



Wind turbine drivetrain condition monitoring
through SCADA data on farm level

Jelle Muller

Restricted

**Wind turbine drivetrains condition monitoring through SCADA data on
farm level
Master thesis report**

For obtaining the degree of Master of Science in:
Offshore and Dredging Engineering at the Delft University of Technology
&
Technology-Wind Energy at the Norwegian University of Science and Technology

As organized by:
The Offshore Engineering track of the European Wind Energy Master program

Jelle Peter Muller

Comittee: Prof. Dr. Simon Watson
Dr. Amir Nejad
Dr. Ir. Jasper Kreeft
Dr. Ir. Eliz-Mari Lourens

Chairman TU Delft
NTNU
Shell
TU Delft

Acknowledgements

After completing my master thesis, I can truly say that I enjoyed to process of it and have grown during this period. Many people are to reason for this and I would like to use this moment to thank them.

My profound gratitude is extended to my supervisor in Shell, Jasper Kreeft, for providing me with the opportunity to carry out my master thesis at Shell and for supporting me with much energy and enthusiasm. I would like to thank both Jasper and Erik for giving me the possibility to explore the different areas of the wind industry, both related and unrelated to my work.

I would like to acknowledge my entire committee: Simon Watson, Amir Nejad, Eliz-Mari Lourens and Jasper Kreeft for giving me the opportunity to work on my thesis, guiding and steering me through the process but also giving me the freedom to move freely. I thank Simon for his thought-provoking remarks and comments, Eliz-Mari for her expertise and support in areas that were new to me and Amir especially for his steering in my literature research during my time at NTNU.

Finally, I would like to thank my family and friends for their unwavering support and interest in my thesis. I would especially like to thank my toughest peer, greatest supporter and lifeline, Ilsa Maria. Finally I would like to thank my friends from EWEM: Wichert, Sara, Matthijs, Sanne and Johan for the journey we have been on the last 2 years.

Jelle Muller
Delft
July 8th 2019

Abstract

The offshore wind industry has grown rapidly over the last decade and drivetrains are increasing in size to reduce the cost of energy. These turbines are operating in a harsh environment. Adopting a preventive maintenance strategy is important to achieve an as high as possible availability of the farm and reduce the cost of maintenance. A well performing condition monitoring system that utilizes SCADA data from the wind farm can enable this strategy without the need in additional cost in hardware.

This master thesis focusses on the development of a framework that can be utilized for this task. This framework can process raw operational SCADA data collected at the Egmond aan Zee offshore wind farm to create a clean dataset to train supervised machine learning models on. This work provides an insight in the correlation between different SCADA signals using a mathematical approach and from a understanding of the system integration of drivetrain components. Bearing temperatures are modelled using a data driven approach to describe the temperatures under healthy conditions. Several models are evaluated for this task and it was concluded that a decision tree supervised machine learning regression model resulted in the lowest error between predicted and measured values. Anomalies are detected and tracked with a normal behaviour model and a Shewart and CUSUM control chart that are applied on the residual error between modelled and measured temperature signals.

4 anomalies could be identified in the gearbox bearings using the developed framework. Abnormal behaviour of the drivetrain could be identified as early as 1 month before the turbine was taken out of productions. This highlights that temperature based condition monitoring that utilizes SCADA data can be used for early detection of faults by combining the accuracy of supervised machine learning methods with different fault detection methods like the CUSUM control chart.

This work also investigates the relation between experienced wake of a wind turbine and the influence on the drivetrain component temperatures. The wake conditions at Egmond aan Zee, modelled with an Ishahara wake model, and the component temperature measurements from the SCADA data are used for this analysis. The bearing temperature distributions under different operational and wake conditions can be compared by clustering over the wind speed and the velocity deficit or turbulence intensity at turbine level. It is concluded from this work that wake effects do not result in a change in drivetrain component temperatures.

The effects of asymmetric wake conditions opposed to wake experienced over the entire rotor is analysed by comparing the temperature distributions under these conditions in a cluster where the turbine is partially waked. A small shift towards higher component temperatures can observed on a limited amount of data for turbines under asymmetric loading conditions.

Table of contents

Acknowledgments	II
Abstract	III
Table of contents	IV
List of Figures	VII
List of Tables	XI
Nomenclature	XII
Abbreviations	XII
Equation Variables	XII
1 Introduction	1
1.1 Industry and technical overview	1
1.1.1 Offshore wind industry	1
1.1.2 Wind turbine drivetrains	3
1.1.3 Condition monitoring	4
1.2 Research motivation	6
1.3 Thesis objective	7
1.4 Thesis outline	7
2 Background	9
2.1 Case Study	9
2.1.1 Farm placement and foundation specification	9
2.1.2 Environmental conditions	10
2.1.3 Vestas V90 sepecification	11
2.2 SCADA system	12
2.3 Introduction to machine learning	13
2.3.1 Types of machine learning	13
2.3.2 Consideration around fitting	14
2.3.3 Empirical risk and error function	15
2.3.4 K-fold cross validation	16
3 Data proceedings	17
3.1 Data collection	17
3.1.1 SCADA system	17
3.1.2 Ambient condition data farm level	18
3.2 Data filtering	19
3.2.1 Time stamp matching	19
3.2.2 General filtering	19
3.2.3 Missing data filtering	20
3.2.4 Mahalanobis filtering	20
3.2.5 Filtering result	22
3.3 Correlation of parameters	23
3.3.1 Mathematical correlation	23
3.3.2 Physical correlation and system integration	23

- 3.4 Training period and parameter selection 26
 - 3.4.1 Period selection 26
 - 3.4.2 Parameter selection 27
 - 3.4.3 Training dataset evaluation 28
 - 3.4.4 Test dataset evaluation 31
- 4 Data driven models 34**
 - 4.1 Farm average signal trending 34
 - 4.2 Thermal difference analysis 35
 - 4.3 Normal behaviour models 37
 - 4.3.1 Simple regression models 37
 - 4.3.2 Decision Trees 39
 - 4.3.3 Support Vector Machine 42
 - 4.4 Fault detection methods 46
 - 4.4.1 Shewhart control chart 46
 - 4.4.2 Cusum control chart 47
 - 4.5 Model evaluation 49
 - 4.5.1 Model accuracy 49
 - 4.5.2 Model speed 52
- 5 Condition monitoring results 54**
 - 5.1 Thermal difference results 54
 - 5.2 Normal Behaviour models under healthy conditions 55
 - 5.3 Anomaly detection 61
 - 5.3.1 Case turbine G 61
 - 5.3.2 Case turbine H 63
 - 5.3.3 Case turbine I 64
 - 5.3.4 Case Turbine J 65
 - 5.3.5 Case turbine K 66
 - 5.4 Discussion on the condition monitoring framework 67
- 6 Wake effects on component temperatures 69**
 - 6.1 Approach 69
 - 6.2 Ishihara wake model 71
 - 6.2.1 Model input parameters 71
 - 6.2.2 Modelled wake conditions 73
 - 6.2.3 OWEZ wake simulations 75
 - 6.3 Wake temperature distribution results 76
 - 6.3.1 Results turbulence intensity cluster 76
 - 6.3.2 Results velocity deficit cluster 77
 - 6.3.3 Results asymmetric loading 79
 - 6.4 Discussion on wake effects on component temperatures 80
- 7 Conclusions and Discussion 82**
 - 7.1 Conclusions 82
 - 7.2 Discussion and recommendations 84
- References 87**

Appendix	91
A Correlation coefficient with respect to Hollow Shaft bearing 1 temperature	91
B Correlation coefficient with respect to High Speed Shaft bearing 1 temperature	92
C Correlation coefficient with respect to generator bearing 1 temperature . .	93
D Correlation coefficient with respect to slipping temperature	94
E Cooling water circuit	95
F Turbine A, training dataset	96
G Turbine A, test dataset	102
H Decision criteria for decision trees	108
I Time series results linear regression model	109
J Time series results 6th order polynomial regression model	110
K Time series results decision tree regression model	111
L Time series results decision tree classification model	111
M Time series results SVM regression model	112
N Time series results SVM classification model	112
O Results error distributions	113
P Cluster Iw/Ia statistical properties	117
Q Cluster Uw/Uinf statistical properties	122

List of Figures

1.1	Cumulative and annual offshore wind energy installation in Europe [1] . . .	1
1.2	Trend of turbine size developments [3]	2
1.3	Expert estimates of median-scenario LCOE [2]	2
1.4	Drivetrain of a geared horizontal axis wind turbine [9]	3
1.5	Indication of failures and response time [9]	5
2.1	OWEZ geographical location and lay out [3]	9
2.2	Annual wind conditions at hub height for the OWEZ [24]	10
2.3	Annual wave conditions for the OWEZ location [24]	11
2.4	Vestas V90 Nacelle [25]	12
2.5	Concept of appropriate fitting of a model	15
2.6	Concept of k-fold cross validation	16
3.1	Vestas V90 SCADA signal sensor locations	17
3.2	Application of the general filtering	19
3.3	Top figure: before application of missing data filter. Bottom figure: after application of missing data filter	20
3.4	Mahalanobis distance distribution per cluster and threshold value	21
3.5	Application of the Mahalanobis distance filtering	22
3.6	Filtering of the training data	22
3.7	Pearson correlation coefficient with respect to delivered grid power	24
3.8	Temperature sensor locations in the gearbox [25]	25
3.9	Temperature sensor locations in the generator [25]	26
3.10	Turbine A in-sample dataset, observations of input conditions	28
3.11	Turbine A in-sample dataset, observations of response conditions	29
3.12	Turbine A in-sample dataset, High speed shaft bearing 1 in relation to the power curve	30
3.13	Turbine A in-sample dataset, High speed shaft bearing 1 in relation to the generator speed and return oil temperature	30
3.14	Turbine A out-of-sample dataset, observations of the input conditions	31
3.15	Turbine A out-of-sample dataset, observations of response conditions	32
3.16	Turbine A out-of-sample dataset, High speed shaft bearing 1 in relation to the power curve	32
3.17	Turbine A out-of-sample dataset, High speed shaft bearing 1 in relation to the generator speed and return oil temperature	33
4.1	Farm average signal trending for the high speed shaft bear 1 temperatures	34
4.2	Parameters for the thermal difference analyses in relation to each other	35
4.3	Thermal difference between generator windings in healthy condition	36
4.4	Normal behaviour model structure	37
4.5	Regression functions fitted to the high speed shaft bearing 1 temperatures	38
4.6	Concept of decision trees data splitting process	39
4.7	Concept of a complexity-cost function for the size of a decision tree	40
4.8	Hyperparameter optimization for decision trees	41
4.9	First 8 nodes and 9 children of the generated decision tree	41
4.10	Overview of the generated decision tree	42
4.11	Concept of support vector machine data separation	42
4.12	Concept of a support vector machine including impurities	43
4.13	Concept of using kernel functions to map in an enhance feature space	43

4.14 validation error of SVM regression with different kernel functions 44

4.15 concept of ϵ insensitive SVM regression 45

4.16 Hyperparameter ϵ sensitivity 46

4.17 Shewart control chart 46

4.18 Concept of 1 and 2 sided CUSUM control charts 48

4.19 Top graph: time series for \hat{y} and y . Middle graph: the Sherhard controlchart on e . Bottom graph: the CUSUM control chart 49

4.20 Distribution of the validation error for predicting the hollow shaft bearing 1 50

4.21 Distribution of the validation error for predicting the high speed shaft bearing 2 50

4.22 Decision tree classification generalization gearbox bearing 52

4.23 Decision tree classification generalization generator bearing 52

5.1 Temperature difference between generator phase windings, turbine E . . . 54

5.2 Temperature difference between generator phase windings, turbine F . . . 55

5.3 Test error distribution for Hollow Shaft Bearing 1 temperatures, turbine A 56

5.4 Out-of-sample time series for hollow shaft bearing 2, turbine A using a decision tree regression model 57

5.5 Test error distribution for the high speed shaft bearing 2 temperatures, turbine A 58

5.6 Out-of-sample time series for generator bearing 1,turbine A using a decision tree regression model 59

5.7 Turbine A training dataset, Generator bearing 2 in relation to the power curve 59

5.8 Turbine A training dataset, Generator bearing 2 in relation to the power curve for a small operational window 60

5.9 Test dataset turbine A, generator bearing 2 validation error for different size training sets 60

5.10 Time series result Turbine G: hollow shaft bearing 2 temperature 62

5.11 Time series result Turbine G: hollow shaft bearing 4 temperature 62

5.12 Time series result Turbine H: hollow shaft bearing 1 temperature 63

5.13 Time series result Turbine I: hollow shaft bearing 1 temperature 64

5.14 Time series result Turbine J: hollow shaft bearing 1 temperature 65

5.15 Time series result Turbine K: High Speed shaft bearing 1 temperature . . 66

5.16 Time series result Turbine K: hollow shaft bearing 2 temperature 67

6.1 2 wake conditions for partially waked conditions 70

6.2 Turbulence intensity vs. wind speed and fitter power law 72

6.3 Vestas V90 Thrust coefficient curve C_T 72

6.4 Gaussian velocity deficit distribution [49] 73

6.5 Gaussian added turbulence intensity distribution [49] 75

6.6 Simulated velocity deficit and turbulence intensity at OWEZ 75

6.7 Temperature distribution over different wake conditions, by normalized turbulence intensity 76

6.8 Temperature distribution over different wake conditions, by normalized velocity deficit 78

6.9 Temperature distributions for the hollow shaft bearing under wake condition 1 and 2 clustered over wind speed: 5.5-7.5 [m/s] & $\frac{\Delta U}{U_{infy}}$: 0.07-0.15 [-] 79

6.10	Temperature distributions for the hollow shaft bearing under wake condition 1 and 2 clustered over wind speed: 7.5-9.5 [m/s] & $\frac{\Delta U}{U_{infty}}$: 0.07-0.15 [-]	80
A.1	Correlation coefficient with respect to Hollow Shaft bearing 1	91
B.2	Correlation coefficient with respect to High Speed Shaft bearing 1 temperature	92
C.3	Correlation coefficient with respect to generator bearing 1	93
D.4	Correlation coefficient with respect to slipping temperature	94
E.5	Two cooling water circuit for drive train components [25]	95
F.6	Turbine A in-sample dataset, Hollow shaft bearing 1 in relation to the power curve	96
F.7	Turbine A in-sample dataset, Hollow shaft bearing 2 in relation to the power curve	96
F.8	Turbine A in-sample dataset, High Speed shaft bearing 1 in relation to the power curve	97
F.9	Turbine A in-sample dataset, High Speed shaft bearing 2 in relation to the power curve	97
F.10	Turbine A in-sample dataset, High Speed shaft bearing 3 in relation to the power curve	98
F.11	Turbine A in-sample dataset, Generator bearing 1 in relation to the power curve	98
F.12	Turbine A in-sample dataset, Generator bearing 2 in relation to the power curve	99
F.13	Turbine A in-sample dataset, Hollow shaft bearing 1 in relation to the generator speed and return oil temperature	99
F.14	Turbine A in-sample dataset, Hollow shaft bearing 2 in relation to the generator speed and return oil temperature	100
F.15	Turbine A in-sample dataset, High Speed shaft bearing 1 in relation to the generator speed and return oil temperature	100
F.16	Turbine A in-sample dataset, High Speed shaft bearing 2 in relation to the generator speed and return oil temperature	101
F.17	Turbine A in-sample dataset, High Speed shaft bearing 3 in relation to the generator speed and return oil temperature	101
G.18	Turbine A out-of-sample dataset, Hollow shaft bearing 1 in relation to the power curve	102
G.19	Turbine A out-of-sample dataset, Hollow shaft bearing 2 in relation to the power curve	102
G.20	Turbine A out-of-sample dataset, High Speed shaft bearing 1 in relation to the power curve	103
G.21	Turbine A out-of-sample dataset, High Speed shaft bearing 2 in relation to the power curve	103
G.22	Turbine A out-of-sample dataset, High Speed shaft bearing 3 in relation to the power curve	104
G.23	Turbine A out-of-sample dataset, Generator bearing 1 in relation to the power curve	104
G.24	Turbine A out-of-sample dataset, Generator bearing 2 in relation to the power curve	105

G.25 Turbine A out-of-sample dataset, Hollow shaft bearing 1 in relation to the generator speed and return oil temperature 105

G.26 Turbine A out-of-sample dataset, Hollow shaft bearing 2 in relation to the generator speed and return oil temperature 106

G.27 Turbine A out-of-sample dataset, High Speed shaft bearing 1 in relation to the generator speed and return oil temperature 106

G.28 Turbine A out-of-sample dataset, High Speed shaft bearing 2 in relation to the generator speed and return oil temperature 107

G.29 Turbine A out-of-sample dataset, High Speed shaft bearing 3 in relation to the generator speed and return oil temperature 107

H.30 Data splitting criteria for decision trees [40] 108

I.31 In-sample time series for hollow shaft bearing 1, turbine A using linear regression model 109

I.32 Out-of-sample time series for hollow shaft bearing 1, turbine A using linear regression model 109

J.33 In-sample time series for hollow shaft bearing 1, turbine A using linear regression model 110

J.34 Out-of-sample time series for hollow shaft bearing 1, turbine A using a 6th order polynomial regression model 110

K.35 Out-of-sample time series for hollow shaft bearing 1, turbine A using decision tree regression model 111

L.36 Out-of-sample time series for hollow shaft bearing 1, turbine A using decision tree classification model 111

M.37 Out-of-sample time series for hollow shaft bearing 1, turbine A using SVM regression model 112

N.38 Out-of-sample time series for hollow shaft bearing 1, turbine A using SVM classification model 112

O.39 Model test error distribution for Hollow Shaft Bearing 1 temperatures, turbine A 113

O.40 Model test error distribution for Hollow Shaft Bearing 2 temperatures, turbine A 113

O.41 Model test error distribution for High Speed Shaft Bearing 1 temperatures, turbine A 114

O.42 Model test error distribution for High Speed Shaft Bearing 2 temperatures, turbine A 114

O.43 Model test error distribution for High Speed Shaft Bearing 3 temperatures, turbine A 115

O.44 Model test error distribution for generator bearing 1 temperatures, turbine A 115

O.45 Model test error distribution for generator bearing 2 temperatures, turbine A 116

List of Tables

3.1	Filter clusters	21
3.2	Generic input and response parameters for the condition monitoring framework	28
4.1	RMSE in °C for different order polynomials and different response signals	38
4.2	Validation error distribution	44
4.3	minimum RMSE in °C after training the model 100 times on the in-sample data	51
4.4	Required run time to train different models	53
5.1	The validation error in °C over the in-sample data and test error in °C over the out-of sample data using different machine learning models . . .	57
6.1	Wind speed clusters	69
6.2	Wind speed clusters	70
6.3	Mean second high speed shaft bearing temperature of clusters from turbine A	77
6.4	Standard deviations second high speed shaft bearing temperature of clusters from turbine A	77
6.5	Mean second high speed shaft bearing temperature of clusters from turbine A	78
6.6	Standard deviations second high speed shaft bearing temperature of clusters from turbine A	78
P.1	Mean bearing temperature of clusters	117
P.2	Number of datapoints per clusters	117
P.3	Number of datapoints per clusters	117
P.4	Mean bearing temperature of clusters	118
P.5	Number of datapoints per clusters	118
P.6	Number of datapoints per clusters	118
P.7	Mean bearing temperature of clusters	119
P.8	Standard deviations bearing temperature of clusters	119
P.9	Number of datapoints per clusters	119
P.10	Mean bearing temperature of clusters	120
P.11	Standard deviations bearing temperature of clusters	120
P.12	Number of datapoints per clusters	120
P.13	Mean bearing temperature of clusters	121
P.14	Standard deviations bearing temperature of clusters	121
P.15	Number of datapoints per clusters	121
Q.16	Mean bearing temperature of clusters	122
Q.17	Standard deviations bearing temperature of clusters	122
Q.18	Number of datapoints per clusters	122
Q.19	Mean bearing temperature of clusters	123
Q.20	Standard deviations bearing temperature of clusters	123
Q.21	Number of datapoints per clusters	123

Nomenclature

Abbreviations

AID	Automatic Interaction Detection
ANN	Artificial neural networks
ARX	Auto regressive model with exogenous input
CUSUM	Cumulative sum
DOWA	Dutch Offshore Wind Atlas
FSRC	Full signal reconstruction
LCOE	Levelized cost of energy
MSL	Mean sea level
O&M	Operation and maintenance
OWEZ	Offshore Windfarm Egmond aan Zee
RMSE	Root mean square error
RPM	Rotations per minute
SCADA	Supervisory control and data acquisition
SVM	Support Vector Machines
WTG	Wind turbine

Equation Variables

ξ	Slack variable for impurities when separating data using SVM
β	Impurity threshold for stopping rule
ΔU	Wind velocity deficit due to wake
ΔU	Wind velocity deficit
ℓ	Loss function
$\hat{R}_n(f)$	Empirical risk function
\hat{y}	Predicted Output
\mathcal{L}	Empirical Risk function
μ	Mean value
μ_0	Target mean of the dataset
ϕ	Spanwise function for velocity deficit distribution

$\rho_{X,Y}$	Pearson correlation coefficient
σ	Standard deviation
ε	Insensitivity hyperparameter for SM regression models
A	Data subset of D
B	Data subset of D
c	Decision criteria for decision tree splitting
C^+	CUSUM parameter above target mean
C^-	CUSUM parameter below target mean
C_T	Thrust coefficient
D	Complete dataset
e	Residual error
$G(x, y)$	Kernel function
H	CUSUM control limit
h	CUSUM control limit parameter
I	Number of impurities in a dataset
I_+	Added turbulence intensity due to wake
I_a	Ambient turbulence intensity
I_w	Turbulence intensity at turbine level
K	Allowance or slack value for the CUSUM control chart
MHD_i	Mahalanobis distance of datapoint i to the cluster mean
n	Order of polynomial
P	Power output
$P(x, y)$	Probability distribution of data x and y
P_{rated}	Rated power output
Q_1	25 th percentile value
Q_3	75 th percentile value
r	rotor radius
$R(f)$	True risk function
t	Pruned back decision tree where $t \in T_{max}$
T_{max}	Maximum size of trained decision tree

u	Input, time dependent
U_{∞}	Wind velocity free stream
U_{cutin}	Cut in wind velocity
U_{cutout}	Cut out wind velocity
U_h	Wind velocity at the hub
WA	Area of the rotor that is experiencing waked conditions
WA_a	Wake covered area of turbine a
y	Measured output from the SCADA system

1 Introduction

In this chapter, first an overview will be provided of the most important background and will be followed by the research motivation and research objective. The last part of this chapter will be dedicated to the outline of the report.

1.1 Industry and technical overview

To understand the research motivation, it is important to first get an insight in the offshore wind industry, the turbine and the common practices around condition monitoring of the turbine's drivetrain.

1.1.1 Offshore wind industry

The offshore wind energy industry is maturing in a full grown market that operates in a merchant environment. The installed capacity of offshore wind energy in Europe has grown rapidly over recent years with a record year in 2017. In this year, 17 new offshore wind farms were taken into operation, resulting in an additional 3,148 MW of installed capacity. With several large wind farms in the pipeline, this record is expected to only hold until this year. This trend is clearly visible in figure 1.1. The increase in offshore capacity is predicted to follow the current trend with a projected offshore capacity in Europe of 25 GW by 2020 [1].

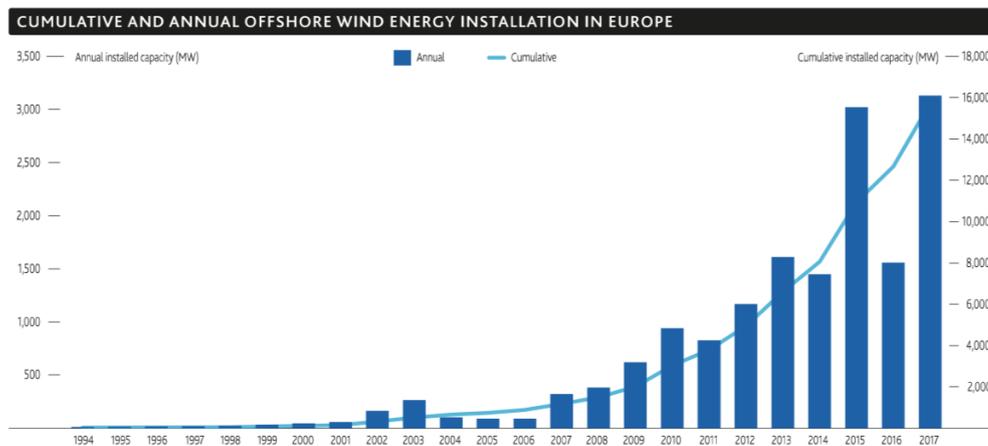


Figure 1.1: Cumulative and annual offshore wind energy installation in Europe [1]

Not only the number of wind farms has grown over the last decade, also the average size of the turbines is increasing exponentially, as can be seen in figure 1.2. In 2017, the average size of a new installed offshore wind turbine drivetrain was 5.9 MW, which represents a 26% increase compared to the previous year. This growth is expected to continue. An example of this is the 12MW prototype that GE renewable energy is testing this year, to be available for installation by 2021. Manufactures constantly push the boundaries to realise larger wind turbine drivetrains. By doing so, the levelized cost of energy (LCOE) is reduced as a smaller number of foundations and installation steps are

needed for the same capacity of an offshore wind farm [2].

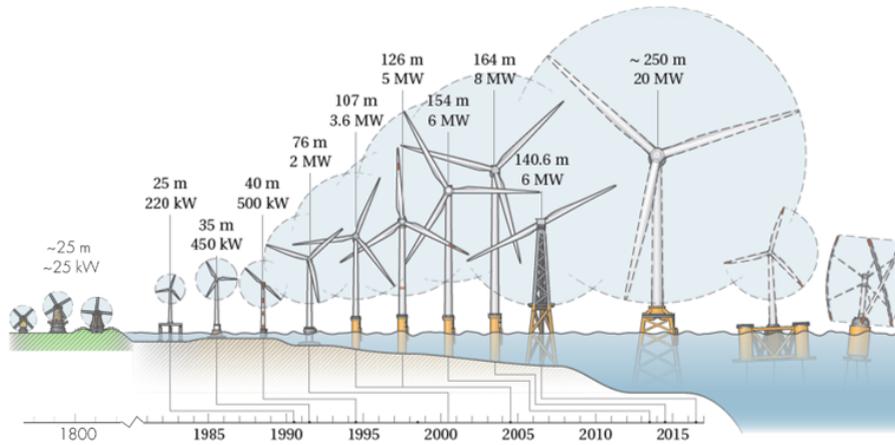


Figure 1.2: Trend of turbine size developments [3]

It is important to note that the reduction of cost for offshore wind energy in more than the experts predicted in 2014, as shown in figure 1.3 [2]. Prove of this, are the tenders that were awarded in 2016 for Borssele I and II, Danish Kriegers Flak project and Borssele III and IV, these tenders will have a strike LCOE of €72.7/MWh, €49.9/MWh and €54.5/MWh respectively. This last named tender is won by a Shell-led consortium. The cost of energy for these wind farms are much lower than the experts predicted only two years earlier and are directly competitive with onshore wind farms. The Hollandse Kust Zuid I and II tender was awarded in 2018 without the need of any government subsidies and in 2019 the Dunkirk tender was awarded with a strike LCOE of €44.7/MWh to highlight the further decreasing cost of offshore wind energy.

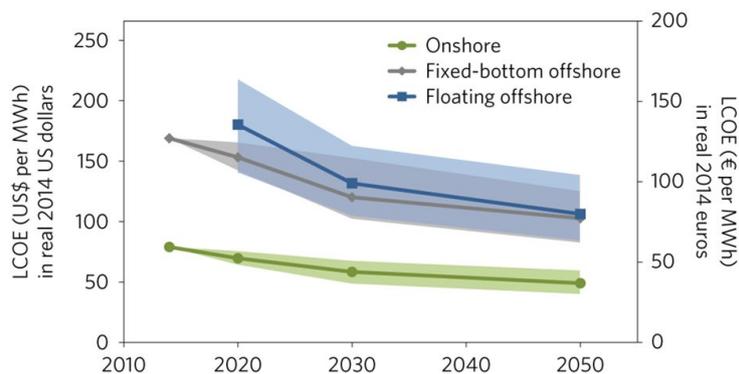


Figure 1.3: Expert estimates of median-scenario LCOE [2]

With the move to offshore and larger drivetrains, multiple challenges arise. The operation and maintenance (O&M) of offshore wind farms can be characterised as; hard to access, low availability of vessels and high cost [4; 5; 6]. These challenges become

more important as the offshore industry tends to move further from shore in harsher environmental conditions [7; 8]. With the increasing size of the drivetrain capacity, components become more expensive and the consequences of downtime of a large capacity turbines will have a large impact on the wind farm efficiency and the lost income.

1.1.2 Wind turbine drivetrains

As this thesis focusses on the condition monitoring of wind turbine drivetrains, it is important to provide an overview of the components that the drivetrain is made up from and which parts are taken into account for this research.

A modern horizontal axis wind turbine consist of the following components: Sub-structure, tower, nacelle and rotor. The drivetrain refers to all the components needed for the energy conversion system and is located in the nacelle. This includes the main shaft, brake, high speed shaft and generator and, if the wind turbine is not a direct drive turbine, this also includes a gearbox. A typical geared wind turbine drivetrain is displayed in figure 1.4.

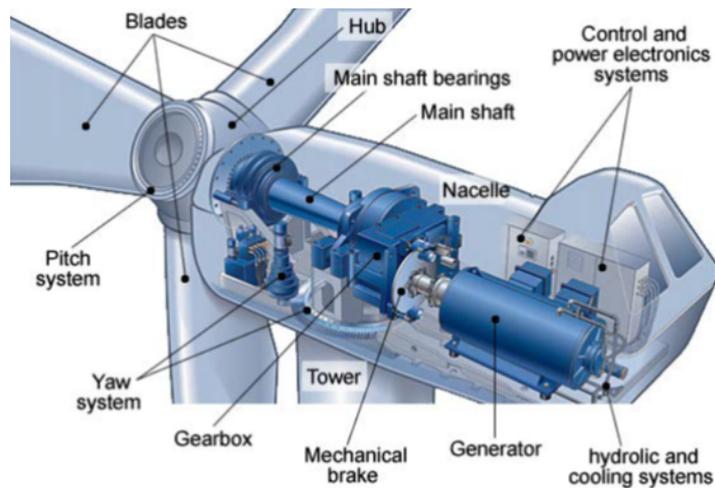


Figure 1.4: Drivetrain of a geared horizontal axis wind turbine [9]

The drivetrain is composed of vulnerable components, failure of these components could lead to downtime. When looking at the entire lifespan of offshore wind farms, it can be concluded that the failure rate of electrical components is the highest. But, drivetrain components lead to the highest annual downtime [6; 10]. For the Horns Rev I offshore wind farm the generator and gearbox contributed to 55% of the downtime of the wind farm [10].

Efficient use of condition monitoring systems can lower downtime and increase the availability of wind farms. The additional cost of developing or using commercial available condition monitoring methods are significantly lower than the savings in O&M cost [9; 10].

1.1.3 Condition monitoring

The IEC standard dictate that the objective of the communication system of wind turbines and wind farms is to provide the actors with information about turbine component or system performance and status [11]. This is deemed essential to control and monitor the wind farm. These communication systems administer components operational status and are able to check whether or not a command has been executed. The information that is exchanged between different components and agents, like the operator, are generally collected by multiple intelligent electronic devices (IED), for example, temperature sensors. The communication of these monitoring and control systems for the turbine can be performed in several ways such as; local or remote supervisory control and data acquisition (SCADA) systems, local real time build-in control systems and energy dispatch centres [11].

The main function of the monitoring system of a wind turbine is to track changes of a system or performance over a certain time interval. This can be done locally off-line by for example taking oil samples that are analysed in a laboratory or it can be done remotely by observation of the data collected by the IEDs [5]. For this research, a system is constructed that relies on the latter approach as it will utilise the SCADA data. Unlike the SCADA systems, most specifically designed or commercial available condition monitoring systems use high resolution data [4], whereas the SCADA systems communicated time period average values.

To prevent shutdown of a turbine due to unexpected failures, condition monitoring can be applied. The most common methods for remote condition monitoring are either: vibrational, acoustic, strain measurements or temperature based. For all these methods, the process of condition monitoring consists of three steps. The first step is data collection by sensors or oil sampling. Secondly the data is proceeded by processing methods and lastly, the data features are analysed to estimate the state of the component.

Vibrational analysis based condition monitoring techniques are the most well-known technique for rotating equipment. This method is able to detect high to very high frequency faults on very specific components. The application of this technique on turbines is standardized in ISO10816-21 [12].

An other method for condition monitoring is by acoustic emission. The propagation of a crack, breaking of fibers or increased friction is registered by transducers and optical fiber displacement sensors. This method is only able to track specific fault that occur in the high frequency range [9]. Research has proven that this method can successfully be applied on the gearbox for the detection of specific faults and is usually performed in combination with vibrational monitoring [13; 14].

Temperature based condition monitoring analyses are used fo detect failures related to temperature changes. This technique is often applied on gears, bearings, oil or electrical components. Usually, temperature systems register the normal operational temperature range and abnormal high temperature can be related to a fault. This method

is considered as very reliable and is not sensitive to false alarms. This approach can however not be used for early detection of faults when only threshold values are applied on the maximum temperature under normal operations [9]. This threshold will only be surpassed if a fault is already present.

Research is done on this approach of condition monitoring to not only link high temperatures outside of the normal operational range, but also within this range. For this, a normal behaviour model is constructed where the temperature under normal conditions is calculated and compared with sensor measurements [15]. Methods of describing the temperature under normal conditions can be done in several many ways and is a common field of research.

Several regression models have been developed using either linear or polynomial regression curves by Schlechtingen [16], Yang developed a regression method using the least square approach and analysing the model parameters [4] and Garlick has proved to be able to describe component temperatures by using an ARX model [17]. Qiu has highlighted the accuracy of applying thermophysics to calculate bearing temperatures if all thermal and system status parameters are known [18].

Applying artificial neural networks (ANN) has become a popular technique for a data-driven approach to model drivetrain component temperatures. This started in 2006 when Garcia constructed an ARX based ANN [19]. Schlechtingen build on this to build two ANN models, utilizing either ARX or FSRC. Both methods prove to be prone to false alarms. For the ARX method, an accuracy of ± 2 °C could be achieved [16]. Tautz-Weinert investigated other machine learning approaches for normal behaviour models such as, ANN, Gaussian process regression, Support Vector Machines (SVM), and Adaptive Neuro-Fuzzy Inference systems. It was concluded from this that Adaptive Neuro-Fuzzy Inference systems and ANN performed the best whilst SVM resulted in the large errors. It was also concluded that ANN take a significant longer time to train compared to SVM [20].

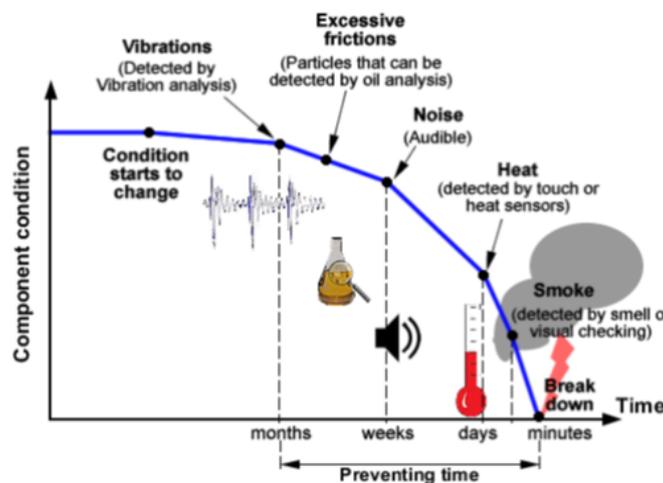


Figure 1.5: Indication of failures and response time [9]

The different indications of failure of a component and the methods to detect these signs of failure at different stages are given in figure 1.5.

Other types of condition monitoring that are commonly used are: Visual inspection which relies on human interpretation or by using an optical fiber to monitor the structural properties and condition of the blades. More research is performed in new state of the art techniques that rely on thermography or radiographic inspection.

1.2 Research motivation

The main approaches that can be taken with regard to O&M are either run-to-failure, time based, condition based or predictive maintenance. The first relies on a corrective approach where maintenance is performed after failure of a component. The time and condition based maintenance both use a preventive approach where the aim of the O&M strategy is to avoid compromise in production due to component failure. By extrapolation of historical data, failures can be predicted before they occur, this approach enables predictive maintenance [9]. This research can enable a shift from a run-to-failure or time based O&M strategy into a condition based strategy. Thus, saving O&M cost by limiting the number of maintenance trips to the wind farm and avoiding downtime of the turbine due to component failure.

The wind farm owner can not only obtain a deduction of O&M costs, but condition monitoring can also result in an increase in the availability of the wind farm. If downtime due to failed components is minimized, the wind farm will have a higher energy output and generate more income. Both these advantages can be realised without any additional investments in new hardware for the Egmond aan Zee offshore wind farm. To do so, a condition based monitoring framework will be developed in this research that utilizes data from the SCADA system.

Lessons learned from this research, where the Egmond aan Zee wind farm is used as case study, can be implemented in the monitoring approach of future wind farms, like Borssele 3 & 4. Since offshore turbines operate in an environment with lots of fluctuating operational conditions, a condition monitoring approach is needed that takes these variations into account. The result of this is a robust general approach, that can also be used for different turbines and/or wind farms.

Since there is no condition monitoring system, vibrational or acoustic data available for the Offshore Wind farm Egmond aan Zee and this research. It is investigated if a stand alone temperature based framework can be utilized for early stage fault detection. To design an approach that can deal with varying environmental conditions, this research will investigate which data driven model can account for this. This model should be able to be used for a large number of signals from within the drivetrain and over a large range of operational conditions. To make this approach fit for the future, it should also be computational quick to accommodate the large data volume that is associated with modern wind farms.

A lot of research has been performed to create methods for condition monitoring methods for fault detection, utilising SCADA data. But research on the effect of turbines placed in offshore condition is missing within these models. In the motivation above, it is mentioned that accommodating for the variations in the operational conditions in the condition based monitoring is of importance. Wake effects in offshore wind farms are a great example of this. This research will address the effects of wake conditions on drivetrain component temperatures to obtain an insight in these effects.

1.3 Thesis objective

Taking the above mentioned motivation into account, two main research objectives can be formulated:

To develop a temperature based condition based monitoring framework for Egmond aan Zee offshore wind farm that utilizes the SCADA data for drivetrain component monitoring.

and:

Use the SCADA data to identify wake effects on drivetrain components by a comparison of component temperature distribution under different wake conditions.

To help and reach these objectives, the following questions are addressed in this research:

1. Can a framework be developed to accounts for all the steps of condition monitoring: Data acquisition, data proceeding, feature extraction and anomaly detection?
2. How should operational SCADA data be proceeded so it can be utilized for (1)?
3. Which data driven model can the best be implemented in the framework of (1), to describe temperatures of drivetrain components under healthy conditions?
4. What analysis should be performed to detect anomalies in the data that could be related to faults in the drivetrain?
5. Can a difference be observed in the temperature probability distributions of component temperatures under different clustered wake and wind conditions?

1.4 Thesis outline

In the background chapter, an overview is presented for the Egmond aan Zee wind farm, as data from this offshore wind farm is used as case study. Also, the fundamentals of machine learning are explained in this chapter.

In the third chapter, all the steps for data collection and proceedings are discussed. The correlation between different drivetrain signal is evaluated from a mathematical and system integration point of view. The consideration around the selection of parameters

and training periods for the used models are discussed after this and the last part of this chapter focusses on the evaluation of the training and test dataset.

The different models that can be utilized for feature extraction in the developed framework with a condition monitoring task are evaluated in chapter four. For this, several simple regression and farm averaging methods are first looked at. After this, supervised machine learning methods with either regression or classification task are covered. The discussed models are then evaluated on accuracy and computational speed. In the last part of this chapter, different control charts are discussed that can be used for anomaly detection.

In the fifth chapter, the performance of the different models is presented for a turbine in healthy conditions. After selecting the best performing model, the framework is used to detect anomalies in the data over a period of one year to detect faults in the drivetrain. The last part of this chapter is used to discuss some remarks on the found results related to the condition monitoring framework.

In the sixth chapter, the approach for investigating the effects wake on component temperature is presented. To describe the wake conditions at OWEZ, the Ishara wake model is evaluated which is implemented in the in-house wake tool of Shell. The temperature distributions for different wake conditions in the same wind speed bin are evaluated in this chapter as well as the results related to effects of asymmetric wake loading over the rotor. The last part of this chapter used for the discuss the results.

In the last chapter, the discussion and conclusion is presented. The research questions are answered to discuss the outcome of this master thesis in respect to the research objectives. The implications of the results are covered and recommendation for future work are presented in this chapter.

2 Background

Offshore Wind farm Egmond aan Zee (OWEZ) is first evaluated in this chapter to provide background on the case study that is used for this research. An overview of the fundamentals of machine learning is provided after for this to provide and understanding of the models used in section 4.

2.1 Case Study

To design and test a SCADA based condition monitoring system, data from the OWEZ is used. This wind farm is owned by the NoordZeeWind consortium, which is a joint venture of the Dutch utility company Nuon Vattenfall and the Royal Dutch Shell.

2.1.1 Farm placement and foundation specification

OWEZ is located 10 to 18 kilometres off the Dutch coast near Egmond aan Zee and construction began in 2006 to become the first Dutch offshore wind farm. The wind farm has been in production since 2007 and supplies energy to around 100.000 Dutch households [21]. The exact location of the wind farm is $52^{\circ} 36' 21.6''$ N, $4^{\circ} 25' 8.4''$ E and the closeness to the Dutch coast is displayed in figure 2.1.

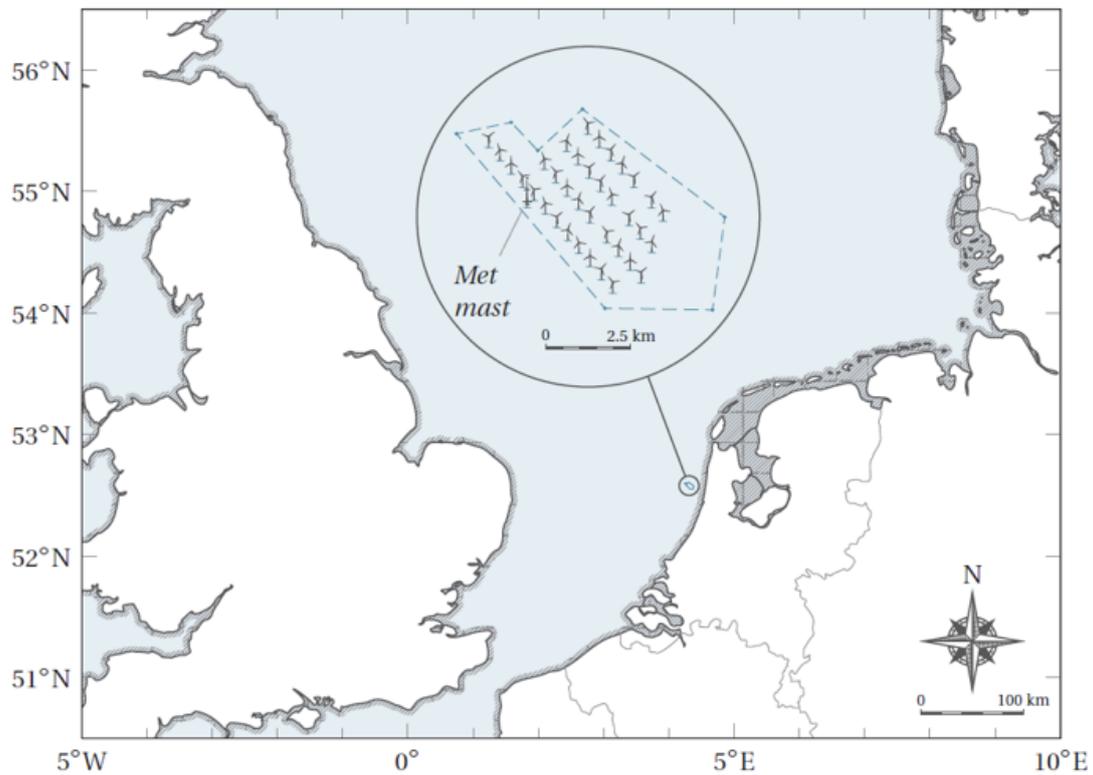


Figure 2.1: OWEZ geographical location and lay out [3]

The lay out of the 36 3.0 MW Vestas turbines in OWEZ is visible in figure 2.1. These turbines are placed with a spacing between the rows of 10.5 time the rotor diameter and 7.3 times the rotor diameter between turbines in the same row [22]. The turbines utilises a monopile foundation with a diameter of 4.6 meter and an average length of 45 meter. Connected to the monopiles are J-tubes, to feed in the power and communication cables from and to the turbine.

The transition piece that connects the monopile and the tower is calculated to have the platform 13.2 meter above mean sea level to prevent wave slamming and water run up. The transition piece is used to level-out small inaccuracies in the vertical angle of the monopile and connected to the transition piece are the boat landings and J-tubes extension pipes.

A meteorological mast is located on the south-west edge of the wind farm. This mast experiences free wind from the dominant wind direction, but due to the small distance to the wind farm, will experience wake effects from different wind directions. This mast was used during a measurement campaign for 2008 until 2010 to collect wind speed and directional measurements as well as measurements on wave and current conditions [23].

2.1.2 Environmental conditions

It can be conclude from the data that is collected over the measurement campaign from 01-07-2005 until 31-12-2008 that the mean annual wind velocity at hubheight is around 8.9 m/s [24]. With data from the Dutch offshore wind atlas, a similar mean wind speed of 8.91 m/s can be found. In figure 2.2 it can be seen that the dominant wind direction is between 250° and 190°. The mean wind velocity of ≈ 9 m/s is significantly lower then the rated wind velocity of 15 m/s, thus most data is expected to be in the operational range between cut-in and rated wind velocity.

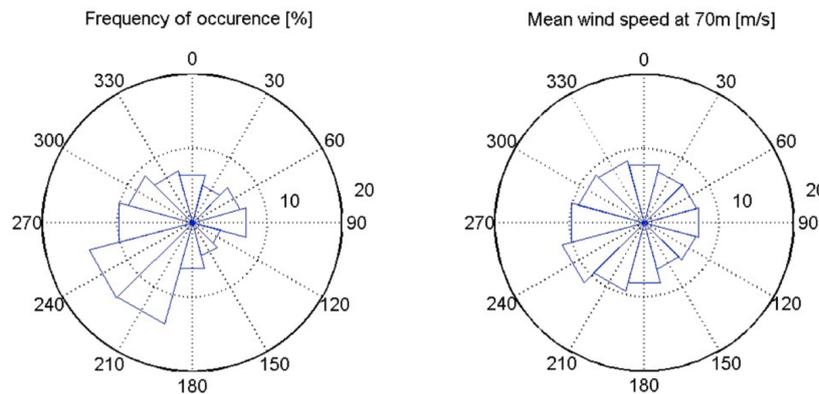


Figure 2.2: Annual wind conditions at hub height for the OWEZ [24]

Over the data, a mean significant wave height of 1.2 meter is recorded and only 11% of the measured significant wave heights are of 1.5 meter or above. From figure 2.3 a clear dominant wave direction from the west is visible. This can be explain by the small

distance to the Dutch coastline. The water depth at the OWEZ site varies between 16 and 20 meter from mean sea level.

