

PREDICTION OF WIND TURBINE BLADE FATIGUE LOADS USING  
FEED-FORWARD NEURAL NETWORKS

Dissertation in partial fulfillment of the requirement for the degree of  
MASTER OF SCIENCE WITH A MAJOR IN WIND POWER  
PROJECT MANAGEMENT



UPPSALA  
UNIVERSITET

Uppsala University

Department of Earth Sciences, Campus Gotland

Mohammad Mehdi Mohammadi

2021

PREDICTION OF WIND TURBINE BLADE FATIGUE LOADS USING  
FEED-FORWARD NEURAL NETWORKS

Dissertation in partial fulfillment of the requirement for the degree of

MASTER OF SCIENCE WITH A MAJOR IN WIND POWER  
PROJECT MANAGEMENT

Uppsala University  
Department of Earth Sciences, Campus Gotland

Approved by: Mr. Henrik Asmuth, PhD Researcher

Supervisor: Mr. Henrik Asmuth, PhD Researcher

Examiner: Dr. Karl Nilsson, Post Doc Researcher

2021

## Abstract

In recent years, machine learning applications have gained great attention in the wind power industry. Among these, artificial neural networks have been utilized to predict the fatigue loads of wind turbine components such as rotor blades. However, the limited number of contributions and differences in the used databases give rise to several questions which this study has aimed to answer. Therefore, in this study, 5-min SCADA data from the Lillgrund wind farm has been used to train two feed-forward neural networks to predict the fatigue loads at the blade root in flapwise and edgewise directions in the shape of damage equivalent loads.

The contribution of different features to the model's performance is evaluated. In the absence of met mast measurements, mesoscale NEWA data are utilized to present the free flow condition. Also, the effect of wake condition on the model's accuracy is examined. Besides, the generalization ability of the model trained on data points from one or multiple turbines on other turbines within the farm is investigated. The results show that the best accuracy was achieved for a model with 34 features, 5 hidden layers with 100 neurons in each hidden layer for the flapwise direction. For the edgewise direction, the best model has 54 features, 6 hidden layers, and 125 neurons in each hidden layer.

For a model trained and tested on the same turbine, mean absolute percentage errors (MAPE) of 0.78% and 9.31% are achieved for the flapwise and edgewise directions, respectively. The seen difference is argued to be a result of not having enough data points throughout the range of edgewise moments. The use of NEWA data has been shown to improve the model's accuracy by 10% for MAPE values, relatively. Training the model under different wake conditions did not improve the model showing that the wake effects are captured through the input features to some extent. Generalization of the model trained on data points from one turbine resulted in poor results in the flapwise direction. It was shown that using data points from multiple turbines can improve the model's accuracy to predict loading on other turbines.

*Keywords: Machine Learning, Feed-Forward Artificial Neural Network, Wind Power, Damage Equivalent Loads, Fatigue, Wind Turbine Blades, NEWA Data, SCADA, Lillgrund Wind Farm*

## Acknowledgments

First, I would like to thank my supervisor Henrik Asmuth for his valuable assistance and guidance through this study. Without his help, this dissertation would not be materialized. I express my gratitude to Ingemar Carlen for providing feedback and insight into the topic. I appreciate the feedback offered by Prabhu Kankatala for improving the methodology used. My special thanks to all my lecturers in the Master's program in wind power project management.

I owe my deepest gratitude to my family for supporting my studies and me through all my life with all means possible. I would not be where I am if It were not for them.

## List of Abbreviations and Symbols

DEL: Damage Equivalent Load

ANN: Artificial Neural Network

GPU: Graphics Processing Unit

SCADA: Supervisory Control and Data Acquisition

RMSE: Root Mean Square Error

SGD: Stochastic Gradient Descent

NN: Neural Network

MLP: Multilayer Perceptron

FFNN: Feed-Forward Neural Network

ReLU: Rectified Linear Unit Function

STD: Standard Deviation

MAPE: Mean Absolute Percentage Error

NRMSE: Normalized Root Mean Square Error

PCE: Polynomial Chaos Expansion

MAE: Mean Absolute Error

FEM: Finite Element Method

PCA: Principal Component Analysis

NCA: Neighbourhood Component Analysis

NEWA: New European Wind Atlas

TKE: Turbulence Kinetic Energy

IDE: Integrated Development Environment

$Re$ : Reynold's number

$\rho$ : fluid density

$U$ : flow speed

$L$ : characteristic length

$\mu$ : dynamic viscosity of the fluid

$V_{rel}$ : relative inflow velocity

$p_N$ : normal force per length

$p_t$ : tangential force per length

$r$ : local radius of blade

$dr$ : incremental length

$R$ : rotor radius/Pearson's correlation coefficient

$M_{flapwise,root}$ : flapwise bending moment at blade root

$M_{edgewise,root}$ : flapwise bending moment at blade root

$S = \sigma$ : stress

$\varepsilon$ : strain

$L_{consumed}$ : consumed fatigue life

$L_{remaining}$ : remaining fatigue life

$M_{eq}$ : equivalent bending moment

$m$ : Wöhler number

$\alpha$ : learning rate

$D$ : rotor diameter

## Table of Contents

<b>Abstract.....</b>	<b>iii</b>
<b>Acknowledgments .....</b>	<b>iv</b>
<b>List of Abbreviations and Symbols.....</b>	<b>v</b>
<b>List of Tables .....</b>	<b>viii</b>
<b>List of Figures.....</b>	<b>ix</b>
<b>1 Introduction.....</b>	<b>1</b>
<b>2 Theory .....</b>	<b>3</b>
<b>2.1 Loadings on Wind Turbine Blades.....</b>	<b>3</b>
<b>2.2 Fatigue.....</b>	<b>5</b>
<b>2.3 Machine Learning.....</b>	<b>9</b>
<b>2.4 Artificial Neural Networks.....</b>	<b>10</b>
<b>3 Literature Review .....</b>	<b>15</b>
<b>3.1 Simulation-Based Neural Networks .....</b>	<b>15</b>
<b>3.2 SCADA-Based Neural Networks .....</b>	<b>16</b>
<b>3.3 Summary and Findings .....</b>	<b>18</b>
<b>4 Data and Methodology.....</b>	<b>20</b>
<b>4.1 Lillgrund Wind Farm .....</b>	<b>20</b>
<b>4.2 NEWA Data.....</b>	<b>30</b>
<b>4.3 Methodology .....</b>	<b>30</b>
<b>5 Results .....</b>	<b>34</b>
<b>5.1 Base Model's Performance.....</b>	<b>34</b>
<b>5.2 Sensitivity Analysis of Features for The Base Model.....</b>	<b>36</b>
<b>5.3 The Effect of Adding Features on The Model's Performance .....</b>	<b>37</b>
<b>5.4 Adjustment of The Number of Hidden Layers and Neurons in Each Layer.....</b>	<b>39</b>
<b>5.5 The Effect of Wake Condition on The Model's Performance.....</b>	<b>40</b>
<b>5.6 Evaluation of The Model Generalization Performance.....</b>	<b>41</b>
<b>6 Discussion.....</b>	<b>50</b>
<b>7 Conclusion .....</b>	<b>56</b>
<b>8 Outlook and Future Work .....</b>	<b>58</b>
<b>9 References.....</b>	<b>59</b>

## List of Tables

Table 1: The value of mean stress and range for closed cycles in the load history of Figure 4.....	8
Table 2: Lillgrund Offshore Windfarm Information.....	20
Table 3: The name of turbines with load measurements.....	22
Table 4: The name of Curtailed Turbines .....	22
Table 5: Used Parameters in the Available Database .....	23
Table 6: The summary of applied filters on data .....	31
Table 7: The performance indicators for ANN trained on B8 Turbine using 14 features .....	34
Table 8: Summary of sensitivity analysis of features on the model performance for <i>Meq flapwise blade 2</i> .....	36
Table 9: Summary of sensitivity analysis of features on the model performance for <i>Meq edgewise blade 2</i> .....	37
Table 10: Summary of effect of adding new features to the model on the ANN's performance for <i>Meq flapwise blade 2</i> .....	38
Table 11: Summary of effect of adding new features to the model on the ANN's performance for <i>Meq edgewise blade 2</i> .....	38
Table 12: The effect of the number of hidden layers on the ANN's performance for <i>Meq flapwise blade 2</i> .....	39
Table 13: The effect of the number of hidden layers on the ANN's performance for <i>Meq edgewise blade 2</i> .....	39
Table 14: The effect of the number of neurons on the ANN's performance for <i>Meq flapwise blade 2</i> .....	39
Table 15: The effect of the number of neurons on the ANN's performance for <i>Meq edgewise blade 2</i> .....	40
Table 16: The wake condition effect on ANN performance.....	40
Table 17: ANN's performance trained on B8 data and tested on C8 and B6: 54 Features, 6 hidden layers, 125 neurons in each layer for <i>Meq edgewise blade 2</i> .....	41
Table 18: ANN's performance trained on B8 data and tested on C8 and B6: 34 Features, 5 hidden layers, 100 neurons in each layer for <i>Meq flapwise blade 2</i> .....	41
Table 19: ANN's performance generalization through the farm for <i>Meq flapwise blade 2</i> .....	45
Table 20: ANN's performance generalization through the farm for <i>Meq edgewise blade 2</i> .....	47



## List of Figures

Figure 1: Aerodynamic forces on an airfoil in one section of the blade .....	4
Figure 2: A typical S-N curve for steel .....	6
Figure 3: Loading history of 10Hz measurement on a wind turbine blade .....	7
Figure 4: A sample load history along with its hysteresis response. [A]: the stress versus time, [B]: the stress versus strain .....	8
Figure 5: The architecture of an MLP neural network with FFNN configuration .....	11
Figure 6: Lillgrund Windfarm Layout: Turbines with load measurements are marked with red circles, Curtailed turbines are marked with blue circles .....	21
Figure 7: The <i>Meq</i> edgewise blade 2 [kN.m] values vs features values, mean power coloured [kW], redline is the <i>Meq</i> mean value for each bin of the studies features: A: <i>Meq</i> edgewise blade 2 vs mean windspeed [m/s], B: <i>Meq</i> edgewise blade 2 vs mean RPM, C: <i>Meq</i> edgewise blade 2 vs STD windspeed [m/s], D: <i>Meq</i> edgewise blade 2 vs STD RPM , E: <i>Meq</i> edgewise blade 2 vs STD power [kW], F: <i>Meq</i> edgewise blade 2 vs STD pitch angle2 [degree], G: <i>Meq</i> edgewise blade 2 vs STD fore-aft nacelle acceleration [ $ms^2$ ], H: <i>Meq</i> edgewise blade 2 vs STD side-side nacelle acceleration [ $ms^2$ ] .....	24
Figure 8: The <i>Meq</i> Flapwise blade 2 [kN.m] values vs features values, mean power coloured [kW], redline is the <i>Meq</i> mean value for each bin of the studies features: A: <i>Meq</i> Flapwise blade 2 vs mean windspeed [m/s], B: <i>Meq</i> Flapwise blade 2 vs mean RPM, C: <i>Meq</i> Flapwise blade 2 vs STD windspeed [m/s], D: <i>Meq</i> Flapwise blade 2 vs STD RPM , E: <i>Meq</i> Flapwise blade 2 vs STD power [kW], F: <i>Meq</i> Flapwise blade 2 vs STD pitch angle2 [degree], G: <i>Meq</i> Flapwise blade 2 vs STD fore-aft nacelle acceleration [ $ms^2$ ], H: <i>Meq</i> Flapwise blade 2 vs STD side-side nacelle acceleration [ $ms^2$ ] .....	25
Figure 9: The distribution of incoming wind direction for turbine B8, each bin includes the degrees within $\pm 15^\circ$ .....	26
Figure 10: The wind rose for turbine B8 .....	26
Figure 11: The value of <i>Meq</i> edgewise blade 2 [kN.m] vs mean wind speed [m/s], mean power coloured [kW]: A: for incoming wind direction of $221^\circ \pm 5^\circ$ B: for incoming wind direction of $41^\circ \pm 20^\circ$ .....	27
Figure 12: The value of <i>Meq</i> edgewise blade 2 [kN.m] vs STD wind speed [m/s], mean power colored [kW]: A: for incoming wind direction of $221^\circ \pm 5^\circ$ B: for incoming wind direction of $41^\circ \pm 20^\circ$ .....	27
Figure 13: The value of <i>Meq</i> flapwise blade 2 [kN.m] vs mean wind speed [m/s], mean power colored [kW]: A: for incoming wind direction of $221^\circ \pm 5^\circ$ B: for incoming wind direction of $41^\circ \pm 20^\circ$ .....	28
Figure 14: The value of <i>Meq</i> flapwise blade 2 [kN.m] vs STD wind speed [m/s], mean power colored [kW]: A: for incoming wind direction of $221^\circ \pm 5^\circ$ B: for incoming wind direction of $41^\circ \pm 20^\circ$ .....	28
Figure 15: The effect of the power curve on the value of <i>Meq</i> edgewise blade 2 [kN.m] vs mean wind speed [m/s], mean power coloured [kW], lines are the <i>Meq</i> mean value for each bin of the mean wind speed: A: Turbine D7 with power curve A, B: Turbine D7 with power curve C, C: Turbine D7 with power curve A and B, the black line is turbine D7 with power curve A, redline is turbine D7 with power	

curve C, D: Turbine B6 under the effect of turbine B7 with power curve A, E: Turbine B6 under the effect of turbine B7 with power curve B, F: Turbine B6 under the effect of turbine B7 with power curve A and B, the black line is turbine B6 under the wake of turbine B7 with power curve A, the red line is turbine B6 under the wake of turbine B7 with power curve B Figure 16: The effect of the power curve on the value of  $Meq$  Flapwise blade 2 [kN.m] vs mean wind speed [m/s], mean power coloured [kW], lines are the  $Meq$  mean value for each bin of the mean wind speed: A: Turbine D7 with power curve A, B: Turbine D7 with power curve C, C: Turbine D7 with power curve A and B, the black line is turbine D7 with power curve A, redline is turbine D7 with power curve C, D: Turbine B6 under the effect of turbine B7 with power curve A, E: Turbine B6 under the effect of turbine B7 with power curve B, F: Turbine B6 under the effect of turbine B7 with power curve A and B, the black line is turbine B6 under the wake of turbine B7 with power curve A, the red line is turbine B6 under the wake of turbine B7 with power curve B ..... 29

Figure 17: The structure of the reorganized database ..... 31

Figure 18: The summary of the used methodology ..... 33

Figure 19: Normalized predictions vs Normalized true values for the  $Meq$  edgewise blade 2, redline linearly fitted to data ..... 34

Figure 20: MAPE distribution for  $Meq$  edgewise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure ..... 34

Figure 21: Normalized predictions vs Normalized true values for the  $Meq$  flapwise blade 2, redline linearly fitted to data ..... 35

Figure 22: MAPE distribution for  $Meq$  flapwise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure ..... 35

Figure 23: Convergence for the trained ANN vs Number of Epochs for  $Meq$  edgewise blade 2 ..... 35

Figure 24: Convergence for the trained ANN vs Number of Epochs for  $Meq$  flapwise blade 2 ..... 36

Figure 25: Normalized Predictions vs Normalized True values for ANN trained on B8 and tested on B6 for  $Meq$  edgewise blade 2, redline linearly fitted to data ..... 42

Figure 26: Percentage of Total vs MAPE for ANN trained on B8 and tested on B6 for  $Meq$  edgewise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure ..... 42

Figure 27: Normalized Predictions vs Normalized True values for ANN trained on B8 and tested on C8 for  $Meq$  edgewise blade 2, redline linearly fitted to data ..... 42

Figure 28: Percentage of Total vs MAPE for ANN trained on B8 and tested on C8 for  $Meq$  edgewise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure ..... 43

Figure 29: Normalized Predictions vs Normalized True values for ANN trained on B8 and tested on C8 for  $Meq$  flapwise blade 2, redline linearly fitted to data ..... 43

Figure 30: Percentage of Total vs MAPE for ANN trained on B8 and tested on C8 for <i>Meq flapwise blade 2</i> , redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure.....	43
Figure 31: Normalized Predictions vs Normalized True values for ANN trained on B8 and tested on B6 for <i>Meq flapwise blade 2</i> , redline linearly fitted to data.....	44
Figure 32: Percentage of Total vs MAPE for ANN trained on B8 and tested on B6 for <i>Meq flapwise blade 2</i> , redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure.....	44
Figure 33: Normalized predictions vs normalized true values for ANN trained on B7, B8, C8, D7, and D8 tested on B6, <i>Meq flapwise blade 2</i> , redline linearly fitted to data.....	46
Figure 34: Percentage of total vs MAPE for ANN trained on B7, B8, C8, D7 and D8 tested on B6, <i>Meq flapwise blade 2</i> , redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure.....	46
Figure 35: Normalized predictions vs normalized true values for ANN trained on B6, B7, B8, C8, D7, and D8 tested on C8, <i>Meq flapwise blade 2</i> , redline linearly fitted to data .....	46
Figure 36: Percentage of total vs MAPE for ANN trained on B6, B7, B8, C8, D7, and D8 tested on C8, <i>Meq flapwise blade 2</i> , redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure.....	47
Figure 37: Normalized predictions vs normalized true values for ANN trained on B6, B7, B8, C8, D7, and D8 tested on C8, <i>Meq Edgewise blade 2</i> , redline linearly fitted to data.....	47
Figure 38: Percentage of total vs MAPE for ANN trained on B6, B7, B8, C8, D7, and D8 tested on C8, <i>Meq Edgewise blade 2</i> , redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure.....	48
Figure 39: Normalized predictions vs normalized true values for ANN trained on B7, B8, C8, D7, and D8 tested on B6, <i>Meq Edgewise blade 2</i> , redline linearly fitted to data.....	48
Figure 40: Percentage of total vs MAPE for ANN trained on B7, B8, C8, D7, and D8 tested on B6, <i>Meq Edgewise blade 2</i> , redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure.....	48
Figure 41: Distribution of <i>Meq Edgewise blade 2</i> for turbine B8: <i>Meq Edgewise blade 2</i> [kN.m] vs Percentage of Total .....	51
Figure 42: MAPE value in each bin vs number of datapoints in each bin% for <i>Meq edgewise blade 2</i> .....	51
Figure 43: Distribution of <i>Meq flapwise blade 2</i> for turbine B8: <i>Meq flapwise blade 2</i> [kN.m] vs Percentage of Total .....	52
Figure 44: MAPE value in each bin vs number of datapoints in each bin% for <i>Meq flapwise blade 2</i> .....	52
Figure 45: The distribution of <i>Meq flapwise blade 2 kN.m</i> for B6 turbine .....	54

Figure 46: The distribution of $Meq$ flapwise blade 2 kN.m for C8 turbine .....	54
---	----

# 1 Introduction

Machine Learning (ML) is a term defined as finding a mathematical algorithm based on a set of data to produce the desired output. This topic has drawn a lot of attention in recent years due to increasing computational power, the introduction of cloud-based solutions, and improvements in data collection. It has benefited a wide variety of sectors including health care, transportation, finance, and so on. Within machine learning, artificial neural network (ANN) has gained great importance due to its ability to solve problems difficult to humans where the relationship between inputs and outputs are complex and non-linear (Burkov, 2019).

Availability of data from different sources such as supervisory control and data acquisition system (SCADA) has enabled the wind power industry to utilize machine learning in various ways. Among others it has been used for fault detection and condition monitoring, forecasting power generation and wind speed, optimization of control systems, and wind turbine/wind farm modeling (Lydia & Kumar, 2020) (Stetco, et al., 2019). Another application of ML that has been studied within the field of research is load predictions for wind turbines using ANNs.

Wind turbines undergo different load conditions throughout their lifetime. One of the major aspects in determining the lifetime of wind turbine components is fatigue considerations (Sutherland, 2000). It refers to the fact that wind turbine components under cyclic loadings with magnitudes less than their yield strength will fail if they have been subjected to enough cycles. Two fatigue critical components in wind turbines are rotor blades and tower base. Therefore, several attempts have been made in the literature to predict the fatigue loads in the shape of damage equivalent loads (DEL) from SCADA or simulation-based data.

For instance, in two separate studies by Zhou et al., (2018) and Vera-Tudela & Kühn, (2017), DEL values for turbine blades in flapwise and edgewise directions have been predicted using 10-min SCADA data by ANNs. Although there are similarities between the methodologies, they have resulted in different results regarding the used features, the model structure, pre-processing steps, and achieved accuracies. Since the number of studies is limited, it is not possible to rely on the conclusions as they might be a result of using different databases with different conditions.

