



Thesis submitted in fulfillment of the requirements for the  
award of the degree of Doctor in Engineering

# **PERFORMANCE MONITORING AND LIFETIME ASSESSMENT OF OFFSHORE WIND FARMS BASED ON SCADA DATA**

**NYMFA NOPPE**  
March 7, 2019

Promotors: Prof. dr. ir. Christof Devriendt  
Dr. ir. Wout Weijtjens  
Faculty of Engineering - Department of Applied Mechanics



Thesis submitted in fulfillment of the requirements for the  
award of the degree of Doctor in Engineering

# **PERFORMANCE MONITORING AND LIFETIME ASSESSMENT OF OFFSHORE WIND FARMS BASED ON SCADA DATA**

**NYMFA NOPPE**  
**March 7, 2019**

Promotors: Prof. dr. ir. C. Devriendt  
dr. ir. W. Weijtjens

Jury: Prof. dr. ir. D. Van Hemelrijck, president  
Prof. dr. R. Vounckx, vice-president  
Prof. dr. ir. J. Helsen, secretary  
Prof. dr. M. Kühn (Carl von Ossietzky University of  
Oldenburg)  
Prof. dr. ir. W. De Waele (UGent)  
Prof. dr. ir. L. Pyl  
Prof. dr. ir. P. Guillaume

Faculty of Engineering - Department of Applied Mechanics

Alle Rechten voorbehouden. Niets van deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook, zonder voorafgaande schriftelijke toestemming van de auteur.

All rights reserved. No part of this publication may be produced in any form by print, photoprint, microfilm, electronic or any other means without permission from the author.

Printed by  
Crazy Copy Center Productions  
VUB Pleinlaan 2, 1050 Brussel  
Tel / fax : +32 2 629 33 44  
crazycopy@vub.ac.be  
www.crazycopy.be

ISBN :  
NUR CODE :

# Acknowledgements

Writing these acknowledgements is literally the final part towards the completion of this PhD thesis. After spending more than 4 years to complete the research presented in this book, I'm almost looking forward to thank all important people directly or indirectly involved in the process.

Let me start with the person who actually convinced me to start a PhD in the first place. Christof, thank you for giving me this opportunity! What started as "continuing the work done for my Master's thesis for one year" ended up in a PhD thesis. The many fruitful discussions along the way, sometimes abroad during conferences, helped me being critical and seeing different opportunities. And last but not least, without you, there wouldn't be as many social events and, maybe even more importantly, they wouldn't be as lively as they usually are.

The next person in line is probably the one I have to thank the most. Wout, I can honestly say that without you, I would have never written this book. I probably would have quit after one or two years. Your enthusiasm for my research made me smile many times and gave me a boost to continue. Your encouragements, continuous support and guidance helped me through the downs. I've learned so much the last 4,5 years and I'm pretty sure a big part is thanks to you.

Thank you to all members of the jury, for taking the time to critically read my PhD thesis and the fruitful comments during the private defense.

I also want to thank all my (former) colleagues from the wind-team and OWI-lab for the collaborations and the good times at the many conferences and social events together.

Many thanks to the colleagues from the entire department to make the department an enjoyable environment to work in. The little, sometimes silly, talks offered often the needed diversion.

I'm really looking forward to stay a little longer at the university and see what the future brings!

Next, I'd like to thank all people from the industry I had the opportunity to meet and collaborate with. Thanks to you, I was able to do this research. Your input and comments made the research presented in this thesis something that can actually serve and help the industry in the future (hopefully).

Daarnaast wil ik ook graag mijn ouders bedanken. In de eerste plaats om zoveel jaar geleden toch toe te stemmen in mijn keuze voor de VUB. Maar nog veel belangrijker, en dat geldt voor iedereen van mijn familie en schoonfamilie, de

vele steun en trots tijdens mijn doctoraat.

