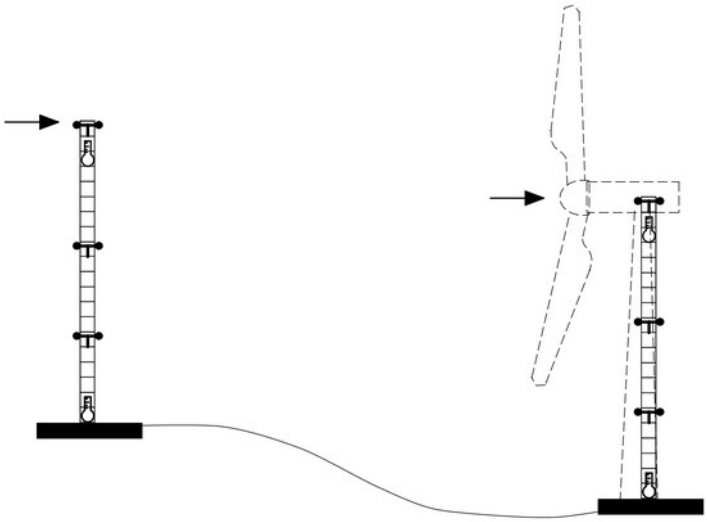
TÉCNICO LISBOA



Improvement of the standard Site Calibration for wind turbines

Implementation of Machine Learning techniques for Power Curve Measurement campaigns

Irma García Galadí

Thesis to obtain the Master of Science Degree in

Engineering and Energy Management

Supervisors: Prof. Ricardo Balbino Santos Pereira
Eng. Andreas Schmitz

Examination Committee

Chairperson: Prof. Duarte de Mesquita e Sousa
Supervisor: Prof. Ricardo Balbino Santos Pereira
Member of the Committee: Prof. João Carlos de Campos Henriques

December 2021

Acknowledgments

I would like to express my great appreciation to Andreas Schmitz and Fernando de Freitas for their valuable and constructive suggestions during the planning and development of this research work. Their willingness to give their time so generously has been very much appreciated.

I also would like to express my deep gratitude to Prof. Ricardo Pereira, my thesis supervisor, for his patient guidance, enthusiastic encouragement, and also his insight, always support and sharing of knowledge that has made this Thesis possible.

I would like to thank my family for their encouragement and caring over all these years and for always being there for me through thick and thin.

Last but not least, to all my friends and colleagues that helped me grow as a person and were always there for me during the good and bad times.

To each and every one of you – Thank you.

Abstract

Site Calibration is the previous step to Power Curve Verification for wind turbines on complex terrains. An uneven Site Calibration can affect the Measured Power Curve and increase the risk of compensation. Nowadays, Site Calibration can be performed by multi-binning methods or using CFD tools as described in the IEC 61400 – Part 12. This Master Thesis proposes a new Site Calibration methodology that consists of using the existing reference met masts data as the input of Machine Learning regression models to better predict the wind speed at the turbine location and hub height. Nine wind turbines at three different locations are the object of this study. Three Machine Learning techniques are implemented: linear modelling Polynomial regression considering both Ridge and Lasso regularization, Artificial Neural Networks and the decision tree-based Extreme Gradient Boosting techniques. The main outcome of this research is that Machine Learning models applied to Site Calibration are more accurate than the current IEC 2005 standards and improve the Measured Power Curve estimation. Extreme Gradient Boosting outperformed in an average difference of more than 30% for the RMSE wind speed error and around 29.3% in power by wind turbine compared to the IEC 2005 baseline. Universal models by wind farm still perform better than the standards. And finally, SHAP values explainability tool points out the most important variables: wind speeds and wind directions at different heights including Turbulence Intensity due to its non-linearity. Furthermore, the most important sensors are the anemometer at the hub height and the ultrasonic anemometer.

Keywords

Site Calibration, Wind turbines, Machine Learning, IEC standards, Power Curve Measurement campaigns, Warranty contracts, Wind Speed prediction, Meteorological masts data.

Resumo

A calibração ao local é a etapa anterior à verificação da curva de Potência para turbinas eólicas em terrenos complexos. Uma calibração ao local incorrecta afecta a curva medida, aumentando o risco de compensação para os fabricantes. Hoje em dia a calibração ao local pode ser realizada por métodos de multi-binning ou usando ferramentas CFD conforme descrito na norma IEC 61400. Esta tese propõe uma nova metodologia que utiliza dados do mastro meteorológico de referência como input para modelos de regressão de aprendizagem automática para prever melhor a velocidade do vento no local da turbina e à altura da Nacelle. Nove turbinas eólicas em três locais diferentes são o objeto deste estudo. Três técnicas de Aprendizagem Automática são implementadas: modelação linear, regressão polinomial considerando a regularização Ridge e Lasso, Redes Neurais Artificiais e Extreme Gradient Boosting baseado em árvore de decisão. O principal resultado desta pesquisa é que todos modelos de aprendizagem automática aplicados à calibração ao local são mais precisos do que os atuais padrões IEC 2005 e melhoram a estimativa da curva de potência medida. O Extreme Gradient Boosting reduziu o erro RMSE em mais de 30% para a velocidade do vento e cerca de 29.3% na potência da turbina eólica em comparação com a norma IEC 2005. Uma análise aos valores SHAP aponta como variáveis mais importantes a velocidade e direcção do vento em diferentes alturas e a intensidade de turbulência, enquanto os sensores mais importantes são os anemómetros da nacelle e ultra-sónico.

Palavras Chave

Calibração ao Local, turbinas eólicas, Aprendizagem Automática, norma IEC, campanhas de medição de Curva de Potência, contratos de garantia, previsão da velocidade do vento, dados de mastros meteorológicos.

Table of Contents

List of Figures

List of Tables

Acronyms

AEP	Annual Energy Production
AI	Artificial Intelligence
ANN	Artificial Neural Network
CFD	Computational Fluid Dynamics
CV	Cross-Validation
DT	Decision Trees
FCF	Flow Correction Factors
IAL	Independent Accredited Laboratories
IEC	International Electrotechnical Commission
LASSO	Least Absolute Shrinkage and Selection Operator
LES	Large Eddy Simulation
LIDAR	LIght Detection And Ranging
MAE	Mean Absolute Error
MAEP	Measured Annual Energy Production
MAPE	Mean Absolute Percentage Error
met	meteorological
ML	Machine Learning
MLP	Multi-Layer Perceptron
MPC	Measured Power Curve
MSE	Mean Squared Error
NSC	Numerical Site Calibration
NTF	Nacelle Transfer Functions
NWPS	Numerical Weather Prediction Systems

PC	Power Curve
PM	Permanent Mast
PPT	Power Performance Test
PR	Pressure
RANS	Reynolds-Averaged Navier-Stokes
RH	Relative Humidity
RMSE	Root Mean Squared Error
RS	Remote Sensoring
SC	Site Calibration
SCADA	Supervisory Control And Data Acquisition
SC-PC	Site Calibration- Power Curve
SGD	Stochastic Gradient Descent
SHAP	Shapley Additive exPlanations
SODAR	SOund Detection and Ranging
T	Temperature
TI	Turbulence Intensity
TM	Temporary Mast
TSA	Turbine Supply Agreement
WAEP	Warranted Annual Energy Production
WAsP	Wind Atlas Analysis and Application Program
WD	Wind Direction
WME	Wind Measurement Equipment
WPC	Warranted Power Curve
WSD	Wind Speed Distribution
WSH	Wind Shear
WSVer	Inflow Angle
XGB	Extreme Gradient Boosting

1

Introduction

Contents

1.1 Motivation

Wind power is one of the fastest growing renewable energy technologies. In only one year, from 2019 to 2020, the global installed wind generation capacity onshore and offshore increased by more than 18% [1] and for 2021 the forecast assumes a further acceleration of wind energy devices despite the global pandemic [2]. However, a high degree of certainty in investments return is the key for the development and execution of renewable projects. Ensuring the profitability of wind farms is the only effective way for new wind projects to be kick-started.

In the wind industry, particular attention is paid to the Annual Energy Production (AEP) since AEP is a fundamental parameter influencing the revenues of a project. A major concern for the AEP is the Power Curve (PC) of the wind turbine which expresses the relationship between the wind speed at the hub height and the power output. In this sense, the most broadly accepted Power Performance Test (PPT) for verifying the correct performance of wind turbines and assessing the Warranted Power Curve (WPC) is the Power Curve Measurement campaign. It consists in measuring the power output of the wind turbine after its commissioning, being the the Measured Power Curve (MPC) the result of this verification.

On top of this, the optimal conditions for wind energy production are high and steady wind velocities together with low turbulence levels. These circumstances can only happen in very specific areas which are usually onshore near the coastline [3]. However, such locations are often found in zones that are densely populated where the development of wind energy is often constrained. Therefore, other optimal locations for the deployment of wind farms are sought.

The onshore alternative for a massive deployment of wind turbines is in complex terrain. However, complex terrain has a significant impact on the local wind climate. Terrain local features strongly influence the wind speed and the wind direction [3]. Therefore, when it comes to complex terrains, the PPT is a two-step procedure that involves the 'Site Calibration (SC)' campaign prior to the Power Curve Measurement campaign. However, conducting a fair SC is not a smooth task and determining which methodology should become standard has become a subject for discussion in the wind industry. Moreover, several internal reports from the turbine supplier Vestas, have demonstrated that an uneven SC can significantly affect the MPC and thus increase risk for compensation due to underperformance on behalf of the wind turbine supplier, which is 5% of the turbine purchase price for a warranty period between 2 and 5 years [4].

1.2 Topic Overview

SC is the method used to enable Power Curve Verification in complex terrains. Obstacles and surface roughness may disturb the airflow between the position of the meteorological (met) mast used as a

reference and the centre of the turbine rotor. The SC approach estimates the impact of such disturbance and its uncertainty.

Fig. 1.1 shows the standard SC setup. It consists in placing two met masts, one in the reference position also called Permanent Mast (PM) and the second in the turbine location, known as Temporary Mast (TM) before the machine is commissioned. During the study period, the met masts record the wind speeds at the same time in each location. Once the recording period ends, the analysis is performed, and the turbine is erected. The SC procedure allows estimating a function or functions that transform the wind speed from the PM into the wind speed at the turbine position.

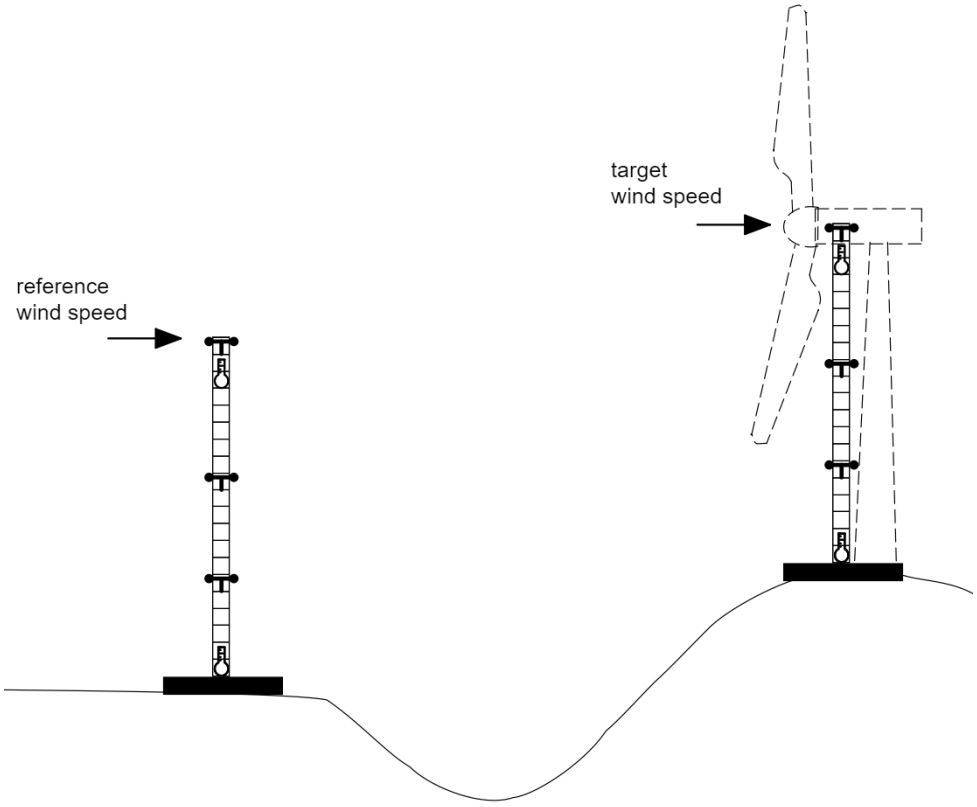


Figure 1.1: Standard Site Calibration setup

Currently, there are many different methodologies for a SC to be conducted. For simplicity, those will be split into two big groups: the International Electrotechnical Commission (IEC) standards-based SC which is nowadays broadly accepted and used in the wind industry. And also, the current main alternative is the physics-based SC, known as Numerical Site Calibration (NSC). Finally, advanced anemometry technology may be a future solution for measuring the wind speed at the turbine position and the hub height.

1.3 Objectives

The main goal of the work to be carried out is to develop a reliable, efficient, and non-expensive method for SC during Power Curve Measurement campaigns. Such a system should perform better than the existing IEC 2005 standards, thus should reduce the regression error on the target wind speed at the hub height in the wind turbine location and therefore should reduce the uncertainty on the MPC. In addition, this method would need to be of less complexity and more convenient for the practical implementation in the wind industry than the most powerful Computational Fluid Dynamics (CFD) modelling techniques.

The proposed methodology is based on the so-called Artificial Intelligence (AI). Specifically, different Machine Learning (ML) techniques are implemented to solve the regression problem which the SC consists of. Fig. 1.2 shows the workflow of this Thesis for the ML modelling.

The idea is to use the met masts data monitored by the Supervisory Control And Data Acquisition (SCADA) system to build a dataset. First, pre-process and condition the data and afterwards input the data to different ML models. The output of the modelling is the predicted wind speed at the turbine location and hub height.

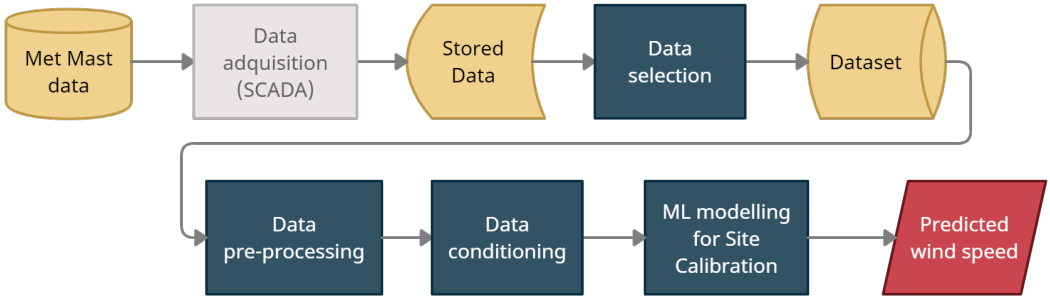


Figure 1.2: Complete Data process

One main difference with the current SC methodologies is that the proposed solution will consider the different signals that the sensors mounted on the PM are usually recording during SC, for instance, Temperature (T), Relative Humidity (RH) and Pressure (PR), and other post-processing variables which can be computed using the recorded signals.

In addition, part of the work carried out during this Thesis may consist also in determining the applicability of the method in the wind industry. The main idea is to address three different topics that would be essential for a real implementation of the proposed methodology: Viability of the models, sensitivity of the models to the site, and a sensor importance analysis.

The viability of the models is crucial and it is the first step for the implementation. Through the estimation of the uncertainty in power and energy terms for the recommended models, the goal is to determine if there is a potential reduction of the risk of compensation on behalf of the wind turbine

supplier, Vestas.

The second step would involve outlining a practical implementation of a Universal model that may allow for a more convenient SC by wind project. By determining the sensitivity of each of the ML techniques to the site, it is possible to estimate the trade-off between complexity and accuracy of the models and thus recommend a SC protocol based on ML.

And finally, a closely related objective of this Master Thesis would be identifying which are the most important sensors for SC. Accomplishing this objective may eliminate the need for installing certain sensors on the met masts, reducing the total cost of Power Curve Measurement campaign.

1.4 Thesis Outline

The present document is divided into 7 chapters:

- **Chapter 2: State-of-the-art** – Presentation of the state-of-the-art methodologies used for Site Calibration including a detailed analysis of prior work regarding Machine Learning techniques applied to the wind field, especially wind speed prediction.
- **Chapter 3: Problem Statement** – Detailed analysis of the limitations and inconsistencies of the standard IEC Site Calibration procedure as well as a presentation of the possible impact of an inaccurate Site Calibration on the final result of a Power Curve Verification test.
- **Chapter 4: Data analysis** – Full analysis of the data used for modelling. In this chapter, the complete procurement process data is explained, from the data collection on the meteorological masts to the data mining including data pre-processing and conditioning.
- **Chapter 5: Machine Learning modelling** – Description of the proposed Machine Learning methodology as well as the three different Machine Learning techniques implemented for modelling the datasets presented in Chapter 4.
- **Chapter 6: Results and implementation** – Detailed review of all the experimental tasks tackled during the modelling phase explained in Chapter 5 of this research project. In this chapter, the results and the explainability of the models proposed are shown. In addition to that, an initial study of the possible implementation according to the wind industry necessities is also presented.
- **Chapter 7: Conclusions and recommendations** – Overall conclusions of this Master Thesis as well as the main recommendations and further work based on the outcomes of the modelling phase and the implementation study.

2

State-of-the-art

Contents

Chapter 2 presents the state-of-the-art methodologies used for Site Calibration including a deep dive in Numerical Site Calibration and an overview of the Remote Sensing devices. This chapter also includes a detailed analysis of prior work regarding Machine Learning techniques applied to the wind field, especially wind speed prediction.

2.1 Numerical Site Calibration

NSC is the alternative method to the 2005 and 2017 IEC standard SC which is currently under research and is based on Multiphysics simulation. Entitled "Numerical site calibration for power performance testing of wind turbines", is the new IEC standard is part 12-4 published in September 2020 [5].

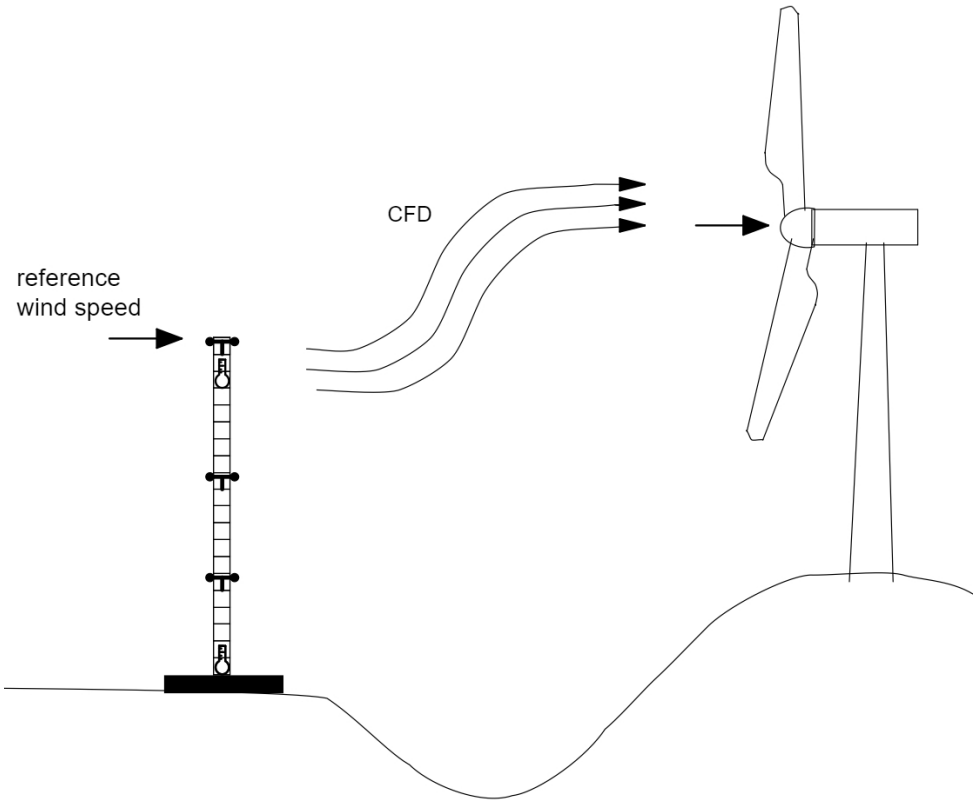


Figure 2.1: Numerical Site Calibration setup (adapted from [6])

2.1.1 Description

NSC is a three-dimensional flow analysis method based on CFD simulations to forecast the wind condition utilizing the correlations of wind characteristics between the reference site and the prospective wind turbine installation site [6]. NSC consists in using numerical simulations to produce the so-called wind

speed Flow Correction Factors (FCF) between the PM and the TM. The combination of the permanent reference mast together with the numerical simulation results will produce the wind speed at the turbine location required by the PPT. Compared to the traditional SC, NSC avoids the need for a TM at the future turbine location and thus eliminates the cost of erecting a second met mast. Fig. 2.1 shows the usual setup for a NSC to be conducted [5, 7].

2.1.2 ALEX17

ALEX17, the acronym for ALaiz EXperiment 2017, is the last full-scale experiment within the NEWA (New European Wind Atlas) project, which presents a utility-scale measurements campaign to characterize the wind flow in complex terrain [8]. The benchmark was open to anonymous participation of the different state-of-the-art modelling techniques of the wind resource until February 2021.

The main goal of this benchmark is to contribute to the evaluation of numerical uncertainty and maturity level by determining if state-of-the-art numerical flow models can, or not, provide reliable SC FCF to support PPT in complex terrain. It also will try to define best practices and summarize the existing guidelines for conducting and documenting NSC that is now included in the IEC 61400–12–4 [5, 9].

2.1.3 Numerical modelling techniques

CFD is a field of application developed through numerical solution techniques for producing quantitative predictions of fluid flow phenomena based on the conservation of mass, momentum, and energy laws [10]. As can be seen in Fig. 2.2, these techniques are typically based on the finite-volume method (mesh) which is a discretized representation of the computational domain by a series of mesh faces at which the flow equations are numerically solved. The resulting grid defines the calculation regions and requires the user to set up a number of parameters. Those parameters may include the mesh generation, boundary conditions, turbulence model, discretization algorithm. In wind energy applications, and particularly for estimating flow field in wind farms, it is also necessary topographical data (slopes, roughness, etc.) among other parameteric values and numerical schemes [11].

For wind energy applications and according to the new IEC Technical Report 61400-12-4 standard, three numerical techniques are employed for modelling the effects of turbulent flow fields: Direct Numerical Simulation (DNS), Large Eddy Simulation (LES) and Reynolds-Averaged Navier-Stokes (RANS). RANS, LES, Hybrid RANS-LES and Wind Atlas Analysis and Application Program (WAsP) (linear) [6]. These numerical techniques are also among the ALEX17 benchmark participants.

2.1.3.A LES

LES is a technique in which the structures of relatively large eddies are directly simulated and smaller eddies are modelled using a sub-grid scale model [12]. It uses high-resolution simulations which require substantial calculation resources to analyze the vortex to simulate the transient turbulence over complex terrain.

Matsushita et al. [13] proposed the application of a turbulence model standard Smagorinsky LES-CFD simulation with the software RIAM-COMPACT to perform a study on a NSC over complex terrain in Ubuyama, Japan.

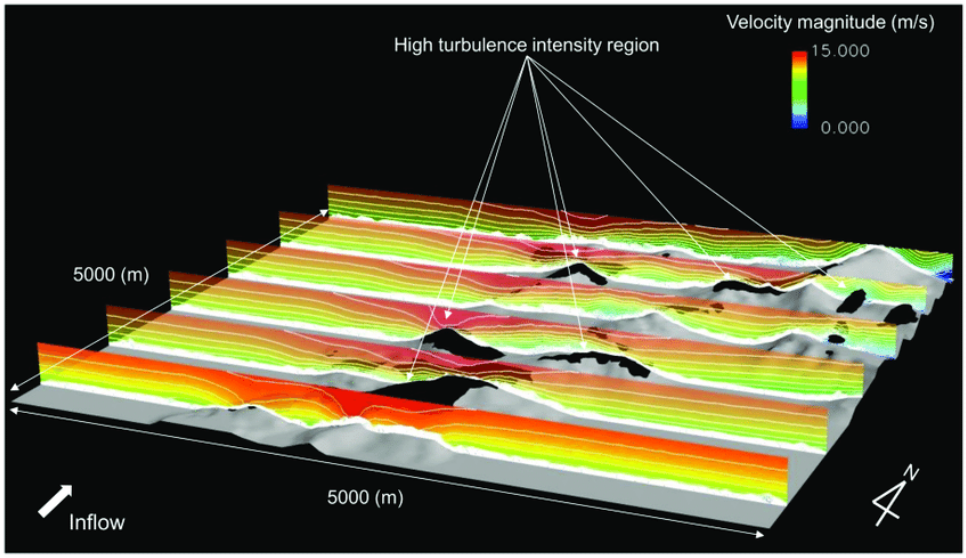


Figure 2.2: Finite-volume method based on a structured computational grid for the computation of Velocity distribution and high turbulence intensity region for NSC [6]

The author highlights the importance of both the matching between the grid and the topographical data in the selection of the simulation boundary parameters, especially the inlet wind direction which has a large effect on the uncertainty of the FCF. Furthermore, in the case of complex terrain sites, the error of estimating the FCF is increased due to flow separation caused by the steep geographical change. The author concludes that using the bin-averaged method, the difference between standard SC and NSC is 13% on average which results in an 8% Measured Annual Energy Production (MAEP) relative error.

2.1.3.B RANS

RANS uses time-based, ensemble-averaged Navier-Stokes ideal gas state equations and focuses on molleding some of the effects from turbulence. Although it yields a lower resolution of analysis than LES, RANS simulations are widely used in engineering applications due to the practical aspect of not

requiring high-resolution calculation grids [6].

Jeong, Jae-Ho and Ha, Kwangtae [6] applied a RANS CFD simulation with turbulence model RNG k-ε to perform a NSC with met mast measurements over complex terrain in Methil, Scotland. The Ansys CFX solver was fed with topographical data coming from WAsP software. The approach was to use the k-ε turbulence model, which accurately analyses turbulence behaviour in free-stream regions. The results show that the uncertainty on the wind speed prediction was about 3.75% concerning the 10 min average single observation of the wind speed at the met mast for the main wind direction.

2.1.3.C WAsP-CFD

WAsP is a program for predicting wind climates, wind resources, and energy yields from wind turbines and wind farms distributed by the DTU (Technical University of Denmark). The software package includes a Map Editor that can transform the Google map data into topographical digitalized input for feeding CFD simulation-based NSC models. It also includes the integrated part 'WAsP CFD' which is an ensemble model approach to compute the FCF that combines WAsP linear wind model calculations based on the Wind Atlas Method [14] for the vertical and horizontal extrapolation of measured wind data and a WAsP non-linear CFD calculation [11].

In WAsP, the role of the flow model is primarily to account for the atmospheric microscales, i.e. scales typically ranging from a few kilometres down to a few meters. Larger scale variations are modelled in WAsP by boundary layer theory to connect wind statistics at different locations and heights (the Wind Atlas Method). The balance between large-scale pressure forces and near-surface winds are assumed to follow the geostrophic drag law. Coriolis force is therefore neglected in the CFD simulations. Also, the different stability conditions are modelled as small perturbations of the neutral stability state, as such, stability effects are omitted from the CFD simulations. Due to the assumption of neutral stability and the omission of Coriolis effects, the modelled wind becomes Reynolds number independent. Thus, the FCF calculated by the WAsP CFD model are dependent on topography only [11].

2.1.4 Numerical Site Calibration assessment

When considering NSC as a wind energy application, the general advantages and drawbacks of CFD simulations can be summarized as follows [11, 16].

2.1.4.A Advantages

- Unlike standard SC, NSC is not dependent on the construction schedule and since it can be conducted at any time, it avoids an early start phase relative to the construction works.

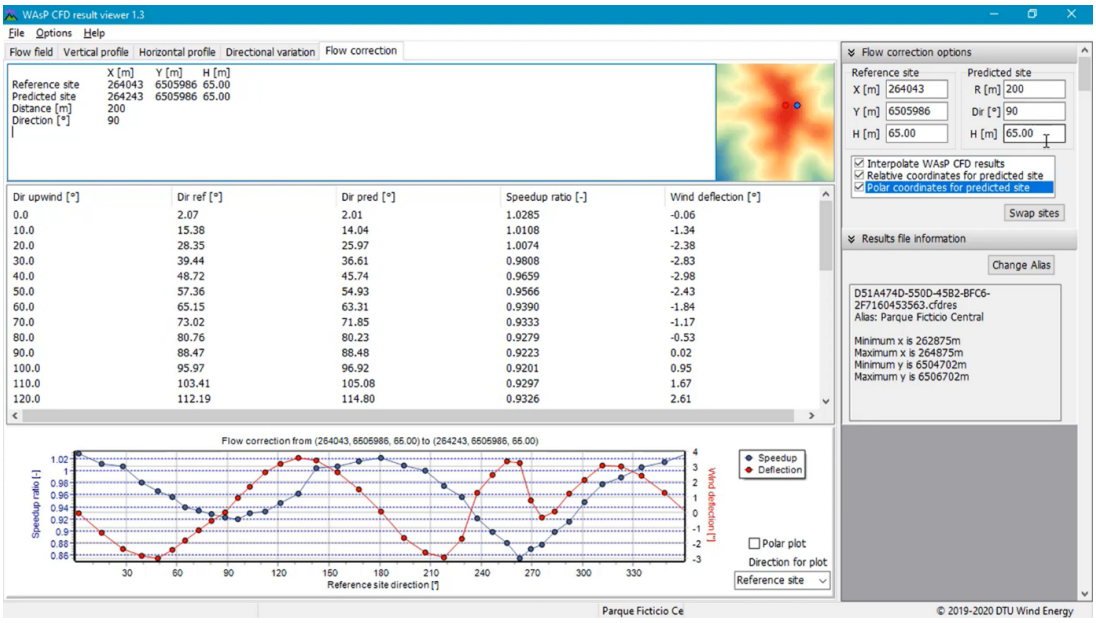


Figure 2.3: WAsP CFD result viewer while computing the Flow Correction Factors [15]

- Although time-consuming, CFD simulations produce results faster than standard SC because they do not require a period of data collection.
- Since no met masts are needed, all complex and time and resource-consuming administrative processes regarding met masts are eliminated, for instance, construction permits and consequence risks and the calibration of the Wind Measurement Equipment (WME).
- Conducting NSC potentially eliminates the need for a TM at the turbine location, CFD simulations are cheaper than a standard SC.

2.1.4.B Drawbacks

- Require good topographical data, mainly terrain and roughness models.
- A mesh-independent solution can be very expensive and time-consuming.
- Since model inputs such as the boundary conditions, can highly affect the final solution, CFD models need to be calibrated, preferably with met masts in the vicinity of the wind farm.
- CFD simulations are expensive from both computational and economical perspectives.
- There is a lack of knowledge on NSC. However, CFD requires trained specialists.
- Conducting NSC leads to high uncertainty in AEP. That makes the Power Curve Verification uninteresting from the customer point of view.

2.2 Remote Sensing

Wind monitoring for energy purposes has today relied on the cup anemometer as the facto standard [17]. A cup anemometer requires mounting on a mast that must be at or near the hub height and is set upwind to the wind turbine at a distance between 2 and 4 rotor diameters [18]. This location must be such that represents the wind flow over the site and ensures no flow distortion caused by obstacles, the wake of other wind turbines, or terrain features [19]. However, with the continuous development of larger wind turbines, the need for higher met masts is adding significant cost and risk to wind projects and thus the need for more flexible methods of wind monitoring.

2.2.1 Remote Sensing description

Remote Sensoring (RS) techniques, such as advanced nacelle anemometry, could also be a solution for accurately measuring the oncoming wind conditions. Nacelle anemometers are devices that are mounted at the top of the nacelle of the wind turbines and thus measure the wind speed and wind direction at the hub height behind the rotor. Although these alternatives are already partly addressed in the draft IEC:2013 standard and IEC:2017 standard where the met mast is still required, they could avoid the need for SC in the future.

The two remote sensing techniques used to date in wind energy applications are LIght Detection And Ranging (LIDAR) or 'laser radar' and SOund Detection and Ranging (SODAR) or 'acoustic radar'. These devices are Doppler effect instruments that utilize, respectively, light and sound transmission and reception from instruments located at ground level. In the case of LIDAR, electromagnetic radiation is reflected off particles, whereas with SODAR a pulse of sound is reflected off the varying temperature structure in the atmosphere [19].

2.2.2 Remote Sensing assessment

For measurements provided by RS to give correct readings, a correction for the effect of rotor disturbance on the oncoming wind field is needed. This requires the use of so-called Nacelle Transfer Functions (NTF). An NTF is an empirical function that relates the undisturbed wind field to the measured wind speed and wind direction at the nacelle. Currently, the NTF is obtained through a series of costly and time-consuming field measurement campaigns, although other numerical techniques based on CFD simulations can be applied [20].

Moreover, because these techniques, by their nature, measure the wind field over a much larger volume than that of a cup anemometer, their application for measuring wind speeds over terrain where large spatial flow variations occur has shown to be less predictable than cup anemometry. Thus, RS devices may not be suitable for measurements in complex terrain [19].