Therefore, there is a need for further contribution to the body of research to provide additional examples of how ANNs can be used for this purpose and the shortcomings and unanswered questions of previous studies can be addressed.

In light of this, the objective of this study is to answer the following questions:

- How accurate can ANNs predict the DEL values for wind turbine blades in flapwise and edgewise direction using the available database?

- What is the contribution of different input features on the accuracy of the model?
- Can mesoscale data be used as an alternative when met mast data is unavailable and to what extent do they contribute to the model's accuracy?
- How do different wake conditions affect the model's accuracy?
- How accurate can a model trained on data from one or multiple wind turbines predict the fatigue loads on other turbines within the farm?

The results of the current study along with other examples in the literature can provide a cheap and fast alternative compared to strain gauge measurements used for load measurements to evaluate the remaining lifetime of wind turbine components within wind farms. This will provide windfarm owners and asset managers with information about how they can take different actions to reduce operational costs. For instance, knowing the remaining lifetime of a component and how it is affected by environmental factors, it is possible to curtail the turbine thereby avoiding or delaying major component failures, and plan the suitable predictive action or replacement.

Moreover, knowing the remaining lifetime of components allows them to make a rational decision regarding the end of life of turbines, for instance, whether they can be used on another turbine as a spare part. The presented approach can also be incorporated into wind turbine control systems. This way, the damage will be treated as one of the decisive factors to adjust the turbine pitch angle and rotor speed. For instance, it is conceivable that in a situation where the electricity price is low, running the turbine at full load condition is not profitable.

Chapter 2 lays out the theoretical dimensions of the research required to understand the current study including loadings on wind turbines, fatigue, machine learning, and artificial neural networks. Chapter 3 reviews the literature related to the topic and summarizes the findings. In Chapter 4, the data used in this study including SCADA data and load measurements from Lillgrund windfarm, and NEWA data are described.

It is followed by explaining the methodology used to develop the models including feature selection and determining the model parameters. Chapter 5 presents the results from various developed models. In Chapter 6, obtained results are discussed and analyzed. Moreover, sources of uncertainty are explained. The findings of the study are summarized in chapter 7. Besides, limitations of the current work are identified and suggestions for future studies are presented in chapter 8.

## 2 Theory

In this part, the theoretical aspects required for following the content of this study are provided. The chapter is divided into four sections. First, different loadings on wind turbine blades are explained. It is followed by introducing how fatigue calculations are performed. Next, the concept of machine learning is touched upon. In the end, the fundamentals of artificial neural networks are presented.

### 2.1 Loadings on Wind Turbine Blades

The wind turbine rotor is responsible for harnessing the energy from the coming wind. This process poses loadings on the turbine components including the rotor blades. The loading refers to forces and moments on different elements resulted from different sources. A turbine needs to withstand these loadings during its lifetime. Therefore, it is necessary to understand the origin of them and how it affects the turbine structural integrity.

Forces can be categorized based on their source into three categories. It includes inertial, aerodynamic, and structural forces (Eggleston & Stoddard, 1987). For offshore applications, hydrodynamic forces also exist. However, their effect is minimal on the rotor (Kuhn, et al., 2012). Inertial forces are caused as the blades have mass and motion and are influenced by gravity. It includes forces such as centrifugal force, actuation due to control system, Coriolis as an effect of dynamic coupling, mass imbalance, gyroscopic due to turbine yawing, and gravity itself (Eggleston & Stoddard, 1987).

The aerodynamic forces are a result of blades' interaction with the surrounding wind field. A wind turbine blade is normally divided into a number of sections radially. Each section is then defined by one single airfoil. The interaction of the relative inflow velocity with the airfoil creates a resultant force acting on the aerodynamic center. The magnitude and direction of this force are dependent on the magnitude and direction of the relative inflow velocity, airfoil shape, pitch angle, and airfoil angle of attack which is the angle between relative inflow velocity and airfoil chord line.

This resultant force can be divided into a tangential and normal component relative to the rotor plane where the normal component contributes to the thrust force and the tangential force results in the torque driving the blades. Also, it is common to divide the resultant force into drag and lift where drag is parallel, and lift is perpendicular to the relative inflow velocity. It is shown in Figure 1. The aerodynamic resultant force therefore the lift and drag forces result from the pressure distribution on the airfoil due to its shape when subjected to the airflow. Two important factors in determining the magnitude of lift and drag forces are the angle of attack and Reynold's number. Reynold's number can be defined as the ratio of inertial to viscous forces and can be written as follows.

$$Re = \frac{\rho UL}{\mu} \quad (1)$$

In which  $\rho$  is the fluid density,  $U$  is the flow speed,  $L$  is the characteristic length, and  $\mu$  is the dynamic viscosity of the fluid. In general, the lift force increases with an increasing angle of attack until it reaches the stall angle where the flow separates on the airfoil. The drag force also increases with increasing angle of attack and increases sharply when flow separation occurs. Regarding Reynold's number, in general, the lift force increases with increasing Reynold's number. In contrast, the drag force increases with decreasing Reynold's number. For conventional horizontal axis wind turbine blades, the blade is designed so the lift force is dominant compared to the drag. Therefore, they are referred to as lift-based wind turbines.

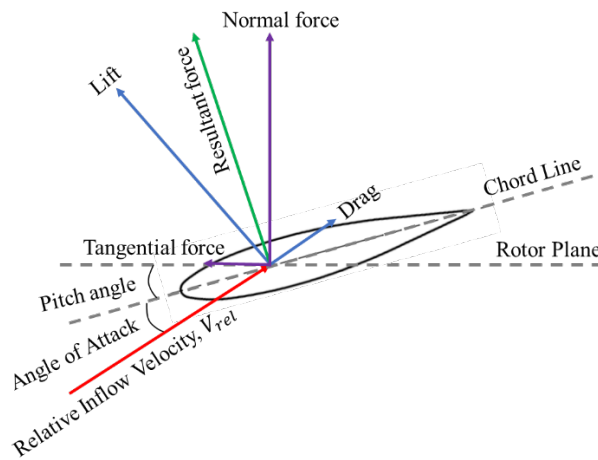


Figure 1: Aerodynamic forces on an airfoil in one section of the blade

The relative inflow velocity seen by each section of the blade varies temporally and spatially. For instance, the change in time can be a result of turbulence, gust, diurnal changes, and weather fronts. The spatial change is among others a result of blade passage in front of the tower structure, wind shear, and wind veer. The latter two are due to the rotor placement in the boundary layer where wind direction and speed change with height (Söker, 2013).

Structural forces result from the dynamic response and deformation of blades under loading. Resonance condition is an example of this force. Turbine blades are not rigid components meaning they deform when subjected to forces and moments. Moreover, loadings create a dynamic response on the blades that need to be considered to increase the accuracy of the analysis. For turbines with a hinge-less hub, the blades can be considered as cantilever beams. Three main dynamic responses exist as follows. The first one is flapping that refers to the blade's back and forth movement in and out of the rotor plane. The second one is lead-lag defined as the relative motion of the blade to its rotational speed, and the third one is torsion referring to the blade's dynamic rotation along its pitch axis (Manwell, et al., 2010).



These mentioned loadings can be further categorized by their duration into steady, transient, cyclic, and stochastic (Manwell, et al., 2010). Steady loadings are those that do not vary in a relatively long time, for instance, the mean thrust force on the blades. Transient loadings are those that occur only during certain events, for instance, the start and stop of the wind turbine. The cyclic loads vary periodically, for instance, the gravity, tower passage, or wind shear effects on blades. The last is stochastic loads that have a somewhat random behavior. Wake effects and atmospheric turbulence are examples of this type of loading and have a significant effect on the loading condition on turbine blades' aerodynamic loading (Garrad, 1990).

Considering the contribution of these forces on turbine blades, the loading condition can be characterized although it is not a trivial task. It is common to evaluate the loadings on blades at the blade root using two bending moments. They are flapwise bending moment referring to the bending moment forcing the blade out of the rotor plane and edgewise bending moment referring to the bending moment acting parallel to the plane of rotation. These two can be calculated as follows in which  $p_N$ ,  $p_t$ ,  $r$ ,  $dr$ , and  $R$  are normal and tangential forces per length, the local radius, incremental length, and rotor radius, respectively (Hansen, 2010).

$$M_{flapwise,root} = \int_0^R r p_N(r) dr \quad (2)$$

$$M_{edgewise,root} = \int_0^R r p_t(r) dr \quad (3)$$

The dominant contributions from different sources for each of these moments depend on the turbine conditions such as the presence of mass imbalance, blades with different pitch angles, and rotor geometry including coning and tilt angle, along with operating and environmental conditions such as turbulence intensity and wake conditions (Eggleston & Stoddard, 1987). However, it can be said that for edgewise moment, the effect of cyclic gravitational loading is significant while for flapwise moment the aerodynamic forces are major contributors (Hansen, 2010).

## 2.2 Fatigue

Turbine blades experience cyclic and fluctuating loading due to rotation, wind shear, and different operational and environmental conditions. Materials subjected to this loading situation are prone to fatigue. It means that if they undergo a certain number of cyclic or fluctuating loads with a certain amplitude much lower than yield strength, they rupture. This number among others depends on the stress amplitude and material properties. For a typical steel specimen under cyclic load with constant amplitude, a so-called  $S - N$  curve can be drawn using results from experiments. It is shown in Figure 2 where  $S$  is the amplitude of stresses and  $N$  is the number of cycles it takes before the specimen fails. With decreasing stress amplitude, the material can withstand a higher number of full cycles until it

reaches a value where it can experience an indefinite number of cycles without failure. This is called the endurance limit. (Beer, et al., 2015)

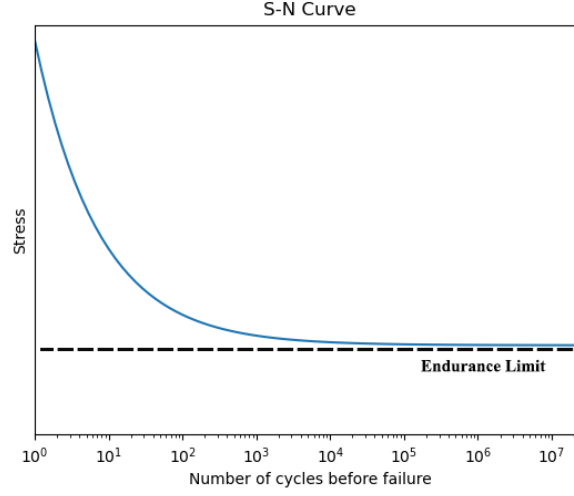


Figure 2: A typical S-N curve for steel

S-N curve is usually drawn for constant cyclic loadings with zero mean value. However, if the mean value is nonzero, the number of cycles before failure is different. If the mean value is in tensile range, the number of cycles decreases and if it is in compressive range, the number of cycles will increase. A Goodman diagram can be used to account for this effect. (Megson, 2017) Now assume that for the stress  $S$ , the specimen can withstand  $N$  number of cycles. Therefore, if it has already been through  $n$  number of cycles, its consumed life can be written as  $L_{consumed} = \frac{n}{N}$  and its remaining life as  $L_{remaining} = 1 - \frac{n}{N}$ .

In practice, the amplitudes in the loading history vary and can be even random due to a variety of reasons. Therefore, the S-N curve cannot be used directly. A solution to this problem was first proposed by Pålmgren in 1924 through what is called nowadays as Miner rule or linear cumulative damage hypothesis. It states for  $i$  different cyclic conditions with stress  $S_i$ , completed cycles of  $n_i$ , and number of cycles before failure of  $N_i$ , the consumed life can be written as:

$$L_{consumed} = 1 - \sum \frac{n_i}{N_i} \quad (4)$$

alternatively, the failure happens when:

$$L_{consumed} = \sum \frac{n_i}{N_i} = 1 \quad (5)$$

Although the Miner rule has its limitations and some discrepancies have been shown experimentally, it has been used widely due to its easy implementation (Schijve, 2009). The problem yet remains that measured loading histories rarely show an ideal cyclic behavior. An example is shown in Figure 3 from

10Hz load measurements from a wind turbine blade. As can be seen, extracting cycles and amplitudes is not a trivial task.

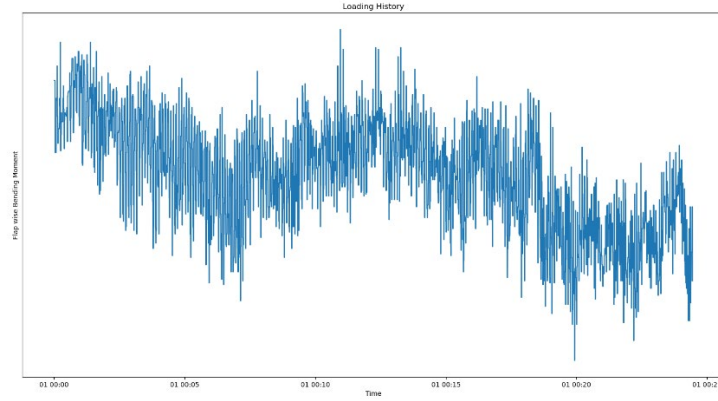


Figure 3: Loading history of 10Hz measurement on a wind turbine blade

One method to extract the range, number of cycles, and mean value from such loading history is Rainflow counting which its fundamentals are explained in Schijve, 2009. A hysteresis response diagram can be used to explain how this method works. It shows how a specimen is deformed when subjected to stress. For viscoelastic materials, when stress is applied, the occurred deformation path under loading is different than the deformation path when the applied stress is lifted. It is due to energy dissipation as a result of internal friction in the specimen.

To exemplify how the cycles, their range, and mean values for such a loading history are determined, consider the load history shown in Figure 4. If the hysteresis response corresponding to this load history is drawn, it can be seen that the deformation path of the specimen has created three closed loops namely ADG, BCD, and EFE. Each of these loops corresponds to a load cycle. Therefore, it is possible to treat them as a cyclic load. It means that similar to a cyclic load, the specimen is loaded to a maximum and unloaded to a minimum stress. After the cycles are determined, their range and mean values are calculated as follows. These values are summarized in Table 1.

$$Range = maximum - minimum \quad (6)$$

$$mean\ value = \frac{maximum + minimum}{2} \quad (7)$$

Rainflow counting method has been adapted for implementation in different programming languages where load history is the input and range, number of cycles, and mean value are outputs (MATLAB, 2021) (PyPi, 2021). Having these values, it is now possible to calculate the consumed or remaining life of the component using the S-N curve.

Table 1: The value of mean stress and range for closed cycles in the load history of Figure 4

Loop	Mean stress	Range
ADG	150	700
BCD	-150	100
EFE	75	250

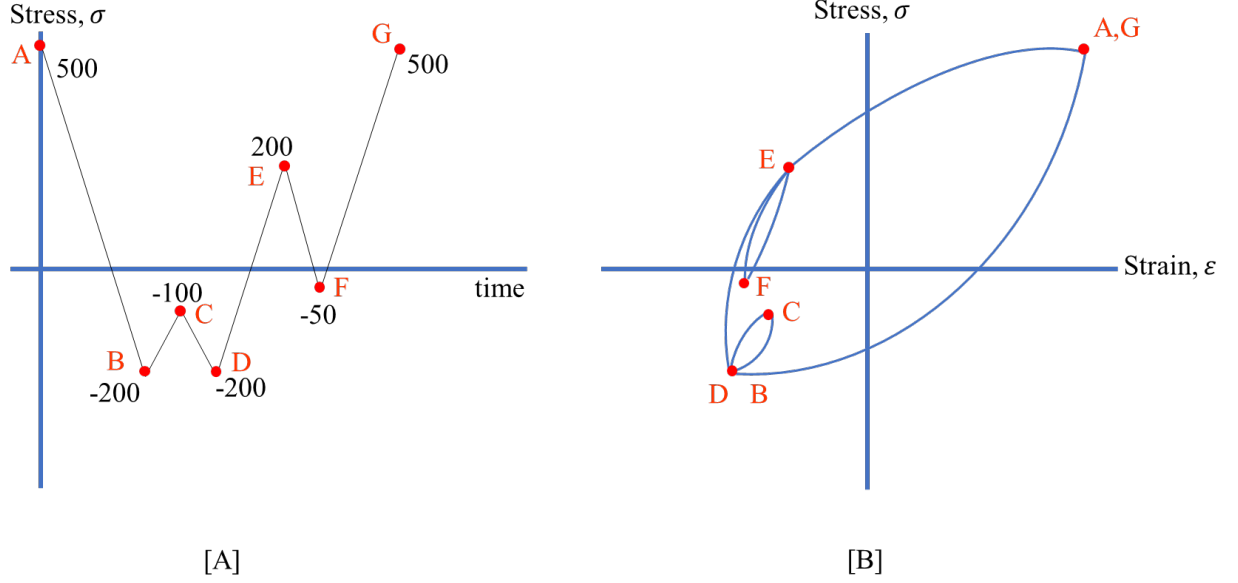


Figure 4: A sample load history along with its hysteresis response. [A]: the stress versus time, [B]: the stress versus strain

The obtained values can be used to calculate the so-called Damage Equivalent Load (DEL) which represents a constant amplitude cyclic load with a fixed frequency that would consume the same amount of component's fatigue life. Therefore, in our case since bending moments are of interest, it can be calculated as:

$$M_{eq} = \left( \frac{\sum_{i=1}^{n_b} n_i M_i^m}{N} \right)^{\frac{1}{m}} \quad (8)$$

In which  $M_{eq}$  is the equivalent bending moment in the interval for which Rainflow is implemented,  $n_b$  is the number of bins with different range and mean value,  $n_i$  is the number of cycles in each bin,  $M_i$  is the range of each bin after accounting for the mean value,  $m$  is Wöhler number and depends on the material, and  $N$  is the assumed number of cycles during the turbine's lifetime (Blasques & Natarajan, 2013).

## 2.3 Machine Learning

The term machine learning has been around since 1960. The term “Learning” is however different than the type of learning seen in humans and animals. Rather it refers to finding a mathematical model that can produce a series of desired outputs (Burkov, 2019). Machine learning can be categorized into three major groups as will be discussed as follows.

1. **Supervised Learning:** In this type of learning, data points are labeled, meaning there is an output associated with a set of features or inputs. For instance, assume that we have a housing database with information about each housing area, postal code, number of rooms, age, and the price of the house and it is desired to develop a model to predict housing price. In this example, area, postal code, number of rooms, and age are inputs or features and the house price is the label, output, or target value. In the model, the features and labels are handled in the shape of vectors where for a single number, the vector has only one component.

Supervised learning can be further categorized into two groups. The first one is a classification problem where labels are discrete values. Developing a model that can determine whether an email is a spam, or no spam is one example. Here, the output is either spam or no spam and no other value is possible for the output. The second one is the regression problem where the output is a real value similar to the example of the housing price.

2. **Unsupervised Learning:** In this type of learning, datapoints are not labeled and the objective is to create a model that can transform the vector of features into another vector. An example of this type of learning is outlier detection in which a real value showing how much the features in a datapoint are similar to other datapoints in the database, is calculated. Then, the data points with the lowest score can be considered as outliers.

3. **Reinforcement Learning:** In this type of learning, the machine exists in an environment as an agent where the state of the environment is stored as features in a vector. The agent is allowed to take actions and every action is followed by a reward and takes the environment into another state. The objective is to develop a model maximizing the reward in the long run. To clarify this, assume a simple videogame as the environment, a player as the agent, and capturing a flag with the least number of steps as the objective. The agent will be rewarded every time it gets closer to the flag. Therefore, the model will learn a policy to capture the flag as soon as possible.

Another term usually used is deep and shallow learning. Shallow models are those that learn the model directly from the features. For instance, fitting a line to a series of data points using the least square method is considered an example of shallow learning. In contrast, in deep models, the vector of features

can be transferred into another vector several times before the last vector is mapped into the output. This concept is explained further in the coming sections.

Here, our objective is to predict the damage equivalent moments from a set of environmental and operational features. Therefore, it falls under supervised regression learning. There are different algorithms developed to create a model to solve such a problem. Using the following criteria, we have selected to use artificial neural networks (ANN) for this purpose. The reasoning is as follows.

First, ANNs have proven to have an acceptable performance to handle a large number of data points and features. Second, they can capture and incorporate non-linear patterns and relationships between features and output. Third, they have been used in several similar studies and have resulted in acceptable results. However, there are some downsides to using ANN.