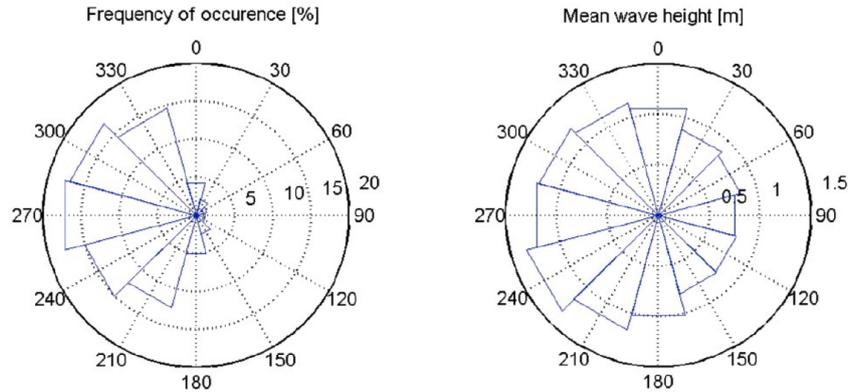


Figure 2.3: Annual wave conditions for the OWEZ location [24]

2.1.3 Vestas V90 sepecification

As mentioned before, the Vestas V90-3.0 MW turbine is used for the OWEZ and SCADA data from this turbine will be used as case study for the development of the condition monitoring framework in this research. The different components inside the nacelle that make up the drivetrain, are discussed to have a good understanding of the system integration.

The gearbox used in the Vestas V90 consists of two planetary stages and one helical stage. The rotor is directly connected to the low speed input of the gearbox, thus eliminating the need for a main shaft. The torque is transmitted through the high speed shaft to the generator, with a composite coupling in between. On this shaft, hydraulic disk brakes are connected as addition to the pitch control brakes [25].

The gearbox relies on a force feed lubrication system, thus no oil sump is needed [26]. The oil that is circulated in the gearbox is cool be an oil-to-water heat exchanger that is placed before the oil distribution block. The water circuit in the oil-to-water heat exchanger is force circulated and cooled water-to-air by two large fans at the air outlet in the back of the nacelle [25]. Between 2008 and 2009, all gearboxes within the wind farm where replaced due to reliability issues [22].

The generator that is part of the Vestas V90-3.0 drivetrain is an asynchronous 4-pole generator with a wound rotor [26]. The water circuit that is used to cool the generator is also used for cooling the converter and hydraulic system and is cooled water-to-air, in a similar way as the cooling water for the gearbox. Since this generator is a double fed induction generator, there are brushes and a slipring present. These components are cooled be an air-duct.

A visualisation of the Vestas V90-3.0 drivetrain that is placed inside the nacelle is displayed in figure 2.4.

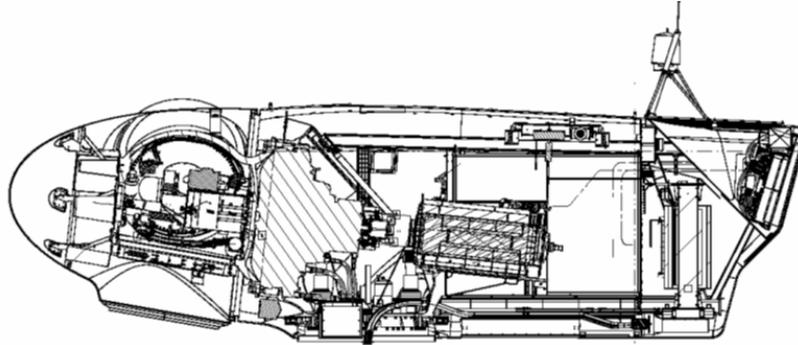


Figure 2.4: Vestas V90 Nacelle [25]

The rotor of the Vestas V90 has a diameter of 90 meter, resulting in a swept area of 6362 m². The operational rotor interval is from 9.0 until 19.0 RPM with a nominal rotor speed of 16.1 RPM. This relates to the cut-in, cut-out and rated wind velocity of 3.5, 25.0 and 15.0 m/s respectively [26]. The hub is located 70 meter above mean sea level. The tower is specifically designed for OWEZ [22].

2.2 SCADA system

In practice, almost all operational wind turbines are equipped with sensors to capture environmental condition surrounding the turbine as well as information on the performance and operational state of the turbine. The collected SCADA data often contains hundreds of signals from different sensors. These sensors outputs are dependent on the turbine manufacturer. But in general contain: Pitch, yaw, active and reactive power output, currents, temperatures from important lubricants and bearings, nacelle temperature, shaft speeds, wind velocity and ambient temperature.

For all these channels, the mean, maximum and minimum value as well as the standard deviation are collected over a given time period. This period is generally 10 minutes, but can differ from this for specific systems. Initially, this system is designed to capture the state of the turbine and provide feedback on the performance of this to the operator. The SCADA system is not able to fully replace the condition monitoring system due to the conflict of frequency. The SCADA system has a very low frequency since it stores a mean value of a 10 minute time interval, whilst many dynamics inside the drivetrain perform at a higher frequency range. Thus many dynamics would be missed when only analysing the SCADA data for condition monitoring.

Nevertheless, applying SCADA data for a condition base system has the advantage that it is available to the operator or wind farm owner without any additional investments in hardware. The SCADA data can be used for health and condition monitoring by multiple different methods. The methods of relevance for this project shall be discussed in section 4.

2.3 Introduction to machine learning

Some machine learning techniques are used in section 4.3. Because of this, it is important to have a fundamental understanding of machine learning.

The definition of machine learning is given by Arthur Samuel in 1959 as [27]:

"Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed".

This explanation of machine learning is well in line with the paradigm of the 60's in the field of statistics. This paradigm is given by Vapnik as [28]:

"What must one know a priori about an unknown functional dependency in order to estimate it on basis of observations?"

Data samples or past experiences are used to train machine learning models by utilizing learning algorithms. These learning algorithms are in essence a form of applied statistics, where computers are used to develop complex functions to describe the behaviour of data, based on statistics [29]. When this functional dependency of the data is found using training data, the model is used to predict outcome based on data outside of the training period. This is called generalisation and this is the main objective of a supervised machine learning model.

2.3.1 Types of machine learning

Most machine learning methods can be considered as supervised learning or unsupervised learning. The difference between these approaches is in how the data is presented to the learning algorithm.

- **Unsupervised learning:** Here, the learning algorithm is presented with the entire dataset without any additional information. Useful features and or properties are left to the algorithm to learn. Some examples of these application are denoising, or clustering of data.
- **Supervised learning:** When a learning algorithm is used with supervised learning, then the dataset is presented with targets or labels. The target y is presented as well as the input x so the learning algorithm knows what to train for.

A crude example of unsupervised learning is to find the probability distribution of a random vector x as $p(x)$, whilst during supervised learning the random vector x is presented together with the according value of y , to calculate $p(y|x)$ [29]. In this research only supervised learning is used to develop the models discussed in section 4.3.

Not only in the way the model is trained, also in the task of the model a distinction can be made. The two main tasks of the machine learning models that are used are either regression or classification. The main difference and application for the normal behaviour model of these types will be discussed below:

- **Classification:** This task assigns to which class a given data point belongs, based on the input. This is done based on a decision function like: $f : \mathbb{R}^n \rightarrow \{1 \dots k\}$. Here, k is the number of different classes [29].

This type of machine learning is an useful tool for pattern or object recognition. Still, this type of machine learning can also be used for predicting bearing temperature in the turbine drivetrain. In that case, k is the range of temperatures on which the model is trained. Since the SCADA signal stores integer data, the difference between the classes will be of 1 °C.

- **Regression:** With this task, a machine learning model formulates a function that can describe values along a continuous trend. The output of the learning algorithm in this case a function like: $f : \mathbb{R}^n \rightarrow \mathbb{R}$ [29].

Temperatures inside the drivetrain are a continuous phenomena. It is possible to predict component temperature along this range with a regression model to also describe the region between the classes of the SCADA signal.

Other types of machine learning task are for example clustering, pattern recognition, transcription or translation. For the application of this research, only the classification and regression models are of interest. More on different machine learning tasks can be found in [29; 30] or [28].

2.3.2 Consideration around fitting

When using learning algorithm to train a model to fit to operational data, a couple of considerations need to be taken into account. A model can be considered as under-fitted, when the model can not describe the functional dependence of the data. This can be seen by a large error of the model on the training data, and the model will also perform poorly when generalizing values for new data [29].

It is imported that the learning algorithm does not result in an over-fitted model. The concept of over-fitting is when, for a given training dataset, a too complex model is developed. This phenomena is visible when the training error of the model is very small whilst the number of errors are still large when the model is tested on new data [28; 29]. To test for this concern, it is important to visualize a training period as well as an unfamiliar dataset, to see if the trained model can be used for both datasets.

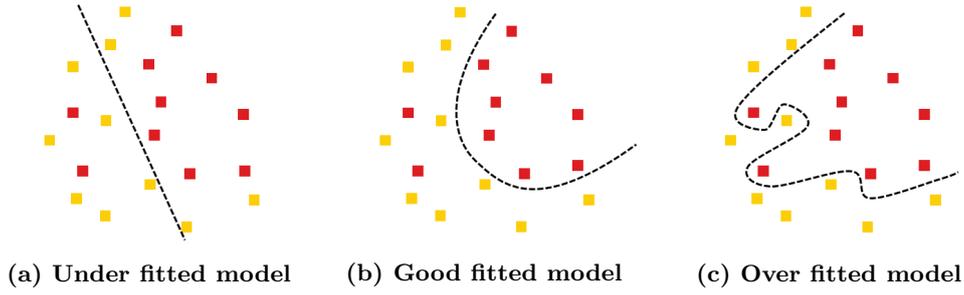


Figure 2.5: Concept of appropriate fitting of a model

In figure 2.5, the different consideration around fitting a model to the training data can be seen.

2.3.3 Empirical risk and error function

The loss function can be used to quantify how well the functional behaviour found by the learning algorithm, $\hat{y} = f(x)$, can be used to describe the data, y . Two examples of loss functions are the absolute loss and quadratic loss functions given by equations 2.1 and 2.2 [31]:

$$L_1 = \ell(f(x), y) = |\hat{y} - y|, \quad (2.1)$$

$$L_2 = \ell(f(x), y) = (\hat{y} - y)^2. \quad (2.2)$$

The expected loss over a dataset can be expressed by the risk function, given in equation 2.3. The goal of a learning algorithm is to construct a function for $\hat{y} = f(x)$ where the expected loss is minimized:

$$R(f) = \int \ell(f(x), y) dP(x, y). \quad (2.3)$$

The risk function of equation 2.3 is referred to as the true risk. The distribution of the data in the validation or test set, given by $P(x, y)$, is not known in advance. $R(f)$ can therefore not be calculated, what can be calculated is the empirical risk over the training data by equation 2.4:

$$\hat{R}_n(f) = \frac{1}{n} \sum_{i=1}^n \ell(f(x), y). \quad (2.4)$$

A L_2 loss function in combination with equation 2.4 is used in this research to calculate the average loss over a given dataset. The root mean square error (RMSE) can be calculated by taking the root of this average loss to obtain an error with the same unit as y . Different models can be evaluated on bases of accuracy by analysing the resulting RMSE. The RMSE can be calculated using equation 2.5:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y} - y)^2}. \quad (2.5)$$

2.3.4 K-fold cross validation

The best way to validate a trained model is by using data that has not been experienced by the model whilst training. Data that is used for training can be considered contaminated and will always result in a low error, this error is referred to as the training error. To eliminate this problem, the concept of k-fold cross validation is used [32].

The selected training dataset is randomly broken up into k number of parts for cross validation. The model is then trained k times where $k-1$ parts are used for training of the model and one part is withheld from the training data to use as validation data. The resulting error is calculated on the validation dataset. This process is repeated k times, rotating the training and validation datasets parts. The final validation error is an average of the k number of validations.

The validation error will be used as indication of how well the model fits to the in-sample dataset. The validation error that is the result of a k-fold cross validation, with $k=5$, and will be used in the remainder of the report when analysing the model error over the in-sample data. The concept of k-fold cross validation is visualised in figure 2.6.

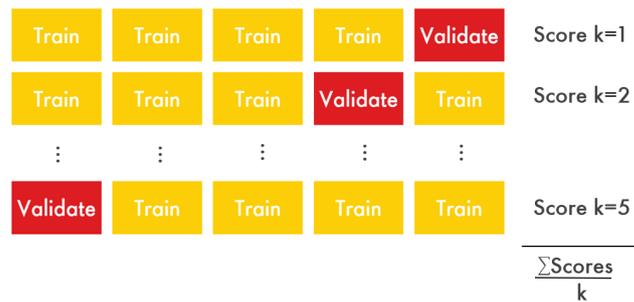


Figure 2.6: Concept of k-fold cross validation

The data that is used in the validation and training process is referred to as the in-sample dataset. An out-of-sample dataset is disjoint from the training and validation data and can be used to calculate the test error. This error can be used to test for over fitting. A too complex model developed for the in-sample data might indicate a low validation error but will not fit over the testing data. Thus, resulting in a large testing error.

3 Data proceedings

A collection of raw operational data as well as modelled meteorological data is used in this research. These different datasets are first covered in this section. To identify anomalies caused by component failure from the raw operational data, first the data needs to be processed to collect valuable information hidden in the smeared out dataset. The different methods and consideration taken in the research will also be covered in the section. The relation for different parameters to each other are covered in the last part of this section.

3.1 Data collection

The data that is used for this project consist of two datasets. The first set is on ambient and operational conditions on turbine level at the OWEZ. This dataset is constructed from data collected by the SCADA system. The second dataset is on the ambient wind condition at the OWEZ location, this dataset is collected from the Dutch Offshore Wind Atlas (DOWA). Data from this project is constructed by KNMI, ECN part of TNO and Wiffle [33].

3.1.1 SCADA system

As explained in section 2.2, the SCADA system stores information on the turbine performances and operation conditions. In this section, the SCADA system for OWEZ is discussed as well as the signals that are interesting for this research.

The SCADA signals that are available for this research are given below. All these signals are available for every turbine in OWEZ over the period of one year. The sensor location in the nacelle are displayed in figure 3.1.

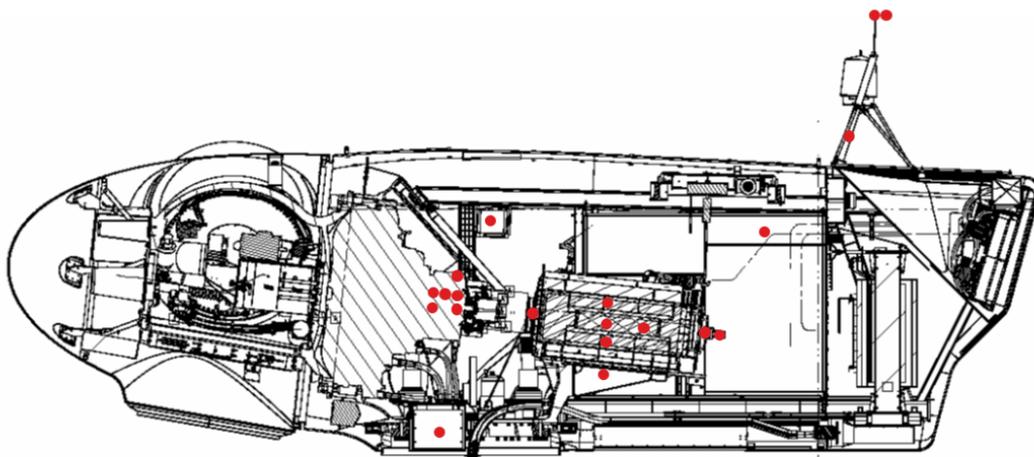


Figure 3.1: Vestas V90 SCADA signal sensor locations

1. Ambient conditions
 - Wind speed
 - Wind direction
 - Ambient temperature
 - Nacelle temperature
2. Power performance
 - Delivered grid power
 - Grid frequency
3. Generator
 - Bearing 1 temperature
 - Bearing 2 temperature
 - Phase 1 winding temperature
 - Phase 2 winding temperature
 - Phase 3 winding temperature
 - Cooling water temperature
 - Slipring temperature
 - Generator rotational speed
4. Gearbox
 - Hollow shaft bearing 1 temperature
 - Hollow shaft bearing 2 temperature
 - High speed shaft bearing 1 temperature
 - High speed shaft bearing 2 temperature
 - High speed shaft bearing 3 temperature
 - Oil inlet temperature
 - Oil return temperature
5. Hydraulic system
 - Hydraulic oil temperature
6. Operational state
 - Alarms code

The available dataset lacks vibrational measurements, commonly used in condition monitoring. Because of this, the methods for condition monitoring will be constrained to temperature based methods. Most of the information stored in this dataset can be used to obtain insights in the condition of the gearbox and generator.

3.1.2 Ambient condition data farm level

To investigate the wake effects in the wind farm, data needs to be used that describes the free stream condition in OWEZ. Since this information is missing from the SCADA dataset, data from the DOWA is utilized.

The data in the DOWA is constructed with numerical prediction based on the HiRLAM model. The wind profile, Weibull distributions and surface roughness, based on wave condition, are included in this model. The outcome of the model is then validated with data from several KNMI measurement point. This results in a mean error of -0.08 m/s between the model and measurement [34].

The data collected from the DOWA is are hourly average values for the wind speed and wind direction at different heights at the geographical coordinates, corresponding to the OWEZ location [33]. The measurement height are 10, 20, 40, 60, 80, 100, 150 and 200. The dataset spans over a period of the 1 January 2014 01:00:00 until 31 January 2019 18:00. From this dataset, only the period that corresponds to the time range of the SCADA dataset is used.

3.2 Data filtering

Different filters and selection methods have been applied at different stages in this research to increase the quality of the dataset. The application and necessity of these filters will be discussed in this section.

3.2.1 Time stamp matching

The SCADA data for this case study is stored with different time intervals, depending on the signal. Some channels store the statistical properties over a time period of approximately 10 minutes, whilst others store 10 seconds values. To obtain a standard signal, all values are converted to ten minute intervals. When less than 50% of this interval contains data, the time stamp is discharged as the statistical quality can not be guaranteed.

3.2.2 General filtering

When evaluating the condition of the turbine drivetrain, it is important to only analyse a period where the turbine is operating and communicating with the SCADA system. For this reason, the following filter criteria are applied [4; 35].

- Filter out data where: $P < 0.01 \cdot P_{rated}$
- Filter out data where: $U < U_{cutin}$ or $U > U_{cutout}$
- Filter out data where one of the input or output signals communicate NaN value

In figure 3.2, an example is given of the application of this filter on data for the one of the high speed shaft bearing temperature. It can be seen from this figure that data outside of the operational range of the turbine is filtered out.

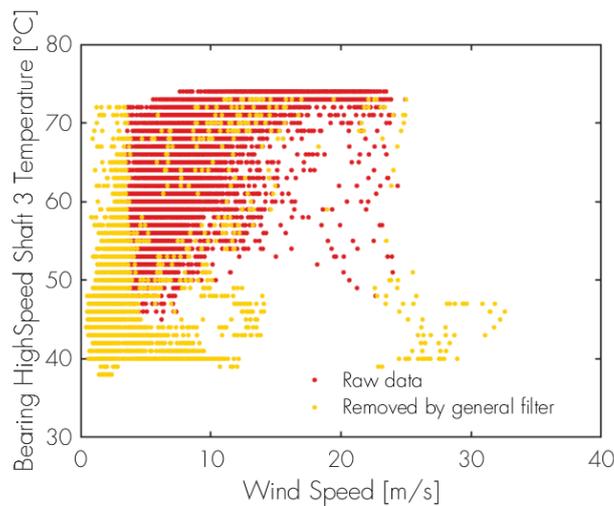
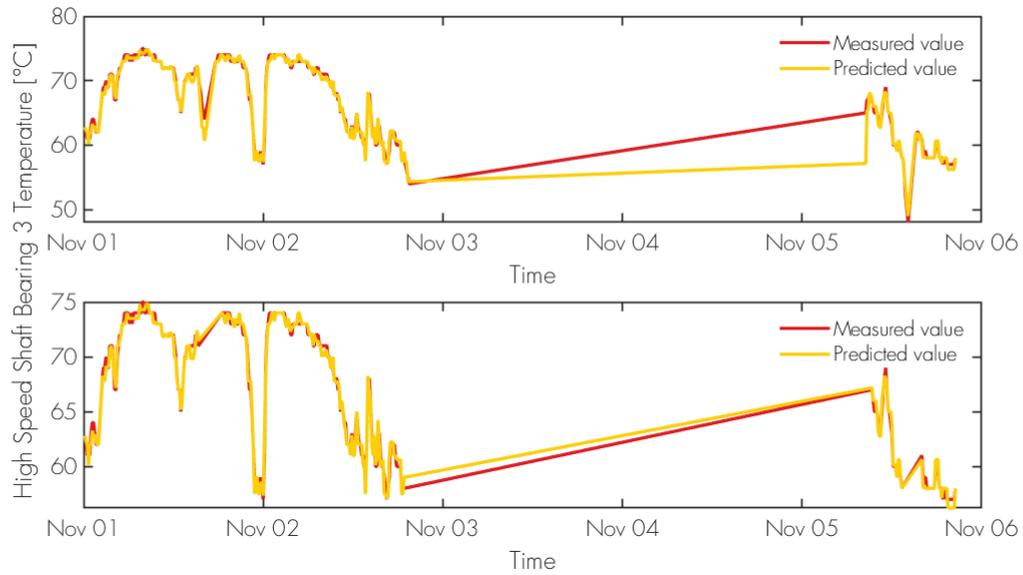


Figure 3.2: Application of the general filtering

3.2.3 Missing data filtering

After a hiatus in the data due to losing communication or due to a stop of the turbine, the model might result in a large error because this uses the last known input and response for predictions. Thus, to filter out this effect the missing data filter is applied. To do so, a constraint is introduced that one hour of continuous data is available, prior to a considered data point for a prediction [35].



**Figure 3.3: Top figure: before application of missing data filter.
Bottom figure: after application of missing data filter**

The application of the filter can be seen in figure 3.3. This time series also highlight the necessity of the filter. In the top figure, the missing filter is not yet applied. Here, a large error is visible after the hiatus from November 2 until November 6. In the lower figure, the filter has been applied and one hour of predictions are discarded, resulting in the removal of the error.

3.2.4 Mahalanobis filtering

To create an as clean as possible dataset to train the model to and reduce the error within the model, the outliers in the training data need to be removed. To do so, a filter is constructed according the following steps:

- (i) The dataset is split in N number of clusters and each data point in dataset D is assigned to a cluster number n where $n \in \{1, 2, \dots, N\}$. The number of clusters is based on the different power-curve areas and the understanding of the turbine behaviour. These clusters are presented in table 3.1.

Cluster	n=1	n=2	n=3	n=4	n=5
Wind Speed [m/s]	0 → 3.5	3.5→7	7→11	11→15	15→25

Table 3.1: Filter clusters

- (ii) For each of the clusters $D^{(n)}$, the centroid $\mu^{(n)}$ is calculated. Next the Mahalanobis distance to the centroid of each point in the dataset $D_i^{(n)}$, is calculated using equation 3.1. Here, C represents the covariance matrix and $i = \{1, \dots, \text{length}(D^n)\}$:

$$MHD_i^{(n)} = \sqrt{\left(D_i^{(n)} - \mu^{(n)}\right) C^{-1} \left(D_i^{(n)} - \mu^{(n)}\right)^T}. \quad (3.1)$$

- (iii) A log-logistic distribution is fitted to the distribution of the MHD values in each cluster [35], as shown in figure 3.4.
- (iv) The improbable outliers can be filtered out by assigning a threshold value of 2.5% to the probability density distribution and removing all data points that exceed this threshold.

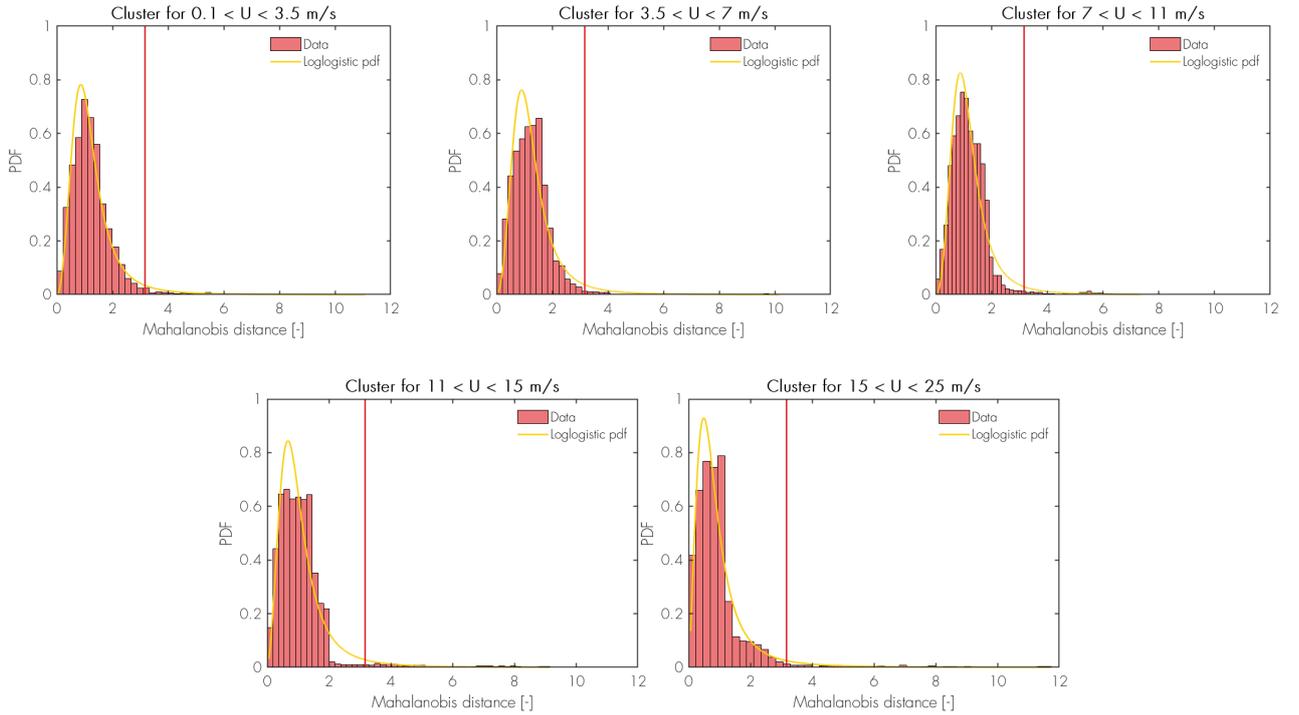


Figure 3.4: Mahalanobis distance distribution per cluster and threshold value

The filtering result is visualized in figure 3.5. It can be seen from this figure that all outliers are removed from the dataset. It must be noted that this filter should only be applied to the training data, as outliers might contain valuable information when generalizing over data where a fault is present in the drivetrain.

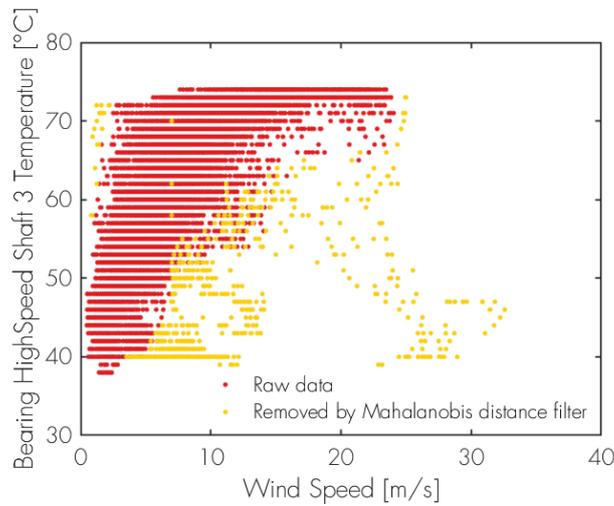


Figure 3.5: Application of the Mahalanobis distance filtering

3.2.5 Filtering result

The resulting data after applying the general filter and Mahalanobis distance filter on the training data can be seen in figure 3.6. Here, it can be seen that the remaining dataset is removed from any noise and improbable outliers and can be considered as a clean dataset to train the models on.

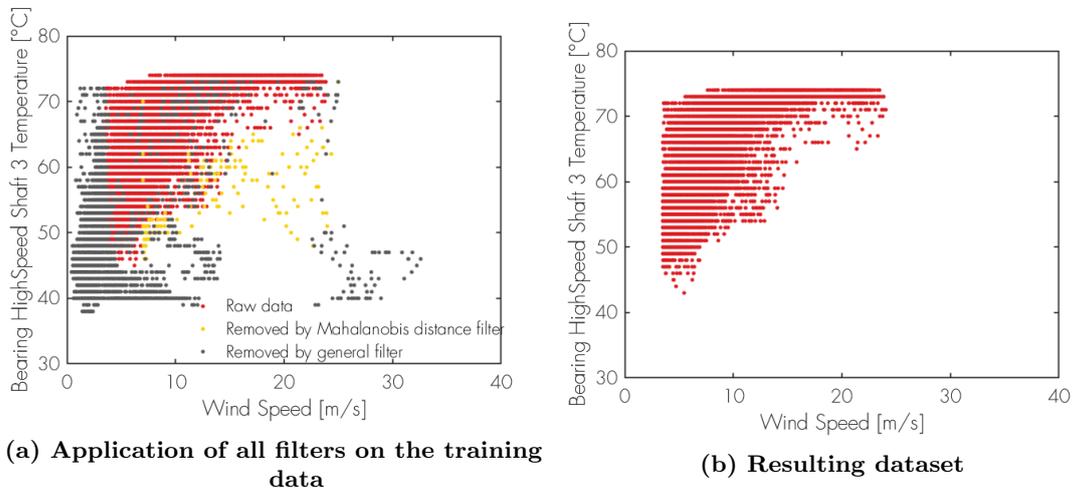


Figure 3.6: Filtering of the training data

The missing data filter is not included in figure 3.6. This filter is not applied on the training period and is only used on generalisations of the model.

For 6 out of the 36 turbines, less than 30% of data is available after the filtering. These turbines are discarded from the condition monitoring as too many hiatus in the data will compromise the ability to identify a training period where the turbine can be considered healthy and is large enough to experience data from a range of operational conditions.

3.3 Correlation of parameters

To select the correct SCADA signals for input and response of the model, it is important to understand how different sensor outputs are related to each other. This is determined by the mathematical correlation to each other and also have a good understanding of the system intergeneration to obtain a insight in the physical correlation to each other. These matter will be covered in the next sections.

3.3.1 Mathematical correlation

To test the a linear correlation between two scores, the Pearson correlation coefficient can be used as given by equation 3.2 [36]:

$$\rho_{X,Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}. \quad (3.2)$$

$\rho_{X,Y}$ will obtain a value of $-1 \leq \rho_{X,Y} \leq 1$, where a value of 0 implies no correlation of the data and -1 or 1 indicates full negative or positive linear correlation. In equation 3.2, σ_X and σ_Y represent the standard deviations of the realisations and μ_X and μ_Y are the mean values of the data.

In figure 3.7, the correlation coefficient of all the SCADA signals are presented with respect to the delivered grid power. It can be seen that some parameters are highly correlated. For example, grid power to: Ambient wind speed, gearbox and generator bearings temperature and generator phase temperatures. Parameters with a lower correlation are for example: ambient and nacelle temperature, wind direction, grid frequency and generator slipping temperatures.

In appendix A until D the correlation of all parameters with respect to the temperatures of the hollow shaft bearing 1, high speed shaft bearing 1, generator bearing 1 and slipping are given. This provides a valuable insight in the correlation of different signals.

3.3.2 Physical correlation and system integration

Now that the statistical relation is determined, it is important to also analyse the correlation from a physical point of view to see if there is a causal relation between different component signals. To fully understand the physical correlations between parameters, the system integration of the components and sensor locations need to be evaluated. The information in this section is taken from the Vestas V90 general specification document and the operation and service document [26] [25].

Gearbox

Ideally, the bearing temperature is measured in all stages of the gearbox and on all the gears. Due to mechanical constrains, all bearing temperature sensors locations are at the stationary part of the bearing in the third stage of the gearbox. In this stage, there are 3 sensor locations on the high speed shaft and two on the hollow shaft that passes

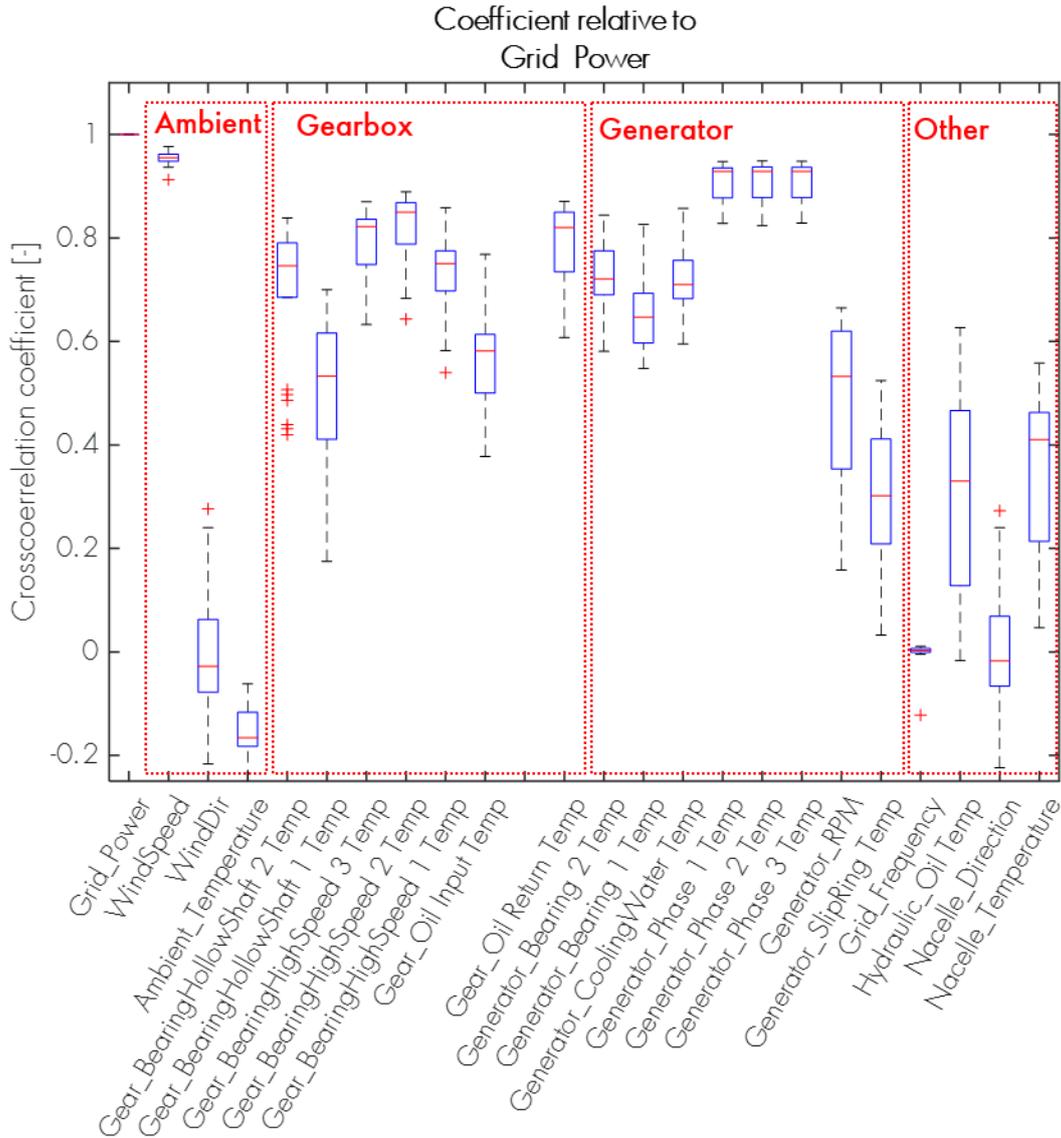


Figure 3.7: Pearson correlation coefficient with respect to delivered grid power

through all stages. These bearings are therefore located very close to each other and high correlation is expected as well as small temperature differences between the bearings on the same shaft. This is because of the heat transfer between these parts due to convection. These bearing locations can also be seen in figure 3.8.

Two other sensor locations for temperature are in the gearbox oil system, one in the oil tank and one in the distribution block right before the oil inlet of the gearbox. An oil-to-water heat exchanger is placed between the oil tank and distribution block. The oil temperature at the distribution block is thus expected to be colder than the one in the tank and less correlated to the bearing temperatures in the gearbox because heat is being taken away from the system.

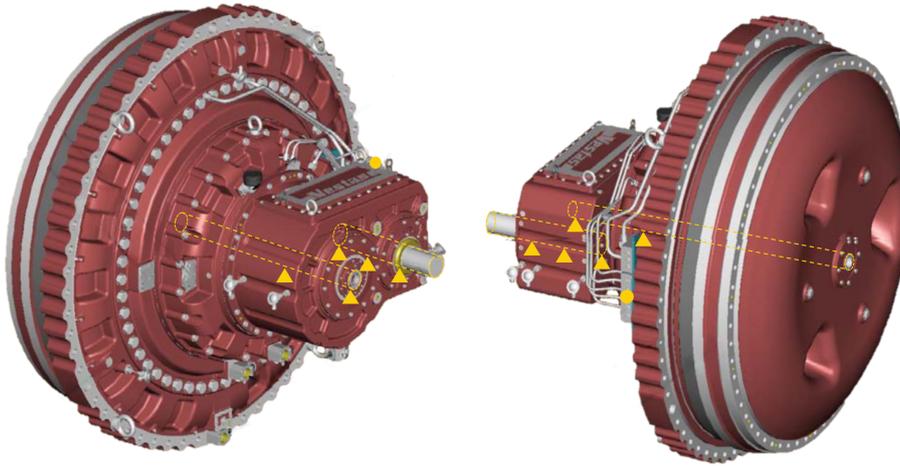


Figure 3.8: Temperature sensor locations in the gearbox [25]

The water circuit in the oil-to-water heat exchanger is force circulated and is cooled water-to-air by two large fans at the air outlet in the back of the nacelle. Valves inside this system are temperature controlled, but the signal is not stored in the SCADA system. This cooling water circuit is displayed in appendix E.

Generator

The generator is cooled by a separate forced flow water-to-air cooling system. The hydraulic system and converter are also connected to this circuit, as can be seen in appendix E. Thus, correlation between these systems is expected. The temperature of the bearings that support the generator are also expected to be correlated to the cooling water as the heat is transferred to this medium.

Since the generator in the Vestas V90, used at OWEZ, is a dubbed fed induction generator, there is a slipring present. The temperature at the slipring is measured and stored in the SCADA signal. To remove dust on the brushes and cool the slipring, there is an air duct on this part of the generator. Because of this, correlation with the measured nacelle temperature is expected.

The remaining temperature signals that are measured in the generator are on the windings of each of the three phases. A spike in the temperature in one of the phases can indicate short circuit and should also be measurable by a voltage drop over the phase. An exploded view of the generator can be found in figure 3.9.

Ambient conditions

All the components inside of the nacelle are affected by the air temperature inside the nacelle, this temperature is measured and controlled. A fan circulates air from the inlet at the bottom of the nacelle to the outlet at the back. The temperature inside the nacelle is correlated to the ambient temperature as it is controlled to always be at least 5 degrees warmer then the ambient temperature. this is to prevent condensation inside the nacelle.

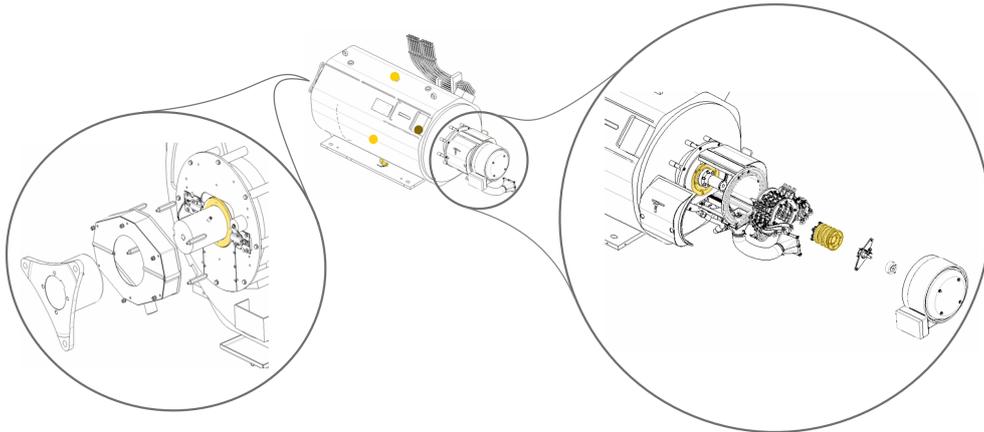


Figure 3.9: Temperature sensor locations in the generator [25]

Hydraulic system

The hydraulic pressure and oil temperature are measured and stored in the SCADA system. As mentioned above, this system is connected to the same water-cooling circuit as the generator and convertor.

3.4 Training period and parameter selection

To make sure that the utilized models can generalize the response signal well, they must be trained on a well defined training period. Supervised machine learning models can only predict values over a range of operational conditions that have experienced during training. [32].

3.4.1 Period selection

To select a period to train the model on, two main considerations are taken: An offshore wind turbine operates in a very varying environment and a machine learning model can only generalize responses over experience that were present in the training dataset. To accommodate all possible operational conditions in the training dataset, the selected trainings period is taken as long as possible.

During this training period, the drivetrain should be in a healthy condition. The trained model will be used to describe the healthy response of a turbine for a disjoint dataset and the residual error should be small if the turbine is performing normally. In addition to this, if maintenance or calibration of the sensors is performed, the model should be trained again. Because the model and the reality will differ from each other.

Since no maintenance logs are present for this research, the above named consideration are taken into account in the following way: The training period for each turbine is taken as the longest period over the dataset that the turbine was in production and not stopped, for other reasons than low wind speeds, for longer than 3 hours. The reason for

taking a period without any stop in production is to make sure that no maintenance is done, or the turbine is stopped because of a fault. The margin of 3 hours is introduced to allow for some gaps in the data due to filtering as is explained in section 3.2.

Depending on the turbine, a training period with a length of approximately two to six weeks is selected. In some cases, a large error is present in the generalizations after a long period of standstill of the turbine. The model is trained again when this phenomena is present, as this is likely due to maintenance been done on the drivetrain and the model is trained on the behaviour of the drivetrain from before the maintenance.

3.4.2 Parameter selection

Several models that are covered in section 4.3, rely on multiple input parameters to describe the functional behaviour of the drivetrain component temperatures. The input and response parameters that are used for this research are collected from the SCADA dataset. The process of selecting these parameters is discussed in this section.

The responses that are selected in this research should be able to be used as indicator for the state of the drivetrain components. It was already discussed in section 3.1.1 that with the available SCADA signals, the monitoring techniques are constrained to temperature based methods. Previous literature has proven that temperature based condition monitoring is possible and bearing temperature can be used as an indicator for a present fault in the drivetrain [4; 9; 37; 38].

Two main considerations are taken for the selection of the input. The first consideration is that the input should be able to describe the operational conditions in which the component of interest operates. Thus, there must a high correlation between the input and response. For this selection section 3.3 can be used.

A very important second consideration is that an anomaly present in the response data, may not influence or be influenced by the input. If this is the case, the model will be biased and might not perform well for fault detection. Two bearings on the same gearbox shaft is an example of this causation. These signals are highly correlated, but if the temperature of one rises due to increased wear, the other bearing will likely also rise in temperature. Because of this, the one can not be used as input and the other as response for fault detection.

A general setting of input parameters is selected that is able to describe the gearbox and generator operational conditions and response without resulting in biased predictions. These parameters are presented table 3.2.

Gearbox		Generator	
Input			
Grid Power		Grid Power	
Wind Speed		Wind Speed	
Generator RPM		Generator RPM	
Nacelle Temperature		Nacelle Temperature	
Oil Return Temperature		Cooling Water Temperature	
Responses			
Hollow Shaft Bearing 1 Temperature		Generator Bearing 1 Temperature	
Hollow Shaft Bearing 2 Temperature		Generator Bearing 2 Temperature	
High Speed Shaft Bearing 1 Temperature		Generator Winding Temperature	
High Speed Shaft Bearing 2 Temperature			
High Speed Shaft Bearing 3 Temperature			

Table 3.2: Generic input and response parameters for the condition monitoring framework

3.4.3 Training dataset evaluation

To obtain an insight in the condition present in one of the training datasets, the data collected over the period from 25 October 17:30:00 until 04 December 14:00:00 of turbine A is highlighted below. This dataset is evaluated as this is used for the evaluation of different data driven models in section 4.3.

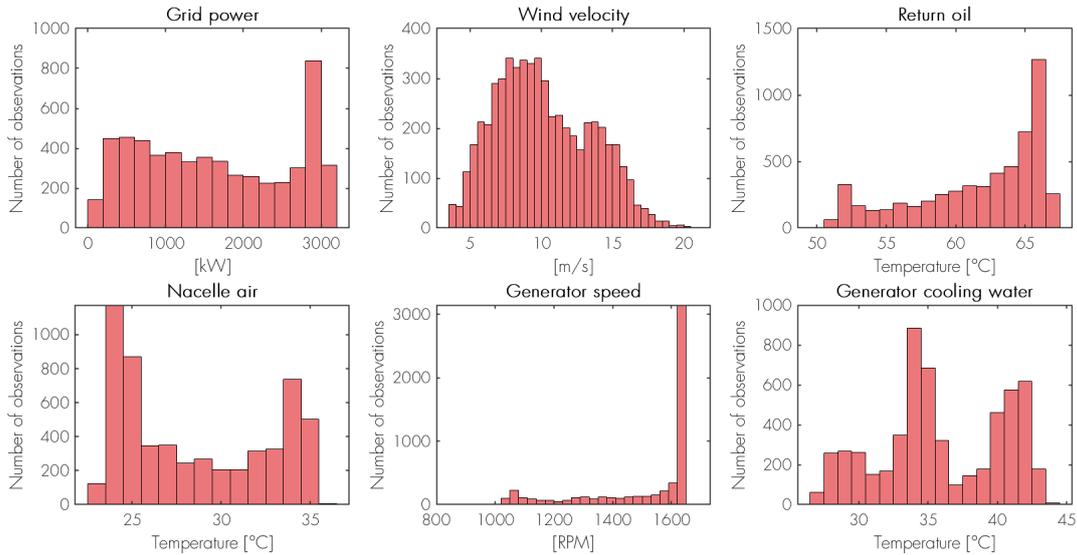


Figure 3.10: Turbine A in-sample dataset, observations of input conditions

The histograms displayed in figure 3.10 show the distributions of the different input parameters over the 5680 number of observations in the in-sample dataset. It can be seen from this figure that the number of observation decreases after $U > U_{rated}$ and that

the generator is operating at maximum RPM for most of the time.

The histograms in figure 3.11 display the number of observations of different component temperatures that are used as model response. It can be seen that the bearings on the high speed shaft have a wider range of operational temperatures compared to the bearing temperatures on the hollow shaft. The widest range of temperatures can be observed over the generator responses.

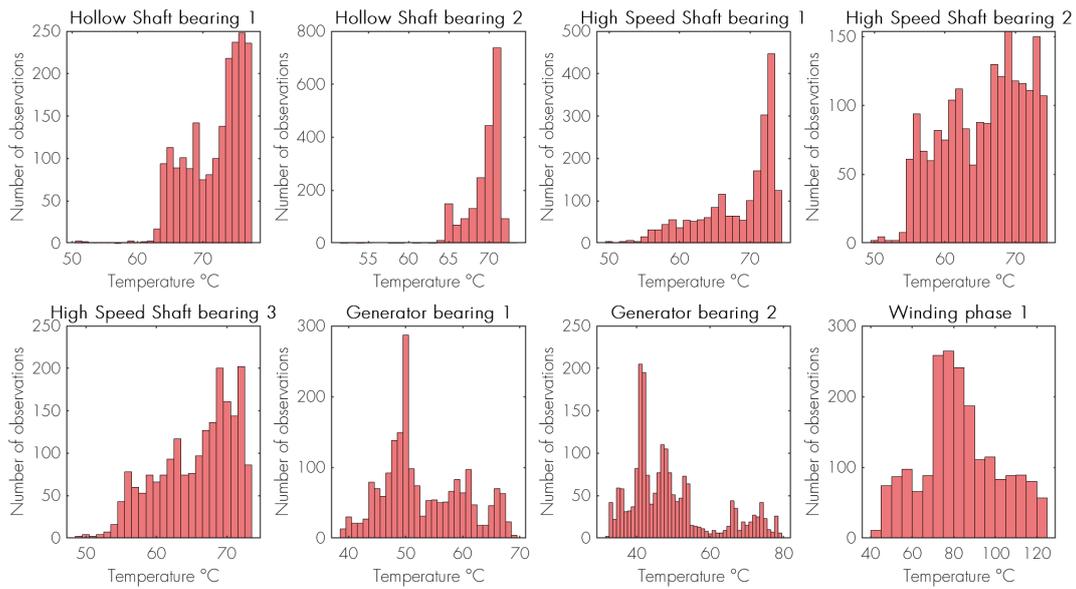


Figure 3.11: Turbine A in-sample dataset, observations of response conditions

The relation of the first high speed shaft bearing temperatures to the power curve can be seen in figure 3.12. It can be seen from this figure that higher wind speeds and higher energy yield can be related to higher bearing temperatures. The temperature of the other response parameters distributed over the power curve can be found in appendix F.