2.3 Machine Learning in the wind field

"The wind industry is drowning in data" [21], since it has an overwhelming volume of sensor measurements (T, PR, rotor speed, wind speed, etc.) and images or videos (blades, welds, cable routes, etc.). And data is one of the main ingredients required to effectively deploy ML.

The second main ingredient is to find the appropriate domain challenges, thus asking the right questions about data. Some of the applications of ML to the wind field are exposed in this subsection as well as the two most relevant topics for this Thesis: wind speed prediction and turbine performance evaluation.

2.3.1 Machine Learning applications to the wind field

ML-based algorithms powered by big data available in SCADA systems have been proven to solve critical problems in many different aspects of the wind industry. M. Lydia and G. Edwin [22] classify the ML applications on wind turbines in three big groups: Modelling, Monitoring and Prediction.

Modelling using ML can be applied to very specific wind farms aerodynamic modelling such as wake models, acoustic emission, and atmospheric turbulence. On the other hand, it can also be applied to wind turbine modelling, for instance, PC models or load flow distribution models.

Using ML techniques for Condition Monitoring (CM) of wind turbines is a vital aspect of operation and maintenance. It has been proved useful to predict the failure of generator bearing or detect blade faults, for instance. Other explored areas for Monitoring are PC and performance monitoring.

Lastly, the Prediction subclass includes predictive control aimed to maximize the turbine's power generation at a minimized thrust force. Wind power prediction is implemented in time intervals to estimate the variability in power system planning and control. Moreover, wind energy generation forecasts may be used for improving the integration of wind power into the grid or reducing the uncertainty in project investments [23–28]. Wind speed prediction is discussed in deep in the following section 2.3.2.

2.3.2 Wind speed prediction with Machine Learning

The challenges caused by non-stationary, non-linearity and non-Gaussianity in wind energy can be effectively handled using ML algorithms [22]. Thus, through an exhaustive literature search, it has been found that different ML techniques have been implemented for wind speed prediction.

The common methodology consists in using Numerical Weather Prediction Systems (NWPS) and met masts data for powering supervised ML regression models. For the input features, the most common atmospheric variables used are wind direction, temperature, pressure and Rain, among others. These data resolutions might be from minute to six-hour-based data. And the main ML techniques used are different types of Artificial Neural Network (ANN) and Ensemble models based on Decision Trees (DT)s.

Among the purposes of implementing supervised ML regression algorithms for wind speed prediction are improving the accuracy of prediction for wind energy systems or reducing the influence of intermittent wind power on system operations and other tasks related to climate classification and severe weather emergency preparedness [29–34].

2.3.3 Site Calibration with Machine Learning

Some documentation found during the literature research was aimed at applying ML techniques for power performance evaluation at the wind turbine level [27]. Regarding the SC problem, earlier in 2001, J.P. Verhoef and G.P. Leendertse [35] had already pointed out the necessity of exploring different ML regression techniques for SC.

The report on Task 8 (A literature survey on theory and practice or Parameter Identification, Specification and Estimation (ISE) techniques) of the project SiteParIden: "Identification of Variables for Site Calibration and Power Curve Assessment in Complex Terrain" (JOR3-CT98-0257) from the ECN (Energy research Centre of the Netherlands) is public and available on TNO (Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek, in English: Netherlands Organisation for Applied Scientific Research) website.

This report states that some ML multivariate regressions models are techniques that should be able to cope with the PC assessment and SC for complex terrain problems. This report specifies Linear Regression, Ridge Regression and the non-linear modelling ANN as highly recommended methods for SC. Other suggested techniques are Factor Analysis and Principal Components, Singular Value Decomposition and Multi-binning.

Despite that ANNs were recommended for solving the SC problem 20 years ago, this Master Thesis researcher could not find any report, article, or document regarding the application of ANN or any other ML technique to SC. Nevertheless, it is important to bear in mind for instance, that although ANN started to be an area of deep research in the middle 70s, it was not until March 2015 when Keras, the open and high-level Deep Learning software for building ANN on python, was released for public use by its developer François Chollet [36].

3

Problem Statement

Contents

Chapter 3 presents a detailed analysis of the limitations and inconsistencies of the standard IEC Site Calibration procedure as well as an analysis of the possible impact of an inaccurate Site Calibration on the final result of a Power Curve Verification test.

3.1 IEC standard Site Calibration

The standard SC is a procedure defined by the IEC which should be applied to any terrain labelled as 'complex' that exceeds a pre-defined maximum slope or a maximum terrain variation from the plane, as described in Annex B of IEC 2005 [18]. The PM should be installed at a distance between 2 and 4 times the rotor diameter (5.2.2 Measurement sector IEC 2005 [18]) and the period of data collection for the SC should be such that that minimum required SC data is collected (C3. Data acquisition and analysis IEC 2005 [18]).

3.1.1 IEC 2005 & 2017

Annex C of IEC 61400 – Part 12-1: 2005 and 2017 standards editions are used as a guide to conduct a SC. The transfer function that converts the reference wind speed into the target wind speed is a table of FCF that are computed according to Eq. 3.1:

$$FCF_k = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{WS_{\text{target}}}{WS_{\text{reference}}}\right)_k \tag{3.1}$$

The FCF are computed as the average wind speed ratio for the n observations may depend on certain a k atmospheric variable bin depending on the IEC edition. For instance, k corresponds only to Wind Direction (WD) bin for IEC 2005 edition.

3.1.1.A Edition 2005 FCFs computation

Section C1 'General' of the IEC-61400-12-1 2005 standards considers that "Terrain and obstacles may cause a systematic difference in wind speed between the position on the met mast where the power performance anemometer is mounted and the centre of the turbine rotor". Thus, SC based on the IEC 2005 consists of a simple linear regression between the reference wind speed and the wind speed at the turbine location at the hub height. The key result is a table of FCF for the relevant WD bins. As presented in Eq.3.1, the FCF for each WD bin is just the average of the ratios between the reference wind speed and the wind speed at the turbine location at the hub height for that WD bin. And thus, the resulting predicted wind speed is computed as shown in Eq. 3.2

$$WS_{\text{target}} = FCF_{WDbin} \cdot (WS_{\text{reference}})_{WDbin} \tag{3.2}$$

3.1.1.B Edition 2017 FCFs computation

The 2017 version of the IEC standards introduces the idea of the ratio being dependent on the Wind Shear (WSH) and the reference wind speed itself. 2017 IEC defines a method to assess the shear conditions at the site. Thus, if this method concludes that the WSH is significant for the behaviour of the ratios at the site, the FCF may be computed dependent on the WD, wind speed and the WSH bins. On the other hand, if the assessment method is finalised and the conclusion is that the WSH is not significant, the FCF are computed dependent only on the WD and the wind speed bins.

3.1.2 IEC inconsistencies

Up to 5 inconsistencies have been found regarding the IEC standard-based SC.

3.1.2.A WD binning

One of the most key steps in 2005 IEC SC might be selected by the engineer from the Independent Accredited Laboratories (IAL) who performs the SC. A different WD binning criteria might be chosen from a wind turbine to wind turbine even if those are on the same wind farm and belong to the same project. The only guide from the IEC 2005 section C.3 'Data acquisition and analysis' regarding how to bin the WDs is "the data sets shall be sorted into wind direction bins. Each bin shall be no larger than 10°. The wind direction bin should be not less than the uncertainty of the wind direction sensor." The methodology on how to determine the reference direction and bin step is neither guided nor recommended.

3.1.2.B WD bin and ratio dependency

The IEC 61400-12-1:2005 standard assumes a direct relationship between the WD bin and the wind speed ratio. However, this has been deeply analysed in this Thesis, and as can be seen in Appendix D, although this correlation might be slightly visible in some cases, it is not always clear enough.

3.1.2.C Seasonality effect

In section C.1 'General' from IEC 2017, the following is ensured: "Seasonal considerations: atmospheric stability, Turbulence Intensity (TI) and WSH can be related to different seasonal conditions. There are also concerns about the effects of changes in roughness due to changes in the vegetation in the testing area or other roughness changes directly caused by different seasonal surface characteristics (water/-land vs. ice/land, snow, crops, etc.). In the light of these considerations, the SC and Power Curve Measurement campaign should be conducted during the same season or seasons". However, this is just an IEC recommendation and due to the project delivery schedule, it is usually not the case.

3.1.2.D Data filtering

Data rejection specified in the standards IEC 2005 section C.3 'Data acquisition and analysis' include icing filter and "special weather conditions". On the one hand, the icing filter is defined as "failure or degradation (e.g. due to icing) of test equipment". Without any more guidance, the IAL might apply their own icing filter criterion. On the other hand, the special weather conditions filter is defined as "any other special atmospheric conditions that will also be used as rejection criteria during the power performance test." Again, since this filtering is not specified, it might change from one site to another and each IAL might apply their own criterion.

3.1.2.E Data quality control

The first step before any data mining procedure is to apply a correct data quality control to the corresponding datasets. The IEC standards do not determine at any instance which kind of filters and for which atmospheric signals they should be applied, thus the corresponding IAL might apply their own data quality control criterion.

3.1.3 IEC limitations

Up to 3 limitations have been found regarding the IEC standard-based SC.

3.1.3.A Multi-binning methodology

Multi-binning itself, of WD or any other atmospheric variable, might represent a limitation for SC. For instance, when discretizing the WD in steps as large as 10°, a significant amount of probably very useful information is being neglected and a higher uncertainty on the FCF is being caused.

A clear side effect of this methodology can be seen in the case when the differences in neighbouring FCF are very high. In C.5 'Selection of final measurement sector' of IEC 2005 recommends the following: "In addition, the correction factors may change abruptly between WD bins. It is recommended that WDs from the measurement sectors be eliminated when flow correction factors change by more than 0.02 between neighbouring sectors".

The elimination of those WD bins might lead to no FCF for a specific wind turbine, and thus, the IAL be obliged to discard the wind turbine from the SC procedure as well as for the Power Curve Measurement campaign. Nonetheless, the investment made by the project owner, required for erecting the two necessary met masts to perform the SC, would be already made.

3.1.3.B Atmospheric effects

In C.1 'General' from IEC 2017, the following is ensured: "The relationship between the reference met mast wind speed and the wind speed at the turbine position may also be affected by changes in atmospheric stability and/or the shear profile". However, IEC 2005 computes the FCF depending only on the WD, not taking into consideration any other atmospheric variable.

Despite IEC standards' new version of 2017, IEC 2005 standard is broadly accepted and is the base of more than 97% of the past and actual PC warranties of the wind turbine supplier Vestas. That is the main reason why this Thesis will use as a reference the IEC 2005 version.

Nevertheless, the 2017 IEC version only includes the WSH bins as a variable for computing the FCF. These standards do not consider any other atmospheric variable for instance TI or Wind Veer, which also affect the atmospheric stability and have a high impact on complex terrain climate.

3.1.3.C Upper tip height met masts

The SC IEC procedure is given for the wind speed defined as the hub height wind speed. This is so that the procedure does not mandate upper tip height met masts because they are expensive and may be impractical. However, a clear limitation for wind speed prediction might be the fact that the PM is only able to record information until the hub height. When computing the transfer function that translates the reference wind speed into the target wind speed there is no data available on what is the wind profile above the actual met masts height.

3.2 Wind projects warranty

In the wind industry, particular attention is paid to the AEP, since AEP is a fundamental parameter influencing the revenues of a project. Thus, the final key variable that is warranted in a Warranty Contract is the Warranted Annual Energy Production (WAEP).

3.2.1 Power Curve Measurement

A major concern for the AEP is the PC of the wind turbine which expresses the relationship between the wind speed at the hub height and the power output. In this sense, the most broadly accepted PPT for verifying the correct performance of wind turbines and assessing the WPC is the Power Curve Measurement campaign. It consists in measuring the power output of the wind turbine after its commissioning and the result is the MPC.

The Power Curve Measurement campaign in a new wind farm is a contractual issue agreed upon between the Turbine Supplier and the owner and consists in measuring the power output of the wind

turbine after its commissioning. On a Turbine Supply Agreement (TSA) between the parties, the PC provided by the manufacturer, which is a standard PC used as a reference, is the actual WPC. The WPC of a turbine is validated through a Power Curve Measurement campaign under a specific methodology as described in the contract and is mainly based on the IEC 61400-12-1:2005 standard and executed by an IAL.

The validation is done by simultaneously recording the wind speed at the PM and the wind turbine power output. Once the data is collected, data normalization to a reference air density should be performed according to the IEC 61400-12-1:2017. The reference density should be the mean of the average of the measured air density of the collected data during the PPT or a pre-defined nominal air density representative of the site. The normalization should be applied to the wind speed according to Eq. 3.3:

$$V_n = V_m \left(\frac{\rho_m}{\rho_o} \right)^{\frac{1}{3}}, \tag{3.3}$$

Where V_n is the normalized wind speed, V_m is the measured wind speed, ρ_m is the measured air density and ρ_o is the reference air density.

On the other hand, the normalization should be applied to the measured power according to Eq. 3.4

$$P_n = P_m \left(\frac{\rho_m}{\rho_o} \right), \tag{3.4}$$

Where P_n is the normalized power output and P_m is the measured power output.

Once the data is normalized, the wind speed is binned and the average power for each wind speed bin is computed. The result of this computation is the MPC.

3.2.2 Warranty Contract

In the wind industry, a PPT is also known as Power Curve Verification, or Power Curve Validation. It is a type of warranty contract through which the performance of individual wind turbines is warranted. Warranty contracts are legally binding assurance that wind turbines meet the standards and play an important role for both parties signing them.

From the owner perspective, a Power Curve Measurement campaign is a verification and evaluation method of the future sustainability of the wind project. It also provides a fine-tuning of the energy production estimations, which substantially removes business uncertainty by clarifying the wind farm profits over its lifetime operation. From the wind turbines manufacturer (e.g. Vestas) point of view, the Power Curve Measurement campaign is first, a requirement for certification but also provides differentiation and a transparent demonstration of their products, which highly strengthens the firms' reliability and enhances their customers' trust.

3.2.2.A Contract filters

Contracts define the individual measurement conditions for which the WPC is valid. Those filters can include any range for any atmospheric variable that the turbine supplier considers. Some of the most common filters are the following:

- **TI filter:** Depending on the topography and the atmospheric stability, the TI can increase or decrease the MPC.
- **WSH filter:** The wind speed profile concerning the height can have a significant impact on the power output.
- **Inflow Angle (WSVer) filter:** In complex terrain, higher inflow angles are expected. Therefore, the WSVer component is the vertical component of the wind speed at the hub height.

Wind turbine rotors are several meters long and when the difference in wind conditions between the upper and the lower tip is significant, the performance of the wind turbine is affected. Since the PC warranties currently only rely on the wind speed at the hub height, all these filters in a warranty contract aim at reducing the accepted range of conditions for which the WPC provided by the turbine supplier is valid.

3.2.2.B Warranty contract formulation

Usually, a warranty contract stipulates that the MAEP of the wind turbine calculated as the product of the MPC multiplied by the Wind Speed Distribution (WSD) shall be equal or greater than WAEP calculated as the product of the WPC multiplied with the WSD and multiplied by the Uncertainty Adjustment factors.

The WSD of the project represents the average hub height WSD as the number of hours over a year. The equations for AEP (Measured and Warranted) are:

$$MAEP = MPC \cdot WSD \tag{3.5}$$

$$WAEP = WPC \cdot WSD \cdot (1 - u_{\text{AEP}}) \tag{3.6}$$

Where u_{AEP} is the uncertainty in the AEP, which although it is usually formulated and computed in warranty contracts, is not addressed in this Thesis. An outline of the entire verification procedure is presented in Fig. 3.1.

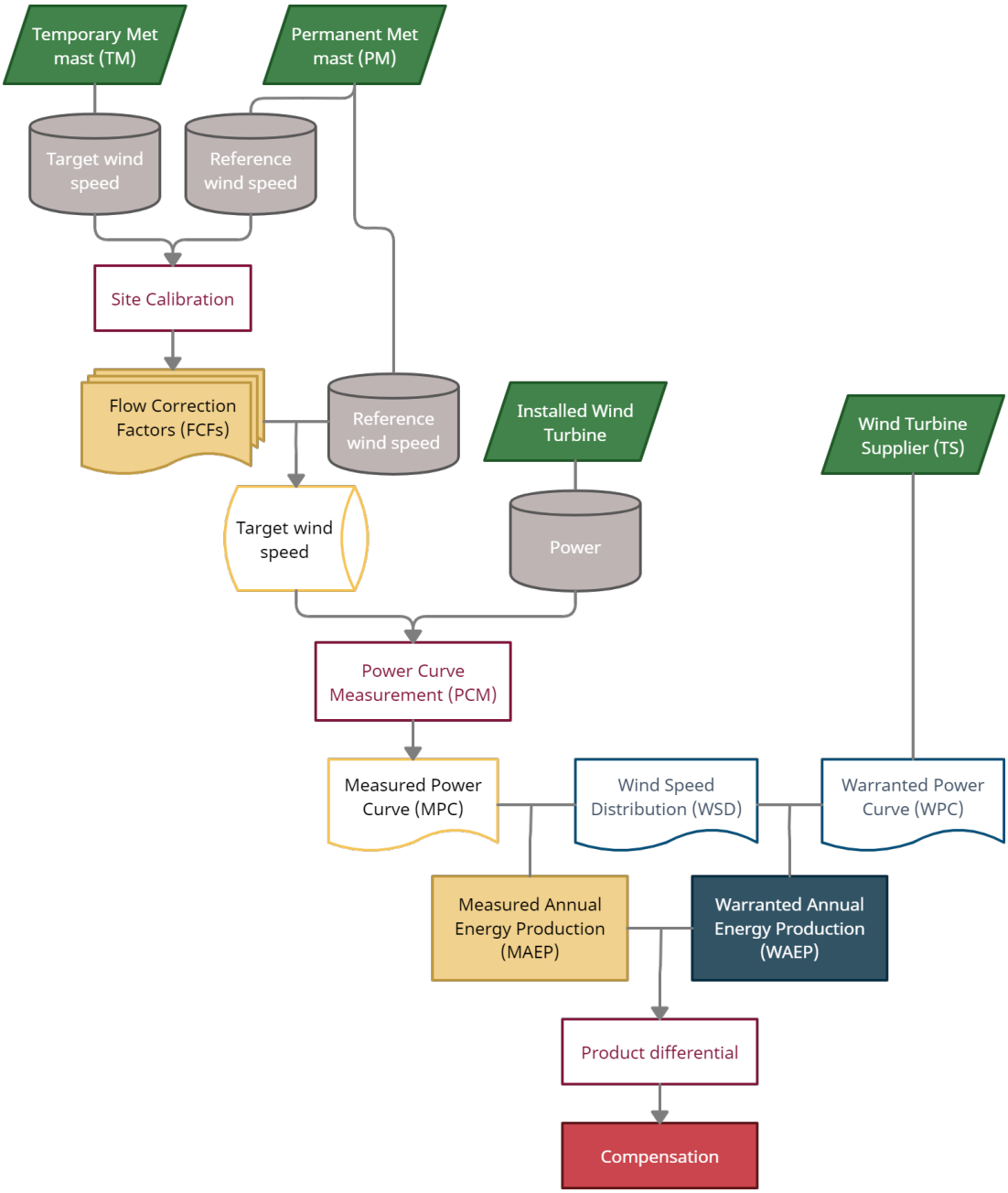


Figure 3.1: Flowchart of the Power Curve Verification procedure (adapted from [37])

Within the contract warranty, the WPC is found and is used as an indicator of the turbine performance. It is multiplied by the WSD to estimate the WAEP. When the WAEP is not met, the compensation rules apply.

3.2.2.C Compensation

In some cases, a PC warranty may stipulate compensation on behalf of the wind turbine supplier in case of wind turbine underperformance. The amount of compensation is also stipulated as to 5% of the wind turbine purchase price. The warranty period for which this compensation applies is in the range of 2 to 5 years.

3.3 Inaccurate Site Calibration

Although many recent articles and technical documentation indicate that there is ongoing research on alternative SC methodologies, no literature was found regarding the impact of an uneven SC on the result of a PPT. Nevertheless, after several years of investigation, the Power Curve Verification department of Vestas has found enough evidence for this, and hence the core problem of this Master Thesis. The present subsection aims to expose the possible impact of an inaccurate SC on the result of a PPT.

3.3.1 Site Calibration impact

As explained in section 3.1, the actual standard SC methodology has some limitations and uncertainties that can generally lead to an increase in the error in the wind speed regression between the PM and the TM.

As seen in Fig. 3.1, SC plays an important role when estimating the MPC in complex terrains. Through the FCF calculated during the SC procedure, together with the reference wind speed measured during the Power Curve Measurement campaign, the target wind speed is estimated.

Thus, if an inaccurate target wind speed is estimated, an uneven SC can transfer a significant prediction error to the wind speed binning used for estimating the MPC. The risk of compensation on behalf of the wind tubine supplier increases significantly when the target wind speed is over-predicted for a wind speed since it will lead to a lower MPC. This increase is especially risky when the wind speed over-estimation is that to the wind speed corresponding to power below the rated power because this error is transferred to the MPC to the third power, 3.7:

$$P = \frac{1}{2} C_p \, \rho \, \pi \frac{D^2}{4} U^3, \tag{3.7}$$

Where C_p is the power coefficient, around 0.593. ρ is the air density in kg/m^3, D is the wind turbine rotor diameter in meters and U is the wind speed at the turbine height in m/s.

3.3.2 Measurement of Site Calibration impact on the Measured Power Curve

A method was developed by the Department of Power Curve Verification of Vestas to quantify the impact of an "uneven" SC procedure in energy terms. It consists in obtaining a 'Site Calibration- Power Curve (SC-PC)' from the SC data.

First, since the methodology consists in computing a MPC, but without actually using the actual power output, both target and predicted wind speeds are normalized regarding the air density as presented in Eq. 3.3.

For the wind speed data, the predicted wind speed at the turbine location is binned and used for the computation of the MPC. For the IEC method, the predicted wind speed is computed as the reference wind speed multiplied by the corresponing FCF depending on the WD bin, as presented in Eq. 3.2.

And for the power values, the target wind speed at the TM is used. For each measured wind speed readings at the turbine location, the power is calculated by the mean of interpolation using the WPC as the reference and it is afterwards binned regarding the wind speed bins.

Finally, combining the predicted wind speed at the turbine location and the power derived from the target wind speed at the turbine location, the SC-PC can be obtained [38]. Fig. 3.2 shows an example of this method for the wind turbine WTG14 analysed on Dataset1, only for the test set of the data (please refer to section 5.2) using the IEC 2005 reference method.

Represented in Fig. 3.2 as the blue scatter on the x-axis, the predicted wind speed by the IEC method is shown, while on the y-axis the power calculated based on the turbine mast location measured wind speed is shown and can be considered as the "expected power". Also, the black dotted line represents the WPC provided by Vestas. While the red dotted line is the SC-PC which represents the SC error.

When the computed FCF over-estimate the wind speed, the blue scatter points are shifted to the right and the red line SC-PC is forced downward. Thus, it can be stated that the standard SC error is transferred to the wind speed prediction binning of the MPC, and for a SC error equal to zero, black (WPC) and red lines (SC-PC) should be equal.

Both Measured and Warranted AEP are computed using Eqs. 3.5 and 3.6 and their difference is used for measuring the error in energy terms through Eqs. 3.8 and 3.9.

$$AEP_{\text{diff}} = MAEP - WAEP \tag{3.8}$$

$$AEP_{\text{percentage}} = \frac{MAEP}{WAEP} \cdot 100 \tag{3.9}$$

For $AEP_{\text{diff}} < 0$ or $AEP_{\text{percentage}} < 100\%$, the SC procedure would increase Vestas' risk of compensation.

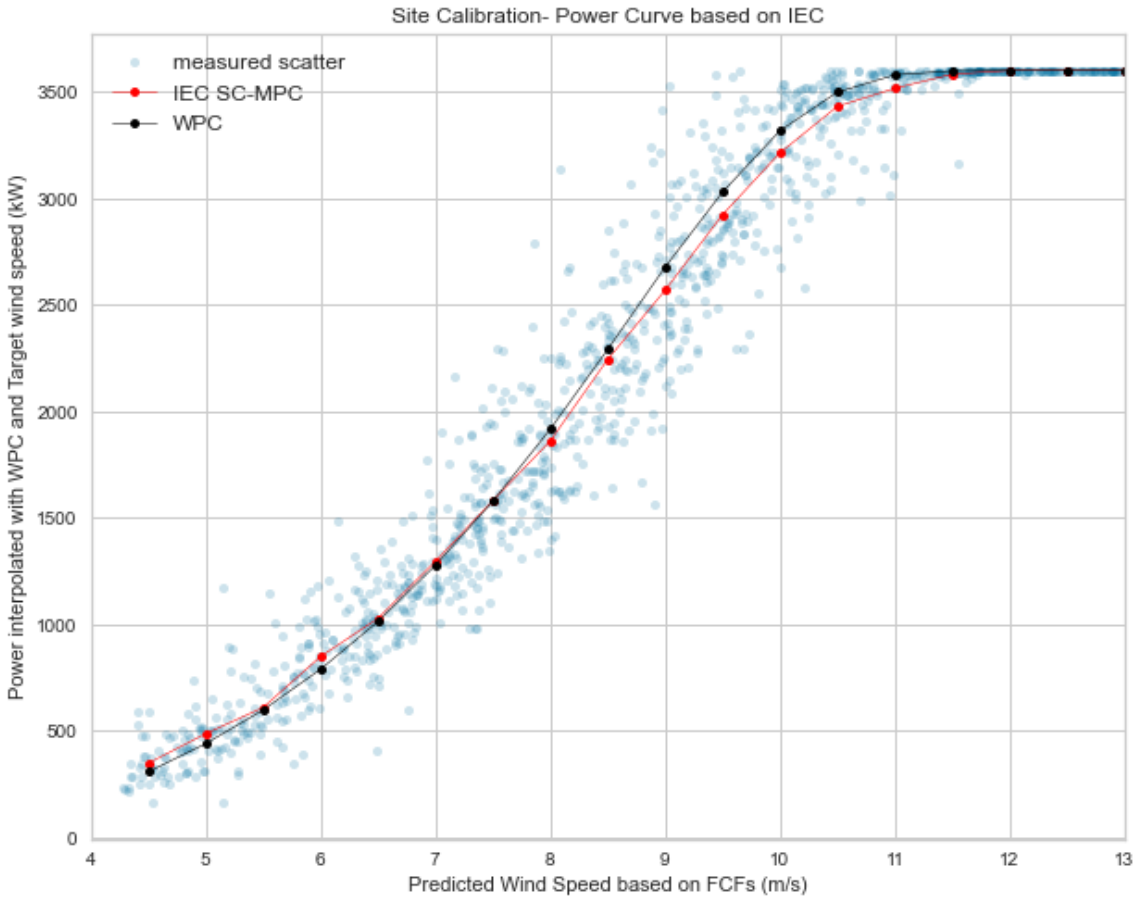


Figure 3.2: Site Calibration-Power Curve of IEC for WTG14

4

Data Analysis

Contents

Chapter 4 presents a full analysis of the data used for modelling. In this chapter, the complete procurement process data is explained. First, how data is collected by the wind measurement equipment mounted on the meteorological masts as well as the computation of the post-processed signals. In addition to that, the methodology for data mining is shown and finally, the basic information regarding the datasets is presented.

4.1 Meteorological data

This section aims to provide all necessary information regarding the meteorological variables collected by the met masts during the SC period which will be used as input variables (from now on, input features) for the ML modelling which corresponds to section 5 of this Thesis.

4.1.1 Wind variables

Fig. 4.1 shows an outline of the different wind variables and a complete description is given as follows:

Figure 4.1: Wind variables sketch [37]

- **Wind speed** - Wind speed, or wind flow speed, is a fundamental atmospheric quantity in meteorology caused by air moving from high to low pressure, usually due to temperature changes. In this Thesis, WS is the acronym for Wind Speed.
- **Air Density** - The density of air or atmospheric density is the mass per unit volume of atmospheric air. Air density, like air pressure, decreases with increasing altitude. However, it also depends on the air pressure, temperature and humidity. In this Thesis, AD is the acronym for Air Density.
- **Turbulence Intensity** - Turbulence Intensity is defined as the ratio of the standard deviation of fluctuating wind velocity to the mean wind speed, and it represents the intensity of wind velocity fluctuation. In this Thesis, TI is the acronym for Turbulence Intensity.
- **Wind Shear** - Wind Shear sometimes may be referred to as wind gradient, is a difference in wind speed over a relatively short distance in the atmosphere. Wind shear is a microscale meteorological phenomenon occurring over a very small distance. In this Thesis, Wind Shear is referred to as

vertical wind shear, thus it is considered as the difference in wind speed with a change in height in a logarithmic scale. In this Thesis, WSH is the acronym for Wind Shear.

- **Wind Veer** - Wind Veer in meteorology terms is a difference in wind direction with a change in height. In this Thesis, WVeer is the acronym for Wind Veer.
- **Inflow angle** - The Inflow angle is the direction in degrees of the vertical wind speed component. That is why Wind Veer is also known as 'vertical wind direction'. In this Thesis, WDVer is the acronym for Inflow angle.

The formulas used for computing the post-processed variables are presented in Appendix B.

4.1.2 Wind Measurement Equipment

The met mast is a tower made of steel on the top of which the WME is mounted. As can be seen in Fig. 4.2, ideally, the met mast is as high as the wind turbines that will be commissioned in the wind farm, because met masts are expensive and thus, higher met masts can increase the cost exponentially.

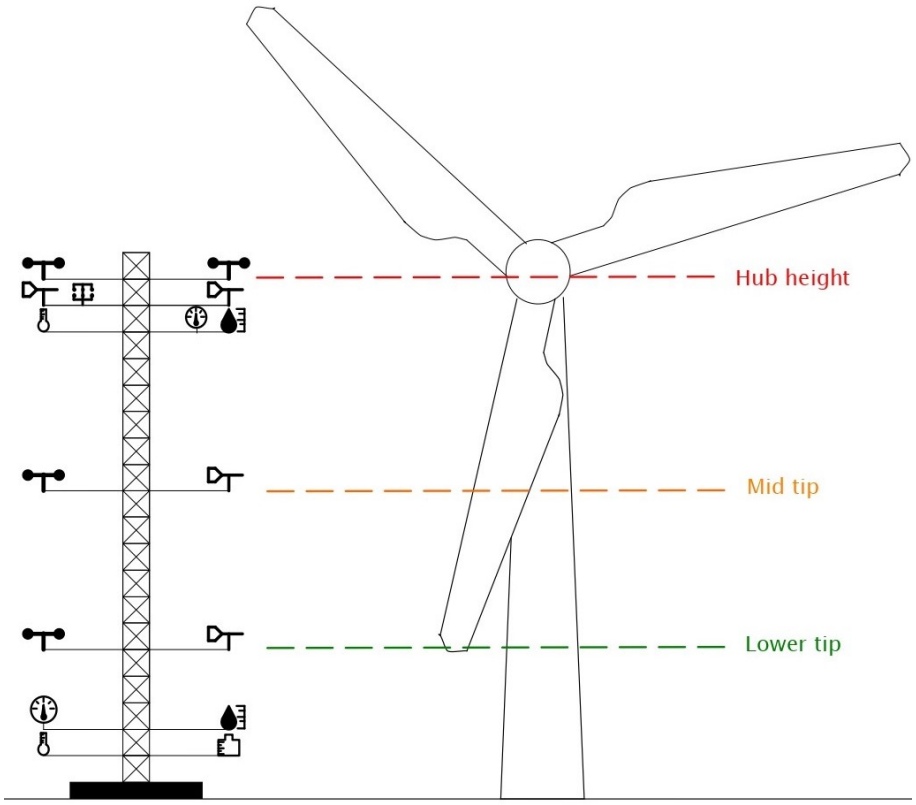


Figure 4.2: Front view of a met mast with the Wind Measurement Equipment and a wind turbine. The three different main levels are identified with colours. (Not at scale)

Each measuring device is connected to a channel of the digital SCADA system which collects the data continuously at a sampling rate of 1 Hz (every second) for most of the signals. Air temperature, air pressure, wind turbine status and precipitation may be sampled at a lower rate [18]. The SCADA system either stores the pre-processed sampled data, or directly stores the following statistics of the datasets in a base of 10-minute periods: mean, standard deviation, maximum and minimum. Each of these points in the dataset will be named “an observation” from now on.

