

Performance Analysis of Wind Turbines at Pen y Cymoedd Wind Farm Using High-Frequency SCADA Data Sets

K.J.M. van Beveren

Performance Analysis of Wind Turbines at Pen y Cymoedd Wind Farm Using High-Frequency SCADA Data Sets

by

K.J.M. van Beveren

Student Name	Student Number
Koen van Beveren	4603281

subject

Wind Farm Post Construction Analysis & Data Science

Supervisor TU Delft: Dr.ir. W.A.A.M. (Wim) Bierbooms
Supervisor Vattenfall: Levent Küçük, PhD
Project Duration: March 2022 - March 2023
Faculty: Faculty of Electrical Engineering, Mathematics & Computer Science



Preface

This document is the product of the master thesis on the performance analysis of a Vattenfall wind farm. This master thesis is written as a final research of the Master's study Sustainable Energy Technology at the TU Delft. The research has been carried out in collaboration with the Wind Energy group from TU Delft and Vattenfall.

During the two-years master's program Sustainable Energy Technology (SET), I learned about designing energy systems by matching multiple energy sources with demand, storage and a grid. I specialised in the fields of wind energy, solar energy and energy storage. I used the opportunity of this master thesis to specialise more in the direction of wind energy. Moreover, I gained valuable insights in the data science field as I worked with big data sets, which required skills I had to learn in the course of this thesis.

This research is written for the Wind Resource Onshore team of Vattenfall, to gain insights in the performance of one of Vattenfall's onshore assets. Moreover, this research is meant for anyone interested in the potential of high-frequency data, or wind farm performance analysis in general.

In the first place, I would like to thank both my supervisor from TU Delft Dr. Wim Bierbooms and my supervisor from Vattenfall Levent Küçük, PhD. Throughout the past year, they have supported me and gave insights and feedback which were of great value for this research. Moreover, I would like to thank all other colleagues within the Vattenfall Wind Resource team, who were always ready to help when I had any kind of question. In general, I would like to thank Vattenfall for the opportunity within the company.

*K.J.M. van Beveren
Delft, April 2023*

Abstract

This research is carried out in collaboration with Vattenfall, to investigate the performance of a 76-turbine onshore wind farm. Since the start of operation, the wind farm has not been reaching P50 production estimates. Over the first five years of operation, the average gap between Annual Energy Production (AEP) and long-term AEP P50 estimates has been substantial. This thesis gives insights into the physical causes of this apparent underperformance. Conventional 10 minutes averaged data as well as data with a higher resolution, named high-frequency data, has been available for this purpose. Therefore, the research question to be answered is: “Can the underlying physical causes for the underperformance of the Pen y Cymoedd wind farm be found by the use of 10 minutes averaged and/or high-frequency data?”.

Answers to this research question were found by following a set-up analysis procedure, partially based on IEC (International Electrotechnical Commission) standards and with an added purpose for high-frequency data. This analysis was carried out for 14 turbines divided over 3 turbine clusters, with each turbine cluster accompanied by a MET mast, performing independent site measurements. Firstly, raw turbine data was filtered to only contain unwaked sectors and normal operational data. This is data where the turbine should not be hindered by other turbines, and should not be limited in its power output. After this filtering step, the turbine data was correlated to independent MET mast measurements at the site, situated close to the investigated turbines. Next to the data preparation, theoretical behaviour of the site-specific turbines was estimated. This theoretical expected behaviour, together with the processed data sets, were input to the analysis. This analysis consisted of two in parallel carried out analyses: a Cloud analysis and an Overall & Directional performance analysis. Within the cloud analysis, deviating turbine state curve behaviour¹ was selected and analysed. The overall & directional analysis investigated the overall turbine performance, and the performance of 18 different wind direction sectors per turbine.

This set-up methodology and analysis led to the following findings: on average, after data filtering, wind turbines perform 95% compared to the performance numbers the manufacturer delivers. There is a big difference in performance between the two turbine types at the site. Overall, the turbulence intensity (TI) class A eastern cluster turbines are performing better than the TI class B western and central cluster turbines. From the directional performance analysis, bad performance (>-5% compared to the warranted performance) was seen for sectors with a complex orography, as well as for sectors with few elevation deviation or forestry. The selected clouds from 10 minutes averaged data as well as from high-frequency data showed losses of potential power in regions such as the cut-in wind speed and storm control region, but also in the more critical partial load region, where the turbine should be performing at maximum power coefficient. Within these critical regions, potential performance losses of 10%-14% compared to the cloud energy production were found. In short, the clouds revealed the performance improvement potential of the 14 investigated turbines. The high-frequency data gave additional unique insights, only visible in higher resolution: it was concluded that there are periods of curtailments in the data after the filtering process, influencing potential performance calculations. Moreover, pitching to feathering position throughout the whole wind speed spectrum was observed. Furthermore, high pitching values during storm control and extreme torque for high wind speeds were visible in high resolution data. These events had a minor influence on turbine AEP than selected 10 minutes averaged clouds, but showed big potential to be improved.

¹The six state curves are described in figure 2.11

To answer the main research question, causes for underperformance can be mapped by the successfully set-up analysis using 10 minutes averaged and high-frequency data sets. High-frequency data shows high-resolution turbine behaviour not visible in 10 minutes averaged data, and clarifies averaged behaviour visible in 10 minutes averaged data. Underperformance compared to long term P50 estimates partially originates from turbine behaviour. High TI at the site, a complex orography, and potential conservative turbine settings have a big influence on turbine performance. The underperformance has the potential to be improved according to the cloud analysis.

Recommendations are based on improving data quality for future post construction analyses. Wind directional and wind speed data was partially available at the turbines and MET masts. Moreover, defective logging was discovered for those parameters. Non-natural offsets between wind directional data sets were observed. Furthermore, recommendations were made on the implementation of HF standard deviation filtering and on machine learning for underperformance detection.

Contents

Preface	i
Summary	ii
Nomenclature	xiv
1 Introduction	1
2 Literature Review	3
2.1 Wind turbine fault detection using a linear model	3
2.2 Previous HF works on PYC wind farm	4
2.3 Use of HF data in research	5
2.3.1 HF data filtering	5
2.3.2 WTG HF data analysis	6
2.3.3 Modelling with HF data sets	6
2.4 Causes for gap between theory and operation	8
2.5 Onshore site conditions	10
2.5.1 IEC standards regarding obstacles	13
2.6 WTG dynamics and control	14
2.7 Site-influence on onshore WTG performance	23
2.8 Turbine cluster analysis requirements	25
3 Methodology	26
3.1 Workflow	27
3.1.1 Filtering	28
3.1.2 Free sectors estimation	32
3.1.3 Data correlation and validation	33
3.1.4 Constructing turbine state curves	33
3.1.5 Overall & Directional performance analysis	35
3.1.6 Cloud analysis	37
3.2 Software and tools	40
3.2.1 Used software	40
4 Site analysis	41
4.1 Turbines at the site	41
4.2 Site conditions and turbine selection	42

4.2.1	Western Cluster	44
4.2.2	Central Cluster	46
4.2.3	Eastern Cluster	48
4.3	Wind farm performance	50
5	Data gathering	51
5.1	Data gathering	51
5.1.1	Pre-filtered data	52
5.2	Post-filtered data	53
5.3	Limitation of data: lack of wind directional measurements	55
6	Results western cluster	57
6.1	Wind resource	57
6.1.1	Influence of site on turbine performance D05	60
6.2	Correlations MET mast and turbines	61
6.3	Overall & directional performance analysis	63
6.3.1	Overall turbine performance	63
6.3.2	Directional turbine performance	64
6.4	Cloud analysis	68
6.4.1	Results 10 minutes averaged cloud analysis	68
6.4.2	Results HF cloud analysis	73
7	Results eastern cluster	80
7.1	Wind resource	80
7.1.1	Influence of site on turbine performance L06	82
7.2	Correlations MET mast and turbines	83
7.3	Overall & directional performance analysis	84
7.3.1	Overall turbine performance	84
7.3.2	Directional turbine performance	84
7.4	Cloud analysis	86
7.4.1	Results 10 minutes averaged cloud analysis	87
7.4.2	Results HF cloud analysis	94
8	Results central cluster	98
8.1	Wind resource	98
8.1.1	Influence of site on turbine performance H05	100
8.2	Correlations MET mast and turbines	101
8.3	Overall & directional performance analysis	103
8.3.1	Overall turbine performance	103
8.3.2	Directional turbine performance	104

8.4	Cloud analysis	106
8.4.1	Results 10 minutes averaged cloud analysis	106
8.4.2	Results HF cloud analysis	108
9	Conclusion	111
10	Discussion	113
11	Recommendations	115
11.1	Data quality & availability	115
11.2	Improving curtailment filtering procedure	116
11.3	Filtering on standard deviation	116
11.4	Use machine learning for outlier detection	116
11.5	Revise wind resource analysis at site	116
A	Defective wind vane measurements at MET masts	119
B	Correlations turbines western cluster	120
B.1	MET-D05 correlations	120
B.2	D04-D03 correlations	121
B.3	D03-D01 correlations	121
B.4	Correlations wind direction - nacelle direction	121
C	Algorithm step detection for correcting offset turbine D04	122
D	Machine learning for detecting outliers on D05 operational active power data	123
E	Example of cloud analysis tool output	124
F	Code	127
F.1	Script for filtering raw turbine data	127
F.2	Two examples of functions from master filtering script	129
F.3	Example of function from analysis script	130
F.4	Turbine class as input to Cloud analysis tool	131
G	Google Earth site analysis examples	134

List of Figures

2.1	FDI Residual-based approach Habibi et al. (2019)	4
2.2	Random Forest Algorithm Gonzalez et al. (2017)	7
2.3	Overview of technical wind turbine losses Mortensen (2012)	8
2.4	Visual representation of wind turbine losses Lee and Fields (2021)	9
2.5	Visual representation of turbine performance subcategory data Lee and Fields (2021)	9
2.6	Power coefficient for varying tip speed ratio and pitch angles (Lio, 2018)	16
2.7	General regions of the wind turbine power curve (Sohoni et al., 2016)	17
2.8	Rotor speed - Torque relation (van Wingerden, 2021)	18
2.9	Active power, rotor speed and pitching during storm control (Markou & Larsen, 2009)	20
2.10	Ramping effects on power curve (Antinio Notaristefano, 2021)	21
2.11	Different state curves of a variable wind speed pitch-controlled wind turbine	22
2.12	Example of a power-to-power visualisation of a reference turbine and a test turbine	25
3.1	Schematic overview of the methodology workflow	27
3.2	Filtering process for turbine data (left) and MET mast data (right)	30
3.3	Post-processing steps before analysis	32
3.4	Description of correlation checks	33
3.5	Datashader visualisation algorithm	40
4.1	PYC wind farm with the three selected clusters	43
4.2	Height profile of western cluster (Openwind)	44
4.3	Free sectors of the MET mast and turbines of the western cluster	45
4.4	Height profile of central cluster (Openwind)	46
4.5	Free sectors of the MET mast and turbines of the central cluster	47
4.6	Height profile of eastern cluster (Openwind)	48
4.7	Free sectors of the MET mast and turbines of the eastern cluster	49
5.1	MET 10 minutes averaged data sampling frequency	52
5.2	MET HF data sampling frequency	52
5.3	D05 10 minutes averaged data sampling frequency	53
5.4	D05 HF data sampling frequency	53
5.5	Filtered out data in a period of curtailment	53
5.6	Filtered out data in a period of curtailment	53
5.7	Multiple MET mast measurement heights	55

5.8	Wind direction - nacelle direction correlation turbine D05	56
6.1	Wind rose western cluster	57
6.2	Wind speed histogram western cluster	57
6.3	10 minutes averaged MET mast wind speed veer correlation.	58
6.4	10 minutes averaged MET mast wind direction veer correlation.	58
6.5	10 minutes averaged MET mast wind speed shear correlation.	59
6.6	10 minutes averaged MET mast wind direction shear correlation.	59
6.7	2020 average wind shear profile at the MET mast	59
6.8	10 minutes averaged MET mast wind speed TI correlation.	60
6.9	10 minutes averaged MET mast wind direction TI correlation.	60
6.10	Influence of veer on D05 active power signal.	61
6.11	Influence of shear on D05 active power signal.	61
6.12	10 minutes averaged wind speed correlation of D05 and D04	62
6.13	10 minutes averaged wind direction correlation of D05 and D04	62
6.14	10 minutes averaged nacelle direction correlation of D05 and D04	62
6.15	Total power curve turbine D05	63
6.16	Total power curve turbine D04	63
6.17	Total power curve turbine D03	63
6.18	Total power curve turbine D01	63
6.19	Binned power curves turbine D05	64
6.20	Binned power curves turbine D04	65
6.21	Binned power curves turbine D03	65
6.22	Binned power curves turbine D01	66
6.23	Relative power gain/loss per wind direction bin for turbine D05	66
6.24	Relative power gain/loss per wind direction bin for turbine D04	66
6.25	Relative power gain/loss per wind direction bin for turbine D03	67
6.26	Relative power gain/loss per wind direction bin for turbine D01	67
6.27	Compared 10 minutes averaged data of western cluster turbines. Data is binned by the MOB.	69
6.28	Wind speed, TI degradation turbines western cluster	70
6.29	Pitching between 3-5 m/s (cloud D05-10-794).	71
6.30	RPM below torque curve (cloud D05-10-794).	71
6.31	Torque behaviour above optimal- c_p region.	71
6.32	Scattering below the power curve (cloud D01-10-174).	72
6.33	Aggressive pitching (cloud D01-10-174).	72
6.34	Time series of power output during curtailed period of cloud D01-10-174.	72
6.35	HF power curve scatter plots of the western cluster turbines.	74

6.36	Time series of power output during curtailed period of cloud D05-HF-099.	74
6.37	Time series of power output during two-hour period of cloud D05-HF-936.	74
6.38	Time series of pitching during two-hour period of cloud D05-HF-936.	75
6.39	4 hour active power close up of 11-day period cloud D05-HF-882.	75
6.40	4 hour pitching close up of 11 day period cloud D05-HF-882.	75
6.41	4 hour TI close up of 11 day period cloud D05-HF-882.	75
6.42	Pitching >30° at turbine D05.	76
6.43	Active power output for pitching >30° at turbine D05.	76
6.44	TI distribution of cloud D05-HF-022.	76
6.45	Cloud selection of pitching above cut-in wind speed.	77
6.46	Active power output for pitching above cut-in wind speed.	77
6.47	Cloud D04-HF-981: time series of wind speeds.	78
6.48	Cloud D04-HF-981: time series of pitching.	78
6.49	Time series of wind speed ramping up.	79
6.50	Time series of rotor speed response to wind speed ramp ups.	79
6.51	Time series of pitching response to rotor speed increase.	79
7.1	Wind rose eastern cluster	80
7.2	Wind speed histogram eastern cluster	80
7.3	10 minutes averaged MET mast wind speed veer correlation.	81
7.4	10 minutes averaged MET mast wind direction veer correlation.	81
7.5	10 minutes averaged MET mast wind speed shear correlation.	81
7.6	10 minutes averaged MET mast wind direction shear correlation.	81
7.7	10 minutes averaged MET mast wind speed TI correlation.	82
7.8	10 minutes averaged MET mast wind direction TI correlation.	82
7.9	Influence of veer on L06 active power signal.	82
7.10	Influence of shear on L06 active power signal.	82
7.11	Total power curve turbine L06	84
7.12	Total power curve turbine L05	84
7.13	Binned power curves turbine L06	85
7.14	Binned power curves turbine L05	85
7.15	Relative power gain/loss per wind direction bin for turbine L06	86
7.16	Relative power gain/loss per wind direction bin for turbine L05	86
7.17	Compared 10 minutes averaged data of eastern cluster turbines. Data is binned by the MOB.	88
7.18	High wind speed power control active power output.	89
7.19	High wind speed power control pitching behaviour.	89
7.20	Time series of power output during high wind speed control.	90

7.21	Wind speed - TI relation cloud L06-10-150. In blue: TI class A turbulence relation. . . .	90
7.22	Rotor speed - pitching relation turbine D05.	91
7.23	Rotor speed - pitching relation turbine L06.	91
7.24	Rotor speed - torque relation turbine D05.	91
7.25	Rotor speed - torque relation turbine L06.	91
7.26	Optimal- c_p region of turbine D05.	92
7.27	Optimal- c_p region of turbine L06.	92
7.28	Power curve with second line of power production.	93
7.29	Wind speed - rotor speed relation with second line of rotor speed.	93
7.30	Wind speed - TI relation cloud L05-10-248. In blue: TI class B turbulence relation. . . .	93
7.31	High wind speed power control active power output.	95
7.32	Pitching and Active power distributions for cloud L06-HF-627.	95
7.33	Cloud L06-HF-632: extreme pitching.	96
7.34	Cloud L06-HF-632: decrease of rotor speeds.	96
7.35	Cloud L06-HF-632: example of power signal around selected data.	96
7.36	Selected high-torque cloud for high pitching.	97
7.37	Wind speed and pitching distributions for cloud L06-HF-678.	97
8.1	Wind rose central cluster	98
8.2	Wind speed histogram central cluster	98
8.3	10 minutes averaged MET02 wind speed veer correlation.	99
8.4	10 minutes averaged MET02 wind direction veer correlation.	99
8.5	10 minutes averaged MET02 wind speed shear correlation.	99
8.6	10 minutes averaged MET02 wind direction shear correlation.	99
8.7	10 minutes averaged MET02 wind speed TI correlation.	100
8.8	10 minutes averaged MET02 wind direction TI correlation.	100
8.9	Influence of veer on H05 active power signal.	100
8.10	Influence of shear on H05 active power signal.	100
8.11	Wind speed correlation of MET02 and turbine H05.	103
8.12	Wind speed correlation of turbines H05 and H06.	103
8.13	Binned power curve of turbine H05.	103
8.14	Binned power curves turbine H05	104
8.15	Relative power gain/loss per wind direction bin for turbine H05.	105
8.16	Compared 10 minutes averaged data of central cluster turbines. Data is binned by the MOB.	107
8.17	2200 kW curtailments cloud H05-HF-473.	109
8.18	1700 kW curtailments cloud H05-HF-082.	109
8.19	500 kW curtailments cloud H05-HF-511.	109

A.1	The origin of wrong wind vane and wind direction measurements/calculations at the MET mast.	119
B.1	10 minutes averaged wind speed correlation of MET and D05	120
B.2	10 minutes averaged wind direction correlation of MET and D05	120
B.3	10 minutes averaged wind speed correlation of D04 and D03	121
B.4	10 minutes averaged wind direction correlation of D04 and D03	121
B.5	10 minutes averaged nacelle direction correlation of D04 and D03	121
B.6	10 minutes averaged wind speed correlation of D03 and D01	121
B.7	10 minutes averaged wind direction correlation of D03 and D01	121
B.8	10 minutes averaged nacelle direction correlation of D03 and D01	121
B.9	Wind direction - nacelle direction correlation turbine D04	121
B.10	Wind direction - nacelle direction correlation turbine D03	121
B.11	Wind direction - nacelle direction correlation turbine D01	121
C.1	Output step detection algorithm turbine D04 before correction.	122
C.2	Output step detection algorithm after correction.	122
D.1	Outlier scores for 10 minute power curve data for turbine D05.	123
D.2	Included and excluded points after applying model.	123
E.1	Histograms output from cloud analysis tool.	124
E.2	Scatter plots output from cloud analysis tool.	125
E.3	Time series output from cloud analysis tool.	126
G.1	Western cluster bad performing sectors in Google Earth ('Google Earth', n.d.).	134
G.2	Western cluster bad performing sectors in Google Earth ('Google Earth', n.d.).	135
G.3	Eastern cluster bad performing sector in Google Earth ('Google Earth', n.d.).	135

List of Tables

2.1	Correlations investigated in research by T.Kaniewski	5
2.2	Filters applied to Horn Rev I Wind Turbine	5
2.3	Properties able to investigate by using HF data sets (Wilkinson, 2016).	6
2.4	TI class specific constants Øistad, 2015	10
2.5	Roughness lengths for different terrains.	11
2.6	Standard shear factor values for different terrains (Ray et al., 2006)	12
2.7	Horizontal wind speed influenced by an object, represented as a percentage from the initial wind speed upfront of the object.	12
2.8	IEC obstacle requirements	13
2.9	Magnitude of investigated properties in Murphy et al. (2020)	23
2.10	Magnitude of investigated properties in Sanchez Gomez and Lundquist (2020)	24
2.11	Magnitude of investigated properties in Stival et al. (2017)	24
2.12	Magnitude of investigated properties in Vahidzadeh and Markfort (2019)	24
2.13	Parameters needed for turbine cluster analysis (Albers, 2014)	25
3.1	Python packages and corresponding logos.	26
3.2	Possible corrections carried out to data sets after filtering.	32
3.3	Tabular overview of cloud turbine analyses.	39
4.1	Properties of respectively SWT-3.0-113 and SWT-3.0-108	41
4.2	Design boundary conditions turbine SWT-3.0-113	42
4.3	Design boundary conditions turbine SWT-3.0-108	42
4.4	Distances between selected turbines from western cluster	44
4.5	Distances between selected turbines from central cluster	46
4.6	Distances between selected turbines from eastern cluster	48
5.1	Must-have tags and their availability in the Wind Web Portal (WWP) and Azure database.	51
5.2	MET mast tags.	52
5.3	Data tags after renaming.	54
5.4	Properties of wind direction - nacelle direction correlations for turbines of the western cluster.	56
6.1	Overview of correlations for the western cluster.	61
6.2	Expected slopes for wind speed correlations western cluster.	62
6.3	Overview of corrections carried out on correlations western cluster.	62

6.4	Efficiencies compared to WPC for turbines western cluster	64
6.5	Worst performing unawaked sectors turbines western cluster.	67
6.6	10 minutes averaged clouds analysed western cluster.	68
6.7	Data collection of selected 10 minutes averaged clouds western cluster.	68
6.8	HF clouds analysed western cluster.	73
6.9	Data collection of selected HF clouds western cluster.	73
7.1	Overview of correlations for the eastern cluster.	83
7.2	Expected slopes for wind speed correlations eastern cluster.	83
7.3	Efficiencies compared to WPC for turbines eastern cluster	84
7.4	Worst performing unawaked sectors turbines eastern cluster.	86
7.5	10 minutes averaged clouds analysed eastern cluster.	87
7.6	Data collection of selected clouds eastern cluster	87
7.7	HF clouds analysed eastern cluster.	94
7.8	Data collection of selected clouds eastern cluster	94
8.1	Overview of corrections carried out on correlations central cluster	101
8.2	Overview of correlations for the central cluster.	102
8.3	Expected slopes for wind speed correlations central cluster.	102
8.4	Efficiencies compared to WPC for turbines central cluster.	104
8.5	Worst performing unawaked sectors turbines central cluster.	105
8.6	10 minutes averaged clouds analysed central cluster.	106
8.7	Data collection of selected clouds central cluster.	106
8.8	HF clouds analysed central cluster.	109
8.9	Data collection of selected clouds central cluster.	109

Nomenclature

Abbreviations

Abbreviation	Definition
ABL	Atmospheric Boundary Layer
ADW	Azure Data Warehouse
AEP	Annual Energy Production
ANN	Artificial Neural Network
CF	Capacity Factor
DWG	Deutsche Wind Guard
FDI	Fault Detection Isolation
HF	High-Frequency
HSS	High Speed Shaft
IEC	International Electrotechnical Commission
KNN	k-Nearest Neighbour
LOF	Local Outlier Factor
LPP	Lost Potential Power
LPPC	Lost Potential Power as a fraction of cloud power production
LPPT	Lost Potential Power as a fraction of annual power production
LSS	Low Speed Shaft
MCP	Measure Correlate Predict
MOB	Method of Bins
NBM	Normal Behaviour Model
NWD	Nominal Wind Distribution
OEM	Original Equipment Manufacturer
OP	Operation Point
OPC	Operational Power Curve
O&M	Operations & Maintenance
PCA	Post-Construction Analysis
PLU	Production, Losses and Uncertainty
PYC	Pen Y Cymoedd
RMSE	Root Mean Squared Error
RPM	Rotations Per Minute
SCADA	Supervisory Control and Data Acquisition
TI	Turbulence Intensity
WF	Wind Farm
WPC	Warranted Power Curve
WRA	Wind Resource Analysis
WT	Wind Turbine
WTG	Wind Turbine Generator
WWP	Wind Web Portal

Symbols

Symbol	Definition	Unit
A	Area	[m ²]
a	Scale factor	[-]
c_P	Power coefficient	[-]
c_Q	Torque coefficient	[-]
c_T	Thrust coefficient	[-]
D	Diameter	[m]
D	Distance	[m]
f	Frequency	[Hz]
H	Height	[m]
I	TI constant	[-]
k	Shape factor	[-]
k	Wake decay coefficient	[-]
l	Length of object	[m]
P	Power	[W]
Q	Torque	[Nm]
R	Radius	[m]
r	Gearbox ratio	[-]
s	Dimensionless distance	[-]
T	Temperature	[°C]
t	Time	[s]
U	Wind speed	[m/s]
u	Friction velocity	[m/s]
v	Wind speed	[m/s]
z	Elevation	[m]
α	Shear factor	[-]
β	Wind directional veer	[°/m]
η	Efficiency	[-]
θ	Wind direction	[°]
θ	Pitching angle	[°]
κ	Von Kármán's constant	[-]
λ	Tip speed ratio	[-]
ρ	Density	[kg/m ³]
σ	Standard deviation	[-]
Ω	Rotor speed	[rad/s]
ω	Rotor speed	[m/s]

1

Introduction

The process of developing a wind farm includes detailed predictions of wind farm yield, with site specific conditions such as site orography and wind properties as inputs. Onshore wind farm yield, in contrast to offshore wind farm yield, is highly influenced by site conditions, which brings challenges when estimation and prediction performance for a certain wind farm layout. The Pen Y Cymoedd (PYC) wind farm is situated in the southern part of Wales, and was constructed in a rough landscape with high elevation differences. This brings challenges when estimating wind farm yield in the design phase of the project. Long-term yield and site condition estimations were made, representing the average of multiple years the farm would be in operation. Annual Energy Production (AEP) numbers of this wind farm do not match the predictions done years before construction, which make PYC one of the underperforming assets of Vattenfall.

The wind industry's interest on high-frequency data has gradually been increasing. The common practice for analysing wind turbine performance in the wind energy sector still relies on 10 minutes averaged Supervisory Control and Data Acquisition (SCADA) data, whereas using high-frequency (HF) data is started to be seen as another opportunity for turbine Post-construction analysis (PCA). There are world-wide published studies, consultant reports and white papers demonstrating that performance issues and anomalies which are not visible in 10 minutes averaged SCADA data can be detected, analysed and reported once the data resolution is increased. For the PYC wind farm, this high-resolution data is available.

The main aim of this thesis is to understand the physical meaning behind the underperformance of the PYC wind farm by the use of 10 minutes averaged data and HF data. To achieve this, three sub-goals are defined. The first goal is developing a structured approach to investigate different aspects of turbine performance, with using 10 minutes averaged data as well as HF data. The second goal is to investigate the added value of HF data sets for this wind farm performance analysis. The final sub goal is to gain insights in the physical causes of the PYC wind farm. An output of this work would be an overview of recommendations for the operations and maintenance (O&M) strategy.

The scope of this thesis is 14 turbines and MET masts, in normal operational conditions. Moreover, the research is focussed on unwaked sectors for all investigated turbines and MET masts. The turbines and MET mast are divided over 3 turbine clusters. One limitation of this study pertains to the unavailability of unaffected site condition measurements at the turbines within the examined turbine clusters. These unaffected measurements are only available at the MET masts. The analysis is set up to cope with this limitation.

This thesis contributes to the understanding in using high-frequency data for wind farm post-construction analysis. It gives additional insights in the added value of this data. Moreover, the thesis provides a structured methodology which can be optimised and repeated in future research to wind farm performance by using high-frequency data sets.

This thesis starts with a literature review of recent studies on high-frequency (HF) data, turbine under-performance and onshore site conditions (chapter 2). This is done to obtain a sufficient background in these topics before doing research in the field of high-frequency data sets and post-construction analyses for onshore wind farms. Also, in this chapter, some IEC guidelines useful for this work are discussed. Secondly, the methodology of this work will be explained in chapter 3. This chapter contains the overall workflow of the thesis, as well as a more in-depth look into the cloud analysis that will be performed among three turbine clusters. In chapter 4, the site is investigated. Geometric properties of the turbine clusters are summarized. Also, unyawed sectors are estimated for all turbines to be investigated. In the next chapter, chapter 5, properties of the retrieved data are shown. On top of that, the filtering process prior to the analysis is listed and visualised. Chapters 6 and 7 contain results from the analysis carried out on the two clusters of the wind farm. This includes site conditions per cluster, as well as results showing turbine performance for the full spectrum of site conditions. Finally, conclusions and recommendations are given in chapter 9.

2

Literature Review

In this chapter, the collected information from the literature review is summarized. This literature review is carried out prior to the data analysis that will be carried out on the PYC wind farm. A literature study has been done on subject which serve as background for the research. Different papers have been studied and an overview of those can be found in the bibliography. In addition, a report about HF data analysis on the PYC wind farm was studied. This report is the result of a research done before this thesis.

2.1. Wind turbine fault detection using a linear model

Wind turbine behaviour is a strongly nonlinear. To make the prediction of wind turbine performance better computable and understandable, nonlinear models have been linearised around different operation points (OP). A linearised model of wind turbine behaviour is shown below Habibi et al. (2019):

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + F_a\mathbf{f}_a + \mathbf{R}V_r \quad (2.1)$$

Where:

- $\dot{\mathbf{x}}$ = Estimated values (derivatives of \mathbf{x})
- \mathbf{x} = Vector with rotor speed, generator speed, pitching and torque values
- \mathbf{u} = Vector with torque and reference pitching values
- \mathbf{f}_a = Vector with offsets of multiple parameters
- A, B, F_a = Matrices with components to link known values to predicted values
- V_r = Effective wind speed

Where the bold terms are all variables describing the wind turbine behaviour, structured in vectors. The explanation of those vectors is elaborated in Habibi et al. (2019). In addition, there is a measurement model, describing measured parameters as a function of theoretical parameters.

$$\mathbf{y} = \mathbf{C}\mathbf{x} + F_s\mathbf{f}_s + \mathbf{D} \quad (2.2)$$

Where:

- \mathbf{y} = Rotor speed, generator speed, pitching and torque measurements
- \mathbf{x} = Vector with rotor speed, generator speed, pitching and torque values
- \mathbf{f}_s = Measurement bias vector
- \mathbf{D} = Measurement noise vector
- C, F_s = Matrices with components to link known values to measured values

A very common way of detecting abnormal behaviour is Fault Detection Isolation (FDI), with a residual-based approach. A general schematic overview of how a model can be used to detect faults is shown below:

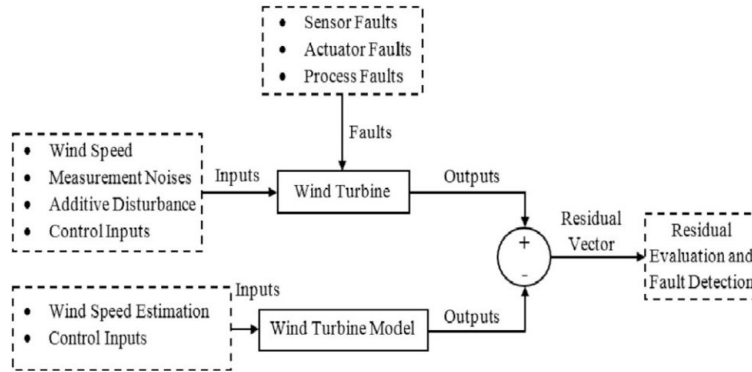


Figure 2.1: FDI Residual-based approach Habibi et al. (2019)

This approach is based on a method which compares a model to wind turbine output. This can be described mathematically as:

$$\mathbf{d} = \mathbf{y}_m - \mathbf{y}_w \quad (2.3)$$

Where:

- \mathbf{d} = Residual vector
- \mathbf{y}_m = Model output
- \mathbf{y}_w = Turbine output

If the residual vector exceeds a certain pre-set threshold, a fault can be assumed.

2.2. Previous HF works on PYC wind farm

The essence of this thesis is to understand the underperformance on a more wind farm wide approach. A preliminary research on this topic was done by a fellow student, T. Kaniewski (Kaniewski, 2022). This research was carried out on turbine level. In the report, the first correlations between different parameters measured by the SCADA instruments are plotted. Possible reasons for underperforming were addressed by analyzation of those plots. Additionally, the high-frequency data is compared to the more commonly used 10 minutes averaged data. This to get better insights in the differences between data from different sampling frequencies.

One of the worst performing turbines (the LO1 turbine) and one of the best performing turbines (the LO65 turbine) were pointed out for analysis. Firstly, MET mast data was analysed to understand wind behaviour, turbulence and other parameters without interference of turbines. Consequently, this MET mast data was compared to the LO1 turbine. Correlation plots between the LO1 turbine and the met mast were made regarding wind direction and wind speed. The next step in this research was the comparison between the LO1 and LO6 turbine. At this stage, two different turbines were analysed side-by-side. Active power and wind speed correlation plots were made between the two turbines. Also, turbulence intensity and torque behaviour was visualised to make a first comparison between the turbines regarding these parameters.

An overview of the visualisations made by T.Kaniewski are given below. A backslash in the table, a relation between two different turbines or a met mast and a turbine, was investigated.

Table 2.1: Correlations investigated in research by T.Kaniewski

	X	Y	Turbines
1	Active power	Active power	L01/L06
2	Wind speed	Active power	L01, L06
3	Wind speed	Wind speed	MET, MET/L06, L01/L06
4	Wind direction	Wind direction	MET, MET/L06
5	Wind speed	TI	MET, L01, L06
6	Wind speed	Torque	L01
7	Wind direction	TI	MET, L01, L06
8	Wind speed	Pitch	L01, L06
9	RPM	Active power	L01, L06

Recommendations of the work from T. Kaniewski will be regarded in this thesis. For more information about the research done, consult the report High-Frequency SCADA Data for Performance Analysis of Multiple Wind Turbines in Complex Terrain Kaniewski (2022).

2.3. Use of HF data in research

2.3.1. HF data filtering

In a paper about machine learning techniques, John Thomas Lyons makes use of filters to clear data from outliers Lyons and Göçmen (2021). This is done to create a data set that can be used to build a model for analysing future data sets. The techniques described by J.T. Lyons in the paper will be shortly discussed below and can be of use for the research on PYC wind farm.

When filtering data, it is important that there is a balance between filtering data, and preserving a detailed sample size. When filtering, outliers can be removed to create a normal behaving dataset, but when over removing points this can cause an unreliably small data set Lyons and Göçmen (2021). For the filtering process for the HF data from Horns Rev I wind farm, two methods have been used. Firstly, the data has been filtered by the use of pre-defined operational maxima and minima of the wind turbine. Secondly, to filter out still existing abnormal behaviour, a detection algorithm called the Local Outlier Factor (LOF) is used. This factor is used to calculate a percentile of outliers to be further removed. By combining these two techniques, a power curve of one of the turbines was filtered successfully. This approach can be applied to more than only power curve data. Filtering requirements will change, but the approach will be the same. An overview of the power curve filters is given in the table below:

Step	Filter name	Filter Type	Filter conditions
1	Curtailements	Operational	ActivePowerSP <Prated
2	Cut In	Operational	Wind speed <minimum Active Power <minimum
3	Cut Out	Operational	Wind speed >maximum
4	Rated	Operational	Active power >Prated
5	LOF	Machine Learning Detection	n-neighbours contamination
6	LOF	Machine Learning Detection	n-neighbours contamination

Table 2.2: Filters applied to Horn Rev I Wind Turbine

In a report on HF data analysis by E. Gonzalez et al., a way of filtering is discussed that can be done preliminary to the filtering steps as discussed above. SCADA data can flag abnormal events at different components. This flagging can be used as a first filter to the HF data Gonzalez et al. (2017).

2.3.2. WTG HF data analysis

In Antinio Notaristefano (2021), the added value of HF data is investigated. In that specific research, turbine dynamics with a smaller resolution than 10 minutes are investigated: upwards and downward ramps in wind speed, as well as turbine yawing. Wind speed ramping and yawing response can not be observed in 10 minutes averaged data, but do have influence on turbine performance and power curves Antinio Notaristefano (2021). In power curves, these events are visualised by scattering around the average operational power curve (OPC). When imposing limitations on the standard deviations regarding wind speed and wind direction, ramping of wind speeds and high-frequent changes in wind direction are (partially) filtered out. As a result, the power curve scattering decreases.

In a report written by M. Wilkinson, another analysis was performed on the potential use of HF data sets (Wilkinson, 2016). According to Wilkinson (2016), HF data sets can potentially give insights in the following turbine control and dynamics:

Table 2.3: Properties able to investigate by using HF data sets (Wilkinson, 2016).

	Property	Lowest frequency [Hz]
1	Speed control	1
2	Yawing strategy	0.2
3	Start/stop policies	0.2
4	Tower/foundation dynamics	1
5	Effects of TI on power	-

In the report, multiple examples of analyses carried out on above properties are briefly discussed. Regarding yawing manoeuvring of the turbine, wind speed dependent yaw thresholds are possible to detect when looking at HF yawing time series. Tower and foundation dynamics also occur at higher frequencies than a frequency belonging to a sampling time of 10 minutes. The first mode of tower frequency in the order of magnitude of 1 Hz.

In Gonzalez et al. (2019), HF data is used to model normal behaviour. This is explained in the next subsection. Before creating these Normal Behaviour Models (NBM's), the HF data has to be pre-processed and filtered (Gonzalez et al., 2019). In the pre-processing phase, Alarm data is gathered and divided per component, to prepare it as input for the filtering process. As a second step, the turbine data is filtered. This is done by using power output data, as well as wind speed data measured at the nacelle. As the nacelle wind speed measurements contain uncertainties, other parameters are used, in combination with wind speed measurements, to ensure a more robust filtering process. Other parameters used for this purpose are pitching of the blades, as well as rotational speed of the rotor (Gonzalez et al., 2019). In the paper Gonzalez et al. (2019), the above prescribed filtering steps are used to create normal behaviour data for the NBM. However, the filtering process above can also be used to detect abnormal behaviour (or underperformance) as a final aim, without producing a NBM afterwards.

2.3.3. Modelling with HF data sets

As described in the previous subchapter, in phase two of the paper by E. Gonzalez et al., filtered HF data is used to do a performance analysis on the wind farm (WF). These analysis techniques can be used to map the performance of multiple wind turbines in the farm, and to compare wind turbines to others.

Before discussing some fundamental techniques to produce a Normal Behaviour Model, it should be noted that doing calculations on a filtered data set with many SCADA data tags, can be very time-consuming Lyons and Göçmen (2021). Therefore, one should only be using the essential data that is needed for the analysis. A selection is based on:

- A number of selected turbines
- A number of SCADA tags considered and used in the analysis

Making an accurate data selection is done to minimise computational time for further analysis and modelling. Below, the fundamental data modelling approaches as introduced at the beginning of this paragraph are discussed.

1. Method of Bins (MOB)

The Method of Bins is a method in which the operational wind speed spectrum is divided into multiple bins. For each bin, a mean power output and standard deviation is calculated by using the data within that bin. The standard deviation represented the uncertainty of the performance prediction of the bin.

A different approach to the Method of Bins is the translated data approach. The first steps of the process is equivalent. To create the reference power curve, first, wind speeds are divided into bins and consequently the average power output of each bin is calculated. The average value of two consecutive bins, $P_{i,i+1}$, is used for calculating the residual. Afterwards, all data points of two consecutive bins are linearly translated towards the centre of the two bins Cambron et al. (2016). These translated values are compared to $P_{i,i+1}$ to calculate the residual Cambron et al. (2016):

$$R_{i,i+1} = P_{i,i+1} - P_{trans} \quad (2.4)$$

Where:

$R_{i,i+1}$ = Residual of a translated point and the average of two wind speed bins

$P_{i,i+1}$ = Average power output of two wind speed bins

P_{trans} = Power output of the translated data point

By translating the measured data points, a normal distribution between two bins is guaranteed.

2. k-Nearest Neighbours

This algorithm uses a dataset to estimate other points within the reach of the dataset. Within a pre-set range, points around the unknown value are collected. The chosen range often relies heavily on the Root Mean Squared Error (RMSE). This average of n selected data points $(X_{(1)}, Y_{(1)}), \dots, (X_{(n)}, Y_{(n)})$ within the region $\|X_{(1)} - x\| \leq \dots \leq \|X_{(n)} - x\|$ is the estimate of the unknown value.

3. Random Forests

By bagging, multiple datasets can be acquired from one dataset. This is done by randomly picking and replacing a value from the original dataset. From these multiple datasets, multiple decision trees are produced. This is done by choosing a decision node based on optimal entropy distribution. A parameter x with unknown value y can be run through all decision trees. The average outcome from all decision trees is the estimation of value y . An overview of this algorithm is shown below:

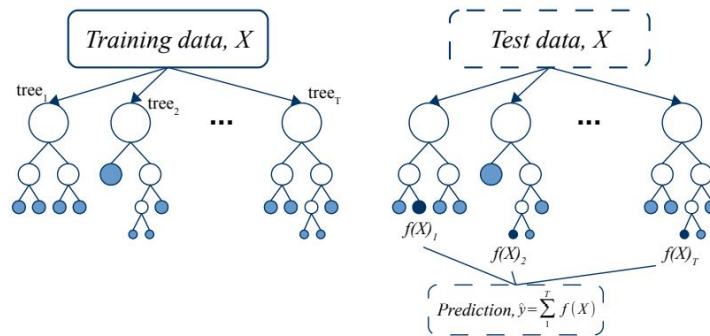


Figure 2.2: Random Forest Algorithm Gonzalez et al. (2017)

These performance modelling techniques can be used to see if the actual performance is deviating from the model made. If this is the case, this can result in a diagnosis of under or overperformance. Analysis of how the turbines should perform is already done. Results from this analysis can be compared to the data collected by the turbines. However, in a further part of the thesis, to better understand and analyse the performance of a turbine, these models might be of use.

2.4. Causes for gap between theory and operation

Before starting an analysis on performance, it is good to get better insights in the possible causes of discrepancy in the first place. An overview of possible causes can help to understand the underperforming behaviour of the farm. Therefore, in this paragraph, the most important causes for this research are listed below Tücer (2016).

- **Data.** For wind farm analysis, data sets are used. These data sets contain information about the site conditions of the farm. This concerns information about the wind, as well as information about the topography of the site.

Data about the wind itself is collected at a certain sampling time. Data is an approximation of reality, but does not fully represent the wind conditions on site. The sampling, but as well the way data is calibrated, may cause differences between theory and operation. On top of that, the way data is processed and validated can have influence on signs of underperformance.

Elevations as well as roughness are included in the topography of a site. Also, obstacles can be considered part of the topology of a site. All these parameters influence the wind properties at the site.

- **Long term correlation.** Because of the variability of wind resource on an annual scale, wind resource data sets are correlated with long term reference data. This long term reference data set is best to stem from a nearby source. By the use of this correlation, future wind data can be predicted, and annual variance is taken into account less in the modelling. This process is called Measure-correlate-predict (MCP) Beltrán et al. (2010).
- **Modelling on wind farm performance** is done in different ways. Regression analysis, but also pre-developed software is used to model wind farm performance such as WindPro or WindSim Tücer (2016). As a simplified version of real conditions, models can cause discrepancy.
- **Measurements.** The sensor type as well as the positioning can influence measurement outcomes Tücer (2016). For example, there can be differences in wind speed measurement depending on the kind of anemometer used Tücer (2016). However, IEC standards are prescribed to minimise the variance coming from measurements.
- **Technical problems and power curve deviation.** This thesis will predominantly focus on the discrepancy coming from underperformance due to technical issues, and thus power curve deviation.
- **Losses.** Losses can come from different origins. A physical phenomenon that causes losses are wake effects, which effects can be wrongly estimated. An overview of other loss mechanisms is given below Mortensen (2012):

Loss category	Additional losses type	Typical values
1 Availability	<ul style="list-style-type: none"> • turbine availability • grid availability 	97% > 99%
2 Electrical efficiency	<ul style="list-style-type: none"> • operational electrical efficiency • wind farm consumption 	1-2%
3 Turbine performance	<ul style="list-style-type: none"> • power curve adjustments • high-wind hysteresis 	1-2%
4 Environmental	<ul style="list-style-type: none"> • blade degradation and fouling • icing or temperature shutdown 	1-2%
5 Curtailments	<ul style="list-style-type: none"> • wind sector management • noise, visual and environmental 	Design dependent

Figure 2.3: Overview of technical wind turbine losses Mortensen (2012)

To give a different overview of discrepancy, J.C.Y. Lee investigated loss mechanisms as well as uncertainty for Wind Resource Analysis (WRA) in wind farms. In an article published by EAWE, Lee firstly visualised loss categories in the bottom figure Lee and Fields (2021). These are the categories as proposed by the IEC 61400-15 standard. For example, the category “generic power curve adjustment” is

the subcategory which stands for the difference between the advertised power curve and actual power performance in standard test conditions (STC).

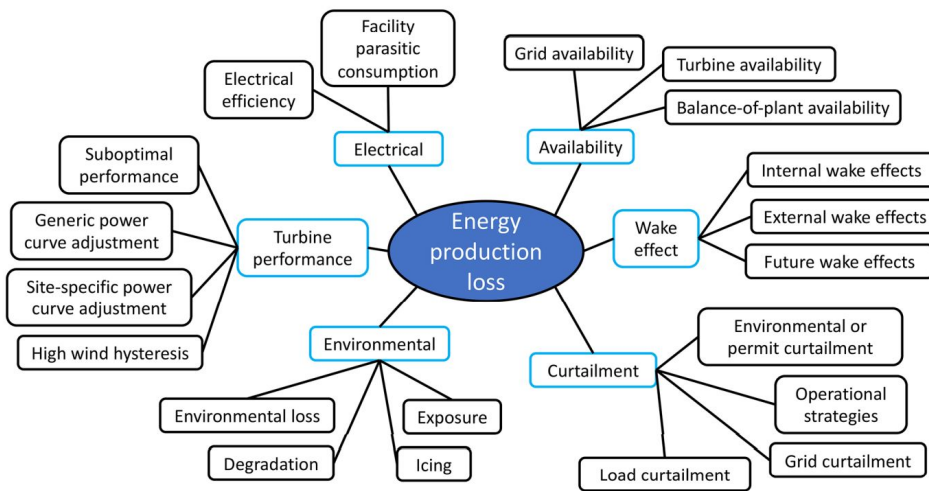


Figure 2.4: Visual representation of wind turbine losses Lee and Fields (2021)

Using these categories, in the same article, causes for energy loss were further investigated and visualised per subcategory. This was done by gathering information from multiple reports in which wind resource at different farms was analysed. The literature review results from the turbine performance category are shown in the figure below. An orange dot represents an observed loss, and a blue dot represents an estimated loss. The numbers on the right of the figure show the number of data points for observations as well as estimations. On the horizontal axis is the fraction of lost Annual Energy Production (AEP) given as a percentage Lee and Fields (2021).

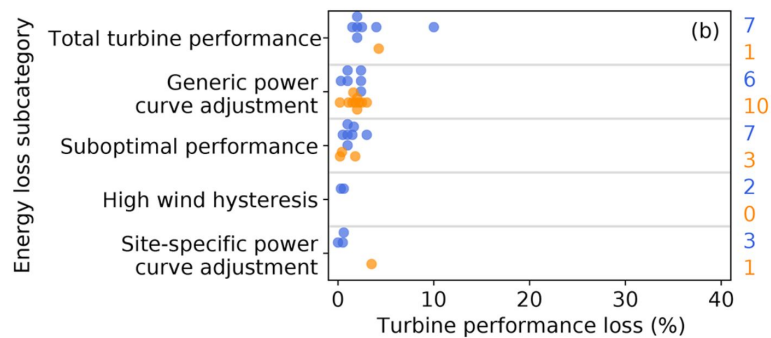


Figure 2.5: Visual representation of turbine performance subcategory data Lee and Fields (2021)

In addition to all categories and subcategories from IEC, Lee proposed two more categories Lee and Fields (2021). Firstly, the **First few years of operation losses**. A wind farm which is in its first operational years often does not produce at maximum capacity. Secondly, the **blockage effect**. Blockage effect is the slowdown of wind speed upwind of the farm.

Besides losses, uncertainties are also discussed in the paper. As the main goal of this thesis will be to contribute to the understanding of underperformance, and therefore losses, uncertainties will not be further elaborated on in this literature review. If information regarding WRA uncertainties will be needed, Lee and Fields (2021) will be consulted.

2.5. Onshore site conditions

The complex terrain of the PYC wind farm makes it important to have a good understanding of the fundamentals of wind farm site conditions. Due to different elevations and forestry around the wind turbines, performance is greatly influenced by the site conditions.

Turbulence Intensity (TI)

Turbulence Intensity is a measure of variation of wind speed over time. For a chosen time interval, TI can be calculated by the fraction of wind speed standard deviation over the mean wind speed measured in that interval. A common time interval to choose is 10 minutes averaged data.

$$TI = \frac{\sigma_u}{U_{10}} \quad (2.5)$$

Where:

σ_u = Wind speed standard deviation
 U_{10} = Average 10 minute wind speed

IEC have developed three standard TI classes to classify turbulence intensity on different sites. Three TI curves are representing the TI behaviour for classes A, B and C. The wind speed standard deviation for calculating TI is mathematically described as follows by IEC (Øistad, 2015):

$$\sigma_1 = I_{ref}(0.75V_{hub} + 5.6) \quad (2.6)$$

Where:

σ_1 = Wind speed standard deviation
 I_{ref} = TI class specific constant
 V_{hub} = Wind speed at hub height
 b = Empirical constant

Table 2.4: TI class specific constants Øistad, 2015

	I_{ref} [-]
Class A+	0.18
Class A	0.16
Class B	0.14
Class C	0.12

Wind shear

Wind velocity increases with increasing height. This phenomenon is described by height profiles. These height profiles consist of two parts: a logarithmic and an exponential part. The mathematical description of the wind profile is given below.

$$U(z) = \frac{u_*}{\kappa} * \ln\left(\frac{z - z_h}{z_0}\right) \quad (2.7)$$

Where:

$U(z)$ = Wind speed at height z
 u_* = Friction velocity
 κ = Von Kármán's constant
 z = Height
 z_h = Zero-plane displacement
 z_0 = Roughness length

The above equation can be rewritten so that the wind speed at height z_2 is given as a function of the wind speed at height z_1 (Holmes, 2007):

$$U(z_2) = U(z_1) * \frac{\ln\left(\frac{z_2 - z_h}{z_0}\right)}{\ln\left(\frac{z_1 - z_h}{z_0}\right)} \quad (2.8)$$

Where:

$U(z_2)$ = Wind speed at height z_2

$U(z_1)$ = Wind speed at height z_1

z_2 = Height level 2

z_1 = Height level 1

Multiple roughness lengths for different terrains are given in the table below (Burton et al., 2011):

Table 2.5: Roughness lengths for different terrains.

Terrain type	Roughness length [m]
Flat terrain, ice	0.00001 - 0.00003
Calm sea	0.0002 - 0.0003
Sand	0.0002 - 0.001
Mown grass	0.001 - 0.01
Low grass	0.01 - 0.04
Fallow field	0.02 - 0.03
High grass	0.04 - 0.1
Forest / Woodland	0.1 - 1
Built-up area	1 - 2
City	1 - 4

The magnitude of the wind speed change over height is described by the shear factor parameter α . In a simplified manner, the wind shear is calculated by the gradient of horizontal wind velocity over height, on a logarithmic scale.

$$\alpha = \frac{\ln\left(\frac{v_2}{v_1}\right)}{\ln\left(\frac{h_2}{h_1}\right)} \quad (2.9)$$

Where:

α = Wind shear coefficient

v = velocity

h = height

According to IEC standards, as an average value, and α value of 0.2 can be taken. For extreme weather conditions with a return period of once every 50 years, a shear factor of 0.11 is used for modelling wind shear profiles (IEC, 2019).

In a different research on onshore wind shear (Ray et al., 2006), standard power law coefficients were given for different terrains. An overview of those is listed below:

Table 2.6: Standard shear factor values for different terrains (Ray et al., 2006)

Terrain Description	Power law exponent, α
Smooth, hard ground, lake or ocean	0.10
Short grass on untilled ground	0.14
Level country with foot-high grass, occasional tree	0.16
Tall row crops, hedges, a few trees	0.20
Many trees and occasional buildings	0.22 - 0.24
Wooded country, small towns and suburbs	0.28 - 0.30
Urban areas with tall buildings	0.4

Wind directional veer

Besides the variation of wind speed over height, wind direction can vary with height as well. This is known as directional veer. Veer can be described as the change of wind direction per unit length:

$$\beta_{bulk} = \frac{\theta_{top} - \theta_{bottom}}{z_{top} - z_{bottom}} \quad (2.10)$$

Where:

θ = Wind direction

z = height

Horizontal wind speed variation

Horizontal influences from surroundings on the wind flow are encountered when the wind is disturbed by obstacles. In front of an obstacle, wind can stagnate while downstream of an obstacle a wake occurs slowing down the wind. A brief overview is given in the table below.

Distance from object	Fraction of initial wind speed
-4H	80-90%
5H	25-50%
8H	50-70%
10H	50-70%
15H	70-80%
20H	100%

Table 2.7: Horizontal wind speed influenced by an object, represented as a percentage from the initial wind speed upfront of the object.

Besides site specific objects like forests and housing, wind turbines itself also influence the wind flow. Downstream of the turbine, this influence is described as wakes. To describe and understand those wakes, multiple models have been developed. A basic model describing the wind velocity as well as the geometry of the wake is the Jensen Model. The wind speed at a certain distance behind the turbine can be calculated as follows:

$$1 - \frac{U_w}{U} = \frac{1 - \sqrt{1 - c_T}}{(1 + 2ks)^2} \quad (2.11)$$

Where:

U_w = Waked wind speed

U = Undisturbed wind speed

c_T = Thrust coefficient

k = Wake decay coefficient

s = Dimensionless distance

The dimensionless distance s is calculated as follows:

$$s = \frac{x}{D} \quad (2.12)$$

Where:

x = Distance behind the turbine

D = Rotor diameter

Wind speeds, when consequently measured at a site, can be collected into wind speed bins. When the measurements are carried out on a long term, and all measurements are collected in the bins, a histogram can be constructed. For wind speeds, this histogram corresponds to a Weibull distribution. With the created histogram, a Weibull distribution can be fitted. The general description of such a Weibull distribution is mathematically described as follows:

$$f(U) = \frac{k}{a} \left(\frac{U}{a} \right)^{k-1} e^{-\left(\frac{U}{a}\right)^k} \quad (2.13)$$

Where:

k = Shape parameter

a = Scale parameter

The shape and scale factors can be found when fitting the Weibull to the site-specific nominal wind distribution (NWD).

2.5.1. IEC standards regarding obstacles

In the IEC standards, there is information given about in- and excluding sectors due to obstacles. In these standards, it is described how to calculate disturbed sectors for obstacles and operating wind turbines. Table 2.8 below shows different object dimensions and according affected sectors. Below, the table is further explained.

Table 2.8: IEC obstacle requirements

Distance	Sector	Maximum obstacle height from terrain surface
< 2L	360°	<1/3 (H - 0,5 D)
≥ 2L and <4L	Preliminary measurement sector	<2/3 (H - 0,5D)
≥ 4L and <8L	Preliminary measurement sector	<(H - 0,5D)
≥ 8L and <16L	Preliminary measurement sector	<4/3 (H - 0,5D)
≥ 2L and <16L	Clearly outside preliminary measurement sector by 40° or more	No limit to height

If an obstacle is in accordance with a certain distance and sector (columns one and two), and it exceeds the height limit as stated in the third column, the object has to be taken into account as a disturbing obstacle. When taken into account, the equivalent diameter D_e of the object should be calculated accordingly:

$$D_e = \frac{2l_h l_w}{l_h + l_w} \quad (2.14)$$

Where:

D_e = Diameter

l_h = Object height

l_w = Object width

This diameter D_e , together with the distance to the source, is then used to make an estimation of affected sectors by the obstacle:

$$\theta_{object} = 1,3 \arctan\left(2,5 \frac{D_e}{L} + 0,15\right) + 10 \quad (2.15)$$

Where:

θ_{object} = Size of sector to exclude
 D_e = Equivalent object diameter
 L = Object distance to source

For wind turbines, the rotor diameter as well as the distance from the turbine to the source is used to approximate the sector influenced by the turbine. When applied to wind turbines, equation 2.15 is changes to:

$$\theta_{turbine} = 1,3 \arctan\left(2,5 \frac{D}{L} + 0,15\right) + 10 \quad (2.16)$$

Where:

$\theta_{turbine}$ = Size of sector to exclude
 D = Rotor diameter
 L = Turbine distance to source

2.6. WTG dynamics and control

In this subchapter, relevant wind turbine dynamics and control theory for this research are discussed. Used theory from books, articles and papers are summarised.