The relation of the first high speed shaft bearing temperature against the generator speed and return oil temperature can be seen in figure 3.13. It can be seen that almost a linear relation can be found between these three scores. This relation can also be observed for the other response parameters, as can be seen in appendix F.

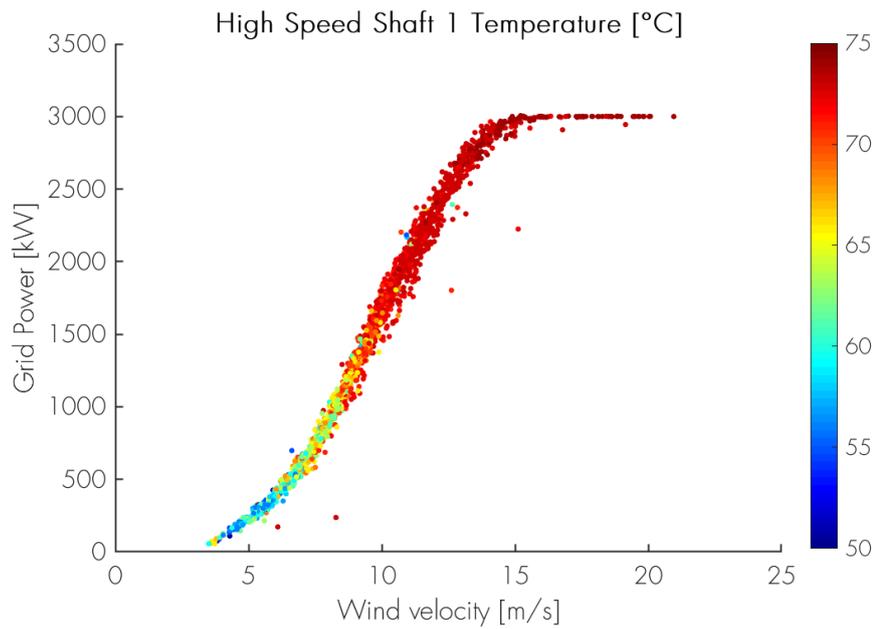


Figure 3.12: Turbine A in-sample dataset, High speed shaft bearing 1 in relation to the power curve

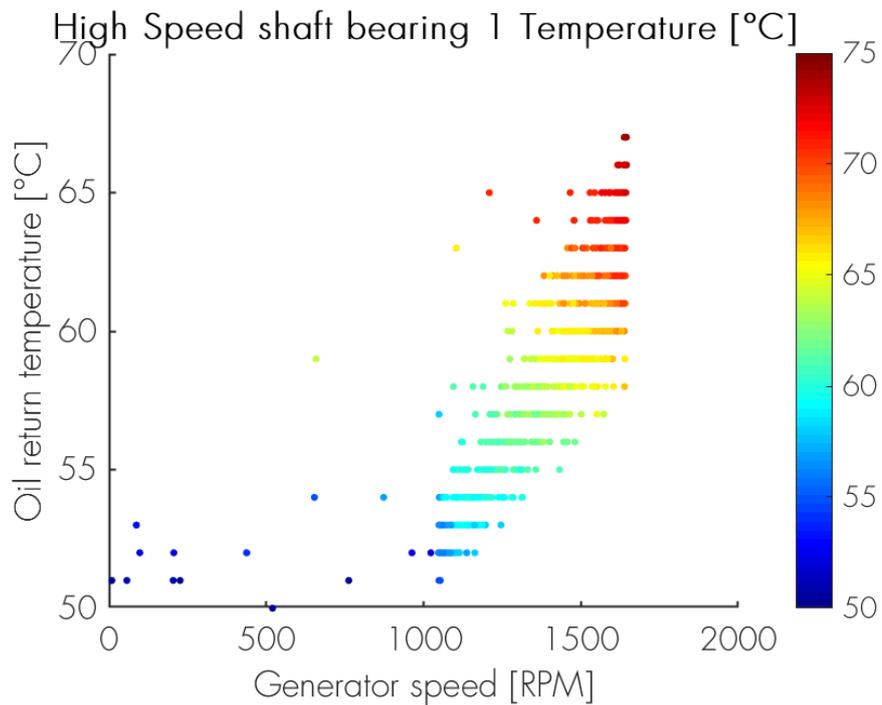


Figure 3.13: Turbine A in-sample dataset, High speed shaft bearing 1 in relation to the generator speed and return oil temperature

3.4.4 Test dataset evaluation

As discussed in section 2.3.4, a disjoint dataset should be used as out-of-sample data to test the models for over fitting. In this section, the test data related to turbine A is highlighted as example. It is important for the test dataset, that the turbine is performing similarly as for the in-sample dataset. For turbine A, a period directly after the in-sample dataset is taken. This dataset from 4 December 14:00:00 until 31 December 23:00:00 contains 3350 data points.

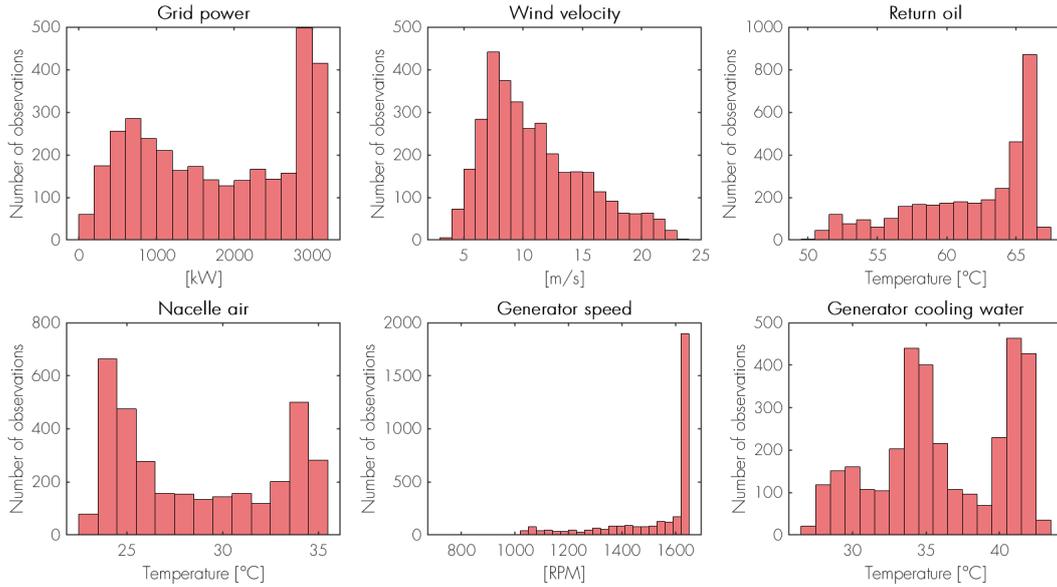


Figure 3.14: Turbine A out-of-sample dataset, observations of the input conditions

The histograms 3.14 and 3.15 display the input and response parameters of the SCADA signal respectively in-sample and the out-of-sample dataset. It can be seen from these distributions that for most the signals a similar spread over the operational range is visual as for the in-sample dataset. It must be noted that for the middle high speed shaft bearing more temperatures related to the highest observed temperature are recorded.

The relation of the first high speed shaft bearing to the power curve and to the generator speed and return oil temperature is presented in figures 3.16 and 3.17 to compare with the presented data in section 3.4.3. The heat-maps for the remaining response parameters are presented in appendix G.

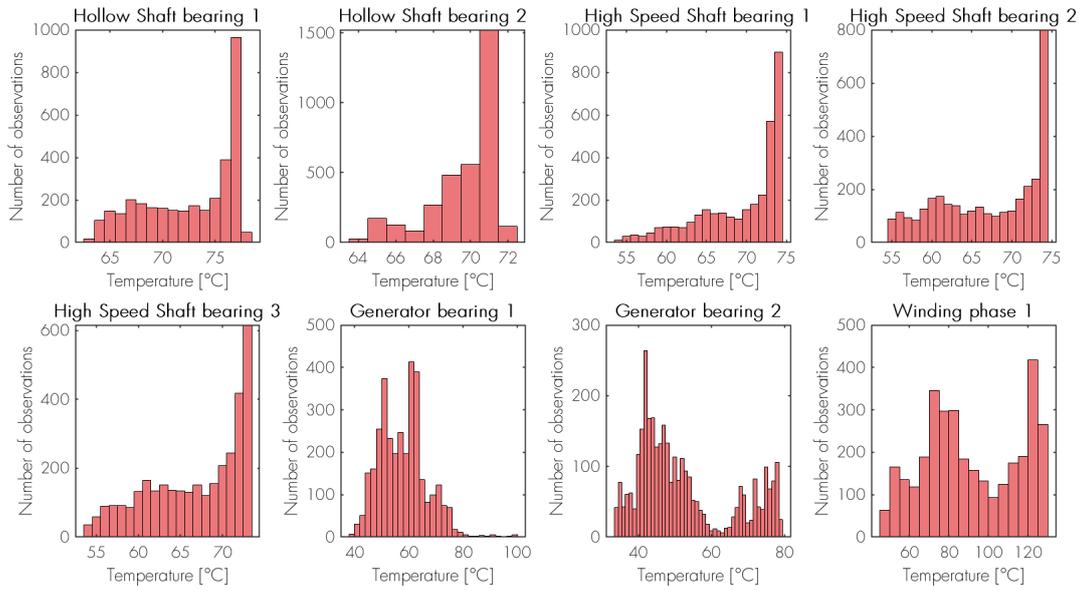


Figure 3.15: Turbine A out-of-sample dataset, observations of response conditions

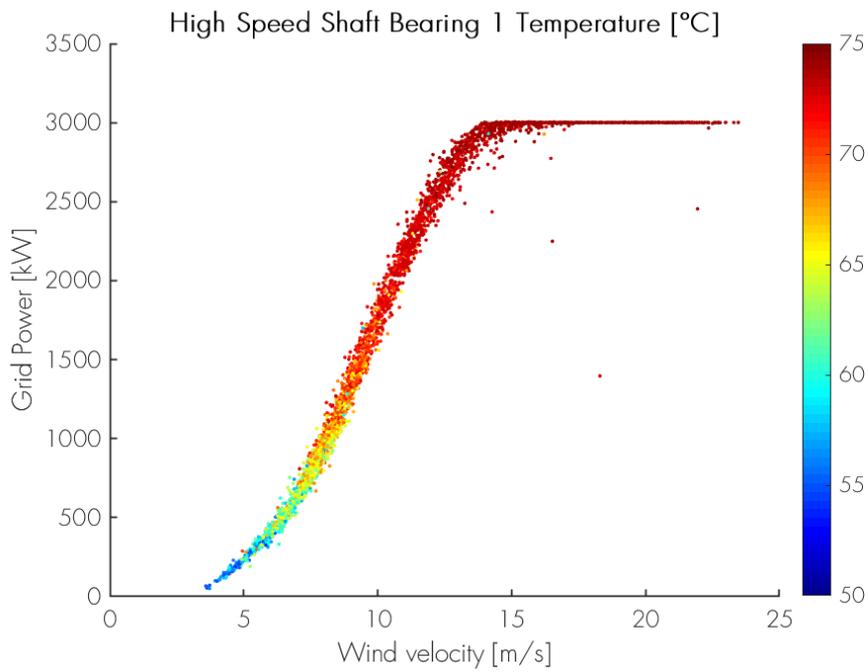


Figure 3.16: Turbine A out-of-sample dataset, High speed shaft bearing 1 in relation to the power curve

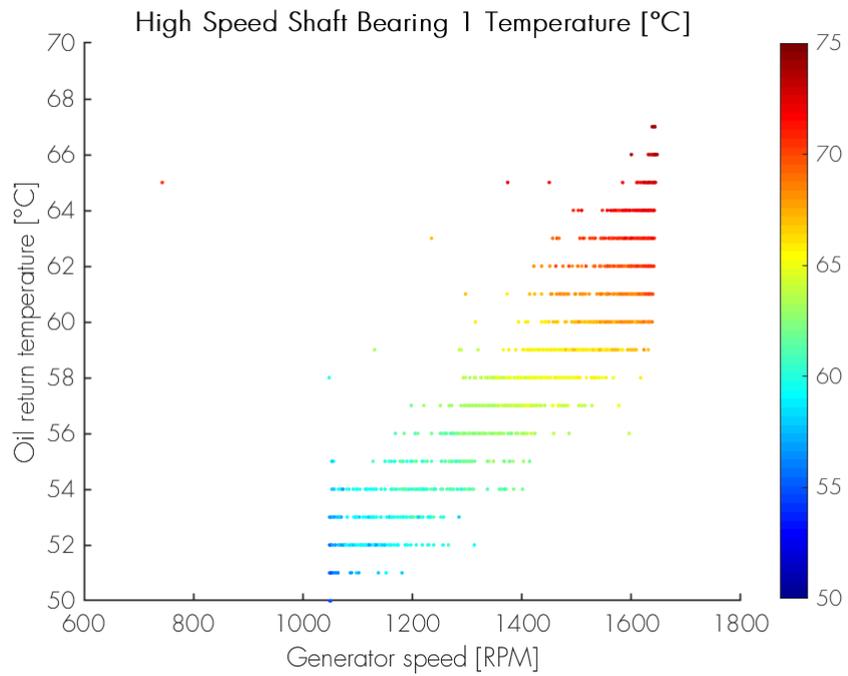


Figure 3.17: Turbine A out-of-sample dataset, High speed shaft bearing 1 in relation to the generator speed and return oil temperature

4 Data driven models

A framework is developed to perform the condition monitoring of the drivetrain components based on the SCADA data. In this framework, several data driven models can be used to describe the healthy behaviour of the drivetrain temperatures. The different implemented models that are used to unveil the valuable information hidden in the SCADA signal are covered in this section. The methods for anomaly detection, based on the model responses are discussed after the different models are covered. The different models are evaluated based on accuracy and computational speed in the last part of this chapter.

4.1 Farm average signal trending

One of the simplest anomaly detection methods is to rely on comparing the behaviour of one turbine against the farm average behaviour. This method provides a quick insight in the expected response under normal conditions and identification of noticeably deviating drivetrains.

By binning the average component temperature over a 10% interval of the rated power, different turbines can be compared against each other. A smaller interval can be taken, but Astolfi states that this does not influence the observations of the results [39].

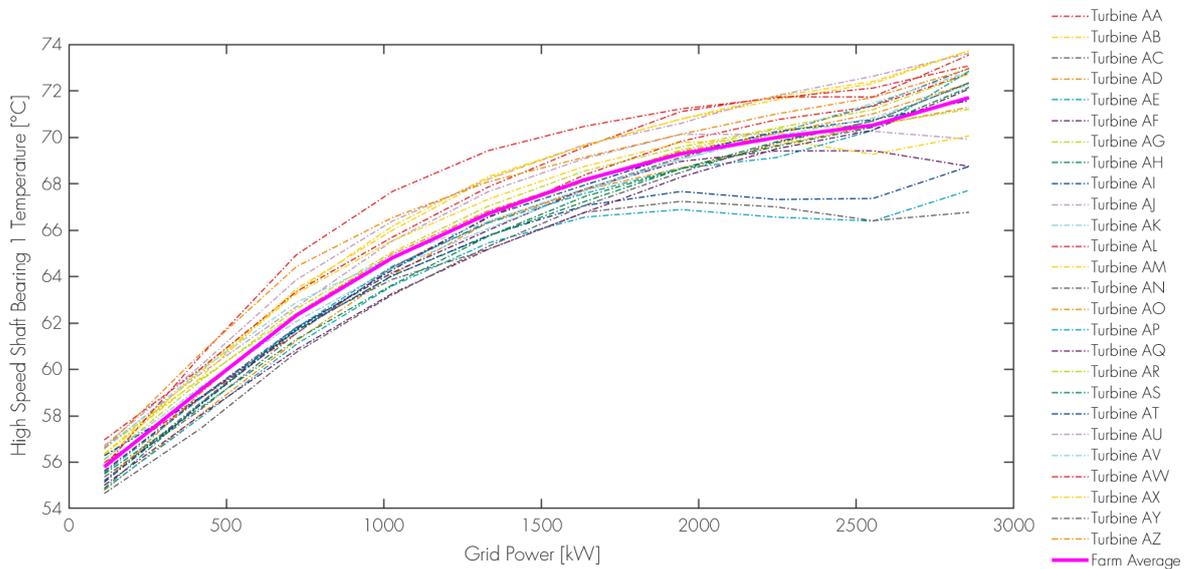


Figure 4.1: Farm average signal trending for the high speed shaft bear 1 temperatures

The temperatures of the first high speed shaft bearing are presented in figure 4.1 as an example of this approach. It can be seen from this figure that there is a large spread over the bearing temperature for different turbines. Abnormal high temperature are usually an indication of additional wear and an increased chance of failure. Turbines above the farm average are especially of interest. Turbine AL shows the most deviating

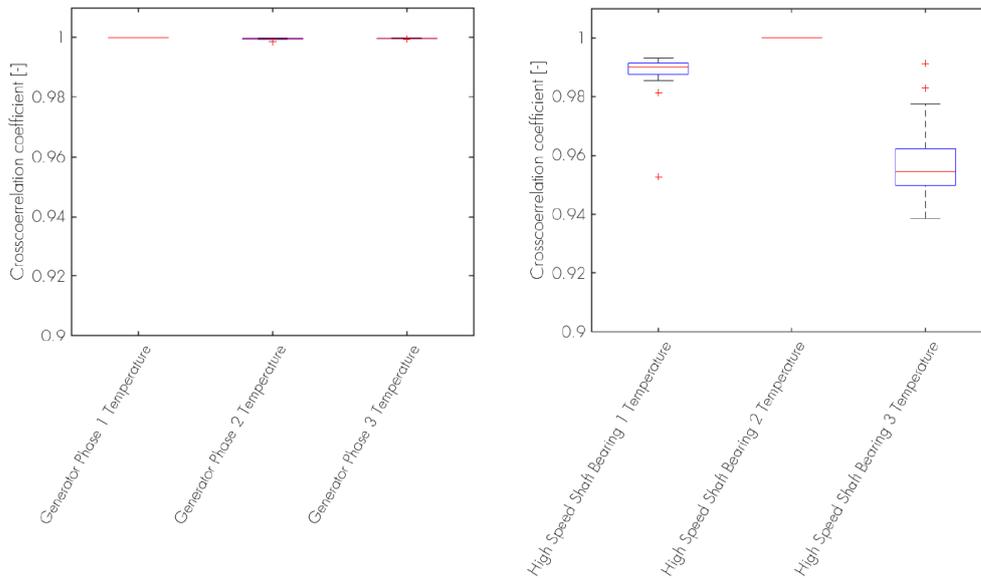
behaviour, but no conclusions can be made on the condition of the turbine at this stage as this can also be the result of badly calibrated sensors.

This method has a couple of disadvantages. Firstly, this method is dependent on visual interpretation of the trends. When analysing a great number of SCADA signals and turbines, this is not desirable. Secondly, this method is only able to detect faults that already result in strong deviations from the farm average behaviour. By analysing the change over a long time interval, it is unlikely that an anomaly can be detected in time to allow for preventive maintenance.

4.2 Thermal difference analysis

Another analysis method, is looking at the difference between different component temperatures that are very closely correlated. From the mathematical correlation and understanding of the system integration in section 3.3, two cases can be identified.

The first is by analysing the temperature difference in the generator windings of the different phases, these are expected to show little to no difference. The second case can be identified for the three bearings on the high speed shaft in the gearbox. These bearings are close to each other and are expected to perform in the same way. If one of these component temperatures start to rise compared to the others, then this can be a signal of additional friction and thus wear and degradation. For the generator, a large difference can indicate a short circuit in the windings of one of the phases.



(a) Cross correlation coefficient compared to Generator phase 1 (b) Cross correlation coefficient compared to the middle high speed shaft bearing

Figure 4.2: Parameters for the thermal difference analyses in relation to each other

The cross correlation coefficient of these parameters to each other is displayed in figure 4.2. It can be seen in figure 4.2(a) that all generator phase temperatures have a correlation coefficient of $\rho \approx 1$. This indicated full correlation. The correlation of the bearing temperatures on the high speed shaft is slightly lower, as can be seen in figure 4.2(b). But can be considered as high enough to expect similar behaviour, thus this analysis can be used.

For a healthy condition of the generator or high speed shaft bearings, near zero difference is expected. An example of healthy behaviour is presented in figure 4.3 for the temperature difference between the generator phases of turbine C. It must be noted that some turbines show a constant temperate difference between the generator phases in the range of 1 to 5°C. When evaluating this difference over the entire dataset, no change in this difference can be observed over one year or indication of faulty behaviour of the generator could be found. This can be considered as a bias due to calibration of the sensors and this phenomena should be taken into account.

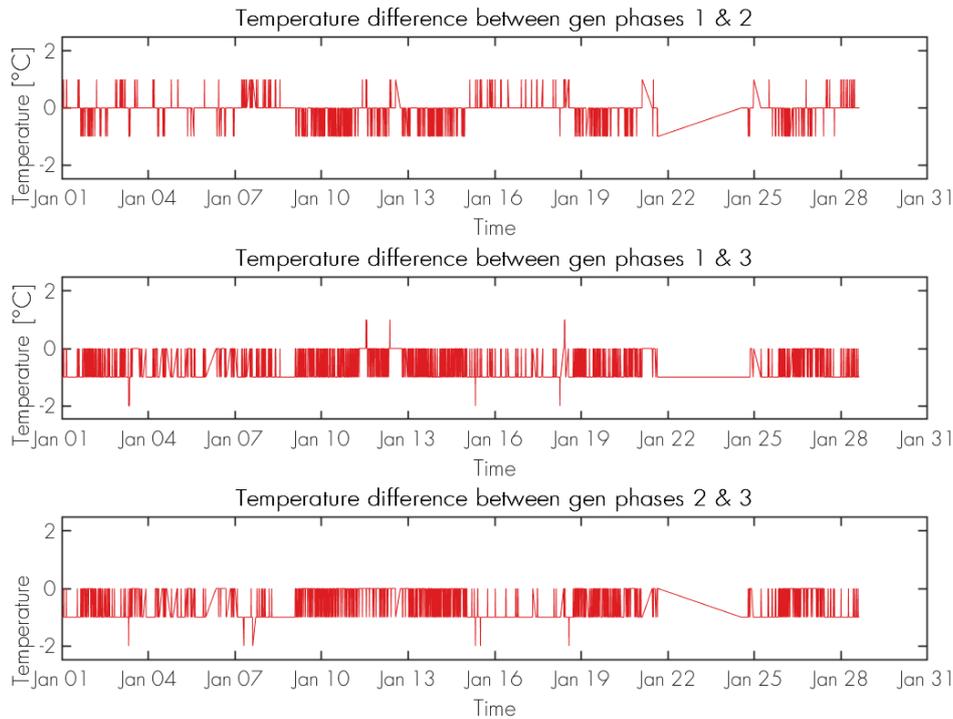


Figure 4.3: Thermal difference between generator windings in healthy condition

4.3 Normal behaviour models

For a normal behaviour model approach, the measured sensor data from the SCADA system is compared with predicted temperature responses of a model and a residual between the two is calculated. The model predicts the response at a given time instance using the operating conditions from the SCADA system as input. This concept is visualised in figure 4.4. The developed framework can utilize different models to predict the temperature response for this task. The ones most interesting for this research are discussed in this section.

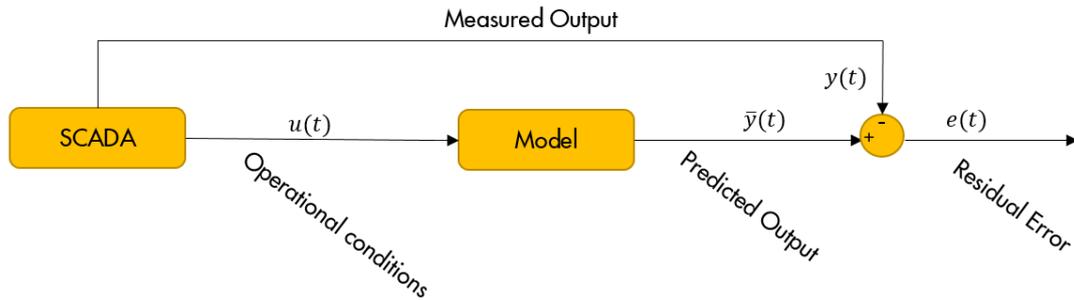


Figure 4.4: Normal behaviour model structure

The residual error, $e(t)$, is expected to be small for a well fitted model and a healthy drivetrain. When components in the drivetrain are subject to wear and degradation, the temperature is expected to increase due to friction. Thus, the predicted outcome is expected to be lower than the measured data, under the same operational conditions. This is because the model does not take the degradation into account. Analysing the residual error can thus be used to draw conclusions on the condition of the drivetrain.

4.3.1 Simple regression models

Linear regression

One of the simplest normal behaviour models is constructed using a linear relation between two parameters. This relation, also referred to as linear regression can be described with equation 4.1. Schelghingen has shown that this model can be used for temperature based condition monitoring for data with a high correlation between y and u [16],

$$\hat{y}(u) = p_1 + p_2 u. \quad (4.1)$$

The bearing temperatures are highly correlated with the delivered grid power of the turbine, thus the linear relation between these parameters is investigated. The linear regression between the bearing temperature on the first high speed shaft in the gearbox and the grid power is shown in figure 4.5. It can be seen from this figure that the linear regression does not properly describe the observed trend within the data. Because of this, this model can be considered as under fitted, as explained in section 2.3.2.

Higher order regression

To improve on the concept of linear regression, the relation between input and response is investigated using higher order polynomials as given by equation 4.2:

$$\hat{y}(u) = p_1 + p_2u + p_3u^2 \cdots p_nu^{n-1}. \quad (4.2)$$

The number of polynomials needed is investigated by looking at the relation between the grid power delivered by the turbine and the different component temperatures in the drivetrain. For each relation, the RMSE is calculated for different order polynomial curves fitted over the data. The cases and errors are given in table 4.1.

Grid Power vs ...	n=1	n=2	n=3	n=4	n=5	n=6	n=7	n=8
Hollow Shaft Bearing 1	1.976	1.603	1.603	1.602	1.601	1.601	1.601	1.602
High Speed Shaft Bearing 1	2.916	2.097	2.044	2.044	2.044	2.043	2.043	2.043
Generator Bearing 1	1.779	1.759	1.745	1.741	1.733	1.714	1.714	1.714
Gearbox oil outlet	2.160	1.574	1.574	1.545	1.539	1.533	1.532	1.532

Table 4.1: RMSE in °C for different order polynomials and different response signals

The improvement from linear regression, $n = 1$, is evident from table 4.1. The selected order of polynomial used for the model is $n = 6$. For this order, the RMSE is not decreasing compared to higher order polynomials, when looking at the different component temperature signals.

The improvement compared to the linear fit through the data is also evident from figure 4.5. Here, it can be seen that the 6th order polynomial is able to follow the observed trend in the data better than the linear regression.

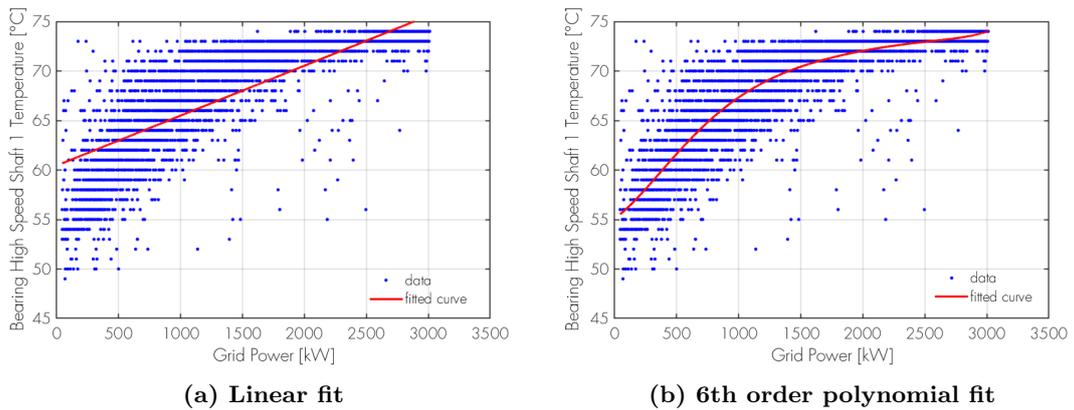


Figure 4.5: Regression functions fitted to the high speed shaft bearing 1 temperatures

4.3.2 Decision Trees

The first decision tree regression algorithm was published by Morgan and Sonquist in 1963. Since that time, this learning algorithm has been widely used in the field of machine learning [40].

The approach for decision trees is to split the data based on decision criteria, for example $X \leq c$. Here is X the original set of data and c the decision criteria. These decision criteria are referred to as nodes and split datasets in the so called children. The node $X \leq c$, creates two children. A and B , where $A \subset X$ and $B \subset X$.

The concept of a decision tree is visualized in figure 4.6. All the coloured squares in figure 4.6 represent elements in the set X . The first node splits the data in yellow children and other colour children. The second decision criteria splits the dataset A again in two subsets, this can be seen as the blue and red squares.

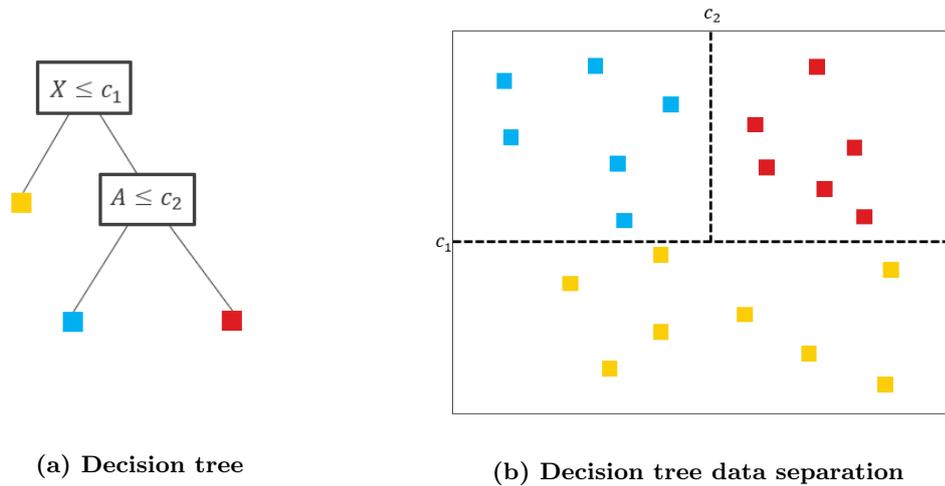


Figure 4.6: Concept of decision trees data splitting process

In reality, data will not always let itself split cleanly due to outliers and impurities in the data. To deal with this, the learning algorithm can optimize the splitting criteria so that the sum of wrongly label elements in the two children is as low as possible [40]. The learning algorithm continues growing the tree by splitting the data until the stopping rule is satisfied. This rule stops the splitting of the data if the number of impurities in the children is lower then a given threshold value, as can be seen in equation 4.3. Here, I is the notation for the number of impurities in dataset A and B while β is the threshold value of the stopping rule [32]:

$$\max(I(A, B)) < \beta. \tag{4.3}$$

The decision criteria $X \leq c$ is referred to as Automatic Interaction Detection (AID) and is developed as the first decision tree by Morgan and Sonquist [41]. There are also other decision tree learning algorithms like: THeta Automatic Interaction Detec-

tion (THAID), CHi-squared Automatic Interaction Detector (CHAID), C4.5, FACT or QUEST. A visualisation of the data splitting using these criteria is provided in appendix H and more information can be found in [40]. For the application of this research, an understanding of AID is sufficient.

The regression learner toolbox and classification learner toolbox used in MATLAB utilizes classification and regression trees based on the Breimans theory of decision trees from 1984 [32]. This method constructs a decision tree based on AID but implements some additional steps. The main addition is in the approach to the right size of the decision tree. The issue with the stopping rule, given by equation 4.3, is that this is prone to over fitting as it can grow to children that only contain one data point.

The addition of Breiman is to not use the stopping rule but grow the tree using AID until it is so large that every subset only contains one data point or a small amount of data points. This large tree is denoted with T_{max} . This is in essence, intentionally over fitting of the model. After this step, the tree is pruned back by removing subtrees to tree t , where $t \in T_{max}$. The removal of these nodes is based on a minimum cost-complexity function [32]. This approach eliminates the issue with over fitting. An example of a cost-complexity function is given in figure 4.7. Here, the cost is defined as the model error and the complexity is defined as the size of the decision tree.

A decision tree can be used for classification or regression tasks, depending if the target

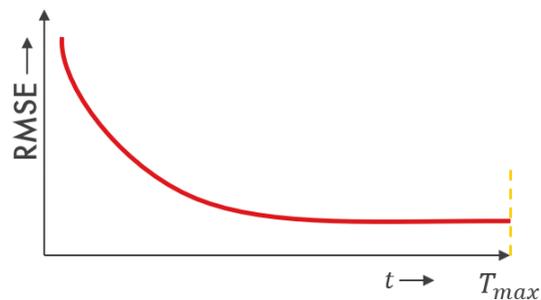


Figure 4.7: Concept of a complexity-cost function for the size of a decision tree

is discrete a continuous value [42]. If the model is performed with a classification task, then it can only generalize over the experienced temperatures. Whilst for regression, also the values in between these classes can be predicted. Both these tasks have been used in the research and will be discussed.

To build on this, an optimization is sought after for the hyperparameter(s) used for the learning algorithm. To only hyperparameter that is used for regression or classification trees, used by the toolboxes, is the constraint for the minimum observations in each subset of data after a node. This will be referred to as the minimum leaf size. The object for the optimization of this hyperparameter is to minimise the loss function. Figure 4.8 is provided as an example of this process, it can be concluded that for the middle high speed shaft bearing on turbine D, the most optimal hyperparameter can be found as a minimum leaf size of 20.

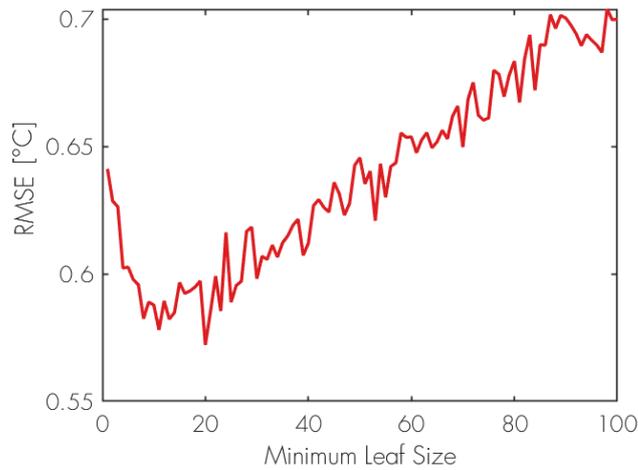


Figure 4.8: Hyperparameter optimization for decision trees

This optimization is only valid for this turbine and response. For different turbines or responses, a different value can be found as optimal minimum leaf size. Because of this, an optimization is build in the toolbox to make use of the most optimal hyperparameter value when training the model.

The first couple of nodes and children and an overview of the entire resulting decision tree for the first high speed shaft bearing temperature of turbine A is visualised in figures 4.9 and 4.10 as an example. The trained decision tree model for figure 4.10 contains 283 children.

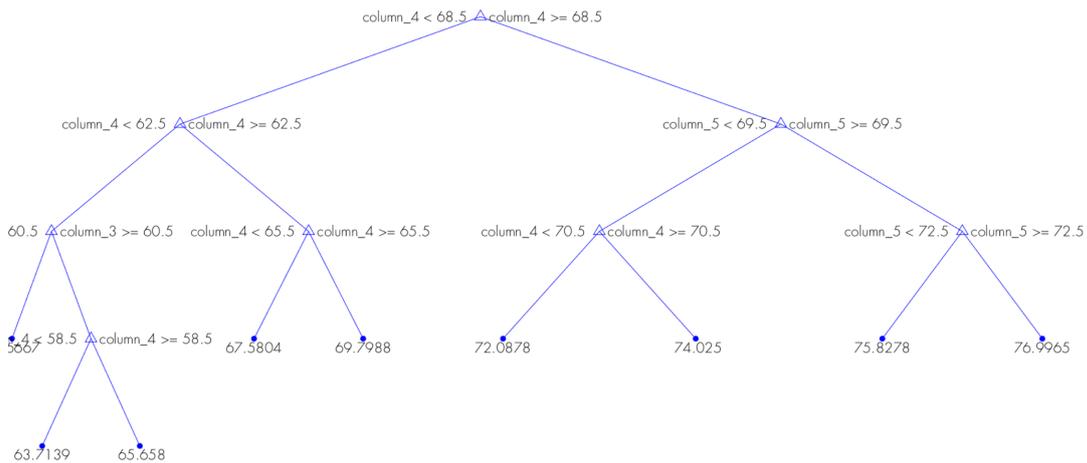


Figure 4.9: First 8 nodes and 9 children of the generated decision tree

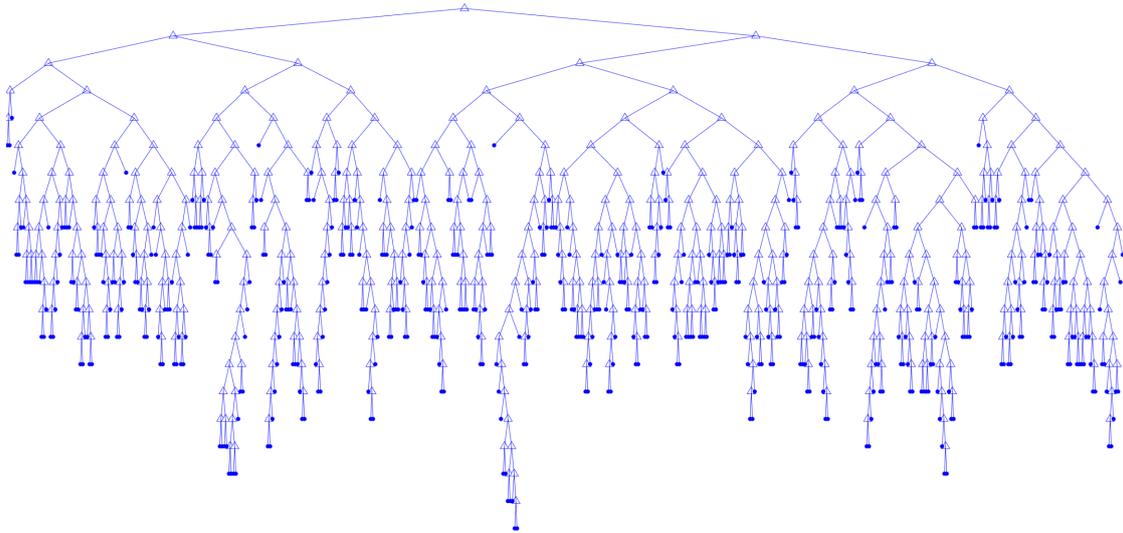


Figure 4.10: Overview of the generated decision tree

4.3.3 Support Vector Machine

The theory of Support Vector Machines (SVM) was developed by Vapnik in the early 60's but only started becoming popular in the early 90's [43]. In this section, the principles behind support vector machines are explained as well as some optimizations and considerations for the application in this research.

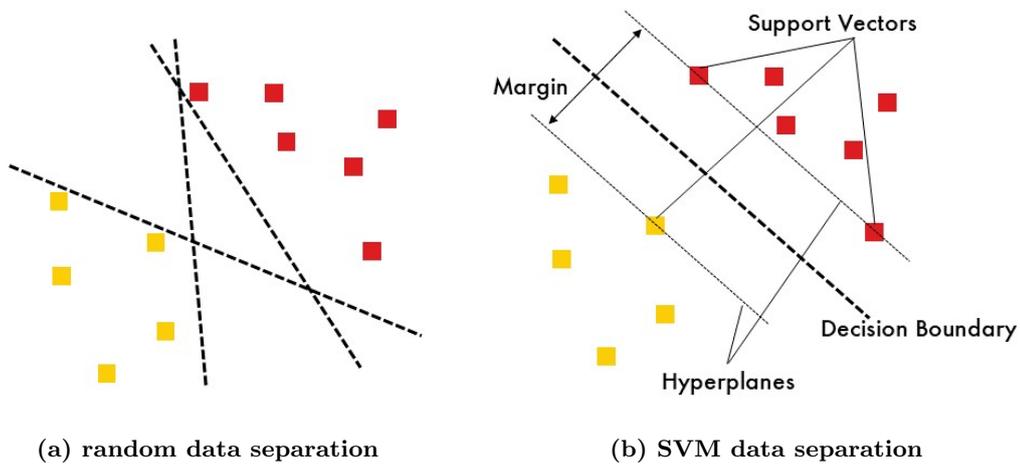


Figure 4.11: Concept of support vector machine data separation

Linear separable and non separable support vector machines

The aim of support vector machines is to linear separate data belonging to different classes by using a hyperplane. With this condition, the linear separation of the data can be done in an infinite number of ways, as shown in figure 4.11(a). A support vector machine does this in the most optimal way, this is defined as the hyperplane that results in the largest margin to both data populations as shown in figure 4.11(b). The data point used to construct the hyperplanes are considered as the support vectors.

In reality, not always a perfect linear separation of the data is possible. This is due to impurities in the data, as can be seen in figure 4.12. Here, ξ represents a slack variable for the distance from the decision boundary to the outlier. By introducing the Karush-Kuhn-Tucker condition in the process for deciding the optimal hyperplane, the number of outliers, $\sum_i \xi_i$, versus the most optimal hyperplane can be taken into account [44]. By implementing the regression learner and classification learner toolbox in MATLAB, the Karush-Kuhn-Tucker condition is met for the lowest possible loss function.

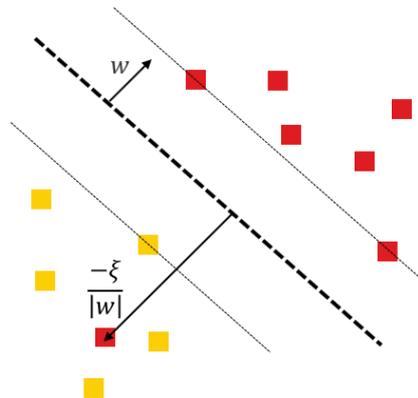


Figure 4.12: Concept of a support vector machine including impurities

Non-linear support vector machines

When it is not possible to separate the data linearly, then the data can be mapped in an enhanced feature space where the data is linearly separable. The mapping in this higher dimensional feature space is done by using kernel functions. This process is illustrated in figure 4.13.

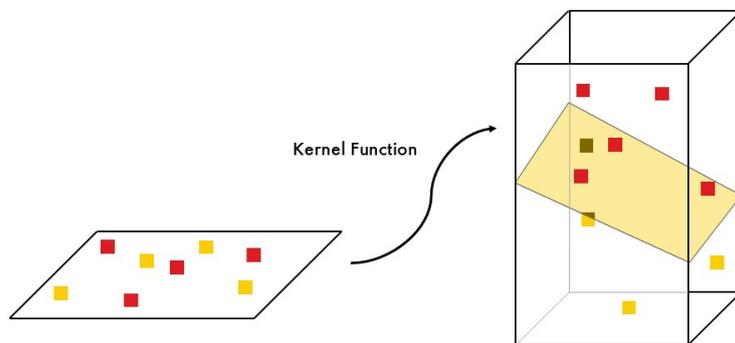


Figure 4.13: Concept of using kernel functions to map in an enhance feature space

The most common kernel functions are the linear, polynomial and radial basis kernel function [28], given by equations 4.4, 4.5 and 4.6. Here, equation 4.4 is a linear kernel function. Equation 4.5 is a polynomial kernel function, thus for $q = 2$ a quadratic kernel function is given and for $q = 3$ a cubic kernel function. Equation 4.6 represent a radial base kernel function,

$$G(x, y) = x'y, \quad (4.4)$$

$$G(x, y) = (1 + x'y)^q, \quad (4.5)$$

$$G(x, y) = e^{-\frac{\|x-y\|^2}{2\sigma^2}}. \quad (4.6)$$

A sensitivity study is performed to determine which kernel function performs best for the application of this research. A support vector machine is trained 50 time for each kernel function to predict the first high speed shaft bearing of turbine A using different kernel functions. The resulting validation error can be found in table 4.2 and the distribution of these errors over 10 bins is shown in figure 4.14. Also the computational speed of the different models is evaluated and presented in table 4.2. For this, the average time is taken for the k-fold cross validation for the 50 times each model is trained.

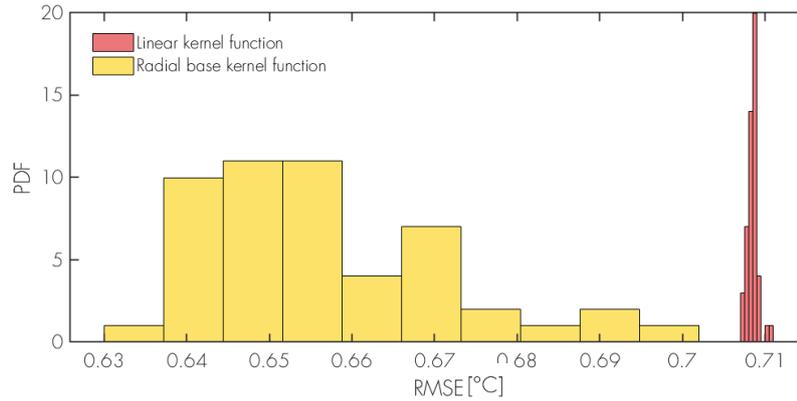


Figure 4.14: validation error of SVM regression with different kernel functions

	Min RMSE	Mean RMSE	Computational Time
Linear Kernel	0.7070 [°C]	0.7085 [°C]	23.52 [-]
Quadratic Kernel*	0.6501 [°C]	0.6528 [°C]	225.89 [-]
Radial Base Kernel	0.6349 [°C]	0.6572 [°C]	4.70 [-]

Table 4.2: Validation error distribution

* trained 5 time in stead of 50.

The computational time is normalized by 1 sec, as this is only meant to be as an indication of the required runtime to train a model. The eventual computational time that is needed to train the model will depend on the length of the training period and

the computational speed of the computer that is used. It must also be noted that the quadratic Kernel is only trained 5 times in stead of 50, due to the required time to train the model, whilst not resulting in a more accurate model then when the radial base kernel function is used.

It can be concluded from this sensitivity study that the radial base kernel function results in the most accurate model. Also, the time that the learning algorithm needs is lower then for the other kernel functions. For the remainder of this research, a radial base kernel function is used when training support vector machines.

Until now, the use of SVM's for classification have been discussed. i.e. assign a class based on the separation of the input vectors by a hyperplane. The SVM can also be used for regression in stead of classification. For this, the data is not being separated, but a linear regression function is fitted in the enhanced feature space with an ϵ -insensitive zone [28]. With this approach, the output can be prediction along a continuous range instead of discrete classes. This process is displayed in figure 4.15.

