22. The simulation is run for 100 seconds and has a time step of 100 μs. The simulation results are presented and discussed.

As an example, some of the outputs from the simulation are plotted in Figure 3.23. From top to bottom, it includes the input wind speed, the pitch angle of the blade, the mechanical torque from the turbine and electromechanical torque from the generator, the mechanical power generated from the turbine, the DC-link voltage and currents, active and reactive power at the common point of coupling (PCC as shown in Fig 3.19), a thyristor firing angle and grid voltages at

different locations as described earlier. A constant wind speed of 13 m/s is simulated first. It can be seen that the pitch angle remained constant as the input wind speed is not above the rated value. Moreover, the first 40s of the plot show the wind turbine's starting-up behaviour.

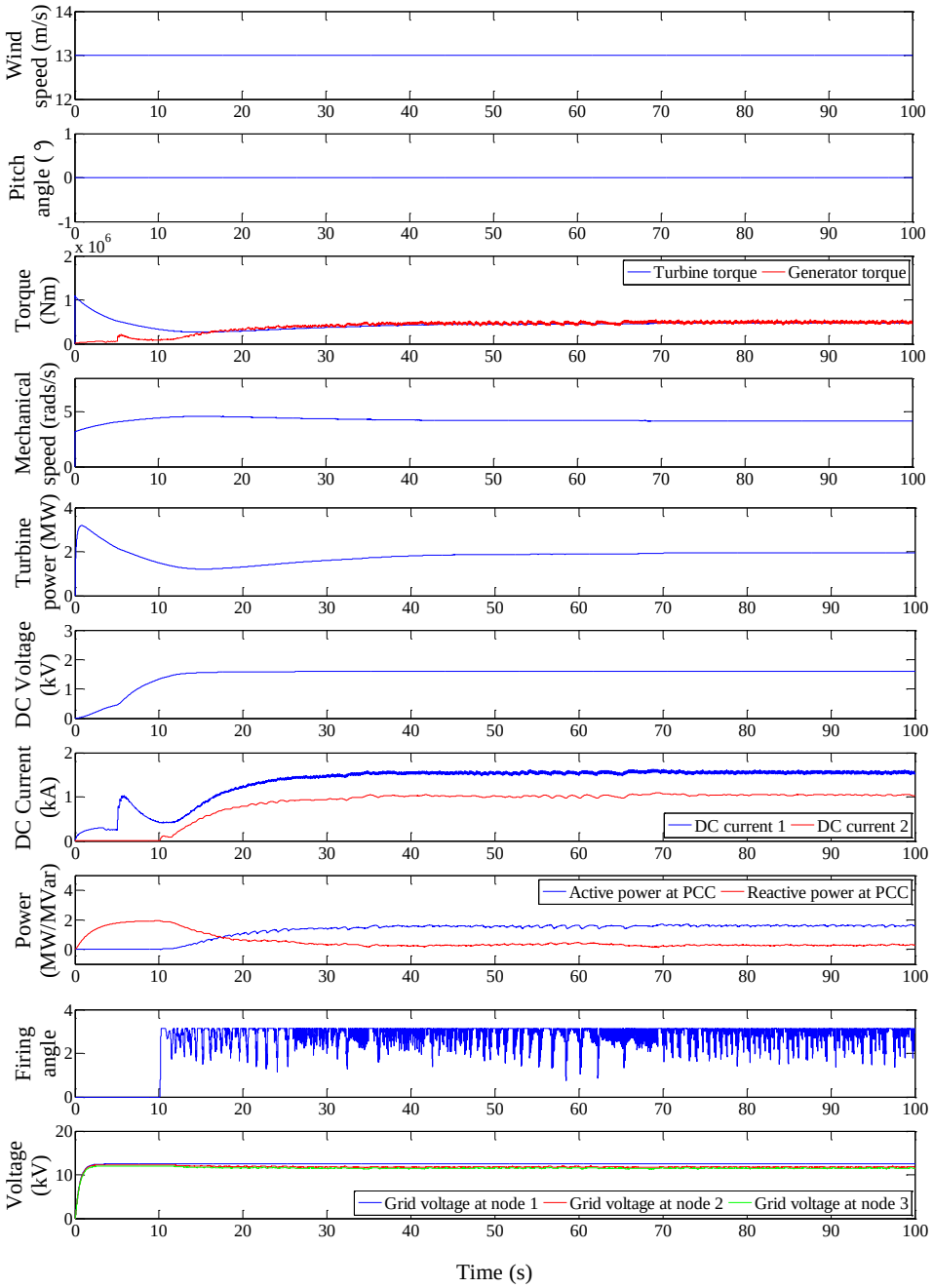


Figure 3.23 - Simulation result of the PMSG WT under constant wind speed

It is evident that the mechanical torque and power have a high initial value and stablise with time. Moreover, between 0-5 s, the DC-link capacitor is under a pre-charging condition and no power is generated. The DC-link voltage starts to increase between 5-10 s, until the voltage stablises and reaches the threshold value and the inverter starts to output power to the grid. It then takes approximately 20 s for the system to reach the steady state and deliver constant power. Results show the active power and reactive power at the grid connection are approximately 2 MW and approximately zero MVar, respectively. It should be noted that this value is smaller than the rated value even though the input wind speed is at 13 m/s. This is because the generator speed control is not possible with uncontrollable rectification, such that MPPT is not included in the model. Hence, the power coefficient for the turbine blades is not optimal. Finally, it can be seen in the last plot that the grid voltages at different nodes have minimal losses, which is determined by equation (3-81).

To further test the simulation, real wind measurements from the Hazelrigg site at Lancaster University have been used. Constant wind speed was used during the start-up period and then switched to real wind data. Results from the simulation are shown in Figure 3.24. It can be seen that the dynamic pitch controller in the model allows blade pitch regulation, such that when the mechanical turbine power is greater than the rated value, the power output is reduced by increasing the pitch angle. It should be noted that when the fluctuation of the input wind speed is too large, the response of the turbine torque and power is not optimised, where overshoot and oscillation can be identified. This is mainly because the pitch controller is not tuned for such high wind speed fluctuation. As the optimisation of the controller is not the focus of this study, and the results are still able to reflect the dynamic behaviour of the wind tubine, the simulation results are assumed to be adequte to validate the proposed variable selection algorithms in later chapters. Moreover, it can be seen from the second and fifth plots in Figure 3.24 that there is a slight delay for the pitch angle. This is due to the presence of a pitch actuator mode in the simulation.

Furthermore, the sixth plot shows that the controller is able to maintain the DC-link voltage constant, resulting in a stable output voltage from the inverter. In

contrast, the mechanical speed, power outputs and DC-link currents (as shown in the 7th and 8th plots in the Figure 3.24) are stongly related to the wind speed with a minor delay. The output active power at the grid connection varies from 2MW to the rated value depending on the wind speed. On the other hand, the reactive power is compensated by the capacitor before connecting to the grid, resulting in minimal reactive power to be injected into the grid.

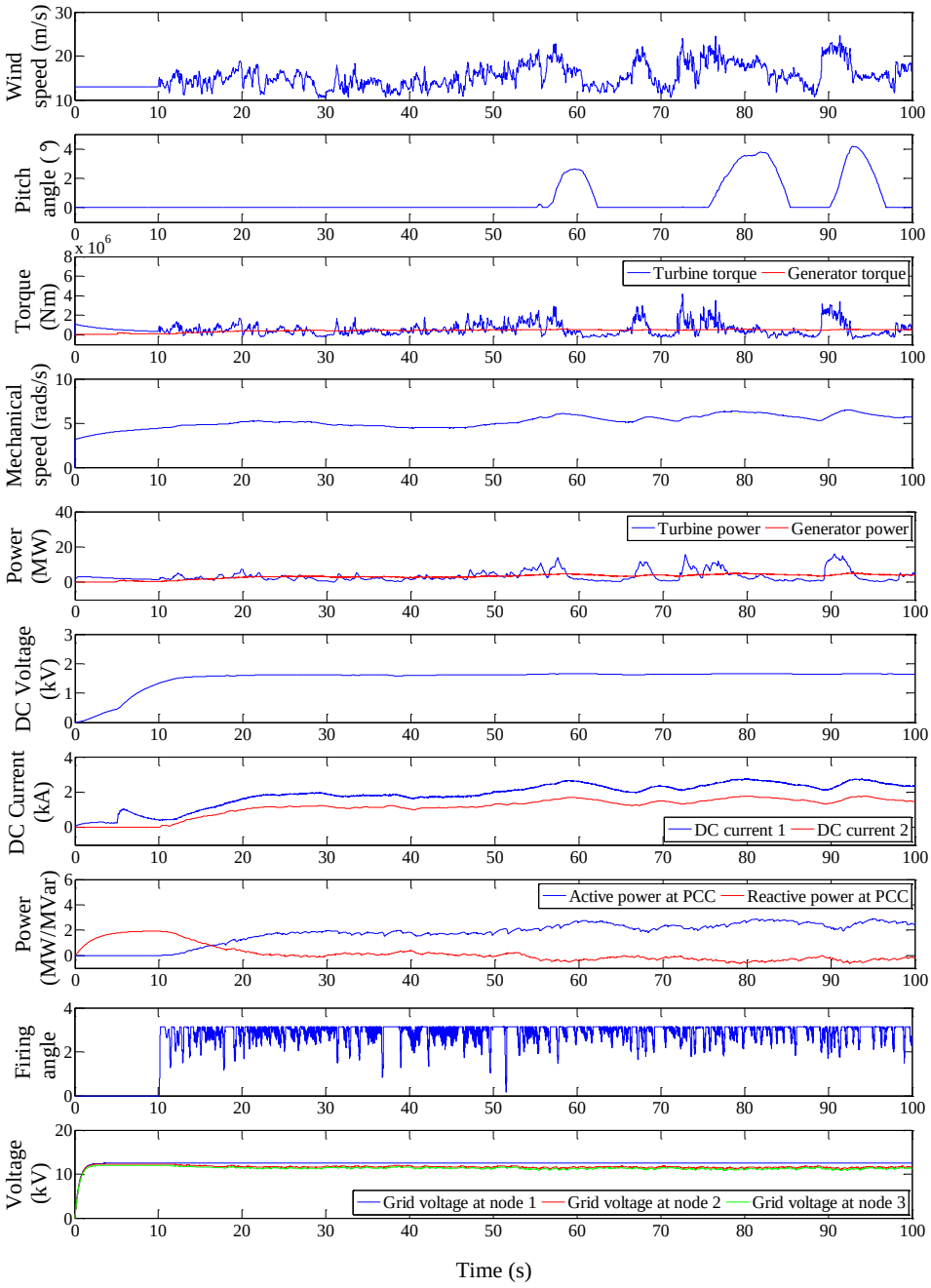


Figure 3.24 - Simulation result of the PMSG WT with real wind speed measurement from the Hazelrigg site at Lancaster University.

3.1.3.2 DFIG wind turbine

With the theory from Section 3.1.2 and a base model from PSCAD, the complete DFIG wind turbine model simulated in PSCAD/EMTDC is shown in Figures 3.25 and 3.26. The simulated system has a rated power of 2 MW at 12 m/s and a rated stator phase to phase voltage of 690 V at 50 Hz. The output of the wind turbine is connected to the grid with a rated power of 500 MW at 20 kV/50 Hz. The parameters of the wound rotor induction machine are shown in the Appendix A.3. As the generator stator and rotor windings have a ratio of 0.3, this causes the rotor voltage at a different rating. Therefore, a transformer with a star-star configuration is used to avoid a difference in phase windings voltage. The back-to-back converter is added in the rotor circuit, adjusting the supply to the rotor with appropriate controls. The detailed control for the rotor side and grid side converters is shown in Figures 3.27 and 3.28.

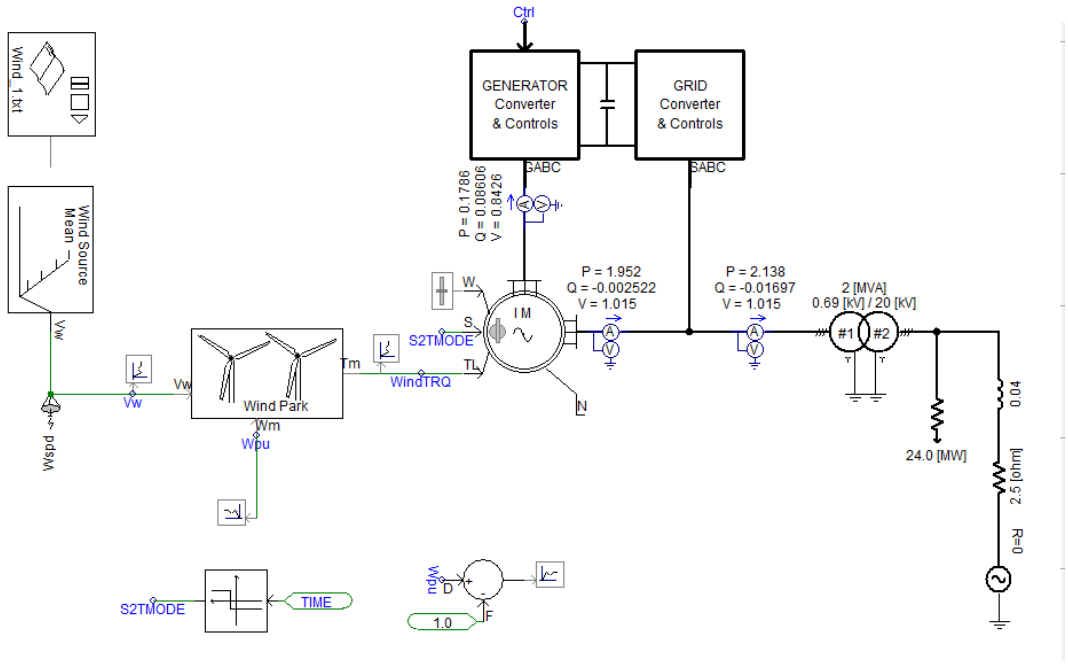


Figure 3.25 - DFIG wind turbine model with an electrical network connection

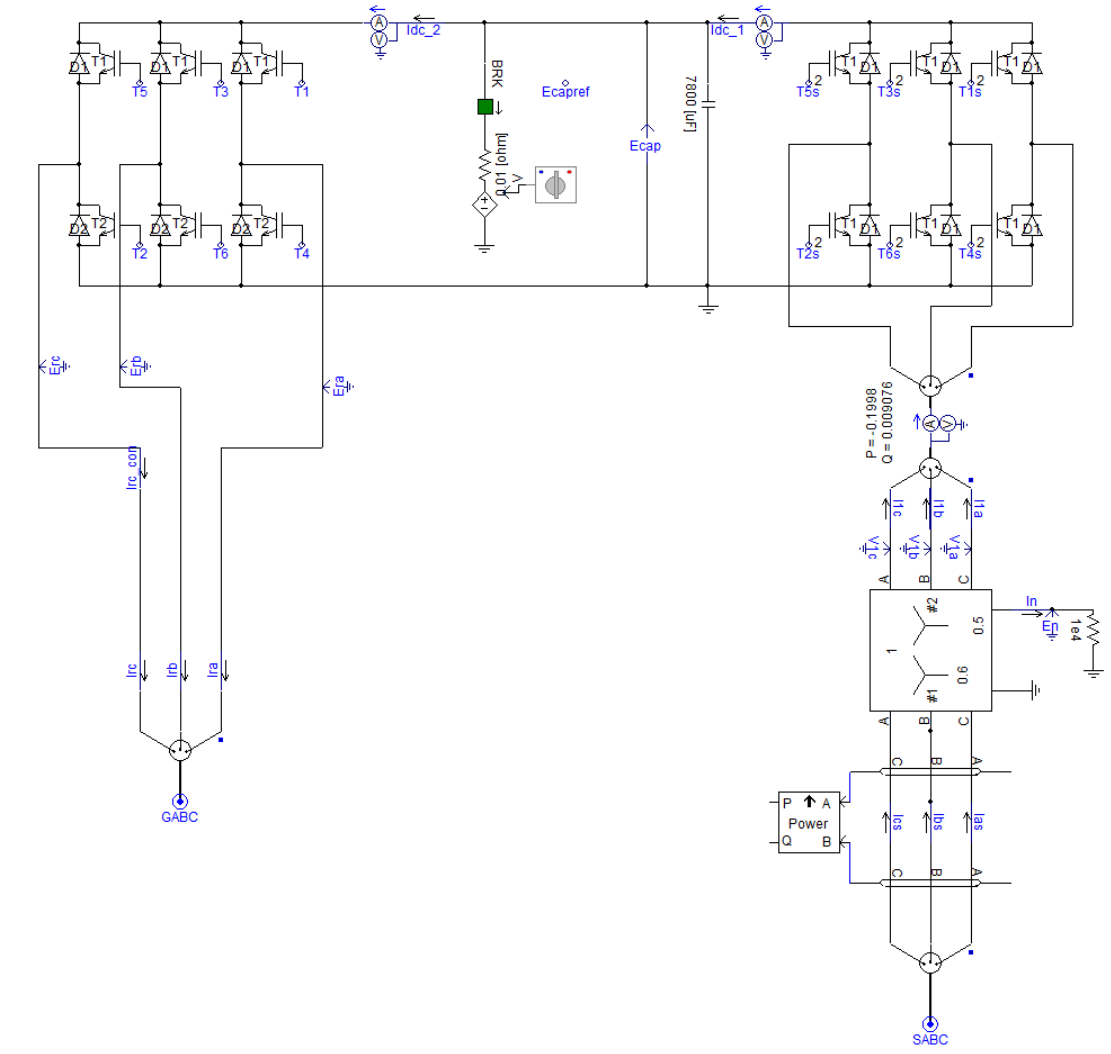


Figure 3.26 - DFIG wind turbine power electronics model

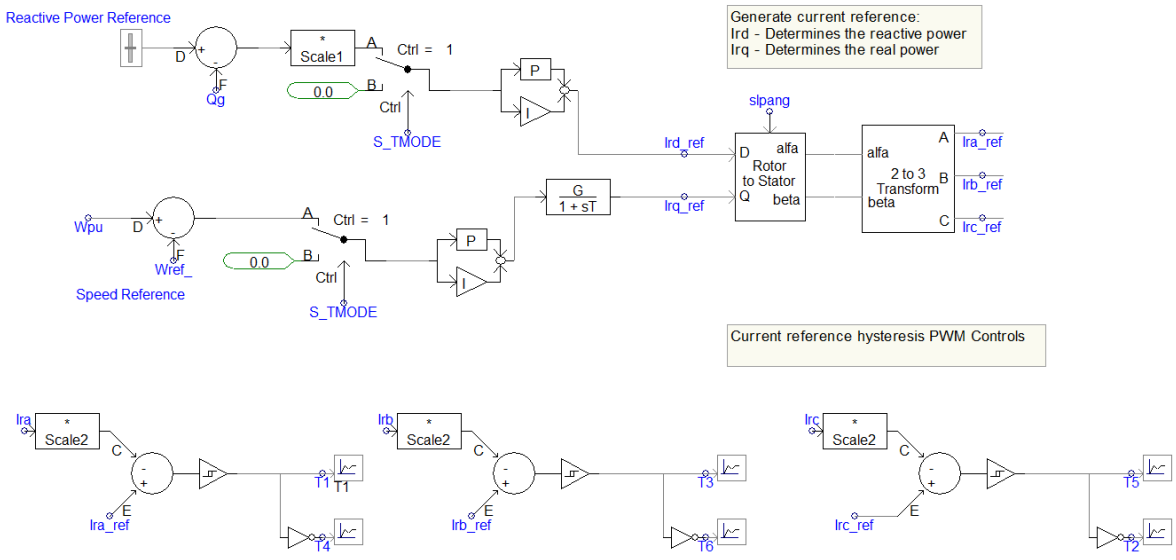


Figure 3.27 - Rotor side inverter control algorithm

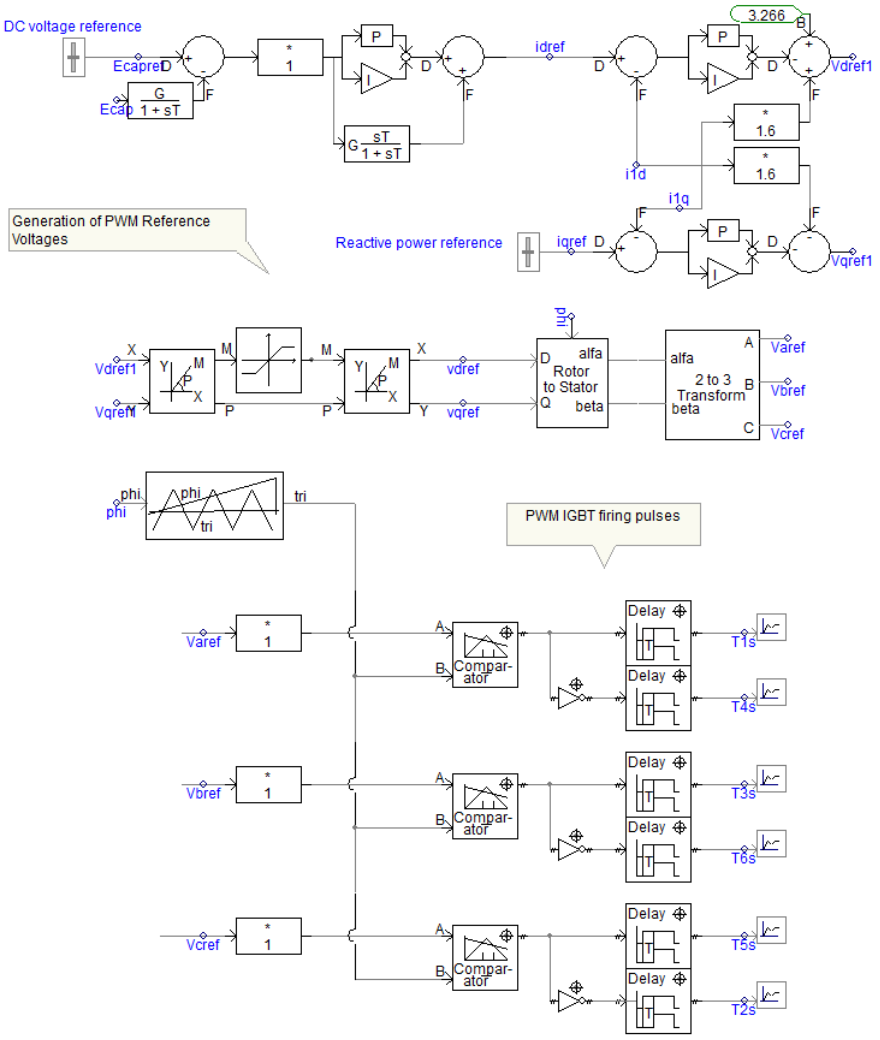


Figure 3.28 - Grid side inverter control algorithm

For the rotor side converter, the controller independently produces the reference current of the rotor in the *dq0* frame based on the reactive power reference and speed reference. In our simulation, the reactive power reference is set to 0, and the speed reference is obtained based on wind speed. The reference current I_{rd_ref} and I_{rq_ref} is then transformed back to the *abc* frame. By incorporating the hysteresis controller, the switching sequences for the rotor side IGBT converter can be generated.

The grid side controller uses PWM to produce the required switching sequences for the IGBT converter with phase lock loop (PLL) integration, where the PLL is aimed to synchronise the grid side converter output to the grid standard. The reference voltage used to produce the PWM waveform is obtained by controlling

the DC-link voltage and the reactive power. The desired DC link voltage is at 0.8 kV, and the reference reactive power is set to 0 MVar.

A constant wind speed of 12 m/s is used in the simulation at a time step of 100 μs for 20 s. Figure 3.29 shows the wind speed, mechanical and generator torque, the mechanical and reference mechanical speed of the generator, the active and reactive power of the stator, the rotor and at point of coupling, the DC-link currents and voltage, the rotor currents, the slip and the voltage at PCC.

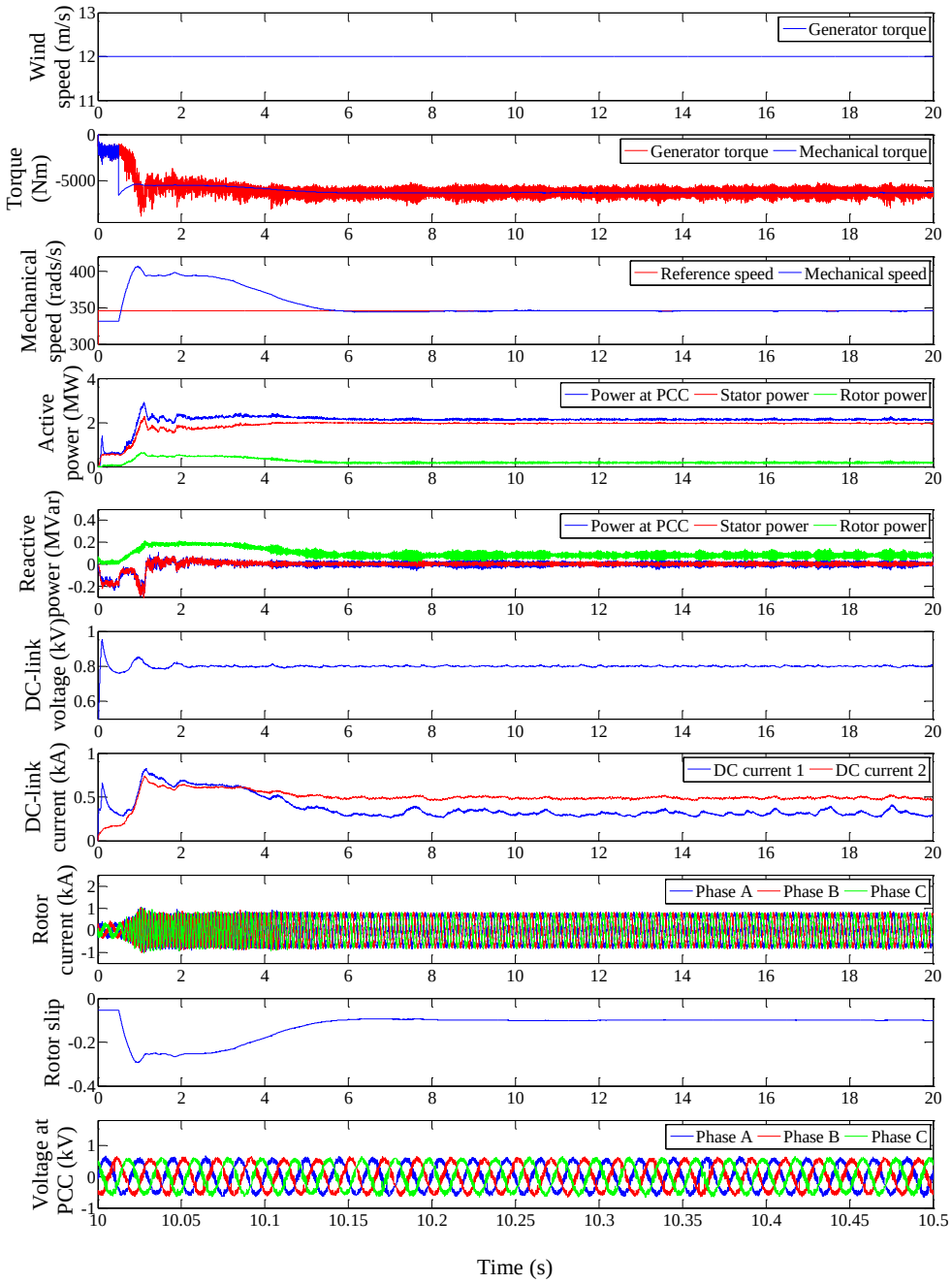


Figure 3.29 - Simulation result of the DFIG WT under constant wind speed

The system stabilises at approximately 5 s. As expected, at 12 m/s, the machine is operating at its rated power, where the active power and reactive power measured at the PCC are 2 MW and 0 Mvar, respectively, and a stable phase voltage at the PCC of 400 V/50 Hz is measured. Moreover, it can be seen that the mechanical speed is above the synchronous speed, which results in a negative slip of -0.1 at the steady state condition.

The DFIG model is simulated again with a wind speed sweep from 8 m/s to 20 m/s. The wind speed in set to 8 m/s for the start up and until it reaches the steady state condition; then, at 10 s, the wind speed sweep is applied. Figure 3.30 shows the dynamic behaviour of the mechanical speed, slip, power and torque of the generator. As can be seen, as the wind speed increases, the angular speed of the generators increases; hence, the slip starts to decrease. The rotor power changes from negative to positive as the slip changes from sub synchronous speed to super synchronous speed, which means the power is supplied to the rotor at low wind speed, and, at high wind speed, the rotor is delivering power to the grid. Moreover, the stator power increases until the rated wind speed is reached, at which time the rated generator power of 2MW is obtained.

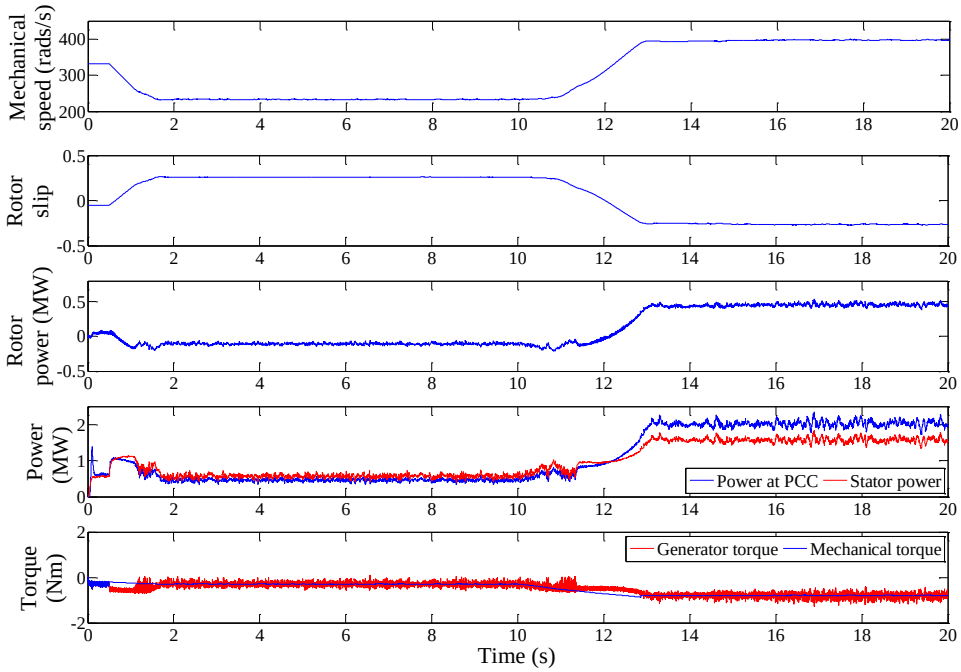


Figure 3.30 - Behaviour of machine power and torque under a varying rotor slip.

To test the model further, actual wind speed data from the Hazelrigg site at Lancaster University is used in the simulation (Figure 3.31). An initial wind speed

of 11 m/s is set and then switched to an actual wind speed at 10 s. As can be seen, as the wind speed changes, the mechanical speed, slip, torque and power vary accordingly. However, the DC-link voltage remains constant at 800 V. Moreover, the frequency of rotor currents changes with respect to the slip. Between 5 to 7 s, the slip is approximately 0, which means the DFIG operates as a synchronised asynchronous generator with a continuous rotor voltage and current.

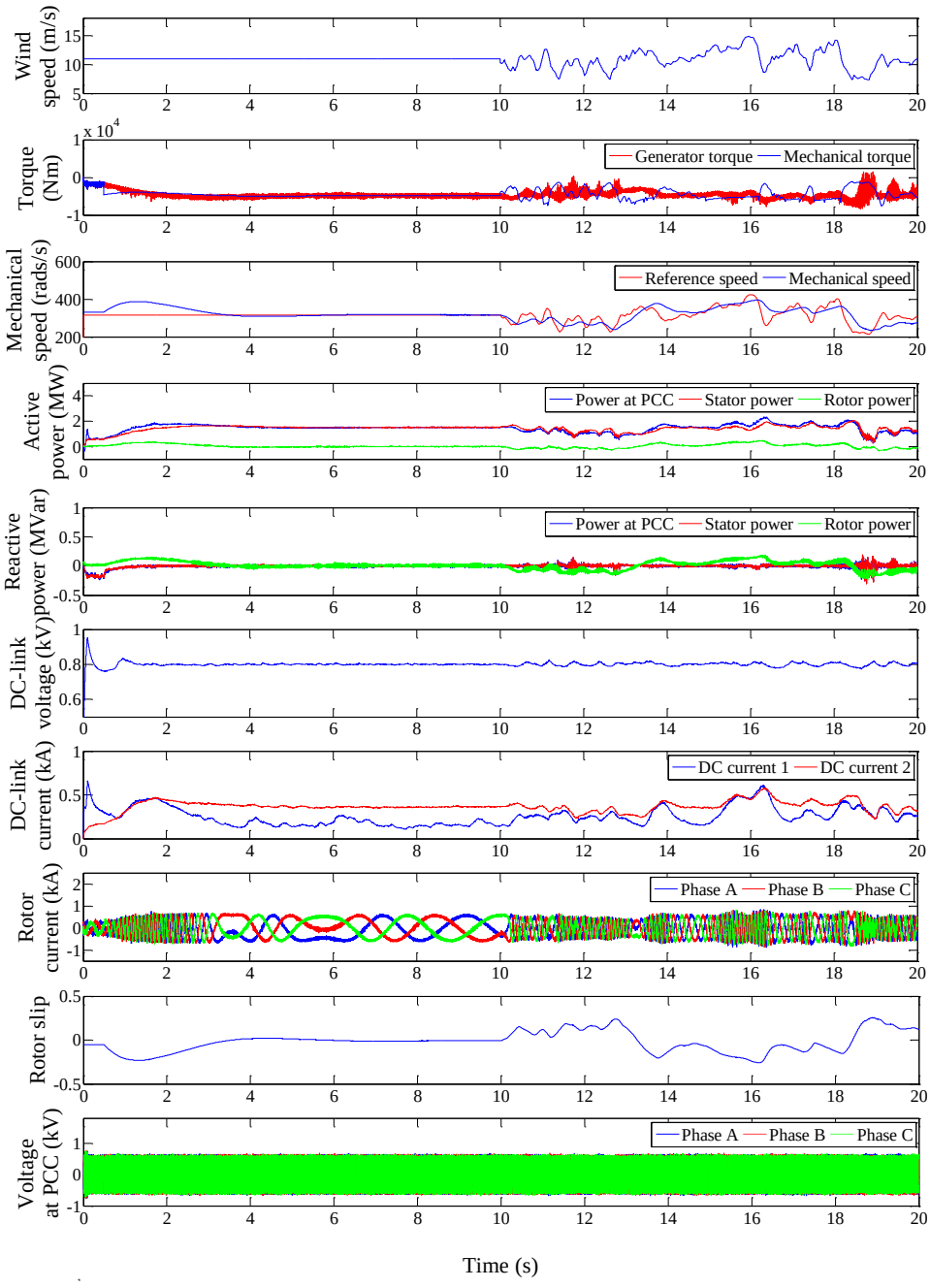


Figure 3.31 - Simulation result of the DFIG WT with real wind speed measurement from the Hazelrigg site at Lancaster University

3.2 Simulation of wind turbine in abnormal conditions

In the previous section, PMSG and DFIG wind turbines under normal operating conditions have been simulated. The results for each model have been discussed in detail. However, normal operation data is less abundant in the condition monitoring perspective. The transient responses and performances of wind turbines under faulty conditions will be more valuable for further condition monitoring and fault detection studies. The common failures of a wind turbine system have been reviewed in Chapter 2. In our study, two electrical faults are simulated to study their impact on the whole system.

3.2.1 DC-link capacitor ageing faults in power converters

For modern multi-megawatt turbines, power converters have become imperative component due to their advantages. Power electronics modules such as IGBT generally have acceptable reliability. On the other hand, DC-link capacitors are considered one of the weakest components used in multi-MW power converters in the wind turbine [114]. The DC link capacitors are required to sustain high ripple currents leading to self-heating. This can result in the deterioration of the electrolyte material and the loss of electrolytes by vapour diffusion, especially in high ambient operating temperatures. When the capacitor is operating at higher temperatures than the rated temperature, the DC voltage will be de-rated. Furthermore, the working life of a capacitor is also dependent upon the operating voltage, current, and frequency. Consequently, failure of the DC-link capacitor can be crucial for reliable wind power generation.

A capacitor ageing fault is introduced by varying the capacitance during the simulation. The simulation is run for 200 s and the fault is injected at 100 s. Results from the simulation are shown in Figure 3.32, where data during the starting-up period (t<50 s) are neglected, as we are solely interested in the steady state condition. The original capacitor has a value of 2.3 F and at t = 50 s, the capacitance is reduced by 5% (2.18 F). It can be seen that there is a transient change when the fault is introduced in the DC link voltage plot. Moreover, this transient change is almost impossible to distinguish in other plots, as it is overlapped with the variation due to variable wind speed.

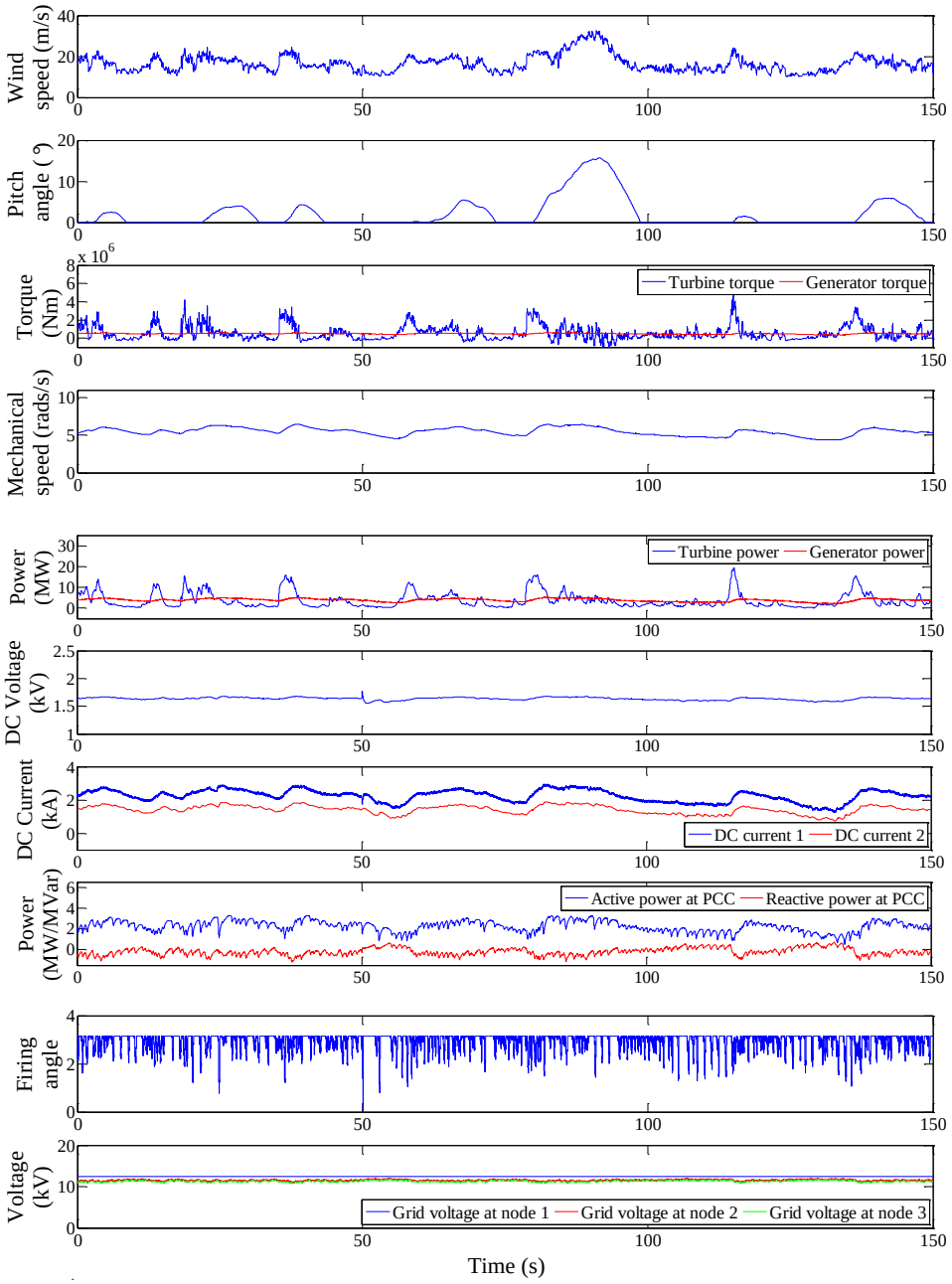


Figure 3.32 - Simulation result of a PMSG WT with real wind speed measurement under a capacitor ageing fault

The simulation is repeated with a varying of severity ranging from a 0% no fault condition to -16% (1.93 F). Any further increase of the fault severity will cause the simulation to be unstable. Figure 3.33 shows the transient change in the DC-link voltage and currents at different severities. It can be seen that both DC-link voltage ($V_{\_DCbus}$) and current 1 (I_{dc_1}) undergo apparent changes with higher severity. In contrast, the DC-link current 2 (I_{dc_2}) remains the same under

different fault level. This behaviour is expected as the DC-link current 2 is measured after the capacitor in the DC-link, as shown in Figure 3.19.

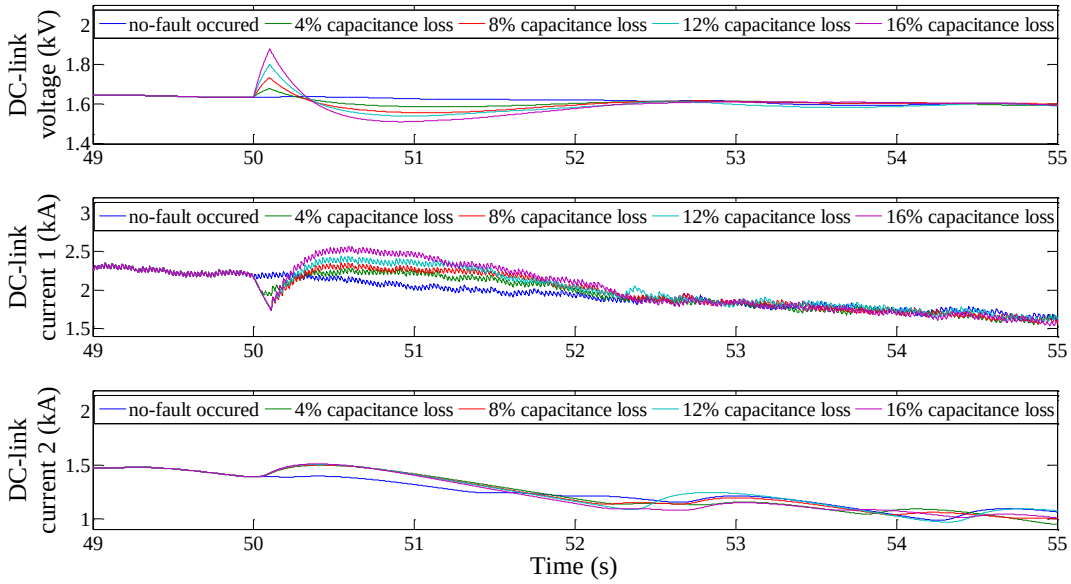


Figure 3.33 - DC-link voltage and current under a capacitor fault at different severities

3.2.2 Grid unbalanced ground short faults

One of the requirements for a wind turbine to have grid connection is that the turbine has to comply with the grid code. This means the turbine should be able to ride through transient faults from the grid. In this study, the unbalanced grid fault is considered, where the mutual effects and the dynamic behaviour of a wind turbine to the power system under a transient fault are studied. The single phase-to-ground short circuit fault, the phase-to-phase short circuit fault and the three phase-to-ground fault are simulated. Moreover, in our study, the aim is not to develop a voltage ride through strategies; rather, our aim is to simulate the fault and acquire data to validate proposed algorithms. Therefore, the severity of the fault is simulated by varying the resistance between the phase and the ground. During a fault-free condition, the resistance is assumed to be 1 MΩ, and the severity is then considered as a decrement of the percentage of resistance at the fault-free condition. Figure 3.34 shows the simulation result of a single phase-to-ground fault located at node 2, as shown in Figure 3.22 at 0.001% severity with a duration of 0.1 s introduced at 60 s. The 0.001% severity represents a fault resistance between a phase and the ground with value of 10 Ω

(0.001% of the normal condition 1 MΩ). A transient change in the grid current at nodes 2 and 3 is identifiable. Because the simulated grid has a large power rating compared to the wind turbine, the grid is stiff enough to withstand the fault without causing any voltage drop.

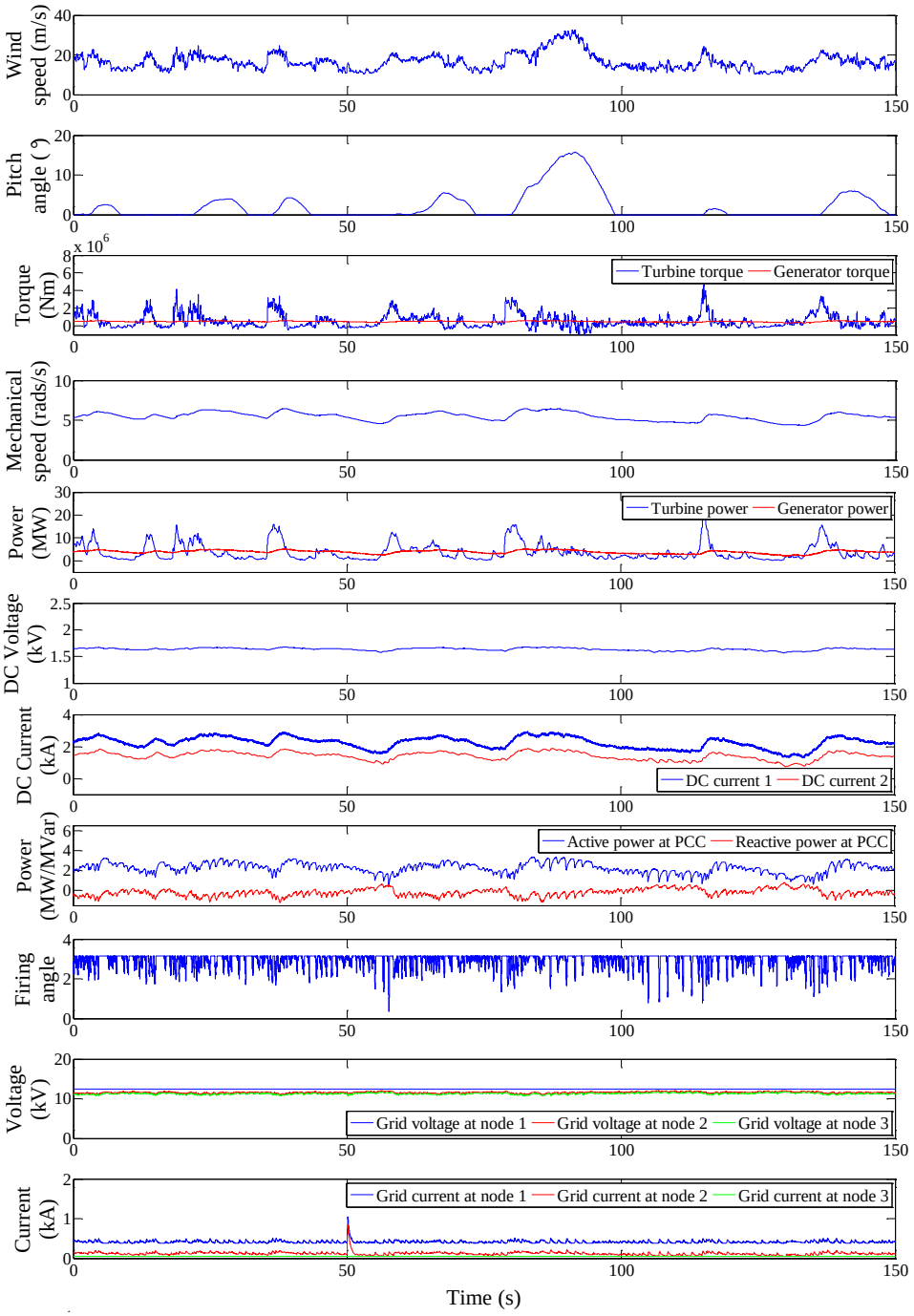


Figure 3.34 - Simulation result of a PMSG WT with a real wind measurement under 0.001% severity (10 Ω) and a 0.1 s duration single phase-to-ground fault.

Moreover, since a single phase-to-ground fault is simulated, only one of the phases has significant variation, as shown in the top plot of Figure 3.35. It can be seen that the fault phase current is tripled compared to the normal condition. Similarly, results for phase-to-phase fault and three phase-to-ground fault are shown in the middle and bottom plot of Figure 3.35, respectively.

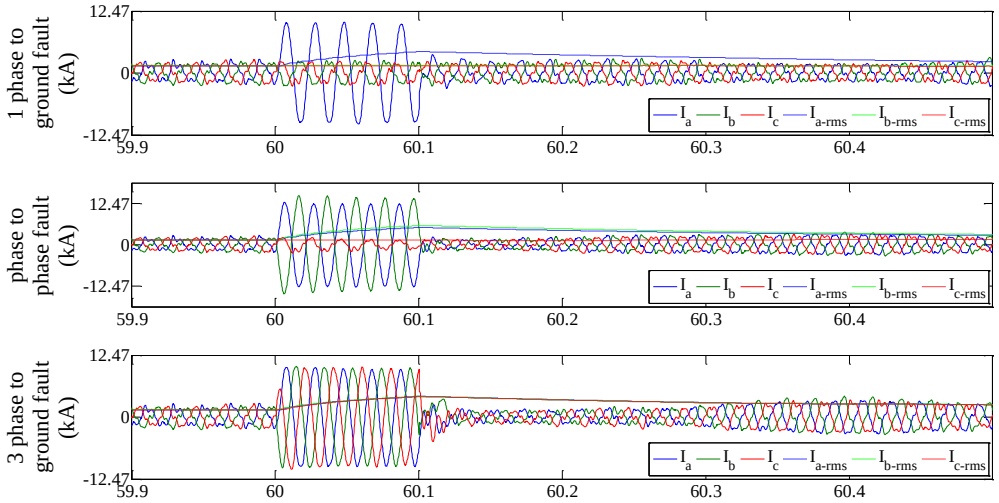


Figure 3.35 - Simulation result of a PMSG WT with a real wind speed measurement under 0.001% (10 Ω) severity and 0.1 s duration; a) single phase-to-ground fault; b) phase-to-phase fault; c) three phase-to-ground fault

Also, Figure 3.36 shows an example of a single phase-to-ground fault with varying resistance from fault free condition (1 MΩ) to 0.1 Ω. The behaviour of the grid current, voltage and active power under different fault severities can be seen.

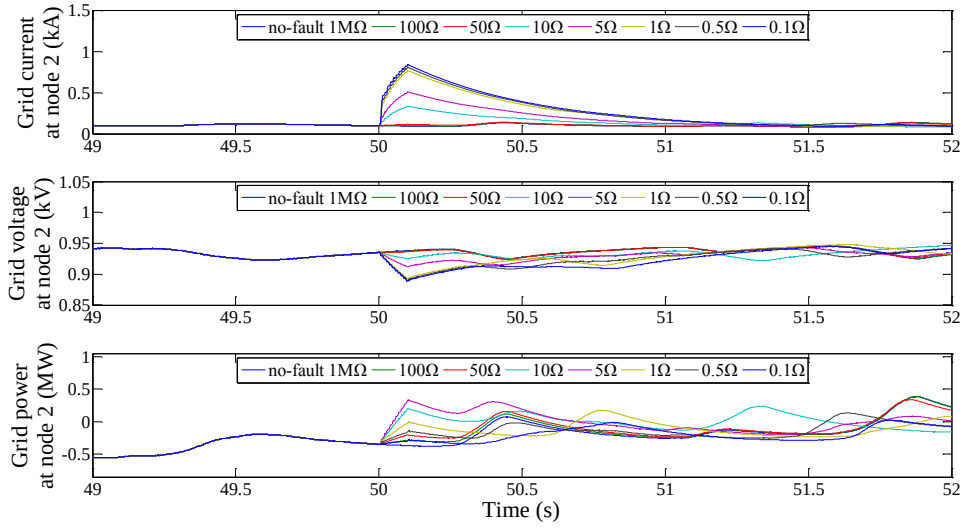
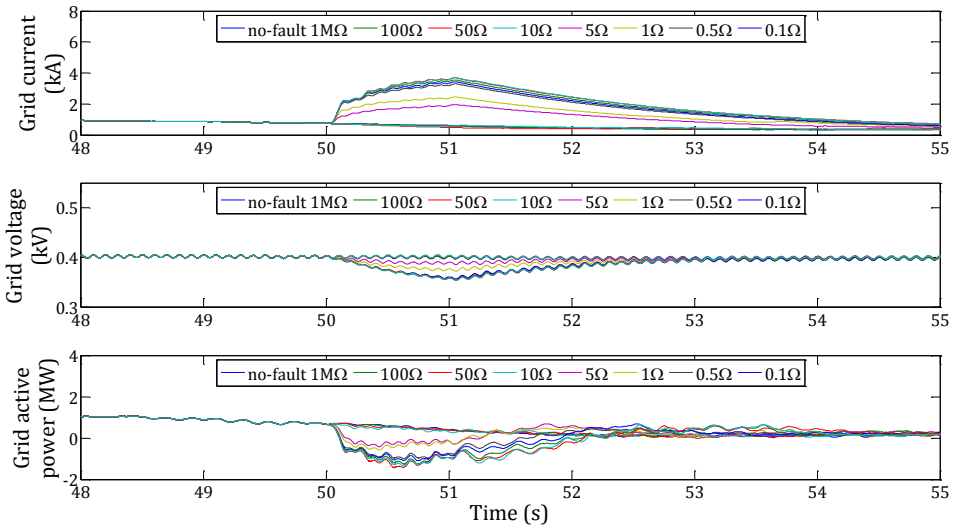


Figure 3.36 - Grid current, voltage and power under a single phase-to-ground fault simulation in a PMSG WT at different severities.