First, the algorithms used in ANNs are more complex compared to other machine learning methods where they are usually considered to be a black box between inputs and outputs. Therefore, understanding how the algorithm works will be more difficult, and explaining it to a non-technical audience such as project managers or policy makers without technical knowledge will be harder. Therefore, this may affect their willingness to utilize such models. Second, the computational time to develop an ANN can be considerably higher than other simpler regression models. However, nowadays, with the advances in cloud solutions, parallel computation, and using GPU processors, it is only a matter of cost rather than time.

Here, these downsides are of less concern to us and the advantages of using ANN and the given proof of concept in the literature presented in Literature Review, have led us to select them as the tool to solve the problem at hand. Previous studies have shown that ANNs can predict damage equivalent loads with acceptable accuracy from SCADA data as discussed in 3.2.

## 2.4 Artificial Neural Networks

To understand how ANN works, it is first necessary to explain how a regression model is built. The simplest form of a regression model is linear regression where a linear relationship is learned to predict the outputs from inputs. Assume a dataset with the shape of  $\{(\vec{x}_i, y_i)\}_{i=1}^N$  where  $\vec{x}_i$  is the vector of inputs with real values and  $y_i$  is a real number output for all data points from  $i$  to  $N$ . The linear regression model can be written as:

$$f_{w,b}(\vec{x}) = \vec{w}\vec{x} + b \quad (9)$$

Where the  $w, b$  index in  $f_{w,b}$  shows that they are the parameters in the function. The  $\vec{w}$  is a vector with the same dimension of  $\vec{x}$  and  $b$  is a real number. The goal is to find the best combination that predicts the output values most accurately. This best combination can be shown with  $(\vec{w}^*, b^*)$ . The accuracy can

be evaluated with a loss function such as root mean square error (RMSE) or any desired loss function. The RMSE can be calculated with the following expression.

$$\frac{1}{N} \sum_{i=1}^N (f_{w,b}(\vec{x}_i) - y_i)^2 \quad (10)$$

To increase the accuracy, it is needed to minimize the loss function. Mathematically, it can be achieved by differentiating the loss function and solving for zero. However, such an analytical solution is only available for some limited cases. Therefore, a more general approach called gradient descent is used for most machine learning applications. This method is iterative and, in each iteration, first, the gradients of the loss function is calculated with respect to the parameters using chain rule (for instance, in our example the parameters are  $\vec{w}, b$ ). Then, parameters are updated in the negative gradient direction using a learning rate. Each iteration through the whole training dataset is called an epoch and the procedure can be stopped when no significant change in the loss function value is seen.

Performing gradient descent on large datasets can be time-consuming. Therefore, improvements have been implemented to increase its speed. One of these improved algorithms is stochastic gradient descent (SGD) where gradients are approximated on the subsets or batches of the training data points. Adam optimizer is a computationally efficient implementation of SGD that is widely used in NN applications that requires little tuning and is suited for large datasets (Kingma & Ba, 2015).

Now turning to neural networks, the mathematical model to be learned can be written as:

$$y = f_{NN}(\vec{x}) = f_3(f_2(f_1(\vec{x}))) \quad (11)$$

In which  $f_{NN}$  is a nested function and  $f_{1,2,3}$  are vector functions. To better explain how the model works, it is required to consider a certain type of architecture and configuration for the neural network and explain some of its basics. There is a variety of architectures and configurations. Here, we consider the one used in this study. The architecture used is called multilayer perceptron (MLP) and the configuration is feed-forward neural network (FFNN). It is shown in Figure 5.

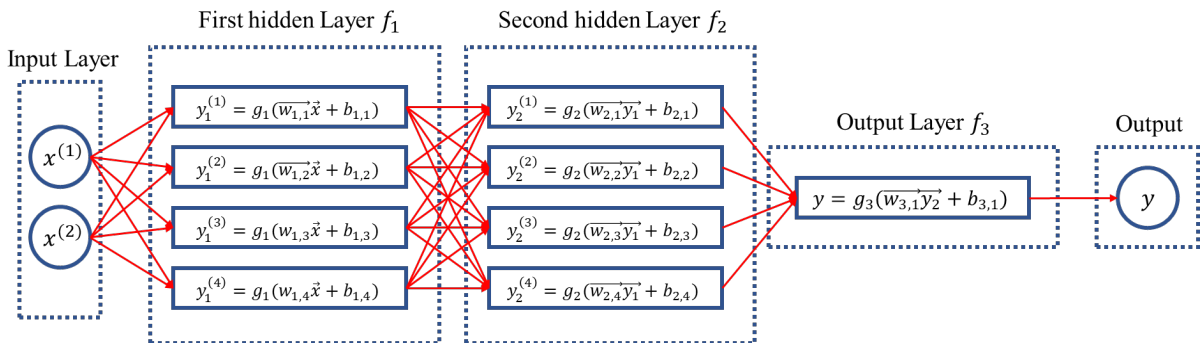


Figure 5: The architecture of an MLP neural network with FFNN configuration

It is assumed that there are two features as seen on the left-hand side in the input layer and there is an output on the right-hand side both being real values. Then, there are three layers in which layer 1 and layer 2 each have four units or neurons, and layer three has one unit. Layers 1 and 2 are commonly called hidden layers as for a non-technical observer the model can be seen as a black box between inputs and the output layer having the same number of neurons as the output number. Layer 3 is called the output layer. The arrows show where the input of each neuron comes from. Note that the value of each item in the previous layer is used as the input of all items in the next layer. This configuration is called fully connected.

Now going back to the mathematical expression of the model,  $f_l(\vec{z}) = g_l(\mathbf{W}_l \vec{z} + \vec{b}_l)$  in which  $l$  is the layer index,  $g_l$  is the activation function, and  $\mathbf{W}_l$  and  $\vec{b}_l$  are similar to  $\vec{w}, b$  in the linear regression example with the difference of  $\mathbf{W}_l$  is a matrix and  $\vec{b}_l$  is a vector that is obtained by gradient descent algorithm where each row of  $\mathbf{W}_l$  is a vector with the same dimension as  $\vec{z}$ . Besides,  $g_l$  is a fixed non-linear vector function selected by the user for layers 1 and 2.

One of the commonly used activation functions for applications such as the problem at hand is rectified linear unit function (ReLU) with the following behavior:

$$ReLU(z) = \begin{cases} 0 & z < 0 \\ z & z \geq 0 \end{cases} \quad (12)$$

The reason why a non-linear function is chosen is that it allows the capture of non-linear behaviors. The output of  $f_l(\vec{z})$  is in the shape of  $[g_l(a_{l,1}), g_l(a_{l,2}), \dots, g_l(a_{l, \text{size } l})]$  in which  $\text{size } l$  equals the number of neurons in layer  $l$  and  $a_{l,u} = \vec{w}_{l,u} \vec{z} + b_{l,u}$  where  $u$  stands for the specific unit or neuron. The last activation function for the last layer is just a linear regression for a regression problem.

In summary, the procedure of calculations is as follows. For each neuron, first, the inputs are gathered in the shape of a vector. This is  $[x^{(1)}, x^{(2)}]$  for the first layer showing the feature vector and  $[y_{l-1}^{(1)}, y_{l-1}^{(2)}, y_{l-1}^{(3)}, y_{l-1}^{(4)}]$  for each hidden layer. Then, a linear transformation is applied to them using the  $\vec{w}_{l,u}$  and  $b_{l,u}$ . Therefore, it can be written for each hidden layer as  $[\vec{w}_{l,1} \vec{z}, \vec{w}_{l,2} \vec{z}, \vec{w}_{l,3} \vec{z}, \vec{w}_{l,4} \vec{z}] \stackrel{\text{def}}{=} \vec{a}_l$  and  $\vec{a}_l + \vec{b}_l = \vec{c}_l$ . Afterward, the activation function is applied resulting in a real value or  $g_l(\vec{c}_l) = \vec{y}_l \stackrel{\text{def}}{=} [y_l^{(1)}, y_l^{(2)}, y_l^{(3)}, y_l^{(4)}]$ . This value is the value of the neuron and the input to all neurons in the next layer.

In each epoch, the gradient descent calculates the partial derivative of the loss function relative to  $\vec{w}, b$  values using the chain rule and update  $\vec{w}, b$  s based on the calculated partial derivatives and a learning rate shown by  $\alpha$ . The learning rate can be either fixed or adaptive where the latter uses the change in the values to adjust itself thereby speeding up the calculations. This process is called back propagation since the derivatives are calculated backward from output to the input.



Now that the model's structure is explained, the procedure of setting up a neural network is presented. Assuming there is a database with a series of different variables, first, it is required to select the features. This can be done based on mainly two approaches. The first approach is using logical reasoning. For instance, in our case, the physical understanding of the loading condition on a wind turbine allows us to select a set of features related to the desired output such as mean and STD of wind speed. Here, it is reasonable to use the experience gained by reviewing the literature within this field.

The second approach is trying different sets of features to check whichever results in better predictions. This can be done using a systematic approach. For instance, a set of base features selected using the first approach can be supplemented with more features step by step while the performance is evaluated. Therefore, the features contributing to the improvement of the model's performance are selected.

Next, the number of hidden layers, the number of neurons in each hidden layer, and the activation function should be determined. Here, a series of trials and errors combining with the suggestions from literature or experts can be used. Although the same activation function can be used for a series of different regression problems, the number of hidden layers and neurons varies from case to case and needs to be optimized for each problem and database with different features.

For the model to learn the optimized  $w, b$  values and evaluate its performance, the database is commonly split into three subsets. It includes a training subset where the model is learned on, a validation subset where the model performance is validated on in each epoch and is used to calculate the loss function to determine when the training should stop, and a test dataset being held aside where it is not seen previously by the model thereby allowing the model to be tested and its performance evaluated. The ratio used to split the database is determined by the user and depends on how large the dataset is. A common ratio is 80 percent for training and validation and 20 percent for test. 20 percent of the selected 80 percent is usually considered for validation.

The model's performance is obtained from comparing the prediction values versus measured values of the desired output. This comparison can be done in various ways. Here, we have chosen four indicators to evaluate the model performance based on the literature review and their relevance. The definition of these values may vary in different literature. Therefore, the definition used in this study are presented as follows (Scikit Learn, 2021) (SciPy, 2021):

$$\text{Mean Absolute Percentage Error} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_{i,prediction} - y_{i,measured}}{y_{i,measured}} \right| * 100 \quad (13)$$

*Normalized Root Mean Square Error =*

$$NRMSE = \frac{1}{\overline{y_{i,measured}}} \sum_{i=1}^N \sqrt{\left( \frac{(y_{i,prediction} - y_{i,measured})^2}{N} \right)} * 100 \quad (14)$$

*Coefficient of Determination =*

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_{i,measured} - y_{i,prediction})^2}{\sum_{i=1}^N (y_{i,measured} - \overline{y_{prediction}})^2} \quad (15)$$

*Correlation of Coefficient = Pearson Correlation Coefficient =*

$$R = \frac{\sum_{i=1}^N (y_{i,measured} - \overline{y_{measured}})(y_{i,prediction} - \overline{y_{prediction}})}{\sqrt{\sum_{i=1}^N (y_{i,measured} - \overline{y_{measured}})^2 \sum_{i=1}^N (y_{i,prediction} - \overline{y_{prediction}})^2}} \quad (16)$$

### 3 Literature Review

In this chapter, some of the recent studies using ANN to predict the loading conditions on turbines are reviewed. Reviewing the literature, it is seen that a wide variety of techniques and approaches have been utilized based on the available data, application, and the developer's opinion. Therefore, there is no consensus on how the ANN should be developed and evaluated and it always depends on the available data and desired outputs. However, it provides valuable insight into how the model performance can be improved.

The developed ANNs reviewed in this section can be categorized into simulation-based and SCADA-based neural networks. The former uses the data directly from simulation software packages to develop the model or simulation results are used supplementary to measurements. The latter utilizes only the sensor readings from turbines to develop the model. The findings are as follows.

#### 3.1 Simulation-Based Neural Networks

Schröder et al. (2018) have evaluated the performance of ANNs compared to other load assessment methods namely polynomial chaos expansion (PCE) and quadratic response surface approach. The ANN is trained using a database resulted from load simulations. The input features are wind speed, wind field variance, vertical wind shear exponent, wind veer, turbulence length scale, and anisotropy factor. The output is DEL values for flapwise root bending moment.

There are two models built. The first one has one hidden layer with 16 neurons and the second model has 2 hidden layers with 11 neurons in each layer. A determination factor of 0.99 is seen in both models. However, the two-layer has resulted in a slightly better performance considering the mean absolute error. The result shows that ANN can generally be a suitable alternative to PCE approach considering both the time and accuracy (Schröder, et al., 2018). However, in the absence of load measurements, it is not possible to evaluate to what extent the developed model matches the reality. Moreover, in real-world scenarios, the used features are not available as part of the data gathered in a wind farm.

Following the approach used by Schröder et al. (2018) that utilizes the aeroelastic simulations, is the study by Natarajan & Bergami (2020). The study uses the mean of 10 min average SCADA data including rotor speed, blade pitch, electrical power, and wind speed is used to train an ANN to estimate the turbulence intensity for different mean wind speed bins. The values of turbulence intensity are obtained from FLEX5 aeroelastic simulation software.

Afterward, another ANN is trained with the 1Hz SCADA time series along with tower top acceleration to estimate the load time series for fore-aft and side-side tower bending moments and blade root flapwise moment from sensor readings for an available turbine. Then, a new ANN is trained using the data from

FLEX5 software to predict the damage equivalent loads in different wind speeds and turbulence intensities. It is concluded that using the presented approach, relative fatigue life consumption of turbines within the farm can be evaluated (Natarajan & Bergami, 2020). However, the model's performance is not quantified.

In a similar study by Dimitrov & Natarajan, (2019), two main cases have been considered. In the first one, the SCADA data is used along with an aeroelastic model using Hawc2 software to produce the loading data in the absence of the load measurements to develop a 3-layer feed-forward neural network with 24 neurons. As a pre-processing step, the SCADA data has been calibrated using the data from met masts to map the SCADA readings including wind speed and wind direction to the free flow condition. The model controller and structural properties are tuned to resemble the investigated turbine. Although the predicted loading values cannot be verified in the absence of measurements, the power output of the farm is evaluated resulting in  $R^2 = 0.95$  and  $NRMSE = 18\%$  where only SCADA data is used.

In the second case, the available SCADA data are used as input to estimate the available load measurements available for a wind turbine. The inputs are 10-min average SCADA of wind speed, blade pitch angle, tower top acceleration, and rotor speed, and the model is a 3-layer FFNN. It has been concluded the model has estimated the load values “sufficiently well” where a 10-15% difference has been seen for blade root flapwise equivalent bending moments (Dimitrov & Natarajan, 2019).

### 3.2 SCADA-Based Neural Networks

Moving to models developed directly from SCADA data to predict load measurements, Zhou et al., (2018) developed an ANN predicting DEL values on 7 turbine components where the number of layers and neurons are selected through trial and error. Furthermore, the features are selected using Pearson's correlation coefficient despite the fact it can only capture linear correlations. The features include nacelle acceleration, blade pitch angle, rotor RPM, generator torque, active power, temperature, and pressure.

The performance of the model is evaluated using  $R^2$  and MAPE. The results range from 1.28% to 15.6% for MAPE and 0.882-0.951 for  $R^2$ . Related to our study, the MAPE value for edgewise and flapwise bending moments are 1.28% and 10.8% respectively. The wind speed is not used as it is believed it is under the rotor wake effect thereby not having an acceptable quality. It is suggested that the accuracy of the model can be increased by using more features available in SCADA data (Zhou, et al., 2018).

The effect of using 1Hz SCADA data is investigated in a study where 1Hz SCADA data including wind speed, pitch angles, nacelle direction, and power from 12 non-consecutive representative days are used to train and test an FFNN to estimate the thrust values extracted from bending moments on the tower

measured by strain gauges. The data is split into an 80-20 ratio for training and test dataset. The trained model is then tested on another turbine within the farm to evaluate the model's performance. The accuracy of the model is measured using MAE and RMSE where the percentage of MAE relative to maximal thrust value is seen to be lower than 2%. Afterward, the predicted 1Hz thrust values are converted to damage equivalent loads leading to severe amplification for some data points (Santos, et al., 2020).

The model developed by Santos et al., (2020), comprises 4 dense hidden layers with 64 to 300 neurons in each layer with a rectified linear activation function. The number of layers and neurons is determined through trial and error. The Adamx optimizer is used to minimize the root square mean error to increase sensitivity to outliers. Moreover, data is normalized as a pre-processing step before being fed into the model. To avoid overfitting, an early stop call-back with 10 epochs patience is used where the maximum epoch is set to 100.

Higher frequency measurements are needed to produce a loading history with a higher resolution. Therefore, in a study by Noppe et al., (2016), data from accelerometers, SCADA, and strain gauges are used to estimate the full load on the foundation of an offshore turbine. The model is an FFNN developed in MATLAB with a tan sigmoid activation function and has 1 hidden layer. Three models are trained for non-operational, below rated speed, and above rated speed where the air density is fixed.

First, a model with 10 min average SCADA data including blade pitch angle, wind speed, rotor speed, and power output as input is trained to estimate the thrust values. Afterward, accelerometer readings are used to capture higher frequency excitations. Superimposing the prediction values from the developed ANN and higher frequency accelerometer readings, the full strain history is achieved. The physical properties of the turbine are obtained using a FEM model tuned to match the modal shapes of the turbine. It is concluded that this approach is satisfactory. However, the quality of the model for estimating thrust can be improved by using 1s data and more parameters to train the model (Noppe, et al., 2016).

Among various attempts to improve the ANN performance, Vera-Tudela & Kühn, (2017) have investigated the effect of flow condition regarding wake on the performance of the model to predict flapwise and edgewise bending moments at the blade root using 10 min average SCADA data including wind speed, power, generator speed, nacelle accelerations, and pitch angle. It is done by binning data into 30-degree bins for incoming wind direction.

It has been shown that mixed wake condition where the wake of more than one turbine affects the investigated turbine, has led to less accurate results. However, the observed difference between models' performances is not significant where the MAPE values range from 0.86% to 2.625% for edgewise bending moment and 6.770% to 10.55% for flapwise bending moment. Furthermore, it has not been

explained why the model trained for the turbine in the platform wake is more accurate than the model for the free-flow condition (Vera-Tudela & Kühn, 2017).

The data used by Vera-Tudela & Kühn, (2017) has been filtered to avoid outliers and only operational data has been used. Load measurements are turned into DEL values. To select the inputs, Pearson coefficient and pair-wise correlation are used to select the parameters with high correlation to loads while removing those inputs that are highly correlated with each other. The model is an FFNN with the logistic sigmoid as the activation function and two hidden layers with 30 neurons. Data is divided into 70-20-10 ratios for training, validation, and test datasets. To avoid overfitting, epochs are limited to 200 in training.

The choice of input features among others depends on the training time, available data, and accuracy of the model. In a structured attempt by Movsessian et al., (2020), 4 feature selection methods are investigated to determine the input of the model. They are namely, principal component analysis (PCA), neighborhood component analysis (NCA), Pearson's correlation, and stepwise regression. 10 min SCADA data including rotor rotational speed, nacelle accelerations, wind speed, wind direction, air density, pitch angle, and power is used. Then, the performance of developed models in predicting the DEL values of tower base fore-aft bending moments is compared.

The available dataset has been pre-processed by filtering outliers, missing data, and out-of-service conditions. Then it is split into a standstill, partial load, and full load scenarios using active power. An FFNN is developed using MATLAB machine learning toolbox with two hidden layers and 10 neurons in each layer with a sigmoid activation function. The training is stopped when no tangible reduction in error is seen after 6 epochs. It is shown that stepwise and NCA have resulted in lower errors while PCA used fewer features thereby increasing the model training speed. The number of features varies between 9-38. However, the accuracy of all models has been in the 9.45%-10.49% range for MAPE when all load cases have been considered while MAPE has been around 2%, 6%, and 23% for the full load, partial load, and standstill, alternatively (Movsessian, et al., 2020).

### **3.3 Summary and Findings**

Reviewing the literature, it is seen that neural networks can predict the loadings from the sensor readings available for wind turbines. In the absence of on-field measurements, simulation results have been used as an alternative or to improve the model's accuracy. There are two major setbacks with the simulation-based models. First, their performance cannot be validated by comparing the results with measured values. Second, they are not applicable to real-case scenarios as not all the used features are common to measure for wind turbines.