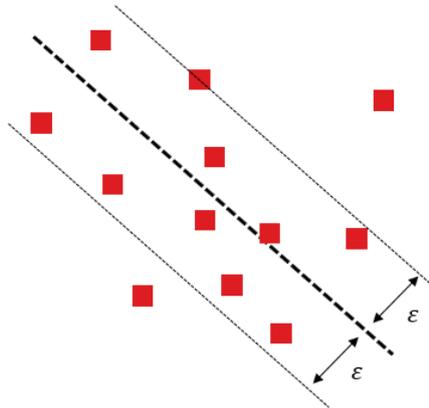


Figure 4.15: concept of ϵ insensitive SVM regression

The data points outside of the insensitivity zone will not be used for the for the construction of the regression line. The default setting for ϵ in the regression learner toolbox is as given in equation 4.7. Here, Q_3 and Q_1 are the value of the 75th and the 25th percentile of the response values:

$$\epsilon = \frac{Q_3 - Q_1}{13.49}. \tag{4.7}$$

To improve on this, the sensitivity of ϵ on the validation error of the model is tested on various turbines and output signals by training the model with different values for ϵ . As an example, the sensitivity for the first high speed shaft bearing of turbine D is presented in figure 4.16. Here, it can be seen that for smaller values of ϵ , a model with a smaller error is constructed. A significant increase in the error is visible for values larger that 0.5 times the normalised ϵ , the default setting of 1/13.49 is well below this. In addition to this, the model was tested for over fitting by applying the model on a

unfamiliar dataset, this also resulted in a small error over the new data. Thus there was no over-fitting with the given default ε .

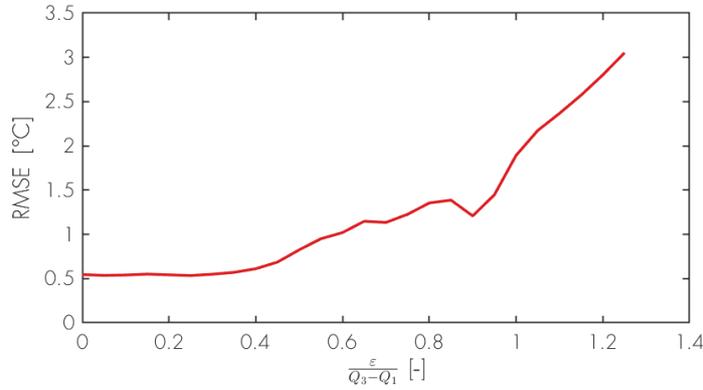


Figure 4.16: Hyperparameter ε sensitivity

4.4 Fault detection methods

Once the residual error between the model response and the measured signal from the SCADA system is calculated, this can be analysed to investigate if the system is in or out-of-control. Two methods will be covered in this section, the first one will be utilized to detect large deviations from normal behaviour and the second method will be used to detect shifts of the mean value of the response.

4.4.1 Shewhart control chart

One of the most common practises in quality control is using a control chart developed by Shewhart [45]. The concept is that when the variation is between a range of normal variations, the system is considered in a state of statistical control. Warning and control limits can be introduced to monitor the state of a component as displayed in figure 4.17.

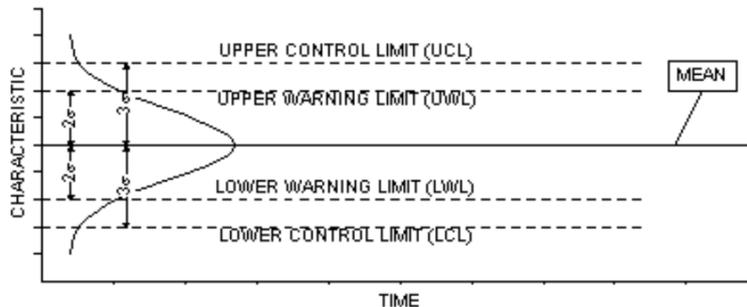


Figure 4.17: Shewart control chart

The control limits that are assigned to the range of normal variations are given by equation 4.8 until 4.10. For the example given in figure 4.17, $A = 3$ is taken as the control

limit and $A=2$ as a warning limit. Schlechtingen proposes a threshold value of $A=5$ [16] and this value is adopted for this research when the Shewhart control chart is applied on the residual error with a target mean of zero.

The advantage of this method over using a predetermined threshold value is that a well trained model will result in a lower control limit and anomalies can be detected as early as possible while false alarms are limited when a model contains more variations under healthy conditions.

$$\text{Lower limit} = \mu_0 - A\sigma, \quad (4.8)$$

$$\text{Central limit} = \mu_0, \quad (4.9)$$

$$\text{Upper limit} = \mu_0 + A\sigma. \quad (4.10)$$

This method is a powerful tool to detect anomalies that result in a large deviation from the normal conditions. A fault in the drivetrain will result in such a large deviations in late stages before failure and is therefore a reliable indicator.

4.4.2 Cusum control chart

The Shewhart control chart is an efficient approach to detect anomalies that result in a large residual error. These large errors are often linked to a significant fault in the drivetrain. To detect the fault as early as possible, it is also important to track smaller shifts from the zero mean of the residual error. For this, the cumulative error sum (CUSUM) analysis is a good substitute, as this is able to track shifts smaller than 1.5σ [46].

The simplest form of a cumulative sum of the residual error, e , is given by equation 4.11. With this analysis, a deviation between the model and the measured signal is tracked. Since a residual error of zero indicate a healthy state for the drivetrain, the target mean, μ_0 , is set to zero. The accumulated C is calculated for every instance i , where $i \in \{1..n\}$ with n being the number of data points in the dataset.

$$C_i = \sum_{i=1}^n (e_i - \mu_0). \quad (4.11)$$

An improvement on equation 4.11, can be made by utilizing the tabular CUSUM developed by Montgomery [46] to allow for some small variations in the residual error and reduce the sensitivity to noise. The tabular CUSUM consisted of two weighted averages, C^+ and C^- and can be calculated with equation 4.12. Here, C^+ is an accumulated weighted average of deviations above the target mean, μ_0 , and C^- represent that same parameter only for accumulated deviations below the target mean [46]:

$$\begin{aligned} C_i^+ &= \max(0, C_{i-1}^+ + e_i - \mu_0 - K), \\ C_i^- &= \min(0, C_{i-1}^- + e_i - \mu_0 + K), \end{aligned} \quad (4.12)$$

with:

$$C_1^+ = C_1^- = 0. \quad (4.13)$$

The benefit of having a 2 sided CUSUM over an one sided CUSUM is to quickly detect a shift in negative or positive variation. For example, if a positive shift is build up first but stays below the control limits, the distance to the negative control limit is larger then when a negative shift starts at 0. Thus, a positive accumulated error followed by a negative shift in the temperature will sooner by marked as critical by a two sided CUSUM then by an one sided CUSUM. This benefit can also be seen in the visualized concept of a CUSUM control chart in figure 4.18.

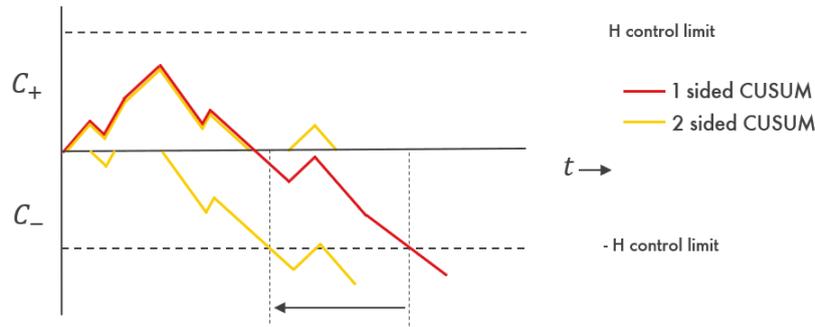


Figure 4.18: Concept of 1 and 2 sided CUSUM control charts

K in equation 4.12 represents the so called allowance or slack value. This value reduces the sensitivity of the analysis to noise in the signal and prevent false alarms. K is usually chosen so it is a fraction of the interval between the target mean μ_o and the control limit. Since the control limit is expressed as a multitude of the standard deviations, so is the slack value. The value of $\delta = 1$ is taken as proposed by Montgomery [46].

$$K = \frac{\delta}{2}\sigma. \quad (4.14)$$

The control limits for the CUSUM analysis are usually given as a multitude of the standard deviation, given by equation 4.15. An out-of-control situation can be considered when $C_i^+ > H$ or $C_i^- < -H$. The value of h should, according to Montgomery, be determined with a known failure case. When the mean of the system is shifted out-of-control, one can count backwards to the first time CUSUM is lifted significantly and use that value as H [46]. Since no information on failure cases is available, the value of h should be estimated.

$$H = h\sigma. \quad (4.15)$$

An example of the two control charts is given in figure 4.19. Here, the time series for the generalized output of the model, \hat{y} , and the measured temperature by the SCADA system, y , are given in the top graph. The middle graph represents the residual error, e , between \hat{y} and y with the Sherhard control limits introduced at five times the standard deviation of the error during the training period. Here, it can be seen that this threshold value is crossed. In the bottom graph, the CUSUM control chart is presented. The shift

in under under and over estimations of the model is monitored by the 2 weighted averages.

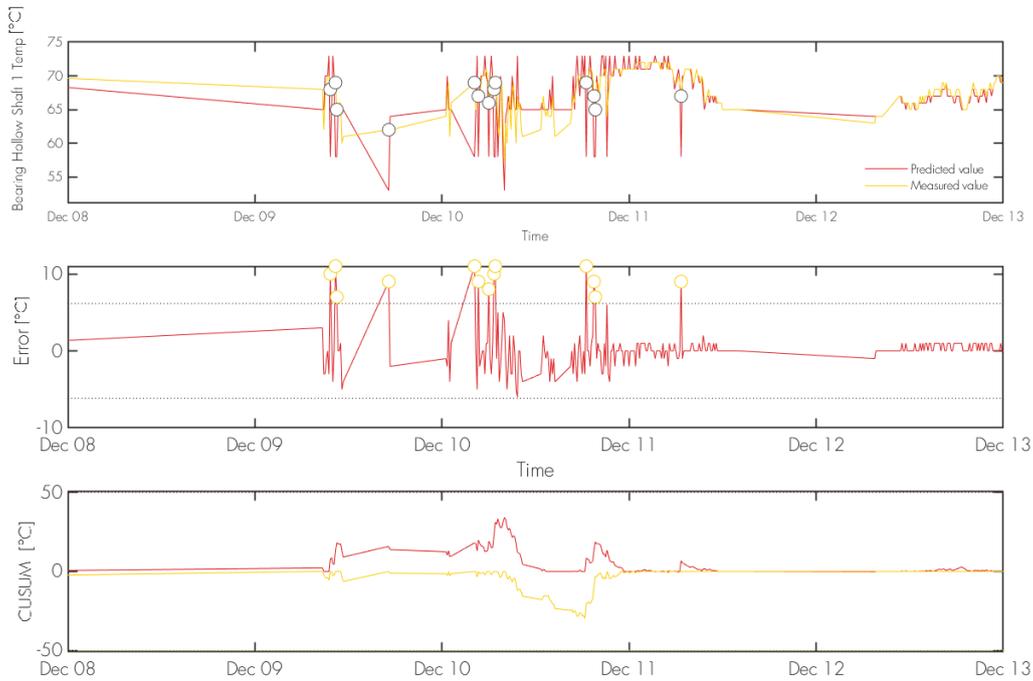


Figure 4.19: Top graph: time series for \hat{y} and y . Middle graph: the Sherhard controlchart on e . Bottom graph: the CUSUM control chart

4.5 Model evaluation

Now that the different models and methods for anomaly detection are covered, the most suitable normal behaviour model can be selected. To do so, two criteria are covered. First, the ability of the different models to describe the functional behaviour in the in-sample data is covered. This is an important indicator for the model accuracy. This is of importance because an accurate model will result in as early detection of a fault as possible and limit the number of false alarms. Secondly the training speed of the different models is discussed. Since a great number of signals and turbines need to be evaluated, this criteria is of importance.

4.5.1 Model accuracy

As discussed before, the most suitable model will be selected on several criteria the most important one of this, is the accuracy of the model. For this the RMSE over the validation dataset using k-fold cross validation is used as indication of how well the model can describe the functional behaviour of the data. More on the RMSE and k-fold cross validation can be found in sections 2.3.3 and 2.3.4.

For the supervised machine learning models, the validation error will differ slightly for each iteration of training over the same dataset. This is due to small difference in weights distribution over the parameters. Because of this, the distribution of the error is looked at after training the model 100 times. The errors related to the models trained on data from the first hollow shaft bearing and the middle high speed shaft bearing of turbine A, are highlighted in figures 4.20 and 4.21. The lowest RMSE after 100 iterations of training for the remaining responses are presented in table 4.3.

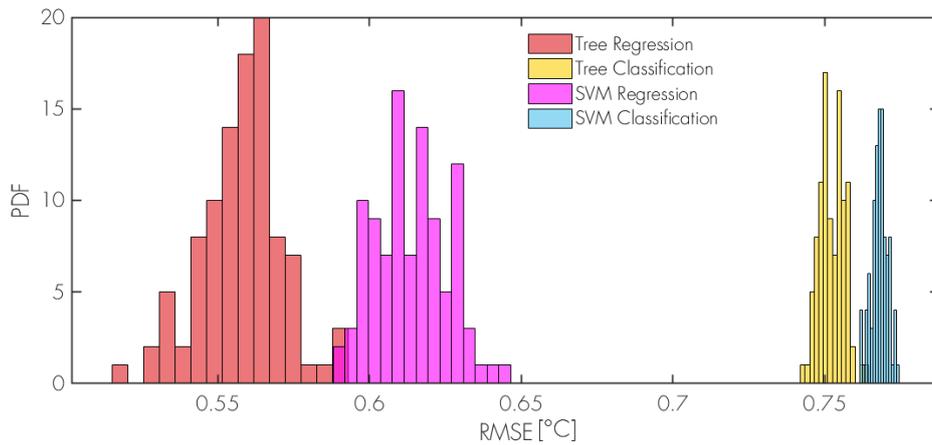


Figure 4.20: Distribution of the validation error for predicting the hollow shaft bearing 1

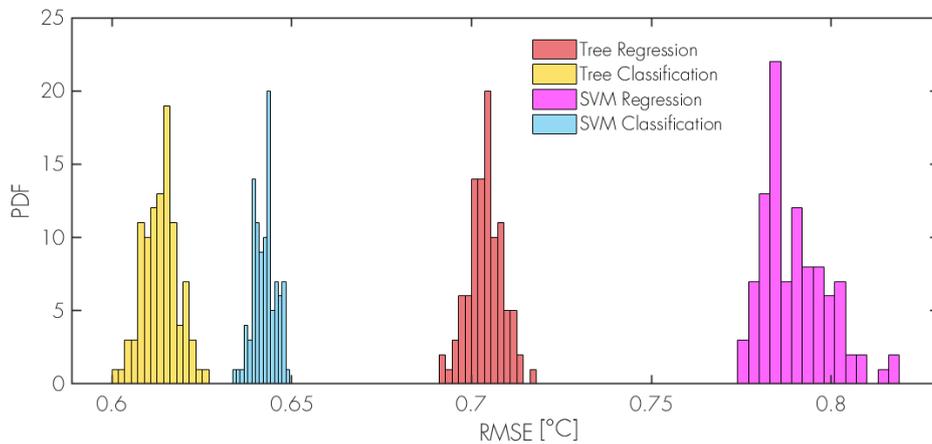


Figure 4.21: Distribution of the validation error for predicting the high speed shaft bearing 2

It can be seen from figures 4.20 and 4.21 that the model with the smallest error can differ per case. It can be seen from table 4.3 that for all the response parameters in the gearbox, the decision tree will result in the validation smallest error. The highlighted case of the middle high speed shaft bearing from figure 4.21, was the only response where the classification outperformed the regression task.

Response	Regression				Classification	
	Linear	Polynomial	Trees	SVM	Trees	SVM
Hollow Shaft 1	4.32	3.77	0.526	0.612	0.745	0.772
Hollow Shaft 2	4.30	3.67	0.499	0.609	0.766	0.756
High Speed Shaft 1	5.04	4.40	0.555	0.637	0.686	0.747
High Speed Shaft 2	5.10	4.49	0.693	0.783	0.603	0.644
High Speed Shaft 3	5.01	4.19	0.552	0.640	0.645	0.705
Generator 1	8.31	8.16	3.61	3.62	0.318	0.463
Generator 2	9.38	9.28	7.14	7.19	0.301	0.397

Table 4.3: minimum RMSE in °C after training the model 100 times on the in-sample data

It can be seen that the validation error associated with both the generator bearing temperature signals is large for the regression models and very small for the classification models. Before deciding on implementing a classification model, the models should be tested for over fitting. To do so, the models are used to generalize over a test dataset opposed to the validation dataset.

For the generalization over a test dataset, two cases are highlighted as an example. The first case is the first hollow shaft bearing temperatures using a decision tree classification model. The small errors over the test dataset, containing 3350 number of data points, can be seen in figure 4.22. The resulting test error is a RMSE of 0.528 °C, and a validation error of 0.744 °C could be seen from table 4.3.

The second highlighted case is a model for the first generator bearing bearing temperatures that also utilized a decision tree classification learning algorithm. It can be seen from both figure 4.23 and the test error with a RMSE of 7.58 °C, that the model is unable to fit over the test data whilst having a RMSE of 0.318 °C over the validation dataset. This indicates that there is over fitting for the generator bearing temperatures and not for the gearbox bearings.

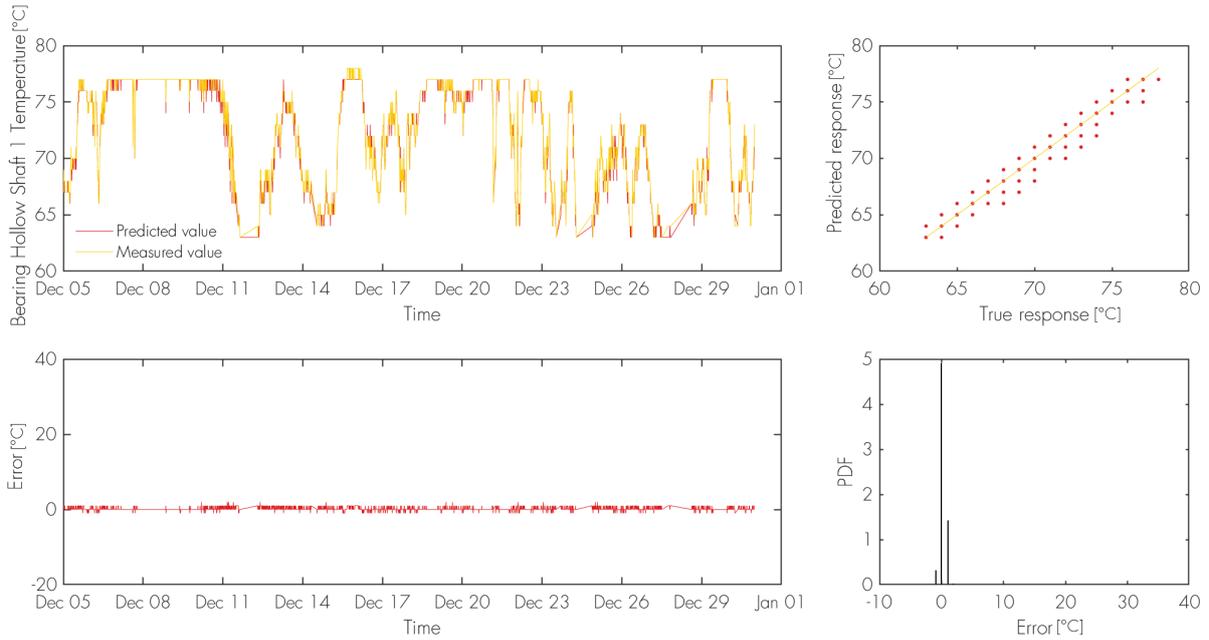


Figure 4.22: Decision tree classification generalization gearbox bearing

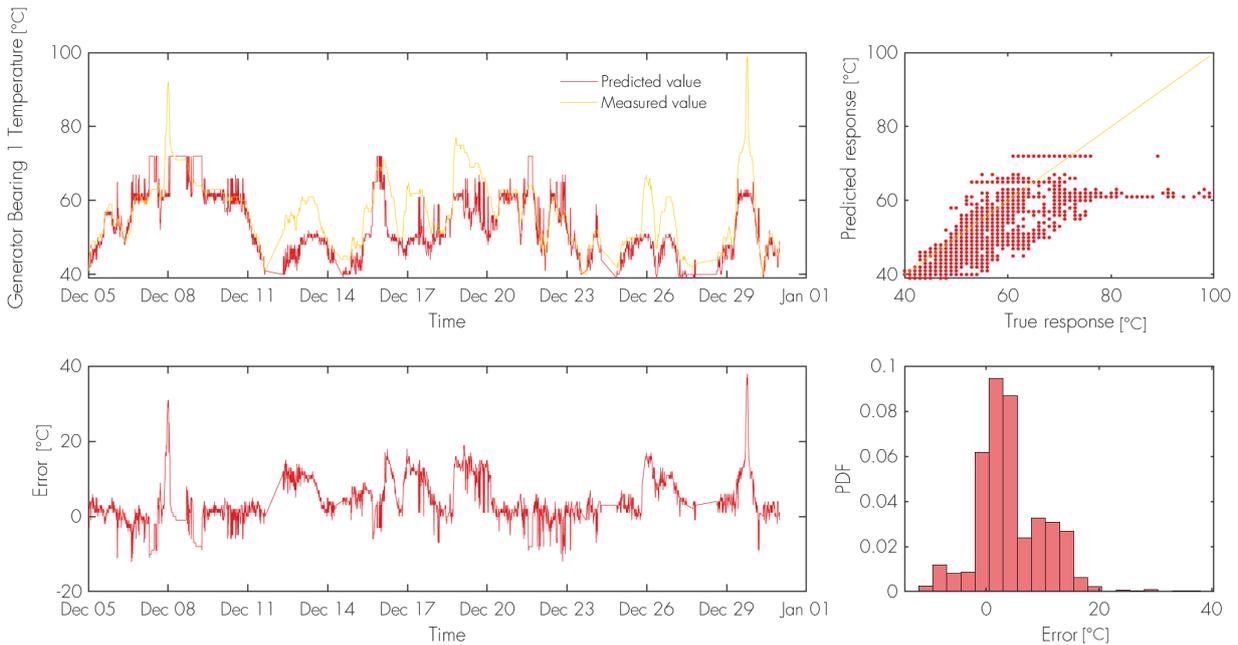


Figure 4.23: Decision tree classification generalization generator bearing

4.5.2 Model speed

In modern offshore wind farms, thousands of signals are usually collected over hundreds of turbines. The framework can be made fit for the future by implementing models that can quickly process all these signals. The time that is needed to train the different supervised machine learning models vary greatly.

An indication to compare the required time that is needed to train different models on the same dataset is presented in table 4.4. These values are obtained by taking the average time required to train a model after 50 iterations and normalizing this by 1 second. The eventual time that is required to train the different models depends on the length of the training period and the computational power of the used computer.

Model	Speed
<i>Training period is 1 month</i>	
Linear regression	0.010
6th order polynomial regression	0.014
Decision Trees Regression	0.41
Decision Trees Classification	0.32
Support Vector Machines Regression	3.64
Support Vector Machines Classification	17.1

Table 4.4: Required run time to train different models

It can be concluded from table 4.4, that the support vector machines take the most time to train and the supervised machine learning model that utilizes decision trees were able to train the model under 1 second for this case. To select the best possible decision tree model, a further optimization can be made. From figure 4.20, it can be seen that the model has a significant spread over the RMSE after training it 100 times. The short training time that is required, enables the implementation of training the model 100 times. The model with the lowest RMSE is selected and the remaining models are discarded.

5 Condition monitoring results

The different models that are implemented in the developed framework are discussed in this section. For this, the models are tested on a turbine in healthy conditions for both the in-sample and out-of-sample dataset. When the best performing model is identified, then this model is used to identify anomalies in the entire dataset. The findings related to the results will be discussed in the last part of this section.

5.1 Thermal difference results

As covered in section 4.2, the difference between the winding temperatures in the three phases of the generator and the thermal difference between the bearings on the high speed shaft in the gearbox are analysed to detect abnormal behaviour.

For healthy conditions, small temperature differences are found, as can be seen for the generator winding temperatures of turbine E in figure 5.1. The data in the SCADA system is stored as integer data, thus the difference of -1°C or 1°C can be the result of a rounding error. Not only is the mean of the temperature difference between the phases near zero, the variance is also small.

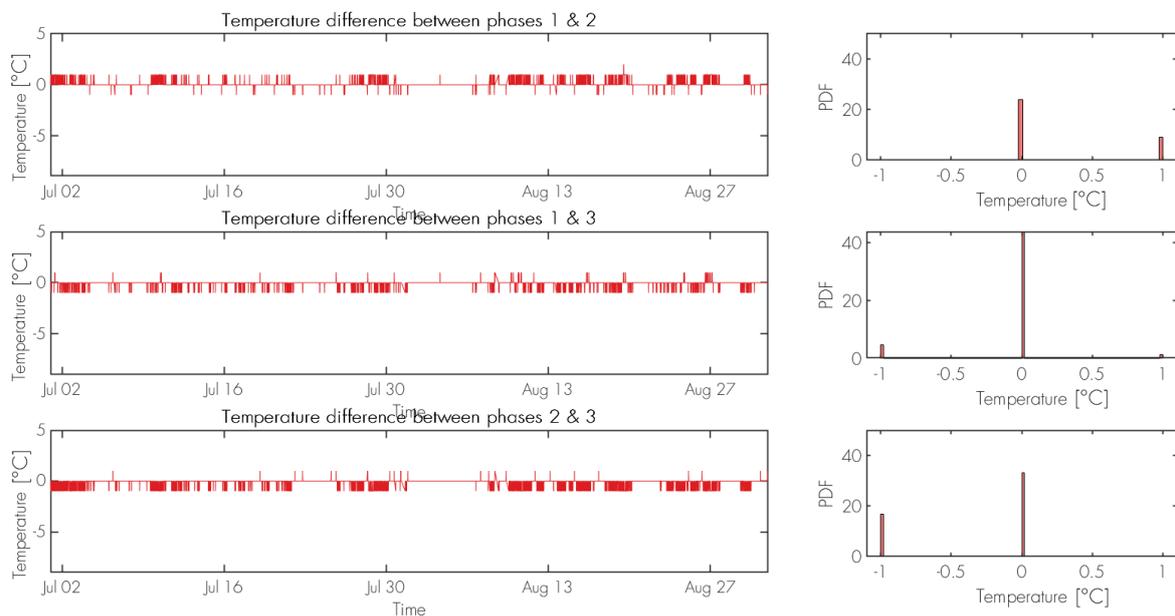


Figure 5.1: Temperature difference between generator phase windings, turbine E

When analysing all the turbines over the period of one year, a couple of interesting results could be found. For example, the thermal difference of the winding temperatures in turbine F is presented in figure 5.2. It can be seen that for this case the thermal difference between the winding temperatures show a larger variability and is not concentrated around zero.

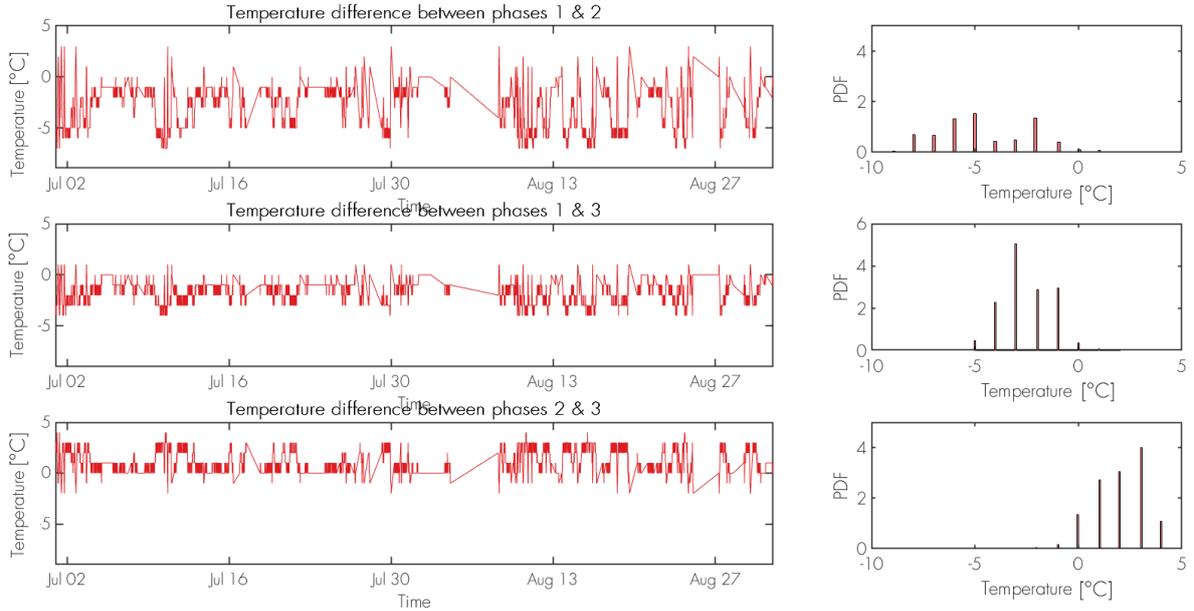


Figure 5.2: Temperature difference between generator phase windings, turbine F

This anomaly can not directly be related to a fault in the generator, because of 2 reasons. The first one is that this behaviour is showing throughout the entire dataset. Thus, if it is related to a fault then this shows no progression or worsening. Secondly, this behaviour was found for 5 other turbines. This implies that if this behaviour is related to a fault, it is not destructive as it does not result in the compromise of the turbines operation. This deviating behaviour can also be the result of badly calibrated sensors or lots of noise in the signal.

5.2 Normal Behaviour models under healthy conditions

The performance of all the normal behaviour models, that are discussed in section 4.3, are evaluated in this section. The validation error over the in-sample dataset is used to analyse how well the model can describe the functional behaviour in the data. The generalization performance is tested on the out-of-sample dataset, this dataset is also used to test the model for over fitting. The in-sample and out-of-sample datasets are discussed in sections 3.4.3 and 3.4.4.

The time series for y , \hat{y} , e and the distribution of e over both the in-sample and out-of-sample dataset are evaluated. These time series results over the out-of-sample data, for all the different implemented models that are tested on the hollow shaft bearing 1 temperatures of turbine A and can be found in appendix I until N. The take away from all these models will be discussed below.

Simple regression models

It can be seen from figure J.33 and I.32 in appendix I that for both the in-sample dataset as well as for the test dataset, the linear regression is not able to describe lower temperature range, this is due to under fitting of the model. As the model is not able to describe the functional dependence in the data. This could already be seen in figure 4.5. Because of this, higher order polynomials were looked at.

For the results in appendix J, it can be seen that an improvement is made for the 6th order polynomial in relation to the linear regression. Still, the residual error is over a range of -10 to 10 °C for the in-sample data. This can not be considered as accurate enough for fault detection as a large error can not only be linked to a fault.

Supervised machine learning models

In section 4.3, classification and regression decision trees as well as SVM for classification or regression are covered. These supervised machine learning models will now be evaluated for a drivetrain in healthy condition. This is done by first analysing the performance on the first hollow shaft bearing temperature and later on the remaining bearing temperatures in the drivetrain.

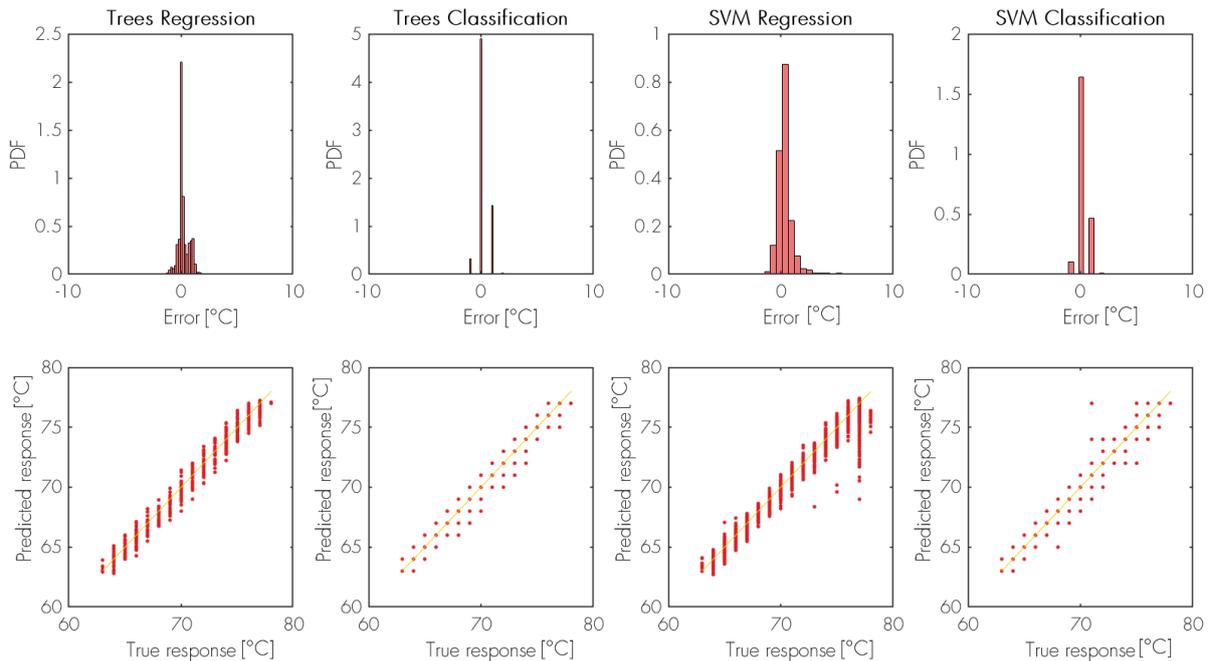


Figure 5.3: Test error distribution for Hollow Shaft Bearing 1 temperatures, turbine A

Figure 5.3 provides the error distribution of the different models applied on the out-of-sample data for the first hollow shaft bearing temperature. It can be seen that the SVM regression model result in a large residual between y and \hat{y} when generalizing over the higher temperature range. This will result in false alarms when using the framework

Bearing Temperature	Trees Regression		Trees Classification		SVM Regression		SVM Classification	
	Validation	Test	Validation	Test	Validation	Test	Validation	Test
Hollow Shaft 1	0.526	0.511	0.742	0.528	0.593	0.777	0.763	0.593
Hollow shaft 2	0.507	0.489	0.769	0.528	0.633	0.501	0.787	0.993
High Speed Shaft 1	0.554	0.589	0.687	0.630	0.640	0.854	0.746	2.43
High Speed Shaft 2	0.688	0.783	0.602	0.867	0.810	1.16	0.642	2.44
High Speed Shaft 3	0.552	0.634	0.645	0.692	0.637	0.911	0.705	2.65
Generator 1	3.758	7.15	0.246	7.45	3.687	7.45	0.529	7.19
Generator 2	7.44	10.0	0.246	11.6	7.60	10.3	0.379	14.3

Table 5.1: The validation error in °C over the in-sample data and test error in °C over the out-of sample data using different machine learning models

for anomaly detection. The validation and test error over the in-sample and out-of-sample datasets can be found in table 5.1 for all the applied models. From this table, it can be seen that the decision tree regression model results in the lowest error but decision tree classification and SVM classification models also yield in accurate models for the first hollow shaft bearing temperature. The resulting time series of y , \hat{y} and e for the decision tree regression model over the test dataset is presented in figure 5.4. The time series results for the remaining supervised machine learning models over the test dataset are presented in appendix K until N.

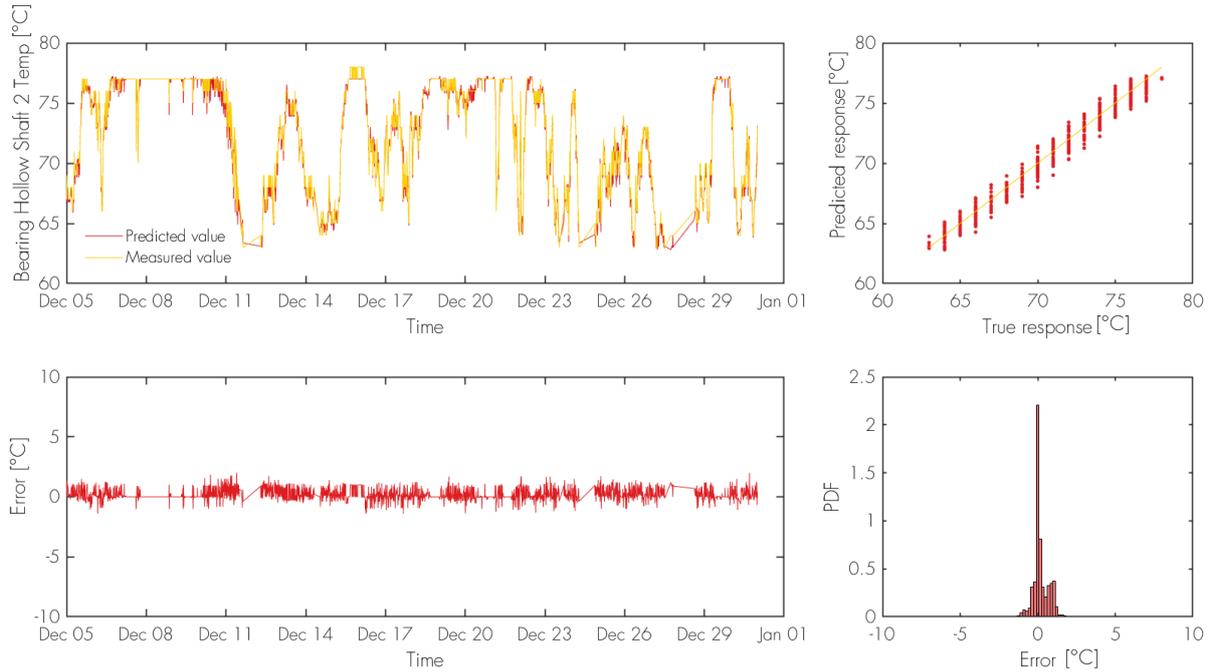


Figure 5.4: Out-of-sample time series for hollow shaft bearing 2, turbine A using a decision tree regression model

When analysing the supervised machine learning models on the remaining responses, it can be noted from table 5.1 that the decision tree model always result in the lowest test error for generalizing the gearbox bearing temperatures. It was found that for the first hollow shaft bearing, the SVM regression model resulted in large test errors in the

higher temperature range, this is also evident for both the regression as well as the classification SVM model when analysing the remaining responses. An example of this is given in figure 5.5, where the error distribution over the test dataset is given for to the middle high speed shaft bearing by using different models.

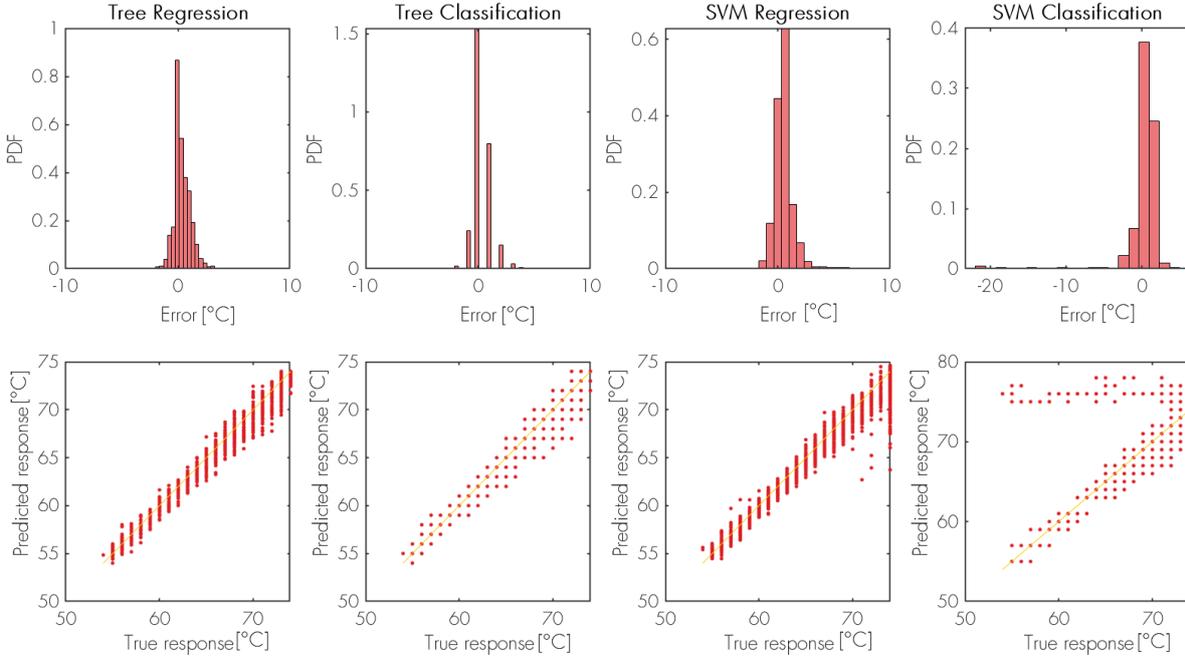


Figure 5.5: Test error distribution for the high speed shaft bearing 2 temperatures, turbine A

In addition to the decision tree models performing the best at describing the functional behaviour of the bearing temperatures and generalizing new results, it is also the computational most efficient model. As was already stated in section 4.5.2, the required runtime to train the model is the lowest for this type of supervised machine learning.

For the generalization of the generator bearing temperature, it can be seen in table 5.1 that none of the implemented models result in a low testing error, indicating that none of the methods result in an accurate model. This can also be seen from the time series for y , \hat{y} and e for the generator bearing 1 temperature using a regression decision tree model, displayed in figure 5.6. It was already concluded in section 4.5.1, that classification task models resulted in over fitting for the generator bearing temperatures.

The in-sample data that is used to train the model for turbine A is discussed in section 3.4.3. Figure 5.7 is presented to highlight the variability in this dataset. The relation of the second generator bearing temperature to the power curve is visible in this figure. The variability in this dataset is more clearly visible when analysing a smaller part of the operational window. The bearing temperatures under a 10 to 15 m/s wind velocity are displayed in figure 5.8. It can be seen from this figure that under similar

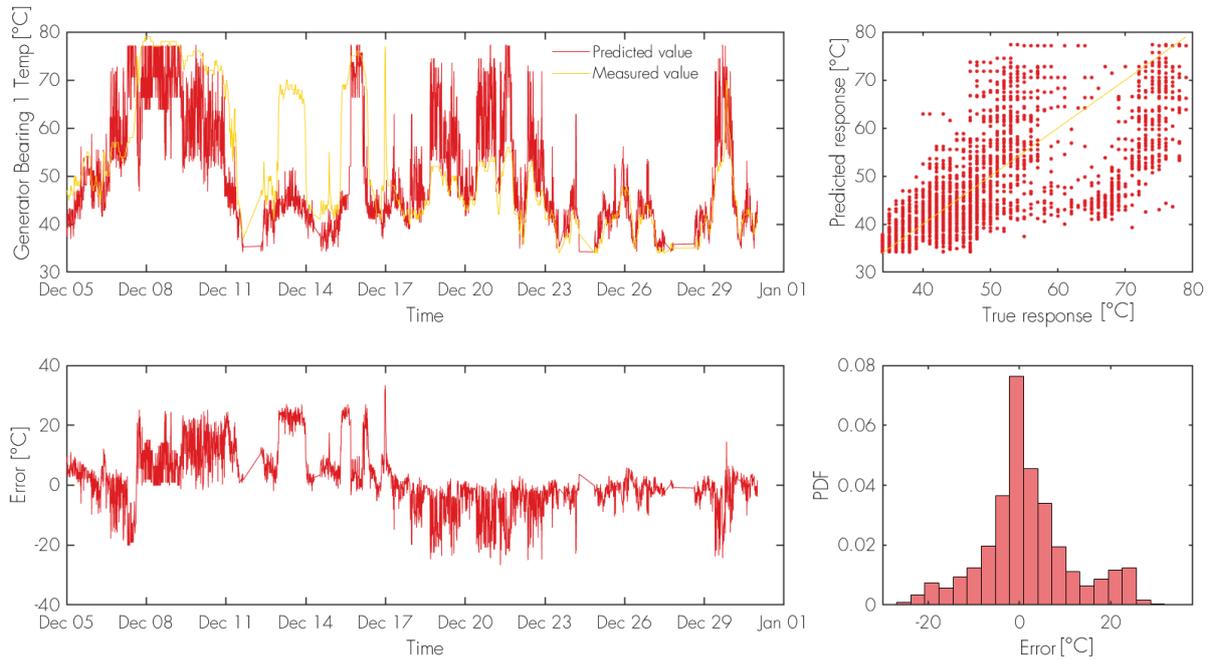


Figure 5.6: Out-of-sample time series for generator bearing 1,turbine A using a decision tree regression model

operational conditions a large range of bearing temperatures can be observed. Making it difficult to fit an accurate model without resulting in an over fitted model.

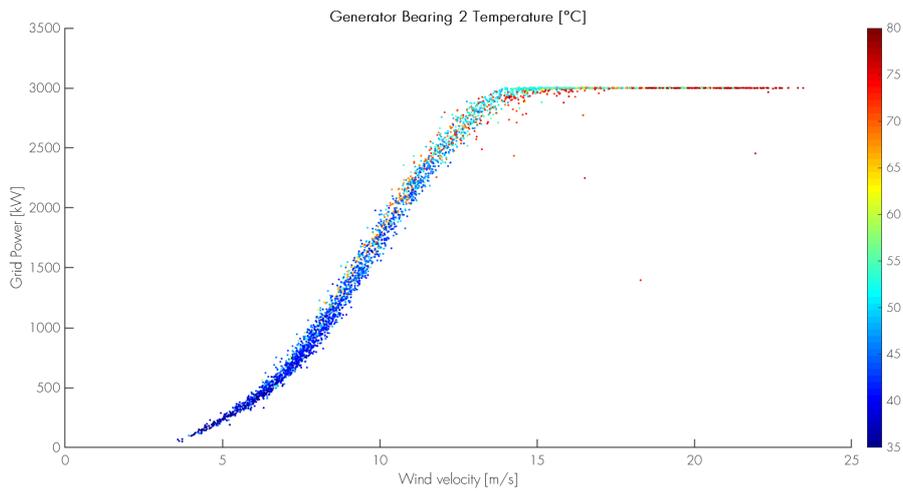


Figure 5.7: Turbine A training dataset, Generator bearing 2 in relation to the power curve

With the risk for a biased model, parameters from inside the generator are considered as input parameters in an attempt to obtain a model that could be used to describe the generator bearing temperatures. Several parameters are tested as additional input

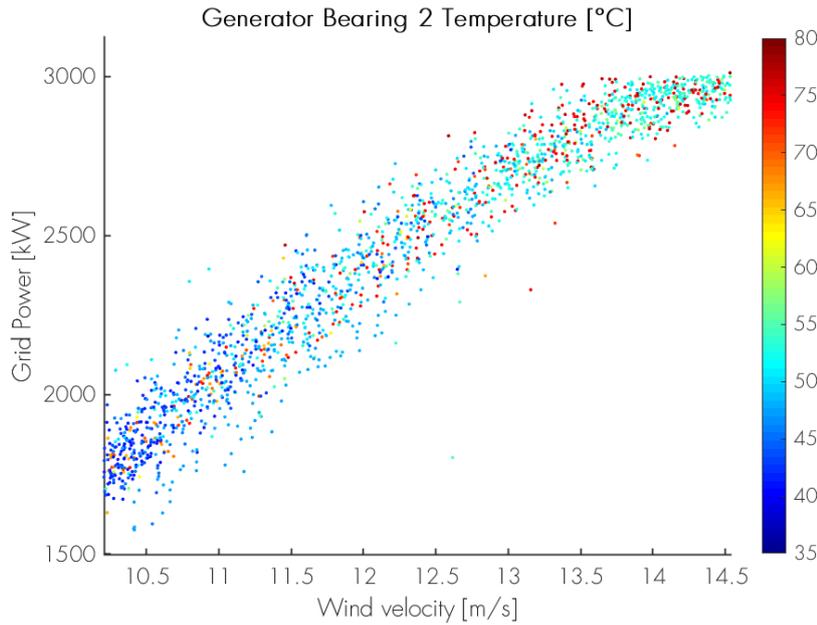


Figure 5.8: Turbine A training dataset, Generator bearing 2 in relation to the power curve for a small operational window

signals, taking the correlation coefficient, presented in appendix C into account. The lowest test error was achieved by introducing generator bearing 1 as an additional input to predict generator bearing 2, a validation error of 6.95 °C is found over the in-sample data and a testing error of 9.80 °C is found for the test dataset opposed to a validation and test error of 7.44 and 10.0 °C respectively. Thus, this does not result in an improved model.