Lastly, as an example, simulation results including the grid current, voltage and active power of the phase-to-ground fault in the DFIG wind turbine model are shown in Figure 3.37, where resistance varied from 1 MΩ (fault-free condition) to 0.1 Ω and these values was used to simulate different fault severities.



Grid current, voltage and power of a single phase-to-ground fault simulation in a DFIG WT at different severities.

3.3 Summary and discussion

In this chapter, an analysis and modelling of both PMSG and DFIG variable speed wind turbines have been carried out. The models are simulated in PSCAD/EMTDC to study their dynamic responses under constant speed and real wind speed. During the modelling, certain assumptions were made to simplify the analysis, as the aim of the simulation was to obtain an adequate turbine model under different operation conditions and acquire relevant data to validate proposed algorithms in later chapters. Results show the simulation models are comprehensive. Moreover, two major electrical related fault conditions have been considered and simulated in our study. These are the DC-link capacitor ageing fault and the grid phase-to-ground fault. More results will be presented in the following chapters during the validation of the proposed algorithms. In the next chapter, the proposed variable selection algorithms are described. Validation of the proposed algorithms will be carried out with the data obtained in this chapter and from SCADA data obtained from an operating wind farm.

Chapter 4. Variable Selection based on Principal Component Analysis

For an operational wind turbine, hundreds of sensors are installed to monitor the status and condition of the system. However, there are inter-dependencies amongst these monitoring variables, which can create information redundancies. This will in turn increase the installation cost and create challenges when developing a condition monitoring system for wind turbines. Therefore, in this chapter, a multivariate feature extraction-based variable selection method is proposed to identify a set of variables contributing to the system most whilst with minimal information redundancies at the system level. Additionally, a modified method to identify a set of variables related to a particular fault signal of interest is presented. A basic theory of the employed methodologies is discussed first, and then the proposed variable selection algorithms and performance measures are described. The proposed algorithms are then validated against simulation data obtained from the previous chapter and SCADA data from an operational wind farm under different operational conditions.

4.1 Methodologies

4.1.1 Basic data mining techniques

Data mining is the analysis and exploration of meaningful patterns or systematic relationships between variables from a large set of data [115]. The concept was formed as a branch of applied artificial intelligence (AI) beginning in the 1960s [116]. With the development of computer systems, new data mining techniques have been introduced, particularly in the past decades. Both Friedman and Hand have demonstrated the relationship between data mining and statistics, placing emphasis on their similarities and differences [117, 118]. Major data mining techniques that have been developed include association rules, characterisation, clustering, decision trees, feature extraction techniques and neural networks. The application of these methodologies can be found in different disciplines, including the social sciences, management science, economics and engineering. Moreover, Liao *et al*. [116] conducted a survey of data mining techniques from 2000 to 2011 and categorised common data mining techniques as follows: neural networks, algorithm architecture, dynamic prediction-based, analysis of systems architecture, intelligence agent systems, modelling, knowledge-based systems, system optimisation and information systems.

As in the application of data mining in wind power, many literatures can be found in wind speed or power prediction, performance monitoring and analysis and fault diagnosis and prognosis. Accurate power prediction can be helpful in the operation management and on-site management of wind farms. Kusiak *et al.* [119] have adopted weather data to predict short-term and long-term power with a neural network, support vector machine regression, classification and regression tree and random forest algorithms. Feature selection is used to reduce the dimension of predictors and enhance model accuracy. The final model is able to predict a maximum short-term power of up to 12 h forecast and a 84 h forecast length for long-term predictions. Moreover, accurate monitoring of the turbine during operation is necessary for convenient maintenance scheduling and performance optimisation. Astolfi *et al.* [120] have proposed performance indicators based on data mining techniques by post-processing the SCADA data from both on-shore and off-shore wind farms sited in southern Italy.

Geographical, meteorological and system dynamic aspects are all considered in establishing the performance metric such as the multifunctioning index, the polar efficiency index, the stationarity index and the misalignment index. The former two indexes evaluate the system at a macro level, and the latter two indexes are dedicated to turbine assessment under wake effects. Finally, research in utilising data- based fault detection can be found in [121, 122]. Prediction models of the wind turbine are proposed, and various measures are used to distinguish normal and faulty operation conditions. In this study, a principal component analysis-based feature extraction technique is used as the basis for the proposed variable selection algorithms.

4.1.2 PCA-based feature extraction

Essentially, PCA is a variant of multivariate analysis that relies on the data-analytic technique and attempts to reveal the multivariate structure of the data. It transforms a set of data in a higher dimension into a set of orthogonal principal components (PCs) in the lower dimensional space. The PCs are found by maximising their variance and ranking them in terms of their magnitude. Because of the orthogonality property, these PCs are linearly uncorrelated [123, 124]. PCA was initially used as a dimension reduction technique in different fields [125]. Due to their unique properties, researchers have shown that by retaining the first few components, the dimension of the data can be reduced dramatically, while minimal information is sacrificed. Further application of PCA in the feature selection technique incorporated into a forecasting model, such as regression and an artificial neural network (ANN), can be found in [126, 127]. Moreover, PCA-based fault detection methods have been applied for wind turbines gear box faults and electrical machine faults [128, 129].

The orthogonal transformation of a given dataset $\mathbf{X}$ ($p \times n$ dimensions) with p variables and n samples is essentially the rotation of the coordinate axes such that the variance of the dataset is maximised. The resulting principal component scores, $\mathbf{Z}$, will have p number of PCs with n samples in the new coordinate. The transformation of the dataset is calculated by singular value decomposition (SVD) of the covariance matrix $\mathbf{S}$ of the dataset $\mathbf{X}$ by optimising the variance. This means that the first principal component to be found will have the highest

variance, and the following PCs have their variances in a descending order. Because the transformation is carried out orthogonally, each PC is linearly uncorrelated to another. Evidently, each of these PCs is considered as a specific feature of the original dataset **X**.

PCs are obtained by satisfying the relationship in (4-1) using SVD and the determinental equation (4-2) of the covariance matrix **S** of dataset **X**,

$$\mathbf{U}'\mathbf{S}\mathbf{U} = \mathbf{L} \tag{4-1}$$

$$|\mathbf{S} - l\mathbf{I}| = 0 \tag{4-2}$$

where **L** (l_1,l_2,...,l_p) is the diagonal matrix known as the characteristic roots (eigenvalue of **S**), and **U** is the corresponding characteristic vectors or eigenvectors (u_1, u_2,..., u_i,..., u_p, where u_i are the columns of **U**). **S** is the covariance matrix of the original dataset and it is found by:

$$s_{ij} = \frac{n\sum_{t=1}^{n} x_{i,t}x_{j,t} - \sum_{t=1}^{n} x_{i,t}\sum_{t=1}^{n} x_{j,t}}{n(n-1)} \tag{4-3}$$

The relationship between the PCs **Z** (z_1, z_2,...,z_p) and the original dataset **X** after normalisation is mathematically expressed below:

$$\begin{aligned} z_1 &= u_{11}x_1 + u_{12}x_2 + \cdots + u_{1p}x_p \\ z_2 &= u_{21}x_1 + u_{22}x_2 + \cdots + u_{2p}x_p \\ &\vdots \qquad\qquad\qquad\qquad \vdots \\ z_p &= u_{q1}x_1 + u_{q2}x_2 + \cdots + u_{qp}x_p \end{aligned} \tag{4-4}$$

A detailed description of finding the principal components through variance optimisation can be found in [9, 10].

4.1.3 Hilbert-Huang transform

Measurements from the sensor are always in the time domain. However, in certain applications, data in other domains such as frequency or time-frequency domains can reveal hidden features from the time domain. We use the fast Fourier transform (FFT) and the Hilbert-Huang transform (HHT) to transform the time series data into frequency and frequency-time domain data. The HHT is a combination of empirical mode decomposition (EMD) proposed by Huang and the Hilbert spectral analysis [130]. EMD can decompose the signal into a series of oscillatory signals known as intrinsic mode functions (IMF). The Hilbert

transform is then used to calculate the instantaneous frequency change with respect to time at different oscillatory frequencies. The HHT is considered to be adaptive and able to analyse data from nonlinear and nonstationary processes, which is able to extract features in time-frequency-energy space.

Zhang has applied the HHT in earthquake motion recordings [131], where it was proved that HHT outperforms the conventional methods such as FFT to analyse non-stationary dynamic earthquake motion recordings. Moreover, the signal reconstructed from certain levels of IMF can be useful for capturing important frequency-time features contained in the original signal. It has also been shown that under certain conditions, the HHT is superior to the short-time Fourier transform (STFT) and wavelet analysis to analyse vibration signals for machine health monitoring and to diagnose localised defects in roller bearings [132].

4.1.3.1 Empirical mode decomposition

EMD decomposes the original signals $x(t)$ into a set of IMFs, each of which represents the intrinsic oscillatory modes of the signal at different frequency bands. According to Huang, the IMF has to satisfy two properties: 1) the number of zero-crossings and extrema must be equal or at most differ by one; 2) the mean value of the envelope defined by local maxima and minima has to be zero at any point. Therefore, the IMF is found by first identifying the local extrema and then by fitting the cubic spline line through all the maxima and minima to obtain the upper envelope $x_{up}(t)$ and lower envelope $x_{low}(t)$. Their mean is defined as $m_{ik}(t)$, and the difference between the original signal and the envelope mean is $h_{ik}(t)$.

$$m_{ik}(t) = \left[x_{up}(t) + x_{low}(t)\right]/2 \tag{4-5}$$

$$h_{ik}(t) = h_{i(k-1)}(t) - m_{ik}(t) \tag{4-6}$$

The process is known as sifting and repeats k times until the $h_{ik}(t)$ satisfies the criteria defined for the IMF, where $h_{i(k-1)}$ is the original signal when k=1. Moreover, a stopping criterion is needed for the sifting process to terminate in order to obtain meaningful IMF. The longer sifting will reduce the amplitude modulation and results in a frequency modulated signal. The most commonly adopted criterion is based on the convergence test, such that the normalised

squared difference (SD) between h_{ik} and $h_{i(k-1)}$ is smaller than a predefined value and is found by:

$$SD_k = \frac{\sum_{t=0}^{n} \left| h_{i(k-1)}(t) - h_{ik}(t) \right|^2}{\sum_{t=0}^{n} h_{i(k-1)}^2} \tag{4-7}$$

Huang *et al.* [130] have shown that the SD threshold value of 0.3 is a sufficient and reasonable choice.

Once an IMF is found, it is then subtracted from the original signal, and a residual signal $r_i(t)$ is obtained. The process repeats i iterations until the final residual is a constant or a monotonic function.

$$c_i(t) = h_{ik}(t) \tag{4-8}$$

$$r_i(t) = x(t) - c_i(t) \tag{4-9}$$

The original signal can be reconstructed by summing all the IMFs and the residual using the formula below.

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{4-10}$$

4.1.3.2 Hilbert spectral analysis

The Hilbert transform calculates the instantaneous frequency of the IMFs obtained through EMD. The original signal can be expressed as the real part $\Re$ of the form $z(t) = x(t) + jy(t)$:

$$x(t) = \Re\left(\sum_{i=1}^{n} a_i(t) e^{j \int \omega_i(t) dt} \right) \tag{4-11}$$

where $y(t)$ is the complex conjugate of $x(t)$, n is the total number of IMFs, a_i is the amplitude of the signal of IMF at level i, $\omega_i(t)$ is the frequency of the signal at level i and j^2=-1.

At level i, the corresponding amplitude $a_i(t)$ and phase $\theta_i(t)$ can be found by:

$$a_i(t) = \sqrt{c_i(t)^2 + y_i(t)^2} \tag{4-12}$$

where $c_i(t)$ is the IMF at level i, and

$$\theta_i(t) = \tan^{-1}\left(\frac{y_i(t)}{c_i(t)} \right) \tag{4-13}$$

Finally, the instantaneous frequency $\omega_i(t)$ at level i can be found by:

$$\omega_i(t) = \frac{d\left(\theta_i(t)\right)}{dt} \tag{4-14}$$

Consequently, the HHT transforms the original time series signal *x(t)* into a new set of instantaneous frequency signals *f(t)*, i.e. the frequency changing with respect to time of *x(t)*. In our study, all IMFs, with the exception of the residual, are used to reconstruct the instantaneous frequency data for sensor selection in order to avoid feature losses due to data transformation.

4.2 Variable reduction algorithm using PCA

Currently, the information maximisation principle is the methodology adopted in wind turbine condition monitoring data acquisition systems [18], which means the number of sensors installed depends on the trade-off between obtaining as much data as possible and the cost. For the measurements collected from sensors installed on the wind turbine, relationships must exist among them, and this can induce redundancies within the data. Thus, an appropriate sensor selection technique is desirable in order to identify and remove these unnecessary redundancies since too many sensors are providing similar information. In the meantime, the method should be able to retain the provision of vital information, which is critical for fault diagnosis, prognosis and maintenance scheduling. Moreover, it is aimed for the actual number of physical sensors to be reduced through estimation of the least significant variables. For wind turbines, the method can be used to reduce the complexity and the cost associated with developing a condition monitoring system.

4.2.1 Variable reduction at system level

Due to the unique properties of variance maximisation and the orthogonality of transformed PCs following PCA, it has been used widely in dimension reduction and feature extraction applications. The variable selection problem can be considered as an extension of the dimension reduction problem. If the problem is simply to preserve most of the variability in the dataset, the solution can be relatively simple. However, for a dataset that contains a high degree of inter-correlation between variables, such as condition monitoring data collected from

wind turbines, there is repeated information amongst variables that needs to be identified and removed. This makes the PCA analysis a suitable candidate in developing the optimal variable selection algorithm for wind turbine condition monitoring systems. By transforming the multi-collinear data through PCA, each principal component corresponds to a particular feature of the data (including noises), and because these components are uncorrelated, there is no redundancy present. Then, PCs with low variance can be discarded to achieve dimension reduction with minimal information lost. Cost functions, which are related to the selected PCs and the original variables, are then used to obtain an optimal set of original variables. The three proposed selection algorithms for variable selection at the system level are described individually.

4.2.1.1 B2 method

For the transformed PCs $\mathbf{Z}$ ($p \times n$), the set of q PCs $\boldsymbol{\mathcal{R}}^{\boldsymbol{pc}}$ to be retained is found by satisfying equation :

$$\boldsymbol{\mathcal{R}}_j^{\boldsymbol{pc}} = \underset{i \in p}{arg\,min}\, l_i \,, j \in q \tag{4-15}$$

where l is the eigenvalue vector found from SVD, i is the index for PCs and j is the index for the q selected PCs. The eigenvectors in equation (4-5) is used to relate the original variable to the selected PC. By examining the eigenvectors, variables satisfying equation (4-16) are removed, and the rest of the variables are retained $\boldsymbol{\mathcal{R}}^{\mathbf{var}}$.

$$\boldsymbol{\mathcal{D}}_j^{\boldsymbol{var}} = \underset{k \in \mathcal{R}^{pc}}{arg\,max}\, u_k \,, j \in q \tag{4-16}$$

where u_k is the corresponding eigenvector for the selected PC. Once a variable is retained, the corresponding PC's eigenvalue and eigenvectors are removed from $\mathbf{L}$ and $\mathbf{U}$ for the next iteration. A stopping criterion is introduced and evaluated at every iteration, and it is used to terminate the selection algorithm until a predefined threshold value is reached to obtain the optimal number of the retained set $\boldsymbol{\mathcal{R}}^{\mathbf{var}}$. A detailed discussion on stopping criteria will be discussed in a later section. The predefined threshold value has a great influence on the number of variables being retained, which in turn indicates the amount of information retained.

4.2.1.2 B4 method

Similar to the B2 method, the selection process is carried out by satisfying equation (4-17) in order to identify the set of PCs to be retained. Then, the retained variable set $\boldsymbol{\mathcal{R}}^{\mathbf{var}}$ from X is obtained using equation (4-16). Again, the algorithms terminate if the predefined threshold value is satisfied.

$$\boldsymbol{\mathcal{R}}_j^{\boldsymbol{pc}} = \underset{i \in p}{arg\,max}\, l_i, j \in q \tag{4-17}$$

As can be seen, the B2 selection method is based on the concept of identifying and discarding variables that have the least variability from the dataset, and the remaining variables then contain the major portion of variance. The B4 method directly identifies variables which preserve a larger proportion of the variability of the dataset.

4.2.1.3 H method

The H method is derived from one of the criteria for the principal variables proposed in [136, 137], where the cost function relies on the optimisation of minimising the squared norm of the original variables. The H method examines the H values, h_1, h_2, ..., h_p, and selects a set of variables which is based on equation (4-18). H values are calculated and ranked in a decreasing order for all original variables. Variables that have the highest H value are retained first. Again, the process stops when the criteria exceed the predetermined threshold.

$$\boldsymbol{\mathcal{R}}_j^{\boldsymbol{var}} = \underset{j \in q}{arg\,max}\, h_j : h_j = \sum_{i=1}^{p} \left(l_i u_{ji}\right)^2, i \in p \tag{4-18}$$

where l and u are the eigenvalue and eigenvector, respectively, as described earlier. By utilising the squared norm, the selection exploits the multivariate structure of the original dataset, where it captures features that are not included in the retained variable set from each iteration. This may result in a better performance over the B2 and B4 methods.

4.2.2 Stopping criterion

Stopping criterion used in the selection algorithms, as described in the previous section, has a significant influence on the performance of the method. Different stopping criterion can be found in literature. Regardless of the type of criteria to

be used, varying the threshold value will always affect the number of retained PCs and variables, hence the amount of information preserved. Therefore, the problem of identifying the right stopping criteria and the optimal threshold value need to be addressed first.

The question of how many principal components should be retained is application dependent [124]. Different rules have been proposed by researchers; however, the choice of a stopping threshold value differs from datasets. The commonly used methods include analysis of the individual eigenvalue, cumulative percentage of total variation, cross-validation and screen plot.

One of the pre-requirements for PCA is that the dataset used has to be standardised, which means a zero mean and unit variance. Assume that the original datasets **X** are independent between variables, and then their corresponding PCs should also have unit variance. Based on this assumption, by analysing the variance of the PCs (eigenvalues, **L**), any PCs having variance greater than 1 should be kept [123, 124]. However, it has been argued that the threshold for this selection criterion is dataset dependent. Jolliffe *et al.* [133, 134] have shown that the threshold value of 0.7 is reasonable. An alternative criterion is to calculate the total cumulative percentage variance of the number of selected PCs with equation:

$$t_m = 100 \times \sum_{k=1}^{q} l_k \Big/ \sum_{k=1}^{p} l_k \tag{4-19}$$

where l_k is the eigenvalues obtained from PCA, q is the number of PCs retained and p is the total number of variables in the original dataset. However, this method requires an additional threshold value in determining the optimal number of PCs to be retained. A sensible threshold value between 70% and 90% is suggested in [124], depending on the number of samples and variables from the dataset.

A more commonly used graphical method known as the screen plot analysis is the plot of eigenvalue l_k against k, where the judgement of the number of PCs retained depends on the steepness of the plot (l_{k-1} to l_k and l_k to l_{k+1}), as shown in Figure 4.1. The method selects q number of PCs such that the change of steepness

of the plot is minimal (red dotted line in Fig 4.1). Nevertheless, this method requires human supervision as it is difficult to solve the problem analytically. The two methods described involve a degree of subjectivity; therefore, a more analytical solution can be found with a cross validation method.

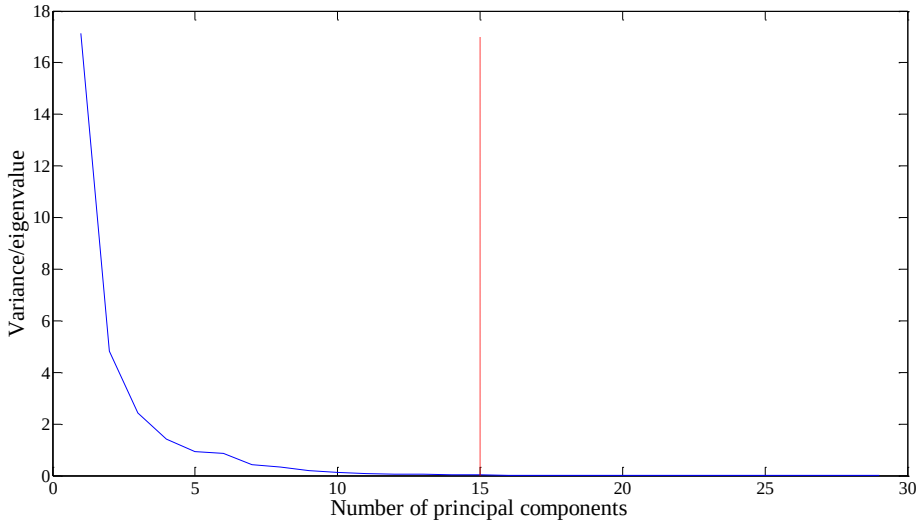


Figure 4.1 - Example of screen plot in determining number of PCs retained

The main purpose of cross validation is to estimate the original dataset $\mathbf{X}$ from q number of PC scores equation (4-20) and determine an optimum value for q based on the predicted residual sum of squares (PRESS) eq. 4-21.

$$\hat{\mathbf{X}}_{\boldsymbol{k}} = \hat{\mathbf{Z}}_{\boldsymbol{k}}\hat{\mathbf{U}}_{\boldsymbol{k}}^{\mathbf{T}} + \boldsymbol{\varepsilon} \tag{4-20}$$

$$\text{PRESS}(k) = \sum_{i=1}^{n}\sum_{j=1}^{p} \hat{x}_{i,j}^{k} - x_{i,j} \tag{4-21}$$

where $\hat{\mathbf{Z}}_{\boldsymbol{k}}$ and $\hat{\mathbf{U}}_{\boldsymbol{k}}$ are the estimated PC scores and eigenvectors and $\boldsymbol{\varepsilon}$ is the error between the estimated and actual data using k PCs. A relationship between the PRESS measure and varying number of PCs retained is used to determine the optimum solution. A main point during cross validation is that when estimating $\hat{x}_{i,j}$, the original data $x_{i,j}$ (the i^{th} variable and j^{th} sample of $\mathbf{X}$) should not be included when establishing the model. A drawback of the method has been pointed out in [138] with the data use in this study, PCs with important information may occasionally be discarded, but it can be used as a reference for the minimal number of PCs desired. In our study, the screen plot and cross

validation are used in combination to determine the number of PCs for the proposed selection algorithms, as described in the previous section.

4.2.3 Performance measure

Once a set of variables are retained, three performance measures are proposed in order to validate the result from the selection algorithms described. Each of these measures analyses a different aspect of the retained dataset. An evaluation is then made using these three measures in combination.

4.2.3.1 Cumulative variance

Variance is usually used to measure the dataset variability. However, because normalisation is a requirement for PCA, such that each variable in the dataset has unit variance, the calculation of cumulative variance from the normalised dataset will not help. Alternatively, cumulative percentage variance based on partial variance is used. It measures the percentage variability of the retained variables with regards to the whole dataset, where the multivariate structure of the dataset is considered. For a dataset $\mathbf{X}$ ($p \times n$) with q variables ($q<p$) being retained and m variables ($m=p-q$), the covariance matrix of the dataset $\mathbf{X}$ can be divided into $\boldsymbol{S}_{rr}$ ($q \times q$), $\boldsymbol{S}_{rd}$ ($q \times m$), $\boldsymbol{S}_{dr}$ ($m \times q$), $\boldsymbol{S}_{dd}$ ($m \times m$), as shown in (22). The subscripts r and d represent the retained set with q number of variables and the discarded set with m number of variables, respectively.

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{S}_{rr} & \boldsymbol{S}_{rd} \\ \boldsymbol{S}_{dr} & \boldsymbol{S}_{dd} \end{bmatrix} \tag{4-22}$$

The partial covariance matrix $\boldsymbol{S}_{rr.d}$ for retained variables is:

$$\boldsymbol{S}_{rr.d} = \boldsymbol{S}_{rr} - \boldsymbol{S}_{rd}\boldsymbol{S}_{dd}^{-1}\boldsymbol{S}_{dr} \tag{4-23}$$

The cumulated percentage partial variance (*cppv*) can then be obtained by the equation below, where *tr* is the trace of the partial covariance matrix, i.e. sum of the elements on the main diagonal.

$$cppv = tr(\boldsymbol{S}_{rr.d})/tr(\boldsymbol{S}) \tag{4-24}$$

4.2.3.2 Average correlation coefficient

For a retained set of variables, *cppv* measures the variability, but it does not explain any repetition/redundancy among retained variables, such as that among

power, current and voltage. This measure only calculates fluctuation of the magnitude of the signal. Thus, an average correlation coefficient r_a based on Pearson's correlation coefficient is introduced to measure the degree of associations between variables in the dataset. Due to the fact that Pearson's correlation coefficients are not additive, the average correlation coefficient cannot be calculated using a simple arithmetic mean. To be able to calculate the average correlation coefficient, Pearson's correlation coefficient is first transformed using Fisher's transformation, and then the arithmetic average of the transformed value is converted back with Fisher's inverse transform. The Fisher's transform and its corresponding inverse transform are given below:

$$r_z = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right) \tag{4-25}$$

$$r = \frac{e^{2r_z}-1}{e^{2r_z}+1} \tag{4-26}$$

where r_z is the transformed correlation coefficient, and r is the Pearson correlation coefficient. This measure considers the multi-collinearity behaviour of the dataset. A higher value indicates high degrees of correlation among the dataset, and a low value indicates less dependency between variables. Moreover, the method not only considers the magnitude of dependency but also includes the direction of the associated variables.

4.2.3.3 Information entropy

The *cppv* and average correlation coefficient are dependent on the sample size of the dataset, and this may lead to a biased result. Therefore, information entropy is used to further validate the results. As a measure of information discrepancy, entropy has been used extensively in communication, data compression and data encoding [139], as well as in feature selection and classification for ANN and fault detection [140]. After the application of PCA, the information entropy of the original dataset and those retained variables from each selection method are calculated individually.

With a given variable **X** and the probability mass function of the variable $p(x)=Pr\{X=x\}$, $x \in \Re$, the information contained or the uncertainty in the variable **X** can be quantified by the information entropy E(**X**):

$$E(X) = -\sum_{i=1}^{n} P(x_i) \log_b P(x_i) \tag{4-27}$$

where $P(x_i)$ is the probability $p(X=x_i)$, and b is the base for each different entropy unit. In this dissertation, the Shannon's entropy is used, where b=2; hence, the unit of the entropy is the bit. Moreover, normalised entropy is also introduced in order to compare different variables, as the normalised entropy is bounded between 0 and 1, which is obtained by:

$$H(X) = -\sum_{i=1}^{n} \frac{p(x_i) \log_b \big(p(x_i)\big)}{\log_b(n)} \tag{4-28}$$

where n is the length of the signal, and $log_b(n)$ is the maximum entropy of the signal. Based on the normalised information entropy, the percentage information entropy η_e of any retained set of variables can be found, where it is the ratio of the normalised entropy of retained variables E_r and the original datasets E_t, where E_r and E_t are the sum of entropy of the retained variables and original data, respectively.

4.2.3.4 Combination performance measure

Since all three measures are bounded within the same range, and each measure focuses on a specific nature of the dataset, a more generalised result can be obtained using a combination performance measure:

$$cpm = (1 - \text{cppv}) + (1 - \eta_e) + |r_a| \tag{4-29}$$

where 1-*cppv* and 1-η_e measure the variance and information entropy losses due to discarded variables, and $|r_a|$ is the magnitude of the overall correlation coefficient of the dataset. As each of these three terms is bounded between 0 and 1, and this result in the *cpm* being bounded between 0 and 3. The smaller the *cpm* the higher the performance, such that the retained variables have minimal information losses, multi-collinearity and redundancies.

4.2.4 Variable reduction based on particular targets

The three selection algorithms described in Section 4.2.1 mainly investigate the variability of the data, and this can be useful and accurate during fault-free conditions. When faulty data is used, the method may result in false outcomes, as

in most cases the variability of transient change due to a fault is insignificant compared to the overall trend of the dataset. Therefore, a targeted selection method is proposed, named the T method. The algorithm maximises the variance and minimises the correlation of the selected variables. It also seeks the underlying features of the predefined fault signal/variable within the dataset and retains a set of variables that most effectively explains these features. The selection of PCs is based on equation (4-30):

$$\boldsymbol{\mathcal{R}_j^{pc}} = \underset{i \in p}{arg\ min}\left(r_{i,j}^2 - r_{tar,j}^2\right), j \in q \tag{4-30}$$

where it selects a set of q PCs $\boldsymbol{\mathcal{R}^{pc}}$ by minimising the squared correlation difference between original variables $r_{i,k}^2$ and PCs $r_{tar,k}^2$. The subscript i, k and tar are the i^{th} variable from the original dataset, the k^{th} principal component and the targeting variable, respectively. The square correlation for each term is found by:

$$r_{i,j}^2 = \left\{ \frac{\sum_{i=1}^{p}(x_i - \bar{x})\left(z_j - \bar{z}\right)}{\sqrt{\sum_{i=1}^{p}(x_i - \bar{x})^2 \sum_{j=1}^{p}\left(z_j - \bar{z}\right)^2}} \right\}^2 \tag{4-31}$$

which is also equivalent to:

$$r_{i,j}^2 = l_i u_{i,j}^2 \tag{4-32}$$

where l_i and $u_{i,j}$ are the eigenvalue and eigenvector obtained from SVD. The selection of original variables based on $\boldsymbol{\mathcal{R}^{pc}}$ is then obtained:

$$\boldsymbol{\mathcal{R}_j^{var}} = arg\ max\ u_k, j \in q, k \in \mathcal{R}^{pc} \tag{33}$$

where j and k are the index for the selected original variables and retained principal components, respectively. The retained variable set $\boldsymbol{\mathcal{R}_j^{var}}$ is updated each iteration, and the same stopping criterion as discussed in Section 4.2.2 is used to terminate the process until the threshold is reached.

4.3 Data pre-processing for validation (simulation and SCADA data)

To validate the proposed algorithms, simulation data obtained from Chapter 3 and SCADA data from an operational wind farm are used. Data pre-processing is necessary for the algorithms to perform accurately and reliably. Moreover, the

original time domain dataset is transformed into frequency and frequency-time domain prior to the selection algorithm.

4.3.1 Data pre-processing

As PCA is a statistical tool, in order to perform the analysis, certain pre-requirements of the datasets are needed. As described in Chapter 2, SCADA contains various measurements from a wind turbine including constant variables, discrete signals (such as control set points) and missing values which may due to sensor malfunction or system downtime. These variables can be ineffective to PCA, hence the selection algorithm. Therefore, in our study, constant variables and discrete signals are excluded prior to the analysis. Moreover, for any missing values at the j^{th} sample from a variable, the corresponding values for the entire row at sample j are removed. Finally, normalisation of the dataset is carried out to allow for comparison between variables or datasets, as the normalised datasets $\overline{\mathbf{X}}$ have zero mean and unit variance:

$$\overline{\mathbf{X}} = \frac{\mathbf{X} - \boldsymbol{\mu}}{\boldsymbol{\sigma}} \tag{4-34}$$

where $\mathbf{X}$ is the original dataset, $\boldsymbol{\mu}$ is a vector of the arithmetic mean and $\boldsymbol{\sigma}$ is the vector of the standard deviation of the dataset.

The constant variables and missing values are not present in the simulation data simply because only useful data is collected during the simulation process, and the only discrete signals are the switching pulses for the inverters. These variables are also excluded before selection algorithms are applied. Therefore, the only pre-processing step needed for the simulation data is the normalisation using equation (4-34). On the contrary, the SCADA data is obtained from an operational wind farm, and the pre-processing processes described earlier are necessary in order to identify and remove data which may lead to biased results such as constant variables and digital signals. The SCADA data consist of sensor measurements of 26 wind turbines across a time span of 16 months. The data is acquired at a sampling rate of two seconds and are processed and stored at 10-minute intervals in order to significantly reduce the amount of data that needs to be processed, while still reflecting the normal and faulty status of wind turbine

operations. Each turbine consist of 128 variables, including various temperatures, pressures, vibrations, power outputs, wind speed and digital control signals. In this study, the data used is the 10-minute averages for each variable.

To understand the SCADA data further, these variables are categorised based on their type of measurement, including different angles, electrical parameters, set points, pressures, vibrations (oscillations), speeds, wind related, temperatures and miscellaneous. Variables that fall under the category of set points are considered as discrete control signals. One example is Pit_PID1_P, which is the proportional parameter for the PID controller. Figure 4.2 from top to bottom shows the example of a discrete signal, the pitch angle set point and the active power before and after the pre-process of missing values from the SCADA data. It can be seen that during the period where no active power is produced (red dots in the middle plots), the pitch angle set points are abnormally high, reaching approximately 90°. This may be because an alarm was raised and the operator intentionally shut down the turbine, thus causing false readings such as pitch angle set point.

A total of 77 variables are obtained after the pre-processing, and an example of monthly normalised active power, generator speed and main line-to-line voltage (L1-L2) is shown in Figure 4.3. These datasets are then used in the later validation processes.

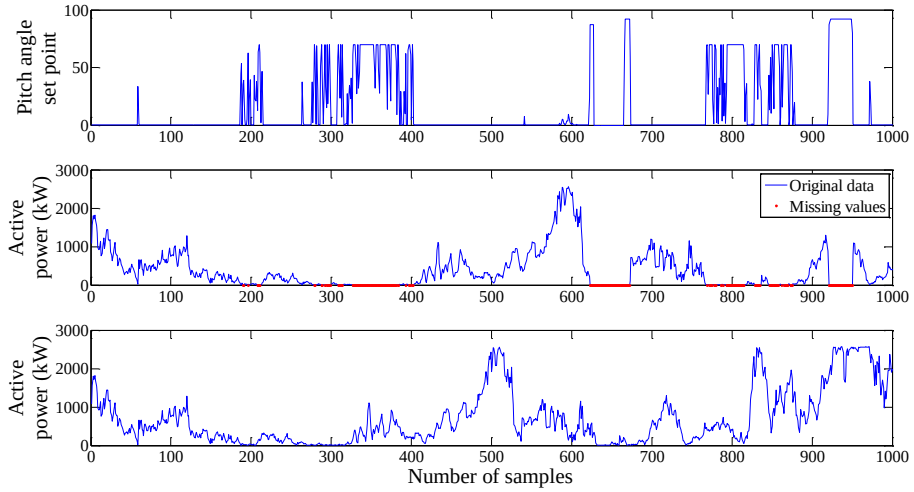


Figure 4.2 - Example of SCADA data. Top: discrete control signal to be removed, pitch angle set point; middle: original active power before pre-processing; bottom: active power after pre-processing

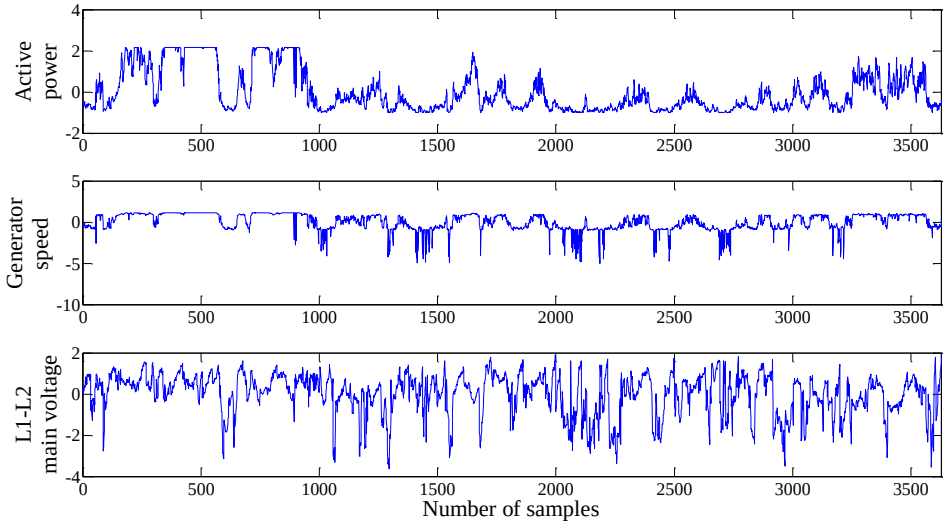


Figure 4.3 - Example of monthly active power (top plot), generator speed (middle plot) and line-to-line main voltage (bottom plot) after pre-processing

4.3.2 Data transformation into frequency and time-frequency domains

The analysis of the measurement data focuses on the transient characteristics not only in the time domain but also in terms of the frequency and instantaneous frequency time domains. The traditional fast Fourier transform is used to obtain data in the frequency domain:

$$F(t) = \frac{1}{n}\sum_{i=0}^{n-1} x_i e^{-j2\pi k\frac{i}{n}}, 0 < k < n-1 \tag{4-35}$$

where n is the total number of samples in the original dataset $\mathbf{X}$, and j is the complex notation.

Figure 4.4 shows an example of active power in the time and frequency domain from the PMSG simulation data using practical wind speed as the input wind profile. As shown in the bottom plot of Figure 4.4, the low frequencies are due to wind speed variation. Moreover, there are several spikes at the frequency of 50, 150 and 300 Hz, respectively, which may represent the fundamental frequency and corresponding harmonics of the electrical power. Finally, noise is also a major contributor in the FFT plot.

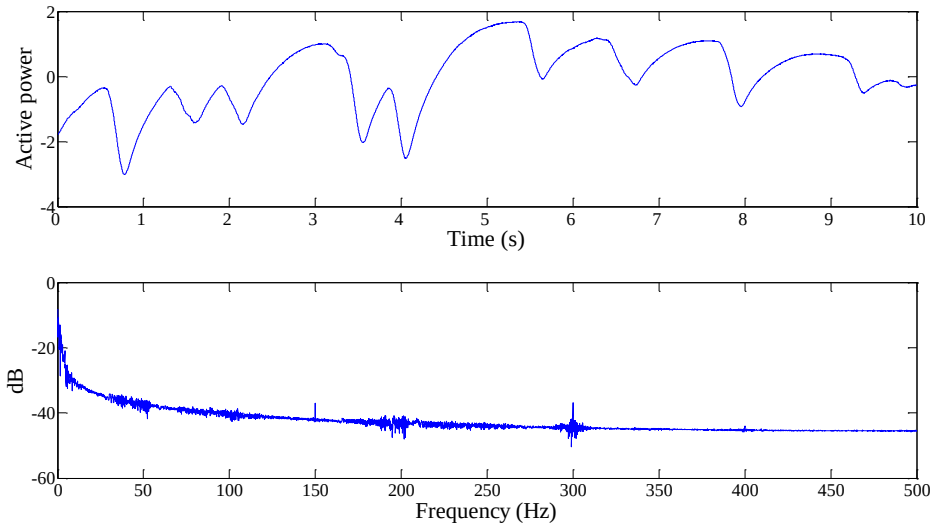


Figure 4.4 - Active power of PMSG based wind turbine simulation data using practical wind speed in time and frequency domain

The frequency time domain data is obtained using HHT, as described in Section 4.1.3. Intrinsic mode functions are first calculated for each variable through empirical mode decomposition. The instantaneous frequency for each IMF is then found using the Hilbert transform except for the last IMF, where it is monotonic. Then, by summing all the instantaneous frequencies, the reconstructed frequency time domain data is obtained. As an example, Figure 4.5 shows the active power from PMSG simulation in the time domain, the extracted IMFs (1-8) and the reconstructed instantaneous frequency time domain data. It can be seen that the frequency of the IMFs decreases until it becomes monotonic. Moreover, it can be seen that there is a sudden change of frequency between IMF3 and IMF4. The high frequency IMFs may relate to the electrical power, while the low frequency IMFs may represent variations of wind speeds. The residual usually represents the trend of the original data as concluded in [130]. The last plot shows the instantaneous frequency variation with respect to time.

The same procedures are applied to all simulation and SCADA datasets to obtain the pre-processed data in frequency and frequency time representations. These newly obtained datasets are then used as inputs to perform the proposed variable selection algorithms. Results are then compared and discussed in later sections.

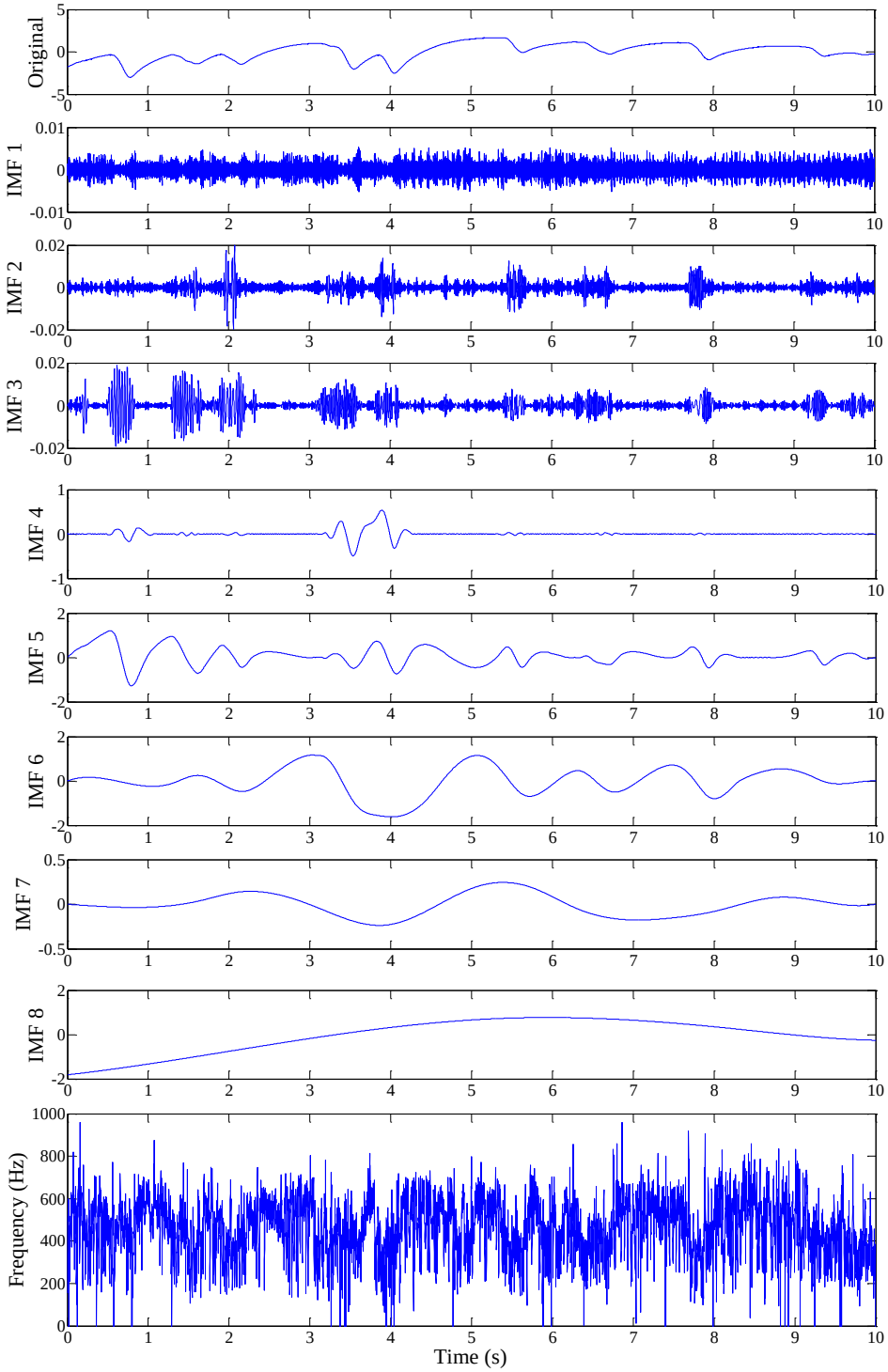


Figure 4.5 - Decomposed intrinsic mode functions of active power of PMSG based on wind turbine simulation data and the reconstructed instantaneous frequency data using HHT

4.4 Verification of the methods

To verify the proposed selection algorithms and their robustness, different types of datasets are used. A summary block diagram of the validation processes at the system level and the targeted method are shown in Figures 4.6 and 4.7. It includes different operational conditions, types of systems and data domain representations. The proposed algorithms are applied to each dataset, and the performance is evaluated with measures described earlier. Based on this process, results are given in later sections.

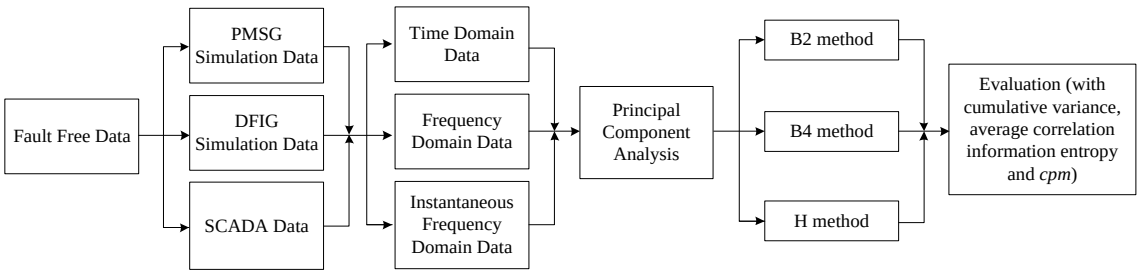


Figure 4.6 - Block diagram of variable selection validation process using different types of datasets at the system level

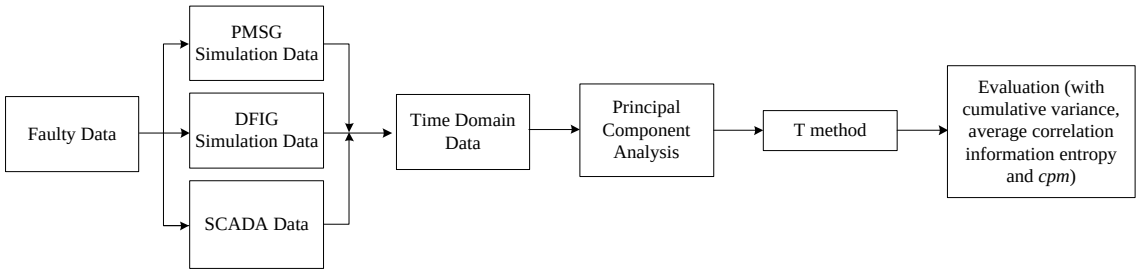


Figure 4.7 - Block diagram of targeted variable selection validation process using different types of datasets

4.4.1 Validation in normal operational condition

During a normal condition, the selection is carried out at the system level, where it focuses on maximising variability and minimising redundancies of the measurements. Thus, a set of variables is selected that represents the general trend of the dataset.

4.4.1.1 Simulation and SCADA data

Both simulation and SCADA datasets are used, where the SCADA data has a total of 77 variables. In the case of simulation data, both PMSG and DFIG-based wind turbines, as described in Chapter 3, are used. A total number of 46 and 36 variables for each system and the variable names and description of the variables are provided in Tables 4.1 and 4.2, respectively. The dataset mainly consists of mechanical power, torque and speed, and electrical power, torque, voltages and currents at different locations within the system as described in Chapter 3.1.3.

No.	Variable Name	Description	No.	Variable Name	Description
1	I_{dc1}	Dc-link current before capacitor	**24**	Q_3	Grid reactive power at node 2
2	V_{dc}	DC-link voltage	**25**	Q_g	Reactive power at PCC (after transformer)
3	I_{dc2}	DC-link current after capacitor	**26**	V_{a1}	Grid phase A voltage at node 1
4	Q_{grid}	Reactive power at PCC (before transformer)	**27**	V_{b1}	Grid phase B voltage at node 1
5	P_{grid}	Active power at PCC (before transformer)	**28**	V_{c1}	Grid phase C voltage at node 1
6	I_{dc_ref}	Reference DC-link current after capacitor	**29**	V_{a2}	Grid phase A voltage at node 2
7	α	Firing angle	**30**	V_{b2}	Grid phase B voltage at node 2
8	β	Pitch angle	**31**	V_{c2}	Grid phase C voltage at node 2
9	Q_{gen}	Generator reactive power	**32**	V_{a3}	Grid phase A voltage at node 3
10	T_t	Turbine torque	**33**	V_{b3}	Grid phase B voltage at node 3
11	P_{gen}	Generator active power	**34**	V_{c3}	Grid phase C voltage at node 3
12	T_{gen}	Generator torque	**35**	I_{a1}	Grid phase A current at node 1
13	P_t	Turbine power	**36**	I_{b1}	Grid phase B current at node 1
14	ω	Turbine mechanical speed	**37**	I_{c1}	Grid phase C current at node 1
15	v_w	Wind speed	**38**	I_{a2}	Grid phase A current at node 2
16	P_{source}	Grid source active power	**39**	I_{b2}	Grid phase B current at node 2
17	P_1	Grid active power at node 1	**40**	I_{c2}	Grid phase C current at node 2
18	P_2	Grid active power at node 2	**41**	I_{a3}	Grid phase A current at node 3
19	P_3	Grid active power at node 3	**42**	I_{b3}	Grid phase B current at node 3
20	P_g	Active power at PCC (after transformer)	**43**	I_{c3}	Grid phase C current at node 3
21	Q_{source}	Grid source reactive power	**44**	I_{ga}	Phase A current at PCC
22	Q_1	Grid reactive power at node 1	**45**	I_{gb}	Phase B current at PCC
23	Q_2	Grid reactive power at node 2	**46**	I_{gc}	Phase C current at PCC

Table 4.1 - Variable names of the PMSG-based wind turbine simulation data

No.	Variable Name	Description	No.	Variable Name	Description
1	V_w	Wind speed	**19**	V_{sa_rms}	Phase A stator rms voltage
2	ω_{ref}	Reference turbine speed	**20**	V_{sb_rms}	Phase B stator rms voltage
3	ω_{mech}	Turbine mechanical speed	**21**	V_{sc_rms}	Phase C stator rms voltage
4	T_{mech}	Turbine torque	**22**	V_{ra_rms}	Phase A rotor rms voltage
5	T_{elec}	Generator torque	**23**	V_{rb_rms}	Phase B rotor rms voltage
6	s	Generator slip	**24**	V_{rc_rms}	Phase C rotor rms voltage
7	I_{dc1}	DC-link current 1	**25**	V_{ga_rms}	Phase A rms voltage at PCC
8	I_{dc2}	DC-link current 2	**26**	V_{gb_rms}	Phase B rms voltage at PCC
9	V_{dc}	DC-link voltage	**27**	V_{gc_rms}	Phase C rms voltage at PCC
10	P_{dc1}	DC-link active power 1	**28**	I_{sa_rms}	Phase A stator rms current
11	P_{dc2}	DC-link active power 2	**29**	I_{sb_rms}	Phase B stator rms current
12	P_t	Turbine power	**30**	I_{sc_rms}	Phase C stator rms current
13	P_g	Active power at PCC	**31**	I_{ra_rms}	Phase A rotor rms current
14	P_s	Generator stator active power	**32**	I_{rb_rms}	Phase B rotor rms current
15	P_r	Generator rotor active power	**33**	I_{rc_rms}	Phase C rotor rms current
16	Q_g	Reactive power at PCC	**34**	I_{ga_rms}	Phase A rms current at PCC
17	Q_s	Generator stator reactive power	**35**	I_{gb_rms}	Phase B rms current at PCC
18	Q_r	Generator rotor reactive power	**36**	I_{gc_rms}	Phase C rms current at PCC

Table 4.2 - Variable names of the DFIG-based wind turbine simulation data

For both PMSG and DFIG simulation data, the fault-free data with actual wind speed is obtained at a sampling rate of 1 ms over a period of 10 s and yield a total of 10000 numbers of samples for each variable. In the case of SCADA data, a month's worth of data is used. In order to obtain healthy turbine data, typical wind speed vs. active power curve (S-curve) is adopted to distinguish the overall health condition of the turbine. S. Gill *et al.* [22] have shown that the S-curve can be used as a metric for determining a turbine's operation conditions. Based on the S-curve metric, healthy turbine data is selected from the 26 wind turbines. Figure 4.8 shows S-curves of different wind turbines from the wind farm. The top plot shows the power curve of a healthy turbine where all alarms registered in the alarm log from the SCADA database are removed. The middle plot gives a reasonable power curve, which comes from a healthy turbine. However, there are still outliners, indicating the existence of anomalies. The timestamp of these outliners coincides with the information recorded in the alarm log in the SCADA

data as shown as red dots. The bottom plot shows the S-curve of a turbine with a known fault, where it can be seen that there is a reduced power output irrespective of the wind speed. The turbine that produces the best S-curve (without alarm) is used.

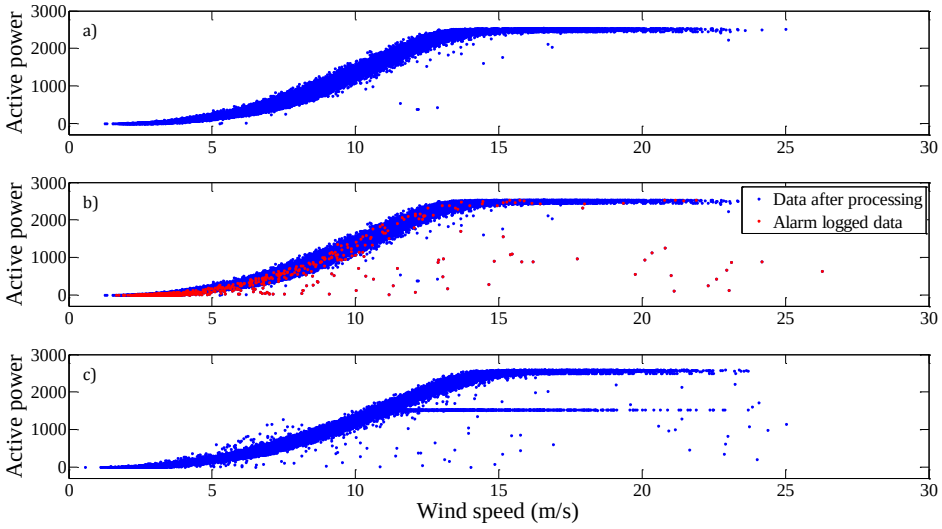


Figure 4.8 - S-curve of SCADA data with different processing steps: a) fault free turbine without alarm, b) fault free turbine with alarm, c) faulty turbine

Before carrying out the selection algorithm, threshold values to terminate the selection process need to be first determined. The screen plot and cross-validation techniques as described in Section 4.2.2 are applied to the three datasets. Figures 4.9 to 4.11 show the results from the three datasets (PMSG, DFIG and SCADA), respectively. The top plot shows the cumulative variance-based SCREEN plot, and the bottom plot shows the squared residual from cross-validation.