Regarding the ANN architecture, the use of densely connected FFNN has been dominant. A pre-processing method has usually accompanied the model's development. It usually includes removing outliers, data normalization, calibration of inputs, and binning the data various criteria such as operational status or inflow condition. The input features are usually selected based on physical understanding of the problem, correlation coefficients, or their contribution to the model's accuracy.

The features commonly used for SCADA-based ANNs include power, wind speed, rotor RPM, generator speed, wind direction, air pressure, temperature, and blade pitch angles. Both DEL values and loadings have been used as the model outputs. The former is used when average data usually - 10-min SCADA- is available while the latter is used with 1Hz measurements as inputs. The number of hidden layers and neurons in each layer have been selected usually by trial and error and based on the developer's criteria which is usually accuracy and time.

Turning now to the model's evaluation, the choice of accuracy indicators varies from one study to another thereby making the comparison of the models harder. The most used indicators are MAPE, NRMSE,  $R$ , and  $R^2$ . Therefore, models developed in this study are evaluated using all these indicators. Related to our study, the performance of the models in the literature predicting flapwise bending moment from SCADA data have been reported to have a MAPE value of 10%-15%, 10.8%, and 6.77%-10.55%. In contrast, the MAPE value for models predicting edgewise bending moment has been lower and has been reported to have a value of 0.86%-2.626%, and 1.28%.

## 4 Data and Methodology

In this part, first, the structure and the origin of the data used in this study are presented. The relationship between DEL values and different features are examined. Afterward, the available data is studied to provide further insights by illustrating the data at hand in different scenarios including different wake condition, and different power curves. Subsequently, the methodology used to utilize the data for the objective of the study is explained.

### 4.1 Lillgrund Wind Farm

In this part, a description of the available database and its origin is presented to provide the reader with the required background to understand the used material. It is followed by a series of qualitative examination on different aspects of the database to give insight into how different variables, wake conditions, and power curves have affected the measured loading values on the blades.

#### 4.1.1 Overview

“Lillgrund” offshore wind farm is located south of Oresund bridge connecting Sweden and Denmark. It comprises 48 Siemens SWT-2.3-93 turbines with a total capacity of 110 MW. This is the largest offshore wind farm in Sweden, and it is owned by Vattenfall (Vattenfall, 2010). The information about the wind farm and its turbines is tabulated in Table 2.

*Table 2: Lillgrund Offshore Windfarm Information*

Name of the Turbine	SWT-2.3-93
Number of Turbines	48
Capacity of Each Turbine	2.3 MW
Total Capacity of the farm	110 MW
Turbine's Rotor Diameter	93 m
Turbine's Hub Height	68.8 m
Turbine's Cut-in Speed	3.5 m/s
Turbine's Rated-Speed	13 m/s
Turbine's Cut-out Speed	25 m/s

The layout of the farm is shown in Figure 6. The location of the turbine A1 is (0,0) and the unit for both x and y axes is the rotor diameter. As can be seen, there is a 3.3D spacing between A-H and 4.4D spacing between 1-8 rows except for D4-D6 and E4-E6 turbines, where the spacing is 8.8 D. The reason is during the construction the vessels could not be operated due to shallow depth in that area. (Sikkeland, 2020)



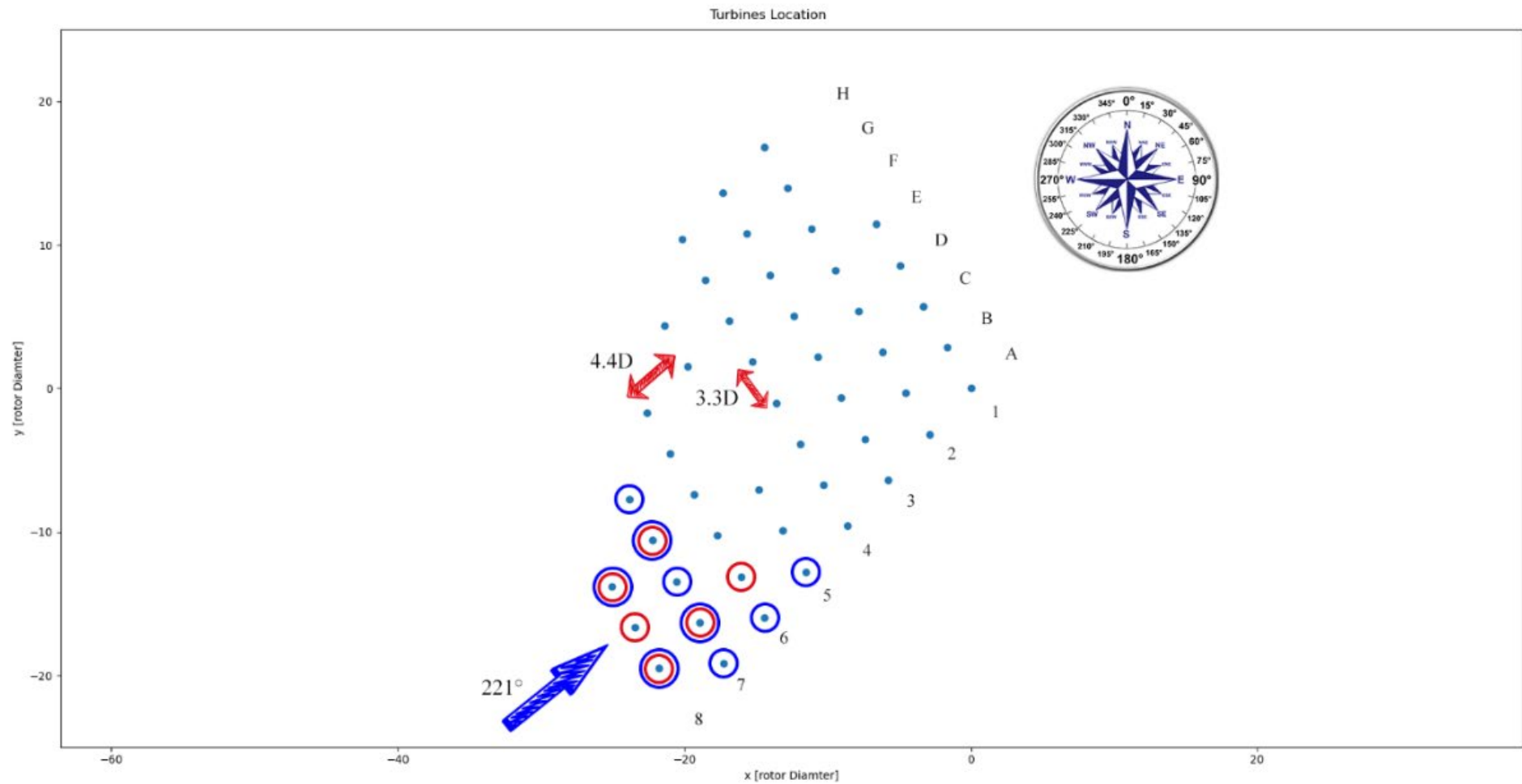


Figure 6: Lillgrund Windfarm Layout: Turbines with load measurements are marked with red circles, Curtailed turbines are marked with blue circles

*Table 3: The name of turbines with load measurements*

Turbine with Load Measurements
B6
B7
B8
C8
D7
D8

*Table 4: The name of Curtailed Turbines*

Name of Curtailed Turbines
A5
A6
A7
B7
C8
C7
D7
D8
E7

In 2010, a set of strain gauges were installed on 6 turbines in the Lillgrund wind farm to measure the bending moments on the tower, and blades along with accelerometers to measure nacelle accelerations with 10Hz frequency. The name of these turbines is tabulated in Table 3 and their location within the farm is shown in Figure 6. Moreover, 9 turbines were subjected to different power curves to perform another study which are referred to as “curtailed turbines” in this study. The effect of this will be discussed further in 4.1.4. Names and locations of these turbines are presented in Table 4 and Figure 6, respectively. The available data for tower bending moments are not calibrated. Therefore, they are disregarded in this study. Processing the above-mentioned data along with 1Hz SCADA data resulted in a 5min averaged database where the parameters used in this study are summarized in Table 5. Nacelle accelerations are available in form of standard deviations of side-side and for-aft components. Blade load measurements are converted into damage equivalent loading (DEL) of flapwise and edgewise components using Rainflow counting algorithm where Whöler coefficient is 4 as the common value for steel components assuming that the loading is exerted on a steel element in the blade structure. (Carlén & Andersson, 2014)

This procedure is implemented by the windfarm owner and the processed data has been gathered in monthly files for all turbines within the farm. The author of the current study has solely relied on the

instruction documents for the database. Therefore, any possible flaws in the process of making the database are not known to the author.

Before attempting to develop the ANN, it is required to gain insight into the available data. Namely, it is of interest to illustrate how the change in different parameters, using different power curves, and wake conditions affect the DEL values. Therefore, in this section, a series of evaluations have been performed and the important findings are presented.

*Table 5: Used Parameters in the Available Database*

Used Parameters in the Available Database	
Date-time	STD/min/max/mean Pitch Angles
STD/min/max/mean Power	STD for-aft Nacelle Acceleration
STD/min/max/mean Nacelle Direction	STD side-side Nacelle Acceleration
STD/min/max/mean Wind Speed	DEL flap wise bending moment blade number 2
STD/min/max/mean rotor RPM	DEL edgewise bending moment blade number 2

#### **4.1.2 Evaluation of The Effect of Different Features on The Moment Values**

First, the possible effect of a series of parameters on the DEL values is investigated for turbine B8 to inspect how well the output values correlate with the features to find which ones contribute to the predictions. The reason why this turbine is selected is that it is placed on one edge of the windfarm, the loading data is available for it, and it is not curtailed. It allows us to check the effect of wake condition without the concern of how utilizing different power curves may have affected the observation. However, the effect of using varying power curves is done separately on other turbines.

The investigation is mainly done qualitatively as the objective here is to improve the understanding of the database at hand. First, based on the physical understanding of measured loads, the 8 most relevant parameters influencing the DEL values are considered as shown in Figure 7 and Figure 8. The outliers are filtered as summarized in Table 6 for the sake of readability of figures and to avoid false readings.

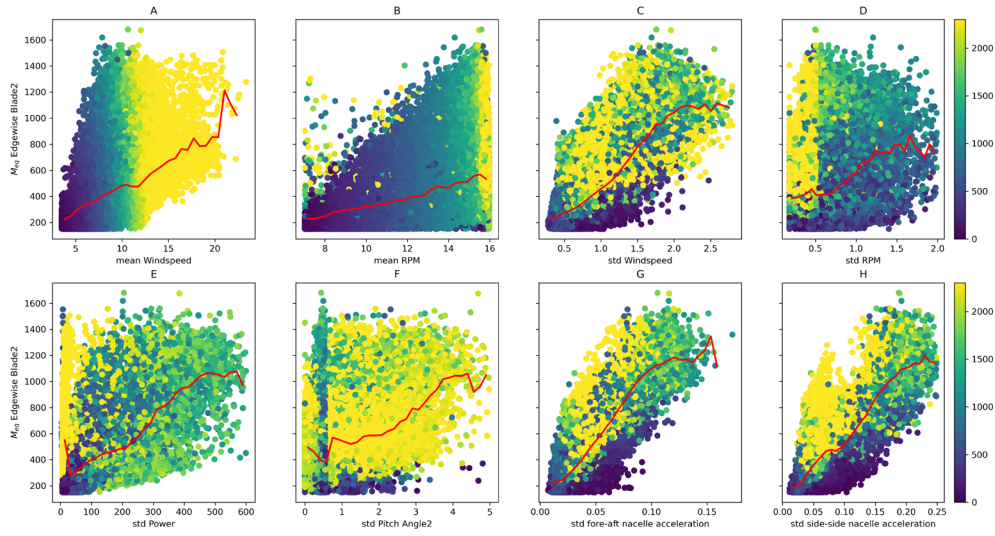


Figure 7: The  $M_{eq}$  edgewise blade 2 [kN.m] values vs features values, mean power coloured [kW], redline is the  $M_{eq}$  mean value for each bin of the studies features: A:  $M_{eq}$  edgewise blade 2 vs mean windspeed [m/s], B:  $M_{eq}$  edgewise blade 2 vs mean RPM, C:  $M_{eq}$  edgewise blade 2 vs STD windspeed [m/s], D:  $M_{eq}$  edgewise blade 2 vs STD RPM, E:  $M_{eq}$  edgewise blade 2 vs STD power [kW], F:  $M_{eq}$  edgewise blade 2 vs STD pitch angle2 [degree], G:  $M_{eq}$  edgewise blade 2 vs STD fore-aft nacelle acceleration [ $\frac{m}{s^2}$ ], H:  $M_{eq}$  edgewise blade 2 vs STD side-side nacelle acceleration [ $\frac{m}{s^2}$ ]

As can be seen, the edgewise moment increases with increasing wind speed and RPM as expected since the generated torque by edgewise moments is higher in higher wind speeds. Also, RPM increases with wind speed before the rated wind speed. Also, there is a strong correlation seen for STD wind speed and similarly nacelle acceleration. STD wind speed is an indication of turbulence intensity. Therefore, it can be concluded that since nacelle accelerations are highly correlated with STD wind speed, they are also excited by turbulence. In contrast, although the increase in STD RPM, STD pitch angle, and STD power increases the edgewise moment, the correlation is less strong. For lower values of STD pitch angle, STD power, and STD RPM, there can be high values for edgewise moment which is due to the change in other feature values.

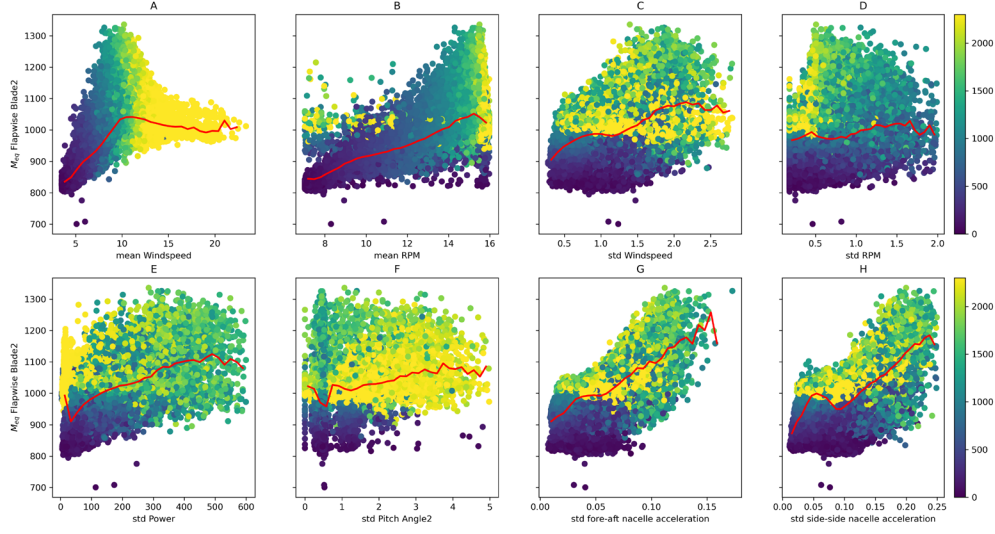


Figure 8: The  $M_{eq}$  Flapwise blade 2 [kN.m] values vs features values, mean power coloured [kW], redline is the  $M_{eq}$  mean value for each bin of the studies features: A:  $M_{eq}$  Flapwise blade 2 vs mean windspeed [m/s], B:  $M_{eq}$  Flapwise blade 2 vs mean RPM, C:  $M_{eq}$  Flapwise blade 2 vs STD windspeed [m/s], D:  $M_{eq}$  Flapwise blade 2 vs STD RPM, E:  $M_{eq}$  Flapwise blade 2 vs STD power [kW], F:  $M_{eq}$  Flapwise blade 2 vs STD pitch angle2 [degree], G:  $M_{eq}$  Flapwise blade 2 vs STD fore-aft nacelle acceleration [ $\frac{m}{s^2}$ ], H:  $M_{eq}$  Flapwise blade 2 vs STD side-side nacelle acceleration [ $\frac{m}{s^2}$ ]

Evaluating Figure 8, it is observed that the increase in windspeed increases the flapwise moment until it reaches rated windspeed. Then, it reduces and plateaus in the region for rated power showing that it is correlated with the regulation done by the control system. This also agrees with our understanding that the thrust coefficient decreases for high values of wind speed. Also, flapwise moment magnitude increases with increasing power or RPM as expected since it increases the thrust value.

Here, it is seen that the correlation of STD RPM and STD pitch angle is weak. It can be due to that the control system does not respond to high-frequency variation and its objective is to increase the power value which is more related to edgewise moment. Moreover, the loading due to actuation captured by STD pitch angle is in the edgewise direction. The correlations seen for nacelle directions and STD wind speed are strong. However, it is not as steep as seen for the edgewise bending moment due to less range of flapwise bending moments.

In summary, it is seen that how the considered variables correlate with moment values. However, it is needed to incorporate all of them in a non-linear model to predict the damage equivalent load in each direction.

#### 4.1.3 Evaluation of Wake Condition Effect on The Moment Values

To evaluate the effect of the wake condition, two different scenarios are considered. First, the DEL values for those time stamps where the incoming wind within  $221^\circ \pm 5^\circ$  where there is no wake on B8 turbine. Second, DEL values for those time stamps with the incoming wind within  $41^\circ \pm 20^\circ$  where the

turbine is under the wake of other turbines. The range of degrees for the second scenario is selected to have enough data points allowing a meaningful comparison. It is due to that the incoming wind direction is more frequent in  $221^\circ \pm 5^\circ$  as can be seen in Figure 9 and Figure 10. The results are shown in Figure 11 to Figure 14.