Aerodynamics

The aerodynamic behaviour of the wind turbine rotor blades are described by the (c_p, λ) - curve. This curve shows the relation between the power coefficient (c_p) and the tip speed ratio λ . The curve is rotor blade specific and determined from blade element momentum theory (Zaaijer & Viré, 2021). This (c_p, λ) - curve is zero for a tip speed ratio of zero, (as no power is produced when the rotor is in stationary conditions), and at the maximum tip speed ratio value of the turbine. The curve has a turbine-specific maximum at which it is controlled to operate between cut-in and rater wind speed. The definition of the power coefficient and tip speed ratio is given below:

$$c_p = \frac{P}{\frac{1}{2} \rho U^3 \pi R^2} \quad (2.17)$$

Where:

c_p = Power coefficient
 P = Possible power generated by blade
 ρ = Air density
 U = Wind speed
 R = Rotor radius

$$\lambda = \frac{\Omega R}{U} \quad (2.18)$$

Where:

λ = Tip speed ratio
 Ω = Rotor speed

This (c_p, λ) - curve can be translated towards a wind speed specific (Q, Ω) curve. This curve has the same shape for different wind speed values, but is scaled differently (Hau, 2013).

The behaviour of the blade upstream of the drive train, is transmitted through the LSS (Low Speed Shaft), and gearbox towards the HSS (High Speed Shaft). At this HSS, the torque and rotational speed behaviour from the blades should match the torque and rotational speed behaviour from the generator-side of the drive train. This is either done by setting the desired torque at the generator, or by setting the desired rotational speed of the generator (Zaaijer & Viré, 2021).

Partial load control

Between cut-in and cut-out wind speed, there exist two different stages of operation: partial load and full load. For both operational conditions, there are different controls. At partial load, the control is set to maximise power output and therefore the power coefficient c_p . This is done by keeping the torque coefficient constant. This comes down to keeping the torque at the HSS (Q_{HSS}) proportional to the rotational speed at the HSS (Ω_{HSS}) squared, with the proportionality constant k_{opt} (Zaaijer & Viré, 2021). This accounts for the HSS as well as for the rotor itself. It may include an additional torque control loop to damp drive-train torsion mode and resonant loads. (Novaes Menezes et al., 2018). A mathematical definition of this control mechanism is shown below.

$$Q_{HSS} = \frac{\frac{1}{2}\rho c_{P,max}\eta_{gearbox}\pi R^5}{r_{gearbox}^3\lambda_{design}^3}\omega_{HSS}^2 \quad (2.19)$$

Where:

- Q_{HSS} = Torque at the high speed shaft
- ω_{HSS} = Rotational speed at the high speed shaft
- $c_{P,max}$ = Maximum power coefficient
- $\eta_{gearbox}$ = Gearbox efficiency
- $r_{gearbox}$ = Gearbox ratio
- λ_{design} = Tip speed ratio for maximum power coefficient

Apart from equation 2.19, there is another characteristic relation active in partial load. As the tip speed ratio is kept at its design value λ_{design} , the rotor speed and wind speed are linearly related in the partial load region. This can mathematically be described as follows:

$$\Omega = U\left(\frac{\lambda_{design}}{R}\right) \quad (2.20)$$

Full load control

During full load control, the turbine is not controlled to maximise the power coefficient, but to maintain at rated power. The power generated by a rotor blade is described by rewriting equation 2.17 towards:

$$P = \frac{1}{2}\rho c_P U^3 A \quad (2.21)$$

Where:

- P = Possible power generated by blade
- A = Rotor swept area

To keep the power constant, the power coefficient and the wind speed cubed should be inversely proportional. This decrease in power coefficient is obtained by pitching the blades of the turbine. Besides decreasing the power captured by the rotor, pitching to vane (pitching towards the apparent wind direction) had the effect of reducing loads on the turbine. The pitching of the blades is controlled by tracking the rotor speed, and correcting for deviations from the reference rotational speed by pitching of the

blades. Keeping the rotational speed at a desired value also keeps the torque at the HSS at a desired pre-defined rated control value.

$$\Omega_{ref} = \frac{P_{rated}}{Q_{rated}(\eta_{generator}\eta_{converter})} \quad (2.22)$$

Where:

- Ω_{ref} = Reference rotational speed for control
- P_{rated} = Turbine rated power
- Q_{rated} = Set torque value
- $\eta_{generator}$ = Generator efficiency
- $\eta_{converter}$ = Converter efficiency

From the above theory it can be concluded that during full load control the torque and rotational speed are predefined for turbine control and used as reference (Zaaijer & Viré, 2021). In full load, an additional control objective is mechanical load reduction. Reducing loads can be needed due to the high wind speeds that can damage the WT structure in the long run (Novaes Menezes et al., 2018).

To better understand the two different control modes of the turbines, the figure below is presented. This figure contains the aerodynamic behaviour of the NREL 5 MW reference turbine. The turbines at PYC wind farm are another type, consisting of other blades, but the general aerodynamic properties are similar.

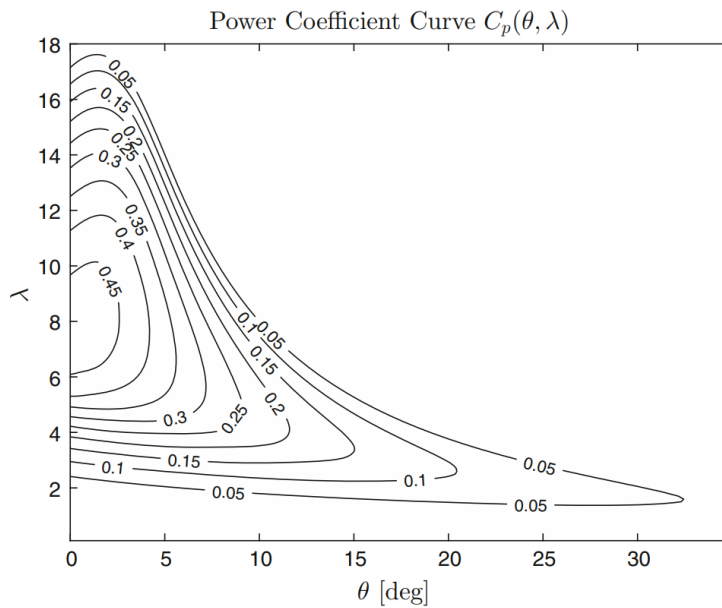


Figure 2.6: Power coefficient for varying tip speed ratio and pitch angles (Lio, 2018).

As can be seen in the figure, optimal C_p is obtained when no pitching is applied to the blades. A desired tip speed ratio λ is strived for when below rated wind speed. This maximum tip speed ratio is obtained by setting the torque by the control mechanism from equation 2.19. In full load control, the rotor speed has reached its maximum value. This means for every incoming wind speed above rated wind speed, there is a matching tip speed ratio λ . Therefore, the pitch angle θ has to be changed in order to control the C_p value to the desired value, as the tip speed ratio is no control variable above rated wind speed (Lio, 2018).

The above described control strategies are parts of the power and torque curves. These two characteristic wind turbine curves are described below.

Power curve

The wind turbine power output for different wind speeds is represented by the power curve. A wind turbine power curve can be described by four regions, of which two are between cut-in and cut-out wind speed. In these two regions (region 2 and 3) the turbine is in partial and full load respectively. These two conditions are explained above. A mathematical description of the power curve is given below:

$$P(t) = \begin{cases} 0 & \text{for } v(t) < v_{ci} \\ P_r \frac{v^3(t) - v_{ci}^3}{v_r^3 - v_{ci}^3} & \text{for } v_{ci} < v(t) < v_r \\ P_r & \text{for } v_r < v(t) < v_{co} \\ 0 & \text{for } v(t) > v_{co} \end{cases} \quad (2.23)$$

Where:

$P(t)$ = Power output
 P_r = Rated wind speed
 $v(t)$ = Wind speed at the turbine rotor
 v_{ci} = Cut-in wind speed
 v_{co} = Cut-out wind speed

The different power curve regions can be described as follows (Novaes Menezes et al., 2018):

- Regions 1 and 4
The wind turbine is out of operation, commanded by the supervisory control.
- Region 2
The partial load region of the turbine, as explained above.
- Region 2.5
The rated speed should be maintained constant, and the torque should be slightly increased until its rated value, ensuring a smooth transition between Regions 2 and 3.
- Region 3
The full load region of the turbine, as explained above.

A visual representation of the power curve and its different regions is given below. Region 2.5 is not visible in the power curve itself, as within this region no wind speed or active power changes occur.

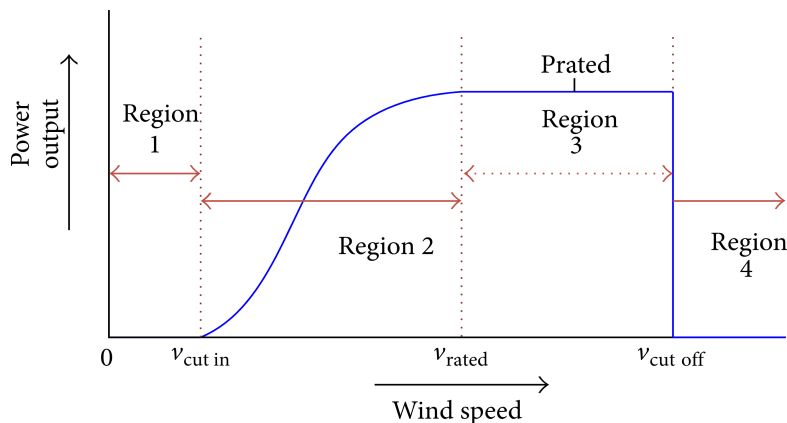


Figure 2.7: General regions of the wind turbine power curve (Sohoni et al., 2016).

When comparing the turbine power curve to the operational data, it is of importance to correct wind speeds measured for the air density deviation from the site compared to the reference air density used by the manufacturer. This has is important to compensate for differences in energy content of the incoming wind (Carullo et al., 2021). In the equation below, the correction step is explained mathematically:

$$v_{cor} = v_{exp} \left(\frac{\rho_{air}}{\rho_{ref}} \right)^{1/3} \quad (2.24)$$

Where:

- v_{cor} = Corrected wind speed
- v_{exp} = Measured wind speed at site
- ρ_{air} = Air density at site
- ρ_{ref} = Reference air density

Torque curve

The torque behaviour is described by a characteristic relation with rotor speed. The relation is made up of different regions, just as for the power curve. The regions are discussed below.

- Region 0-A
In this region, no power is produced, and no torque is generated. The turbine is being curtailed, or is in idling conditions below cut-in wind speed.
- Region A-B
Transition region from cut-in rotor speed towards optimal- c_p torque control. The turbine starts producing power and torque is being built up towards the optimal torque-rotor speed relation (as was described by equation 2.19).
- Region B-C
Optimal- c_p torque control as described by equation 2.19.
- Region C-D
Transition region between full load control and optimal- c_p control. When maximum rotor speed is reached (point C), the torque increases, and behaviour starts deviating from equation 2.19 again. Torque builds up towards its maximum value (point D). This point coincides with rated power, and covers the whole rated region of the power curve. In practice, line C-D has a slope, just like in region A-B.

A visual representation of the torque curve is given below in figure 2.8. The black quadratic relation is the optimal- c_p curve. The black dashed lines are torque relations for different wind speeds. The optimal- c_p curve is a collection of all the optima from the torque relations at individual wind speeds.

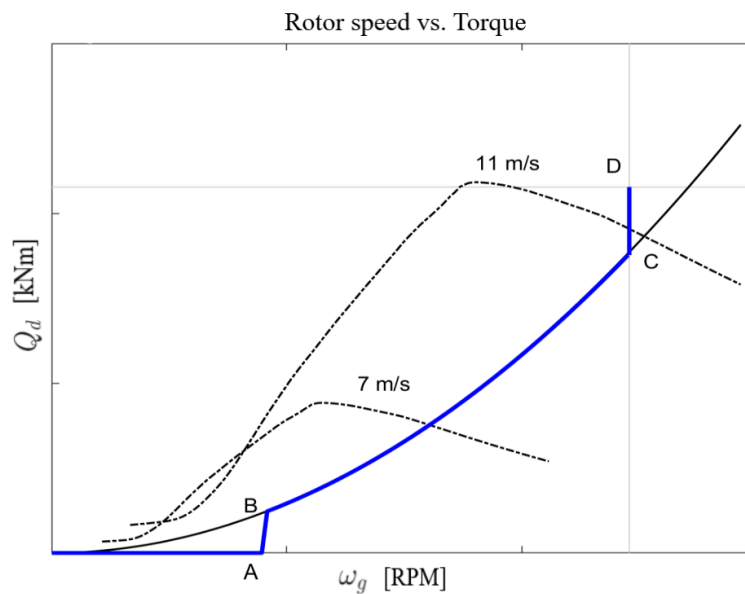


Figure 2.8: Rotor speed - Torque relation (van Wingerden, 2021).

Derated power control

When turbine power output is being reduced and turbine behaviour is not corresponding with the operational power curve and torque curve any more, the turbine is within derated conditions (Lio et al., 2018). Within derated conditions, wind turbine power output, $P_{derated}$, is regulated to a value lower than rated power, independent of wind speeds.

For this research, normal turbine operation is investigated. Curtailments are filtered out of the data. Therefore, in the first place, derated conditions are no part of the analysis performed in this thesis. However, derated turbine conditions can still be present in post-filtered data, when curtailments are not logged correctly within the raw data. It is important to understand these occurrences when present in the post-filtered data. For that reason, derated conditions are discussed.

The derated wind speed is the minimum wind speed needed for producing the derated power output. The derated rotor speed is the rotor speed reached when the wind turbine reaches derated power at derated wind speed. Both can be derived from the power curve within the partial load region (Lio et al., 2018).

$$U_{derated} = \frac{P_{derated}}{\frac{1}{2}\rho\pi r^2 C_{p,max}\eta} \quad (2.25)$$

Where:

$U_{derated}$ = Derated wind speed
 $P_{derated}$ = Derated power
 $C_{p,max}$ = Maximum power coefficient
 η = Generator efficiency

$$\omega_{derated} = \frac{v_{derated}\lambda_{design}}{R} \quad (2.26)$$

Where:

$\omega_{derated}$ = Derated rotor speed

When wind speeds are above $v_{derated}$, the turbine can set rotor speed at an arbitrary value, the set point rotor speed ω_{sp} . This can be explained by the fact that, above $v_{derated}$, the desired power coefficient can be obtained by any point on the corresponding isoline (figure 2.6)¹. The fact that rotor speed can be set at any arbitrary value within derated conditions, results in multiple derated control strategies. The chosen strategy is often related to the thrust coefficient C_T , which is tip speed ratio, loads and pitch angle dependent. For a given power coefficient, there is a range of thrust coefficients a turbine can operate in (Lio et al., 2018).

The most widely used derating strategy within the wind industry is the maximum rotor speed strategy, where the set point rotor speed is equal to the rated rotor speed. This is because a high rotor speed causes more kinetic energy, which can provide a higher possible inertia response to the grid when not in balance.

This is also the derating strategy applied at the PYC wind farm. Other derating strategies are constant tip speed ratio control, constant rotor speed control ($\omega_{sp} = \omega_{derated}$) or constant pitch control.

Storm control

Turbines can be equipped with a control algorithm that enlarges the wind speed range the turbine is able to operate in. Where a wind turbine without this additional control is limited to its cut-out wind speed, a turbine with storm control is able to exceed this cut-out wind speed. It produces less power inversely proportional to the increasing wind speed. This also accounts for the rotor speed. Below,

¹This in contrast to full load or partial load control, where the turbine is controlled to maintain a tip speed ratio or rotor speed value.

three diagrams are describing the development of active power, rotor speed and pitching when wind speed exceeds cut-out.

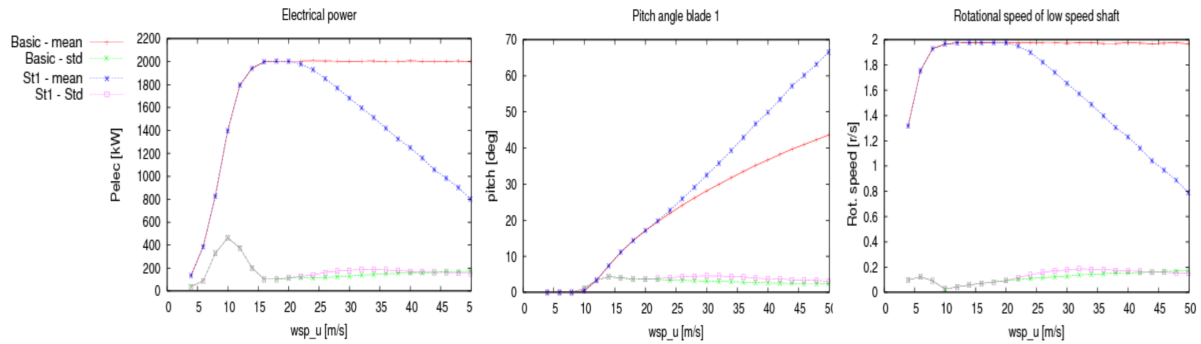


Figure 2.9: Active power, rotor speed and pitching during storm control (Markou & Larsen, 2009).

Ramping behaviour

Ramping behaviour of a wind turbine is of big importance for the variation of power production through time. It represents the turbine power output response to positive or negative wind speed fluctuations in the measurements performed by the turbine. Below, three different ramping properties are defined. These are ramping magnitude, duration and ramping change rate respectively. A change in power is considered a ramp if the properties are exceeding a pre-defined threshold (Ahn & Hur, 2022). The ramping magnitude is given as follows:

$$|P_{(t+\Delta t)} - P_{(t)}| > P_{val} \quad (2.27)$$

Where:

$$\begin{aligned} P_{(t+\Delta t)} &= \text{Power output after ramp} \\ P_{(t)} &= \text{Power output before ramp} \\ P_{val} &= \text{Threshold} \end{aligned}$$

The ramping change rate is the change of power output over time, as defined below.

$$\frac{|P_{(t+\Delta t)} - P_{(t)}|}{\Delta t} > PT_{val} \quad (2.28)$$

Where:

$$\begin{aligned} \Delta t &= \text{Ramp duration} \\ PT_{val} &= \text{Threshold} \end{aligned}$$

As a threshold for duration of a ramp, often a period of 4 hours is used. Regarding up- and down ramping, respectively 20% and 15% are used.

Ramping behaviour of wind and power output of a turbine are of relevance for understanding scattering around average turbine behaviour, and especially power curves. When wind speeds ramp up or down for a long enough time period, a turbine reacts and an increase in power output is the result. Between a wind speed increase and a turbine response, a delay exists due to rotor inertia. This delay is visible as scattering in the power curve. When wind speed ramps up and a turbine has not responded yet, this results in scattering below the power curve (Marked as (B) in the figure below). When wind speed ramps down, this results in scattering above the power curve (Marked as (A) in the figure below) (Antinio Notaristefano, 2021).

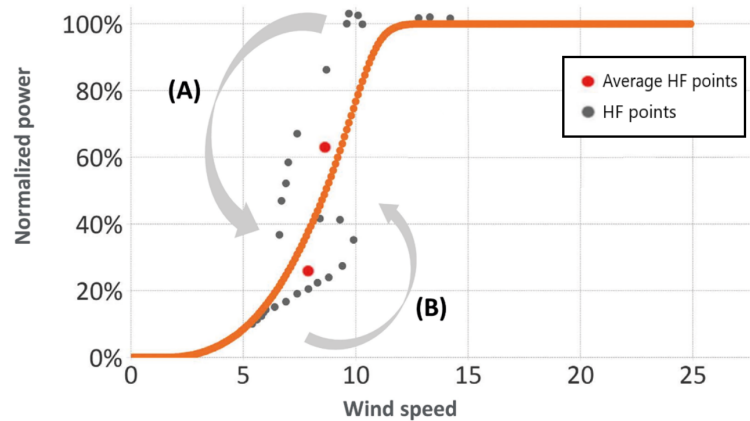


Figure 2.10: Ramping effects on power curve (Antinio Notaristefano, 2021).

Especially in the high-frequency domain of this performance analysis, this scattering will have big influence on the power curve. In the 10 minutes averaged domain, scattering will also be present, but less dominant. This because ramping behaviour is averaged together with more wind speed stationary data. This is also shown in figure 2.10. The red data points represent averaged data, while the grey data points represent HF data.

Pitching

Pitching behaviour for different wind speeds can be derived from the turbine behaviour during partial and full load conditions, as described earlier in this chapter. With wind speeds below cut-in, a turbine is not in operating conditions. At wind speeds close to zero, the turbine is in feathered position: the blades are pitched into the wind by applying a pitching angle close to 90 degrees. When wind speeds are close to cut-in wind speed, pitching is applied to gain momentum and get the turbine close to cut-in rotational speed. During partial load, the power output is maximised by controlling the torque to maximise the power coefficient. In this region, no pitching is applied, as pitching reduces the aerodynamic properties of the blades and therefore the c_p -value. Above rated wind speed, the power is stabilised at its rated value by controlling the rotor speed. This is done by active pitching of the blades. By pitching in this region, the power coefficient is decreased to its desired value to obtain rated power wind speeds above rated. Figure 2.6 shows how pitching results in a decrease in power coefficient.

Wind turbine state curves

The above described turbine control and dynamics can be summarised into so-called wind turbine state curves. These state curves show the theoretical behaviour of a variable wind speed pitch-controlled wind turbine during normal operational conditions for multiple relations. Below, the most widely used state curves are shown. These were described in (Sun et al., 2019), as well as in the course Wind Turbine Design at TU Delft (van Wingerden, 2021).

Added to the five state curves described in (Sun et al., 2019), is the expected relation of torque and pitching (figure 2.11e). When pitching is set to a non-zero value, either the torque is controlled at its maximum value, or no torque is experienced due to no production below the cut-in wind speed region (the two horizontal regions). In between, there is a vertical region of no applied pitching and torque build up to its maximum value. Two of the six state curves are the power curve and torque curve. These were described in detail earlier in this section.

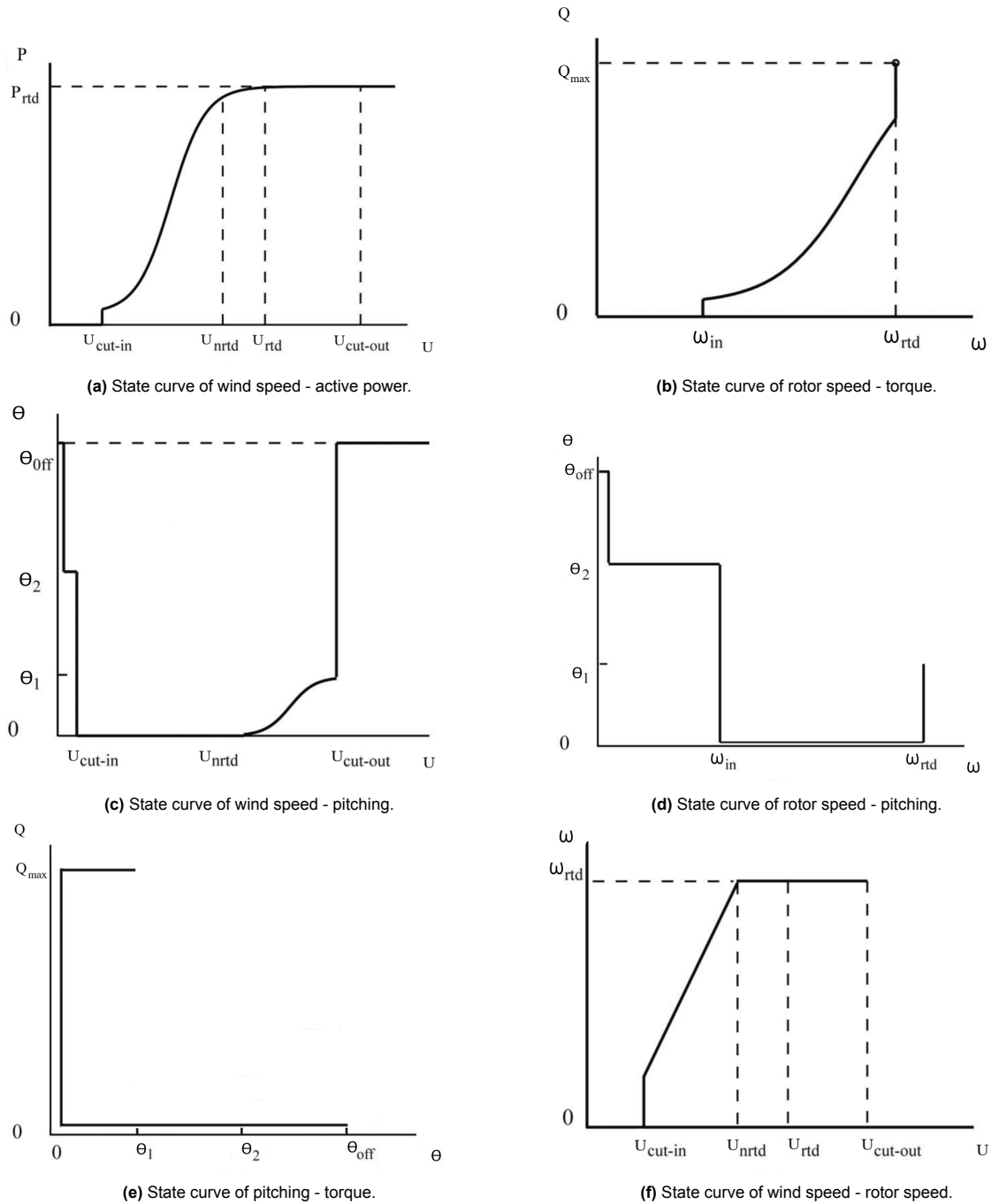


Figure 2.11: Different state curves of a variable wind speed pitch-controlled wind turbine.

Where:

- U_{nrtd} = Wind speed at which maximum rotor speed is reached
- Q_{max} = Torque level at rated power
- ω_{rtd} = Rated rotor speed
- ω_{in} = Rotor speed at cut-in wind speed
- θ_{off} = Feathered pitching value
- θ_1 = Pitching at cut-out wind speed
- θ_2 = Pitching below cut-in wind speed

2.7. Site-influence on onshore WTG performance

In this subsection, an overview of relevant research on the site-influence on WTG's is given. This overview will be a reference for the work in this thesis. When investigating site conditions at the PYC wind farm, only conclusions can be drawn when compared to similar research from other onshore wind farms.

In Murphy et al. (2020), the influence of wind speed shear as well as directional veer on wind turbine performance is given. The paper is investigating a wind farm in North America. Multiple parameters defining wind shear and wind veer in different ways are correlated with power under- or over performance.

One of the parameters analysed in this research is $\Delta REWS$ (Rotor Equivalent Wind Speed). This parameter is the difference between the integrated momentum through the rotor sector (which includes wind shear variation), measured hub height wind speed. When negative, the integrated rotor momentum is underestimated by the nacelle, and when positive it is overestimated by the nacelle. Another $\Delta REWS_{\theta}$ takes into account the different wind directions flowing into the rotor disk as well. A high $\Delta REWS$ value generally goes together with high shear values (Murphy et al., 2020).

Besides these $\Delta REWS$ values, also shear and veer as described in respectively equation 2.9 and equation 2.10 are tested for over- and underperformance (Murphy et al., 2020). These two parameters will also be used in this thesis.

The paper concluded that high $\Delta REWS$ values cause higher power production. A gain of between 3% and 4% is seen as maximum power increase. This occurs between 3 m/s and 13 m/s. Vice versa, low $\Delta REWS$ values causes significant production losses (up to 2% of rated power). Regarding wind shear (α), there is no clear correlation between shear and power production. A counterintuitive result was obtained for wind speeds larger than 8 m/s: high shear resulted in lower power production. This can be the result of complicated wind profiles due to the complex terrain. The latter was also concluded by Vanderweijde and Lundquist Vanderwende and Lundquist (2012). The conclusions for directional veer were that positive veer values (veering) caused power production to increase, while negative veer (backing) caused a decrease in power production.

Magnitudes for the different parameters from this research were the following:

Table 2.9: Magnitude of investigated properties in Murphy et al. (2020)

	Unit	Min	Max	Mean
Shear factor	-	-0.2	1	0.15
Directional veer	degrees/m	-0.5	0.5	0
$\Delta REWS$	m/s	-2	2	0

In another research on the influence of site conditions on turbine performance by Gomez and Lundquist (Sanchez Gomez & Lundquist, 2020), directional veer and wind speed shear were investigated as well. It was found that under- and over performance was separated by a relation between shear and veer. Underneath this line, the turbine was predominantly over performing. In this research, over performance meant performing better than the mean power output for the concerning wind speed bin.

$$\beta = \frac{2}{3}\alpha - 0.1 \quad (2.29)$$

Where:

α = Shear factor

β = Veer

Power law values between 0 and 0.33 have shown in past research that this leads to underperformance (Antoniou et al., 2009) (Bardal et al., 2015). In the research of Gomez and Lundquist it was found that combinations of shear and veer below the correlation of equation 2.29, for the wind shear region between 0 and 0.33 the turbines seemed to over perform.

In accordance with the linear boundary from equation 2.29, The turbines investigated seemed to over perform for high shear values in combination with low directional veer, while under performing for high directional veer values in combination with low shear. Even for shear values higher than 0.7 the turbines were performing better than average performance.

Magnitudes for the different parameters from this research were the following:

Table 2.10: Magnitude of investigated properties in Sanchez Gomez and Lundquist (2020)

	Unit	Min	Max	Mean
Shear factor	-	-0.3	0.8	-
Directional veer	degrees/m	-0.1	0.6	-

In a research done by Stival et al. (2017), the influence of wind speed shear and TI was investigated. From this research, it was concluded that high TI values cause a turbine to under perform for higher wind speed regions, and over perform for moderate wind speed regions. TI reached high values (of more than 0.1) only for moderate wind speeds. Overall, it was concluded, high values of TI can cause a turbine to perform less than situations of low TI. Wind shear caused the investigated turbines to slightly under perform when compared to moderate shear circumstances. The latter was also concluded by Bardal, L. M. et al. (Bardal et al., 2015).

Magnitudes for the different parameters from this research were the following:

Table 2.11: Magnitude of investigated properties in Stival et al. (2017)

	Unit	Min	Max	Mean
TI	-	0	> 0.15	-
Shear factor	degrees/m	< 0	> 0.4	-

Finally, in a research done by Vahidzadeh and Markfort (Vahidzadeh & Markfort, 2019), power curve models based on high-frequency data including TI, wind shear and wind veer were compared to the standard 10 minutes averaged power curve. Results from the research were that all models including the above-mentioned parameters predicted the WTG power output better than the conventional power curve. Measurements used in this research were obtained at a flat site with trees and buildings. An overview of those measurements is given in the table below.

Table 2.12: Magnitude of investigated properties in Vahidzadeh and Markfort (2019)

	Unit	Min	Max	Mean
Shear factor	-	-0.1	1.1	0.3
Directional veer	m	0	24	-
TI	-	0	0.45	-

2.8. Turbine cluster analysis requirements

A consultant closely related to Vattenfall did a description of ways how to compare two neighbouring wind turbines regarding performance. In this section, useful analysis tools from this report are discussed. Parts of the analysis can be of use when setting up the methodology for this thesis.

According to this consulting firm, a number of parameters are needed for comparing two turbines within a turbine cluster. A list of those parameters is given below. Something that stands out is that no wind speed measurements are required within this method. This is an advantage as wind speed measurements at the turbine are affected by rotor movements.

Table 2.13: Parameters needed for turbine cluster analysis (Albers, 2014)

Parameter	Required
Active power	X
Nacelle direction	X
Status operational/idle	X
Rotor speed	X
Pitching angle	
Air temperature	
Wind speed	

Two steps as proposed by the consultant which are of potential relevance for this thesis are given below:

- **Wind directional filtering**
Highly manipulated by other turbines need to be excluded from a performance analysis. A method for excluding this contaminated data per turbine given by the consultant is based on the ratio between the wind speeds measured at the two compared turbines. This ratio is calculated for all wind direction sectors. The directional sectors at which this ratio crosses a certain threshold are deleted from the data sets (Albers, 2014).
- **Power-to-power relation analysis**
For different directional bins, the power of the two investigated turbines are plotted against each other. This can be done by dividing the total power range into different bins. By default, 5 bins can be generated. By creating these power-to-power relations for different wind speed bins, a first model is created for comparing the performance of the two turbines. This power-to-power relation does not contain any time-dependency and is therefore a good tool for comparison of two turbines which are not experiencing exactly the same winds. The goal of this analysis is to detect and highlight less performing timestamps of one turbine compared to another (Albers, 2014).

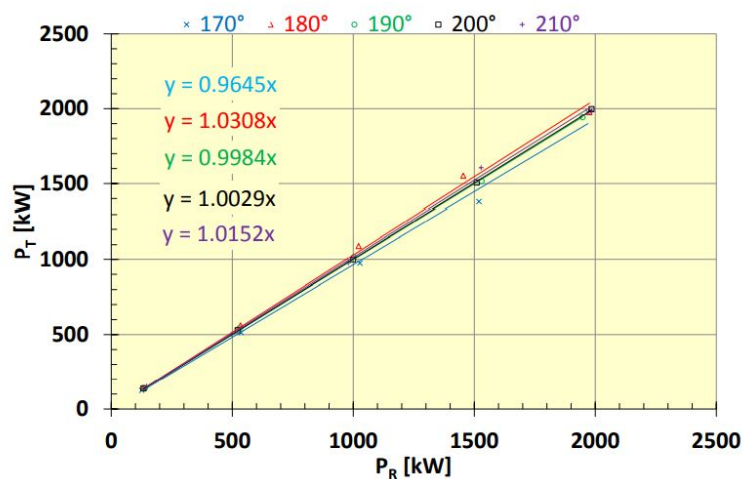


Figure 2.12: Example of a power-to-power visualisation of a reference turbine and a test turbine Albers (2014)








3

Methodology

In this chapter, the methodology of this research is explained. The different steps are elaborated on, and the different tools used are described. In the figures, the programming packages, tools and data formats used are displayed as symbols. The symbols and the corresponding packages/tools are given in the table below. This table also serves as a legend for the flow charts including these symbols.

Important parts of coding written for executing the set-up methodology, as described further on in this chapter, can be found in appendix F.

Table 3.1: Python packages and corresponding logos.

	Logo	Type
Spyder		Python package/tool
Jupyter Notebook		Python package/tool
Bokeh		Python package/tool
Datashader		Python package/tool
Pandas		Python package/tool
Excel (.csv)		File format
Pickle (.pkl)		File format

The aim of this methodology is gaining insights in the underperformance of the PYC wind farm. To work towards this aim, multiple processes are carried out. An overview of the total methodology of this research is given in figure 3.1 below. The multiple processes of the methodology are visualised as the white boxes. The blue parallelograms are representing data streams which is input/output to the different processes. Each of the processes is described in more detail in the sections below.

3.1. Workflow

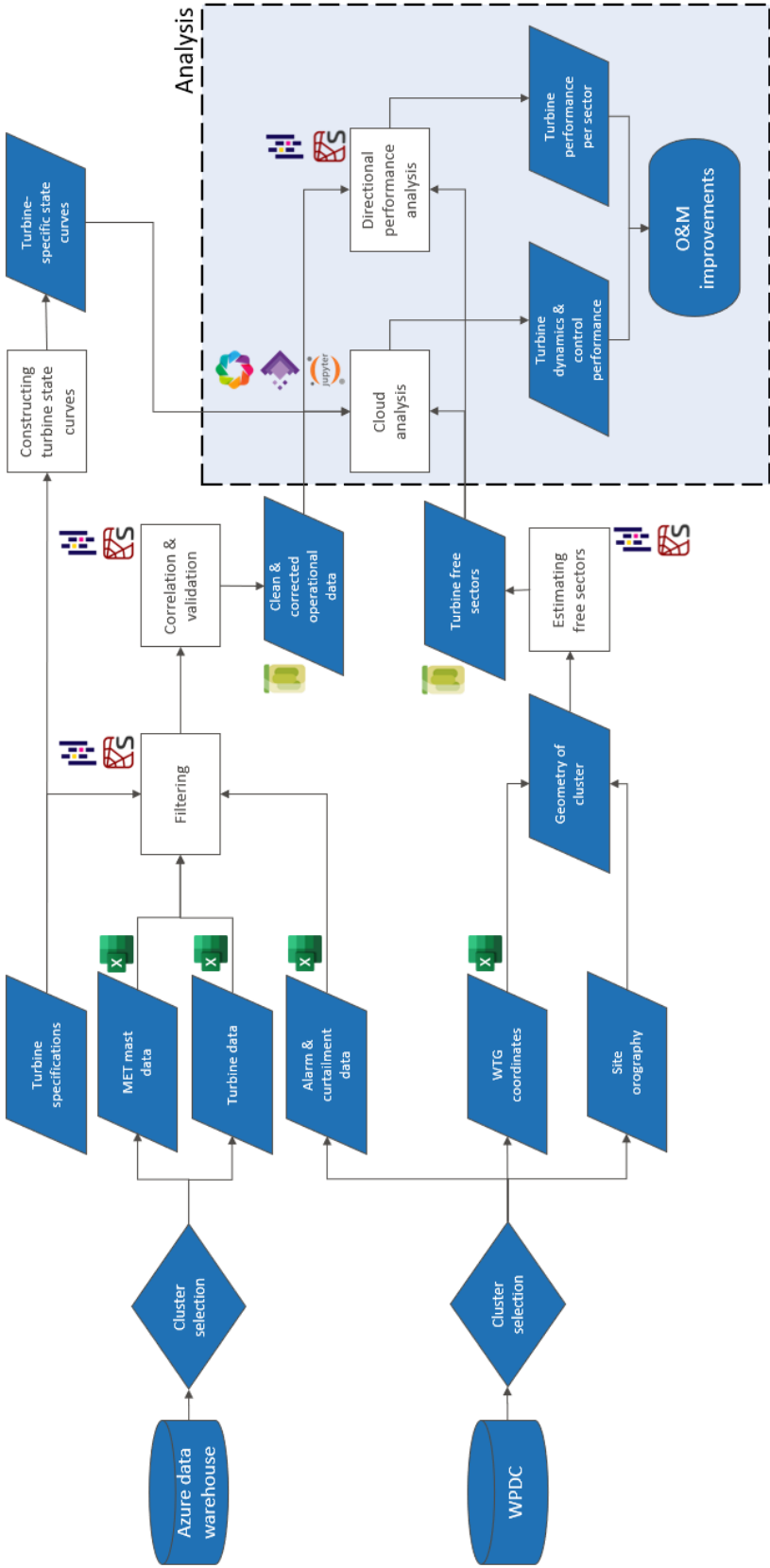


Figure 3.1: Schematic overview of the methodology workflow.

3.1.1. Filtering

Main data filtering

As can be seen in figure 3.1, the turbine data will be filtered to ensure all data used for analysis will be data without containing moments of malfunctioning. This is done by using alarm and curtailment signals from the turbines. The whole filtration process is described step-by-step below:

1. Loading in raw 10 minutes averaged data.
The raw 10 minutes averaged data is loaded in for the MET mast as well as for selected turbines. The files loaded in are significantly smaller than the HF data loaded in.
2. Loading in high-frequency data sets.
This data has a significantly higher resolution than the 10 minutes averaged data. Therefore, it takes more computational power and time to load in and manipulate this data. The HF data sets are the bottleneck of the filtering process.
3. Loading in curtailment and alarm data.
Part of the data filtering concerns filtering out timestamps where the turbine is curtailed, or where an alarm is logged. Curtailment data gives information about periods in time that a turbine or park's power output is reduced. As the power output is limited by a control mechanism during these periods, this data shall not be used for further analysis. Alarm data gives information about when possible faults on a turbine occur. These timestamps are discarded from the turbine data as well.
4. Reshaping 10 minute and HF data.
All csv data files (10 min and HF) are loaded into python as Pandas DataFrames. This data structure is convenient for data analysis, because of the wide possibilities of the Panda's API ('API reference — pandas 1.5.3 documentation' (n.d.)). Both 10 minutes averaged and HF data time columns are changed into pandas DateTime objects, to let python understand this column are dates. Moreover, the HF data set structure is changed to the format of time in the index, and the different SCADA tags as column names. By doing so, the HF data is structured in the same way as the 10 minutes averaged data.
5. Filtering data from alarms and curtailments.
Alarms and curtailments as loaded in at step 1 of the filtering processed are filtered out of the turbine data by filtering out timestamps that contain curtailments and/or alarms. For the HF data sets, a buffer of 30 minutes is applied on both sides of the filtered out timestamps to have a more complete curtailment filtering procedure¹.
6. Filtering timestamps without logged wind speeds.
Timestamps without a wind speed measurement logged are filtered out, as wind speed is a crucial parameter for the analyses to be carried out. This is step is more important for the HF data sets.
7. Filtering of icing periods.
Low temperatures (below and around zero degrees) can cause water on the measurement equipment to freeze. This freezing can disturb the measurements carried out by the equipment, by blocking movements from the equipment. As a result, during moments of freezing, the wind directional data can be stuck at a certain level, while the anemometer ² is not rotating and logging wind speed values of 0 m/s. When a combination of these two phenomena occurs, the timestamps are filtered out of the data.
8. Filtering of frozen logging.
Frozen logging is a phenomenon where all parameters are measured correctly, but are not logged properly. This results in non-changing data over time, over the whole range of parameters. When frozen logging is noticed, the timestamps are deleted from the data.
9. Velocity filtering for MET mast and turbine data.
Velocity filtering is done to make sure the data sets only contain operational data. Data below cut-in wind speed and above cut-out wind speed is therefore filtered out.

¹The 30-minutes buffer has been chosen after an iterative process of applying buffers to the raw data sets. The effect of the applied buffer is used for the decision for increasing/decreasing the buffer size.

²Wind speed meter.

10. Calculating yaw misalignment.
The difference between wind direction and nacelle direction can be useful for showing misalignment between nacelle and wind direction. It is calculated by subtracting both tags from each other, and adding the results in an additional column.
11. TI calculations.
HF data can be used to calculate Turbulence Intensity. The standard deviation of the HF signal of each 10-minute period is divided by the 10 minutes averaged wind speed of that period. The calculated turbulence intensities are saved as a new column within the 10 minutes averaged data sets.
12. Torque calculations.
Torque is a measure of the loads on a turbine. It is a parameter that can be easily computed by dividing the active power output of the turbine by the rotational speed. This can be done for both the 10 minutes averaged and HF data sets.
13. Veer calculations (MET mast).
Wind directional veer ³ can only be calculated at the MET mast. At the MET mast, wind directions and wind speeds are measured at multiple heights. For the veer calculation, the difference between the hub height and lowest tip height wind direction is calculated, and divided by the elevation difference between the two levels (0.5D). This gives the wind directional veer in degrees per meter (see equation 2.10).
14. Shear calculations (MET mast).
By using the same measurement levels as for the veer calculation, wind shear can be calculated at the MET mast. This is done by using equation 2.9.
15. Wind speed air density correction.
To compare velocity measurements at the site with wind speed-included data provided by the manufacturer (e.g. power curve data), the downloaded wind speed data needs to be corrected towards the same air density as used in the technical turbine documentation of the manufacturer (1.183 kg/m³). For correcting the wind speeds, the location-specific air densities are needed (equation 2.24). To calculate those, the average temperature and elevation at the location are needed⁴. The air density corrected wind speeds are saved in an additional column for both the 10 minutes averaged and HF data sets.
16. Saving DataFrames.
After the 10 minutes averaged and HF DataFrames have been filtered by all steps above, they are saved as csv files again in the right cluster folder. The above process resulted in data sets with additional columns (torque, shear, veer, TI), and fewer rows (as a result of the filtering process).

A visualisation of the filtering and calculation steps are shown in the figure below. The left column contains steps performed on turbine data, and the right contains steps performed on the MET mast. The blue and grey rows represent filtering steps, and calculations performed on the data, respectively. Because of wind directional and wind speed data being available for multiple altitudes, directional veer and wind shear could be calculated with the MET mast data.

³Wind direction changes over a vertical distance

⁴It would be more accurate to correct the data for every 10 minutes. For the scope of this work, enough accuracy is obtained when correcting the data as a whole with the average yearly temperature.



Figure 3.2: Filtering process for turbine data (left) and MET mast data (right)

Further details and pieces of code of the filtering process can be found in appendix F. For the purpose of this thesis and for other similar cloud analyses, the layout of the filtering process has been changed. The filtering code is applicable for any string containing of a MET mast and multiple turbines to be analysed. Besides the layout, also the way of coding is changed in certain functions, to speed up the code. A function for calculating the turbulence intensity (TI) was re-coded with the `@njit` package. This package makes it possible to do simple calculations more efficiently than standard Python. As a result, it is possible to run the filtering script for a string of turbines combined with a MET mast, within a time period of 1-2 hours.

Post-filtering

After the collected data sets have been filtered according to the steps of figure 3.2, the data sets are further processed before starting analysing the data. This post-processing does not include deleting rows of data as is done in the filtering process, but is mainly meant for manipulating the data to make it reliable and easy-to-use. Below, the different steps from this post-processing are described. Moreover, in figure 3.3 an overview of this process is given.

1. Renaming data.

The data loaded in is named the same as the tags visible in section 5.1.1. To make the data usable for analysis, more appropriate names are chosen for data selection and plotting more downstream in the process. The turbine-specific first part of the names is deleted, and names representing the parameter of that specific tag are chosen.

2. Creating raw data set.

In the analysis, deviating behaviour of one turbine compared to the other turbines in that string is selected. When further analysing a cloud, it can be valuable to understand how that cloud of points would behave in the not-processed raw data of that turbine. Therefore, of the turbine closest to the MET mast, a raw data set is loaded in and saved. The tags of this data set are also renamed, as described in the previous step.
3. Resampling HF data for correlations.

10 minutes averaged data is used for correlating the data sets of two turbines (or a MET mast and a turbine). When a correlation with a reliable data set is strong, the other 10 minutes averaged data set of the correlation can be trusted and analysed. To be able to trust the HF data sets of the different turbines as well, the HF data per turbine is resampled to a sampling rate of once every 10 minutes, and correlated to the corresponding 10 minutes averaged data set. It is expected that this correlation is very strong, as the definition of 10 minute data is averaged higher resolution data. When these correlations do not show surprising results, the HF data can be trusted if their corresponding 10 minutes averaged data sets can be trusted.
4. Create data sets for comparisons.

When correlating the data of two neighbouring turbines, it is convenient to work with one data set including the data of the two turbines. For this purpose, data sets of two neighbouring turbines are combined.
5. Correlating 10 min and HF data.

The resampled HF data of a turbine is correlated against the corresponding 10 minutes averaged data set of that turbine. Expected is that correlations between the two data sets are very strong, as explained above.
6. Correlating wind speed, wind direction and nacelle direction.

After renaming and combining data sets, these are used for correlations. Correlations for three parameters are checked: wind speed, wind direction and nacelle direction. When correlations turn out reliable (a slope between 0.95 and 1.05), the tested data set can be trusted and used for the analysis. Nacelle directions are correlated, as the wind direction parameter is only available for the last third of 2020. Nacelle direction measurements can be used as an alternative.
7. Correlating active power output.

Active power output data of two neighbouring turbines is checked to confirm turbines are not weirdly aligned. In data, an offset in directional data of one turbine can be visible with respect to a neighbouring turbine. This can be the results of misalignment between turbines. To check if this offset is caused by wrong data logging or by actual misalignment, active power outputs are correlated. When the two active power signals follow each other, it is concluded the offset is caused by wrong data logging.
8. Wind direction / nacelle direction offset correction.

When a constant wind direction and/or nacelle direction offset is observed, this means that for every wind/nacelle direction, there is the same offset between the directional measurements of neighbouring turbines. This is no natural behaviour, and is caused by wrong data logging. Therefore, the data is corrected for the found offset. This is done for a found offset of greater than 10 degrees, because a smaller offset can also be caused by general data scattering. These offsets should not be corrected.
9. Jump correction.

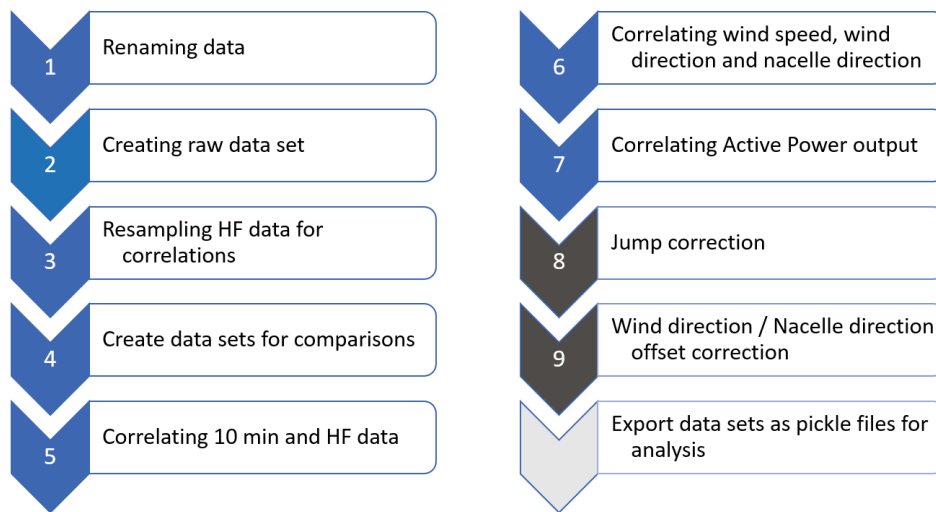
Directional data measured on the nacelle of a turbine can show shifts in a specific moment in time. At this moment, the directional data undergoes a jump. From this jump onwards, data is logged with an offset compared to the data before the jump. This can be caused by for example frost to the wind vanes at the turbines or MET mast. A detection code was written for investigating the occurrence of these internal jumps, as well as the moment in time. When detected, data from this timestamp on is counter corrected to match the average direction before the timestamp. In appendix C, an example of the output of the code for the correction of an internal jump is given. In the table below, the two correction mechanisms are summarised:

Table 3.2: Possible corrections carried out to data sets after filtering.

Step	Operation	Involved tags	Applied when
8.	Correct for directional offset compared to neighbouring turbine	Nacelle direction Wind direction	Offset > 10°
9.	Correct for sudden offset change in time series (internally)	Nacelle direction Wind direction	Detected by code

10. Export data sets as pickle files for analysis.

The data was loaded in as csv files. To make the data easily importable for further analysis, the data sets are saved at pickle files. Pickle files save the data type and structure of Python objects. When loading in pickle files into new Python projects, the data type and structure is imported correctly and no manipulation has to be done to the imported data before being able to use it.

**Figure 3.3:** Post-processing steps before analysis.**3.1.2. Free sectors estimation**

Besides data filtering, free sectors are calculated prior to the analysis. As the scope of this research is the investigation of unwaked sectors, the waked sectors shall be excluded from the data prior to analysis. To do this, latitude and longitude coordinates of the PYC turbines are collected from WWP. Moreover, coordinates of neighbouring smaller farms (also consisting of 3 MW turbines) are estimated by the use of Google Earth ('Google Earth', n.d.). Moreover, of all turbines the rotor diameters are collected, as these are of importance in the selections of turbines to include for the free sector estimates.

A turbine is considered as influential to another turbine when closer than 20D to the turbine in question. Wake expansion is modelled by a relation as provided by IEC. This relation is described in equation 2.16. When a turbine is closer than the threshold of 20D, the distance from the turbine to the investigated turbine is inserted into equation 2.16, together with the rotor diameter of the turbine. This gives the sector which is influenced. When subtracting all disturbed sectors for all turbines closer than 20D, what is left are the unwaked sectors for a specific turbine. This procedure is repeated for each of the 14 investigated turbines. Results are shown in chapter 4.

3.1.3. Data correlation and validation

Correlating measurement data of the MET mast and neighbouring turbines is of great importance for carrying out an analysis using data measured at the turbines. When one wants to use the measurements further up the turbine cluster, these should be calibrated with measurements more upstream of the turbine cluster analysed. When a reliable correlation can be found between a MET mast and a turbine (or a turbine and a turbine), measurements of the tested turbine can be used for comparisons of the specific turbine. A correlation is tested on slope (α) and the fit to a first order polynomial (R^2). A correlation is considered reliable if it meets the following two conditions:

1. $R^2 > 0.7$

A coefficient of determination (R^2) of greater than 0.7 is typically regarded as being sufficient for data modelling and forecasting in terms of regression analysis.

2. $0.95\alpha_{expected} < \alpha_{expected} < 1.05\alpha_{expected}$

The expected slope α depends on the height differences between the two correlated turbines. When heights are comparable, the expected slope of the linear fit to the wind speed data is 1. A 5% deviation is accepted regarding wind speed correlations.

The procedure and sequence of the correlations is shown in the figure below. Important note: correlations are checked with unwaked data, to make sure wakes are not disturbing correlation results.

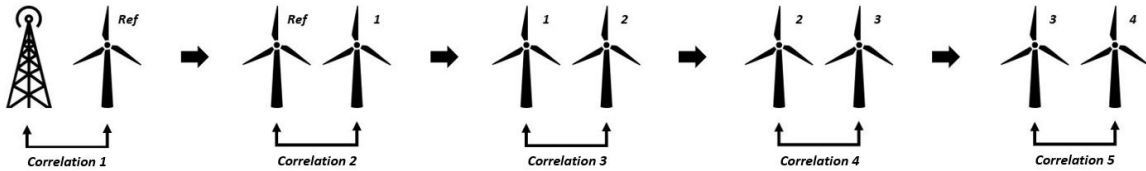


Figure 3.4: Description of correlation checks.

3.1.4. Constructing turbine state curves

Within this analysis, groups of turbines are investigated with relatively small inter-turbine distances. Because of this, behaviour between turbines can be compared by looking at the difference in state curve relations between the different cluster turbines (state curve relations are explained in section 2.6). To make the analysis not only reliable on measured behaviour, theoretical state curves are constructed for both of the turbines existing at PYC wind farm. These state curve relations are based on turbine properties. Below, the procedure for the estimation of those curves is discussed.

Power curve (Wind speed - Active power relation)

The theoretical power curve consists is a non-rated and a rated part. The non-rated part follows the power relation of equation 2.21. Input tot this relation are the rotor swept area, air density (reference value used by the manufacturer and power coefficient. The power coefficient was calculated from a point on the WPC in the constant- λ part of the power curve: in this part, the power coefficient is theoretically constant and at its maximum. The power coefficient can then be calculated accordingly:

$$c_P = \frac{P}{\frac{1}{2}\rho U^3 A} \quad (3.1)$$

At the point where the turbine reaches nominal power, the partial load region switches to a full load region with a constant power output of 3 MW.

The above described theoretical state curve is not that relevant: as a WPC is provided by the manufacturer, it is more relevant to use this power curve as reference. For completeness, the theoretical power curve as described above is still included in the state curve comparisons within chapters 6, 7 and 8.

Rotor speed - Wind speed relation

The rotor speed - wind speed relation based on two principles. Up and until U_{nrtd} , the rotor speed is linearly related to wind speed by the tip speed ratio λ (equation 2.20). From U_{nrtd} onwards, the rotor

speed is constant at its maximum value. Both U_{nrtd} and λ are extracted from the wind speed - rotor speed relation itself, as no information about these parameters is known for the PYC turbines. The λ is also input for the torque curve.

Torque curve (Rotor speed - Torque relation)

The torque curve consists of three regions, as explained by 2.8: In theory, one would expect sharp regions between the three different regions. Regions A-B and C-D are described as vertical segments within state curves. Especially the quadratic part of the torque curve tells a lot about optimal turbine performance. This relation is directly derived from the optimal power curve of the turbine (equation 2.21). Inputs to this equation are site air density, power coefficient, design tip speed ratio and rotor radius. The design tip speed ratio is estimated from the slope of the wind speed-rotor speed data (in the partial load region).

Pitching - Torque relation

The theoretical pitching - torque behaviour is based on three basic principles:

1. At zero torque, a wind turbine is not producing power. This happens when the turbine is in idling conditions, or the turbine being pitched to get the turbine up to speed. In short, pitching can range from 0 until 90 degrees (θ_{off}).
2. Between 0 and Q_{max} , the turbine blades are not pitched in the ideal case. The control is within the optimal- c_p region. Torque control is used to keep the turbine at optimal- c_p . For this turbine, Q_{max} is calculated as follows:

$$Q_{max} = P_{rated}/(\omega_{max} * 2\pi/60) \quad (3.2)$$

3. At Q_{max} , the turbine is producing at nominal power. The turbine blades are pitch controlled up to values of θ_1 to keep the turbine at maximum rotor speed. For this turbine, θ_1 is 23 degrees.