For the PMSG data, it can be seen that the optimal number of PCs from the two techniques are 25 and 8, respectively; the cumulative variance and PRESS are 99.92% and 3.43×10^{4}, respectively, as shown in Figure 4.9 (PRESS is a measure of the total absolute residual, thus a smaller PRESS indicates a higher model fit). As already discussed in Section 4.2.2, the cross-validation technique may occasionally eliminate important information. However, it is useful to provide a benchmark for the minimal limit. As for the SCREEN plot, the curve is practically flat with 25 numbers of PCs, and a very high cumulative variance is obtained. Any further increase in the number of PCs will have an insignificant increase in the

cumulative variance. Therefore, the threshold value of 25 PCs is used for the selection algorithm using PMSG data. Similarly for the data obtained from DFIG based simulation and SCADA data, the cumulative variance and PRESS are (99.93%, 5.96×10^4) and (99.7%, 3.59×10^4) respectively; the optimal threshold values are 20 and 35 for each dataset, as shown in Figures 4.10 and 4.11.

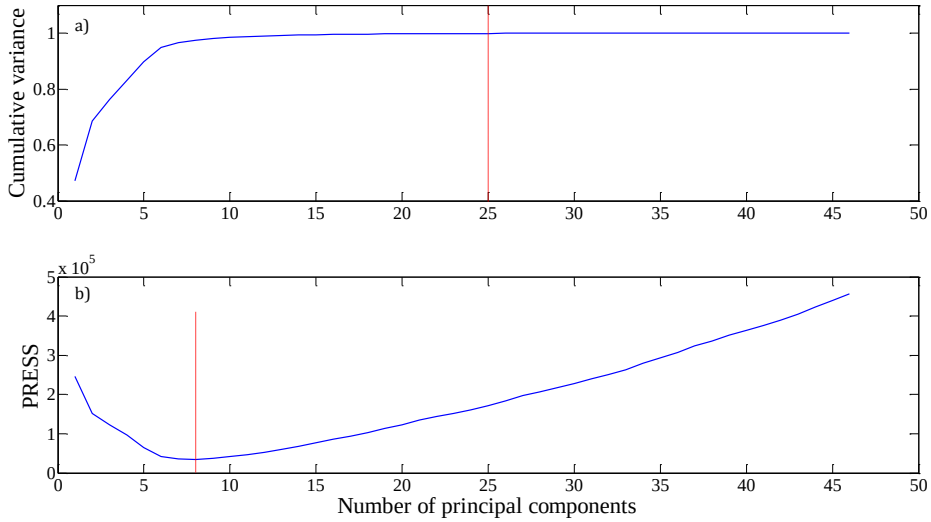


Figure 4.9 - Optimal threshold value for selection algorithm using PMSG data: a) SCREEN plot of cumulative variance; b) PRESS from cross-validation

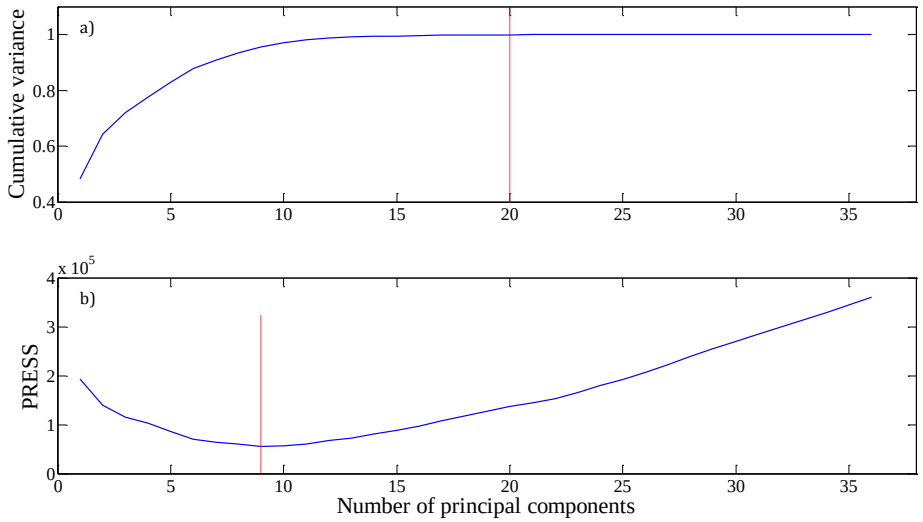


Figure 4.10 - Optimal threshold value for selection algorithm using DFIG data: a) SCREEN plot of cumulative variance; b) PRESS from cross-validation

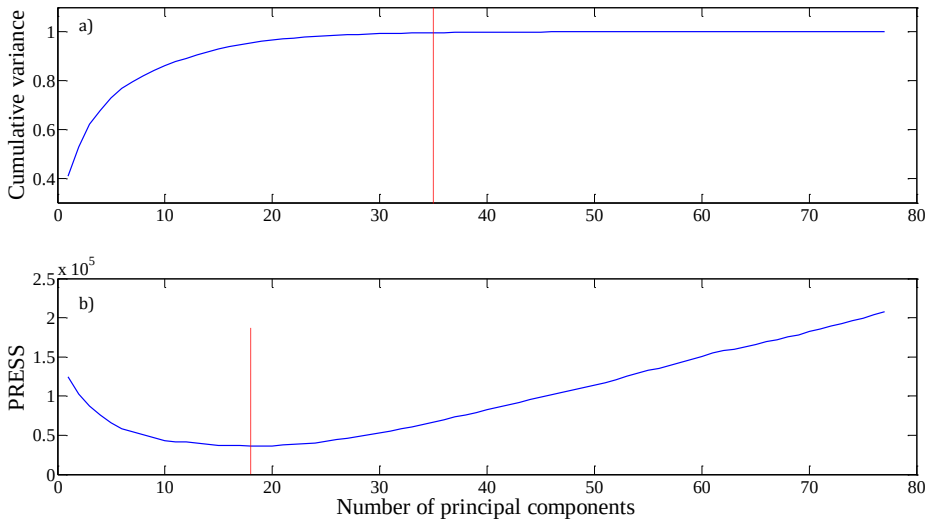


Figure 4.11 - Optimal threshold value for selection algorithm using SCADA data: a) SCREEN plot of cumulative variance; b) PRESS from cross-validation

With the pre-processed data and optimal threshold value, the proposed selection algorithms are carried out for each dataset. The performance measures for each input dataset using different selection algorithms are calculated and shown in Table 4.3. By looking at each measure individually, the cumulative variances from the time and frequency domains have similar values and are considerably higher than those from instantaneous frequency data for all three types of datasets (PMSG, DFIG and SCADA). This might be because during the reconstruction of the instantaneous frequency data from IMFs in HHT, certain features of the dataset are lost. Moreover, the instantaneous frequency data represents the localised oscillatory phase change with respect to time; thus, it only contains a portion of the total features (only oscillatory phase change) within the dataset. Hence, during variable selection, the algorithms focusing on the instantaneous frequency might ignore certain information from the overall dataset. Furthermore, results show that the H method has the lowest performance among the three selection methods, and the B2 method has the highest cumulative variance for all three domains.

As for the average correlation coefficients, values of 0.1484, 0.3394 and 0.3588 are found for the three original datasets. These values serve as the basis when examining the performance of the selection algorithms. When comparing to these benchmarks, the results show that the B2 and B4 methods produces a lower average correlation coefficient in all three domains. This indicates the selected

variables have low interdependencies. In addition, it should be noted that a higher average correlation coefficient than the original dataset can be obtained, implying the presence of a higher degree of redundancies among the retained variables, such as results using the H method. In addition, since the *cppv* and average correlation coefficient may lead to biased results, the percentage entropy is shown in Table 4.3. A similar trend can be observed, such that the B2 and B4 have the highest percentage entropy, and the H method has the lowest. Moreover, the percentage entropies obtained in time, frequency and instantaneous frequency domain data have a descending order. These findings agree with the results from the cumulative percentage variance and average correlation measures.

To visualise the results from each selection algorithm, scatter plots of the cumulative percentage variance and average correlation coefficients of each variables are plotted. Figures 4.12 to 4.14 shows all the results using three different datasets from three selection algorithms in different domains, where the blue crosses are all the variables from the original datasets, and the red circles represent the retained variables from the corresponding selection algorithm and domain. Overall, the scatter plots show positive relationships between the *cppv* and the average correlation coefficients, where variables having high *cppv* tend to also have a higher absolute correlation coefficient. This behaviour is the most obvious for the SCADA dataset (Fig 4.13), as it contains actual measurements from a physical system rather than simulations. Moreover, it can be seen from the figures that the H method tends to retain variables which have high cumulative variance, and this is expected as it is depicted by its cost function. However, some of these high variance variables also have high correlations among each other, which will increase information redundancy considerably in the retained variables. This behaviour also explains why the H method has poor results in Table 4.3, as high redundancies are involved in the retained set of variables. On the contrary, the retained variables from the selection methods which produce high performance, such as B2 and B4, are spread across the entire area (both low and high cumulative variances). The majority of the retained variables have low average correlation coefficients and selectively retain variables with high correlation and variance.

		Time			Frequency			Instantaneous frequency		
	Original dataset	B2	B4	H	B2	B4	H	B2	B4	H
PMSG data										
Cumulative variance	100%	99.19%	99.60%	90.21%	98.34%	97.33%	92.35%	69.35%	67.75%	57.71%
Average correlation	0.1484	0.0518	0.0565	0.1164	0.0589	0.1062	0.2824	0.0242	0.0408	0.0781
Percentage entropy	100%	86.86%	89.50%	48.99%	82.29%	70.35%	68.26%	84.62%	81.20%	67.27%
DFIG data										
Cumulative variance	100.00%	99.67%	97.67%	83.70%	98.31%	97.61%	87.81%	78.28%	76.84%	58.23%
Average correlation	0.3676	0.0317	0.0801	0.4139	0.1181	0.1290	0.4444	0.1417	0.1888	0.2102
Percentage entropy	100.00%	84.82%	79.97%	45.53%	75.60%	78.95%	49.58%	67.20%	71.21%	76.70%
SCADA data										
Cumulative variance	100%	97.17%	98.42%	83.94%	95.94%	94.78%	68.77%	66.60%	64.08%	49.77%
Average correlation	0.3588	0.0777	0.0650	0.4983	0.0821	0.1063	0.3098	0.1213	0.1245	0.4124
Percentage entropy	100%	80.37%	81.71%	37.95%	78.36%	76.41%	48.33%	71.34%	68.40%	25.72%

Table 4.3 - Results from selection methods B2, B4 and H in the time, frequency and instantaneous frequency domains using simulation and SCADA data

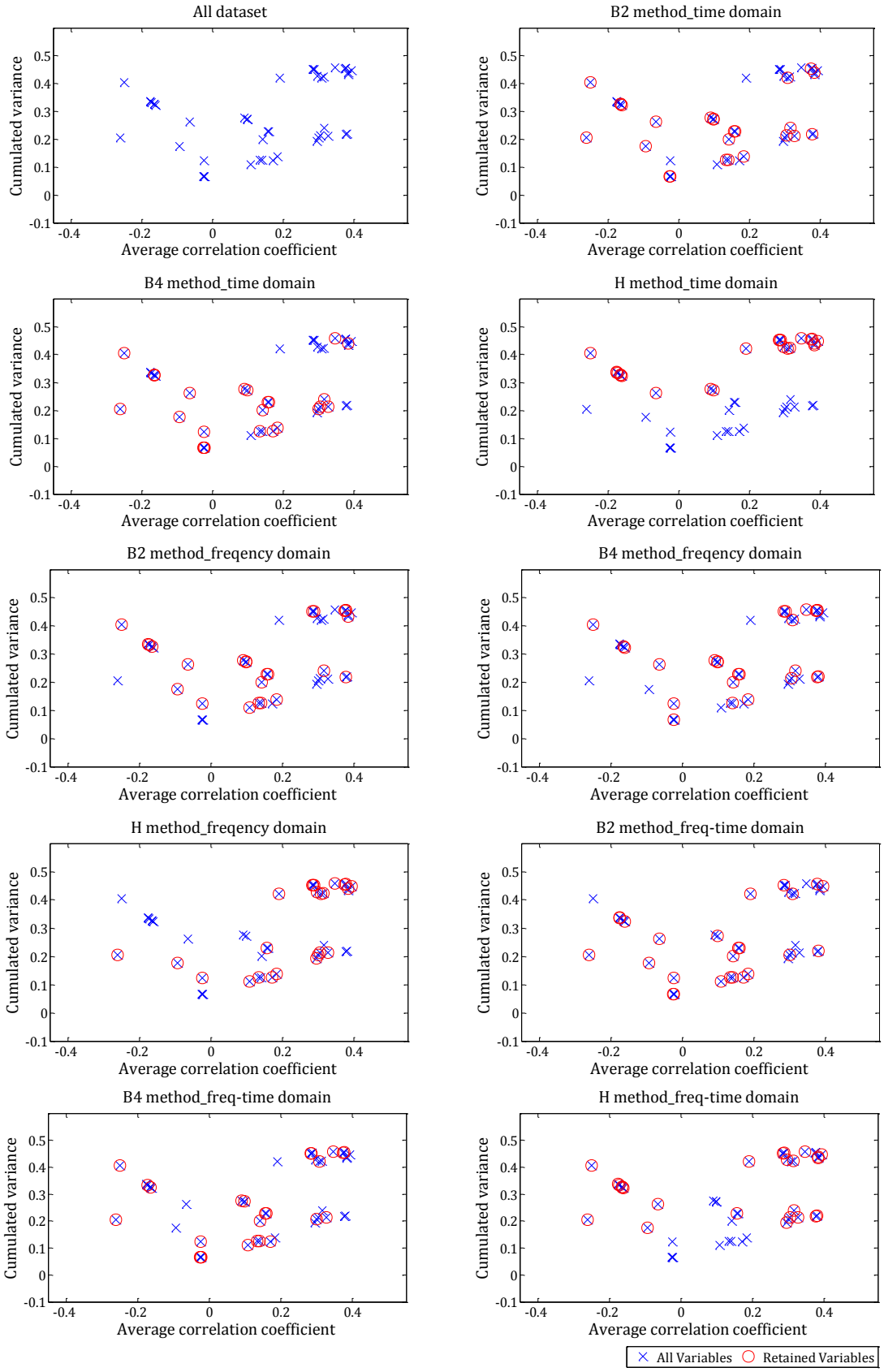


Figure 4.12 - Average correlation coefficient vs. cumulative variance with three selection methods in the time, frequency and instantaneous frequency domain using PMSG-based simulation data.

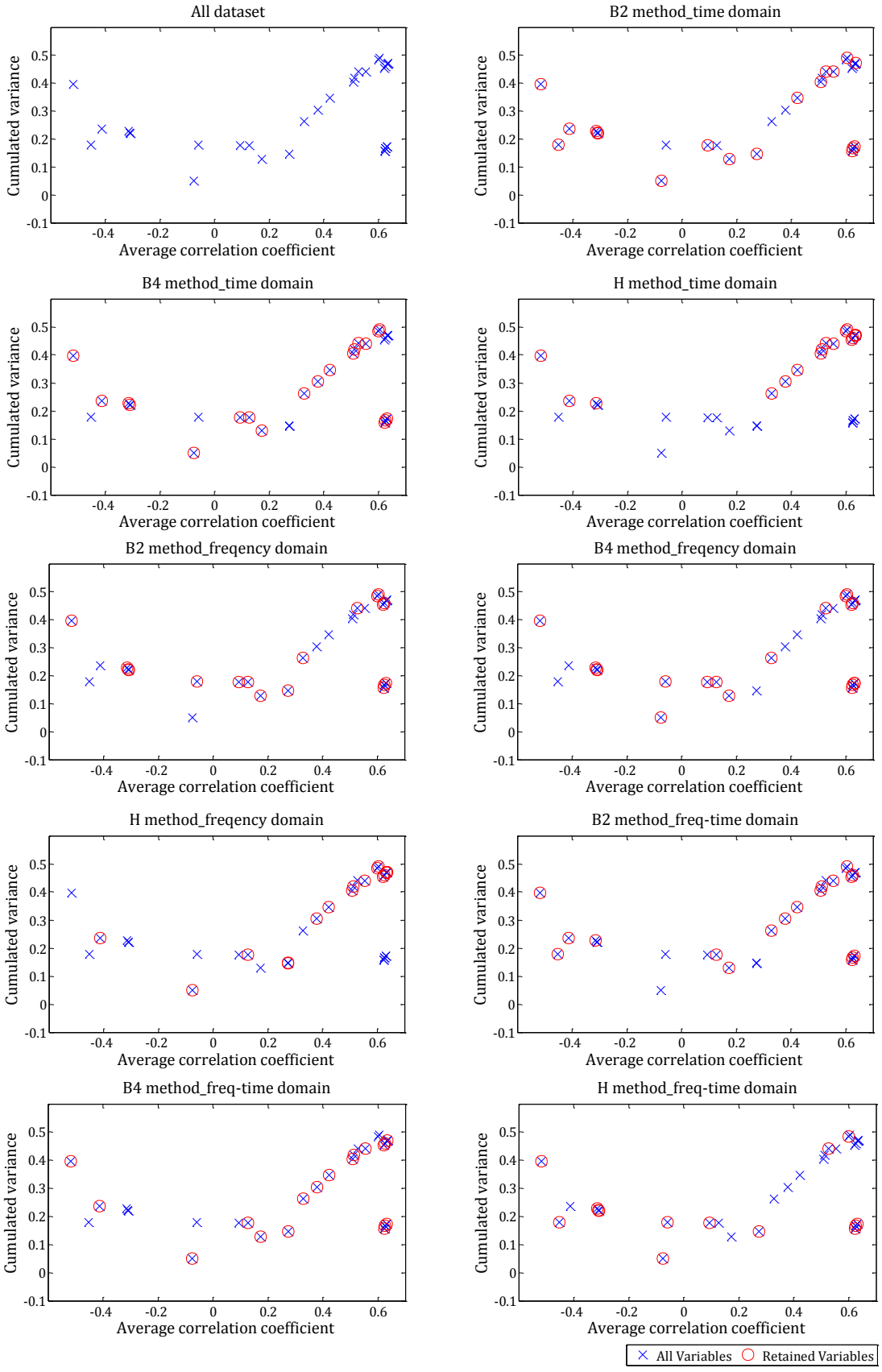


Figure 4.13 - Average correlation coefficient vs. cumulative variance with three selection methods in the time, frequency and instantaneous frequency domain using DFIG-based simulation data

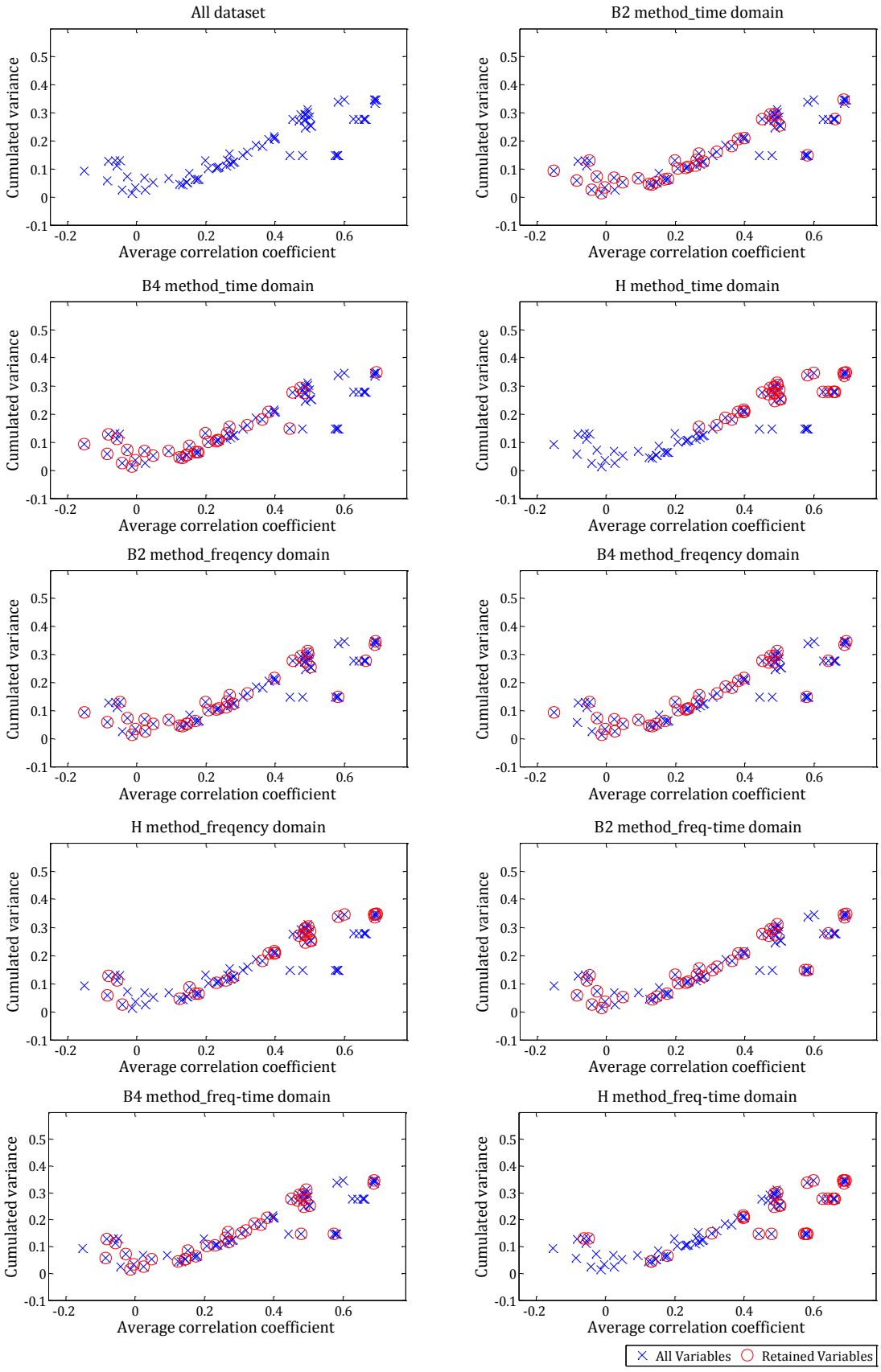


Figure 4.14 - Average correlation coefficient vs. cumulative variance with three selection methods in the time, frequency and instantaneous frequency domain using SCADA data

This finding implies that critical information will also exist in variables with low variances, which is consistent with the result found by Hawkins in his research [141]. As for the selection in different domains, the behaviour, as described earlier, still holds but is less observable in the frequency and instantaneous frequency domains, which further explains the relative low performance.

Finally, Figure 4.15 shows the combination performance measure calculated using equation (4-29). Results from three selection methods using different data in all domains are shown in the three plots, respectively. It can be seen that the B2 and B4 methods in the time domain have the best performance, as they have the lowest *cpm* (described in Section 4.2.3.4). The *cpm* of the B2 and B4 methods in the frequency domain has an increasing behaviour with PMSG, DFIG and SCADA data. This is because the PMSG dataset might have relatively less frequency variations compared to the other two datasets, as the system is connected through a direct AC-DC-AC inverter. In contrast, for the DFIG dataset, the system's frequency domain features are more abundant, due to the constant variation of the rotor current and frequency as a result of the inverter control. Similarly, for the SCADA dataset, more measurements are included which enriches the frequency domain features such as vibration and environmental signals. Overall, the results show that the B2 and B4 selection methods in the time domain outperform the others.

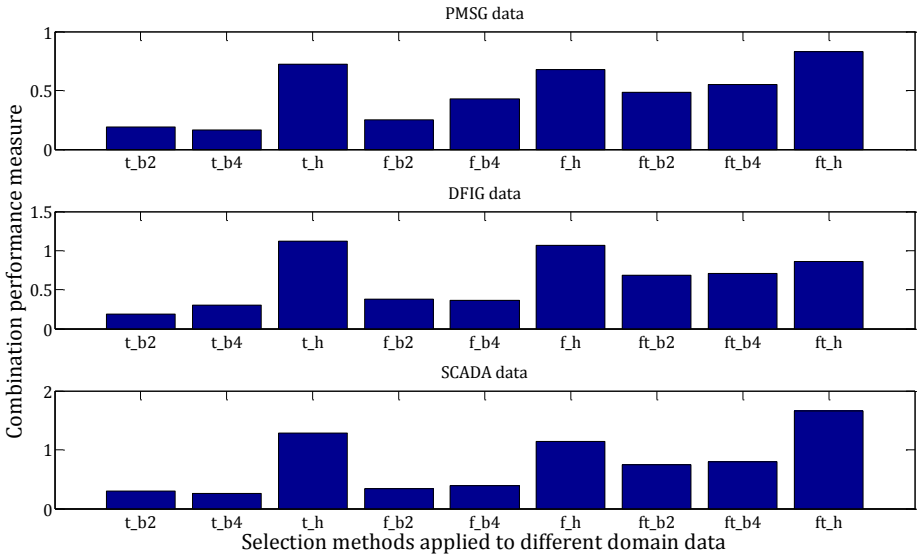


Figure 4.15 - Combination performance measure of different selection algorithms using different data

In terms of the variables retained, for the PMSG simulation data, variables such as pitch angle, torque and active/reactive power at the grid coupling point are always selected. For variables with high dependency among them, such as the active/reactive power, voltage and currents at different locations of the grid, as well as DC bus voltage and currents, at least one variable is retained. As for the DFIG simulation data, apart from torque and active/reactive power at grid coupling point that are retained, parameters related to the machine rotor are also selected. The simulation data is relatively simple in structure, as dominant variables are electrical. On the contrary, the SCADA data have a more complex data composition, as data consist of more signal variability, including various temperatures and environmental parameters. Apart from the general parameters such as gearbox and generator temperatures, oscillations and vibrations, most variables retained are electrical related. Variables such as pitch angle or environmental conditions (e.g. air pressure, ambient temperature) are always selected because they are less likely dependent on other parameters. Furthermore, for variables with a high degree of dependencies such as temperature generators L1 and L2, only one is selected. Finally, it is worth noting that with the wind speed or related variables, at least one is always selected. A complete table of retained variables for each selection method can be found in the Appendix B.1.

4.4.1.2 Effect of sampling frequency on the selection algorithm

In order to assess the robustness of the proposed selection algorithm, the PMSG simulation is carried out at varying sampling frequencies from 50μs to 2000 μs. Figure 4.16 shows an example of the phase current at the point of coupling (as shown in Fig 3.19) using different sampling rates. The blue, green and red plots represent waveforms with sampling rates of 50, 1000 and 2000 μs, respectively. It can be seen that at 2000 μs, certain harmonics, especially during the failing edge, are not captured as in data sampled at 1000 or 50 μs. Moreover, the waveform with sampling rates between 1000 and 50 μs still exhibit a minor difference.

A higher sampling frequency can ensure that more transient features within the signals are captured. The selection algorithms are tested with the data collected

at varying sampling rates, and the combination performance measures are shown in Figure 4.17. The *cpm* calculated from three selection algorithms are plotted individually in the figure. It can be seen that the *cpm* remains almost constant until the sampling rate increases and reaches 1000 μs. As the sampling rate increases above 1000 μs, there is a minor drop of *cpm*. However, overall, the performance of the selection algorithms is relatively insignificant from the sampling rate of the data.

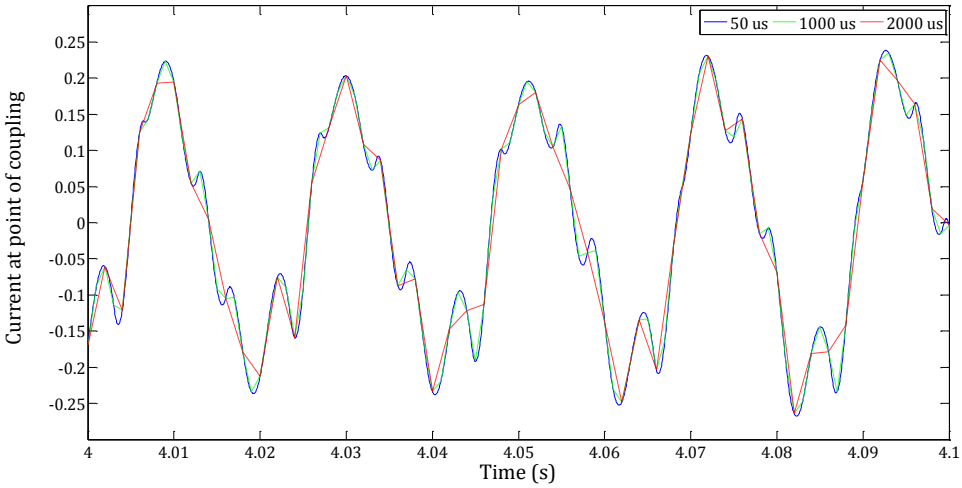


Figure 4.16 - PMSG simulation phase current at point of coupling with varying sampling frequencies

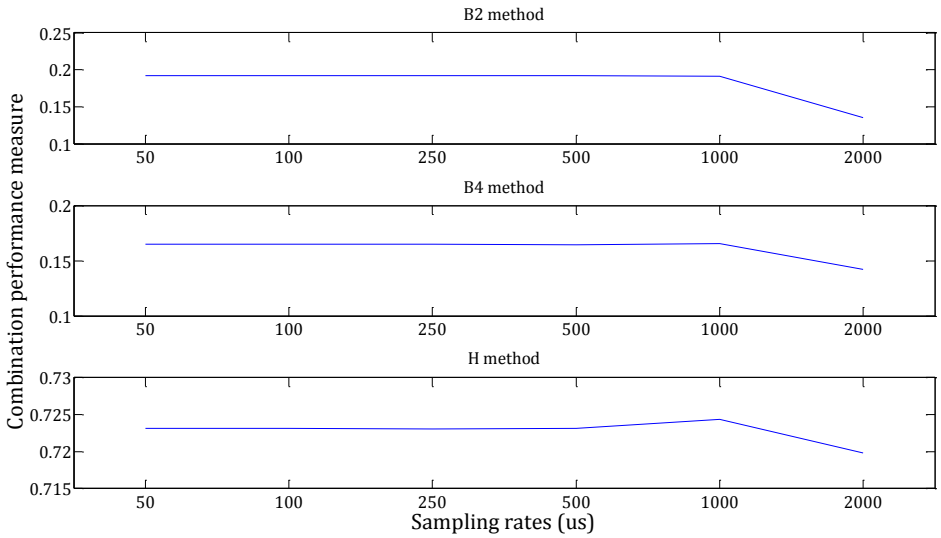


Figure 4.17 - Combination performance measure of three selection methods using PMSG simulation data with varying sampling frequencies

Moreover, Figure 4.18 shows the retained variables using three selection algorithms with data using different sampling rates. The x axis shows the variables in the dataset (Table 4.1), and the y axis shows the number of occasions when each variable is selected. For each dataset, the retained variable set will have a value of 1, and those discarded will have a value of 0. With a total number of 6 datasets used (at different sampling rates), if the retained variables remain the same for each case, the resulting bar chart should have a value of 6 for all retained variables, which indicates the sampling rate has no impact on the selection algorithm. It can be seen that only the H (bottom plot of Fig 4.18) method has the same set of variables retained. On the contrary, for the B2 (top plot) and B4 (middle plot) methods, there is minor difference in the retained variables (variables with a bar value of 1). However, the majority of the retained variables are the same. This can prove that regardless of the sampling rate, the selection algorithms are robust and their effects are insignificant.

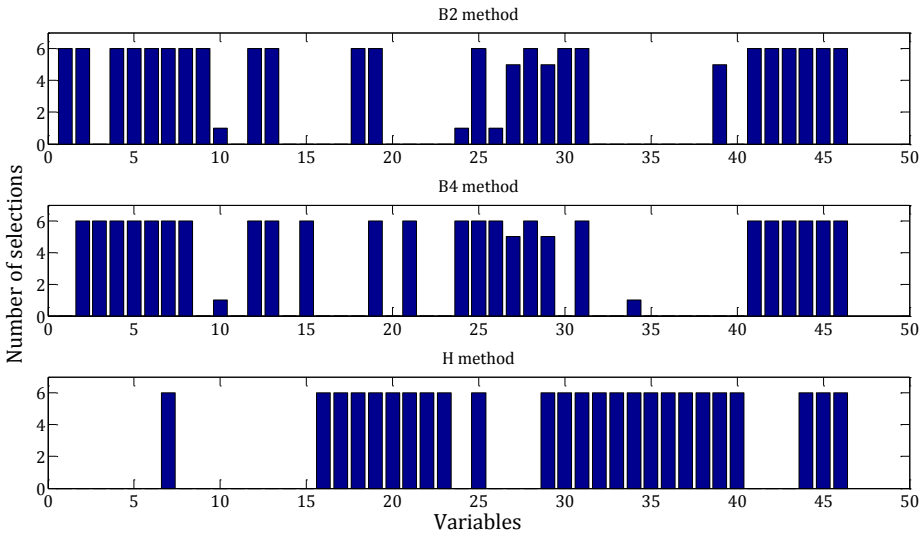


Figure 4.18 - Retained set of variables from three selection methods using PMSG simulation data with varying sampling frequencies

In conclusion, results show that the B2 and B4 methods in the time domain have the ability to identify redundancies and reduce the dimension of the dataset significantly without losing much information.

4.4.2 Validation in faulty condition

During faulty conditions, the variability and redundancies are considered, and the proposed algorithm is able to select variables that are targeted to a particular

fault signal of interest. As proven in the previous section, the proposed selection algorithms using PMSG, DFIG simulation data and SCADA data in the time domain outperforms the other domains. Therefore, in a later section, only time domain data is considered in evaluating the proposed targeted selection algorithm as described in Section 4.2.4.

For the PMSG and DFIG simulation, data with the capacitor ageing fault, single phase-to-ground fault, phase-to-phase fault and three phases-to-ground fault, as described in Chapter 3.2, are used. The data is collected for a period of 10 s at a sampling rate of 250 μs, and the faults are introduced at 5 s. The capacitor ageing fault for PMSG and DFIG data has a severity of capacitance change of -1% and -35% with respect to the fault free condition (0%), respectively. The negative sign indicates that the capacitance value is decreased by a certain percentage compared to the healthy condition. It is worth mentioning that the PMSG system is more susceptible to a DC-link capacitor ageing fault, where the DFIG system is able to withstand a relative high variation of the DC-link capacitance. For the single phase-to-ground, phase-to- phase and three phase-to-ground faults, the severity of fault is simulated with a resistance change of 0.0001% (from the fault free condition (1 MΩ) for both systems. Examples of the waveforms during faulty conditions were presented in the simulation chapter (Figures 3.32 to 3.37). Finally, for the SCADA data, based on the alarm log and investigation of the dataset, turbines with a generator fault and a gearbox fault have been identified and used in a later validation process. Figures 4.19 and 4.20 plot some of the variables from the SCADA data for turbines with the generator fault and gearbox fault, respectively.

To perform the proposed targeted selection algorithm, it is necessary to define a targeting signal for a particular fault. For a capacitor ageing fault, the DC-link voltage is used, as any changes of the DC-link capacitance will have a direct impact on the voltage. In the case of the phase-to-ground faults, the active power at the point of coupling is selected. The targeting signals for the two types of faults are the same for both PMSG and DFIG simulation data. Moreover, the gearbox bearing temperature and the generator winding temperature are nominated as the targeting signal for the two types of faults.

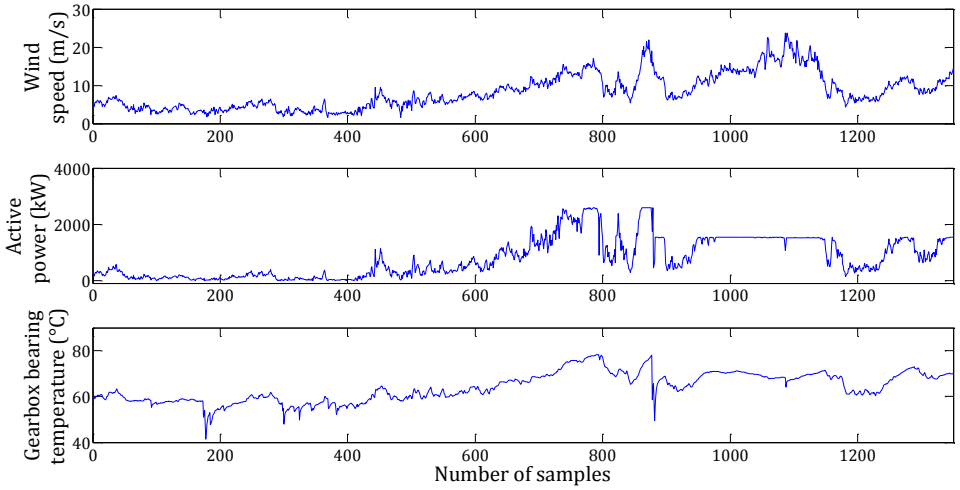


Figure 4.19 - Example of SCADA data with gearbox fault showing wind speed (top), active power (middle) and gearbox bearing temperature (bottom)

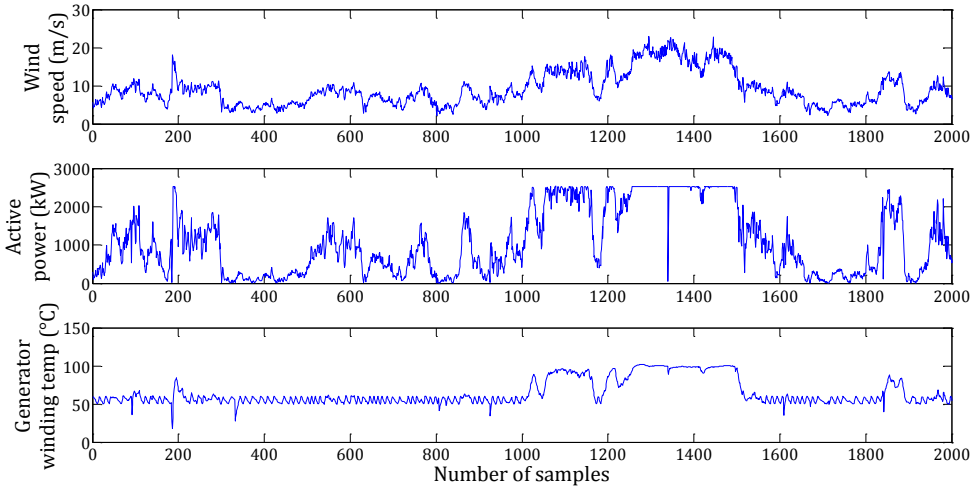


Figure 4.20 - Example of SCADA data with generator fault showing wind speed (top), active power (middle) and generator winding temperature (bottom)

With the targeting variable defined, the T selection algorithm proposed in Section 4.2.4 is applied with the same procedure used in earlier sections. The performance measures as described in Section 4.2.3 are calculated for each dataset. Table 4.4 lists all the performance measures of all the datasets under different faulty conditions. The targeting variable for each case is also included. It can be seen that all datasets have a *cppv* above 97%, indicating the retained variable set obtained from the T selection algorithm has a high percentage of variance of the original dataset. Moreover, there are significant reductions of the average correlation compared to the original data, which implies low

redundancies within the retained variable set. Finally, reasonable percentage entropy is also obtained.

PMSG simulation data					
Type of data	Original data	Capacitor ageing fault	Single phase-to-ground fault	Phase-to-phase fault	Three phase-to-ground fault
Target signal		*DC-link voltage*	*Active power at PCC*	*Active power at PCC*	*Active power at PCC*
Cumulative variance	1	99.61%	99.47%	99.45%	98.82%
Average correlation	0.1435	0.0618	0.0377	0.0373	0.0862
Percentage entropy	1	87.50%	85.83%	81.30%	82.79%
DFIG simulation data					
Type of data	Original data	Capacitor ageing fault	Single phase-to-ground fault	Phase-to-phase fault	Three phase-to-ground fault
Target signal		*DC-link voltage*	*Active power at PCC*	*Active power at PCC*	*Active power at PCC*
Cumulative variance	1	96.15%	98.32%	97.59%	98.44%
Average correlation	0.3596	0.0616	0.0176	0.0878	0.0754
Percentage entropy	1	76.67%	80.16%	80.58%	76.79%
SCADA data					
Type of data	Original data	Gearbox fault	Generator fault		
Target signal		*Gearbox bearing temperature*	*Generator winding temperature*		
Cumulative variance	1	97.11%	97.42%		
Average correlation	0.3412	0.0677	0.0588		
Percentage entropy	1	75.91%	78.09%		

Table 4.4 - Results from targeted selection algorithm using simulation and SCADA data with various types of faults in time domain

The retained variables are very similar to those obtained using data under a normal operating condition. In the case of PMSG data, variables such as pitch angle, torque and power at PCC are always retained with data under different

faulty conditions. However, for variables with high dependency, the selection is determined by the targeting variable rather than the variance of the data as in the B2 or B4 method. Thus, the algorithm selects the variable which shares a common feature with the targeting variable, e.g. for PMSG data under a capacitor ageing fault, between DC-link currents 1 and 2 (as shown in Fig 3.19), which are currents before and after the DC-link capacitor; the DC-link current 1 is selected using the T selection algorithm. This is because the current before the capacitor would have more transient change during the fault than the DC-link current 2, as can be seen on the seventh plot in Figure 3.32.

Similar results are obtained for the DFIG data, where wind speed, mechanical speed and electrical power are always selected, even with data under different faulty conditions. Moreover, the generator rotor voltage and current are always retained as well, which is because the high variability of the rotor voltage and current due to inverter control. Overall, the retained variables for data with a capacitor ageing fault and grid phase-to-ground faults are similar. This can be a result of the system configuration of the DFIG wind turbine, as any fault in the grid will also have an impact on the DC-link of the WT, hence the rotor current and voltage. This result also suggests that the composition of the simulation data is relatively simple, as the majority of the data is electrical related. A complete list of the retained variable sets for simulation data can be found in the Appendix B.2.

Lastly, for the SCADA data, as the targeting variable used for the gearbox and generator faults are both temperature based, an obvious difference for the retained variable set from those obtained with B2 or B4 is that less electrical related parameters are selected. Moreover, parameters such as pitch angle or environmental conditions (e.g. pressure, wind direction) and vibrations are always selected, similar to the results in the previous section (B2 or B4 method). It should be noted that the retained variables selected by the T selection algorithm should share common features with the targeting variable in the lower dimensional space, but this does not necessary mean the selected variables must be physically close to it. For example, the gearbox bearing temperature is used for the data with a gearbox fault; this does not mean all variables related to the

gearbox should be retained. In fact, if that was the case, the degree of redundancy of the retained dataset could potentially increase.

With these results, it can be determined that the proposed selection algorithms are able to identify and reduce redundancies even during faulty conditions, yet still persevering dominant information and maintaining low inter-correlation between the retained datasets. However, whether or not the fault features are present in the retained variable set is a critical question. This issue will be discussed and further validated in the next chapter using the retained variable sets obtained from this section.

4.5 Summary and discussion

In this chapter, several PCA-based variable selection algorithms have been proposed. The objective of the algorithms is to identify redundancies within a dataset such that the dimensionality of the dataset can be reduced by eliminating these redundancies, but without sacrificing a high amount of information lost. Three selection algorithms (B2, B4 and H) have been proposed to carry out variable selection at the system level based on the general variability of the dataset during normal operation conditions. Results from these selection algorithms have been evaluated with three performance measures: the *cppv*, the average correlation coefficient and the percentage entropy. The algorithms have been tested against datasets from different systems and representation domains. Results show that the B2 and B4 algorithms using data in the time domain are able to achieve the objective, and the results are consistent and robust for different types of datasets. Moreover, the T method is also proposed for targeted selection during a faulty condition. The method is not only able to identify an optimal set of variables with minimal redundancies and maximising information but also seeks the underlying relationship with the predefined target signal. Positive results are obtained based on the performance measures. However, further validation is needed to prove the fault signatures are preserved in the retained variable set. This will be discussed in detail in the next chapter.

Chapter 5. Evaluation of Variable Selection using Artificial Neural Network

The objective of the target selection method is to select a set of variables that have minimal information repetition and still maintain vital information relating to the targeted fault signal. In the previous chapter, information redundancy of the retained variable set has been assessed. The question as to whether vital information relating to the particular fault is still present in the retained variable set using a proposed selection algorithm is discussed and evaluated in this chapter. Two methods are adopted, which are feature-based and ANN (artificial neural network) prediction model- based. The evaluation process associated with each method is described. Thereafter, an analysis of the results from each method is provided. In order to test the robustness of the results, different datasets, as described in the previous chapters, are used in the examination of the individual methods.

5.1 Feature-based evaluation method

5.1.1 Methodology description

As the transformed principal components are uncorrelated and each PC represents a particular feature within the dataset, for any given dataset which contain fault signals, the fault signature should be reflected in the PCs as a unique feature. Based on this, the feature-based evaluation method is proposed, and it is divided into two steps. The first step is the identification of the principal component which may reveal the fault signature in the original dataset. The second step involves a comparison between the identified PC from the original dataset and the PCs obtained from the retained set of variables. Pearson's correlation coefficient is used as a measure to test and quantify the result.

Figure 5.1 shows the block diagram of the evaluation process. In order to identify the PC containing the fault feature (step 1), an original dataset with n number of samples and p variables of a particular fault at different severity levels α_1, α_2, ... , α_m is obtained and transformed into principal components. As the fault level increases, it will have a higher effect on the system, hence resulting in a larger change of waveforms of the measurements. Because entropy measures the information content based on the probability distribution of the data, any variation in the measurements would result in a change of the distribution. Therefore, instead of visually inspecting the principal components for any fault signatures, normalised entropy is used to quantify the information content of each PC, consequently identifying the PC that contains the fault feature. Once the fault feature is identified, correlation coefficients between the identified PC from the original dataset and the PCs obtained from the retained dataset at fault level α_k are calculated (step 2). If the fault feature is present in the retained variable set, there should be a high correlation coefficient between them. Thus, it can be determined that critical information associated with the fault is maintained. The proposed method is tested against simulation data and selection results from previous chapters under different types of faults. Moreover, it should be noted that this method requires obtaining data of a particular fault at different levels in order to identify the fault feature (PC); hence, the method is not applicable for

SCADA data. An alternative method is proposed and discussed for SCADA data in later sections.

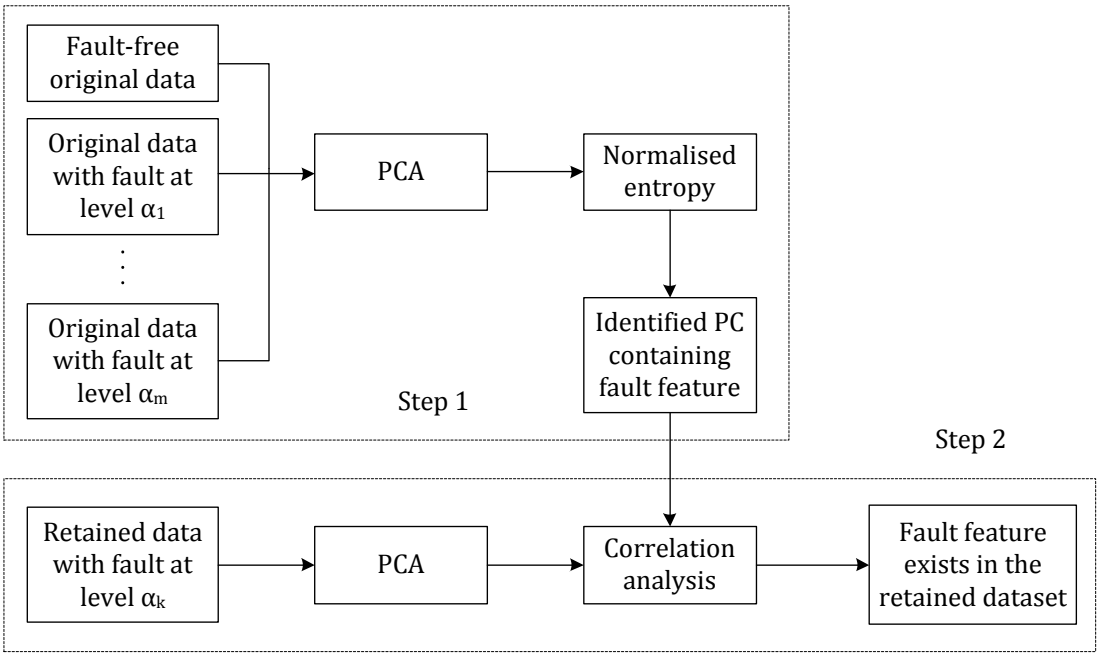


Figure 5.1 - Block diagram for feature-based evaluation procedures

5.1.2 Results and discussion

The DC-link capacitor ageing fault of a PMSG wind turbine is simulated with an increasing severity from no-fault (2.3F) occurred with 18% of capacitance loss at a step of 1%. Data is collected for 10 s at a sampling rate of 250 μs, and the fault is triggered at 5 s for a period of 0.1 s. Figure 5.2 shows the identified PC 7 which contains the fault signature with varying fault intensities between sample points 19800 and 21500 (4.95 s to 5.375 s). It can be seen that when a no fault has occurred (blue line), the waveform remains unchanged, and an obvious change of the waveform can be found during fault conditions. This feature represents the DC-link voltage drop and increase of the DC-link current 2 (as shown in Fig 3.19) when the fault occurs. With increasing fault severity, the peak magnitude of the waveform increases. The blue dots in Figure 5.3 show the change of normalised entropy at varying fault levels, and the red line represents the fitted curve. The results clearly show a decreasing trend of the normalised entropy. Consequently, a more severe fault will result in a larger change of waveform during the fault, which in turn leads to a larger change in the normalised entropy.

Figure 5.4 shows the comparison of features for the DC-link capacitor fault found in the original dataset and the PCs of the retained variables set obtained with the T selection method at a 1% capacitance loss. The top plot in Figure 5.4 plots the 7th PC obtained from the original dataset, and it represents the fault feature as proven earlier based on the normalised entropy. The middle plot shows the Pearson's correlation coefficients between the identified fault feature (7th PC) from the original dataset and all the PCs obtained from the retained dataset. It can be seen that only the 7th PC from the retained dataset has a significant correlation coefficient of 0.9562 and the remaining PCs are all close to 0. The PC found from the retained dataset is shown in the bottom plot. It can be seen that the trends of the fault feature found in the original (top plot) and the retained dataset (bottom plot) are the same. This proves that the proposed selection algorithm is able to select a set of variables that maintains vital information about the fault.