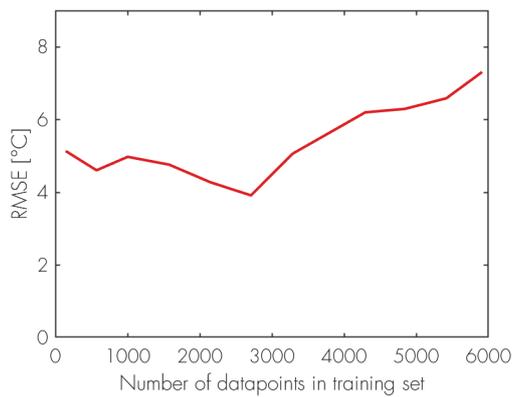


Figure 5.9: Test dataset turbine A, generator bearing 2 validation error for different size training sets

To be able to describe the functional dependence of the generator bearing, the influence of the size of the training dataset is investigated to see if additional training data could resolve the problem. The original training dataset contained 5680 number of data points. It can be seen from figure 5.9 that a model trained to a smaller sample of data points results in a lower validation error, this example indicated that the model shows signs of over fitting after a training dataset larger than 2750 data points. Increasing the training dataset will thus not result in a well fitted model.

Thus, for the generator in this study, only the winding temperatures signal can be used for condition monitoring. For this, the difference between the temperatures in the phases is used in stead of a normal behaviour model, as explained in section 4.2.

5.3 Anomaly detection

Based on the results from section 5.2, it can be concluded that the decision tree regression models perform best and will be used for anomaly detection. In this section, the detected anomalies over a period of one year will be discussed. Also the model limitations and abilities will be highlighted by showcasing the detected anomalies and false alarms.

5.3.1 Case turbine G

When analysing the middle bearing on the high speed shaft of turbine G, a couple of interesting cases could be identified. The first is illustrated in figure 5.10. From this figure, it can be seen that the turbine is stopped six times. These stops could not be coupled to anomalies detected previous to these stops, but result in a large error between \hat{y} and y shortly prior to a stop. This surpasses the threshold value for the control limits, but does not indicate a fault in the drivetrain, as the nature of the stop is unfamiliar.

These anomalies have in common that the model generalizes a value of $\hat{y} = 62$ °C. From the top right plot in figure 5.10, it can be seen that the model can't generalize values below 62 °C. This indicated that temperatures below this value were not present in the training data, this could be the result of filtering out all data when the turbine was not in operating conditions. These false alarms can be identified for a great number of turbines, and can easily be recognized as this sort of false alarm by their characteristics. The error coincide with a stop of the turbine directly after. Also, a bottom limit for the values of \hat{y} is visible in the true versus predicted value plot.

The second case that is identified for turbine G, is related to the stop at December 21st. Analysing the results presented in figure 5.11, the control charts result in the following findings: The Sherhard control chart identify 3 anomalies, all of these anomalies are combined with a generalized response of 62 °C. This value corresponds with the limit of the generalization range of the model, as discussed above. An other motivation to discard these errors as false alarms is that the error is the result of a large shift in the predicted value and not in the measured signal.

More interesting is that the CUSUM control chart indicated a positive shift of the

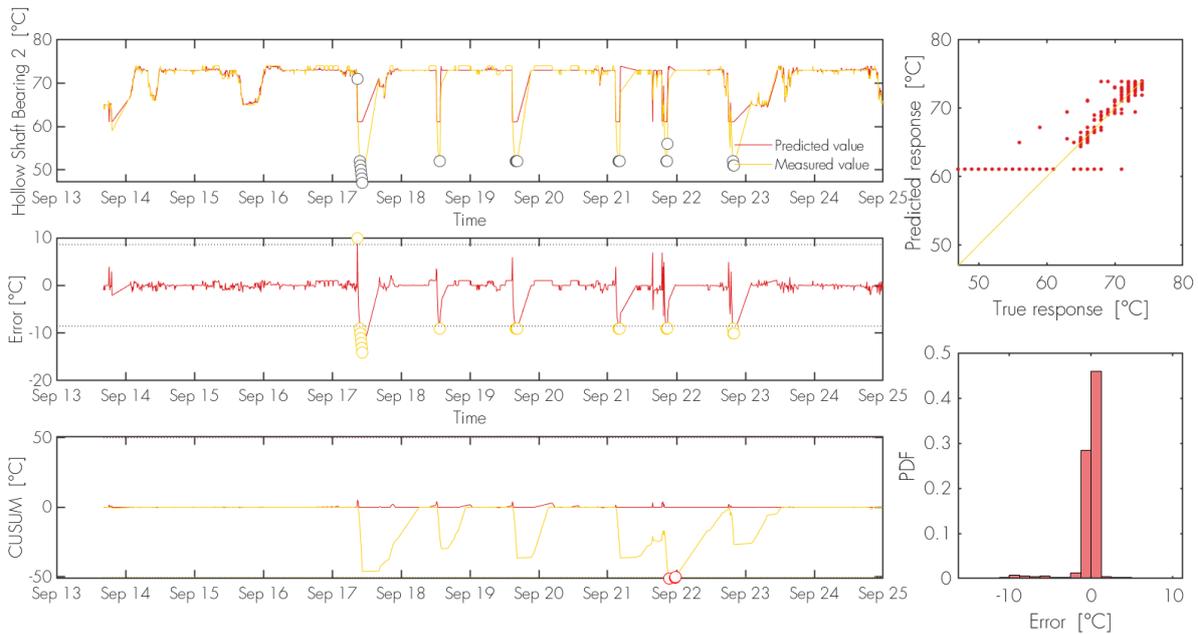


Figure 5.10: Time series result Turbine G: hollow shaft bearing 2 temperature

mean value of the middle hollow shaft bearing temperature, starting November 21. This is one month before the turbine is stopped for the remainder of the year. From figure 5.11, it could also be seen that between 11 and 13 December the turbine was stopped and the behaviour was shortly back to normal after this period before it started rising again. This indicated that the fault present prior to the 11th was not mended properly and ultimately led to the stop at December 21.

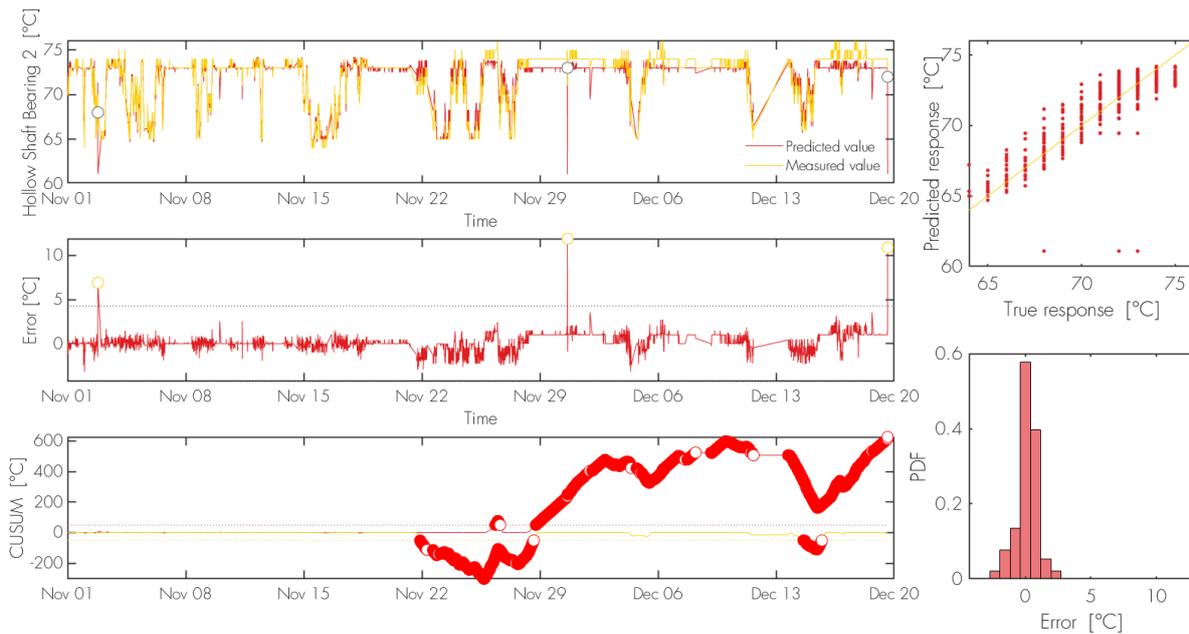


Figure 5.11: Time series result Turbine G: hollow shaft bearing 4 temperature

5.3.2 Case turbine H

In figure 5.12, it can be seen that the hollow shaft bearing results in a large error on December 4th 10:00, this is directly after a period where the bearing was experiencing high temperatures. This coincided with a alarm that was communicated to the SCADA system and stopped operation of the turbine for 9 hours. From the CUSUM control chart, a shift in the mean of the residual error was already picked up November 29. The model is able to describe the behaviour of the turbine with a small residual error after this stop. Thus, indicating the problem in the drivetrain is resolved and that model does not have to be retrained due to large changes to the drivetrain.

It is also interesting to note that the model for turbine H, that is highlighted in figure 5.12, is trained over a period from January 18 until January 31. This shows that, if no changes to the drivetrain are made, the training and prediction period can be far apart from each other. It can also be noted that for this example a training period of approximately 2 weeks provides enough data to train an accurate model on.

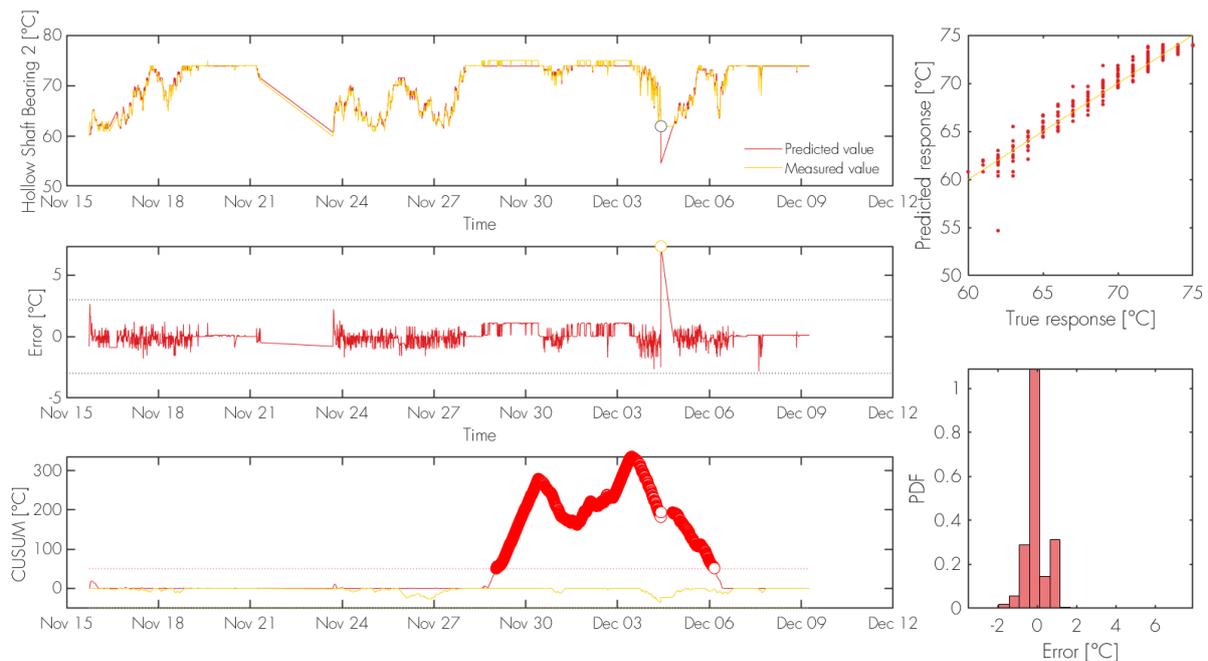


Figure 5.12: Time series result Turbine H: hollow shaft bearing 1 temperature

5.3.3 Case turbine I

The case presented for turbine H, highlighted that a well trained model can be used over a long period after the training period, if it is trained correctly and no changes to the drivetrain are made. To show the need of retraining the model, figure 5.13 is provided.

The model for this case is trained from November 6 until November 28 and is able to describe the temperature response of the gearbox bearing well for the period after the standstill from October 12 until November 5. The model is however not able to generalize the behaviour before this standstill with a small residual error, this indicates that it is likely that maintenance has been done to the gearbox of the turbine and now has a different functional behaviour.

To describe the behaviour of the drivetrain before the 12th of October, the model should be trained within this period.

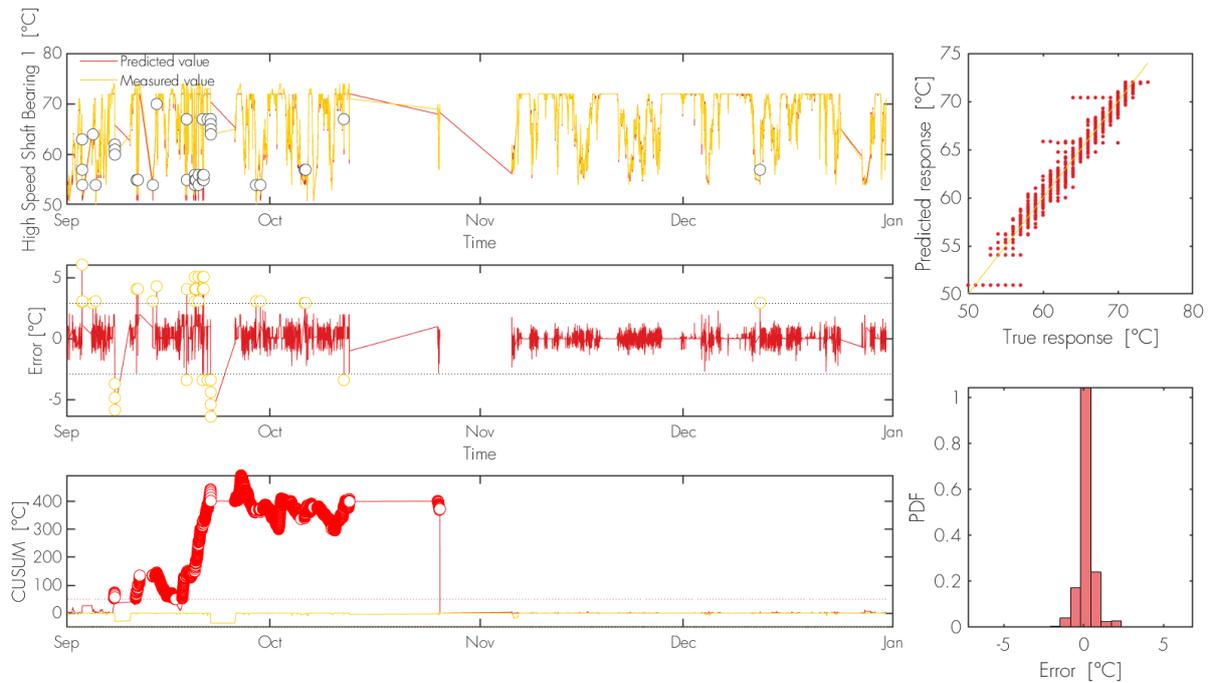


Figure 5.13: Time series result Turbine I: hollow shaft bearing 1 temperature

5.3.4 Case Turbine J

Some anomalies in the data show that the turbine is not operating as normal, but do not have to be linked to a fault present in the drivetrain. An example of that is given in figure 5.14.

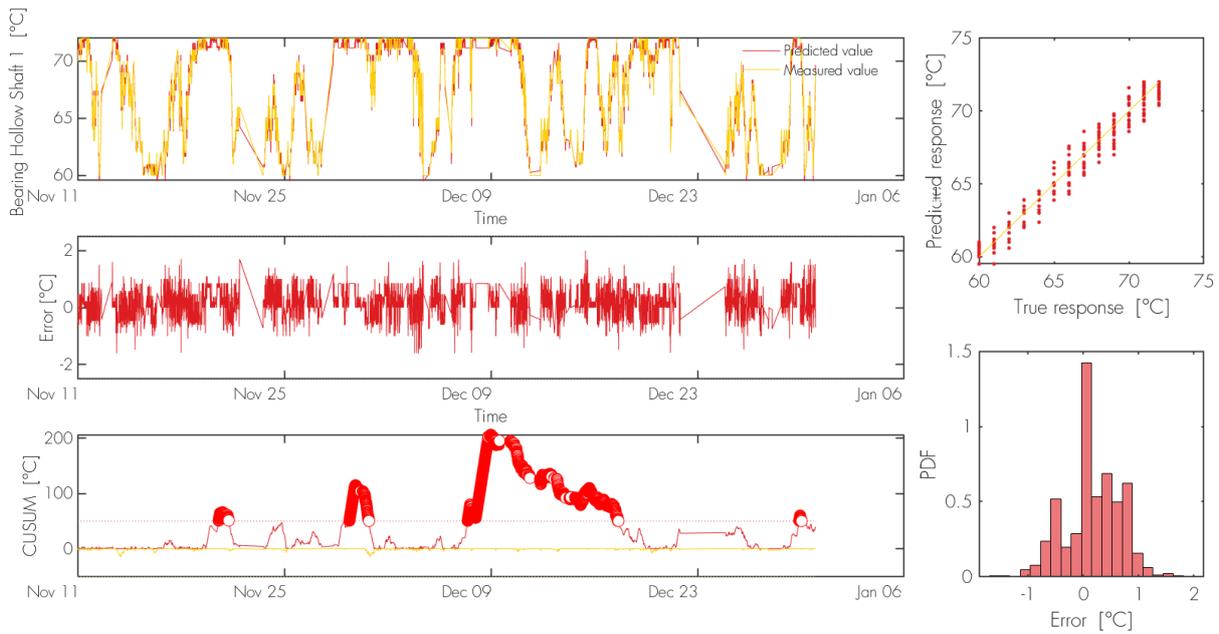


Figure 5.14: Time series result Turbine J: hollow shaft bearing 1 temperature

It can be seen from figure 5.14 that the average temperature of the first hollow shaft bearing was on average hotter than normal after December 8th, but returned to normal conditions after 5 days. This anomaly was only picked up by the CUSUM control chart and would not have been detectable otherwise. It can not be connected to any failure or unexpected stopping of the turbine. Also, no large deviations were registered with the Shewhart control chart and the turbine was not taken out of production during this 5 day period.

5.3.5 Case turbine K

Turbine K was taken out of production from a period 4 days after the stop at 1 December. It can be seen from figure 5.15 that one day before the stop, the high speed shaft bearings show large deviations from the expected behaviour. During this period, all bearings on the high speed shaft show abnormal high temperatures over the same 15 hours before the stop.

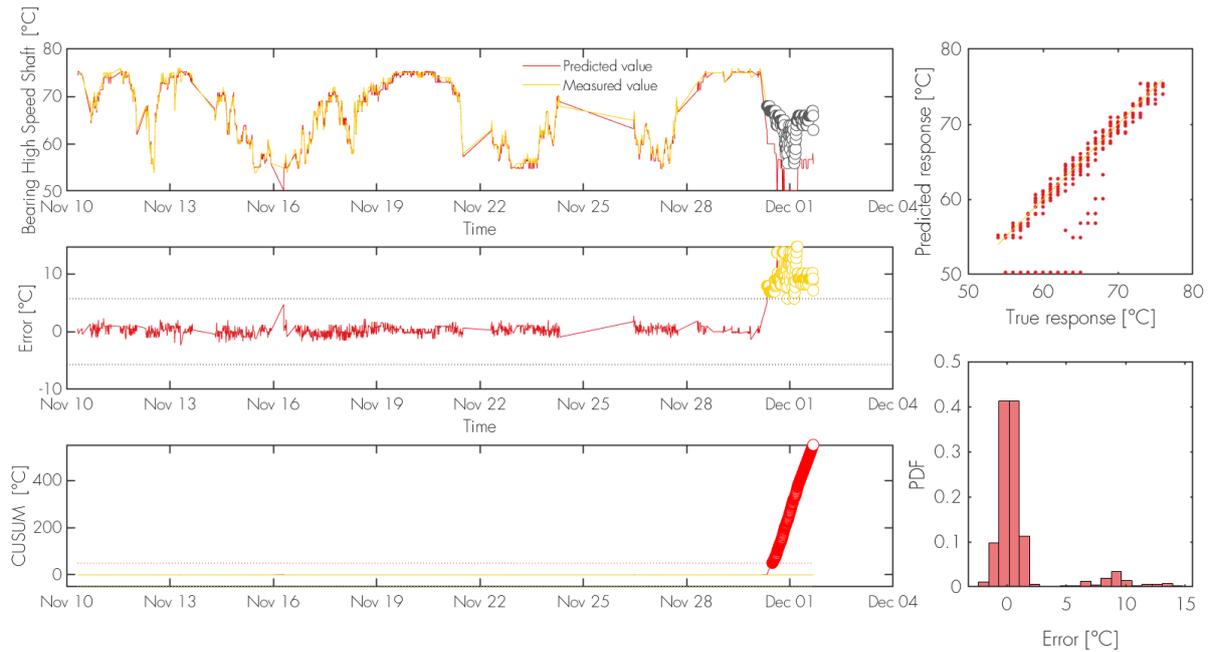


Figure 5.15: Time series result Turbine K: High Speed shaft bearing 1 temperature

Prior to this stop, no deviating behaviour could be detected by the CUSUM control chart. However, a smaller anomaly was detected on the hollow shaft bearings at the same time. This can be seen in figure 5.16. The bearings seem to be colder than expected. This deviation could be tracked from 20 November until the turbine was stopped at 1 December. This could indicate that the control system of the turbine was registering an increasing in the temperature on the high speed shaft bearing and increased the active cooling. This cooling system also effect the hollow shaft and ultimately creating the anomaly in these bearings that could be identified and tracked with the developed framework.

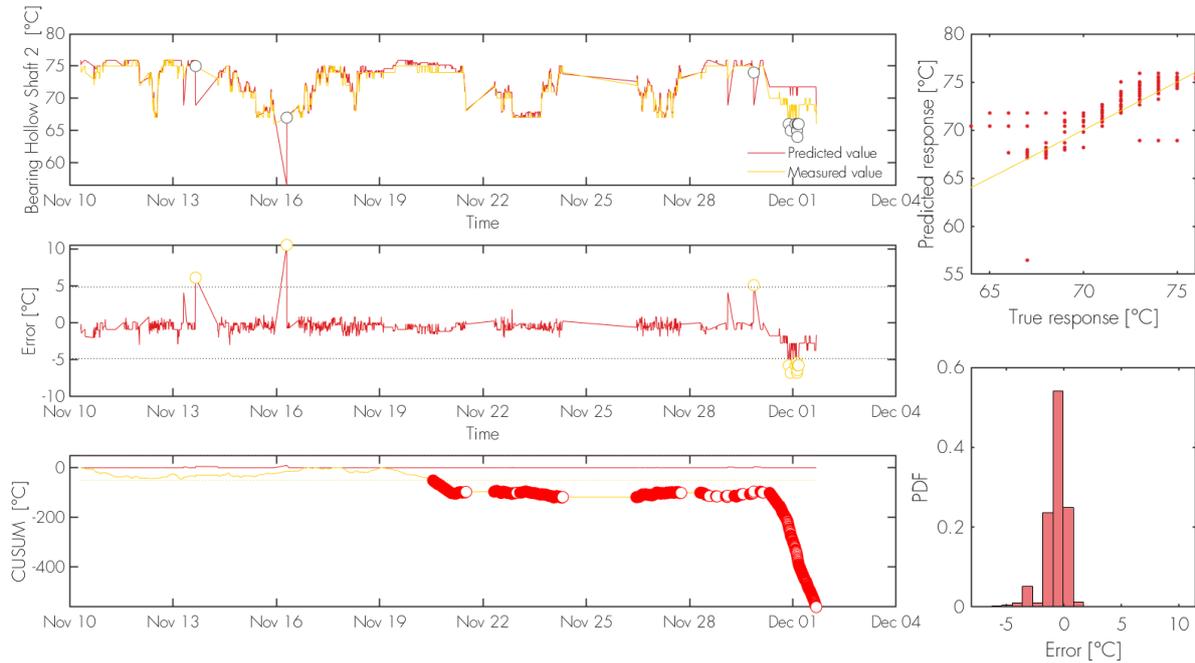


Figure 5.16: Time series result Turbine K: hollow shaft bearing 2 temperature

5.4 Discussion on the condition monitoring framework

Different anomalies were detected with the temperature based condition monitoring framework that is developed to answer the first research question. Different models are evaluated in section 5.2 to answer the third research question. The findings and limitations related to the results presented in chapter 5 will be discussed in this section. The overall discussion in chapter 7 will focus on the relation of the findings in respect to previous research and discuss the implications of the results.

For the results presented in this chapter, turbine A has been used primarily been used to highlight the findings and draw conclusions. Other turbines have been used to investigate if this changed the observations or findings. It was found that this was not the case and the decision trees regression model could be applied of all turbines and gearbox signals.

In the machine learning domain, only 4 supervised machine learning methods have been thoroughly evaluated: A decision trees model using regression, A decision trees model using classification, A SVM using regression and a SVM using classification. The regression learner and classification learner toolbox also contained the following methods: higher dimension linear regression, Gaussian process regression, random forest or K-nearest neighbour classifiers. These methods were also looked at but it could be seen in the initial phases that these models resulted in a higher validation error o the training dataset compared to the decision tree or support vector machine models.

Several false alarms can be observed when the decision tree model is applied on a new dataset to detect anomalies that are related to a fault in the components. These false alarms coincided with the turbine been taken out of operation and the model not being able to generalize over these conditions as they were not present in the training dataset. These false alarms can be filtered out by developing an additional filter in further developments.

No access to maintenance logs or other ways to validate the found anomalies was available in this research. Also, no information on known failure cases could be used to benchmark the framework. Further validation is needed before this tool could be implemented at OWEZ and used for condition monitoring.

With the models that were covered in this research, it was not able to generalize the generator bearings with a normal behaviour model. The variability in the training data shows that under similar operational conditions, a range of temperature responses could be observed. Thus, making it hard to fit a model without resulting in over fitting. Additional input signals did not resolve this problem, neither did enlarging the training dataset. It was even found that a smaller dataset resulted in a lower error.

The large variability in the data for the generator could be a result of other drivetrain components influencing the bearing temperatures of the generator. The cooling water of the generator is also used to cool the hydraulic system and the convertor, it could be that these component temperatures also influence the generator bearing temperatures through the cooling water.

6 Wake effects on component temperatures

Wake effect can result in a loss of power production as well as additional fatigue loading on the turbine [47]. Churchfield has shown with large eddy simulations that downwind turbines experience higher fluctuations in several drivetrain component moments and that result in increased damage equivalent loads [48]. The influence of this additional loading on component temperatures is investigated to answer the second thesis objective.

6.1 Approach

To identify the relation between wake and component temperatures, the following work flow is applied in this research.

First, the wake conditions on turbine level are modelled. An in-house developed wake tool of Shell is used for this. This tool calculates the velocity deficit and turbulence intensity on turbine level based on the Ishihara wake model. The fundamental theory of this model will be covered in section 6.2. This model utilizes wind velocity, wind direction and ambient turbulence intensity as input. For this, the DOWA dataset is utilized, as discussed in section 3.1.2.

Next, the timestamps of the waked conditions time series and the data in the SCADA dataset need to be matched. Because the DOWA data contains hourly data, the component temperature data needs to be converted from ten min average values to hourly mean values.

Data can now be clustered over both the wind speeds as well as over the waked conditions. To decide on the cluster criteria, the operational behaviour of the turbine are taken into account. Because of this the wind speed cluster criteria are given as by table 6.1. First, the dataset, D , is clustered over the wind speeds given by table 6.1 so that $d_i \subset D$ with $i = \{1, \dots, n\}$. Here, d_i is a clustered subset of D and n is the number of wind speed cluster.

Cluster	n=1	n=2	n=3	n=4	n=5	n=6	n=7	n=8
U [m/s]	0 → 3.5	3.5→5.5	5.5→7.5	7.5→9.5	9.5→11.5	11.5→13.5	13.5→15	15→25

Table 6.1: Wind speed clusters

Next, each dataset d_i is clustered over the different wake conditions to create subsets $w_{i,j}$, with $w_{i,j} \subset d_i$ and j is the number of wake clusters. When clustering over the waked conditions, both the normalized turbulence intensity or normalized velocity deficit can be used as indication of experience wake by the turbine. These parameter are normalized with the ambient turbulence intensity and free stream wind velocity respectively. The applied clustering criteria are presented in table 6.2.

Cluster	n=1	n=2	n=3	n=4	n=5	n=6	n=7
$\frac{I_w}{I_g}$	1 → 1.1	1.1 → 2	2 → 3	3 → 4	4 → 6	6 → 8	8 → 12
$\frac{U_w}{U_\infty}$	0 → 0.002	0.002 → 0.03	0.03 → 0.07	0.07 → 0.15	0.15 → 0.30		

Table 6.2: Wind speed clusters

For each dataset of $w_{i,j}$, the distribution of the component temperature is constructed. These distributions can be compared against every subset of d_i . Thus here temperature distributions of different wake conditions are compared in the same wind speed cluster.

It must be noted, that for a good distribution, each dataset of $w_{i,j}$ should have sufficient data. If there are less than 50 data points, then the distribution is discharged due to too little information to make a reliable distribution.

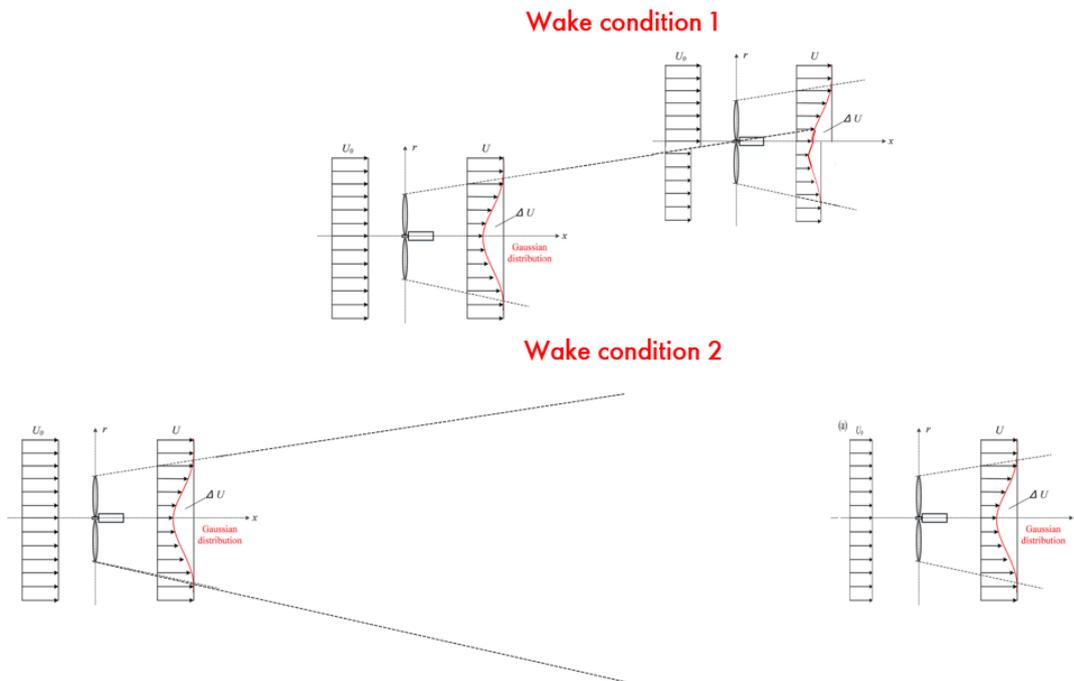


Figure 6.1: 2 wake conditions for partially waked conditions

Turbines that are partially waked can be the result of two wake conditions. The first is when the rotor is experiencing wake over the entire rotor area but the turbine is located relatively far upstream. The second wake condition is when a turbine is experiencing asymmetric loading due to the wake area of the upstream turbine not completely covering the full swept area of the rotor. These different conditions for a partially wake turbine are visualized in figure 6.1.

The effects of asymmetric wake loading on component temperatures are investigated by dividing the cluster $w_{i,j}$ in the two wake conditions that are discussed above. The cluster containing the effects of wake condition 1 is selected on the following criteria:

- $WA_a > 0.25$ & $WA_a < 0.75$,
- $\Delta u_{b,a} > 0.1$

The physical interpretation of these criteria is that the cluster for wake condition 1, the wake covered area over the rotor area (WA) of turbine a is 25% to 75% as a result of turbine b and this results in a velocity deficit larger than 10%, only as a result of turbine b ($\Delta u_{b,a}$).

6.2 Ishihara wake model

As mentioned before, the Ishihara wake model [49] is used to model the wake condition at OWEZ. The fundamentals of this model and input as well as responses are discussed in this section.

6.2.1 Model input parameters

Ambient Turbulence intensity

Since the Ishihara wake model takes the ambient turbulence intensity into account, this should also be an input for the tool. This signal is missing in the dataset obtained from the DOWA as discussed in section 3.1.2, thus this needs to be constructed.

To do so, the relation between wind speed and turbulence intensity at OWEZ is determined using the mean and standard deviation of the ten-minute average wind speed in the SCADA system. Using equation 6.1, the turbulence intensity at instance i is calculated. From the data in figure 6.2, a power law function is derived for the relation between turbulence intensity and wind speed for the OWEZ location. This relation is given by equation 6.2:

$$I_i = \frac{\sigma_i}{\mu_i}, \quad (6.1)$$

$$I_a(u) = a \cdot u^b + c = 0.143 \cdot u^{-1.463} + 0.1. \quad (6.2)$$

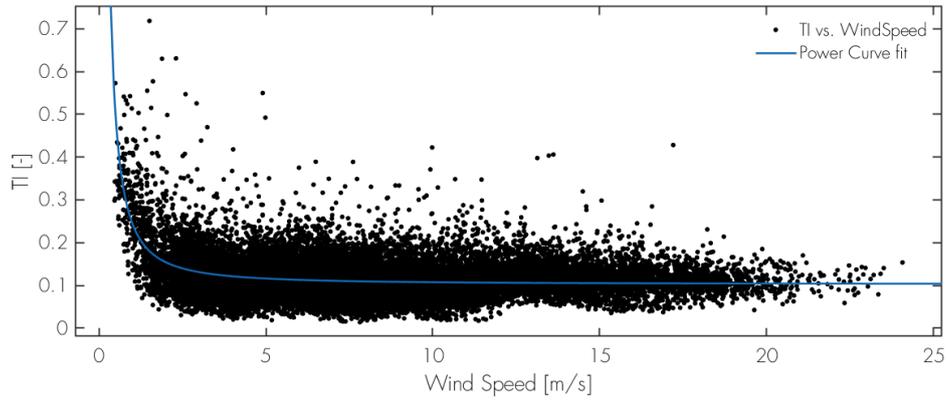


Figure 6.2: Turbulence intensity vs. wind speed and fitter power law

Thrust coefficient

The turbine characteristics are taken into account in the wake model. This is done by scaling the parameters for wake development with the thrust coefficient at given incoming wind speed.

The thrust coefficient, C_T , can be calculated with equation 6.3. Here, the delivered thrust and incoming wind velocity are given by T and U , the swept area of the rotor is denoted as A_D and ρ is the parameter for the air density.

$$C_T = \frac{T}{0.5\rho U^2 A_D}. \tag{6.3}$$

The thrust coefficient for the Vestas V90 3.0 is provided by the manufacturer and can be seen in figure 6.3. It is assumed that this thrust curve can be used for any air density ρ . Since no information is available on the change of ρ over the dataset, and no other thrust coefficient curves are available.

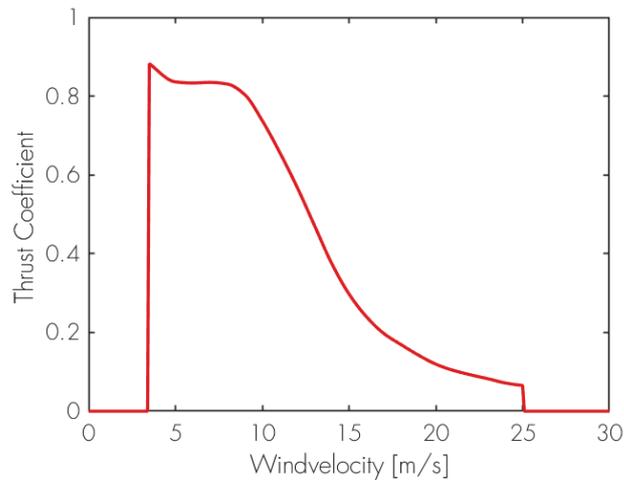


Figure 6.3: Vestas V90 Thrust coefficient curve C_T

6.2.2 Modelled wake conditions

As discussed before, the output of the in-house developed wake tool are the wind velocity deficit and the added turbulence intensity. How the Ishihara wake model calculates these parameters is discussed in this section.

Velocity deficit

The Ishihara model can calculate the velocity deficit by means of the product of the streamwise function and a spanwise function. As shown in equation 6.4 [49]:

$$\frac{\Delta U}{U_h} = F\left(C_T, I_a, \frac{x}{D}\right) \phi\left(\frac{r}{\sigma}\right). \quad (6.4)$$

The streamwise function $F\left(C_T, I_a, \frac{x}{D}\right)$ is given as the maximum velocity deficit, normalized with the mean wind speed at the hub, U_h . This function is given by equation 6.5:

$$F\left(C_T, I_a, \frac{x}{D}\right) = \frac{1}{\left(a + b\frac{x}{D} + c\left(1 + \frac{x}{D}\right)^{-2}\right)^2}. \quad (6.5)$$

Next, Ishihara used large eddy simulations to fit data to equation 6.5 and find empirical expressions for a , b and c . These expressions are given in relation to the thrust coefficient and ambient turbulence intensity, as can be seen in equation 6.6 [49]:

$$a = 0.93C_T^{-0.75}I_a^{0.17}, \quad b = 0.42C_T^{0.6}I_a^{0.5}, \quad c = 0.15C_T^{-0.25}I_a^{-0.7}. \quad (6.6)$$

The Ishihara model assumes a Gaussian distribution of the velocity deficit behind the rotor [49], as shown in figure 6.4. This distribution can be found in the spanwise function given in equation 6.7:

$$\phi\left(\frac{r}{\sigma}\right) = \exp\left(-\frac{r^2}{2\sigma^2}\right). \quad (6.7)$$

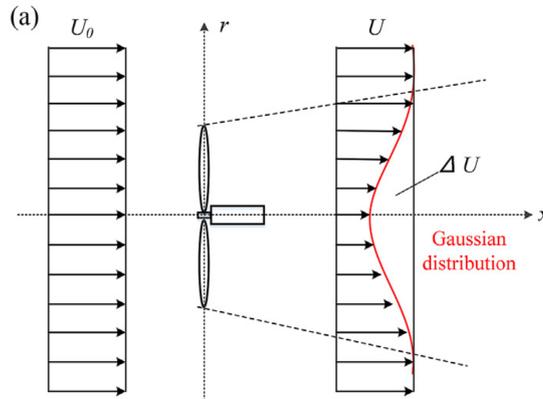


Figure 6.4: Gaussian velocity deficit distribution [49]

Ishihara also assumes that the wake behind the turbine expands linearly, as given by equation 6.8, For this expansion, the large eddy simulation are used for a empirical expression [49]:

$$\frac{\sigma}{D} = 0.11C_T^{1.07}I_a^{0.20}\frac{x}{D} + 0.23C_T^{-0.25}I_a^{0.17}. \quad (6.8)$$

Added turbulence intensity

To model the added turbulence intensity, Ishihara takes a similar approach as for the velocity deficit. The added turbulence intensity is spoken of, as the experienced turbulence intensity is a sum of the ambient and the added turbulence intensity due to wake. The added turbulence intensity, I_+ , can be calculated by equation 6.9. This is again the product of a streamwise and a spanwise function and the additional term for $\delta(z)$ is to account for lower turbulence intensity in the lower areas of the wake stream. Similar as to the velocity deficit, the added turbulence intensity is assumed to be axial symmetrical [49].

$$I_+ = G\left(C_t, I_a, \frac{x}{D}\right)\phi\left(\frac{r}{\sigma}\right) - \delta(z). \quad (6.9)$$

The streamwise function is given by equation 6.10. Similar as to the streamwise function of the velocity deficit, an empirical expression is found for the function parameters by large eddy simulation. These expressions can be found in equation 6.11 [49]:

$$G\left(C_t, I_a, \frac{x}{D}\right) = \frac{1}{d + e\frac{x}{D} + f\left(1 + \frac{x}{D}\right)^{-2}}, \quad (6.10)$$

$$d = 2.3C_T^{-1.2}, \quad e = I_a^{0.1}, \quad f = 0.7C_T^{-3.2}I_a^{-0.45}. \quad (6.11)$$

For the added turbulence intensity, Ishihara assumes that the maximum added turbulence intensity will occurs around the tip of the blades. This results in a Gaussian shape distribution around the tip of the blades, given by equation 6.12 and a double Gaussian shape for the added turbulence intensity over the rotor, as can be seen in figure 6.5,

$$\phi\left(\frac{r}{\sigma}\right) = \exp\left(-\frac{r'^2}{2\sigma^2}\right). \quad (6.12)$$

Here, r' is given as $r' = r - D/2$. To describe the area behind the rotor, a superposition of the two Gaussian distributions needs to be taken. Because of this, equation 6.12 can extended to equation 6.13 to describe the spanwise function [49]:

$$\phi\left(\frac{r}{\sigma}\right) = k_1\exp\left(-\frac{(r - D/2)^2}{2\sigma^2}\right) + k_2\exp\left(-\frac{(r + D/2)^2}{2\sigma^2}\right). \quad (6.13)$$

The parameters k_1 and k_2 can be described by equation 6.14 [49]:

$$k_1 = \begin{cases} \cos^2\left(\frac{\pi}{2}\frac{r}{D-0.5}\right) & r/D \leq 0.5 \\ 1 & r/D > 0.5 \end{cases}, \quad k_2 = \begin{cases} \cos^2\left(\frac{\pi}{2}\frac{r}{D+0.5}\right) & r/D \leq 0.5 \\ 1 & r/D > 0.5 \end{cases}. \quad (6.14)$$

The asymmetry of the added turbulence intensity over the vertical direction is taken into account by the correction term $\delta(z)$. This accounts for the lower added turbulence intensity in the lower region of the wake flow. This is given by equation 6.15 [49]:

$$\delta(z) = \begin{cases} 0 & (z \geq H) \\ I_a \sin^2\left(\pi\frac{H-z}{H}\right) & (z < H) \end{cases}. \quad (6.15)$$

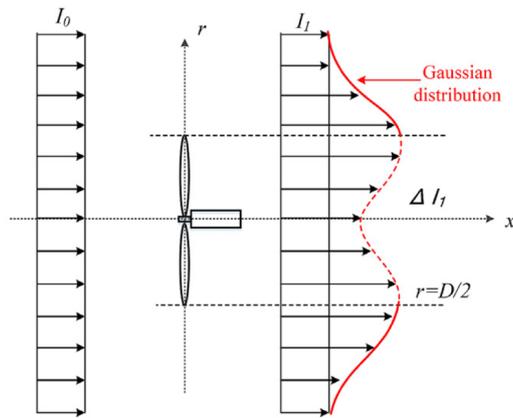


Figure 6.5: Gaussian added turbulence intensity distribution [49]

6.2.3 OWEZ wake simulations

Now that the fundamentals and assumptions of the Ishihara wake model are covered, the in-house developed wake tool can be used to simulate the wake conditions on turbine level at OWEZ.

For this the time series for the wind velocity, wind direction and ambient turbulence intensity as well as the turbine characteristics and the placement in the wind farm are taken as input for the tool. The output of the tool are the velocity deficit, ΔU , and turbulence intensity, I , as $I = I_a + I_+$. To investigate the effects of experienced wake, a ratio can be made by normalizing these outputs. Thus, clustering the temperature distributions over $\Delta U/U_h$ or I/I_a and the wind velocity at turbine level.

An example of the wake conditions at the OWEZ wind farm is given in figure 6.6. Here, the wind farm experience an incoming wind velocity of 8 m/s under a 330° angle with a turbulence intensity of 7.7%.

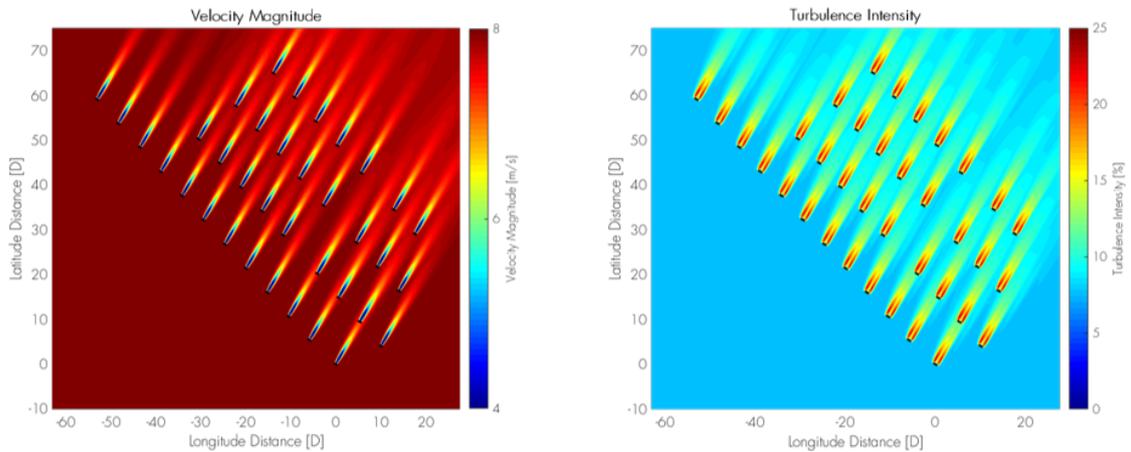


Figure 6.6: Simulated velocity deficit and turbulence intensity at OWEZ

6.3 Wake temperature distribution results

To investigate the wake effects on component temperatures, the temperature distribution for the gearbox bearings is visualized for the hourly mean data from the SCADA data and the modelled wake conditions at OWEZ.

If the hypothesis is true that, wake effects have influence on the drivetrain component temperatures, then a shift in the peak or width of the temperature distributions is expected under different wake conditions. This could also be interpreted by the statistical properties of these distribution as a shift in the mean or standard deviations.

6.3.1 Results turbulence intensity cluster

The bearing temperatures clustered over the incoming wind speed at turbine level and normalized turbulence intensity is first looked at. $\frac{I_w}{I_a}=1$ represent the turbine performing under free stream conditions.

In figure 6.7, the second high speeds shaft bearing temperature distribution of turbine L is presented as an example. 5 temperature distributions are presented under different wake conditions over the wind speed cluster, from 11.5 until 13.5 m/s. The mean value and standard deviations of these distributions are given in tables 6.3 and 6.4. The statistical properties as well as the number of elements per distributions for the first hollow shaft bearing and the first high speed shaft bearing of turbines A and G are highlighted in appendix P as additional results.