Rotor speed - Pitching relation

The rotor speed - pitching relation is based on three rotor speed regions:

1. Before cut-in rotor speed, the pitching is controlled to 1 (or in the case of PYC 2) pitching values. This is done to get the rotor up to speed for production from cut-in wind speed onwards. In the case of PYC, these two values are 10 and 30 degrees.
2. Between cut-in and rated rotor speed, no pitching is expected to be seen: within this region, the turbine is partial load, where the turbine is torque-controlled to operate at optimal- c_p
3. At rated rotor speed, the turbine is experiencing wind speeds between rated and cut-out wind speed. Therefore, pitching values are expected between 0 and the pitching angle for full load control just below cut-out wind speed. For the PYC turbines, this pitching angle is around 23 degrees. Unfortunately, there is no pitching data available from the turbines. Therefore, this maximum pitching value can not be compared to theoretical behaviour.

Pitching curve (Wind speed - Pitching relation)

As discussed above, there is no expected theoretical pitching behaviour to be discussed, as there is no pitching information available from the Siemens turbines at the site. Although there is no theoretical reference for this relation, the pitching curve is still a relation that is of value for this thesis performance analysis.

The above described curves are described mathematically in table 3.3.

3.1.5. Overall & Directional performance analysis

One of the two analyses carried out is the overall & directional performance analysis. This analysis is carried out in parallel to the cloud analysis. The aim of the analysis is to gain insights in the variation of turbine performance over wind directions, as well as in the total performance figures.

For both the overall and the directional turbine performance analyses, 10 minutes averaged data is used. Another important note is that only air density corrected wind speeds are used, as these wind speeds are corresponding to the environmental conditions of the warranted power curve as described in section 3.1.1.

The performance estimates made in this analysis are based on power curves. Active power output (kW) is used for this purpose, as production [kWh] estimates are made for the long term. Moreover, active power is wind speed dependent and therefore a good input for potential power. Therefore, to get an estimation of wind turbine performance, the binned operational power curves of the year 2020 are compared to the warranted power curve as provided by the manufacturer. Binning of the active power data is done by means of MOB, where the active power signals within each wind speed bin are averaged. Moreover, the standard deviation for each bin is calculated. When appending these averaged active power values for every bin, the operational power curves can be constructed.

Overall performance calculations

For the overall performance calculations, two methods are used. The first performance calculation is based on comparing these binned operational power curves (OPC) to the warranted power curves by using equation 3.3.

$$\eta_{power,1} = \frac{OPC \cdot NWD * 24 * 366}{WPC \cdot NWD * 24 * 366} \quad (3.3)$$

Where:

OPC = Operational power curve
 WPC = Warranted power curve
 NWD = Nominal wind distribution

The second method used to calculate overall performance of the turbines is by using the potential power signals from the turbines. Potential power signals are logged with the same sampling time as the active power signals (10 min). The potential power signal resembles the potential power output for a certain moment in time. The signal is based on neighbouring turbine data, as well as nacelle wind speed measured at the turbine. Results from both calculations are shown in table 6.4 ⁵.

$$\eta_{power,2} = \frac{(\sum_{t=1}^n P_{active,t}) * \frac{1}{6}}{(\sum_{t=1}^n P_{potential,t}) * \frac{1}{6}} \quad (3.4)$$

Where:

$P_{active,t}$ = Active power signal on timestamp t
 $P_{potential,t}$ = Potential power signal on timestamp t

The output of both calculations is a fractional performance number between 0 and 1.

Directional performance calculations

To understand the turbine performance per wind direction sector, the data is divided into 18 equal 20 degrees wind direction sectors. For each sector, the active power data is then again binned and averaged as described in the section above.

⁵It is expected that both estimates of performance are close to each other, as the calculation in the denominator of equation 3.3 is one of the estimates used for calculating the potential power signal.

With the operational power curves (OPC) of the turbines, the performance per directional bin can be calculated by comparing the weighed OPC to the weighed WPC. The power curves are weighed by the normal wind speed distribution at the specific cluster.⁶

$$\eta_i = \left(\frac{OPC_i \cdot NWD_i}{WPC \cdot NWD_i} * 100 - 100 \right) \quad (3.5)$$

Where:

- η_i = Percentage of power gain/loss compared to total power output for sector i
- OPC_i = Operational power curve for sector i
- NWD_i = Normal wind speed distribution (over wind directions) for sector i

When calculating performance gains and/or losses compared to total power output of a turbine, an option is to take into account the relative power production for each specific wind direction. This relative production is multiplied by the ratio of operational power curve and warranted power curve, for which the wind speed distribution is accounted for. This gives the loss in a specific sector relative to the total production of a turbine. Within this research, the focus will be on equation 3.5 instead of equation 3.6, as equation 3.5 gives more information about the unweighted over-or underperformance of a wind direction sector.

$$\eta_{i,rel} = \eta_i * P_{rel,i} \quad (3.6)$$

Where:

- $\eta_{i,tot}$ = Percentage of power gain/loss compared to total power output of turbine
- η_i = Percentage of power gain/loss compared to total power output for sector i
- $P_{rel,i}$ = Fractional power production of sector i

$P_{rel,i}$ is calculated as follows:

$$P_{rel,i} = \frac{P_i}{\sum_{i=1}^n P_i} \quad (3.7)$$

Where:

- P_i = Power production of sector i

The output of equation 3.5 is a performance gain/loss percentage compared to WPC. This is done for visualising purposes. Figures showing an overview of performance for every wind sector per turbine are given in the results chapters.

Site influence

For sectors with a negative mismatch compared to the WPC of bigger than 5%, the site conditions are further investigated. The descriptions of the site for these worst performing sectors give an insight in the influence of the site to the performance of the selected turbines.

⁶Important note at equation 3.5: in the numerator of the equation, the cross product of (pre-calculated) operational power curve and nominal wind distribution is defined. Using the sum of active power signals in the numerator will lead to the same efficiency outcome, as the OPC is calculated from those same active power signals.

3.1.6. Cloud analysis

After filtering and correlating, the data sets (10 minutes averaged and HF) are input to the cloud analysis to investigate turbine behaviour⁷. Moreover, the theoretical state curves as described in figure 2.11 are inputs. For all six state curve relations, the turbine data is plotted against the theoretical expected turbine state curve behaviour of section 3.1.4. From these state curve relations, clouds can be selected to further analyse and understand deviating turbine behaviour. This is done for HF data as well as for 10 minutes averaged data.

10 minutes averaged data clouds are selected for one of two reasons:




1. **Turbine behaviour deviating from theoretical turbine state curves**
When 10 minutes averaged or HF scattering is differing from the turbine-specific state curves as described in section 3.1.4, this is reason for analysing this cloud in more detail.
2. **Deviating turbine state curve behaviour compared to other turbines from the cluster**
When the 10 minutes averaged data of one turbine does not match the behaviour of the other turbines within that cluster, this can cause differences in performance between that turbine and the others. Therefore, it is worth analysing that particular deviating data cloud in more detail.

HF data clouds are selected for the following two motives:

1. **Differences in HF turbine behaviour compared to 10 minutes averaged turbine behaviour**
Apart from wider scattering, more differences can be put to the light in HF data compared to 10 minutes averaged data. Understanding these differences can be done by selecting HF clouds.
2. **Deviating turbine state curve behaviour compared to other turbines from the cluster**
When the HF data of one turbine does not match the behaviour of the other turbines within that cluster, this can cause differences in performance between that turbine and the others. Therefore, it is worth analysing that particular deviating data cloud in more detail.

Important note on HF cloud selection: Scattering within HF relations can be of two types. The first type is caused by wind speed up- and down ramping. The turbine control system does not respond directly to these fluctuations in wind speed. This delay in response is visible in HF data sets, because of the short sampling time of measurements. The second type of scattering is caused by actual turbine behaviour, deviating from the expected behaviour. Understanding which type of scattering is observed is important in judging the significance of selected clouds.

The selection of clouds from the state curve relations can be done in three ways using the Bokeh interface. More about the functionalities of Bokeh is given in section 3.2.

-  **BoxSelectTool**
This tool makes it possible to select an area from a plot by a left and right boundary for the x-axis, as well as for the y-axis. A square selection is the result ('Plot tools', n.d.).
-  **PolySelectTool**
By using this selected, it is possible to create a polygon with an arbitrary number of sides, depending on the number of corners selected. This selection tool is slower than the BoxSelectTool, but still usable for HF data sets ('Plot tools', n.d.).
-  **LassoSelectTool**
The lasso selection tool brings the opportunity to create a custom shape to select. It collects the coordinates of the points the cursor moves when selecting the cloud. The collection of these points provide the boundaries of the selected cloud. This is the cloud selection method demanding most computational power, as the created selection criterium is most complex. It is too computationally slow for HF data sets ('Plot tools', n.d.).

After selecting a data cloud from a scatter plot, the following (interactive) outputs are generated by the tool. Every output contains the full turbine data set in grey, with the selected cloud data highlighted in blue.

⁷Parallel to the overall & directional performance analysis.

1. Histograms of: wind speed, nacelle direction, active power, pitching, rotor speed, generator speed, torque, temperature, yaw misalignment and TI.
2. Turbine data shown within the six state curve relations of figure 2.11. For every relation, the turbine data is plotted together with the expected theoretical behaviour (section 3.1.4).
3. Yaw misalignment and TI plots.
4. Time series of: wind speed, nacelle direction, active power, pitching, rotor speed, generator speed, torque, temperature, yaw misalignment and TI.
5. Performance data of the selected cloud including the following parameters:
 - Cloud number of points [-]
 - Cloud share in total number of data set points [%]
 - Cloud energy production [kWh]
 - Cloud share in annual energy production [%]
 - Cloud lost potential power (LPP) [kWh]
 - Lost potential power share in annual energy production (LPPT) [%]
 - Lost potential power share in cloud energy production (LPPC) [%]

The potential energy loss of a selected cloud is calculated as follows:

$$E_{pot,loss} = (WPC \cdot NWD_{cloud}) * t_{cloud} - E_{out,cloud} \quad (3.8)$$

Where:

$E_{pot,loss}$ = Cloud potential energy lost
 t_{cloud} = Total summed up duration of cloud
 $E_{out,cloud}$ = Cloud energy production
 NWD_{cloud} = Cloud normal wind speed distribution

Besides these cloud performance calculations, the cloud is saved after selection. By saving a cloud, the data is stored in an Excel file, collecting all clouds selected within the analysis. Moreover, a pickle file of the cloud is saved. This pickle file can be loaded into the python tool in a later stage for further analysis of the cloud. This gives the tool better reproducing properties. In appendix E, an example of the outputs of the cloud analysis tool is given.

To create structure when selecting multiple clouds, a unique tag marks each selected cloud. This tagging system gives structure when labelling and discussing visualisations of clouds. The tagging can be comprehended as follows: [turbine]-[data type]-[tag ID]. The first part of the cloud tag is the code of the turbine from which the cloud is selected. The data type is either 10 or HF⁸. The tag ID is a unique three-integer code. The data type '10d' resembles the use of raw, unfiltered data. This unfiltered data is used to clarify clouds selected from the filtered data.

For this cloud analysis, Datashader, HoloViews and Bokeh were combined to create a powerful high-frequency (and 10 minutes averaged) data analysis tool in Jupyter Notebook. More information about the advantages and applications of those packaged is given in the section below. The tool makes fast visualisation possible, and produces interactive plots in which data clouds can be selected and further investigated.

A tabular overview of the analyses carried out is represented below. This table includes the expected characteristics and behaviour per analysis ⁹. These characteristics are based on the turbine control theory discussed in section 2.6.

⁸resembling 10 minutes averaged and high-frequency data, respectively.

⁹Note that this expected turbine behaviour only accounts for normal operational conditions, which is the scope of this thesis

Analysis	X	Y	Frequency	Unit(s)	Expected characteristics
Active power	Hub height wind speed	Active power	10 min, HF	All turbines	$U < U_{rated} : P \sim U^3$ $U > U_{rated} : P = P_{rated}$
Torque	Rotor speed	Torque	10 min, HF	All turbines	$\omega_{rotor} < \omega_{rotor,max} : Q \sim \omega_{rotor}^2$ $\omega_{rotor} = \omega_{rotor,max} : Q \rightarrow Q_{rated}$
Pitching	Hub height wind speed	Pitching	10 min, HF	All turbines	$U < U_{cut-in} : 0 < \theta < \theta_{idling,max}$ $U_{cut-in} < U < U_{rated} : \theta = 0$ $U > U_{rated} : 0 < \theta < \theta_{full\ load,max}$
	Rotor speed	Pitching	10 min, HF	All turbines	$\omega_{rotor} < \omega_{rotor,max} : \theta = 0$ $\omega_{rotor} = \omega_{rotor,max} : 0 < \theta < \theta_{full\ load,max}$
Pitching/Torque	Pitching	Torque	10 min, HF	All turbines	$\theta = 0 : Q \in [0, Q_{rated}]$ $\theta > 0 : Q = 0 Q_{rated}$
Rotor speed	Hub height wind speed	Rotor speed	10 min, HF	All turbines	$U < U_{rated} : \omega_{rotor} \sim U$ $U > U_{rated} : \omega_{rotor} = \omega_{rotor,max}$
TI	Hub height wind speed	TI	10 min	All turbines	$TI \sim U^{-1}$
	Wind direction	TI	10 min	All turbines	-
Yawing	Wind direction	Yaw misalignment	10 min	All turbines	-
Veer/Pitching	Directional veer	Pitching	10 min, HF	MET, closest turbine	-

Table 3.3: Tabular overview of cloud turbine analyses.

The results from the analyses are discussed in chapters 6, 7 and 8. The expected theoretical behaviour for each of the state curve relations described in section 3.1.4 and table 3.3 guides as a reference when analysing the scatter plots of the different state curve relations.

3.2. Software and tools

In this section, the software used for this research is discussed. The main applications of the used software are processing the obtained raw data from the Azure Data Warehouse (ADW), as well as performing multiple analyses on the processed data. Finally, the tool is described that was built with the described software and python libraries.

3.2.1. Used software

For the purposes as mentioned above, the programming language Python[®] is used within two Python-compliant interfaces: Spyder and Jupyter Notebook. Spyder is used for heavy computational code, such as the filtering of high-frequency data sets. These specific pieces of code include lots of parameters which change continuously throughout the code. The Spyder interface is suited for this purpose, as it includes a window with an overview of all parameters, as well as a separate console window, in which outputs of the code are displayed.

For a part of the cloud analyses, which consists of dynamic programming and visualisation, Jupyter Notebook is used. This interface gives fewer options compared to the Spyder interface. However, it is suitable for dynamic visualisations, which are a key part of the cloud analysis. For these visualisations, a combination of Python libraries Bokeh and Datashader is used. Both are part of the HoloViews library. Bokeh is a library making dynamic visualisations possible, whereas Datashader is the right tool for big data visualisations.

The HoloViews library is built to make data processing as well as visualising easier and faster, by combining different python libraries effectively. It makes raw data and its visualization equally accessible at all times. The essence of HoloViews is to describe your data within a small amount of information, and to give minimal additional information to customise visualisations. Besides the coding efficiency, HoloViews makes combining different python packages possible. One of those combinations is Bokeh together with Datashader, as already introduced above.

As high-frequency data sets consist of (way) more information compared to 10 minutes averaged data, it can be a challenge to extract all useful information from visualisations. Another challenge is to do this effectively. Both challenges are solved by Bokeh and Datashader respectively. The dynamic interface Bokeh makes it possible to look at data sets from different perspectives and at different scales. These dynamic properties are particularly useful when analysing data sets of the size used within this research.

In figure 3.5, the Datashader algorithm is described. Datashader rasterizes data before plotting. After a raster is created, the data to be visualised is divided into the pre-defined raster. The amount of points within each area of the raster is represented by a shade ¹⁰. The fact that Datashader produces plots with a resolution of a pre-defined raster while not plotting every single data point separately, makes plotting more efficient and easy to work with.

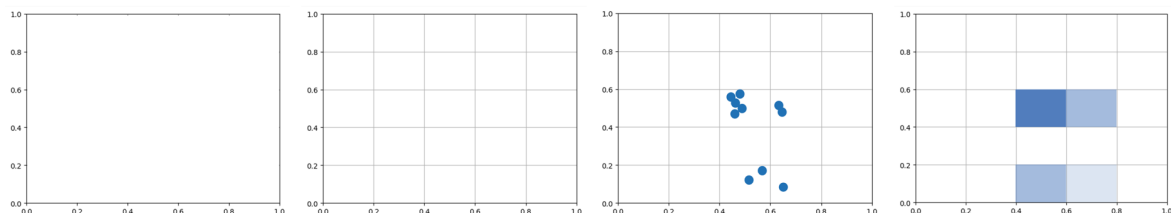


Figure 3.5: Datashader visualisation algorithm

Essential pieces of code used within this research are listed in appendix F.

¹⁰The shade pattern can be created manually, or can be imported from the HoloViews database.

4

Site analysis

For the cloud analysis that will be carried out as described above, multiple turbines are selected from the total wind farm (consisting of 76 WTGs). The choice for the turbines is based on their location within the farm and site conditions. In this chapter, the turbine type, as well as the site conditions and turbine selection will be covered.

4.1. Turbines at the site

The Pen Y Cymoedd wind farm has two different turbine types installed. Both are Siemens turbines, with a different rotor diameter. Both turbines have a rated power of 3 MW. The properties of both turbines are tabulated below. These properties will be of use when analysing the data.

Power	
Rated power	3 MW
Cut-in wind speed	4 m/s
Rated wind speed	12.5 m/s
Cut-out wind speed	25 m/s
Rotor	
Diameter	113 m
Swept area	10000 m ²
Max. rotor speed	15.5 RPM
Max. tip speed	92 m/s
Gear box	
Type	Direct drive
Generator	
Voltage	690 V
Frequency	50/60 Hz

Power	
Rated power	3 MW
Cut-in wind speed	3 m/s
Rated wind speed	12 m/s
Cut-out wind speed	25 m/s
Rotor	
Diameter	108 m
Swept area	9144 m ²
Max. rotor speed	16 RPM
Max. tip speed	90 m/s
Gear box	
Type	Direct drive
Generator	
Voltage	690 V
Frequency	50/60 Hz

Table 4.1: Properties of respectively SWT-3.0-113 and SWT-3.0-108

Moreover, site specific power curves were provided by the manufacturer. These power curves show how the turbines should perform regarding the conditions of the site where the turbines are placed. These curves are used as a reference for turbine performance. For all clusters, the site specific power curves are different as the site conditions at both clusters are unique. Another reason for this difference is the placement of SWT-3.0-113 at the western cluster, and SWT-3.0-108 at the eastern cluster. Both turbines have different properties regarding power output over the wind speed spectrum. For confidentiality reasons, the site specific power curves are not shown within this thesis. However, they were used as input to the performance modelling.

The design boundary conditions applicable on turbines SWT-3.0-113 and SWT-3.0-113 respectively are shown in the two tables below. When parameters are outside the design values of the table, application

of the turbines might be possible, but additional information about the site is needed to draw conclusions on normal operation of the turbines.

Table 4.2: Design boundary conditions turbine SWT-3.0-113

Subject	Issue	Unit	Value
Wind, operation	IEC class	-	IIB
	Mean wind speed	m/s	8.5
	Weibull scale parameter (A)	m/s	9.6
	Weibull shape parameter (k)	-	2
	Wind shear exponent	-	0.20
	TI at 15 m/s	-	0.14
	Standard deviation of wind direction	Deg	7.5
	Maximum flow inclination	Deg	8
Wind, extreme	Maximum hub height wind speed	m/s	42.5
	Maximum 3 s hub height gust wind speed	m/s	59.5
	Maximum hub height power law index	-	0.11
Trees	If within 500 m, maximum height of $1/3(H - D/2)$		

Table 4.3: Design boundary conditions turbine SWT-3.0-108

Subject	Issue	Unit	Value
Wind, operation	IEC class	-	1A
	Mean wind speed	m/s	10
	Weibull scale parameter (A)	m/s	11.3
	Weibull shape parameter (k)	-	2
	Wind shear exponent	-	0.20
	TI at 15 m/s	-	0.16
	Standard deviation of wind direction	Deg	7.5
	Maximum flow inclination	Deg	8
Wind, extreme	Maximum hub height wind speed	m/s	50
	Maximum 3 s hub height gust wind speed	m/s	70
	Maximum hub height power law index	-	0.11
Trees	If within 500 m, maximum height of $1/3(H - D/2)$		

Besides the turbine properties given in table 4.1, there is one turbine specific control feature to be discussed. The turbine control is equipped with a control algorithm, making it possible to operate at higher wind speeds. This is done by pitching the turbine blades for power reduction, and operating the turbine at a decreased rotor speed. This mechanism has influence on the power curve, as for wind speeds higher than 20 m/s a deviation from rated power should be observed, as well as more extreme pitching and lower rotor speeds.

4.2. Site conditions and turbine selection

One important reason for the selection of all clusters is their position compared to the MET mast. What can be said for all clusters, is that there is one turbine relatively close to the MET mast. The other chosen turbines are relatively close to the reference turbine. As the turbines from the selected cluster are relatively close to each other, the turbine behaviour and performance can be better compared.

Below, an overview of the full site of PYC is given. The three selected turbine clusters are encircled. An individual description of each cluster is given in sections 4.2.1 until 4.2.3.

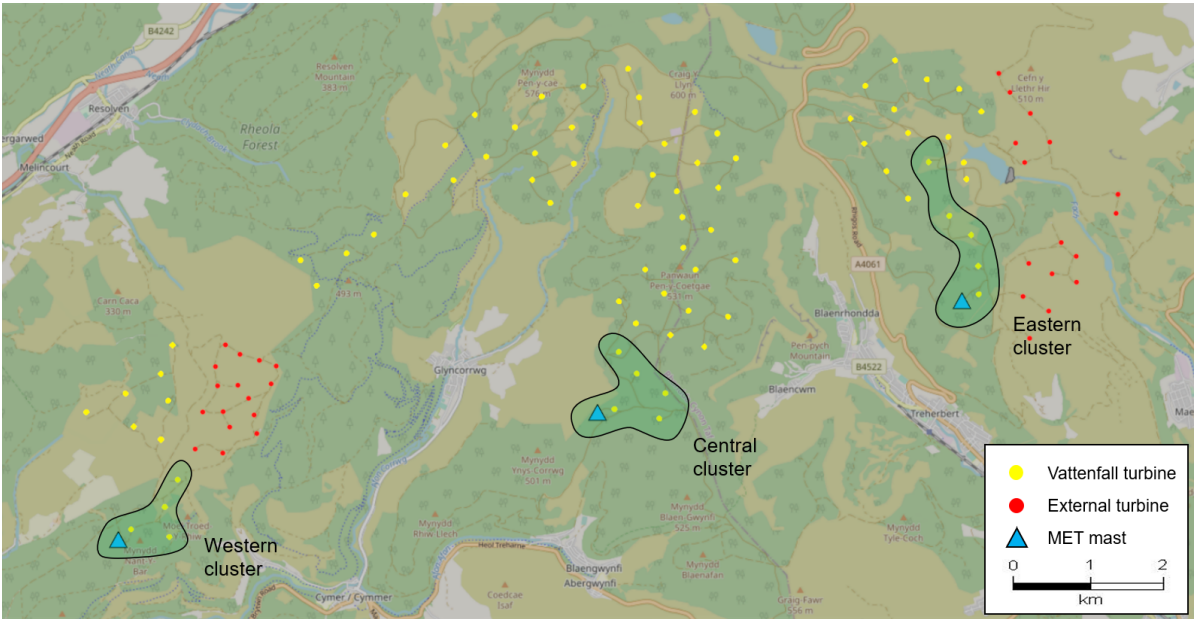


Figure 4.1: PYC wind farm with the three selected clusters.

4.2.1. Western Cluster

The selected turbines and MET mast in the western cluster are presented below. As can be seen, one turbine is particularly close to the MET mast at a distance of 2.2D. Also, all turbines including MET mast are within the same contour lines (between 360 and 370 meters), showing the heights of the MET mast and turbines are comparable.

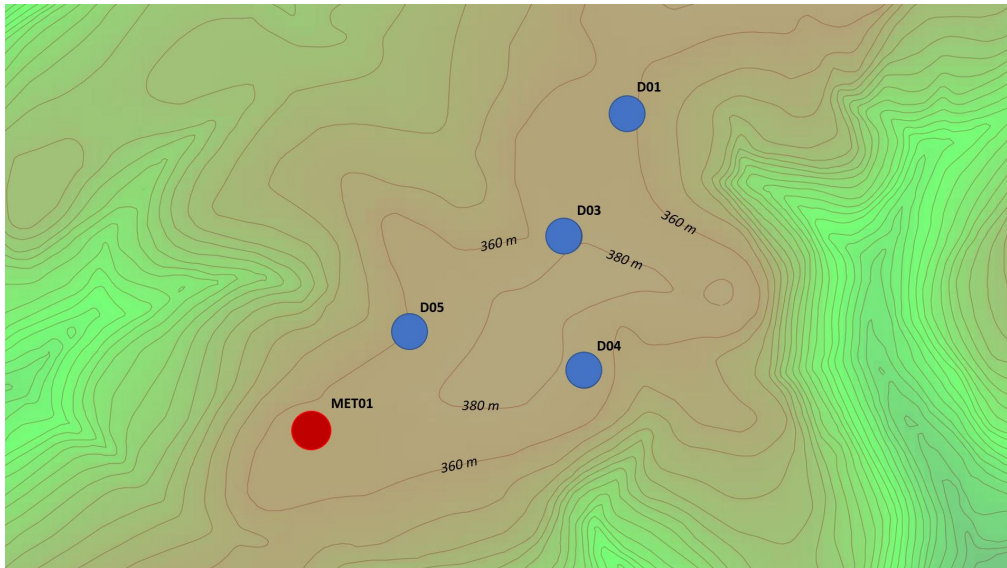


Figure 4.2: Height profile of western cluster (Openwind)

Turbine	Height [m]	Distance [-]				
		MET01	D01	D03	D04	D05
MET01	361	0	10.2D	6.8D	5.9D	2.2D
D01	369	10.2D	0	3.6D	6.8D	8.1D
D03	370	6.8D	3.6D	0	3.6D	4.8D
D04	370	5.9D	6.8D	3.6D	0	4.6D
D05	364.6	2.2D	8.1D	4.8D	4.6D	0

Table 4.4: Distances between selected turbines from western cluster

A range of free sectors for the MET mast and turbines of the western cluster are given in the polar plots below. The green sectors in the images represent approximations of free sectors. The plots are generated by a Python code which uses the latitude and longitude coordinates of turbines as inputs, as well as a threshold regarding the distance for taking into account wakes of neighbouring turbines. For filtering out any influence by wakes, a threshold of 20D (20 times rotor diameter) is used. The wakes of a turbine within this distance are taken into account in the free sector model. Wake expansion is modelled using equation 2.16.

The wind directions shown in the diagrams below will therefore be used for the performed unwaked sector analysis. The small unwaked sectors (narrow green areas) that can be seen in some polar plots will not be taken into account, as these areas are too sensitive for wakes in a complex terrain as PYC.

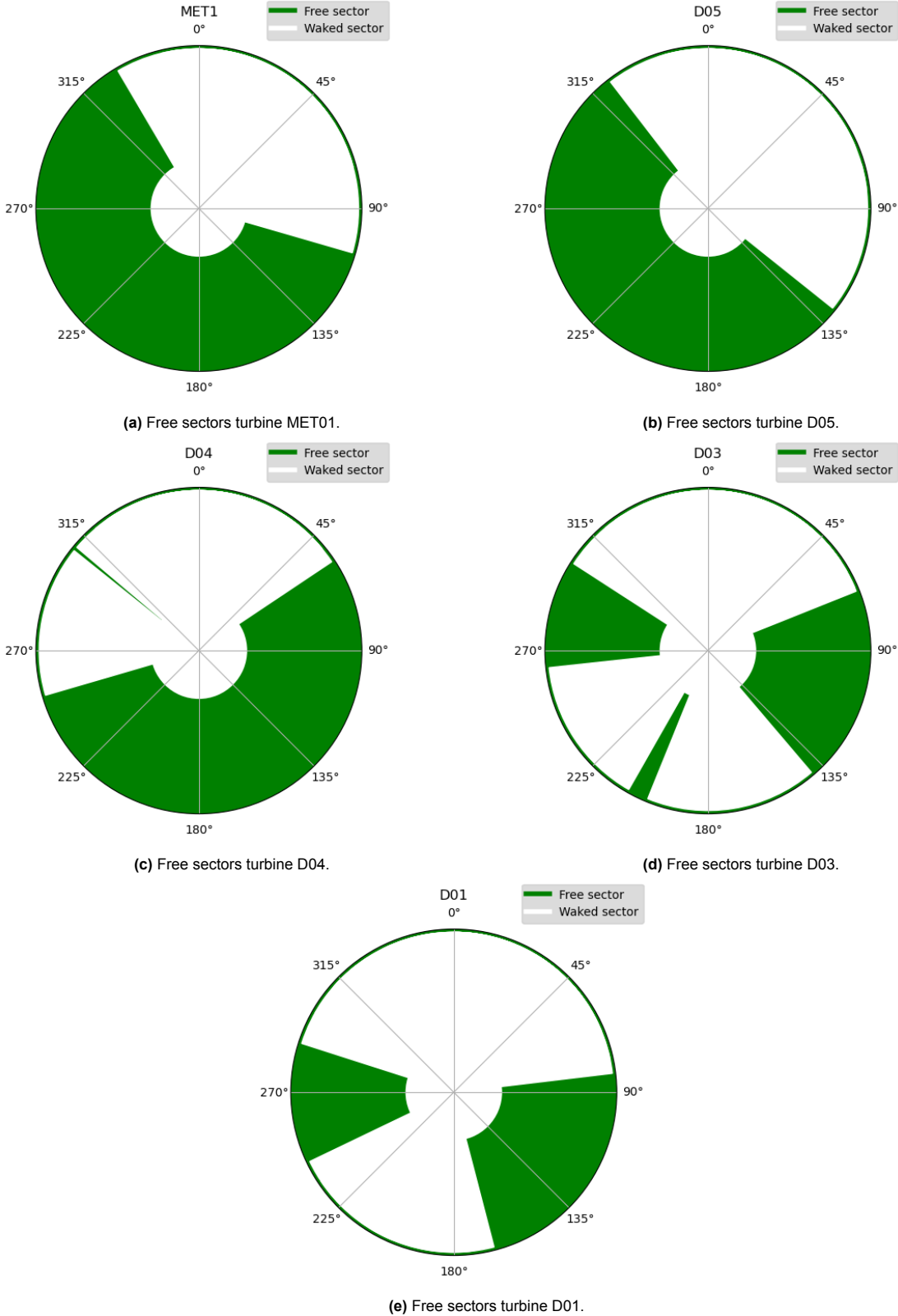


Figure 4.3: Free sectors of the MET mast and turbines of the western cluster.

4.2.2. Central Cluster

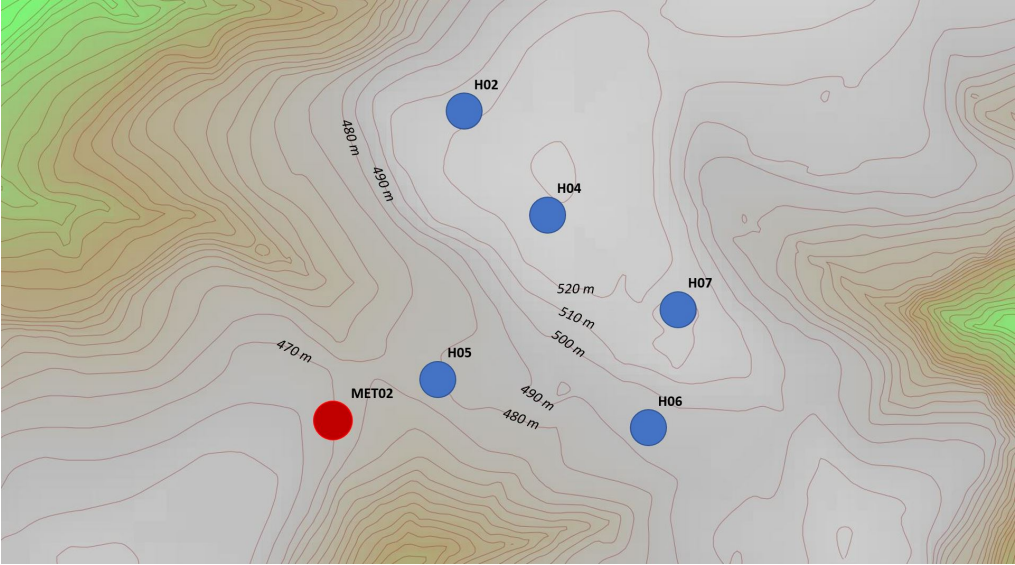


Figure 4.4: Height profile of central cluster (Openwind)

Turbine	Height [m]	Distance [-]					
		MET02	H02	H04	H05	H06	H07
MET02	463	0	7.8D	6.7D	2.0D	7.3D	8.4D
H02	511.7	7.8D	0	3.3D	6.8D	9.2D	7.4D
H04	520	6.7D	3.3D	0	5.0D	6.0D	4.1D
H05	471.2	2.0D	6.8D	5.0D	0	5.4D	6.3D
H06	488.3	7.3D	9.2D	6.0D	5.4D	0	3.1D
H07	510	8.4D	7.4D	4.1D	6.3D	3.1D	0

Table 4.5: Distances between selected turbines from central cluster

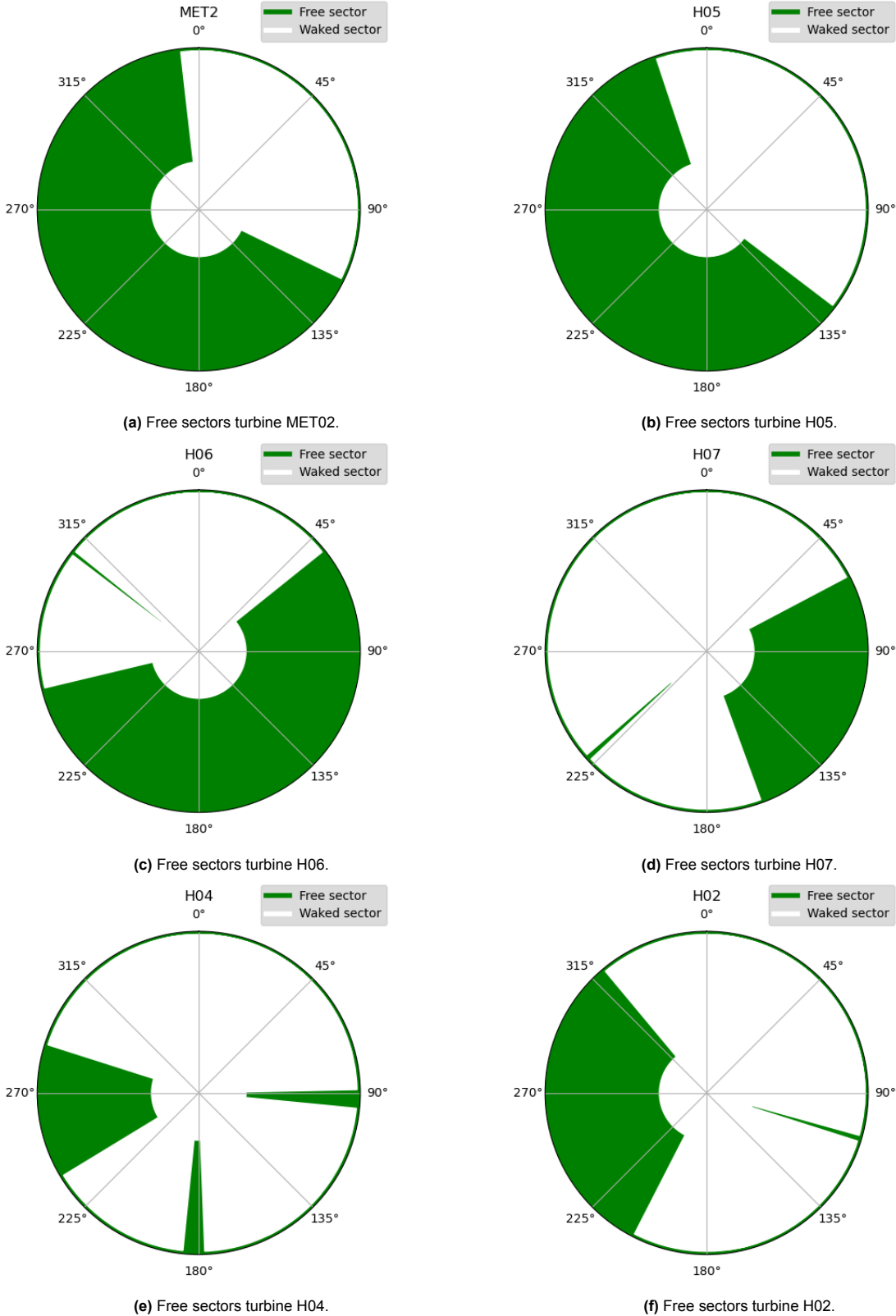


Figure 4.5: Free sectors of the MET mast and turbines of the central cluster.

4.2.3. Eastern Cluster

For the eastern cluster, five turbines and one MET mast were selected. As also accounts for the other cluster, one turbine is relatively close to the MET mast: at 2.0D. The turbines are close in elevation, but there are bigger differences compared to the western cluster. Especially turbine L01, which is one of the worst performing turbines, is situated lower compared to the rest of the eastern string.

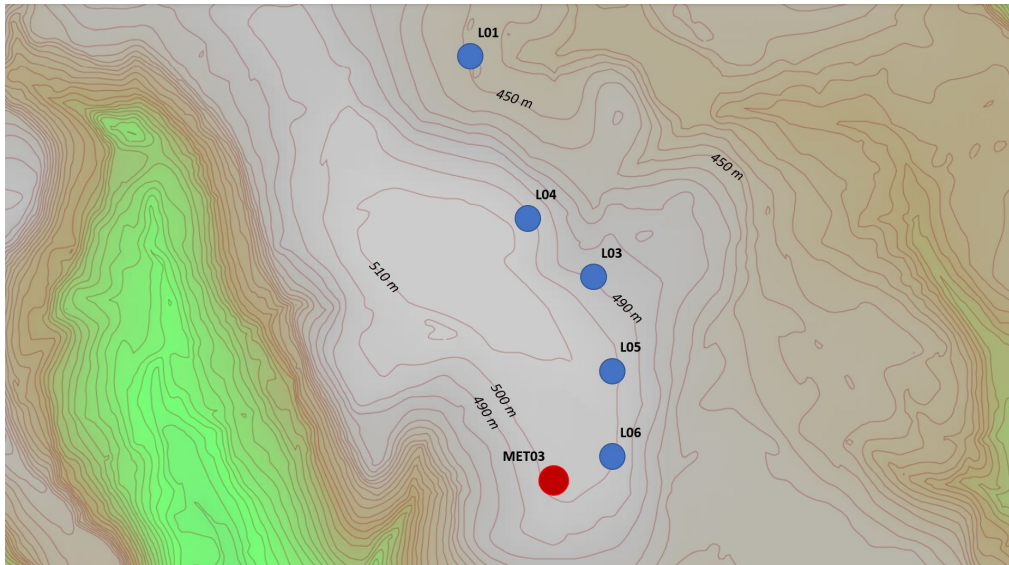


Figure 4.6: Height profile of eastern cluster (Openwind)

Turbine	Height [m]	Distance [-]					
		MET03	L01	L03	L04	L05	L06
MET03	504	0	17.8D	8.3D	10.8D	4.7D	2.0D
L01	449.3	17.8D	0	10.4D	7.1D	14.3D	17.5D
L03	496.1	8.3D	10.4D	0	3.6D	3.9D	7.4D
L04	504.3	10.8D	7.1D	3.6D	0	7.2D	10.4D
L05	503.2	4.7D	14.3D	3.9D	7.2D	0	3.5D
L06	500	2.0D	17.5D	7.4D	10.4D	3.5D	0

Table 4.6: Distances between selected turbines from eastern cluster

The free sectors for the MET mast and the five turbines from the eastern cluster is given below. As can be seen, the further up the string, the less unwaked sectors a turbine experiences. This is due to the turbine cluster north-west of the eastern string. There are relatively more sectors affected by wakes in this eastern string compared to the western string. This is due to the relatively remote location of the western string compared to the rest of PYC.

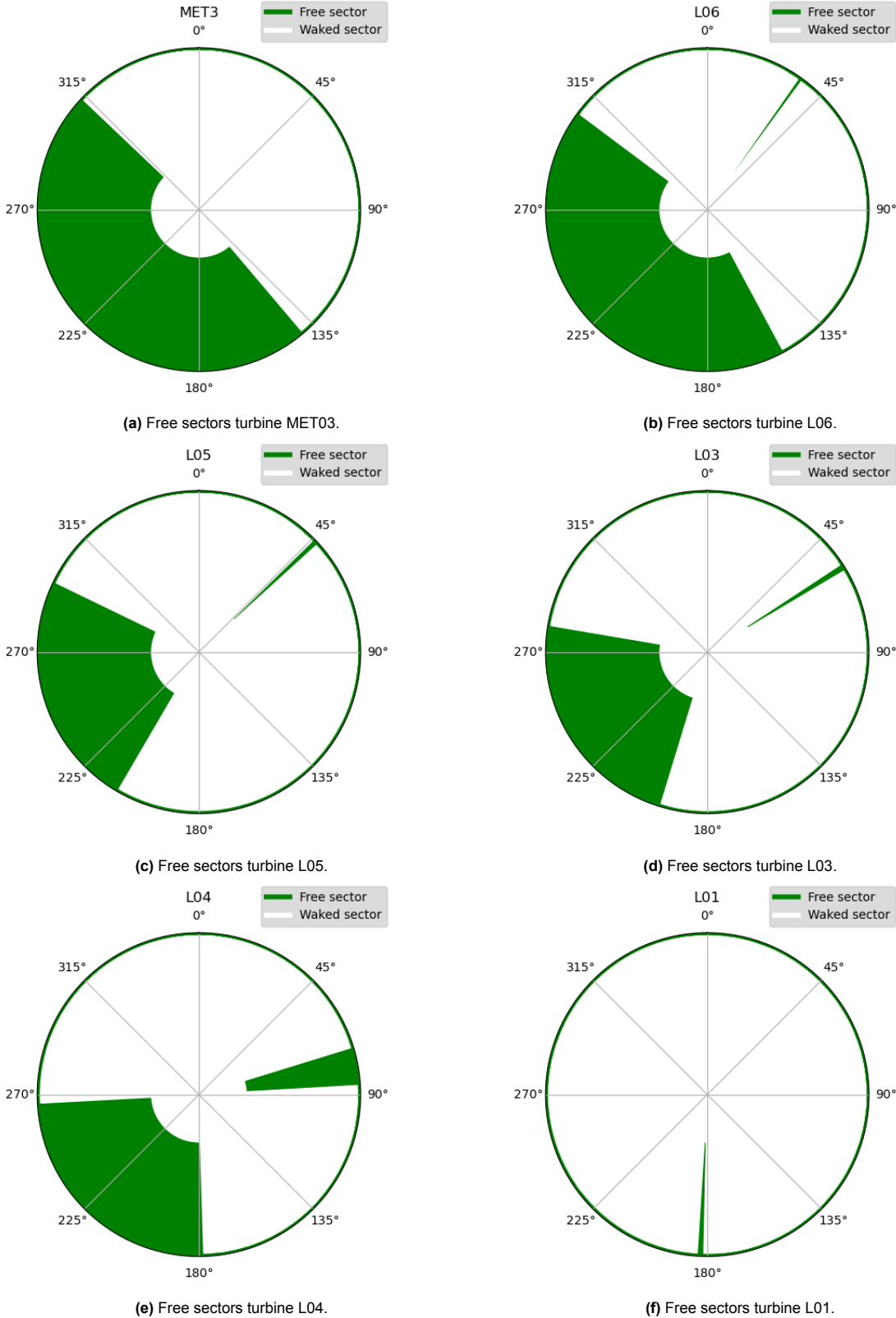


Figure 4.7: Free sectors of the MET mast and turbines of the eastern cluster.

4.3. Wind farm performance

To get insights into wind farm energy production compared to its capacity, the capacity factor can be calculated. This is the ratio of energy production over a period, and the production if the farm had been producing at rated power for that full period (van Wingerden, 2021).

For the five year of operation, this capacity factor can be calculated as follows:

$$CF = \frac{AEP}{365 * 24 * n_{turbines} * P_{rated}} \quad (4.1)$$

Where:

- CF = Capacity Factor
- AEP = Annual Energy Production
- $n_{turbines}$ = Number of turbines
- P_{rated} = Rated power

Using an average AEP of 631 kWh, 76 turbines and a rated power of 3000 kW, a capacity factor of 32% is calculated. This 32% represents all five years of operation. The year 2020¹ has a capacity factor of 35%. On average, the UK reached capacity factors of 35% that year. Only Norway reached higher onshore capacity factors (37%) (IEA, 2020). Capacity factors give information about the harvested power compared to the theoretical maximum. It does not give an insight into turbine performance. However, what can be said is that PYC is in line with average UK capacity factors for the year 2020. Important note is that site location has big influence on capacity factors: a higher located site reaches more favourable capacity factors. PYC is a relatively high site compared to other UK wind farms.

¹The investigated year within this thesis.

5

Data gathering

This part of the report describes the process of data gathering and filtering. The chapter starts with a description of the initially collected pre-filtered data, and consequently elaborates on the filtering process for this data set, which is needed to work with the data properly.

5.1. Data gathering

To perform the data analysis on the PYC wind farm, data has to be gathered from the Azure database. This database contains all 10 minutes averaged as well as HF data used by Vattenfall for a wind resource analyses. The data is saved in a structured way by IEC tags. These tags have been analysed by a previous intern, Luuk Schouten (Schouten, 2021). He made an overview of the IEC tags and definitions, and described which tags are needed for performing a PCA (principal component analysis). His documentation will be used for the selection of the needed IEC tags to collect from the Azure database. According to Schouten (2021), must-needed tags for a performance analysis are the tags listed in table 5.1 below.

Parameter	Name in WPDC / SCADA tag	WWP		Azure		Description	Unit
Active Power	WTUR11.W.mag	X	X	X	X	Active Power	kW
Wind Speed	WMET11.HorWdSpd.mag	X	X	X	X	Horizontal Wind Speed	m/s
Wind Direction	WMET11.HorWdDir.mag	X	X	X	X	Horizontal Wind direction, absolute	°
Nacelle Direction	WNAC11.Dir.mag	X	X	X	X	Nacelle Direction	°
Pitch Angle	WROT11.BIPthAngVal.mag	X		X		Pitch angle at all blades	°
	WROT11.BIPthAngVal1.mag		X		X	Pitch angle for blade A	°
	WROT11.BIPthAngVal2.mag		X		X	Pitch angle for blade B	°
	WROT11.BIPthAngVal3.mag		X		X	Pitch angle for blade C	°
Temperature	WMET11.EnvTmp.mag	X	X	X	X	Temperature of environment	°C
RPM	WGEN11.RotSpd.mag	X	X	X	X	Rotational Speed Generator	rpm
	WROT11.RotSpd.mag	X	X	X	X	Rotational Speed Rotor	rpm

Table 5.1: Must-have tags and their availability in the Wind Web Portal (WWP) and Azure database.

In addition to those tags that are used for turbines, there are tags that contain MET mast data. An overview of the tags for each MET mast are given below:

Parameter	Name in WPDC / SCADA tag	Description	Unit
Humidity	MMET0X.EnvHum.mag	MET X Humidity	%
Pressure	MMET0X.EnvPres.mag	MET X Pressure	mBar
Temperature	MMET0X.EnvTmp.mag	MET X Temperature	°C
Wind Direction	MMET0X.HorWdDirHub5.mag	MET X Wind Direction Hub - 5m (83.725 m)	°
	MMET0X.HorWdDirMWay.mag	MET X Wind Direction Midway (59.75 m)	°
	MMET0X.HorWdDirRot.mag	MET X Wind Direction Hub (88 m)	°
	MMET0X.HorWdDirTipLow.mag	MET X Wind Direction Tip-Low (35.5 m)	°
Wind Speed	MMET0X.HorWdSpdHub2.mag	MET X Wind Speed Hub - 2.5m (85.2 m)	m/s
	MMET0X.HorWdSpdHub5.mag	MET X Wind Speed Hub - 5 m (83.725 m)	m/s
	MMET0X.HorWdSpdMWay.mag	MET X Wind Speed Midway (59.75 m)	m/s
	MMET0X.HorWdSpdRot.mag	MET X Wind Speed Hub (88 m)	m/s
	MMET0X.HorWdSpdTipLow.mag	MET X Wind Speed Tip-Low (31.5)	m/s

Table 5.2: MET mast tags.

Apart from the tags describing the measured site conditions at the turbine, information is needed about the curtailment data and alarms recorded on the MET mast and turbines. This information is also logged and collected in different tags. Furthermore, an overview of alarms can be downloaded from the Wind Power Data Center (WPDC) database.

5.1.1. Pre-filtered data

Before filtering and processing the collected data, it is important to comprehend the size and properties of the data. Per cluster, there is data collected for the MET mast as well as for the turbines. This data is HF data (Sampling time < 10 minutes) and 10 minutes averaged data. To visualise the size of the different tags collected, average sampling frequencies for multiple data sets are shown below.

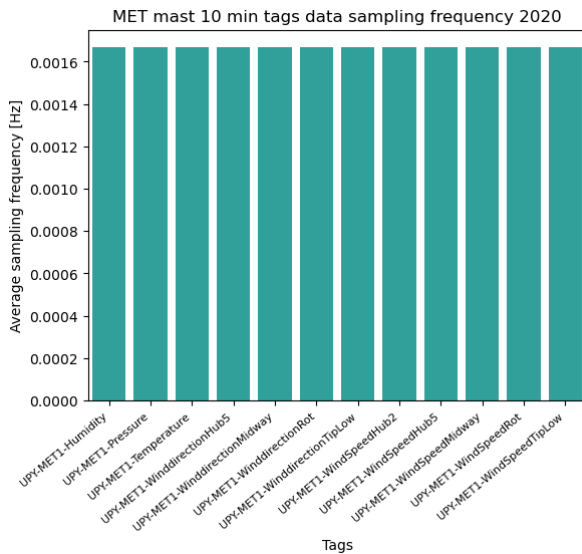


Figure 5.1: MET 10 minutes averaged data sampling frequency

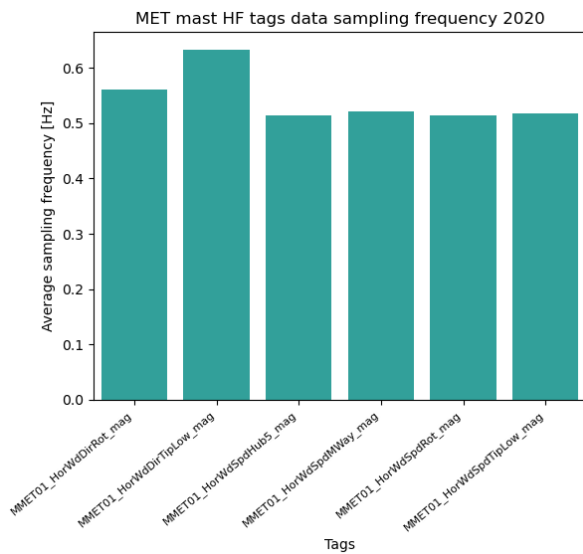


Figure 5.2: MET HF data sampling frequency

For the MET mast, the 10 minutes averaged collected data fulfils the expectations regarding sampling frequency. An average sampling frequency of 0.0017 Hz matches a sampling time of 10 minutes. For the high-frequency data, sampling frequencies and therefore sizes of the different tags are of the same order of magnitude. The HF data collected for the MET mast is collected at a rate around one sample every two seconds (Figure 5.2).

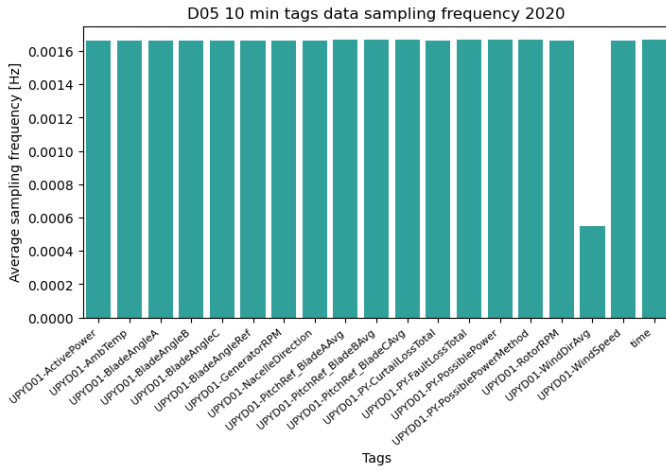


Figure 5.3: D05 10 minutes averaged data sampling frequency

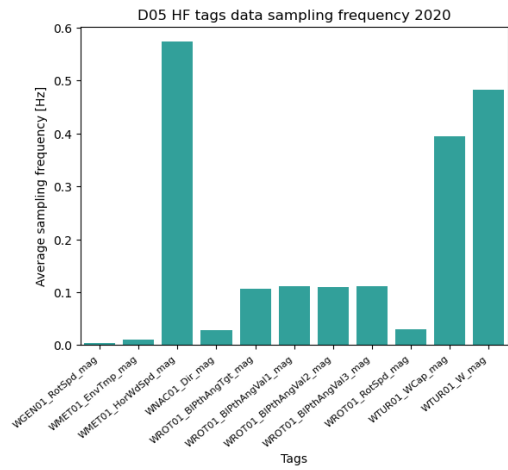


Figure 5.4: D05 HF data sampling frequency

Above, an example of collected HF and 10 minutes averaged data for the turbines is shown. As a reference, the D05 turbine is chosen. The patterns for this turbine are the same as for the other turbines of the eastern string. For the 10 minutes averaged data of the turbines, the sampling frequencies are the expected value, except for the wind direction tag. This is because from the data it can be seen that turbine wind direction measurements are only collected from 02-09-2020 onwards. This limits wind directional analyses that can be performed at the PYC wind farm.

For the HF data collected at the turbines, a wide spread in sampling frequencies is seen. Wind speed and active power output are measured at an average frequency of around one sample every two seconds. Also, pitch data is collected at a relatively high rate. One sample is collected every 10 seconds on average.

5.2. Post-filtered data

Below, filtering checks were carried out on turbines from the western cluster. The post-filtered data was checked for periods of curtailments, and the presence of data within these periods. If data is filtered out correctly, no data is available for periods of curtailment. This is visible in the figures below.

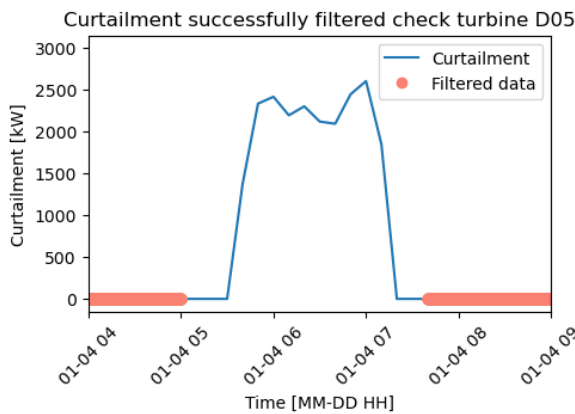


Figure 5.5: Filtered out data in a period of curtailment.

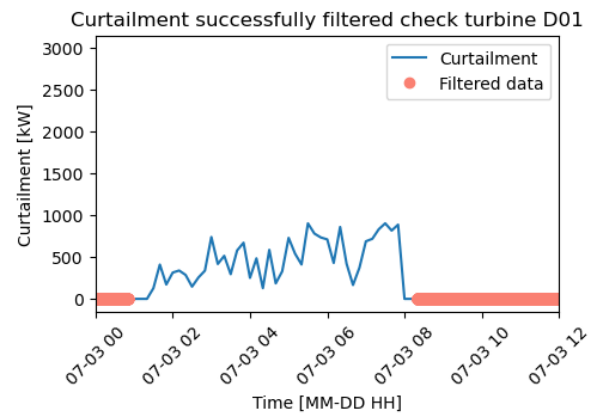


Figure 5.6: Filtered out data in a period of curtailment.

In the table below, an overview is given of the different parameters logged by the SCADA system, as well as the calculated parameters during the filtering process. Moreover, the naming given to these parameters during the renaming process is shown in the table. This naming is used during the analysis.

Table 5.3: Data tags after renaming.