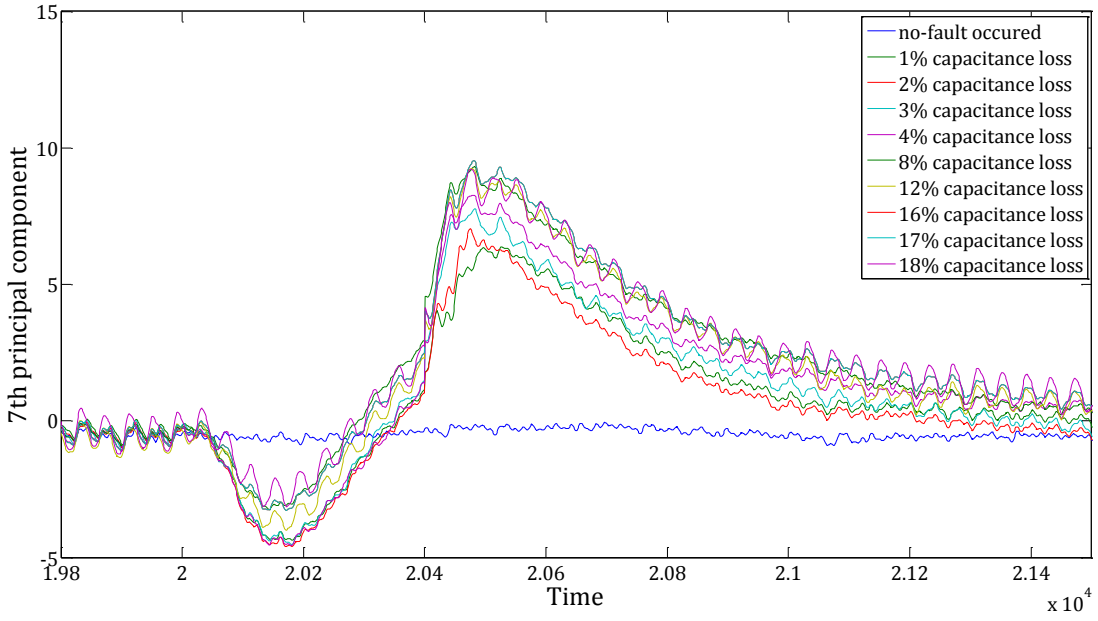


Figure 5.2 - The featured principal components of the DC-link capacitor fault at different ageing levels from the PMSG data

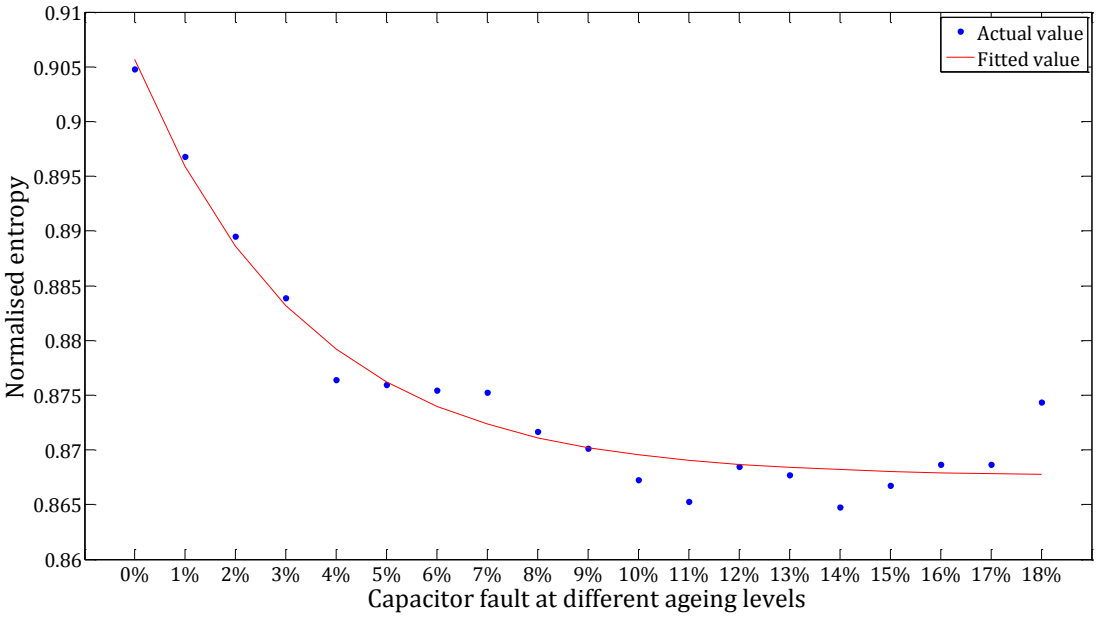


Figure 5.3 - Normalised entropy of the 7th principal components of the DC-link capacitor fault at different ageing levels from the PMSG data

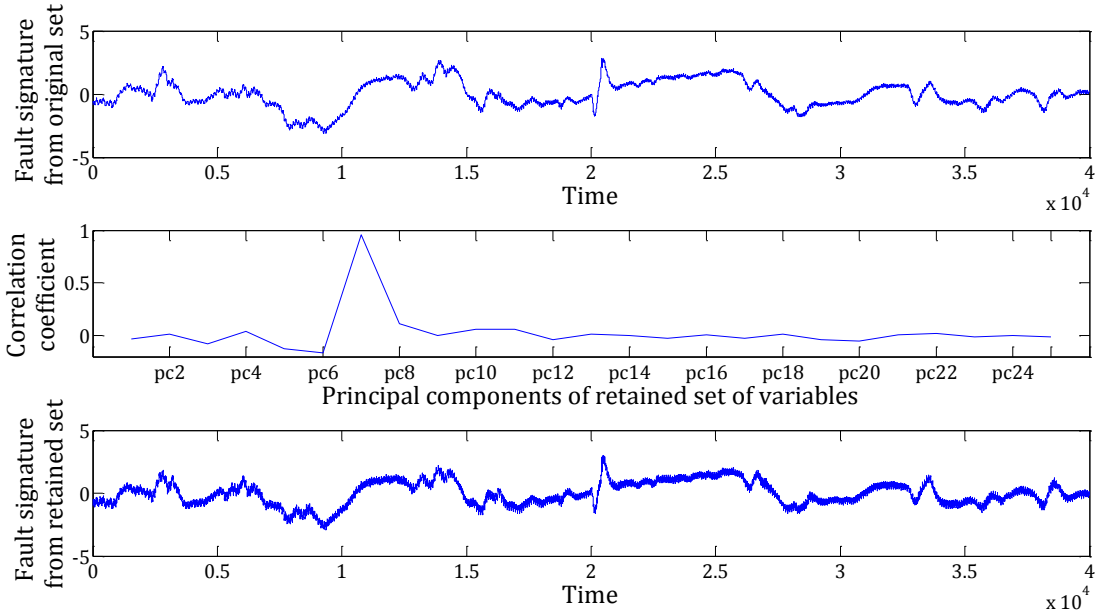


Figure 5.4 - Comparison of DC capacitor fault between the original dataset and retained variables from the PMSG data

PMSG datasets with representative faults including a single phase-to-ground fault, a phase-to-phase fault and a three phase-to-ground fault are also evaluated. As described in Chapter 3, the fault severity is simulated with resistance change. Fault severity ranges from no-fault occurred (1 MΩ) to 10kΩ, 1kΩ, 100Ω, 50Ω, 10Ω, 5Ω, 1Ω, 0.5Ω, 0.1Ω, 0.05Ω and 0.01Ω, which are used in the evaluation process. Again, data is collected for a 10 s simulation with a sampling frequency of 250 μs, and the fault is applied at 5 s for a duration of 0.1 s. It can be seen from Figures 5.5, 5.8 and 5.11 that the patterns of the fault signature occurred at sample points 19800 to 21800 (4.95 s to 5.3 s) with varying severities that are similar for all three types of faults.

During fault conditions, a large magnitude of a fault current is present in the system. Transients are observed in those short-circuit currents. The transient currents consist of an AC component with a frequency equal to the grid frequency and a DC component that decays exponentially in several cycles. The magnitude and shape of transient current are determined by the fault location in the network, the type of the fault and the phase angle between the initiation of the fault and zero voltage. This is expected as the three fault types have the same intrinsic nature.

Figures 5.6, 5.9 and 5.12 show the normalised entropy for the three phase-to-ground faults. From the figures, there is the same trend of entropy change at different fault levels. The plots show that the normalised entropy has minimal change from the no-fault occurrence (1 MΩ) to 10Ω and 1Ω to 0.01Ω. Additionally, between 10Ω to 1Ω, the rate of change of normalised entropy of the fault signature is the greatest across all three fault types. However, there are two main differences between the entropy plots for the three fault types. One difference concerns the slop change during the dominant variation region, where the gradient decreases when the fault changes from a single phase ground fault to a more detrimental phase-to-phase short circuit fault and a three phase-to-ground fault. In regard to the second difference, the magnitude of the normalised entropy also differs for the three fault types. These may imply the number of faulty phases occurring.

Finally, a comparison of fault features found in the original dataset and PCs of the retained variables set for each type of fault is shown in Figures 5.7, 5.10 and 5.13. It is worth noting that the fault feature in the original dataset for each fault is not found on the same principal component, as the impact of each fault to the system is different. The 3rd, 1st and 1st featured PC are found to include the fault signature of each fault type, respectively, as shown in the top plot of Figures 5.7, 5.10 and 5.13. Moreover, the middle plot in these figures shows the correlation coefficients between the identified fault feature and the PCs of the retained variables. For the single phase-to-ground fault and phase-to-phase fault, the high correlation coefficient values of 0.9619 and 0.8845 are found on the PC 3 and PC 1 of the retained dataset, respectively, and the remaining PCs all have values close to 0. For the three phase-to-ground fault, the highest value of 0.7811 is found on PC 2. Also, the first PC from the retained variables set has a reasonable high correlation of 0.567. This may because the first PC also contains a certain amount of fault signature. However, the dominant signature is still present in PC 2. Lastly, the identified PCs (PC 3, PC 1 and PC 2) from the retained dataset for each fault are plotted in the bottom of Figures 5.7, 5.10 and 5.13. Again, it can be seen that these PCs are practically the same with the fault features (top plots) found from the original dataset.

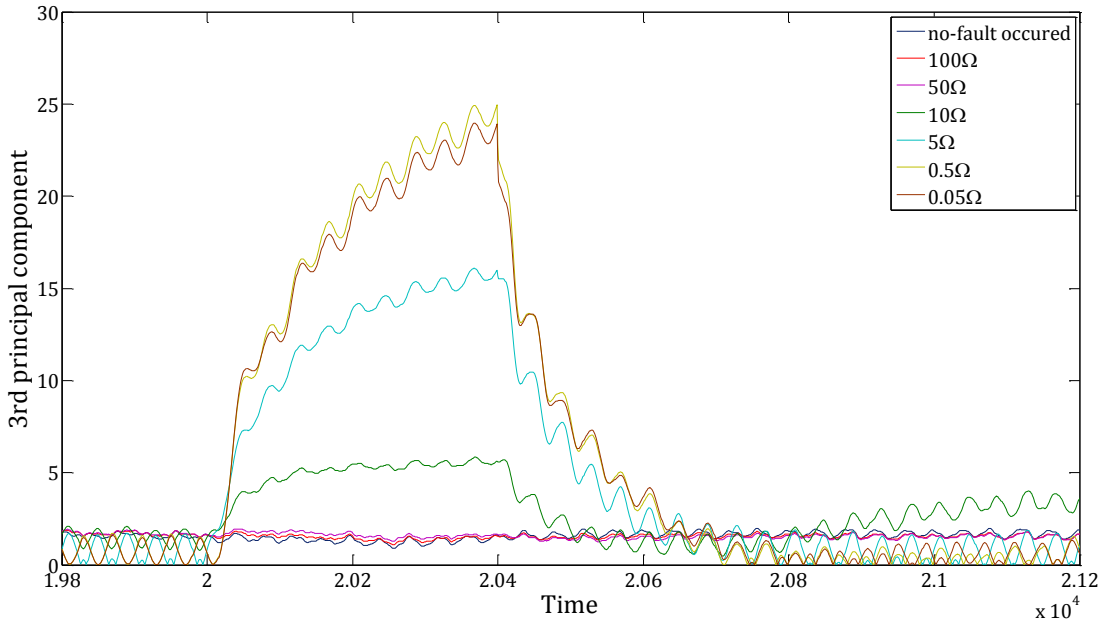


Figure 5.5 - The featured principal components of the single phase-to-ground fault at different levels from the PMSG data

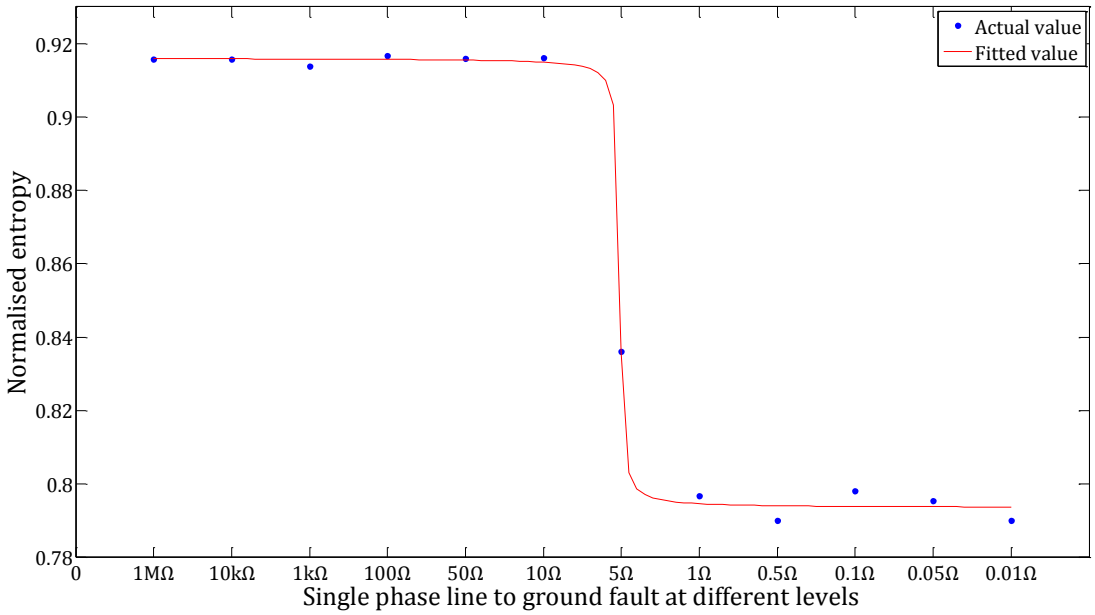


Figure 5.6 - Normalised entropy of the 3rd principal component of the single phase-to-ground fault at different levels from the PMSG data

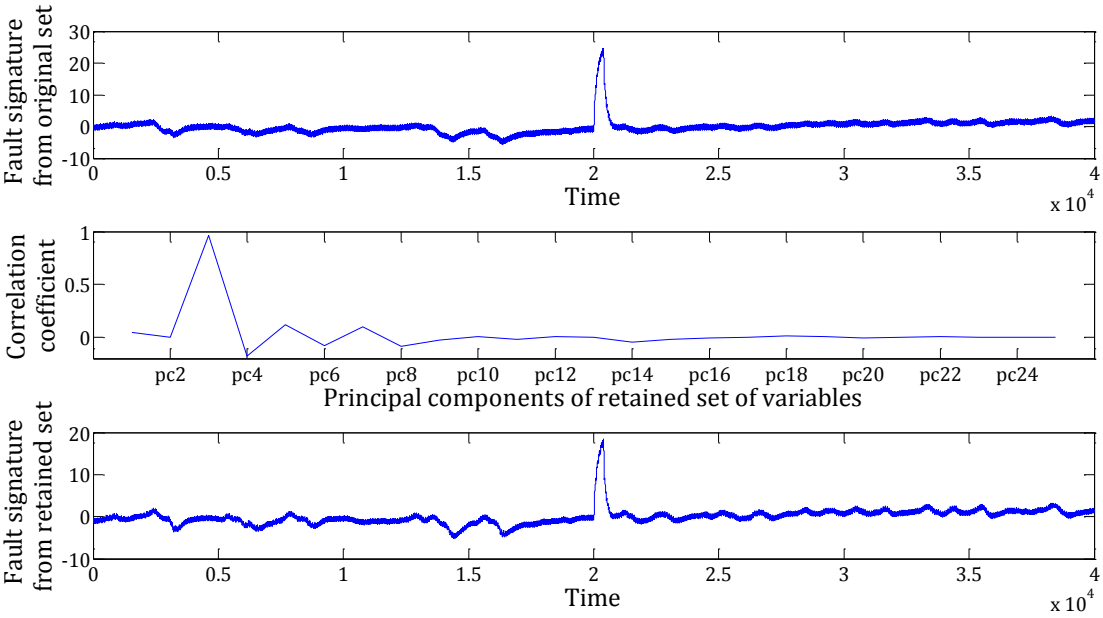


Figure 5.7 - Comparison of single phase-to-ground fault between the original dataset and the retained variables from the PMSG data

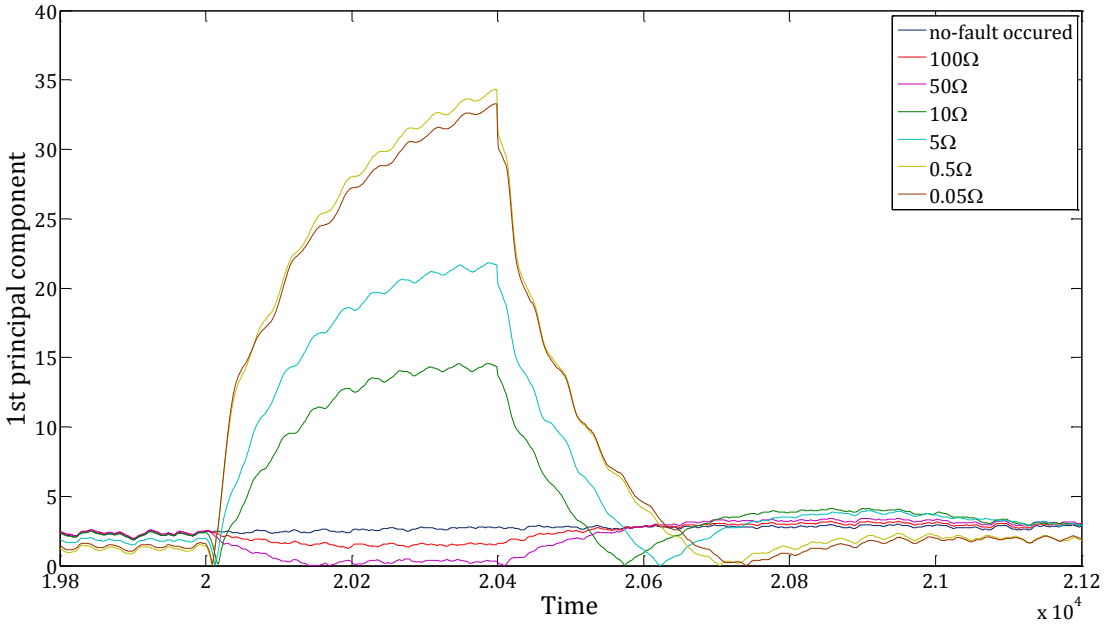


Figure 5.8 - The featured principal components of the phase-to-phase fault at different levels from the PMSG data

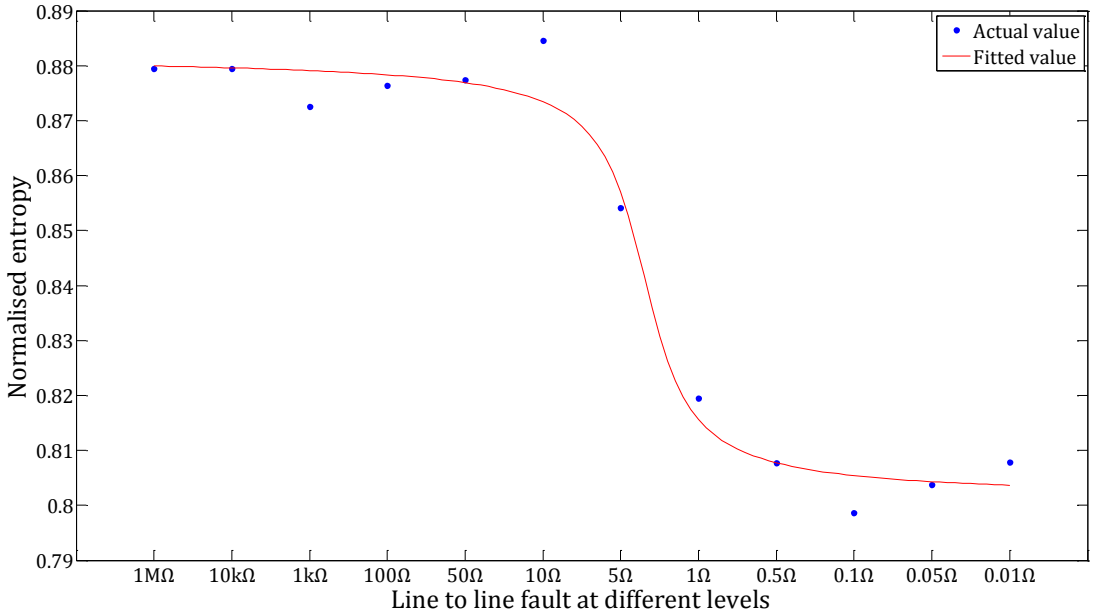


Figure 5.9 - Normalised entropy of the 1st principal component of the phase-to-phase fault at different levels from the PMSG data

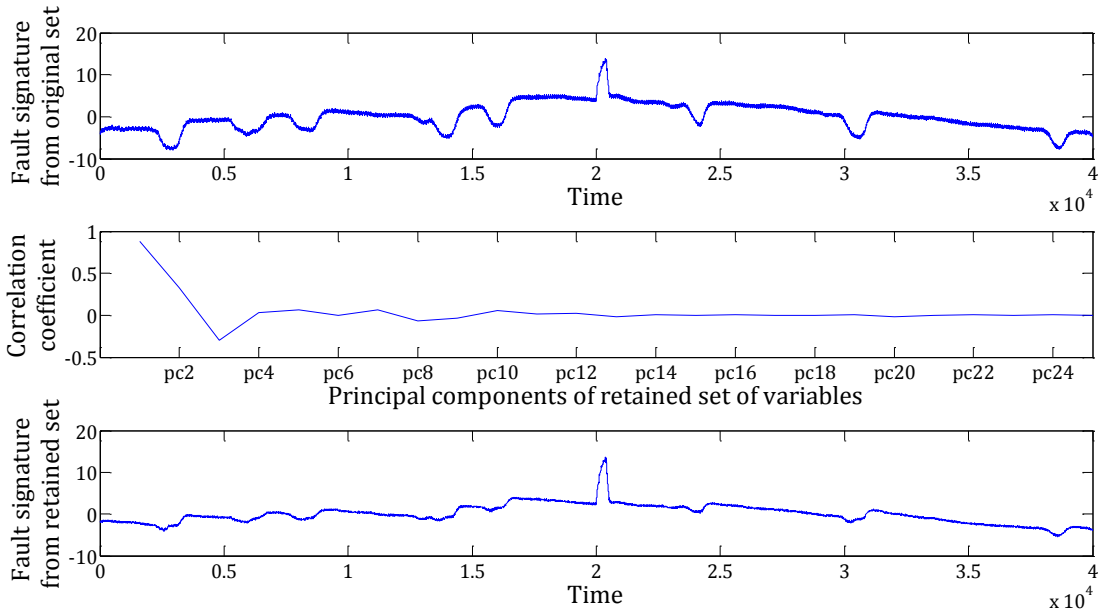


Figure 5.10 - Comparison of the phase-to-phase fault between the original dataset and the retained variables from the PMSG data

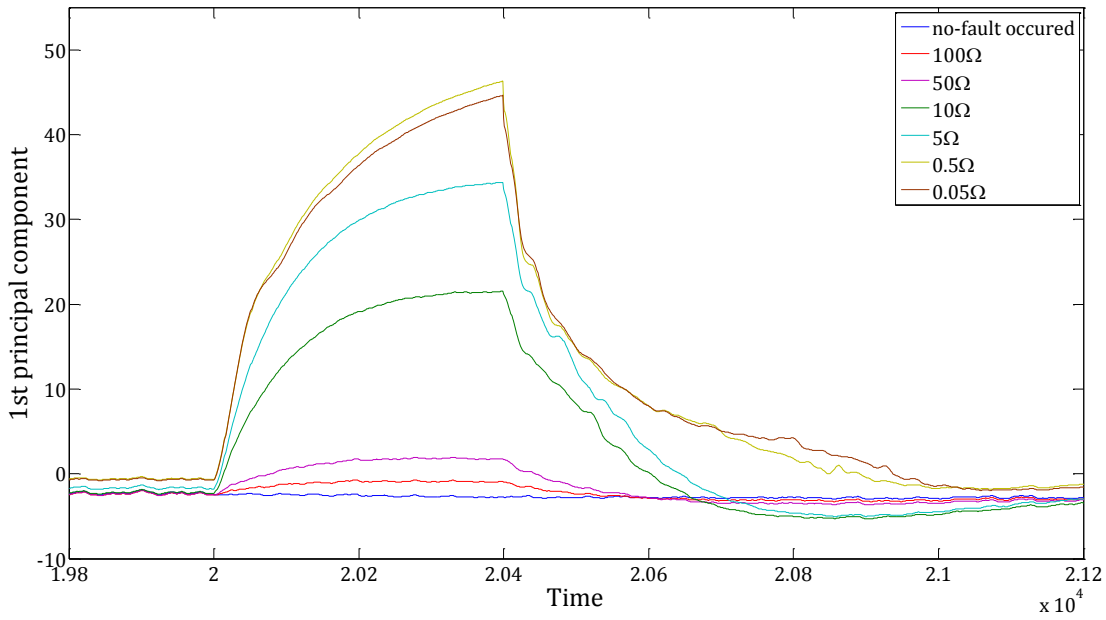


Figure 5.11 - The featured principal components of the three phase-to-ground fault at different levels from the PMSG data

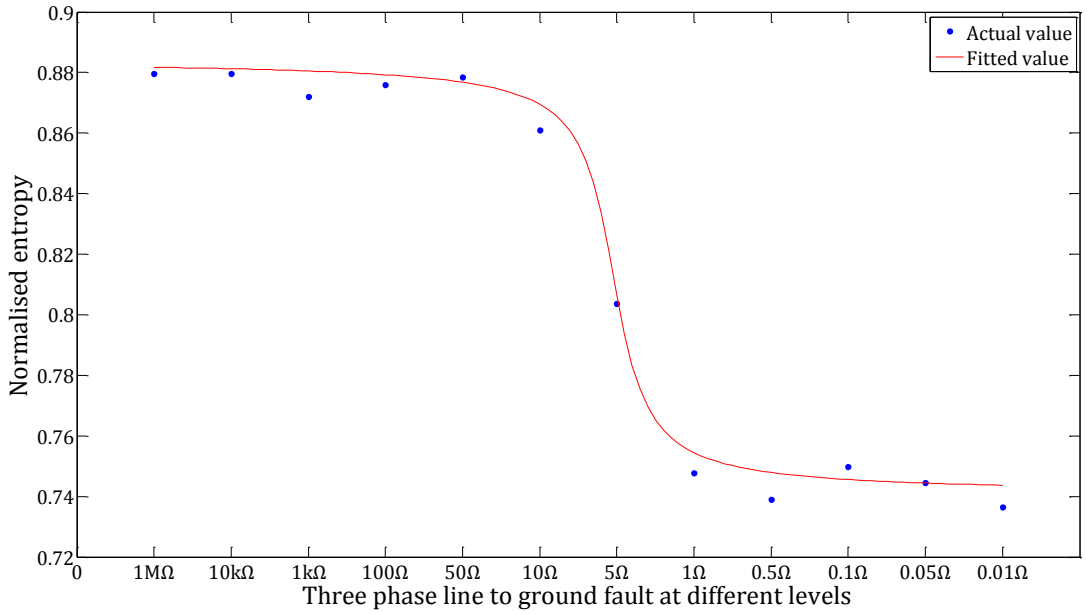


Figure 5.12 - Normalised entropy of the 1st principal component of the three phase-to-ground fault at different levels from the PMSG data.

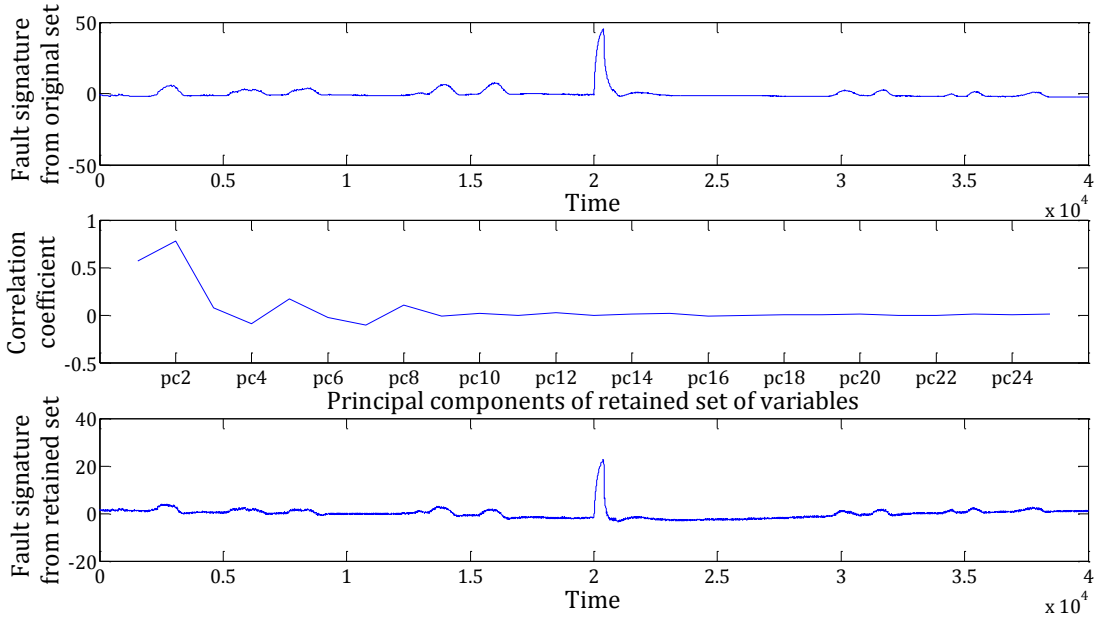


Figure 5.13 - Comparison of three phase-to-ground fault between the original dataset and the retained variables from the PMSG data.

In order to validate the robustness of the method, the DC-link capacitor ageing fault for the DFIG wind turbine is also evaluated. The above methodology should be applicable in spite of the system, as the method explores the fault feature at a

lower dimension of the measurements data. Because the DFIG has a partial AC-DC-AC converter, the system is able to withstand more capacitance losses compared to the PMSG wind turbine. Therefore, an increasing fault severity from the no-fault occurred to 50% with a step of 5% simulated for 10 s at the sampling rate of 250 μs. The fault occurs between sample points 20000 (5 s) for a duration of 0.1 s. Results are shown in Figures 5.14 - 5.16. Figure 5.14 plots the fault signatures from the original dataset at different fault severities between sample points 19500 and 21500 (4.865 s to 5.25 s). It can be seen that there is still an increasing trend of the peak amplitude, but it is different from the PMSG data. The two spikes at sample points 20000 and 20400 are related to the increase of DC-link currents due to the capacitance change. Figure 5.15 shows a similar trend of the normalised entropy compared to the normalised entropy plot from PMSG data, where the normalised entropy decreases as the fault severity increases. Finally, the top plot in Figure 5.16 shows the fault signature (PC 8) found from the original dataset with a loss of capacitance of 35%. The correlation coefficients between the identified fault signature from the original data and the PCs from the retained dataset obtained in Chapter 4.4.2 are shown in the middle plot of Figure 5.15. A highest correlation of 0.7879 is found on the 6^{th} PC from the retained dataset, and the remaining PCs are insignificant. It is worth noting that the fault feature (PC) found in the original dataset and the retained dataset does not necessarily need to be the same, as the dimensions of these datasets are different. Finally, the fault feature (PC 6) found in the retained variable set is plotted in the bottom plot.

The results clearly demonstrate that the variables retained by the targeted selection algorithm are able to maintain vital information associated with the fault. However, the feature-based evaluation method requires measurements from the same fault at different fault levels in order to identify the corresponding principal component. Therefore, SCADA is not able to be evaluated with such a method. In the next section, the ANN prediction model-based method is proposed and further investigates whether dominant features relating to the fault are present in the retained variables using SCADA data.

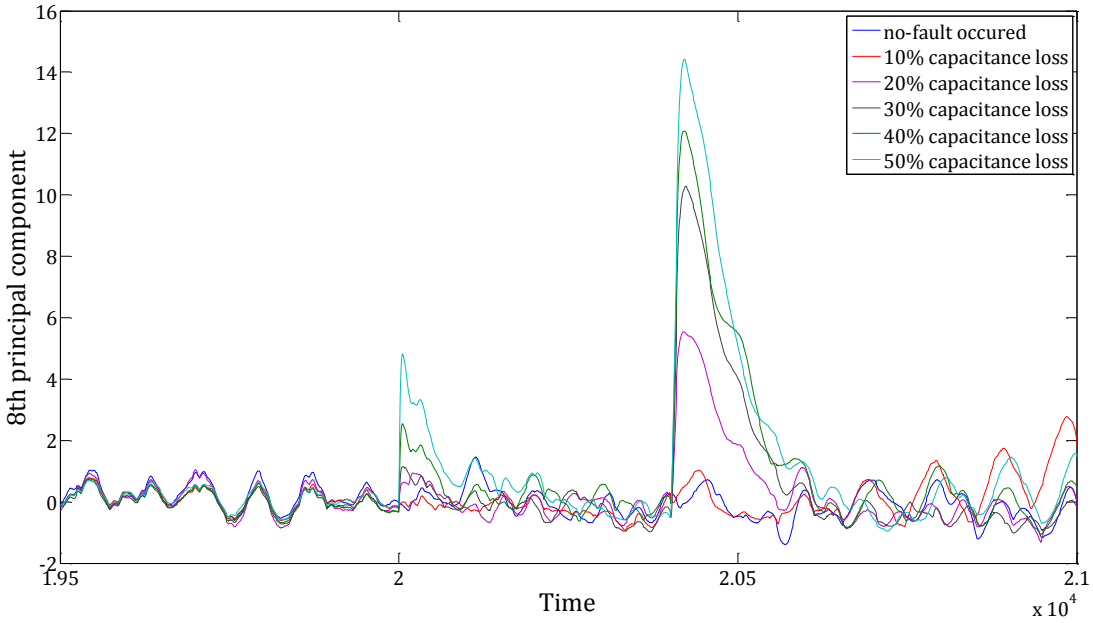


Figure 5.14 - The featured principal components of the DC-link capacitor fault at different ageing levels from the DFIG data

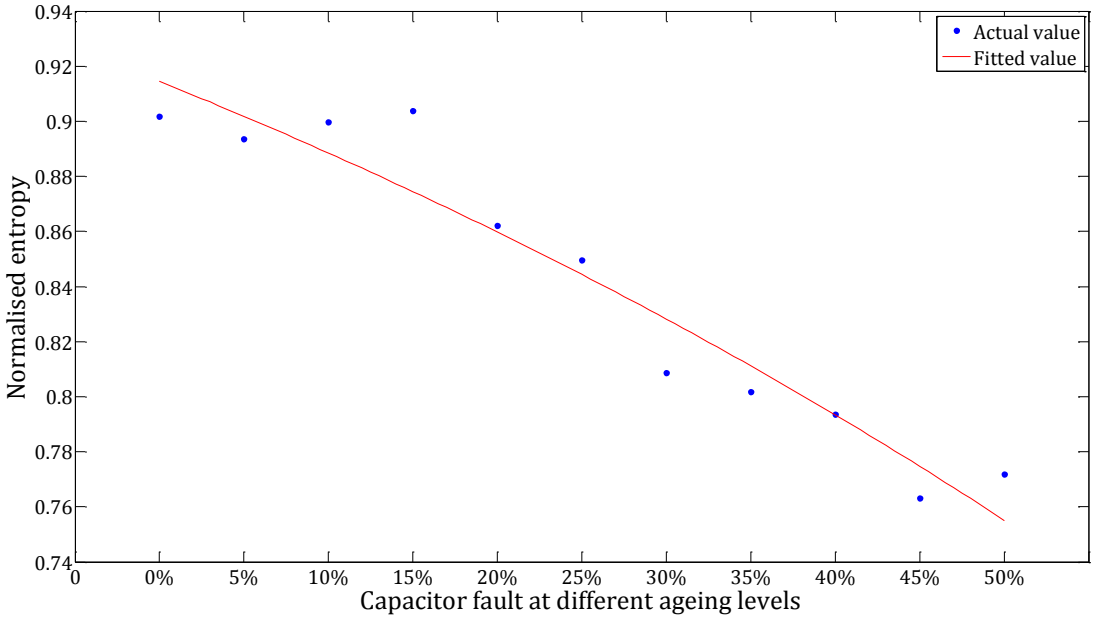


Figure 5.15 - Normalised entropy of the 8th principal components of the DC-link capacitor fault at different ageing levels from the DFIG data

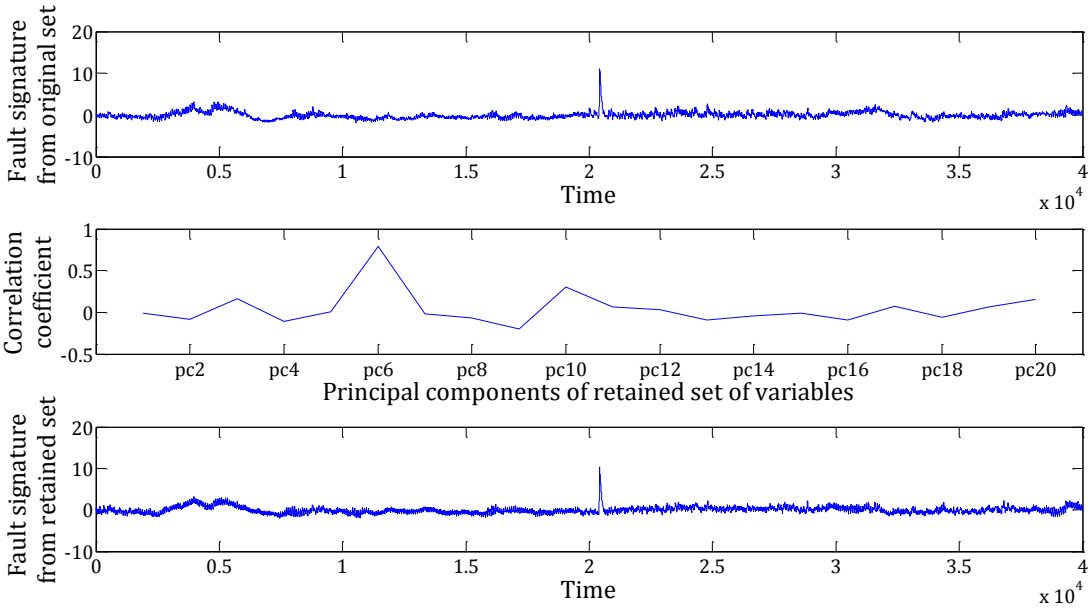


Figure 5.16 - Comparison of DC-link capacitor ageing fault between original dataset and retained variables from the DFIG data

5.2 NARX (nonlinear autoregressive exogenous) ANN model

The ANN model is dynamic, non-parametric and nonlinear; thus, it has a vast application in wind power forecasting, classification, anomaly detection and other elements. [142-144]. The neural network is a mathematical model which imitates a human's biological neuronal system. It has a structure of numerous nodes and layers as shown in Figure 5.17. Generally, for a typical three-layer model, it comprises an input layer with p number of input variables, x_1, x_2, …, x_p, and neurons of n_{i1}, n_{i2}, …, n_{ip}. The output layer has m number of outputs, y_1, y_2, …, y_m, and neurons of n_{o1}, n_{o2}, …, n_{om}. For the hidden layer, it has j number of neurons, n_{h1}, n_{h2}, …, n_{hj}. The inputs **u** and outputs O_j for each neuron in any layer are associated through weights w_1, w_2, …, w_j, a bias parameter b and an activation function as shown in Figure 5.17(b). The mathematical representation of the relationship is:

$$O_j = f\left(\sum_{k=1}^{j} w_k u_k\right) \tag{5-1}$$

where f is the activation function. Then, the relationship between inputs **X** and outputs **Y** of the model is given:

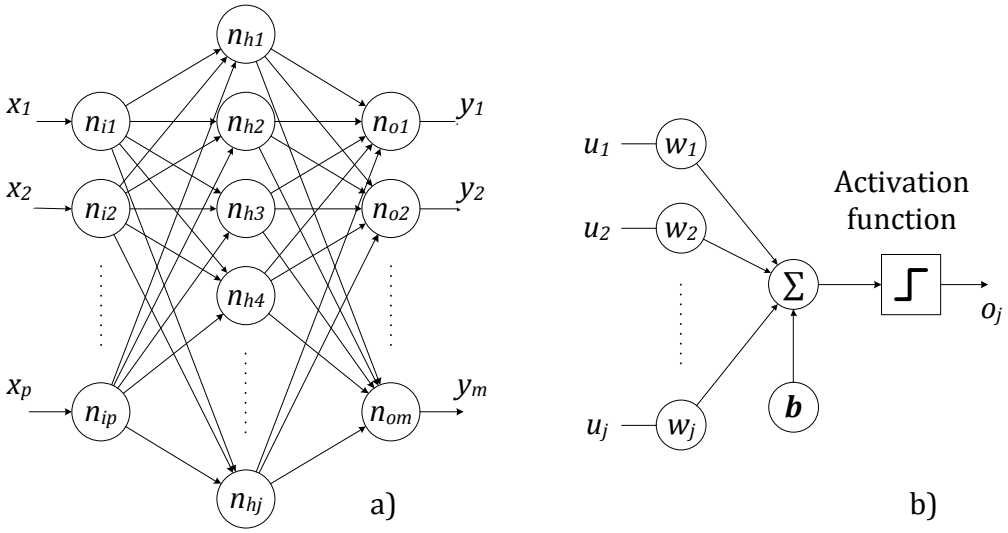


Figure 5.17 - Structure of ANN: a) typical three-layered ANN structure; b) structure for each neuron within the network

$$\mathbf{Y} = f^o\left(\mathbf{W^o} f^h\left(\mathbf{W}^h f^i\left(\mathbf{W}^i\mathbf{X} + \mathbf{b}^i\right) + \mathbf{b}^h\right) + \mathbf{b}^o\right) \tag{5-2}$$

where **X** and **Y** are the inputs and outputs of the model in matrix form, $\mathbf{W}^i$, $\mathbf{W}^h$ and $\mathbf{W}^o$ are the weight matrix for the input, hidden and output layers with a dimension of $(1\times p)$, $(p\times j)$ and $(j\times m)$, respectively. **b** is the scalar bias matrix which enables the position of the activation function to shift. The superscript *i*, *h* and *o* represent the input, hidden and output layer respectively. In this study, the hyperbolic tangent sigmoid activation function is adopted, as it is one of the most commonly used functions.

The model seeks a relationship between the inputs and outputs under a supervised learning algorithm known as backpropagation learning. The weights are updated during the 'training' phase, where the model tries to explore the correct weights such that the sum squared error (SSE) between the actual and the predicted outputs are minimised. The SSE described as a function of inputs X and weight **W** is given:

$$\mathrm{SSE}(\mathbf{X}, \mathbf{W}) = \frac{1}{2}\sum_{1}^{n}\sum_{1}^{m} e_{n,m}^2 \tag{5-3}$$

where, $e_{n,\mathrm{m}}$ is the error between the actual y and predicted $\hat{y}$ output, and it is:

$$e_{n,m} = y_{n,m} - \hat{y}_{n,m} \tag{5-4}$$

The new weight matrix after each training iteration can be found by several techniques which depend on the gradient function of the SSE error function. Commonly used algorithms include gradient descent, Newton's method and Levenberg-Marquardt. In nonlinear function fitting, the Levenberg-Marquardt algorithm generally has the best performance [145, 146]; the update rule for the weights at each iteration is:

$$\mathbf{W}_{l+1} = \mathbf{W}_l - \left(\mathbf{J}_l^{\mathrm{T}}\mathbf{J}_l + \mu\mathbf{I}\right)^{-1}\mathbf{J}_l\mathbf{e}_l \tag{5-5}$$

where **W** is the weight matrix, **J** is the Jacobian matrix of the SSE function, μ is the combination coefficient and can be interpreted as a learning rate as $1/\mu$, **I** is the identity matrix, **e** is the actual error of the model depicted by equation (5-4) and the subscript *l* represents the index for the training iteration.

Since the model is under supervised learning, the data are normally split into the training set, the validation set and the testing set, such that the model is validated with data unseen during the training process. The training process can then be terminated with multiple criteria, including the *SSE*, the number of training iterations and the gradient change of the error matrix. Once the acceptable ANN model is obtained, it is used to forecast the output with testing data.

In addition, in this study, the nonlinear autoregressive exogenous neural network model (NARX ANN) is used. Different from other ANN models, the NARX ANN model is a recurrent dynamic model dedicated to time-series data. Both of the current and past values of the data are considered during the model training as:

$$y_t = \mathrm{f}\left(y_{t-1}, y_{t-2}, \ldots, y_{t-n}, x_{t,} x_{t-1}, \ldots, x_{t-n}\right) \tag{5-6}$$

where the subscript *t-n* is the time delay of the input signal.

5.2.1 Evaluation process description

Due to the powerful learning characteristic of ANN, it is possible to establish a nonlinear relationship between some inputs and outputs based on historical data and produce an accurate output prediction. This makes the model suitable for wind turbine fault diagnosis applications, where the prediction model is first trained using data from a fault-free condition, after which the corresponding outputs can be estimated during a faulty condition. Assuming an accurate

prediction model is achieved, by comparing the predicted output to the actual output, the anomalies can then be identified based on the residual. The larger the residual during the fault period, the easier it is for the model to identify the fault. Moreover, the choice of input variables during model building has a direct impact on the fault detection performance, such that effective detection can be achieved for inputs which are independent of the fault. For example, Cross *et al*. in [47] developed an ANN model using wind speed and active power to detect a generator fault, where the two input variables contain information of the wind speed variation and the demand from the grid network.

On the contrary, if the input variables are closely related to the fault, the fault signatures are certainly included in the model training process. This implies that the predicted output has increased prediction accuracy even during a faulty period, but this can reduce the residual during the fault period as well as the possibility of identifying anomalies. Based on this characteristic, the ANN prediction model is used to verify if the retained variables using the T selection algorithm contain critical information relating to the targeting fault signal. The model prediction accuracy measure is used to evaluate different input datasets.

Figure 5.18 shows the block diagram of procedures for the evaluation process. It is divided into two steps: the model training and validation and the prediction and evaluation. During the first step, for any dataset under a healthy condition, it is split into input and output, where the output is the variable that relates to the signal of interest which may reveal faults or failures of the system. It is also the target variable in the targeted selection algorithm in the previous chapter. As for the inputs, three cases are considered using different variable sets, including the original dataset (without any reduction), the first q number of principal components with its cumulative variance greater than 0.95 [123] and the retained variable set from the targeted selection algorithm obtained in Chapter 4.4.2. The model established for the original dataset is used as a benchmark, where comparisons are then carried out to determine the performance of the proposed algorithm.

At the second step, using the ANN models built in step 1, the outputs from each model using turbine data with known faults are estimated. The performance of

the ANN mode is evaluated with two model fit measures, the squared correlation coefficient R^2 and the relative root mean squared error (RMSE):

$$R^2 = 1 - \frac{SSE}{TSS} = 1 - \frac{\sum(Y - \hat{Y})^2}{\sum(Y - \bar{Y})^2} \tag{5-7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{5-8}$$

where *SSE* stands for the sum of squared errors, *TSS* is the total sum of squares, Y, $\hat{Y}$ and $\bar{Y}$ are the actual output, predicted output and actual output mean, respectively, and n is the number of samples in the dataset.

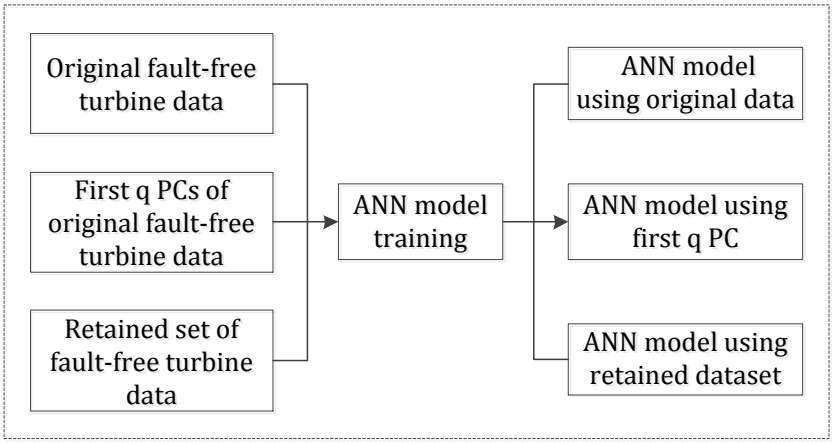


Figure 5.18 - Block diagram for the ANN prediction model-based evaluation procedure.

By comparing the performance measures, it can be determined as to whether the vital features associated with the fault are present in the retained variable set using the proposed selection algorithm.

5.2.2 Model parameter selection and optimisation

The accuracy of the ANN model is greatly dependent on its architecture and parameters involved. Therefore, the optimisation of the selection of the model parameters for the data used in this study is carried out in this section. The parameters of interest include the number of hidden layers, the number of neurons and the initial weightings, which will be investigated independently.

5.2.2.1 Choosing the number of hidden layers

The hidden layer has a significant effect on the learning power of the ANN model. Generally, if the relationship between the inputs and outputs is linearly related, no hidden layer is needed. However, if a nonlinear relationship is present, at least one hidden layer is needed. A higher number of hidden layers may improve the model accuracy depending on the data, but it will certainly increase the training time dramatically. Moreover, there exists the possibility of overtraining when a higher number of hidden layers is adopted, where the model is actually 'memorising' the data rather than learning the inter-relationship. This behaviour will cause poor performance when the new data is used for estimation. Researchers have shown that one number of hidden layers is sufficient in determining the nonlinear relationship of the data accurately for most applications [143, 144, 147, and 148. Hence, in our study, the ANN model will consist of three layers: an input layer, an output layer and a hidden layer.

5.2.2.2 Choosing the number of neurons

The numbers of neurons for the input and output layers are straightforward, as they are dependent on the model configuration, which means the number of neurons for the input and output layers should be exactly the same as the number of input and output variables.

The selection of the number of hidden layer neurons can be more complex. Theoretically, a model with too few numbers of hidden neurons will have a poor learning power and experience under-fitting, such that the model is insufficient

to interpret the inputs and outputs relationship. In contrast, a sizable number of neurons in the hidden layer may cause over-fitting, where the model may produce high accuracy for the training dataset, but the ability to obtain a generalised prediction using testing data is compromised. Moreover, the training time will increase dramatically with the number of neurons. A common rule of thumb is that the optimal number of neurons should be between the number of input and output neurons [149].

In our study, the cross-validation method is employed, where the ANN model is trained with an increasing number of neurons in the hidden layer and the remaining parameters are unchanged. A square correlation coefficient is then used to evaluate the model performance. Figure 5.19 shows an example of the change of the square correlation coefficient with the number of neurons using the healthy SCADA data, as described in Chapter 4.4.1.1. The number of neurons increases from 2 to 20 with a step of 2 neurons. Results show that there is an insignificant change of prediction accuracy for models with different numbers of neurons. Also, when taking the execution time into account, 10 number of neurons are selected in the hidden layer for later evaluations.

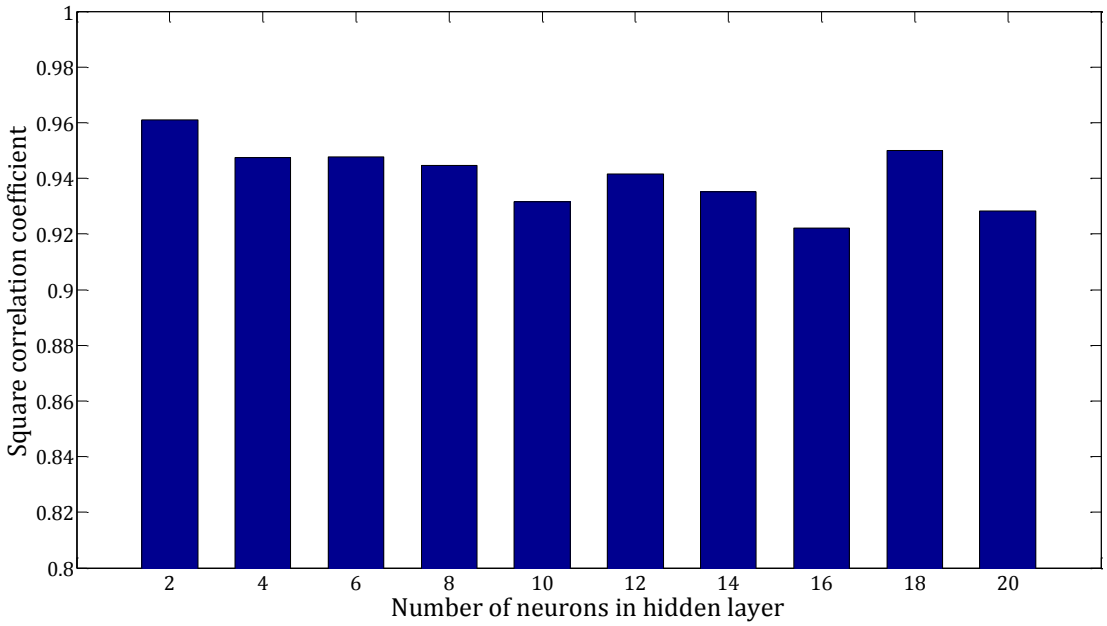


Figure 5.19 - Square correlation coefficient for the ANN model with a varying number of neurons in the hidden layer

5.2.2.3 Effect of number of feedback delays on the model

The feedback delays allow the model to have the ability to include the cross-correlation behaviour between inputs and outputs during training, where cross-correlation represents the relationship of two signals with respect to time delays. Hence, the resulting ANN model will have better accuracy. However, this is only true when the optimal numbers of delay are provided. Again, the cross-validation technique is used to examine the best number of feedback delays. Figure 5.20 shows the R^2 of several models created with a varying number of feedback delays using SCADA data. It can be seen that with an increasing number of delays, all R^2 have similar values, and they are approximately 0.93. Therefore, in a later study, a feedback delay of two is selected.

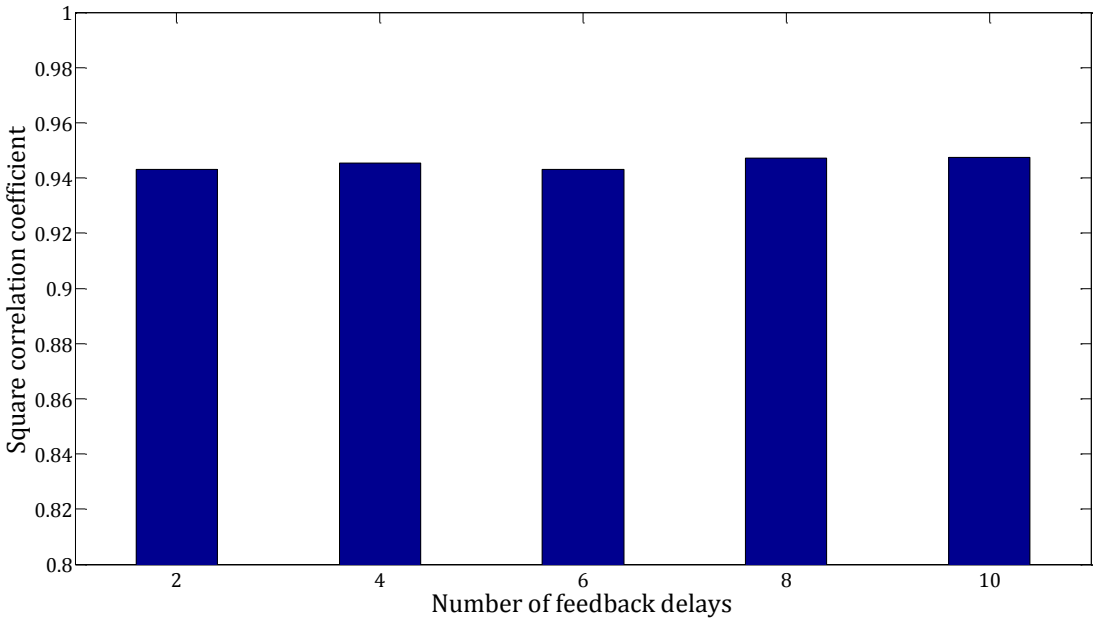


Figure 5.20 - Square correlation coefficient for the ANN model with a varying number of feedback delays

5.2.2.4 Effect of number of input delays on the model

During model training, different from feedback, the input delays allow the model to have the capability to incorporate the relationship of the input signals with itself at different time stamps, known as auto-correlation behaviour, thus achieving a more accurate estimation during testing. Figure 5.21 shows the R^2 using a cross-validation technique of different numbers of input delays. It can be seen that there is no significant change of the R^2 as the delays increases from 2 to

8 at a step of 2. The R^2 with 10 number of delays have the least value of 0.89, which may be due to over-fitting. Based on the result, a delay of 2 is chosen.