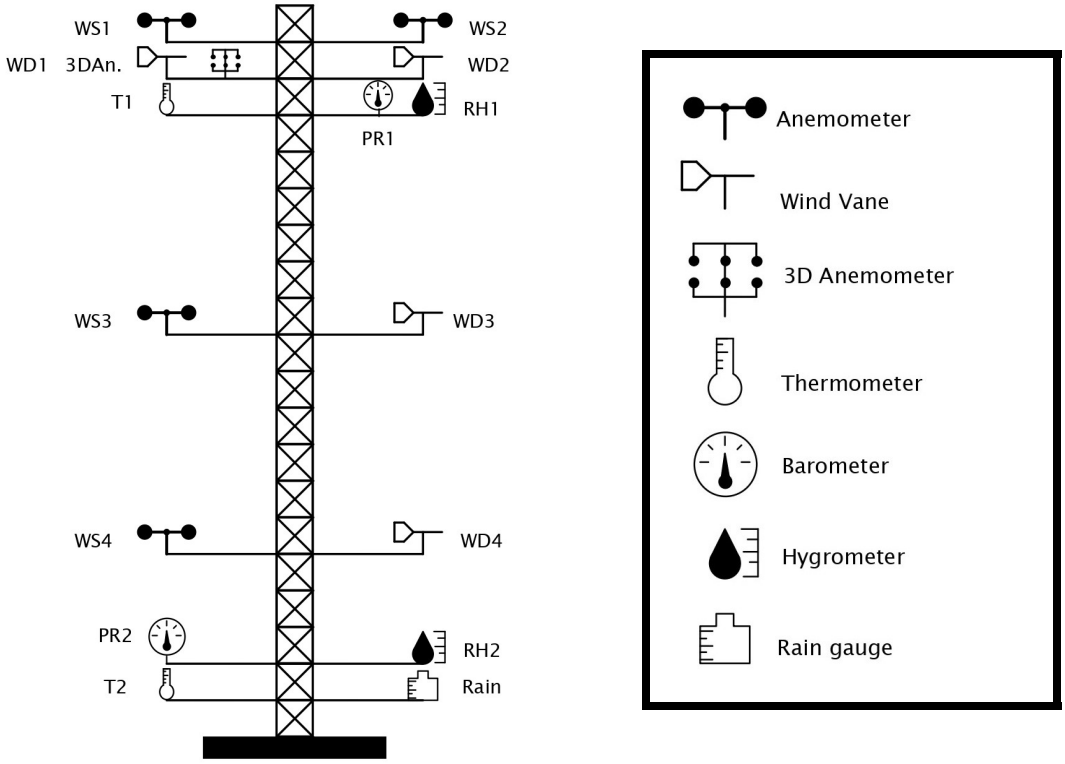


Figure 4.3: Meteorological mast with the Wind Measurement Equipment (Not at scale)

The WME mounted on the met masts is usually the following:

- **Anemometers:** Usually, cup anemometers, measure wind speed and are installed at 3 different levels: hub height, mid tip, and lower tip levels. At hub height, two different anemometers are mounted. The primary anemometer and a second anemometer which is named control anemometer.
- **Wind vanes:** Measure wind direction from 0 to 360° and are installed together with their corresponding anemometers.
- **3D anemometers:** also called ultrasonic anemometers, they measure the horizontal and vertical components of the wind speed and wind direction.

- **Thermometers:** that measure the temperature of the air are usually close to the highest point of the met mast (located within 10 m of the hub height) and at the ground level. The thermometer at the hub height should represent the temperature at the wind turbine rotor centre-line.

- **Barometers:** that measure the air pressure are usually installed at two different levels and together with the thermometers.

- **Hygrometers:** measure the relative humidity in the air and are installed at two different levels and together with the thermometers.

- **Rain gauge:** measures the daily volume of the precipitation and is usually installed at the ground level.

4.1.3 Input features

The wind variables and other meteorological variables measured by the WME on the PM will be used as the input data for the different ML models implemented. In ML, each variable used as input for modelling is known as 'feature'. This Master Thesis adopts the ML terminology and from now on, the meteorological variables used for modelling will be named as input features. Table 4.1 shows those input features and the resource sensors.

Table 4.1: Sensors and Input features (variables)

Sensors	Input features (variables)
WS1	WS1, TI, WSH
WS3	WS3
WS4	WS4, WSH
WD1	WD1, WD_bin, WVeer
WD3	WD3
WD4	WD4, WVeer
T1	T1, AD1
T2	T2, AD2
PR1	PR1, AD1
PR2	PR2, AD2
RH1	RH1, AD1
RH2	RH2, AD2
3D anemometer	WSHor, WSVer, WDHor, WDVer
tod	tod (time of day)
Rain	Rain

Some of the variables may have been discarded for certain wind turbines due to measurement inconsistencies.

4.2 Data mining

This section aims to describe the methodology used for pre-processing and preparing the data for modelling. First, data pre-processing is a data mining technique that involves transforming raw data into an understandable format, thus, after pre-processing, the data is ready to be explored. Secondly, data filtering consists in applying the corresponding filters recommended by IEC. Finally, data conditioning is the previous step prescribed in the good practises of ML modelling.

4.2.1 Data pre-processing

Figure 4.4 shows the main steps in data pre-processing:

Figure 4.4: Flowchart for data pre-processing

4.2.1.A Data correction

Some of the variables may need to undertake some data corrections before any further pre-processing is performed. The following corrections were applied for uniformity reasons:

- **Inflow angle:** in case the vertical wind direction was ranged between 0 and 360°, it would be corrected to range between -180° and +180°.
- **Turbulence Intensity:** in case the TI was scaled to 100, it is corrected as to be ranged between 0 and 1.

4.2.1.B Data cleaning

Data cleaning is the pre-processing technique that deals with the missing data. For each of the variables in the dataset, the data availability for the p variables is computed using the Eq. 4.1:

$$availability_j = \frac{\sum_{j=1}^{p} observations \text{ (obs=null)}}{\sum_{j=1}^{p} observations} \tag{4.1}$$

The threshold of 0.95 of data availability is set. Lower than 0.95 will be considered as low availability. Thus, data cleaning consists in either dropping those observations with missing data or, in case of low

availability of one or more variables, the missing values will be replaced by an estimation. These estimations are computed through interpolation methods by the use of other variables that can be considered equivalent. This small correction is performed by the correction factor for the target variable which is computed as the average ratio between the target variable and the auxiliary variable. The missing values of the target variable are substituted by the auxiliary value observation multiplied by the correction factor. (e.g., if PR1 has a data availability of 0.88, the average ratio between PR2 and PR1 will be used to fill the missing values).

4.2.1.C Data quality control

Data quality control may include several steps for error detection depending on the field to which this method is applied. In this case, data quality control will rely on detecting the faulty values for each variable defined by a range of quality and dropping the corresponding observations. As shown in table 4.2 he ranges of quality are defined for each of the variables, based on the criterion of the basic physics laws:

Table 4.2: Data quality control applied to raw met masts data

Variable	Unit	Quality range	Signals
Wind Speed	m/s	[0, 40]	WS1, WS3, WS4, WSHor, Target
Wind Speed component	m/s	[-40, 40]	WSVer
Wind Direction	°	[0, 360]	WD1, WD3, WD4, WDHor
Wind Direction component	°	[-180, 180]	WDVer
Temperature	℃	[-40, 60]	T1, T2
Pressure	hPa	[500, 1300]	PR1, PR2
Relative Humidity	%	[10, 102]	RH1, RH2

4.2.1.D Outliers' detection and removal

In statistics, an outlier is an observation point that is distant from the other observations. The outliers can be a result of a mistake during data collection, or it can be just an indication of variance in the data. However, in ML methodology, it is usual to apply an outliers' detection and removal procedure, since ML algorithms might be sensitive to data's noise originated by outliers.

The outliers' method applied is the z-score method [39]. The z-score is the signed number of standard deviations by which the value of an observation is above or below the mean value of what is being measured. The usual threshold is 3 in absolute terms. If an observation has a higher z-score than 3, it is identified as an outlier and it is dropped from the dataset [40].

4.2.1.E Data exploration

A complete data exploration is performed. These are the selected plots:

- **Wind rose:** A wind rose is plotted for WD1. It helps to identify the predominant wind directions [41].
- **Histograms:** histograms of the meteorological variables help on the identification of the range of each variable and its distribution. These plots are also relevant for SC since very often, SC is not performed on the same season that Power Curve Measurement campaign. However, one important ML assumption is that those variables used for the training set (SC dataset) need to reflect the usual distribution of the datasets that will be used for predicting (Power Curve Measurement campaign dataset). Thus, these histograms should be compared to those of the Power Curve Measurement campaign and possibly determine the uncertainty of the model due to the distribution of each variable [42].
- **Daily cycle plot:** A daily cycle plot for the meteorological variables may help to detect outliers as well as for identifying the patterns of each variable [42].

The results of the exploration for the 3 datasets can be seen in Appendix C.

4.2.2 Data filtering

Different filters may be applied in the actual SC procedure. Two major groups can be distinguished: IEC recommended, and turbine supplier determined. The filters determined by the turbine supplier, which are included in the TSA clauses, will not be applied to the reference method since a pure IEC SC is intended. The summary of the filtering applied is presented in table 4.3.

Table 4.3: Summary of data filtering applied by method

Filter	Variable	Range	Vestas' method	Reference method
Measurement sector	WD1	site-dependent [1]	✓	✓
Operational range	WS1	[4 , 16] m/s	✓	✓
Logger counts	HeartBeats	600	✓	✓
Icing	RH1, T1	T1=2℃ or RH1=80 %	✓	✓
Uncomplete WD_bins	WD_bin	[2]	✓	✓
Turbulence Intensity	TI	[6, 12] %	✓	✗
Wind Shear	WSH	[0, 0.3]	✓	✗
Inflow angle	WDVer	[-2, 2] °	✓	✗
Rain	Rain	0	✓	✗

[1] For more information, please check section 4.3.1

[2] IEC 61400-12-1:2005 suggests that in order to apply the FCFs, a minimum site calibration data set shall consist of 24 h of data for each non-excluded wind direction bin. Also, of these, each bin shall have at least 6 h of data where winds are above 8 m/s and at least 6 h of data where winds are below 8 m/s. This filter is implemented just to obtain exactly the same dataset as the IEC method.

4.2.3 Data conditioning

Figure 4.5 shows the main steps in data conditioning:

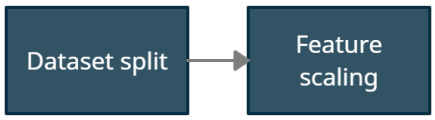


Figure 4.5: Flowchart for data conditioning for modelling

4.2.3.A Dataset split

The ML methodology recommends splitting the raw dataset into two different datasets: the training set and the test set. This method is also known as "Hold-out" method. The ML models are trained and validated with the training set while are assessed with the test set, which is basically a complete unseen data series for the trained model.

The test set is set to be the 30% of the raw dataset. The split is not performed randomly but through a stratification of the target in 1 m/s step. The test set is selected such that the distribution of the target is as similar as possible to the raw dataset. This method ensures that the set used for assessing the model (test set) obtains a similar target to that from the set used for training the model.

The results of the dataset split for the 3 datasets can be seen in Appendix B.

4.2.3.B Feature scaling

The feature scaling is a very important step before ML modelling, since with few exceptions, ML algorithms do not perform well when the input numerical attributes have different scales. The scaling method used is the Normalization (also known as min-max scaling). It consists in subtracting the min value and dividing by the maximum minus the min. The features end-up ranging from 0 to 1.

In order to meet the ML best practises, only the training set is used to fit the scaler (and thus to obtain the statistical values of the features). Then both the training and test set are scaled. This is done to avoid leaking information from the test set to the modelling in any way. Thus, it is ensured that the test set is completely unseen by the algorithms. It is not necessary to scale the target.

4.3 Datasets basic information

This subsection aims to briefly present the sites where the datasets were collected and the corresponding wind turbines. Please, bear in mind that some information might be missing due to confidentiality

reasons.

In this Thesis, three different datasets are analysed. Each dataset corresponds to a different project and wind farm all of them located in Australia. The first dataset includes data for two wind turbines. For Dataset2, three wind turbines are included in the study and finally, for the third dataset four different wind turbines are analysed. Those nine wind turbines are located, thus, in three different wind farms in Australia.

4.3.1 Project specifications

Table 4.4 summarizes the information by wind turbine regarding the three different wind projects under study. The table shows the measurement sectors by wind turbine considered on the filtering phase from subsection 4.2.2. It also includes the period for which each turbine was under the SC procedure.

Table 4.4: Project specifications for the different datasets

Project specifications		Wind sectors	SC period - Start date	SC period - End date
Dataset1	WTG14	90° - 180°	10/10/2019	21/01/2020
	WTG15	95° - 195°	10/10/2019	29/01/2020
Dataset2	T11	20° - 155°	04/05/2018	02/08/2018
	T17	100° - 120°	24/02/2018	30/07/2018
	T22	10° - 170°	14/05/2018	06/09/2018
Dataset3	WTG18	75° - 195° & 315° - 345°	03/04/2020	29/08/2020
	WTG20	80° - 180° & 330° - 350°	03/04/2020	10/08/2020
	WTG43	20° - 140°	05/04/2020	26/08/2020
	WTG46	35° - 175°	05/04/2020	26/08/2020

4.3.2 Wind turbines

Table 4.5 presents the wind turbines models, their hub height, rotor diameter and rated power.

Table 4.5: Wind turbines specifications for the different datasets

Wind turbines specifications		Vestas WT model	Hub height (m)	Rotor Diameter (m)	Rated Power (kW)
Dataset1	WTG14	V136-3.6 MW®	91	136	3600
	WTG15	V136-3.6 MW®	91	136	3600
Dataset2	T11	V117-3.45 MW®	90	117	3450
	T17	V112-3.45 MW®	84	112	3450
	T22	V117-3.45 MW®	90	117	3450
Dataset3	WTG18	V150- 4.20 MW®	105	150	4200
	WTG20	V150- 4.20 MW®	105	150	4200
	WTG43	V150- 4.20 MW®	105	150	4200
	WTG46	V150- 4.20 MW®	105	150	4200

4.3.3 Met masts

Table 4.6 shows the identification the temporary and the permanent met masts as well as their height and elevation.

Table 4.6: Meteorological masts specifications for the different datasets

Met masts		Temporary Met mast	Permanent Mast	Height (m)	Elevation (m)
Dataset1	WTG14	TM14	PC02	75	514
	WTG15	TM15	PC02	75	507
Dataset2	T11	TM11	PM1112	88	868
	T17	TM17	PM17	82	930
	T22	TM22	PM22	88	882
Dataset3	WTG18	TM18	PCV02	103	256
	WTG20	TM20	PCV02	103	255
	WTG43	TM43	PCV01	103	239
	WTG46	TM46	PCV01	103	221

5

Machine Learning modelling

Contents

Chapter 5 describes the proposed Machine Learning methodology in detail as well as the three different Machine Learning techniques implemented for modelling the datasets presented in Chapter 4. This chapter also includes an estimation of the IEC baseline error.

5.1 Modelling methodology

This section is split into the four main steps of any ML task: hyperparameter tuning, model training and generalization as well as a model explainer. Fig. 5.1 shows the complete flowchart of the ML methodology proposed:

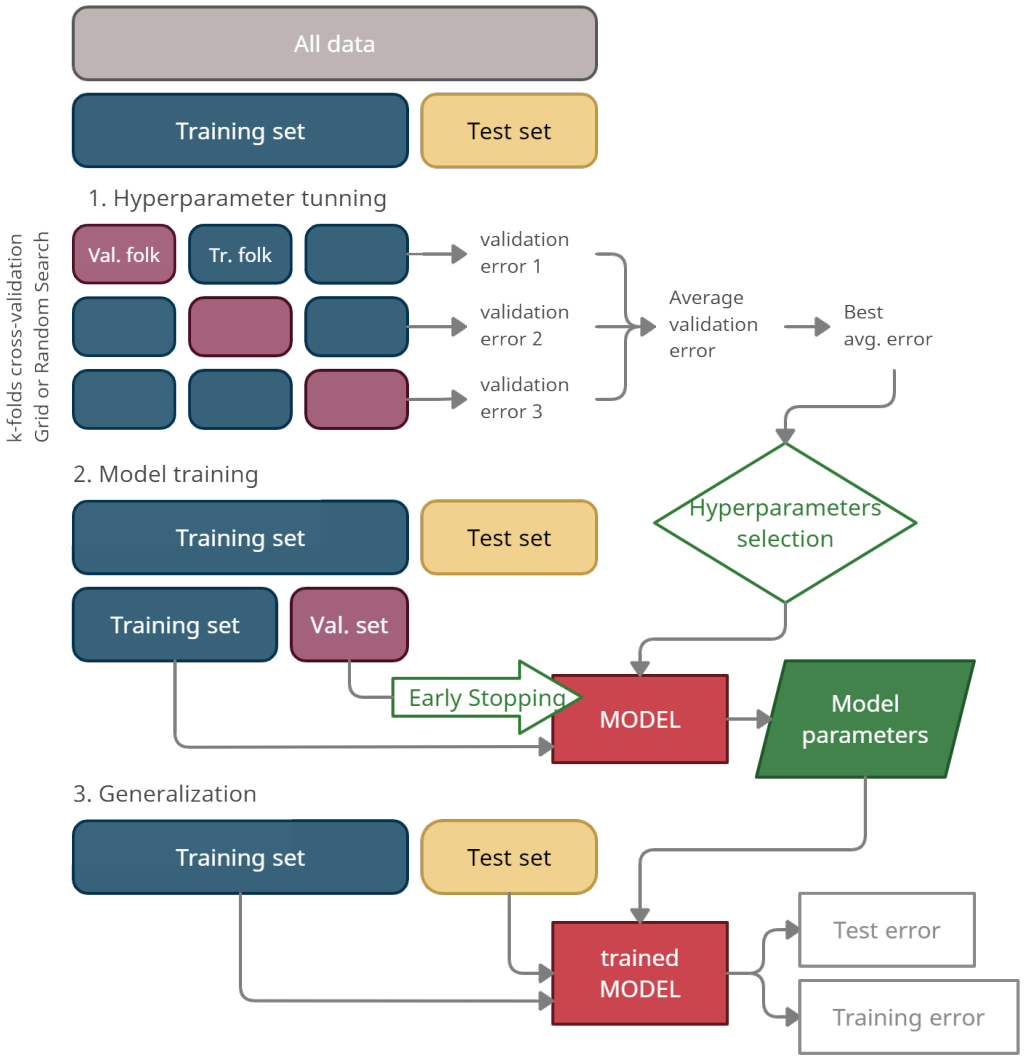


Figure 5.1: Machine Learning methodology proposed

The work done on this Master Thesis is mainly about ML modelling based on python code. The scripts were developed using Jupyter Notebook [43] and the main pyhton libraries, packages and functions used are referenced during this section but also summarized on Appendix E.

5.1.1 Hyperparameters tuning

To tune the model's hyperparameters is the very first step in any ML problem. The hyperparameters are sets of information that are used to control the way of learning from an algorithm. Their tuning affects the performance, stability and interpretation of the model. Each algorithm requires a specific set of hyperparameters that need to be adjusted according to the task. The list of hyperparameters to tune according to the implemented ML can be found in section 5.3.

Two different approaches for hyperparameter tuning are implemented: Grid Search and Random Search. Grid Search consists in building a grid of possible hyperparameters' values and trying out all possible combinations. This method is suitable when there is a shortlist of possible combinations. However, when the search space is large, Random Search is preferred. It evaluates a given number of random combinations by selecting a random value for each hyperparameter at every iteration [36].

In both cases, a Cross-Validation (CV), also known as k-fold Cross-Validation, is used to determine the best hyperparameters. It consists in splitting the training set into k distinct subsets called folds and then training and evaluate the model k times, picking a different fold for evaluation every time and training on the other k-1 folds [36].

Some of the hyperparameters that need to be tuned in a model are regularization hyperparameters. Regularization consists in constraining a model to make it simpler and reduce the risk of overfitting. Overfitting happens when a model performs well on the training data, but it does not generalize well [36].

5.1.2 Model training

Training a model in ML is like solving an optimization problem. Thus, the model learns from the training set of inputs by adjusting the models' parameters (also named coefficients) so that the difference between the predicted output and the target value is minimum. This difference is called 'loss' and it is expressed through the Mean Squared Error (MSE) as can be seen in Eq. 5.1.

$$\text{minimize } \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{5.1}$$

Where y_i is the target value and $\hat{y}_i$ is the predicted value for each of the n observations.

ML involves using an algorithm to learn and generalize from historical data to make predictions on new data. The variables that are entered into the model are called features and the output of the model is called a prediction.

5.1.3 Generalization

The generalization is the last step in the methodology. It consists in using the hold-out set, thus the test set, to assess how well the model can generalise to new and unseen data. For that, different metrics are defined so that the models are comparable.

These metrics will be applied to both wind and power. Wind speed is the target of the regression and to compute the errors based on power measures, the predicted power is estimated from the predicted wind speed by simple 2D interpolation from the corresponding PC.

The three different types of errors that are estimated for wind and power are the following:

5.1.3.A Mean Absolute Error

The Mean Absolute Error (MAE) is the simplest one and the one which is easier to understand. It is just the average error.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|^2 \tag{5.2}$$

5.1.3.B Root Mean Squared Error

The Root Mean Squared Error (RMSE) is the second metric for generalization assessment. It is a variant of the MSE but it is interpretable in the y units. It punishes large errors, thus it is highly affected by outliers, but tends to be useful when modelling real data.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{5.3}$$

5.1.3.C Mean Absolute Percentage Error

The Mean Absolute Percentage Error (MAPE) is the last metric assessed. It is intuitive and gives an idea of the magnitude of the absolute error.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{5.4}$$

5.1.4 Model explainer

Shapley Additive exPlanations (SHAP) [44, 45] is the state of the art in ML models' explainability. SHAP values are a series of parameters that quantify the contribution of each feature to the prediction of a given complex model. In a nutshell, predictive models answer the "how much" and SHAP answers the "why". Model explainers are found to be useful to increase ML the model's transparency, although some

of the model explainability techniques may be seen as a black box. SHAP values computation is not difficult to understand, however. By way of example, take Fig. 5.2 as a model that is trained with three input features: wind speed at the hub height, TI and WSH. SHAP requires training a distinct predictive model for each possible input feature combination. For instance, if the total number of input features is 3, the number of models to train is 2^3.

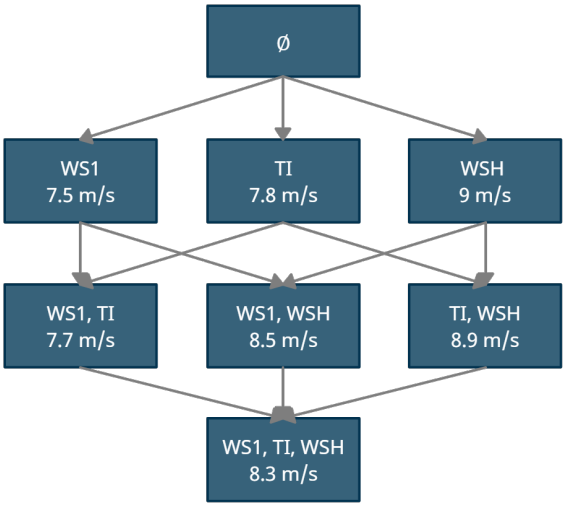


Figure 5.2: SHAP values explanatory example for a single prediction of a model with three variables

These models, however, are completely equivalent to each other, since the only thing that changes is the subset of input features used for training. The outcomes for each of the training runs are saved. The gap between predictions of two connected nodes can be imputed to the effect of that additional feature. This is called the "marginal contribution" of a feature for a specific model. A weighted average is computed to obtain the marginal contribution of each feature in average. Please note that SHAP values are a local measure (they are computed for each feature and each observation). Thus, this Thesis will consider the average SHAP values for the 10% of the test set observations for each of the input features.

5.2 IEC baseline

Despite IEC standards' new version of 2017, IEC 2005 standard is broadly accepted and is the base of more than 97% of the past and actual PC warranties of the wind turbine supplier Vestas. That is the main reason why this Thesis uses as a reference the IEC 2005 version.

5.2.1 IEC Reference methodology

One of the main objectives of this Thesis is to be able to compare the results of the proposed ML methodology with the IEC methodology as a baseline. Table 5.1 shows a complete comparison of the Current IEC & Vestas methodology, the reference IEC methodology that is used to compute the FCF and obtain the reference error and the ML methodology proposed:

Table 5.1: Summary of methodologies for performing a Site Calibration

Steps of the Methodology		Current IEC-Vestas	IEC reference	ML methodology
Data pre-processing & preparation	Data Correction	✓	✓	✓
	Data Cleaning	✓	✓	✓
	Data Quality Control	✓	✓	✓
	Outliers detection and removal	✗	✓	✓
	Data clustering	✗	✓	✓
	Data exploration	✗	✓	✓
	IEC filters	✓	✓	✓
	Warranty Contract filters	✓	✗	✗
	Complete WD sectors	✓	✓	✓
Inputs	WS1	✓	✓	✓
	Other atmospheric variables	✗	✗	✓
Regression	WD binning	✓	✓	✗
	FCFs computation by WD_bin	✓	✓	✗
	Linear regression	✓	✓	✓
	Polynomial regression	✗	✗	✓
	Non-linear regression	✗	✗	✓
Split & Error	All dataset for training and error	✓	✓	✓
	Model training on train set only	✗	✓	✓
	Error computation on test set only	✗	✓	✓

With the aim to finally obtain the same observations in each dataset so that the methods are comparable, some changes are added to the IEC reference method. For instance, the outliers' detection and removal is included as an important ML data pre-processing step. On the other hand, although the IEC filters and the complete WD sectors filter are maintained, the warranty contract filters are neglected. Finally, the dataset is split into training and test set, as the original ML methodology, thus, for the FCF computation the training set is used and for the IEC reference error estimation, only the test set is used. For more detailed information regarding the FCF for the wind turbines and the different datasets, please check Appendix E.

5.2.2 IEC Reference errors

Table 5.2 shows the reference errors based on the IEC 2005 method presented in subsection 3.1 for both wind power and energy for all wind turbines under study.

Table 5.2: IEC baseline errors for the different wind turbines

IEC baseline errors		MAPE		MAE		RMSE	
		wind speed %	power output %	wind speed m/s	power output kW	wind speed m/s	power output kW
Dataset1	WTG14	4.05%	8.98%	0.349	134.0	0.467	214.0
	WTG15	5.88%	14.03%	0.49	193.6	0.655	305.4
Dataset2	T11	12.16%	-	0.725	207.8	1.009	373.0
	T17	6.82%	14.80%	0.635	207.3	0.807	310.3
	T22	12.13%	35.94%	0.965	375.7	1.232	542.9
Dataset3	WTG18	5.26%	13.14%	0.41	207.9	0.57	318.2
	WTG20	4.94%	12.29%	0.383	213.2	0.516	317.9
	WTG43	6.30%	14.31%	0.522	235.3	0.672	371.0
	WTG46	4.30%	11.37%	0.315	168.5	0.425	258.3

5.3 Machine Learning techniques

Three different ML techniques are applied to the three different datasets. The first one is indeed a set of models which are based on linear relationships among variables. The second is the widely known ANN and finally, the Extreme Gradient Boosting (XGB) regressor based on DTs will also be implemented. To get a bigger picture of which ML models have been selected, please check Appendix E.

5.3.1 Linear models

A linear model in ML is any model that assumes a linear relationship between the input features and the output and those can be expressed explicitly. Since met masts record more than one signal during SC, the models will consider several input variables and thus multivariate linear models are implemented.

5.3.1.A Optimization

Two types of linear models are implemented: simple linear regression and polynomial regression. The most simple ML linear model is the simple linear regression. The equation of the simple linear model only considers the set of values x_{ij} for each of the p inputs, for each of the n observations, in their original form to predict the target value $\hat{y}_i$, through the optimization of the linear coefficients for each feature β_j and the independent coefficient β_o.

$$\hat{y}_i = \beta_o + \sum_{j=1}^{p} (\beta_j + x_{ij}) \tag{5.5}$$

On the other hand, the polynomial regression inputs those features in a polynomial shape. For instance, if the problem had the features (a, b, c) the input Polynomial Features function of Scikit-Learn [46] for degree equal to 2 would be $(1, a, b, c, a^2, b^2, c^2, ab, bc, ca)$ [47].

5.3.1.B Regularization techniques

Ridge and Least Absolute Shrinkage and Selection Operator (LASSO) regression are two simple techniques to reduce model complexity and prevent over-fitting which may result from linear models. These regularization techniques consist in adding a term to the error to reduce the value of the coefficients in the model. The level of penalty is controlled by the hyperparameter α. When $\alpha = 0$, the penalty is zero and the result is equivalent to that of a linear model by Ordinary Least Squares (OLS). As α increases, the greater the penalty and the lower the value of the predictors.

- **LASSO:** The added penalty in the cost function is equivalent to the absolute magnitude of the coefficients [48]. This magnitude is known as L1 and allows the algorithm to force the coefficients to zero and for this reason, it performs feature selection. The cost function for LASSO can be written as in Eq. 5.6.

$$\text{minimize } \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j + x_{ij}\right)^2 + \alpha\sum_{j=1}^{p}|\beta_j| \tag{5.6}$$

- **Ridge:** The added penalty in the cost function is equivalent to the square of the magnitude of the coefficients [48]. This magnitude is known as L2 and provides a proportional reduction of the value of all the coefficients in the model, but preventing them from reaching a zero value. Eq. 5.7 is the expression for the Ridge regression cost function:

$$\text{minimize } \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j + x_{ij}\right)^2 + \alpha\sum_{j=1}^{p}\beta_j^2 \tag{5.7}$$

5.3.1.C Hyperparameters

For both Ridge and LASSO regressions, the only hyperparameter to tune is α. To find the optimal α, a Grid Search with CV is implemented [46]. The distribution of α is:

$$\{0.0003; 0.0004; 0.0005; 0.001; 0.003; 0.005; 0.01; 0.03; 0.05; 0.1; 0.3; 0.5; 10; 50; 80; 120; 300; 500\}$$

5.3.2 Artificial Neural Networks

ANN is the second method implemented for solving the SC regression task. It is a type of non-linear ML modelling. For more information regarding the type of ML models, please check Appendix E.

ANN have a biological inspiration on the human brain biological neurons, which with a large number of highly connected neurons, can perform complex tasks effortlessly. Fig. 5.3 shows the analogy between a biological and an artificial neuron. Each neuron receives several inputs x_i that are weighted w_i into a sum and are finally transformed into an output. The activation function defines how this transformation is performed.

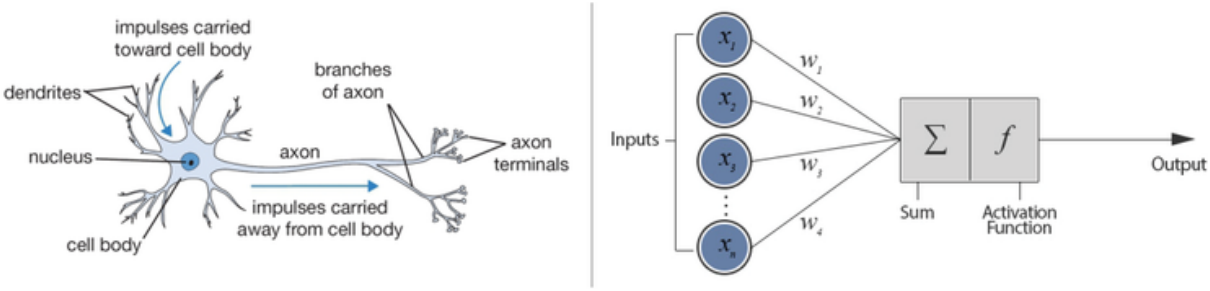


Figure 5.3: Biological and Artificial neurons from [49]

An ANN is then an interconnected group of nodes, known as Artificial Neurons that through an iterative process can learn from a dataset to predict an output.

5.3.2.A Architecture

There are several types of ANN. In this case, we will be using what is called a Multi-Layer Perceptron (MLP). The basic architecture of a MLP consists of the inputs, the input layer, hidden layers, the output layer and the output. In each of the layers, one or several neurons can be placed.

Since the SC is translated into a ML regression task, the architecture used for ANN is the following:

- The number of neurons in the input layer will be the same as the number of input features, which will depend upon the wind turbine.
- The number of neurons in the output layer is 1 because the SC task is to estimate one single target, the wind speed at the wind at the hub height at the turbine location.
- The activation function for the hidden layers will be ReLU or their variants.
- The activation function for the output layer will be none (thus linear).

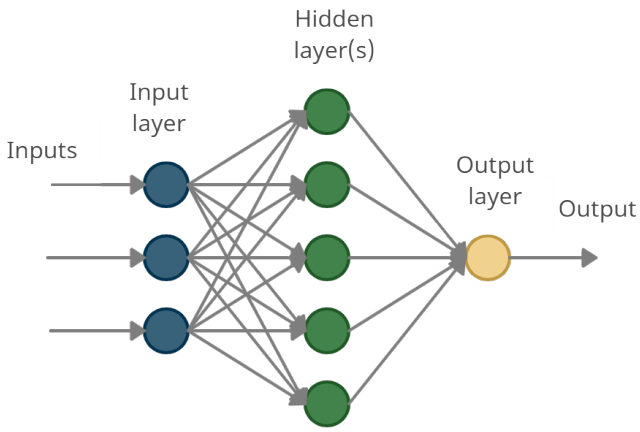


Figure 5.4: Basic Architecture of an Artificial Neural Network

5.3.2.B Optimization

The process of training an ANN is indeed iterative as shown in Fig. 5.5 and the optimization problem is reduced to finding the optimal values for the weights w_{ii} for each neuron.

The inputs x_{ii} together with the activation function and the weights w_{ii} determine the output through the forward propagation. Just that on the first iteration, the weights are initialized using an initialization technique specified by the user. All observations are passed through the network in parallel. The network's output is then compared with the target and the loss is computed. Through the backward propagation, the gradient is computed and together with the learning rate, which is either tuned and or adapted through an optimizer, the weights are updated. Then the iteration is repeated. An entire pass forward and backward of the entire dataset through the ANN is called "Epoch".

5.3.2.C Regularization techniques

There are mainly four types of regularization techniques applied to Those are the following:

- **L1 and/or L2:** these regularization techniques corresponds to LASSO and Ridge regularization techniques applied also for linear modelling in subsection 5.3.1.B
- **Dropout:** This regularization technique consists in training the network ignoring at each step some of the neurons. This method can avoid overfitting because it prevents neurons from co-adapting with their neighbours and they are forced to be as useful as possible on their own.
- **Early Stopping:** This regularization technique consists in computing the loss for every epoch, for both training error and the validation error (which represents 20% of the original training set). The training will be stopped whenever the validation error stops decreasing.