Parameter	Tag name	Unit	Availability	
			MET	Turbine
Air humidity	Humidity	g/m ³	X	
Air pressure	Pressure	hPa	X	
Air temperature	Temperature	°C	X	X
Wind direction at hub height	Wind direction hub	°	X	X
Wind direction 5m below hub height	Wind direction hub -5m	°	X	
Wind direction at blade midway height	Wind direction midway	°	X	
Wind direction at lowest blade tip height	Wind direction tip low	°	X	
Wind speed at hub height	Wind speed hub	m/s	X	X
Wind speed 2m below hub height	Wind speed hub -2m	m/s	X	
Wind speed 5m below hub height	Wind speed hub -5m	m/s	X	
Wind speed at blade midway height	Wind speed midway	m/s	X	
Wind speed at lowest blade tip height	Wind speed tip low	m/s	X	
Corrected wind speed at hub height	Wind speed hub (c)	m/s	X	X
Time	Time	y-m-d h:m:s	X	X
Wind directional veer	Veer	°/m	X	
Wind shear	Shear	-	X	
Turbulence intensity	TI	-	X	X
Active power output signal	Active power	kW		X
Pitching of blade 1	Pitching blade 1	°		X
Pitching of blade 2	Pitching blade 2	°		X
Pitching of blade 3	Pitching blade 3	°		X
Power lost due to curtailment	Curtailements	kW		X
Power lost due to alarm	Alarms	kW		X
Rotational speed of generator	Generator speed	RPM		X
Rotational speed of rotor	Rotor speed	RPM		X
Direction of the turbine nacelle	Nacelle direction	°		X
Misalignment nacelle and wind direction	Yaw misalignment	°		X

As can be seen in the table, wind speed and wind directional measurements are collected at multiple heights. To understand the naming of the different height levels, figure 5.7 below shows the different measurement heights (projected on the turbine) including the naming. Please note that the measurements at those different heights are measured at the MET mast only. At the turbines, wind speed and wind direction is only measured at hub height level. In figure 5.7, the heights Hub -2m and Hub -5m are not drawn to scale. This is done to make these 2 measurement levels visible in the figure. In reality, these measurement heights are respectively 2 meters and 5 meters below hub height.

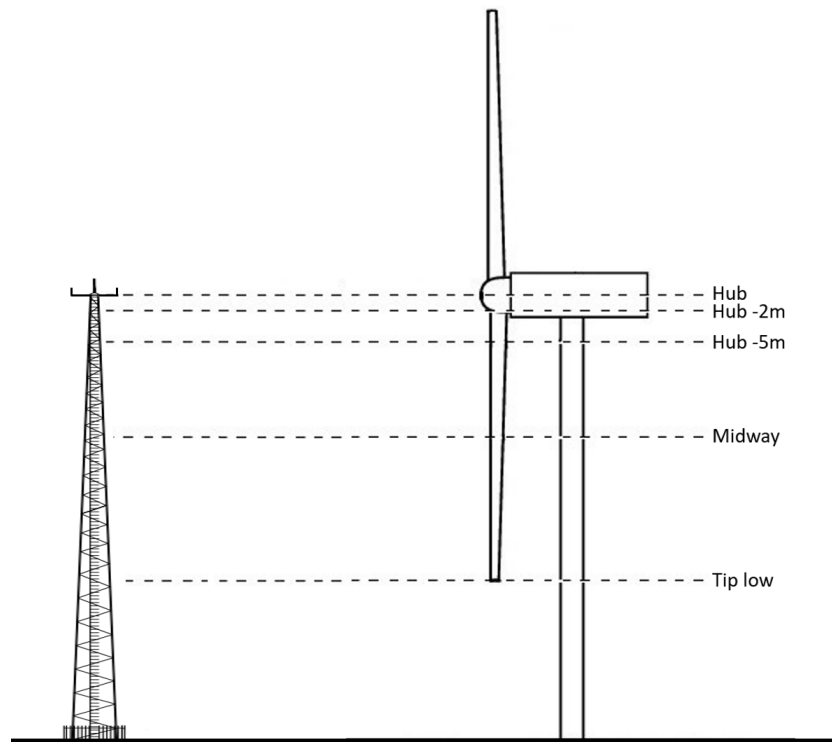


Figure 5.7: Multiple MET mast measurement heights.

5.3. Limitation of data: lack of wind directional measurements

As can be seen in figure 5.3, the wind direction parameter is sampled at a lower annual average frequency than the other parameters. This is because this parameter is only collected from the 2nd of September 2020 onwards. This means directional data is only available for three months of the analysed year. As wind directional data is of immense importance for a wind site analysis, it would influence the quality of this work if directional analyses could only be carried out for the last quarter of the year.

Therefore, the correlation of nacelle direction and wind direction was investigated ¹. When nacelle direction and wind direction are showing a steady correlation, the nacelle consequently follows the direction from the wind. If this is the case, the nacelle direction can be assumed to represent the wind direction in a precise enough way. To have directional data available for the whole year of 2020, the nacelle direction will serve as wind directional data ².

In the figure below, nacelle direction and wind direction are correlated for 10 minute data for turbine D05. Correlations of the other correlations are represented in appendix B.4. Other turbines from the western cluster show the same strong correlation.

¹This was done after filtering the data (section 3.1.1, as solely filtered data is used during the cloud analysis)

²Within the scope of this thesis, using nacelle direction as a representative of wind direction is precise enough. When one is investigating directional influences on turbine performance on the scale of single degrees, this assumption can cause wrong results. This work is based upon wind direction bins of 10 degrees minimum

10 min wind direction - nacelle direction correlation D05

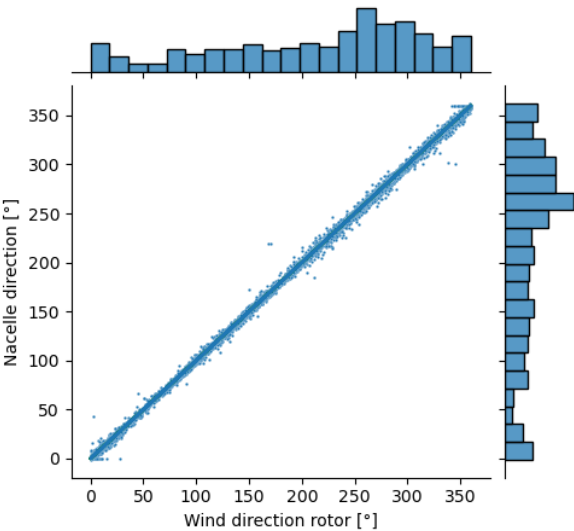


Figure 5.8: Wind direction - nacelle direction correlation turbine D05

Properties of the correlation above, and other correlations from the western cluster, are given in the table below:

Table 5.4: Properties of wind direction - nacelle direction correlations for turbines of the western cluster.

	Slope	Intercept	R^2 score
D05	1.00	0	1.00
D04	1.00	0	1.00
D03	1.00	0	1.00
D01	1.00	0	1.00

The perfect linear correlation with maximum r-squared values confirm the assumption that nacelle direction can be a good representative for the wind direction at the western cluster. In the cloud analysis, nacelle direction will therefore be used to represent the wind direction.

6

Results western cluster

Within this chapter of the report, the results are discussed of the analyses performed on the filtered data of the western cluster. The approach of this analysis is described earlier this report in chapter 3. At first, the processed MET mast data is used to understand the site conditions at the western cluster. Secondly, correlations of wind speed and wind directional data are checked in a side-by-side approach to validate the data. At last, a performance analysis is performed on the validated turbine data, including a cloud analysis.

6.1. Wind resource

Firstly, the MET mast is used to understand the site influence on the incoming wind at the eastern cluster. In addition, the MET mast is used to calibrate the reference turbine closest to the mast. This is done in the next section. Although each turbine experiences other specific site conditions, the MET mast gives a good general perception of these conditions. Moreover, it is the only location within the cluster where veer and shear are measured. The wind speed and wind directional distributions are set up by respectively a wind rose and a Weibull fitted histogram. On top of that, veer shear and TI are investigated. In chapter 7, where MET mast data of the eastern cluster is analysed, the site behaviour of the western cluster will be used as a comparison.

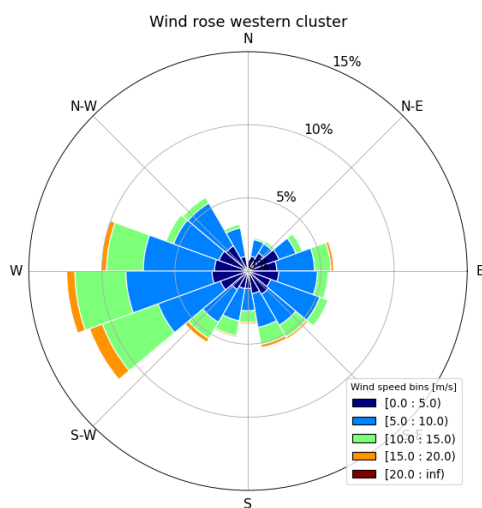


Figure 6.1: Wind rose western cluster

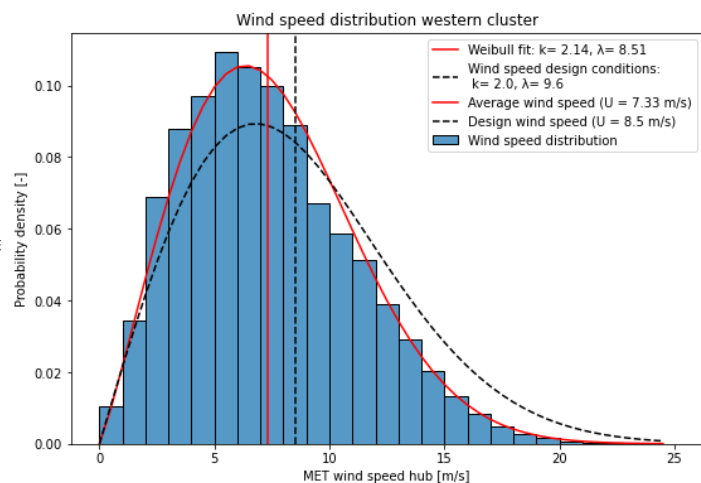


Figure 6.2: Wind speed histogram western cluster

In the figures above, the wind speed and wind direction distributions of the western cluster are presented. As can be seen in figure 6.1, the prevailing wind direction is a W / S-W wind. Occurrence of wind from other directions is more evenly spread. The north-eastern direction is less occurring. Regarding

wind speed, the distribution follows a standard Weibull distribution, with the 5-6 m/s bin as prevailing wind direction bin. The Weibull distribution of the year 2020 is less spread than the pre-defined turbine design conditions by the manufacturer. The mean wind speed of the collected data is also lower than the design mean wind speed.

Next to wind speed and direction, wind veer is analysed at the MET mast. Wind veer is a parameter that can only be calculated at the MET mast, as the MET mast has wind vanes placed at multiple heights. High veer tells a rotor experiences a wide spread in wind directions over the length of the blade. This can increase loads on the blade. Moreover, turbine performance can be negatively affected by high veer. As the turbine only measures the incoming wind direction at hub height, high veer can cause this not being a representative wind direction for the full rotor swept area. Within this thesis, the wind veer is calculated by the difference of measured wind direction per meter length of the rotor [$^{\circ}/m$]. The difference in directional measurements between hub height and lowest tip height is divided by the height difference between those installed wind vanes (56.6 m).

When analysing the wind veer measurements, it was discovered the wind vanes record the wind direction wrongly when the wind speeds are coming from the north when compared to the neighbouring turbine D05. Wind vanes placed on turbine D05 do not show this behaviour. As the wind vane data from the MET mast is used for calculating veer, the veer data contains errors as well when wind is coming from the northern direction. Therefore, for analysing veer, the MET data is filtered for the northern wind direction $\pm 60^{\circ}$. A graph visualising these errors can be found in Appendix A.

Below, wind veer is plotted against wind direction and wind speed. Wind Veer at the MET mast location reaches high values, while following an expected downwards trend for increasing wind speeds. The wind veer shows a symmetric profile, but not around the horizontal axis: it runs towards a veer of $0.25^{\circ}/m$, whereas it is expected to run towards $0^{\circ}/m$ for high wind speeds, as the wind directional profile over height is more neutral for high wind speeds. A wind veer of $0.25^{\circ}/m$ means a directional difference of more than 25 degrees over the full rotor span of the turbine. This had a big influence on the inflow angle of the wind on the turbine, and can have a substantial effect on power production. Looking at figure 6.4, it can be seen the extreme values of veer are mostly measured at a wind direction of 150 degrees. This occurs for wind speeds below cut-in wind speed. From cut-in wind speed, the more extreme veer cases are a minor part of power production.

In the diagrams below, red-coloured sectors are present in the plots including wind directions (e.g. figure 6.4). Those red sectors represent the waked sectors the MET mast experiences. The data itself is represented by the scatter plots. The dark blue lines are the binned averaged lines of the data.

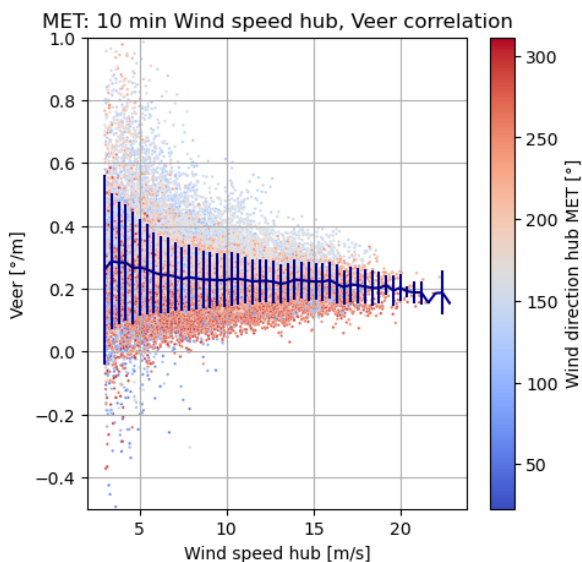


Figure 6.3: 10 minutes averaged MET mast wind speed veer correlation.

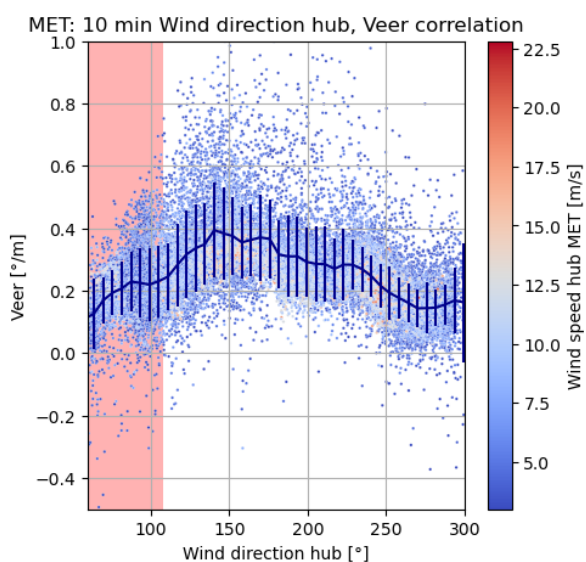


Figure 6.4: 10 minutes averaged MET mast wind direction veer correlation.

The wind shear at the western cluster is represented by the shear factor. This shear factor is calculated in the same manner as is shown by equation 2.9. At the MET mast, negative shear occurs at a wind direction of ± 50 degrees. This could correspond with wind coming from turbine D03. At 150 degrees inflow angle, high shear is detected. Also, the average shear from this wind direction is higher than compared to the other unwaked MET mast sectors. On average, a wind shear factor of 0.5 is measured at the site. This is a high value for a land covered in forestry. According to table 2.6, the shear factor at a wooded landscape is expected to be around 0.3.

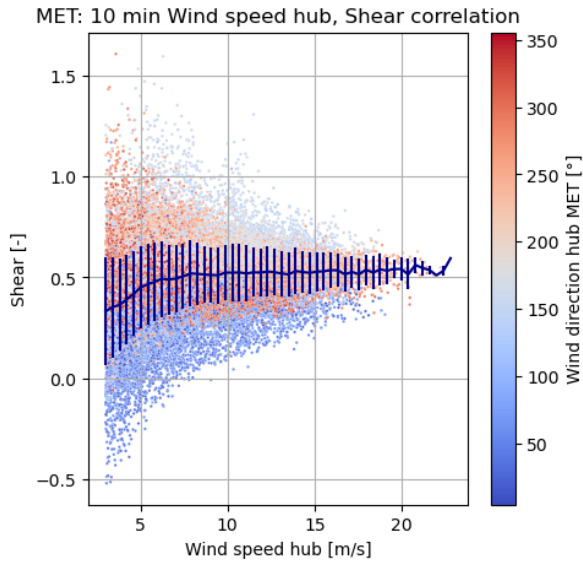


Figure 6.5: 10 minutes averaged MET mast wind speed shear correlation.

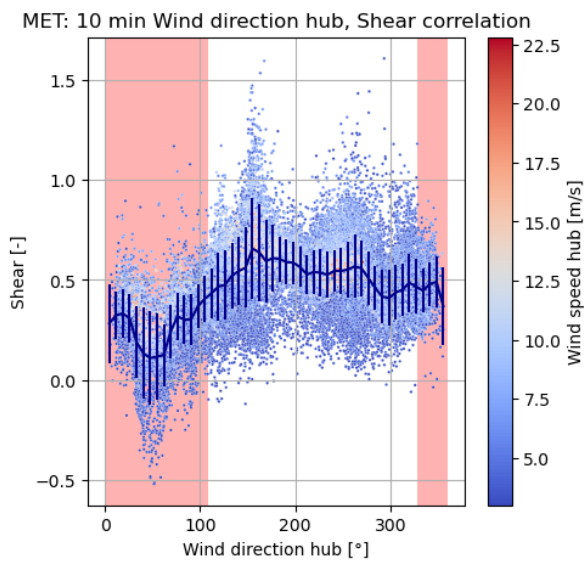


Figure 6.6: 10 minutes averaged MET mast wind direction shear correlation.

To understand the reliability of wind speed measurements at the lowest tip height and hub height, a wind shear profile is fitted to the five wind speed measurement heights. If the other measurements at independent height confirm the shear behaviour from figures 6.5 and 6.6, the shear factor values found can be trusted. A Python curve solver from the SciPy package is used to solve the power law to the average wind speeds measured at the five heights described by table 5.2. In the figure below, results from this fit are shown.

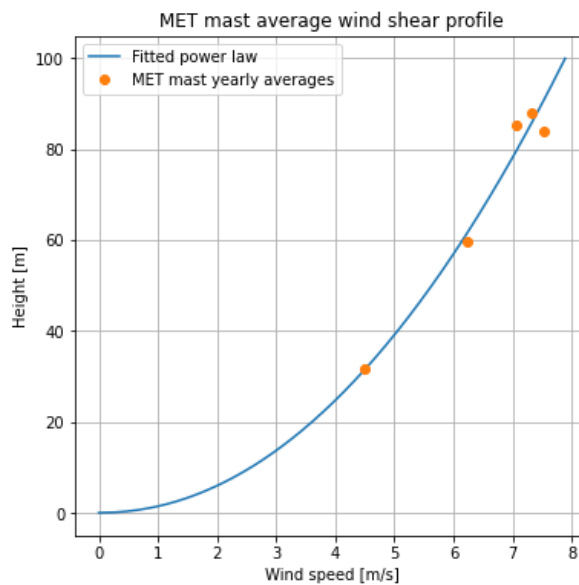


Figure 6.7: 2020 average wind shear profile at the MET mast

From this fitted curve, an average wind shear factor (α) of 0.47 was found. The average wind speeds at other mounted heights match the fitted profile. Only the two other wind speed measurements around the hub (-5 m and -2 m) are a little off the expected trend. However, the measurements are confirming the magnitude of wind speed values measured at hub height.

Turbulence intensity measurements at the site are plotted against IEC standard TI classes A, B and C in the left figure below. Moreover, the design turbulence intensity at 15 m/s (TI_{15}) is plotted by the intersection of the grey dotted lines. The SWT-3.0-113 turbines at the western cluster are IEC TI class B turbines, which means the TI should not exceed the orange curve from figure 6.8. As can be seen from the blue average line, this on average is the case up and until 15 m/s.

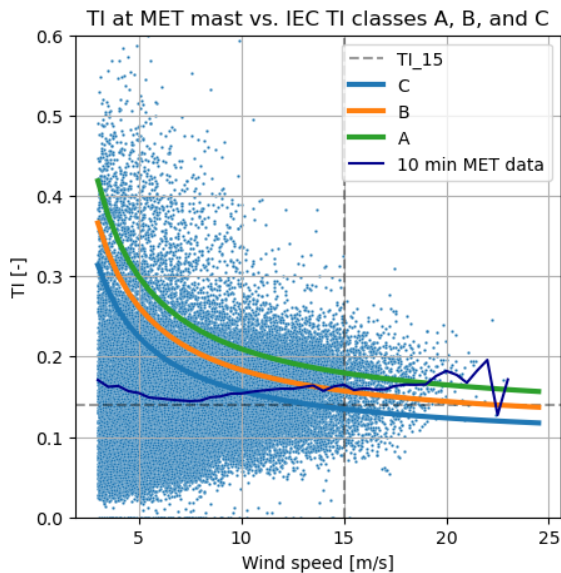


Figure 6.8: 10 minutes averaged MET mast wind speed TI correlation.

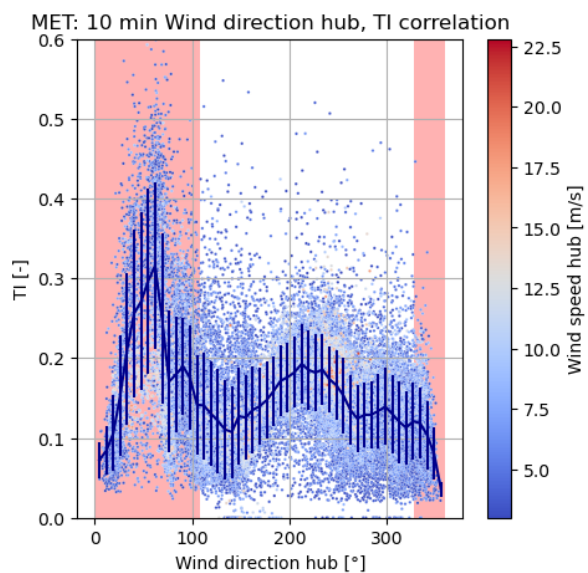


Figure 6.9: 10 minutes averaged MET mast wind direction TI correlation.

The influence of shear, veer and TI on turbine performance will be further investigated in the next sections. The neighbouring turbine of the MET mast will be used to understand veer and shear influence on turbines, as this turbine is within the most comparable condition of the MET masts. TI is analysed at every turbine individually, as this property is measurable at every turbine.

6.1.1. Influence of site on turbine performance D05

There are three parameters possible to extract at the MET mast which describe the site conditions apart from wind speed and direction. Those parameters are shear, veer and TI as discussed in the previous section. MET mast measurements can be used for performance analyses on turbines with a distance of 2-4D to the MET mast (IEC, 2019). For the western cluster, turbine D05 is within these limits. As wind veer and shear are not measured at other turbines, it is valuable to see how veer and shear (measured at the MET mast) influence performance of turbine D05.

To do that, according to IEC, only the undisturbed sectors may be taken into account. This concerns the unwaked sectors of the MET mast as well as the neighbouring turbine (IEC, 2019). In the visualisations below, the disturbed sectors have been filtered out of the combined MET mast and turbine data set.

Below, power production efficiency is plotted for different veer and shear bins measured at the MET mast. Only the veer and shear bins contributing to more than 5% of the data are included. This efficiency is calculated according to equation 3.3. It resembles the performance compared to warranted power curve behaviour. By doing this, it is possible to visualise influence from veer and shear on turbine performance.

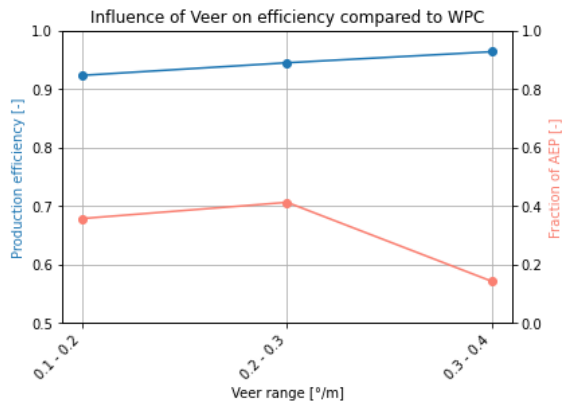


Figure 6.10: Influence of veer on D05 active power signal.

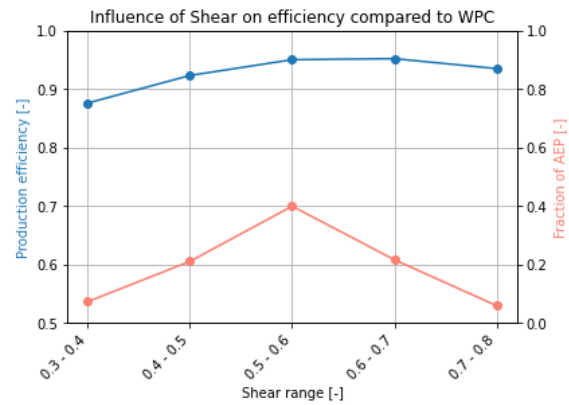


Figure 6.11: Influence of shear on D05 active power signal.

From figures 6.10 and 6.11, it is observed an increase in directional veer and wind shear are both positively effecting turbine performance. For this western cluster, shear reaches relatively high values compared to the other clusters. The same accounts for wind veer.

Remarkable from these findings is the positive effect of wind veer to turbine performance. As veer is measure for wind directional variation over height, high veer means parts of the rotor have a larger misalignment with wind direction. The diagrams above give an indication that the high veer and shear values at the western cluster do not have a negative effect on turbine performance.

6.2. Correlations MET mast and turbines

Below, properties of the correlations between the western cluster's MET mast and turbines are given.

Using equation 2.8, the expected slopes of the wind speed correlation between a reference turbine and test turbine can be estimated. The fraction on the right side of the equation is equal to the slope between two wind speeds at different heights. For calculating the expected slope, two heights of the turbines in question are needed. Moreover, a roughness length of 0.1-1 was assumed (table 2.5). When the wind speed slope is close to the expected slope within a 5% margin, the data of the tested turbine is assumed to be reliable. The expected slopes are listed in the bottom row of table 6.2.

Table 6.1: Overview of correlations for the western cluster.

Correlation	Intercept	Property	MET-D05	D05-D04	D04-D03	D03-D01
Wind speed	No	Slope	1	0.97	0.99	0.97
	No	R^2 score	0.93	0.87	0.84	0.87
Wind dir	No	Slope	1.03	0.97	0.98	1.33
	No	R^2 score	1.00	1.00	1.00	1.00
	Yes	Slope	1.02	0.99	1.00	0.99
	Yes	Intercept	1.50	-5.6	-4.8	-5.2
	Yes	R^2 score	1.00	1.00	1.00	1.00
Nac dir	No	Slope	-	0.97	0.98	1.33
	No	R^2 score	-	1.00	1.00	1.00
	Yes	Slope	-	0.99	1.00	0.99
	Yes	Intercept	-	-5.6	-5.3	-5.3
	Yes	R^2 score	-	1.00	1.00	1.00

Table 6.2: Expected slopes for wind speed correlations western cluster.

	MET-D05	D05-D04	D04-D03	D03-D01
Height difference [m]	3.6	5.4	0	-1
Expected slope ($\alpha = 0.1$)	1	1	1	1
Expected slope ($\alpha = 1$)	1	1.01	1	1
Average	1	1.01	1	1
Accepted slopes (5% deviation)	0.95-1.05	0.96-1.06	0.95-1.05	0.95-1.05

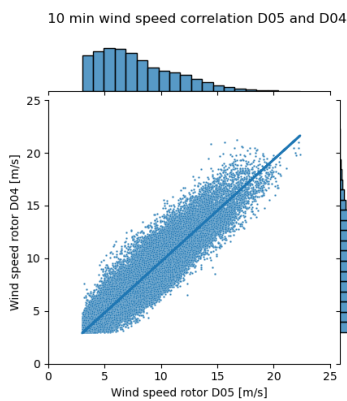
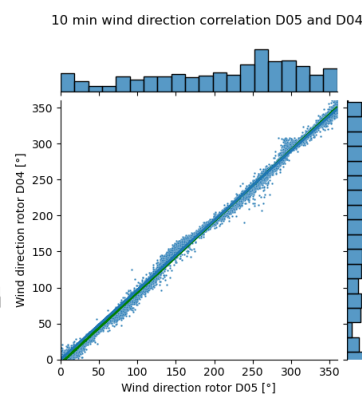
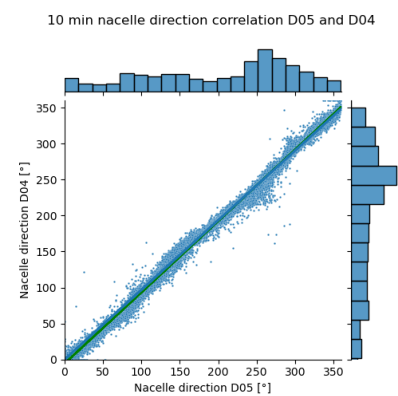
Per property, two correlations were done: one with, and one without intercept (as shown in the second column of table 6.1).

This table contains results after correcting the data for incorrect offsets between turbines, and within turbine data. Before correcting for these false offsets, the data was not usable as data more downstream or the turbine cluster could not be trusted. Two errors were corrected to come to the properties as shown in table 6.1. The corrections carried out are summarised in the table below. Explanations about the thresholds chosen for the corrections can be read in section 3.1.3.

Table 6.3: Overview of corrections carried out on correlations western cluster.

	Operation	Applied on turbine	Properties	Size of offset
1	Correct for offset >10 degrees compared to neighbouring turbine	D04	Nacelle direction Wind direction	86.5 degrees
2	Correct for sudden offset change in time series	D01	Nacelle direction Wind direction	22.5 degrees

To visualise the data shown in table 6.1, the three properties correlated between turbine D05 and D04 are shown in the figures below. Correlation plots for wind speed, wind direction and nacelle direction for the other turbine pairs of the western cluster are shown in appendix B.

**Figure 6.12:** 10 minutes averaged wind speed correlation of D05 and D04**Figure 6.13:** 10 minutes averaged wind direction correlation of D05 and D04**Figure 6.14:** 10 minutes averaged nacelle direction correlation of D05 and D04

From tables 6.1 and 6.2, it can be concluded for all turbines the wind speed correlations are within the expected margins and can be used for further analysis.

6.3. Overall & directional performance analysis

In this section, the results from the performance analysis are shown. This analysis is carried out in parallel to the cloud analysis (described in the next section). The methods to come to the results presented in this section are represented in chapter 3.

6.3.1. Overall turbine performance

Below, the binned power curves of the western cluster turbines are displayed for the full wind direction spectrum. In blue, the power curve for the full wind direction spectrum is shown. In yellow, the power curve for only unwaked sectors is given. These power curves are inputs to equations 3.3 and 3.4 as described in the methodology (chapter 3).

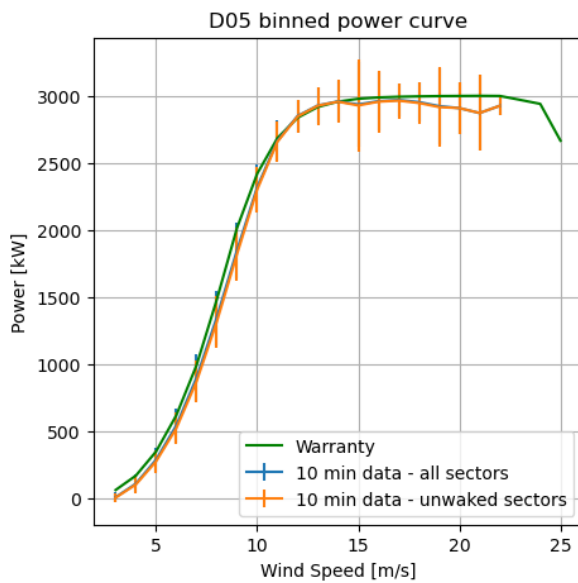


Figure 6.15: Total power curve turbine D05

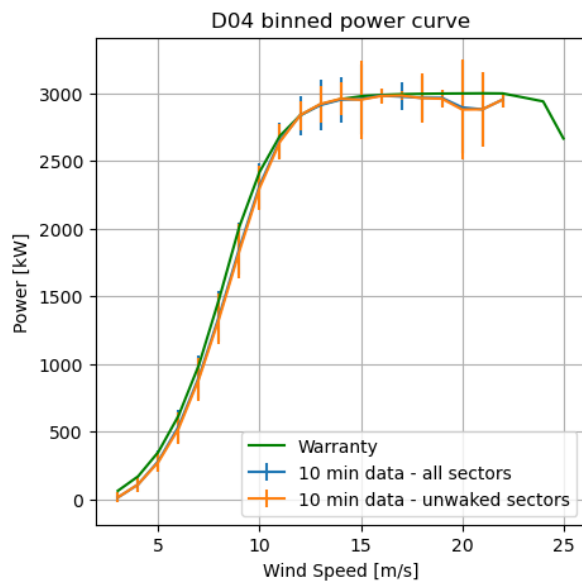


Figure 6.16: Total power curve turbine D04

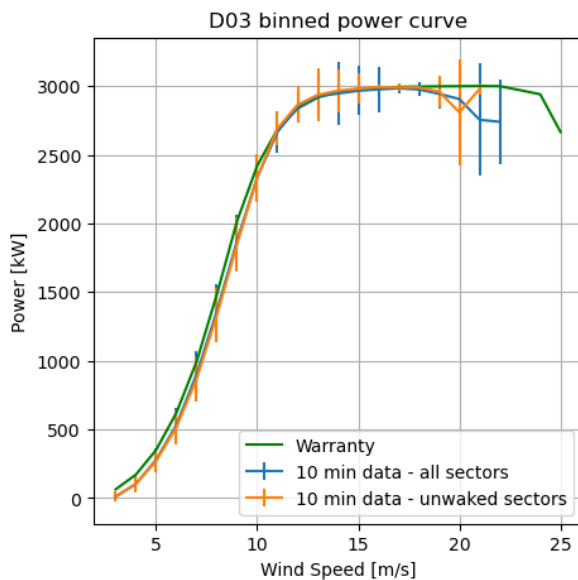


Figure 6.17: Total power curve turbine D03

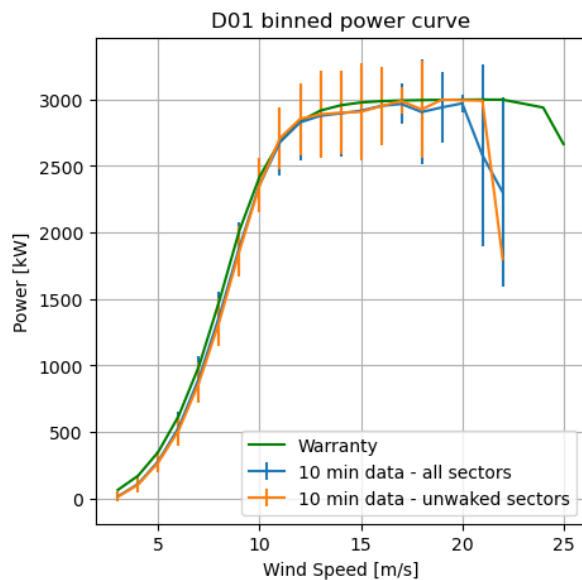


Figure 6.18: Total power curve turbine D01

Looking at the total power curve figures, it can be seen the operational power curve (again, with curtailments, alarms and other defects filtered out) is reaching less power than the warranted power curve for the total wind speed spectrum, except for the knee of the power curve. This is the case as well

for the power curve, where wake effected sectors are not taken into account (blue curve). The small, almost negligible, difference between the blue and yellow curve can be explained by the fact that the prevailing wind directions are not affected by wakes. Below, the overall performance numbers of the western cluster turbines are presented. This is the output of the equations 3.3 and 3.4.

Table 6.4: Efficiencies compared to WPC for turbines western cluster

	Based on	D05	D04	D03	D01
$\eta_{power,1}$	Warranted power curve	0.95	0.95	0.95	0.94
$\eta_{power,2}$	Potential power signal	0.95	0.95	0.94	0.94

From this table it can be seen that based on the two performance indicators $\eta_{power,1}$ and $\eta_{power,2}$, the turbines are performing relatively similar. D01 shows the lowest performance of the four turbines of the western cluster, differing 1% from the better performing turbines from this cluster.

6.3.2. Directional turbine performance

For every turbine, two graphs are plotted with each 9 of the 18 wind direction sectors. These operational power curves per wind direction sector are used to calculate the turbine performance per wind directional sector.

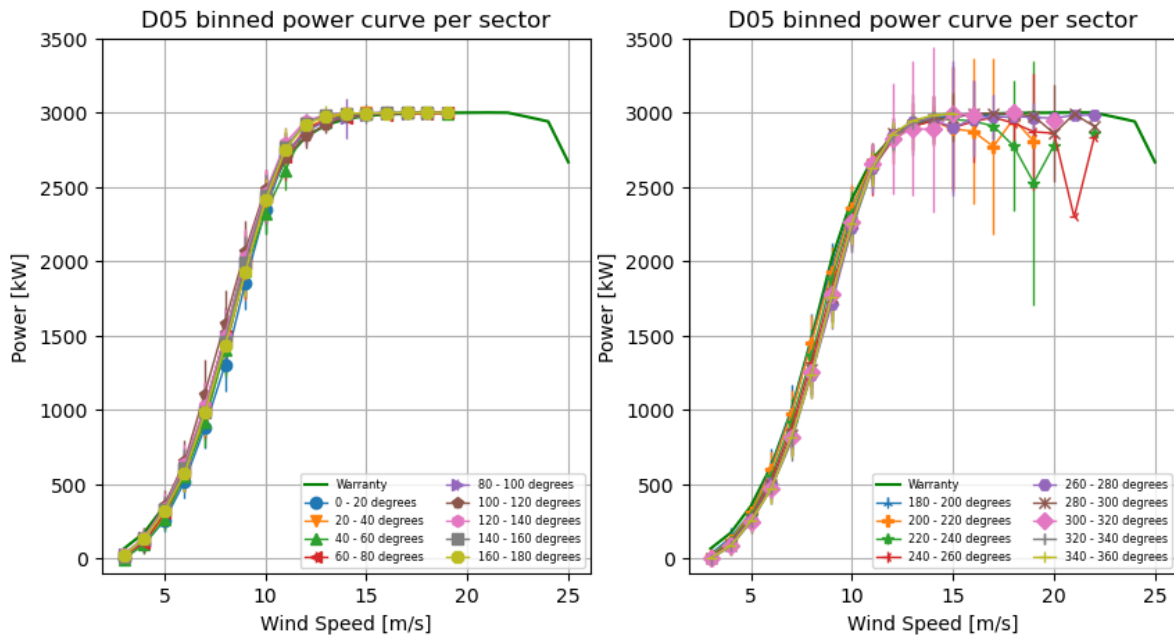


Figure 6.19: Binned power curves turbine D05

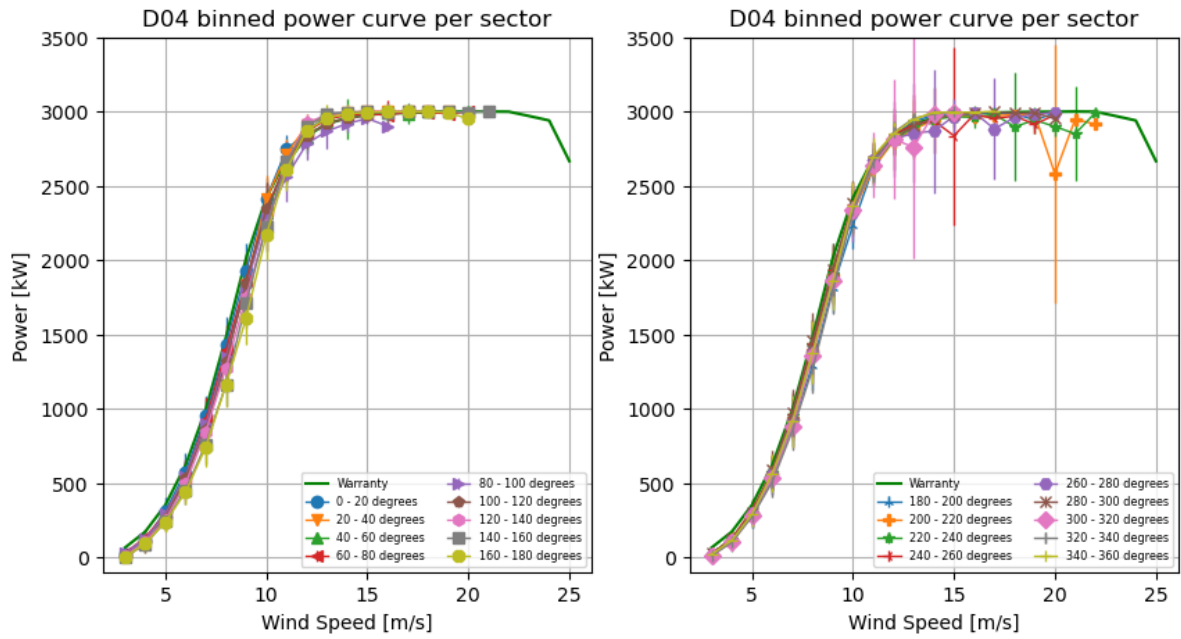


Figure 6.20: Binned power curves turbine D04

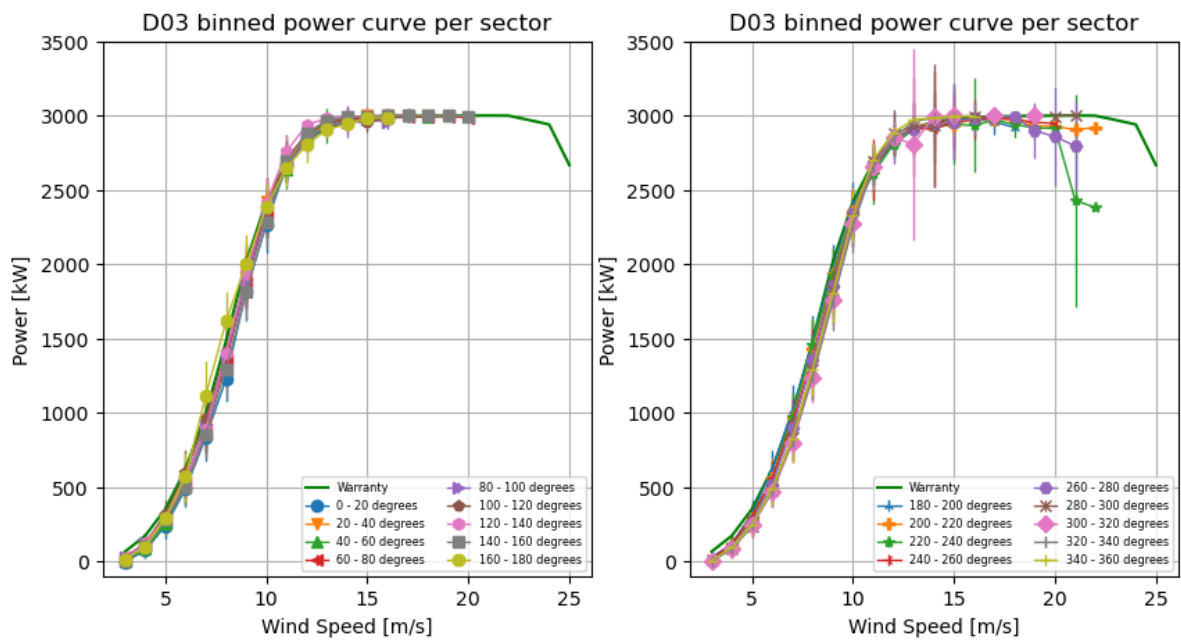


Figure 6.21: Binned power curves turbine D03

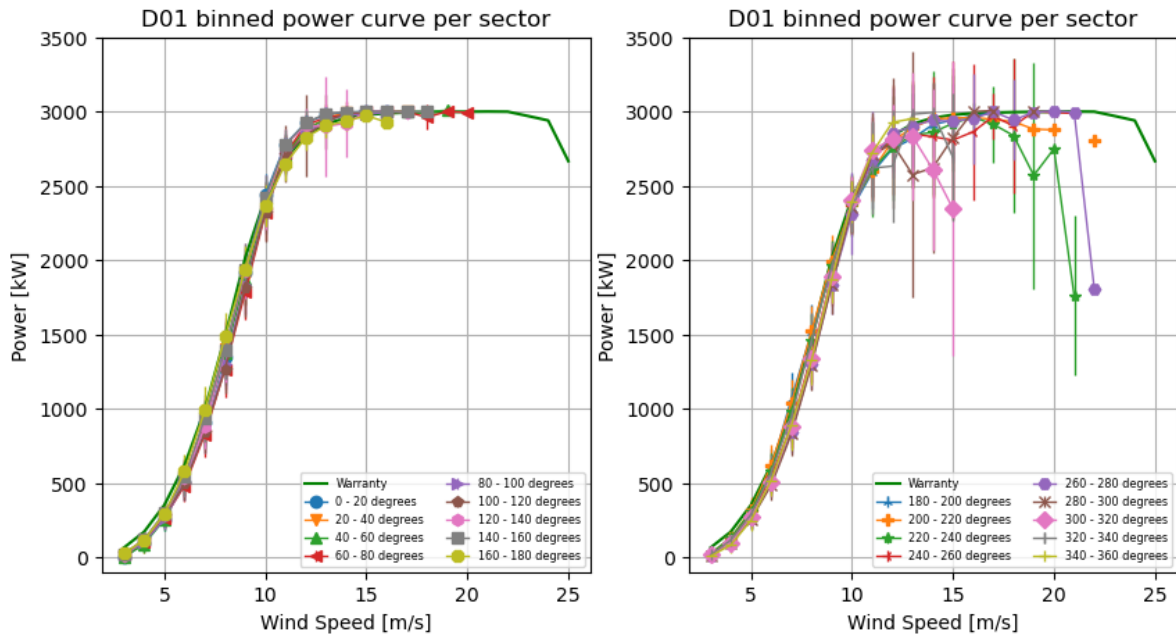


Figure 6.22: Binned power curves turbine D01

From the figures above, it can be seen there is a wide spread in turbine performance for the different wind directions, especially for the rated part of the power curve. This accounts for all four turbines from the western cluster. From the ankle to the knee of the power curve, differences between the different sectors are smaller. Below, the results from equation 3.5 for every 20 degrees directional sector is given. In red, the wake-affected sectors is given, as also illustrated in chapter 4. As this thesis focuses on unwaked sectors, the red sectors will not be further discussed.

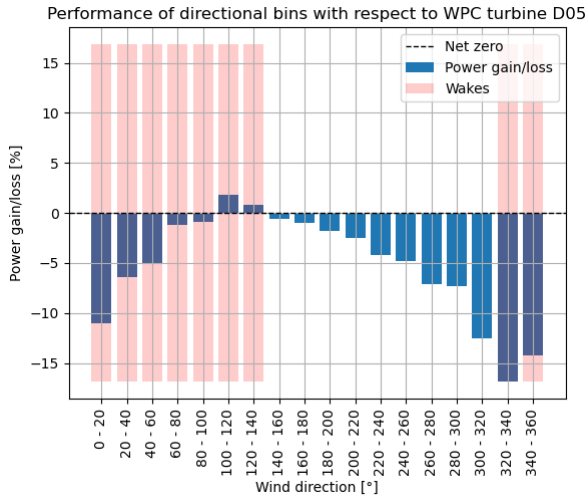


Figure 6.23: Relative power gain/loss per wind direction bin for turbine D05

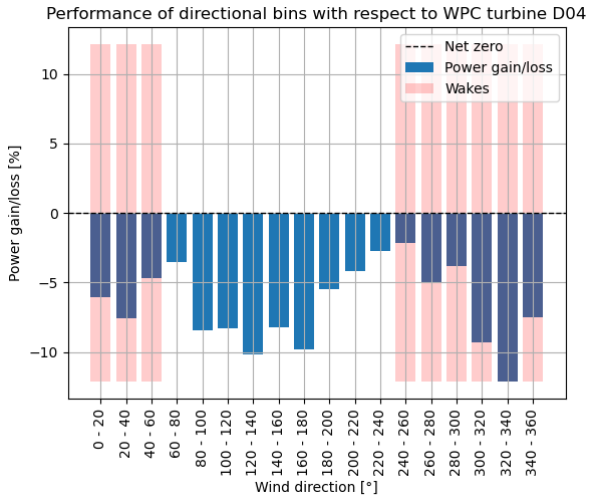


Figure 6.24: Relative power gain/loss per wind direction bin for turbine D04

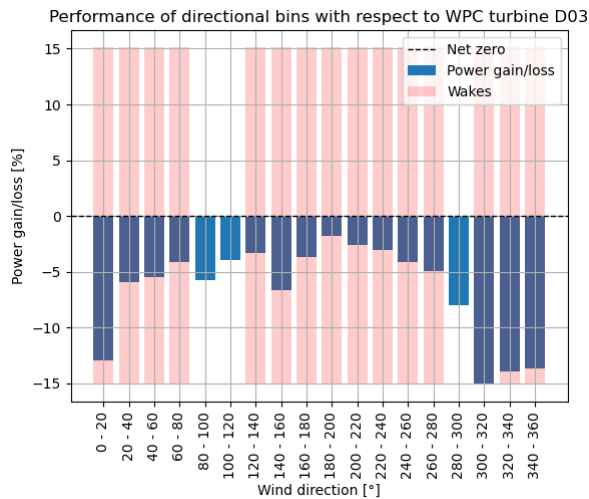


Figure 6.25: Relative power gain/loss per wind direction bin for turbine D03

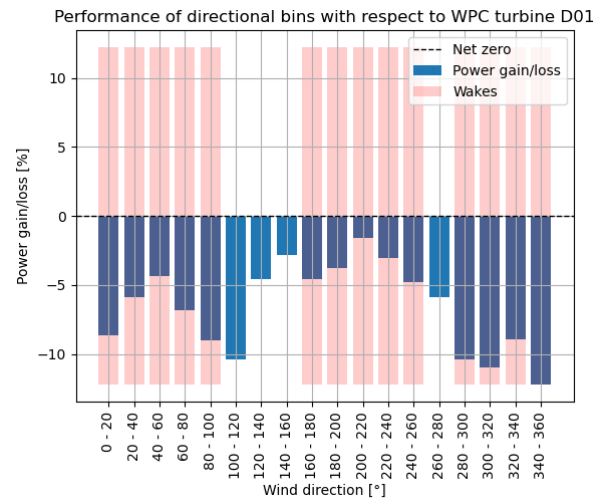


Figure 6.26: Relative power gain/loss per wind direction bin for turbine D01

An overview of the worst-performing unwaked sectors per turbine are presented in the table below. The descriptions of the sectors with a higher loss than 5% compared to WPC tell that the bad performing sectors for the gross part originate from sectors influenced by valleys or elevation run-ups in the orography. 3 out of 10 investigated sectors correspond to relatively flat terrain.

The site at the worst-performing sectors was further investigated using Google Earth ('Google Earth', n.d.). By projecting the sectors on the Google 3D model, it was possible to get insights into the site. Two examples of the Google Earth sector analysis are presented in appendix G.

Table 6.5: Worst performing unwaked sectors turbines western cluster.

Turbine	Sectors (unwaked) [°]	Performance gain/loss [%]	Site description
D05	300-320	-13	Left of valley. Trees at 200 m.
	280-300	-7	Valley run up towards turbine. Trees at 140 m.
	260-280	-7	Right of valley. Trees at 200 m.
D04	120-140	-10	Relatively flat between valleys. 700 m of forest towards turbine.
	160-180	-10	Valley run up towards turbine. Trees at 60 m.
	80-100	-8	Slope run up towards turbine. Trees at 60 m.
D03	280-300	-8	Relatively flat. Trees at 200 m.
	80-100	-6	Slope run up towards turbine. Trees at 380 m.
	100-120	-4	-
D01	100-120	-10	Valley run up (same as D03). Forestry at 90 m.
	260-280	-6	Slowly elevating. Trees at 250 m.
	120-140	-4	-

6.4. Cloud analysis

In this section, the cloud selections leading to conclusions regarding turbine performance of the western cluster are shown, including outputs of the cloud selection tool. First, the 10 minutes averaged cloud analysis results are discussed. Afterwards, the HF cloud analysis. The methods and tools used to get to the results of this chapter can be found in section 3.1.6 of the methodology chapter.

6.4.1. Results 10 minutes averaged cloud analysis

In this section, the results from the 10 minutes averaged data cloud analysis are discussed. The tables below contain information of all clouds selected, leading to results. Moreover, the state curves with binned 10 minutes averaged data of all cluster turbines are given. As explained in section 3.1.6, these state curve relations are inputs for the cloud selections.

Table 6.6: 10 minutes averaged clouds analysed western cluster.

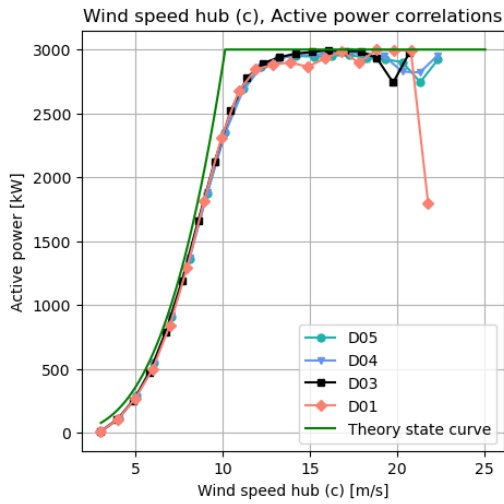
Cloud tag	X	Y	Turbine/ Cluster	Description
D01-10-174	Wind speed hub	Active power	Turbine	Scattering below power curve
D05-10d-662	Wind speed hub	Active power	Cluster	Unfiltered scattering below power curve
D05-10-760	Wind speed hub	Pitch	Cluster	Pitching between 3 and 5 m/s
D05-10d-138	Wind speed hub	Pitch	Cluster	Unfiltered low-wind speed pitching behaviour
D05-10-096	Rotor speed	Torque	Cluster	14 RPM: no clear vertical line
D05-10-794	Rotor speed	Torque	Cluster	6-8 RPM: below optimal torque control curve
D05-10-990	Rotor speed	Torque	Cluster	Performance optimal- c_p region

Table 6.7: Data collection of selected 10 minutes averaged clouds western cluster.

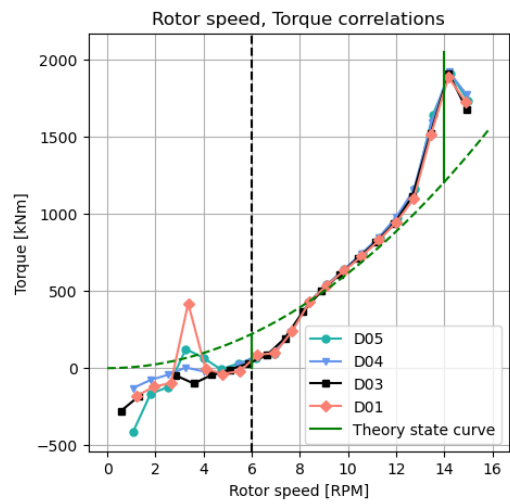
Cloud tag	Data points		Production		LPP		
		%	kWh	%	kWh	% LPPT ¹	% LPPC ²
D01-10-174	476	1.54	155346	2.64	60020	1.02	38.64
D05-10d-662	2158	4.09	292481	2.98	590788	6.02	201.99
D05-10-760	877	2.46	-339	0	13923	0.18	-
D05-10d-138	5515	10.46	-17163	0	54931	0.56	-
D05-10-096	13302	44.23	5620741	78.06	104162	1.45	1.85
D05-10-794	5738	16.08	71638	0.9	80199	1.01	111.95
D05-10-990	6101	20.29	829259	11.52	114721	1.59	13.83

¹LPPT: Lost Potential Power as a fraction of annual power production

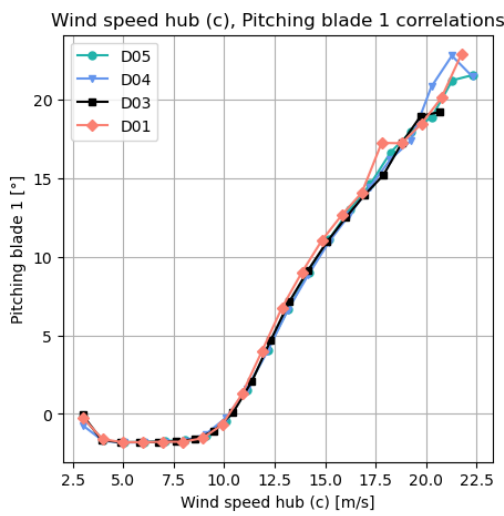
²LPPC: Lost Potential Power as a fraction of cloud power production



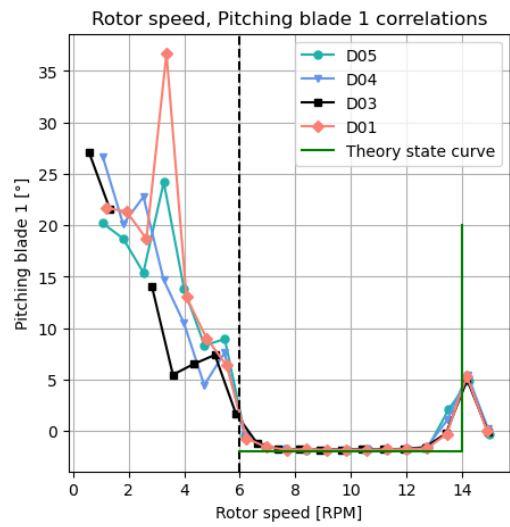
(a) Western cluster wind speed - active power average comparisons.