Ook wil ik graag mijn vrienden bedanken. Om zo vaak interesse te tonen in mijn onderzoek, al zal het niet altijd even duidelijk geweest zijn wat ik precies deed. Om mee te leven bij de belangrijke momenten: "Hoe lukt het met de paper?", "Klaar voor die presentatie?", "Veel succes met het schrijven!", "Wanneer is je verdediging ook al weer?". Vast en zeker om af en toe mijn gezaag te aanhoren wanneer het toch niet allemaal zo goed ging. En om voor de vele ontspanningsmomenten tussendoor te zorgen! JJ, ik was dan misschien wel een beetje moe, ik ben heel blij dat ik de goede afloop van mijn privé verdediging met jou heb kunnen vieren! Glenn, merci om op het laatste moment toch nog eens te willen nalezen, zo zullen er (hopelijk) toch nog iets minder fouten op te merken zijn.

En als laatste, maar minstens even belangrijk als alle voorgaande personen samen: Quentin. Jij bent waarschijnlijk de enige die alle momenten waarop ik het even niet meer zag zitten, heeft meegemaakt. Jij bent waarschijnlijk ook de enige die alle kleine momenten waarop eindelijk wél iets gelukt was, heeft gezien. De rollercoaster die dit doctoraat toch was, heb jij van dichtbij gezien. Mee genoten van de goede momenten, naar mij geluisterd en, in de mate van het mogelijke, advies gegeven op de mindere momenten. Zo vaak onverstaanbare monologen aangehoord over de dingen die fout liepen. En zeker de laatste maanden zoveel huishoudelijk werk uit mijn handen genomen, zodat ik toch maar kon doorwerken, min of meer mijn deadlines kon halen en liefst ook nog voldoende kon rusten. Zonder jou had ik de laatste maanden waarschijnlijk vooral fastfood gegeten, was ik wel nog wat (extra) kilo's verdikt en was het huis een stort geweest. Dankjewel om zo goed voor mij (eigenlijk ons) te zorgen!

# Abstract

As the existing wind farms are growing older and the future wind turbines are growing bigger, fatigue assessments are gaining importance to the wind energy industry. Accurate fatigue analyses can help operators in the decision making to extend the lifetime of existing wind farms when they are close to their designed end-of-life. Moreover, it can help wind farm designers to optimize the designs of fatigue-driven substructures.

All operating wind turbines are equipped with a SCADA system, necessary during operation for the controlling of the plant. However, this SCADA data can be used for other purposes as well. In this thesis the possibility to optimize the use of the available SCADA data (1s and 10 minute statistics) is explored. Although the focus of the thesis lies in fatigue and lifetime assessment, some possibilities for performance monitoring are considered too. It is shown that under-performing turbines in a wind farm can be detected by calculating and comparing power curves based on 10 minute statistics of SCADA data. Moreover, changes in performance over time can be detected as well. However, results are very dependent on the quality of the data and more specific the quality of the wind speed measurement. To overcome this problem, the concept of Rotor Effective Wind Speed is implemented. Here, the wind speed is estimated by using measured produced power, rotor speed and blade pitch angle.

For the remainder of the thesis, the transition towards fatigue is made. An important input for any fatigue assessment is the estimation of acting loads on a structure. In the particular case of a wind turbine, various loads can be detected. The main research question in this chapter is to estimate these loads as accurate as possible using SCADA data only. Given the low frequent data (1Hz) and the non-existence of wave-related parameters in the available SCADA subset, the only reproducible load by SCADA is the thrust load. It is proposed to use a neural network to reproduce the thrust load. This neural network was trained with a combination of SCADA data and measured thrust loads, obtained using strain sensors installed on the substructure. The method was validated on multiple datasets of different turbines at different wind farms, including the cross validation of a trained model on different wind turbines of the same type in a single offshore wind farm.