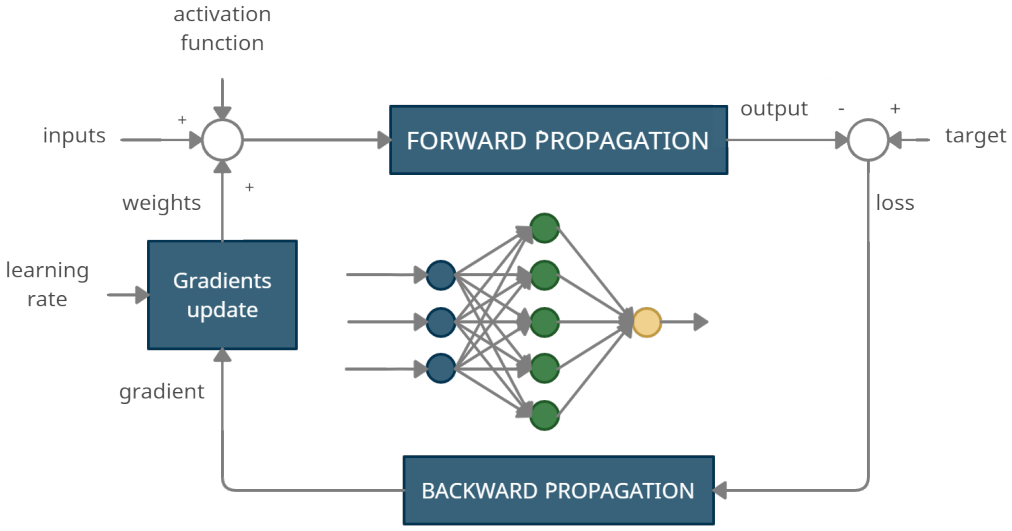


Figure 5.5: The iterative process of Artificial Neural Network model training

5.3.2.D Hyperparameters

There 6 types of hyperparameters that are tuned through a Random Search for the ANN modelling on the Keras function implemented [50]. The number of hidden layers and the number of neurons per layer are part of the architecture of an ANN. Also, the regularization techniques were explained in the previous subsection. Finally, other hyperparameters need also to be tuned:

- `learning_rate` is the step of optimization.
- `optimizer` is the algorithm that allows the neural network to find the minimum of the loss as a function of the weights. Several optimizers will be included in the randomized search, basically two types: The ones based on the Gradient Descent and its corrections: Stochastic Gradient Descent (SGD), Momentum, Nesterov. The adaptative (learning rate) optimization methods: Adam, Nadam, RMSProp.
- `activation_function` is the function that defines the transformation of the output of a neuron. The reason for including these other variants is because of the "dying ReLU" problem: that it is when during training, some neurons "die" since they stop outputting anything other than 0. This may happen to a neuron if its weighted sum of inputs is negative. However, the problem comes if many neurons output zero, then part of the network dies.

At this point, it is important to mention some assumptions made to be aligned with the ML good practices:

- The number of neurons per hidden layer will be the same for all hidden layers.

Table 5.3: Artificial Neural Network hyperparameters distribution for Random Search tuning

Hyperparameters	Distribution
# hidden layers	[0; 1; 2; 3]
# neurons per hidden layer	0 - 100, step:10
learning rate	[0.0001; 0.0003; 0.001; 0.003; 0.005; 0.01; 0.03]
optimizer	[SGD; Momentum; Nesterov; Adam; Nadam; RMSProp]
activation function	ReLU; Leaky ReLU; eLU; SeLU
regularization	None; Early Stopping; Dropout; L1; L2

- The layers will be fully connected.
- The maximum number of epochs is 100. However, the Early Stopping regularization technique may be applied.
- The weights will be initialized through the “he normal” technique.

Table 5.3 shows the hyperparameters and their possible distribution for the Randomized Search.

5.3.3 Extreme Gradient Boosting

The third and final ML implemented is the XGB DT-based, which is an ensemble model. The idea behind ensemble modelling is that a single algorithm, on its own, might not be able to capture all relations in a given dataset. However, a group of algorithms trained with different parts of a dataset might be able.

As can be seen in Fig. 5.6, Boosting trains models sequentially, each new model is trained to correct the errors of the previous ones. At each iteration, the outcomes predicted correctly are given a lower weight for the next model, and the ones wrongly predicted a higher weight. It then uses a weighted average to produce an outcome.

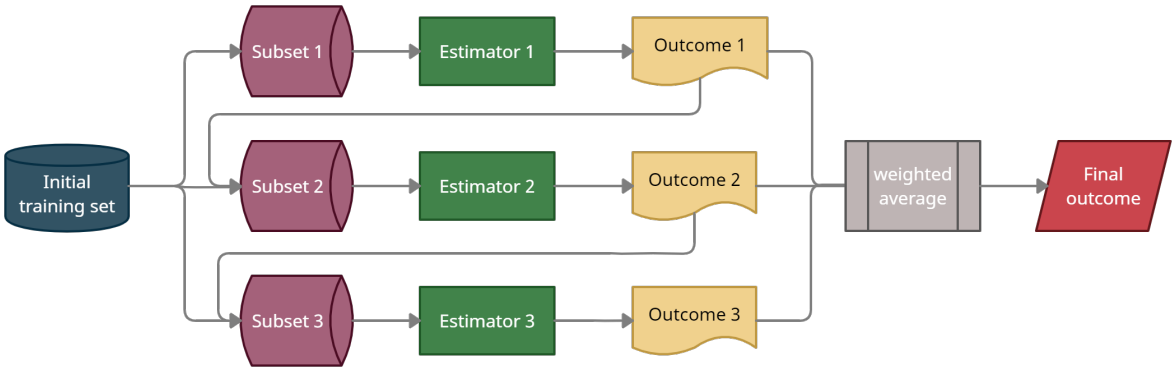


Figure 5.6: Ensemble Boosting modelling basic process of sequential training

XGB is a Boosting method where errors are minimized using a Gradient Descent algorithm and

which has several advantages over regular Gradient methods. For instance, it has a built-in CV method, a regularization technique and it is much faster.

5.3.3.A Architecture

Since the XGB modelling implemented is based on DTs, the corresponding architecture is for DTs. DTs are composed of nodes, branches, and leaves. The Root Node is the initial node that represents the entire sample and may get split further into further nodes. Each Internal Node represents a feature of the dataset and the branches represent the decision rules. Finally, the Leaf Nodes represent the outcome. A 'child' is any node coming from another 'father' node, either a leaf or an internal node. The depth of a DT is defined by the number of levels, not including the root node.

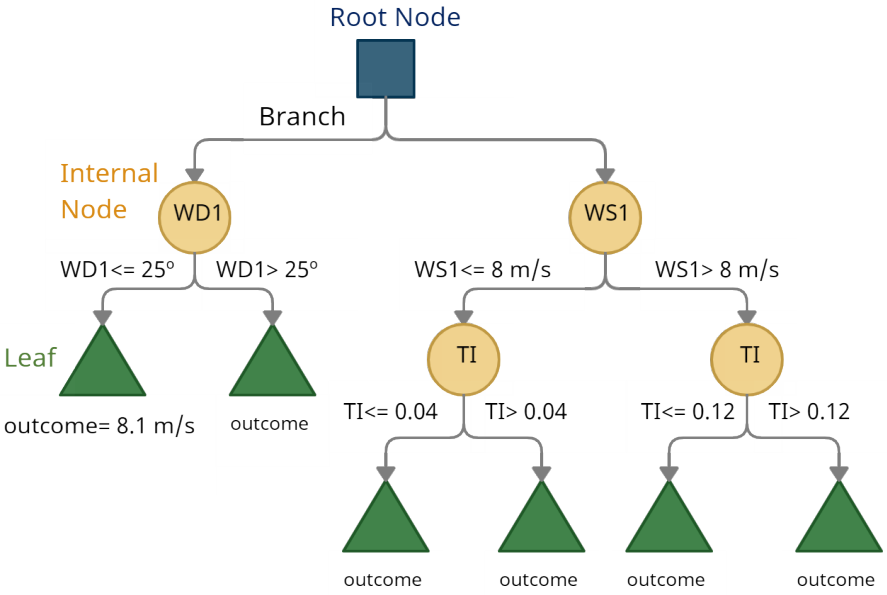


Figure 5.7: Basic architecture of a 3 levels Decision Tree

5.3.3.B Optimization

DTs apply a top-down approach to data so that for a given dataset, the algorithm splits each region in a way that makes most training observations as close as possible to that predicted value. With a particular data point, it is run completely through the entire tree by answering true or false questions till it reaches the leaf node. The final prediction is the average of the value of the dependent variable in that particular leaf node. Through multiple iterations, the Tree can predict a proper value for the data point.

5.3.3.C Regularization techniques

Several regularization techniques can be applied to XGB. Two different methods are applied:

- **Pruning:** This regularization technique consists in growing each decision tree without restrictions and then pruning (deleting) unnecessary nodes. A node whose children are all leaf nodes is considered unnecessary if the purity improvement is not statistically significant (through p-value probability test). The pruning will continue until all unnecessary nodes have been pruned. XGB makes splits up to the `max_depth` hyperparameter specified and then starts pruning the tree backwards and removes splits where the loss is lower than the gamma hyperparameter value.
- **L2 regularization:** This regularization is analogue as for Ridge regression and is controlled through `reg_lambda` hyperparameter.

5.3.3.D Hyperparameters

There are three main types of hyperparameters that can be tuned on the implemented XGB modelling based on xgboost function [51]. The research for the best hyperparameters is based on a stepwise Grid Search for the XGB modelling as Aarshay Jain describes [52]. They are tuned in order: ensemble type hyperparameters, tree-specific hyperparameters and regularization parameters.

Ensemble type hyperparameters:

- `n_estimators` is the number of DTs. Usually, the higher the number of trees, the better to learn the data. However, adding a lot of trees can slow down the training process considerably. Increasing the number of estimators may result in overfitting.
- `learning_rate` is the step size shrinkage used in the update of the weights of the new features to prevent overfitting by making the boosting process more conservative.

Tree specific hyperparameters:

- `max_depth` is the maximum number of levels in each decision tree. The deeper the tree, the more splits it has, and the more information about the data can capture. By reducing the value of this hyperparameter, the model is being regularized and the risk of overfitting is reduced.
- `min_child_weight` defines the minimum sum of weights of all observations required in a child. Higher values prevent a model from learning relations that might be highly specific to the particular sample selected for a tree and thus avoid overfitting.
- `gamma` is a pseudo-regularization parameter and depends on the other parameters. A node is split only when the resulting split gives a positive reduction in the loss function. `gamma` is the minimum

loss reduction required to make a further partition on a leaf node of the tree. The larger `gamma` is, the more conservative the algorithm will be.

- `subsample` is the subsample ratio of the training instances that are inputted to the model. Setting it to 0.5 means that XGB would randomly sample half of the training data before growing trees, and this will prevent overfitting. Subsampling will occur once in every Boosting iteration. Lower values might make the algorithm more conservative and prevent overfitting, but too small values might lead to underfitting.

- `colsample_bytree` denotes the fraction of features to be randomly sampled for each tree. It might reduce overfitting.

Regularization hyperparameters:

- `reg_lambda` is the L2 regularization of the weights (Ridge Regression). Increasing this value will make the model more conservative. It might help to reduce overfitting.

Table 5.4 shows the hyperparameters and their possible distribution for the Grid Search.

Table 5.4: Extreme Gradient Boosting hyperparameters grid for step-wise Grid Search tuning

Hyperparameters	Grid
learning rate	[0.01; 0,03; 0.05; 0.1; 0.2; 0.3; 0.5; 1]
# estimators	[10; 50; 100; 150; 200; 300; 400; 500; 600; 700; 800; 1000]
max_depth	[3; 5; 7; 10; 15; 20 ;25]
min_child_weight	[1; 2; 3; 6; 10]
gamma	[0.1; 0.2; 0.3; 0.4; 0.5]
subsample	[0.5; 0.6; 0.7; 0.8; 0.9; 1]
colsample_bytree	[0.4; 0.5; 0.6; 0.7; 0.8; 0.9; 1]
reg_lamda	[0.4; 0.5; 0.6; 0.7; 0.8; 0.9; 1]

6

Results and implementation

Contents

Chapter 6 gives a detailed review of all the experimental tasks tackled during the modelling phase of this research project. In this chapter, the results and the explainability of the models proposed in this Master Thesis are shown. In addition to that, an initial study of the possible implementation according to the wind industry necessities is also included in this chapter.

6.1 Modelling results

In this subsection, the ML models' errors are compared to the IEC baseline, the main outcomes from the hyperparameter tunning for each of the model's types are explained and also, the most important meteorological variables used as model inputs are ranked by ML technique.

6.1.1 Error analysis

The main goal of the present Master Thesis was to prove the validity of a data-driven model to perform a SC for wind turbines. In this case, three different ML techniques were implemented: Polynomial Regression, ANN and XGB. The results for the three different models are compared to the IEC baseline error for the 9 wind turbines in 3 different locations. Fig. 6.1 shows the MAPE error for wind speed and Fig. 6.2 the MAPE error for power.

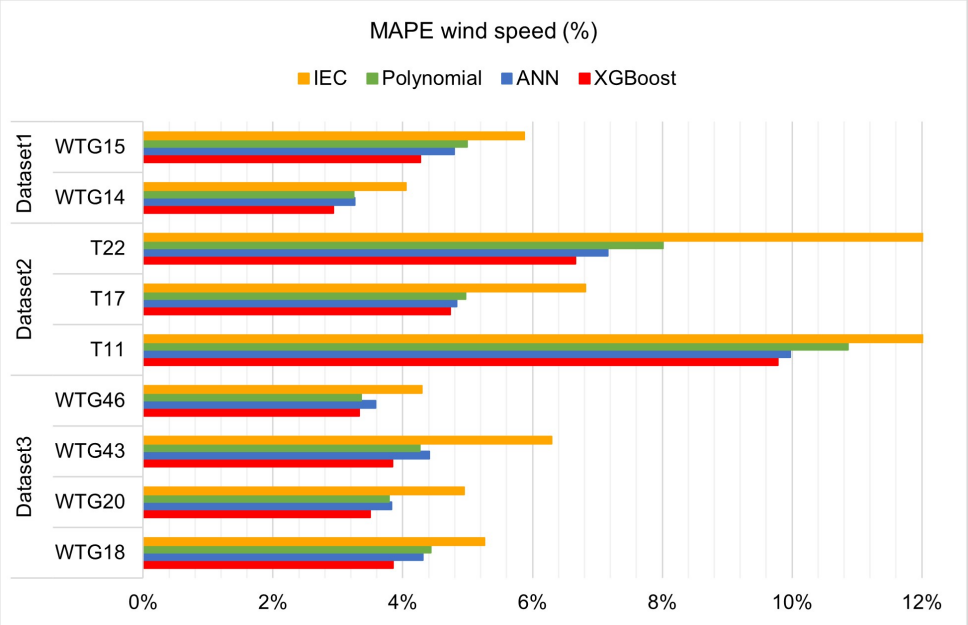


Figure 6.1: Wind speed MAPE by Machine Learning technique for the different wind turbines

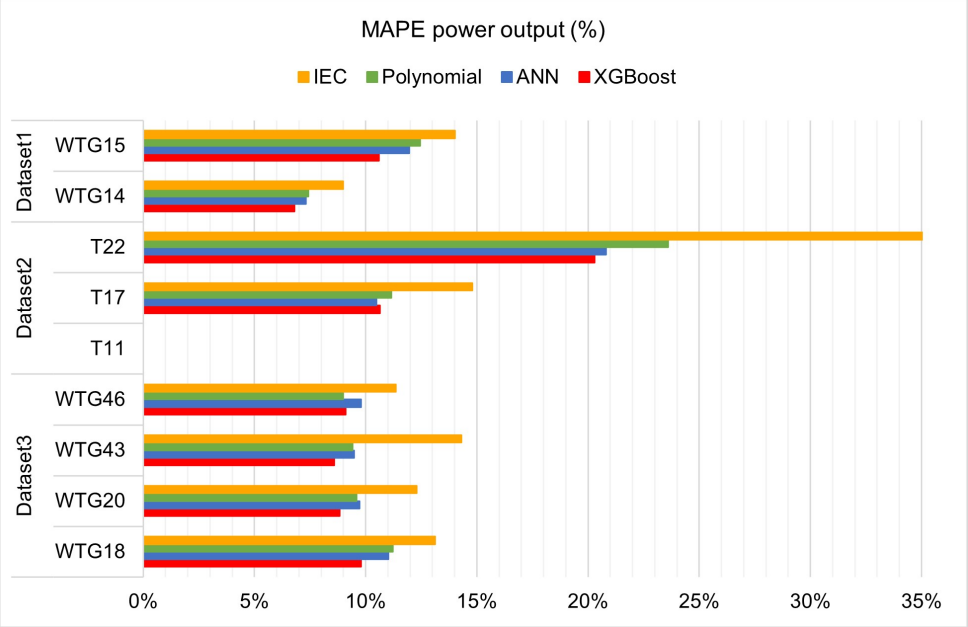


Figure 6.2: Power MAPE by Machine Learning technique for the different wind turbines

The main outcome from this comparison is that in all cases, ML models present a lower error than the IEC standard methodology. The error reduction by wind turbine in the case of wind was from 0.7% up to 5.4% while for the power, the reduction was from 1.5% up to 15.6%. The average reduction by ML technique is summarized in Table 6.1.

Table 6.1: Average MAPE reduction compared to the IEC standards

MAPE reduction (%)	wind speed	power output
Linear modelling	1.54%	3.87%
Artificial Neural Network	1.74%	4.28%
Extreme Gradient Boosting	2.11%	5.02%

As can be seen in Table 6.1, in general, the ML technique with better performance was XGB followed by ANN, being the linear models the less effective in predicting the wind speed. As will be shown in subsection 6.1.3, this is due to the main limitation of linear models which is that those only consider linear relationships between the features and the target, while the non-linear models such as ANN and XGB can find the most relevant non-linear relationships between the meteorological variables and the target wind speed. On the other hand, these types of models require a much higher computational time than linear models. While a linear model can be tuned and trained in seconds, the hyperparameter search for an ANN for the datasets used in this project took several minutes and a complete XGB tunning took up to 2.5 hours.

6.1.2 Best hyperparameters

Hyperparameter tuning is a key step in any ML procedure. It is a required previous step to the model training. Depending on each dataset, the hyperparameters that can lead to better performance can a priori differ significantly. In this sense, one of the main points of the modelling was to assess the importance of hyperparameter tuning itself. That is why during the research for the best hyperparameters the results were always compared among datasets, to finally obtain the best hyperparameters among all searches.

6.1.2.A Linear models best hyperparameters

In the case of linear models, apart from the hyperparameters, some other interesting insights have been obtained from the type of linear model to use, either simple linear regression or polynomial regression, as well as the type of regularization.

The first conclusion regarding linear models is that polynomial regressions perform better than linear regressions either with or without regularization. For instance, using the same subset of variables for the same wind turbine and target, a second-degree polynomial regression reduced in average the MAPE for wind by around 0.8%. The main difference between a linear and a polynomial regression is that, while the former inputs the regular dataset to the model, the latter transforms the input subset to consider several combinations of two input variables. Since the combinations are formed by two multiplying variables, this can be seen as an imitation of non-linear relationships.

On the other hand, for a polynomial regression, a degree greater than 2 generally does not improve the accuracy of the model significantly. For higher degrees, the computational time increases exponentially due to a large number of coefficients to check and in some cases, those higher degrees can lead to model overfitting. Moreover, the general trend for the polynomial regression is not selecting many coefficients of high degrees. Thus, the main conclusion is that a polynomial regression of the degree of 2 is good enough for linear modelling.

Regarding regularization, for wind turbines, the best performing technique was Ridge regression. Although LASSO regression was also performing better than IEC standards, Ridge was in all cases outperforming by more than -0.2% in the average MAPE for wind. This outcome could be translated as feature selection in linear modelling not being so relevant. On the other hand, although a simple linear regression can suffer lower performance due to multicollinearity issues, for some of the wind turbines, a polynomial regression without any kind of regularization has been proved to be a better choice.

6.1.2.B Artificial Neural Networks best hyperparameters

ANN technique had a total of 6 hyperparameters to tune and the procedure behind the ANN tunning was to compute 3 different Random Searches and compare the validation error between the hyperparameters selected by the random search and the previous best hyperparameters from the other wind turbines. The result is that the combination of hyperparameters found for Dataset1 could not be improved by any subsequent random search for Dataset2 or Dataset3. Thus, the best hyperparameters are all the same for all wind turbines and all sites except for the regularization technique.

Four different types of regularization techniques were implemented on the random search: L1, L2, Dropout and Early Stopping. The Early Stopping technique has proved to be the most efficient in all cases. However, in some cases, not applying any regularization technique at all lead to a better ANN modelling performance.

It is also worth mentioning, that in the case of the optimizer, the best performing one is Adam, which is a type of adaptative learning rate optimizer, followed by Nesterov and Momentum, two Gradient Descent-based optimizers.

For the rest, the best combination of hyperparameters was as follows: 3 hidden layers, 80 neurons per layer, a `learning_rate` of 0.001 and eLU as activation function.

6.1.2.C Extreme Gradient Boosting best hyperparameters

The procedure behind XGB modelling is a stepwise Grid search. A Grid Search works fine and it makes sure that the best possible combination of hyperparameters is selected. XGB was the ML technique which best hyperparameters differed the most from wind turbine to wind turbine. However, some general trends could be observed for some specific hyperparameters from the different Grid searches implemented. On the one hand, `n_estimators` tended to be great which can be interpreted as the boosting method trying to convert more weak features into strong ones, and on the contrary, the learning rate tended to be small-medium, which means the boosting process tends to be more conservative. Regarding regularization, gamma is generally 0, which means that the model is free to adapt to each data point by making further partitions on a leaf node of the tree at a low loss reduction and on the contrary, `reg_lambda` is generally large, which means that the model tends to be more conservative regarding the cost function by applying some Ridge regularization.

6.1.3 Feature importance

When looking at the feature importance of the different ML models, several conclusions were obtained. Firstly, feature importance depends on the ML technique. When grouped by model, the SHAP values for each feature are similar among the different wind turbines. Taking this a first clustering, it was also

observed that each of the ML models implemented tended to select roughly the same atmospheric variables for all wind turbines of the same site, especially for the most and the less relevant features. Fig. 6.3 thereupon shows the feature importance by ML technique computed as the SHAP values for wind variables averaged by site first and then averaged for the different sites. It is important to mention that SHAP values were rescaled in the first step in order to obtain comparable results among models.

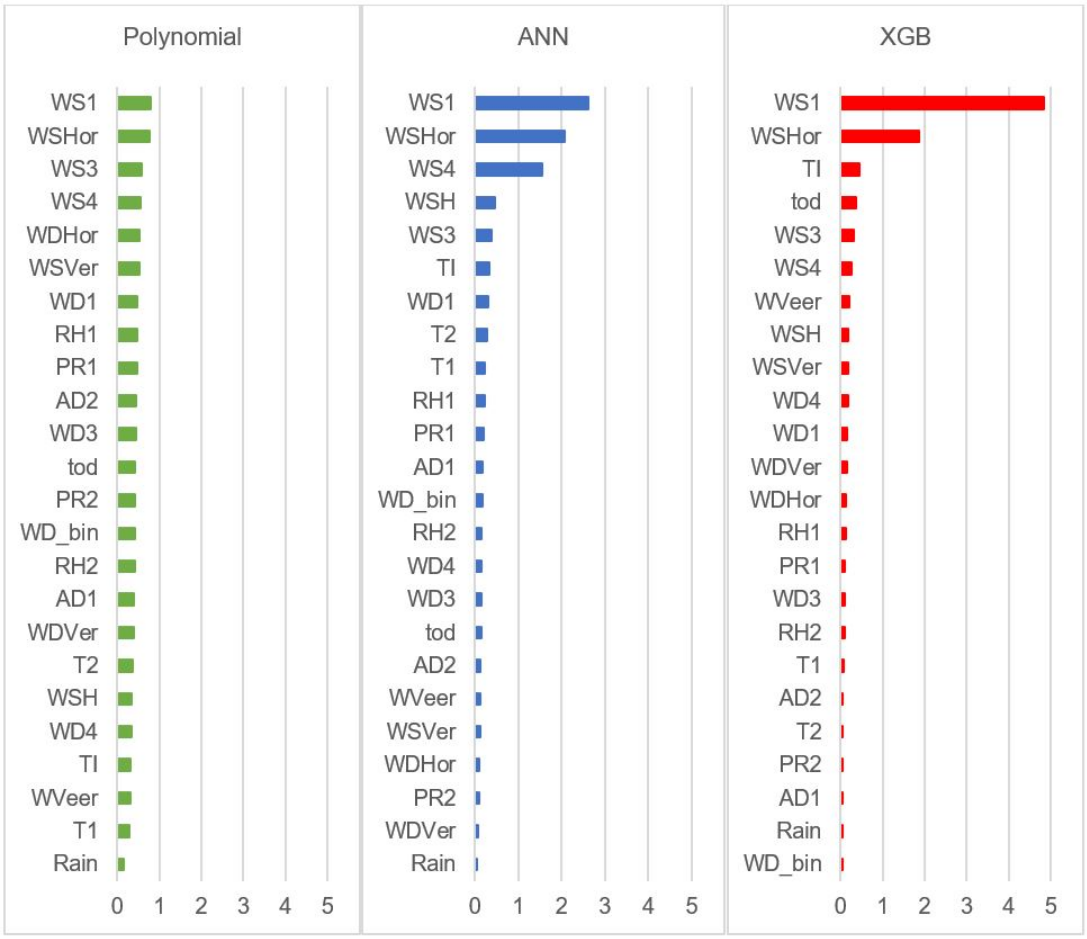


Figure 6.3: Average feature importance Ranking by Machine Learning technique

As seen in Fig. 6.3, for all three models, the two most important features are the wind speed at the hub height, as expected, but also its horizontal component captured by the ultrasonic or 3D anemometer.

Another important insight is the key difference between linear and non-linear models. Linear models are much less selective, thus these models tend to give importance to more variables. While non-linear models can better select the most important variables, especially XGB, which is very consistent regarding feature importance among the different wind turbines.

When taking a closer look at the ranking in Fig. 6.3, Linear models prefer wind variables, relative humidity and air density and ANN also considers temperature and relative humidity. However, XGB, the

best performing technique, top 10 variables are only wind speed and wind direction related, including TI.

Moreover, it has also been observed that even if not in the same order, many meteorological variables are ranked to be around the same position for all three models. This is, for instance, the case of WS3 and WS4. From all wind variables, the case of TI is worth to be mentioned. TI has a highly non-linear relationship with the target, that is why it is at the end of the ranking for the linear models while it is at the top for non-linear models, especially for XGB which TI can be found in the third place.

Now, a short comparison of these results with the IEC standards. On one side, the standards only consider the WD1 variable as the Wind Direction bin. The results of the modelling show that WD1 was ranked in all three cases as more important than Wind Direction bin. Indeed, for the best performing technique, XGB modelling, Wind Direction binning is precisely the last variable in importance according to the ranking in Fig. 6.3. This probably means that the multi-binning methodology of the standards, in the case of the Wind Direction, results in a loss of information. On the other hand, TI was ranked very close to Wind Shear for Polynomial and ANN techniques and in the case of XGB, in a much higher position. Conversely, although TI is an available variable in all the datasets, the IEC standards only consider Wind Shear for the SC procedure.

6.2 Wind industry implementation study

Three critical points from the wind industry perspective have been considered in this implementation study. First, an analysis of the improvement of the Measured Power Curve estimation and thus the AEP estimation. Second, the possibility of approaching the SC task through a Universal ML model by wind farm. And finally, an analysis on the sensor importance for ML modelling for SC.

6.2.1 Improvement of the Measured Power Curve

The main concern raised by the wind industry is to find a SC that accurately predicts the wind speed at the hub height and at the turbine location and thus, correctly estimates the MPC.

Appendix D presents the resulting SC-PC for all wind turbines under study based on the IEC multi-binning method. As can be noted, IEC systematically overestimates the wind speed around the rated power. The two main consequences are a significant decrease in the MPC at its elbow and, in some cases, a reduction in the Measured AEP by default. In addition, in some cases, the wind speed is underpredicted on the lowest part of the PC and thus, the MPC tends to increase.

The methodology developed by the Power Curve Verification department of Vestas, presented in subsection 3.3.2, provides a fair analysis and measurement of the impact of an inaccurate SC on the MPC. This method consists in computing the 'SC-PC' for a specific SC methodology and wind turbine. By the implementation of this methodology, the improvement that ML techniques applied to SC can provide to

the MPC is presented. Following the example of subsection 3.3.2, Fig. 6.4 shows the improvement of the SC-PC for wind turbine WTG14.

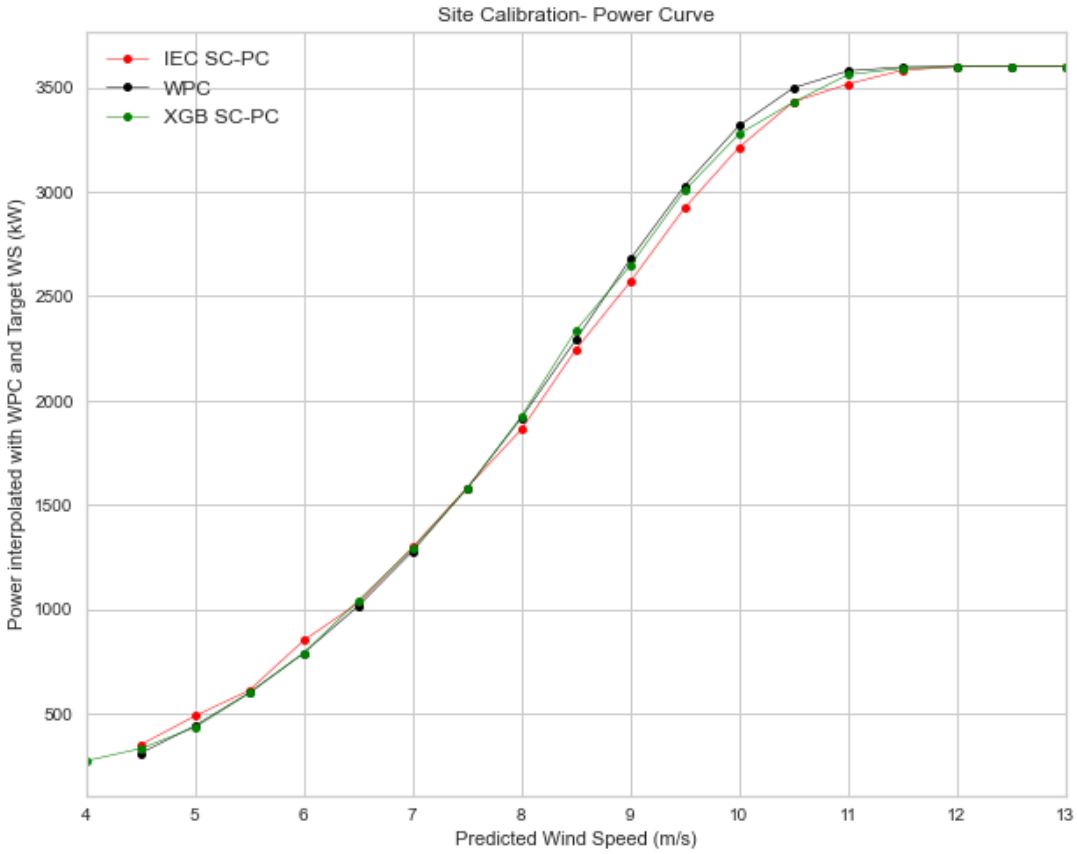


Figure 6.4: Site Calibration Power Curve comparison for WTG14

Generally, ML techniques applied to SC improve the accuracy in estimating the MPC. In most of the cases, the systematic errors introduced in the curve by the IEC standard method, are corrected by the ML models. As seen in Fig. 6.4, the curve based on XGB is closer to the expected WPC than the IEC curve, especially around the elbow, but also on the lowest part of the curve. Some other examples of XGB modelling curves compared to IEC curves for different wind turbines can be found in Appendix E.

It is also possible to see the improvement of the ML models considering the scatter. Fig. 6.5 and Fig. 6.6 represent the scatter of the test set points used to compute the 'Site Calibration Power Curve' shown in Fig. 6.4, for both IEC standards and XGB modelling respectively. As it is noticeable, the XGB scatter is more accurate than the IEC since the points are closer to the WPC.