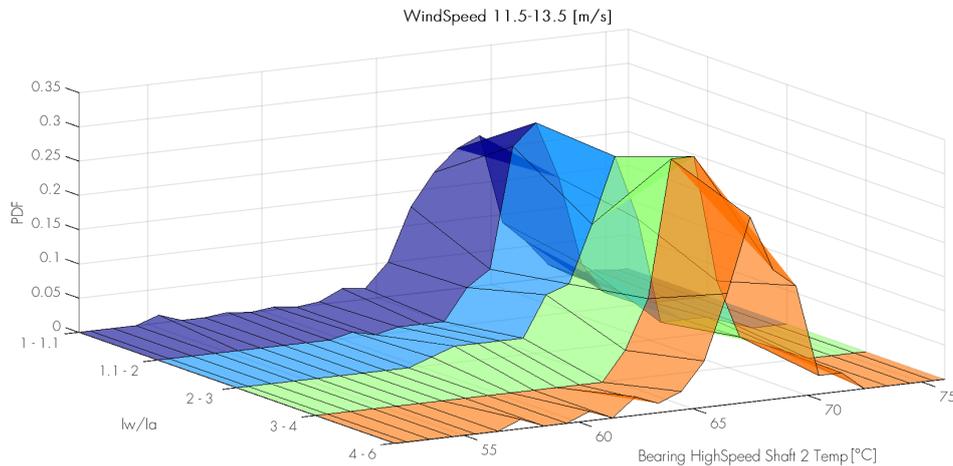


Figure 6.7: Temperature distribution over different wake conditions, by normalized turbulence intensity

It can be concluded from figure 6.7 and tables 6.3 and 6.4, that no noticeable difference between the temperature distribution over different wake conditions can be found. The statistical properties of the distributions, given in the tables 6.3 and 6.4, show negligible differences. It must be noted that a distribution is only included in the analysis

if the cluster $w_{i,j}$ has at least 50 data point. In appendix P, the data points per cluster can be found.

IwIa	Wind Speed						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	55,2	57,9	62,7	65,8	68,1	69,9	71,5
1.1 - 2	54,9	58,7	63,1	66,8	68,1	69,7	71,7
2 - 3	55,7	58,3	62,4	65,3	67,6	69,3	71,1
3 - 4	55,9	57,8	62,1	65,9	68,1	69,6	69,8
4 - 6	55,4	58,8	62,1	65,3	67,3	68,6	69,8
6 - 8	56,2	58,6	62,4	64,7	66,7	70,3	70,1
8 - 12	56,2	58,9	61,1	65,1	67,5	NaN	NaN

Table 6.3: Mean second high speed shaft bearing temperature of clusters from turbine A

IwIa	Wind Speed						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	2,3	2,9	2,8	2,5	2,4	1,9	1,9
1.1 - 2	3,4	4,7	2,9	2,0	1,7	2,4	1,5
2 - 3	2,3	3,4	2,9	2,4	2,4	3,3	3,0
3 - 4	2,9	3,5	3,0	2,0	1,4	1,2	4,0
4 - 6	3,0	3,5	3,0	2,2	2,1	1,5	2,5
6 - 8	2,5	2,9	2,8	2,7	4,1	0,9	0
8 - 12	2,7	3,3	3,4	3,2	0,3	NaN	NaN

Table 6.4: Standard deviations second high speed shaft bearing temperature of clusters from turbine A

6.3.2 Results velocity deficit cluster

The results over the clusters when using the velocity deficit at turbine level as indicator of the wake conditions are presented in this section. The velocity deficit is normalized with the free stream velocity. Thus, 0 indicates that the turbine is experiencing free stream conditions.

The different temperature distribution can be found in figure 6.8 and tables 6.5 and 6.6 for the highlighted example of the third high speeds shaft bearing temperature of turbine M. It can be seen from this that the difference over the temperature distributions with different experienced wake conditions are also relatively small.

The statistical properties of the different distributions using different turbines and signals can be found in appendix Q. It can be concluded from this that no significant change could be found between the temperature distributions over different wake conditions.

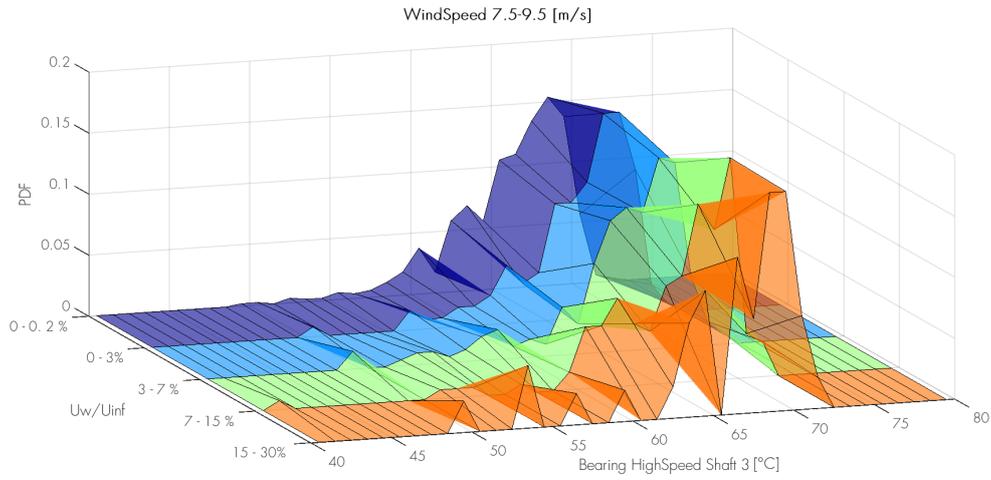


Figure 6.8: Temperature distribution over different wake conditions, by normalized velocity deficit

IwIa	Wind Speed						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 0.002	55,1	60,4	66,0	69,4	70,5	71,1	71,4
0.002 - 0.03	56,3	61,9	66,8	69,3	70,5	71,5	71,8
0.03 - 0.07	56,1	60,5	65,9	69,8	70,3	70,8	72,7
0.07 - 0.15	55,1	59,9	65,8	69,0	70,5	71,5	72,7
0.15 - 0.3	56,1	60,4	65,7	69,4	72,3	72,5	NaN

Table 6.5: Mean second high speed shaft bearing temperature of clusters from turbine A

IwIa	Wind Speed						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 0.002	3,56	4,41	3,48	2,37	1,65	2,23	1,36
0.002 - 0.03	4,85	5,10	3,59	2,95	2,55	1,63	1,44
0.03 - 0.07	3,63	4,43	4,38	1,89	2,32	3,47	1,03
0.07 - 0.15	3,32	4,38	4,31	2,45	1,29	2,49	0,800
0.15 - 0.3	4,40	4,22	4,95	3,91	1,25	0,289	NaN

Table 6.6: Standard deviations second high speed shaft bearing temperature of clusters from turbine A

6.3.3 Results asymmetric loading

To highlight the effects of asymmetric loading on component temperatures, figure 6.9 and 6.10 are presented. Clusters $w_{i,j}$ with the incoming wind speed at turbine level of $5.5 < U < 7.5$ m/s and $7.5 < U < 9.5$ m/s and a normalized velocity deficit of $7\% < \frac{\Delta U}{U_\infty} < 15\%$ are split in data for wake condition 1 and 2. It must be noted that few cases could be identified where wake effects could be identified as asymmetric loading by the divined constraints. This results in temperature distributions over a small data samples, as can be seen in the legends of figures 6.9 and 6.10. The cases where enough data could be found to make a temperature distribution will be presented below.

Three turbines are presented in figure 6.9, these cases are the result of an incoming wind speed of 5.5 - 7.5 m/s and a normalized velocity deficit of $0.07 < \frac{\Delta U}{U_{infty}} < 0.15$. It can be seen from figure 6.9 that for the three highlighted clusters, no large change in the temperature distribution can be observed for the different wake conditions. The orange bars in figures 6.9 and 6.10 represent an overlap of the distributions of the two different wake conditions.

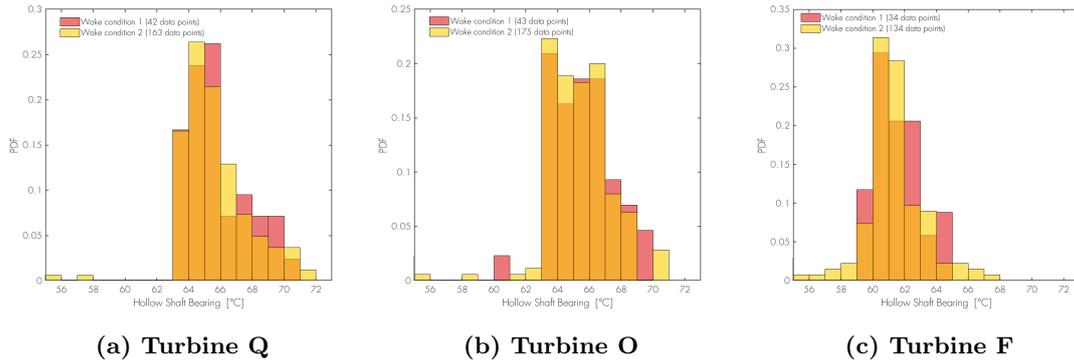


Figure 6.9: Temperature distributions for the hollow shaft bearing under wake condition 1 and 2 clustered over wind speed: 5.5-7.5 [m/s] & $\frac{\Delta U}{U_{infty}}$: 0.07-0.15 [-]

Five cases could be identified in the cluster with an incoming wind speed of 7.5 - 9.5 m/s and a normalized velocity deficit of $0.07 < \frac{\Delta U}{U_{infty}} < 0.15$. It can be seen from figure 6.10 that for turbine N and B, the temperature distributions is shifted to higher temperatures. This is less evident for turbines O and L but it can be observed that the majority of the data is located in the higher temperature range. Turbine G contradicts these findings by indicating a small shift of the peak of the distribution towards the lower temperatures.

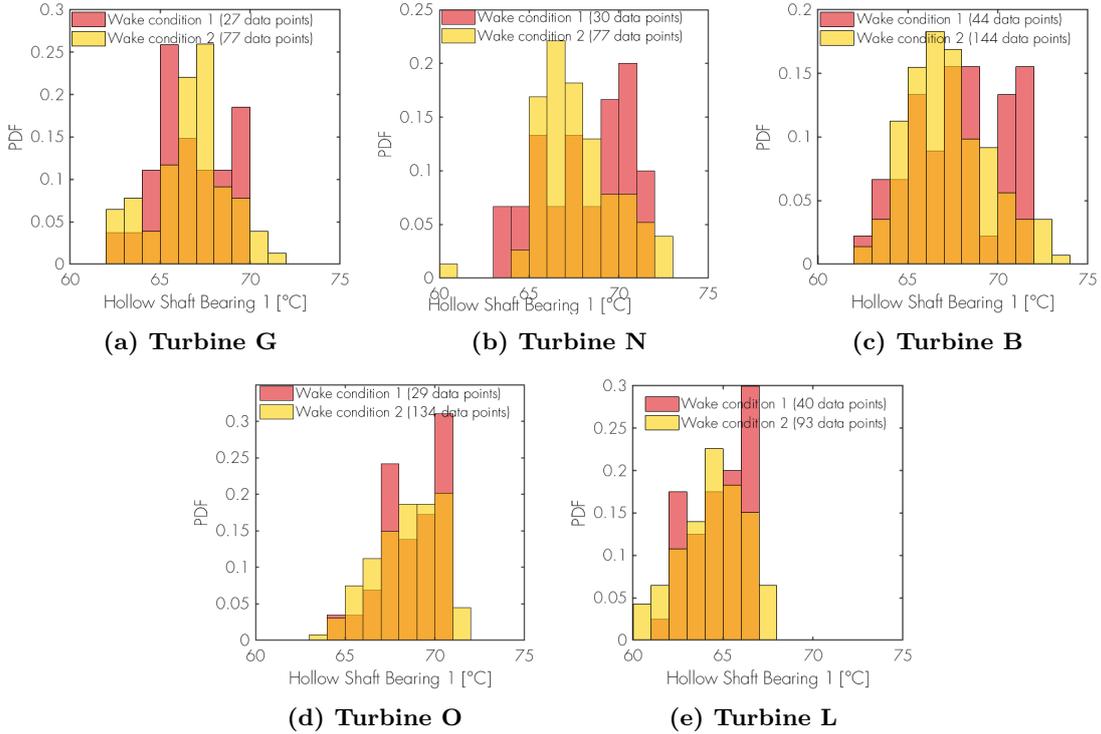


Figure 6.10: Temperature distributions for the hollow shaft bearing under wake condition 1 and 2 clustered over wind speed: 7.5-9.5 [m/s] & $\frac{\Delta U}{U_{infty}}$: 0.07-0.15 [-]

6.4 Discussion on wake effects on component temperatures

The results found in the chapter above are used to answer the last research question and the implications and limitations of these results are discussed in this section.

From the results presented in section 6.3.1 and 6.3.2, it can be seen that no change in the component temperature could be observed when different drivetrain component temperatures are clustered over either the normalized turbulence intensity or the velocity deficit due to wake effects in the OWEZ.

When analysing the different turbines, in the front, middle or back of the wind farm, no significant difference in temperature distributions could be detected for any of the temperature signals from the drivetrain in the SCADA signal. Both bearing, oil and generator winding temperatures signals are analysed. Also, when clustering over the not normalized turbulence intensity or velocity deficit or when varying the cluster bins, no change in the observations can be made.

There are several other possible causes that still need to be eliminated before concluding that wake effects do not influence component temperatures. First of all, the data that is used is hourly mean data. This is due to the DOWA dataset being of this frequency. It is possible that this frequency is too crude to detect short temperature spikes

due to eddies passing through the rotor.

A more plausible hypothesis is that the drivetrain components are too much temperature controlled to detect changes. No information is available on the status of these cooling systems, thus it can not be investigated if more active cooling is present during waked conditions to restrain increasing bearing temperatures related to additional loading.

The last factor that needs to be eliminated before concluding that the temperatures are independent of wake effects is the elimination of thermal inertia. It was found by Churchfield that the spectral peaks of the moment loading on drivetrain components due to wake are related to the 3P frequency [48], what would correspond with 0.81 Hz for rated wind speed conditions of the Vestas V90. Due to the mass of the drivetrain components, the temperature changes are of a much lower frequency. The clustering has not taken time into account or the gradient of temperatures between different data points.

The findings presented for the investigation into the effects of asymmetric wake loading on component temperatures, are made on a limited amount of data. It is interesting to note that either no change of in the temperature distribution could be observed, or a small shift towards higher temperatures. Only one case could be identified where the asymmetric loading resulted in a shift towards lower bearing temperatures.

7 Conclusions and Discussion

Two objectives are formulated for this master thesis. To reach these objectives, 5 research questions are addressed. Chapter 3 and 4 provide answers to research question 2, 3 and 4. The results in chapter 5 present the performance of the developed condition monitoring framework and highlight the answer to research question 1. The last research question, related to the second objective is investigated in chapter 6. The synthesis of the results will be discussed in this chapter. What conclusions can be drawn from the different parts of the condition monitoring framework and the investigation of wake effects on component temperatures? How can Shell and OWEZ utilize the results of this research for condition monitoring of offshore wind farms? And what recommendations can be made for further studies and of the further development of the framework? These questions are answered in this chapter.

7.1 Conclusions

In this research, an anomaly detection tool was successfully developed that can be utilized for condition monitoring and fault detection of the gearbox of the offshore wind turbines at the Egmond aan Zee wind farm. For this, a framework has been developed that accounts for the following steps of condition monitoring: Data proceeding, feature extraction and anomaly detection. Only the data acquisition was not included in the scope as the SCADA data collected at the Egmond aan Zee offshore wind farm is used as a case study. The conclusions presented in this section provide the answers to the research questions and highlight the most important findings.

To proceed the operational SCADA data, 2 filters have been applied on the training data and one to account for errors after hiatus in the data. To create a clean dataset to train the model on, a filter is applied to only consider data from operational conditions and a filter is applied to remove improbable outliers based on their Mahalanobis distance to a cluster mean. The proceeding of data also included a study in the correlation between the different SCADA signals. The Pearson correlation coefficient is used for this as well as an understanding of the system integration to determine the input parameters that can describe the operational conditions of the gearbox and generator without resulting in a biased model.

Different models have been evaluated to describe the normal behaviour of different bearing temperatures in the drivetrain for the feature extraction. After analysing linear regression, polynomial regression, decision trees with either classification or regression task and SVM models with either classification or regression task, it could be concluded that the decision trees with a regression task resulted in generalizing the gearbox components with the smallest error in healthy conditions. The found test error was approximately 0.5-0.6 °C depending on the gearbox signal and turbine. The SVM regression model resulted in generalising responses with a test error of 0.5-1.2 °C depending on the gearbox signal and turbine, but could also result in large errors when generalizing over the higher temperature ranges in the test dataset.

With the evaluated models, it was not possible to fit a model to the generator bearing temperature signals that resulted in a test error lower than 7.15°C for the first generator bearing and 10.0°C for the second bearing temperature signal. This could be assigned to the large variability in the data under similar operational conditions and the influence of different components through the cooling system. The error is distributed around zero but can be as large as 25°C under healthy conditions. This is not considered as accurate enough to be used for anomaly detection as the response ranges over 50°C and a 25°C uncertainty is too large. It is also concluded that providing additional data to the training dataset or introducing additional input signals did not result in a model with a low validation and or test error.

This research was successful in monitoring the condition in the generator winding temperatures through an analysis on the difference between the temperature of the winding in the different phases. This analysis can also be used for bearings on the same shaft in the gearbox.

For the anomaly detection, two different methods have been implemented. The Shewhart control chart is capable in detecting large anomalies shortly prior to the turbine been taken out of production. The CUSUM control chart can be utilized for to detection of variation in the mean of a temperature signal. This resulted in the detection of anomalies as early as one month before the turbine was taken out of production.

With the developed framework, it was possible to identify 4 anomalies for the gearbox bearings in the dataset for 36 turbines over the period of one year. Although, these anomalies could not be validated and linked with known failure cases, it can be concluded that the gearbox was operating in an out-of-control state and coincided with the turbine been taken out of production for a significant period of time.

To model the wake conditions at OWEZ, the wind velocity deficit and added turbulence intensity are calculated using modelled ambient wind conditions from the DOWA and an in-house developed wake tool that utilizes the Ishihara wake model. After clustering the component temperatures over wind speed bins and wake conditions, the difference in the probability density distribution of the component temperatures could be investigated. Visualizing these distributions and the statistical properties led to the conclusion that wake effects do not influence the drivetrain component temperatures.

More interesting are the effects of asymmetric wake loading. It is found that small variations could be detected in the distributions of component temperatures when the turbine is experiencing asymmetric wake loading opposed to wake experienced over the entire rotor area. If there is a shift in the temperature distribution present, then this is most of the cases towards the higher temperature ranges.

7.2 Discussion and recommendations

This section will address the implications of the findings, put the findings in perspective to other research and present recommendations for further research. The limitation of the findings and unexpected results related to the development of the condition monitoring framework have been covered in section 5.4 and the findings related to the second research objective has been discussed in section 6.4.

The first research objective of this masters thesis is phrased as:

To develop a temperature based condition based monitoring framework for Egmond aan Zee offshore wind farm that utilizes the SCADA data for drivetrain component monitoring.

The detected anomalies for turbines G, H, J and K implicate that this objective is reached successfully. A normal behaviour model is used for this approach and this is in line with results found in literature [19; 17; 20] where normal behaviour models were used. The model that is used to describe the healthy temperature response under given operational conditions of the drivetrain components differs from these researches. The results of this research show that a regression model using a decision tree learning algorithm can be used to describe the gearbox bearings with great accuracy.

Simple regression models utilizing one input to describe a temperature response in the drivetrain could be used to obtain an insight in the relation of the temperature to other responses but did not result in an accurate model as was found by Schlechtingen [16]. It was found by Tautz-Weinert that SVM models resulted in a large error [20]. It can be seen from the results that for this dataset the SVM only resulted in a large error over the higher temperature ranges. This illustrates that the model performance is very dependent on the data used for each study. Because of this, it would be interesting for further research to implement this frame work on a different wind farm to see if other models perform better. This would be able to achieve with minimal adjustments due to the parametric set-up of the framework.

The application of ANN for temperature based normal behaviour models has been proven in many researches [19; 16; 20]. It would be interesting to also implement an ANN in this framework and apply on the same dataset for fair comparison and test the performance on an ANN against a decision tree model. The first steps towards the implementation of an ANN in the framework are made to make this comparison. Due to limited time, this was not included in the report or tested for all signals and turbines. The initial findings of this are that an ANN can outperform a decision tree model based on the validation error.

Physics based models are left out of this research due to limited physical information, this would results in too many uncertainties. By creating a model for every turbine and SCADA signal, it was possible to account for different biasses due to calibration of the sensors or deterioration of the drivetrain and increased roughness area of the blades after 12 years of use. This would be hard to account for with a physical model.

Applying the CUSUM control chart developed by Montgomery [46] has resulted in early detection of developing faults. This approach has been used for fault detection based on vibrations in the gearbox of a floating wind turbine [50] but the combination of a CUSUM fault detection with temperature based condition monitoring was not found in past literature to the best of the authors ability.

In further research, an improvement could be made with regard to the threshold values used by the Sherward control chart. The assumption, based on literature [16], was made in this research that a threshold value of $A\sigma$ with $A=5$ was sufficient to detect anomalies. The value should be benchmarked with known failure cases and lower values of A could be tested to allow for earlier detection of faults. Other fault detections threshold designs that are less conservative are interesting to look at for further research. An example of this is a time-varying threshold value based on a boundary that fits the residual error of healthy responses over different operational conditions, as proposed by Ferrari [51].

For turbines that are equipped with either accelerometer or acoustic sensors, it would be interesting for further research to see at which state the anomalies would be detected by vibrational or acoustic based condition monitoring techniques compared to the developed framework that relies on temperature based analyses.

The minimum required training period is not addressed in this research and a conservative approach is taken by selecting the longest continuous period where the turbine was showing normal behaviour to include all possible variations in the operational conditions of the training dataset. The result of this was a training period of approximately 2 to 6 weeks, depending on the turbine. The model trained on the smallest training datasets show a well performing model, this highlights that a training period of 2 weeks containing 10 minute average values is sufficient to train a decision tree regression model on. It would be interesting to further investigate the minimum required size of the training dataset for an accurate model.

The implication of the results of this research is that Shell can utilize the developed framework for the condition monitoring of OWEZ. The framework can be used to provide the wind farm operator with an insight in abnormal behaviour of the gearbox in all of the turbines. Utilizing the different control charts, anomalies can be track as early as possible allowing the operator to perform preventive maintenance. However, before this framework can be implemented at OWEZ or other wind farms, several additional steps need to be taken. The framework needs to be validated with known failure cases as has been mentioned in section 5.4. Furthermore, additional steps need to be taken to allow for real time monitoring opposed to hindsight analyses based on historical datasets.

The second research objective of this masters thesis is phrased as:

Use the SCADA data to identify wake effects on drivetrain components by a comparison of component temperature distribution under different wake conditions.

When analysing the different temperature distributions, it could be seen that no influence of wake effects on drivetrain component temperatures could be observed. The number of elements per cluster indicate that most of the time the turbines at OWEZ are experiencing conditions close to free stream conditions. It would be interesting to investigate the effects of wakes on component temperatures in a different wind farm where the spacing between turbines is smaller.

With respect to asymmetric loading, only a small amount of data points could be identified when a turbine was experiencing asymmetric wake loading over the rotor area and also experience a high velocity deficit from the turbine that created these wake conditions. From the temperature distributions, small variations could be observed between the asymmetric loading and wake experienced over the entire rotor. One case is highlighted where this variation resulted in a shift towards lower temperatures, this distribution was made on the lowest amount of data points for all the considered cases.

All the other cases showed either no shift in temperatures or a shift of the distribution towards higher bearing temperatures. Turbine N and B showed the most deviating behaviours with a large shift towards higher temperatures when under asymmetric wake loading. This indicates that asymmetric loading does influence component temperatures more than wake that is experienced over the entire rotor plane but more research in this field is needed to validate these findings. These results are presented with a note of causation as the results are obtained over a very limited amount of data. This however indicates that further research into these effects would be interesting.

References

- [1] WindEurope. Offshore wind in europe key trends and statistics 2017. WindEurope, 2018. URL <https://windeurope.org/wp-content/uploads/files/about-wind/statistics/WindEurope-Annual-Offshore-Statistics-2017.pdf>.
- [2] Ryan Wiser, Karen Jenni, Joachim Seel, Erin Baker, Maureen Hand, Eric Lantz, and Aaron Smith. Expert elicitation survey on future wind energy costs. *Nature Energy*, 1(10):16135, 2016.
- [3] R Bos. *Extreme gusts and their role in wind turbine design*. PhD thesis, TU Delft, 2017.
- [4] W. Yang. *Offshore Wind Farms Technologies, Design and Operation*, chapter Condition monitoring of offshore wind turbines. OWoodhead Publishing, Oxford, 2016. doi: <https://doi.org/10.1016/C2014-0-00763-0>.
- [5] Shuangwen Sheng and Paul S Veers. *Wind turbine drivetrain condition monitoring-an overview*. National Renewable Energy Laboratory Golden, Colo, USA, 2011.
- [6] Sebastian Pfaffel, Stefan Faulstich, and Kurt Rohrig. Performance and reliability of wind turbines: A review. *Energies*, 10(11):1904, 2017.
- [7] Stefan Faulstich, Berthold Hahn, and Peter J Tavner. Wind turbine downtime and its importance for offshore deployment. *Wind energy*, 14(3):327–337, 2011.
- [8] IRENA. Renewable power generation costs in 2017. International Renewable Energy Agency (IRENA), 2018. ISBN 978-92-9260-040-2.
- [9] Pierre Tchakoua, René Wamkeue, Mohand Ouhrouche, Fouad Slaoui-Hasnaoui, Tommy Tameghe, and Gabriel Ekemb. Wind turbine condition monitoring: State-of-the-art review, new trends, and future challenges. *Energies*, 7(4):2595–2630, 2014.
- [10] Allan May, David McMillan, and Sebastian Thöns. Economic analysis of condition monitoring systems for offshore wind turbine sub-systems. *IET Renewable Power Generation*, 9(8):900–907, 2015.
- [11] NEK IEC 61400-25-1. Iec 61400 part 25-1: Communications for monitoring and control of wind power plants - overall description of principles and models. Standard, July 2017.
- [12] ISO 10816-21 Wind Turbine. Iso 10816-21 mechanical vibration - evaluation of machine vibration by measurements on non-rotating parts part 21: Onshore wind turbines with gearbox. Standard, International Organization for Standardization (ISO), Geneva, Switzerland, July 2017.
- [13] Slim Soua, Paul Van Lieshout, Asanka Perera, Tat-Hean Gan, and Bryan Bridge. Determination of the combined vibrational and acoustic emission signature of a wind turbine gearbox and generator shaft in service as a pre-requisite for effective condition monitoring. *Renewable Energy*, 51:175–181, 2013.

- [14] TH Loutas, G Sotiriades, I Kalaitzoglou, and V Kostopoulos. Condition monitoring of a single-stage gearbox with artificially induced gear cracks utilizing on-line vibration and acoustic emission measurements. *Applied Acoustics*, 70(9):1148–1159, 2009.
- [15] Jan Helsen, Christof Devriendt, Wout Weijtjens, and Patrick Guillaume. Condition monitoring by means of scada analysis. In *Proceedings of European Wind Energy Association International Conference Paris*, 2015.
- [16] Meik Schlechtingen and Ilmar Ferreira Santos. Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. *Mechanical systems and signal processing*, 25(5):1849–1875, 2011.
- [17] William G Garlick, Roger Dixon, and Simon J Watson. A model-based approach to wind turbine condition monitoring using scada data. ICSE, 2009.
- [18] Yingning Qiu, Yanhui Feng, Juan Sun, Wenxiu Zhang, and David Infield. Applying thermophysics for wind turbine drivetrain fault diagnosis using scada data. *IET Renewable Power Generation*, 10(5):661–668, 2016.
- [19] Mari Cruz Garcia, Miguel A Sanz-Bobi, and Javier Del Pico. Simap: Intelligent system for predictive maintenance: Application to the health condition monitoring of a windturbine gearbox. *Computers in Industry*, 57(6):552–568, 2006.
- [20] Jannis Tautz-Weinert and Simon J Watson. Comparison of different modelling approaches of drive train temperature for the purposes of wind turbine failure detection. In *Journal of Physics: Conference Series*, volume 753, page 072014. IOP Publishing, 2016.
- [21] NoordzeeWind. Noordzee wind webpage. <http://www.noordzeewind.nl>. 2019-05-01, .
- [22] NoordzeeWind. Offshore wind farm egmond aan zee general report, .
- [23] H.A. Larsen G.C. Hansen K.S. Larsen, T.J. Madsen. Validation of the dynamic wake meander model for loads and power production in the egmond aan zee wind farm. *WIND ENERGY*, 16:605–624, 2013. ISSN 1752-1416. doi: DOI:10.1002/we.1563.
- [24] Jan Willem Wagenaar and Pieter Jan Eecen. *Measurements of Wind, Wave and Currents at the Offshore Wind Farm Egmond aan Zee*. ECN, 2009.
- [25] Vestas. Mechanical operating and maintenance manual v90 – 3.0 mw, vcrs 60 hz. Vestas, 2007. Appendix H.
- [26] Vestas. General specification for v90 – 3.0 mw. Vestas, 2004. URL http://www.gov.pe.ca/photos/sites/envengfor/file/950010R1_V90-GeneralSpecification.pdf.
- [27] A.L. Samuel. Some studies in machine learning using the game of checkers. *IBM Journal*, 3(3):211–229, 1959.
- [28] Vladimir Vapnik. *The nature of statistical learning theory*. Springer, 1995.

- [29] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [30] Ethem Alpaydin. *Introduction to machine learning*. MIT press, 2009.
- [31] S Lall and S. Boyd. Lecture slides: Supervised learning via empirical risk minimization.
- [32] L Breiman, J Friedman, R Olshen, and C Stone. *Classification and regression trees (monterey, california: Wadsworth)*, 1984.
- [33] DOWA. Dutch offshore wind atlas webpage. <https://www.dutchoffshorewindatlas.nl/>. 2019-05-01.
- [34] JAJ Donkers, AJ Brand, and PJ Eecen. Offshore wind atlas of the dutch part of the north sea. *EWEA 2011 Annual Event, Brussels*, 2011.
- [35] P Bangalore, S Letzgus, D Karlsson, and M Patriksson. An artificial neural network-based condition monitoring method for wind turbines, with application to the monitoring of the gearbox. *Wind Energy*, 20(8):1421–1438, 2017.
- [36] Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. Pearson correlation coefficient. In *Noise reduction in speech processing*, pages 1–4. Springer, 2009.
- [37] B. van Delft T. Harman K Wilkinson, M. Darnell. Comparison of methods of wind turbine condition monitoring with scada data. *IET Renewable Power Generation*, 8(4):390–397, 2014. ISSN 1752-1416. doi: Doi10.1049/ioet-rpg.2013.0318. URL [test](#).
- [38] Francesco Castellani, Alberto Garinei, Ludovico Terzi, Davide Astolfi, and Mario Gaudiosi. Improving windfarm operation practice through numerical modelling and supervisory control and data acquisition data analysis. *IET Renewable Power Generation*, 8(4):367–379, 2014.
- [39] F. Tarzi L. Astolfi, D. Castellani. Fault prevention and diagnosis through scada temperature data analysis of an onshore wind farm. *Diagnostyka*, 15(2):71–78, 2014. ISSN test. doi: Dtest. URL [test](#).
- [40] Wei-Yin Loh. Fifty years of classification and regression trees. *International Statistical Review*, 82(3):329–348, 2014.
- [41] James N Morgan and John A Sonquist. Problems in the analysis of survey data, and a proposal. *Journal of the American statistical association*, 58(302):415–434, 1963.
- [42] Biswajeet Pradhan. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using gis. *Computers & Geosciences*, 51:350–365, 2013.
- [43] Alexey Ya Chervonenkis. Early history of support vector machines. In *Empirical Inference*, pages 13–20. Springer, 2013.
- [44] Christopher JC Burges. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2):121–167, 1998.

- [45] Armand Vallin Feigenbaum. *Total quality control*. McGraw-Hill.
- [46] Douglas C Montgomery. *Introduction to statistical quality control*. John Wiley & Sons, 2007.
- [47] LJ Vermeer, Jens Nørkær Sørensen, and Antonio Crespo. Wind turbine wake aerodynamics. *Progress in aerospace sciences*, 39(6-7):467–510, 2003.
- [48] Sang Lee, Matthew Churchfield, Patrick Moriarty, J Jonkman, and J Michalakes. Atmospheric and wake turbulence impacts on wind turbine fatigue loadings. In *50th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, page 540, 2012.
- [49] Takeshi Ishihara and Guo-Wei Qian. A new gaussian-based analytical wake model for wind turbines considering ambient turbulence intensities and thrust coefficient effects. *Journal of Wind Engineering and Industrial Aerodynamics*, 177:275–292, 2018.
- [50] Mahdi Ghane, Amir R Nejad, Mogens Blanke, Zhen Gao, and Torgeir Moan. Statistical fault diagnosis of wind turbine drivetrain applied to a 5mw floating wind turbine. In *Journal of Physics: Conference Series*, volume 753, page 052017. IOP Publishing, 2016.
- [51] Francesca Boem, Riccardo MG Ferrari, Christodoulos Keliris, Thomas Parisini, and Marios M Polycarpou. A distributed networked approach for fault detection of large-scale systems. *IEEE Transactions on Automatic Control*, 62(1):18–33, 2016.

Appendix

A Correlation coefficient with respect to Hollow Shaft bearing 1 temperature

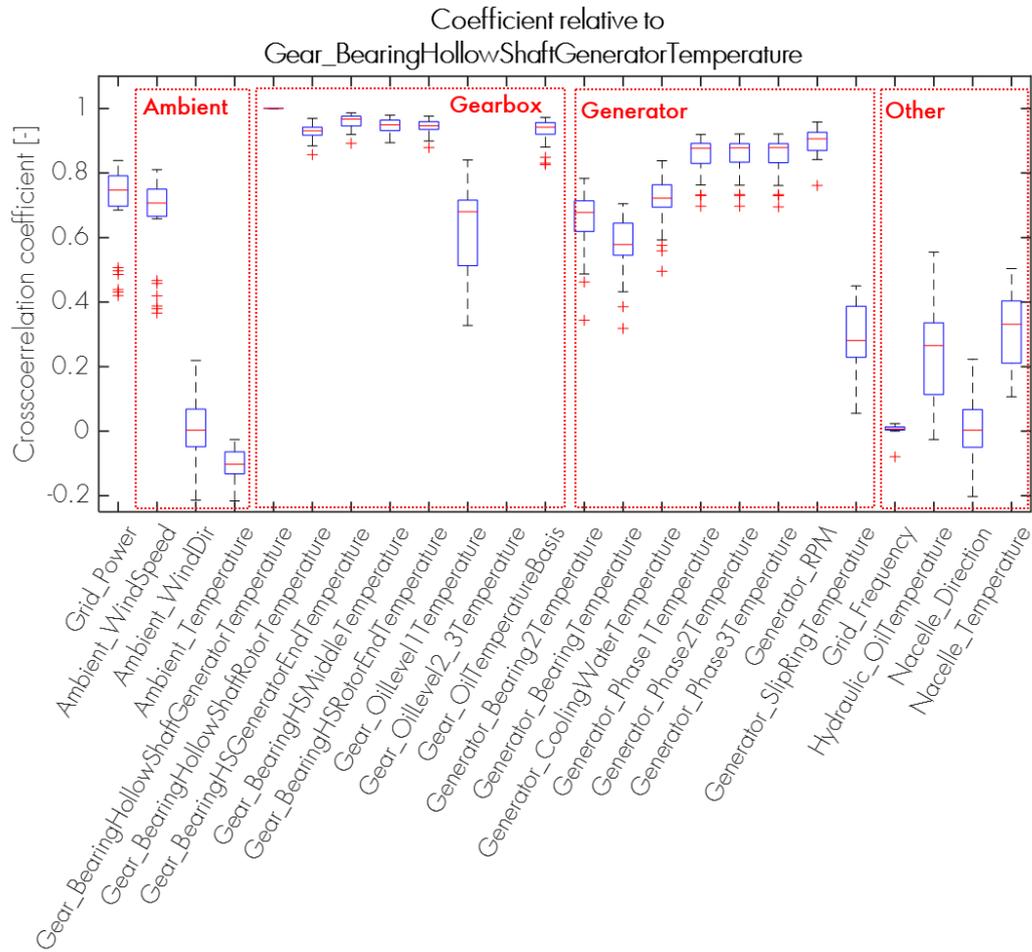


Figure A.1: Correlation coefficient with respect to Hollow Shaft bearing 1

B Correlation coefficient with respect to High Speed Shaft bearing 1 temperature

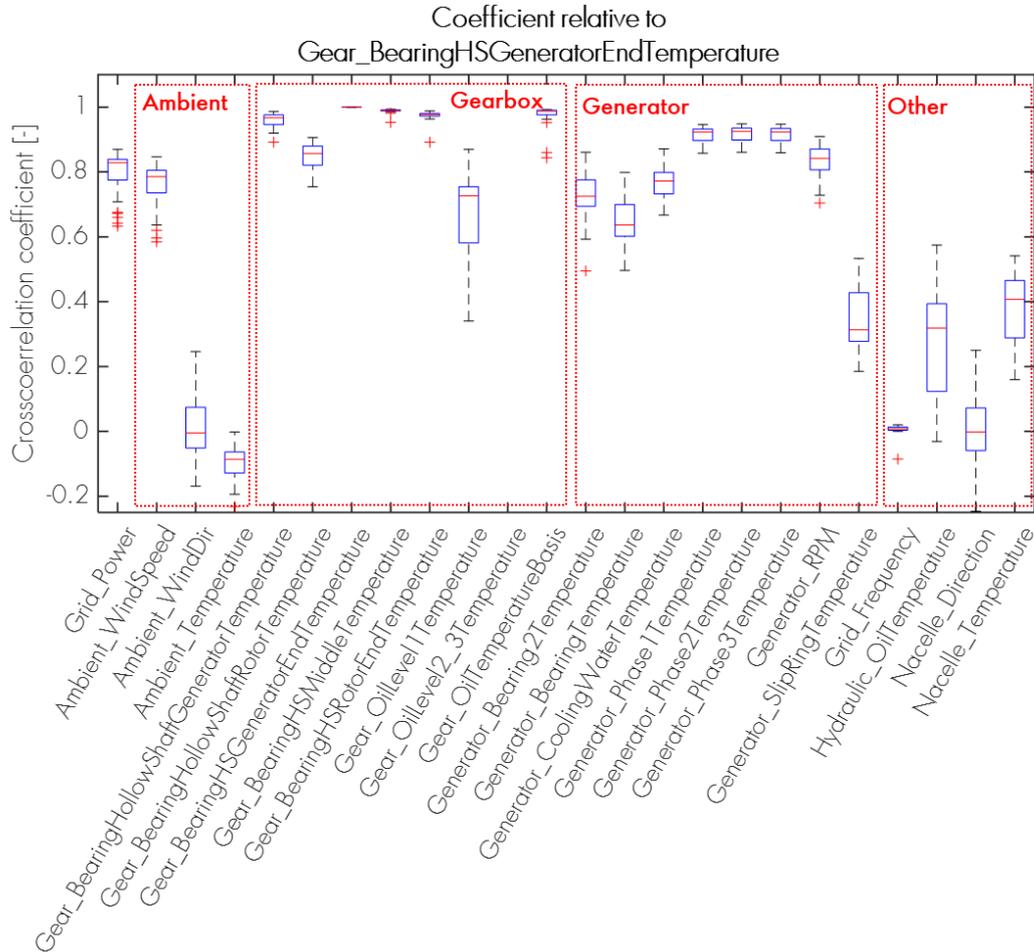


Figure B.2: Correlation coefficient with respect to High Speed Shaft bearing 1 temperature

C Correlation coefficient with respect to generator bearing 1 temperature

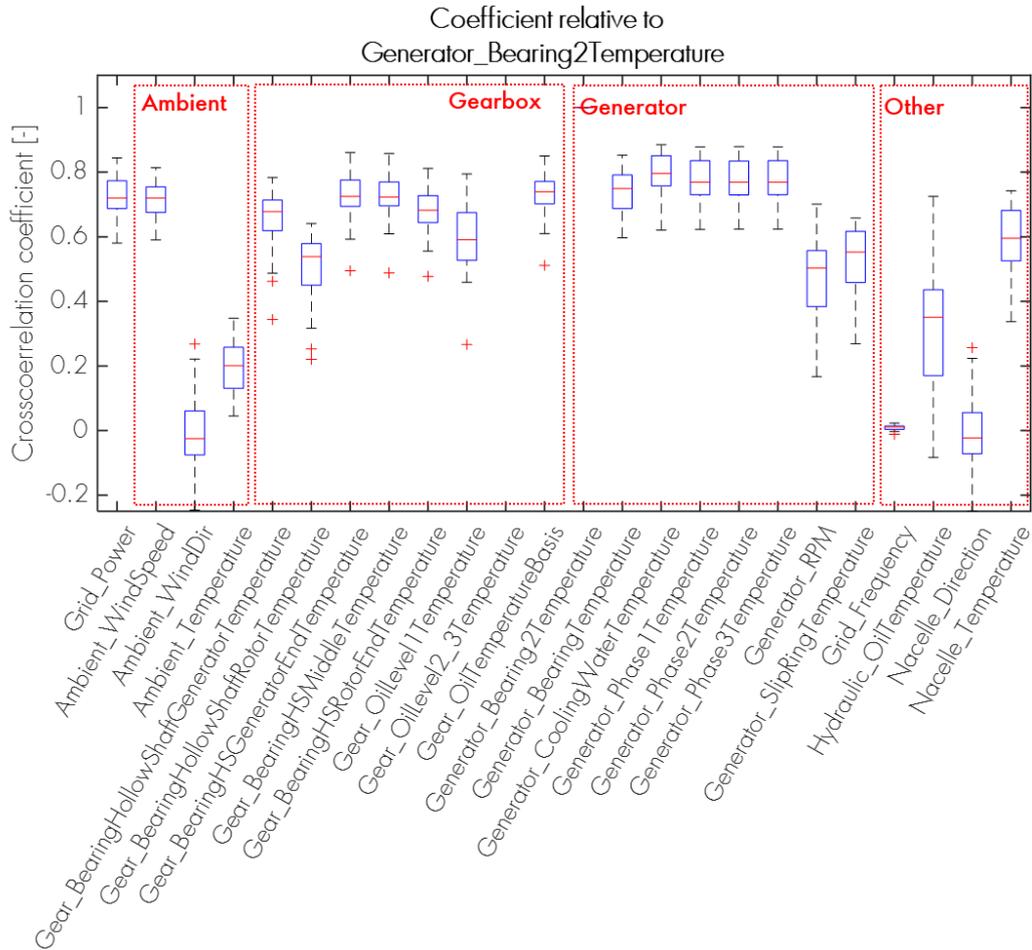


Figure C.3: Correlation coefficient with respect to generator bearing 1

D Correlation coefficient with respect to slipping temperature

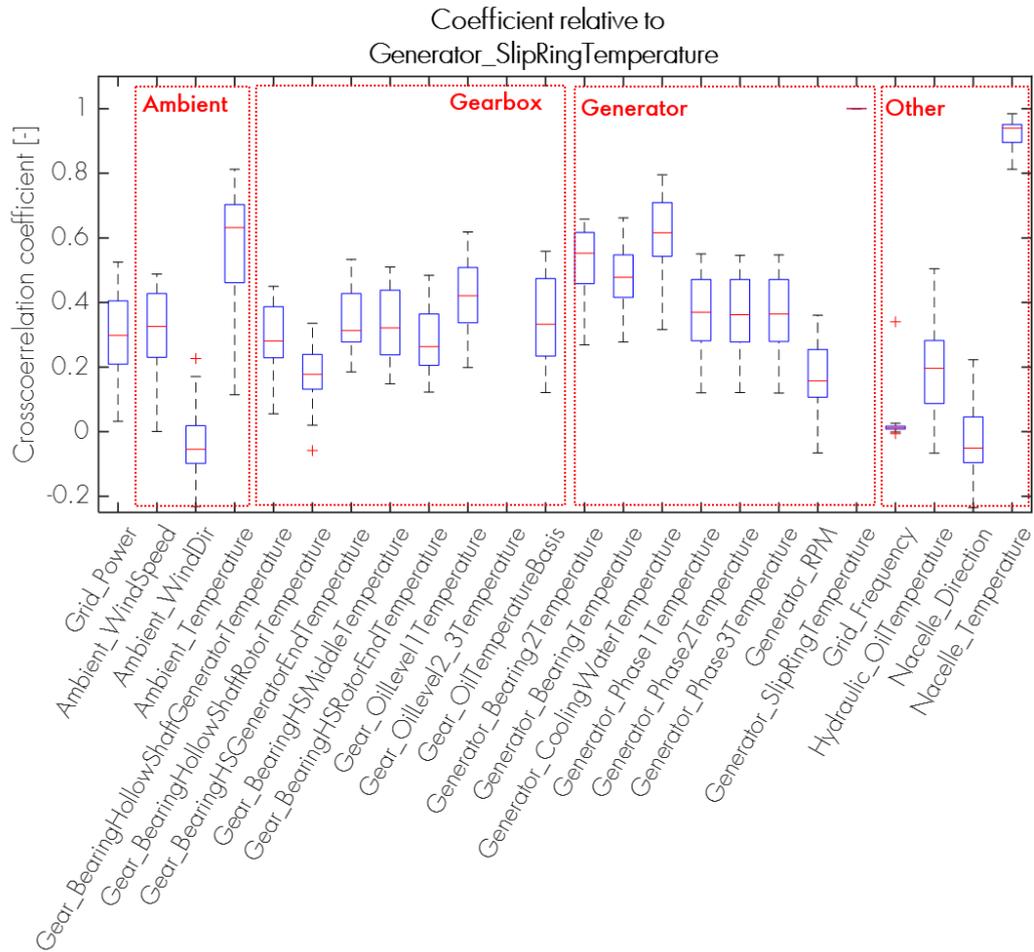


Figure D.4: Correlation coefficient with respect to slipping temperature

E Cooling water circuit

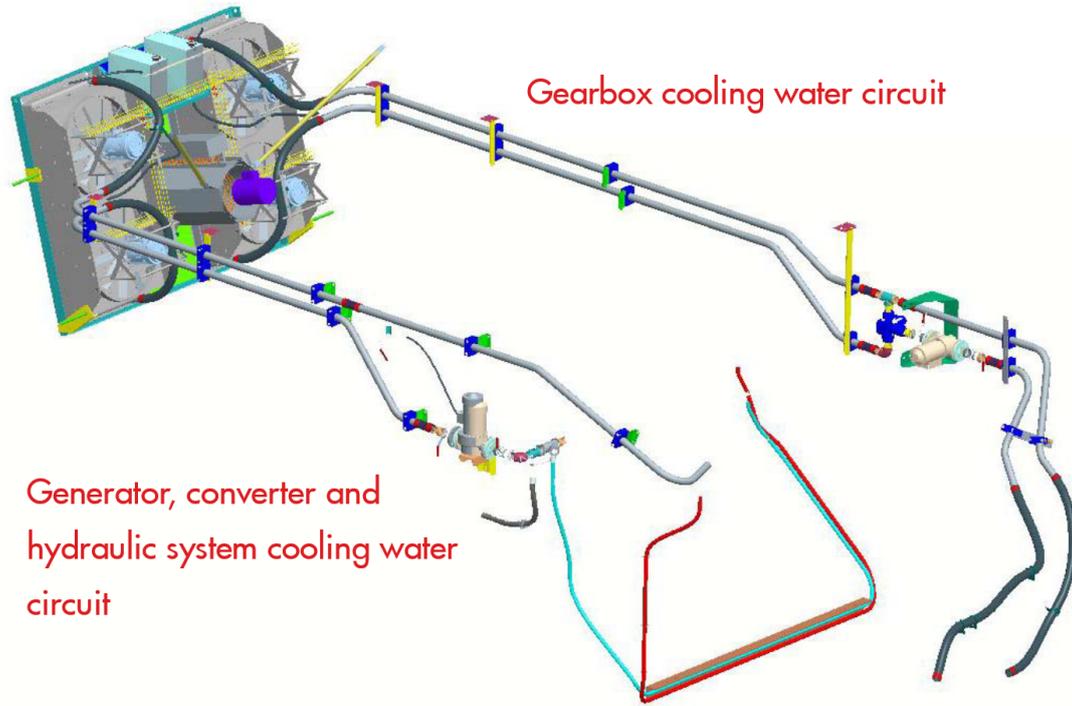


Figure E.5: Two cooling water circuit for drive train components [25]