Since the SCADA data can only be used for the reconstruction of the low frequent thrust load, additional techniques are necessary to reconstruct a full load history caused by all loads acting on the structure. For a fatigue assessment, the stress

history caused by all loads at fatigue critical locations in the structure is highly valuable. Two different techniques to reconstruct the stress history at any location in the structure, including the fatigue hotspots, are presented and compared. Both techniques use a combination of low frequent thrust load estimations and accelerometers installed at accessible locations in the tower. One of the techniques is based on modal decomposition and expansion, while the other makes use of a Kalman filter.

The next step in this thesis is to perform a lifetime assessment of multiple fatigue critical locations in the structure of the offshore wind turbine, including the monopile and the transition piece. Therefore the state-of-the-industry techniques are implemented for cycle counting and damage calculation. Moreover design documents of the wind turbine are consulted to decide on safety factors and specific S-N curves. To extend the calculated damage towards a lifetime, probability distributions regarding environmental conditions are needed. This thesis proposes a strategy for life time assessment and elaborates on several choices that can be made to perform such an assessment. Moreover this strategy is applied on a fictional offshore wind turbine and possible decisions made for the lifetime assessment are compared to each other.

Lastly, the correlation between measured damage and environmental conditions is shown by comparing the results of multiple instrumented turbines. The possibility of extrapolating measured damages at one turbine to other turbines within the same farm using environmental probability distributions is explored and commented upon. Moreover differences in measured damage are shown and explained. In the final chapter the main findings of this research are concluded and future perspectives are discussed.

# Samenvatting

Nu de bestaande windparken steeds ouder worden en de nieuwere turbines almaar groter, wordt er in de industrie meer en meer belang gehecht aan vermoeiingsanalyses van bestaande wind turbines. Deze kunnen operatoren helpen bij beslissingen over levensduurverlenging van oudere windparken. Bovendien kunnen zulke analyses ook ontwerpers helpen om de (vermoeiings-gedreven) ontwerpen van de funderingen te optimaliseren.

Alle operationele wind turbines zijn uitgerust met een standaard meetsysteem, SCADA genaamd, nodig voor de bediening van de installatie. Deze SCADA data kan echter ook nog gebruikt worden voor andere doeleinden. Deze thesis toont de mogelijkheid om het gebruik van de beschikbare SCADA data (1s en 10 minuten statistieken) te optimaliseren. Hoewel de focus in deze dissertatie op vermoeiing en levensduur ligt, worden ook mogelijkheden om de prestaties van de turbines te evalueren bekeken. Het is aangetoond dat turbines die onderpresteren opgemerkt worden door individuele vermogen curves te berekenen en vergelijken. Ook kunnen veranderingen in de tijd gedetecteerd worden. De kwaliteit van de SCADA data en meer specifiek de windsnelheidsmeting heeft echter een grote invloed op de resultaten. Hierdoor is ook het concept Rotor Effective Wind Speed geïmplementeerd, waarbij de windsnelheid bepaald wordt op basis van de metingen van andere parameters beschikbaar in de SCADA dataset.

Vanaf het volgende hoofdstuk wordt de transitie naar levensduur en vermoeiing gemaakt. Vaak wordt een vermoeiingsanalyse gestart op basis van betrouwbare metingen of inschattingen van de werkende krachten op de structuur. In het geval van een wind turbine, kunnen meerdere krachten onderscheiden worden. De hoofdonderzoeksvraag van dit hoofdstuk is om zo'n betrouwbare, accurate inschatting van deze krachten te maken op basis van enkel SCADA data. Gezien de laag frequente datasignalen (1Hz) en het gebrek aan golf-gerelateerde parameters in de SCADA subset, blijft enkel de drukbelasting door de wind over als reproduceerbaar op basis van SCADA data. Hiervoor wordt een neuraal netwerk getraind met een combinatie van SCADA data en gemeten belastingen. De methode werd gevalideerd aan de hand van meerdere datasets, komende van verschillende turbines in verschillende windparken. Hierbij werd ook een kruisvalidatie uitgevoerd, waarbij een model getraind op 1 turbine gevalideerd werd op een andere turbine van hetzelfde type binnen hetzelfde park.