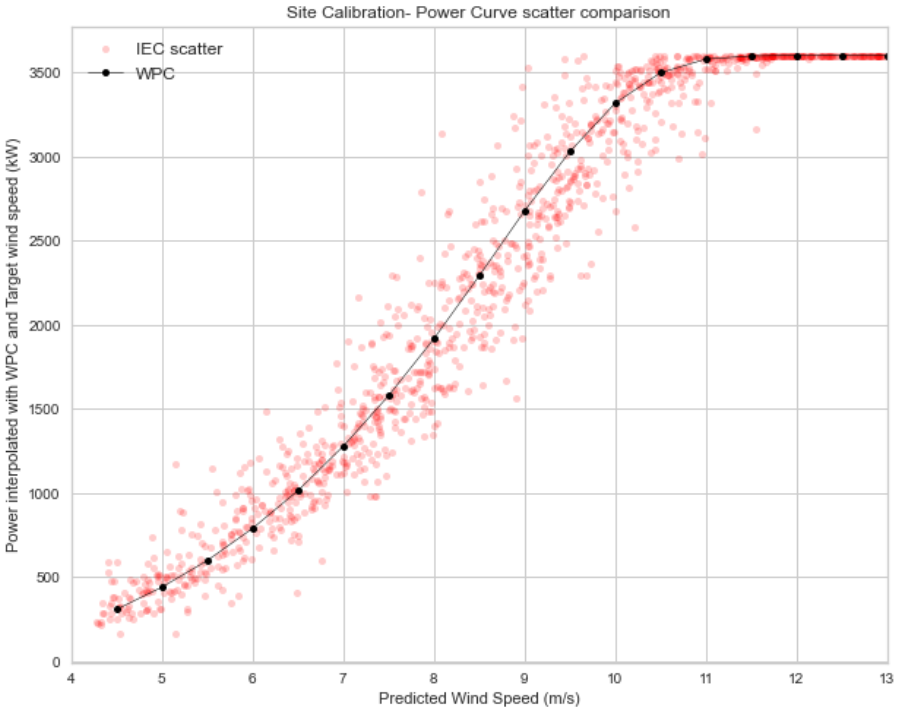


Figure 6.5: Site Calibration Power Curve scatter IEC WTG14

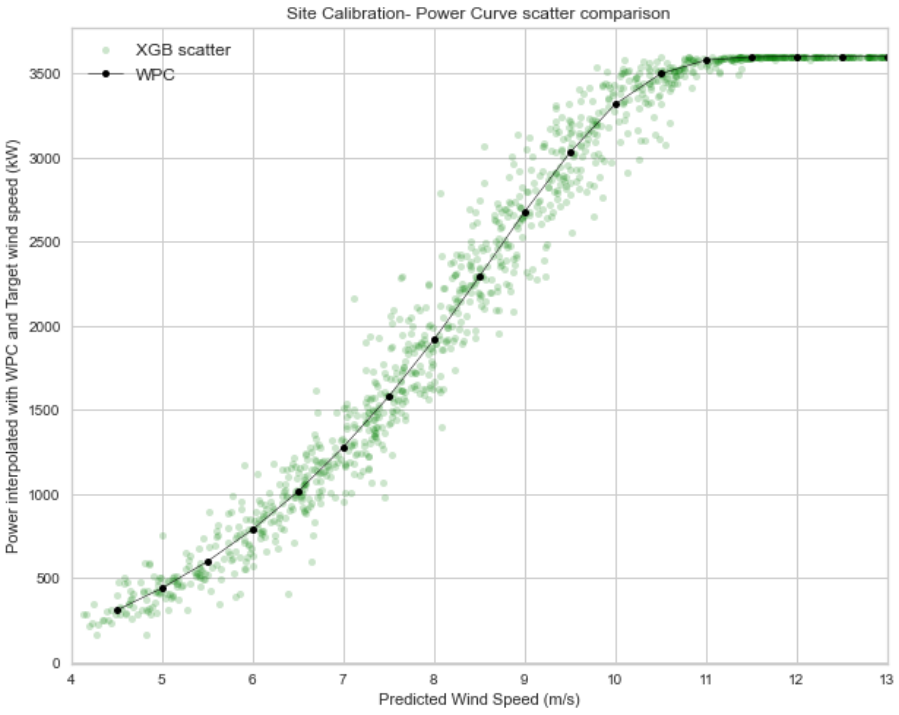


Figure 6.6: Site Calibration Power Curve scatter XGB WTG14

This MPC improvement in some cases leads to an improvement of the AEP estimation. The results of the models for the two different AEP metrics presented in subsection 3.3.2 are shown in Tables 6.2 and 6.3.

Table 6.2: Annual Energy Production percentage by Site Calibration model

AEP (%)		IEC	Linear	ANN	XGB
Dataset1	WTG14	99.17%	100.28%	101.86%	100.09%
	WTG15	99.22%	99.53%	99.40%	99.29%
Dataset2	T11	99.70%	99.92%	96.92%	99.89%
	T17	-	-	-	-
	T22	-	-	-	-
Dataset3	WTG18	99.87%	99.78%	99.27%	99.87%
	WTG20	100.36%	100.20%	101.13%	100.06%
	WTG43	100.40%	100.25%	101.47%	99.83%
	WTG46	99.94%	99.95%	99.34%	99.98%

Table 6.3: Annual Energy Production difference by Site Calibration model

AEP (MWh)		IEC	Linear	ANN	XGB
Dataset1	WTG14	-118.417	40.600	267.813	12.241
	WTG15	-112.892	-67.253	-86.239	-102.004
Dataset2	T11	-40.847	-11.484	-418.767	-14.418
	T17	-	-	-	-
	T22	-	-	-	-
Dataset3	WTG18	-27.157	-46.535	-153.724	-27.685
	WTG20	75.065	42.284	240.789	12.154
	WTG43	84.422	52.978	311.333	-36.857
	WTG46	-12.416	-9.605	-139.094	-4.205

Again, XGB modelling outperforms. XGB technique improves the AEP estimation compared to the IEC method for all cases with the exception of wind turbine WTG18, for which the difference with the expected AEP is similar. The average improvement is around 62% of the AEP baseline and in the best case up to a 89.7% improvement for wind turbine WTG14.

6.2.2 Universal modelling study

Another main inquiry from the wind industry is about the possibility of obtaining a universal methodology that performed accurately for all wind farm sites. Applied to ML modelling, a universal methodology could consist of performing a hyperparameter tuning by wind turbine for a specific ML technique. This perfectly works for SC as proved in section 6.1. However, performing a hyperparameter tunning by wind turbine for all ML techniques and then choosing the most accurate option by wind turbine may be very complex and might be seen as inconvenient from the industry point of view.

The ML modelling accuracy-complexity trade-off is exactly the point this section wants to address. The idea here is to generate a 'Universal model' by site in order to reduce the complexity in terms of number of models. Thus instead of obtaining three or four different wind turbine models, the SC based on ML models would result in one single model by wind farm. Moreover, further action has been taken to reduce the methodology complexity by avoiding hyperparameter tuning. Table 6.4 shows the Universal hyperparameters selected for all sites from the resulting best hyperparameters presented in section 6.1.2.

Table 6.4: Universal hyperparameters selected

Universal model hyperparameters		value
Linear	polynomial	Yes
	degree	2
	regularization	None
ANN	# hidden layers	3
	# neurons	80
	learning rate	0.001
	optimizer	Adam
	activation function	eLU
	regularization	None
XGB	# estimators	2000
	learning rate	0.01
	max_depth	10
	subsample	0.6
	colsample_bytree	0.7
	min_child_weight	10
	gamma	0
	reg_lambda	1

This new methodology would be aligned with both wind industry and ML perspectives. On the industry side, performing one single model training by wind farm is more convenient for the wind farm owner. And, since it is assumed that the correlations between meteorological variables are similar among wind turbines from the same site, it is considered that a ML model can adapt to a SC regression by wind farm.

Nevertheless, Universal modelling is slightly different from the one described in section 5.1. First, regarding data conditioning, wind turbine individual datasets are appendend to an overall dataset by site. Aftwerards, the datasets are split and normalized. The Universal model is trained with the overall dataset from the wind farm, which includes the data from all wind turbines from the site. However, when it comes to Generalization, the Universal model is evaluated on the individual test sets by wind turbine.

The main outcome from the Universal modelling study is that a Universal model by wind farm is not as accurate as a regular individual model by wind turbine. However, generally, a Universal model still performs better than the IEC standards for all ML techniques. This applies to all wind turbines except for the wind turbine T11. Fig. 6.7 shows average MAPE wind speed error for the regular individual modelling

by wind turbine and the Universal modelling by wind farm compared to the IEC standards methodology. Although Universal Polynomial is the technique that generalizes best compared to its regular modelling version, XGB is the Universal model which performs best in average.

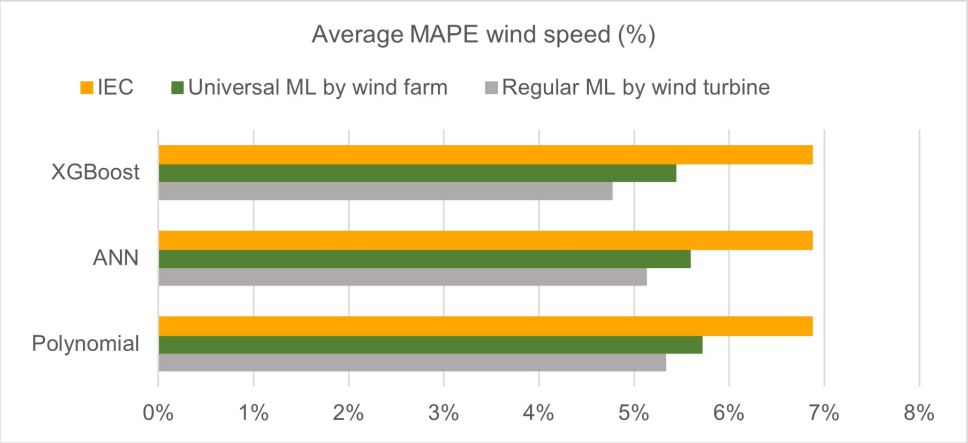


Figure 6.7: Average MAPE wind speed error by Machine Learning technique

6.2.3 Sensor importance

Nowadays, several types of WME at different heights are being installed in the met masts for the prior SC study phase of any wind project in complex terrain. Anemometers, wind vanes, thermometers, barometers, and hygrometers among others are being installed for several months to collect huge amounts of data per wind turbine.

From all those sensors, only the anemometer and the wind vane at the hub height signals are being considered by the IEC standards. Even if only 3 variables are used, the complete WME is still being installed and paid for by the wind farm owner. Thus, this section of the Thesis tries to shed some light on which of those sensors help to improve SC regression and which sensors cost could be avoided in a near future.

SHAP values is also the tool used in this Thesis for ranking the sensor importance. This ML model explainability tool provides the user with feature importance. However, from the wind industry insights learned, the importance is required to be measured by sensor.

One sensor may provide the user with more than one variable, and the other way around, one variable may have more than one required source of sensing. On the other hand, if one sensor is being installed in a met mast, all signals coming provided by this sensor will be available. Please check Table 4.1 to understand the relation between sensors and variables.

Thus, one simple methodology has been developed to transform feature importance to sensor importance using Eq. 6.1:

$$importance_{\text{sensor}} = \sum_{i=1}^{n} \left(\frac{importance_{\text{feature}}}{\sum sensors} \right)_i \tag{6.1}$$

The sensor importance is computed as the sum of the n variables' importance which the sensor provides. In case one variable is obtained through the combination of more than one sensor signal, the feature SHAP value is equally distributed to the sum of all those sensors. For instance, Wind Shear is obtained from WS1 and WS4 signals. Thus, the feature importance of Wind Shear is divided in two and summed for WS1 and WS4 sensors importance.

Fig. 6.8 shows the average sensor importance by ML technique. This average is computed using the same procedure as for the feature importance described in subsection 6.1.3.

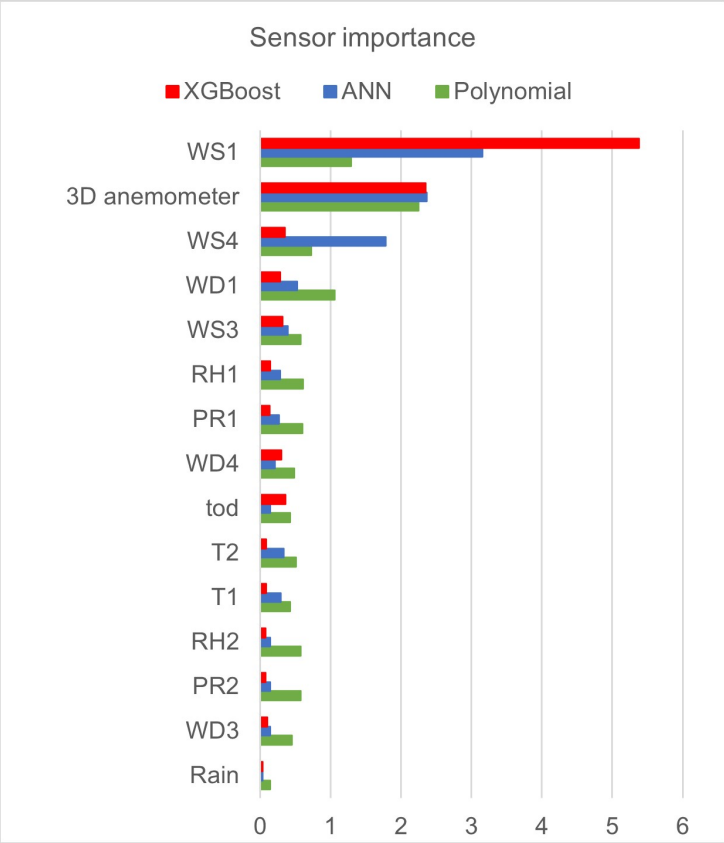


Figure 6.8: Sensor importance averaged for all sites by Machine Learning technique

7

Conclusions and recommendations

Contents

Chapter 7 presents the main outcomes and conclusions of this Master Thesis. Provides some recommendations and suggestions regarding the procurement of a Site Calibration methodology. It also offers a detailed review of the additional work that would be necessary to carry out around the subject dealt with in this Thesis.

7.1 Conclusions

The work of this Master Thesis has been focused on improving the standard SC for wind turbines on Power Curve Measurement campaigns. The most broadly accepted methodology for pursuing a SC procedure is the IEC 2005 standards, and thus that methodology is used as a baseline.

Thanks to the SC-PC methodology developed by the Power Curve Verification Team of Vestas, described in section 3.3.2, it has been proved that an uneven SC can negatively affect the MPC estimation pursued on Power Curve Measurement campaigns. In this sense, the IEC 2005 SC standard method systematically overestimates the wind speed around the rated power. The two main consequences are a significant decrease in the MPC at its elbow and in some cases, a reduction in the MAEP by default. Moreover, the wind speed is underpredicted on the lowest part of the PC and thus, the MPC tends to increase around that area of the curve. Therefore, the inaccuracy of wind speed prediction for SC procedures in some cases increases the risk of compensation on behalf of the turbine supplier.

ML supervised learning tools, specifically applied to regression tasks, are more accurate than the IEC in predicting the wind speed at the hub height and wind turbine location when performing a SC procedure. In most cases, the systematic errors introduced in the MPC by the IEC standard method are corrected by the ML models implemented. Thus, including other meteorological variables rather than the wind speed and wind direction at the hub height in the SC regression task improves the accuracy of the wind speed prediction. Especially, non-linear ML models, which can adapt to the non-linear relationships between the input data and the target for instance Turbulence Intesity, outperform when pursuing a SC task.

Generally, ML techniques applied to SC improve the accuracy of the MPC and the MAEP estimation compared to the IEC 2005 standards. Especially, XGB is the most accurate of the ML techniques implemented. This main conclusion can lead to other two relevant outcomes namely the assurance of development of new renewable projects as well as the integration of wind farms into the network.

First, regarding the main motivation of this Master Thesis, ML models applied to SC can potentially reduce the risk of compensation on behalf of the turbine supplier due to wind turbine underperformance. Moreover, it also can ensure the certainty of the investment return and the profitability for the wind farm owner which is the key for the development and execution of renewable projects.

On the other hand, unlike conventional power plants, wind farm power production is entirely depen-

dent on environmental conditions at each wind turbine location. However, wind speed at the hub height and turbine location not only depends on an intermittent and variable wind resource at the reference location but also on complex non-linear atmospheric interactions. These reference met masts data is later used for power production prediction for the wind farm daily operation. Thus, the increase in the accuracy of the power prediction thanks to a more accurate SC procedure is recognized as a major contribution to reliable large-scale wind power integration.

Model explainability tools can be of great support for implementing ML modelling in the wind industry. One of the main weaknesses of ML models is the poor transparency that this mathematical modelling may provide to its user. For many engineers and data scientists, ML may be a black box tool into which inputting data sometimes leads to an incomprehensible good result. Not understanding the functioning of these models can be a handicap for their implementation. Nevertheless, modern explainability tools for ML models, such as SHAP values, can increase the transparency of the models. This can be translated to increasing acceptance among the wind industry members.

Moreover, model explainers applied to SC can also be useful for optimizing the budget of Power Curve Measurement campaigns which are paid by the wind farm owner. Nowadays, several types of WME at different heights are being installed in the met masts for the prior SC study phase of any wind project in complex terrain. Anemometers, wind vanes, thermometers, barometers, and hygrometers among others are being installed for several months to collect huge amounts of data per wind turbine. From all those sensors, only the anemometer and the wind vane at the hub height signals are being considered by the IEC standards. Even if only three variables are used, the complete WME is still being installed and paid for by the wind farm owner. However, model explainers can rank the importance of sensors for SC and acknowledging the most relevant equipment can potentially reduce the cost of Power Curve Measurement campaigns.

Finally, it is concluded that Universal ML models by farm can perfectly be an option for pursuing a SC. The ML modelling accuracy-complexity trade-off is one of the main inquiries from the wind industry. Thus, although Universal models have not proved to be as accurate as regular individual models by wind turbine, they still perform better than the IEC standards. The idea behind 'Universal models' is to increase convenience regarding ML techniques, since Universal models reduce the complexity in terms of the number of models to tune and train.

7.2 Recommendations

It is recommended to abandon the multi-binning linear regression methods and to adopt the multivariate non-linear regression models. It is especially recommended to consider the XGB modelling implementation due to its outperforming results shown in this Master Thesis research. XGB is recommended also

because it is the most selective ML technique which means that fewer sensors are required for obtaining a more accurate result.

It is also recommended to include other wind and meteorological variables for SC rather than the wind speed and wind direction at the hub height, namely other wind-related variables at the middle tip and lower tip levels. It is especially important to include Turbulence Intensity in the non-linear modelling, as it has been proved to be as important or more than Wind Shear for the different ML techniques. In addition, temperature, pressure, relative humidity and air density at the hub height may be relevant input data for SC.

Regarding the WME, since the ultrasonic anemometer is the second-ranked sensor in the list of more important sensors for SC, it is recommended to install this sensor in the reference met masts and consider its signals as ML models inputs for SC. On the other hand, it is suggested to further study the necessity of measuring temperature, relative humidity, pressure and rain at the ground level so that the purchase and installation cost of those sensors could be avoided in wind projects.

7.3 Further work

Some further work that could be carried out around the possibility of a new standard SC methodology based on ML modelling is listed as follows:

- Since this Master Thesis only considered three different wind farms, future work could be to perform a further study of the effect of using Universal models by wind farm with more wind farm projects to obtain a more general picture and try to prove the validity of Universal models. This study could also include an analysis of the impact on the MPC estimation accuracy of Universal models applied to SC.

- Pursue a complete study with more datasets to try to find out the optimal number of sensors (if one) required for conducting a SC using different ML techniques. The research should be seen as a study of the trade-off between the number of meteorological variables available as the input data of the ML models and the cost of obtaining required data, namely the purchase and installation cost of the necessary sensors. The outcome of the experimental study may be a final list of sensors that should be able to provide enough information for the input dataset of the ML model and be economically sustainable.

- Wind speed and wind direction related variables registered at different heights are of great importance for SC regression task. However, information regarding the environmental conditions at the upper part of the wind turbine rotor is not available due to met masts height limitations. Thus, a

possible future work could be developing experimental research on improving the SC regressions by using upper rotor sensors.

- Since AEP metrics although being a good estimate of the impact of a SC method, are not always aligned with the presented results of MPC improvement, as is the case of the ANN modelling. This is probably because WSD does not provide the same relevance to all wind speed bins, and thus the resulting AEP metrics not only depend on the SC method used but also on the site. Thus, another possible analysis could consist of making research in pursuit of a non-site dependent AEP metric.

Bibliography

[1] I. R. E. A. (IRENA), "Wind energy data," 2021. [Online]. Available: https://www.irena.org/wind

[2] I. E. A. (IEA), "Renewables 2020 – Analysis," 2020. [Online]. Available: https://www.iea.org/reports/renewables-2020

[3] P. H. Alfredsson and A. Segalini, "Introduction Wind farms in complex terrains: an introduction," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 375, no. 2091, p. 20160096, Apr. 2017. [Online]. Available: https://royalsocietypublishing.org/doi/10.1098/rsta.2016.0096

[4] A. Albers, "Critical limitations of wind turbine power curve warranties," *Deutsche WindGuard GmbH*, 2012.

[5] Meteodyn, "BENCHMARK: Numerical Site Calibration – The Alaiz Case." [Online]. Available: https://meteodyn.com/news/renewable-energy/177-benchmark-numerical-site-calibration-the-alaiz-case.html

[6] J.-H. Jeong and K. Ha, "Evaluation of wind flow characteristics by rans-based numerical site calibration (nsc) method with met-tower measurements and its application to a complex terrain," *Energies*, vol. 13, p. 5121, 10 2020.

[7] G. Huang, "Numerical site calibration benchmark: the alaiz case. can flow models provide reliable site calibration factors to support power performance testing in complex terrain?" jul 2020. [Online]. Available: https://thewindvaneblog.com/numerical-site-calibration-benchmark-the-alaiz-case-b3767918d812?gi=fe78504027fe#:~:text=Numerical%20site%20calibration%20(NSC)%20consists,at%20a%20future%20turbine%20location.

[8] E. Cantero, F. Guillén, J. Rodrigo, P. De Azevedo Santos, J. Mann, N. Vasiljevic, M. Courtney, D. Martínez-Villagrasa, B. Martí, and J. Cuxart, *Alaiz Experiment (ALEX17): Campaign and Data Report: NEWA Deliverable Report D2.21*. NEWA - New European Wind Atlas, 2019.

[9] J. Sanz Rodrigo and G. Huang, “GitHub - windbench/alaizNSC: Numerical Site Calibration Benchmark: The Alaiz case.” [Online]. Available: https://github.com/windbench/alaizNSC

[10] ScienceDirect, “Topics & science materials. computational fluid dynamics,” https://www.sciencedirect.com/topics/materials-science/computational-fluid-dynamics.

[11] A. Bechmann, “Wasp cfd: A new beginning in wind resource assessment.” Technical University of Denmark- DTU, Tech. Rep., 2012.

[12] T. Uchida and G. Li, “Comparison of rans and les in the prediction of airflow field over steep complex terrain,” *Open Journal of Fluid Dynamics*, vol. 08, pp. 286–307, 01 2018.

[13] D. Matsushita, H. Matsumiya, Y. Hara, S. Watanabe, and A. Furukawa, “Studies on numerical site calibration over complex terrain for wind turbines,” *Science in China Series E: Technological Sciences*, vol. 53, 01 2010.

[14] I. Troen and E. Lundtang Petersen, *European Wind Atlas*. Risø National Laboratory, 1989.

[15] T. U. of Denmark, “ShowCFD demo 1 – Introduction,” dec 2020. [Online]. Available: https://www.youtube.com/watch?v=4a7rVz56dDc&list=PLrqli3B3J9m5I9gfw7I7thjiibKTJy9Em&index=1

[16] A. Schmitz, personal communication.

[17] *IEC 61400 – Part 12-1: Power Performance measurements of electricity producing wind turbines*, 2nd ed., International Electrotechnical Commission (IEC), 2017.

[18] *IEC 61400 – Part 12-1: Power Performance measurements of electricity producing wind turbines*, International Electrotechnical Commission (IEC), 2005.

[19] S. Lang and E. Mckeogh, “Lidar and sodar measurements of wind speed and direction in upland terrain for wind energy purposes,” *Remote Sensing*, vol. 3, pp. 1871–1901, dec 2011.

[20] R.-E. Keck, “A numerical investigation of nacelle anemometry for a hawt using actuator disc and line models in cfx,” *Renewable Energy*, vol. 48, pp. 72–84, 2012.

[21] C. Soraghan, “Part 2: Why don’t we see more Machine Learning in the wind industry?” July 2020. [Online]. Available: https://ore.catapult.org.uk/blog/part-2-why-dont-we-see-more-machine-learning-in-the-wind-industry/

[22] M. Lydia and G. Edwin Prem Kumar, “Machine learning applications in wind turbine generating systems,” *Materials Today: Proceedings*, vol. 45, pp. 6411–6414, 2021, international Conference on Mechanical, Electronics and Computer Engineering 2020: Materials Science. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214785320388775

[23] L. Fugon, J. Juban, and G. Kariniotakis, “Data mining for wind power forecasting,” *Proceedings European Wind Energy Conference & Exhibition EWEC 2008*, vol. 5, 03 2008.

[24] M. Optis and J. Perr-Sauer, “The importance of atmospheric turbulence and stability in machine-learning models of wind farm power production,” *Renewable and Sustainable Energy Reviews*, vol. 112, pp. 27–41, 09 2019.

[25] J. Nielson, K. Bhaganagar, R. Meka, and A. Alaeddini, “Using atmospheric inputs for artificial neural networks to improve wind turbine power prediction,” *Energy*, vol. 190, p. 116273, 10 2019.

[26] H. Demolli, A. S. Dokuz, A. Ecemis, and M. Gokcek, “Wind power forecasting based on daily wind speed data using machine learning algorithms,” *Energy Conversion and Management*, vol. 198, p. 111823, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0196890419308052

[27] X. Wen, “Modeling and performance evaluation of wind turbine based on ant colony optimization-extreme learning machine,” *Applied Soft Computing*, vol. 94, p. 106476, 2020.

[28] A. Clifton, L. Kilcher, J. Lundquist, and P. Fleming, “Using machine learning to predict wind turbine power output,” *Environmental Research Letters*, vol. 8, p. 024009, 06 2013.

[29] N. Mandal and T. Sarode, “Prediction of wind speed using machine learning,” *International Journal of Computer Applications*, vol. 176, pp. 34–37, 06 2020.

[30] S. Paramasivan, “Improved prediction of wind speed using machine learning,” *EAI Endorsed Transactions on Energy Web*, vol. 6, p. 157033, 07 2018.

[31] D. Vassallo, R. Krishnamurthy, and H. Fernando, “Utilizing physics-based input features within a machine learning model to predict wind speed forecasting error,” *Wind Energy Science*, vol. 6, pp. 295–309, 03 2021.

[32] M. P. Emhart, “Exploring machine learning models for wind speed prediction,” Master Thesis, Universitat Politècnica de Catalunya (UPC), jan 2018.

[33] M. A. E. Bhuiyan, F. Begum, S. Ilham, and R. S. Khan, “Advanced wind speed prediction using convective weather variables through machine learning application,” *Applied Computing and Geosciences*, vol. 1, p. 100002, 10 2019.

[34] A. F. Antor and E. Wollega, “Comparison of machine learning algorithms for wind speed prediction,” in *5th NA International Conference on Industrial Engineering and Operations Management Detroit, Michigan, USA*, 11 2020.

[35] J. Verhoef and G. Leendertse, “Identification of variables for site calibration and power curve assessment in complex terrain. task 8: A literature survey on theory and practice of parameter identification, specification and estimation (ise) techniques,” Energy research Centre of the Netherlands, ECN 7.4062 project report JOR3-CT98-0257, apr 2001, project Name: SiteParIden.

[36] G. Aurélien, *Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow*, 2nd ed. O’REILLY, sep 2019, powered by Jupyter.

[37] C. J. Marchena, “Assessment of the validity of the warranted power curve in nonstandard wind conditions,” Master Thesis, Delf University of Technonolgy (TUDelft), June 2017.

[38] P. C. V. Team, “Theoretical power performance estimation based on site calibration data.” Vestas, Internal Technical Report, jul 2020.

[39] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python,” *Nature Methods*, vol. 17, pp. 261–272, 2020.

[40] N. Sharma, “Ways to Detect and Remove the Outliers,” May 2018. [Online]. Available: https://towardsdatascience.com/ways-to-detect-and-remove-the-outliers-404d16608dba

[41] P. Bachant, B. Blaylock, R. Brown, S. Celles, F. Fernandes, D. Garver, F. Maussion, J. McCann, I. Ogasawara, J. Quick, L. Roubeyrie, B. R. D. Pinho, and S. Weber. (2018) Windrose. [Online]. Available: https://github.com/python-windrose/windrose

[42] D. O. Michael Waskom, Olga Botvinnik and P. Hobson. (2017, Sep.) mwaskom/seaborn: v0.8.1 (september 2017). [Online]. Available: https://doi.org/10.5281/zenodo.883859

[43] T. Kluyver, B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. Hamrick, J. Grout, S. Corlay, P. Ivanov, D. Avila, S. Abdalla, and C. Willing, “Jupyter notebooks – a publishing format for reproducible computational workflows,” in *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, F. Loizides and B. Schmidt, Eds. IOS Press, 2016, pp. 87 – 90.

[44] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds. Curran

Associates, Inc., 2017, pp. 4765–4774. [Online]. Available: http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf

[45] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.-I. Lee, "From local explanations to global understanding with explainable ai for trees," *Nature Machine Intelligence*, vol. 2, no. 1, pp. 2522–5839, 2020.

[46] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

[47] M. J. Grover and M. Maldonado, "Polynomial Features and Regularization Demo - Part 3 - Regression with Regularization Techniques: Ridge, LASSO, and Elastic Net." [Online]. Available: https://www.coursera.org/lecture/supervised-learning-regression/polynomial-features-and-regularization-demo-part-3-MwgnF

[48] S. Bhattacharyya, "Ridge and Lasso Regression: L1 and L2 Regularization," Sep. 2020. [Online]. Available: https://towardsdatascience.com/ridge-and-lasso-regression-a-complete-guide-with-python-scikit-learn-e20e34bcbf0b

[49] K. Willems, "Keras tutorial: Deep learning in python," dec 2019. [Online]. Available: https://www.datacamp.com/community/tutorials/deep-learning-python

[50] F. Chollet *et al.* (2015) Keras. [Online]. Available: https://github.com/fchollet/keras

[51] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: ACM, 2016, pp. 785–794. [Online]. Available: http://doi.acm.org/10.1145/2939672.2939785

[52] A. Jain, "XGBoost Parameters | XGBoost Parameter Tuning," Mar. 2016. [Online]. Available: https://www.analyticsvidhya.com/blog/2016/03/complete-guide-parameter-tuning-xgboost-with-codes-python/

[53] S. J. v. d. W. Charles R. Harris, K. Jarrod Millman and P. V. Ralf Gommers, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020. [Online]. Available: https://doi.org/10.1038/s41586-020-2649-2

[54] W. McKinney *et al.*, "Data structures for statistical computing in python," in *Proceedings of the 9th Python in Science Conference*, vol. 445. Austin, TX, 2010, pp. 51–56.

[55] J. D. Hunter, “Matplotlib: A 2d graphics environment,” *Computing in science & engineering*, vol. 9, no. 3, pp. 90–95, 2007.

[56] B. Bengfort and R. Bilbro. Yellowbrick. [Online]. Available: http://www.scikit-yb.org/

Appendix

Table of Contents

Literature Research

Table A.1: Legend of Literature Research tables- Input features

Acronym	Description
WS	Wind speed
WD	Wind direction
WG	Wind Gust
T	Temperature
PR	Pressure
RH	Relative Humidity
AD	Air Density
TI	Turbulence Intensity
WSH	Wind shear
PTG	Potential Temperature Gradient
TKE	Turbulent kinetic energy
OL	Obukhov length
SWR	Shortwave Radiation
Snow	Snow Depth
time	Datetime or time of the day
THF	Turbulent Heat Flux
Ri	Richardson number
LLJ	Low Level Jet
ULJ	Upper Level Jet
AV	Absolute Vorticipy
DP	Dewpoint
LapR	Lapse Rate
CC	Cloud Cover
Rain	Precipitation
Ozone	Ozone
UV	UV index
V	Visibility
SL	Sea Level
GS	Generator Speed
GOT	Gear Oil Temperature
GBT	Gearbox Bearing Temperature
GST	Generator Stator Temperature

Table A.2: Legend of Literature Research tables- Data source

Acronym	Description
NWP Systems	Numerical Weather Prediction Systems
Met mast	Meteorological mast

Table A.3: Legend of Literature Research tables- Metrics

Acronym	Description
RMSE	Root Mean Squared Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
CRMSE	Centered root mean square error
UR	Uncertainty Ratio

Table A.4: Legend of Literature Research tables- Machine Learning Techniques

Acronym	Description
LR	Linear Regression
LASSO	Lasso Regression
RIDGE	Ridge Regression
PolyR	Polynomial Regression
RF	Random Forest
ERT	Extremely Randomized Trees
SVM	Support Vector Machines
Bagg	Bagging-based models
k-NN	K Nearest Neighbour
GAM	Generalized Additive Model
GBM	Gradient Boosting Models
GPR	Gaussian Process Regression
FFBPN	Feed Forward Back Propagation Network
RBF	Radial Basis Function
NARX	Nonlinear AutoRegressive model with eXogenous inputs
MIFS	Mutual Information Feature Selection
BART	Bayesian Additive Regression Trees
QRF	Quantile Regression Forests
TSC	Time Series Clustering (Unsupervised)
ELM	Extrem Learning Machine

Table A.5: Literature research on Wind Speed prediction using Machine Learning Techniques

Reference	Year	Target	Data resolution	Data source	Input features	ML techniques	Evaluation metrics	Best model
[29]	2020	Average wind speed	-	Met mast	T, PR, WD, Rain	LR, FFBPN, RF	RMSE, MAE, MSE	RF
[30]	2019	Average wind speed	Minute	Met mast	SWR, Snow, T, PR, Rain, WD, SL, Others	FFBPN, RBF, NARX	RMSE, MAE	NARX with MIFS
[31]	2020	Wind speed forecasting error	Hourly	NWP Systems	WD, T, WS, time, TI, THF	RF	RMSE	RF
[32]	2018	Average wind speed	Six-hourly	Met mast & NWP Systems	T, PR, zonal WS, WSVer	k-NN, FFBPN, RF, Others	RMSE, MAE	NN & Ensemble
[33]	2019	Maximum wind speed	Three-hourly	NWP Systems	LLJ, ULJ, AV, DP, LapR, WSH, Others	BART, QRF	CRMSE, MAPE, UR, Others	QRF
[34]	2020	Average wind speed	Hourly	NWP Systems	CC, T, PR, DP, RH, Rain, UV, Ozone, WG, V, Others	RIDGE, PolyR, FFBPN	RMSE	FFBPN

Table A.6: Literature research on Wind power prediction using Machine Learning techniques

Reference	Year	Purpose	Target	Data resolution	Data source	Input features	ML techniques	Evaluation metrics	Best model
[23]	2008	Improving the integration of wind power through 2-3 day-ahead prediction.	Wind farm power production	Hourly	NWP Systems + turbine SCADA	WS, WG	LR, RF, Bagg, FFBPN, SVM, Others	RMSE, MAE	RF
[24]	2019	Modelling wind farm power with atmospheric variables as inputs	Wind farm power production	Hourly	Met mast + turbine SCADA	WS, WD, T, PR, PTG, TI, TKE, OL	FFBPN, GAM, GBM, ERT, SVM	RMSE, MAE	GBM, ERT
[25]	2020	Improving the accuracy of monthly energy estimates.	Wind turbine power production	Monthly	Met mast + turbine SCADA	WS, AD, TI, Ri, WSH	RBF, FFBPN, RF,	MAE	FFBPN
[26]	2019	Improving of wind farm projects development efficiency	Wind turbine power production	Daily	Met mast + turbine SCADA	WS	LASSO, k-NN, GBM, SVM, RF	RMSE, MAE	RF, SVM
[27]	2020	Performance evaluation of a wind turbine	Wind turbine power production	One second	Met mast+ turbine SCADA	WS, GS, GOT, GBT, GST	ELM	RMSE, MAE	ELM

B

Datasets detailed information

B.1 Met mast sensors' heights

The following subsection of the Appendix B shows the height of the wind measurement equipment devices for the different PM.