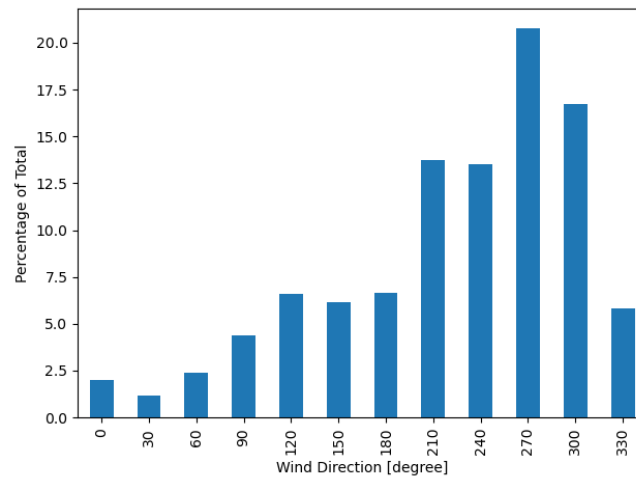


Figure 9: The distribution of incoming wind direction for turbine B8, each bin includes the degrees within  $\pm 15^\circ$

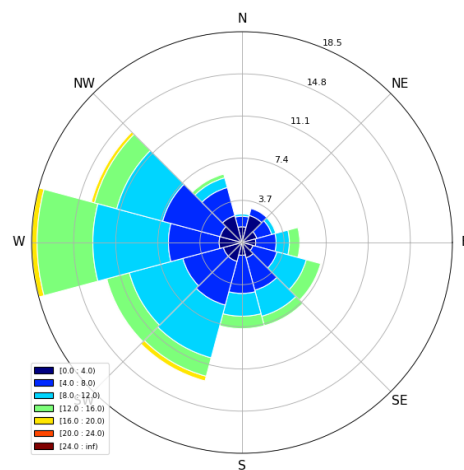


Figure 10: The wind rose for turbine B8

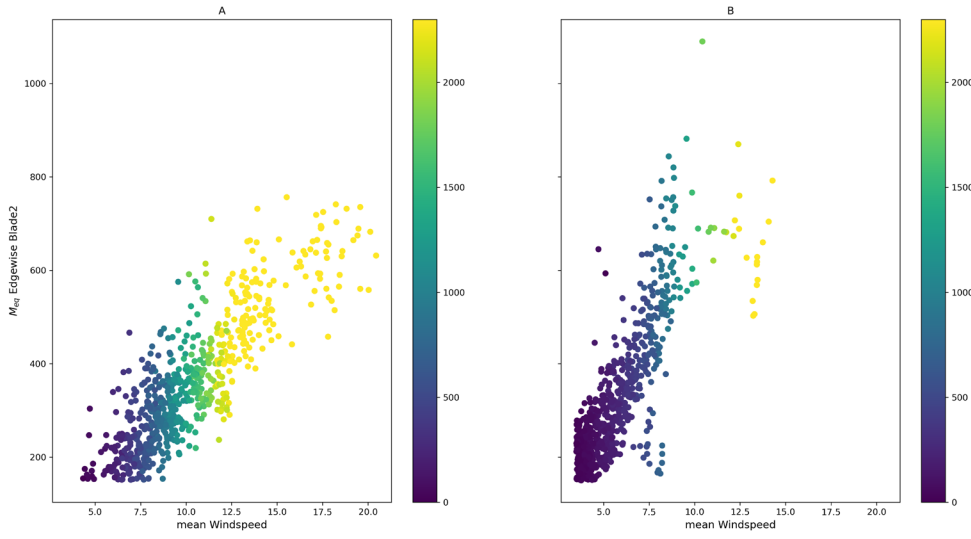


Figure 11: The value of  $M_{eq}$  edgewise blade 2 [kN.m] vs mean wind speed [m/s], mean power coloured [kW]: A: for incoming wind direction of  $221^\circ \pm 5^\circ$  B: for incoming wind direction of  $41^\circ \pm 20^\circ$

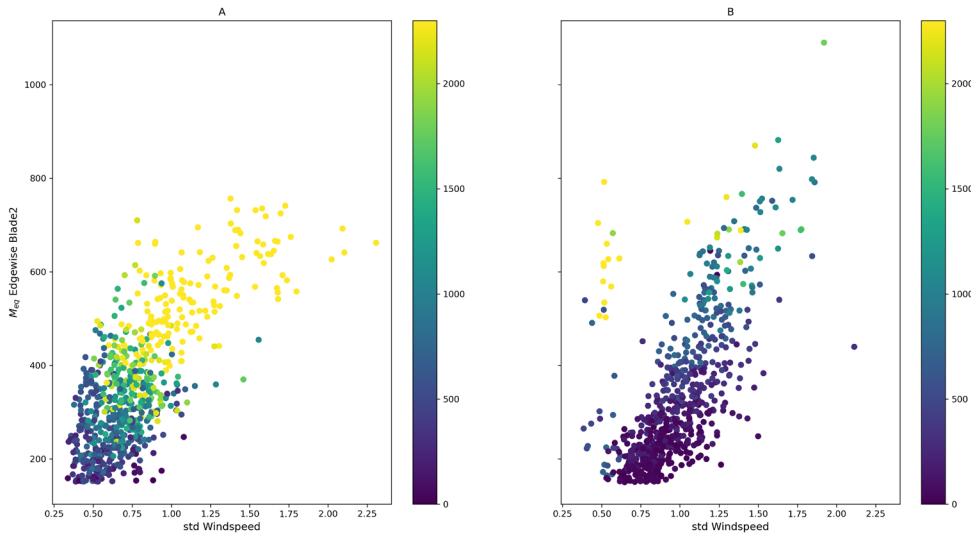


Figure 12: The value of  $M_{eq}$  edgewise blade 2 [kN.m] vs STD wind speed [m/s], mean power colored [kW]: A: for incoming wind direction of  $221^\circ \pm 5^\circ$  B: for incoming wind direction of  $41^\circ \pm 20^\circ$

As can be seen, the edgewise moments are lower for the situation where the turbine is operating under no-wake condition. Putting it in numbers, the same value of edgewise moment obtained at around  $20 \left[ \frac{m}{s} \right]$  for the turbine operating in the no-wake condition is obtained around  $17.5 \left[ \frac{m}{s} \right]$  when the turbine is operating is under the wake of upstream turbines. The reason for the difference can be partially seen in Figure 12 where the value of STD wind speed is higher in lower power values for the turbine under wake condition resulting in an exacerbated loading condition. Therefore, the wake condition is captured by STD wind speed to some extent.

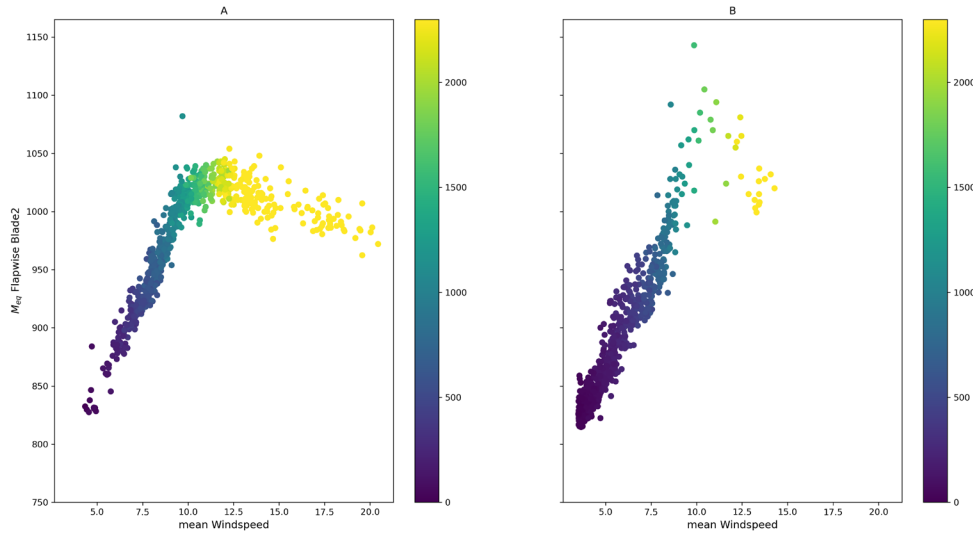


Figure 13: The value of  $M_{eq}$  flapwise blade 2 [kN.m] vs mean wind speed [m/s], mean power colored [kW]: A: for incoming wind direction of  $221^\circ \pm 5^\circ$  B: for incoming wind direction of  $41^\circ \pm 20^\circ$

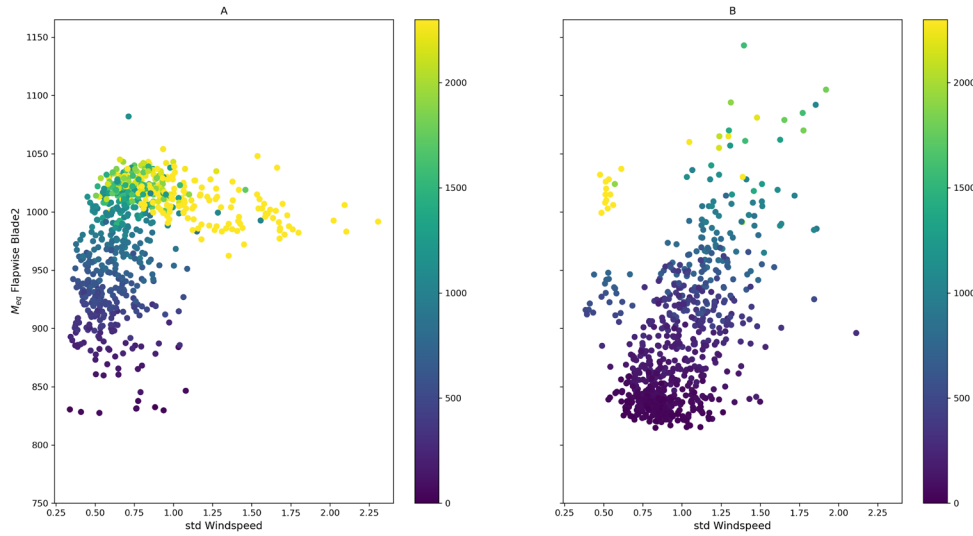


Figure 14: The value of  $M_{eq}$  flapwise blade 2 [kN.m] vs STD wind speed [m/s], mean power colored [kW]: A: for incoming wind direction of  $221^\circ \pm 5^\circ$  B: for incoming wind direction of  $41^\circ \pm 20^\circ$

The same situation is seen for the flapwise bending moment meaning that for the turbine under wake conditions the higher loading is obtained for lower mean wind speeds. Also, higher STD wind speed values are seen for lower power values for the turbine under no wake. However, as seen in Figure 8, the correlation of STD wind speed with the flapwise moment is less strong compared to the edgewise bending moment.



#### 4.1.4 Evaluation of Alternating Power Curves on The Moment Values

Here, the effect of using different power curves or curtailed turbines on DEL values are considered where different power curves correspond to different pitch angle settings. To do so, first, turbine D8 and D7 are considered where turbine D7's power curve alternates between two power curves every 30 minutes, while the power curve for turbine D8 is fixed to check how changing the power curve affects the loading on its blades. Second, turbine B7 and B6 are investigated where turbine B7 is using two different power curves and the power curve for turbine B6 is fixed, to check how changing the power curve for the upstream turbine affects the loading on the down wind turbine. This is illustrated in Figure 15 and Figure 16.

After overlaying the scatter plots of two conditions and plotting the line of mean values of the equivalent moment for each wind speed bin, it is seen that changing the power curve on the upstream turbine does not have any significant effect on the downstream turbine. However, using different power curves on turbine D7 has been shown to affect the loadings where in this case, power curve "C" has resulted in lower loading values for values of mean wind speed from 3.5 m/s to 12.5 m/s. Assuming that statistically, the other environmental conditions are the same, the observed reduction in loading values can be associated solely to the change in the power curve.

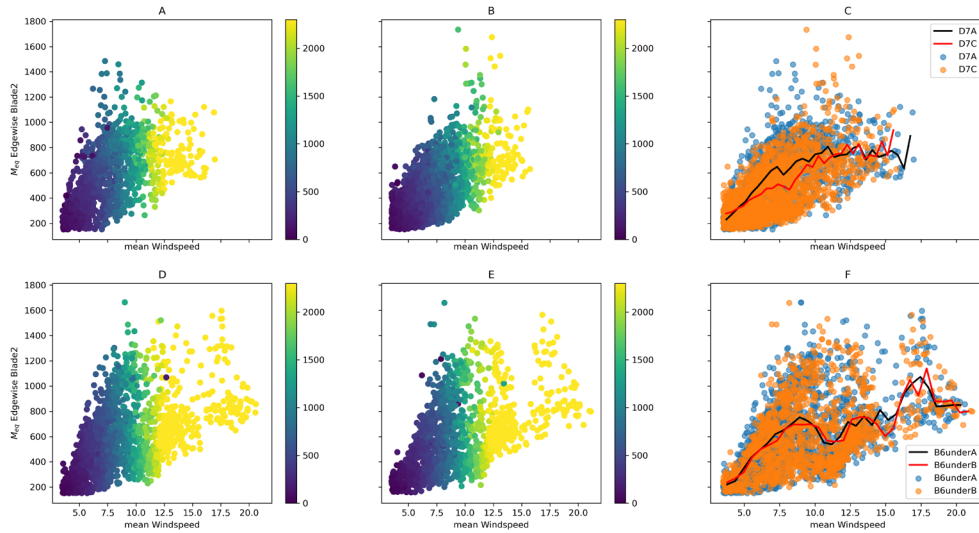


Figure 15: The effect of the power curve on the value of  $M_{eq}$  edgewise blade 2 [kN.m] vs mean wind speed [m/s], mean power coloured [kW], lines are the  $M_{eq}$  mean value for each bin of the mean wind speed: A: Turbine D7 with power curve A, B: Turbine D7 with power curve C, C: Turbine D7 with power curve A and B, the black line is turbine D7 with power curve A, red line is turbine D7 with power curve C, D: Turbine B6 under the effect of turbine B7 with power curve A, E: Turbine B6 under the effect of turbine B7 with power curve B, F: Turbine B6 under the effect of turbine B7 with power curve A and B, the black line is turbine B6 under the wake of turbine B7 with power curve A, the red line is turbine B6 under the wake of turbine B7 with power curve B

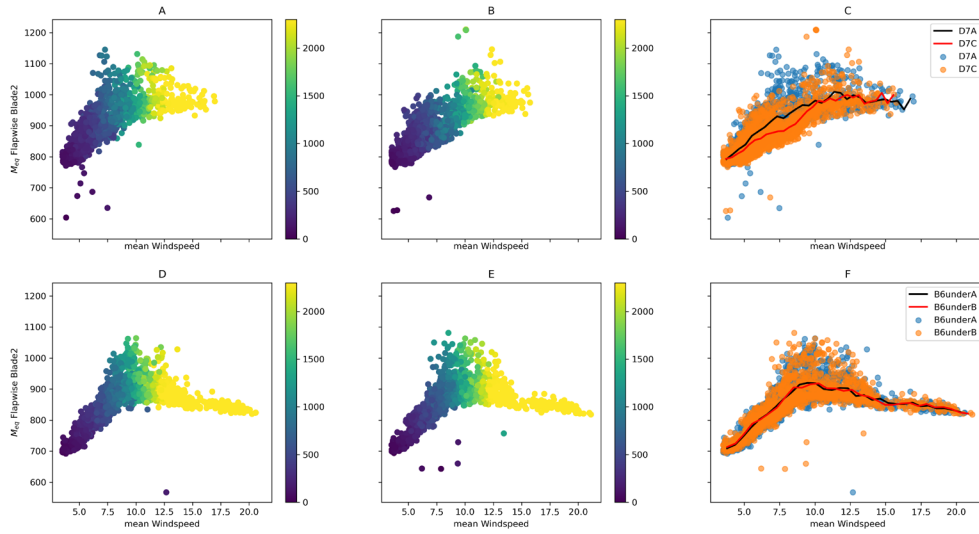


Figure 16: The effect of the power curve on the value of  $M_{eq}$  Flapwise blade 2 [kN.m] vs mean wind speed [m/s], mean power coloured [kW], lines are the  $M_{eq}$  mean value for each bin of the mean wind speed: A: Turbine D7 with power curve A, B: Turbine D7 with power curve B, C: Turbine D7 with power curve A and B, the black line is turbine D7 with power curve A, red line is turbine D7 with power curve B, D: Turbine B6 under the effect of turbine B7 with power curve A, E: Turbine B6 under the effect of turbine B7 with power curve B, F: Turbine B6 under the effect of turbine B7 with power curve A and B, the black line is turbine B6 under the wake of turbine B7 with power curve A, the red line is turbine B6 under the wake of turbine B7 with power curve B

## 4.2 NEWA Data

It is desired to have the free flow conditions from the site met mast to supplement the model inputs. However, since this data is not available, the mesoscale NEWA data at the site is obtained for the concurrent period when measurement data is available. This data is a result of the WRF model for 30 years. It is open source and can be downloaded online. The temporal and spatial resolution of the data is 30min and 3km, respectively.

NEWA data provides information about wind shear, wind veer, turbulence kinetic energy (TKE), the temperature at different heights, surface temperature, and surface pressure. Wind shear, wind veer, TKE, and temperature values are presented through wind speed, wind direction, TKE, and temperature values, respectively, at 50, 75, 100, 150, and 200 meters. Since the temporal resolution of NEWA data is different than SCADA and loading data, they are up sampled meaning the values of different variables are constant within six 5-min time stamps.

## 4.3 Methodology

The original dataset is in a specific format hard to understand by the user without a separate thorough instruction. Therefore, first, the data for different months are compiled into one file, the available data for each turbine is placed together, and the headers for each column of data were given names. The structure of the reorganized database is shown in Figure 17.

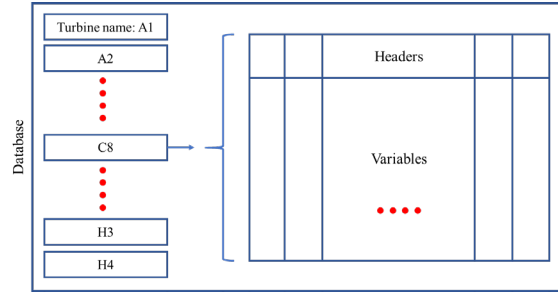


Figure 17: The structure of the reorganized database

Afterward, using the insight gained through reviewing the data for the relevant variables, they are filtered to remove outliers. Besides, only operational data have been used. It means only the data points with a mean wind speed between the turbine's cut-in and cut-out speed are considered. Besides, the non-operational data are removed. It is done using the turbine's electrical power output in the absence of the turbine's status flag. The summary of the applied filters is tabulated in Table 6.

Table 6: The summary of applied filters on data

Variable	Considered range
Mean wind speed	3.5-25 m/s
Mean power	$Power > 0$
$M_{eq}$ edgewise blade 2	150 – 3000 kN.m
$M_{eq}$ flapwise blade 2	500 – 1500 kN.m

In the next step, Python programming language is selected to develop the neural network. It is done using the Keras framework as part of Tensorflow library. The Spyder integrated development environment (IDE) is used to write the scripts. An FFNN configuration and MLP architecture are selected for the ANN as it is suitable for the objective of the study (Fine, 1999). Also, its development does not require high-level expert knowledge.

The data is divided pseudo-randomly into an 80-20 ratio for training and test dataset. It means that a series of random selections are performed. However, with every run, the selection results in the same subsets. It allows meaningful comparison between results. Validation is done by 20 percent of the training dataset. An early stop call-back with a patience of 10 epochs is defined to stop the training when no considerable improvement in the validation loss is seen to avoid overfitting.

Since the model uses randomness during the initialization of weights, the results vary to some extent each time it is trained on the same data. Therefore, here, each model is trained 5 times, and the mean value for each of the accuracy indicators is considered. Besides, for each model, the convergence criteria are checked. A normalization layer is applied to the features to have all variables having the same range.

It will prevent the model to prioritize one feature due to its larger numerical value when doing the predictions.

14 features including mean, min, max, and STD of RPM, power, and wind speed, along with STD of side-side and fore-aft nacelle accelerations are selected based on the obtained insight from data and their physical relevance. Then, it is required to determine the number of hidden layers and the number of neurons in each layer. It is done by trial and error. Therefore, the base model uses three hidden layers and 100 neurons in each hidden layer. Besides, the ReLU activation function is selected for each hidden layer. Adam optimizer is used to minimize the mean absolute error during the training. The learning rate is set as 0.0001. The mean absolute error is calculated as follows.

$$mean\ absolute\ error = \frac{1}{N} \sum_{i=1}^N |y_{measured} - y_{predicted}| \quad (17)$$

Next, a sensitivity analysis is done on how using a different combination of these 14 selected features influences the model's accuracy. Four indicators have been used to quantify the model's performance including *MAPE*, *R*, *R<sup>2</sup>*, and *NRMSE*. Afterward, more features are added sequentially to the model to evaluate whether they result in the model's improvement. Therefore, the features with positive influence will be kept in the model. Subsequently, the number of neurons and hidden layers will be revised to adjust for the increase in the number of features.

Then, the performance of the developed models is studied for a series of different scenarios including the effect of the wake condition on the model's performance, and model generalization. The latter is done by training the model on the data from one or more turbines and testing its performance on another turbine absent in the training. The outcome of each of the aforementioned steps are presented in 5. The summary of these steps is illustrated in Figure 18.

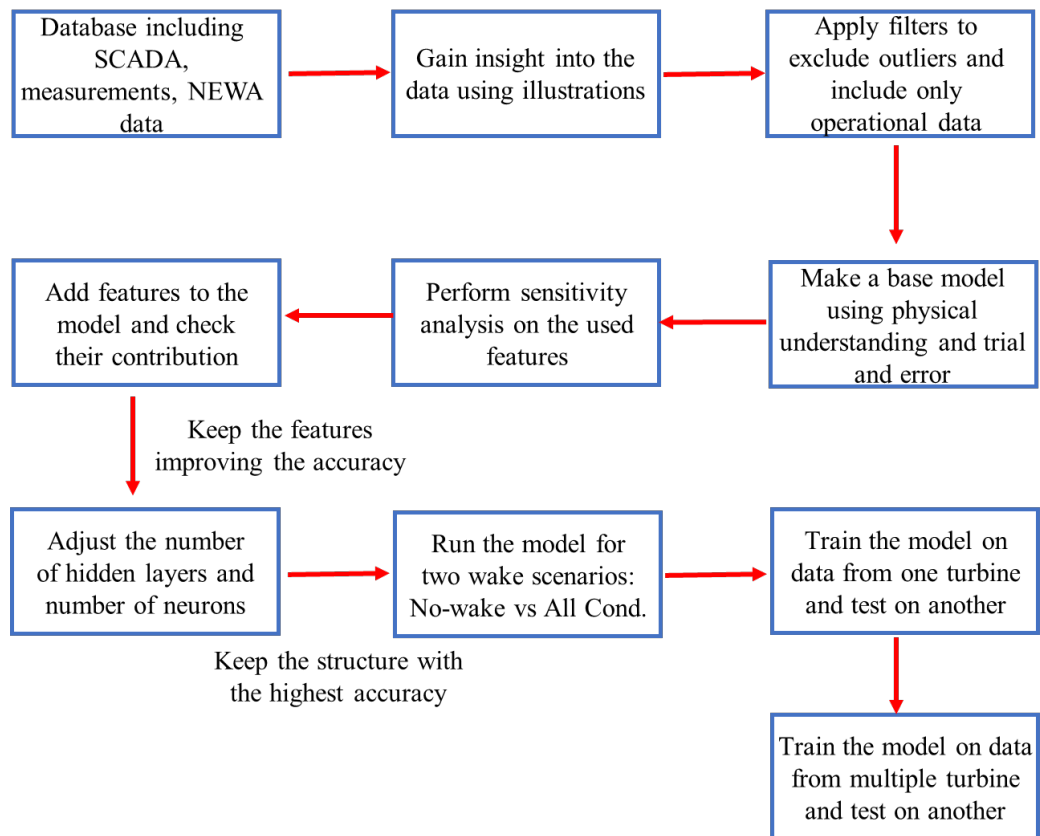


Figure 18: The summary of the used methodology

## 5 Results

In this section, the results of the steps mentioned in 4.3 and the accuracy of the model for each case is presented. It begins with presenting the performance of the base model. Next, the effect of removing subsets of base model's features is investigated. Afterward, it is examined adding which features improves the model's accuracy. Next, the number of hidden layers and neurons are adjusted to account for the increase in the number of features. It is followed by checking whether the developed model can capture the wake effect using the available features. In the end, the performance of the model trained on data from one or multiple turbines to predict the loading on two other turbines in the farm is evaluated.