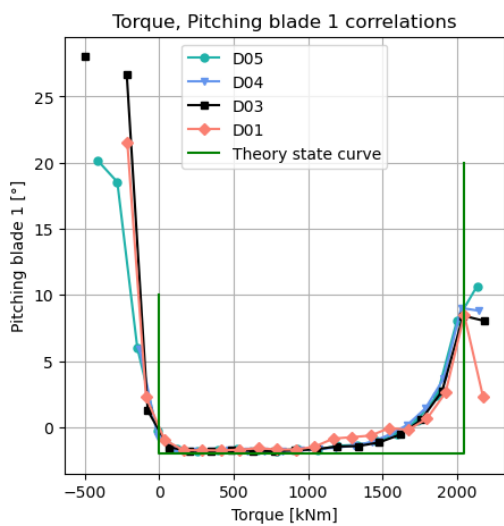
(b) Western cluster rotor speed - torque average comparisons.



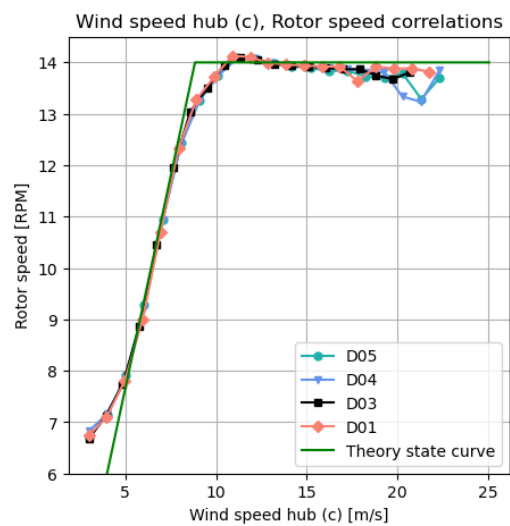
(c) Western cluster wind speed - pitching average comparisons.



(d) Western cluster rotor speed - pitching average comparisons.



(e) Western cluster pitching - torque average comparisons.



(f) Western cluster wind speed - rotor speed average comparisons.

Figure 6.27: Compared 10 minutes averaged data of western cluster turbines. Data is binned by the MOB.

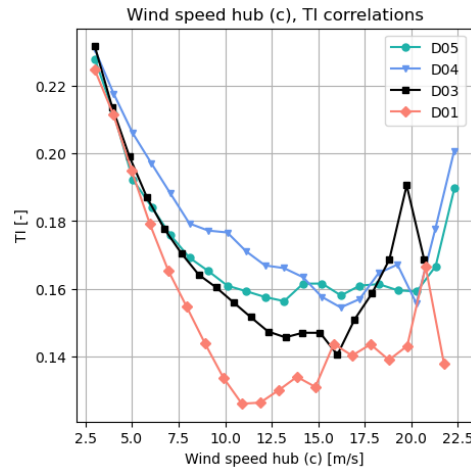


Figure 6.28: Wind speed, TI degradation turbines western cluster

As discussed in section 3.1.6, a comparison is made between the averaged binned state curves of the turbines and the theoretical expected turbine state curves. Observations from these comparisons are input for the cloud selection tool. From the overview of binned 10 minutes averaged state curves in figure 6.27, small deviations between especially turbine D01 and the other turbines can be observed. Turbine D01 seems to underperform compared to its neighbouring turbine in the rated region of the power curve, and seems to pitch more aggressively compared to the other turbines from the cluster.

Moreover, big differences between the TI variation over wind speed between the four turbines were observed (figure 6.28). This did not result in direct performance differences between the different turbines of the western cluster.

The following section examines low wind speed pitching behaviour, as well as torque behaviour within the transition region from optimal- c_p to full load. Additionally, the variation in turbine D01 performance in comparison to its neighbouring turbines is analysed. For the cluster-wide turbine behaviour, turbine D05 is used as it is closest to the MET mast.

Pitching at low wind speeds

At first, pitching behaviour below 5 m/s was investigated by selecting cloud D05-10-760. As the cut-in wind speed of SWT-3.0-113 turbines ranges from 3-5 m/s, the data was filtered with a minimum wind speed threshold of 3 m/s to include the cut-in region in the filtered data. Below 5 m/s the rotor speed is regulated between 6 and 8 RPM. Pitching in this low wind speed region is set at values of 10 and 30 degrees to control the rotor speed at a favourable rate prior to wind speeds increasing above 5 m/s.

As the rotor speed is very low / close to zero, the turbine experiences high angles of attack. By increasing pitching, the angle of attack is reduced to positively affect the aerodynamic properties of the blades. This can help in decreasing the duration of an idling period (Ebert & Wood, 1997). From the wind speed - rotor speed scatter plot of cloud D05-10-760 (figure 6.30), it was noticed for 6-8 RPM the rotor speed is not linearly related to the wind speed³. As the turbine is getting up to speed in this RPM region and is not producing at optimal power coefficient, it is expected not to show linear behaviour for this rotor speed region⁴. The low wind speed / rotor speed behaviour of cloud D05-10-794 makes 16.08% of total D05 filtered data, and production numbers can potentially be improved 1% (LPPT⁵) within this region.

From this analysed pitching behaviour, the added value of HF data became clear, as it gave insights into the control of the turbine (pitching at 10 and 30 degrees), where the 10 minutes averaged data did

³At the left of figure 6.30, figure 6.29 is shown. This figure shows the pitching behaviour corresponding to the selected cloud D05-10-794 in the torque curve.

⁴The selection of cloud D05-10-760 corresponded to the selected cloud D05-10-794. This cloud was selected to investigate the rotor speed behaviour from 6-8 RPM from the wind speed-rotor speed plots.

⁵LPPT: Lost Potential Power as a fraction of cloud power production
LPPT: Lost Potential Power as a fraction of annual power production

not show this behaviour. Moreover, the relative bad performance of this region was exposed with an LPPC value of over 100%, meaning the cloud underperforms 50% compared to WPC.

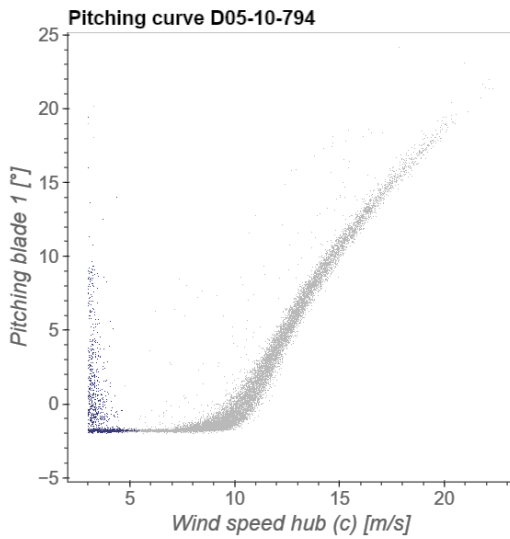


Figure 6.29: Pitching between 3-5 m/s (cloud D05-10-794).

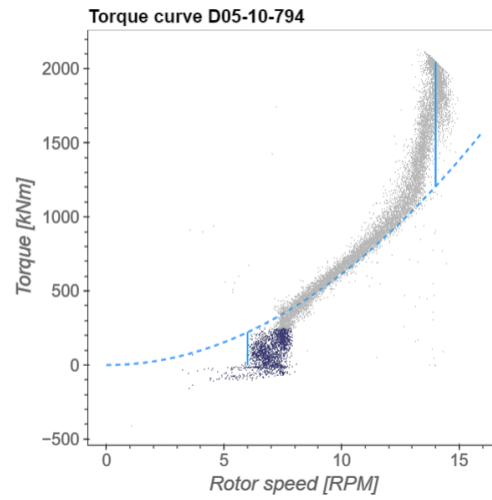


Figure 6.30: RPM below torque curve (cloud D05-10-794).

Torque curve: transition region between optimal- c_p and full load region

In addition, the torque development for high rotor speed regions was investigated with selected cloud D05-10-096 (figure 6.31). This selected region, deviating from the $Q \sim \omega_{rotor}^2$ curve, is a transition region between optimal turbine performance in partial load, and the full load region. This region corresponds to power curve data which is deviating from the optimal- c_p power curve (figure 2.11a). As the selected region contains two controller strategies, the selected cloud is built up of two sub clouds: one torque-controlled transition region, and one pitch-controlled full load region. The selected region is 44.23% of total data, and is showing behaviour not matching the WPC. The LPPC value of this cloud is 1.85%, meaning the energy production within this cloud is 1.85% less than warranted behaviour would deliver. A big contributor to this -1.85% is the region where the turbine blades are already pitching before rated rotor speeds are reached. Theoretically, this is not expected to happen, as the turbine is still in the optimal- c_p region below rated rotor speed. This region has a LPPC value of 6%.

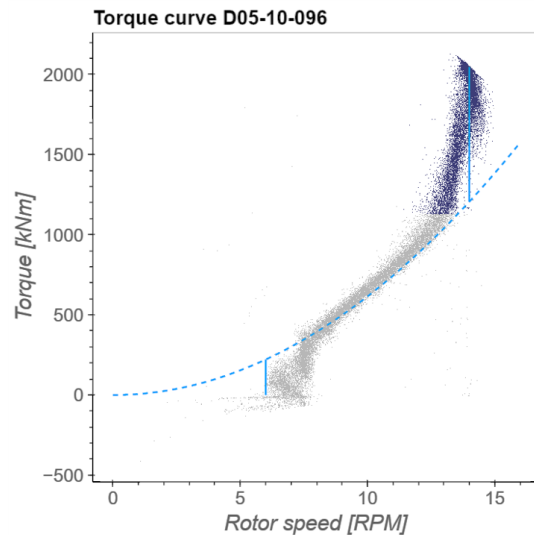


Figure 6.31: Torque behaviour above optimal- c_p region.

The deviating behaviour with an LPPT of 1.85% is a major part of underperformance of the western cluster turbines. Moreover, it is interesting to see this is originating from a cloud with the size of almost half of the data, all not matching the theoretical torque curve behaviour.

Other clouds analysed and shown in table 6.6 are selected from raw data. These raw clouds had the purpose of clarifying the filtered cloud behaviour. Therefore, these clouds will not be discussed further on in this thesis, as they did not lead to extra conclusions and did not light out other behaviour. They were included in tables 6.6 and 6.7 for completeness.

Turbine D01: scattering below power curve

To understand the deviating behaviour of turbine D01 within the state curve relations of figures 6.27a and 6.27c, scattering below the power curve of turbine D01 was selected as a cloud using the cloud selection tool produced for this analysis (D01-10-174). This scattering below the power curve corresponded to more aggressive pitching in the pitching data, which made the binned pitching curve of figure 2.11c deviate from the other turbines. The cloud originated from a time period of 11 days, where the turbine was curtailed to power values of around 500 and 1200 kW, while keeping rotor speed maximum for quickest inertia response to grid imbalances (Lio, 2018). It was concluded this cloud is curtailed data, which is not logged as curtailed data. It belonged to a small 3% of power production of the year 2020 for turbine D01, where 1% of potential power was lost. Note: the fact this cloud is not flagged as curtailments, can lead to misleading performance calculations when including this data.

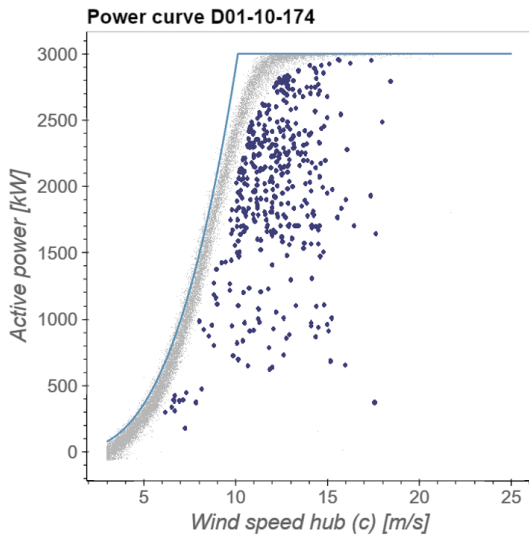


Figure 6.32: Scattering below the power curve (cloud D01-10-174).

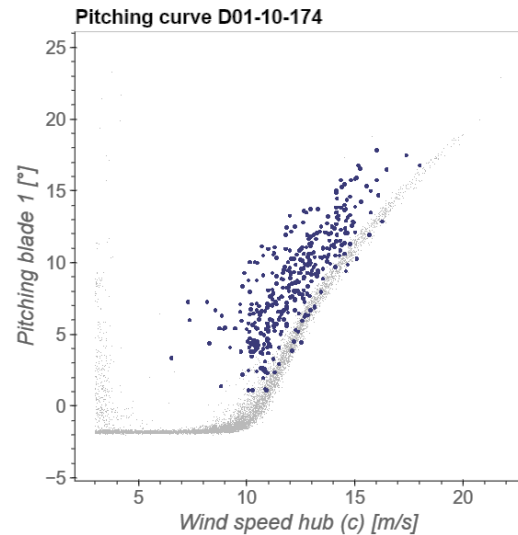


Figure 6.33: Aggressive pitching (cloud D01-10-174).

In the figure below, a zoomed in active power signal of the selected cloud is given. It shows fluctuation of active power which occurs for a number of hours. For this time period, wind speeds were above rated and there was no clear reason for the turbine to produce below rated power. As this behaviour is a part of the filtered data, it is important to get understanding in the fact that this curtailed behaviour is not flagged as curtailments in the 10 minutes averaged curtailment tag. This is passed on to the OEM. The difference from this cloud compared to the others, it that the data originates from a single period. The other clouds from this section is turbine behaviour, caused by the turbine controller. The selected cloud D01-10-174 does not give information about turbine performance.

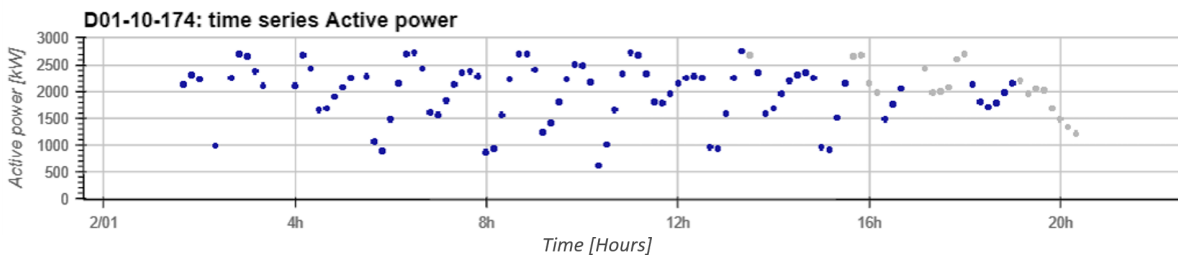


Figure 6.34: Time series of power output during curtailed period of cloud D01-10-174.

6.4.2. Results HF cloud analysis

HF clouds are selected for reasons as explained in section 3.1.6. In the tables below, an overview of the selected clouds is given, together with relevant cloud information and data.

In this section, reasons for selecting the tabulated clouds as well as results obtained from the multiple cloud analyses are addressed. Differences between HF power curves, HF pitching behaviour and rotor speed behaviour will be discussed.

Table 6.8: HF clouds analysed western cluster.

Cloud tag	X	Y	Turbine/ Cluster	Description
D05-HF-407	Wind speed hub	Active power	Cluster	General scattering below HF power curves
D05-HF-936	Wind speed hub	Active power	Turbine	Curtailed cloud around 1700 kW
D05-HF-099	Wind speed hub	Active power	Cluster	Curtailed down to 2200 kW
D05-HF-022	Wind speed hub	Pitch	Turbine	Pitching above 30 degrees
D05-HF-392	Wind speed hub	Pitch	Cluster	Pitching behaviour below 7 m/s
D04-HF-981	Rotor speed	Pitch	Cluster	Pitching around 10 deg above 5 m/s wind speeds
D05-HF-949	Wind speed hub	Pitch	Cluster	Pitching values up to 30 degrees at non-rated rotor speeds
D01-HF-882	Wind speed hub	Active power	Turbine	Curtailed down to 500 kW.
D05-HF-954	Rotor speed	Torque	Cluster	Rotor speed higher than maximum of 14 RPM.

Table 6.9: Data collection of selected HF clouds western cluster.

Cloud tag	Data points		Production		LPP		
	Points	%	kWh	%	kWh	% LPPT ⁶	% LPPC ⁷
D05-HF-407	16025	0.13	12397	0.16	7971	0.11	64.3
D05-HF-936	4375	0.03	2495	0.03	1818	0.02	72.85
D05-HF-099	8834	0.07	8727	0.12	2248	0.03	25.76
D05-HF-022	504	0	15	0	352	0	2269.34
D05-HF-392	198321	1.55	1180	0.02	9347	0.12	791.92
D04-HF-981	4120	0.04	1552	0.03	-754	-0.01	-48.57
D05-HF-949	3007	0.02	2486	0.03	1062	0.01	42.74
D01-HF-882	19428	0.17	6025	0.11	11748	0.21	194.97
D05-HF-954	327629	3	430298	6.27	-23914	-0.35	-5.56

Differences in HF power curves between turbines

One of the observations looking at the 10 minutes averaged state curves was apparent underperformance in the rated region of turbine D01 compared to the other turbines of the cluster. This was caused by scattering below the power curve of D01, which was not flagged as curtailment. When looking at the HF wind speed - active power behaviour of the four western cluster turbines, clear scattering can be observed at all four turbines. For all turbines, curtailments until 2200 kW can be observed (cloud D05-HF-099), as well as curtailments down to 1700 kW (D05-HF-936). Only for turbine D01, curtailments get down to 600 kW (D01-HF-882). To give an insight in the scattering behaviour of the four turbines, an overview of the different power curve scattering is given below. The colouring in the plots is used to show the number of points within a Datashader raster point⁸.

⁶LPPT: Lost Potential Power as a fraction of annual power production

⁷LPPC: Lost Potential Power as a fraction of cloud power production

⁸More information about Datashader is given in section 3.1.6

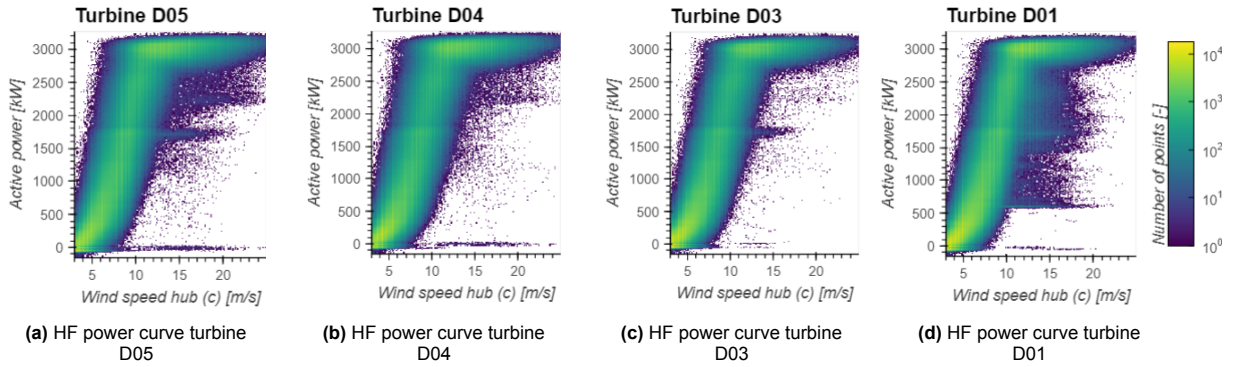


Figure 6.35: HF power curve scatter plots of the western cluster turbines.

1. Curtailments 2200 kW.

These curtailments occur at moments where wind speeds do not ramp down or up. Active power output is fluctuating around 3000 kW. At certain moments in time, power drops down to 2200 kW. This happens for periods of multiple seconds to minutes. Below, an example of a drop to 2200 kW is shown that lasts for 30-40 seconds.

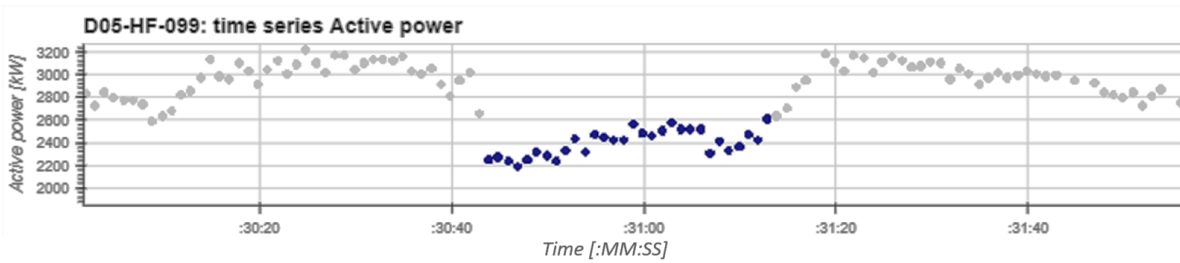


Figure 6.36: Time series of power output during curtailed period of cloud D05-HF-099.

2. Curtailments 1600 kW.

This cloud of scattering is very clear visible for turbines D05, D03 and D01. Part of the cloud can be explained by small peaks of wind speed data during a period of below rated average wind speed. This can also be explained as more extreme type 1 scattering.

The major part of the selected cloud consists of data belonging to a time period of two hours, where the turbine has been curtailed to 1600 kW. This can be seen from the pitching behaviour as well. Moreover, RPM is kept at maximum in this period, which was also seen in the 10 minutes averaged scattering below the D01 power curve. Below, the power output and pitching for the two-hour period described are shown. The contribution of data outside this two-hour period is negligible, as two hours of HF data consists of more data than the cloud contains.

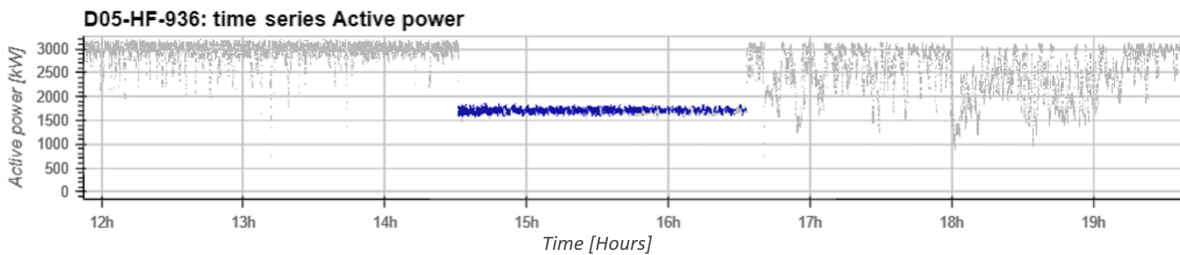


Figure 6.37: Time series of power output during two-hour period of cloud D05-HF-936.

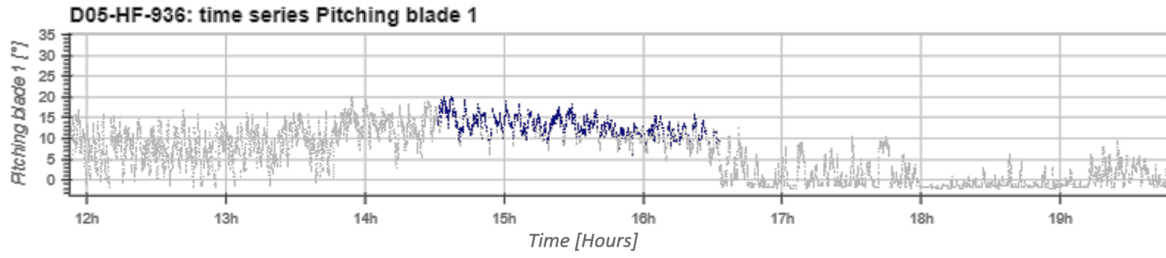


Figure 6.38: Time series of pitching during two-hour period of cloud D05-HF-936.

3. Curtailments 600 kW.

This curtailment level only occurred at turbine D01. The cloud selected to investigate this behaviour is cloud D01-HF-882. This cloud consists of scattering below 1600 kW down to 600 kW. It corresponds to the scattering below the 10 minutes averaged power curve of D01. This was a period of 11 days in which the turbine was curtailed down to 600 kW. Rotor speed was at its rated value during this period. Other points within this cloud are single data points where during a period of expected behaviour the wind speed ramped up or down significantly, leading to some scattering deviating more extremely from the power curve than the general HF scattering. Below, a close up of 4 hours of this 11 days period is shown. The time series below clearly show the oscillating power output in combination with high pitching.

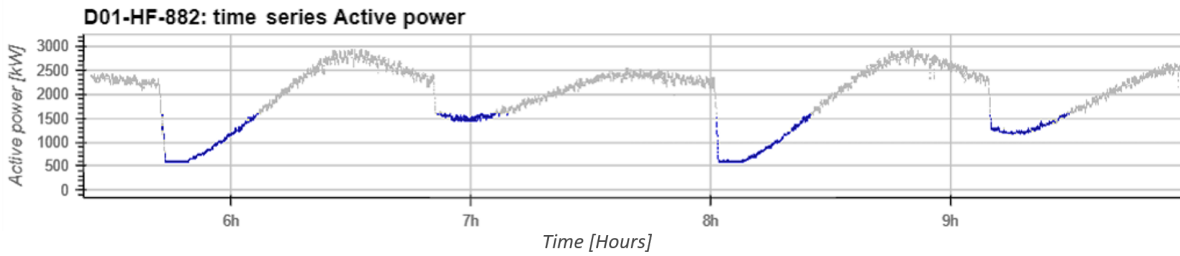


Figure 6.39: 4 hour active power close up of 11-day period cloud D05-HF-882.

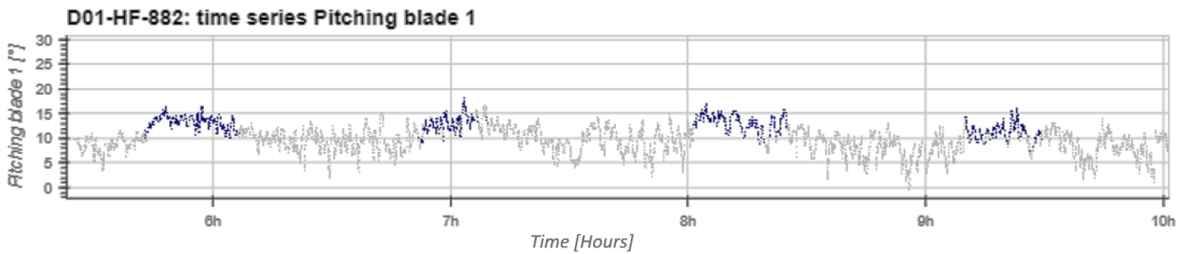


Figure 6.40: 4 hour pitching close up of 11 day period cloud D05-HF-882.

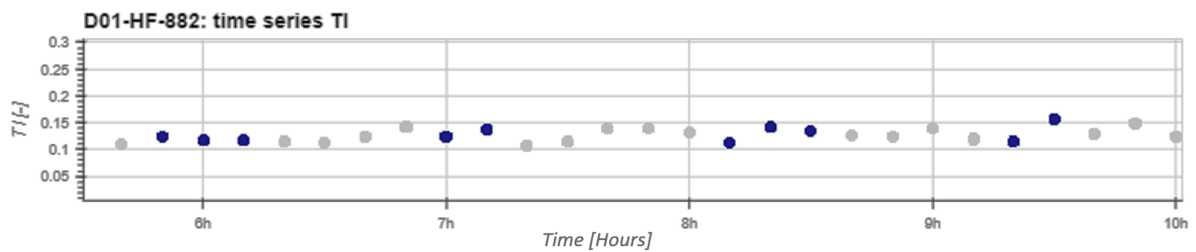


Figure 6.41: 4 hour TI close up of 11 day period cloud D05-HF-882.

A HF cloud like the 1600 kW curtailment cloud D05-HF-936 shows the strengths, and at the same time the weakness, of the HF data. In HF data, behaviour lasting for a short duration is visible. Short

deviations from expected performance are therefore possible to detect and analyse. On the other hand, these short time periods have a big influence on the scatter plots, but do not contribute substantially to the yearly performance of the turbine.

When comparing figure 6.39 to figure 6.32, the value of HF data for understanding turbine behaviour can be seen. In the HF active power time series of the 11 day period, the levels of curtailment are better visible, as well as the power recovery after curtailment.

Just like for the 10 minutes averaged power curve scattering, the selected clouds do not give information about turbine underperformance, but give insights in the wrong logging of curtailed data. For the HF data, a buffer of 30 minutes was applied to each filtered curtailment period to not include curtailments in HF filtered data. Even with this buffer, curtailments are part of the filtered data. Cumulatively, it makes up less than 0.3% of total 2020 data.

Pitching > 30°

Another parameter for which there was unexpected behaviour, observed within the HF data, is pitching. For a filtered data set, one would expect pitching not to exceed the pitching applied near cut out wind speed to keep the turbine at P_{rated} . However, different behaviour is observed. Scattering from 30 degrees of pitching up to feathering position is present for the whole wind speed range of turbine D05. Turbine D04 shows less of this same scattering. Turbine D03 and D01 have fewer data occurring in the 30-90 pitching range. D01 still has some pitching above 30 degrees, occurring from 15-20 m/s.

To understand this high pitching still visible in the HF data sets, the scattering above 30 degrees of pitching for turbine D05 was selected as a cloud (D05-HF-022). This because for turbine D05 there is most data available in this region. Below, respectively, the pitching and power curves with the selected cloud is given.

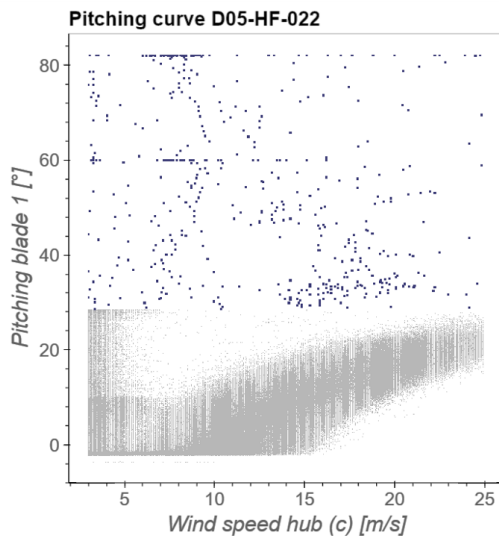


Figure 6.42: Pitching >30° at turbine D05.

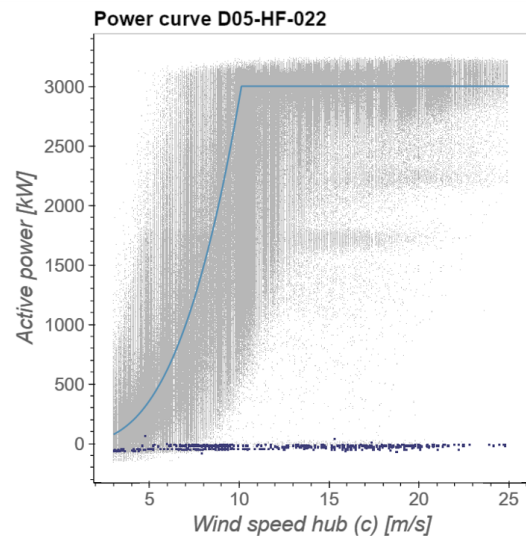


Figure 6.43: Active power output for pitching >30° at turbine D05.

This pitching behaviour corresponds to handful moments in time where turbines are not producing power. One example is a moment right after a filtered out period. Important to note, as said before, is the margin applied to any curtailment filtering action: Right and left of each filtered out period, 30 minutes extra were extracted from the data. Therefore, to still see this pitching behaviour at the edge of filtered out periods, is unexpected. Another instance of this pitching to feathering that was found, was for a moment of high TI. This caused the control system to pitch the turbine

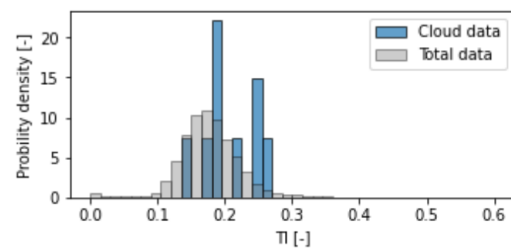


Figure 6.44: TI distribution of cloud D05-HF-022.

blades into the wind to reduce loads.

The pitching behaviour selected is negligible when looking at its contribution to the total data set and power production, as can be seen from the data in table 6.9. Rounded off, it does not contribute to 2020 data (0%).

Pitching < 5 m/s wind speeds

The pitching behaviour in the cut-in region was selected from the 10 minutes averaged data by cloud D05-10-794. This cut-in region is observed in more detail in the HF spectrum by cloud D05-HF-392. From the HF data, it is possible to see what happens to the turbine behaviour for the two distinct pitching settings of 10 and 30 degrees. At a pitching setting of 10 degrees, the turbine rotor speed is regulated at 5 RPM. When the pitching is increased to 30 degrees, the rotor speed is minimised to 0 RPM. These two pitching settings are only used for wind speeds between 3 and 5 m/s. Pitch in this region is increased to decrease the high angle of attack, as also explained in section 6.4.1.

Lost potential of this selected HF cloud is 0.12% of the annual power output. As it is the cut-in region of the turbine, it is expected to not instantaneously perform at maximum power coefficient, as the turbine needs time to get up to speed. However, as also shown by cloud D05-10-794, the turbine does under perform heavily in this region compared to WPC.

10° pitching above cut-in wind speed

As can be seen for all turbines, within HF data, pitching unexpectedly still occurs above the cut-in wind speed of 5 m/s. At turbines D05 and D04, this is denser scattering than for turbines D03 and D01. To analyse this unexpected pitching behaviour, cloud D04-HF-981 is selected. The cloud boundaries are chosen with the attempt to avoid selecting general HF scattering originating from the right curved part of the pitching relation. Below, the selected cloud is shown within the pitching and power curve.

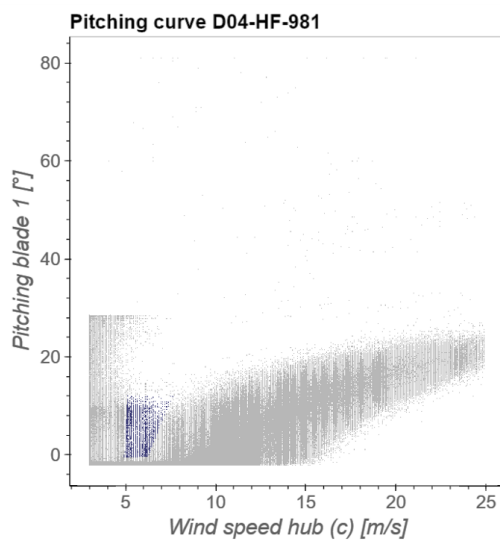


Figure 6.45: Cloud selection of pitching above cut-in wind speed.

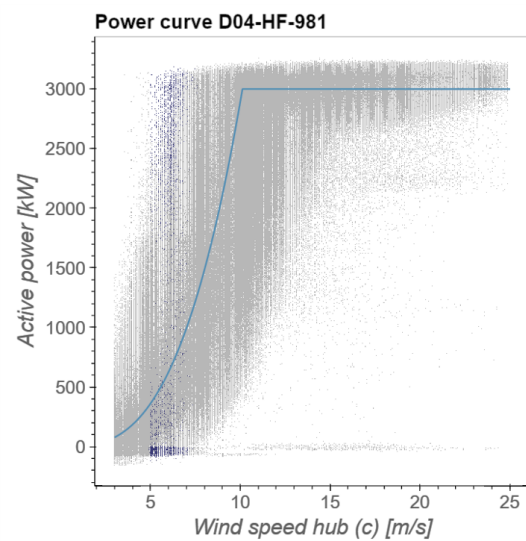


Figure 6.46: Active power output for pitching above cut-in wind speed.

In figure 6.46, the cloud selected contains both zero active power data, and active power data close to rated power. This, while the selected wind speed range is above cut-in and 4 m/s below rated wind speed.

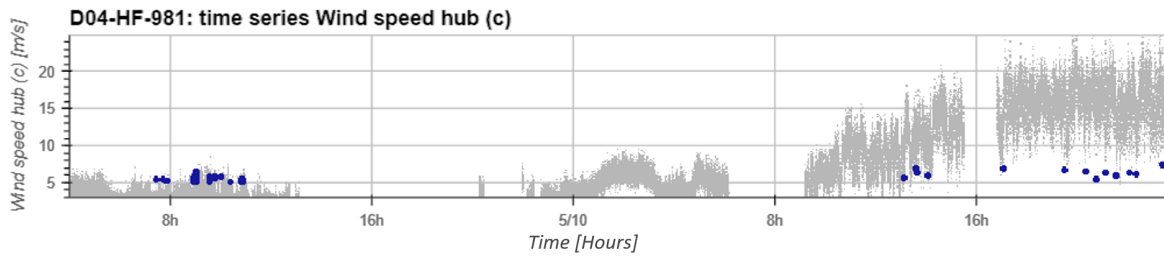


Figure 6.47: Cloud D04-HF-981: time series of wind speeds.

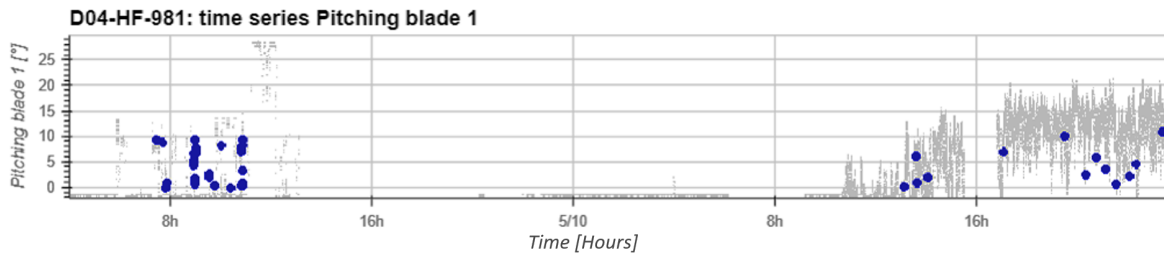


Figure 6.48: Cloud D04-HF-981: time series of pitching.

Within the time series of figure 6.47 above, two different wind speed conditions are shown (in grey). Both include data of the cloud D04-HF-981 (blue dots). The left part of the grey time series is a period of low average wind speed, between 3 and 5 m/s. The right part of the grey time series contains above rated average wind speed data⁹. In both situations, the pitching behaves as shown within other clouds (figure 6.48). For low wind speeds, pitching is regulated between 0, 10 or 30 degrees. For the above-rated wind speeds, pitching reaches values up and until 20 degrees to keep the rotor speed at the desired level.

Although for both situations the mean wind speeds are respectively below 5 m/s and above rated, there exist wind speed measurements with a deviation from the mean. For those deviations, the pitching strategy is not changed, as it are singular occurrences. Some of those singular occurrences are shown by the blue dots within the two figures 6.47 and 6.48, and are part of the cloud D04-HF-981. For this reason, pitching can be seen in the HF pitching curve for values of between cut-in and rated wind speed (cloud D04-HF-981).

The selection of this cloud shows a downside of HF data. Due to the turbine (correctly) not responding to a single wind speed up/down ramp causing a temporary deviation from the average wind speed, pitching is seen in the wind speed region between cut in and rated wind speed. This turbine robust control in combination with volatility of the wind cause scattering in unexpected parts of state curves, making the HF plots less clear to read and analyse.

High rotor speed

From the six theoretical state curve relations of figure 2.11, three contain rotor speed on one of the axes. The HF rotor speed signal is a discrete signal. It is logged with resolution of 0.5 RPM. This discrete logging makes cloud selection more difficult, as a scatter plot is a combination of vertical lines. Density is more difficult to extract from the plot. Moreover, this discrete signal does not represent the rotor speed in the most realistic way, as rotor speed is a continuous parameter.

Despite the downsides of the rotor speed signals, useful information was still extracted from the state curves, including rotor speed. Cloud D05-HF-954 was selected to understand occurrences of rotor speed exceeding the maximum turbine rotor speed. This cloud originates from wind speed ramp ups, which temporarily increases the rotor speed. The turbine responds by increasing pitching and driving the rotor speed back to its desired value of 14 RPM.

⁹In between the two described conditions, wind speed is lower than 3 m/s. Data below 3 m/s has been filtered out during the filtering process, as these are no conditions within which the PYC turbines are operational.

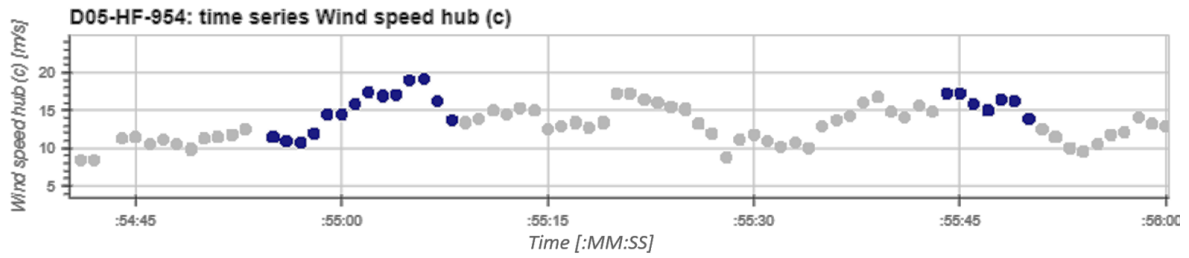


Figure 6.49: Time series of wind speed ramping up.

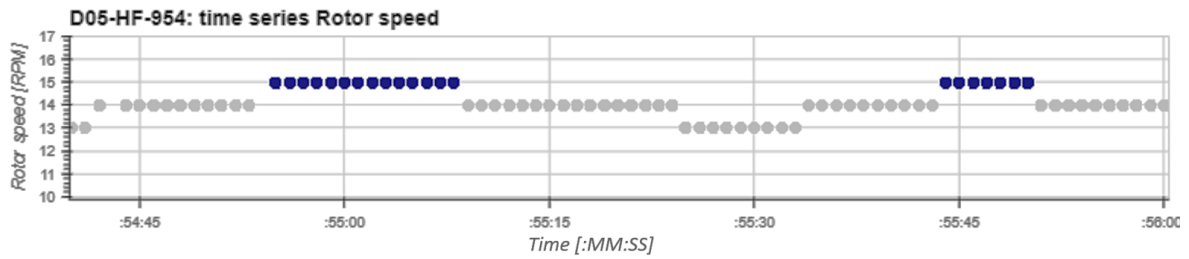


Figure 6.50: Time series of rotor speed response to wind speed ramp ups.

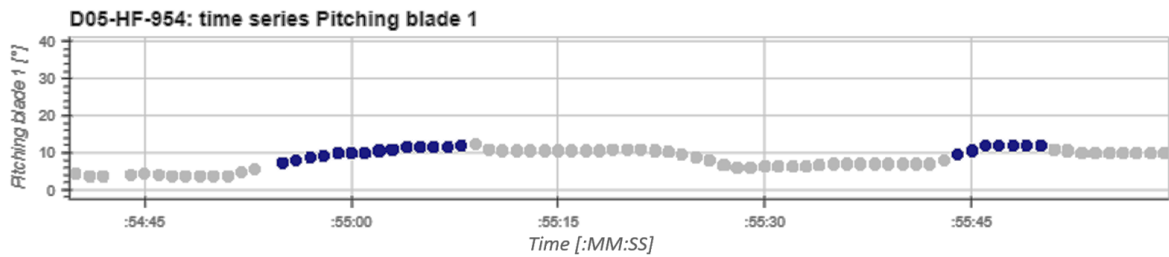


Figure 6.51: Time series of pitching response to rotor speed increase.

This increase in rotor speeds results in a temporary increase in active power. This behaviour does not negatively affect the turbine production. However, if wind speed ramps at these high wind speeds occur frequently, it can have a long term effect on the lifetime of the turbines. Rotor speeds above rated make up 3% of the total data.

7

Results eastern cluster

For the eastern cluster, the same procedures were followed as for the western cluster. In this chapter, the deviations in the results compared to the other clusters are discussed.

7.1. Wind resource

To comprehend the site conditions not influenced by turbine rotor movements or wake effects, MET mast data is used, as was also done for the western cluster. The conditions for the eastern cluster are compared with the site specific conditions at the western cluster, as differences in site conditions are important to understand when analysing turbine performance. The distributions and correlations studied at the western MET mast were reproduced for the eastern cluster.

Below, the wind speed and wind directional distributions of the eastern cluster are given. For fitting a Weibull distribution, the curve-fit class from Scipy was used, as the Weibull fitting function gave a local optimum as a solution, which did not match the shape of the wind speed distribution.¹

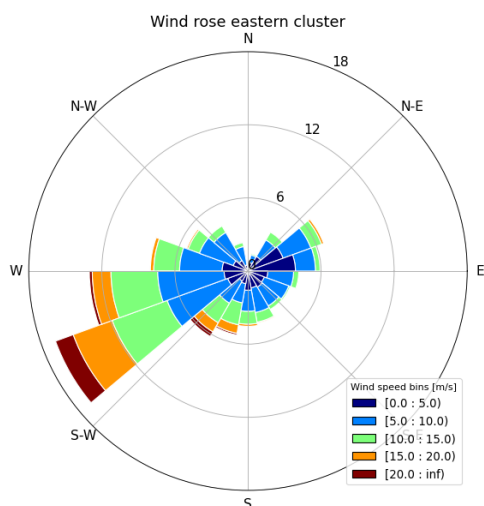


Figure 7.1: Wind rose eastern cluster

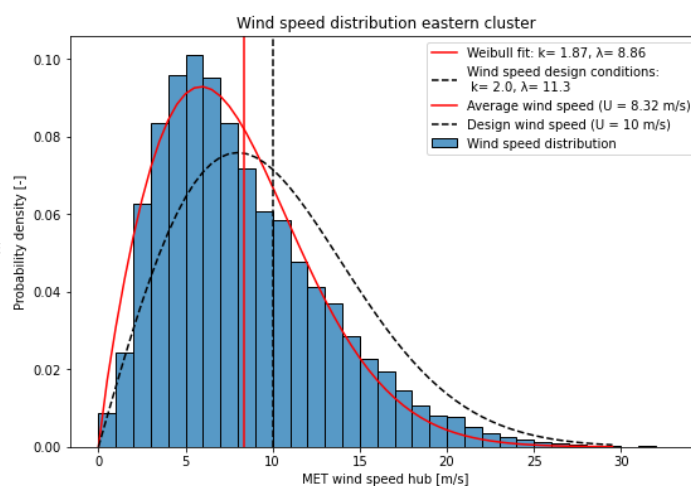


Figure 7.2: Wind speed histogram eastern cluster

Just as for the western cluster, other wind characteristics of the site were investigated: directional veer, wind shear and turbulence intensity.

Characteristic wind veer relations are visualised below. For the same reasons are described in 6.1, the northerly direction ± 60 degrees are filtered out. Differences that can be spotted from the western cluster is the veer at high wind speeds: for the western cluster the veer runs towards $0.25^\circ/\text{m}$, whereas for

¹The found shape parameter k was smaller than 1, which gives an exponential distribution.

the eastern cluster it runs more to around $0.06^\circ/\text{m}$. This difference of approximately $0.2^\circ/\text{m}$ between the western and eastern high-speed veer measurements is also the difference between the mean veer of both clusters. The average spread and development of standard deviation over wind speed is comparable between both clusters.

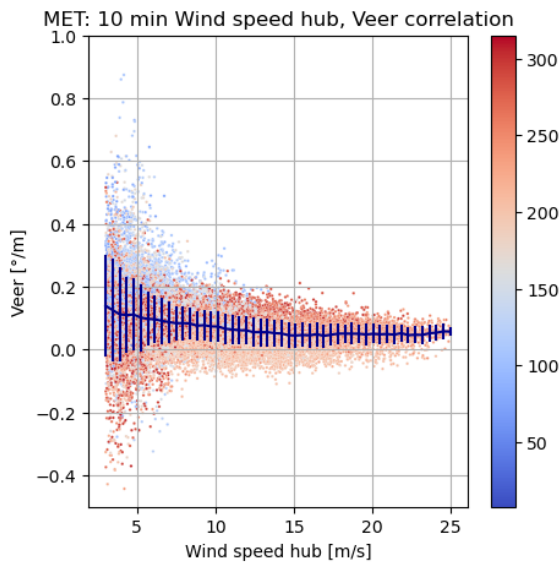


Figure 7.3: 10 minutes averaged MET mast wind speed veer correlation.

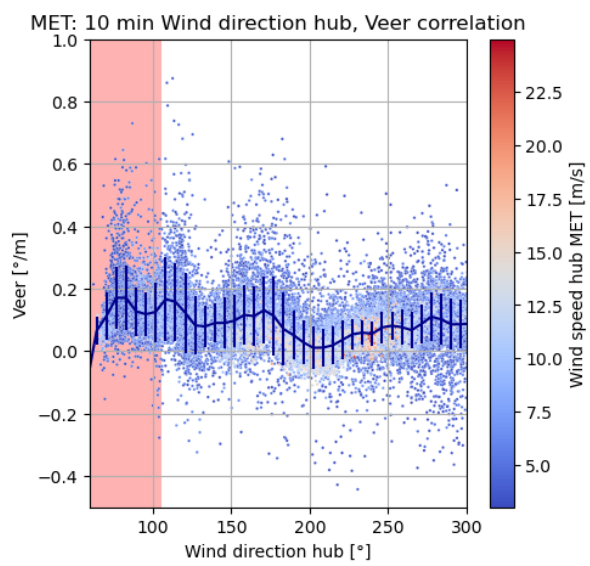


Figure 7.4: 10 minutes averaged MET mast wind direction veer correlation.

When comparing wind shear measured at the western and eastern cluster, the wind shear development with wind speed and the spread are comparable. However, on average, the western cluster measures higher wind shear than the eastern cluster. The difference in wind shear factor on average is 0.2. Just like the wind directional veer, the shear is more extreme for the western cluster.

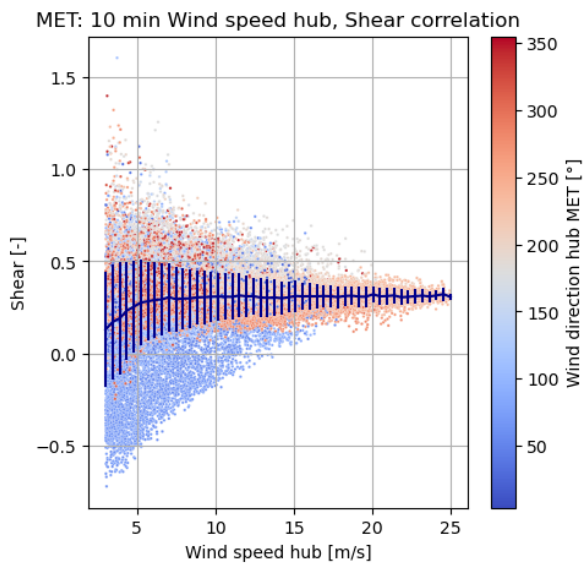


Figure 7.5: 10 minutes averaged MET mast wind speed shear correlation.

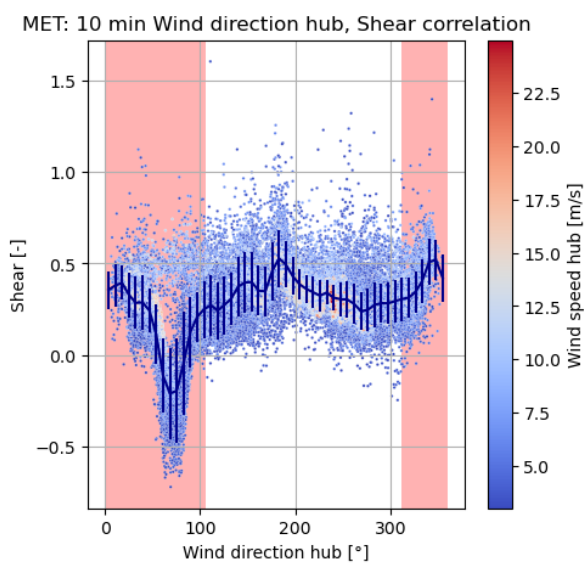


Figure 7.6: 10 minutes averaged MET mast wind direction shear correlation.

At last, the TI experienced by the MET mast. For both clusters, extreme TI values are experienced within the waked sectors of the MET mast. This behaviour is measured for low wind speeds. Whereas for the eastern cluster the TI is more stable for the unwaked sectors, for the western cluster higher values of TI are measured around a sector of 200-230 degrees.

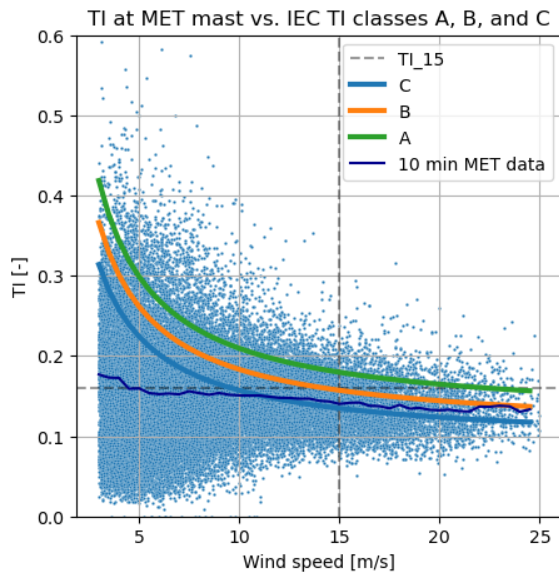


Figure 7.7: 10 minutes averaged MET mast wind speed TI correlation.

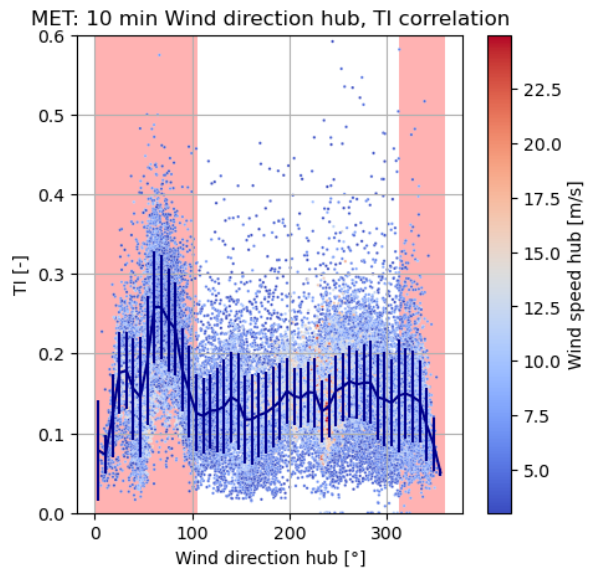


Figure 7.8: 10 minutes averaged MET mast wind direction TI correlation.

7.1.1. Influence of site on turbine performance L06

The effect of wind directional veer and wind shear on the turbine performance of the closest MET mast is examined by comparing measurements from MET02 and turbine L06. The distance from turbine L06 to the MET mast is 2D, and therefore within the 2-4D margin. At the eastern cluster, average veer and shear values are lower when compared to the western cluster.

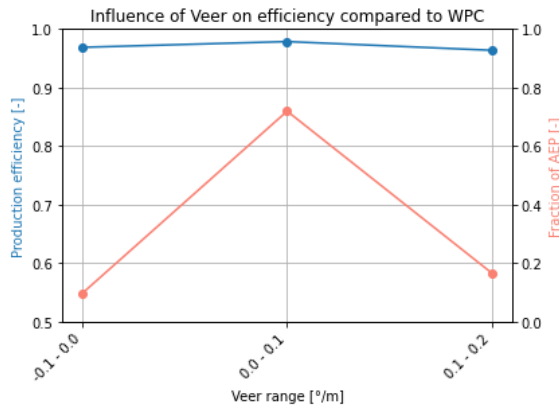


Figure 7.9: Influence of veer on L06 active power signal.

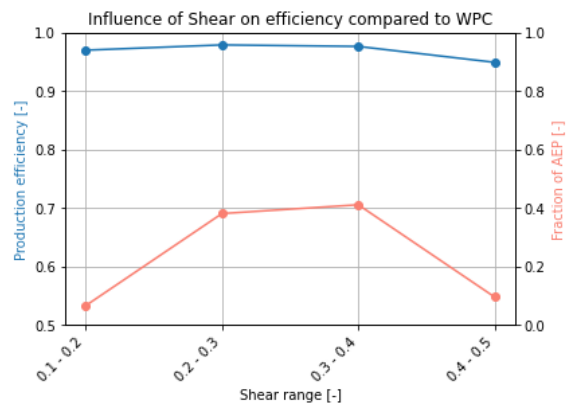


Figure 7.10: Influence of shear on L06 active power signal.