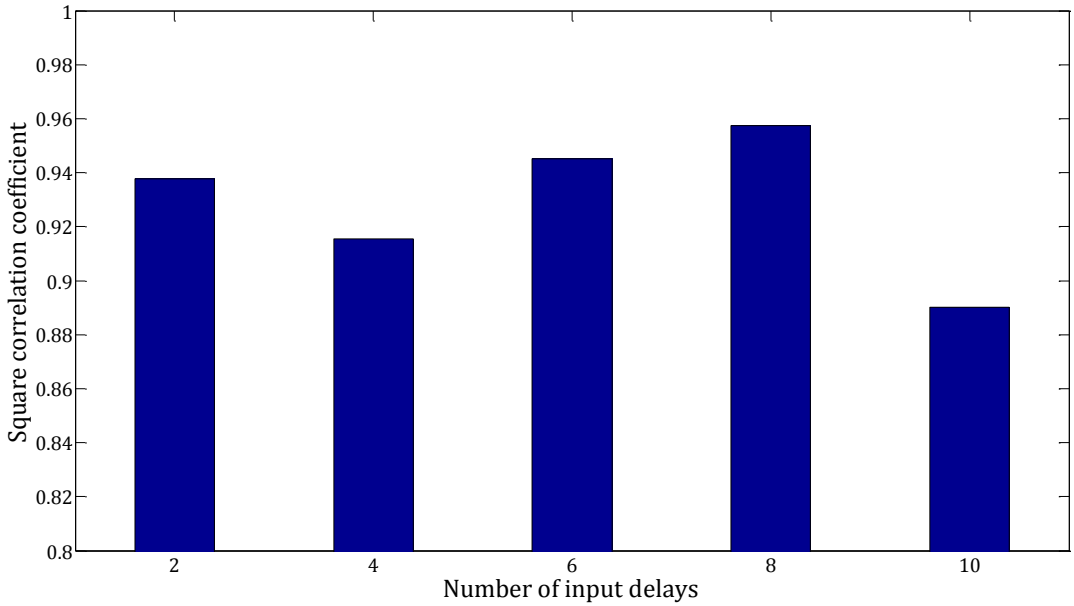


Figure 5.21 - Square correlation coefficient for the ANN model with a varying number of input delays

5.2.2.5 Selection of the learning algorithm

The learning algorithm is responsible for the neural network to study the nonlinear relationship of datasets. As described earlier, in function mapping, the Levenberg-Marquardt algorithm has the best performance. To validate this statement, nine learning algorithms, which are presented in Table 5.1, are used to train the model, and the square correlation coefficient for each model is found with testing data. More details concerning these learning algorithms can be found in [146]. Although some of these algorithms are not designed for function mapping, for comparison purposes, they are also included. R^2 as the result of each learning algorithm is shown in Figure 5.22. It can be seen that the Levenberg-Marquardt algorithm has the highest R^2 of 0.95. Thus, it will be used throughout later studies.

5.2.2.6 Choosing the initial weightings

The initialisation of weights is necessary for the training process to begin. Generally, random numbers are assigned for the weight matrix. However, this may cause the model to converge to a false local minimal of the error function during training, which results in poor prediction accuracy. Different techniques

in determining the optimal initial weighting for the ANN model can be found in literatures [150, 151]. In our study, the ensemble averaging method is adopted, where multiple ANN models have been created with random initial weight, and the actual output is obtained by taking the mean from all the models. The resulting model often has a superior performance compared to an individual model, as the errors from each model are minimised by taking the average. There are two main disadvantages of this method, which are the considerably increased training time and the need for several models during the testing stage. In a later evaluation process, 30 ANN models are trained for each dataset, and the ensemble result is then used.

Abbreviation	Training algorithm
trainlm	Levenberg-Marquardt
trainbfg	BFGS Quasi-Newton
trainrp	Resilient Backpropagation
trainscg	Scaled Conjugate Gradient
traincgb	Conjugate Gradient with Powell/Beale Restarts
traincgf	Fletcher-Powell Conjugate Gradient
traincgp	Polak-Ribiére Conjugate Gradient
trainoss	One Step Secant
traingdx	Variable Learning Rate Backpropagation

Table 5.1 - Types of learning algorithms.

5.2.2.7 Choosing the learning rates

As can be seen from equation (5-5), the learning rate is directly related to the updated weight matrix. This implies the value will affect the step size in which the weight matrix is modified in the multidimensional weight space. A large learning rate may cause the algorithm to overstep the local minimum and can be difficult to converge. On the other hand, too small of a learning rate will increase the learning time dramatically. An adaptive learning rate is therefore used when building the function in the Matlab Neural Network Toolbox with the Levenberg-Marquardt training algorithm. Instead of changing the learning rate, the algorithm directly controls μ in equation (5-5) to achieve optimal training. An

initial value is assigned for μ, and, based on the gradient change of the error function, the value is adjusted by a scaling factor.

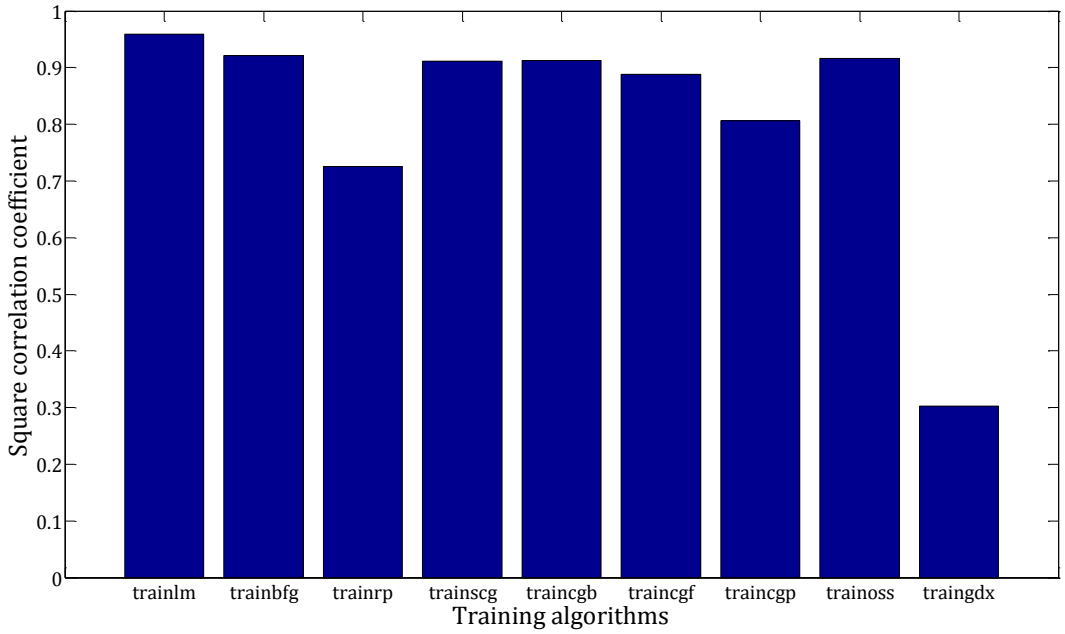


Figure 5.22 - Square correlation coefficient for the ANN model with different training algorithms

5.2.3 Results and discussion

With the optimised parameters obtained for the ANN models, an evaluation of data with various fault and retained variables obtained from Chapter 4.4.2 is carried out based on the procedures as described in Section 5.2.1. The results from each case are presented and discussed.

5.2.3.1 Simulation data

Results for the evaluation using PMSG data with a -1% capacitor ageing fault are shown in Figure 5.23. The top three plots show the actual DC-link voltage in the blue lines, and the red dotted lines are the predicted value using inputs from all datasets, the PCA reduced dataset and the dataset obtained from the targeted selection algorithm as described in Chapter 4.4.2. The bottom three plots in Figure 5.23 show the residual for the three cases, respectively. The capacitor ageing fault occurred at 5 s (sample point 20000). It can be seen from the plots that all three methods have a strong prediction capability. However, there are always residuals during the fault period even when all datasets are used, but these residuals are very small at the order of 1×10^{-03}.

The RMSE and R^2 for each prediction is calculated and listed in Table 5.2. Results show that there is an increasing trend for the RMSE and R^2 from all datasets, as well as from the PCA reduced dataset and targeted dataset, where the targeted dataset has the highest performance measure [3×10^{-04}, 99.991%] and the all dataset has the lowest [4×10^{-04}, 99.983%]. This outcome stems from the correlations among the variables in the dataset. In the case of all datasets, there is a high degree of correlation and causing of over-fitting during the prediction. For the PCA reduced dataset, variables are not correlated; however, these variables are chosen simply based on the variance maximisation principal. Hence, some information may be lost, thus reducing the prediction accuracy. Finally, the datasets obtained from targeted selection algorithms are aimed to have minimal correlation and maximising information relating to the fault. Therefore, it has the highest performance measure. The high prediction performance indicates that a fault signature is present in the input dataset from all three cases.

To further validate the variable selection, the evaluation procedures are applied to the PMSG data with a single phase-to-ground fault, a phase-to-phase fault and a three phase-to-ground fault. Table 5.2 also shows the RMSE and R^2 for these fault types. The same trend has been obtained, where the model with all of the data has the lowest accuracy and the targeted selection model has the highest.

		All data	PCA reduction	Targeted selection
PMSG - capacitor ageing fault	RMSE	0.00004	0.00004	0.00003
	R^2	0.99983	0.99984	0.99991
PMSG - 1 phase-to-ground fault	RMSE	0.00300	0.00241	0.00111
	R^2	0.99909	0.99987	0.99994
PMSG - phase-to-phase fault	RMSE	0.00389	0.00347	0.00283
	R^2	0.99887	0.99988	0.99882
PMSG - 3 phase-to-ground fault	RMSE	0.00505	0.00309	0.00407
	R^2	0.99816	0.99982	0.99951

Table 5.2 - RMSE and R^2 for PMSG simulation data with different types of faults using all datasets, a PCA reduced dataset and a targeted selection dataset

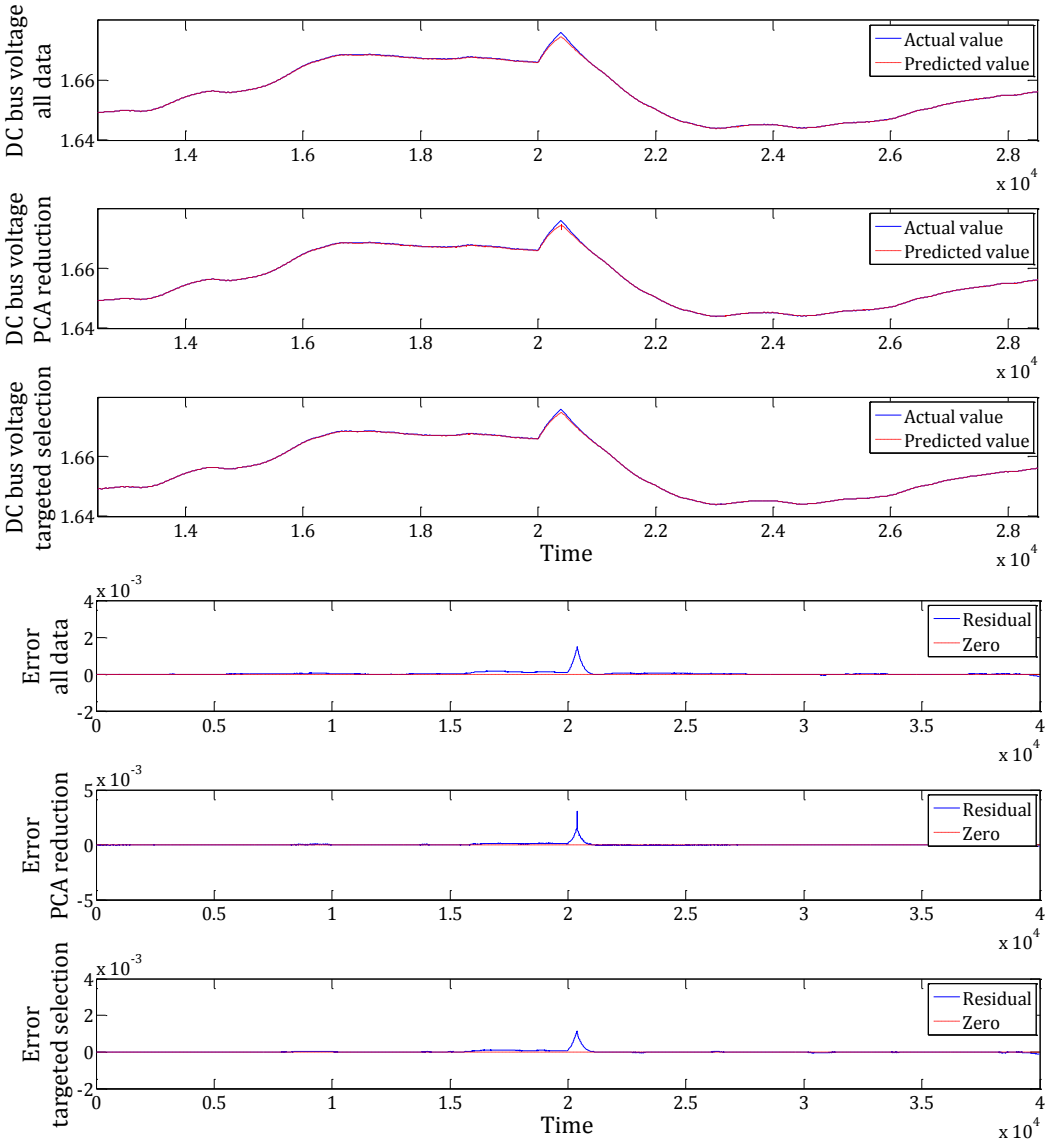


Figure 5.23 - Actual DC-link voltage from the PMSG simulation data and ANN predictions using different input datasets. Top three plots: model predictions using all dataset, PCA reduced dataset and targeted selection dataset. Bottom three plots: residual plots of each input dataset

In the case of DFIG simulation data, ANN models are established for each type of fault, and the predicted outputs are obtained. The performance measures are shown in Table 5.3. Again, very high prediction accuracies are achieved for all three input datasets, which proves that a fault signature is present in the dataset obtained with the targeted selection algorithm. However, there is a different trend of the performance measure from the capacitor ageing fault data, where the RMSE and R^2 using all data and the targeted selected datasets are similar [3×10^{-04}, 99.984% and 4×10^{-04}, 99.985%, respectively], and the PCA reduced

dataset has the lowest performance [1.5×10^{-03}, 99.958%]. This may be because the effect of the loss of capacitance (at -35%) for a DFIG system is very small, which means the data variability due to the capacitor ageing fault with respect to the whole dataset is minimal. Therefore, certain fault signatures may be removed during PCA reduction, resulting in a lower prediction performance. Figure 5.24 shows the model prediction (top three plots) and residual (bottom three plots) for each method. As the composition of simulation data is relatively simple, the fault signature amount being removed is minimal; hence, there is almost no visual difference amongst the three cases in the residual plot. However, using the two performance measures, it can be seen that results using data obtained from targeted selection algorithms outperform the others. This implies that fault features are present within the retained variables sets.

		All data	PCA reduction	Targeted selection
DFIG - capacitor ageing fault	RMSE	0.00003	0.00015	0.00004
	R^2	0.99984	0.99959	0.99985
DFIG - 1 phase-to-ground fault	RMSE	0.00477	0.00427	0.00366
	R^2	0.99659	0.99830	0.99737
DFIG - phase-to-phase fault	RMSE	0.00496	0.00487	0.00321
	R^2	0.99793	0.99859	0.99877
DFIG - 3 phase-to-ground fault	RMSE	0.00255	0.00363	0.00157
	R^2	0.99951	0.99900	0.99957

Table 5.3 - RMSE and R^2 for DFIG simulation data with different types of faults using all datasets, a PCA reduced dataset and a targeted selection dataset

5.2.3.2 SCADA data

As SCADA data is obtained from an operating wind farm and has a more complex data composition, it is used to further examine if the fault signature is retained in the dataset obtained from the targeted selection algorithm. As already described in the previous chapter, two types of faults have been identified from the different turbines: a gearbox related fault and a generator winding related fault.

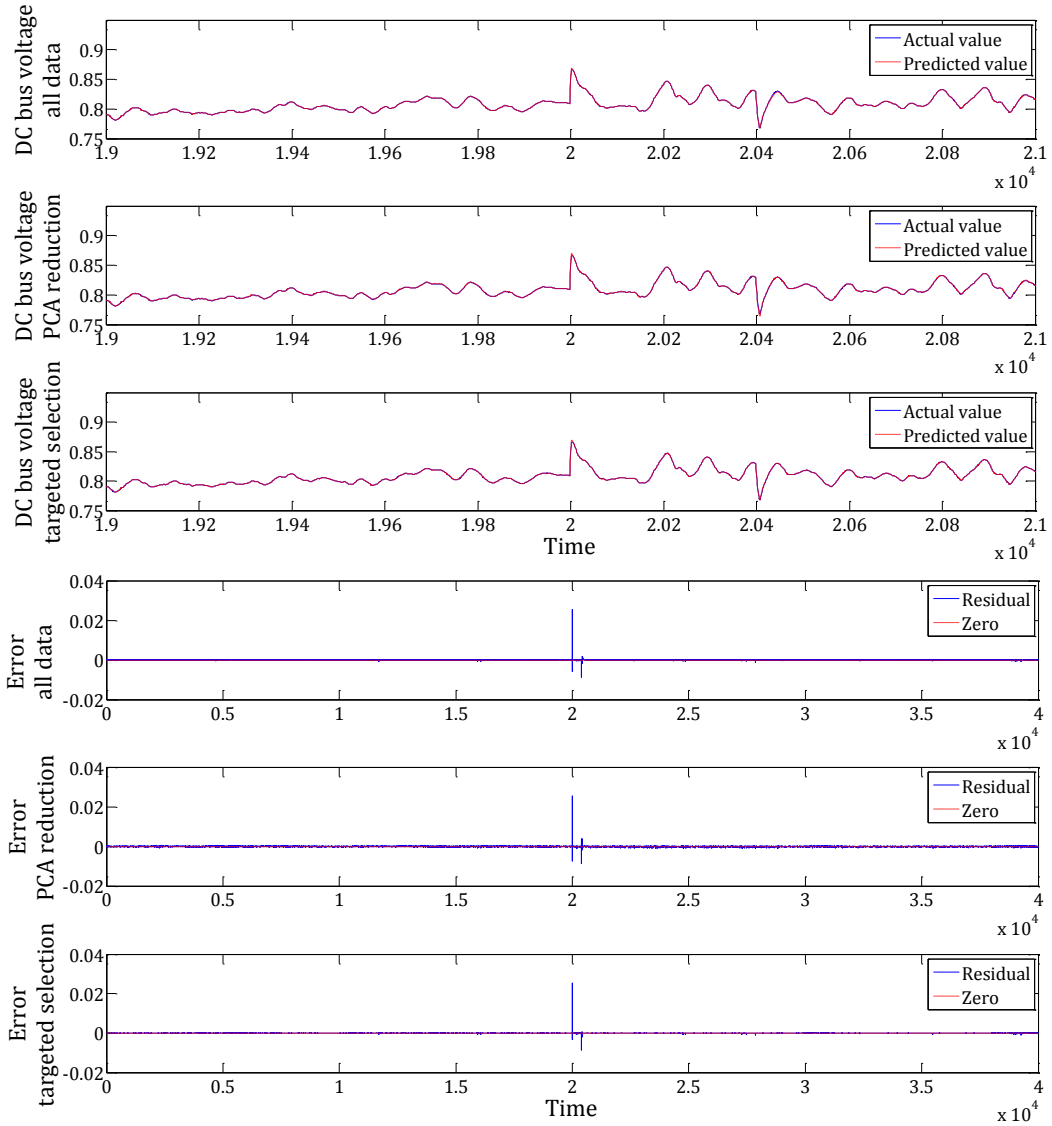


Figure 5.24 - Actual DC-link voltage from the DFIG simulation data and ANN predictions using different input datasets. Top three plots: model predictions using all dataset, a PCA reduced dataset and a targeted selection dataset. Bottom three plots: residual plots of each input dataset

Figure 5.25 shows the model prediction and residual plots for the gearbox bearing temperature using different input datasets, and the performance measures are given in Table 5.4. The fault occurred approximately at sample point 720, where the gearbox bearing temperature started to increase to an abnormal level. It can be seen that the model prediction with all of the data is very close to the actual value with RMSE and R^2 of 0.276 and 99.5%, respectively. Similarly, in the case of the targeted selection data obtained in Chapter 4, a high model prediction is also obtained [0.397, 99.2%]. However, the residual plot still

shows a minor deviation during the fault period. Finally, the PCA reduced data show the worst performance, where an obvious difference between the actual and predicted gearbox bearing temperature is found. Because the SCADA data composition is more complex, when 95% of the data variability is maintained, the fault signature causing the increase of the gearbox bearing temperature may be excluded in the PCA reduced dataset. On the other hand, the targeted selection algorithm is able to capture this signature during variable selection, thus resulting in better prediction accuracy.

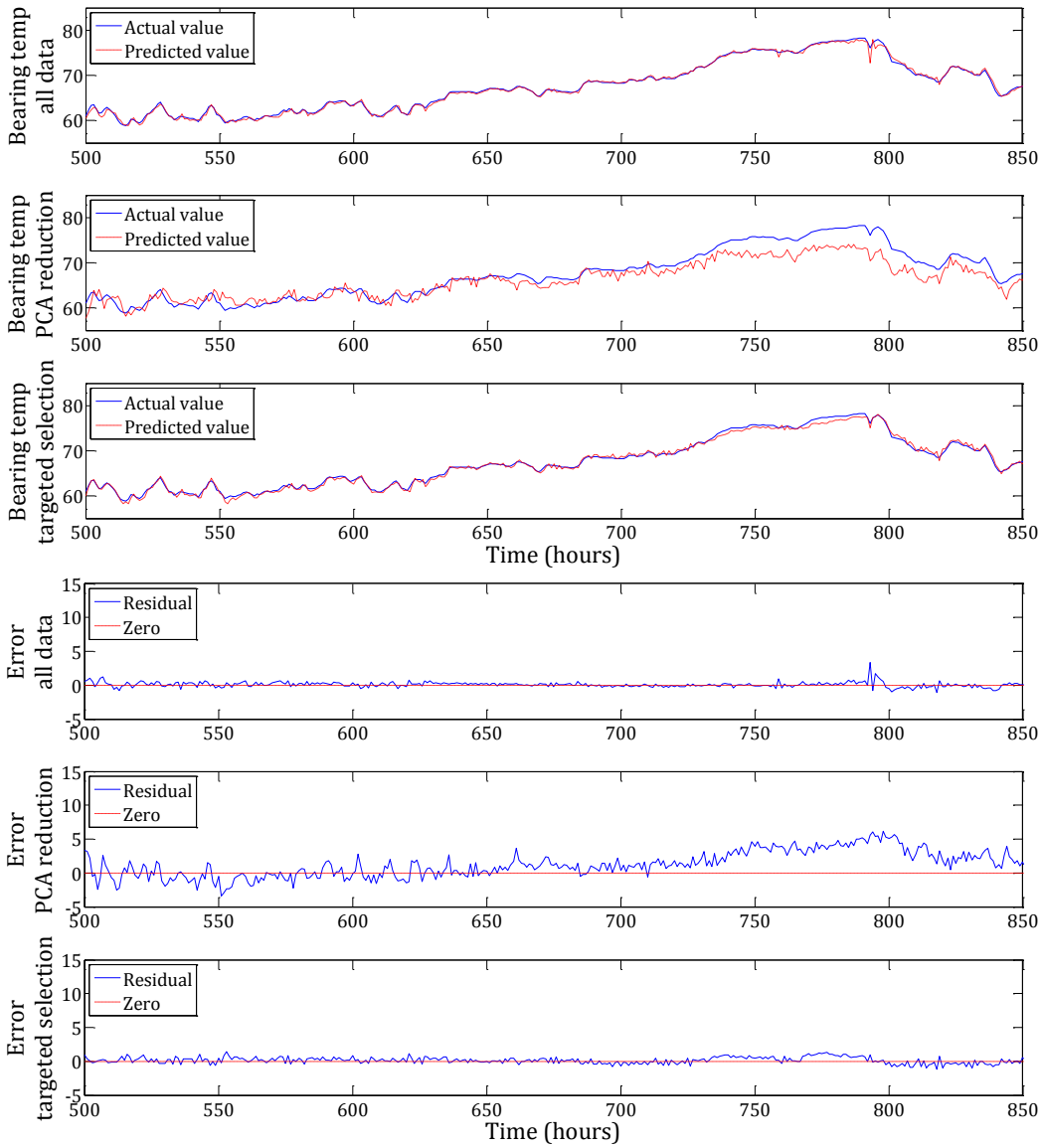


Figure 5.25 - Actual gearbox bearing temperature from the SCADA data and ANN predictions using different input datasets. Top three plots: model predictions using all dataset, a PCA reduced dataset and a targeted selection dataset. Bottom three plots: residual plots of each input dataset.

		All data	PCA reduction	Targeted selection
SCADA - gearbox fault	RMSE	0.276	1.793	0.397
	R^2	0.995	0.823	0.992
SCADA - generator fault	RMSE	4.513	8.466	2.556
	R^2	0.914	0.629	0.972

Table 5.4 - RMSE and R^2 for SCADA data with different types of fault using all dataset, PCA reduced dataset and targeted selection dataset.

Finally, SCADA data with a generator fault, as described in Chapter 4.4.2, is evaluated. The generator winding temperature is used as a targeting signal, and the model prediction and residual of the generator winding temperature are shown in Figure 5.26. It can be seen that the fault starts at sample point 1000, and the performance measures of the predictions are given in Table 5.4 Results show that the targeted selection dataset has the best performance of RMSE and R^2 of 2.556 and 97.2%, respectively, followed by the model using all data [4.513, 91.4%]. The PCA reduced data has the worst model prediction of [8.466, 62.9%]. It is worth mentioning, as with all other cases, that the predictions during the fault-free portion are very similar. The difference between the actual and the predicted value becomes obvious as the fault starts. The aim of the ANN model is to prove if a fault signature is present in the input dataset rather than developing accurate prediction model for fault detection. Therefore, based on the results, it can be concluded that the dataset retained by the targeted selection algorithm is able to capture a fault signature relating to a particular fault.

5.3 Summary and discussion

The evaluation of the presence of vital fault information relating to a particular fault in the retained set of variables obtained from targeted selection algorithm has been carried out in this chapter. Two types of evaluation methods have been proposed: the feature-based method and the ANN prediction model method. The former method is based on the correlation test of fault signatures identified from the original dataset. The latter method predicts the fault signal of interest using different input datasets, such that the prediction accuracy is used as a measure to

evaluate if a fault feature is preserved in the retained dataset. Both simulation and SCADA data with different types of faults are evaluated. Results from both methods have shown that the retained variable set obtained from the targeted selection algorithm is able to maintain sufficient information to assess the system's faults. Moreover, with the evaluation using different datasets, it also demonstrates that the targeted selection algorithm is robust. In the next chapter, detection methods are proposed, allowing for the identification of a fault and severities estimation using datasets selected with the T algorithm in Chapter 4.

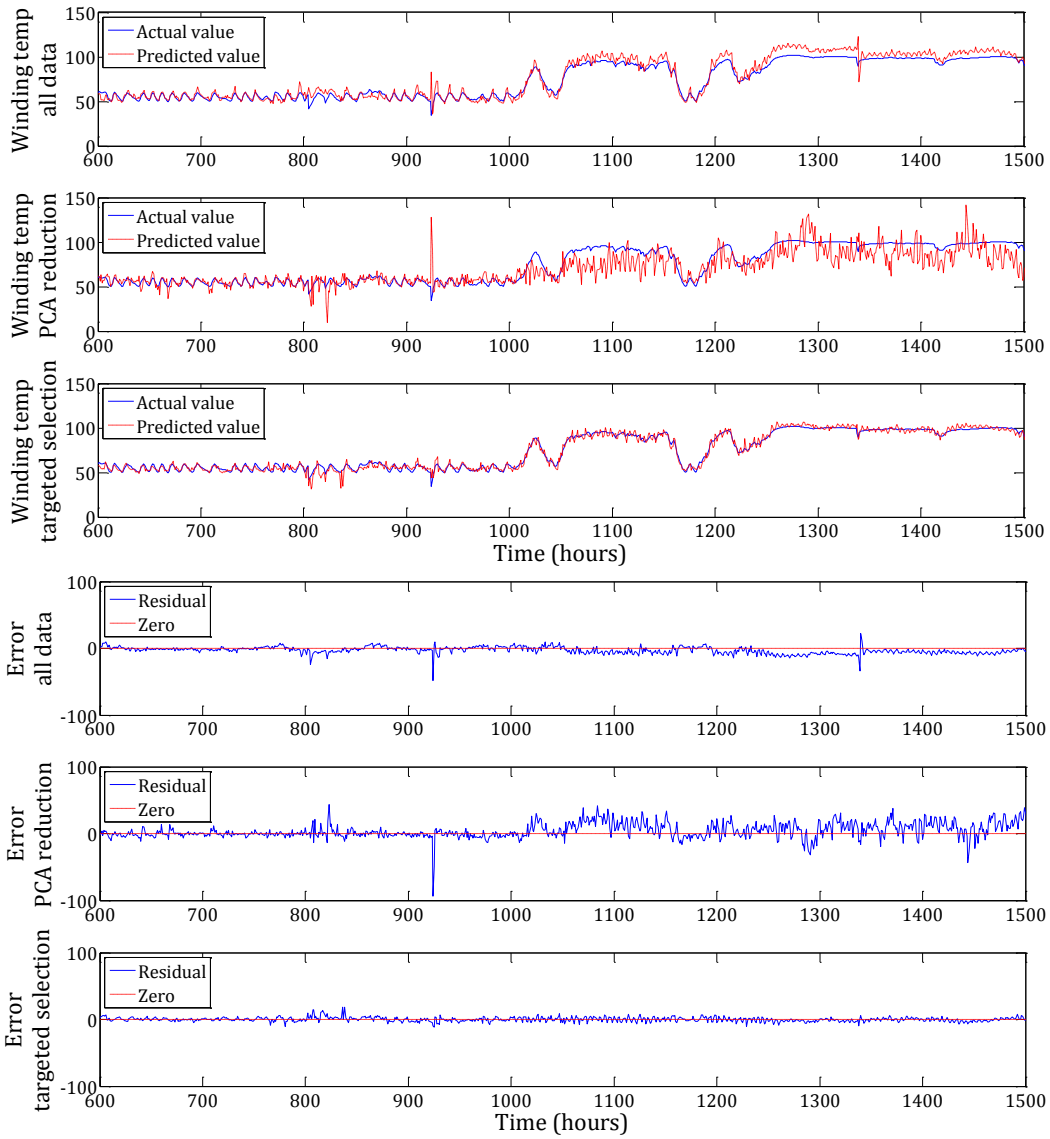


Figure 5.26 - Actual generator winding temperature from the SCADA data and ANN predictions using different input datasets. Top three plots: model predictions using all dataset, a PCA reduced dataset and a targeted selection dataset. Bottom three plots: residual plots of each input dataset.

Chapter 6. Feature Extraction-based Fault Detection and Identification

The previous chapters have shown that the proposed targeted selection algorithm is able to retain a set of variables that minimises redundancy and maintains vital information relating to a particular fault signal. In line with this result, in this chapter, two fault detection methodologies are proposed using the retained variable set from the selection algorithm. The former method detects faults based on the Hotelling's T2 statistic, and the method is able to tackle the identification problem using the instantaneous energy of PCs.The latter feature-based method permits the determination of a fault and its severity. Finally, different datasets are used to validate the detection algorithms.

6.1 Instantaneous PC energy density method

The application of a principal component analysis is often found in multivariate statistical process monitoring. Due to the uncorrelation property of the transformed PCs, it has shown a superior performance compared to the traditional process monitoring control chart, which is generally used in quality control [123, 152, 153]. The basic theory of the PCA-based process monitoring is presented first.

As stated earlier, for any given dataset $\mathbf{X}$ ($p \times n$) with p number of variables and n number of samples, the transformed principal component scores $\mathbf{Z}$, eigenvalues $\mathbf{L}$ and eigenvectors $\mathbf{U}$ can be found by using SVD of its covariance matrix $\mathbf{S}$ (Eq. 4-1 to 4-3). The relationship between the original variables and the PC scores in matrix form is:

$$\mathbf{Z} = \mathbf{U}^{\mathbf{T}}[\mathbf{X} - \bar{\mathbf{X}}] \tag{6-1}$$

where, $\bar{\mathbf{X}}$ is the matrix of the mean of the original variables. Then, from the p number of PCs, q number will be selected (q<p) such that it is adequate to describe dominant variability of the dataset. Using this q number of PCs, it is possible to estimate a new set of original variables $\hat{\mathbf{X}}$,

$$\hat{\mathbf{X}} = \bar{\mathbf{X}} + \mathbf{U}_q\mathbf{Z}_q \tag{6-2}$$

where $\mathbf{U_q}$ and $\mathbf{Z_q}$ are the corresponding eigenvector and PC scores containing the first q columns of $\mathbf{U}$ and $\mathbf{Z}$. The PCA model then becomes:

$$\mathbf{X} = \hat{\mathbf{X}} + \mathbf{E} \tag{6-3}$$

such that $\mathbf{E}$ is the residual signals and signifies the amount of information unexplained by the PCA model. Usually, the residual represents the noises present in the dataset. However, sometimes fault signature may exist in the residual [141].

In the perspective of fault detection using PCA-based process monitoring, two statistical measures, the Hotelling's T^2 statistics and the Q statistics, are commonly used. For the Hotelling's T^2 statistics, it is found by:

$$\mathbf{T^2} = [\mathbf{X} - \bar{\mathbf{X}}]^T\mathbf{U}_q\mathbf{L}_q^{-1}\mathbf{U}_q^T[\mathbf{X} - \bar{\mathbf{X}}] \tag{6-4}$$

$$= [\mathbf{X} - \bar{\mathbf{X}}]^T \mathbf{W}_q \mathbf{W}_q^T [\mathbf{X} - \bar{\mathbf{X}}]$$

$$= \mathbf{Y}_q^T \mathbf{Y}_q$$

where $\mathbf{W_q}$ and $\mathbf{Y_q}$ are the scaled eigenvector and PC scores of first q columns. The residual **Q** statistics can be calculated with:

$$\mathbf{Q} = \mathbf{E}^T \mathbf{E} \tag{6-5}$$

The two statistics are constantly monitored, and if any of these two statistics is above a threshold value, it may indicate the occurrence of an abnormality. Otherwise, the process is considered as normal. The threshold values for the T^2 is given [123]:

$$T_{q,n,\alpha}^2 = \frac{q(n-1)}{n-q} F_{q,n-q,\alpha} \tag{6-6}$$

where q is the number of selected PCs in the PCA model, and $F_{q,n-q,\alpha}$ is the critical point of the F distribution (which is a continuous probability distribution notably for variance analysis; more information can be found in [123]) with q and n-q degrees of freedoms at the α confidence interval. The threshold value for the **Q** statistics at the α confidence level is obtained from approximated distribution by [123]:

$$Q_\alpha = \theta_1 \left[\frac{c_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} + 1 \right]^{1/h_0} \tag{6-7}$$

with

$$\theta_i = \sum_{j=q+1}^{p} l_j^i \tag{6-8}$$

and

$$h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2} \tag{6-9}$$

where c_α is the value from normal distribution at the α confidence level. It should be noted that the **Q** statistics are not considered in our study because it represents the error term, where it mainly contains various types of noise from the original dataset. It has already proven that the variable set retained by the

targeted selection method contains dominant variability of the dataset. Therefore, only the T^2 statistics is used for fault detection and identification.

6.1.1 Anomaly detection and identification

Fault diagnoses utilising the two statistics are widely found in literatures. The authors from [154, 155] have shown effective fault detection and diagnosis of sensors in wastewater treatment. However, the application in wind turbine fault detection is limited. The employment of the two statistics in conjunction with ANN-based clustering and classification in determining turbine failures using SCADA data are studied in [156, 157]. The work focused on the anomaly detection, though, and the identification problem was not addressed. Moreover, the authors in [158] have proposed a fault identification technique relying on the relative contribution index d_i of the T^2 statistic from the original measurements using:

$$d_i = T^2 - T_i^2 \tag{6-10}$$

where T_i^2 is the T^2 statistic excluding the i^{th} variable from the original dataset. The proposed method is tested against SCADA data, where it contains 120 variables. Instead of using all the variables, the author selected a subset of variables based on experience and the investigation of the alarm log, including wind speed, pitch angle, gear bearing temperature, gear oil temperature, active power output, generator bearing temperature, generator speed and rotor speed. This step requires an excessive amount of time in analysing the SCADA data, and the variables used may not be optimum due to being selected based on experience.

In our study, there are two improvements of the method in the anomaly detection and identification based on the T^2 statistics approach. First, instead of selecting a set of variables based on experience as in [158], the dataset used in the detection and identification is selected using the targeted selection algorithm proposed in Chapter 4. As already proven in previous sections, the targeted selection method is able to retain a set of variables which maximises variability of the dataset, and they are highly related to the targeted signal of interest. Thus, by implementing this method, it has the benefits of reducing the excessive amount of time spent analysing the alarm log and data. Additionally, because the

selected variable set is fault dependent, it has the potential to increases accuracy in fault detection and identification. In addition, the method would require minimal knowledge of the applied system because of the incorporation of the variable selection algorithm.

Secondly, a PC energy-based method is proposed for decomposing the T^2 statistic in order to carry out fault identification. As can be seen from equation (6-6), the T^2 statistic is a multivariate measure, and it is the linear combination of the scaled PC scores **Y**. Therefore, the proposed identification process is divided into two stages: the first stage is to identify the PC that has the highest contribution to the T^2 based on the sum of instantaneous energy of the PCs. When monitoring the T^2 statistic, any T^2 statistic above the threshold value is considered abnormal, and the relative contribution of each PC to the T^2 statistic during this period can be found by:

$$\mathrm{TC}_i = (|z_i|)^2 \tag{6-11}$$

where z_i is the i^{th} unscaled PC. Once the PC with the highest contribution is found, the variables contributing to the fault can be identified using the corresponding eigenvector loadings.

The complete process of anomaly detection and identification is shown in Figure 6.1. The process is divided into training and testing stages. During the training stage, the original dataset under a normal operation condition is normalised to zero mean and unit variance, and the variables used are selected via the application of the targeted selection algorithm for the particular fault (i.e. the DC-link capacitor ageing fault). The PCA is then carried out to obtain the corresponding eigenvalues, eigenvectors and PCs. Using the parameters, the T^2 statistic and threshold value T_α^2 are calculated using equation (6-4), and q is set as the total number of variables.

Thereafter, during the testing stage, the process contains the anomaly detection phase and the identification phase. First, the testing data is normalised using the healthy turbine data. Then, a PCA model is created and the T^2 statistics are calculated. If any of the T^2 statistics exceed the threshold value T_α^2 as calculated from the normal operational data, the measurement is considered to be an anomaly. Once the anomalies are identified, the identification procedure based

on the decomposition of the T^2 statistics (Equation 6-11), as described earlier, is implemented in order to determine which variables have the highest contribution for the anomalies.

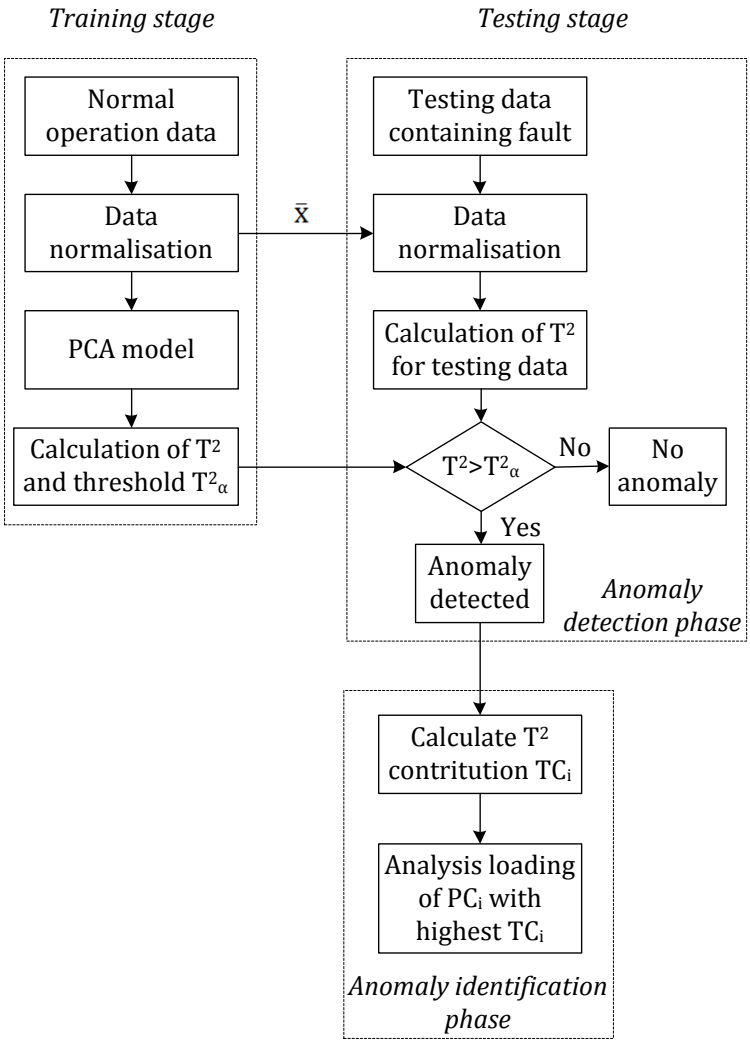


Figure 6.1 - Block diagram of the fault detection and identification process

6.1.2 Evaluations of PC energy density method

Before carrying out an evaluation of the proposed method, two aspects need to be addressed. First, several methods for determining the optimal number of variables to be retained have been discussed in Chapter 4. Results from the cross-validation method and the traditional SCREE plot have shown different optimal number. In previous chapters, because information maximisation is a main objective for the selection algorithm, a larger optimal number of variables is chosen. However, in the context of fault detection, the interest is focused on the fault signature rather than on maximising information of the dataset. Hence, the

optimal number of variables obtained from the cross-validation method is used, as the method examines the number of PCs needed to reconstruct accurate original data. Second, one of the main assumptions of the T^2 statistic is that the original data should have a normal distribution or an approximate normal distribution [152]. Therefore, an additional pre-processing step is carried out for all the normal operating and testing (with particular fault) datasets, where the data is normalised to have an approximate normal distribution using the Box-cox transformation:

$$x_i^{(\lambda)} = \begin{cases} \dfrac{x_i^{\lambda} - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \ln(x) & \text{if } \lambda = 0 \end{cases} \tag{6-12}$$

where x is the original data, and λ is the coefficient optimised through the maximum likelihood function such that the resulting data is approximately normal distributed. An example of the wind speed distributions before and after the processing is shown in Figure 6.2. The top plot show the histogram of the normalised wind speed (blue bars) and the fitted distribution (red line). It can be seen the data is skewed to the left. The histogram and fitted distribution of the wind speed data after transformation is plotted in the bottom plot. It can be seen that the data has a very close shape to the normal distribution (bell shape).

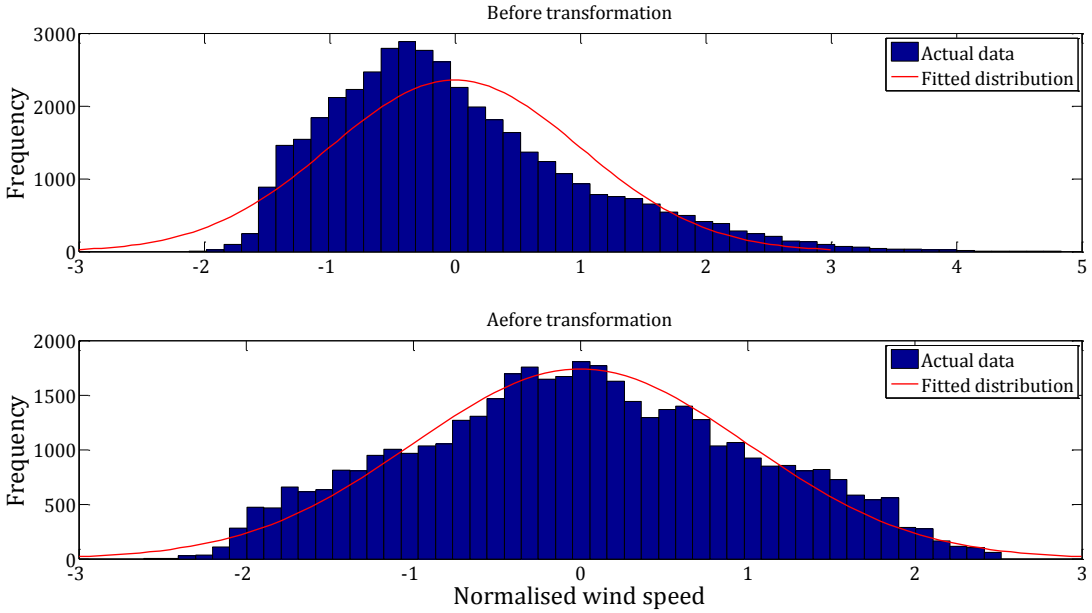


Figure 6.2 - Example of the histogram of wind speed from SCADA data before (top) and after (bottom) the Cox-box transformation

The proposed method is first tested using the PMSG simulation data with a capacitor ageing fault. The eight variables to be retained are considered adequate according to the cross-validation method, and they are the DC bus current 2 (I_{dc2}), the active power at point of coupling (P_g), the reference DC bus current (I_{dc_ref}), the pitch angle, the turbine reactive power, the generator torque, the turbine power and the phase A current of the grid at node 1 (I_{a1}).

Using these variables, the T^2 statistics and the threshold value at a 0.99 confidence level are calculated and shown in Figure 6.3. The top plot shows the T^2 statistic from the normal operation data appears as the blue line, and the dotted red line is the threshold value, as calculated using equation (6-6). The bottom plot is the T^2 statistic calculated using the testing data (simulation data with fault). As can be seen, the T^2 from the normal operation data are well below the threshold value. Additionally, in the case of faulty data, there is an obvious overshoot of T^2 from the faulty data between sample points 20000 and 20400. The data is sampled at 250 μs; thus, the fault is triggered at 5 s for 0.1 s. This indicates that an abnormal behaviour is found during that period. The contribution index TC_i for this period is then calculated using equation (6-11) as shown in the top plot of Figure 6.4. It can be seen that PC 1 has the highest contribution to the T^2 statistic. Therefore, the corresponding loading of PC 1 is shown in the bottom plot of Figure 6.4.

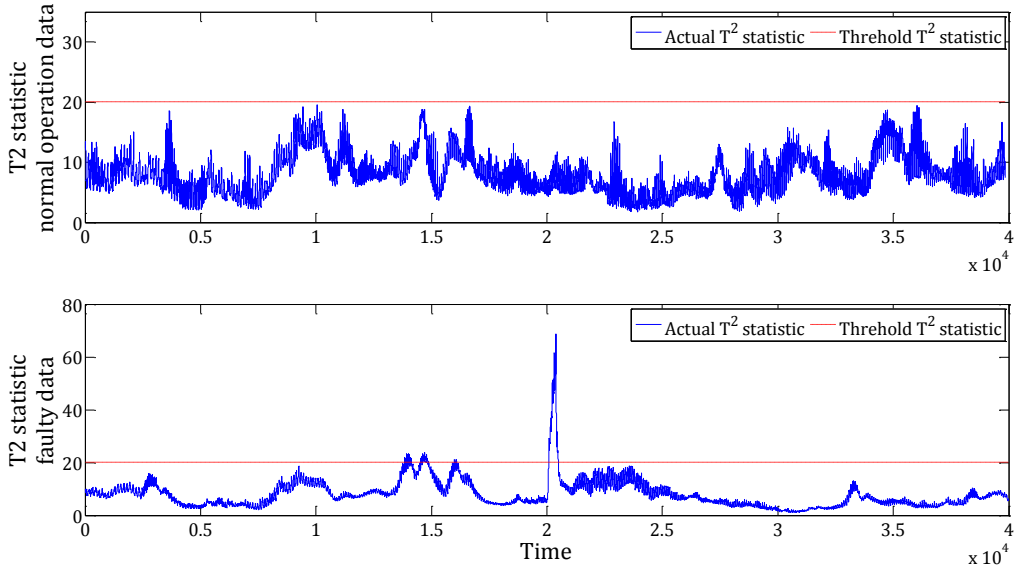


Figure 6.3 - T^2 statistic from the PMSG simulation data. Top plot: normal operation data; Bottom plot: data with a capacitor ageing fault at -1%

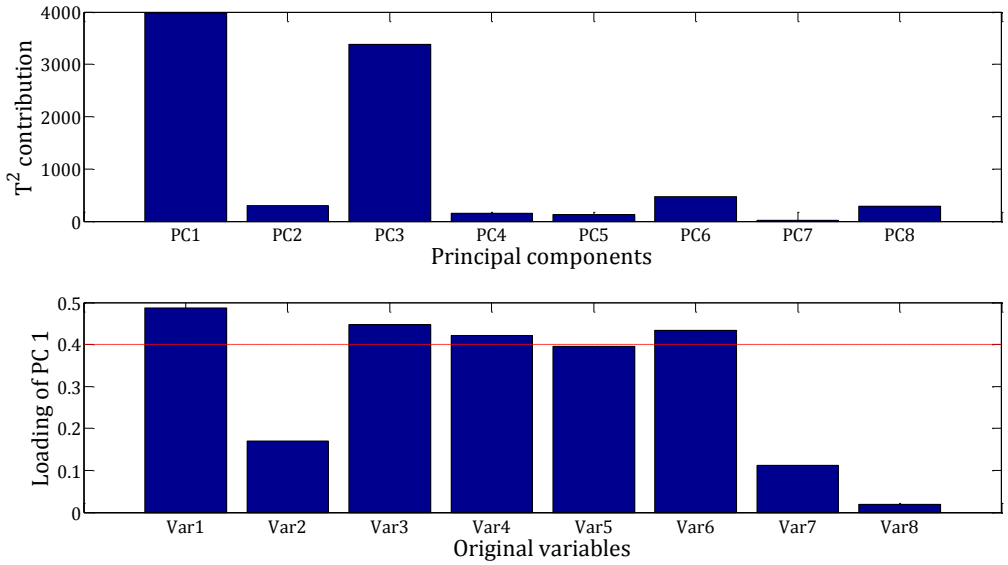


Figure 6.4 - Fault identification of the PMSG simulation data with a capacitor ageing fault. Top plot: T^2 contribution from each PC; Bottom plot: loadings of the PC with the highest contribution

The significance of PC loadings (**U** matrix) is dependent on the dataset, where if high correlation is present in the dataset, the significant loadings should be higher and vice versa. However, generally loadings with absolute value larger than 0.3-0.5 are considered significant [123, 124]. Because the simulation data contains electrical variables with a limited number of mechanical variables such as speed and torque, the composition of the data is assumed to be relatively simple. Therefore, the significant loading threshold of 0.4 is used. The bottom plot of Figure 6.4 shows that variable 1 (I_{dc_2}), 3 (I_{dc_ref}), 4 (pitch angle) and 6 (generator torque) are dominant contributors to PC1 with I_{dc_2} with the highest loading. Because the targeting variable during the selection was the DC bus voltage and because the variables obtained from the identification process with the highest loadings are also related to the DC bus current, it is possible to conclude that the origin of the anomaly is localised to the DC bus of the AC-DC-AC converter. This is expected because during the fault period where the DC-link capacitance is changed, the transient changes of the DC bus voltage and current have the most impact on the variability of the dataset. The proposed method is able to identify this variation and traces back to individual variables in terms of data variability.

The method is further tested using PMSG data with a three phase-to-ground fault with a grounding resistance of 1 Ω. The fault is placed at the grid at node 2, as shown in Figure 3.33. The retained variables obtained from the selection algorithm are the reactive power of the generator, turbine torque, turbine power, phase B grid voltage at node 1 (V_{b1}), grid currents I_{c1}, I_{a3}, I_{c3} and phase B current at the point of coupling before the transformer (I_{gb}), with active power at the point of coupling (P_g) being the targeting variable. A clear abnormality is found between sample points 20000 and 20400 (5 - 5.1 s) from the testing data (Figure 6.5 bottom), which is the period during which the fault occurs. Further investigation of the T^2 contribution plot (Figure 6.5 top) shows that the 5th principal component has the highest TC_i. The bottom plot of Figure 6.6 shows that variables 1 (reactive power), 5 (I_{c1}) and 7 (I_{c3}) are the most significant contributors for PC 5. In this example, with the active power at PCC chosen as the targeting variable, it was found that the grid currents at different nodes are the dominant contributors to the detected abnormality. Therefore, the anomaly is somewhat related to the grid. Moreover, variables with insignificant loadings may be indicative in determining the location of fault originated, as it shows how each variable is related to the corresponding PC in terms of variability. Hence, it may also be helpful for further root cause analysis.