### 5.1 Base Model's Performance

In the first step, the base model explained in 4.3 is used to train the data from turbine B8 using 14 selected features. This will function as the basis to allow a comparison between the performance of the models developed. The results are summarized in Table 7.

Table 7: The performance indicators for ANN trained on B8 Turbine using 14 features

	<i>MAPE</i>	<i>NRMSE</i>	<i>R</i>	<i>R</i> <sup>2</sup>
<i>M<sub>eq</sub> edgewise blade 2</i>	11.57	16.57	0.96	0.93
<i>M<sub>eq</sub> flapwise blade 2</i>	1.21	2.46	0.95	0.91

To show how accurate the model has been able to regenerate the labels in the test dataset, the normalized prediction values are plotted versus the normalized true values as seen in Figure 18 and Figure 20. Moreover, the distribution of MAPE values as a percentage of the total number of datapoints in the test data is presented in Figure 19 and Figure 21. It is seen the error values are centered around zero and

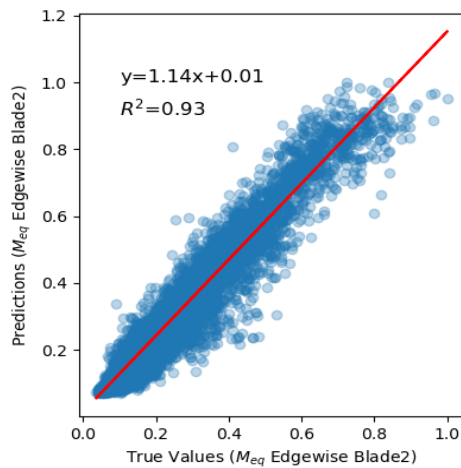


Figure 19: Normalized predictions vs Normalized true values for the  $M_{eq}$  edgewise blade 2, redline linearly fitted to data

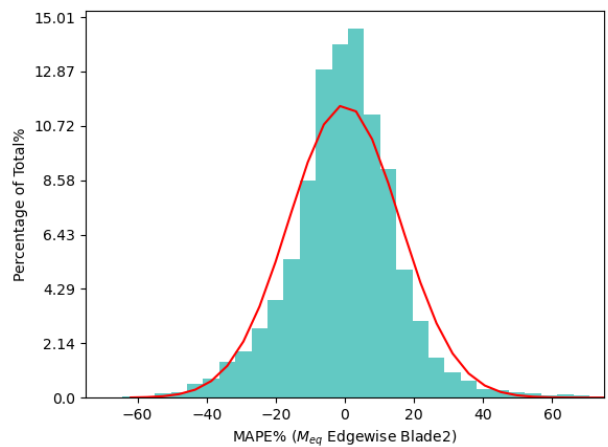


Figure 20: MAPE distribution for  $M_{eq}$  edgewise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure

follow a bell curve shape. However, the error values are higher for the edgewise bending moment compared to flapwise bending moment. The model has converged in both cases as can be seen in Figure 22 and Figure 23 as the loss value of training and validation datasets has not changed significantly for 10 epochs. The overfitting is avoided as there is no considerable difference between the loss value of training and validation datasets.

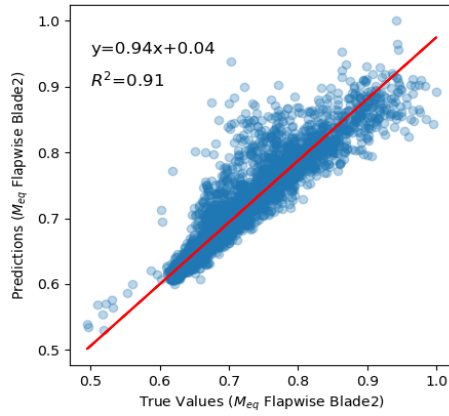


Figure 21: Normalized predictions vs Normalized true values for the  $M_{eq}$  flapwise blade 2, redline linearly fitted to data

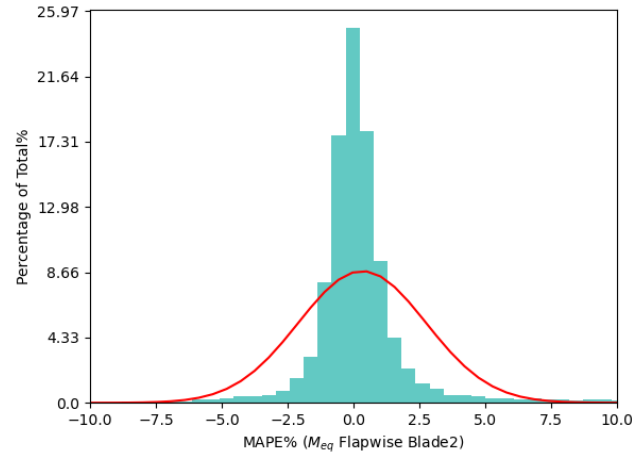


Figure 22: MAPE distribution for  $M_{eq}$  flapwise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure

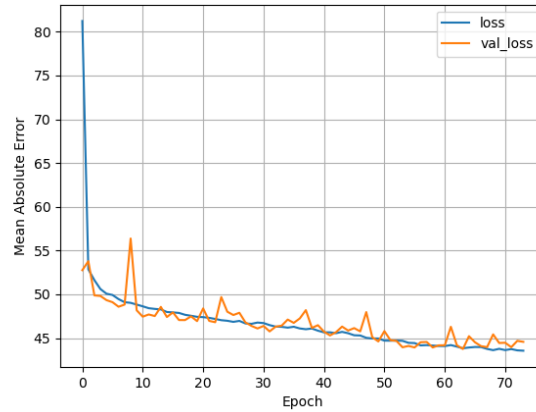


Figure 23: Convergence for the trained ANN vs Number of Epochs for  $M_{eq}$  edgewise blade 2

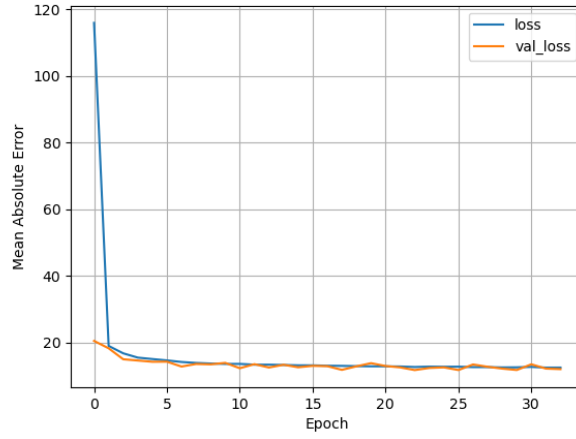


Figure 24: Convergence for the trained ANN vs Number of Epochs for  $M_{eq}$  flapwise blade 2

## 5.2 Sensitivity Analysis of Features for The Base Model

The ANN's performance sensitivity to the features is of interest as it shows the contribution of each of the features to the model's accuracy. Therefore, keeping the same structure for the model i.e., the same number of layers, neurons, etc, the number of features is changed to see how it affects the results. Here, 8 different scenarios are considered. The summary of this sensitivity analysis is presented in Table 8 and Table 9.

Table 8: Summary of sensitivity analysis of features on the model performance for  $M_{eq}$  flapwise blade 2

Description	No. of Features	MAPE	NRMSE	R	$R^2$
Base model	14	1.21	2.46	0.95	0.91
Without min values	11	1.21	2.45	0.95	0.91
Without max values	11	1.21	2.45	0.95	0.91
Without mean values	11	1.25	2.51	0.95	0.90
Without STD values	9	1.50	2.93	0.93	0.85
Without windspeed	10	1.27	2.50	0.95	0.90
Without RPM	10	1.29	2.51	0.95	0.90
Without power	10	1.21	2.50	0.95	0.91
Without nacelle accelerations	12	1.48	2.89	0.93	0.86



Table 9: Summary of sensitivity analysis of features on the model performance for  $M_{eq}$  edgewise blade 2

Description	No. of Features	MAPE	NRMSE	R	$R^2$
Base model	14	11.57	16.57	0.96	0.93
Without min values	11	11.86	17.01	0.96	0.92
Without max values	11	11.83	16.90	0.96	0.92
Without mean values	11	11.82	17.11	0.96	0.92
Without STD values	9	17.21	26.61	0.91	0.79
Without windspeed	10	12.44	18.84	0.95	0.90
Without RPM	10	12.37	18.41	0.96	0.91
Without power	10	11.94	17.17	0.96	0.92
Without nacelle accelerations	12	15.54	23.49	0.93	0.84

Starting with the model predicting flapwise bending moments, as can be seen, it is not required to use both min and max values in the model as removing them has not deteriorated the model's accuracy. In fact, removing one of them has slightly improved the NRMSE. In contrast, removing the STD values and mean values has affected the results negatively where the effect of STD values has been greater. Considering the features, removing each feature has led to less accurate results where nacelle accelerations have the greatest effect.

Reviewing the results for edgewise bending moment, removing any subset of features has worsened the model's performance. The difference between min, max, and mean values is minimal meaning removing any of them has the same effect on the model. Similar to flapwise bending moment, the greatest effect is associated with nacelle accelerations and STD values. Removing STD values and nacelle directions have increased the MAPE value by 48% and 34%, respectively.

### 5.3 The Effect of Adding Features on The Model's Performance

Next, the effect of adding new features on the model's performance is investigated to examine whether they can improve the model's accuracy by providing more information to the input vector. The added features include mesoscale NEWA data, min/max/mean/STD of nacelle direction, and blade pitch angles.

The results are summarised in Table 10 and Table 11. Note that in every step, the previous features considered in the previous step remain as features, and new features are added to them. Moreover, the structure of the model is left unchanged.

Table 10: Summary of effect of adding new features to the model on the ANN's performance for  $M_{eq}$  flapwise blade 2

Description	No. of Features	MAPE	NRMSE	R	$R^2$
Base model	14	1.21	2.46	0.95	0.91
With nacelle direction	18	1.05	1.99	0.97	0.94
With pitch angles	34	1.12	2.15	0.96	0.93
With wind shear	39	1.06	1.95	0.97	0.94
With wind veer	44	1.07	1.99	0.97	0.94
With TKE	49	1.07	1.97	0.97	0.94
With temperature	54	0.98	1.83	0.97	0.95
With surface temperature	55	0.95	1.71	0.97	0.95
With surface pressure	56	1.08	2.00	0.97	0.94

Table 11: Summary of effect of adding new features to the model on the ANN's performance for  $M_{eq}$  edgewise blade 2

Description	No. of Features	MAPE	NRMSE	R	$R^2$
Base model	14	11.57	16.57	0.96	0.93
With nacelle direction	18	11.27	16.14	0.96	0.93
With pitch angles	34	11.10	16.13	0.96	0.93
With wind shear	39	10.61	14.72	0.97	0.94
With wind veer	44	10.47	14.39	0.97	0.94
With TKE	49	10.19	13.87	0.97	0.95
With temperature	54	9.87	13.34	0.97	0.95
With surface temperature	55	9.90	13.49	0.97	0.95
With surface pressure	56	9.94	13.63	0.97	0.95

In the first step, the mean/STD/min/max nacelle direction is added. Next, the mean/STD/min/max values of pitch angle for three blades and the reference pitch angle value coming from the controller are fed to the model. It is followed by mean values of wind speed, wind direction, TKE, and temperature values at 5 mentioned heights in each step. In the last two steps, the effect of the surface temperature and surface pressure is evaluated.

As can be seen, adding pitch angles, wind veer, and surface pressure has not improved the model for  $M_{eq}$  flapwise blade 2. Also, surface pressure and the temperature has not contributed to the model's accuracy for  $M_{eq}$  Edgewise blade 2. Therefore, they are not used in the next step. The greatest improvement for flapwise bending moment has been seen with adding nacelle direction, temperature, and wind shear values. For the edgewise bending moment, the contribution of wind shear, nacelle direction, and TKE is more than other parameters. Here, the pitch angles have improved the model in contrast to flapwise bending moment where pitch angles decreased the model's accuracy.

## 5.4 Adjustment of The Number of Hidden Layers and Neurons in Each Layer

At this point, it is possible that the number of layers is not enough as the number of features has increased significantly. Therefore, in this step, the number of hidden layers and neurons in each layer, respectively, is adjusted to ensure the model structure is kept optimal. The previously obtained number of features for each target value is considered as the base model. For  $M_{eq}$  flapwise blade 2, it has 3 layers and uses 34 features, while the base model for  $M_{eq}$  edgewise blade 2 uses 3 layers and 54 features. The summary of results from increasing the number of layers is presented in Table 12 and Table 13.

Table 12: The effect of the number of hidden layers on the ANN's performance for  $M_{eq}$  flapwise blade 2

Description	No. of Features	MAPE	NRMSE	R	$R^2$
Base model, 3 layers	34	0.91	1.63	0.98	0.96
4 layers	34	0.87	1.54	0.98	0.96
5 layers	34	0.78	1.40	0.98	0.97
6 layers	34	0.84	1.51	0.98	0.96

Table 13: The effect of the number of hidden layers on the ANN's performance for  $M_{eq}$  edgewise blade 2

Description	No. of Features	MAPE	NRMSE	R	$R^2$
Base model, 3 layers	54	9.87	13.34	0.97	0.95
4 layers	54	9.83	13.26	0.97	0.95
5 layers	54	9.52	12.89	0.98	0.95
6 layers	54	9.41	12.77	0.98	0.96
7 layers	54	9.58	12.96	0.98	0.95

It is seen that the MAPE value for the model with flapwise bending moment as the target improves by adding two more layers and decreased afterward. Therefore, 5 hidden layer model is selected. For the model predicting edgewise bending moment, adding 3 more layers resulted in the improvement of results. Hence, it will be considered for the next step.

Next, the effect of adding more neurons to the model with 34 features and 5 layers for  $M_{eq}$  flapwise blade 2 and the model with 54 features and 6 layers for  $M_{eq}$  edgewise blade 2 is investigated. The results are tabulated in Table 14 and Table 15.

Table 14: The effect of the number of neurons on the ANN's performance for  $M_{eq}$  flapwise blade 2

No. Hidden Layers	No. of neurons in each layer	MAPE	NRMSE	R	$R^2$
5	75	0.87	1.54	0.98	0.96
5	100	0.78	1.40	0.98	0.97
5	125	0.81	1.50	0.98	0.96

Table 15: The effect of the number of neurons on the ANN's performance for  $M_{eq}$  edgewise blade 2

No. Hidden Layers	No. of neurons in each layer	MAPE	NRMSE	R	$R^2$
6	75	9.91	12.73	0.97	0.95
6	100	9.41	12.77	0.98	0.96
6	125	9.31	12.73	0.98	0.96
6	150	9.53	12.92	0.98	0.95

Trying out with 25-neuron steps, the model with 34 features, 5 hidden layers, and 100 neurons in each layer for  $M_{eq}$  flapwise and the model with 54 features, 6 hidden layers, and 125 neurons in each layer for  $M_{eq}$  edgewise are selected.

### 5.5 The Effect of Wake Condition on The Model's Performance

It is desired to check whether the ANN performance can be improved further if the wake condition is constant. The reason is that it is not certain whether the wake condition is captured by the used features. Moreover, by checking how dependent the model is on the specific wind direction, it is tested whether for this specific turbine the model learns which direction refers to different wake conditions or whether it remains a general predictor.

Therefore, by filtering the data by an incoming wind direction range where there is no wake on the studied turbine, the possible effect of wake condition on the model is eliminated. To do so, the model structures obtained previously are trained on data points in the range of incoming wind direction of  $221^\circ \pm 30^\circ$  that meets the no wake requirement while provides enough data points to train the model. The results are presented in Table 16.

Table 16: The wake condition effect on ANN performance

Label Name	Incoming wind direction range	MAPE	NRMSE	R	$R^2$
$M_{eq}$ edgewise blade 2	$0^\circ$ - $360^\circ$	9.31	12.73	0.98	0.96
$M_{eq}$ edgewise blade 2	$221^\circ \pm 30^\circ$	9.91	13.32	0.95	0.89
$M_{eq}$ flapwise blade 2	$0^\circ$ - $360^\circ$	0.78	1.40	0.98	0.97
$M_{eq}$ flapwise blade 2	$221^\circ \pm 30^\circ$	0.95	1.56	0.97	0.94

As seen, filtering the data points based on the incoming wind direction did not influence the ANN performance. The slight difference in results is believed to be due to reducing the number of available data points in the training dataset after applying the filter. It possibly shows that the wake effect is being captured by the features in the model structure to some extent.

## 5.6 Evaluation of The Model Generalization Performance

In this section, first, the accuracy of the model trained on data from turbine B8 to predict the loadings on turbine B6 and C8 is investigated. Next, the model is trained with data from all turbines with load measurements within the farm excluding the test turbine to examine whether using data from multiple turbines can improve the ability of the model to predict the loadings on a turbine absent in the training process.

### 5.6.1 Single Turbine

In this part, the performance of the models built previously using the data from B8 turbine will be evaluated on the data from two other turbines in the farm to investigate the possibility of generalizing the model trained on one turbine to other turbines within the farm. One of the turbines is C8 and the other one is B6. The reason why we have chosen these turbines is that first, the loading data is available for both and second, they represent two different wake conditions in the farm. C8 has a similar wake condition to B8 while B6 is almost always under the wake of other turbines to some extent. The results are tabulated in Table 17 and Table 18. The error distribution and comparison of predictions versus true values are shown from Figure 24 to Figure 31.