Image 7.9 and 7.10 show different relations compared to the western cluster. Wind veer and shear both have a negative, but very small, influence on power production. For the western cluster, when comparing the same veer and shear bins, both parameters had a positive effect on power production.

From the above figures, it can be concluded the veer and shear measured at the eastern cluster compared to the western cluster do not have an influence on the efficiency numbers of tables 7.3. In fact, from comparing figures 6.10 and 6.11 with figures 7.9 and 7.10, it can be concluded the lower veer and shear at the eastern cluster compared to the western cluster has a negative effect on production of the eastern cluster.

An important note is that these conclusions are based on the active power signal of one turbine, and veer/shear measurements of one MET mast. Influence of these parameters further on in the turbine clusters is not possible to determine.

7.2. Correlations MET mast and turbines

The correlations regarding wind speed and direction between the MET mast and turbines of the eastern cluster showed different results. Wind direction as well as nacelle direction correlations seemed as reliable for the eastern cluster as for the western cluster. Wind speed correlations, however, showed relations deviating from a slope of 1. This was expected, as the height differences between the individual turbines of the eastern cluster are larger compared to the relatively flat western cluster. According to equation 2.8, wind speeds at different heights are related with a slope deviating from 1, depending on the height difference.

As can be seen in the table below, correlations with intercepts were introduced for the wind speed correlations, on top of the correlations without intercept. This was added compared to table 6.1 as wind speed correlations were less strong, and introducing an intercept could potentially give answers regarding those deviating correlations.

Using equation 2.8, the expected slopes of the wind speed correlations between different cluster units can be estimated. For these estimations, a roughness length of 0.1-1 was assumed (table 2.5). When the eventual slope is within the margin at the bottom of table 7.2, is assumed trustable. From turbine L03 onwards, the wind speed correlations do not meet the pre-set margins. Therefore, only turbine L06 and L05 will be further analysed.

Table 7.1: Overview of correlations for the eastern cluster.

Correlation	Intercept	Property	MET-L06	L06-L05	L05-L03	L03-L04	L04-L01
Wind speed	No	Slope	0.99	0.96	0.91	1.07	0.88
	No	R^2 score	0.99	0.95	0.91	0.94	0.86
	Yes	Slope	1	0.94	0.84	1.04	0.87
	Yes	Intercept	-0.03	0.22	0.8	0.31	0.1
Wind dir	Yes	R^2 score	0.99	0.95	0.92	0.94	0.86
	No	Slope	0.98	1	0.97	1	0.99
	No	R^2 score	1	1	1	1	1
	Yes	Slope	1.01	1	1	0.99	0.97
Nac dir	Yes	Intercept	-6.28	1.34	-7.79	1.3	4.57
	Yes	R^2 score	1	1	1	1	1
	No	Slope	-	1	0.97	1	0.99
	No	R^2 score	-	1	1	1	0.99
	Yes	Slope	-	1	1	0.99	0.98
	Yes	Intercept	-	1.15	-7.77	1.31	4.02
	Yes	R^2	-	1	1	1	1

Table 7.2: Expected slopes for wind speed correlations eastern cluster.

	MET-L06	L06-L05	L05-L03	L03-L04	L04-D01
Height difference [m]	-4	3.2	-7.1	8.2	-55
Expected slope ($\alpha = 0.1$)	0.99	1	0.99	1.01	0.86
Expected slope ($\alpha = 1$)	0.99	1.01	0.98	1.02	0.79
Average	0.99	1	0.98	1.02	0.82
Accepted slopes (5% deviation)	0.95-1.05	0.96-1.06	0.95-1.05	0.97-1.07	0.78-0.87

7.3. Overall & directional performance analysis

In this section, the results from the performance analysis are shown. This analysis is carried out in parallel to the cloud analysis (described in the next section). The methods to come to the results presented in this section are represented in chapter 3.

7.3.1. Overall turbine performance

Below, the binned power curves of the eastern cluster turbines are displayed for the full wind direction spectrum. In blue, the power curve for the full wind direction spectrum is shown. In yellow, the power curve for only unwaked sectors is given. These power curves are inputs to equations 3.3 and 3.4 as described in the methodology (chapter 3). Results from those calculations are given in table 7.3.

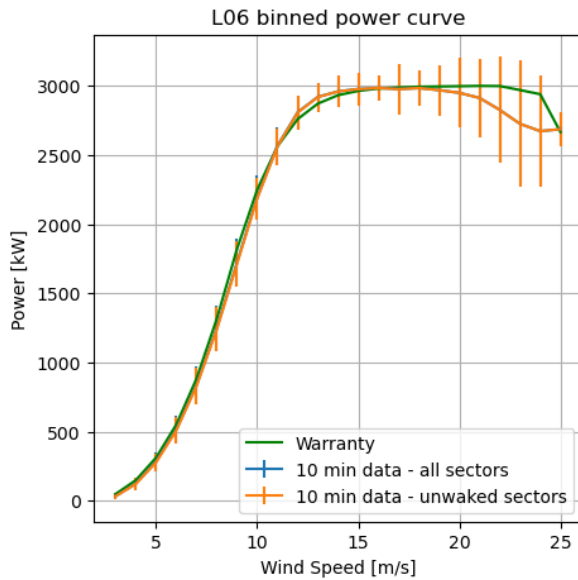


Figure 7.11: Total power curve turbine L06

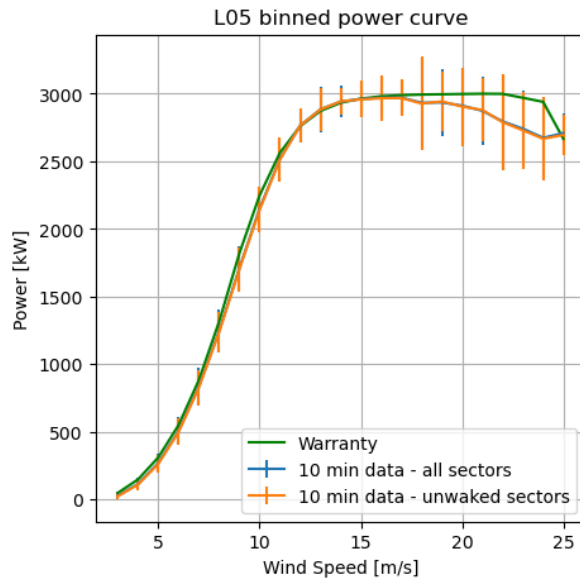


Figure 7.12: Total power curve turbine L05

Table 7.3: Efficiencies compared to WPC for turbines eastern cluster

	Based on	L06	L05
$\eta_{power,1}$	Warranted power curve	0.98	0.96
$\eta_{power,2}$	Potential power signal	0.97	0.96

7.3.2. Directional turbine performance

In this section, the turbine performance per directional sector is calculated and visualised. The wind direction spectrum is divided into bins of 20 degrees for this purpose. Below, the power curves for the 18 directional sectors are displayed.

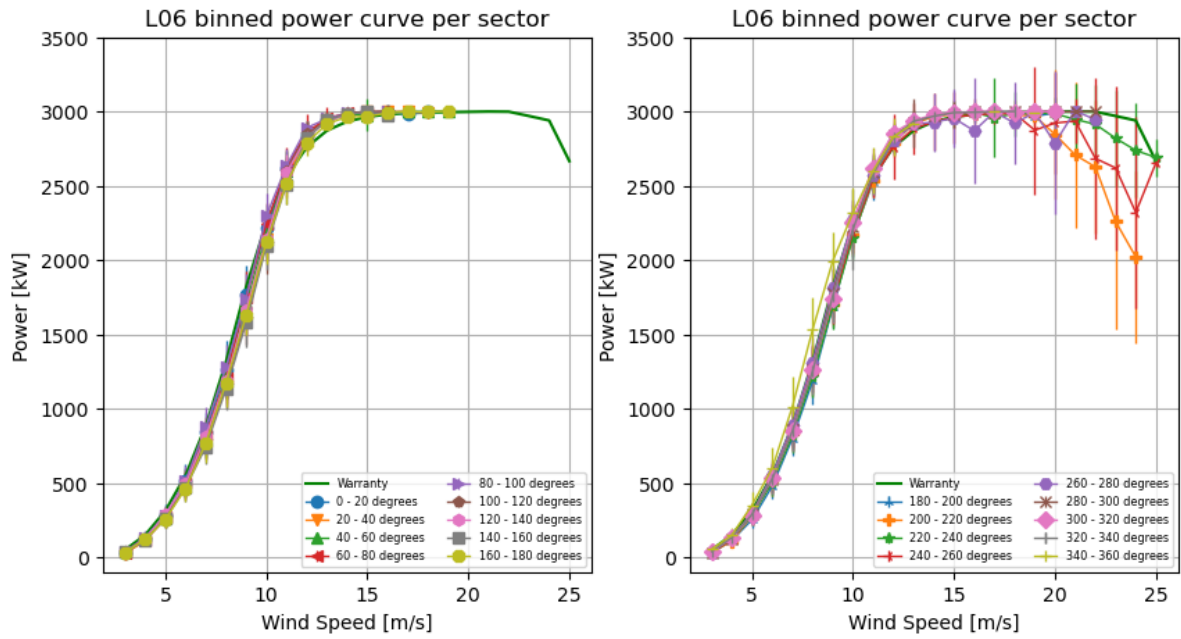


Figure 7.13: Binned power curves turbine L06

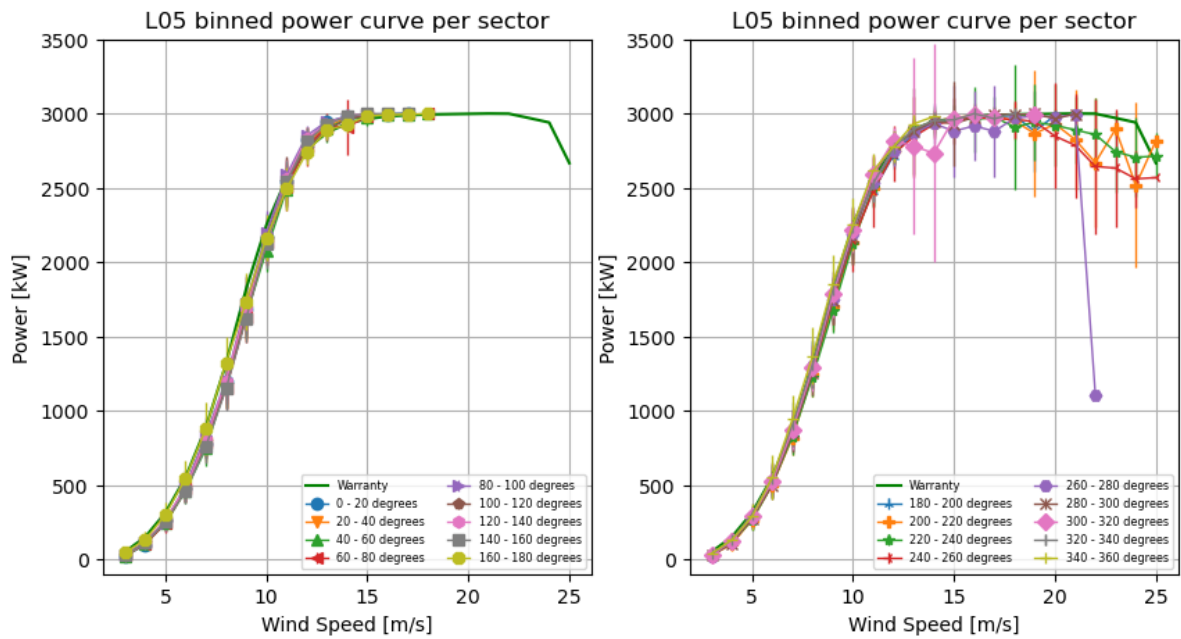


Figure 7.14: Binned power curves turbine L05

For each power curve per sector, the performance compared to the warranted power curve can be calculated by using equation 3.5. This relative performance per sector is given by a percentage in the diagrams below.

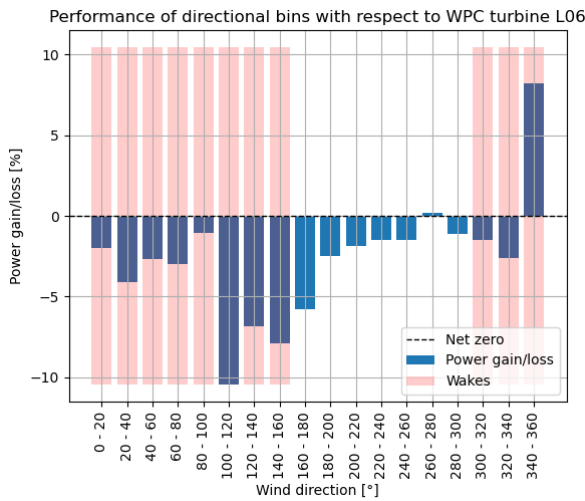


Figure 7.15: Relative power gain/loss per wind direction bin for turbine L06

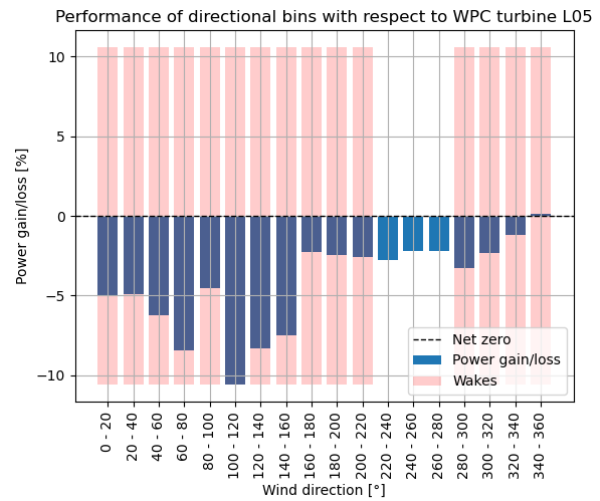


Figure 7.16: Relative power gain/loss per wind direction bin for turbine L05

The three worst-performing unawaked sectors from the figure above are extracted and tabulated in the figure below. The sectors with a negative mismatch to WPC of more than 5% are selected for further site investigation, resulting in a short site description in the last column of the table. For the eastern cluster, only one sector performs worse than 5% compared to WPC. This is a sector with the ridge of a hill in direction towards the turbine. This sector is also shown in figure G.3 in appendix G.

Table 7.4: Worst performing unawaked sectors turbines eastern cluster.

Turbine	Sectors (unwaked) [°]	Performance gain/loss [%]	Site description
L06	160-180	-6	Ridge in length direction towards the turbine. Dense trees at 600 m.
	180-200	-2	-
	200-220	-2	-
L05	220-240	-3	-
	260-280	-2	-
	240-260	-2	-

7.4. Cloud analysis

In this section, the cloud selections leading to conclusions regarding turbine performance of the eastern cluster are shown, including outputs of the cloud selection tool. First, the 10 minutes averaged cloud analysis results are discussed. Afterwards, the HF cloud analysis. The methods and tools used to get to the results of this chapter can be found in section 3.1.6 of the methodology chapter.

Looking at the total performance compared to WPC (table 7.3), the turbines from the eastern cluster show better performance than the turbines from the western cluster. The relation of the HF and 10 minutes averaged scatter data to this performance difference are put to the light.

In this chapter, only the results **deviating from the other clusters** are discussed. The tables may present more clouds than discussed in the section. The clouds not discussed were selected, but did not show deviating behaviour from other clusters during the analysis.

7.4.1. Results 10 minutes averaged cloud analysis

In this section, the results from the 10 minutes averaged data cloud analysis are discussed. The tables below contain information of all clouds selected, leading to results. Moreover, the state curves with binned 10 minutes averaged data of all cluster turbines are given. As explained in section 3.1.6, these state curve relations are inputs for the cloud selections.

Table 7.5: 10 minutes averaged clouds analysed eastern cluster.

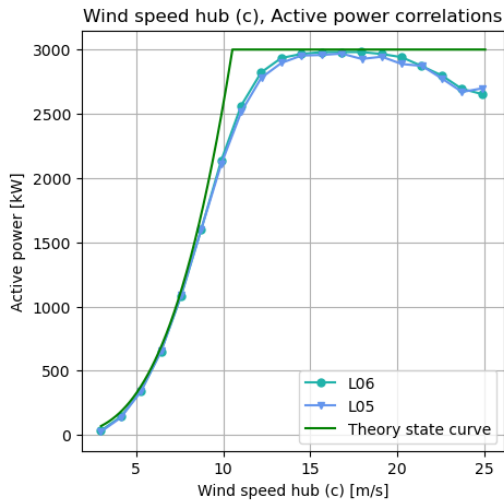
Cloud tag	X	Y	Turbine/ Cluster	Description
L06-10-657	Wind speed hub	Active power	Cluster	Scattering below PC for wind speeds higher than 20 m/s
L06-10-365	Wind speed hub	Active power	Cluster	Scattering below PC
L06-10-211	Wind speed hub	Pitch	Cluster	Pitching at low wind speeds
L06-10-648	Rotor speed	Torque	Cluster	Between 6-8 RPM torque not torque-rotor speed relation
L06-10-350	Rotor speed	Torque	Cluster	Sharper transition region compared to western cluster
L05-10-248	Wind speed hub	Rotor speed	Turbine	Second line of pitch/power-/RPM.
L06-10-275	Wind speed hub	Pitch	Cluster	Scattering above pitching curve.
L06-10-150	Wind speed hub	Active power	Cluster	High wind ride through
L06-10-588	Rotor speed	Torque	Cluster	Behaviour in optimal- c_P region

Table 7.6: Data collection of selected clouds eastern cluster

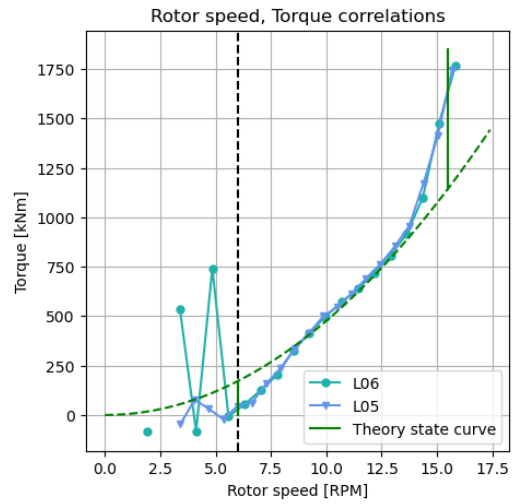
Cloud tag	Data points		Production		LPP		
	Points	%	kWh	%	kWh	% LPPT ²	% LPPC ³
L06-10-657	257	0.96	110821	1.47	15660	0.21	14.13
L06-10-365	164	0.61	60240	0.8	17258	0.23	28.65
L06-10-211	669	2.5	5504	0.07	3600	0.05	65.42
L06-10-648	1916	7.17	22960	0.3	12542	0.17	54.62
L06-10-350	13390	50.1	6040601	80.23	-10119	-0.13	-0.17
L05-10-248	256	1.32	118608	2.25	7359	0.14	6.2
L06-10-275	125	0.47	36546	0.49	23641	0.31	64.69
L06-10-150	224	0.84	104866	1.39	5318	0.07	5.07
L06-10-588	4167	15.59	421845	5.6	41994	0.56	9.95

²LPPT: Lost Potential Power as a fraction of annual power production

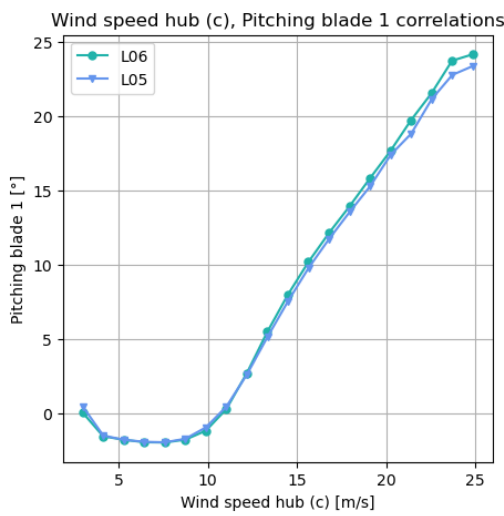
³LPPC: Lost Potential Power as a fraction of cloud power production



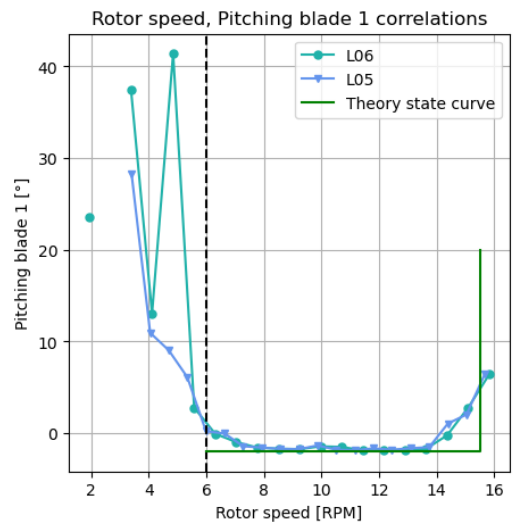
(a) Eastern cluster wind speed - active power average comparisons.



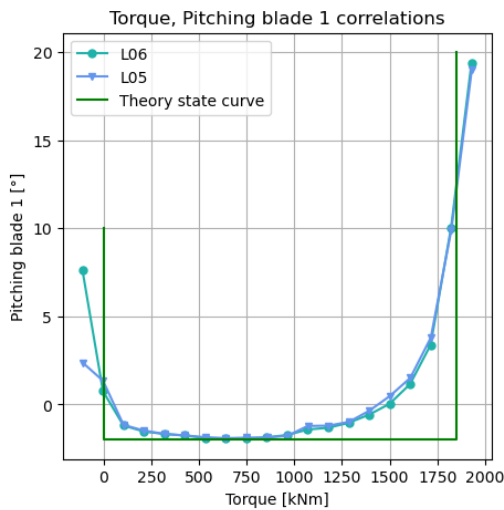
(b) Eastern cluster rotor speed - torque average comparisons.



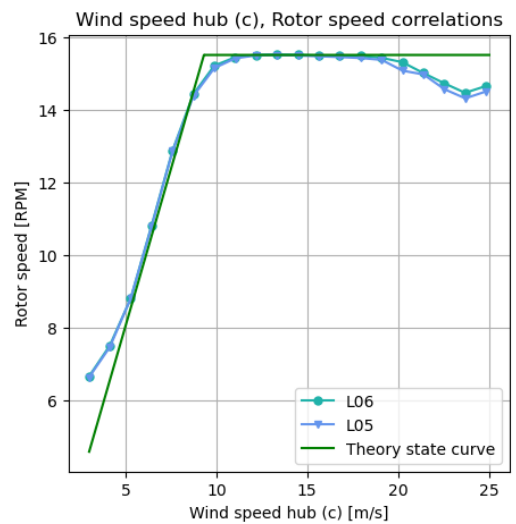
(c) Eastern cluster wind speed - pitching average comparisons.



(d) Eastern cluster rotor speed - pitching average comparisons.



(e) Eastern cluster pitching - torque average comparisons.



(f) Eastern cluster wind speed - rotor speed average comparisons.

Figure 7.17: Compared 10 minutes averaged data of eastern cluster turbines. Data is binned by the MOB.

As discussed in section 3.1.6, a comparison is made between the averaged binned state curves of the turbines and the theoretical expected turbine state curves. Observations from these comparisons are input for the cloud selection tool. From the overview of binned 10 minutes averaged state curves in figure 7.17, the storm control is well visible in the power curve of figure 7.17a. Moreover, higher average pitching behaviour is observed at turbine L06 compared to turbine L05. Also, the optimal- c_p region of the torque state curve (figure 7.17b) shows a slight different fit to theoretical behaviour than the western cluster turbines. The binned power curves in figure 7.17a show an underperformance of turbine L05 compared to L06 in the rated region of the power curve⁴.

Therefore, in the below section, the following observations are discussed: high wind speed power control, stability of the rotor speed - torque transition region, performance of the optimal- c_p region and second lines of pitching/RPM.

High wind speed power control

As the wind speed distribution of the eastern cluster has a longer right tail compared to the western cluster (figure 7.2), the eastern turbines are experiencing higher wind speeds. The eastern turbines are exposed to wind speeds ranging from 20 to 25 m/s. For the western cluster, 10 minutes averaged wind speed of 20 m/s or more was a rarer event. From 20 m/s, the turbine control is switched to a mechanism which makes production possible at those high wind speeds (explained further at section 4.1). This behaviour is also included in the WPC of the turbine. Above 20 m/s, the active power decreases. This can be seen in the averaged active power state curves (figure 7.17) and in the scatter plot below (figure 7.18).

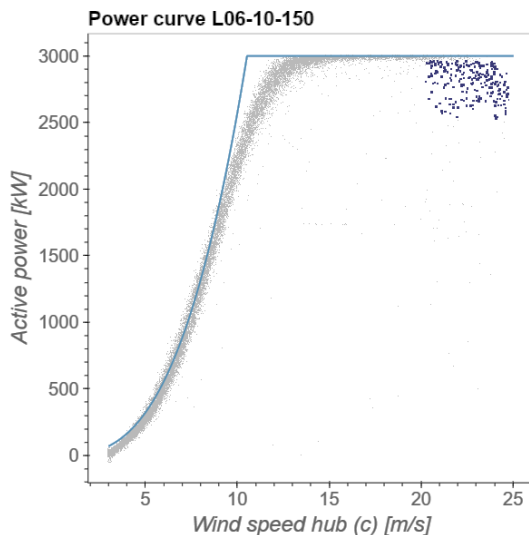


Figure 7.18: High wind speed power control active power output.

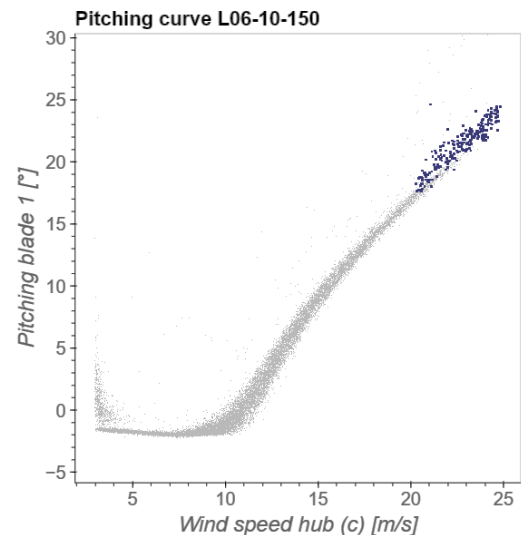


Figure 7.19: High wind speed power control pitching behaviour.

To see what the active power output signal of this high wind speed control looks like, the HF data for the specific timestamps of cloud L06-10-150 is consulted. In the figure below, this power output is shown with respect to power outputs of rated wind speeds below 20 m/s (in grey).

⁴This is also visible in the binned power curves of figure 7.11 and 7.14

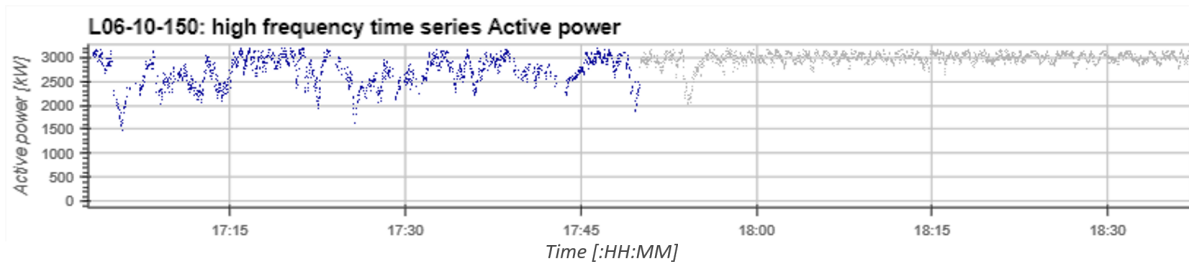


Figure 7.20: Time series of power output during high wind speed control.

The above rated wind speed turbine control (right side of the time series, grey) is showing a relatively stable active power signal around 3000 kW. Small fluctuations are visible due to the wind speed up and down ramping. On the left side of the time series (in blue), the active power output for storm control is shown. This active power output is fluctuating more, most likely to cope with higher loads that come together with high wind speeds.

In table 7.5, the statistics of the selected cloud L06-10-150 are given. The high wind speed control cloud has a LPPC value of 5%, meaning the storm control is performing 5% under the WPC behaviour. The lost amount of power production does not contribute much to the annual power production that year (0.07%).

The 5% underperformance compared to WPC can be partially explained by a minor amount of data exceeding the turbine TI class A conditions (figure 7.21). The cloud is shown in dark blue. The maximum TI class A turbulence conditions is given in light blue.

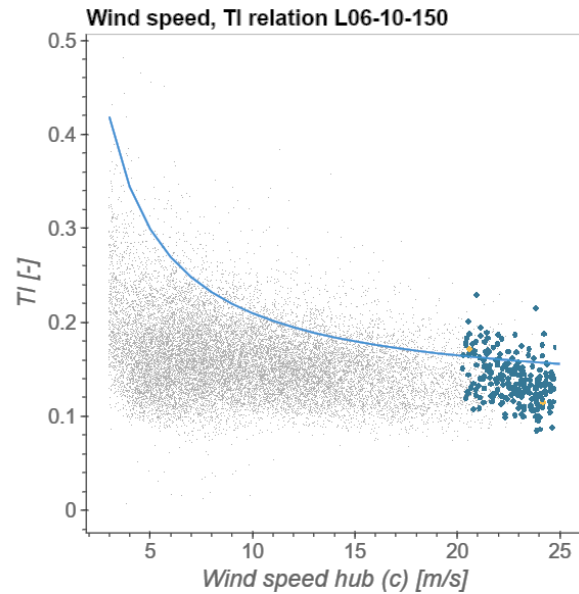


Figure 7.21: Wind speed - TI relation cloud L06-10-150. In blue: TI class A turbulence relation.

Rotor speed - torque region in transition region between optimal- c_p and full load

As discussed above, the turbines of the eastern cluster are exposed to higher wind speeds compared to the western cluster. Nevertheless, the 10 minutes averaged rotor speed signal exceeds the maximum rotor speed less compared to the western cluster. Below, a comparison between the wind speed - rotor speed relations is given. For the comparison, the two turbines closest to the MET mast were used.

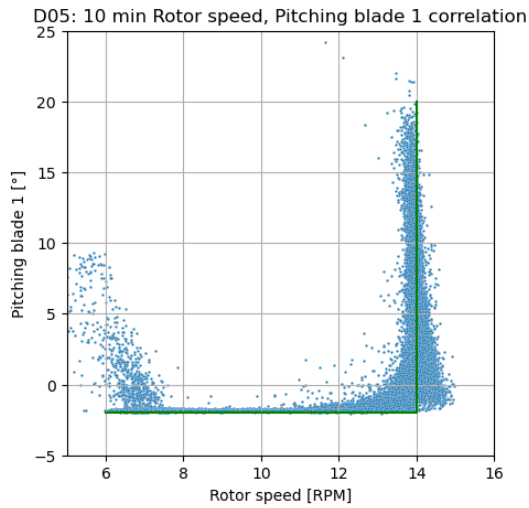


Figure 7.22: Rotor speed - pitching relation turbine D05.

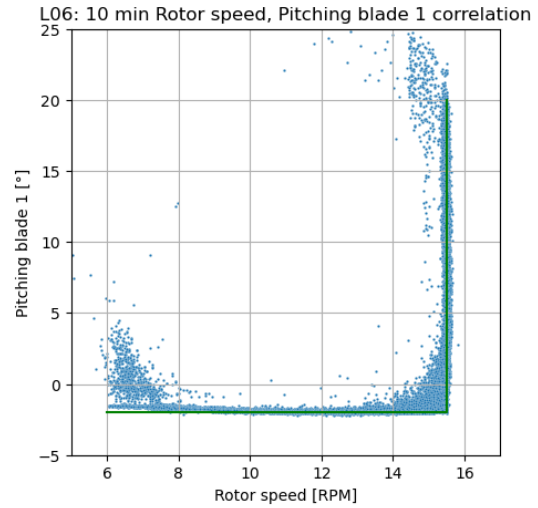


Figure 7.23: Rotor speed - pitching relation turbine L06.

For the western string turbines, it can be observed that where maximum rotor speed is reached (U_{nrtd} , the right bottom of the scatter plot) the rotor speed has a relatively high standard deviation and can increase up to 15 RPM. This was also visible in the HF data.

When looking at figure 7.23, this behaviour is not seen for the eastern cluster. The wind speed - rotor speed relation fits the pre-calculated state curve better and consists of less scattering around the maximum rotor speed.

This same difference in rotor speed behaviour also shows differences in the rotor speed - torque relation for both clusters. For the western cluster, the rotor speed - torque relation did not show the expected behaviour from the point U_{nrtd} onwards. U_{nrtd} is the point from which the rotor speed - torque relation is not following the $Q \propto \omega^2$ trend any more. The eastern cluster fits the expected behaviour better, with a more stable rotor speed for high wind speeds. The expected behaviour is a transition region evolving in a vertical line of constant rotor speed (see light blue line). Below, for both the western and eastern reference turbine, the rotor speed - torque relations are given. In dark blue are the two selected clouds which resemble the data from U_{nrtd} onwards.

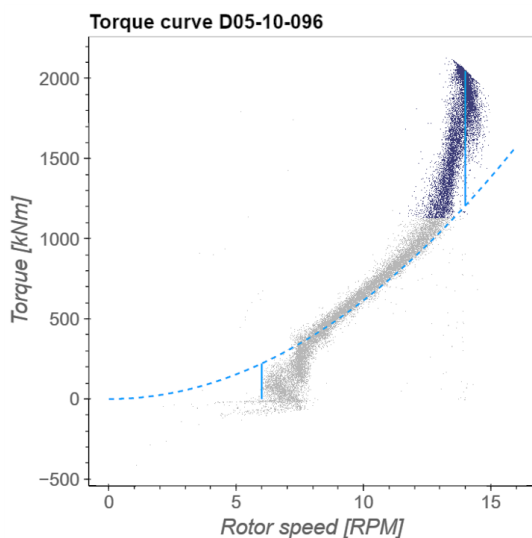


Figure 7.24: Rotor speed - torque relation turbine D05.

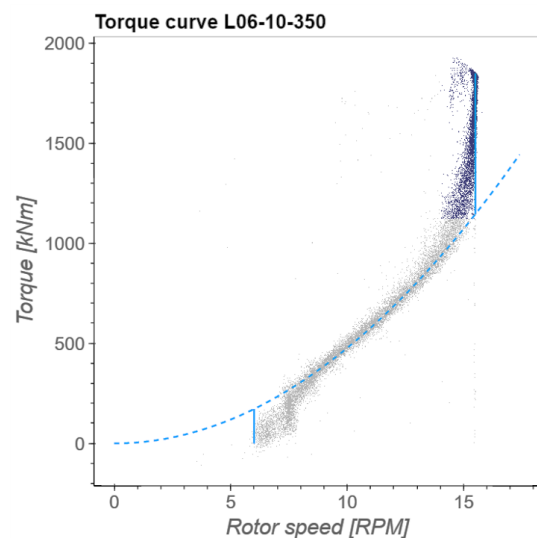


Figure 7.25: Rotor speed - torque relation turbine L06.

When comparing the two selected clouds, they consist of a comparable amount of data points (± 13000).

For the western cloud, there is a 1.85% power loss compared to annual production, whereas for the eastern cluster the cloud is performing slightly better than WPC. This western cluster power loss can be caused by the torque and pitch control, which seems not to match the theoretical rotor speed - torque state curve.

Although the eastern cluster experiences higher wind speeds, TI values measured at the eastern MET mast are less extreme compared to the western cluster. Moreover, as discussed before, the western cluster are TI class B turbines, where the eastern cluster are class A turbines. A combination of these two can be an explanation of the difference in rotor speed control and turbine performance in this transition region between optimal- c_p and full load.

Torque curve: performance optimal- c_p region

When comparing the torque curve scatter plots of the western and eastern clusters (7.17b and 6.27b) it is concluded there is a difference in the fit to the maximum power coefficient part of the theoretical torque curve (dotted green line) for the SWT-3.0-113 turbines compared to the SWT-3.0-108 turbines. For turbine D05, the data is slightly below the expected theoretical behaviour. Looking at the eastern cluster fit to the torque curve in this region, it shows a better fit to this map c_p region.

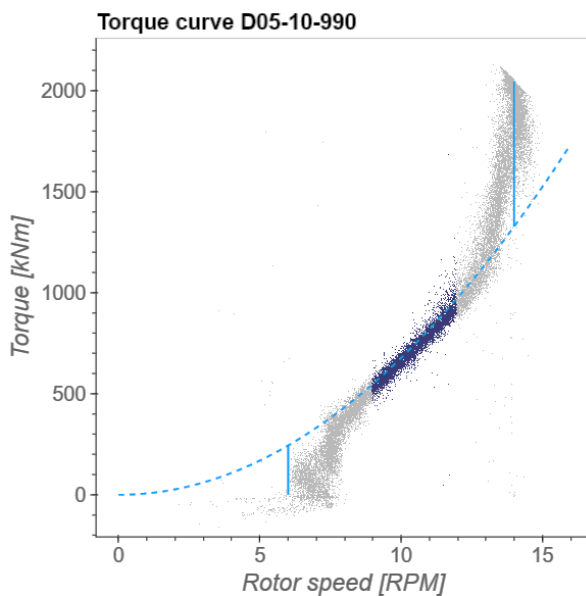


Figure 7.26: Optimal- c_p region of turbine D05.

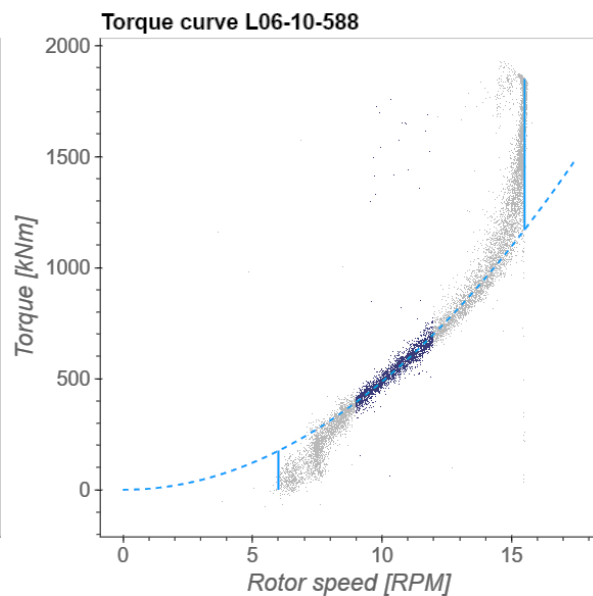


Figure 7.27: Optimal- c_p region of turbine L06.

When looking at the calculated performance parameters for each cloud, it is seen that the optimal- c_p region of turbine D05 has a LPPC of 14%, while turbine L06 of the eastern cluster had an LPPC of 10%. The difference in fit to the theoretical state curve is thus expressed within the performance of this area of the torque curve. This optimal- c_p region is an essential part of turbine performance. The fact that these regions are performing worse compared to overall turbine performance, is enough incentive to start a discussion with the OEM.

Second line of pitch, RPM and power

Another difference that is seen within most of the state curve scatter plots of turbine L05 is a cloud of data on which more extreme pitching is applied. This is visible in three ways:

1. A lower horizontal line within the rated region of the power curve.
2. A second line of pitching higher than the average pitching curve for wind speeds above U_{nrtd} .
3. Rotor speed measurements lower than the maximum rotor speed within the rated wind speed region.

The cloud selected for investigating this is cloud L05-10-248. The cloud is 256 data points (1.32% of total data), and the potential production lost is less than 0.15% of annual power production. Below, the power curve and wind speed - rotor speed relation with the selected cloud are given.

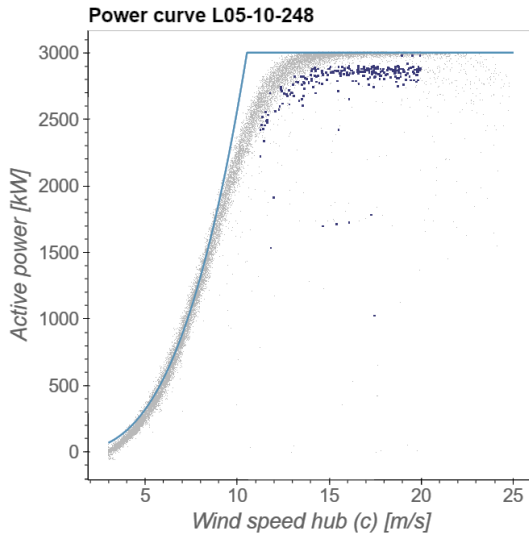


Figure 7.28: Power curve with second line of power production.

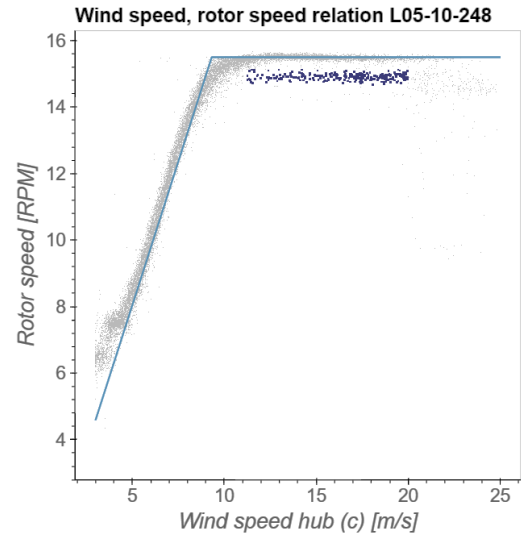


Figure 7.29: Wind speed - rotor speed relation with second line of rotor speed.

This cloud explains the relative underperformance of turbine L05 compared to L06 in the rated region of the power curve, as was observed from the binned state curve relations (figure 7.17a). Reasons for this derated behaviour can potentially be caused by TI exceeding turbulence class A thresholds. In figure 7.30, the TI values of the cloud L05-10-248 are shown. Half of the cloud data is exceeding turbulence class A conditions.

Moreover, this cloud selection reveals the impact of extreme site conditions on production of TI class A turbines. These losses are unavoidable and part of the rough site.

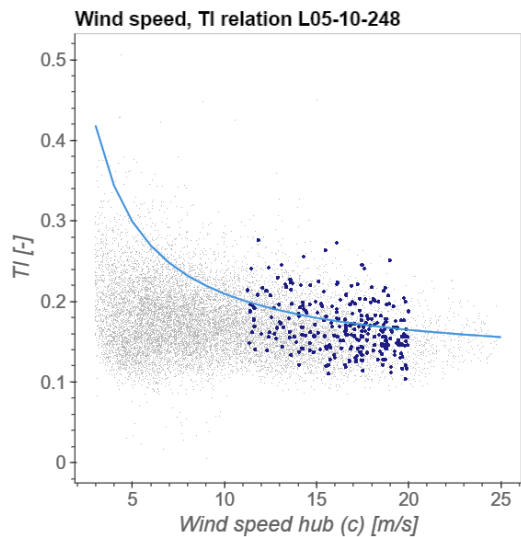


Figure 7.30: Wind speed - TI relation cloud L05-10-248. In blue: TI class B turbulence relation.

7.4.2. Results HF cloud analysis

Below, an overview of the selected clouds are given in tables 7.7 and 7.6. The results from these clouds showing different behaviour compared to the other clusters are discussed. Scattering below the power curves, extreme pitching and torque variation for high wind speeds are reviewed in this section.

Table 7.7: HF clouds analysed eastern cluster.

Cloud tag	X	Y	Turbine/ Cluster	Description
L06-HF-627	Wind speed hub	Active power	Cluster	General scattering below PC
L06-HF-632	Wind speed hub	Pitch	Cluster	Scattering above pitching curve
L06-HF-678	Pitching	Torque	Cluster	High torque for high wind speeds and pitching
L06-HF-552	Pitching	Torque	Cluster	Low torque for high wind speeds and pitching
L06-HF-742	Rotor speed	Pitch	Cluster	Pitching values up to 30 degrees at non-rated rotor speeds
L06-HF-137	Wind speed hub	Pitch	Cluster	Pitching behaviour below 7 m/s

Table 7.8: Data collection of selected clouds eastern cluster

Cloud tag	Data points		Production		LPP		
	Points	%	kWh	%	kWh	% LPPT ⁵	% LPPC ⁶
L06-HF-627	19629	0.2	17179	0.24	7394	0.1	43.04
L06-HF-632	4792	0.05	3961	0.06	3357	0.05	84.74
L06-HF-678	9077	0.09	10784	0.15	-72	0	-0.66
L06-HF-552	11904	0.12	12200	0.17	2488	0.03	20.39
L06-HF-742	7965	0.08	7850	0.11	2998	0.04	38.19
L06-HF-137	63206	0.65	12	0	2399	0.03	20809.53

Different scattering below power curve

Below the power curve of the eastern turbine clusters, a different scattering can be seen compared to the western cluster HF power curve scattering (figure 6.35). For this eastern cluster, the scattering is also dense for high wind speeds, whereas for the western cluster, the scattering density decreased for increasing wind speeds. To investigate this scattering below the power curve of eastern cluster turbines, cloud L06-HF-552 was selected. Scattering up to 20 m/s was selected to keep the scattering separated at 20 m/s. By doing so, it was easier to investigate the influence of wind speeds higher than 20 m/s on the selected cloud.

The reason for the different looking scattering is the cause: For the eastern cluster, the majority of the scattering (below a wind speed of 20 m/s) are data points where wind speeds have ramped down from the high wind speed control region. This was observed from when zooming in to samples of the selected cloud L06-HF-552. The active power signal does not respond or responds later to those wind speed ramps. Therefore, scattering is visible. For example, the HF time series shown in figure 7.20 is a part of cloud L06-HF-627. As can be seen from the active power distribution, there is a high occurrence of active power signals around 1600 kW. This can be explained by the time series of figure 7.20. This is an example of the storm control active power output signal. One of the derated power levels of the storm control is around 1600 kW.

⁵LPPT: Lost Potential Power as a fraction of annual power production

⁶LPPC: Lost Potential Power as a fraction of cloud power production

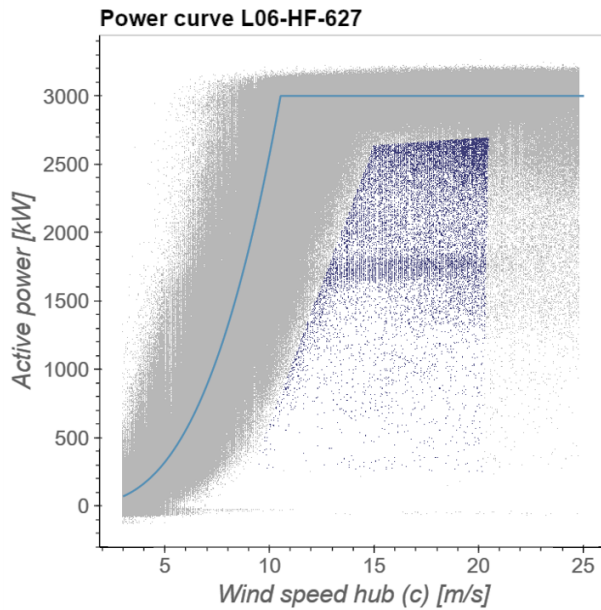
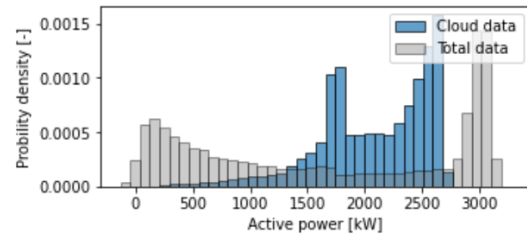
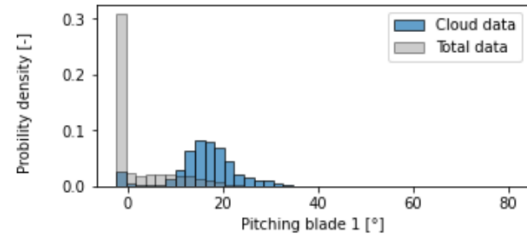


Figure 7.31: High wind speed power control active power output.



(a) Active power distribution of cloud L06-HF-627.



(b) Pitching distribution of cloud L06-HF-627.

Figure 7.32: Pitching and Active power distributions for cloud L06-HF-627.

Cloud L06-HF-627 is a minor part of the total HF data set (0.2% of total data points). The potential power lost makes up only 0.1% of annual power production. This cloud shows how HF data should be dealt with carefully when analysing. At the first sight, the turbine seems to be produce below rated for a significant amount of data points. However, the eastern cluster scattering below the power curve does not cause any underperformance, as it originates from unavoidable wind speed down ramping.

Extreme pitching during storm control

The extreme scattering above the pitching curve is collected within cloud L06-HF-632. This pitching behaviour originates from data close to cut-out wind speed. At cut-out wind speed, the turbine is curtailed to prevent the turbine from suffering from the extreme wind conditions. Close to these cut-out wind speeds, the PYC turbines are in storm control mode. The turbine exceeds the pitching thresholds for keeping the turbine at rated power (blue data in figure 7.33), to derate the turbine to an operational condition with lower loads.

In figure 7.34, the rotor speed response to this high pitching is represented. Rotor speeds are decreased to make the turbine cope with the loads at high wind speeds. For comparison: during curtailments due to grid limitations, rotor speeds are not decreased to have a high inertia response to the grid.

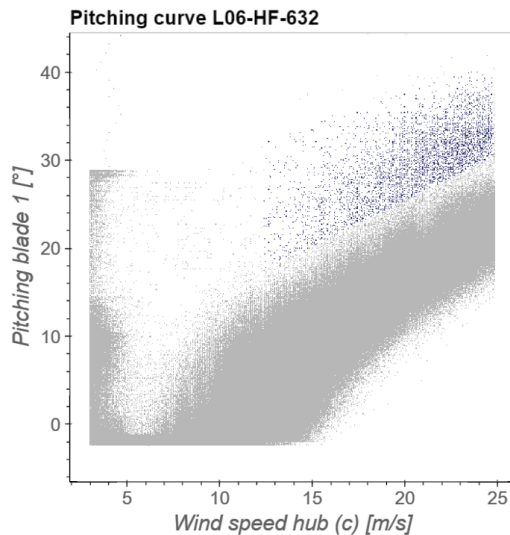


Figure 7.33: Cloud L06-HF-632: extreme pitching.

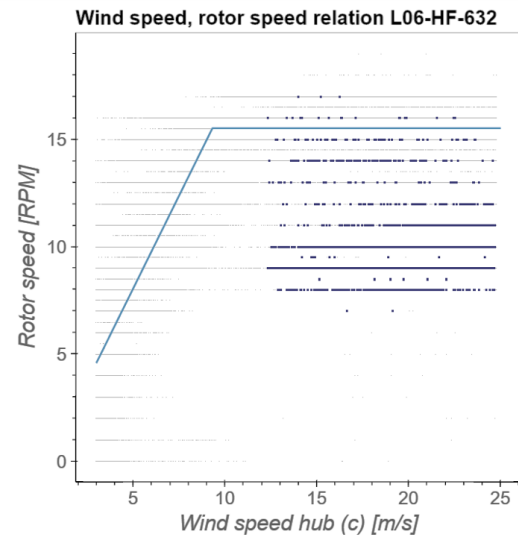


Figure 7.34: Cloud L06-HF-632: decrease of rotor speeds.

In the time series below, segments of data are filtered out. These are the sections where no grey data can be seen. This is due to turbine curtailments during high average wind speed periods. The extreme pitching cloud occurs at moments of high average wind speeds, close to cut-out. The data of cloud L06-HF-632 is shown as the circled dark blue data. The pitching selected is only 0.05% of total data points, and potential losses are the same percentage of annual power production.

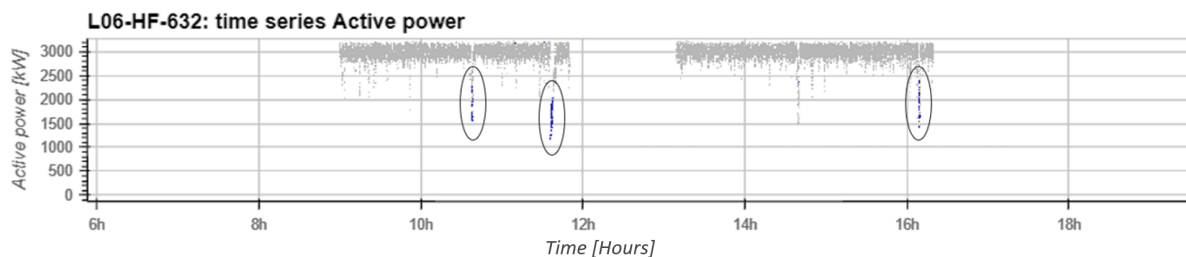


Figure 7.35: Cloud L06-HF-632: example of power signal around selected data.

This HF cloud shows the exact pitching strategy for storm control. When looking at the same timestamps in the 10 minutes averaged data, an averaged out behaviour is observed, showing pitching marginally deviating from the pitching curve. The HF resolution makes it possible to strictly select the pitching behaviour in storm control, without including non-storm control data. This has a big effect on the LPPC value: a 84% lost power potential as a fraction of cloud production is observed, whereas the 10 minute averaged cloud showed a LPPC of 5%. Cloud L06-HF-632 therefore shows the potential of more strictly selecting instances of underperformance.

Torque variation for high wind speeds

During high wind speed control, rotor speed is not controlled at its maximum value of 15.5 RPM, but is used as a variable together with the active power signal as explained in section 2.6.

As both rotor speed and active power are variable above 20 m/s wind speeds, resulting in torque also becoming a varying parameter. Clouds L06-HF-678 and L06-HF-552 were selected to analyse respectively high and low torque behaviour at high wind speeds and pitching. Each cloud is 0.1% of total data.

In the HF data, this wide variation in torque becomes visible. Cloud L06-HF-678 represents the high torque, exceeding the general HF scattering. Below, the high-torque cloud that deviates from the general torque scattering is displayed (figure 7.36), together with distributions of some key properties in

figure 7.37.

Cloud L06-HF-678 originates from the active power exceeding P_{rated} . Moreover, rotor speeds dropping below w_{rated} for wind speeds in the full load region⁷. The low-torque data for high wind speeds is from data where the active power has dropped to values around 2500 kW as a result of the high wind speed control, and for rotor speeds exceeding w_{rated} . The high torque is problematic for turbine lifetime. Moreover, it is an incentive for applying storm control, or for complete shut down of the turbine.

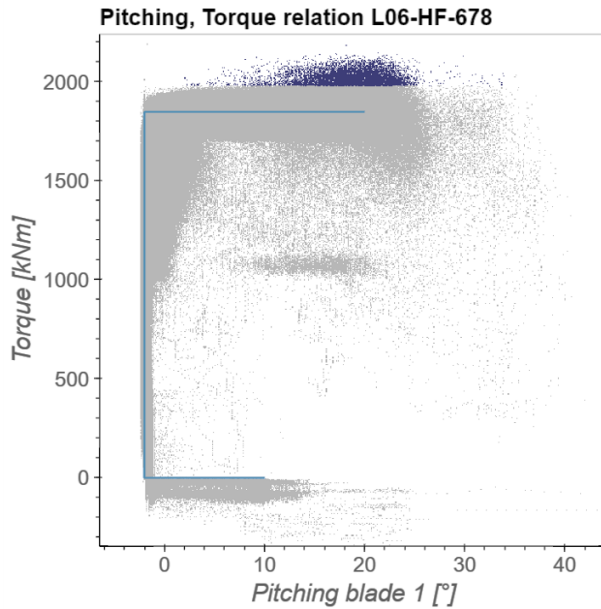
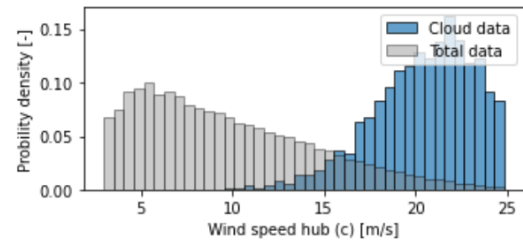
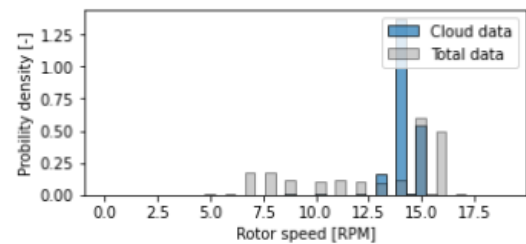


Figure 7.36: Selected high-torque cloud for high pitching.