Table B.1: Sensors heights for the different Permanent Masts

Sensors' height (m)	Dataset1	Dataset2			Dataset3	
	PC02	PM1112	PM17	PM22	PCV02	PCV01
WS1	91	90	84	90	104.9	104.9
Target	91	90	84	90	104.9	104.9
WS2	91	90	84	90	104.9	104.9
WS3	57	60.75	56	60.75	67.2	67.2
WS4	23	31.5	84	31.5	30.2	30.2
WD1	88	88	82	88	103.1	103.1
WD2	88	-	-	-	-	-
WD3	57	-	-	-	67.2	67.2
WD4	23	29.5	26	29.5	30.2	30.2
T1	87	82	82	88	102.6	102.2
T2	3	-	-	-	1.8	7.7
PR1	87	82	82	88	102.6	102.2
PR2	9	-	-	-	1.8	1.8
RH1	87	82	82	88	102.6	102.2
RH2	9	-	-	-	9.5	7.7
3D anemometer	88	88	82	88	103.1	103.1
Rain gauge	3	3	3	3	4	4

B.2 Post-processing variables formulas

The following subsection presents the formulas used to compute the post-processing variables during data pre-processing.

B.2.1 Air Density [kg/m^3]

Air Density is computed using several signals: temperature, air pressure and relative humidity.

$$AD1 = \frac{1}{T1 + 273.15}\left[\frac{PR_{\text{hub}}100}{R_{\text{o}}} - \frac{RH1}{100}\, p_{\text{vapor}}\left(\frac{1}{R_{\text{o}}} - \frac{1}{R_{\text{w}}}\right)\right] \tag{B.1}$$

- Where PR_{hub} is the pressure at the hub height [hPa], which may not correspond to the height at which the pressure sensor is set. Thus:

$$PR_{\text{hub}} = PR1 - \Delta z\, \Delta PR \tag{B.2}$$

  - Delta height [m], $\Delta z = z_{\text{hub}} - z_1$

  - Delta pressure [hPa], $\Delta PR = \left[PR_{\text{o}} - PR_{\text{o}}\, e^{\left(\frac{-9.8}{R_{\text{o}}(15+273.15)}\right)}\right] 0.01$
  - Standard Pressure, $PR_{\text{o}} = 101325$ Pa

- Dry air constant

$$R_{\text{o}} = 287.05\ J/(kgK) \tag{B.3}$$

- Water vapor gas constant

$$R_{\text{w}} = 461.5\ J/(kgK) \tag{B.4}$$

- Vapor pressure [Pa]

$$p_{\text{vapor}} = 0.0000205\, e^{(0.0631846\ (T1+273.15))} \tag{B.5}$$

B.2.2 Turbulence Intensity [-]

Turbulence Intensity is computed only for wind speed at the hub height. It is the ratio between the wind speed standard deviation and the average.

$$TI = \frac{WS1_{\text{std}}}{WS1_{\text{avg}}} \tag{B.6}$$

B.2.3 Wind Shear [-]

Wind Shear is computed for the wind profile described by the wind speed at lower tip level and that at the hub height.

$$WSH = \frac{\log\left(\frac{WS1}{WS4}\right)}{\log\left(\frac{z_1}{z_4}\right)} \tag{B.7}$$

B.2.4 Wind Veer [°]

Wind Veer is computed for the difference in wind direction for that at the hub height and at the lower tip level.

$$WVeer = \begin{cases} WD1 - WD4 - 360, & \text{if } WD1 - WD4 > 180 \\ WD1 - WD4 + 360, & \text{if } WD1 - WD4 < -180 \\ WD1 - WD4, & \text{otherwise.} \end{cases} \tag{B.8}$$

B.2.5 Inflow Angle [°]

Also known as Vertical Wind Direction (WDVer), it is computed as the angle between the horizontal and the vertical wind speed measured by the ultrasonic anemometer.

$$WDVer = \frac{-180}{\pi} \arctan\left(\frac{WSVer}{WSHor}\right) \tag{B.9}$$

B.3 Power Curves

The corresponing Power Curves for the different wind turbine Vestas models are presented in this section. These curves are normalized for the reference air density depending on the site of the project. For Dataset1 site the reference air density is 1.15 kg/m^3. For Dataset2, 1.225 kg/m^3 and for Dataset3, 1.175 kg/m^3.

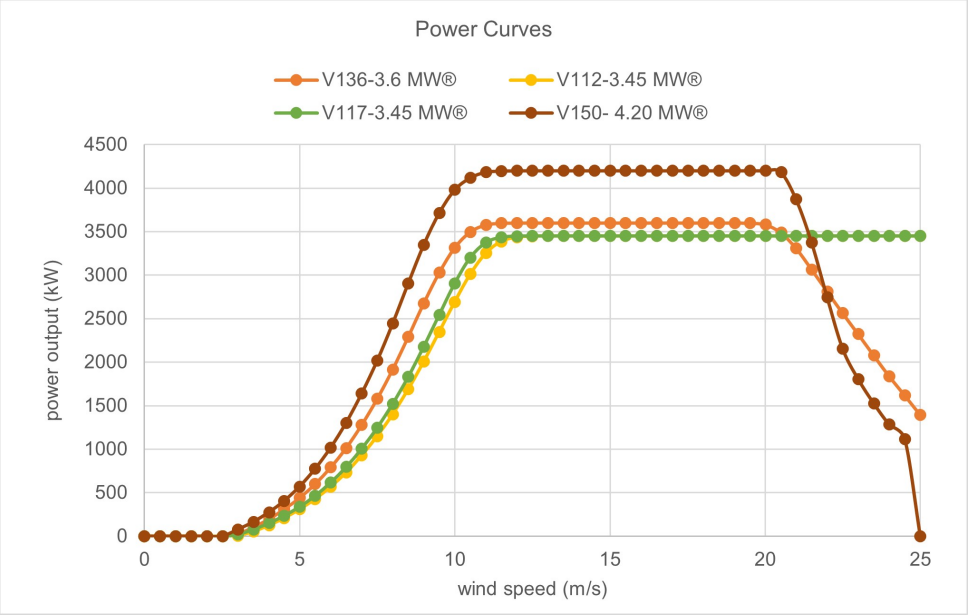


Figure B.1: Power Curves for the different wind turbine Vestas models

B.4 Annual Wind Speed Distributions

The corresponing Annual Wind Speed Distributions for the three different sites are presented in this section. The distribution is measured in hours per year.