Table 17: ANN's performance trained on B8 data and tested on C8 and B6: 54 Features, 6 hidden layers, 125 neurons in each layer for  $M_{eq}$  edgewise blade 2

Description	MAPE	NRMSE	R	$R^2$
Trained on C8 Tested on C8	9.53	13.13	0.96	0.92
Trained on B8 Tested on C8	13.28	20.12	0.93	0.80
Trained on B6 Tested on B6	8.81	12.25	0.97	0.94
Trained on B8 Tested on B6	11.92	0.87	0.94	16.88

Table 18: ANN's performance trained on B8 data and tested on C8 and B6: 34 Features, 5 hidden layers, 100 neurons in each layer for  $M_{eq}$  flapwise blade 2

Description	MAPE	NRMSE	R	$R^2$
Trained on C8 Tested on C8	0.98	1.64	0.97	0.95
Trained on B8 Tested on C8	8.83	9.19	0.95	-0.07
Trained on B6 Tested on B6	1.83	2.77	0.95	0.91
Trained on B8 Tested on B6	18.98	19.85	0.93	-1.21

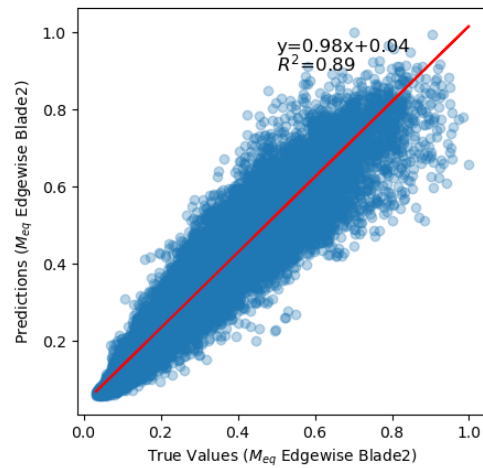


Figure 25: Normalized Predictions vs Normalized True values for ANN trained on B8 and tested on B6 for  $M_{eq}$  edgewise blade 2, redline linearly fitted to data

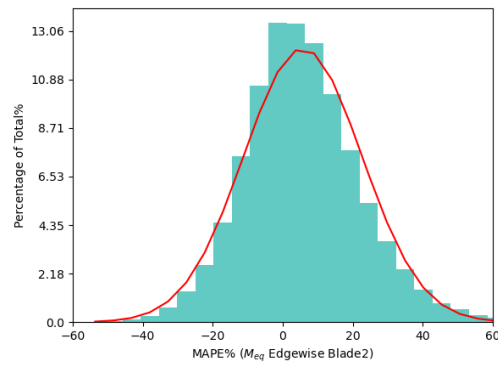


Figure 26: Percentage of Total vs MAPE for ANN trained on B8 and tested on B6 for  $M_{eq}$  edgewise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure

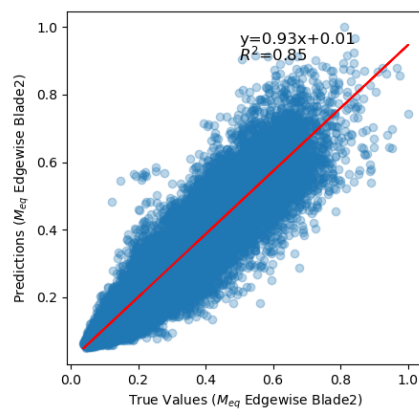


Figure 27: Normalized Predictions vs Normalized True values for ANN trained on B8 and tested on C8 for  $M_{eq}$  edgewise blade 2, redline linearly fitted to data

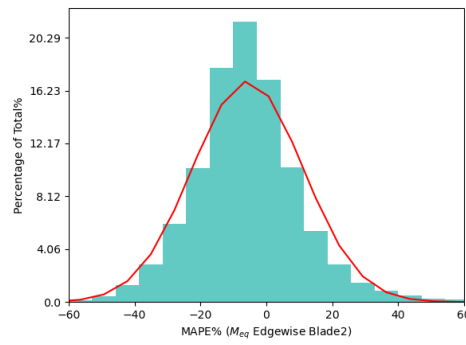


Figure 28: Percentage of Total vs MAPE for ANN trained on B8 and tested on C8 for  $M_{eq}$  edgewise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure

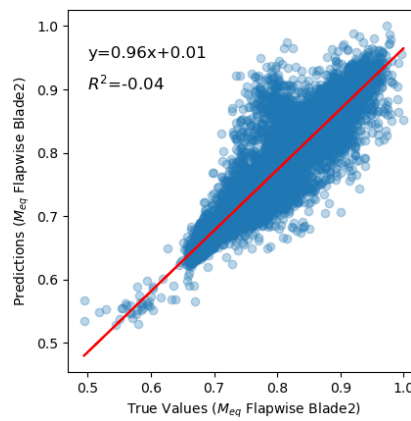


Figure 29: Normalized Predictions vs Normalized True values for ANN trained on B8 and tested on C8 for  $M_{eq}$  flapwise blade 2, redline linearly fitted to data

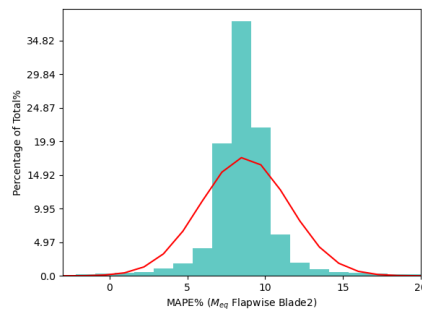


Figure 30: Percentage of Total vs MAPE for ANN trained on B8 and tested on C8 for  $M_{eq}$  flapwise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure

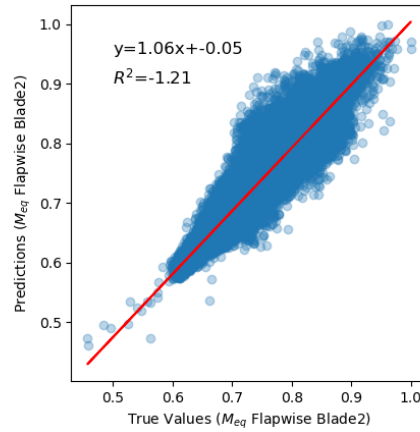


Figure 31: Normalized Predictions vs Normalized True values for ANN trained on B8 and tested on B6 for  $M_{eq}$  flapwise blade 2, redline linearly fitted to data

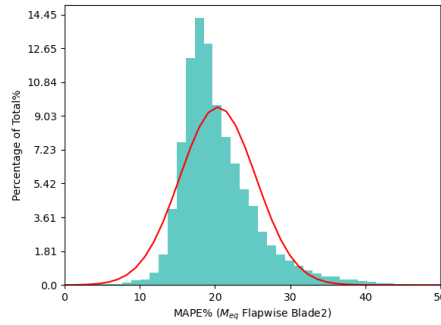


Figure 32: Percentage of Total vs MAPE for ANN trained on B8 and tested on B6 for  $M_{eq}$  flapwise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure

As can be seen, the performance of the ANN trained and tested on the same turbine is within the range achieved for B8 previously except for turbine B6 for flapwise bending moment where the error value is relatively higher. However, when the model trained on B8 is tested on B6 and C8, it results in significantly poorer performance for  $M_{eq}$  flapwise blade 2 while the degree of change is lower for  $M_{eq}$  edgewise blade 2 and of the same order compared to the case where the model is trained and tested on the same turbine.

Note that, as can be seen in Figure 29 and Figure 31, the error values for flapwise bending moment are centred around a value different than zero meaning that the error is systematic. It means one or more variables not considered have led to this condition. Also, the model is predicting poorly on both turbines with different wake conditions. Moreover, although it is seen that the ANN predicts better for turbine C8 with the same wake condition as B8 compared to turbine B6 for flapwise moment, the ANN prediction is less erroneous for turbine B6 for edgewise moment.



### 5.6.2 Multiple Turbines

It is conceivable that if the turbine operational condition is another feature, the higher error seen in the previous section should dampen if more than one turbine is used to train the model. It assumes that the operational condition of each turbine over the farm should be closer to the average operational condition of all turbines. The term “operational condition” includes those properties of turbine and measurement systems that were not available in database such as calibration errors, poor performance of subsystems, structural health, and so on. Therefore, in this section, an ANN is trained on all the turbines with loading measurements except the turbine to be tested.

Here, the model is once trained on the B6, B7, B8, D7, and D8 and is tested on C8. The next time B7, B8, C8, D7, and D8 are used for training, and B6 is tested. Moreover, it is useful to evaluate the performance of the ANN trained on all turbines with loading data and tested on itself along with ANN trained and tested using the data from the set of 5 turbines considered in the abovementioned scenarios. The results are tabulated in Table 19 and Table 20. The error distribution and comparison of predicted values versus true values are illustrated from Figure 32 to Figure 39.

*Table 19: ANN's performance generalization through the farm for  $M_{eq}$  flapwise blade 2*

Description	MAPE	NRMSE	R	$R^2$
Trained and tested on B6, B7, B8, D7, D8	2.16	4.73	0.94	0.87
Trained on B6, B7, B8, D7 and D8 tested on C8	5.55	6.15	0.93	0.47
Trained and tested on B7, B8, C8, D7, and D8	2.10	4.61	0.93	0.87
Trained on B7, B8, C8, D7 and D8 tested on B6	9.06	9.99	0.93	0.23
Trained and tested on all turbines with loading	2.08	4.14	0.94	0.87

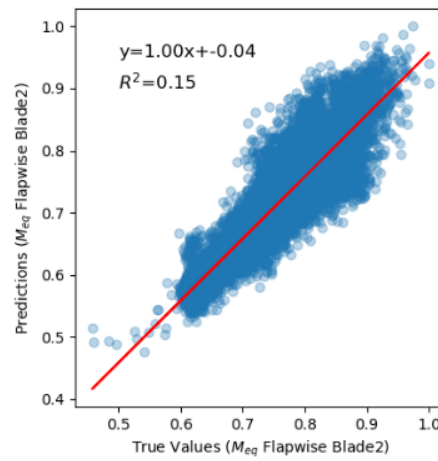


Figure 33: Normalized predictions vs normalized true values for ANN trained on B7, B8, C8, D7, and D8 tested on B6,  $M_{eq}$  flapwise blade 2, redline linearly fitted to data

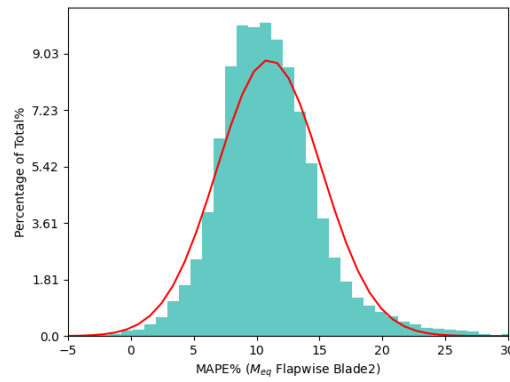


Figure 34: Percentage of total vs MAPE for ANN trained on B7, B8, C8, D7 and D8 tested on B6,  $M_{eq}$  flapwise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure

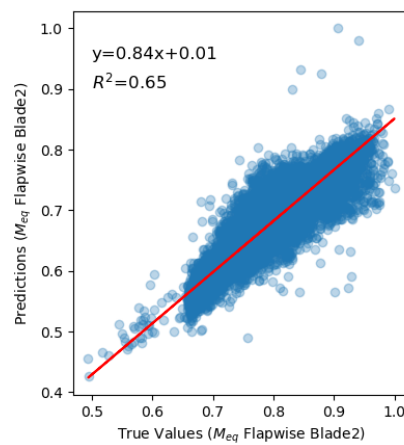


Figure 35: Normalized predictions vs normalized true values for ANN trained on B6, B7, B8, C8, D7, and D8 tested on C8,  $M_{eq}$  flapwise blade 2, redline linearly fitted to data

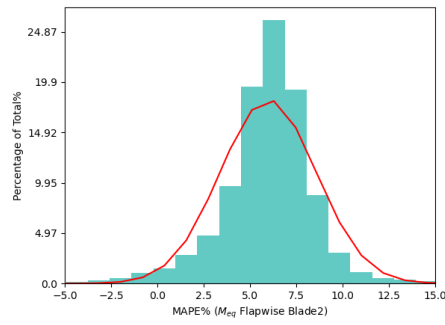


Figure 36: Percentage of total vs MAPE for ANN trained on B6, B7, B8, C8, D7, and D8 tested on C8,  $M_{eq}$  flapwise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure

Table 20: ANN's performance generalization through the farm for  $M_{eq}$  edgewise blade 2

Description	MAPE	NRMSE	R	$R^2$
Trained and tested on B6, B7, B8, D7, D8	9.01	12.19	0.97	0.94
Trained on B6, B7, B8, D7 and D8 tested on C8	13.47	19.21	0.92	0.86
Trained and tested on B7, B8, C8, D7, and D8	10.90	14.28	0.96	0.92
Trained on B7, B8, C8, D7 and D8 tested on B6	12.96	16.61	0.96	0.90
Trained and tested on all turbines with loading	10.30	13.70	0.96	0.92

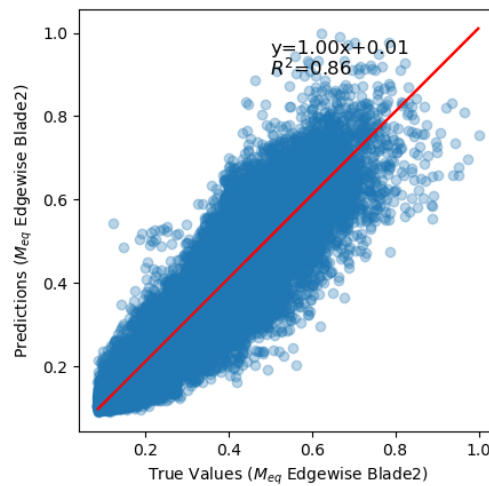


Figure 37: Normalized predictions vs normalized true values for ANN trained on B6, B7, B8, C8, D7, and D8 tested on C8,  $M_{eq}$  Edgewise blade 2, redline linearly fitted to data

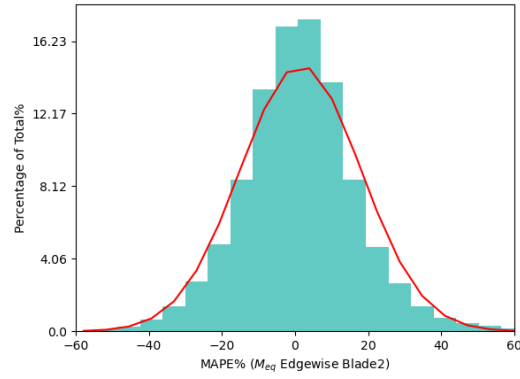


Figure 38: Percentage of total vs MAPE for ANN trained on B6, B7, B8, C8, D7, and D8 tested on C8,  $M_{eq}$  Edgewise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure

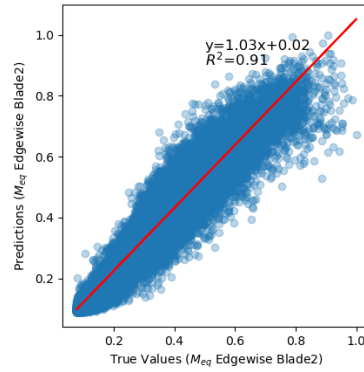


Figure 39: Normalized predictions vs normalized true values for ANN trained on B7, B8, C8, D7, and D8 tested on B6,  $M_{eq}$  Edgewise blade 2, redline linearly fitted to data

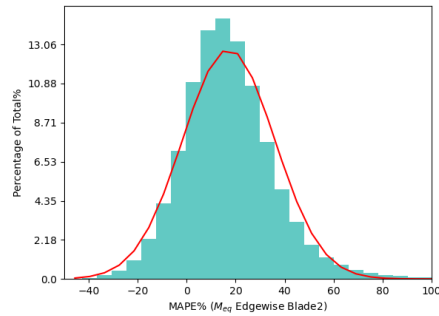


Figure 40: Percentage of total vs MAPE for ANN trained on B7, B8, C8, D7, and D8 tested on B6,  $M_{eq}$  Edgewise blade 2, redline normally fitted to data: it does not represent the range of all observations to increase the readability of the figure

It is seen that using more than one turbine for training, has improved the predictions in three cases and it has led to similar results for the other one. Namely, MAPE values for predictions on turbine C8 and B6 are improved from 8.83% to 5.55% and from 18.98% to 9.06% for flapwise moment, and from 11.92% to 10.30% for turbine B6 and edgewise moment. In contrast, the MAPE value for edgewise moment predictions is changed from 13.28% to 13.47% for turbine C8 which is presumably almost similar.

Regarding the models trained and tested using the same dataset from multiple turbines, the results are less accurate compared to the models trained and tested using one turbine for flapwise bending moment. For instance, the MAPE value for the model predicting flapwise bending moment is around 2.1% for multiple turbine scenarios compared to 0.78% for the model developed using only B8 turbine. In contrast, the MAPE value for models predicting edgewise bending moment has been around 9.5% in all cases.

## 6 Discussion

In this part, a discussion is made around the obtained results. It investigates the sources of uncertainty in the process, reasons why different models have produced different results, how the results correlate to the physical understanding of the problem, and a comparison between the results of this study and those found in the literature. It is as follows.

The process of building a database from measurement sensors can introduce uncertainty. For instance, measurement values commonly are stored in signal forms and need to be calibrated using gain and offset values to obtain physical values. During the measurement period, the sensors may degrade or be affected by various operational conditions. Therefore, previously established calibration values represent the reality with less accuracy.

In the absence of further information about the status of different sensors, it is not possible for the author to take these into account. It is possible to increase the quality of the database by applying filters. However, it results in having fewer data points thereby decreasing the model training performance. Therefore, a balance needs to be established. For this study, the filters are only applied on mean power, mean wind speed, and DEL values as described in 4.3.

Moreover, by using the statistical indicators within a 5-min period instead of high-frequency measurements some of the information will be lost. This is especially important when DEL values are used as Rainflow counting algorithm results are based on the load history of 10 Hz measurements and each load measurement depends on its concurrent feature values. For instance, for two sets of relatively similar min/max/STD/mean values of a feature, having different DEL values are possible. It is however possible that using higher-order indicators such as skewness and flatness can improve the model's accuracy by reducing the amount of lost information.

Another source of uncertainty is due to using mesoscale NEWA data in the absence of met mast measurements. The data has a temporal and spatial resolution of 30min and 3km, respectively. Therefore, for all 6 5-minute within 30 minutes, the values are the same. Moreover, the exact location of the reference point where values are available is not clear. Furthermore, since the methodology used in this study for feature selection and determining the number of layers and neurons is sequential, it is possible that a better accuracy can be achieved with different sets of layers, neurons, and features. Accounting for this shortcoming requires a complete grid search. Such an approach is highly time-consuming. Therefore, it is not used in this study.

A considerable difference can be seen between the MAPE values of flapwise and edgewise bending moments. For the case considered in Table 7, the MAPE values for flapwise and edgewise bending moment are 1.21% and 11.57%, respectively. As can be seen in Figure 40 and Figure 42, the range of

change for flapwise bending moment is less than the edgewise bending moment. Moreover, data points are distributed more evenly in each bin for flapwise compared to edgewise bending moment. Therefore, it is conceivable that the model is less trained for such situations than for those with more data points.

Since the test data is selected randomly from the training data, it has the same distribution. It means the number of data points in the test data from the bins with fewer data points is also in minority. Therefore, their contribution to the final MAPE is less. However, the degree of increase in the MAPE value due to the decrease in the number of data points is not clear as the model is non-linear. Therefore, the decrease in the number of data points does not necessarily level out the increase in MAPE values.

To investigate this, MAPE versus the number of data point in percentage for each bin is drawn for both models as shown in Figure 41 and Figure 43. It is seen that for both directions, MAPE value converges to the achieved MAPE value in 5.1 where more than around 2% of data points are available in each bin. Hence, the seen difference cannot be associated with the number of data points in different bins.