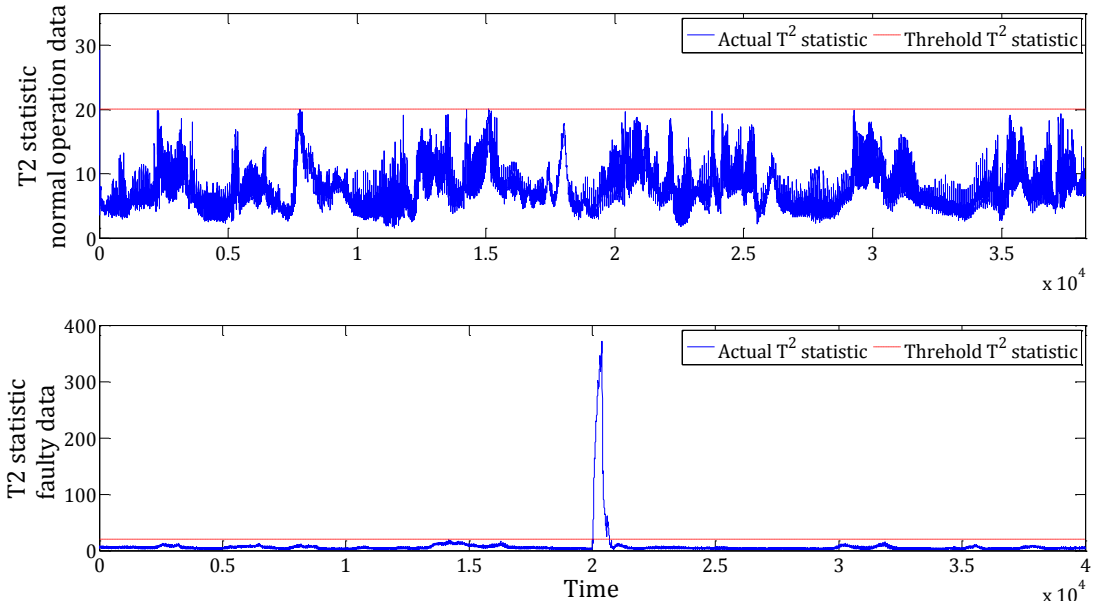


Figure 6.5 - T^2 statistic from the PMSG simulation data. Top plot: normal operation data; Bottom plot: data with three phase-to-ground fault at 1 Ω

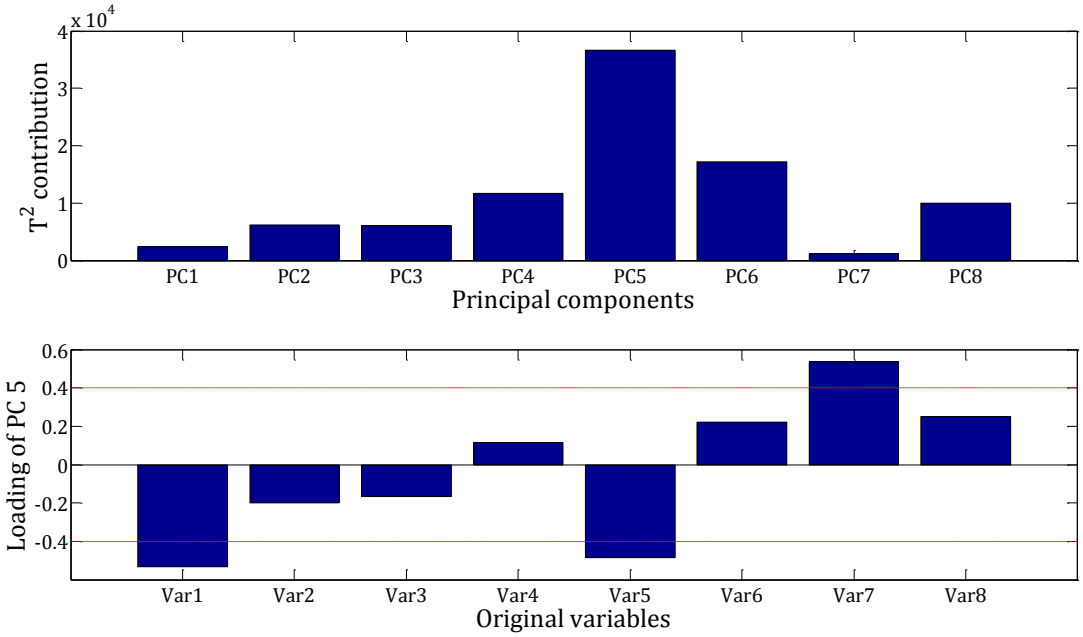


Figure 6.6 - Fault identification of the PMSG simulation data with three phases-to-ground fault at 1 Ω. Top plot: T^2 contribution from each PC; Bottom plot: loadings of the PC with the highest contribution

Results using DFIG simulation data with a capacitor ageing fault at -35% is shown in Figures 6.7 and 6.8, where nine variables are used: turbine torque (T_{mech}), generator torque (T_{elec}), DC-link current 1 (I_{dc_1}), DC-link power 2 (P_{dc2}), active power at point of coupling (P_g), stator reactive power (Q_s), phase C stator voltage (V_{sc_rms}), rotor voltage of phase A (V_{ra_rms}) and phase C (V_{rc_rms}). In the bottom plot of Figure 6.7, the T^2 statistic from the testing data has a clear indication of a fault between sample points 20000 and 20400 (5 - 5.1 s). By decomposing the T^2 statistic using equation (6-11), it was found that PC2 is the main contributor to this fault (Fig 6.8 top). In line with this result, the loadings of PC 2, as shown in the bottom plot of Figure 6.8, show that I_{dc_1} and P_{dc2} have the highest loading values. Moreover, the loading values of V_{ra_rms} and V_{rc_rms} are also very close to the threshold point. All these variables suggest that the cause or main contributor of the anomaly is related to the AC-DC-AC converter in the rotor circuit.

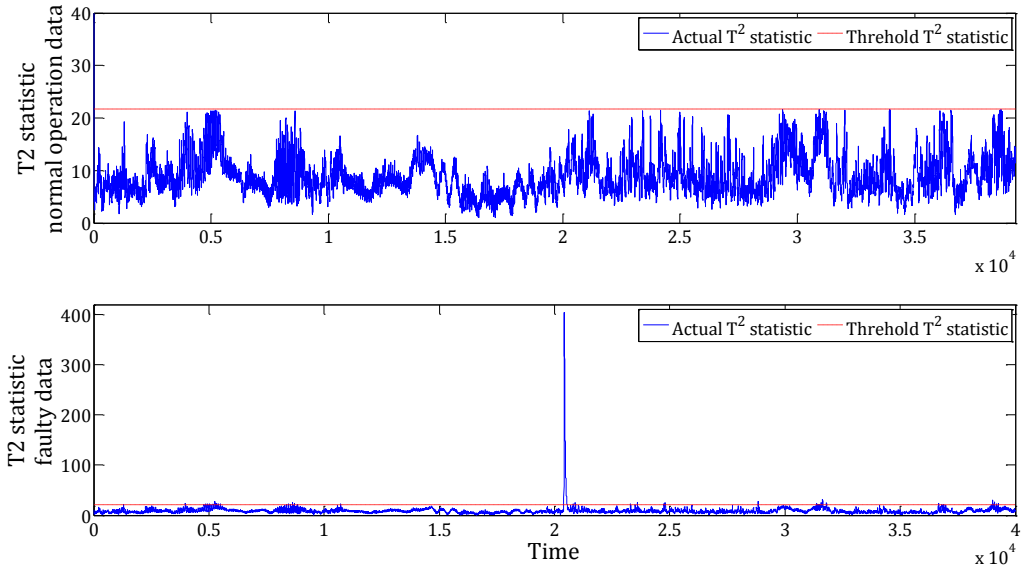


Figure 6.7 - T^2 statistic from the DFIG simulation data. Top plot: normal operation data; Bottom plot: data with a capacitor ageing fault at -35%

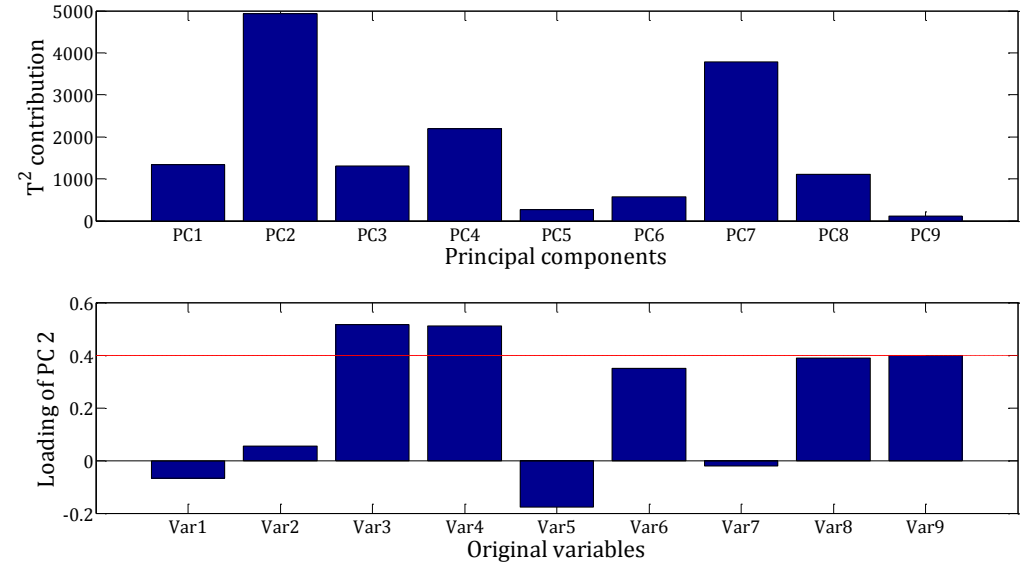


Figure 6.8 - Fault identification of the DFIG simulation data with a capacitor ageing fault at -35%. Top plot: T^2 contribution from each PC; Bottom plot: loadings of the PC with the highest contribution

The proposed method is further validated with SCADA data, where two sets of testing data containing gearbox fault and generator fault are used. The 18 variables selected for each testing dataset are shown in Table 6.1. The T^2 statistic and the threshold of the normal operating (top) and the test data (bottom) with gearbox fault are shown in Figure 6.9. The bottom plot shows that between

sample points 1400 to 1450 and 1510 to 1545, the T^2 statistic is well above the threshold. As the SCADA data has a sampling rate of 10 minutes, the detected anomalies lasted for a period of 8 hours. By decomposing the T^2 statistic, PC 1 has the highest contribution index (Fig 6.10, top). The loading values for PC 1 are then shown in the bottom plot of Figure 6.10. The composition of SCADA data is much more complex compared to the simulation data, where signals such as vibration, temperatures and environmental related parameters are also contained. Therefore, in order for the loadings to be significant, the threshold value of 0.3 is chosen.

	SCADA data with gearbox fault	SCADA data with generator fault
1	'Twist'	'Redundant pitch position 2'
2	'Redundant pitch position 3'	'Net Frequency'
3	'Active power'	'Pitch Inverter Voltage 1'
4	'Reactive power'	'System pressure'
5	'Power factor'	'Oil pressure gearbox'
6	'Pitch Inverter Voltage 2'	'Oscillation Signal Eff Z'
7	'System pressure'	'Vibration z-direction'
8	'Oscillation Signal Filt Eff Z'	'Pitch Motor 3 RPM'
9	'Vibration y-direction'	'Wind speed right-hand'
10	'Vibration z-direction'	'Wind direction'
11	'Pitch Motor 1 RPM'	'Temperature main bearing 1'
12	'Pitch Motor 3 RPM'	'Temperature hydraulic fluid'
13	'Wind direction right-hand'	'Temperature cooling water generator return'
14	'Temperature hydraulic fluid'	'Temperature generator bearing BS'
15	'Temperature generator bearing BS'	'Temperature Topbox 1'
16	'Temperature Gearbox Oil Sump'	'Temperature Topbox 2'
17	'Medium VoltageTrafo Temperature'	'LS Temperature Control Switchboard'
18	'Temperature Gearbox Bearing 2'	'Hygrometer 2 Temperature'

Table 6.1 - Variable names for the testing SCADA dataset with gearbox and generator faults

Figure 6.10 shows that the active power, temperature gearbox bearing 2 and temperature gearbox oil sump have the top three loading values. Other variables with significant loadings are the temperature generator bearing BS, power factor and pitch motor 1 RPM. The result shows that the root cause of the fault might occurs at the cooling system of the gearbox; hence, the temperature of the gearbox bearing is also increased. Furthermore, the signal shows a reduced active power output during that period, and a warning of a high gearbox

temperature is found in the alarm log. This might be because the operator intentionally controlled the turbine to operate at a lower power rating because of the warning appearing, thus avoiding damage to the turbine. Evidently, it can be determined that the anomaly is related to the gearbox.

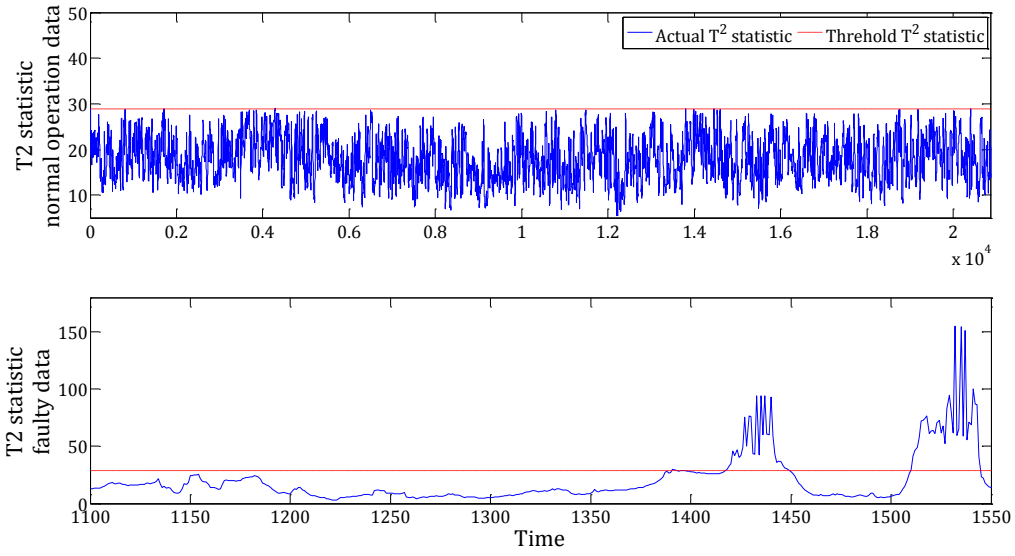


Figure 6.9 - T^2 statistic from the SCADA data. Top plot: normal operation data; Bottom plot: data with a gearbox fault

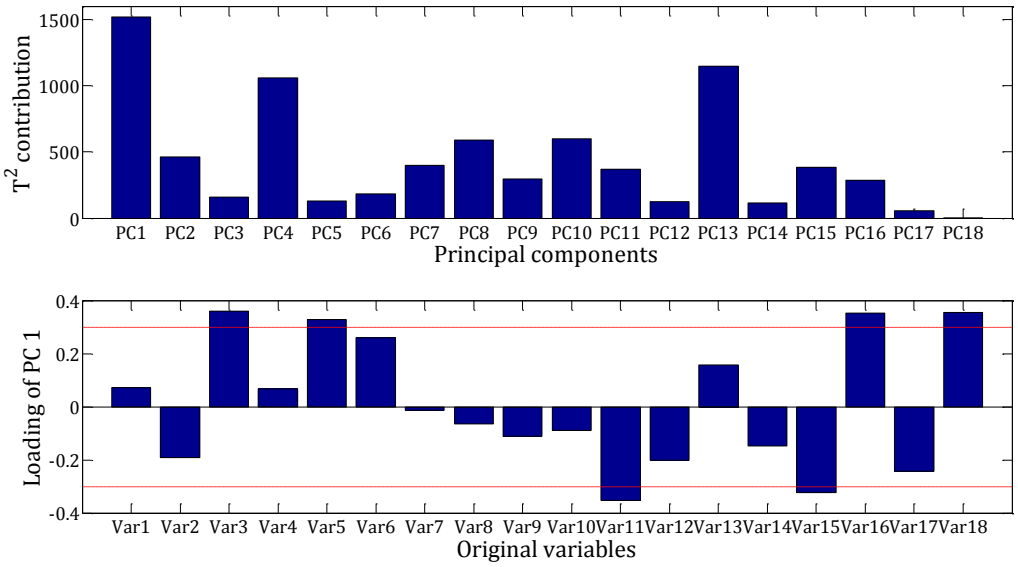


Figure 6.10 - Fault identification of the SCADA data with a gearbox fault. Top plot: T^2 contribution from each PC; Bottom plot: loadings of the PC with the highest contribution

Lastly, the SCADA data containing a generator fault is examined. For the detected anomaly, as shown in Figure 6.11, the T^2 contribution index and the loading values for the main contributor PC1 are shown in Figure 6.12. Parameters such as temperature cooling water generator return, generator bearing temperature, vibration in z direction and temperature main bearing 1 have significant loading values. The result shows that the anomaly is caused by the abnormal temperature change from these locations which are localised to the generator. This result coincides with the findings from the analysis of data and the alarm log, where a warning of a high generator bearing temperature has been flagged. Since both the main bearing and the generator bearing temperature are high contributors to this anomaly, and since the vibration in z direction is also involved, the root cause of the anomaly might be wear of the generator bearing, which results in excessive vibration of the shaft. Again, more information is needed to determine the root cause of the anomaly, as it is not provided from the SCADA data.

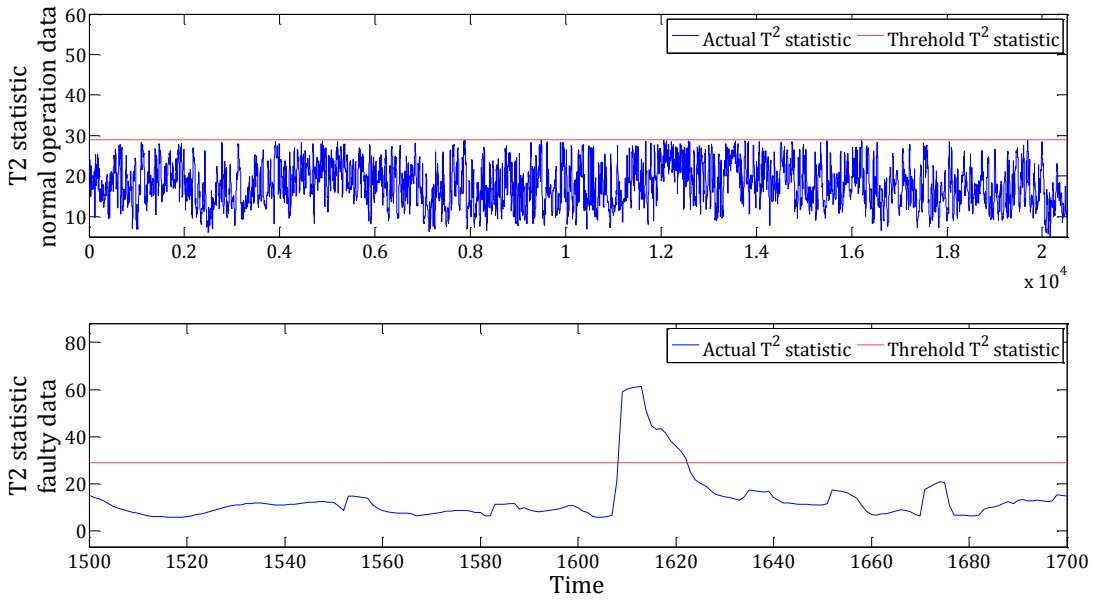


Figure 6.11 - T^2 statistic from the SCADA data. Top plot: normal operation data; Bottom plot: data with a generator fault

To conclude, the method is able to detect an anomaly in the dataset and narrow down the variables contributing to the anomaly, thereby identifying and locating the fault based on the loading values of the most contributed PC. However, the method can be sensitive to the targeting variable used in the T selection

algorithm, as different targeting variables may result in a different initial dataset. Therefore, the incorporation of prior knowledge of the system would be helpful in enhancing the accuracy of fault detection and identification.

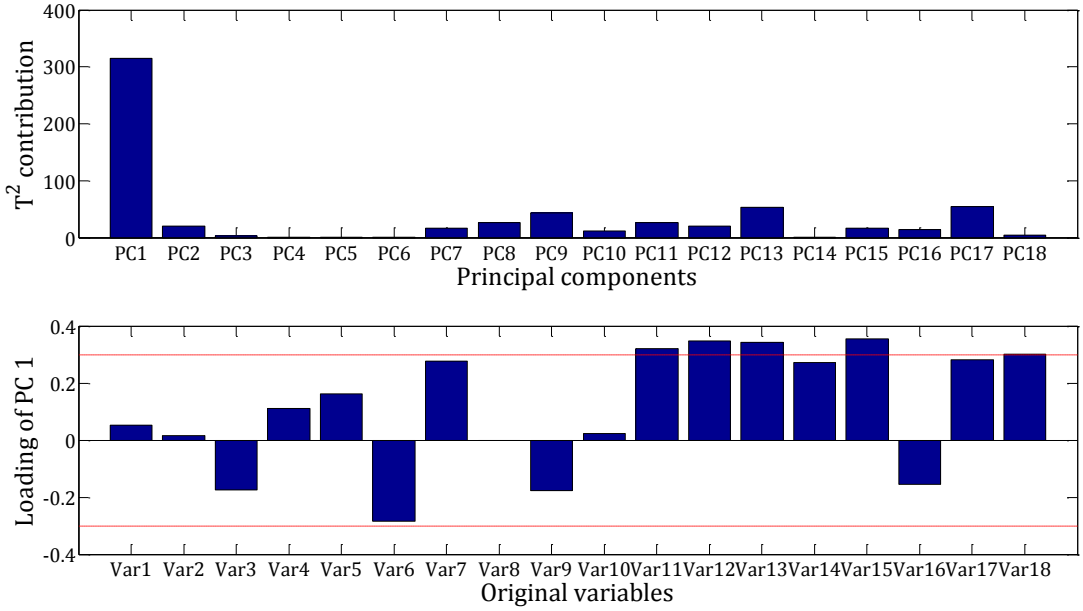


Figure 6.12 - Fault identification of the SCADA data with a generator fault. Top plot: T^2 contribution from each PC; Bottom plot: loadings of the PC with the highest contribution

6.2 Feature-based fault detection

6.2.1 Fault severity detection

In this section, a model-based method is proposed in order to identify the fault and its severity, which targets specific types of faults. It uses variables that are localised or have a great impact on the fault as a starting point. Thus, the variable set retained by the targeted selection algorithm is the ideal candidate.

Suppose the retained variable set obtained from the targeted selection algorithm $\mathbf{X}$ with dimension ($q \times n$) are related to a particular type of fault. To build the detection model, computer simulation data for these variables are collected multiple times $\mathbf{X}^r$ at different fault severities, where r is the severity index for each dataset. PCA is carried out for all these datasets, where r eigenvector matrices $\mathbf{U}$ ($q \times q$) and r eigenvalue vectors $\mathbf{L}$ ($q \times 1$) are discovered. It was found that there is a relationship between the ratio of eigenvalue and eigenvector and the fault severity S_v where:

$$S_v = f\left(U_{i,j}, L_j\right) \tag{6-13}$$

The relationship function *f* is dependent to the nature of the fault and the system. Once this model is built, it can be implemented in order to identify the fault and its severity for any testing dataset. This is achieved by transforming the testing dataset with PCA and by obtaining the corresponding eigenvalues and eigenvectors. The fault severity can then be identified by using the inverse relationship of equation (6-13).

6.2.2 Results and analysis

The PMSG with a DC-link capacitor ageing fault is considered first. The capacitance is varied from the normal condition (2.3 F) to -18% of the normal capacitance with an increment of -1%, where fault signals can be obtained while still maintaining a stable system. PCA is applied to the each dataset, where the corresponding eigenvalues $\mathbf{L}^{\mathrm{r}}$ and eigenvectors $\mathbf{U}^{\mathrm{r}}$ are obtained. It has been found that there is an obvious relationship between the fault severity and the ratio $r_{l/u}$ between $\mathbf{L}_1^r$ and $\mathbf{U}_{2,1}^r$ (the first characteristic value and the second row and first column of the characteristic vector matrix). Figure 6.13 plots this relationship, where the x axis is the fault severity and the y axis is the magnitude of the ratio. The red dots are the actual value, and the blue line is the fitted function. The function $f(u,l)$ is:

$$r_{l/u} = \left|\frac{l_1^r}{u_{2,1}^r}\right| = \frac{a}{S_v + b} \tag{6-14}$$

where a, and b are the coefficients of 78.39 and -0.9254, respectively. The fitted function has a model fit of 0.9581. It can be seen that the $r_{l/u}$ ratio is maximised when the severity is at 0%, which is where the no-fault condition and the ratios start to drop with increasing fault severity. However, it should be noted that the gradient of the plot decreases as the fault severity increases until approximately -7%. Thereafter, a more gradual gradient is obtained for the rest of the capacitance values. The result complies with the working principle of the DC-link capacitor for the PMSG system, such that all the power is delivered through the DC-link, and it is aimed to stabilise the DC-link voltage and current. Thus, the DC-link voltage and current are very sensitive to capacitance change. However, any further increase of capacitance loss will not lead to a significant change.

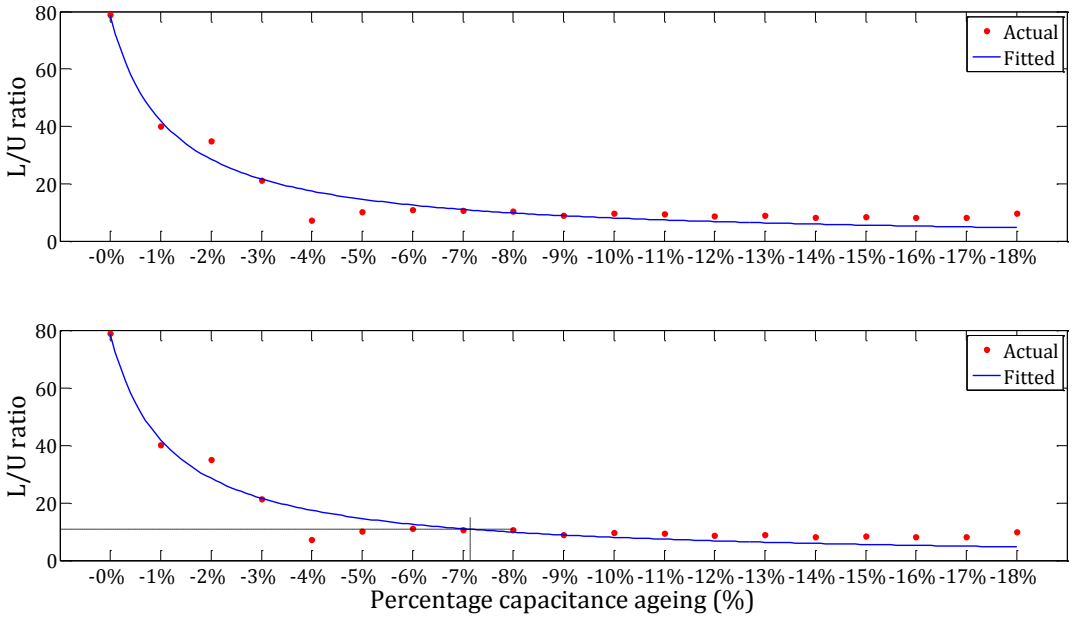


Figure 6.13 - Fault severity plot of the PMSG data with a capacitor ageing fault through PCA. Top: plot of actual and fitted model of $r_{l/u}$ ratio with severity; Bottom: estimation of unknown severity from established model

To test this model, the simulation is carried out one more time with an unknown severity. The data is used to find the $r_{l/u}$ coefficient for this particular case. Then, using the model created, the severity of the fault is identified using the inverse function f:

$$S_v = \frac{a}{r_{l/u} - b} \tag{6-15}$$

The bottom plot in Figure 6.13 shows the identification of the unknown severity from the $r_{l/u}$ ratio. The black dotted lines are the estimated $r_{l/u}$ ratio and the corresponding severity. The ratio is found to be 10.8647, and with solving equation (6-15) the fault severity is estimated to be -7.14%. Compared to the actual severity of -6.9%, there is an error of 3.4%. This proves that the method is perfectly capable of determining faults and their severity level within a reasonable margin.

The model is further tested using PMSG data with a grid phase-to-phase fault. The severities are modelled with varying resistance between the phases and ranges from no-fault occurred (1 MΩ), to 10 kΩ, 1 kΩ, 100 Ω, 50 Ω, 10 Ω, 5 Ω, 1 Ω, 0.5 Ω, 0.1 Ω, 0.05 Ω and 0.01 Ω. Using these datasets and the retained variables,

the ratio $r_{l/u}$ between $\mathbf{L}_1^r$ and $\mathbf{U}_{7,1}^r$ (same notion as described in the previous example) is shown in the top plot of Figure 6.14. The *x* axis is the logarithms of the resistance values with a base of 10 and the fitted relationship function *f* is given:

$$r_{l/u} = \left| \frac{l_1^r}{u_{1,7}^r} \right| = \text{atanh}(bS_v + c) + d \tag{6-16}$$

where the coefficients *a*, *b*, *c* and *d* are 21.15, 5.786, -4.718 and 33.31, respectively. The fitted function has an R^2 of 95.33%, which proves an accurate model fit is obtained. For this example, the no-fault condition has the logarithm value of 6 on the right side of the plot. It can be seen that the $r_{l/u}$ has the greatest change between 50 Ω (log value of 1.6989) and 0.5 Ω (log value of -0.301). This is expected as the grid's voltage is fairly stable, and too large of a resistance connected between the phases would cause a negligible effect. Moreover, as the resistance reduces to a certain value (within a stable operation condition), any further decrease will have an insignificant effect.

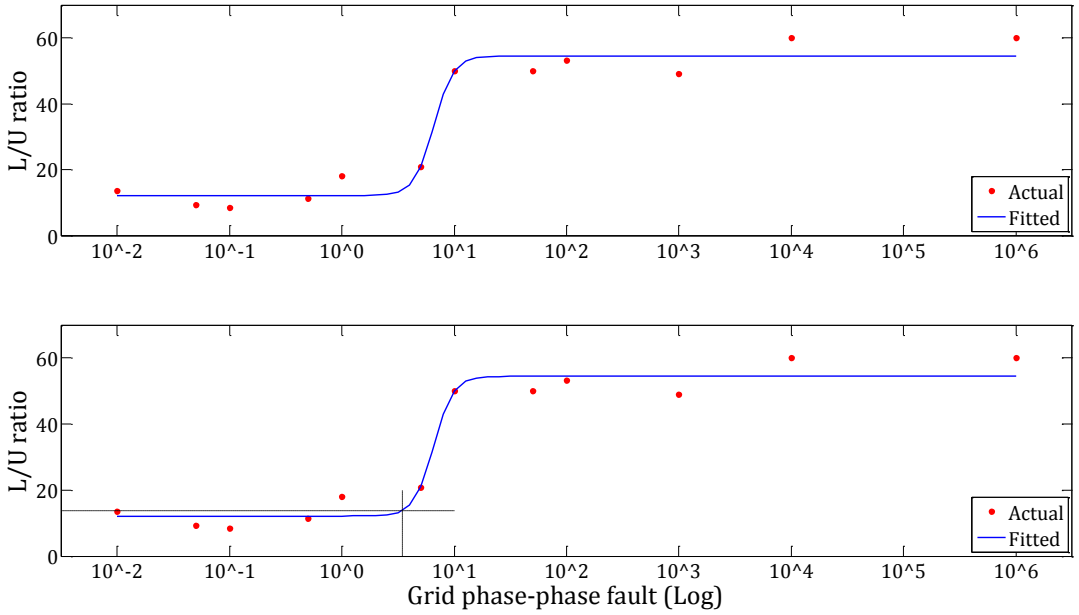


Figure 6.14 - Fault severity plot of the PMSG data with a grid side phase-to-phase fault through PCA. Top: plot of actual and fitted model of $r_{l/u}$ ratio with severity; Bottom: estimation of unknown severity from the established model.

Again, the model is tested with other simulation data with unknown fault severity. The $r_{l/u}$ ratio for the data is found to be 13.844, and the actual severity is found using the inverse *f*:

$$S_v = \frac{\operatorname{arctanh}\left(\frac{r_{l/u} - d}{a}\right) - \mathrm{c}}{b} \tag{6-17}$$

The calculated S_v is 0.5404 and the actual resistance is found to be 3.4706 Ω (Fig 6.14 bottom). Therefore, compared to the actual resistance of 3.2 Ω, there is an error of 8.4%. This again shows that the severity is found within an acceptable limit.

Finally, the method is tested with a dataset from the DFIG turbine, where it has a different system architecture. Therefore, a different relationship function is expected. Since the AC-DC-AC converter is also an essential component for the DFIG system, the capacitor ageing fault is considered. The fault severity is simulated with a percentage of capacitance losses from the normal operation condition of 7800 μF at a step of -5% until -50%. The result is shown in Figure 6.15. As can be seen in the top plot, the $r_{l/u}$ ratio between $\mathbf{L}_1^r$ and $\mathbf{U}_{1,1}^r$ has a different relationship function of:

$$r_{l/u} = \frac{l_1^r}{u_{1,1}^r} = a\tanh(bS_v + c) + d \tag{6-18}$$

where the coefficients are 3.234, 0.9597, -5.7903 and 19.06, respectively. The fitted curve has an R^2 of 94.49% which indicates an accurate fit. It can be seen that the $r_{l/u}$ has an increasing trend as the capacitance losses increase. It has the most rapid increase between -15% up to -35%. The reason for this behaviour is due to the presence of the chopper protection circuit in order to control the DC-link voltage.

Finally, the $r_{l/u}$ ratio for the dataset with an unknown fault level is found to be 19.7805. Using the inverse function (6-17), the estimated severity S_v is -26.362%, and the actual fault level is -27% as shown in the bottom plot of Figure 6.15. An error of 2.4% is found.

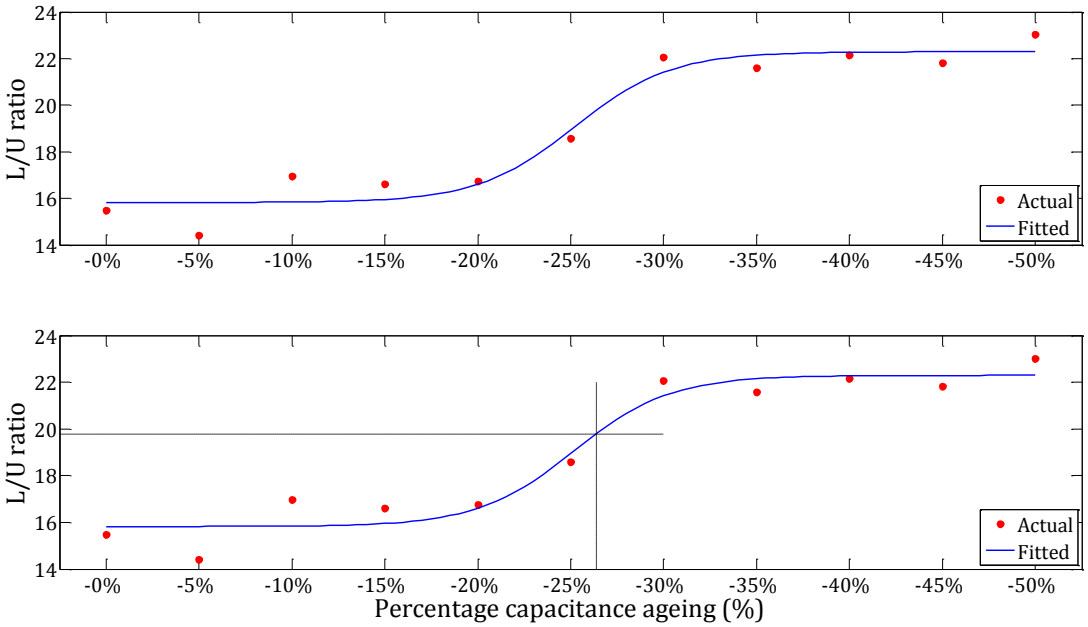


Figure 6.15 - Fault severity plot of the DFIG data with a capacitor ageing fault through PCA. Top: plot of actual and fitted model of $r_{l/u}$ ratio with severity; Bottom: estimation of unknown severity from the established model.

The results show that the proposed method is capable of estimating the fault severity within an acceptable limit. However, there are a few limitations. First, the model is built using variables selected from the T algorithm, and this requires a targeting variable. Thus, the feasibility and the accuracy of the method is dependent on selecting the correct set of variables, which means that it is crucial to carefully select a targeting variable. Second, to build the model, data from a fault with different severities are needed, and the results are based on the trial and error method. Finally, expert supervision is needed during the model building process. However, once the model is created, the method can be implemented autonomously.

6.3 Summary and discussion

In this chapter, two fault detection methods have been proposed. The methods are not only able to detect an anomaly but also have the capability of carrying out the identification and estimation of fault severity. Moreover, to simplify the detection process, the proposed method utilises the variables retained from the T selection algorithm as described in Chapter 4. This can enhance the possibility of

detecting an anomaly as only a targeting variable is required. The PC energy-based detection and identification method is validated using simulation data with different system architecture and faults, and the SCADA data from operational wind farm. Promising results have been obtained, where a fault can be identified and the location of the fault is narrowed down. In order to carry out further root cause analysis, understandings of the system structure and sensor placement are needed. Furthermore, it is feasible for the method to be tested with simulation data only because datasets from different severities are required for estimating the severity based on the features extracted. Results show accurate detection and estimation of fault severity. In the next chapter, an experimental test rig is designed and constructed, which is used to conduct experiments with faults and collect data to further validate the feature-based fault detection method.

Chapter 7. Wind Turbine Test Rig Design and Experimental Validation

In order to further validate the proposed feature-based fault detection method, experimental data obtained from a wind turbine testing platform is used. In this chapter, the design and the construction of the testing platform are first addressed. Additionally, an explanation of the experiment procedure associated with phase-to-phase short circuit fault is given, which is followed by a discussion of the data acquisition.

7.1 Description of wind turbine test rig

The aims of the experimental test rig are to simulate the dynamic operation of a wind turbine under different operation conditions and to acquire measurement data. The test rig has several advantages over the computer simulation data and SCADA data from an operational wind farm. First, the physical measurements are more realistic compared to computer simulation, as the simulation results are dependent on the mathematical model of the wind turbine. Second, measurements with higher sampling rates can be obtained. This is crucial for condition monitoring because the transient behaviour of faults may be lost when the sampling rates are low, as in the case of SCADA data. Therefore, the design and construction of the wind turbine experimental test rig is necessary.

A 3 kW PMSG-based wind turbine test rig is developed. The overall layout of the wind turbine test rig and the major components are shown in Figure 7.1 and Table 7.1, respectively. The test rig consists of the wind turbine emulator, which is responsible for simulating the rotation of the wind turbine drive train. The energy conversion from the blades and the aerodynamics are simulated in the computer; an induction motor controlled by a frequency drive is used to emulate the simulated mechanical rotation. This mechanical rotation is then coupled with a permanent magnet synchronous generator (PMSG), where the mechanical power is converted into electrical power. Moreover, the generated electrical power is regulated by an AC-DC-AC converter, such that a stable voltage and frequency can be achieved. Finally, the generated electricity is dissipated through an isolated three-phase resistive load bank. Transducers are placed within the system to collect data for monitoring and controlling purposes. The control and data acquisition of the test rig are performed by a desktop computer (PC), where all signals are interfaced through a signal condition module and digital data acquisition cards. Peripheral components such as circuit breakers, contactors, and a power supply are also included. They are used either to assist the operation of other components or for safety purposes. Extra efforts have been applied to guarantee the safe operation of the test rig and the operator's safety during the design. In the following sections, each of these components is discussed in detail.

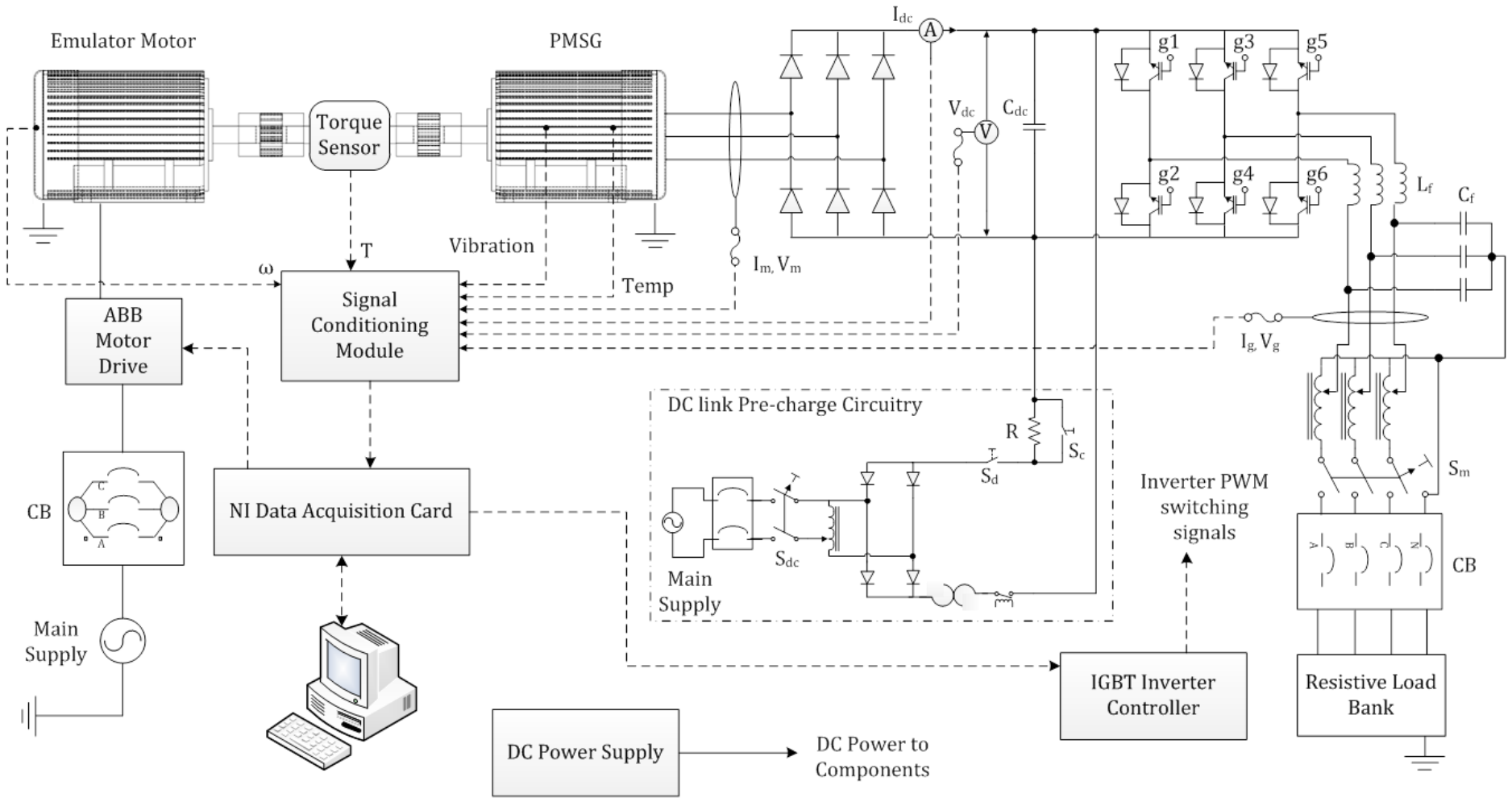


Figure 7.1 - Overall schematic block diagram of the PMSG-based experimental wind turbine test rig

Main functionality	Components
Wind turbine emulation	*Induction motor*
	Motor frequency drive
Power generation	*Permanent magnet synchronous generator*
Power regulation	*Power rectifier*
	IGBT power inverters
	IGBT inverter controller
	DC-link capacitors
	DC-link pre-charge circuitry
	Inductance filter
Power consumption	*Resistive load bank*
Measurements and signal interface	*Transducers and sensors (Voltage, current, torque, speed, vibration and temperature)*
	Signal conditioning modules
	NI data acquisition card
Controller	*Desktop computer*
Miscellaneous	*DC power supplies, contactors, circuit breakers, fuses, etc.*

Table 7.1 - Components list of the test rig based on main functionality

7.1.1 Emulation of wind turbine rotation

The emulation of a turbine's mechanical rotation is implemented by means of a squirrel-cage induction motor with a torque control. The dynamics of the turbine are simulated in the computer as described in Chapter 3.1.1. Using equations (3-2) and (3-7), the optimal power is found for any given wind speed, and the corresponding torque reference is then calculated for the frequency drive to control the induction motor.

The squirrel-cage induction motor has an IE2 standard with a rated power of 11 kW and rated speed of 1463 rpm. The size selection of the emulation motor depends on the generator power and load capacity. The test rig is designed for a rating of 3kW; thus, the PMSG used has a rated power of 3kW at the rated rpm, and the rated torque is 69.03Nm. The emulation motor has to be able to provide the required torque at the nominal speed of the generator. This can be achieved by incorporating the variable frequency drive, where the motor is able to produce the rated torque at a lower speed. Hence, the rated power required for the emulation motor is $69.03*1463*(2*\pi/60) = 10.57$kW. The next available rating is 11 kW. Therefore, an 11 kW squirrel-cage induction motor is chosen as the emulation motor. The specification of the motor is shown in Table 7.2.

Motor type		M2AA160MLA4
Rated power	kW	11
No. of poles		4
Rated speed	rpm	1463
Rated voltage	V	380-420
Rated current	A	20.7
Power factor		0.85
Inertia	kg.m2	0.064

Table 7.2 - Specification of the 11 kW emulation motor from ABB

In order for the induction motor to operate at the required torque calculated from the computer simulation at low speed, a frequency drive is used. Basically, the frequency drive adjusts the voltage and frequency supplied to the rotor winding to achieve variable speed control. A detailed explanation of the variable frequency drive can be found in [160]. When selecting the frequency drive, parameters such as power rating, full load current (FLA), supply voltage and types of control that the drive provides have to all be taken into consideration. Thereby, the ACS-550 frequency drive provided by ABB with 15kW, Y415V/50Hz and FLA of 31A is selected. Moreover, the frequency drive is able to accept an external torque reference from the computer and controls the motor to operate at this desired torque value with the built-in PI controller. Finally, the controller also consists of several safety protection functionalities, such as maximum speed limit, maximum torque and current feeding to the motor. Hence, the motor will never operate above these limits.

In addition, the emulation motor is powered from the grid through a series of protection functions, including a circuit breaker and an emergency stop button. The circuit breaker prevents excessive current from being drawn from the grid to the motor through the controller, and the emergency stop button is used to disconnect the whole system from the grid in case of any anomalies.

7.1.2 Power generation

The permanent magnet synchronous generator EOGEN 150/16 used has a power rating of 3kW from the manufacturer Mecc Alte. The specification of the generator is shown in Table 7.3.

Mechanical coupling is used to physically link the motor and generator shafts together. When selecting the coupling, apart from basic physical requirements such as bore sizes of the coupling, two main factors can affect the choice. First, the coupling has to be able to withstand the maximum torque transmitted from the motor to the generator at the rated speed. Second, the coupling should be able to tolerate a certain amount of shaft misalignments, as these misalignments will have a direct impact on the system, such as increasing stresses to the components. Based on these factors, the RADEX NN42-30 laminate coupling manufactured by KTR Couplings Limited is used. The component has a rated torque of 180 Nm with a maximum speed of 10,000 rpm. Moreover, the flexible laminated coupling is able to withstand an axial displacement of 1.4 mm and angular displacement of 1 mm.

Motor type		EOGEN150/16
Rated power	kW	3
No. of poles		16
Rated speed	rpm	415
Rated voltage	V	400
Rated current	A	4.35
Power factor		0.86
Inertia	$kg.m^2$	0.115

Table 7.3 - Specification of the 3 kW PMSG from Mecc Alte

7.1.3 Power regulation

The AC-DC-AC converter is responsible for the regulation of the power generated from PMSG due to the variable wind profile. It consists of an uncontrollable AC-DC rectifier, a DC-link and a DC-AC inverter. The selection of each of these components is discussed individually in the following sections. Moreover, it is worth mentioning that certain components referenced in later sections have been over dimensioned. The reason for this is that the selection of over-dimensioned components allows fault simulation to be carried out on the test rig without damaging the components.

The VISHAY VS-26MT80 three-phase power rectifier is selected for the AC-DC conversion. It has a maximum repetitive peak reverse voltage of 800V and a maximum forward voltage of 1.26V. The rectifier is able to handle a forward

current of 25A. As can be seen, all these values are well above the rated voltage and current of the generator. Moreover, when the system is operating at the rated condition, a large amount of heat is being generated. A heat sink is attached to the rectifier and ventilated with a fan. According to the specification, the maximum operating temperature of the rectifier is 150°C, and the maximum power dissipation is 55 W. Assuming the ambient temperature to be 50°C, and the equivalent thermal resistance of the rectifier-heat sink combination from the datasheet is 1.78°C/W. This will result in a dissipation capacity of (150-50)/1.78=56.2 W and shows that the rectifier is able to operate at the full rate.

The DC-link capacitor is used to reduce the ripple of the rectified DC voltage. Three electrolytic capacitors rated at 3900μF, 3900μF, and 2700μF, 400V are connected in a series to produce an equivalent capacitance of 1130μF, with a 1200V rating. When the system is operating at the rated condition, the maximum DC-link voltage is found to be 540.2V with equation (7-1):

$$V_{dc} = \frac{3\sqrt{3}}{\pi} V_p \tag{7-1}$$

where V_{dc} is the maximum DC-link voltage, and V_p is the peak value of the phase voltage. Then, using equation 7-2, the peak-peak ripple voltage on the DC link is calculated to be 23.1 V or 4.2% [160].

$$\Delta V_{dc} = \frac{P}{2V_{dc}C_{dc}\omega_{sync}} \tag{7-2}$$

where P is the power of the motor, C_{dc} is the DC link capacitance and ω_{sync} is the synchronous angular frequency.

A DC-link capacitor pre-charging circuit is also added. It is aimed to charge and discharge the capacitor (Fig 7.1). Power is taken from the grid and rectified into DC to perform pre-charging of the capacitor. A power resistor of 2kΩ with 225W is connected to limit the inrush current flow, and it has a charging time constant of 2.26 s. The charging and discharging of the capacitor is realised by two single pole single throw DC contactors (S_c and S_d in Figure 7.1) from TE connectivity with a maximum 900 V_{dc} and 100 A rating. Moreover, a variable transformer is included in order to adjust the DC-link voltage level, and the charging circuit can be switched on and off by a normally open contactor S_{dc} from Schneider with

voltage and a current rating of 400 Vac and 25 A, respectively. In addition, in case of any over current due to a short circuit, an overload relay from Schneider with a maximum current rating of 5A is used to prevent this by disconnecting the circuit.

The DC-AC conversion is performed with SKiiP 24AC12T4V1 three-phase IGBT inverter module, where it consists of 6 IGBTs. The component has a rated voltage of 1200V and a maximum current of 35 A. The switching gate-emitter voltage required to turn-on and off the IGBT is +15 V and -15 V respectively, where the IGBT typically has a turn-on voltage of 5.8V. Moreover, the collector to emitter voltage (V_{ce}) has a maximum saturation value of 2.1 V and 2.45 V at a junction temperature of 25°C and 125°C. In the case of a fault, the module is able to withstand a short-circuit current of 210 A at the rated voltage for 10 μs; otherwise, the component will experience a thermal breakdown. Furthermore, a maximum number of 1000 short circuit events are permitted, and the time between short circuits must be greater than 1 s. If the current increases above the nominal value (maximum 6 times the rated value), de-saturation of the IGBT will occur, and the collector-emitter current is limited with the increase of V_{ce} [161, 162]. Finally, the junction temperature of the module should never exceed 150°C.

In terms of heat dissipation, the total power losses can be divided into conduction losses and switching losses, and it can be found by equation (7-3):

$$P_{tot} = \mathrm{n} \times \left(P_{cond(igbt)} + P_{cond(d)} + P_{sw(igbt)}\right) \tag{7-3}$$

where *n* is the number of IGBTs and freewheeling diodes in the module, P_{tot} is the total power losses of the module, $P_{cond(igbt)}$ and $P_{cond(d)}$ are the conduction losses due to IGBTs and diodes and $P_{sw(igbt)}$ is the losses from the switching of the IGBT. Each of these losses for the PWM inverter application can be found by the following equations [161]:

$$P_{cond(igbt)} = \frac{1}{2}\left(V_{ce0}\frac{I}{\pi} + \frac{r_{ce0}I^2}{4}\right) + \mathrm{m}\cos\varphi\left(V_{ce0}\frac{I}{8} + \frac{r_{ce0}I^2}{3\pi}\right) \tag{7-4}$$

$$P_{cond(d)} = \frac{1}{2}\left(V_{d0}\frac{I}{\pi} + \frac{r_{d0}I^2}{4}\right) - \mathrm{m}\cos\varphi\left(V_{d0}\frac{I}{8} + \frac{r_{d0}I^2}{3\pi}\right) \tag{7-5}$$

$$P_{sw(igbt)} = \frac{\sqrt{2}}{\pi} IVf_{sw}K \quad (7\text{-}6)$$

where V_{ce0}, V_{d0}, r_{ce0} and r_{d0} are the reference collector-emitter voltage and resistance of the component that can be found from the datasheet, I and V are the operating current and voltage of the module, m is the modulation index of the PWM wave with a switching frequency of f_{sw}, φ is the phase difference between the voltage and current and K is the energy coefficient to turn the IGBT on and off. Using the manufacturer's recommendation and values from the datasheet, the total power losses is 6×(18.4+16.3+24.3)=354W. According to the datasheet, the thermal resistance of the IGBT, the diode and the heat sink is 0.85, 1.2 and 0.167 °C/W, respectively. Assuming the ambient temperature to be 50 °C (enclosed in the distribution box as shown in Figure 7.6), the dissipation power is found to be 400W. This shows the system is able to operate within the recommended temperature range. Moreover, the thermal behaviour of the system is dynamic due to the internal thermal capacities, where the thermal resistance varies with the temperature nonlinearly. A detailed discussion of the dynamic thermal analysis can be found in [161, 162]. In our study, the selection of the component and heat sink is further validated using the online thermal analysis tool provided by the manufacturer (SEMIKRON), where the dynamic thermal behaviour is also considered.