(a) Wind speed distribution of cloud L06-HF-678.

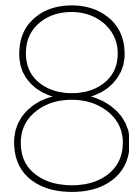


(b) Rotor speed distribution of cloud L06-HF-678.

Figure 7.37: Wind speed and pitching distributions for cloud L06-HF-678.

This 0.1% high-torque contribution is visible in HF data, where hidden in 10 minutes averaged data. It is valuable to understand this HF high-torque phenomenon when implementing turbine control improvements, such as a better responsive rotor speed control.

⁷The lower rotor speeds are shown by the highest peak in figure 7.37b.



Results central cluster

In this chapter, deviations in the central cluster results compared to the eastern and western cluster are discussed.

8.1. Wind resource

Central cluster site conditions are compared to the eastern and western cluster site conditions to understand differences. The site conditions can be related to performance differences.

Below, the wind speed and wind directional distributions of the central cluster are given. Just as for the eastern cluster, Scipy was used to fit a Weibull function to the data. As can be seen from the Weibull distribution, average wind speeds are low compared to the more extreme eastern cluster. Wind speeds of the central cluster are comparable to the wind speeds measured at the central cluster.

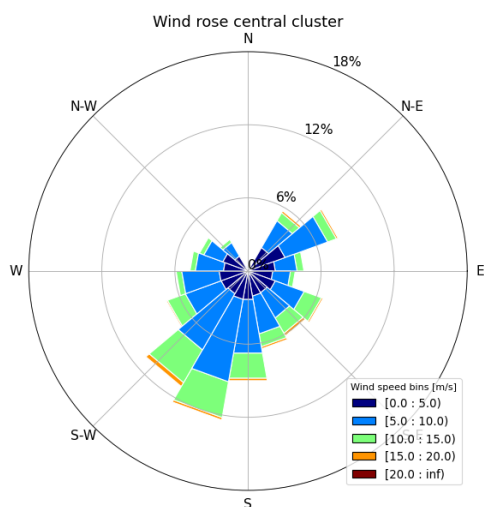


Figure 8.1: Wind rose central cluster

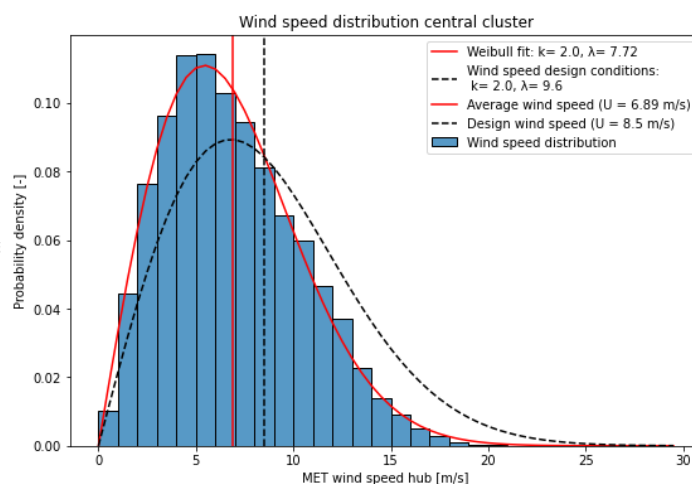


Figure 8.2: Wind speed histogram central cluster

Regarding wind veer at the central cluster site, again the northerly direction ± 60 degrees was filtered out. For all investigated MET masts at the site, the northerly wind directional data is corrupt. The wind speed-veer relation at the central cluster does not contain an offset, where for the western and eastern cluster an offset of $0.25^\circ/\text{m}$ and respectively $0.06^\circ/\text{m}$ was observed. As can be seen in figure 8.4, average wind speed is above zero for wind directions effected by neighbouring turbines. For the unawaked sectors, wind veer shows stable results.

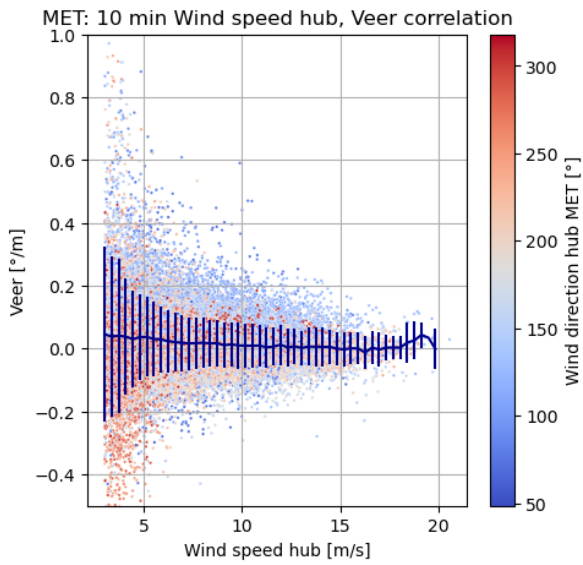


Figure 8.3: 10 minutes averaged MET02 wind speed veer correlation.

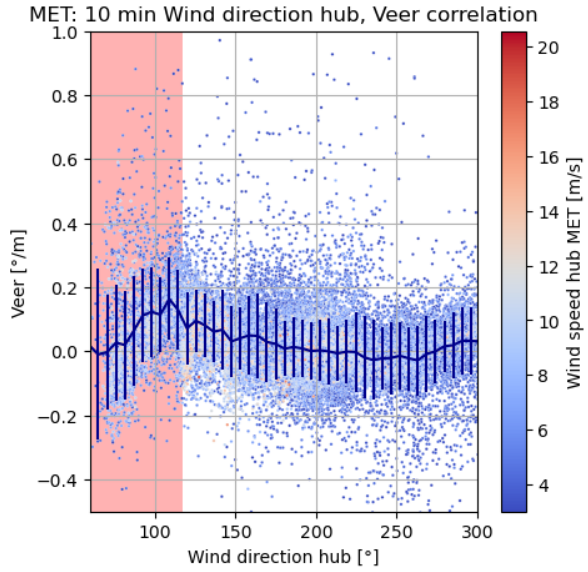


Figure 8.4: 10 minutes averaged MET02 wind direction veer correlation.

Where wind directional veer shows less extreme values for the central cluster, wind shear is more extreme. These extreme wind shear occurrences result in average wind shear values of 0.5. This is comparable to the shear measured at the western cluster, and higher than the shear measured at the eastern cluster (on average 0.3).

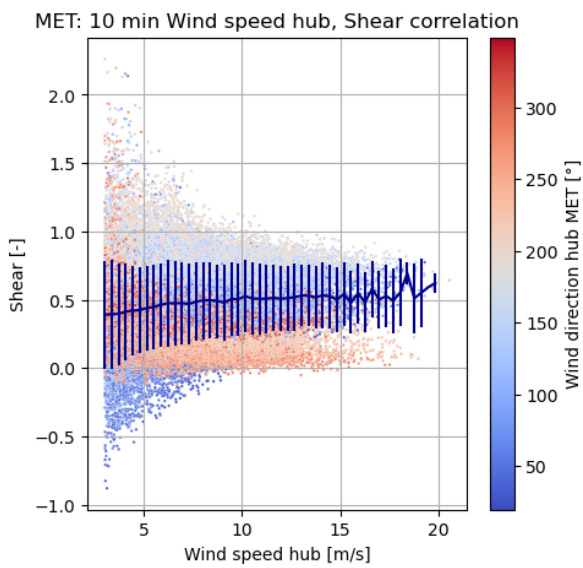


Figure 8.5: 10 minutes averaged MET02 wind speed shear correlation.

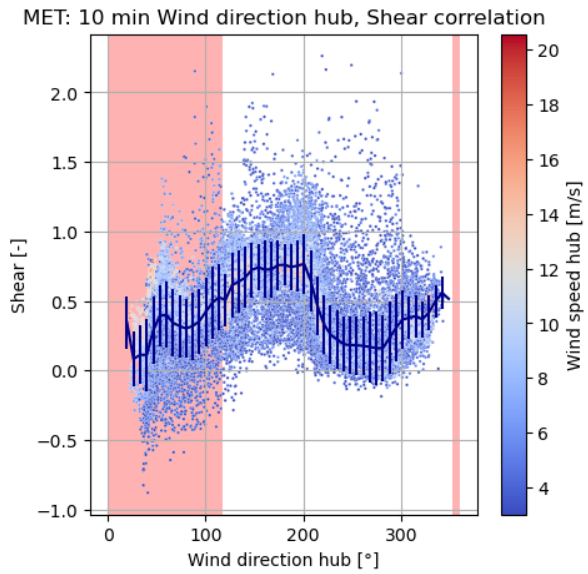


Figure 8.6: 10 minutes averaged MET02 wind direction shear correlation.

When comparing TI measured at MET02 to the measurements at the other two clusters, the central cluster experiences higher TI values than the western and eastern cluster. This is also expressed in the average TI over wind speed: Up and until a wind speed of 15 m/s the average TI is close to 0.18 whereas for the western and eastern cluster it decreases below 0.15 from 7 m/s onwards. Just as for the western and eastern cluster, the highest TI occurs within the waked MET mast sectors.

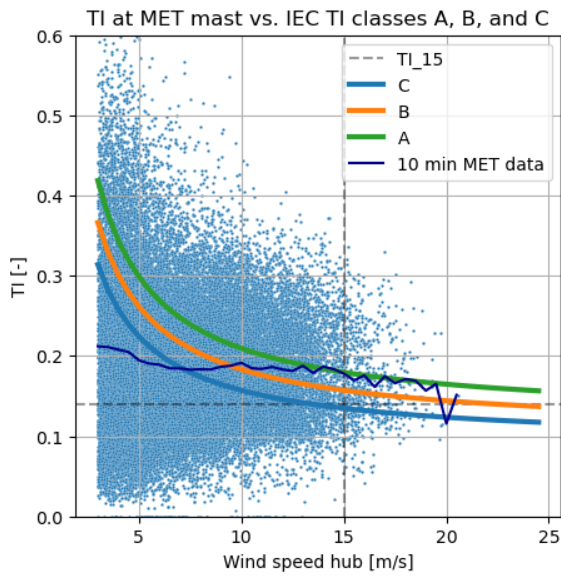


Figure 8.7: 10 minutes averaged MET02 wind speed TI correlation.

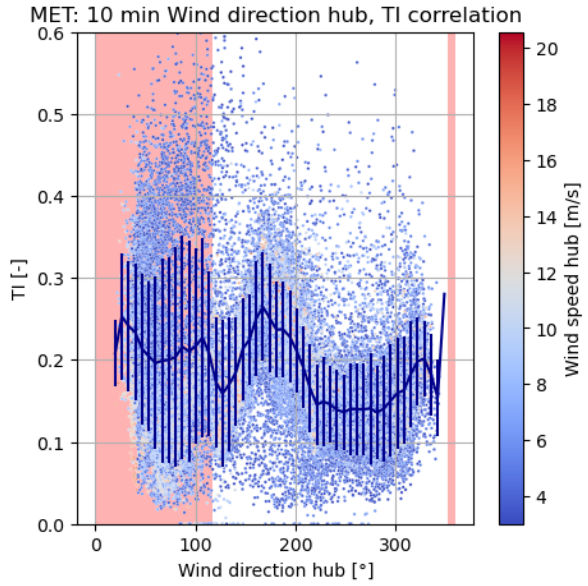


Figure 8.8: 10 minutes averaged MET02 wind direction TI correlation.

8.1.1. Influence of site on turbine performance H05

The effect of wind directional veer and wind shear on the turbine performance of the closest MET mast is examined by comparing measurements from MET03 and turbine H05. The distance from turbine H05 to the MET mast is 2D, and therefore within the 2-4D margin. At the central cluster, average veer values are lowest compared to the other two clusters. Average shear values are comparable to the western cluster. However, the shear does have a higher standard deviation for all wind speed bins, resulting in more extreme shear measurements at the central cluster¹.

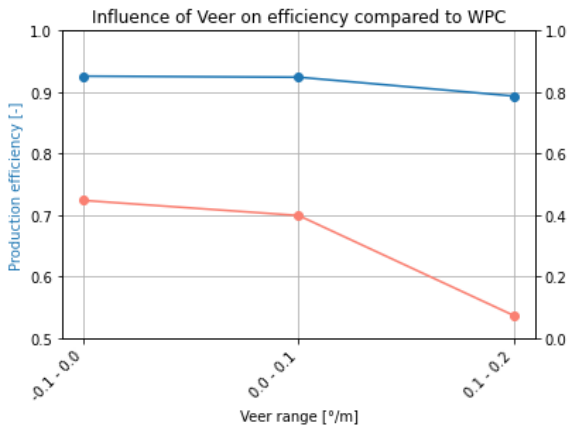


Figure 8.9: Influence of veer on H05 active power signal.

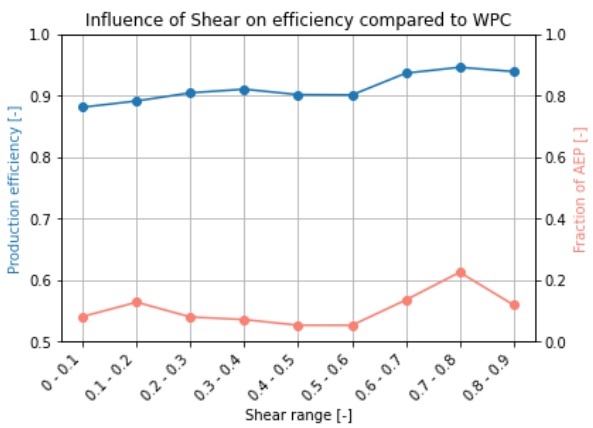


Figure 8.10: Influence of shear on H05 active power signal.

The same observations are done as for the western cluster regarding wind shear: an increase in shear positively affects the turbine performance. As for wind veer, there is a slight negative influence of 2% for the 0.1-0.2 °/m veer bin compared to the 0.0-0.1 °/m bin. However, this bin is of far less influence to power production compared to the two lower veer bins.

To conclude, the (extremely) high shear measurements are of positive effect on the turbine performance estimates. This was also concluded for the western cluster. Regarding veer, a negligible negative effect was noticed for veer above 0.1 °/m.

¹This is observed from comparing figures 6.5 and 7.5

An important note is that these conclusions are based on the active power signal of one turbine, and veer/shear measurements of one MET mast. Influence of these parameters further on in the turbine clusters is not possible to determine.

8.2. Correlations MET mast and turbines

After the first iteration of the side-by-side correlation process, it became clear the directional data of all turbines was offsetted for all turbines of the central cluster. This phenomenon was also observed for turbine D04 of the western cluster (table 6.3). The difference from the western cluster, is the fact that this offset occurs at all turbines. This makes the western cluster data set more uncertain, as more manipulations were carried out prior to the analysis of this cluster.

As the wind speed correlations only consider free sector data, and wind directional data is needed to select on these criteria, having good calibrated wind directional data is important for wind direction as well as wind speed correlations.

Another change, applied to the selected central cluster data, is the elimination of turbine H07 from the 5 selected turbines. This turbine was eliminated because the free sectors of the turbine did only match a minor part of the free sectors of the consecutive turbine. This harms the reliability of the correlations. Moreover, exclusion of this turbine does not have a negative influence on the inter-turbine distances of the 4 remaining turbines. In the table below, an overview of the corrections carried out is given:

Table 8.1: Overview of corrections carried out on correlations central cluster

	Operation	Applied to turbine	Properties	Size of offset
1	Correct offset >10 degrees compared to neighbouring turbine	H05	Nacelle direction Wind direction	55.4
2	Correct offset >10 degrees compared to neighbouring turbine	H06	Nacelle direction Wind direction	14.7
3	Correct offset >10 degrees compared to neighbouring turbine	H07	Nacelle direction Wind direction	23.6
4	Correct offset >10 degrees compared to neighbouring turbine	H04	Nacelle direction Wind direction	-10.6
5	Correct offset >10 degrees compared to neighbouring turbine	H02	Nacelle direction Wind direction	8.4

The correlations from this table were applied because of unnatural offsets between the directional measurements of the turbines. The size and presence of the offsets at every single turbine of this cluster is a reason for discussing the offsets with the O&M (Operations & Maintenance) team on site. After correcting the data as described by table 8.1, results of the side-by-side correlations of the turbines are given after the corrections listed in table 8.1 were carried out. The bottom table consists of expected ranges of slopes, based upon 2.8, a roughness length of 0.1-1 (table 2.5), and a 5% deviation.

Table 8.2: Overview of correlations for the central cluster.

Correlation	Intercept	Property	MET-H05	H05-H06	H06-H04	H04-H02
Wind speed	No	Slope	1.05	1.06	0.98	1.22
	No	R^2 score	0.93	0.71	0.93	0.75
	Yes	Slope	1.04	0.98	1.02	0.98
	Yes	Intercept	0.14	0.77	-0.42	1.56
	Yes	R^2 score	0.93	0.71	0.93	0.8
Wind dir	No	Slope	0.99	0.97	1	1.01
	No	R^2 score	0.97	0.99	0.99	1
	Yes	Slope	0.97	0.99	0.98	1.01
	Yes	Intercept	0.95	-3.7	3.16	0.18
	Yes	R^2 score	0.97	0.99	1	1
Nac dir	No	Slope	-	0.97	1	1.01
	No	R^2 score	-	0.99	0.99	1
	Yes	Slope	-	1	0.98	1
	Yes	Intercept	-	-5.74	3.2	0.87
	Yes	R^2	-	0.99	1	1

Table 8.3: Expected slopes for wind speed correlations central cluster.

	MET-H05	H05-H06	H06-H04	H04-H02
Height difference [m]	8.2	17.1	31.7	-8.3
Expected slope ($z_0 = 0.1$)	1.01	1.03	1.05	0.99
Expected slope ($z_0 = 1$)	1.02	1.04	1.07	0.98
Average	1.02	1.03	1.06	0.98
Accepted slopes (5% deviation)	0.97-1.07	0.98-1.08	1.01-1.11	0.93-1.03

The correlations of table 8.2, combined with the expectations of table 8.3 and the visualisation of the multiple wind speed and wind direction correlations, are used to validate the multiple turbine data sets of the central cluster. The MET mast and turbine H05 are correlating as expected: The slope is within the limits of table 8.3, and R^2 values are well above the limit of 0.7. Wind speed correlations of turbine H05 and H06 are showing different results. The slope of this wind speed correlation is within the expected regions. However, R^2 values are just above the 0.7 limit. This can be explained by the occurrence of two clouds, both having a different average slope. Below, both the wind speed correlations of MET-H05 and H05-H06 are shown.

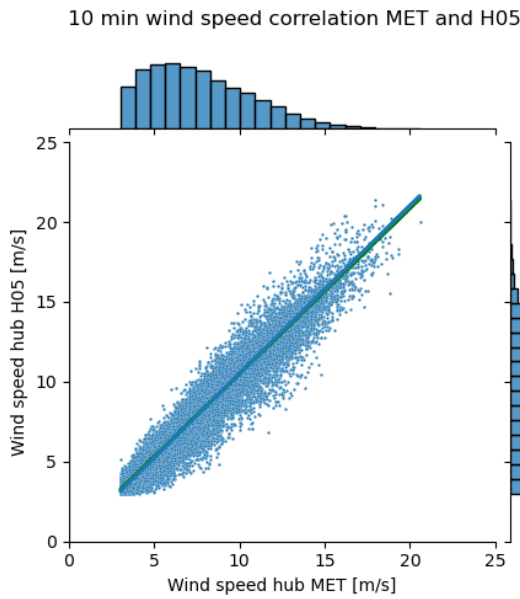


Figure 8.11: Wind speed correlation of MET02 and turbine H05.

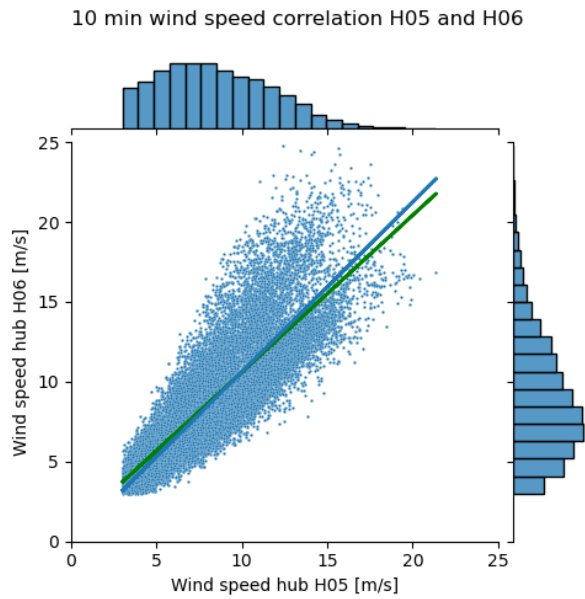


Figure 8.12: Wind speed correlation of turbines H05 and H06.

Because of the inconsistent wind speed behaviour of turbine H06 compared to turbine H05, the data of turbine H06 can not be trusted. Wind speed measurements seem to differ substantially for different orientations of turbine H06. As turbine H04 is a neighbouring turbine of H05 as well, the wind speed correlation between H05 and H04 was checked as well. The correlation showed the same behaviour as the correlation of figure 8.12. The steeply correlated data has a slope of >1.4 . This high slope can not be explained by the height difference between H05 and its neighbouring turbines. Because of the observed behaviour, turbines H06 and H04 can not be used for further analysis. As turbine H02 is to be validated by turbine H04, also turbine H02 is not used within the analysis of the central cluster. As a result, for the central cluster, only turbine H05 is investigated.

8.3. Overall & directional performance analysis

As discussed above, turbine H05 will be analysed as the only turbine from the central cluster. Please see the above section for reasoning. In this section, the results from the performance analysis are shown. This analysis is carried out in parallel to the cloud analysis (described in the next section). The methods to come to the results presented in this section are represented in chapter 3.

8.3.1. Overall turbine performance

The power output is shown as a binned power curve, compared to the warranted power curve of the central cluster site. The turbine is performing below WPC for the non-rated part of the power curve, as well as for the rated part. In the transition region from non-rated to rated (knee of the power curve) the turbine is performing close to WPC. The relatively bad performance in the rated part of the power curve coincides with turbine D05, D01 and L05. The underperformance in the non-rated part of the power curves was observed at the western cluster turbines.

As was also seen for the other cluster turbines, the difference between unwaked and waked power curves is small: this is because the pre-

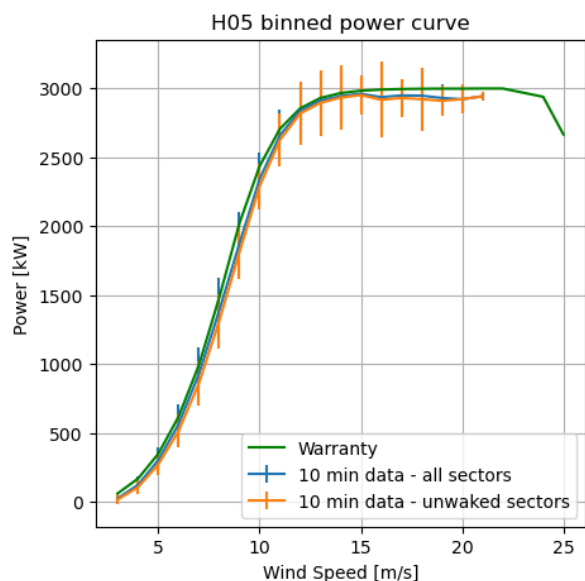


Figure 8.13: Binned power curve of turbine H05.

vailing wind directions are often the unwaked sectors. Therefore, waked sectors make up a minor part of the turbine data.

Table 8.4: Efficiencies compared to WPC for turbines central cluster.

	Based on	H05
$\eta_{power,1}$	Warranted power curve	0.93
$\eta_{power,2}$	Potential power signal	0.94

8.3.2. Directional turbine performance

For turbine H05, two graphs are plotted with each 9 of the 18 wind direction sectors. These operational power curves per wind direction sector are used to calculate the turbine performance per wind directional sector.

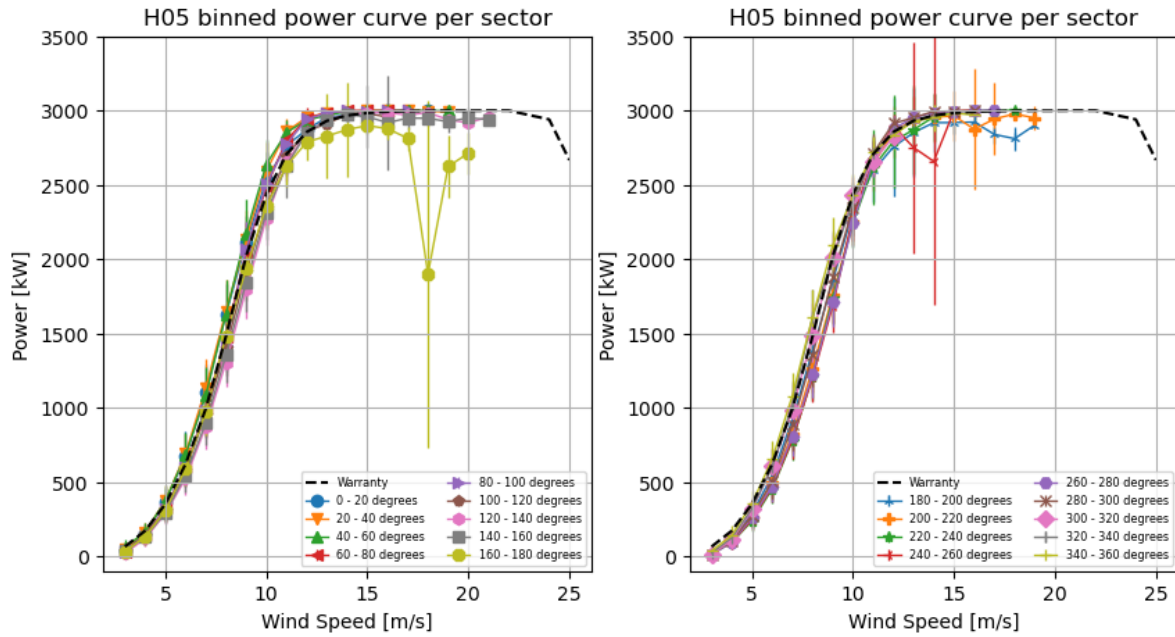


Figure 8.14: Binned power curves turbine H05

For each power curve per sector, the performance compared to the warranted power curve can be calculated by using equation 3.5. This relative performance per sector is given by a percentage in the diagrams below.

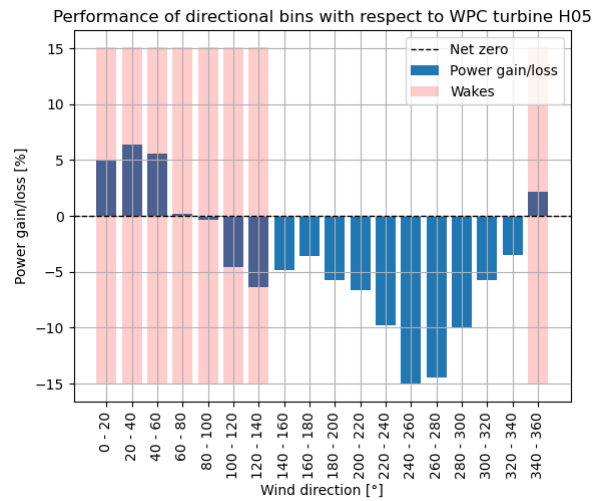


Figure 8.15: Relative power gain/loss per wind direction bin for turbine H05.

The sectors with a negative mismatch to WPC of more than 5% are selected for further site investigation, resulting in a short site description in the last column of the table. The three worst performing sectors of turbine H05 are from relatively flat terrain with no forestry. A rough valley is at the edge of the 280-200 degrees sector. The orography of the three tabulated sectors of this central cluster is less extreme than orography at sectors from tables 6.5 and 7.4.

Table 8.5: Worst performing unawaked sectors turbines central cluster.

Turbine	Sectors (unawaked) [°]	Performance gain/loss [%]	Site description
H05	240-260	-15	Forestry at 340 m. MET mast at 230 m.
	260-280	-14	No forestry. Slowly elevating.
	280-300	-10	No forestry. Side of valley towards turbine.

8.4. Cloud analysis

In this section, the cloud selections leading to conclusions regarding turbine performance of the central cluster are shown, including outputs of the cloud selection tool. First, the 10 minutes averaged cloud analysis results are discussed. Afterwards, the HF cloud analysis. The methods and tools used to get to the results of this chapter can be found in section 3.1.6 of the methodology chapter.

8.4.1. Results 10 minutes averaged cloud analysis

In this section, the results from the 10 minutes averaged cloud analysis of the central cluster are discussed. The different findings compared to the eastern and western cluster are discussed. In the tables below, an overview is given of the selected clouds. Below the tables in figure 8.16, the state curves of turbine H05 are shown. Deviation of turbine behaviour from these theoretical state curves were input to the selection of clouds.

In this section, two regions of the torque curve are compared between the different clusters. Moreover, pitching for low wind speeds and scattering below the power curve are discussed.

Table 8.6: 10 minutes averaged clouds analysed central cluster.

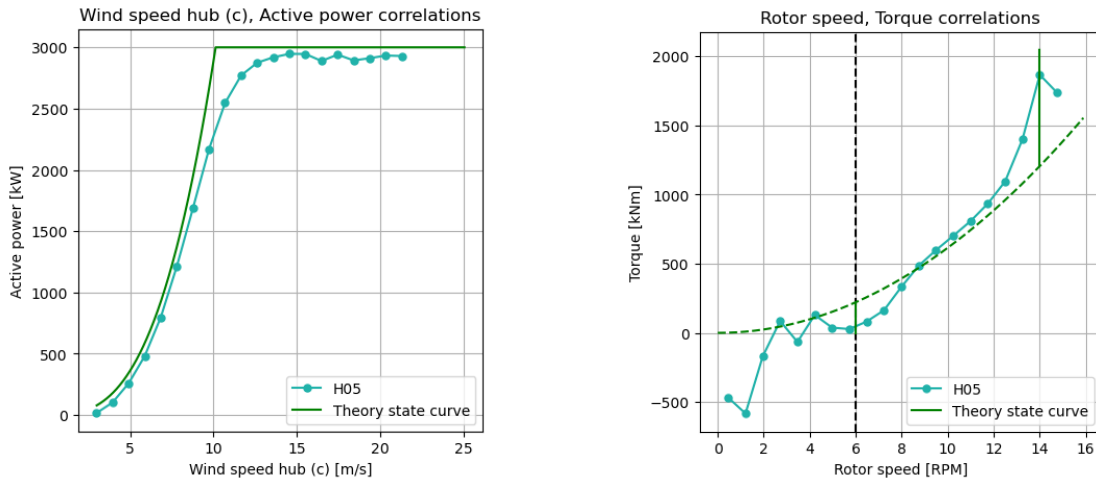
Cloud tag	X	Y	Turbine/ Cluster	Description
H05-10-355	Wind speed hub	Active power	Turbine	Scattering below power curve
H05-10-144	Wind speed hub	Pitch	Turbine	Pitching between 3 and 5 m/s
H05-10-466	Rotor speed	Torque	Turbine	6-8 RPM: below optimal torque control curve
H05-10-118	Rotor speed	Torque	Turbine	Transition region from optimal- c_p to maximum rotor speed
H05-10-352	Wind speed hub	Pitch	Turbine	Scattering above pitching curve
H05-10-572	Rotor speed	Torque	Turbine	Performance optimal- c_p region

Table 8.7: Data collection of selected clouds central cluster.

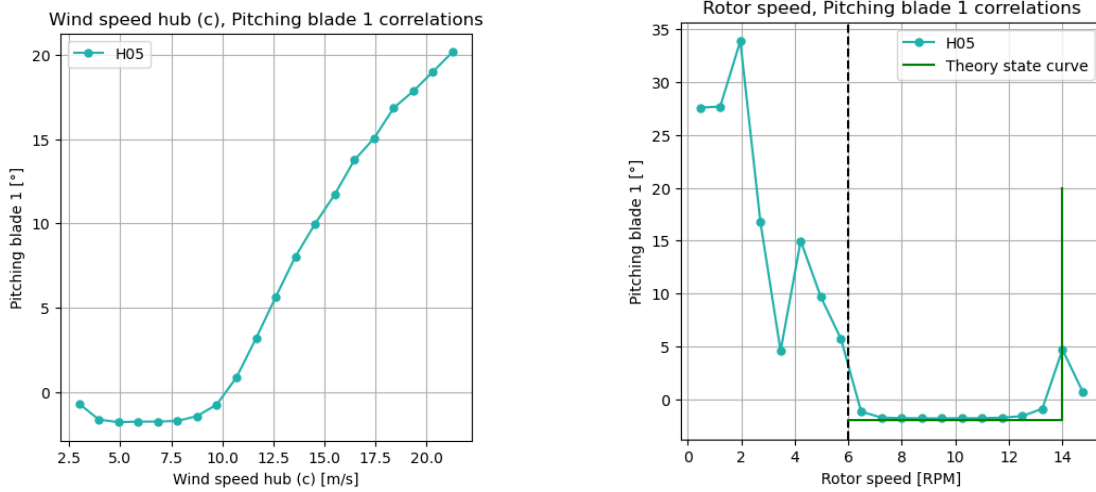
Cloud tag	Data points		Production		LPP		
	Points	%	kWh	%	kWh	% LPPT ²	% LPPC ³
H05-10-355	675	2.28	266803	4.1	45742	0.7	17.14
H05-10-144	620	2.09	1920	0.03	8507	0.13	443.09
H05-10-466	4833	16.32	68054	1.04	59789	0.92	87.86
H05-10-118	12148	41.03	4949955	75.99	165632	2.54	3.35
H05-10-352	174	0.59	49712	0.76	29996	0.46	60.34
H05-10-572	6086	20.55	836083	12.83	139305	2.14	16.66

²LPPT: Lost Potential Power as a fraction of annual power production

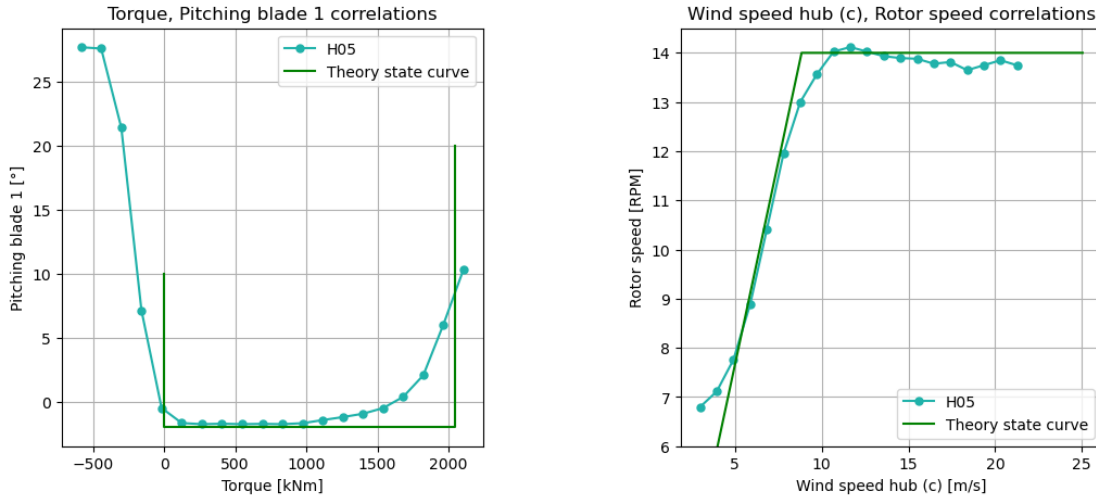
³LPPC: Lost Potential Power as a fraction of cloud power production



(a) Central cluster wind speed - active power average comparisons. (b) Central cluster rotor speed - torque average comparisons.



(c) Central cluster wind speed - pitching average comparisons. (d) Central cluster rotor speed - pitching average comparisons.



(e) Central cluster pitching - torque average comparisons. (f) Central cluster wind speed - rotor speed average comparisons.

Figure 8.16: Compared 10 minutes averaged data of central cluster turbines. Data is binned by the MOB.

When investigating the state curve relations of central cluster turbine H05, it was concluded there is no visual differing behaviour compared to the western cluster. Turbine H05 matched the state curves behaviour of the turbines in that string. One cause for this is the fact that both clusters contain the same

SWT-3.0-113 TI class B turbines. To understand the differences that can not visually be seen, the same clouds as were selected for turbine D05 are selected for the central cluster turbine H05. Differences in performance calculations of those clouds is given below.

Torque curve: Transition from optimal- c_p region to rated power

In the transition region from optimal- c_p to rated power (vertical part of the theoretical torque curve in e.g. figure 8.16f), the western cluster turbines underperformed compared to WPC. 2.54% of total annual production is the amount of LPP for this area of the torque curve (cloud H05-10-118). For the western cluster, this relative lost potential was 1.85%. There is a potential difference of 0.6% lost power in this region between the two turbines. More interesting is that this underperformance also detected for the eastern cluster, apart from the western cluster.

It is expected to see losses in this area of the torque curve, as this area makes up the majority of annual turbine data, and the western and central cluster are overall underperforming turbines. However, it is interesting to see those losses in this area for the SWT-3.0-113 turbines and not for the SWT-3.0-108 turbines.

Torque curve: performance maximum c_p region

The western cluster is, as described in chapter 6, losing a power potential of 14% compared to the production of the cloud. For the eastern cluster, this lost potential is 16% compared to cloud power production (cloud H05-10-572). Also, in this region of the torque curve, the western cluster is performing slightly better than the central cluster. The most important finding is the underperformance occurring at both of the western and central turbine clusters.

Pitching curve: pitching at low wind speeds

Both clouds H05-10-466 and H05-10-144, although separately selected, represent the pitching behaviour at low wind speeds (below 7 m/s). The selected wind speed region of this cloud is a part of the WPC provided by the original equipment manufacturer (OEM). This caused an LPPT value of 1% for the western cluster turbine D05. For the central cluster, LPPT values in this region are comparable.

Scattering below power curve

LPP below the power curve of turbine H05 is less than this was for turbine D01. Within the western cluster, turbine D01 was selected for power curve scattering as it was most visible for that turbine. LPPT was 1%. For turbine H05, LPPT due to scattering below the power curve was 0.7%.

In general, turbine H05 performs worse than the western cluster turbines. Clouds selected within crucial parts of turbine behaviour confirmed this observation when compared to similar western cluster clouds.

8.4.2. Results HF cloud analysis

HF clouds are selected for reasons as explained in section 3.1.6. In the tables below, an overview of the selected clouds is given, together with relevant cloud information and data.

In this section, reasons for selecting the tabulated clouds as well as results obtained from the multiple cloud analyses are addressed. No unique HF turbine behaviour was observed for the central cluster compared to the western cluster turbines. Therefore, the same approach as for the central cluster 10 minutes averaged cloud analysis is chosen: differences in performance statistics of comparable relevant HF clouds will be discussed. These are power curve scattering clouds, and rotor speed HF clouds exceeding the limit of 14 RPM.

Table 8.8: HF clouds analysed central cluster.

Cloud tag	X	Y	Turbine/ Cluster	Description
H05-HF-473	Wind speed hub	Active power	Cluster	Curtailments down to 2200 kW
H05-HF-082	Wind speed hub	Active power	Cluster	Curtailments cloud around 1700 kW
H05-HF-511	Wind speed hub	Active power	Cluster	Curtailments down to 500 kW
H05-HF-813	Wind speed hub	Pitch	Cluster	Pitching behaviour below 7 m/s
H05-HF-804	Rotor speed	Pitch	Cluster	Pitching at non-rated rotor speeds
H05-HF-530	Rotor speed	Torque	Cluster	Rotor speed above maximum of 14 RPM

Table 8.9: Data collection of selected clouds central cluster.

Cloud tag	Data points		Production		LPP		
	Points	%	kWh	%	kWh	% LPPT ⁴	% LPPC ⁵
H05-HF-473	11220	0.11	10754	0.17	2864	0.05	26.63
H05-HF-082	5352	0.05	3120	0.05	2196	0.04	70.38
H05-HF-511	4829	0.05	1784	0.03	3525	0.06	197.64
H05-HF-813	118886	1.16	2099	0.03	4986	0.08	237.58
H05-HF-804	4048	0.04	3555	0.06	1531	0.02	43.08
H05-HF-530	284780	2.77	385818	6.2	-22009	-0.35	-5.7

Scattering below the power curve

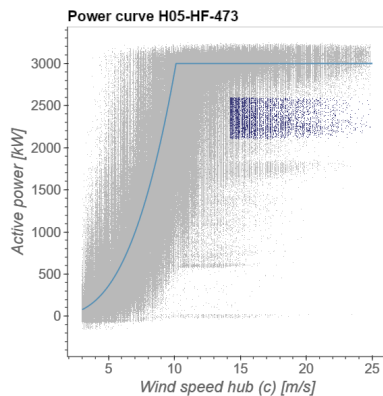


Figure 8.17: 2200 kW curtailments cloud H05-HF-473.

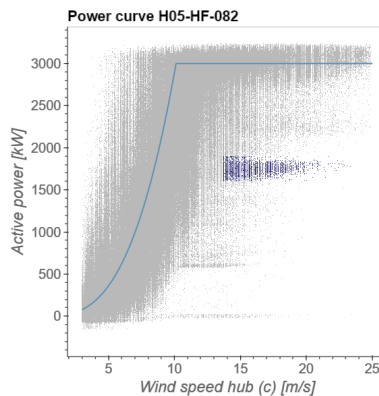


Figure 8.18: 1700 kW curtailments cloud H05-HF-082.

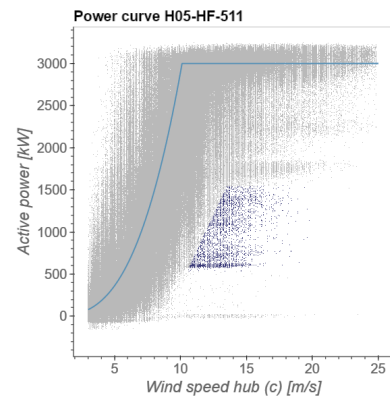


Figure 8.19: 500 kW curtailments cloud H05-HF-511.

Just as for the western cluster turbines, curtailments were observed at three different levels in HF data. By looking at the time series of the selected clouds, it was concluded the curtailments at each level had the same causes as the corresponding curtailments of the western cluster.

⁴LPPT: Lost Potential Power as a fraction of annual power production

⁵LPPC: Lost Potential Power as a fraction of cloud power production

1. Curtailments 2200 kW.
These curtailments have a LPPT value of 0.05%. This is 0.02% more than similar curtailments of the western cluster turbines.
2. Curtailments 1700 kW.
The curtailment cloud visible at 1700 kW (figure 8.18) have a LPPT of 0.04%. This is again 0.02% more compared to the similar western cluster cloud selected for turbine D05.
3. Curtailments 500 kW.
This cloud was only visible for turbine D01 of the western cluster. LPPT of that cloud is 0.21% (one of the larger HF cloud LPPT values found in this research). For turbine H05, this cloud has a LPPT value of 0.06%.

An important takeaway from these comparisons is that contributions of HF curtailment clouds to total data are comparable for western and central cluster. The 0.02% difference can be within the uncertainties of the manual selection process. Only for turbine D01, more curtailments down to 500 kW were observed (0.15% difference).

Rotor speeds exceeding maximum

From the 10 minutes averaged data, it is concluded both western and central cluster are exceeding maximum rotor speed limits set by the OEM. For both clusters, HF clouds of this rotor speed limit exceeding data were selected.

For both clusters, turbines are not underperforming in this region. The turbines meet the WPC. moreover, occurrence of these high rotor speeds are comparable. Annually, rotor speeds exceed 14 RPM for around 3% of time for both clusters.

9

Conclusion

In this chapter, the findings of the research aimed at exploring the potential causes of underperformance at the Pen y Cymoedd (PYC) wind farm are presented. The research question, “Can the underlying physical causes for the underperformance of the Pen y Cymoedd wind farm be found by the use of 10 minutes average and/or High-Frequency data?” was addressed through a comprehensive analysis of data collected from both 10 minutes average and high-frequency data sources. The results of this analysis are discussed and evaluated, ultimately providing insights into the effectiveness of using different data sources for identifying physical causes of underperformance in wind farms. Results are discussed by answering the goals from the three main sub questions as described in the introduction. At the end of this chapter, the main research question is answered.

The structured approach set up for performing an analysis including 10 minutes averaged and HF data, consists of different processes: At first, unwaked sectors for the investigated turbines are estimated. Unwaked sector wind turbine data is the input for the filtering and correlation & validation process. Correlations and validations are carried out to be able to use data measured at the turbines independently. The filtered and validated data, together with free sector data and turbine specific theoretical state curves, are input to the cloud analysis. This cloud analysis gives insights into the turbine performance in the 10 minutes averaged spectrum and high-frequency spectrum. Parallel to this cloud analysis, an overall & directional analysis is carried out to investigate the worst performing wind direction sectors of each of the selected turbines. The output of these two parallel carried out analyses is an understanding of turbine underperformance, with an introduction to the influence of the site.

Selected high-frequency clouds revealed the added value from high-frequency data sets for the understanding of wind turbine performance. Specific processes which are not highlighted in 10 minutes averaged data are highlighted in high-frequency data. High torque experienced during storm control at the eastern cluster turbines as well as derated power control at multiple power levels are examples of turbine behaviour which emerged in high-frequency data sets, while being hidden in 10 minutes averaged data sets. Secondly, high-frequency data can clarify processes which are partially visible in 10 minutes averaged data. Cut-in wind speed pitching control is an example of a high-resolution process which could only be understood by looking at data at a sampling rate of multiple seconds.

The contribution of high-frequency clouds to turbine underperformance was observed by analysing Lost Potential Power (LPP) of selected high-frequency data clouds. As high-frequency clouds are often minor parts of existing (less visible) 10 minutes averaged clouds, the LPPT¹ of high-frequency clouds is smaller than that of 10 minutes averaged clouds. However, contributions of up to 0.21% LPPT for HF clouds still show the influence of HF clouds to annual production. Moreover, of the selected HF clouds, high LPPC² values were reached (of over 100%³). This shows that the selected data has the potential of gaining production improvements. These high LPPC values are the result of the possibility to more critically select instances of underperformance within high resolution data.

Insights into the physical causes of PYC underperformance were gained by the cloud analysis in combination with the overall & directional performance analysis. To start with, turbine performance deviated between the selected clusters: the SWT-3.0-113 turbines (western and central cluster) perform worse than the investigated SWT-3.0-108 turbines (eastern cluster). This can partially be explained by more conservative settings of the SWT-1.0-113 turbine compared to the SWT-3.0-108 turbine. Apart from the differences between SWT-3.0-113 and SWT-3.0-108 performance, performance improvements of multiple percentages can potentially be won for all investigated turbines. Within the cut-in region and within storm control conditions, potential for performance improvement was observed. Moreover, Within more critical regions, such as the optimal power coefficient region of the torque curve, potential performance losses of 10%-14% compared to the cloud energy production were found. When performance can be improved in these regions, this has a big influence on annual energy production (AEP)⁴. On average, turbine performance was 95% of the Warranted Power Curve (WPC) potential production for the year 2020. This is substantially less than the gap between long-term P50 AEP estimates and average annual energy production. This means the wind resource modelling of the site possibly over-estimated the potential wind energy at the site.

Regarding insights in the site influence on turbine performance, high variety in turbine performance between different (unwaked) sectors was observed. Worst-performing sectors of -16% compared to WPC behaviour were found. Moreover, high turbulence intensity (TI) caused SWT-3.0-113 turbines to be curtailed, having negative influence on the performance of the these turbines. This occurred more at the SWT-3.0-113 turbines, as these are TI class B turbines compared to the TI class A SWT-3.0-108 turbines. High shear at the western cluster had a slight positive influence on turbine performance. Veer did not show big influences on turbine performance for all clusters.

To answer the research question “Can the underlying physical causes for the underperformance of the Pen y Cymoedd wind farm be found by the use of 10 minutes averaged and/or high-frequency data?”, it can be concluded that an approach with 10 minutes averaged and high-frequency data as inputs successfully gave insights in the underperformance of the Pen Y Cymoedd wind farm. Site influences as well as turbine behaviour and control mechanisms are compared in 10 minutes averaged resolution as well as in high-frequency resolution. Moreover, turbine behaviour and control can be compared with manufacturer information and theoretical expected behaviour. The total analysis gives better understand individual turbine and wind farm broad underperformance causes.

¹LPPT: Lost Potential Power as a fraction of annual power production

²LPPC: Lost Potential Power as a fraction of cloud power production

³Meaning that potential power output of a selected cloud can be doubled.

⁴A broad overview of the potential points of improvements are given in the results chapters of the different clusters.

10

Discussion

Within this research, 14 turbines are included in a performance analysis. After filtering and correlations and possible validations of all data sets, eventually 7 turbines are investigated in the cloud analysis and overall & directional performance analysis. This because not all turbine data could be correlated correctly to a reference data set. The conclusions of this research are strictly accountable for the turbines of this wind farm. For investigating turbines of other wind farms, the analysis steps have to be reproduced to expectedly come to a different set of results.

Results show extreme site conditions. Veer ranges to values up to $0.5^\circ/\text{m}$, whereas average wind shear factors of around 0.5 are not uncommon for the central cluster turbines. In short, both parameters reach extreme values at the MET masts. Such high average values were not found in other research and theory. According to table 2.6, urban areas experience wind shear factors of 0.4. Wind veer values are high, but are found in research performed at other complex sites.

With the results from the cloud analysis regarding turbine performance, it was shown HF data has the potential to play a role in future more critical turbine performance analysis. Moreover, it was shown that for the 14 investigated turbines, turbine underperformance can partially be responsible for the gap between production estimates, and AEP's of the operational years. Moreover, the results show that a difference in TI class can have a significant effect on turbine controls when in limiting conditions. Finally, by collecting results from this analysis for three separate clusters, it was shown the methodology is reproducible and can be applied to multiple turbine configurations.

Limitations of this research are mainly originating from the correlation process that is an unavoidable part of the research. Turbine data is validated by MET mast data, as the majority of the investigated turbines are not within a 4D distance from the MET mast. Therefore, MET mast data is the reference data in this research, and are assumed to be reliable as stand-alone measurements. Veer and shear measurements are only collected at the MET mast, and have to be assumed as reliable. The impossibility to check those measurements is a limitation of this research.

The fact turbine data is used for performance analysis is another limitation of this research. To be able to trust turbine wind speed and wind directional data, it has been correlated to a turbine-specific reference data set. This makes the turbine data usable. However, in contrast to MET mast data, turbine data is still influenced by rotor movements, adding an uncertainty to the wind speed measurements.

Moreover, the python tool created is based on manual selection of abnormal behaviour. This manual selection is a limitation on the research, as it is not as consistent as a machine learning approach, where abnormal behaviour is selected automatically by a trained model. This makes clouds comparisons between turbines not one to one, as there can be slight differences between clouds selection thresholds.

The last point regarding limitations is the filtering process. Performing a full IEC compliant filtering process was not the scope of this thesis as this is very time-consuming, and not an innovating procedure. However, due to this hybrid filtering process, site specific influences have not been filtered out of the data. This makes the site influence an extra variable in the turbine performance analysis that can not be excluded from the final results.

The implications of this research can be conversations with the OEM of the investigated turbines regarding turbine performance and possible improvements on the turbines. Another effect of this thesis is a better understanding of HF data and the usage and processing of this data. This better understanding can lead to an increase in usage of HF data for PCA work.

11

Recommendations

In this chapter, recommendations are given for measurement campaigns as well as for follow-up research.

11.1. Data quality & availability

When loading in the downloaded data sets prior to filtering and correction, not all pre-selected tags contained data. To start with, wind direction data (measured at the turbines) was only partially available for the 10 minutes averaged turbine data. Only from September 2020 on, 10 minutes averaged wind direction data was available. HF wind direction data was not available throughout the whole year. As wind direction is an important parameter for the understanding of site influence to turbine performance, it is interesting to check the availability of HF and 10 minutes averaged wind direction signals.¹

When 10 minutes averaged wind directional data was available, the data was not logged correctly for many turbines. The data consisted of a non-natural offset compared to the directional data of the reference turbine/MET mast. For example, this was the case for all turbines of the central cluster. These offsets were detected and corrected by the post-filtering code and did not cause problems in further analysis. Although this was the case, it is important to automatically flag these offsets, as these should not be part of the wind directional data tags. Flagging and correcting these tags automatically prior to the processing decreases sensitivity to errors, and saved time more downstream of the data processing.

On top of that, wind speed data was corrupted as multiple anemometer heights for the central cluster at the MET mast. For this research this did not result in bottlenecks as wind speed data at the MET mast is available at two height for all three MET masts. Two heights are of importance to validate the wind speed measurements, and to be able to correlate turbine wind speed measurements to the MET mast wind speed measurements. As the wind speed measurements for the central cluster were partially corrupted, checking these wind speed signals frequently can be of big value for PCA work.

More data that showed unexpected behaviour was the wind directional data measured at the MET mast. From the wind speed roses at the beginning of the results chapter, it is visible that no data is logged from the northerly wind direction. This is the result of wrong averaging of wind speed measurements, alternating between very small and very large angles. As this behaviour is detected for all MET masts, this is a systematic problem that might be solved by the same solution for all MET masts.

¹Wind direction signals are moreover of use for yaw misalignment studies and the filtering of waked sectors from the raw data sets.

Regarding HF data, rotor and generator speed signals are logged as discrete signals², while these are continuous signals. As this signal is discrete, it is difficult to analyse state curve relations, including a rotor or generator speed signal. When visualising a state curve relation as a scatter plot, an image appears consisting out of horizontal and vertical lines. As this is difficult to interpret, few clouds were selected, including rotor or generator speed in the HF spectrum. This is unfortunate, as relations including rotor/generator speed are often wind speed independent and can give valuable insights to turbine behaviour.

11.2. Improving curtailment filtering procedure

When filtering the raw data sets from MET masts and turbine, curtailments and alarms were one of the filtering criteria. As the scope of this research was normal operational data, curtailments and alarms have to be filtered out. Nevertheless, after the filtering process, curtailments were still clearly visible in the HF turbine data. When curtailments are hidden in data which is used for calculating turbine performance, wrong performance estimations can be the effect. For next iterations of the performance analysis as described in this research, it can be of value to investigate more resources giving information about turbine curtailments. This might lead to better curtailments filtering and more accurate performance analysis.

11.3. Filtering on standard deviation

As was also proposed in (Antinio Notaristefano, 2021), when using high resolution data, standard deviation filtering can decrease the scattering of high-frequency scatter plots. So-called stationarity conditions are applied to the data, limiting wind speed and wind direction standard deviation over a certain time delta. As a result, only stationary conditions can be studied. The wind speed and direction ramps have been filtered out, which makes non-representative clouds like L06-HF-627 be deleted from the data. Moreover, this may highlight important stationary behaviour. Applying these filters in a parallel analysis can be of value when analysing HF data sets.

11.4. Use machine learning for outlier detection

In this research, a python cloud selection tool was engineered to detect abnormal turbine performance manually. By the selection of clouds, it is possible to investigate a number of abnormal data points in various turbine relations and time series. Moreover, performance statistics of the cloud as well as parameter histograms are produced. With a machine learning approach, this manual selection could potentially be replaced by an automatised detection algorithm. This reduces time as well as repetitiveness of the analysis. A try-out by applying the k-nearest neighbours approach to a PYC power curve is shown in appendix D. For applying this to multiple turbines and wind farm conditions, a model has to be trained with different site conditions. The set-up theoretical state curves for this research can be a useful input for machine learning training purposes. A research paper describing other approaches using machine learning for underperformance detection is given by (Lyons & Göçmen, 2021).

11.5. Revise wind resource analysis at site

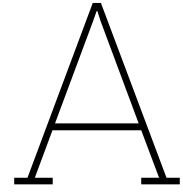
The performance of the 14, and eventually 7, investigated wind turbines was 5% less than WPC, which is a substantially smaller gap than the gap between P50 production estimates and average annual energy production of the past 5 operational years. This difference can potentially be explained by too optimistic site conditions assessments that were performed by Vattenfall and third parties. Revising these models and calculations can be of importance for understanding the unexplained gap. Important note is that the 5% underperformance is based on 1 year data for 7 turbines, and does not represent the whole operational period for all turbines. Before revising any site assessment models, it is of importance to analyse the performance of other PYC turbines (where possible), to validate the 5% performance loss found for the 7 investigated turbines.

²A signal with a step size of 0.5.