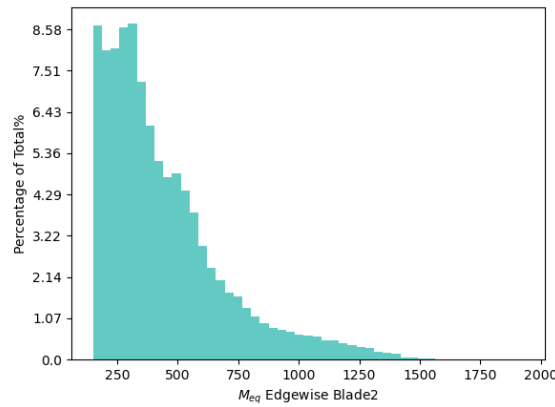


Figure 41: Distribution of  $M_{eq}$  Edgewise blade 2 for turbine B8:  $M_{eq}$  Edgewise blade 2 [kN.m] vs Percentage of Total

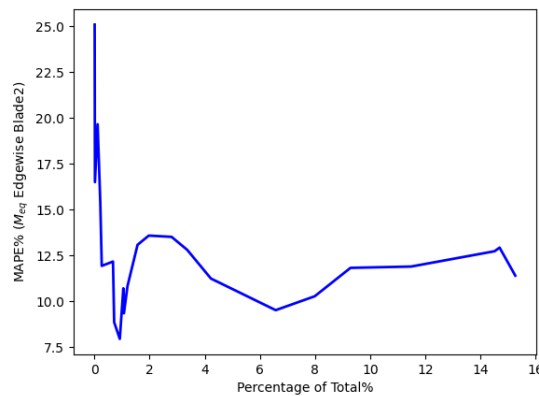


Figure 42: MAPE value in each bin vs number of datapoints in each bin% for  $M_{eq}$  edgewise blade 2

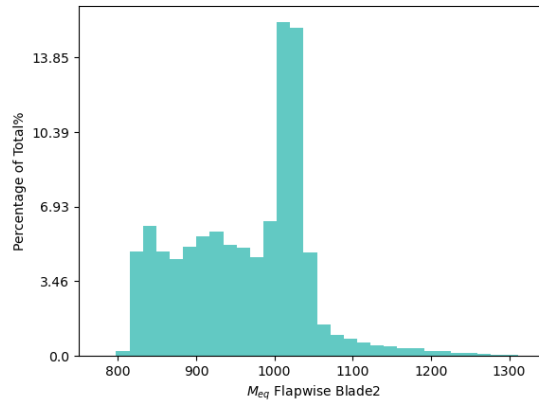


Figure 43: Distribution of  $M_{eq}$  flapwise blade 2 for turbine B8:  $M_{eq}$  flapwise blade 2 [kN.m] vs Percentage of Total

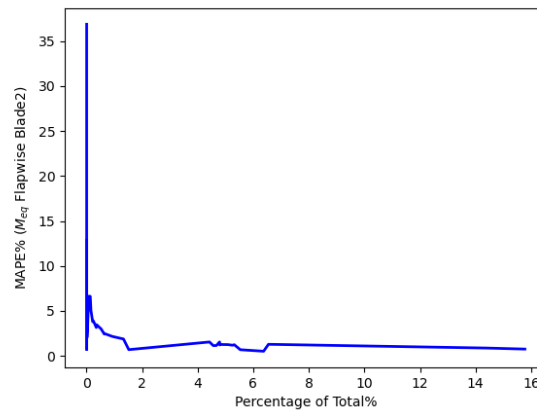


Figure 44: MAPE value in each bin vs number of datapoints in each bin% for  $M_{eq}$  flapwise blade 2

Another possibility is that there are not enough features to predict the DEL values in the edgewise direction as accurate as the flapwise direction. However, it is noteworthy that other studies have achieved better accuracies for edgewise bending moment using the same input features (Zhou, et al., 2018; Vera-Tudela & Kühn, 2017). Therefore, the reason why the model is performing better in the flapwise direction should be investigated in further research.

Based on the sensitivity analysis on the 14 features used in 5.1, removing one of wind speed, rotor RPM, and power as features did not change the model's accuracy significantly. It means that by having two of them the other one is redundant. This is in line with our physical understanding that these features are strongly correlated with each other. Moreover, it is shown that STD values, especially nacelle accelerations are the most determining for the model. Therefore, there is information correlating to the DEL values that cannot be captured using environmental features. In fact, nacelle accelerations are the only features related to the turbine structure. Besides, STD values can represent the load history behaviour by considering the variations from the mean values and account for turbulence intensity to some extent.



In 5.3, it is seen that the nacelle direction indicators have improved the model's accuracy. STD nacelle direction is an indicator of whether the turbine has yawed during the 5-min period. Yawing the turbine creates gyroscopic forces and for a turbine with coning angle, this leads to edgewise and flapwise components. Moreover, it can correspond to small changes in the wind direction which might be related to large scale turbulence. Furthermore, although it is not presented in 5.3 for the sake of conciseness, it was seen that adding the mean nacelle direction solely, cannot improve the model. This indicates that the location of the turbine used for training regarding wake conditions does not influence the learned model and wake effects are captured by other features. Adding pitch angles has not contributed to the model for flapwise moment while it improved the model for the edgewise bending moment. STD pitch angle relates to the actuation loading which is mainly in the edgewise direction. Moreover, the difference between the pitch angle of blades can be a source of an aerodynamic force.

Adding wind shear has contributed to both models as it causes cyclic loads in both directions. Wind veer only improved the accuracy of the edgewise model slightly and did not contribute to the model in the flapwise direction. The effect of wind veer is mainly important for large turbines where the wind direction changes significantly throughout the rotor swept area. TKE values contributed slightly to both models where the contribution to the edgewise model is higher. However, the relative effect is of the same order of magnitude in both directions. The reason can be that the turbulence effects are already captured through other features in the database such as STD wind speed.

Temperature values have improved both models as they provide information related to both density and atmospheric stability. The surface temperature did not improve the model for edgewise moment and its effect was minimal for flapwise moment. Moreover, although having the pressure is necessary for computing the density, the surface pressure did not improve any of the two models. The reason is since in NWEA data these values are constant during the day, the changes are disregarded by the model.

The wake condition effect on the model's accuracy is investigated in 5.5. No improvement is seen when the model is only trained with data points where the turbine is under no wake of upstream turbines. This agrees to some extent with the result of the study by Vera-Tudela & Kühn (2017). Therefore, it is concluded that the wake condition has already been captured to some extent by used features in the model such as STD wind speed which is an indicator of turbulence intensity. Moreover, the results show that the model does not learn a specific relationship between the nacelle direction and loading values but rather remains a general predictor.

Regarding the model's generalization, starting with the scenario considered in 5.6.1, all three turbines are using the same power curve thereby putting aside the possibility that the difference is due to this reason. The remaining reasons are the turbines themselves and the quality of readings from sensors. A difference in calibration gains and offsets or different turbine conditions, in particular the studied blade,

can lead to the observed difference. However, the author can't check the quality of calibration values or the turbines' operation status.

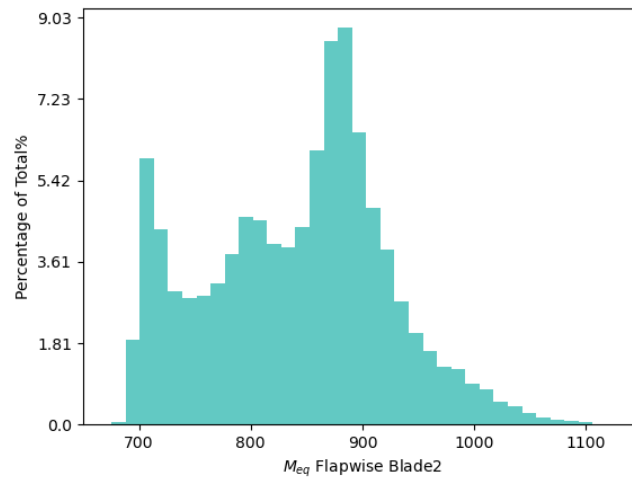


Figure 45: The distribution of  $M_{eq}$  flapwise blade 2 [kN.m] for B6 turbine

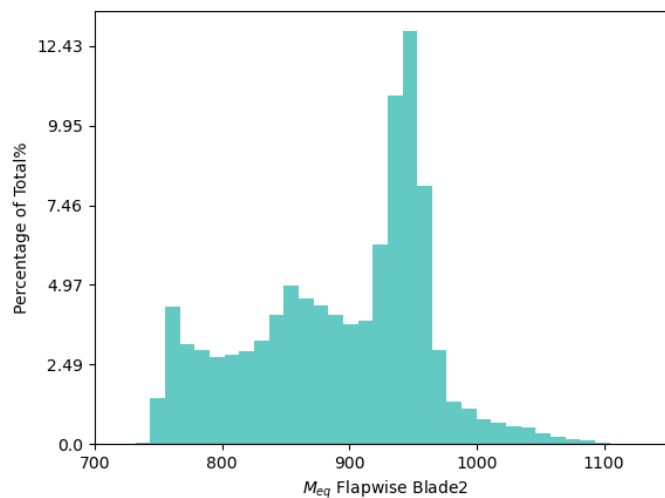


Figure 46: The distribution of  $M_{eq}$  flapwise blade 2 [kN.m] for C8 turbine

Looking at the distribution of the flapwise bending moment for the investigated turbines in this section as seen in Figure 42, Figure 44, and Figure 45, It can be seen that the range of  $M_{eq}$  flapwise for B8 is from 800 to 1200 kN.m with the mode value of just above 1000, while for C8 with the same wake condition and power curve, it ranges from around 750 to 1100 with the mode value between 900 and 1000. For B6, the values range from below 700 to 1100 and mode value below 900 kN.m. It shows that although C8 and B8 share a somewhat similar wake condition, they have been under different loading conditions due to differences that were not captured using the used features.

As shown in 5.6.2, using the data points from more than one turbine, it is possible to improve the model's generalization ability considerably regardless of what has caused the poor model's prediction in the single turbine case. Therefore, the model trained with all turbine data is simply trained on a broader range of loading conditions thereby performing better on the test turbine.

Comparing the obtained results from this study and those discussed in 3, it is seen that compared to the accuracy of the other SCADA-based models in the literature, the accuracy of our model for flapwise bending moment has been superior while it is less accurate for edgewise bending moment. For the edgewise bending moment, the reported MAPE values have been in the range of 0.86%-2.625% while our model resulted in a MAPE value of 9.31%. For the flapwise bending moment, our model resulted in 0.78% MAPE where this value has been reported between 6.77%-15% in the literature reviewed in this study (Vera-Tudela & Kühn, 2017; Zhou, et al., 2018).

Since each database is different and the details of the model development are not usually mentioned in the literature, it is not possible to pinpoint the origin of this difference easily. For instance, the details of the applied filters on the database, measurement duration, and wind farm environmental conditions are examples of what can cause the difference. In return, the superior performance of the flapwise model can be associated with using 5-min data instead of 10-min which is usually the case in previous studies. Therefore, less information is lost through averaging. Moreover, the number of features, neurons, and the number of hidden layers used in this study are determined solely to increase the accuracy of the model. Considering other criteria such as training time could change these parameters and the results. Other possible ways to improve our model can be using the met mast data, higher resolution readings, and applying more filters to the database. However, the latter should be done only if enough amount of data will be available after filters are applied.

## 7 Conclusion

This research aimed to investigate the possibility of developing an artificial neural network to predict the damage equivalent load values in the edgewise and the flapwise directions for wind turbine blades within Lillgrund offshore windfarm using 5-min data statistical indicators including min, max, mean, and STD. The contribution of various features including mesoscale NEWA data on the model's accuracy is evaluated and whether different wake conditions can affect the model's performance is examined. The current study also investigated the possibility of using the developed model to predict the DEL values for turbine blades where load measurements are not available.

This study has shown that it is feasible to use a densely connected FFNN neural network to train a model with acceptable accuracy. Training the model using the data from one turbine, a model with 6 hidden layers and 125 neurons in each layer with 54 features achieved a 9.31% MAPE value for predicting edgewise DEL values while a model with 5 hidden layers and 100 neurons in each layer with 34 features resulted in a 0.78% MAPE value for DEL values in the flapwise direction. The seen difference between MAPE values could not be explained within the performed investigation in this study. However, it was shown that the difference is not due to a lower number of data points in some bins. Therefore, this is yet to be explored in further research.

Investigating the contribution of different features on the model's accuracy in both directions, it is shown that having wind speed, rotor RPM, and power, only using statistical indicators of two features is enough. STD values emerged as the main contributors in predicting the DEL values where the effect of nacelle accelerations is dominant. Using nacelle direction proved to increase the accuracy as it includes the gyroscopic loads induced by yawing. Pitch angle indicators contributed only to the model predicting edgewise moment as the direction of loading caused by actuation is edgewise.

This study has found that in the absence of met mast data, it is possible to use mesoscale data to provide the model with information regarding the free flow conditions. Using mesoscale NEWA data for the concurrent period, it was seen that providing the model with this information decreased the MAPE values from 11.10% to 9.87% and from 1.05% to 0.95% in the edgewise and flapwise direction, respectively. The effect of adding wind speed and temperature values at different heights was dominant as they provide information about wind profile and atmospheric stability.

The investigation of wake condition on the model's accuracy showed that no significant difference is seen when the model is trained only using the data points representing no-wake condition where the turbine is not under the wake of upstream turbines. Therefore, it is concluded the features used in the model including STD wind speed, and nacelle accelerations can capture the wake conditions influencing the turbine blade loadings to some extent.

A significant finding to emerge from this study is that using only the data from one turbine in the farm to train a model to evaluate the fatigue life consumption of other turbines within the farm can result in considerably poor predictions. For the flapwise bending moment, the MAPE values increased from 0.98% and 1.83% to 8.83% and 18.98%, respectively. For the edgewise bending moment, an increase in MAPE values is seen from 9.53% and 8.81% to 13.28% and 11.92%, respectively.

It is suggested that poor results are due to differences in the structural condition of various turbines or faulty calibration sensors which cannot be captured using the available information. However, it is shown that using the data from multiple turbines to train the model can alleviate this problem. In this case, an improvement for the flapwise MAPE values was seen from 8.83% and 18.98% to 5.55% and 9.06%, respectively. For edgewise moment, the MAPE value was improved from 11.92% to 10.30% and remained similar for another case.

These findings suggest that in general, it is feasible to estimate the loadings on the wind turbine blades from SCADA data with acceptable accuracy. The analysis of the contribution of features to the model's accuracy has extended our knowledge about which data is needed for developing more accurate models. Furthermore, the results of this study contribute to the current literature by providing a comparison basis for other SCADA-based load estimations using neural networks.

## 8 Outlook and Future Work

A limitation of this study is that the methodology used for feature selection and determining the number of hidden layers and neurons in each layer cannot guarantee the best configuration has been achieved. A complete grid search would be advisable. However, due to the computational cost and lengthy process of it, it has been disregarded in this study. This study was limited to operational data. Therefore, standstill and transition events such as starts, and stops are not included.

The findings will be of interest to wind farm owners and asset managers to make a wise choice regarding the end of the life decision for the turbines by providing them with insight into the consumed and remaining life of wind turbine blades. It also can be used as an indication of the necessity of performing predictive maintenance, or corrective actions such as changing the control strategy of one or more turbines to reduce the operational costs and prevent failures. Besides, compared to using strain gauges to measure the loads which is time-consuming and labor extensive, this method provides a fast and cheap alternative to predict the loadings on turbine blades.

A natural progression of this work is to analyze the effect of using higher frequency data on the model's accuracy to reproduce the DEL values or high-frequency load history. Using high-frequency data increases the computational time significantly. Therefore, a procedure to select a representative dataset for model training is conceivable. A further study can develop other models to include standstill and transition events. Moreover, taking advantage of other ANN architectures and configurations can be investigated.

Further research needs to examine different possible sets of features along with the number of hidden layers and neurons to quantify the added computational time and the possible gain for accuracy. Moreover, the generalization of the model between different windfarms using the same turbine can be investigated. Furthermore, it is worth comparing the results of SCADA and simulation-based models to quantify the uncertainties associated with using simulations or find out which information cannot be captured through SCADA reading which is required for a more accurate prediction of loadings on turbine blades.

## 9 References

- Beer, F. P., Jr., E. R. J., DeWolf, J. T. & Mazurek, D. F., 2015. Repeated Loadings and Fatigue. In: 7th, ed. *Mechanics of Materials*. New York: McGraw-Hill, p. 67.
- Blasques, J. P. A. A. & Natarajan, A., 2013. Mean load effects on the fatigue life of offshore wind turbine monopile foundations. *Computational Methods in Marine Engineering -Proceedings of the 5th International Conference on Computational Methods in Marine Engineering*, Volume V, pp. 818-829.
- Burkov, A., 2019. *The Hundred Page Machine Learning Book*. Quebec: Andriy Burkov.
- Carlén, I. & Andersson, H., 2014. *Load assessment at Lillgrund*, s.l.: Vattenfall.
- Dimitrov, N. & Natarajan, A., 2019. From SCADA to lifetime assessment and performance optimization: how to use models and machine learning to extract useful insights from limited data. *Journal of Physics: Conf. Series*, Volume 1222 012032.
- Eggleston, D. M. & Stoddard, F. S., 1987. Wind Turbine Load Specification. In: *Wind Turbine Engineering Design*. New York: Van Nostrand Reinhold Company Inc., pp. 262-300.
- Fine, T. L., 1999. Preface. In: *Feedforward Neural Network Methodology*. New York: Springer, pp. v-vii.
- Garrad, A., 1990. Forces and Dynamics of Horizontal Axis Wind Turbines. In: L. Freris, ed. *Wind Energy Conversion Systems*. New York: Prentice Hall, pp. 119-144.
- Hansen, M. O. L., 2010. Introductions to Loads and Structures. In: *Aerodynamics of Wind Turbines*. London: James & James, pp. 103-106.
- Hofemann, C., Bussel, G. v. & Veldkamp, H., 2010. *Forecasting of wind turbine loads based on SCADA data*. Trondheim, EAWC.
- Kingma, D. P. & Ba, J., 2015. *Adam: A Method for Stochastic Optimization*. San Diego, 3rd International Conference for Learning Representations.
- Kuhn, M., Gasch, R. & Sundermann, B., 2012. Structural Dynamics. In: R. Gasch & J. Tvele, eds. *Wind Power Plants*. London: Springer, pp. 272-304.
- Luna, J., Falkenberg, O., Gros, S. & Schild, A., 2020. Wind turbine fatigue reduction based on economic-tracking NMPC with direct ANN fatigue estimation. *Renewable Energy*, Issue 147, pp. 1632-1641.

Lydia, M. & Kumar, G. E. P., 2020. Machine learning applications in wind turbine generating systems. *Materials Today: Proceedings*.

Manwell, J. F., McGowan, J. G. & Rogers, A. L., 2010. Mechanics and Dynamics. In: *Wind Energy Explained: Theory, Design and Application*. Chichester: John Wiley & Sons, pp. 157-202.

MATLAB, 2021. *Rainflow*. [Online]

Available at: <https://se.mathworks.com/help/signal/ref/rainflow.html>

[Accessed 2021].

Megson, T. H. G., 2017. *Aircraft Structures for Engineering Students*. 6th ed. NEW YORK: Elsevier Aerospace Engineering.

Movsessian, A., Schedat, M. & Faber, T., 2020. *Modelling tower fatigue loads of a wind turbine using data mining techniques on SCADA data*, s.l.: EAWC.

Natarajan, A. & Bergami, L., 2020. Determination of wind farm life consumption in complex terrain using ten-minute SCADA measurements. *Journal of Physics: Conference Series*, Volume 1618 022013.

Noppe, N., Iliopoulos, A., Weijtjens, W. & Devriendt, C., 2016. Full load estimation of an offshore wind turbine based on SCADA and accelerometer data. *Journal of Physics: Conference Series*, Volume 753 072025.

PyPi, 2021. *Rainflow*. [Online]

Available at: <https://pypi.org/project/rainflow/>

[Accessed 2021].

Santos, F. d. N., Noppe, N., Weijtjens, W. & Devriendt, C., 2020. SCADA-based neural network thrust load model for fatigue assessment: cross validation with in-situ measurements. *Journal of Physics: Conference Series*, Volume 1618 022020.

Schijve, J., 2009. *Fatigue of Structures and Materials*. 2nd ed. Amsterdam: Springer.

Schröder, L., Dimitrov, N. K., Verelst, D. R. & Sorensen, J. A., 2018. Wind turbine site-specific load estimation using artificial neural networks calibrated by means of high fidelity load simulations. *Journal of Physics: Conference series*, 1037(6).

Scikit Learn, 2021. *R<sup>2</sup> score, the coefficient of determination*. [Online]

Available at: [https://scikit-learn.org/stable/modules/model\\_evaluation.html#r2-score](https://scikit-learn.org/stable/modules/model_evaluation.html#r2-score)

[Accessed 20 April 2021].



SciPy, 2021. *scipy.stats.pearsonr*. [Online]

Available at: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.pearsonr.html>

[Accessed 20 April 2021].

Sikkeland, J. S., 2020. *Power Prediction in Wind Farms with the DWM model*, Ås: NMBU.

Söker, H., 2013. Loads on wind turbine blades. In: P. Brøndsted & R. P. Nijssen, eds. *Advances in Wind Turbine Blade Design and Materials*. s.l.:Woodhead Publishing Limited, pp. 29-58.

Stetco, A. et al., 2019. Machine learning methods for wind turbine condition monitoring: A Review. *Renewable Energy*, Issue 133, pp. 620-635.

Sutherland, H. J., 2000. A Summary of the Fatigue Properties of Wind Turbine Materials. *Wind Energy*, Volume 3, pp. 1-34.

Vattenfall, 2010. *Lillgrund – the largest offshore wind farm in Sweden*. [Online]

Available at: <https://powerplants.vattenfall.com/lillgrund/>

[Accessed 2021].

Vera-Tudela, L. & Kühn, M., 2017. Analysing wind turbine fatigue load prediction: The impact of windfarm flow conditions. *Renewable Energy*, Volume 107, pp. 352-360.

Zhou, S., Xue, Y., Ma, X. & Wang, W., 2018. Prediction of wind turbine key load based on SCADA data. *Transactions of the Chinese Society of Agricultural Engineering*, 34(2), pp. 219-225.