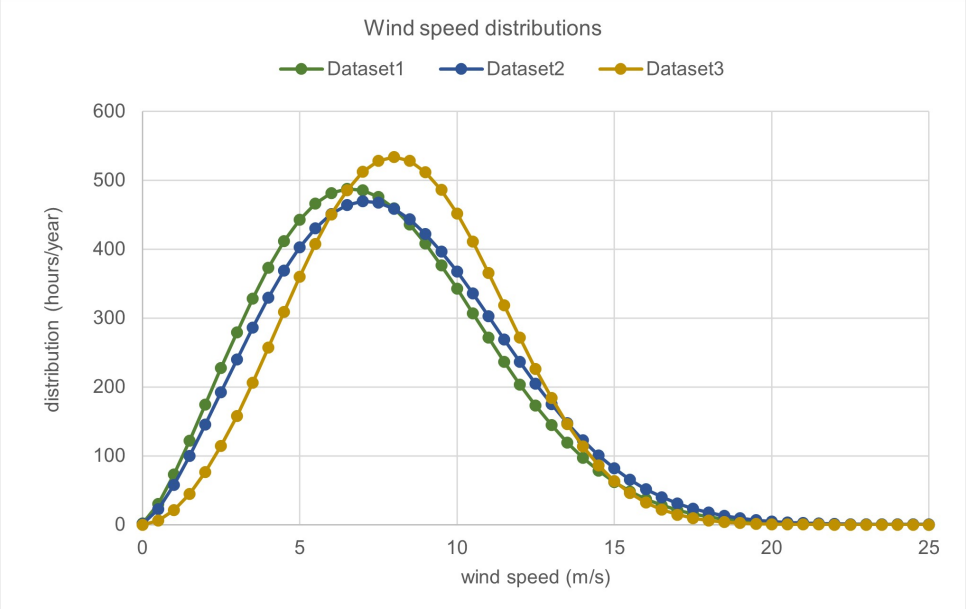


Figure B.2: Wind speed distributions for three different sites

C

Data exploration

C.1 Dataset1

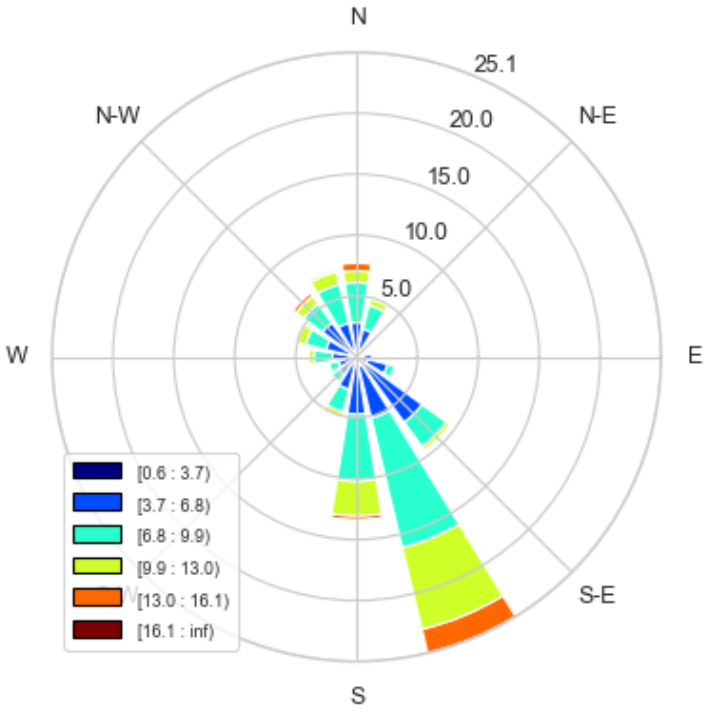


Figure C.1: Wind Rose of WD1 for wind turbine WTG14

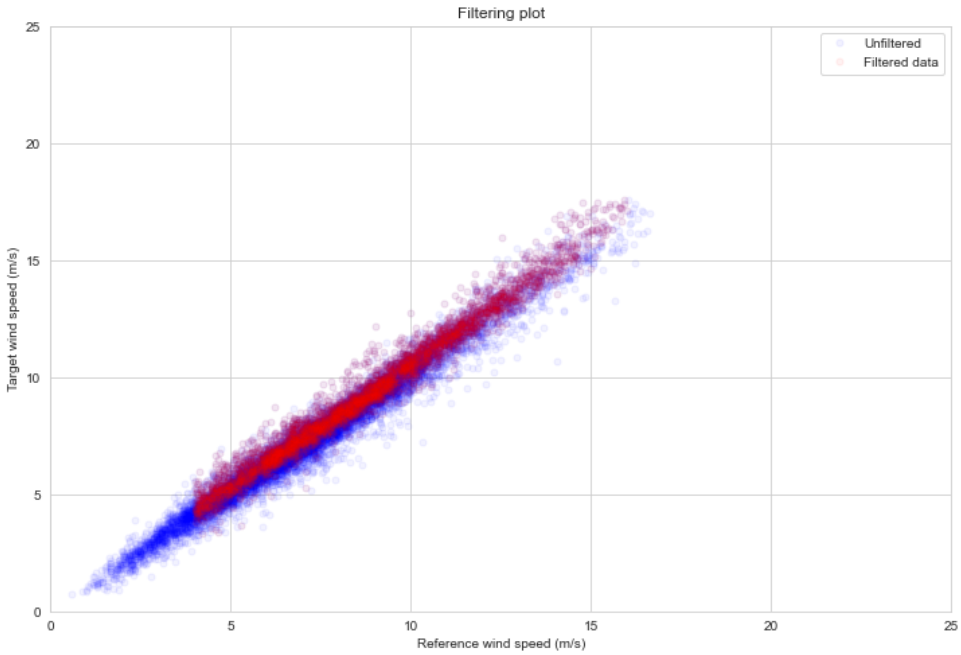


Figure C.2: Filtering comparison for wind turbine WTG14

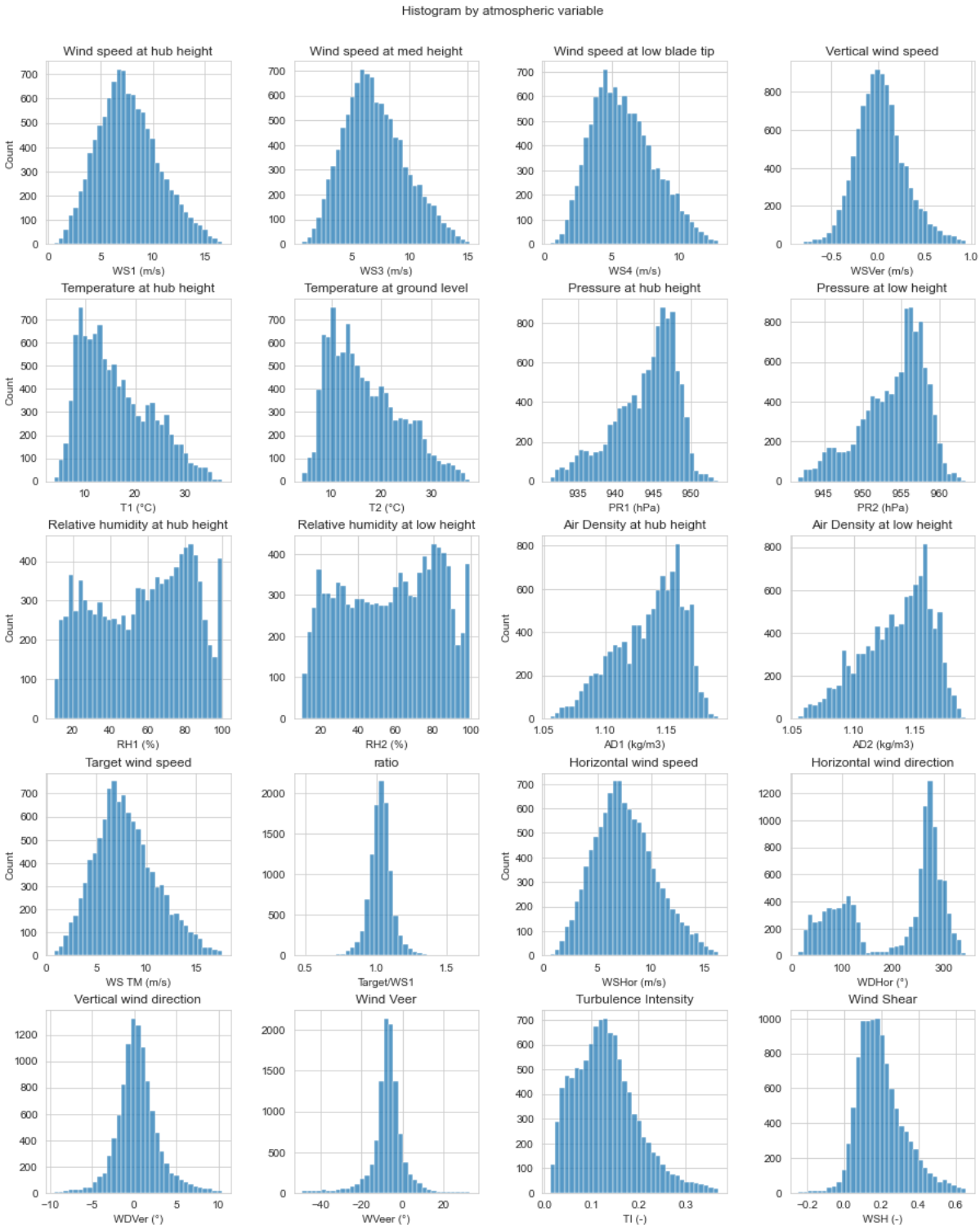


Figure C.3: Histograms for wind turbine WTG14

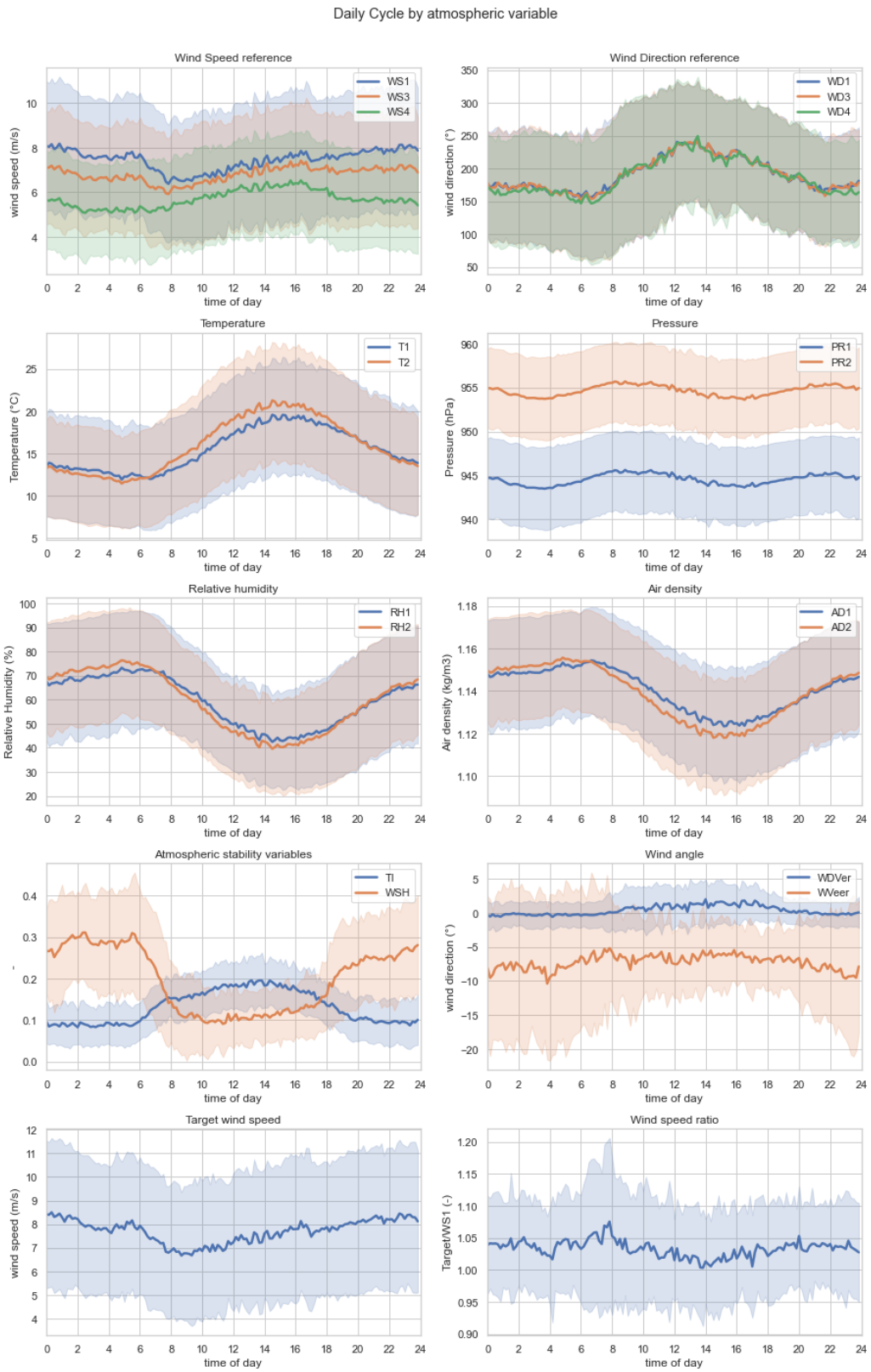


Figure C.4: Daily Cycle plot for wind turbine WTG14

C.2 Dataset2

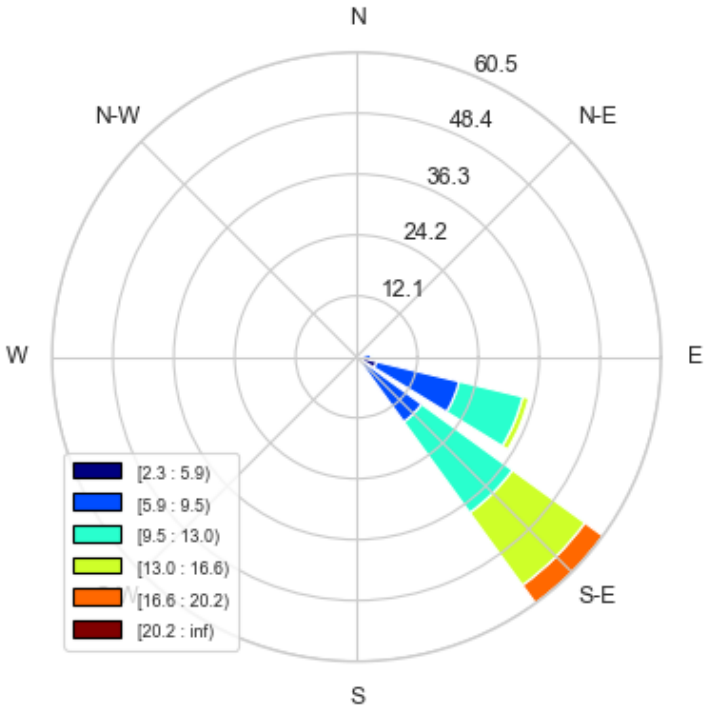


Figure C.5: Wind Rose of WD1 for wind turbine T11

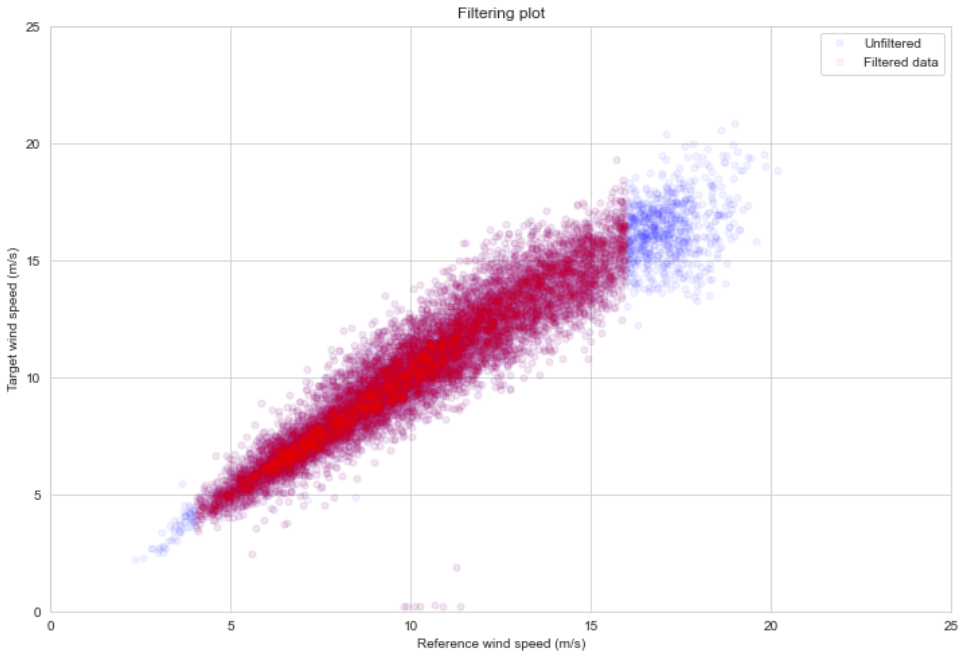


Figure C.6: Filtering comparison for wind turbine T11

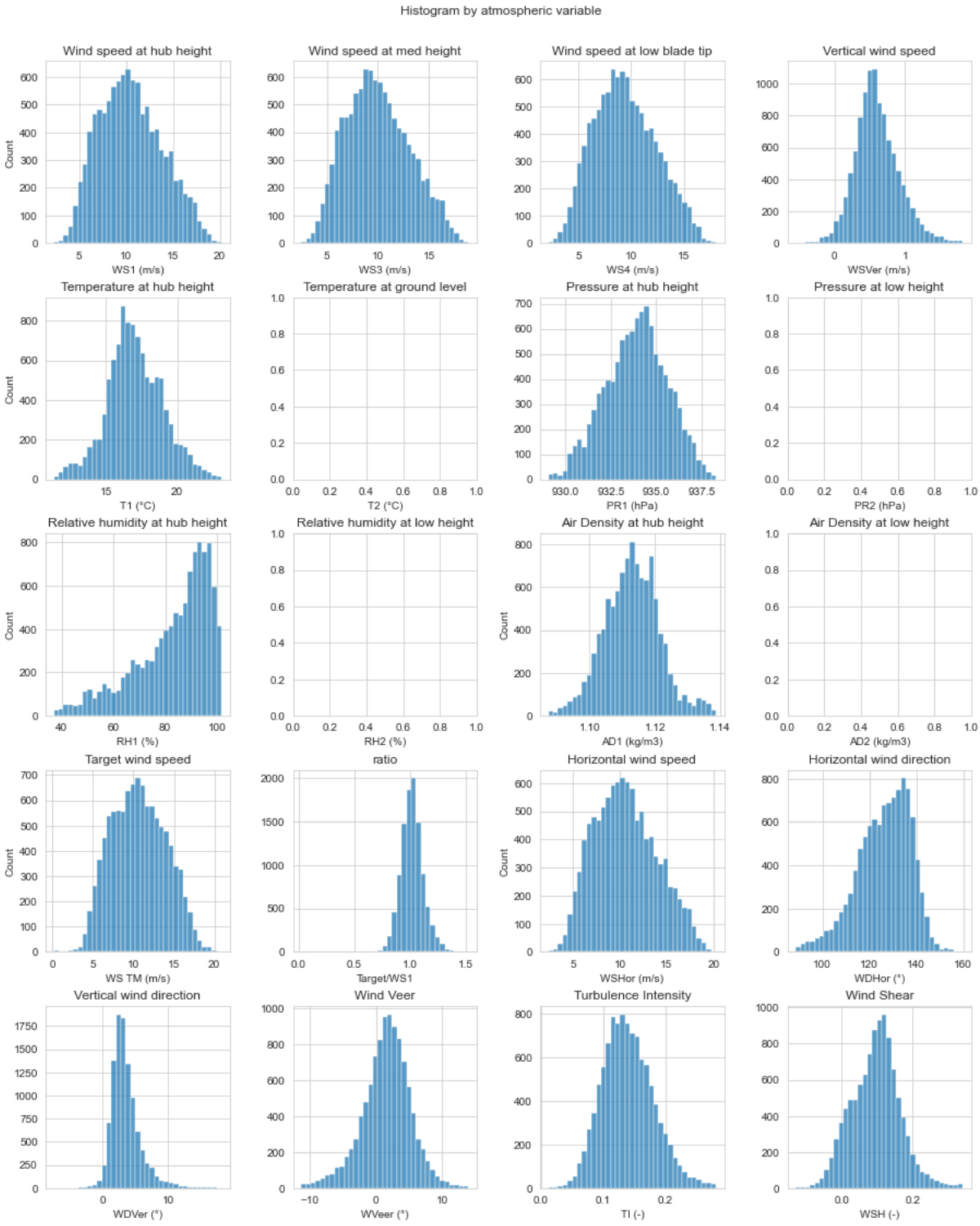


Figure C.7: Histograms for wind turbine T11

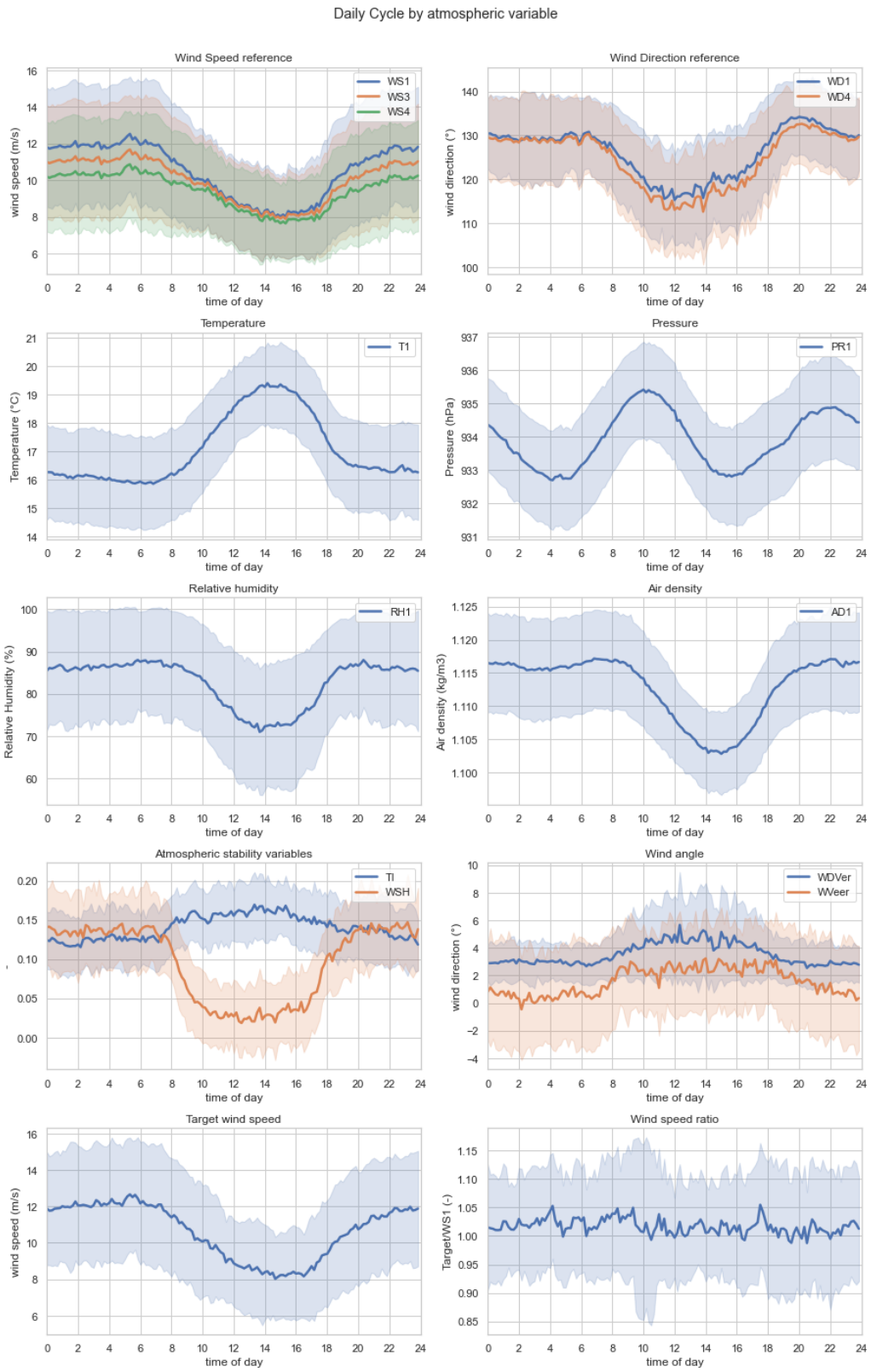


Figure C.8: Daily Cycle plot for wind turbine T11

C.3 Dataset3

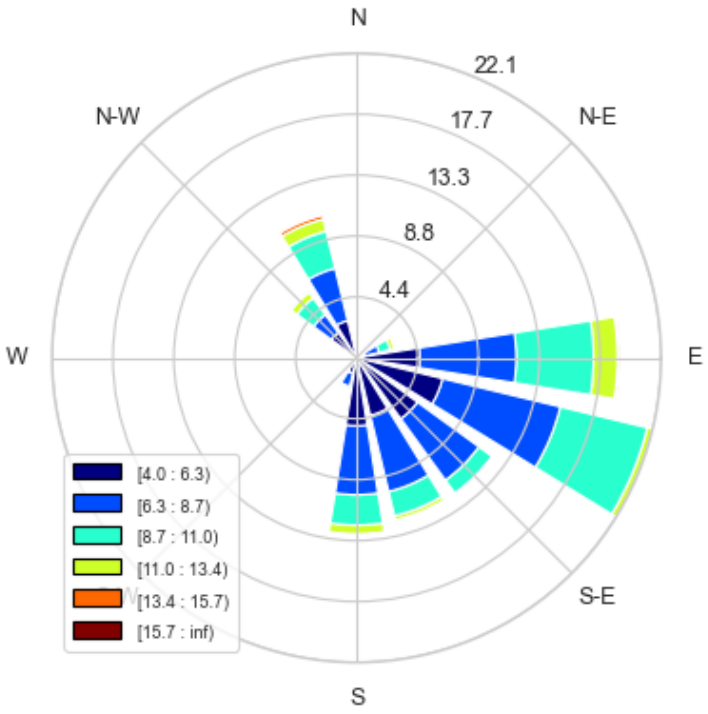


Figure C.9: Wind Rose of WD1 for wind turbine WTG18

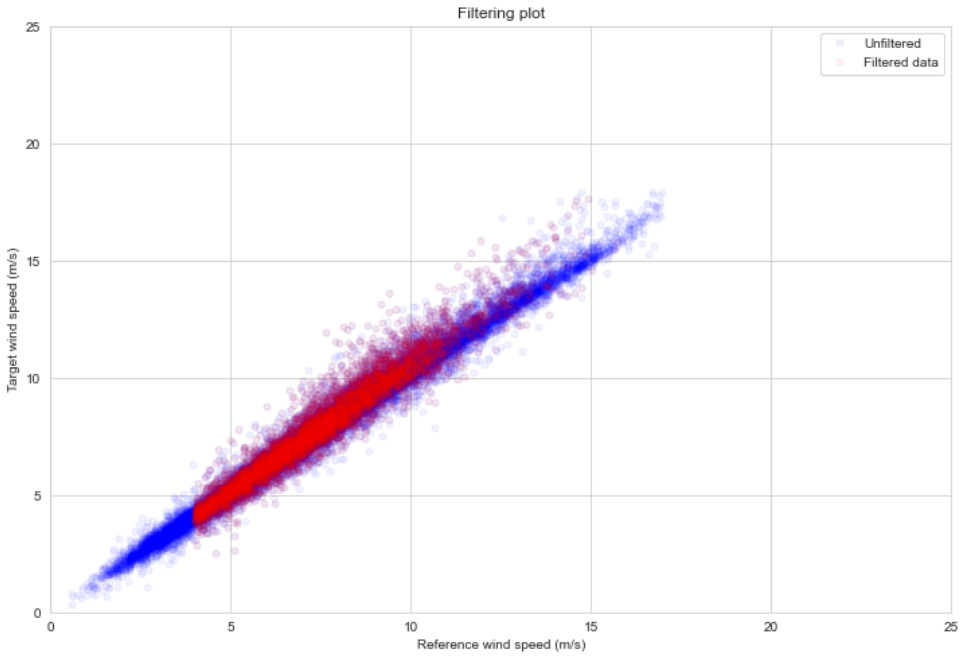


Figure C.10: Filtering comparison for wind turbine WTG18

Figure C.11: Histograms for wind turbine WTG18

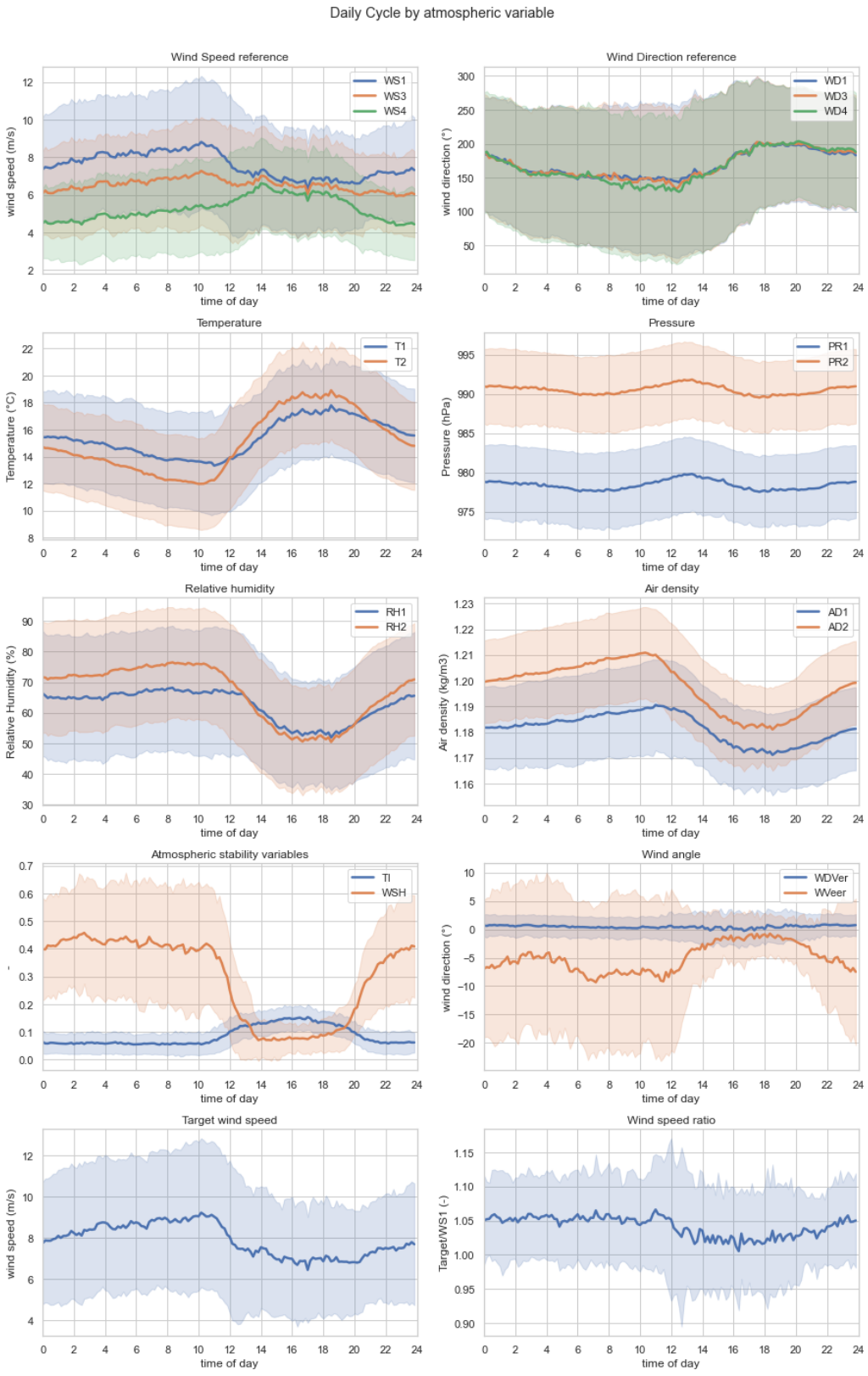


Figure C.12: Daily Cycle plot for wind turbine WTG18

D

IEC baseline

D.1 Dataset1

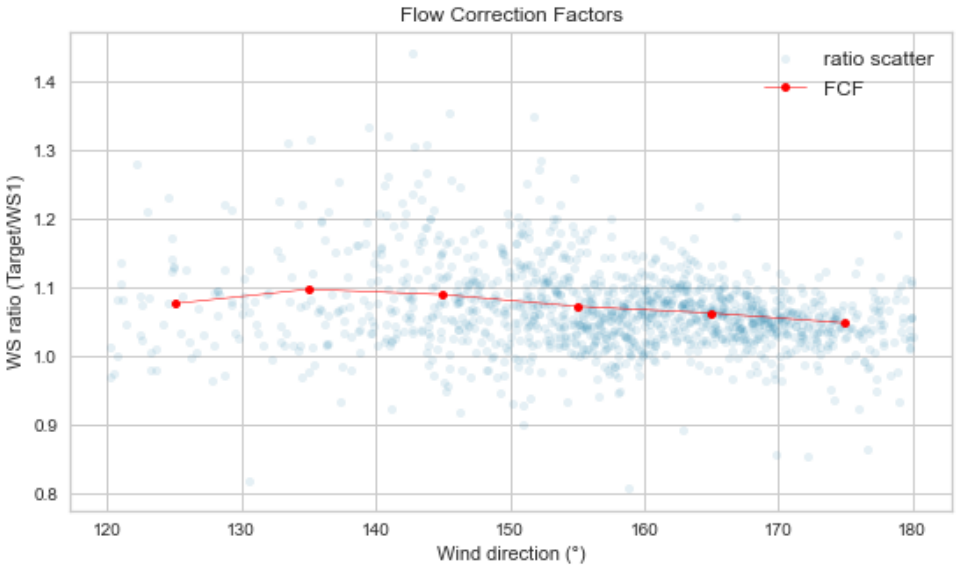


Figure D.1: Flow Correction Factors plot for wind turbine WGT14

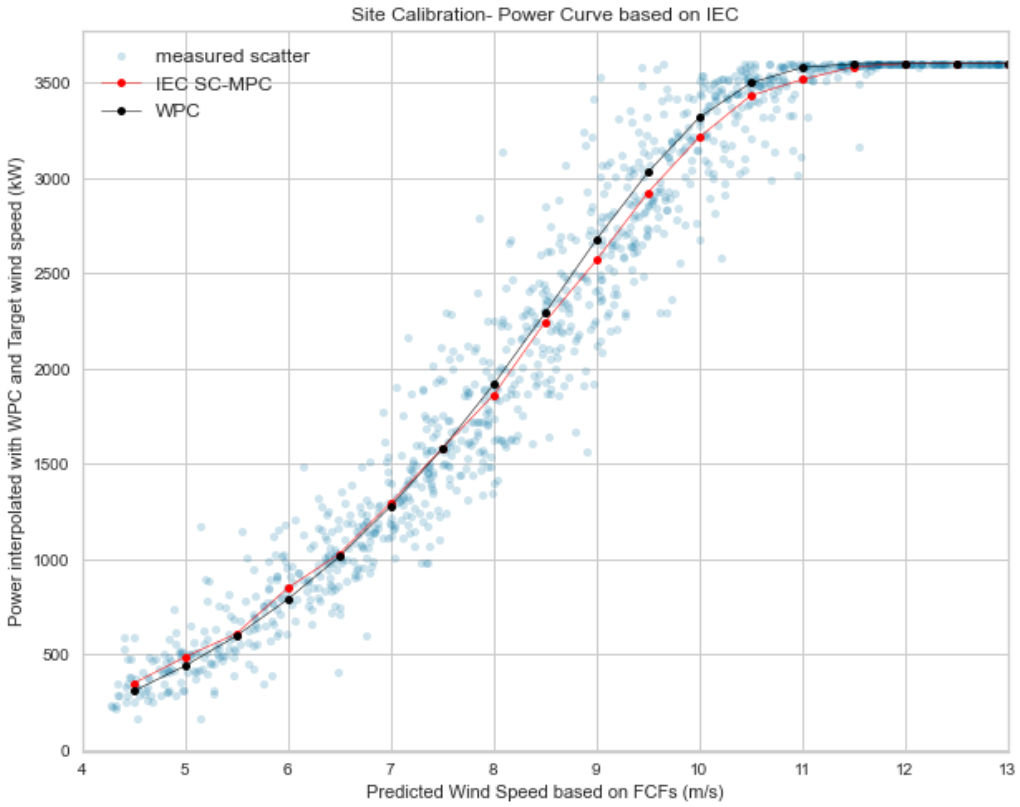


Figure D.2: Site Calibration Power Curve based on IEC FCF plot for wind turbine WGT14

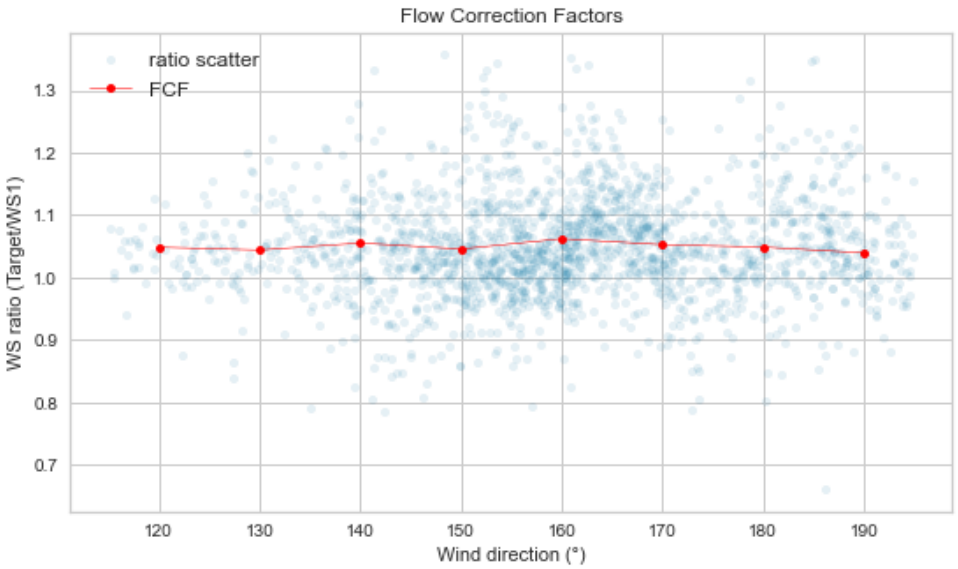


Figure D.3: Flow Correction Factors plot for wind turbine WGT15

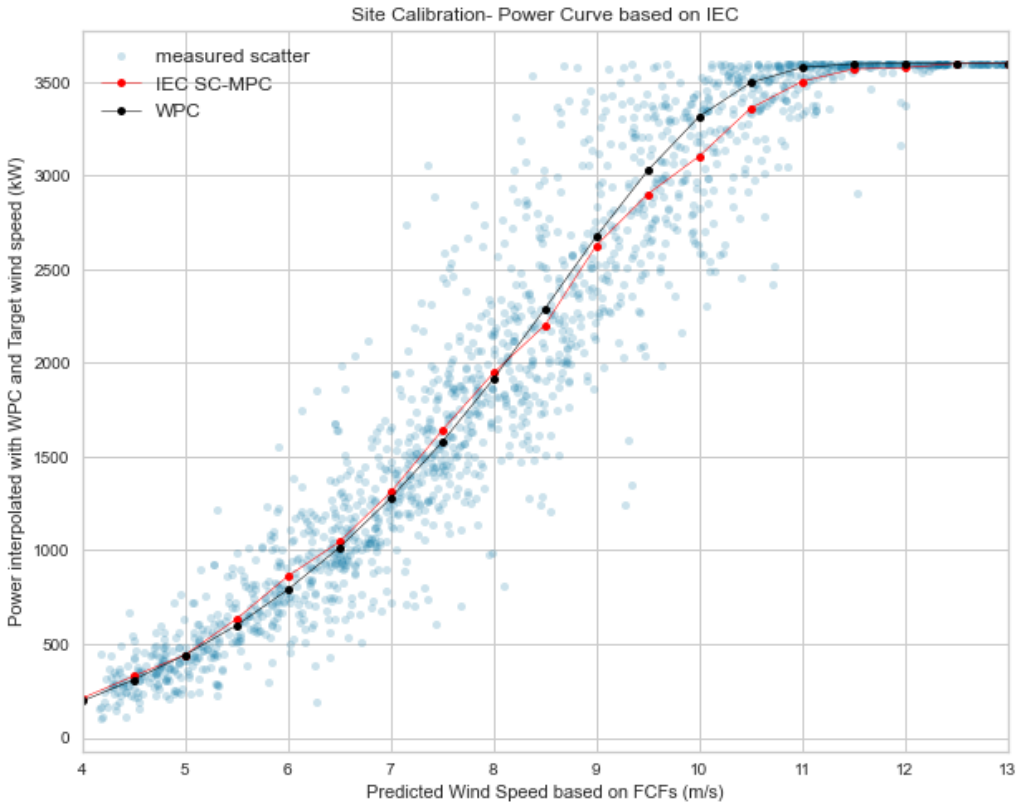


Figure D.4: Site Calibration Power Curve based on IEC FCF plot for wind turbine WGT15

D.2 Dataset2

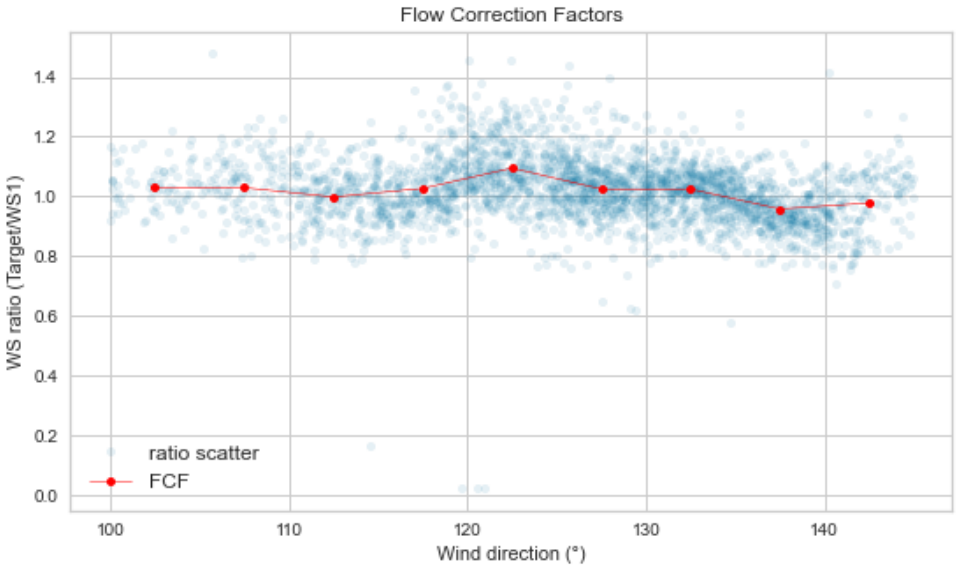


Figure D.5: Flow Correction Factors plot for wind turbine T11

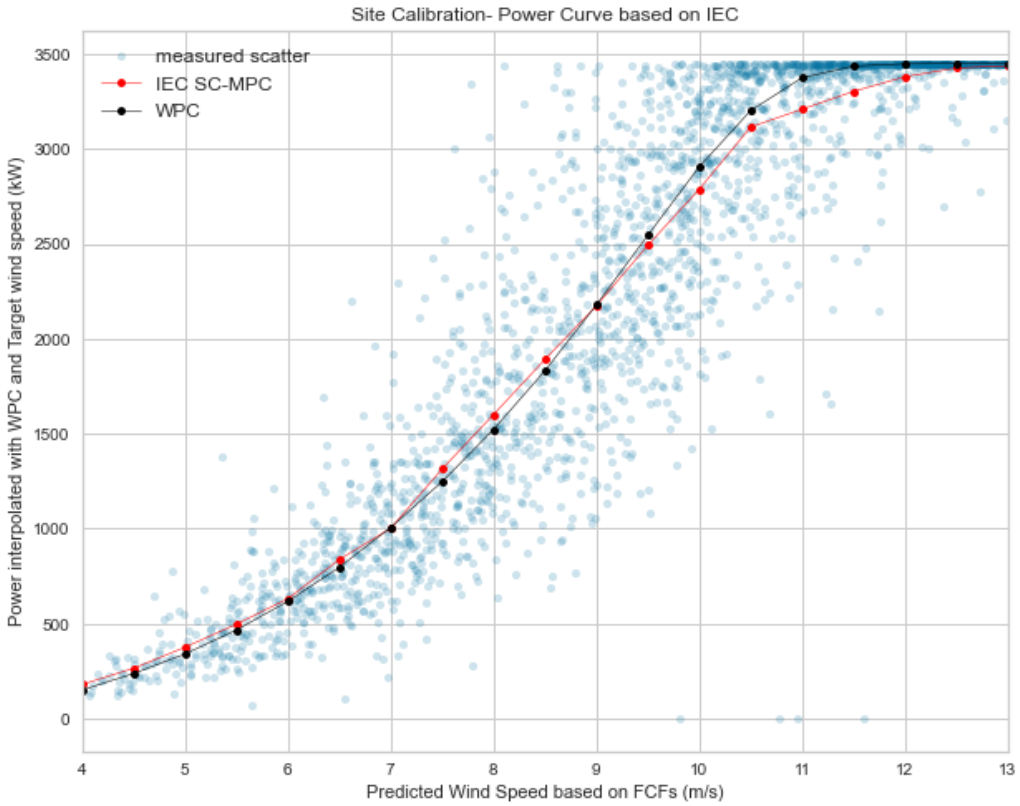


Figure D.6: Site Calibration Power Curve based on IEC FCF plot for wind turbine T11

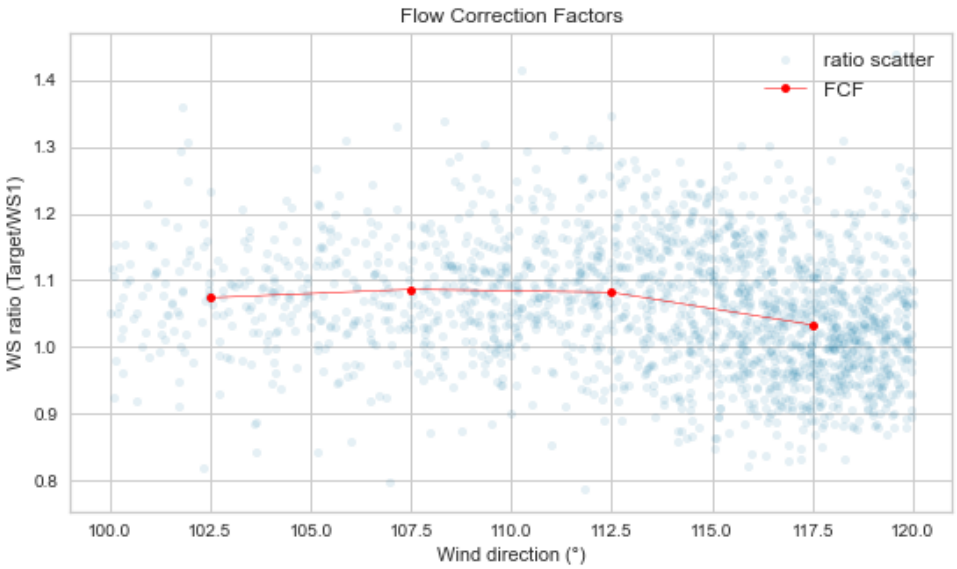


Figure D.7: Flow Correction Factors plot for wind turbine T17

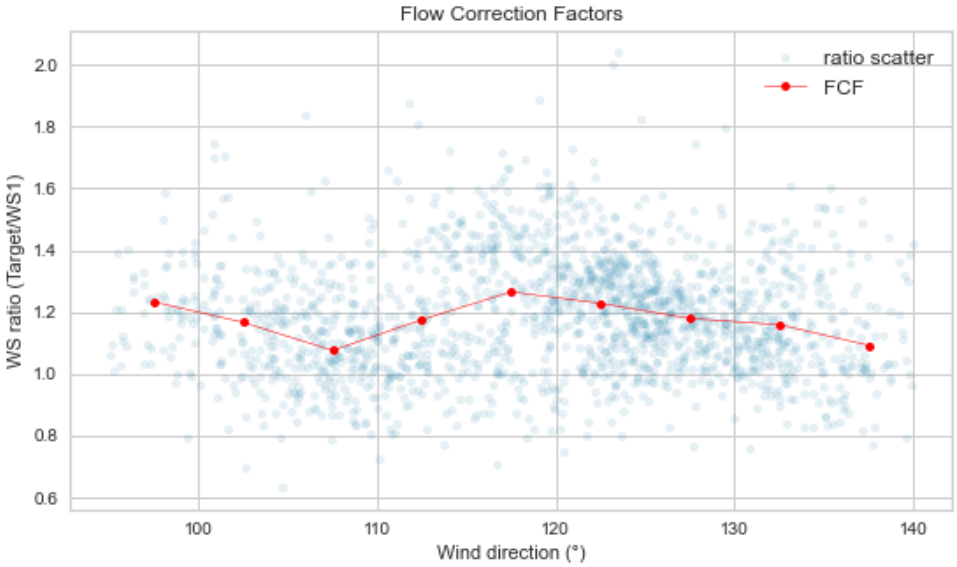


Figure D.8: Flow Correction Factors plot for wind turbine T22

D.3 Dataset3

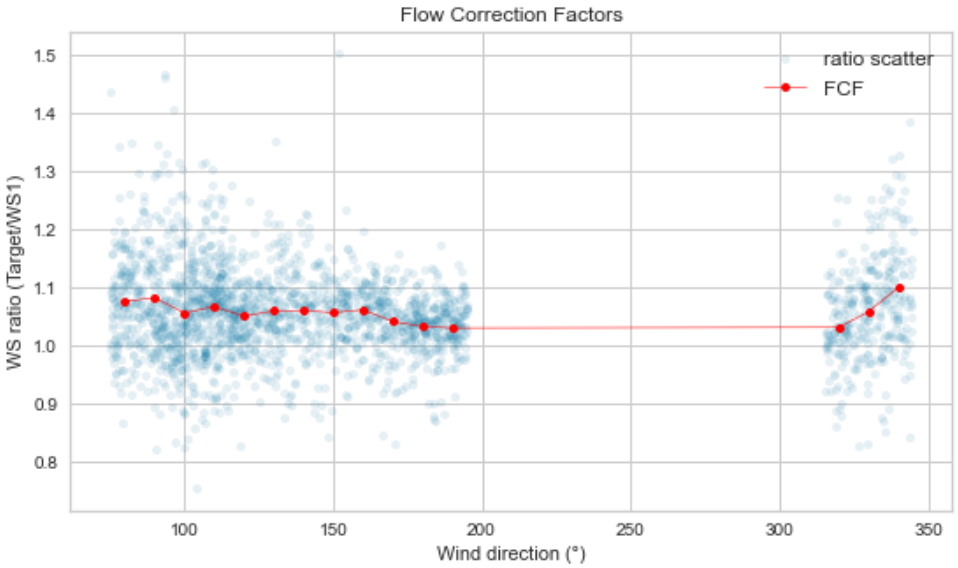


Figure D.9: Flow Correction Factors plot for wind turbine WTG18

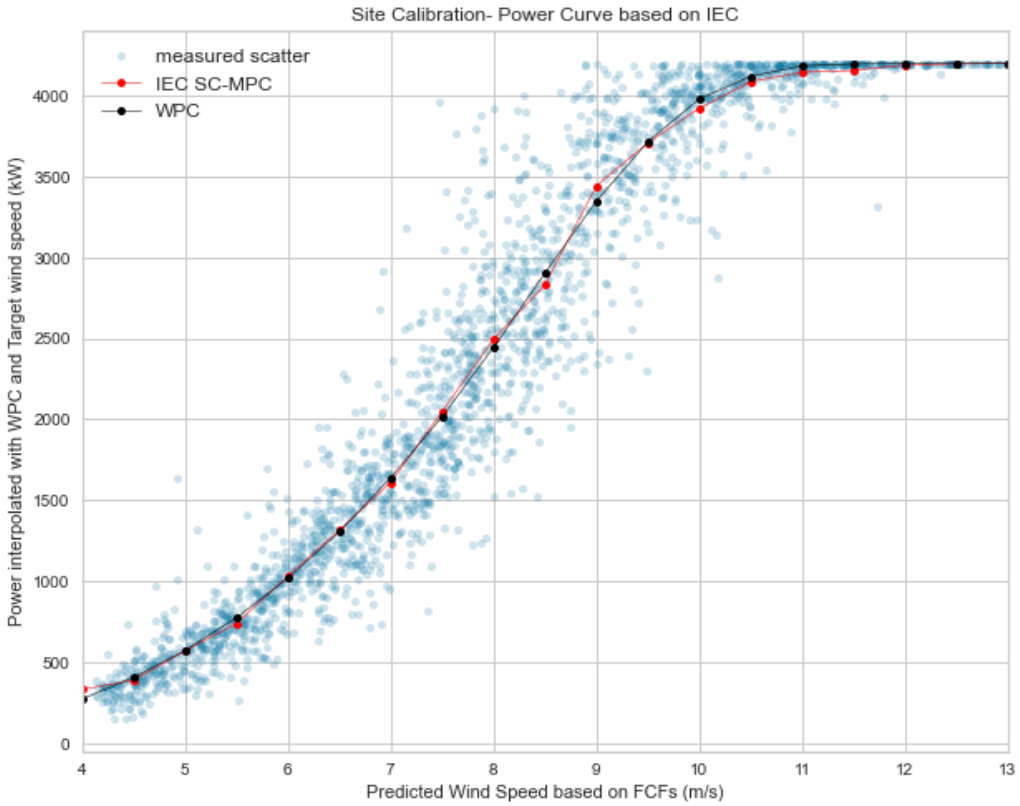


Figure D.10: Site Calibration Power Curve based on IEC FCF plot for wind turbine WGT18

Figure D.11: Flow Correction Factors plot for wind turbine WTG20

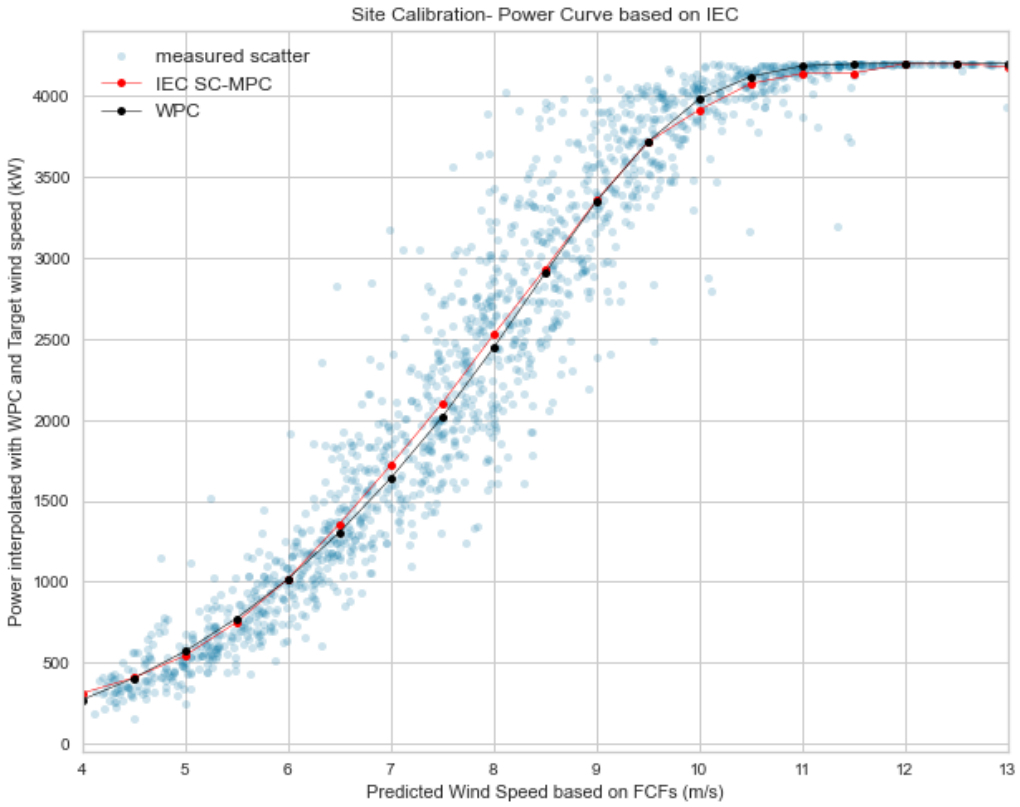


Figure D.12: Site Calibration Power Curve based on IEC FCF plot for wind turbine WGT20