To operate the IGBT inverter, an SKHI 71 driver from SEMIKRON is used. The driver accepts an input voltage level of 0V and 5 V to turn the IGBT on and off. The output from the driver has a peak current of 2 A with +14.9 V and -6.5 V, respectively. When a gate resistor of 15 Ω is used and the driver is supplied with 15.6 V, it consumes an average current of 20 mA per IGBT. The driver is able to achieve a maximum switching frequency of 50 kHz. Moreover, several protection circuitries are also included in the driver. First, an interlock dead time circuitry is included in order to protect any short circuit due to the simultaneous switching on of the IGBTs in the same phase. The default interlock dead time for the driver is 4.1 µs. Second, the driver is able to provide an error signal through the active push-pull buffer, where a +5 V signal will be generated in the presence of a fault. To reset the driver, no error should be detected, and all input signals (switching pulses) have to set to logic low for at least 9 µs. Additional circuits are needed for

the driver to operate, and the complete AC-DC-AC converter is shown in Figure 7.6 a).

The output of the IGBT consists of the low fundamental frequency 50 Hz and the carrying switching frequency f_{sw}. An induction filter is used to filter out the high frequency components. The sinusoidal filter from REO Ltd is selected. It has a rated inductance, current and voltage of 4.7mH, 12A and 400V, respectively.

7.1.4 Isolated load

The current system operates in isolation mode, where all the generated power is dissipated to an off-shelf resistive load bank. The load bank is manufactured by Crestchic, and it has a maximum power rating of 10 kW and a rated voltage of 415V in star connection. The power rating of the load bank can be controlled through a series of $240V_{ac}$ contactors at a step of 1kW. Additional relays are used to control these AC contactors via signals from the desktop computer. Moreover, the load bank has several built-in protection systems, including over current protection, short circuit protection and over temperature protection as the power is dissipated as heat and ventilated by the fan installed. Furthermore, a 10kW variable transformer is installed to permit voltage adjustment from the AC-DC-AC converter. Finally, a three-phase circuit breaker with a rated current of 25A is used to prevent any overcurrent. An auxiliary contactor is mechanically attached to the circuit breaker. If the circuit breaker is tripped, the auxiliary contactor will terminate the main power supply to the system; hence, it will stop the emulator motor from running continuously without a load connected.

7.1.5 Measurements and signal interface

Various transducers are installed in the system to collect data either for control or monitoring purposes. Electrically, AC current and voltage transduces are used to monitor the voltage and current before and after the AC-DC-AC converter, and the DC-link voltage and current are also measured with DC sensors (Fig 7.1). For the AC measurements, both transducers are from ABB. It requires a differential supply voltage of ±12 V and measures the instantaneous voltage and current. The voltage transducer has a maximum measuring range of ±750 V (peak), which corresponds to a current output of ±50 mA. It has a bandwidth of 13 kHz, and the response time is less than 6 μs. Moreover, the linearity and accuracy of the

transducer is less than 0.3% and ±0.9%, respectively. For the current transducer, it has an output current of ±100 mA for an input current range of ± 100 A. Similarly, the bandwidth, response time, linearity and accuracy are 100 kHz, 50µs, 0.1% and ±0.5%, respectively. For the DC voltage transducers, the EDCV-VX from Omni Instruments, it has a measuring range of ± 600 V and produces an output current signal of ±20 mA. The accuracy of the sensor is ± 0.25%. Moreover, the LEM HO 25-NP current sensor is used to measure the DC-link current, where it has a measuring range of ± 25 A and a voltage output of ±2.5 V. The bandwidth, response time, linearity and accuracy are 250 kHz, 3.5 µs, 0.5% and ±1%, respectively. These sensors all have high bandwidth which allows for a faster response to an instantaneous change in a physical signal. Hence, all transient changes can be acquired. Lastly, the measuring range of the transducer is dependent on the rated value of the generator and the AC-DC-AC converter as described in Sections 7.1.2 - 7.1.3. However, it can be observed that the current transducers are over dimensioned. This is done intentionally to have a wider measuring range, which allows for measurement under a faulty condition, e.g. the phase-to-phase short circuit fault.

Mechanical properties such as the torque, speed and generator shaft vibration are also monitored. The RWT411 torque transducer from the ST technology has a measuring range of 0-300 Nm and an output signal of 0-5 V with a supply voltage of 24 V. The transducer also has a built-in speed measurement function up to 12,000 rpm. Moreover, the sensor has a bandwidth of 10 kHz and an accuracy of 0.25%. As for the shaft speed, an additional incremental encoder RI32 from Hengstler with 360 pulses per revolution is used to obtain a more accurate speed measurement. The sensor has a maximum speed of up to 3000 rpm and two TTL (transistor-transistor logic) output channels. Finally, the HS-420 accelerometer from Omni Instruments is used to measure the vibration of the generator. The vibration transducer has a 4-20 mA current output that corresponds to an input range of 0-50 mm/s. In addition, the temperature of the generator is measured with the HBS-6000 temperature transducer from Omni Instruments. The transducer is able to take measurements from 0 to 100°C and outputs a 4-20 mA current signal. The accuracy and stability of the component is ±0.25% and ±0.02%/°C, respectively.

An NI (National Instruments) USB-6229 M series multifunctional DAQ (data acquisition card) is used for measuring data logging. The USB 6229 is a high performance data acquisition module designed for fast and accurate sampling. The module also comprises the new NI signal streaming which allows sustained bidirectional high-speed streams via USB. The DAQ card consists of 32 analogue inputs (16 bits) with a sampling rate of up to 250 kS/s, 4 analogue outputs (16 bits) with a sampling rate of up to 833 kS/s and up to 48 TTL digital input/output lines, where 32 of the channels are hardware-timed at 1 MHz. The timed digital channel is ideal for PWM-based control purposes.

Analogue inputs are responsible for all the transducer measurements. For the analogue input channels, it has a maximum input range of 0-10 V (single ended measurement) and a high input impedance of 10 GΩ. As described earlier, all the transducers have different types of output signals (voltage or current source) and ranges. Therefore, additional signal conditioning and interface circuitry is required to match the measurement signal and the DAQ card.

The conditioning and interface circuitry mainly consists of four stages, as shown in Figure 7.2. The first stage is used to convert the current signal to a voltage signal through resistor R_1. As an example, at the maximum measuring range ±750 V (peak), the voltage sensor is outputting a current of ±50 mA. The recommended value of R_1 from the manufacturer is between 0 and 188 Ω. For the test rig, a 100 Ω resistor with 1% tolerance is used, and it results in a voltage output of ±5 V. Moreover, a fuse F_1 is included for voltage sensors acting as overcurrent protection. The second stage is aimed to amplify and shift the signal voltage level for each sensor to the appropriate range matching the DAQ card. The output signal can be adjusted by choosing the correct resistor values of R_2-R_4 as:

$$-V_{out} = \frac{R_4}{R_2} V_{in} + \frac{R_4}{R_3} V_{shift} \tag{7-7}$$

where V_{in} is the voltage from the sensor, and V_{shift} is the biasing voltage. In the case of the voltage sensor, a 5 V biasing voltage is used, and R_2 - R_4 are all 1 kΩ. The output voltage for stage 2 is then 0 to -10 V correspond to ±750 V (peak). A low pass filter is then used to reduce noise from the measurement signal in stage

3, where the resistor R_5 is 3.3 kΩ and the capacitor is 100 nF. The cut-off frequency is 482.3 Hz, which is very close to the designed frequency of 500 Hz. Furthermore, an inverter amplifier is used to invert the signal to 0-10 V and is followed with a unity voltage follower to reduce the output impedance of the circuit. It can be seen that the signal conditioning circuitry is able to manipulate the signal in order to fulfil the requirement of the DAQ card and to maximise the full range to enhance the resolution. Similar calculations are carried out for other sensors.

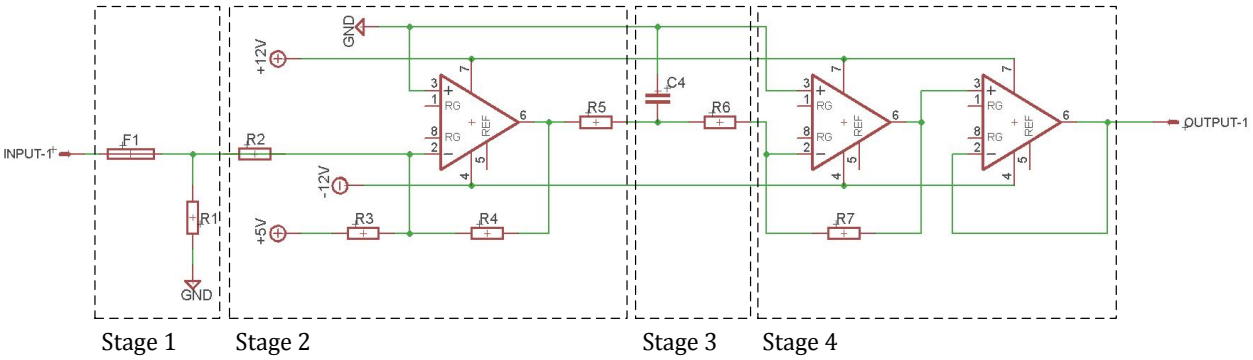


Figure 7.2 - Example of signal conditioning circuitry for a voltage sensor

7.1.6 Controller

The data acquisition and system control are realised using LabVIEW 2013 (professional version) on the desktop computer. The software is developed by National Instruments for a wide range of applications in data acquisition and system control. Graphical programming language is used to develop functions dedicated to specific tasks known as VI. These VIs can then be used as sub-functions and build the overall control for the test rig. An example of the open loop inverter control developed in a LabVIEW environment is shown in Figure 7.3. By adjusting parameters such as the output frequency, PWM frequency and modulation index, the firing PWM pulses for the IGBT inverter are generated and outputted as TTL signals through the digital lines of the DAQ card.

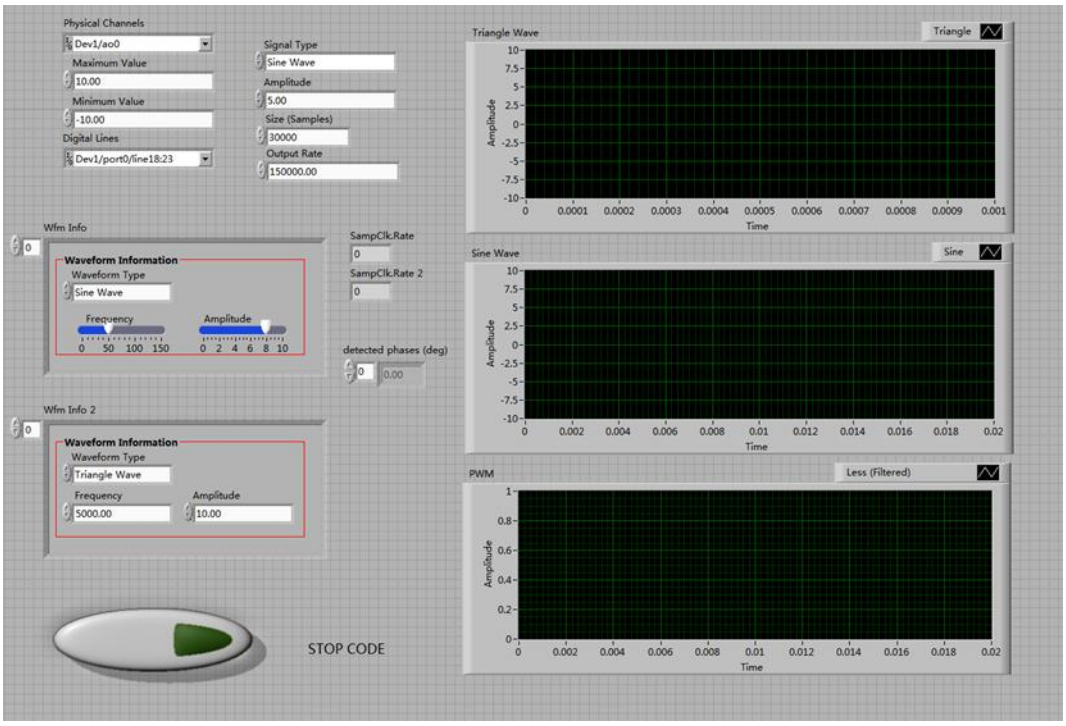


Figure 7.3 - Open loop IGBT inverter control developed in LabVIEW environment

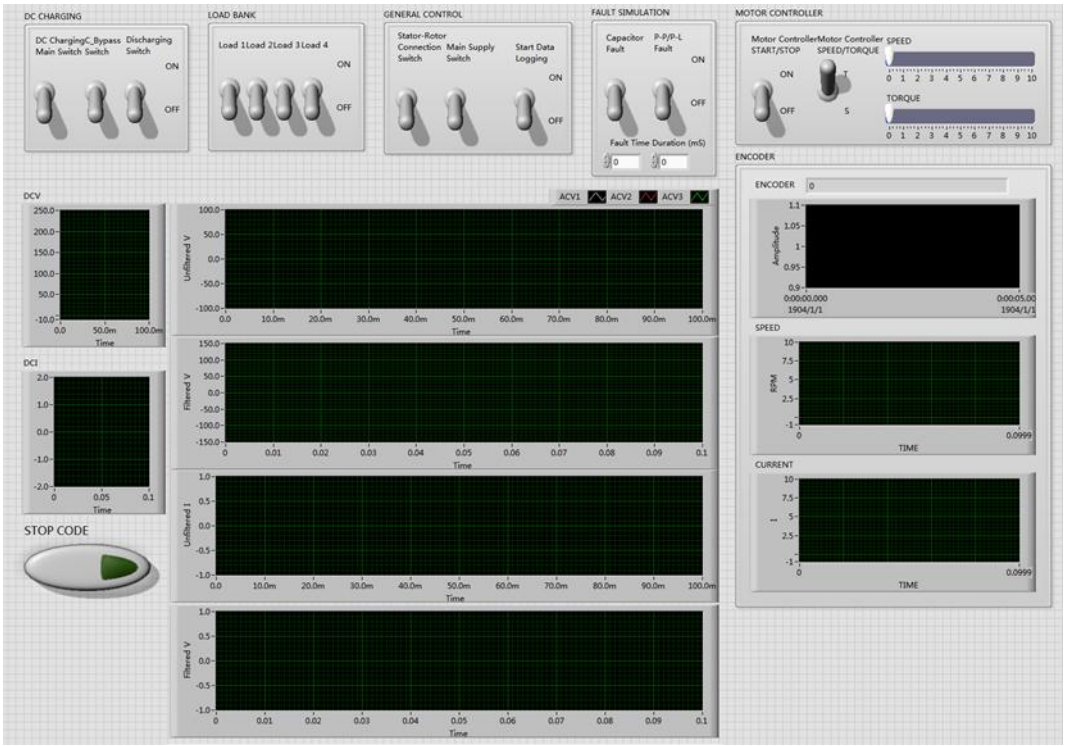


Figure 7.4 - Overall control user interface of the test rig developed in LabVIEW

Moreover, Figure 7.4 shows the overall control user interface of the test rig developed in LabVIEW. Several sub-functions are included: the DC-link pre-charging control, activation of resistive load, contactors control, emulation motor

control, fault simulation and activation of data logging. Signals such as the DC-link voltage, current and shaft speed can be viewed in real time from the user interface. The stop button at the bottom left of the user interface in Figure 4 can terminate the test rig from the operation. It can also serve as an emergency stop button. Furthermore, parameters such as the sampling rate and sensor calibration can be adjusted accordingly in the relevant sub-VI.

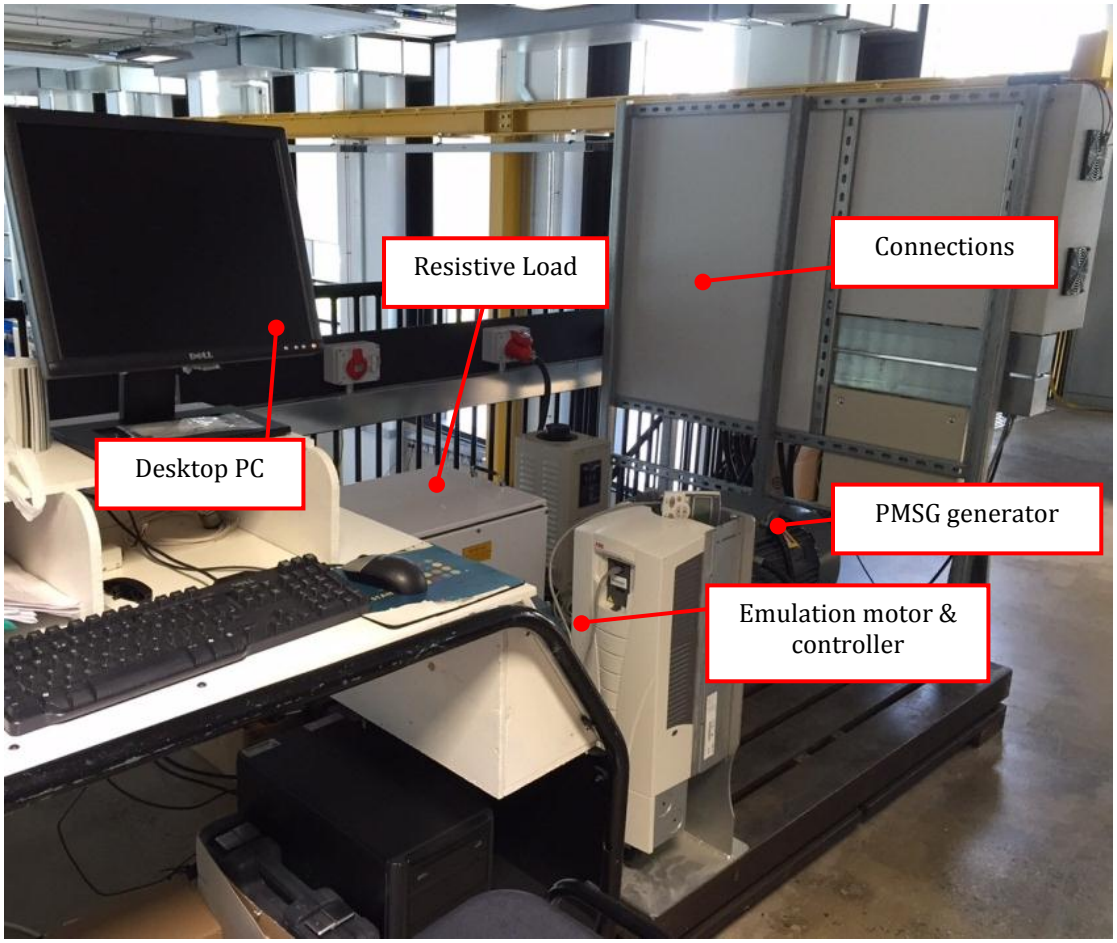


Figure 7.5 - Hardware system layout of the PMSG wind turbine test rig

The hardware layout of the complete test rig is shown in Figure 7.5, where both the emulation motor and PMSG generator are mounted on a machined steel bedplate. During operation, the system is covered by a metal fence for safety purposes, which is not shown in the figure. The desktop computer is located on the left of the test rig, where an operator should never stay in-line with the rotating motion of the system and the ventilation of the resistive load bank. Finally, all electrical connections are enclosed within three terminal boxes. The terminal boxes contain the high voltage AC-DC-AC converter, the controller and

the DC-link pre-charge circuitry (Fig 7.6 a), the main distribution box where all high voltage connections and electrical transducers are installed (Fig 7.6 b) and the low voltage signal conditioning and interface and DAQ card (Fig 7.6 c).

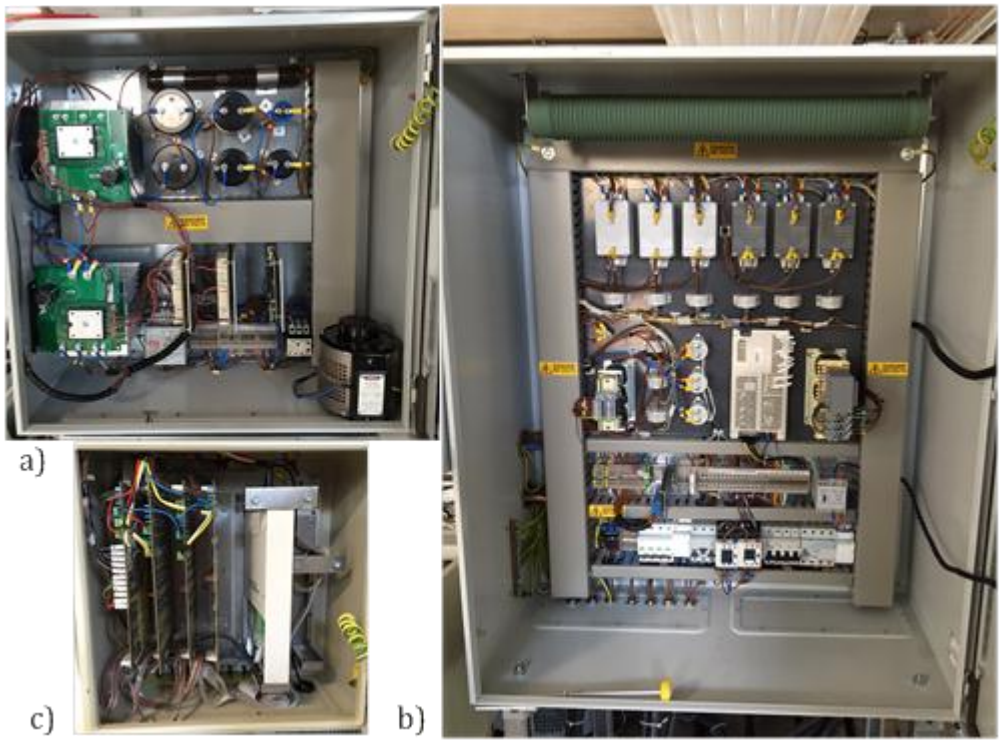


Figure 7.6 - Test rig connections: a) AC-DC-AC converter and controller; b) Main distribution box; c) Signal conditioning and interface and DAQ card

7.2 Experimental validation

7.2.1 Experimental arrangement and data acquisition

The aim of the test rig is to acquire measurement data, especially during a faulty condition. In this study, the phase-to-phase short circuit fault is simulated on the test rig and measurements are taken. Figure 7.7 shows the block diagram of the fault simulation setup. The fault is simulated by shorting two phases of the line output after the inductor filter, and controlled via a contactor. A resistor R_f is added between the two phases, where it limits the current flow between phases during a fault. Also, it is used to simulate different fault severities by adjusting the resistor value. The contactor can be controlled from the PC, where the duration of the fault can also be defined. The phase voltages and currents are measured after the inductor filter, as indicated in the figure. For safety purposes,

the experiment is carried out at a reduced voltage level. A constant power is fed to the AC-DC-AC converter, where a constant DC-link voltage of 100 V is obtained. For the PWM inverter, the switching frequency and modulation index used are 8 kHz and 0.78, respectively, which result in the best output AC waveform. The inverted AC output is then dissipated through a resistive load of 86.1 Ω per phase.

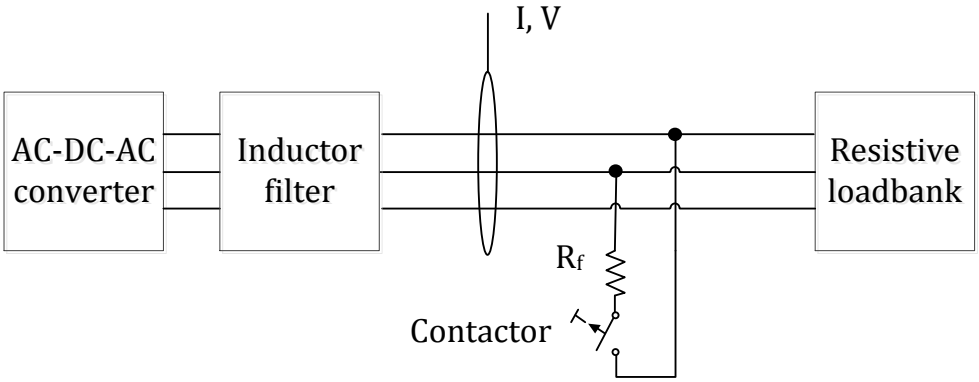


Figure 7.7 - Block diagram of phase-to-phase short circuit fault simulation on the wind turbine test rig

Figure 7.8 shows the oscilloscope measurement of the PWM switching signals of two phases outputted from the DAQ card. The signals from top to bottom are the switching signal *g1* to *g4* before the IGBT controller, as shown in Figure 7.1. It can be observed that the pulses switch between 0 and 5 V. Moreover, Figure 7.9 shows the example of the actual switching signals for one phase after the IGBT controller. The Figure 7.9 shows that the on and off voltage of the switching signals for one bridge of the IGBT (top and bottom as shown in Figure 7.1) are approximately +15 V and -8 V, respectively, and it is represented by the cyan and yellow lines. It can be seen that the on and off signals are in an opposite state. In addition, a dead time of 9 μs can be identified in the figure between the switching pulses for the top and bottom IGBT. This is crucial because the top and bottom IGBT should never be switched on at the same time to avoid short circuit.

An example of the oscilloscope measurement of the line voltage (yellow line) and phase current (blue line) is shown in Figure 7.10, where a phase-to-phase fault is simulated for 0.3 s. The voltage scale for the voltage and current readings are 50 V/div and 50mV/div, respectively. The output range for the current probe is 100mV/A. During the normal operation period, the peak line voltage and phase current are approximately 44 V and 0.3 A, respectively, and the corresponding

rms values are 31.11 V and 0.2121 A. For an ideal inverter with a sinusoidal PWM switching signal, the relationship between the DC-link voltage and the AC line voltage can be expressed as [98]:

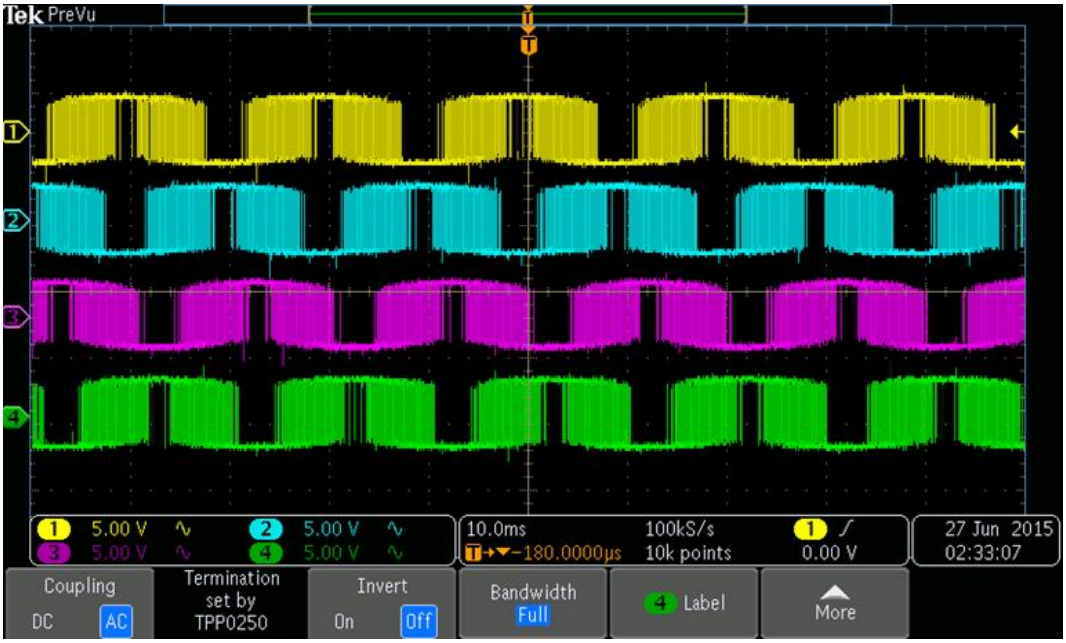


Figure 7.8 - Oscilloscope measurement of a PWM switching signal outputted from an NI DAQ card

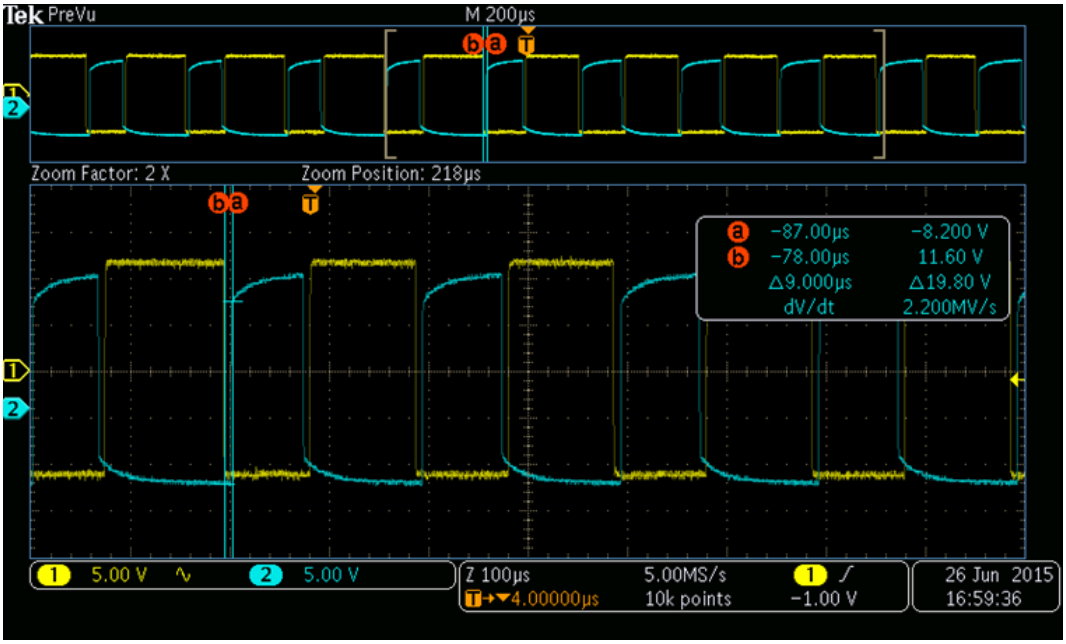


Figure 7.9 - Oscilloscope measurement of a PWM switching signal after an IGBT controller with dead time

$$V_{l,1} = m\frac{\sqrt{3}V_{dc}}{2\sqrt{2}} \tag{7.8}$$

where V_{dc} is the DC-link voltage, m is the modulation index and $V_{l,1}$ is the rms line voltage of the fundamental component. A detailed explanation of the sinusoidal PWM inverter can be found in [98]. Using equation (7-8), the ideal output rms line voltage should be 47.8 V. Compared to the oscilloscope measurement, there is a difference of 16.69 V. This is due to the losses of the IGBT and inductor filter as described in Section 7.1.3. This voltage difference will be less significant if the inverter is operated at the rated level. Furthermore, with a resistive load of 86.1 Ω, the expected peak current to be drawn is 0.295 A. The value is almost identical to the oscilloscope measurement.

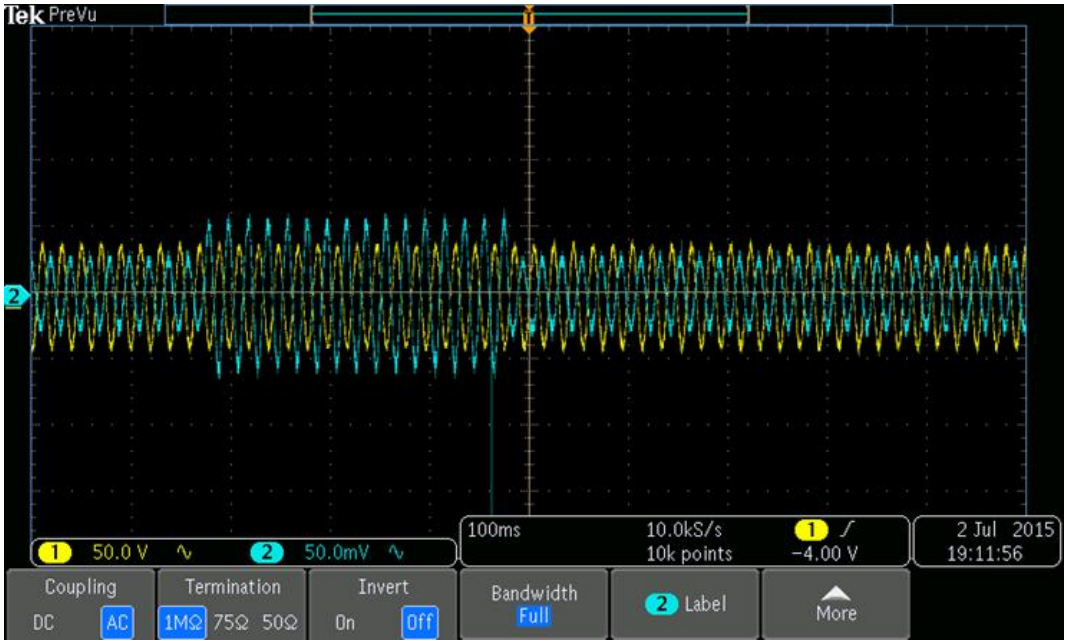


Figure 7.10 - Example of phase-to-phase fault experimental measurement from oscilloscope

Finally, an obvious increase of the phase current can be seen when the phase-to-phase short fault is simulated. The resistor R_f has a value of 351 Ω, and the peak current is approximately 0.6 A. Moreover, Figure 7.11 shows the line voltage (top plot) and phase current (bottom plot) obtained from transducers, where the data is collected with a sampling frequency of 10 kHz. It can be seen that the peak voltage and current during normal operation is approximately 52 V and 0.5 A, respectively. This is slightly larger than the measurements captured on the

oscilloscope. This is because the actual measurement value is only a small percentage of the transducer's measuring span, and relatively high noise contents are included in the measurements. As mentioned earlier, due to safety purpose, the system is operated at a low power rating. With the system operating under rated values, the noise effect should be less significant. However, the dynamic behaviour of the signal due to fault simulation can still be identified clearly, and this is more important for the study. Hence, the measurement data can be used to validate the proposed fault detection algorithm.

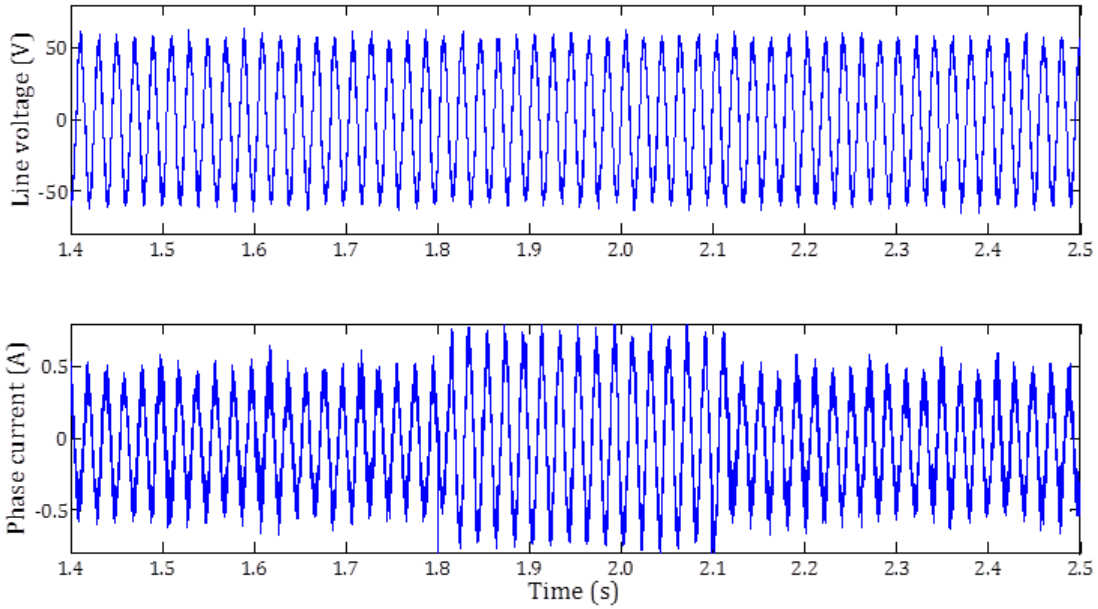


Figure 7.11 - Example of phase-to-phase fault simulation from sensor measurement

7.2.2 Fault detection and severity estimation

The phase-to-phase fault experiment is carried out under different fault severities with a resistance value of no-fault (1MΩ), to 2000 Ω, 351 Ω, 135 Ω and 27 Ω, respectively, as shown in Figure 7.12. Using the fault detection and fault severity estimation methods described in Chapter 6.2, a detection model is developed using the six line voltages and phase currents measurement. The top plot in Figure 7.13 shows the $r_{l/u}$ ratio of $\mathbf{L}_1^r$ and $\mathbf{U}_{2,2}^r$ (the first eigenvalue and second row and column of eigenvectors) for measurement at different severities as red dots, where the x axis is the logarithm scale of the resistance value with base ten (6, 3.3, 2.55, 2.13 and 1.43). The relationship function is:

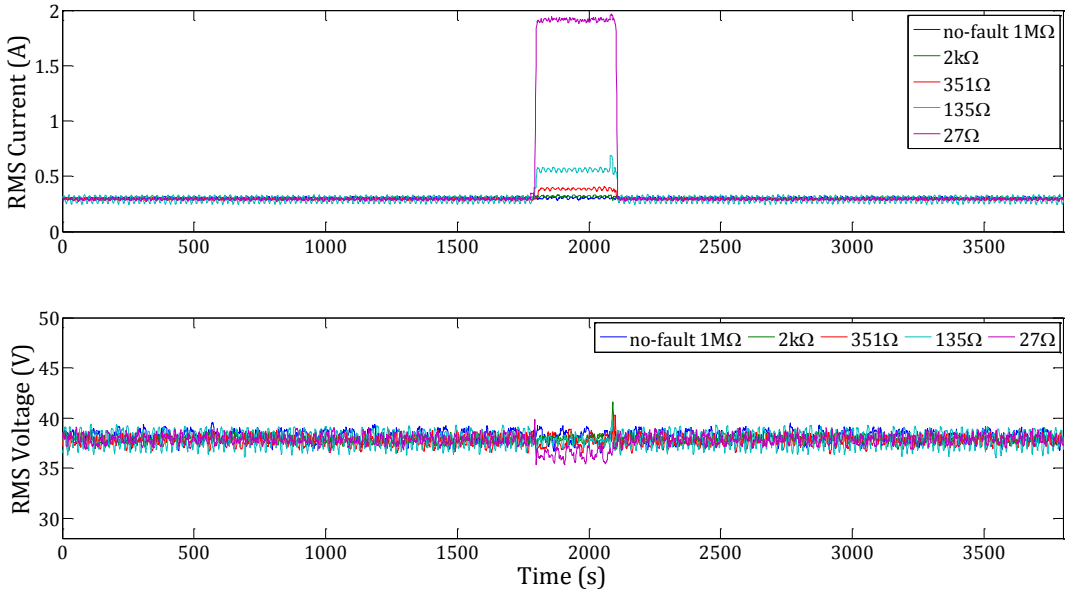


Figure 7.12 - Measured rms voltage and current under phase-to-phase fault with different severities

$$r_{l/u} = \frac{l_1^r}{u_{2,2}^r} = a \times e^{\frac{b}{S_v}} \tag{7-9}$$

where a and b are the coefficients of 0.2857 and 8.685, respectively, and S_v is the fault severity (resistance value). The fitted curve has an R^2 of 97.41%, which indicates an accurate fit. It can be seen that the $r_{l/u}$ ratio increases exponentially as the fault resistance gets smaller, hence implying higher fault severity. During testing, the dataset of an unknown fault severity (resistance value) is used to calculate the $r_{l/u}$ ratio of 10.8789. Then, using the inverse relationship function (7-10):

$$S_v = b \times ln\left(\frac{r_{l/u}}{a}\right)^{-1} \tag{7-10}$$

The estimated severity S_v is found to be 2.3863, as shown in the bottom plot of Figure 7.13. This corresponds to a resistance value of 243.4 Ω. Compared to the actual resistance value of 243 Ω, there is an error of 0.2%. The results show that a fault can be identified, and the severity can be estimated accurately.

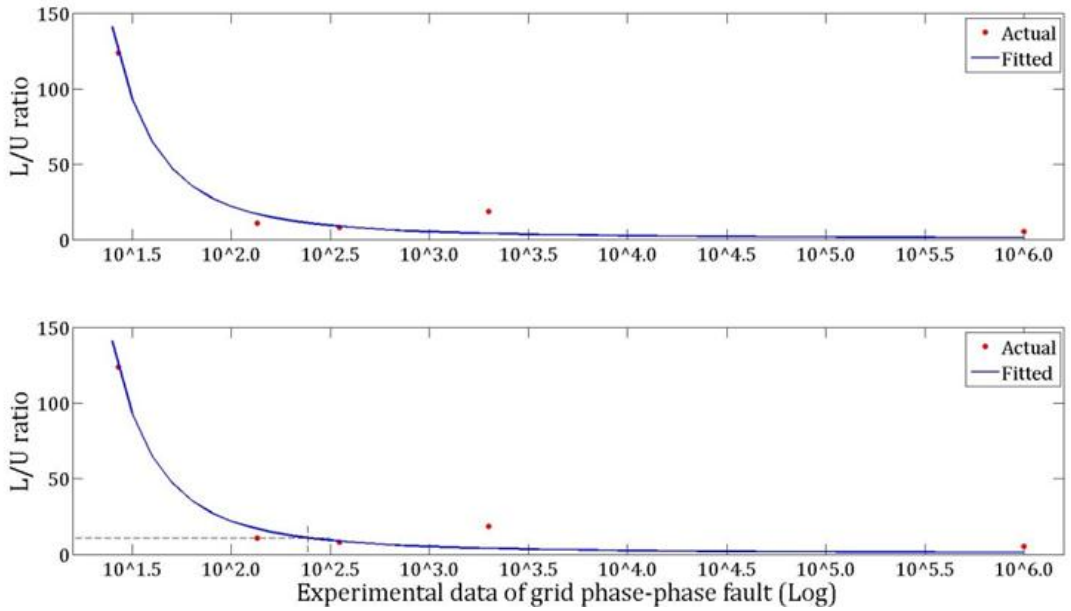


Figure 7.13 - Fault severity plot of test rig data with phase-to-phase fault through PCA. Top: plot of actual and fitted model of $r_{l/u}$ ratio with severity; Bottom: estimation of unknown severity from established model

7.3 Summary and discussion

The design and construction of an experimental PMSG wind turbine test rig has been described in this chapter. The selection and dimensioning of components involved in the test rig have been discussed in detail. The aim of the test rig was to obtain experimental data under various operation conditions for the dedicated fault; hence, the phase-to-phase short circuit fault was simulated. The experimental procedures have been described, and results have been provided. Using the experimental data, the fault detection and fault severity estimation method proposed in the previous chapter has been further validated. The results confirmed that the accurate detection and estimation of severity have been achieved. In the next chapter, a summarisation of the achievements and contributions of this research project is given, and possible future research is discussed.

Chapter 8. Conclusion and Future Work

This chapter summarises the achievements of this research and explains how the objectives stated in Section 1.3 are met by the research. Thereafter, the contributions of knowledge arising from the research are presented. Finally, the chapter concludes with recommendations for future research in the area of study.

8.1 Overview of the research objectives and achievements

With the increasing demand for a renewable energy supply, the research and development of reliable and cost effective wind power generation technology is necessary, especially for offshore wind energy. A condition monitoring system is considered to be one of the viable solutions for enhancing the performance and reliability of the wind turbines through condition-based monitoring maintenance. However, a substantial amount of monitoring data is collected and needs to be transmitted and processed, particularly for instances where high resolution data is needed for the diagnosis and prognosis of a fault. This large volume of data can become a burden for the CMS, as the demand for the processing speed and the cost of the CMS undoubtedly increases. Hence, it may reduce the reliability and performance of the CMS. However, among these monitoring data, certain degrees of redundancy unquestionably exist. Therefore, the development of a generic technique is needed to identify these redundancies. Such a technique retains a set of variables whilst still maintaining sufficient information to detect the faults and hence assess the system's conditions is desirable.

In this thesis, selection algorithms based on a multivariate principal component analysis are proposed at the system level and tested against simulation data and SCADA data from an operational wind farm. Performance measures based on data variability, information entropy and an average correlation coefficient are applied. An extension of the selection algorithm targeting a particular fault signal is also proposed, where the retained variable set not only has a strong relationship to the fault signal but also minimises redundancy among the retained set of variables. Data from simulations and SCADA under different types of faults are considered when evaluating the proposed techniques.

To further validate if vital information of the data is retained in the selected variable set, the techniques based on a correlation test of the features and an ANN prediction model are adopted. The former technique requires the identification of features of a particular fault by applying PCA to data obtained at different severities of the fault. Then, by comparing the principal components produced from the original dataset and the retained variable set, it is possible to

demonstrate if the fault feature is retained by using Pearson's correlation coefficient as a measure. On the other hand, the neural network is optimised for the ANN prediction model-based technique. Thereafter, the correlation of determination (R^2) is used to measure the model accuracy, which is then used to demonstrate if a fault feature is present in the retained variable set. Both methods have shown that a fault signature is always present in the retained variable set using various datasets with different types of faults (e.g. a capacitor ageing fault).

Using the retained variable set, two fault detection methods are proposed in order to perform detection and diagnosis: Hotelling's T^2 statistic-based method, and the feature-based detection and severity estimation method. Methods are tested with data under faulty conditions including simulation and SCADA data from an operational wind farm. For the T^2 statistic method, the detection of an anomaly between the healthy and unhealthy turbine data is first carried out by comparing the T^2 statistics. Afterwards, the principal component having the highest contribution to this anomaly is identified. The anomaly to the original measurement variables is inferred through correlating the magnitude of the characteristic vector from the PCA model to the variables, thus achieving the identification of the faults. Similarly, in the case of the feature-based detection and severity estimation method, an empirical model for a particular fault is established through the ratios of the characteristic root and vector ($r_{u/l}$) found with datasets at different fault levels. By using the empirical model, the fault can be detected, and the severity can be estimated.

Finally, the design and construction of a wind turbine experimental test rig are included for the collection of experimental data. The phase-to-phase short circuit fault is simulated at different fault severities represented by a resistor between the phases. The experiment process and data acquisition are discussed, where a high sampling frequency of 10 kHz is used to capture a transient change of the fault. The experimental data is then used to further evaluate the proposed feature-based detection and severity estimation method, and the result demonstrates the effectiveness of the method. The main achievements of this

research, which are associated with the original objectives given in Section 1.3, are as follows:

Objective 1 - Objective 1 – To conduct a theoretical review of the fundamental working principles, the parameters involved and the common failure modes of a wind turbine system.

Achievement 1 - An overview of current wind turbines with different operational conditions are reviewed in Section 2.1. The common failure modes of the key components of wind turbine are described in Section 2.2. Moreover, theoretical study of turbine reliability is also reviewed in Section 2.3.

Objective 2 - To review current researches in wind turbine condition monitoring algorithms and techniques.

Achievement 2 - An overview of a condition monitoring system is given in Section 2.3. A detailed review of current condition monitoring techniques for wind turbines is given in Section 2.5, where it is mainly categorised into statistical, time domain, frequency domain and time-frequency domain analyses. Moreover, researches in sensor selection techniques are reviewed and limited studies of the application for a wind turbine condition monitoring system is found.

Objective 3 – To model and simulate a wind turbine system with different configurations and to further investigate the dynamic behaviour of a wind turbine with different fault scenarios.

Achievement 3 - The theoretical analysis and modelling of both PMSG and DFIG wind turbine systems are carried out in Section 3.2. The wind turbines are simulated in PSCAD/EMTDC. Moreover, the DC-link capacitor ageing fault and the grid phase-to-ground short circuit fault are simulated for both models for the purpose of understanding the dynamic behaviour of faults to the system and obtaining useful simulation data to validate the proposed algorithms in this research.

Objective 4 - To propose appropriate variable selection methodologies in order to minimise redundancies in the measurement data, whilst still maintaining sufficient information to detect faults.

Achievement 4 - The PCA-based selection algorithms at the system level are proposed in Chapter 4, where the related stopping criterion and performance measures are also discussed. Moreover, the pre-processing of the data used for validation is described in Section 4.3, which includes simulation data obtained in Chapter 3 and SCADA data from an operational wind farm. In addition, both the time domain data and the frequency domain (using FFT) and instantaneous frequency domain (using HHT) data are considered. Hence, the theory of the data transformation is also presented.

Objective 5 - To validate and analyse the proposed algorithms for the particular target signal with simulation data and SCADA data from the operational wind farm.

Achievement 5 - The proposed selection algorithm is further developed so that it now targets the particular fault signal, where the selected variable sets not only minimise information redundancy but also have a strong relationship to the targeting fault signal, as described in Section 4.2.4. The algorithm using different datasets is verified, and discussions of the results are given in Section 4.4.2.

Objective 6 - To further evaluate if vital information relating to the particular fault feature is maintained in the reduced variable set obtained from the proposed variable selection algorithms.

Achievement 6 - Two methods are proposed to prove that vital fault information is present in the selected variable set. The feature-based evaluation method is described in Section 5.1, where the method is validated with simulation data. The ANN prediction model-based method is verified with both simulation and SCADA data, which is given in Section 5.2.

Objective 7 – To locate the fault and identify the severity of the fault through an analysis of a data anomaly utilising variables retained from the selection algorithm.

Achievement 7 - Hotelling's T^2-based fault detection method is given in Section 6.1. The feature-based detection method is described and tested with simulation data in Section 6.2. Both of these detection techniques have adopted data obtained from the targeted selection algorithm as described in Chapter 4. Results

have demonstrated the successful detection of the fault and estimation of the fault severity.

Objective 8 - To design and construct a wind turbine experiment test rig to further validate the proposed algorithm for the detection and severity estimation of the fault.

Achievement 8 - The purpose of the test rig development is to acquire experimental data, especially during a fault condition, and to then use the data to validate the fault detection methods proposed in Section 6.2. The description of the test rig is given in Section 7.1. Fault simulation, data acquisition and validation of the proposed methodology are presented in the following sections in Chapter 7.

8.2 Knowledge contributions arising from the research

Contribution 1 - In this study, redundancies within the measurement dataset at the system level are identified through principal component analysis, which focuses on the general data variability of the dataset. The measures combined with the cumulative percentage partial covariance, information entropy and average Pearson's correlation are proposed. The proposed selection algorithm is able to maximise dataset variability and information contained in the retained variables; in the meantime, the average correlation amongst the retained variables are kept at a minimum. The results demonstrate successful dimension reduction with minimal information loss.

Contribution 2 - The variable selection algorithm is further developed to the target particular fault signal. The method not only identifies information redundancy with the measurement dataset but also seeks variables which have the greatest relationship with the targeting variable. The correlation test of fault feature and an ANN prediction model-based method are the two techniques proposed to demonstrate that within the reduced variable set, vital fault signatures are always present for further diagnosis and prognosis. The first technique is achieved by evaluating the Pearson's correlation coefficient between the principal components in the retained variable set and the predetermined fault feature. The second technique is conducted by predicting the target signal of

interest using different input datasets with an ANN model. The predicted outputs with different input datasets (including the original dataset and the retained variable set obtained from the proposed selection algorithm) are compared and measured by prediction accuracy.

Contribution 3 - Utilising the dataset obtained from the proposed selection algorithm, fault detection and identification are carried out based on Hotelling's T^2 statistic in order to detect an anomaly and infer variables that are contributing the most to this anomaly. Different from the traditional methods, where the selection of measurement data for fault detection is based more on experience, the variable set obtained from the proposed targeted selection algorithm is used. Moreover, the variables that contribute the most to the anomaly are determined by the decomposition of T^2 statistics through PC energy, thus achieving the fault identification.

Contribution 4 - The relationship between the fault severity and the ratio of PC eigenvalue and eigenvector has been identified. Based on this finding, empirical models are developed for the DC-link capacitor ageing fault, grid phase-to-ground short circuit fault and phase-to-phase short circuit fault for different wind turbines (e.g. PMSG and DFIG). Respective empirical functions are established for the different fault types, thus allowing the identification of the fault and the estimation of fault severity.

Contribution 5 - Designed and constructed an experiment wind turbine test rig in order to study the dynamic behaviour of the turbine under a fault condition and to obtain the experimental data to validate the proposed algorithms. The experimental tests are made and consistent results are obtained to identify the fault severity of the dedicated fault, such as the phase-to-phase short circuit fault, with the proposed fault severity identification algorithm.