F Turbine A, training dataset

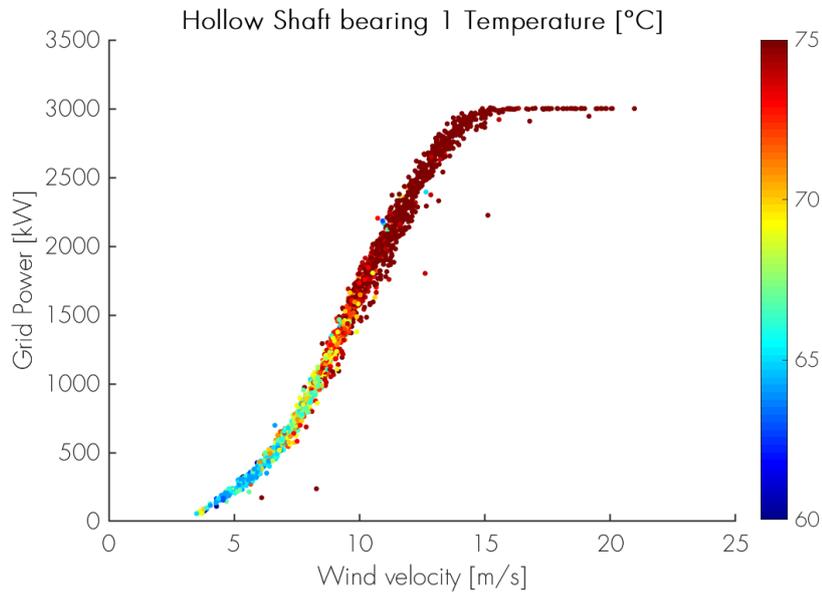


Figure F.6: Turbine A in-sample dataset, Hollow shaft bearing 1 in relation to the power curve

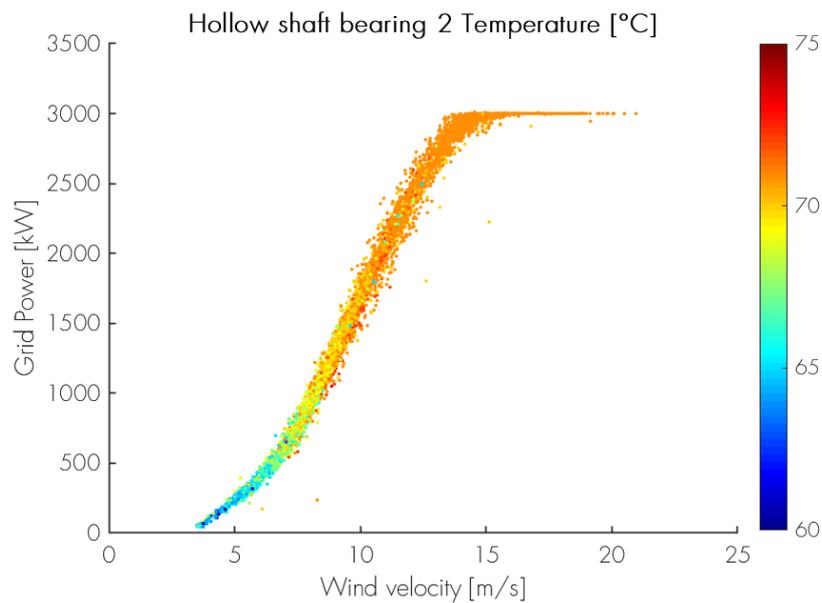


Figure F.7: Turbine A in-sample dataset, Hollow shaft bearing 2 in relation to the power curve

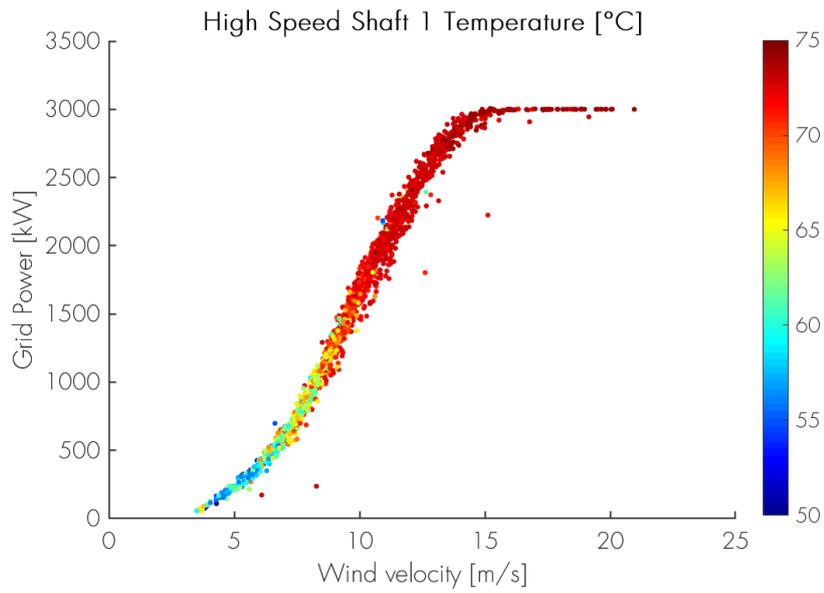


Figure F.8: Turbine A in-sample dataset, High Speed shaft bearing 1 in relation to the power curve

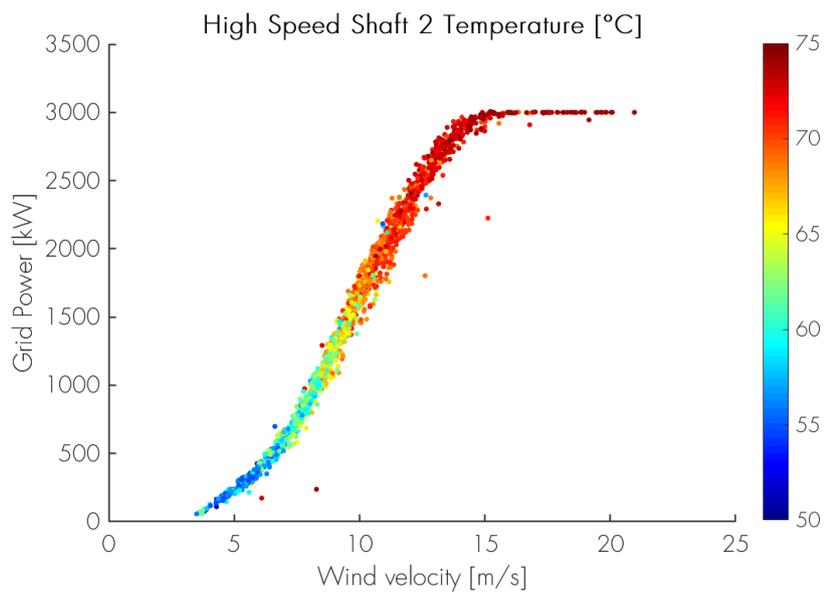


Figure F.9: Turbine A in-sample dataset, High Speed shaft bearing 2 in relation to the power curve

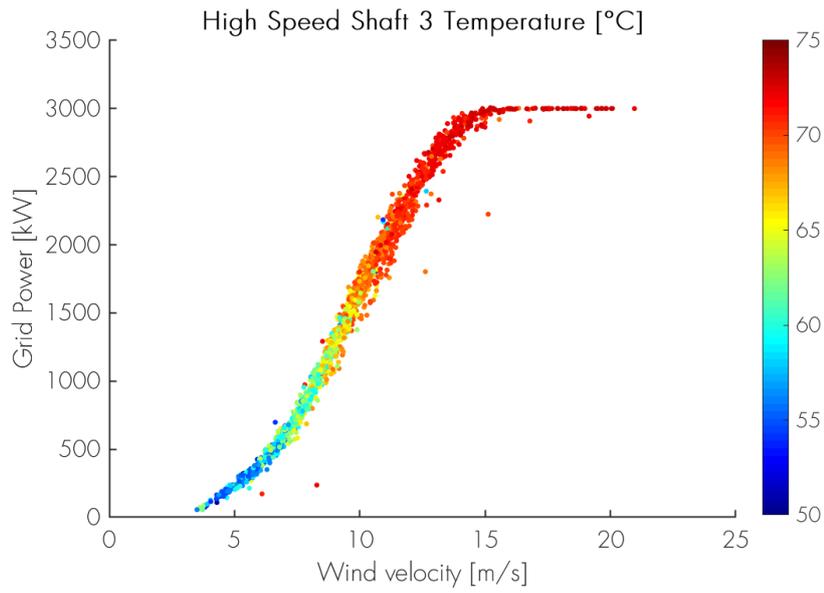


Figure F.10: Turbine A in-sample dataset, High Speed shaft bearing 3 in relation to the power curve

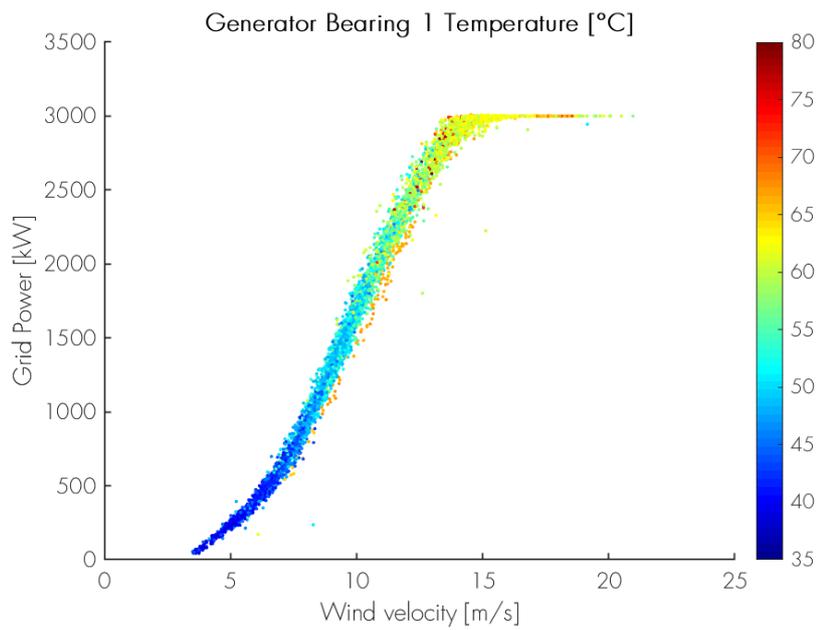


Figure F.11: Turbine A in-sample dataset, Generator bearing 1 in relation to the power curve

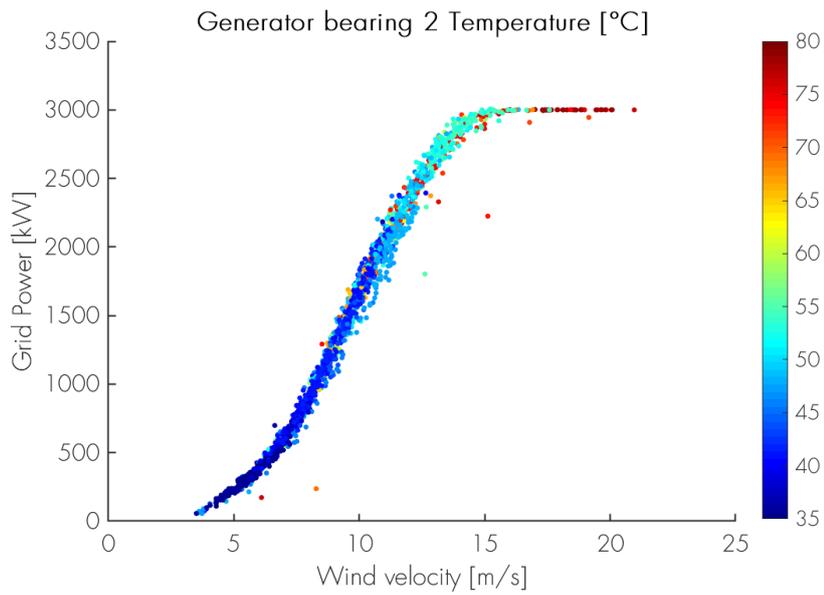


Figure F.12: Turbine A in-sample dataset, Generator bearing 2 in relation to the power curve

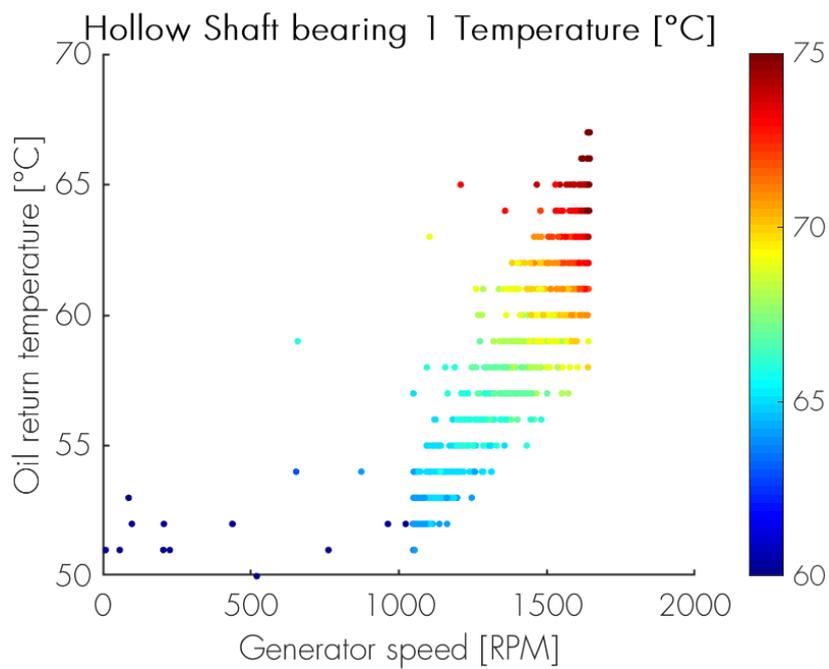


Figure F.13: Turbine A in-sample dataset, Hollow shaft bearing 1 in relation to the generator speed and return oil temperature

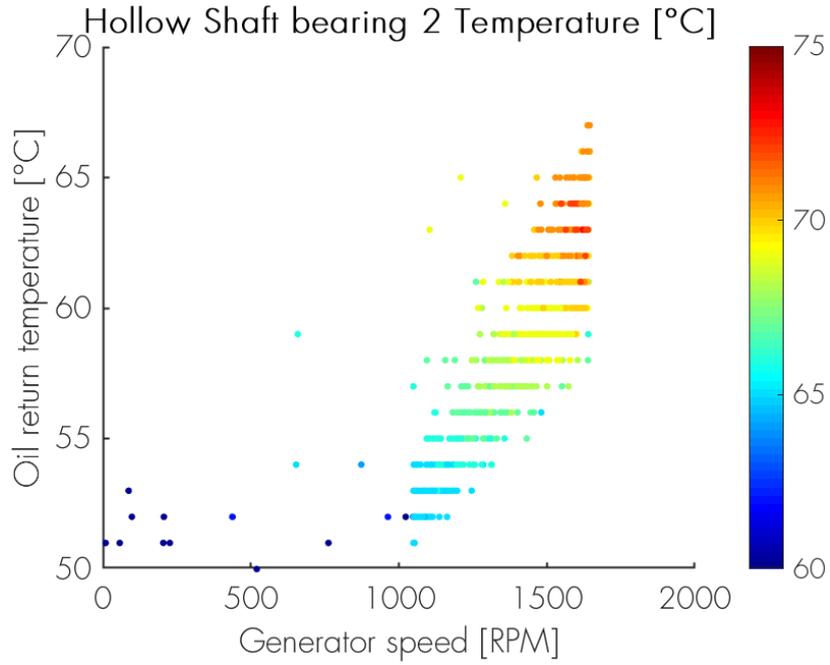


Figure F.14: Turbine A in-sample dataset, Hollow shaft bearing 2 in relation to the generator speed and return oil temperature

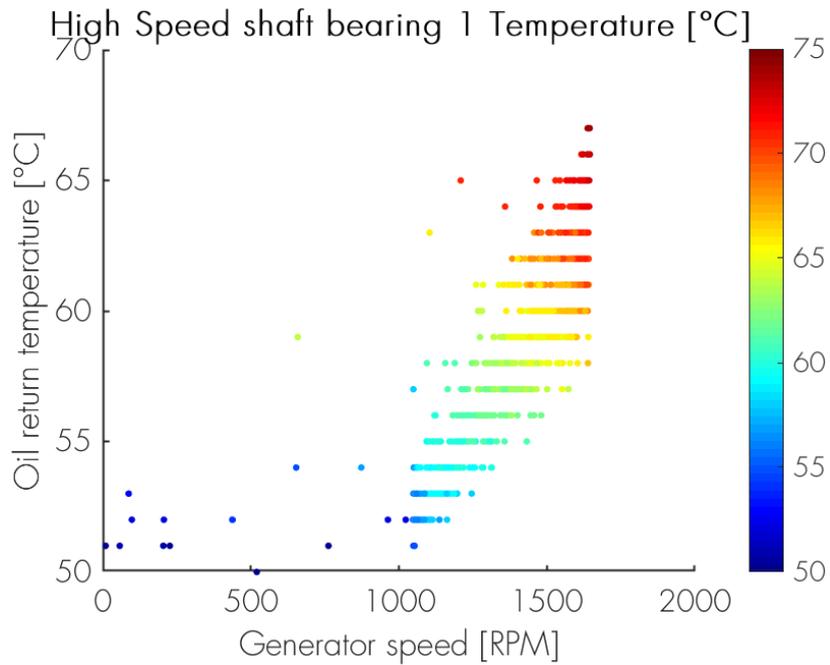


Figure F.15: Turbine A in-sample dataset, High Speed shaft bearing 1 in relation to the generator speed and return oil temperature

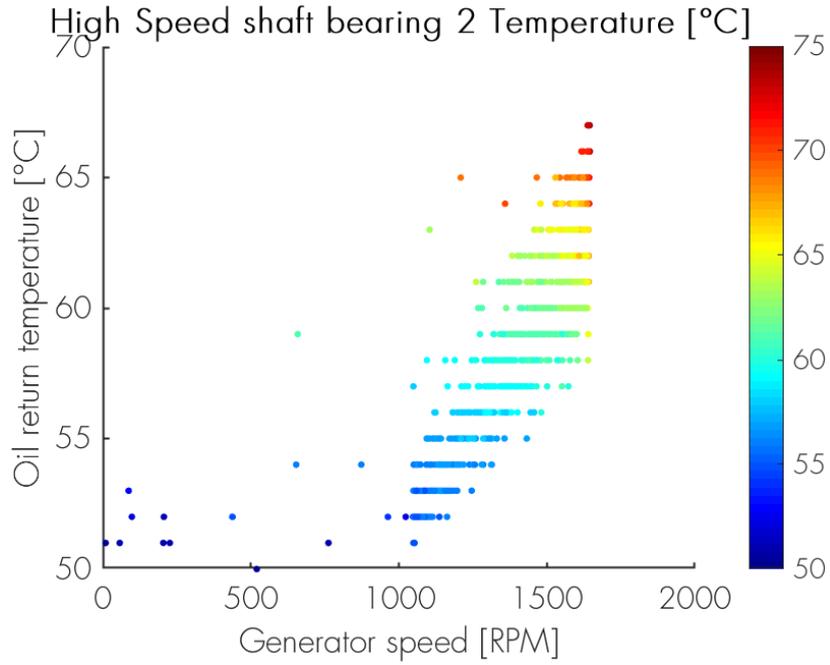


Figure F.16: Turbine A in-sample dataset, High Speed shaft bearing 2 in relation to the generator speed and return oil temperature

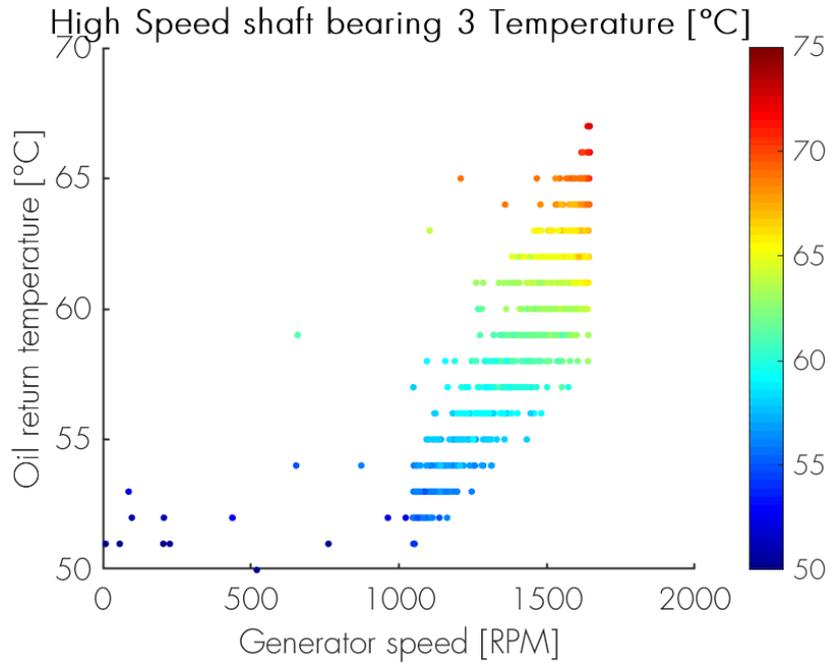


Figure F.17: Turbine A in-sample dataset, High Speed shaft bearing 3 in relation to the generator speed and return oil temperature

G Turbine A, test dataset

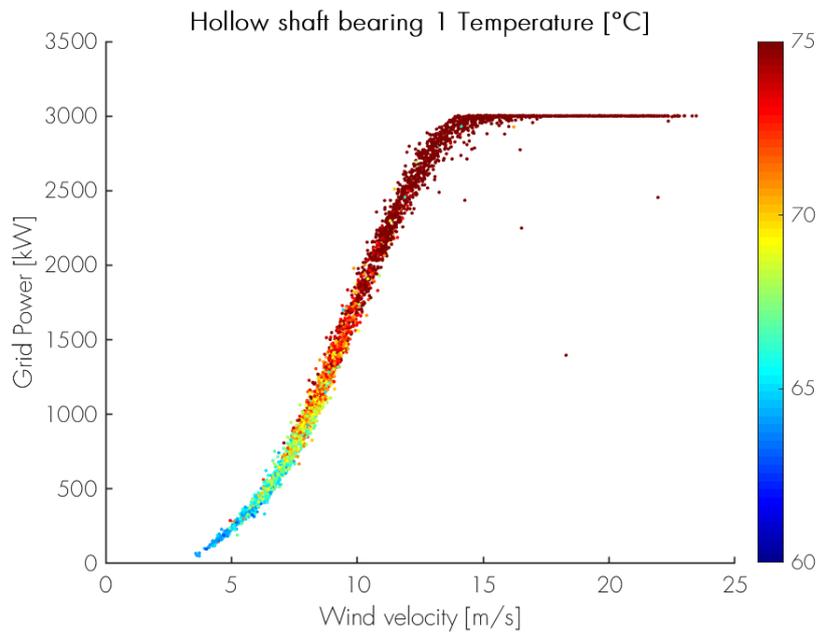


Figure G.18: Turbine A out-of-sample dataset, Hollow shaft bearing 1 in relation to the power curve

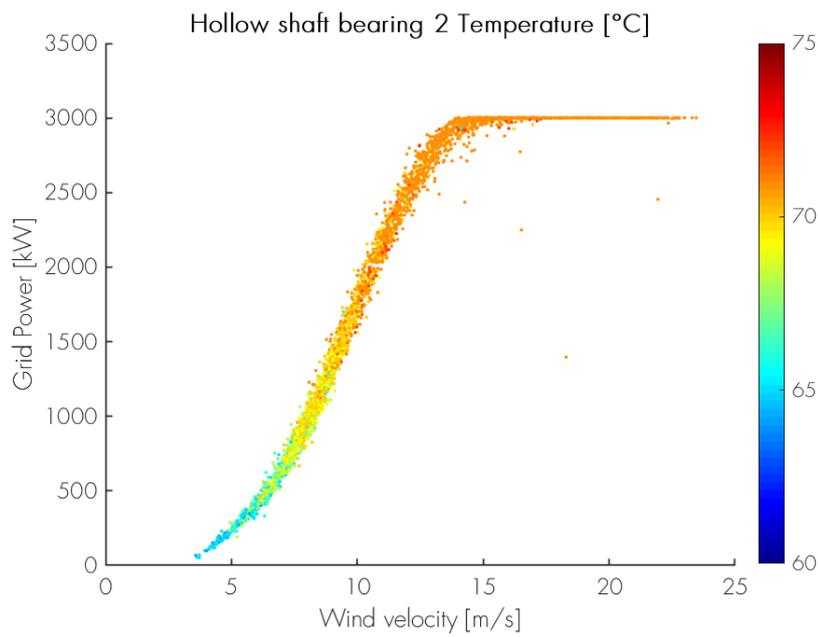


Figure G.19: Turbine A out-of-sample dataset, Hollow shaft bearing 2 in relation to the power curve

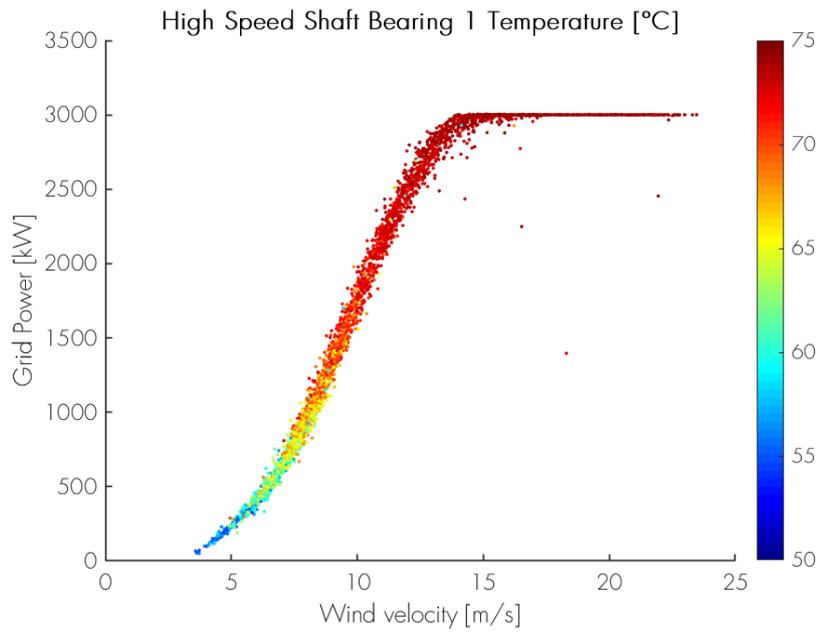


Figure G.20: Turbine A out-of-sample dataset, High Speed shaft bearing 1 in relation to the power curve

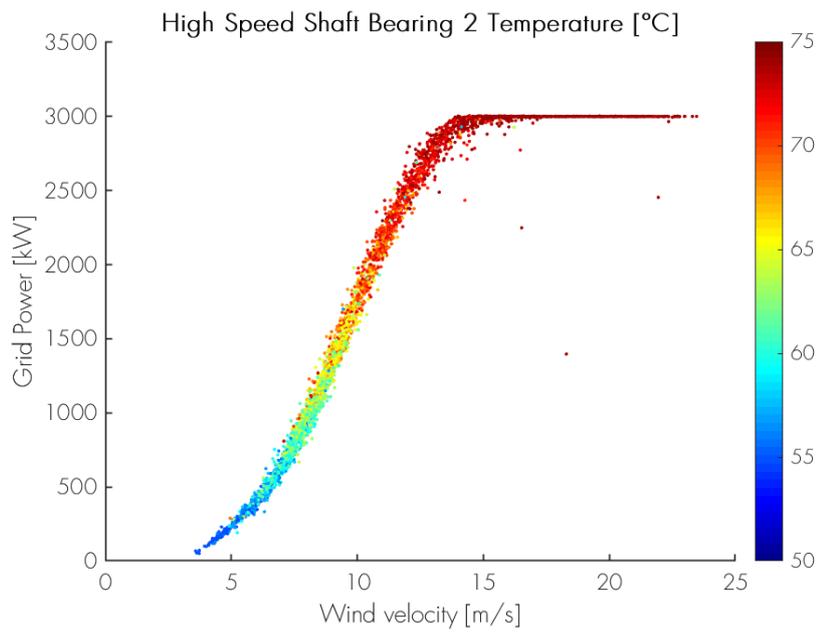


Figure G.21: Turbine A out-of-sample dataset, High Speed shaft bearing 2 in relation to the power curve

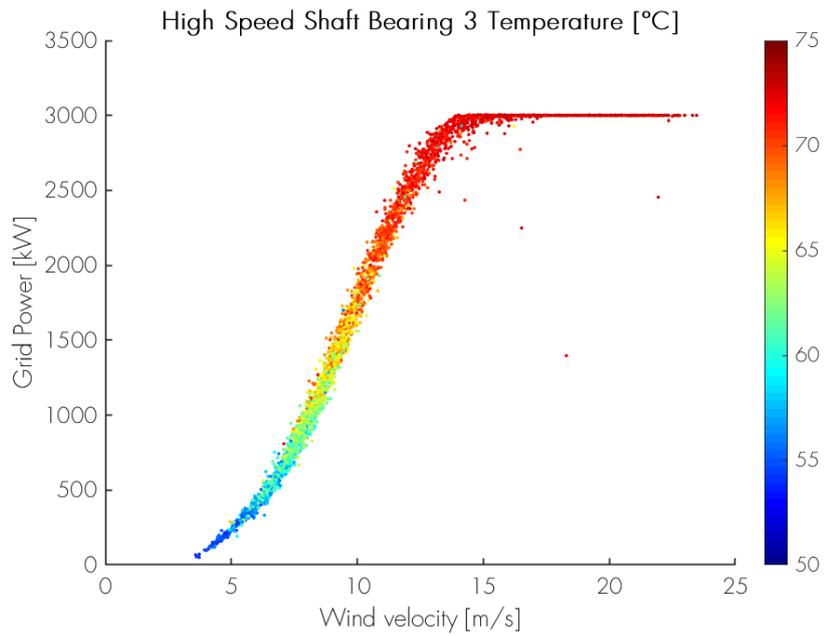


Figure G.22: Turbine A out-of-sample dataset, High Speed shaft bearing 3 in relation to the power curve

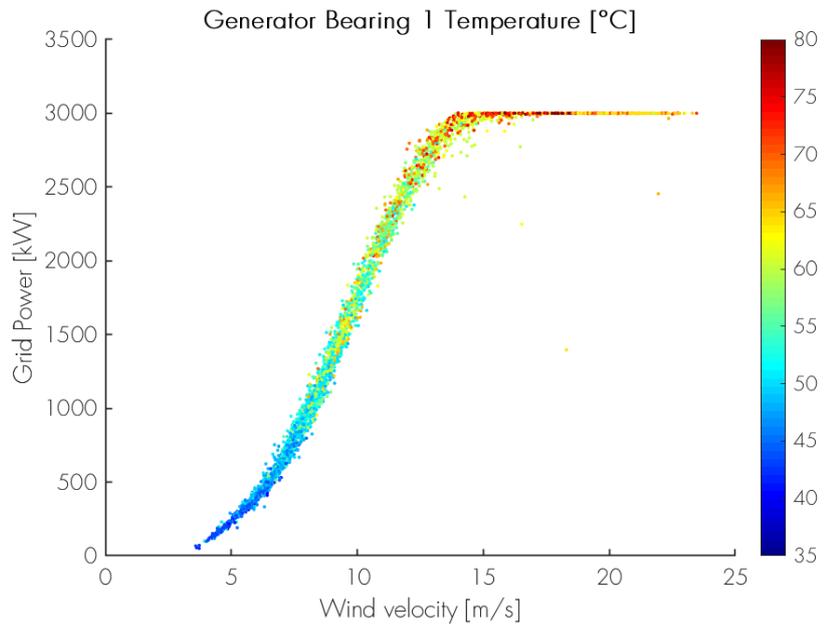


Figure G.23: Turbine A out-of-sample dataset, Generator bearing 1 in relation to the power curve

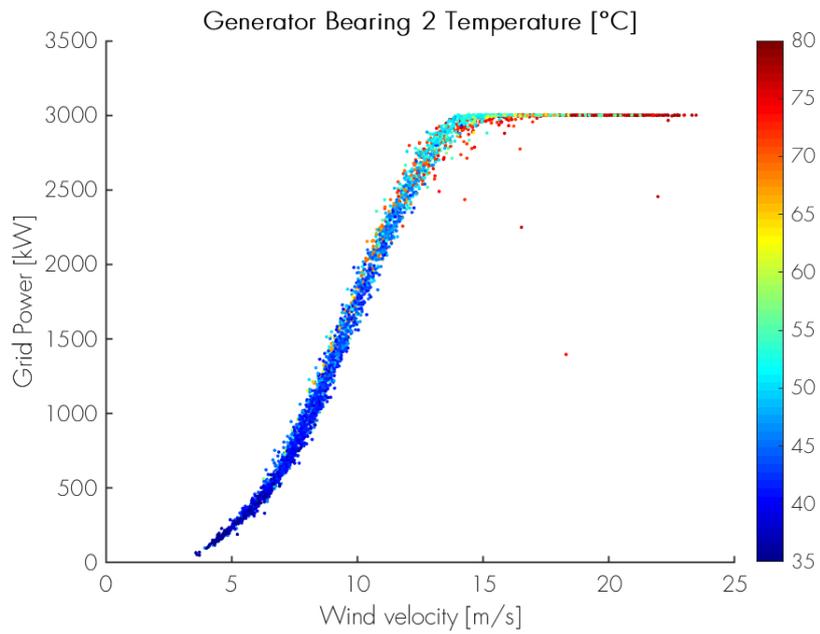


Figure G.24: Turbine A out-of-sample dataset, Generator bearing 2 in relation to the power curve

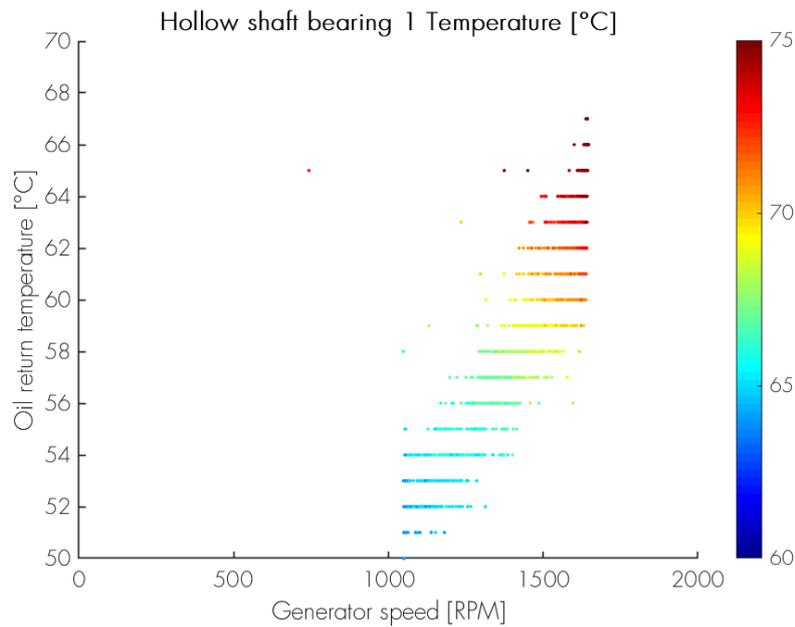


Figure G.25: Turbine A out-of-sample dataset, Hollow shaft bearing 1 in relation to the generator speed and return oil temperature

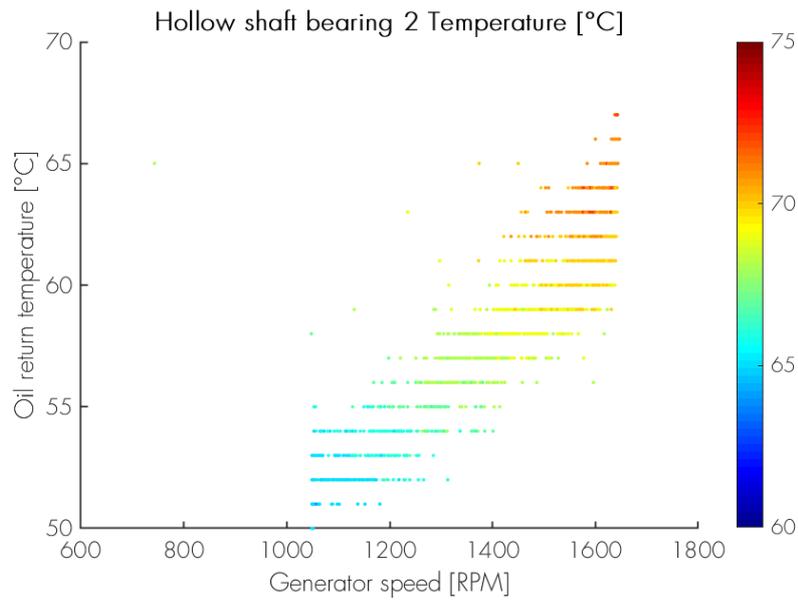


Figure G.26: Turbine A out-of-sample dataset, Hollow shaft bearing 2 in relation to the generator speed and return oil temperature

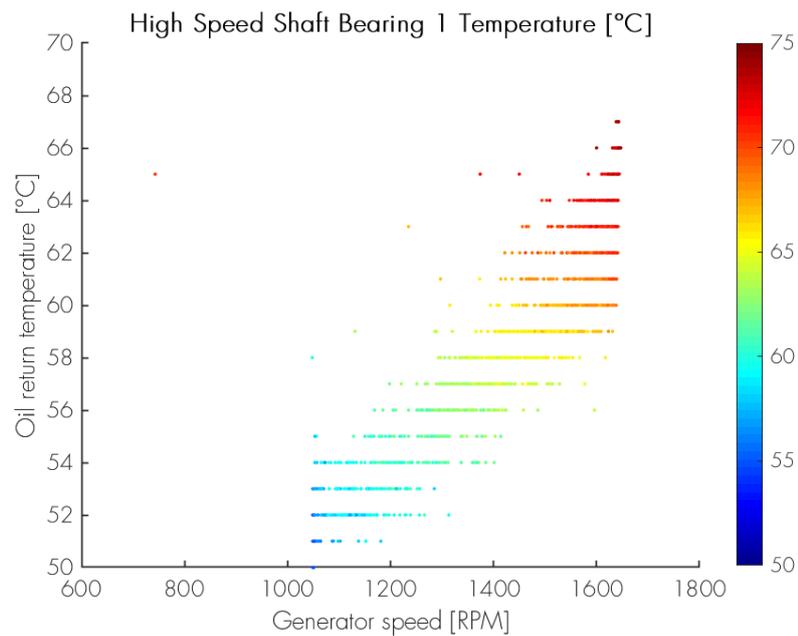


Figure G.27: Turbine A out-of-sample dataset, High Speed shaft bearing 1 in relation to the generator speed and return oil temperature

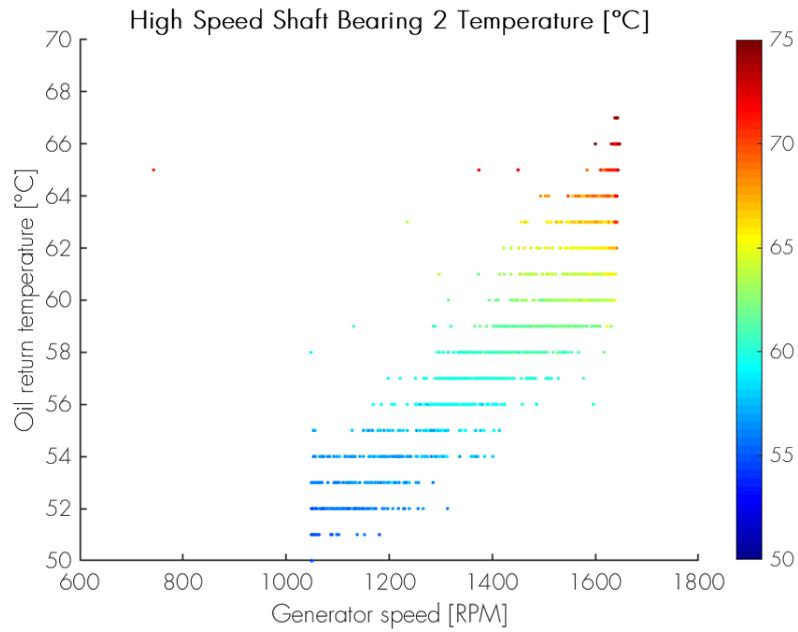


Figure G.28: Turbine A out-of-sample dataset, High Speed shaft bearing 2 in relation to the generator speed and return oil temperature

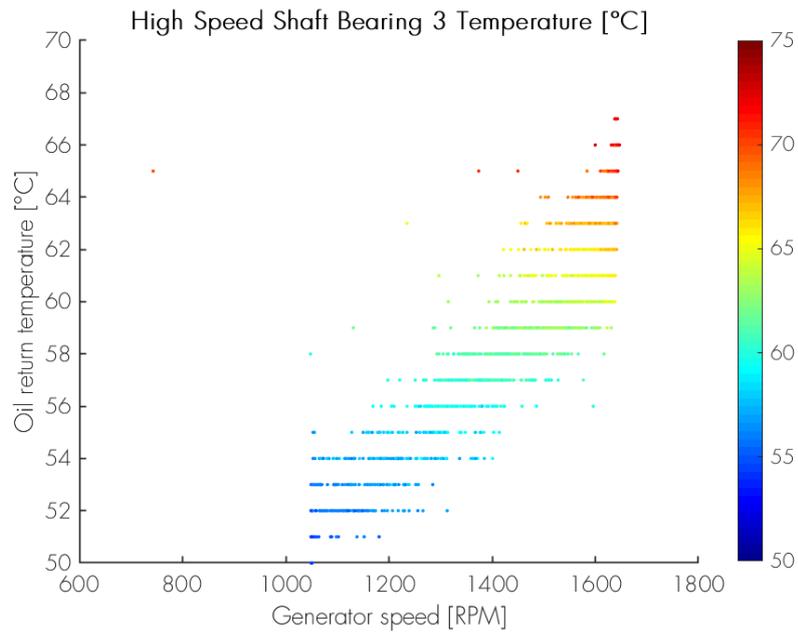


Figure G.29: Turbine A out-of-sample dataset, High Speed shaft bearing 3 in relation to the generator speed and return oil temperature

H Decision criteria for decision trees

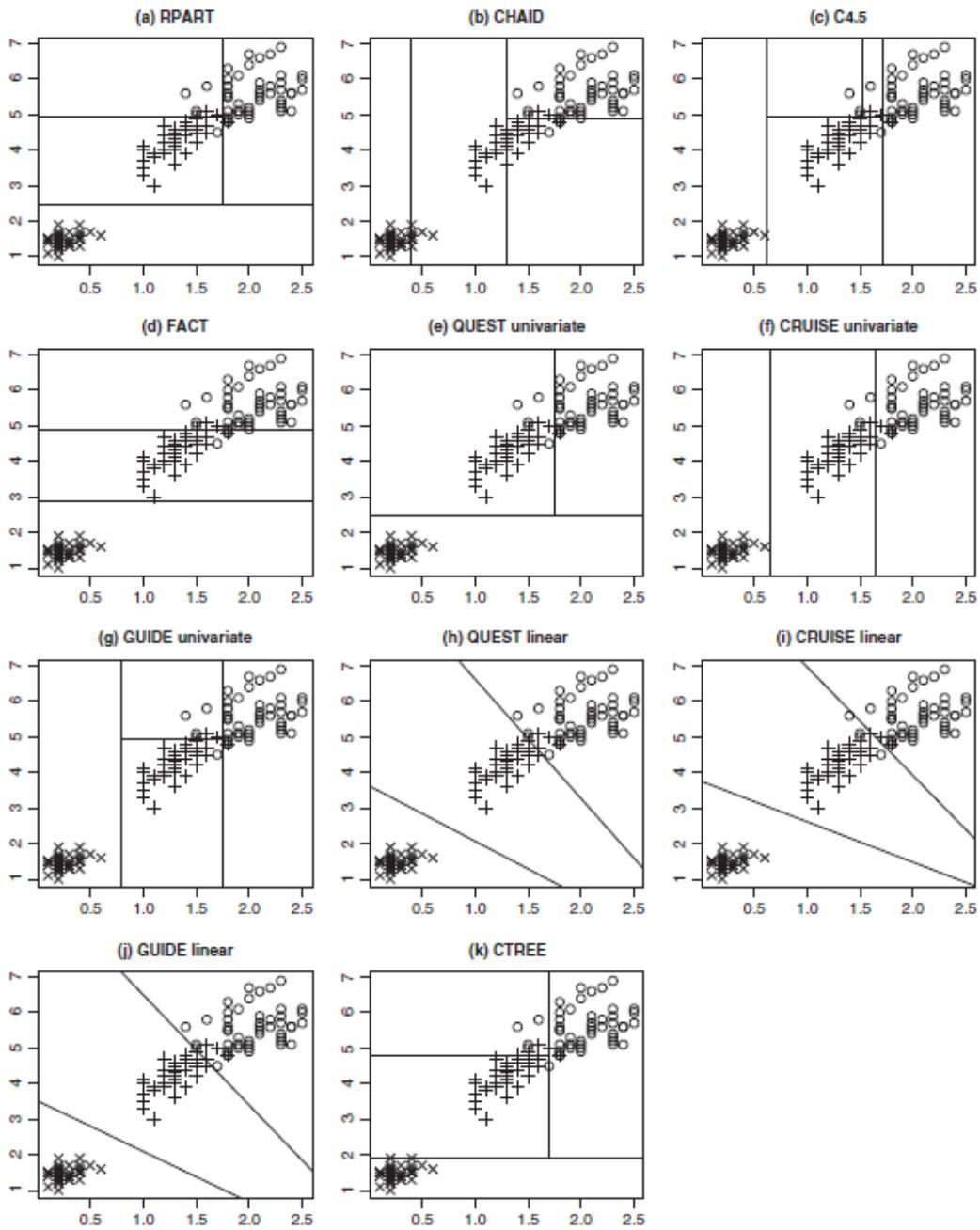


Figure H.30: Data splitting criteria for decision trees [40]