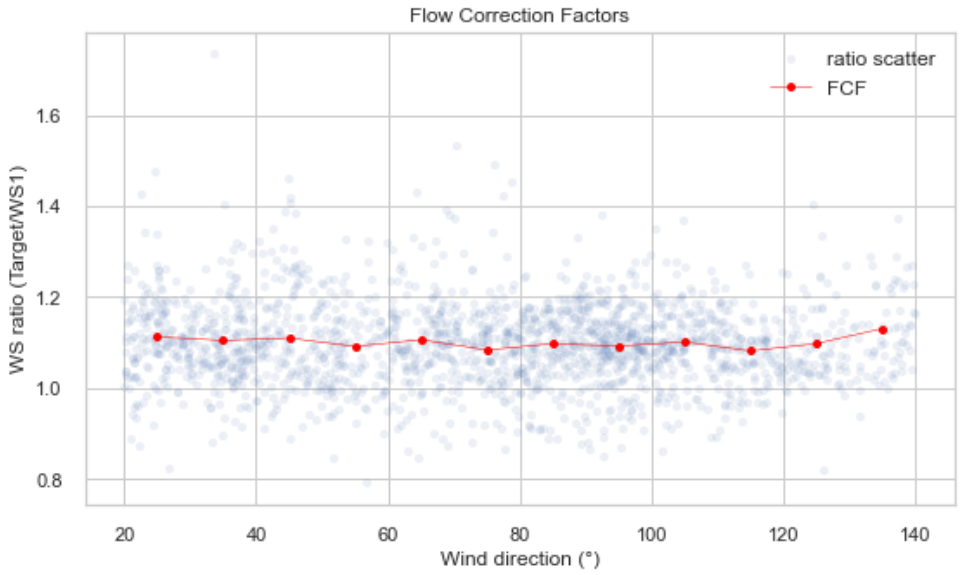


Figure D.13: Flow Correction Factors plot for wind turbine WTG43

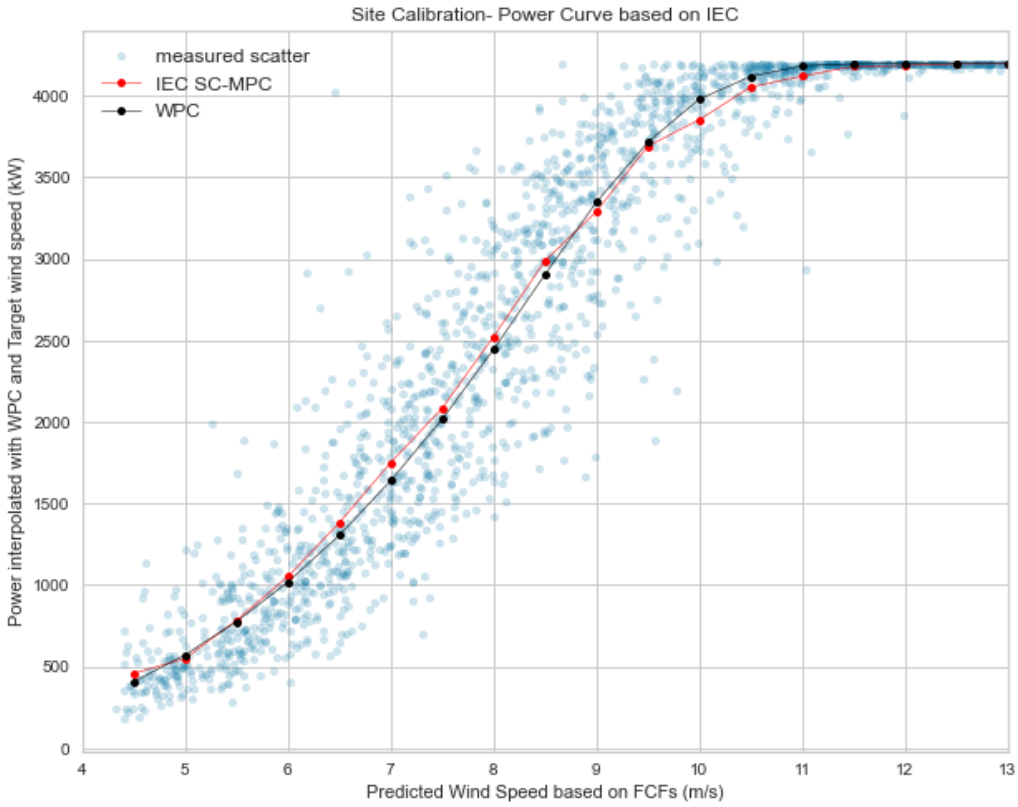


Figure D.14: Site Calibration Power Curve based on IEC FCF plot for wind turbine WGT43

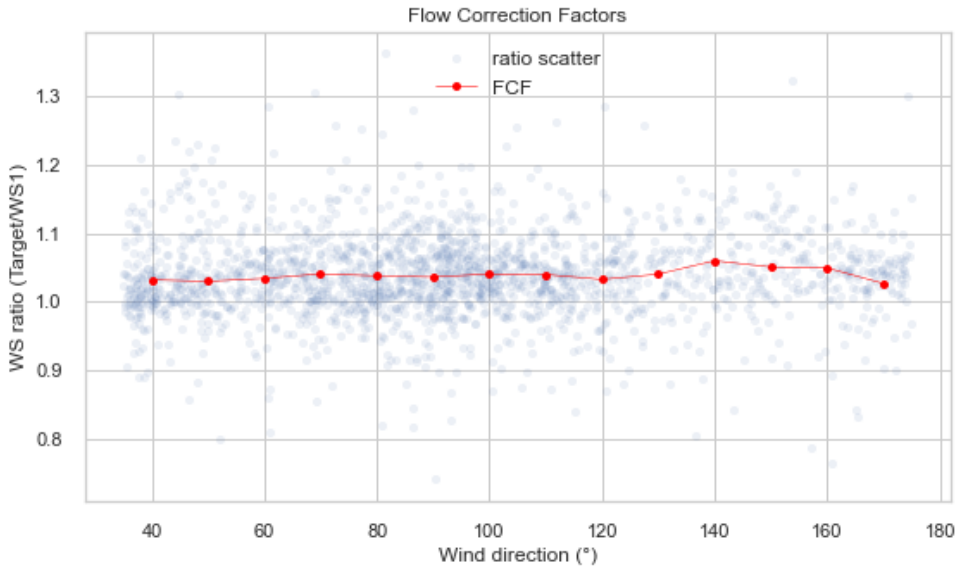


Figure D.15: Flow Correction Factors plot for wind turbine WTG46

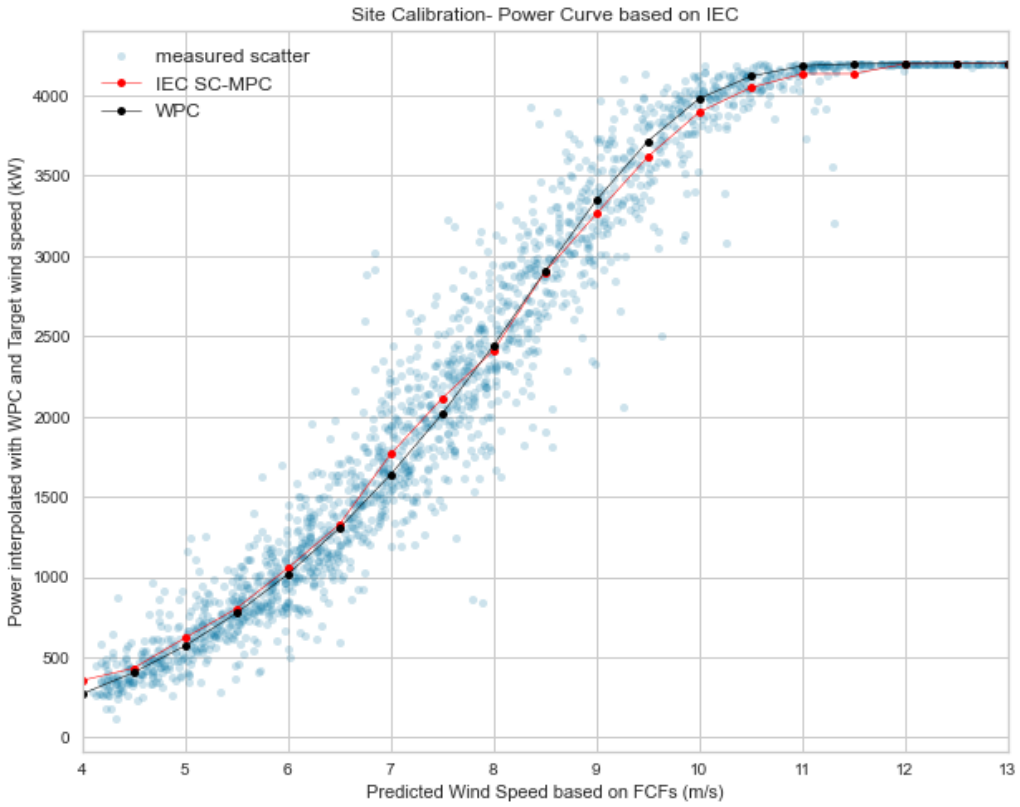


Figure D.16: Site Calibration Power Curve based on IEC FCF plot for wind turbine WGT46

E

Models analysis

E.1 Machine Learning mind map

Three different ML techniques are implemented in this Master Thesis in order to perform a SC regression for nine wind turbines. Fig. E.1 gives an overview of the selected ML techniques.

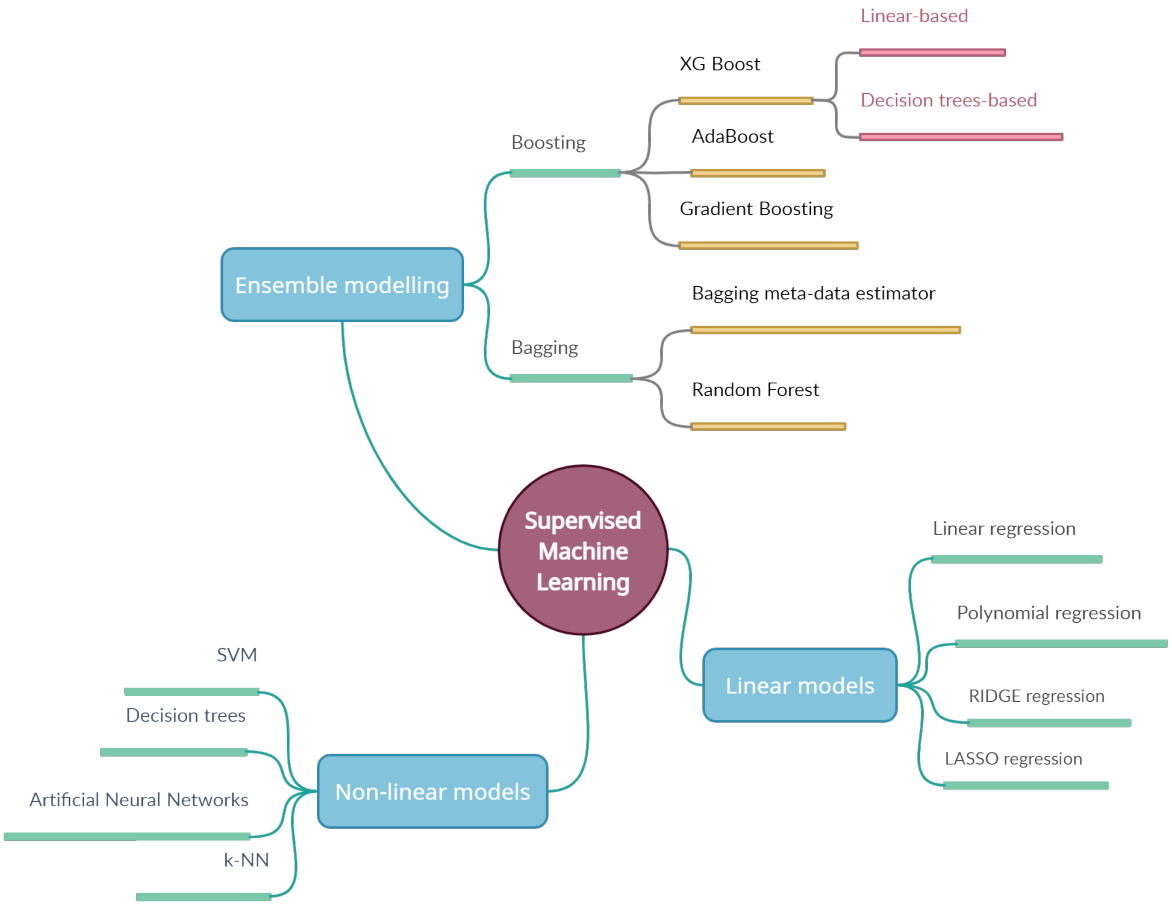


Figure E.1: Machine Learning supervised learning techniques mind map

E.2 Python packages and functions list

This section of Appendix E summarizes the different Python packages and the corresponding functions included in the scripts. Jupyter Notebook [43] is the web-based interactive computational environment use to code all the scripts needed for the development of the work shown in this Thesis. Table E.1 references the packages implemented in the scripts, lists the functions used and decribes the selected libraries.

Table E.1: Python packages and functions list and libraries description

Python package	Functions	Library description
NumPy [53]	sqrt, array	Powerful n-dimensional arrays library
Pandas [54]	DataFrame, read_csv	Python software library for data structure manipulation and analysis
SciPy [39]	zscore	Fundamental library for scientific and technical computing
Scikit-learn [46]	train_test_split, StratifiedShuffleSplit, MinMaxScaler, RandomizedSearchCV, GridSearchCV, mean_absolute_percentage_error, mean_absolute_error, mean_squared_error, r2_score, LinearRegression, Ridge, Lasso, RidgeCV, LassoCV, PolynomialFeatures	Python software module for Machine Learning built on top of SciPy.
Matplotlib [55]	pyplot.subplots, figure, legend, gca	Plotting library for Python and its numerical mathematics extension NumPy.
Seaborn [42]	lineplot, histplot, scatterplot, boxplot, colorbar	Python data visualization library based on Matplotlib.
Windrose [41]	WindroseAxes	Python library to manage wind data, draw windroses and fit Weibull probability density functions.
Yellowbrick [56]	PredictionError, ResidualsPlot	Python library for visual analysis and diagnostic tools.
Keras [50]	models.Sequential, wrappers.KerasRegressor	Open-source Neural Network library written in Python.
xgboost [51]	XGBRegressor	Open-source software Python library which provides a Gradient Boosting framework.
shap [44, 45]	Explainer, KernelExplainer, TreeExplainer, utils.sample, shap_values, summary_plot, plots.beeswarm	SHAP (SHapley Additive exPlanations) is a game theoretic approach to explain the output of any Machine Learning model.

E.3 Modelling errors results

This section shows the results of the ML modelling by wind turbine and by technique compared to the IEC baseline error.

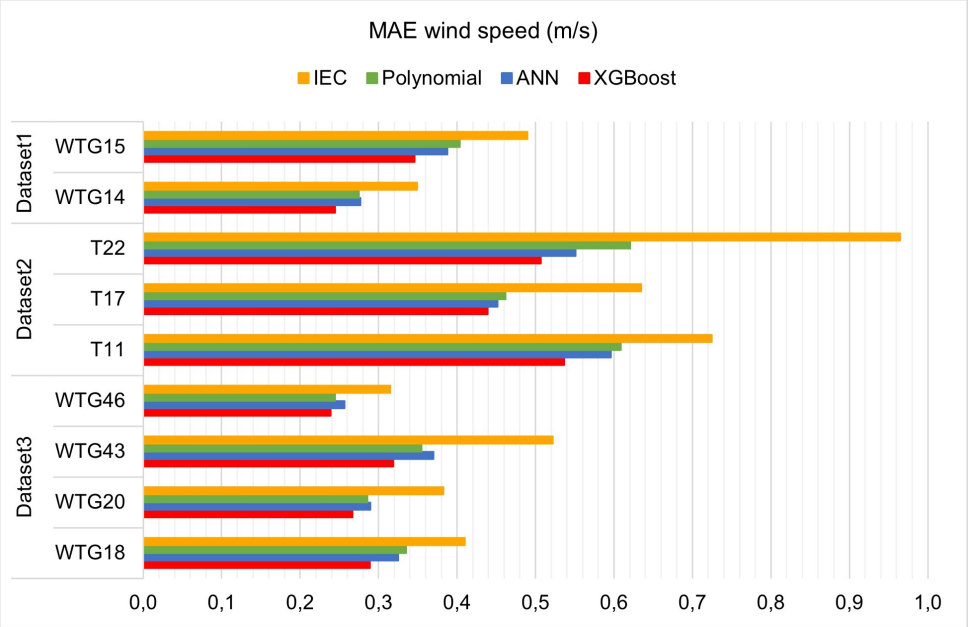


Figure E.2: MAE error for wind speed

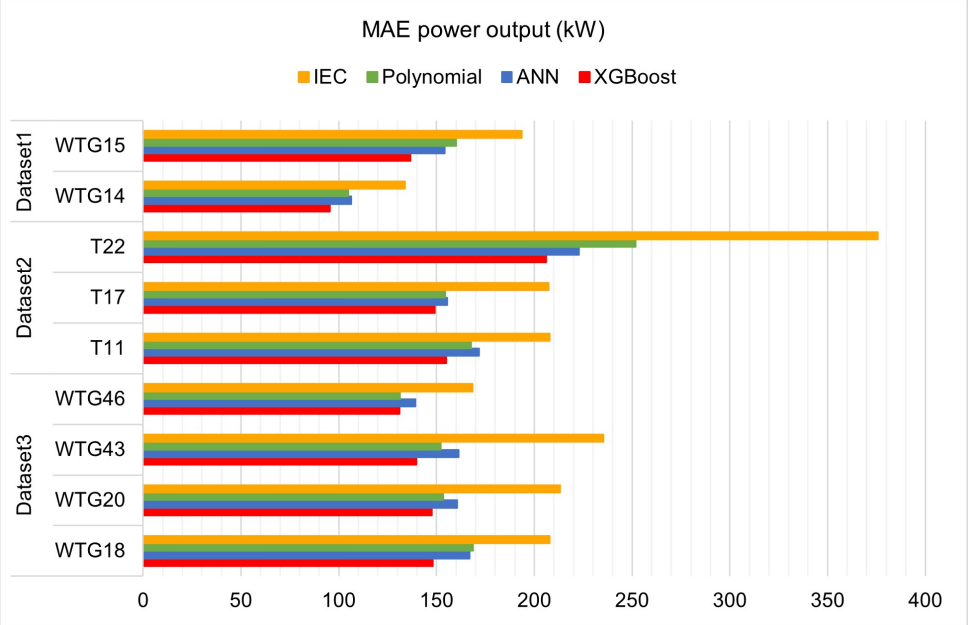


Figure E.3: MAE error for power output

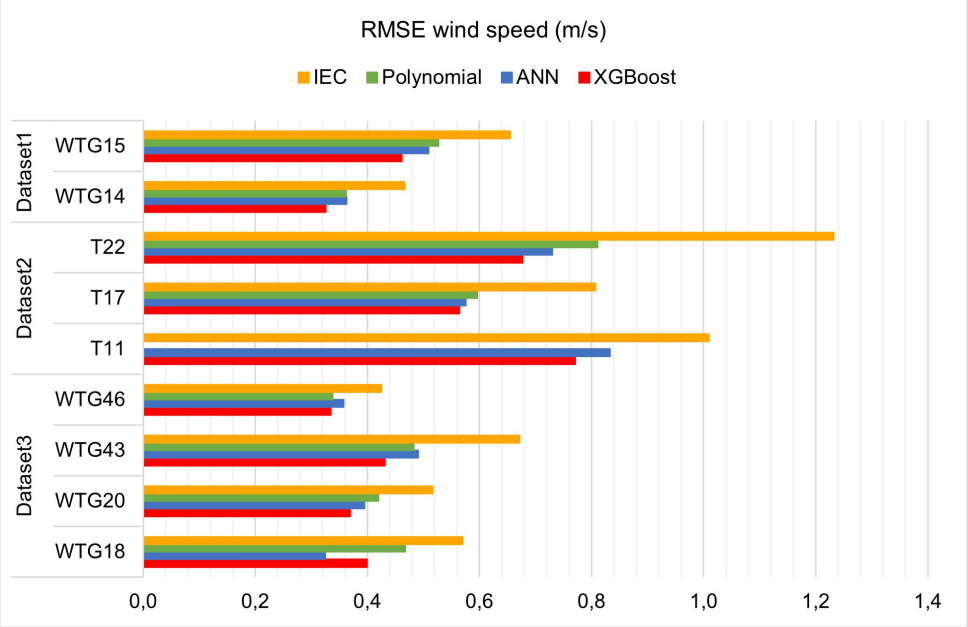


Figure E.4: RMSE error for wind speed

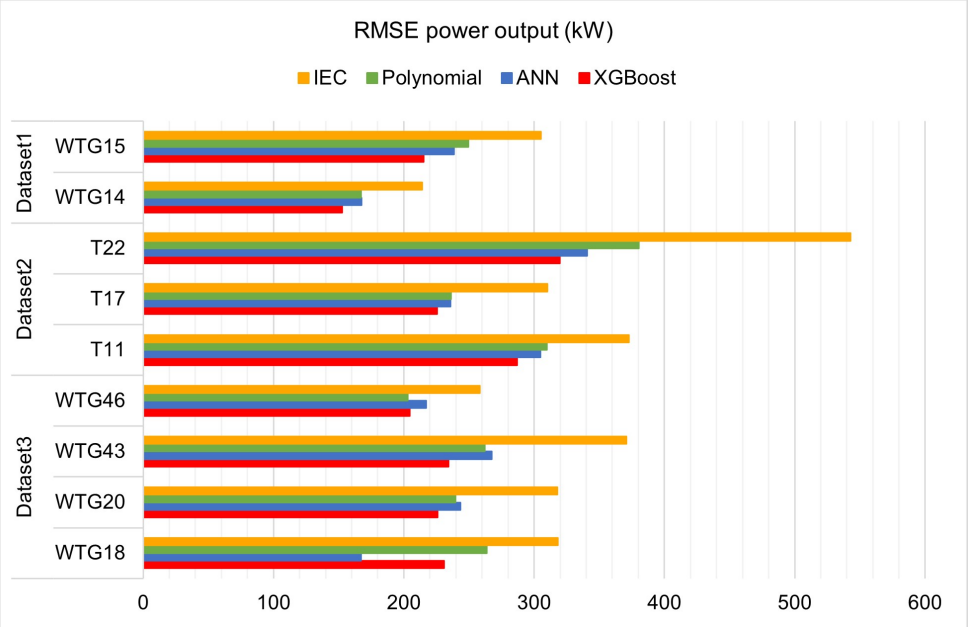


Figure E.5: RMSE error for power output

E.4 Site Calibration-Power Curves comparison

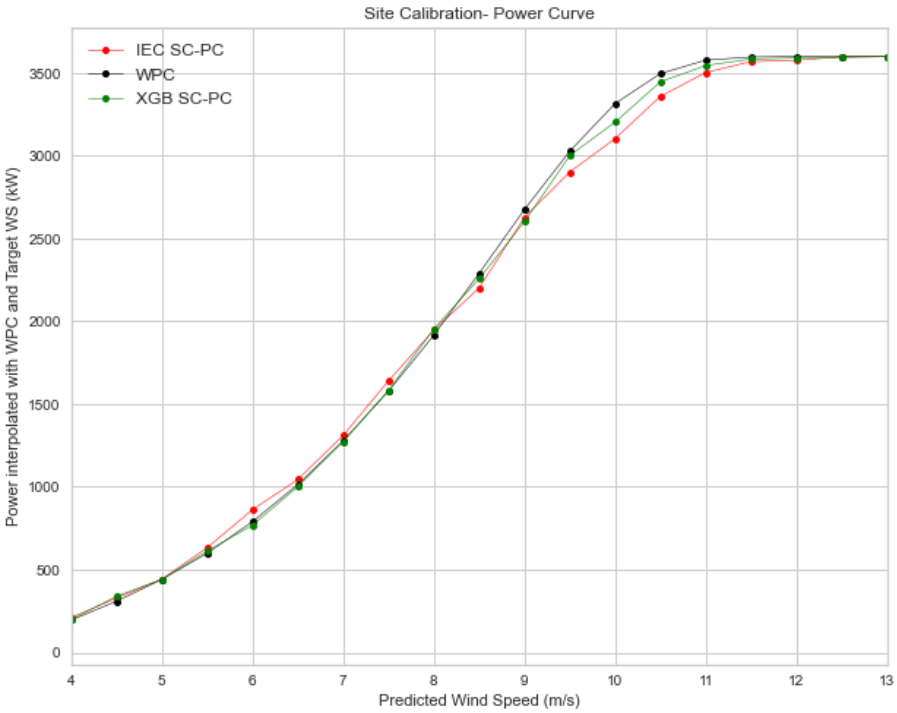


Figure E.6: Site Calibration-Power Curve comparison for wind turbine WTG15

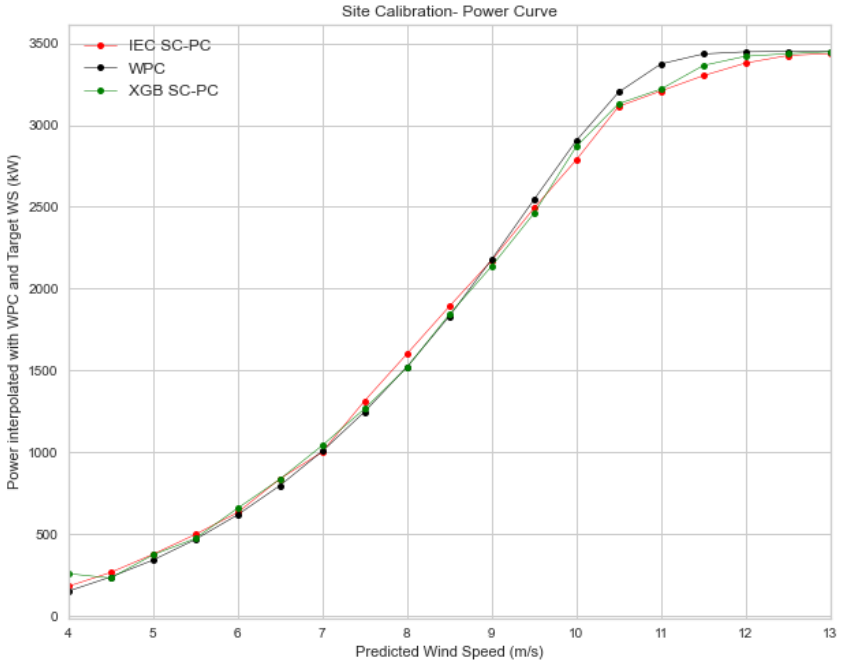


Figure E.7: Site Calibration-Power Curve comparison for wind turbine T11

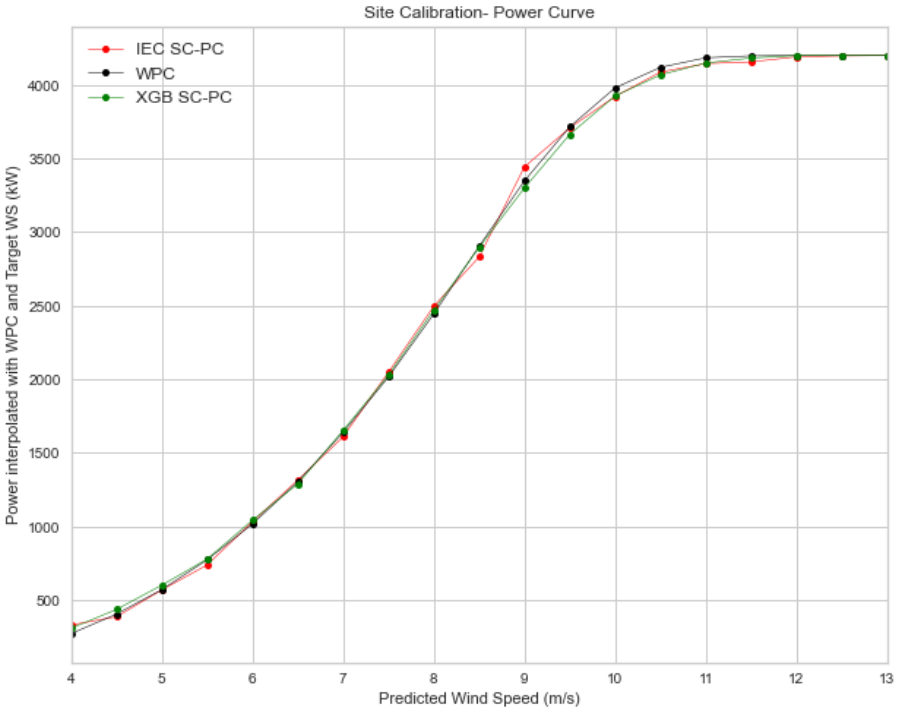


Figure E.8: Site Calibration-Power Curve comparison for wind turbine WTG18

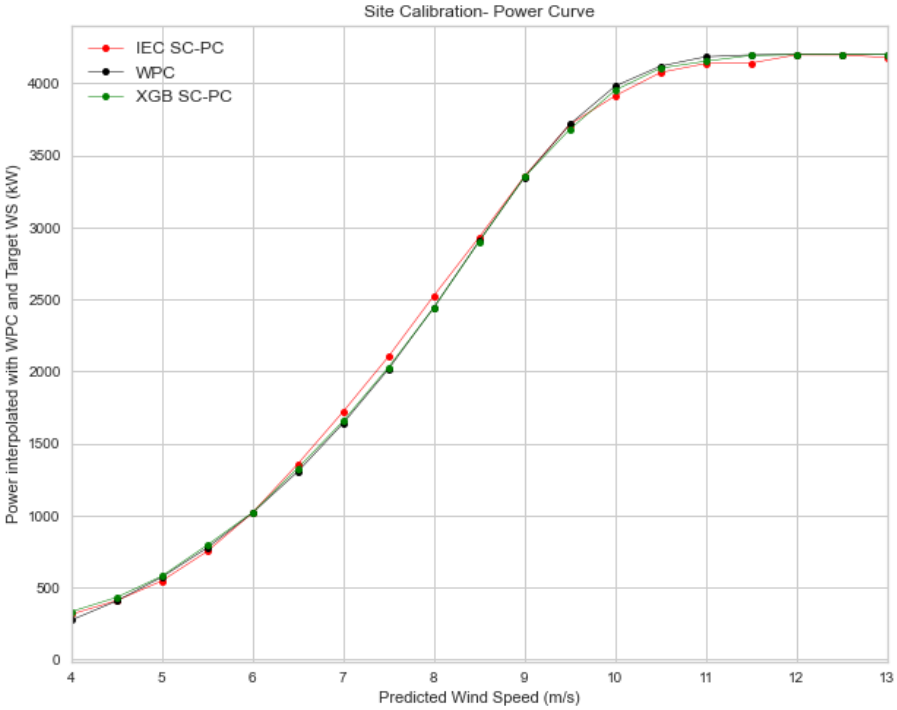


Figure E.9: Site Calibration-Power Curve comparison for wind turbine WTG20

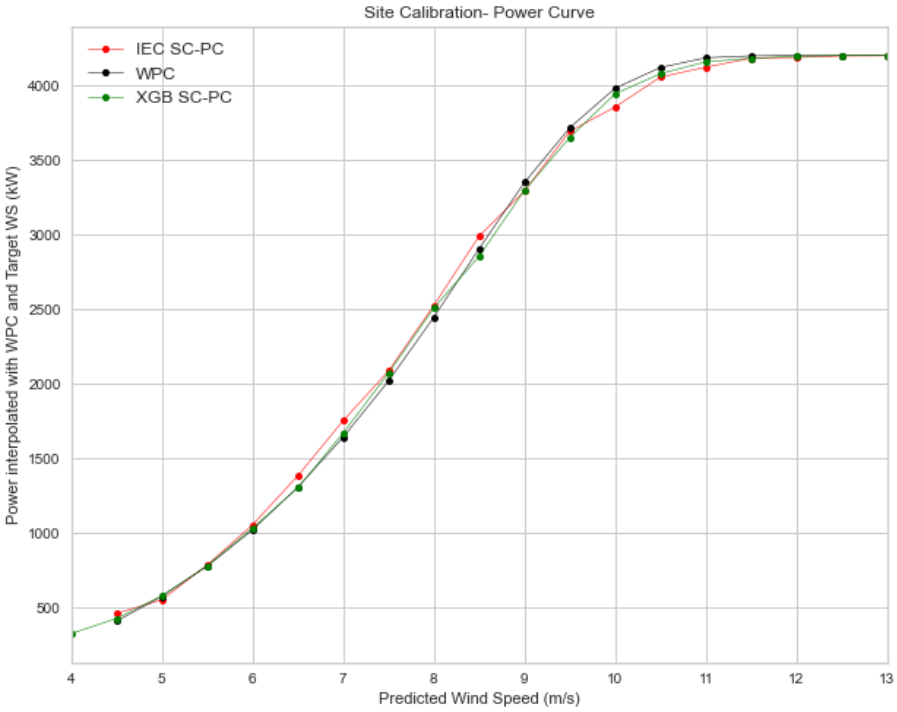


Figure E.10: Site Calibration-Power Curve comparison for wind turbine WTG43

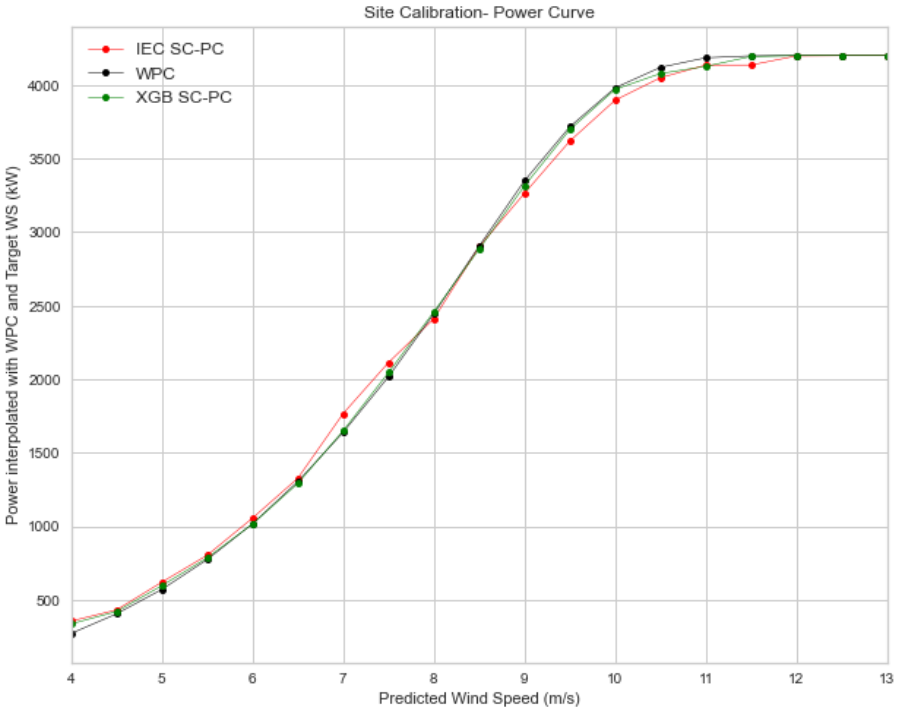


Figure E.11: Site Calibration-Power Curve comparison for wind turbine WTG46