8.3 Future work for improvement

The limitations of the proposed techniques and recommendations concerning future work that could potentially improve the condition monitoring and fault diagnosis of wind turbines are detailed below:

- The selection algorithms have been validated with time, frequency and instantaneous frequency domain data in order to reveal more features of the dataset and to minimise information loss during variable selection. However, the results indicate that time domain data have shown the best performance, and an explanation of this finding has been provided. One of the future works can be focused on the study of the time-frequency domain data for the selection algorithms, as time-frequency data in 2D may reveal more abundant information. However, the fact that time-frequency domain data have an increased data dimension where signals from different frequency levels are extracted imposes difficulties on the current selection algorithm. Therefore, improvements for the proposed selection algorithms are needed.
- In our study, a linear PCA transformation is adopted to extract features based on the variance. However, many processes associated with wind turbines are nonlinear. Therefore, the use of nonlinear principal component analysis (NPCA) in the proposed variable selection algorithms may produce a more accurate feature extraction and minimises information loss during variable selection. However, the types of nonlinear transformation of the original data would have a significant impact on the result, e.g. the nonlinear relationship might be quadratic or exponential. Also, further development of the selection algorithms is required to adapt for the nonlinearity introduced in the PCA transformation, as the U matrix (characteristic vector) can be different.
- Several variable selection and fault detection algorithms have been proposed in this dissertation. Although human intervention are required during the parameter estimation or establishing models, once the model is established, there is the potential to perform variable selection and fault detection autonomously and for possibly implement in real time. This could be achieved by integrating artificial intelligence and should be investigated in the future.
- Due to time limitations, the construction of the experimental test rig has been simplified. For example, an open loop inverter control is implemented. The experimental test rig needs to be further developed,

where a turbine system with multi closed-loop control should be carried out. Moreover, instead of uncontrolled rectification for the AC-DC conversion, a controllable rectifier (IGBT converter) should be considered where advanced control algorithms such as generator speed control, turbine maximum power tracking and reactive power compensation can be realised.

- For safety purposes, although the dynamic behaviour of the fault can be reflected in the data collected, the fault simulation experiments are performed under a low power rating. This can result in several discrepancies in the experimental results. Firstly, the quality of measurements is poor due to the fact that only portion of the measuring range is used, and there is a high low signal to noise ratio. Secondly, with the system operating under a low power rating, certain signatures of the test rig may be excluded (i.e. pitch control). Therefore, further work to enhance the safety level of the test rig is required so that the experiments can be performed under the rated power.
- Apart from the low signal to noise ratio due to the fact that system is operating under a lower power rating, sensor conditioning circuitry should be further improved. By doing so, a higher quality of signals can be acquired for fault diagnosis and prognosis purposes.
- In our study, only the phase-to-phase short circuit fault was considered. Other electrical and mechanical faults should be further simulated on the test rig, such as the DC-link capacitor ageing fault or drive train shaft misalignment. The measurement data will then be used to further validate the proposed variable selection and fault detection algorithms.

Bibliography

[1] EC. (2009). Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of the use of energy from renewable sources and amending and subsequently repealing Directives 2001/77/EC and 2003/30/EC; *Official Journal of the European Communities L140 of 05/06/2009*, p.16-62. Available at http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32009L0028, last accessed 29th Jan 2016.

[2] EC. (2014). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - A Policy framework for climate and energy in the period from 2020 to 2030. *Commission of the European Communities COM(2014) 015 final*. Available at http://eurlex. europa.eu/legal-content/EN/TXT/?uri=CELEX:52014DC0015, last accessed 29th Jan 2016.

[3] DECC. 2011. UK Renewable Energy Roadmap. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/48128/2167-uk-renewable-energy-roadmap.pdf, last accessed 29th Jan 2016.

[4] Pineda, I., and Wilkes, J. (2015). Wind in power 2014 European Statistics. *European Wind Energy Association (EWEA)*. Available at http://www.ewea.org/fileadmin/files/library/publications/statistics/EWEA-Annual-Statistics-2014.pdf, last accessed 29th Jan 2016.

[5] Ho, A., and Pineda, I. (2015). Wind Energy Scenarios for 2030. *European Wind Energy Association (EWEA)*. Available at http://www.ewea.org/fileadmin/files/library/publications/reports/EWEA-Wind-energy-scenarios-2030.pdf, last accessed 29th Jan 2016.

[6] Ma, X. (2011). Novel early warning fault detection for wind turbine-based DG systems. *Proceedings of 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies (ISGT Europe)*.

[7] Spera, D. (2009). Wind turbine technology: Fundamental concepts of wind turbine engineering. *New York, NY: ASME Press*. ISBN: 978-0-79180-260-1.

[8] Jin, X., Zhao, G., Gao, K. and Ju, W. (2015). Darrieus vertical axis wind turbine: Basic research methods. *Renewable and Sustainable Energy Reviews*, 42, pp.212-225.

[9] Hansen, A., Iov, F., Blaabjerg, F. and Hansen, L. (2004). Review of contemporary wind turbine concepts and their market penetration. *Wind Engineering*, 28(3), pp.247-263.

[10] Spinato, F., Tavner, P., van Bussel, G. and Koutoulakos, E. (2009). Reliability of wind turbine subassemblies. IET Renew. Power Gener., 3(4), p.387.

[11] EN ISO 14224:2006, Petroleum, petrochemical and natural gas industries - collection and exchange of reliability and maintenance data for equipment

[12] Faulstich S., Durstewitz M., Hahn B., Knorr K., and Rohrig K. (2008). Windenergie Report, *Institut fu¨r solare Energieversorgungstechnik*, Kassel, Deutschland.

[13] Windstats (WSD & WSDK) quarterly newsletter, part of WindPower Weekly, Denmark.

[14] Ribrant, J., and Bertling, L. (2007). Survey of failures in wind power systems with focus on Swedish wind power plants during 1997-2005. *IEEE Transactions on Energy Conversion*, 22(1), pp.167-173.

[15] Hahn, B., Durstewitz, M., and Rohrig, K. (2007). Reliability of wind turbines – experiences of 15 years and 1,500 WTs. *Wind Energy: Proceedings of the Euromech Colloquium*, S. 329-332, Springer-Verlag, Berlin.

[16] Ensslin, C., Durstewitz, M., Hahn, B., Lange, B., and Rohrig, K. (2005). German Wind Energy Report 2005. *ISET*, Kassel

[17] Faulstich, S., Lyding, P., and Hahn, B. (2011). Component reliability ranking with respect to WT concepts and external environmental conditions, UpWind Report.

[18] Tavner, P. (2012). Offshore Wind turbines: reliability, availability and maintenance. *Institution of Engineering and Technology (IET)*, London, UK. ISBN: 978-1-84919-229-3.

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

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

[21] Jiang, Z., Karimirad, M. and Moan, T. (2014). Dynamic response analysis of wind turbines under blade pitch system fault, grid loss, and shutdown events. *Wind Energy*, 17(9), pp.1385-1409.

[22] Gill, S., Stephen, B. and Galloway, S. (2012). Wind turbine condition assessment through power curve copula modelling. *IEEE Transaction Sustainable Energy*, 3(1), pp.94-101.

[23] Zappalá, D., Tavner, P.J., Crabtree, C.J. and Sheng, S. (2014). Side-band algorithm for automatic wind turbine gearbox fault detection and diagnosis. I*ET Renewable Power Generation*, 8(4): 380-389.

[24] Tavner, P. (2008). Condition monitoring of rotating electrical machines. Institution of Engineering and Technology (IET), London, UK. ISBN: 978-0-86341-739-9.

[25] Polinder, H., Van Der Pijl, F., De Vilder, G. and Tavner, P. (2006). Comparison of Direct-Drive and Geared Generator Concepts for Wind Turbines. *IEEE Transactions on Energy Conversion*, 21(3), pp.725-733.

[26] Arabian-Hoseynabadi, H., Oraee, H. and Tavner, P. (2010). Failure Modes and Effects Analysis (FMEA) for wind turbines. *International Journal of Electrical Power & Energy Systems*, 32(7), pp.817-824.

[27] Kahrobaee, S., and Asgarpoor, S. (2011). Risk-based Failure Mode and Effect Analysis for wind turbines (RB-FMEA), *North American Power Symposium (NAPS)*. August 4-6, pp 1-7, Boston, USA.

[28] Shafiee, M. and Dinmohammadi, F. (2014). An FMEA-based risk assessment approach for wind turbine systems: A comparative study of onshore and offshore. *Energies*, 7(2), pp.619-642.

[29] Wilkinson, M., Harman, K., van Delft, T. and Darnell, B. (2014). Comparison of methods for wind turbine condition monitoring with SCADA data. *IET Renewable Power Generation*, 8(4), pp.390-397.

[30] Hameed, Z., Ahn, S. and Cho, Y. (2010). Practical aspects of a condition monitoring system for a wind turbine with emphasis on its design, system architecture, testing and installation. *Renewable Energy*, 35(5), pp.879-894.

[31] Jardine, A., Lin, D. and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing, 20(7), pp.1483-1510.

[32] Guo, P. and Bai, N. (2011). Wind Turbine Gearbox Condition Monitoring with AAKR and Moving Window Statistic Methods. *Energies*, 4(12), pp.2077-2093.

[33] Zhang, Z., Verma, A. and Kusiak, A. (2012). Fault Analysis and Condition Monitoring of the Wind Turbine Gearbox. *IEEE Transactions on Energy Conversion*, 27(2), pp.526-535.

[34] Popa, L.M., Jensen, B.B., Ritchie, E., and Boldea, L. (2003). Condition monitoring of wind generators. Proceedings of the 38th IAS Annual Meeting, Conference Record of the Industry Applications Conference. October 12-16, pp. (3)1839-1846, Salt Lake City, UT, USA,

[35] Yuan, X. and Cai, L. (2005). Variable amplitude Fourier series with its application in gearbox diagnosis—Part I: Principle and simulation. *Mechanical Systems and Signal Processing*, 19(5), pp.1055-1066.

[36] Yuan, X. and Cai, L. (2005). Variable amplitude Fourier series with its application in gearbox diagnosis—Part II: Experiment and application. *Mechanical Systems and Signal Processing*, 19(5), pp.1067-1081.

[37] Saidi, L., Fnaiech, F., Henao, H., Capolino, G. and Cirrincione, G. (2013). Diagnosis of broken-bars fault in induction machines using higher order spectral analysis. *ISA Transactions*, 52(1), pp.140-148.

[38] Jeffries, W., Chambers, J. and Infield, D. (1998). Experience with bicoherence of electrical power for condition monitoring of wind turbine blades. *IEE Proceedings - Vision, Image, and Signal Processing*, 145(3), p.141.

[39] Yang, W., Tavner, P. and Wilkinson, M. (2009). Condition monitoring and fault diagnosis of a wind turbine synchronous generator drive train. *IET Renewable Power Generation*, 3(1), p.1.

[40] Kar, C. and Mohanty, A. (2008). Vibration and current transient monitoring for gearbox fault detection using multiresolution Fourier transform. *Journal of Sound and Vibration*, 311(1-2), pp.109-132.

[41] Lu, D., and Qiao, W. (2013). Current-based fault detection for wind turbine systems via Hilbert-Huang transform. *Proceeding IEEE PES General Meeting*, July, pp. 1-5, Vancouver.

[42] Chen, J., and Hao, G. (2012). Research on the Fault Diagnosis of Wind Turbine Gearbox Based on Bayesian Networks. *Practical Application of Intelligent Systems*, 124, 217–223.

[43] Yang, S., Li, X. and Liang, M. (2011). Bearing fault diagnosis of a wind turbine using maximum likelihood detection. *Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology*.

[44] Yang, W., Tavner, P.J., Sheng, S., and Court, R.S. (2012). Information entropy: An effective approach for wind turbine condition monitoring. *EWEA conference*.

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

[46] Garcia, M., Sanz-Bobi, M. and del Pico, J. (2006). SIMAP: Intelligent System for Predictive Maintenance. *Computers in Industry*, 57(6), pp.552-568.

[47] Cross, P., Ma, X. (2014). Model-based and fuzzy logic approaches to condition monitoring of operational wind turbines. *International Journal of Automation and Computing*.

[48] Hameed, Z., Hong, Y., Cho, Y., Ahn, S. and Song, C. (2009). Condition monitoring and fault detection of wind turbines and related algorithms: A review. *Renewable and Sustainable Energy Reviews*, 13(1), pp.1-39.

[49] Lu, B., Li, Y., Wu, X. and Yang, Z. (2009). A review of recent advances in wind turbine condition monitoring and fault diagnosis. *2009 IEEE Power Electronics and Machines in Wind Applications,* June 24-16, pp. 1-7, Lincoln.

[50] Verbruggen, T. Wind Turbine Operation & Maintenance based on Condition Monitoring WT-Ω; ECN-C-03-047; *Energieonderzoek Centrum Nederland (ECN): Petten*, The Netherlands, 2003

[51] Yang, S., Li, W., and Wang, C. (2008). The intelligent fault diagnosis of wind turbine gearbox based on artificial neural network. *International Conference on Condition Monitoring and Diagnosis*, April 21-24, pp. 1327-1330, Beijing.

[52] Huang, Q., Jiang, D., Hong, L., and Ding, Y. (2008). Application of wavelet neural networks on vibration fault diagnosis for wind turbine gearbox. Advances in Neural Networks - ISNN 2008, (5264)313-320, Beijing.

[53] Lekou, D.J., Mouzakis, F., Anastasopoulos, A., and Kourousis, D. (2009). Emerging techniques for health monitoring of wind turbine gearboxes and bearings. in *Proc. EWEC 2009, Scientific Track-Operation and Maintenance*, March 16-19, Marseille, France.

[54] Yang, Z., Wang, B., Dong, X., and Liu, H. (2012). Expert system of fault diagnosis for gear box in wind turbine. *Systems Engineering Procedia*, 4, pp.189-195.

[55] Watson, S., Xiang, B., Yang, W., Tavner, P. and Crabtree, C. (2010). Condition monitoring of the power output of wind turbine generators using wavelets. *IEEE Transactions on Energy Conversion*, 25(3), pp.715-721.

[56] Wilkinson, M. R., Spinato, F., Tavner, P. J. (2007). Condition monitoring of generators & other subassemblies in wind turbine drive trains. *Proceedings of 2007 IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives (SDEMPED 2007)*. September 6-8, pp. 388-392, Cracow, Poland.

[57] Bennouna, O., Héraud, N., Camblong, H., and Rodriguez, M. (2005). Diagnosis of the doubly-fed induction generator of a wind turbine. *Wind Engineering*, 29(5), pp.431-447.

[58] Bennouna, O., Héraud, N., Camblong, H., Rodriguez, M. and Kahyeh, M. (2008). Diagnosis and fault signature analysis of a wind turbine at a variable speed. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 223(1), pp.41-50.

[59] Bennouna, O., Heraud, N., Rodriguez, M. and Camblong, H. (2007). Data reconciliation and gross error detection applied to wind power. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 221(3), pp.497-506.

[60] Simani, S., Farsoni, S. and Castaldi, P. (2014). Residual generator fuzzy identification for wind turbine benchmark fault diagnosis. *Machines*, 2(4), pp.275-298.

[61] Cross, P., Ma, X., and Wang. Y. (2013). Feature selection for artificial neural network model-based condition monitoring of wind turbines. I*n Proceedings of the 10th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies (CM 2013 & MFPT 2013)*. Bindt.

[62] Yang, S., Xiang, D., Bryant, A., Mawby, P., Ran, L. and Tavner, P. (2010). Condition Monitoring for Device Reliability in Power Electronic Converters: A Review. *IEEE Transactions on Power Electronics*, 25(11), pp.2734-2752.

[63] Lu, B., and Sharma, S. (2009). A Literature Review of IGBT Fault Diagnostic and Protection Methods for Power Inverters. *IEEE Transactions on Industry Applications*, 45(5), pp.1770-1777.

[64] Ko, Y., Kim, J., Lee, K. and Lee, D. (2012). Fault diagnosis of three-parallel voltage-source converter for a high-power wind turbine. *IET Power Electronics*, 5(7), pp.1058-1067.

[65] Ko, Y., Lee, K. (2010). Fault diagnosis of a voltage-fed PWM inverter for a three-parallel power conversion system in a wind turbine. *Journal of Power Electronics*, 10(6), pp.686-693.

[66] You, X. and Zhang, W. (2012). Fault Diagnosis of Frequency Converter in Wind Power System Based on SOM Neural Network. *Procedia Engineering*, 29, pp.3132-3136.

[67] Wang, Z. and Guo, Q. (2007). The Diagnosis Method for Converter Fault of the Variable Speed Wind Turbine Based on the Neural Networks. *Second International Conference on Innovative Computing, Informatio and Control (ICICIC 2007)*, Kumamoto.

[68] Duan, Q., Zhang, L., and Zhang, L. (2010). A fault detection and tolerant scheme for back-to-back converters in DFIG-based wind power generation systems. *3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*, Chengdu.

[69] Ciang, C., Lee, J. and Bang, H. (2008). Structural health monitoring for a wind turbine system: a review of damage detection methods. *Measurement Science and Technology*, 19(12), p.122001.

[70] Tsai, C., Hsieh, C. and Huang, S. (2006). Enhancement of damage-detection of wind turbine blades via CWT-based approaches. *IEEE Transactions on Energy Conversion*, 21(3), pp.776-781.

[71] Rumsey, M.A., and Musial, W. (2001). Application of infrared thermography nondestructive testing during wind turbine blade tests, *Transactions of the ASME, Journal of Solar Energy Engineering*, 123, p.271.

[72] Rodriguez, L., Garcia, E., Morant, F., Correcher, A. and Quiles, E. (2008). Application of latent nestling method using Coloured Petri Nets for the Fault Diagnosis in the wind turbine subsets. *IEEE International Conference on Emerging Technologies and Factory Automation*, Hamburg.

[73] Echavarria E, Tomiyama T, Huberts H, and van Bussel G. (2008). Fault diagnosis system for an offshore wind turbine using qualitative physics. *In Proc. EWEC*, Brussels, Belgium.

[74] Kammer, D. (1991). Sensor placement for on-orbit modal identification and correlation of large space structures. *Journal of Guidance, Control, and Dynamics*, 14(2), pp.251-259.

[75] Kincaid, R. and Padula, S. (2002). D-optimal designs for sensor and actuator locations. *Computers & Operations Research*, 29(6), pp.701-713.

[76] Rogova, G., and Nimier, V. (2004). Reliability in information fusion: Literature survey. *in Proceeding 7th International Conference on Information Fusion*, pp. 1158-1165, Stockholm, Sweden.

[77] Elouedi, Z., Mellouli, K. and Smets, P. (2004). Assessing sensor reliability for multisensor data fusion within the transferable belief model. *IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics*, 34(1), pp.782-787.

[78] Guo, H., Shi, W. and Deng, Y. (2006). Evaluating sensor reliability in classification problems based on evidence theory. *IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics*, 36(5), pp.970-981.

[79] Wang, H., Yao, K., Pottie, G., and Estrin, D. (2004). Entropy-based sensor selection heuristic for target localization. *in Proc. 3rd Int. Symp. Information Processing in Sensor Networks*, pp. 36–45, Berkeley, CA.

[80] Onel, T., Ersoy, C. and Delic, H. (2009). Information content-based sensor selection and transmission power adjustment for collaborative target tracking. *IEEE Transactions on Mobile Computing*, 8(8), pp.1103-1116.

[81] Shen, X. and Varshney, P. (2014). Sensor selection based on generalized information gain for target tracking in large sensor networks. *IEEE Transactions on Signal Processing*, 62(2), pp.363-375.

[82] Rowaihy, H., Eswaran, S., Johnson, M., Verma, D., Bar-Noy, A., Brown, T., and Porta, T.L. (2007). A survey of sensor selection schemes in wireless sensor networks. *in Proc. SPIE*, vol. 6562.

[83] Santi, L., Sowers, T. and Aguilar, R. (2005). Optimal Sensor Selection for Health Monitoring Systems. *41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, Tucson.

[84] Hovland, G. and McCarragher, B. (1997). Dynamic sensor selection for robotic systems. *Proceedings of International Conference on Robotics and Automation,* pp. 272-277, Albuquerque.

[85] Lynn, P. (2012). Onshore and offshore wind energy. *Oxford: Wiley*. ISBN: 978-0-470-97608-1.

[86] Manwell, J., McGowan, J. and Rogers, A. (2002). Wind energy explained: Theory, design and application. Chichester: Wiley. ISBN: 0-470-84612-7.

[87] de Vries, O. (1979). Fluid dynamic aspects of wind energy conversion. *Advisory Group for Aerospace Research and Development*, North Atlantic Treaty Organization, AGARD-AG-243.

[88] Wilson, R. E., Lissaman, P. B. S. (1974). Applied aerodynamics of wind power machine. Oregon State University.

[89] Van, T., Nguyen, T. and Lee, D. (2015). Advanced pitch angle control based on fuzzy logic for variable-speed wind turbine systems. *IEEE Transactions on Energy Conversion*, 30(2), pp.578-587.

[90] Muyeen, S., Tamura, J. and Murata, T. (2009). Stability augmentation of a grid-connected wind farm. L*ondon: Springer*. ISBN: 978-1-84800-316-3.

[91] Wasynczuk, O., Man, D. and Sullivan, J. (1981). Dynamic behaviour of a class of wind turbine generators during random wind fluctuations. *IEEE Power Engineering Review*, PER-1(6), pp.47-48.

[92] Girsang, I., Dhupia, J., Muljadi, E., Singh, M. and Pao, L. (2014). Gearbox and drivetrain models to study dynamic effects of modern wind turbines. *IEEE Transactions on Industry Applications*, 50(6), pp.3777-3786.

[93] Ramtharan, G., Jenkins, N., Anaya-Lara, O. and Bossanyi, E. (2007). Influence of rotor structural dynamics representations on the electrical transient performance of FSIG and DFIG wind turbines. *Wind Energy*, 10(4), pp.293-301.

[94] Muyeen, S., Hasan Ali, M., Takahashi, R., Murata, T., Tamura, J., Tomaki, Y., Sakahara, A. and Sasano, E. (2007). Comparative study on transient stability analysis of wind turbine generator system using different drive train models. *IET Renewable Power Generation*, 1(2), p.131.

[95] Lubosny, Z. (2003). Wind turbine operation in electric power systems. *Berlin: Springer*. ISBN: 978-3-540-40340-1.

[96] Kenneth E.O. (2012). Effects of drive train model parameters on a variable speed wind turbine. I*nternational Journal of Renewable Energy Research*, vol.2, no.1.

[97] Kim, S. and Kim, E. (2007). PSCAD/EMTDC-based modelling and analysis of a gearless variable speed wind turbine. *IEEE Transactions on Energy Conversion*, 22(2), pp.421-430.

[98] Krause, P., Wasynczuk, O., Sudhoff, S. and Pekarek, S. (2013). Analysis of electric machinery and drive systems. Third Edition, John Wiley & Sons, Inc., ISBN: 978-1-118-02429-4.

[99] Iov, F., Soerensen, P., and Blaabjerg, F. (2004). Wind turbine blockset in Matlab/Simulink: General overview and description of the model. Aalborg University: Institut for Energiteknik, Aalborg Universitet. ISBN: 8789179463.

[100] Barakat, A., Tnani, S., Champenois, G. and Mouni, E. (2010). Analysis of synchronous machine modelling for simulation and industrial applications. *Simulation Modelling Practice and Theory*, 18(9), pp.1382-1396.

[101] Hansen, A. and Michalke, G. (2008). Modelling and control of variable-speed multi-pole permanent magnet synchronous generator wind turbine. *Wind Energy*, 11(5), pp.537-554.

[102] Cheah-Mane, M., Liang, J. and Jenkins, N. (2014). Permanent magnet synchronous generator for wind turbines: Modelling, control and inertial frequency response. 49th International Universities Power Engineering Conference (UPEC), Cluj-Napoca.

[103] Haque, M., Muttaqi, K. and Negnevitsky, M. (2008). Control of a stand alone variable speed wind turbine with a permanent magnet synchronous generator. *IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, Pittsburgh.

[104] Chinchilla, M., Arnaltes, S. and Burgos, J. (2006). Control of permanent-magnet generators applied to variable-speed wind-energy systems connected to the grid. *IEEE Transactions on Energy Conversion*, 21(1), pp.130-135.

[105] Li, S., Haskew, T., Swatloski, R. and Gathings, W. (2012). Optimal and direct-current vector control of direct-driven PMSG wind turbines. *IEEE Transactions on Power Electronics*, 27(5), pp.2325-2337.

[106] Conroy, J. and Watson, R. (2007). Low-voltage ride-through of a full converter wind turbine with permanent magnet generator. I*ET Renewable Power Generation*, 1(3), p.182.

[107] Michalke, G., Hansen, A., and Hartkopf, T. (2007). Control strategy of a variable speed wind turbine with multi-pole permanent magnet synchronous generator. *in Proc. European Wind Energy Conference and Exhibition 2007 (EWEC 2007)*, May, Milan, Italy.

[108] Rabelo, B., and Hofmann, W. (2010). Doubly-fed induction generator drives for wind power plants. *Wind Power*, *InTech*. ISBN: 978-953-7619-81-7, Available at http://www.intechopen.com/books/wind-power/doubly-fed-induction-generator-drives-for-wind-power-plants, last accessed 29th Jan 2016.

[109] Fletcher, J., and Yang, J. (2010). Introduction to the doubly-fed induction generator for wind power applications. *Paths to Sustainable Energy, InTech*. ISBN: 978-953-307-401-6, Available at http://www.intechopen.com/books/paths-to-sustainable-energy/introduction-to-the-doubly-fed-induction-generator-for-wind-power-applications, last accessed 29th Jan 2016.

[110] Ekanayake, J., Holdsworth, L., XueGuang Wu, and Jenkins, N. (2003). Dynamic modelling of doubly fed induction generator wind turbines. *IEEE Transaction on Power Systems*, 18(2), pp.803-809.

[111] Farhad shahnia – PSCAD/EMTDC based simulation of double fed induction generator for wind turbines

[112] Wind turbine application technical paper, (2006). PSCAD version 4.2, Power system simulation, CEDRAT. Available at http://www.cedrat-technologies.com/fileadmin/user_upload/cedrat_groupe/Publications/Publications/2006/06/2006_Wind_turbine_PSCAD_V42_ref.pdf, last accessed on 29th Jan 2016.

[113] EMTDC User's Guide, (2004). *Manitoba HVDC Research Center Inc.*, Winnipeg, Manitoba, Canada. Available at https://hvdc.ca/uploads/ck/files/reference_material/EMTDC_User_Guide_v4_2_1.pdf, last accessed on 29th Jan 2016.

[114] Lee, K., Kim, M., Yoon, J., Lee, S., and Yoo, J. (2008). Condition monitoring of dc-link electrolytic capacitors in adjustable-speed drives. I*EEE Transactions on Industry Applications*, 44(5), pp.1606-1613.

[115] Colak, I., Sagiroglu, S. and Yesilbudak, M. (2012). Data mining and wind power prediction: A literature review. *Renewable Energy*, 46, pp.241-247.

[116] Liao, S., Chu, P. and Hsiao, P. (2012). Data mining techniques and applications - A decade review from 2000 to 2011. *Expert Systems with Applications*, 39(12), pp.11303-11311.

[117] Friedman, J.H. (1998). Data mining and statistics: What's the connection. *29th Symposium on the interf*ace: *Computing Science and Statistics,* May, Houston.

[118] Hand, D. (1999). Statistics and data mining. ACM SIGKDD Explorations Newsletter, 1(1), pp.16-19, New York.

[119] Kusiak, A., Zheng, H. and Song, Z. (2009). Wind farm power prediction: a data-mining approach. *Wind Energy*, 12(3), pp.275-293.

[120] Astolfi, D., Castellani, F., Garinei, A. and Terzi, L. (2015). Data mining techniques for performance analysis of onshore wind farms. *Applied Energy*, 148, pp.220-233.

[121] Kusiak, A. and Verma, A. (2012). A Data-Mining Approach to Monitoring Wind Turbines. *IEEE Transaction on Sustainable Energy*, 3(1), pp.150-157.

[122] Cross, P., Ma, X. (2014). Nonlinear system identification for model-based condition monitoring of wind turbines. *Renewable Energy*, 71, pp.166-175.

[123] Jackson, J., (1999). Users guide to principal components. *A Wiley-Interscience Publication*. ISBN 0-471-62267-2.

[124] Jolliffe, I. (2002). Principal component analysis. *New York: Springer*. ISBN: 978-0387954424

[125] Skittides, C. and Früh, W. (2014). Wind forecasting using Principal Component Analysis. *Renewable Energy*, 69, pp.365-374.

[126] Jeffers, J. (1967). Two case studies in the application of principal component analysis. *Applied Statistics,* 16(3), p.225.

[127] Malhi, A., Gao, R.X. (2004). PCA-based feature selection scheme for machine defect classification. *IEEE Transactions on Instrumentation and Measurement*. 53, pp.1517-1525.

[128] Lu, Y., Tang, J. and Luo, H. (2012). Wind turbine gearbox fault detection using multiple sensors with features level data fusion. *Journal of Engineering for Gas Turbines Power*, 134(4), p.042501.

[129] He, Q., Yan, R., Kong, F. and Du, R. (2009). Machine condition monitoring using principal component representations. *Mechanical Systems and Signal Processing*, 23(2), pp.446-466.

[130] Huang, N., Shen, Z., Long, S., Wu, M., Shih, H., Zheng, Q., Yen, N., Tung, C. and Liu, H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 454(1971), pp.903-995.

[131] Zhang, R., Ma, S., Safak, E. and Hartzell, S. (2003). Hilbert-Huang transform analysis of dynamic and earthquake motion recordings. *Journal of Engineering Mechanics*, 129(8), pp.861-875.

[132] Yan, R. and Gao, R. (2006). Hilbert-Huang transform-based vibration signal analysis for machine health monitoring. *IEEE Transactions on Instrumentation and Measurement*, 55(6), pp.2320-2329.

[133] Jolliffe, I. (1972). Discarding variables in a principal component analysis. I: Artificial data. *Applied Statistics*, 21(2), p.160.

[134] Jolliffe, I. (1973). Discarding variables in a principal component analysis. II: real data. *Applied Statistics*, 22(1), p.21.

[135] Beale, E., Kendall, M. and Mann, D. (1967). The Discarding of Variables in Multivariate Analysis. *Biometrika*, 54(3/4), p.357.

[136] Cumming, J. and Wooff, D. (2007). Dimension reduction via principal variables. *Computational Statistics & Data Analysis*, 52(1), pp.550-565.

[137] McCabe, G. (1984). Principal Variables. *Technometrics*, 26(2), pp.137-144.

[138] Camacho, J. and Ferrer, A. (2012). Cross-validation in PCA models with the element-wise k-fold (ekf) algorithm: theoretical aspects. *Journal of Chemometrics*, 26(7), pp.361-373.

[139] Grunwald, P., and Vitanyi, P. (2010). Shannon information and kolmogorov complexity. IEEE Transaction of Information Theory. Available at http://arxiv.org/abs/cs/0410002, accessed on 29th Jan 2016.

[140] Tang, B., Song, T., Li, F. and Deng, L. (2014). Fault diagnosis for a wind turbine transmission system based on manifold learning and Shannon wavelet support vector machine. *Renewable Energy*, 62, pp.1-9.

[141] Hawkins, D. and Fatti, L. (1984). Exploring multivariate data using the minor principal components. *The Statistician*, 33(4), p.325.

[142] Lei, M., Shiyan, L., Chuanwen, J., Hongling, L. and Yan, Z. (2009). A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews*, 13(4), pp.915-920.

[143] Kalogirou, S. (2001). Artificial neural networks in renewable energy systems applications: A review. *Renewable and Sustainable Energy Reviews*, 5(4), pp. 373-401.

[144] Catalao, J., Pousinho, H. and Mendes, V. (2009). An Artificial Neural Network Approach for Short-Term Wind Power Forecasting in Portugal. *15th International Conference on Intelligent System Applications to Power Systems*, Nov. 8-12, Curitiba.

[145] Hao. Y., and Wilamowski, B.M. (2011). Levenberg–Marquardt Training. *Industrial Electronics Handbook*, vol. 5 - Intelligent Systems, 2nd Edition, chapter 12, pp. 12-1 to 12-15, CRC Press 2011.

[146] Mathworks. (2015). Neural Network Toolbox: User's Guide (R2015b). Available on http://uk.mathworks.com/help/pdf_doc/nnet/nnet_ug.pdf, accessed on 29th Jan 2016.

[147] Abraham, A. (2005). Artificial neural networks. *Handbook of Measuring System Design*. 8(3), pp. 129.

[148] Karsoliya, S. (2012). Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture. *International Journal of Engineering Trends and Technology*, 3, 714-717.

[149] Sheela, K. and Deepa, S. (2013). Review on methods to fix number of hidden neurons in neural networks. *Mathematical Problems in Engineering*, pp.1-11.

[150] Chang, Y., Lin, J., Shieh, J. and Abbod, M. (2012). Optimization the initial weights of artificial neural networks via genetic algorithm applied to hip bone fracture prediction. *Advances in Fuzzy Systems*, pp.1-9.

[151] Beliakov, G. and Abraham, A. (2002). Global optimisation of neural networks using a deterministic hybrid approach. *Hybrid Information Systems*, pp.79-92.

[152] MacGregor, J. and Kourti, T. (1995). Statistical process control of multivariate processes. *Control Engineering Practice*, 3(3), pp.403-414.

[153] Montgomery, D.C. (2013). Introduction to Statistical Quality Control, 7th ed., *Wiley: New York*, NY, USA.

[154] Tao, E., Shen, W., Liu, T. and Chen, X. (2013). Fault diagnosis based on PCA for sensors of laboratorial wastewater treatment process. *Chemometrics and Intelligent Laboratory Systems*, 128, pp.49-55.

[155] Garc´ıa-Alvarez, D., Funente, M.J., Vega, P., and Sainz, G. (2009). Fault detection and diagnosis using multivariate statistical techniques in a wastewater treatment plant. *7th IFAC International Symposium on Advanced Control of Chemical Processes*.

[156] Mirin, S. and Abdul Wahab, N. (2014). Fault Detection and Monitoring Using Multiscale Principal Component Analysis at a Sewage Treatment Plant. *Jurnal Teknologi*, 70(3).

[157] Kim, K., Sheng, S., and Fleming, P. (2011). Use of SCADA data for failure detection in wind turbines. *Energy Sustainability Conference and Fuel Cell Conference*, August 7-10, Washington, D.C.

[158] Kusiak, A. and Verma, A. (2013). Monitoring wind farms with performance curves. *IEEE Transactions of Sustainable Energy*, 4(1), pp.192-199.

[159] Yang, H., Huang, M., and Yang, S. (2015). Integrating auto-associative Neural networks with Hotelling T^2 control charts for wind turbine fault detection. *Energies*, 8(10), 12100-12115.

[160] Kretschmar, K. (1998). An AC converter with a small DC link capacitor for a 15 KW permanent magnet synchronous integral motor. *7th International Conference on Power Electronics and Variable Speed Drives*, September 21-23, London.

[161] SEMIKRON. (2015) Application manual power semiconductors. *Semikron International GmbH.* ISBN: 978-3-938843-83-3. Available at http://www.semikron.com/dl/service-support/downloads/download/semikron-application-manual-power-semiconductors-english-en-2015, accessed on 29th Jan 2016.

[162] SEMIKRON. (2014). MiniSKiiP Generation II Technical Explanations Ver.3.4. Available at http://www.semikron.com/dl/service-support/downloads/download/semikron-technical-explanation-miniskiip-generation-ii-en-2014-07-rev-3-4, accessed on 29th Jan 2016.

Appendix A

A.1 Reference frame transformation

The transformation of variables in a three phase electrical machine into an arbitrary reference frame can be expressed as:

$$\mathbf{f}_{dq0} = \mathbf{K}\mathbf{f}_{abc} \tag{A-1}$$

where $\mathbf{f}_{dq0}$ and $\mathbf{f}_{abc}$ are vectors of parameters involved in the electrical machine including resistance, inductance and voltage in the *dq0* reference and *abc* natural reference respectively. **K** is the transformation matrix:

$$\mathbf{K} = \frac{2}{3}\begin{bmatrix} \cos(\theta) & \cos\left(\theta - \frac{2\pi}{3}\right) & \cos\left(\theta + \frac{2\pi}{3}\right) \\ \sin(\theta) & \sin\left(\theta - \frac{2\pi}{3}\right) & \sin\left(\theta + \frac{2\pi}{3}\right) \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \tag{A-2}$$

Also the inverse transformation is:

$$\mathbf{K}^{-1} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 1 \\ \cos\left(\theta - \frac{2\pi}{3}\right) & \sin\left(\theta - \frac{2\pi}{3}\right) & 1 \\ \cos\left(\theta + \frac{2\pi}{3}\right) & -\sin\left(\theta + \frac{2\pi}{3}\right) & 1 \end{bmatrix} \tag{A-3}$$

The transformed reference frame has an angular velocity of ω and an angular displacement of θ as shown in Figure 3.15,

$$\theta = \int \omega \, dt \tag{A-4}$$

A.2 Manufacturer parameters for synchronous machine

Relationships between the machine parameters as described in Table 3.1 and machine reactances are given:

$$X_d = X_{\sigma D} + X_{mD} \tag{A-5}$$

$$X_q = X_{\sigma Q} + X_{mQ} \tag{A-6}$$

$$\dot{X}_d = X_{\sigma D} + \frac{X_{mD}X'_{\sigma f}}{X_{mD}+X'_{\sigma f}} \tag{A-7}$$

$$\ddot{X}_d = X_{\sigma D} + \frac{X'_{\sigma d}X'_{\sigma f}}{X'_{\sigma d}+X'_{\sigma f}} \tag{A-8}$$

$$\ddot{X}_q = X_{\sigma Q} + \frac{X'_{\sigma q}X_{mQ}}{X'_{\sigma d}+X_{mQ}} \tag{A-9}$$

$$\dot{T}_d = \frac{1}{\omega_r r'_f}\left(X'_{\sigma f} + \frac{X_{mD}X_{\sigma D}}{X_{mD}+X_{\sigma D}}\right) \tag{A-10}$$

$$\dot{T}_{do} = \frac{X_{mD}+X'_{\sigma f}}{\omega_r r'_f} \tag{A-11}$$

$$\ddot{T}_d = \frac{1}{\omega_r r'_d}\left(X'_{\sigma d} + \frac{X_{\sigma D}X'_{\sigma f}}{X_{\sigma D}+X'_{\sigma f}}\right) \tag{A-12}$$

$$\ddot{T}_{do} = \frac{1}{\omega_r r'_d}\left(X'_{\sigma d} + \frac{X_{mD}X'_{\sigma f}}{X_{mD}+X'_{\sigma f}}\right) \tag{A-13}$$

$$\ddot{T}_q = \frac{1}{\omega_r r'_q}\left(X'_{\sigma q} + \frac{X_{mQ}X_{\sigma Q}}{X_{mQ}+X_{\sigma Q}}\right) \tag{A-14}$$

$$\ddot{T}_{qo} = \frac{X_{mQ}+X'_{\sigma q}}{\omega_r r'_q} \tag{A-15}$$

where the superscript ' represents the equivalent parameters of the rotor side in the stator side. X_{mD} and X_{mQ} are the main reactances for the direct axis and the quadrature axis; $X_{\sigma D}$, $X_{\sigma Q}$, $X_{\sigma d}$, $X_{\sigma q}$ are the leakage reactance of stator and damper windings in the *dq0*-axes, and the permanent magnet the equivalent resistance of r_f' and equivalent leakage reactance of X_f respectively.

A.3 PSCAD simulation parameters

The parameters involved in the simulation of the PMSG and DFIG mode is given in Table A.1 and A.2 respectively.

Parameter name	Value
f_{rated}	50 Hz
Poles	100
P_{rated}	3 MW
Inertia	6.3 s
$\boldsymbol{X_d}$	0.4 pu
$\boldsymbol{X_q}$	0.51 pu
$\boldsymbol{\dot{X}_d}$	0.3 pu
$\boldsymbol{\ddot{X}_d}$	0.22 pu
$\boldsymbol{\ddot{X}_q}$	0.29 pu
$\boldsymbol{\dot{T}_{do}}$	10 s
$\boldsymbol{\ddot{T}_{do}}$	0.001 s
$\boldsymbol{\ddot{T}_{qo}}$	0.034 s

Table A.1 - Ratings and parameters of the permanent magnet synchronous generator

Parameter name	Value
f_{rated}	50 Hz
P_{rated}	2 MW
Inertia	0.85 s
r_s	0.0054 pu
r_r	0.00607 pu
X_{ms}	4.5 pu
$X_{\sigma s}$	0.1 pu
$X_{\sigma r}$	0.11 pu

Table A.2 - Ratings and parameters of the doubly-fed induction generator

Appendix B

B.1 Retained variable sets at the system level

The retained variables from B2, B4 and H selection methods in time, frequency and instantaneous frequency domain using PMSG and DFIG data are show in Table B.1 and B.2 respectively. The description of variable name for PMSG and DFIG simulation data can be found in Table 4.1 and 4.2 respectively.

Time			Frequency			Instantaneous frequency		
B2	*B4*	*H*	*B2*	*B4*	*H*	*B2*	*B4*	*H*
'I_{dc1}'	'V_{dc}'	'α'	'I_{dc1}'	'I_{dc1}'	'V_{dc}'	'I_{dc2}'	'V_{dc}'	'I_{dc1}'
'V_{dc}'	'I_{dc2}'	'P_{source}'	'P_{grid}'	'I_{dc_ref}'	'I_{dc2}'	'Q_{grid}'	'I_{dc2}'	'V_{dc}'
'Q_{grid}'	'Q_{grid}'	'P_1'	'I_{dc_ref}'	'α'	'Q_{grid}'	'P_{grid}'	'Q_{grid}'	'Q_{grid}'
'P_{grid}'	'P_{grid}'	'P_2'	'α'	'β'	'P_{grid}'	'β'	'α'	'P_{grid}'
'I_{dc_ref}'	'I_{dc_ref}'	'P_3'	'β'	'Q_{gen}'	'β'	'Q_{gen}'	'Q_{gen}'	'I_{dc_ref}'
'α'	'α'	'P_g'	'Q_{gen}'	'P_{gen}'	'T_t'	'T_t'	'T_t'	'α'
'β'	'β'	'Q_{source}'	'T_t'	'T_{gen}'	'T_{gen}'	'P_{gen}'	'P_t'	'P_{gen}'
'Q_{gen}'	'T_{gen}'	'Q_1'	'P_t'	'P_2'	'P_t'	'P_t'	'v_w'	'T_{gen}'
'T_{gen}'	'P_t'	'Q_2'	'P_3'	'P_3'	'ω'	'v_w'	'P_2'	'ω'
'P_t'	'v_w'	'Q_g'	'Q_2'	'Q_{source}'	'v_w'	'P_{source}'	'Q_{source}'	'P_{source}'
'P_2'	'P_3'	'V_{a2}'	'Q_3'	'Q_3'	'P_{source}'	'P_2'	'Q_3'	'P_1'
'P_3'	'Q_{source}'	'V_{b2}'	'V_{c2}'	'V_{b1}'	'P_1'	'P_3'	'V_{a1}'	'P_3'
'Q_g'	'Q_3'	'V_{c2}'	'V_{a3}'	'V_{a2}'	'P_2'	'Q_1'	'V_{b1}'	'P_g'
'V_{b1}'	'Q_g'	'V_{a3}'	'V_{b3}'	'V_{b2}'	'P_g'	'Q_3'	'V_{c1}'	'Q_{source}'
'V_{c1}'	'V_{a1}'	'V_{b3}'	'I_{a1}'	'I_{a1}'	'Q_{source}'	'V_{b1}'	'V_{a2}'	'Q_1'
'V_{a2}'	'V_{b1}'	'V_{c3}'	'I_{b1}'	'I_{b1}'	'Q_1'	'V_{c1}'	'V_{b3}'	'Q_2'
'V_{b2}'	'V_{c1}'	'I_{a1}'	'I_{a2}'	'I_{a2}'	'Q_3'	'V_{b2}'	'I_{b1}'	'Q_g'
'V_{c2}'	'V_{a2}'	'I_{b1}'	'I_{b2}'	'I_{b2}'	'Q_g'	'V_{a3}'	'I_{c1}'	'V_{a2}'
'I_{b2}'	'V_{c2}'	'I_{c1}'	'I_{c2}'	'I_{c2}'	'I_{a1}'	'V_{c3}'	'I_{b2}'	'V_{b2}'
'I_{a3}'	'I_{a3}'	'I_{a2}'	'I_{a3}'	'I_{a3}'	'I_{b1}'	'I_{b1}'	'I_{c2}'	'V_{c2}'
'I_{b3}'	'I_{b3}'	'I_{b2}'	'I_{b3}'	'I_{b3}'	'I_{c1}'	'I_{a2}'	'I_{a3}'	'V_{b3}'
'I_{c3}'	'I_{c3}'	'I_{c2}'	'I_{c3}'	'I_{c3}'	'I_{a2}'	'I_{a3}'	'I_{b3}'	'V_{c3}'
'I_{ga}'	'I_{ga}'	'I_{ga}'	'I_{ga}'	'I_{ga}'	'I_{b2}'	'I_{b3}'	'I_{c3}'	'I_{a1}'
'I_{gb}'	'I_{gb}'	'I_{gb}'	'I_{gb}'	'I_{gb}'	'I_{c2}'	'I_{c3}'	'I_{ga}'	'I_{c1}'
'Igc'	'I_{gc}'	'I_{gc}'	'I_{gc}'	'I_{gc}'	'I_{a3}'	'I_{gb}'	'I_{gc}'	'I_{a3}'

Table B.1- Retained variables from PMSG simulation data in time, frequency and instantaneous frequency domains using B2, B4 and H selection methods

Time			Frequency			Instantaneous frequency		
B2	*B4*	*H*	*B2*	*B4*	*H*	*B2*	*B4*	*H*
'ωref'	'Vw'	'Vw'	'Telec'	'ωmech'	'Vw'	'Vw'	'Vw'	'ωmech'
'ωmech'	'ωref'	'ωref'	's'	'Telec'	'ωref'	'ωref'	'ωref'	'Telec'
'Tmech'	'Tmech'	'ωmech'	'Idc1'	'Idc1'	'ωmech'	'ωmech'	'ωmech'	's'
'Telec'	'Telec'	'Tmech'	'Idc2'	'Idc2'	'Tmech'	'Tmech'	'Tmech'	'Idc2'
'Idc1'	'Idc1'	'Telec'	'Pdc1'	'Vdc'	's'	'Telec'	'Telec'	'Vdc'
'Idc2'	'Idc2'	's'	'Pdc2'	'Pdc1'	'Vdc'	'Idc1'	'Idc1'	'Pdc1'
'Vdc'	'Vdc'	'Idc2'	'Pt'	'Pdc2'	'Pdc2'	'Pdc2'	'Vdc'	'Ps'
'Pg'	'Pdc2'	'Pt'	'Pg'	'Pt'	'Pg'	'Pt'	'Pdc2'	'Pr'
'Pr'	'Pt'	'Pg'	'Ps'	'Pg'	'Ps'	'Pg'	'Pt'	'Qg'
'Qg'	'Pg'	'Ps'	'Qg'	'Ps'	'Qg'	'Pr'	'Qg'	'Qs'
'Qr'	'Ps'	'Vrc_rms'	'Qr'	'Qr'	'Qs'	'Vsb_rms'	'Vsb_rms'	'Qr'
'Vsa_rms'	'Qr'	'Isa_rms'	'Vsa_rms'	'Vsa_rms'	'Isa_rms'	'Vsc_rms'	'Vsc_rms'	'Vsa_rms'
'Vsb_rms'	'Vsa_rms'	'Isb_rms'	'Vra_rms'	'Vsb_rms'	'Isb_rms'	'Vrc_rms'	'Vga_rms'	'Vsb_rms'
'Vra_rms'	'Vsb_rms'	'Isc_rms'	'Vrb_rms'	'Vra_rms'	'Isc_rms'	'Vga_rms'	'Isa_rms'	'Vsc_rms'
'Vrb_rms'	'Vrb_rms'	'Ira_rms'	'Vrc_rms'	'Vrb_rms'	'Ira_rms'	'Isa_rms'	'Isb_rms'	'Vra_rms'
'Vrc_rms'	'Vrc_rms'	'Irb_rms'	'Vgb_rms'	'Vrc_rms'	'Irb_rms'	'Isb_rms'	'Isc_rms'	'Vrb_rms'
'Vgc_rms'	'Vgc_rms'	'Irc_rms'	'Vgc_rms'	'Vgb_rms'	'Irc_rms'	'Isc_rms'	'Ira_rms'	'Vrc_rms'
'Ira_rms'	'Ira_rms'	'Iga_rms'	'Isa_rms'	'Vgc_rms'	'Iga_rms'	'Ira_rms'	'Irb_rms'	'Vga_rms'
'Irc_rms'	'Irb_rms'	'Igb_rms'	'Isb_rms'	'Isb_rms'	'Igb_rms'	'Irb_rms'	'Irc_rms'	'Vgb_rms'
'Igc_rms'	'Irc_rms'	'Igc_rms'	'Isc_rms'	'Isc_rms'	'Igc_rms'	'Irc_rms'	'Igb_rms'	'Vgc_rms'

Table B.2- Retained variables from DFIG simulation data in time, frequency and instantaneous frequency domains using B2, B4 and H selection methods

B.3 Retained variable sets using T selection algorithms

The retained variables from targeted selection methods PMSG, DFIG and SCADA data with various types of fault are show in Table B.3 and B4 respectively. The targeting signal used during the selection is also listed in the tables.

Type of data	Capacitor ageing fault	Single phase-to-ground fault	Phase-to-phase fault	Three phase-to-ground fault
Target signal	*DC-link voltage*	*Active power at PCC*	*Active power at PCC*	*Active power at PCC*
	'I_{dc1}'	'I_{dc1}'	'V_{dc}'	'I_{dc1}'
	'Q_{grid}'	'I_{dc2}'	'Q_{grid}'	'V_{dc}'
	'P_{grid}'	'Q_{grid}'	'α'	'I_{dc2}'
	'I_{dc_ref}'	'I_{dc_ref}'	'β'	'Q_{grid}'
	'α'	'α'	'Q_{gen}'	'I_{dc_ref}'
	'β'	'β'	'T_t'	'α'
	'T_t'	'Q_{gen}'	'T_{gen}'	'β'
	'P_{gen}'	'T_t'	'P_t'	'Q_{gen}'
	'T_{gen}'	'P_t'	'v_w'	'T_t'
	'P_t'	'ω'	'P_{source}'	'P_{gen}'
	'v_w'	'v_w'	'P_g'	'T_{gen}'
	'P_{source}'	'P_{source}'	'Q_{source}'	'P_t'
	'P_3'	'P_3'	'Q_2'	'ω'
	'Q_{source}'	'Q_3'	'Q_g'	'P_3'
	'Q_3'	'V_{a1}'	'V_{a1}'	'V_{b1}'
	'V_{a1}'	'V_{b1}'	'V_{b1}'	'V_{c1}'
	'V_{b1}'	'V_{c1}'	'V_{a3}'	'V_{a2}'
	'V_{b2}'	'V_{c2}'	'V_{b3}'	'V_{b3}'
	'I_{b2}'	'V_{a3}'	'I_{c2}'	'I_{c1}'
	'I_{a3}'	'I_{a2}'	'I_{a3}'	'I_{a2}'
	'I_{b3}'	'I_{b2}'	'I_{b3}'	'I_{a3}'
	'I_{c3}'	'I_{a3}'	'I_{c3}'	'I_{b3}'
	'I_{ga}'	'I_{b3}'	'I_{ga}'	'I_{c3}'
	'I_{gb}'	'I_{c3}'	'I_{gb}'	'I_{ga}'
	'I_{gc}'	'I_{ga}'	'I_{gc}'	'I_{gb}'

Table B.3- Retained variables from PMSG simulation data with various types of fault in time domain using targeted selection method

Type of data	Capacitor ageing fault	Single phase-to-ground fault	Phase-to-phase fault	Three phase-to-ground fault
Target signal	*DC-link voltage*	*Active power at PCC*	*Active power at PCC*	*Active power at PCC*
	'V_w'	'V_w'	'V_w'	'V_w'
	'ω_{ref}'	'ω_{ref}'	'T_{mech}'	'ω_{ref}'
	'ω_{mech}'	'T_{mech}'	'T_{elec}'	'T_{mech}'
	'T_{mech}'	'T_{elec}'	'I_{dc1}'	'T_{elec}'
	'T_{elec}'	'I_{dc1}'	'I_{dc2}'	'I_{dc1}'
	'I_{dc1}'	'I_{dc2}'	'V_{dc}'	'I_{dc2}'
	'I_{dc2}'	'V_{dc}'	'P_{dc2}'	'V_{dc}'
	'P_{dc2}'	'P_{dc2}'	'P_g'	'P_{dc2}'
	'P_g'	'P_g'	'Q_r'	'P_g'
	'P_s'	'P_r'	'V_{sb_rms}'	'P_s'
	'Q_g'	'Q_r'	'V_{sc_rms}'	'Q_g'
	'Q_s'	'V_{ra_rms}'	'V_{ra_rms}'	'Q_r'
	'Q_r'	'V_{rb_rms}'	'V_{rb_rms}'	'V_{sa_rms}'
	'V_{sb_rms}'	'V_{rc_rms}'	'V_{rc_rms}'	'V_{ra_rms}'
	'V_{ra_rms}'	'V_{gb_rms}'	'V_{gb_rms}'	'V_{rc_rms}'
	'V_{rb_rms}'	'V_{gc_rms}'	'I_{sb_rms}'	'V_{gb_rms}'
	'V_{rc_rms}'	'I_{ra_rms}'	'I_{ra_rms}'	'V_{gc_rms}'
	'V_{gc_rms}'	'I_{rb_rms}'	'I_{rb_rms}'	'I_{sa_rms}'
	'I_{ra_rms}'	'I_{rc_rms}'	'I_{rc_rms}'	'I_{ra_rms}'
	'I_{rc_rms}'	'I_{ga_rms}'	'I_{gc_rms}'	'I_{rb_rms}'

Table B.4- Retained variables from DFIG simulation data with various types of fault in time domain using targeted selection method