Bibliography

- Ahn, E., & Hur, J. (2022). A Practical Metric to Evaluate the Ramp Events of Wind Generating Resources to Enhance the Security of Smart Energy Systems. *Energies*, 15(7), 2676. <https://doi.org/10.3390/en15072676>
- Albers, A. (2014). *Side-by-side testing to verify improvements of power curves* (tech. rep.). DWG.
- Antinio Notaristefano, G. P., Mirko Anoja. (2021). Power curve analysis using high-resolution scada data. *Wind Europe*.
- Antoniou, I., Pedersen, S. M., & Enevoldsen, P. B. (2009). Wind Shear and Uncertainties in Power Curve Measurement and Wind Resources. *Wind Engineering*, 33(5), 449–468. <https://doi.org/10.1260/030952409790291208>
- API reference — pandas 1.5.3 documentation. (n.d.). Retrieved February 21, 2023, from <https://pandas.pydata.org/docs/reference/index.html>
- Bardal, L. M., Sætran, L. R., & Wangsness, E. (2015). Performance Test of a 3MW Wind Turbine – Effects of Shear and Turbulence. *Energy Procedia*, 80, 83–91. <https://doi.org/10.1016/j.egypro.2015.11.410>
- Beltrán, J., Cosculluela, L., Pueyo, C., & Melero, J. (2010). Comparison of measure-correlate-predict methods in wind resource assessments. *Proc. EWEC, Warsaw*.
- Burton, T., Jenkins, N., Sharpe, D., & Bossanyi, E. (2011). *Wind Energy Handbook* (1st ed.). Wiley. <https://doi.org/10.1002/9781119992714>
- Cambron, P., Lepvrier, R., Masson, C., Tahan, A., & Pelletier, F. (2016). Power curve monitoring using weighted moving average control charts. *Renewable Energy*, 94, 126–135. <https://doi.org/https://doi.org/10.1016/j.renene.2016.03.031>
- Carullo, A., Ciocia, A., Malgaroli, G., & Spertino, F. (2021). An Innovative Correction Method of Wind Speed for Efficiency Evaluation of Wind Turbines. *ACTA IMEKO*, 10(2), 46. https://doi.org/10.21014/acta_imeko.v10i2.1037
- Ebert, P., & Wood, D. (1997). Observations of the starting behaviour of a small horizontal axis wind turbine. *Renewable Energy*, 12(3), 245–257. [https://doi.org/10.1016/S0960-1481\(97\)00035-9](https://doi.org/10.1016/S0960-1481(97)00035-9)
- Gonzalez, E., Stephen, B., Infield, D., & Melero, J. J. (2017). On the use of high-frequency SCADA data for improved wind turbine performance monitoring. *Journal of Physics: Conference Series*, 926, 012009. <https://doi.org/10.1088/1742-6596/926/1/012009>
- Gonzalez, E., Stephen, B., Infield, D., & Melero, J. J. (2019). Using high-frequency SCADA data for wind turbine performance monitoring: A sensitivity study. *Renewable Energy*, 131, 841–853. <https://doi.org/10.1016/j.renene.2018.07.068>
- Google Earth. (n.d.). Retrieved July 24, 2022, from <https://earth.google.com/web/>
- Habibi, H., Howard, I., & Simani, S. (2019). Reliability improvement of wind turbine power generation using model-based fault detection and fault tolerant control: A review. *Renewable energy*, 135, 877–896.
- Hau, E. (2013). *Wind Turbines*. Springer. <https://doi.org/10.1007/978-3-642-27151-9>
- Holmes, J. D. (2007). *Wind loading of structures* (2nd ed) [OCLC: ocm68786667]. Taylor & Francis.
- IEA. (2020). *Iea wind tcp annual report 2020* (tech. rep.). IEA.
- IEC. (2019). *Wind energy generation systems - part 1: Design requirement* (tech. rep. IEC 61400). The International Electrotechnical Commission. 3, rue de Varembé, Case postale 131, CH-1211 Genève 20, Switzerland.
- Kaniewski, T. (2022). High frequency scada data for performance analysis of multiple wind turbines in complex terrain.
- Lee, J. C., & Fields, M. J. (2021). An overview of wind-energy-production prediction bias, losses, and uncertainties. *Wind Energy Science*, 6(2), 311–365.
- Lio, W. H. (2018). *Blade-Pitch Control for Wind Turbine Load Reductions*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-75532-8>

- Lio, W. H., Mirzaei, M., & Larsen, G. C. (2018). On wind turbine down-regulation control strategies and rotor speed set-point. *Journal of Physics: Conference Series*, 1037, 032040. <https://doi.org/10.1088/1742-6596/1037/3/032040>
- Lyons, J. T., & Göçmen, T. (2021). Applied machine learning techniques for performance analysis in large wind farms. *Energies*, 14(13), 3756.
- Markou, H., & Larsen, T. J. (2009). Control strategies for operation of pitch regulated turbines above cut-out wind speeds. *2009 European Wind Energy Conference and Exhibition, EWEC*, 16–19.
- Module: Restauration. (n.d.). Retrieved October 7, 2022, from https://scikit-image.org/docs/stable/api/skimage.restoration.html#skimage.restoration.denoise_tv_chambolle
- Mortensen, N. (2012). *Planning and development of wind farms: Wind resource assessment and siting*.
- Murphy, P., Lundquist, J. K., & Fleming, P. (2020). How wind speed shear and directional veer affect the power production of a megawatt-scale operational wind turbine. *Wind Energy Science*, 5(3), 1169–1190. <https://doi.org/10.5194/wes-5-1169-2020>
- Novaes Menezes, E. J., Araújo, A. M., & Bouchonneau da Silva, N. S. (2018). A review on wind turbine control and its associated methods. *Journal of Cleaner Production*, 174, 945–953. <https://doi.org/10.1016/j.jclepro.2017.10.297>
- Øistad, I. S. (2015). *Analysis of the turbulence intensity at skipheia measurement station* (Master's thesis). NTNU.
- Plot tools. (n.d.). Retrieved March 9, 2023, from https://docs.bokeh.org/en/latest/docs/user_guide/interaction/tools.html
- Ray, M., Rogers, A., & McGowan, J. (2006). Analysis of wind shear models and trends in different terrains. *University of Massachusetts, Department of Mechanical and Industrial Engineering, Renewable Energy Research Laboratory*.
- Sanchez Gomez, M., & Lundquist, J. K. (2020). The effect of wind direction shear on turbine performance in a wind farm in central Iowa. *Wind Energy Science*, 5(1), 125–139. <https://doi.org/10.5194/wes-5-125-2020>
- Schouten, L. (2021). *Data structure for the pca* (tech. rep.). Vattenfall.
- Sohoni, V., Gupta, S. C., & Nema, R. K. (2016). A Critical Review on Wind Turbine Power Curve Modelling Techniques and Their Applications in Wind Based Energy Systems. *Journal of Energy*, 2016, e8519785. <https://doi.org/10.1155/2016/8519785>
- Stival, L. J. L., Guetter, A. K., & de Andrade, F. O. (2017). The impact of wind shear and turbulence intensity on wind turbine power performance. *Espaço Energia*, 27, 11–20.
- Sun, Q., Liu, C., & Zhen, C. (2019). Abnormal Detection of Wind Turbine Operating Conditions Based on State Curves. *Journal of Energy Engineering*, 145(5), 06019001. [https://doi.org/10.1061/\(ASCE\)EY.1943-7897.0000612](https://doi.org/10.1061/(ASCE)EY.1943-7897.0000612)
- Tücer, R. (2016). Investigation of potential reasons to account for the underperformance of an operational wind farm.
- Vahidzadeh, M., & Markfort, C. D. (2019). Modified Power Curves for Prediction of Power Output of Wind Farms. *Energies*, 12(9), 1805. <https://doi.org/10.3390/en12091805>
- Vanderwende, B. J., & Lundquist, J. K. (2012). The modification of wind turbine performance by statistically distinct atmospheric regimes. *Environmental Research Letters*, 7(3), 034035. <https://doi.org/10.1088/1748-9326/7/3/034035>
- van Wingerden, J.-W. (2021). *Control of variable speed wind turbines* [Slides of the course Wind Turbine Design at TU Delft].
- Wilkinson, M. (2016). *Use of higher frequency scada data for turbine performance optimisation* (tech. rep.). DNV.
- Zaaijer, M., & Viré, A. (2021). *Introduction to wind turbines: Physics and technology*. TU Delft.



Defective wind vane measurements at MET masts

By the figure below, the cause of the defective northerly wind direction measurements and veer calculations is explained. When high resolution wind vane measurements at the turbine alternate from 0 to 360, the wind vanes of the MET mast (blue and yellow dots) show random values between 0 and 360. These values are not representing the wind direction well. The values do not represent the wind direction correctly, because the averaged wind direction of an array with values close to 0 and 360 degrees is not calculated properly at the MET mast. As wind veer is calculated from these defective directional measurements, wind veer is not reliable as well these wind directions.

To not be hindered by this defective logging, MET mast data is filtered for turbine wind directions coming from the (true) north, $\pm 50^\circ$ when investigating wind veer at the site.

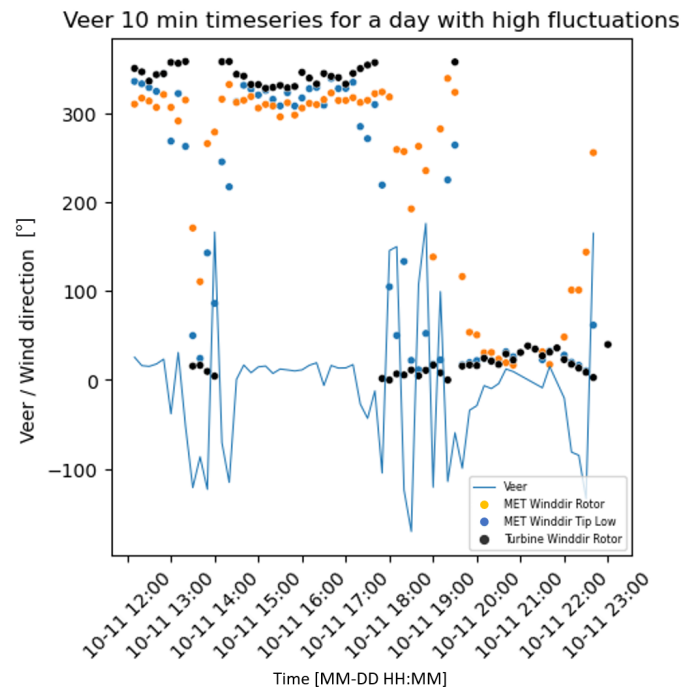


Figure A.1: The origin of wrong wind vane and wind direction measurements/calculations at the MET mast.

B

Correlations turbines western cluster

In section 6.2, the correlations between the MET mast and neighbouring turbine D05 were shown to give a visual representation of their correlations. In this appendix, visualisations of the other western cluster correlations are given for completeness. For the other clusters, the correlations will be described by slopes, intersections and R^2 -values in tables.

B.1. MET-D05 correlations

10 min wind speed correlation MET and D05

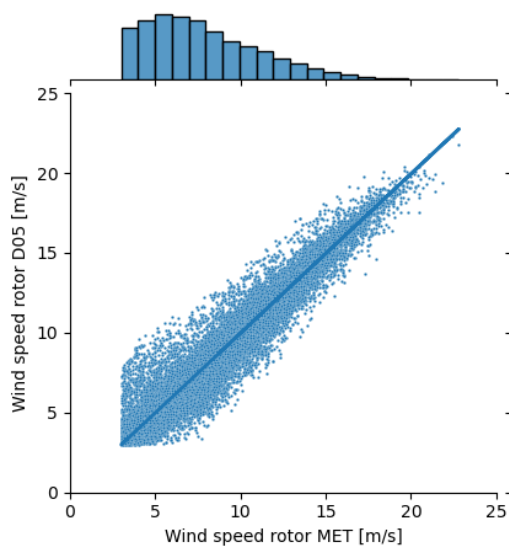


Figure B.1: 10 minutes averaged wind speed correlation of MET and D05

10 min wind direction correlation MET and D05

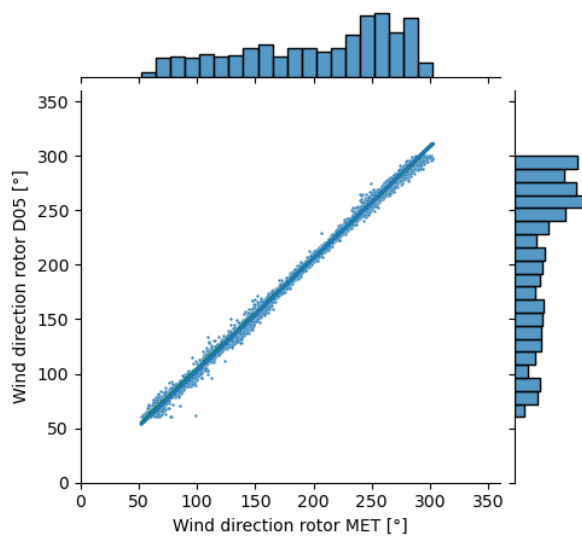


Figure B.2: 10 minutes averaged wind direction correlation of MET and D05

B.2. D04-D03 correlations

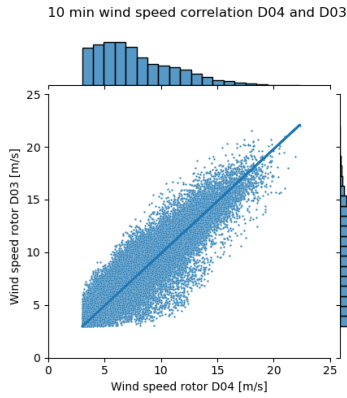


Figure B.3: 10 minutes averaged wind speed correlation of D04 and D03

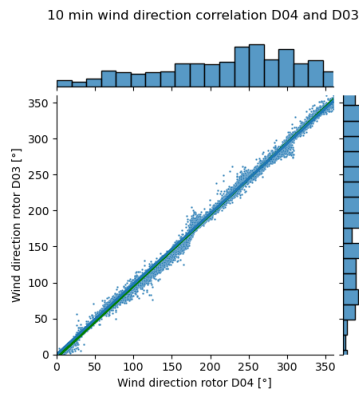


Figure B.4: 10 minutes averaged wind direction correlation of D04 and D03

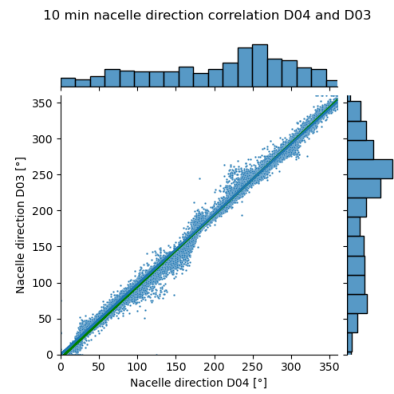


Figure B.5: 10 minutes averaged nacelle direction correlation of D04 and D03

B.3. D03-D01 correlations

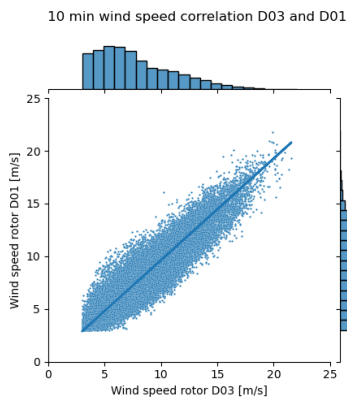


Figure B.6: 10 minutes averaged wind speed correlation of D03 and D01

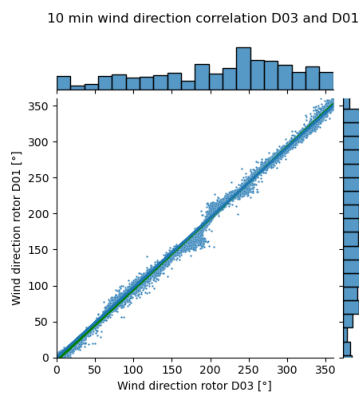


Figure B.7: 10 minutes averaged wind direction correlation of D03 and D01

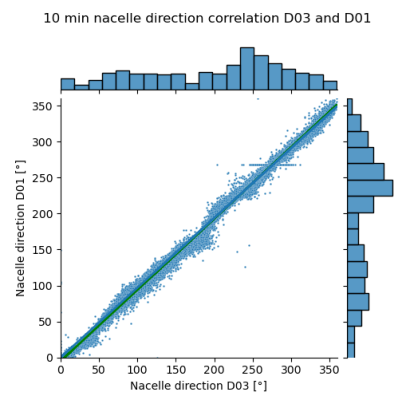


Figure B.8: 10 minutes averaged nacelle direction correlation of D03 and D01

B.4. Correlations wind direction - nacelle direction

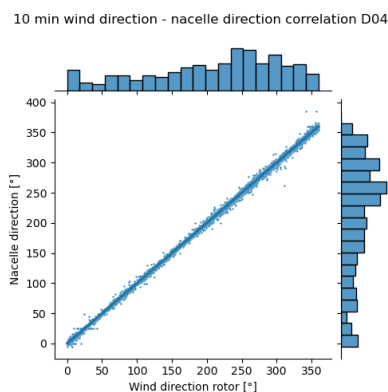


Figure B.9: Wind direction - nacelle direction correlation turbine D04

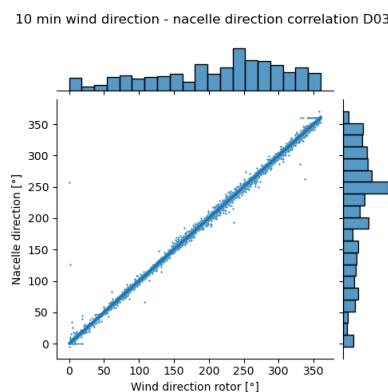


Figure B.10: Wind direction - nacelle direction correlation turbine D03

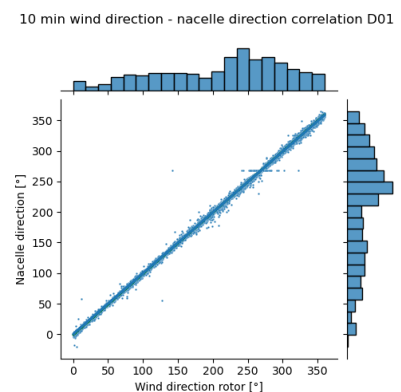
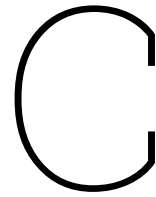


Figure B.11: Wind direction - nacelle direction correlation turbine D01



Algorithm step detection for correcting offset turbine D04

The effect of the jump detection algorithm applied in section 6.2 is shown in figures C.1 and C.2. The algorithm used for detecting the step is the Chambolle denoising algorithm from the Scikit API ('Module: restauration', n.d.). The wind direction difference between turbines D04 and D05 for the full year 2020, including jump, is shown on the left. On the right, the data has been corrected from the detected step (blue vertical line) onwards.

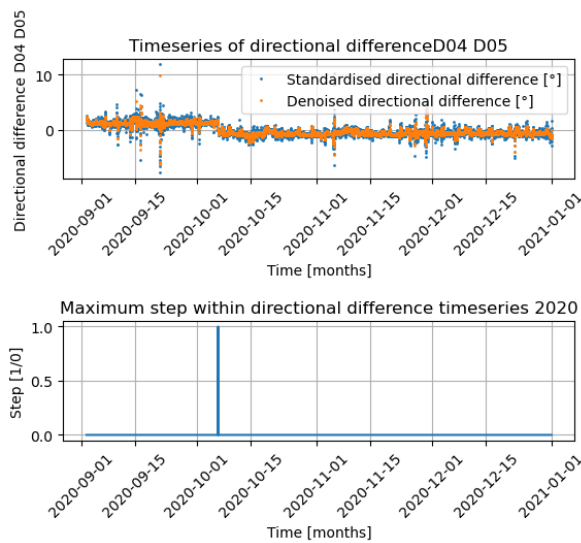


Figure C.1: Output step detection algorithm turbine D04 before correction.

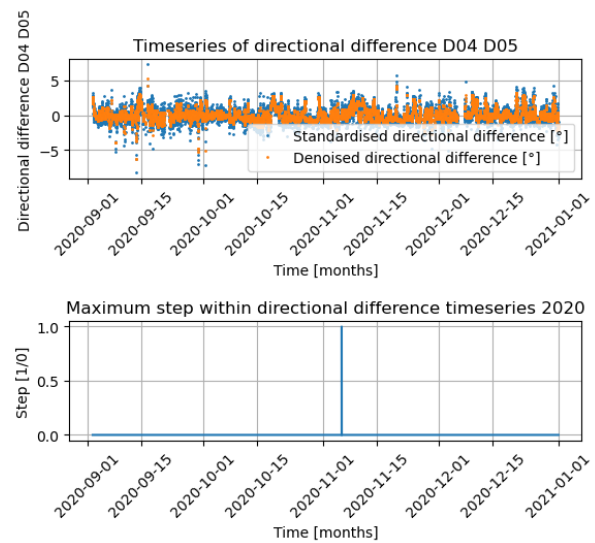


Figure C.2: Output step detection algorithm after correction.

D

Machine learning for detecting outliers on D05 operational active power data

Not part of the report, but of the recommendations (section 11.4) is the topic of using machine learning for detecting underperformance. Literature consulted for this purpose can be found in chapter 2. Out of curiosity about the applicability of machine learning, the Local Outlier Factor algorithm was applied on the operational power curve data of turbine D05. This algorithm is based on the k-nearest neighbours algorithm. A first training in detecting outliers from an operational power curve of turbine D05 was carried out. To get to the outliers of the figure below, an iterative process was executed to get to the optimal number of neighbours, leading to the best outlier detection results.

In the figures below, on the left, the red circles represent the outlier score from the algorithm. On the right, the detected outliers with a score above a certain threshold are shown in black. This try-out is a first and quick iteration of the total process. It does not show any usable results.

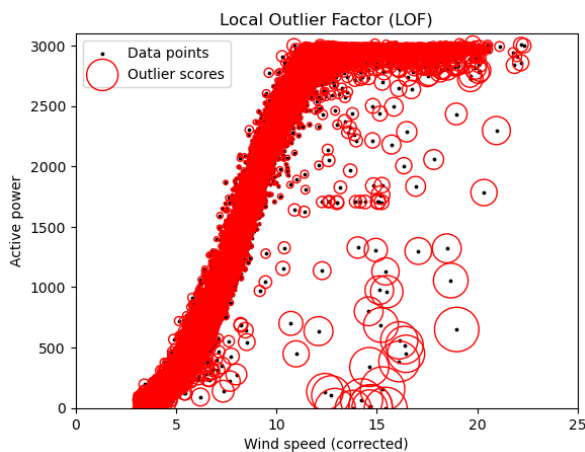


Figure D.1: Outlier scores for 10 minute power curve data for turbine D05.

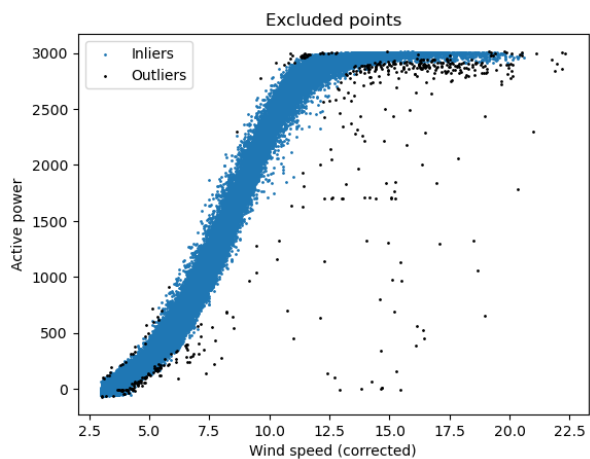


Figure D.2: Included and excluded points after applying model.

E

Example of cloud analysis tool output

In the images below, an example of the Cloud analysis tool outputs is given. The cloud analysis tool is described in section 3.1.6. In the plots, the cloud is represented by blue(yellow), and the total annual 2020 data is represented by gray. The time series and state curve relations are dynamic plots¹. The histograms are not. It is possible to choose between density colouring of the cloud, or represent the cloud by a dark blue colour.

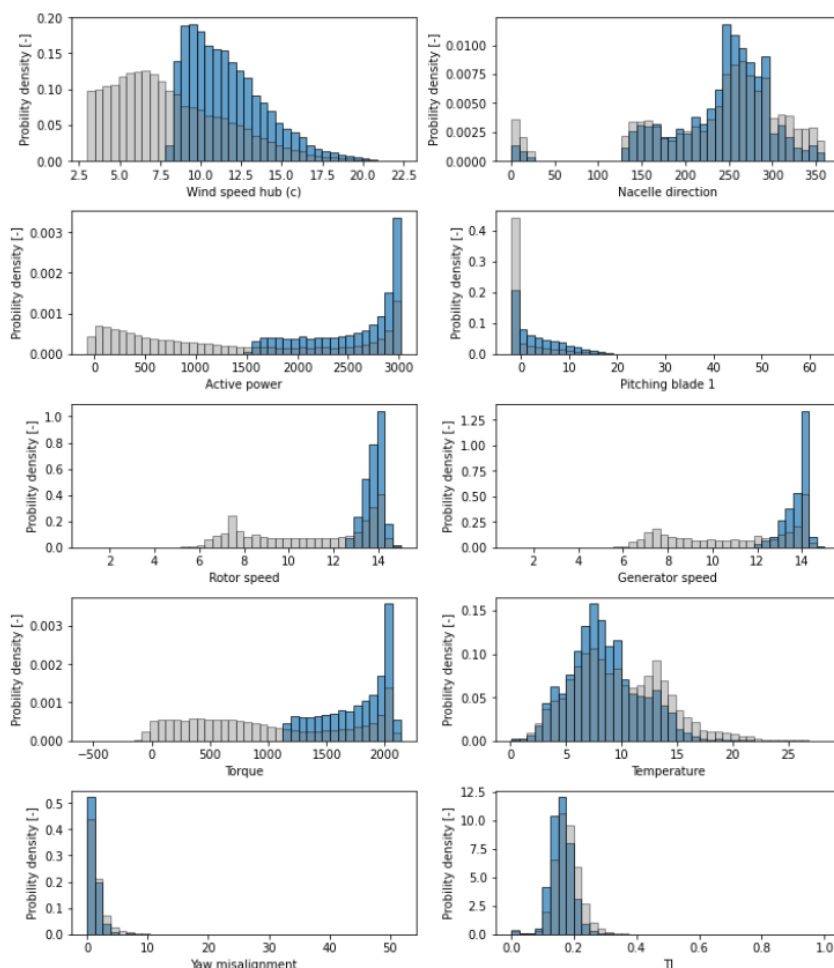


Figure E.1: Histograms output from cloud analysis tool.

¹Zooming inside an image is possible. All images containing the same data change accordingly.

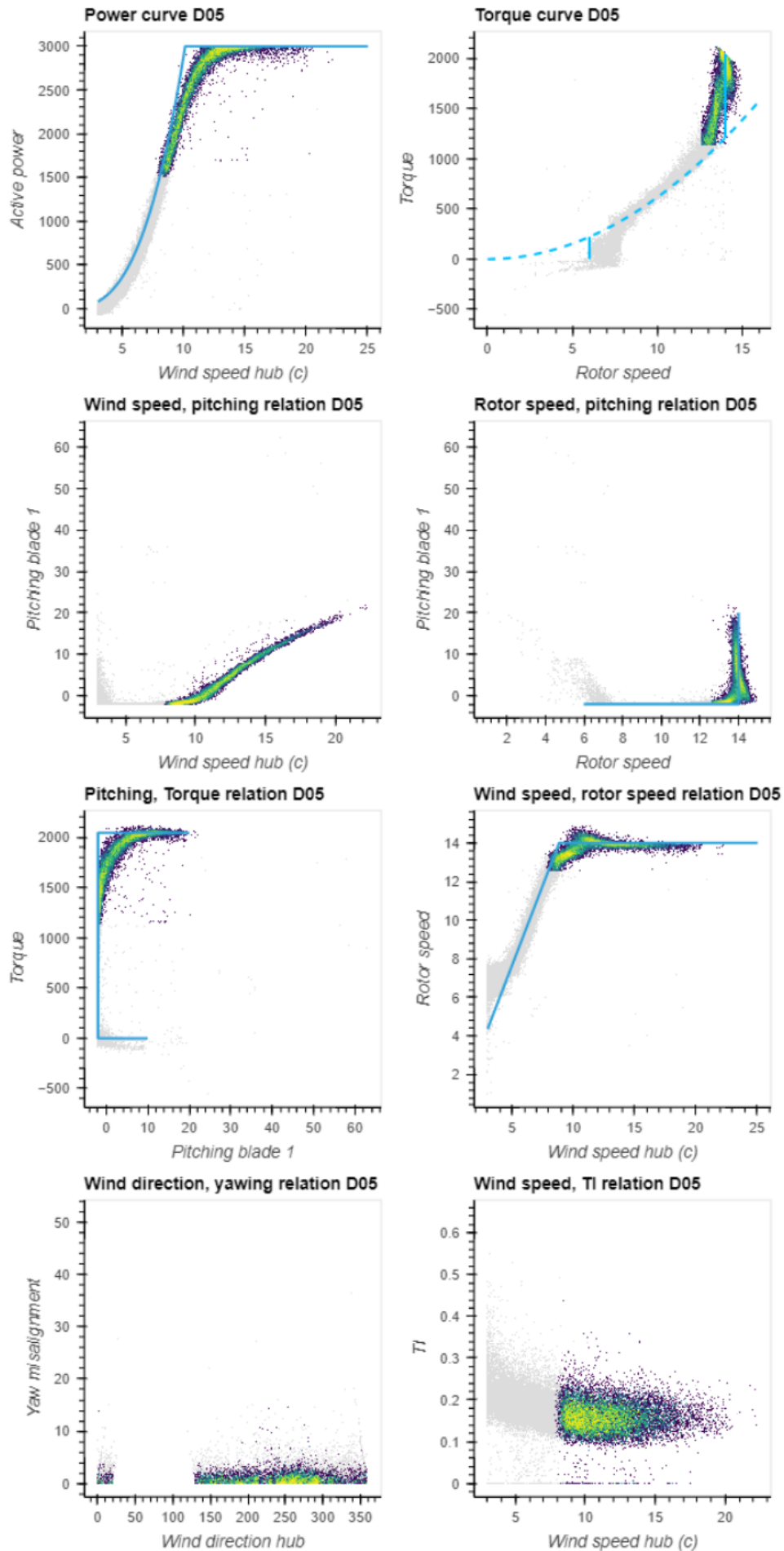


Figure E.2: Scatter plots output from cloud analysis tool.

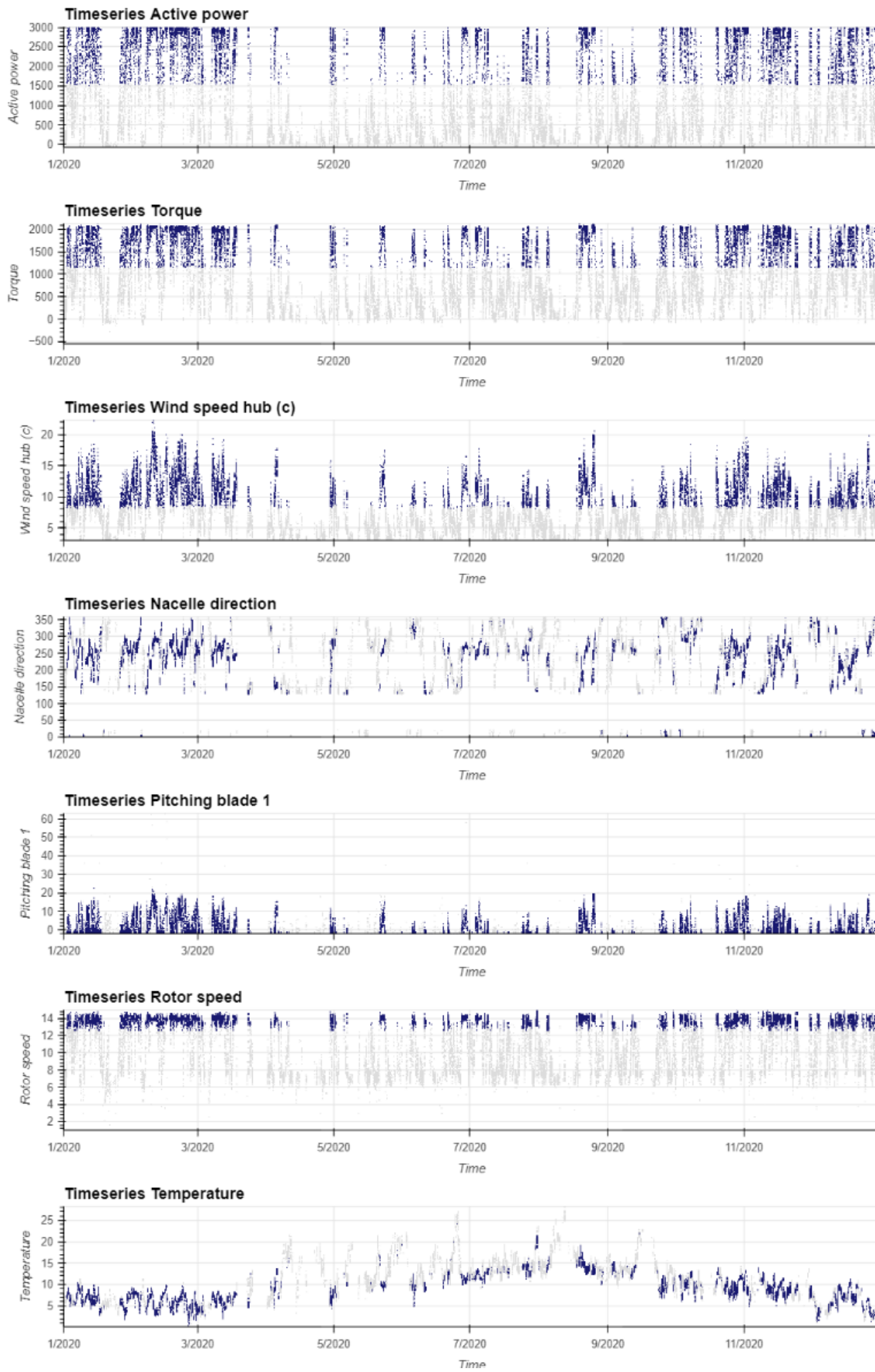


Figure E.3: Time series output from cloud analysis tool.

F

Code

As can be seen in figure 3.1, Python is the programming language used for writing the programming model for carrying out all steps of the methodology. This is done by a combination of the Spyder interface, and Jupyter Notebook interface. Below, important pieces of code are highlighted, as the full code is too large to fit into the appendix.

F.1. Script for filtering raw turbine data

This code filters raw 10 minutes averaged and HF turbine data as described by the 16 steps in chapter 3.1. It imports created functions from a master script to complete all filtering steps.

```
1 start = time.time()
2
3 data_10_turb = pd.read_csv(filename+"\\\\"+raw_data_name)
4
5 data_HF_1=pd.read_csv(filename+"\\UPY"+wtgID+"_HF.csv")
6 data_HF_1[['Timestamp']] = data_HF_1[['Timestamp']].apply(pd.to_datetime)
7 data_HF_1['Timestamp'] = data_HF_1['Timestamp'].dt.round('S')
8 data_HF_1['Timestamp'] = data_HF_1['Timestamp'].dt.tz_localize(None)
9
10 print('loading turbine '+wtgID+' HF data completed')
11
12 col=list(filter(lambda k: wtgID in k, data_10_turb.columns.tolist())) # from all columns,
    filter out the ones not equal to the turbine chosen.
13 col.append('time') #time column added
14 data_10_turb.rename(columns = {'Unnamed: 0':'time'}, inplace = True)
15 data_10_1 = data_10_turb[col]
16 data_10_1.time = pd.to_datetime(data_10_1.time)
17 data_10_1.time=data_10_1.time.dt.tz_localize(None)
18
19 data_HF_1 = pd.pivot_table(data_HF_1,index='Timestamp', columns='Tag',values='Value')
20 data_HF_1 = data_HF_1.ffill()
21 data_HF_1['dT'] = data_HF_1.index.to_series().diff()
22 data_HF_1['dT'] = data_HF_1['dT'].dt.total_seconds()
23 data_HF_1['dT'][:-1] = data_HF_1['dT'][1:]
24 data_10_1 = TI_2(data_HF_1, 'WMET01_HorWdSpd_mag', data_10_1, wtgID )
25
26 del data_10_turb
27
28 data_10_1 = power_curve_correction(data_10_1, wtgID, rho_wp, 'UPY'+wtgID+'-WindSpeed') #
    function from master file
29
30 data_10_1_clean = filterig_a(data_10_1, allarms_raw, 'UPY'+wtgID)
31 data_10_1_clean = data_10_1_clean.loc[(data_10_1_clean['UPY'+wtgID+'-PY-CurtailLossTotal'] ==
    0)] #changed
32 data_10_1_clean = data_10_1_clean.loc[(data_10_1_clean['UPY'+wtgID+'-PY-FaultLossTotal'] ==
    0)] #changed
33 data_10_1_clean = data_10_1_clean.dropna(axis = 0, how = 'any', subset=['UPY'+wtgID+'-
    WindSpeed'])
```



```

34 data_10_1_clean = data_10_1_clean[(data_10_1_clean['UPY'+wtgID+'-PitchRef_BladeAAvg'] < 90) |
    (np.isnan(data_10_1_clean['UPY'+wtgID+'-PitchRef_BladeAAvg']))]
35
36 data_10_1_clean = filter_icing(data_10_1_clean)
37 data_10_1_clean = frozen_logging(data_10_1_clean)
38
39 data_10_1_clean.loc[:, wtgID+'_Dir_diff'] = wdir_diff(data_10_1_clean['UPY'+wtgID+'-
    WindDirAvg'], data_10_1_clean['UPY'+wtgID+'-NacelleDirection'])
40
41 data_HF_1_clean = filterig_HF_a(data_HF_1, allarms_raw, 'UPY'+wtgID,30)
42 data_HF_1_clean = filtering_c2(data_HF_1_clean, data_10_1, wtgID, 30)
43 data_HF_1_clean = filtering_a2(data_HF_1_clean, data_10_1, wtgID, 30)
44 data_HF_1_clean = data_HF_1_clean[(data_HF_1_clean['WROT01_B1PthAngTgt_mag'] < 90) | (np.
    isnan(data_HF_1_clean['WROT01_B1PthAngTgt_mag']))]
45
46 data_HF_1_clean = filter_icing(data_HF_1_clean)
47 data_HF_1_clean = frozen_logging(data_HF_1_clean)
48
49 del data_HF_1
50
51 data_HF_1_clean = power_curve_correction(data_HF_1_clean, wtgID, rho_wp, '
    WMET01_HorWdSpd_mag')
52
53 data_10_1_clean['Torque'] = torque_calc(data_10_1_clean, 'UPY'+wtgID+'-RotorRPM', 'UPY'+wtgID
    +'-ActivePower')
54 data_HF_1_clean['Torque'] = torque_calc(data_HF_1_clean, 'WROT01_RotSpd_mag', 'WTUR01_W_mag')
55
56 data_HF_1_clean = data_HF_1_clean.reset_index()
57
58 data_HF_1_clean_vel = velocity_filter(data_HF_1_clean, "WMET01_HorWdSpd_mag", cut_in =3 ,
    cut_off=25)
59 data_10_1_clean_vel = velocity_filter(data_10_1_clean, "UPY" + wtgID + "-WindSpeed", cut_in
    =3 , cut_off=25)
60
61 del data_HF_1_clean
62
63 data_HF_1_clean_vel.to_csv(filename_save+'\\'+wtgID[:-2]+'\\data_ff30_HF_'+wtgID+'_clean_vel.
    csv')
64 data_10_1_clean_vel.to_csv(filename_save+'\\'+wtgID[:-2]+'\\data_10_'+wtgID+'_clean_vel.csv')
65
66 # data_10_1_clean_vel.to_parquet(filename_save+'\\'+wtgID[:-2]+'\\data_10_1_clean_vel_parquet
    ')
67
68 del data_HF_1_clean_vel
69
70 end = time.time()
71
72 print('Time for filtering turbine '+wtgID+' = ',(end-start)/60,' minutes')

```

F.2. Two examples of functions from master filtering script

The two functions below were written to calculate TI efficiently from HF data, and detect icing periods, respectively. The functions correspond to step 11 and 7 of the 16 filtering steps in chapter 3. The @njit decorator tries to perform calculations in a non-python mode, which increases computational speeds.

```

1 def TI(data, tag_hf, data_10, tag_10, wtg_id):
2     dataHF = data.copy()
3     data10 = data_10.copy()
4
5     #creating timestamps out of datetime series for @njit
6     dataHFtime = [time.mktime(t.timetuple()) for t in dataHF.index]
7     dataHFtime = np.array(dataHFtime)
8     data10time = [time.mktime(t.timetuple()) for t in data10.time]
9     data10time = np.array(data10time)
10    dataHF_tag = dataHF[tag_hf].to_numpy()
11
12    @njit
13    def selecting():
14        std_total = np.zeros(len(data10time))
15        x = 1
16        for i in range(len(data10time)): #for testing
17            # for i in range(len(data10time)):
18                coll = ((dataHFtime >= data10time[i]) & (dataHFtime <= data10time[i+1]))
19                collection = dataHF_tag[coll]
20                collection = collection[collection>0]
21                if len(collection) == 0:
22                    std = 0
23                else:
24                    std = collection.std()
25                print(std)
26                std_total[x] = std
27                x = x+1
28                print(str(i+1)+' out of '+str(len(data10time)))
29        return std_total
30    std_total = selecting()
31    data10['TI_'+wtg_id] = std_total/data10[tag_10]
32    data10['TI_'+wtg_id] = np.where(data10['TI_'+wtg_id]==0, np.nan, data10['TI_'+wtg_id]) #
33    change all 0's in nan
34    return data10

```

```

1 def filter_icing(dataset):
2     dir_cols = np.array([col for col in dataset.columns if any(x in col for x in ['WindDir',
3     direction', 'WdDir'])])
4     spd_cols = np.array([col for col in dataset.columns if any(x in col for x in ['Speed',
5     WdSpd'])])
6     #print(dir_cols)
7     for i in range(len(dir_cols)):
8         dataset['diff_'+str(i)] = abs(dataset[dir_cols[i]].diff())
9     diff_cols = np.array([col for col in dataset.columns if any(x in col for x in ['diff_'])
10    ])
11    boolean_diff = dataset[diff_cols] == 0
12    boolean_spd = dataset[spd_cols] == 0
13    boolean_diff = np.array(boolean_diff.any(axis='columns'))
14    boolean_spd = np.array(boolean_spd.any(axis='columns'))
15    boolean_diff = ~boolean_diff
16    boolean_spd = ~boolean_spd
17    dataset = dataset[np.array(boolean_diff) & np.array(boolean_spd)]
18    dataset = dataset.drop(diff_cols, axis=1)
19    return dataset

```

F.3. Example of function from analysis script

Figures 6.23 - 6.26 are the outputs of the function `performance_sectors()` for the western cluster turbines. This function is part of the analysis script of this thesis.

```

1
2 def performance_sectors(data_turb, turbine, WPC, mean_data, dirbinsbinary, reference):
3
4     PPI1 = [] #second method for calculating PPI based on OPC
5     PPI2 = [] #second method for calculating PPI based on Active power signal (same outcome)
6
7     dhist, dbins = np.histogram(data_turb["Wind direction hub"], bins=int(360/binsize), range
8         =(0,360), density=True)
9     dhist = dhist*binsize
10
11     phist = []
12     for i in DirBins:
13         data = data_turb[(data_turb["Nacelle direction"]>=i) & (data_turb["Nacelle direction"
14             ]<i+binsize)]
15         power = np.sum(data['Active power']*(1/6))
16         phist.append(power)
17     phist_rel = phist/(np.sum(data_turb['Active power']*(1/6)))
18
19     for count, j in enumerate(DirBins):
20
21         #first calculation of power production per sector (based on OPC)
22         v_hist, v_bins = np.histogram(data_turb["Wind speed hub (c)"][(data_turb["Nacelle
23             direction"]>j) & (data_turb["Nacelle direction"]<j+binsize)], bins=30, range
24             =(-0.5,29.5), density=True)
25         prod1 = mean_data[count,:]*v_hist[3:26]
26         warr1 = WPC*v_hist[3:26]
27         rel_perf1 = (prod1/warr1)*100 - 100
28         PPI1.append(rel_perf1)
29
30         #second calculation of power production per sector (based on active power)
31         data_sect = data_turb[(data_turb["Nacelle direction"]>j) & (data_turb["Nacelle
32             direction"]<j+binsize)]
33         prod2 = np.sum(data_sect['Active power'] * (1/6) )
34         warr2 = (WPC*v_hist[3:26]) * len(data_sect) * (1/6)
35         rel_perf2 = (prod2/warr2)*100 - 100
36         PPI2.append(rel_perf2)
37
38     PPI_tot = PPI1*phist_rel
39
40     #make array compatible for wake sector plotting
41     plt.figure()
42     plt.grid()
43     plt.bar(Dir_array, PPI1)
44     plt.bar(Dir_array, np.array(dirbinsbinary)*max(abs(np.array(PPI1))), alpha=0.2, color='r'
45         )
46     plt.bar(Dir_array, np.array(dirbinsbinary)*-1*max(abs(np.array(PPI1))), alpha=0.2, color=
47         'r')
48     plt.axhline(0, color='black', linestyle='--', linewidth=1)
49     plt.xticks(rotation = 90)
50     plt.title('Performance of directional bins with respect to WPC turbine '+turbine)
51     plt.xlabel('Wind direction [°]')
52     plt.ylabel('Power gain/loss [%]')
53     plt.legend(['Net zero', 'Power gain/loss', 'Wakes'], loc='upper right')
54
55     return PPI1, PPI_tot

```

F.4. Turbine class as input to Cloud analysis tool

The code below is a class describing the turbine, and it's key properties, including theoretical state curve behaviour. These curves are created for a static plotting interface (Matplotlib) as well as for a dynamic plotting interface (Bokeh).

```

1 class Turbine:
2     def __init__(self, P_rated, x_p, y_p, A, rho_design, D,
3                 U_rated, Urotor_rated, U_cut_out, TSR_real):
4         #Get x_p and y_p from partial load region of the power curve (around 8 m/s).
5         #TSR_real: not calculated from rated wind speed, as there is no max. power production
6             in that point.
7
8         self.P_rated = P_rated #[W]
9         self.A = A #[m2]
10        self.rho_design = rho_design #[kg/m3]
11        self.R = D/2
12        self.U_rated = U_rated #[m/s]
13        self.W_rated_rpm = Urotor_rated #[RPM]
14        self.W_rated_rad = self.W_rated_rpm * (2*np.pi/60) #[rad/s]
15        self.TSR = self.W_rated_rad * self.R / self.U_rated #[-]
16        self.cp = y_p * 1000 / (0.5 * self.rho_design * self.A * (x_p**3)) #[-]
17        self.U_co = U_cut_out #[m/s]
18        self.Torque_rated = self.P_rated / (self.W_rated_rad * 1000)
19        #calculate theoretical rated power point.
20        self.U_switch = (self.P_rated / (0.5 * self.rho_design * self.A * self.cp)) ** (1/3)
21        #Values maching with rated region in plots
22        self.TSR_real = TSR_real #TSR according to plots in rated region
23        self.corr_fact = (0.6 * 2 * np.pi / 60)
24        #Rated wind speed according to TSR_real
25        self.U_rated_real = (self.W_rated_rad + self.corr_fact) * self.R / TSR_real
26
27        #STATIC
28
29        def plot_U_Power(self):
30
31            x1 = np.linspace(3, self.U_switch, 1000)
32            x2 = np.arange(self.U_switch, 25.1, 0.1)
33            x = np.concatenate((x1, x2))
34
35            y1 = 0.5 * self.cp * self.A * self.rho_design * (x1**3)
36            y2 = np.zeros(len(x2)) + self.P_rated
37
38            y1 = y1 / 1000 #[kW]
39            y2 = y2 / 1000 #[kW]
40            y = np.concatenate((y1, y2))
41
42            plt.plot(x, y, 'g', label='Theory state curve')
43
44            # e = self.P_rated / 1000 * 0.04
45            # plt.fill_between(x, y-e, y+e, alpha=0.2, facecolor='g', edgecolor='g',
46                # linewidth=2, linestyle='dashed', antialiased=True)
47
48        def plot_W_Torque(self):
49            k_opt = 0.5 * self.rho_design * self.cp * np.pi * (self.R**5) / (self.TSR_real**3)
50            W_range = np.arange(0, self.W_rated_rpm + 2, 0.1) #[RPM]
51            Torque1 = k_opt * ((W_range * 2 * np.pi / 60)**2) / 1000 #[kNm]
52            plt.plot(W_range, Torque1, 'g', linestyle='dashed')
53
54            # plot behaviour at Wmin and Wmax
55            p1 = k_opt * ((6 * 2 * np.pi / 60)**2) / 1000
56            p2 = k_opt * ((self.W_rated_rpm * 2 * np.pi / 60)**2) / 1000
57            plt.plot([6, 6], [0, p1], 'g')
58            plt.plot([self.W_rated_rpm, self.W_rated_rpm], [p2, self.Torque_rated], 'g', label='
59                Theory state curve')
60
61        def plot_U_Torque(self):
62            U_range = np.linspace(3, self.U_switch, 1000) #[m/s]
63            factor = 0.5 * self.rho_design * self.cp * np.pi * (self.R**3) / (self.TSR_real)
64            Torque2 = factor * (U_range**2) / 1000
65            plt.plot(U_range, Torque2, 'g', label='Theory state curve')
66
67

```

```

64     def plot_W_Pitch(self):
65         x = [6,self.W_rated_rpm,self.W_rated_rpm]
66         y = [-2,-2,20]
67         plt.plot(x,y, 'g', label='Theory state curve')
68
69     def plot_Pitch_Torque(self):
70         x = [10,-2,-2,20]
71         y = [0,0,self.Torque_rated,self.Torque_rated]
72         plt.plot(x,y, 'g', label='Theory state curve')
73
74     def plot_Torque_Pitch(self):
75         x = [0,0,self.Torque_rated,self.Torque_rated]
76         y = [10,-2,-2,20]
77         plt.plot(x,y, 'g', label='Theory state curve')
78
79     def plot_U_W(self):
80         x1 = np.linspace(3,self.U_rated_real,1000)
81         x2 = np.arange(self.U_rated_real,self.U_co+0.1,0.1)
82         x = np.concatenate((x1,x2))
83
84         y1 = x1*(self.TSR_real/self.R)*(60/(2*np.pi))
85         y1 = y1 - 0.6 #correct for offset
86         y2 = np.zeros(len(x2)) + self.W_rated_rpm
87         y = np.concatenate((y1,y2))
88
89         plt.plot(x,y, 'g', label='Theory state curve')
90
91     #DYNAMIC
92
93     def plot_U_Power_dyn(self):
94
95         x1 = np.linspace(3,self.U_switch,1000)
96         x2 = np.arange(self.U_switch,25.1,0.1)
97         xA = np.concatenate((x1,x2))
98
99         y1 = 0.5 * self.cp * self.A * self.rho_design * (x1**3)
100        y2 = np.zeros(len(x2)) + self.P_rated
101
102        y1 = y1/1000 #[kW]
103        y2 = y2/1000 #[kW]
104        yA = np.concatenate((y1,y2))
105
106        points = [(i, j) for i,j in zip(xA,yA)]
107        plotA = hv.Curve(points)
108
109        return plotA
110
111     def plot_W_Torque_dyn(self):
112        k_opt = 0.5 * self.rho_design * self.cp * np.pi * (self.R**5) / (self.TSR_real**3)
113        W_range = np.arange(0,self.W_rated_rpm+2,0.1) #[RPM]
114        Torque1 = k_opt * ((W_range*2*np.pi/60)**2) /1000 #[kNm]
115
116        points1 = [(i,j) for i,j in zip(W_range,Torque1)]
117        plot1 = hv.Curve(points1).opts(line_dash='dashed', color='deepskyblue')
118
119        # plot behaviour at Wmin and Wmax
120        p1 = k_opt * ((6*2*np.pi/60)**2) /1000
121        p2 = k_opt * ((self.W_rated_rpm*2*np.pi/60)**2) /1000
122
123        points2 = [(6,0), (6,p1)]
124        plot2 = hv.Curve(points2).opts(color='deepskyblue')
125        points3 = [(self.W_rated_rpm,p2), (self.W_rated_rpm,self.Torque_rated)]
126        plot3 = hv.Curve(points3).opts(color='deepskyblue')
127
128        plotB = plot1 * plot2 * plot3
129
130        del plot1,plot2,plot3
131        return plotB
132
133     def plot_U_Torque_dyn(self):
134        U_range = np.linspace(3,self.U_switch,1000) #[m/s]

```

```
135     factor = 0.5 * self.rho_design * self.cp * np.pi * (self.R**3) / (self.TSR_real)
136     Torque2 = factor * (U_range**2) / 1000
137
138     points = [(i,j) for i,j in zip(U_range,Torque2)]
139     plotC = hv.Points(points)
140
141     return plotC
142
143     def plot_W_Pitch_dyn(self):
144         xD = [6,self.W_rated_rpm,self.W_rated_rpm]
145         yD = [-2,-2,20]
146
147         points = [(i,j) for i,j in zip(xD,yD)]
148         plotD = hv.Curve(points)
149
150         return plotD
151
152     def plot_Pitch_Torque_dyn(self):
153         xE = [10,-2,-2,20]
154         yE = [0,0,self.Torque_rated,self.Torque_rated]
155
156         points = [(i,j) for i,j in zip(xE,yE)]
157         plotE = hv.Curve(points)
158
159         return plotE
160
161     def plot_U_W_dyn(self):
162         x1 = np.linspace(3,self.U_rated_real,1000)
163         x2 = np.arange(self.U_rated_real,self.U_co+0.1,0.1)
164         xF = np.concatenate((x1,x2))
165
166         y1 = x1*(self.TSR_real/self.R)*(60/(2*np.pi))
167         y1 = y1 - 0.6 #correct for offset
168         y2 = np.zeros(len(x2)) + self.W_rated_rpm
169         yF = np.concatenate((y1,y2))
170
171         points = [(i,j) for i,j in zip(xF,yF)]
172         plotF = hv.Curve(points)
173
174         return plotF
```

G

Google Earth site analysis examples

Below, two screenshots are provided of the site analysis performed in Google Earth. This results from this site analysis are discussed in sections 6.3.2, 7.3.2 and 8.3.2. The projected yellow areas are the worst-performing sectors, as given in table 6.5. Yellow landmarks are Vattenfall turbines, whereas the red marks are external turbines close to the PYC site.



Figure G.1: Western cluster bad performing sectors in Google Earth ('Google Earth', n.d.).



Figure G.2: Western cluster bad performing sectors in Google Earth ('Google Earth', n.d.).

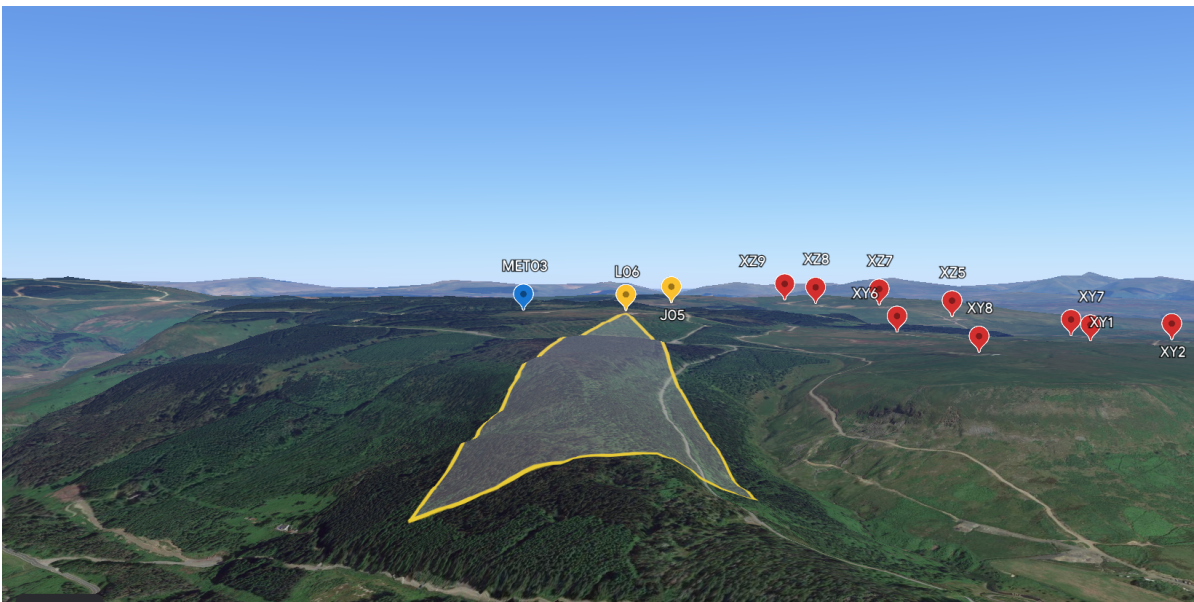


Figure G.3: Eastern cluster bad performing sector in Google Earth ('Google Earth', n.d.).