I Time series results linear regression model

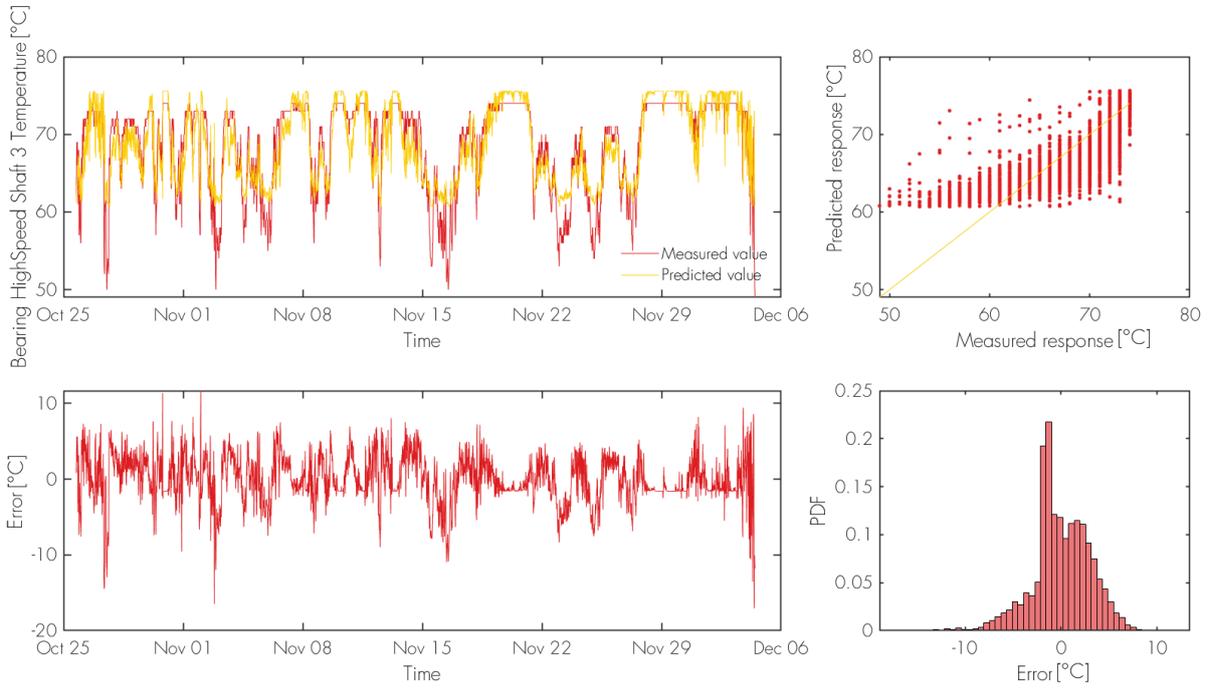


Figure I.31: In-sample time series for hollow shaft bearing 1, turbine A using linear regression model

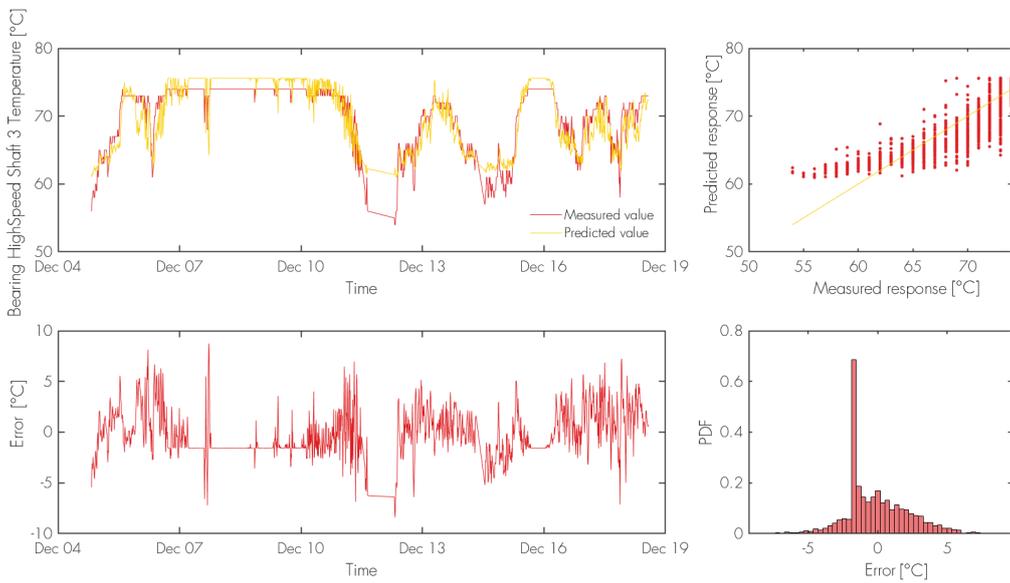


Figure I.32: Out-of-sample time series for hollow shaft bearing 1, turbine A using linear regression model

J Time series results 6th order polynomial regression model

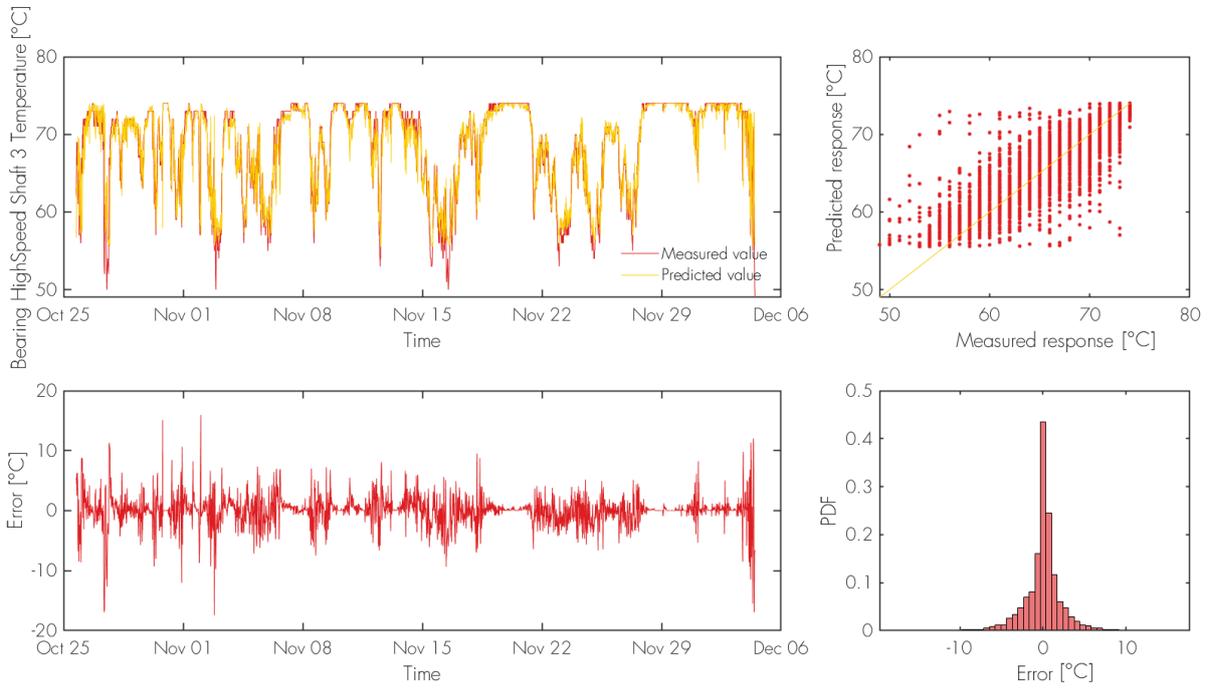


Figure J.33: In-sample time series for hollow shaft bearing 1, turbine A using linear regression model

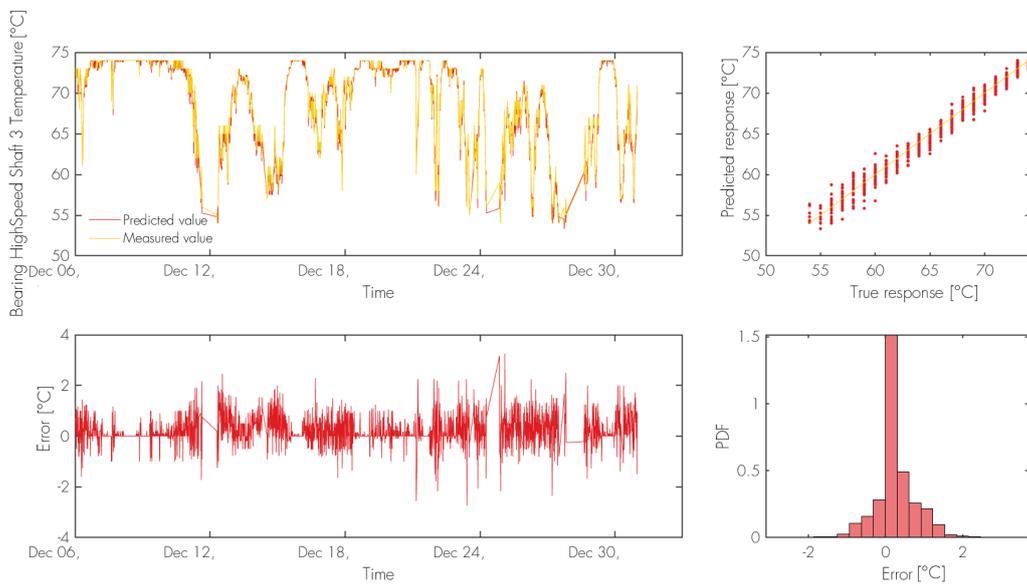


Figure J.34: Out-of-sample time series for hollow shaft bearing 1, turbine A using a 6th order polynomial regression model

K Time series results decision tree regression model

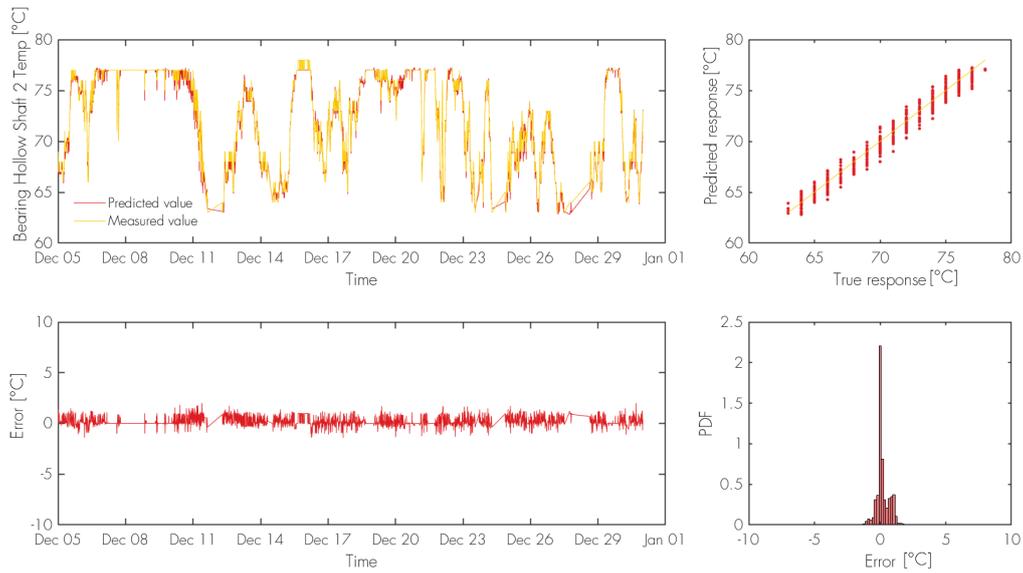


Figure K.35: Out-of-sample time series for hollow shaft bearing 1, turbine A using decision tree regression model

L Time series results decision tree classification model

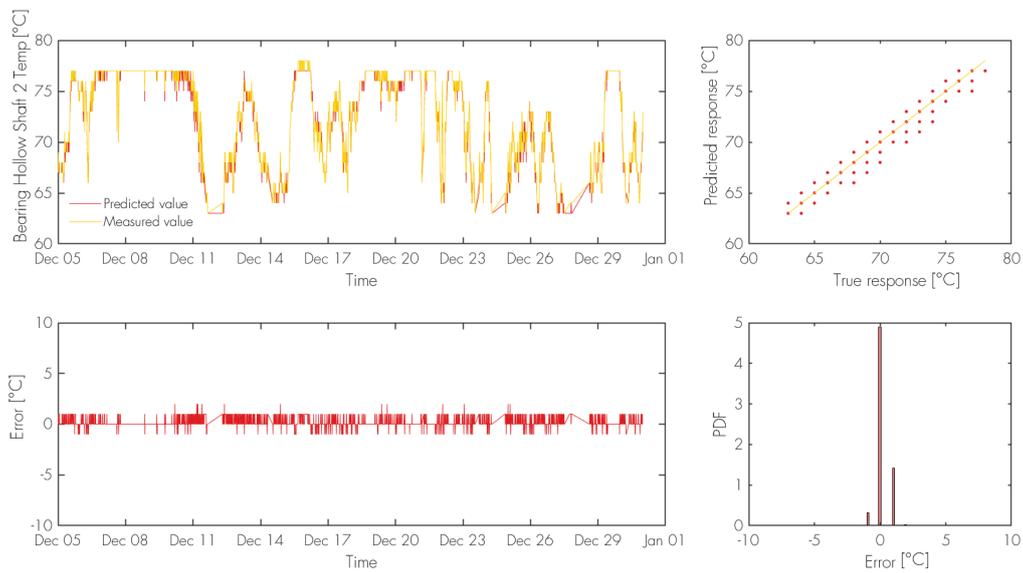


Figure L.36: Out-of-sample time series for hollow shaft bearing 1, turbine A using decision tree classification model

M Time series results SVM regression model

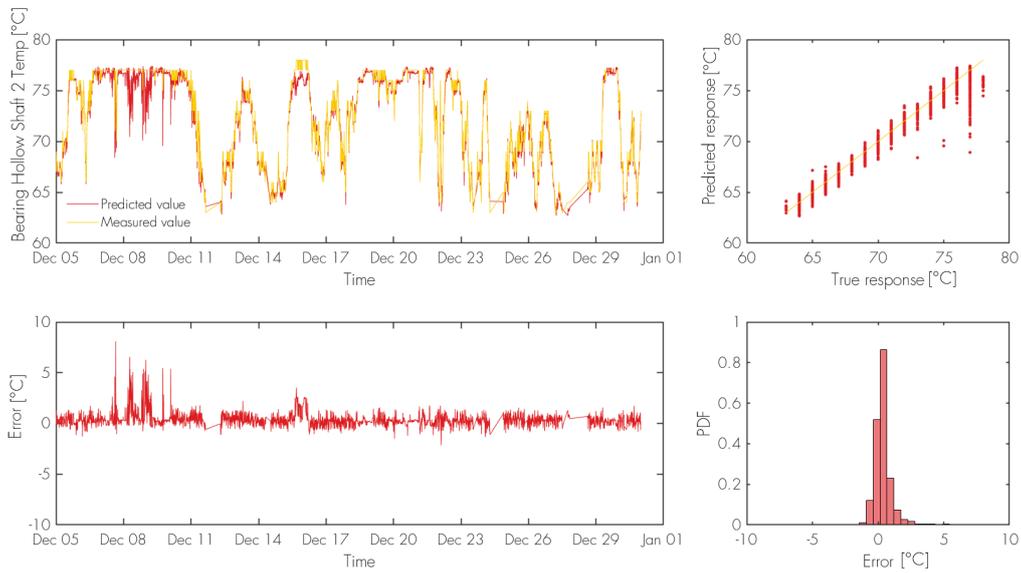


Figure M.37: Out-of-sample time series for hollow shaft bearing 1, turbine A using SVM regression model

N Time series results SVM classification model

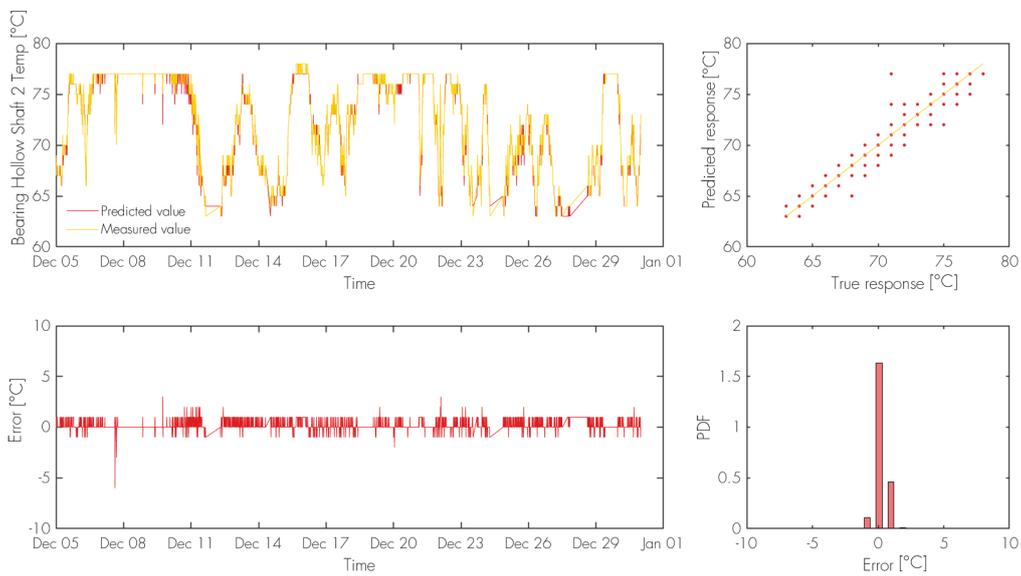


Figure N.38: Out-of-sample time series for hollow shaft bearing 1, turbine A using SVM classification model

O Results error distributions

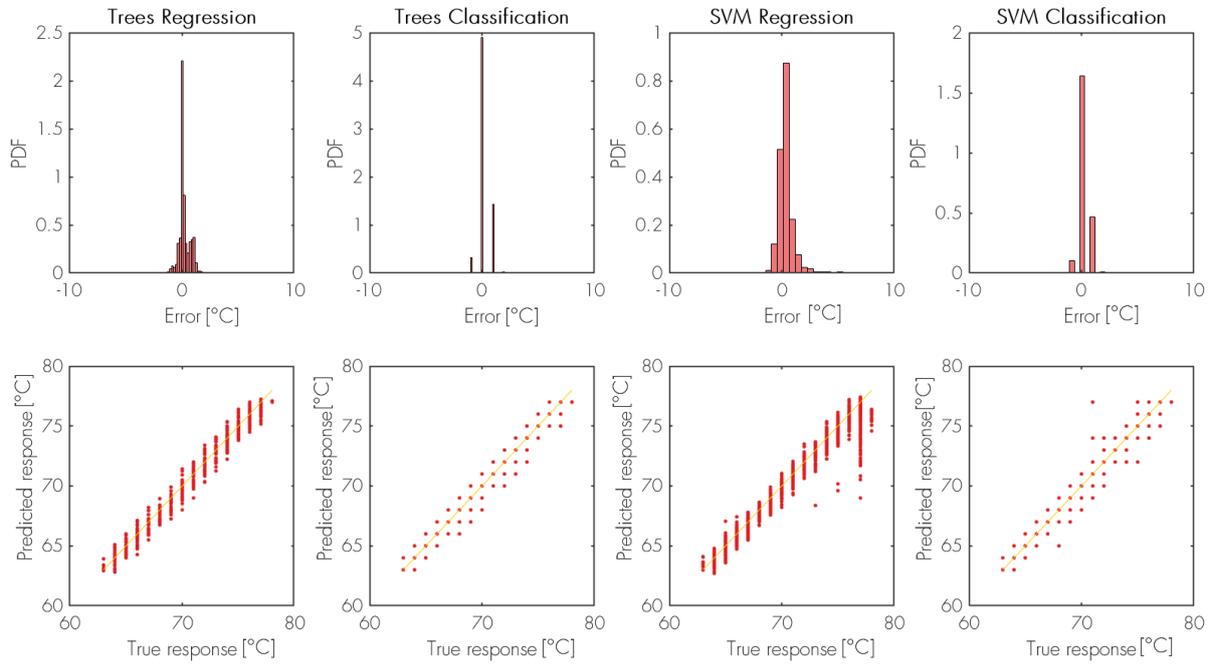


Figure O.39: Model test error distribution for Hollow Shaft Bearing 1 temperatures, turbine A

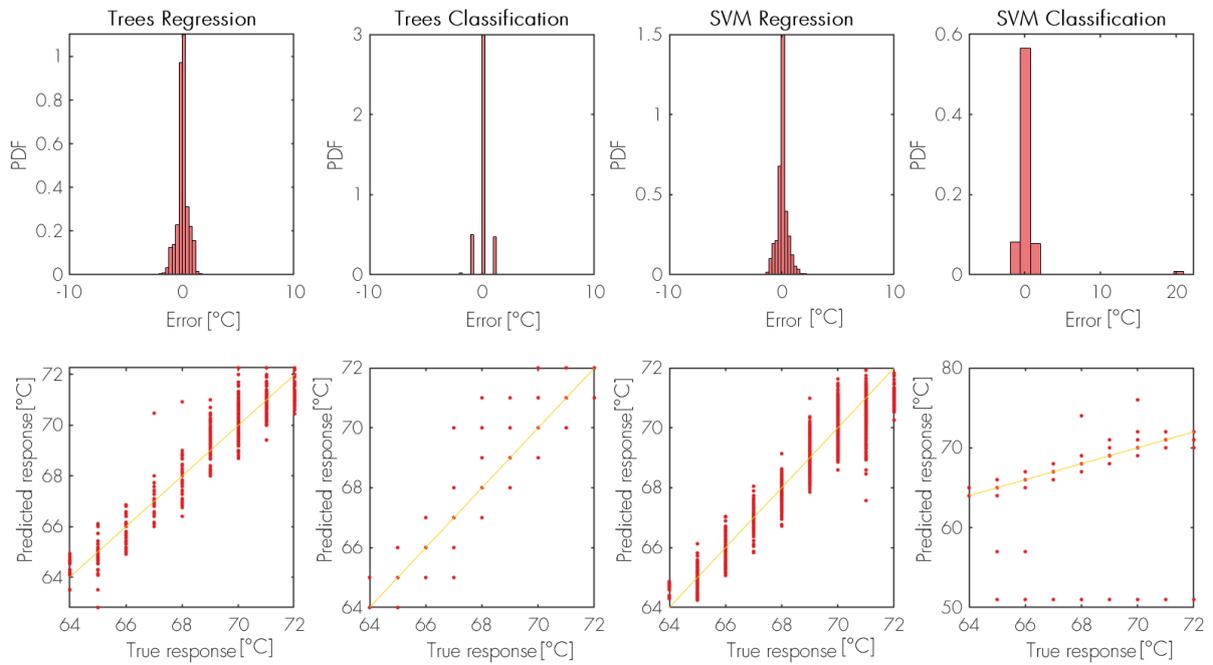


Figure O.40: Model test error distribution for Hollow Shaft Bearing 2 temperatures, turbine A

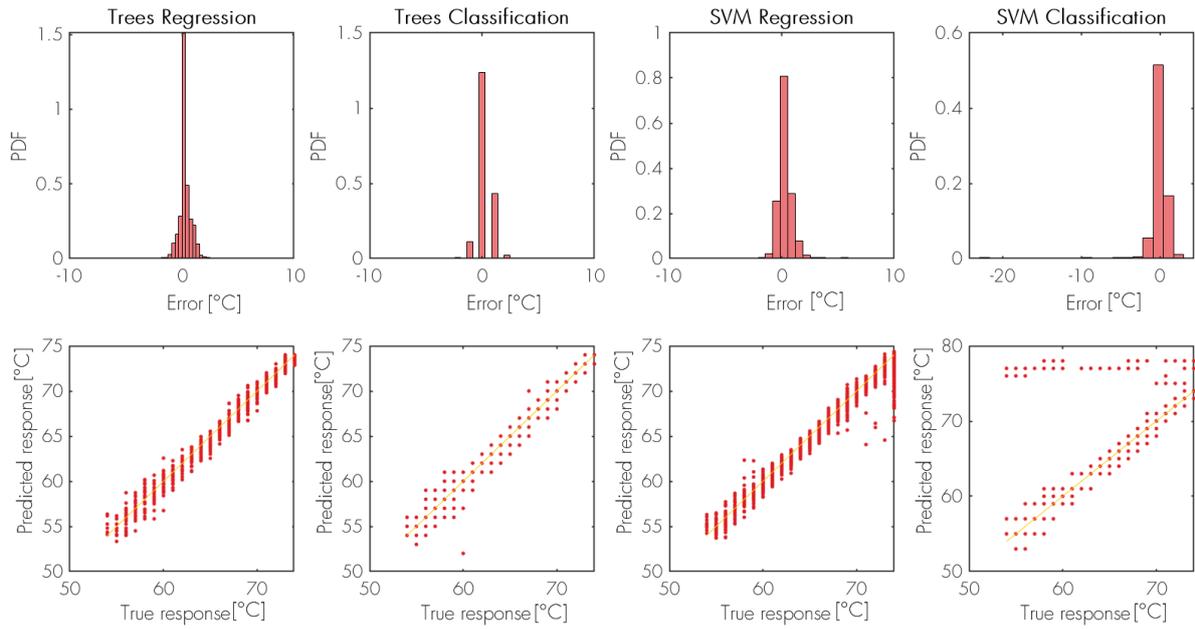


Figure O.41: Model test error distribution for High Speed Shaft Bearing 1 temperatures, turbine A

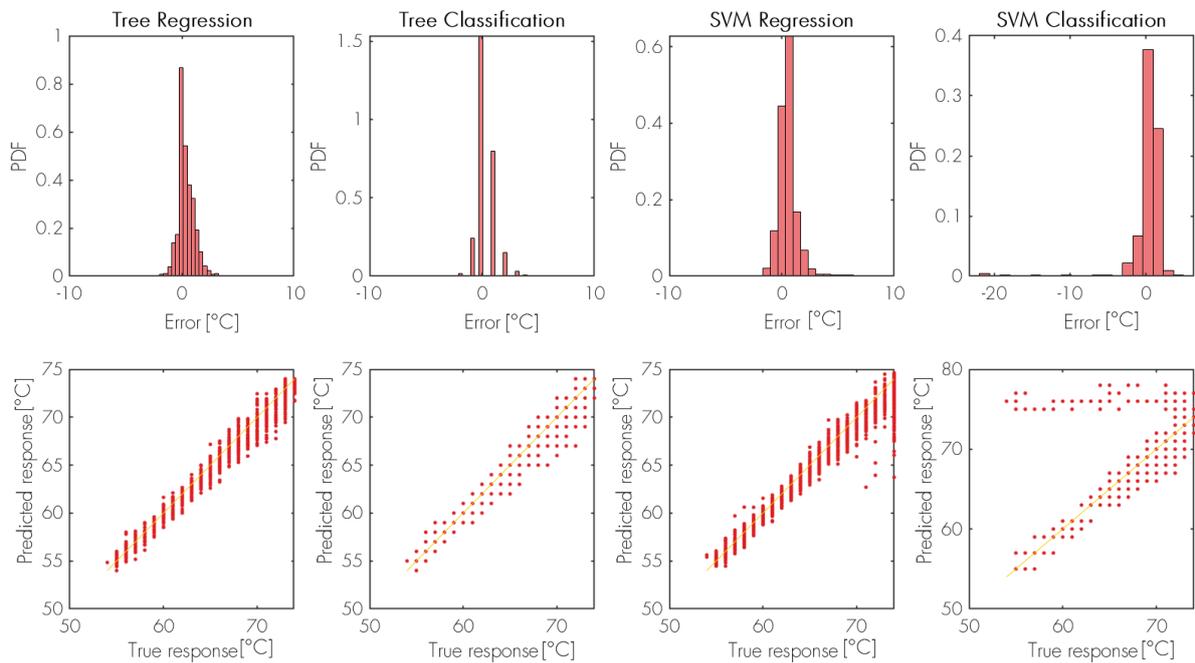


Figure O.42: Model test error distribution for High Speed Shaft Bearing 2 temperatures, turbine A

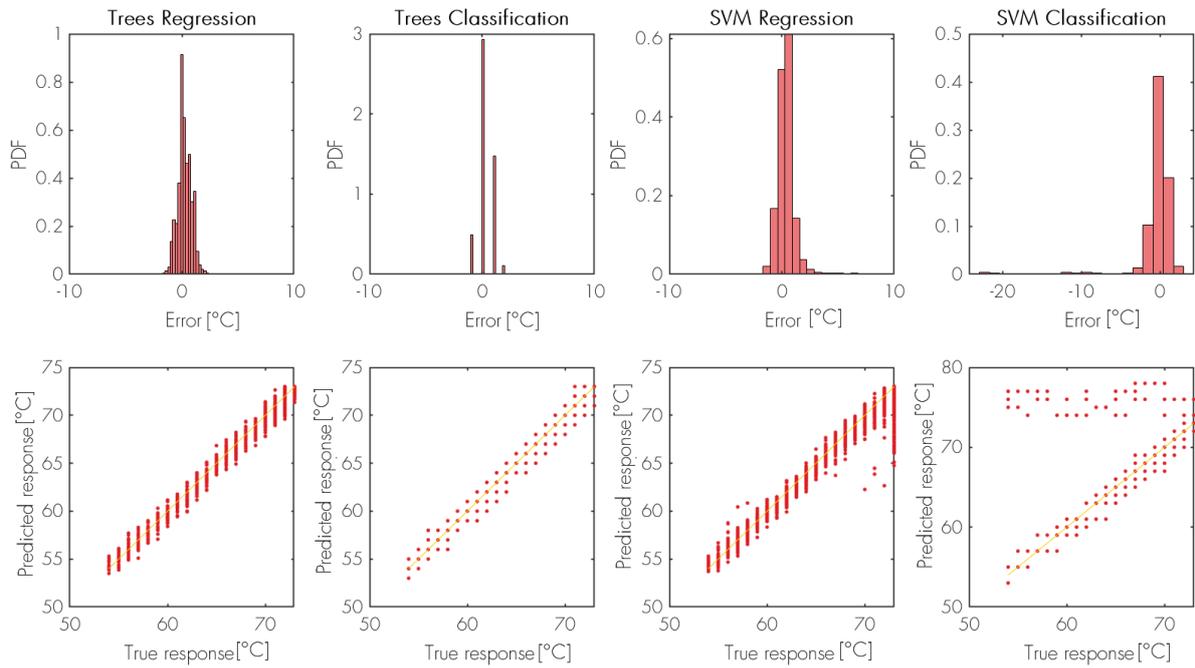


Figure O.43: Model test error distribution for High Speed Shaft Bearing 3 temperatures, turbine A

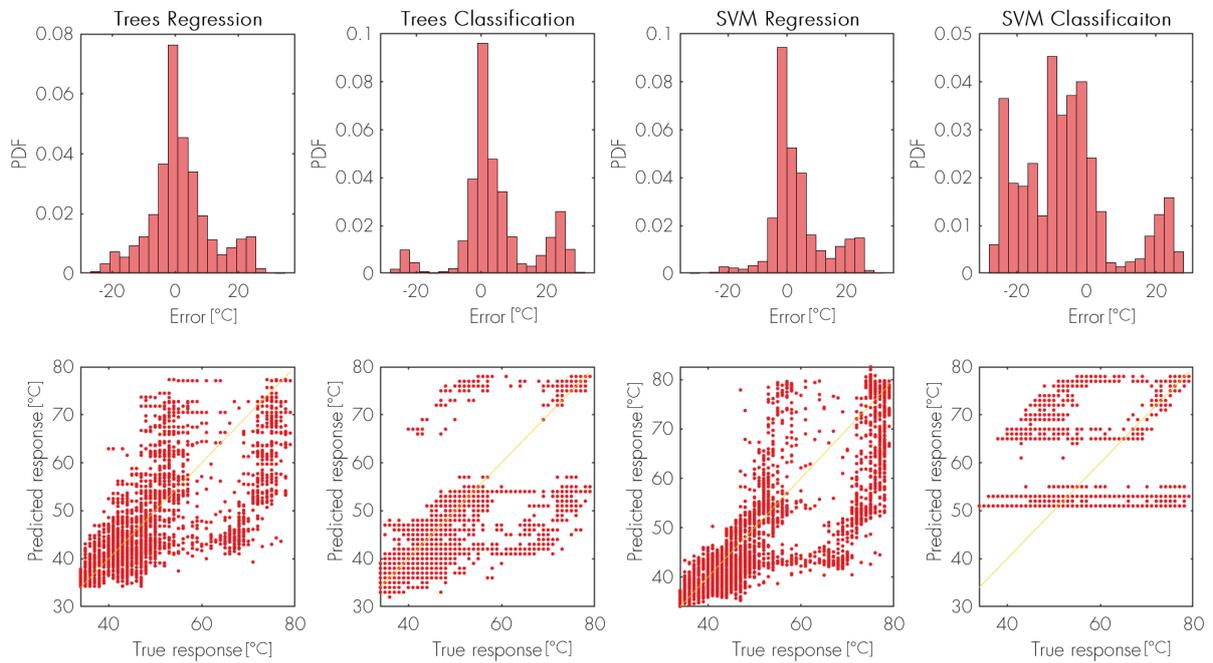


Figure O.44: Model test error distribution for generator bearing 1 temperatures, turbine A

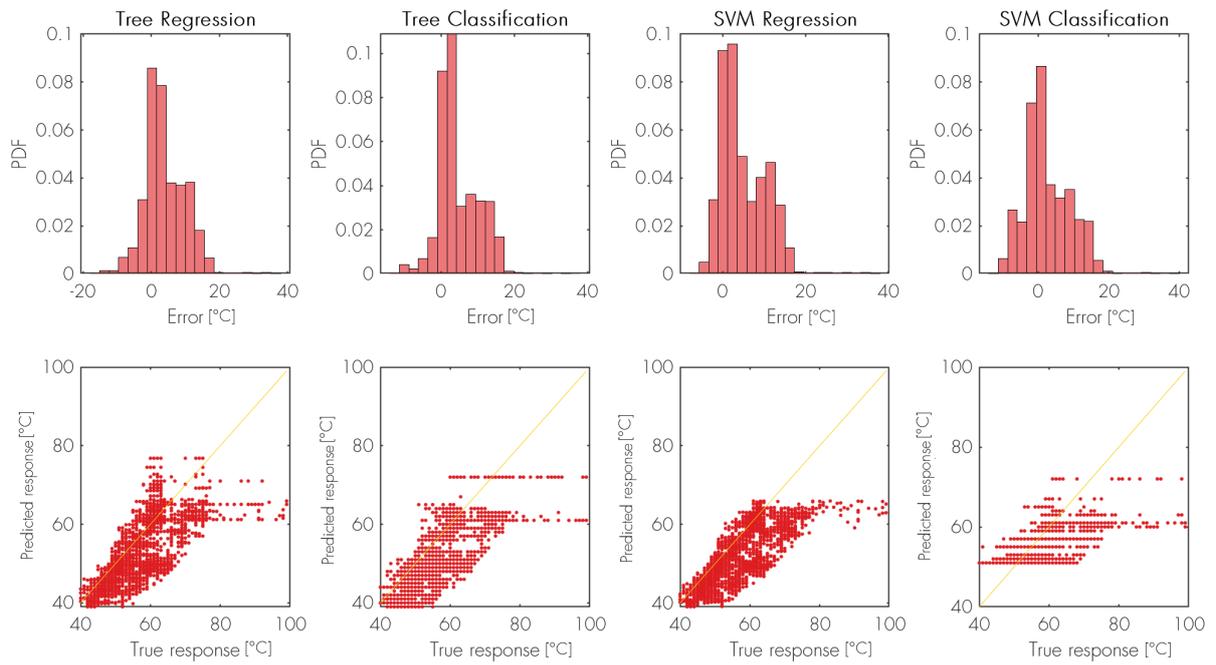


Figure O.45: Model test error distribution for generator bearing 2 temperatures, turbine A

P Cluster I_w/I_a statistical properties

Turbine = A

Response = Hollow shaft bearing 1 temperature

$\frac{I_w}{I_a} \downarrow$	Wind Speed [m/s]						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	62,6	66,0	70,3	73,5	75,2	76,4	76,5
1.1 - 2	64,4	66,8	71,5	73,8	75,3	76,2	76,2
2 - 3	63,2	66,6	70,5	73,6	74,6	76,0	76,1
3 - 4	62,8	66,1	69,2	72,8	73,4	75,3	76
4 - 6	62,9	66,25	69,75	73,0	74,7	75,7	76
6 - 8	64,45	67,41	70,7	72,9	75,4	74,8	NaN
8 - 12	63,2	66,6	70,8	71,9	68,5	NaN	NaN

Table P.1: Mean bearing temperature of clusters

$\frac{I_w}{I_a} \downarrow$	Wind Speed [m/s]						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	3,54	2,71	2,63	1,94	1,48	0,961	0,613
1.1 - 2	4,18	3,17	2,61	1,63	1,41	0,816	0,620
2 - 3	4,04	3,36	2,55	1,68	1,67	0,741	0,620
3 - 4	3,54	2,72	2,59	2,46	3,32	1,78	0
4 - 6	2,39	2,76	3,06	2,15	1,11	0,404	0
6 - 8	2,38	3,42	2,66	2,04	0,894	1,07	NaN
8 - 12	2,66	3,99	2,22	3,15	1,41	NaN	NaN

Table P.2: Number of datapoints per clusters

$\frac{I_w}{I_a} \downarrow$	Wind Speed [m/s]						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	372	755	773	516	311	181	173
1.1 - 2	30	55	79	139	142	150	240
2 - 3	41	80	175	152	82	45	12
3 - 4	86	172	166	96	47	10	1
4 - 6	50	89	82	71	24	6	3
6 - 8	4	15	37	36	5	4	0
8 - 12	39	52	70	22	2	0	0

Table P.3: Number of datapoints per clusters

Turbine = A
 Response = High speed shaft bearing 1 temperature

$\frac{I_w}{I_a} \downarrow$	Wind Speed [m/s]						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	56,0	61,4	67,7	71,1	72,4	73,4	73,4
1.1 - 2	58,2	62,3	69,1	71,4	72,5	73,2	73,2
2 - 3	56,9	62,1	67,9	71,2	71,9	73,0	73,0
3 - 4	56,2	61,4	66,2	70,3	70,4	72,5	73
4 - 6	56,0	61,6	66,8	70,6	72,1	72,9	73
6 - 8	58,1	63,1	68,5	70,7	72,6	72,4	NaN
8 - 12	56,3	62,9	68,2	69,4	64,3	NaN	NaN

Table P.4: Mean bearing temperature of clusters

$\frac{I_w}{I_a} \downarrow$	Wind Speed [m/s]						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	3,61	4,29	3,51	1,91	1,27	0,802	0,540
1.1 - 2	5,72	5,14	3,09	1,47	1,12	0,666	0,536
2 - 3	4,18	5,12	3,19	1,58	1,56	0,538	0,334
3 - 4	3,70	4,23	3,65	2,80	4,51	1,53	0
4 - 6	3,15	4,53	4,04	2,40	0,845	0,267	0
6 - 8	5,12	5,14	3,06	2,30	0,55	0,643	NaN
8 - 12	3,14	4,40	2,70	4,37	2,00	NaN	NaN

Table P.5: Number of datapoints per clusters

$\frac{I_w}{I_a} \downarrow$	Wind Speed [m/s]						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	372	755	773	516	311	181	173
1.1 - 2	30	55	79	139	142	150	240
2 - 3	41	80	175	152	82	45	12
3 - 4	86	172	166	96	47	10	1
4 - 6	50	89	82	71	24	6	3
6 - 8	4	15	37	36	5	4	0
8 - 12	39	52	70	22	2	0	0

Table P.6: Number of datapoints per clusters

Turbine = G
 Response = Hollow shaft bearing 1 temperature

$\frac{I_w}{I_a} \downarrow$	Wind Speed [m/s]						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	61,6	63,6	66,4	68,8	70,2	70,9	71,9
1.1 - 2	61,8	63,6	66,8	69,2	70,6	71,6	72,4
2 - 3	62,0	63,6	66,8	69,2	70,8	71,6	72,5
3 - 4	61,5	63,8	66,3	69,	70,6	71,7	72,3
4 - 6	61,5	63,9	66,2	69,2	70,5	71,1	68,7
6 - 8	62,4	63,3	67,0	68,6	70,4	70,7	71,1
8 - 12	62,8	63,6	66,8	69,1	69,5	70,6	71,2

Table P.7: Mean bearing temperature of clusters

$\frac{I_w}{I_a} \downarrow$	Wind Speed [m/s]						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	1,86	2,73	2,40	1,97	1,02	1,67	1,56
1.1 - 2	1,31	1,96	2,07	1,34	1,03	1,38	2,03
2 - 3	1,27	2,22	2,26	1,61	1,33	1,30	1,01
3 - 4	2,42	1,92	2,04	1,96	1,73	1,34	0,445
4 - 6	1,30	2,24	2,18	1,61	1,65	0,875	8,20
6 - 8	1,30	2,31	3,52	3,09	1,01	1,32	2,08
8 - 12	1,75	2,53	1,81	0,973	1,76	1,76	0

Table P.8: Standard deviations bearing temperature of clusters

$\frac{I_w}{I_a} \downarrow$	Wind Speed [m/s]						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	252	393	368	251	216	152	266
1.1 - 2	30	117	148	165	131	146	216
2 - 3	64	136	141	91	60	51	30
3 - 4	47	126	84	56	28	15	8
4 - 6	31	50	47	57	36	19	6
6 - 8	15	30	27	28	27	11	4
8 - 12	25	84	90	79	18	7	1

Table P.9: Number of datapoints per clusters

Turbine = G
 Response = High speed shaft bearing 1 temperature

$\frac{I_w}{I_a} \downarrow$	Wind Speed [m/s]						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	55,1	58,5	62,9	67,1	69,8	71,4	73,3
1.1 - 2	55,3	58,4	63,7	67,8	70,5	72,4	73,7
2 - 3	55,6	58,5	63,8	67,7	70,5	72,3	73,7
3 - 4	55,1	58,7	62,9	67,4	70,3	72,1	73,9
4 - 6	54,9	58,8	62,8	67,8	70,2	71,7	69,5
6 - 8	56,3	58,1	64,5	67,3	70,1	70,8	72,5
8 - 12	57,1	58,5	63,6	67,4	68,5	70,5	72

Table P.10: Mean bearing temperature of clusters

$\frac{I_w}{I_a} \downarrow$	Wind Speed [m/s]						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	2,30	3,90	3,74	3,38	1,83	2,49	1,90
1.1 - 2	2,25	3,29	3,43	2,29	1,83	1,64	2,14
2 - 3	2,09	3,26	3,21	2,79	2,17	2,18	1,41
3 - 4	2,20	2,97	3,37	3,14	2,86	1,89	0,58
4 - 6	2,26	3,65	3,62	2,64	2,76	1,54	8,69
6 - 8	2,36	3,12	4,33	3,29	1,96	2,33	2,35
8 - 12	2,86	3,35	3,00	1,59	2,94	2,90	0

Table P.11: Standard deviations bearing temperature of clusters

$\frac{I_w}{I_a} \downarrow$	Wind Speed [m/s]						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	252	393	368	251	216	152	266
1.1 - 2	30	117	148	165	131	146	216
2 - 3	64	136	141	91	60	51	30
3 - 4	47	126	84	56	28	15	8
4 - 6	31	50	47	57	36	19	6
6 - 8	15	30	27	28	27	11	4
8 - 12	25	84	90	79	18	7	1

Table P.12: Number of datapoints per clusters

Turbine = L
 Response = High speed shaft bearing 2 temperature

IwIa	Wind Speed						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	55,2	57,9	62,7	65,8	68,1	69,9	71,5
1.1 - 2	54,9	58,7	63,1	66,8	68,1	69,7	71,7
2 - 3	55,7	58,3	62,4	65,3	67,6	69,3	71,1
3 - 4	55,9	57,8	62,1	65,9	68,1	69,6	69,8
4 - 6	55,4	58,8	62,1	65,3	67,3	68,6	69,8
6 - 8	56,2	58,6	62,4	64,7	66,7	70,3	70,1
8 - 12	56,2	58,9	61,1	65,1	67,5	NaN	NaN

Table P.13: Mean bearing temperature of clusters

IwIa	Wind Speed						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	2,3	2,9	2,8	2,5	2,4	1,9	1,9
1.1 - 2	3,4	4,7	2,9	2,0	1,7	2,4	1,5
2 - 3	2,3	3,4	2,9	2,4	2,4	3,3	3,0
3 - 4	2,9	3,5	3,0	2,0	1,4	1,2	4,0
4 - 6	3,0	3,5	3,0	2,2	2,1	1,5	2,5
6 - 8	2,5	2,9	2,8	2,7	4,1	0,9	0
8 - 12	2,7	3,3	3,4	3,2	0,3	NaN	NaN

Table P.14: Standard deviations bearing temperature of clusters

IwIa	Wind Speed						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 1.1	238	441	504	481	260	215	270
1.1 - 2	37	46	108	114	131	120	334
2 - 3	88	163	194	118	84	47	55
3 - 4	47	62	95	78	70	26	19
4 - 6	50	117	167	163	74	24	2
6 - 8	40	64	82	45	8	3	1
8 - 12	21	42	57	16	3	0	0

Table P.15: Number of datapoints per clusters

Q Cluster Uw/Uinf statistical properties

Turbine = M

Response = High speed shaft bearing 3 temperature

IwIa	Wind Speed						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 0.002	55,1	60,4	66,0	69,4	70,5	71,1	71,4
0.002 - 0.03	56,3	61,9	66,8	69,3	70,5	71,5	71,8
0.03 -0.07	56,1	60,5	65,9	69,8	70,3	70,8	72,7
0.07 - 0.15	55,1	59,9	65,8	69,0	70,5	71,5	72,7
0.15 - 0.3	56,1	60,4	65,7	69,4	72,3	72,5	NaN

Table Q.16: Mean bearing temperature of clusters

IwIa	Wind Speed						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 0.002	3,56	4,41	3,48	2,37	1,65	2,23	1,36
0.002 - 0.03	4,85	5,10	3,59	2,95	2,55	1,63	1,44
0.03 -0.07	3,63	4,43	4,38	1,89	2,32	3,47	1,03
0.07 - 0.15	3,32	4,38	4,31	2,45	1,29	2,49	0,800
0.15 - 0.3	4,40	4,22	4,95	3,91	1,25	0,289	NaN

Table Q.17: Standard deviations bearing temperature of clusters

IwIa	Wind Speed						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 0.002	292	584	532	288	131	94	122
0.002 - 0.03	36	80	126	158	220	202	317
0.03 -0.07	76	134	205	156	90	50	27
0.07 - 0.15	68	135	138	88	38	11	7
0.15 - 0.3	23	47	39	33	7	3	0

Table Q.18: Number of datapoints per clusters

Turbine = L
 Response = High speed shaft bearing 3 temperature

IwIa	Wind Speed						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 0.002	52,7	55,4	58,7	61,3	62,7	63,5	63,9
0.002 - 0.03	53,2	56,3	59,2	61,4	62,8	63,9	64,3
0.03 -0.07	53,1	55,3	58,8	61,6	62,5	63,4	65,2
0.07 - 0.15	52,6	55,0	58,6	61,1	62,6	64,0	65,6
0.15 - 0.3	53,2	55,3	58,7	61,5	64,4	64,3	NaN

Table Q.19: Mean bearing temperature of clusters

IwIa	Wind Speed						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 0.002	1,73	2,50	2,10	1,71	1,38	1,72	1,25
0.002 - 0.03	2,64	3,29	2,23	2,00	1,84	1,36	1,05
0.03 -0.07	1,77	2,36	2,59	1,52	1,50	2,63	0,938
0.07 - 0.15	1,51	2,31	2,71	1,71	1,22	1,83	1,13
0.15 - 0.3	2,31	2,29	2,74	2,49	1,52	0,833	NaN

Table Q.20: Standard deviations bearing temperature of clusters

IwIa	Wind Speed						
	3.5 - 5.5	5.5 - 7.5	7.5 - 9.5	9.5 - 11.5	11.5 - 13.5	13.5 - 15.5	15.5 - 25
0 - 0.002	292	584	532	288	131	94	122
0.002 - 0.03	36	80	126	158	220	202	317
0.03 -0.07	76	134	205	156	90	50	27
0.07 - 0.15	68	135	138	88	38	11	7
0.15 - 0.3	23	47	39	33	7	3	0

Table Q.21: Number of datapoints per clusters

The copyright of this document is vested in Shell Global Solutions, B.V. The Hague, The Netherlands. All rights reserved.

Neither the whole nor any part of this document may be reproduced, stored in any retrieval system or transmitted in any form or by any means (electronic, mechanical, reprographic, recording or otherwise) without the prior written consent of the copyright owner. Shell Global Solutions is a trading style used by a network of technology companies of the Shell Group.