Aangezien SCADA data enkel kon gebruikt worden om de drukbelasting van de wind op de turbine te reconstrueren, zijn bijkomende technieken nodig om ook

een inschatting te kunnen maken van de andere krachten die op de structuur werken. Voor vermoeïingsanalyses is de spanning veroorzaakt door al deze krachten op de meest kritische locaties in de structuur van groot belang. Twee verschillende technieken om deze spanning te reconstrueren op elke locatie in de structuur zijn beschreven, geïmplementeerd en vergeleken. Beide technieken combineren de ingeschatte drukbelasting en versnellingsmeters die geïnstalleerd werden op toegankelijke plaatsen in de toren. De ene techniek is gebaseerd op een modale decompositie en expansie, terwijl de andere techniek gebruik maakt van een Kalman filter.

De volgende stap in deze dissertatie is het uitvoeren van een levensduur analyse voor meerdere vermoeïings-kritische locaties in de structuur van een wind turbine in zee. Hiervoor werden vaak gebruikte technieken in de industrie geïmplementeerd om het aantal cycli in een tijdssignaal te tellen en hieruit de opgelopen schade te berekenen. De nodige parameters werden bepaald op basis van standaarden uit de industrie en de ontwerpdocumenten. Om de berekende opgelopen schade uit te breiden naar een levensduur zijn waarschijnlijkheidsverdelingen voor meteorologische omstandigheden, bv. windsnelheid, nodig. Een mogelijke strategie is voorgesteld in deze thesis, waarbij aandacht besteed werd aan de verschillende keuzes die hiervoor gemaakt moesten worden. Deze strategie is toegepast op een fictieve wind turbine, waarbij ook verschillende mogelijkheden voor de te maken keuzes voor zo'n levensduur analyse vergeleken werden.

Als laatste is ook de correlatie tussen gemeten schade en meteorologische condities vergeleken voor verschillende geïnstumenteerde turbines. De mogelijkheid om deze gemeten schade op één turbine te extrapoleren naar andere turbines binnen hetzelfde park op basis van meteorologische waarschijnlijkheidsverdelingen is onderzocht en besproken. Verder werden ook de verschillen in gemeten schade getoond en verklaard.

In het laatste hoofdstuk van deze thesis worden de belangrijkste bevindingen samengevat en mogelijke toekomstige onderzoeksmogelijkheden besproken.

# Nomenclature

## List of operators

- $\#\{\bullet\}$  : count
- $\bar{\bullet}$  : average
- $s_{\bullet}$  : standard deviation

## List of symbols

- $\Delta\epsilon$  : relative error
- $\epsilon$  : strain (-)
- $\theta$  : blade pitch angle ( $^{\circ}$ )
- $\lambda$  : tip speed ratio (-)
- $\rho$  : air density ( $kg/m^3$ )
- $\sigma$  : stress ( $Pa$ )
- $\Omega$  : rotor speed ( $rad/s$ )
- $\psi$  : yaw angle ( $^{\circ}$ )
- $A_{cs}$  : surface area of the cross section ( $m^2$ )
- $A_R$  : swept area of rotor ( $m^2$ )
- $B$  : air pressure ( $mbar$ )
- $c_P$  : power coefficient (-)
- $c_T$  : thrust coefficient (-)
- $D$  : Damage (-)
- $E$  : Young's modulus ( $Pa$ )

- $F_N$  : normal load ( $N$ )
- $F_T$  : thrust load ( $N$ )
- $F_{T,m}$  : measured thrust load ( $N$ )
- $F_{T,s}$  : simulated thrust load ( $N$ )
- $\hat{F}_T$  : modeled thrust load ( $N$ )
- $\hat{F}_{T,corr}$  : air density corrected modeled thrust load ( $N$ )
- $H$  : height ( $m$ )
- $I_c$  : area moment of inertia ( $m^4$ )
- $M$  : bending moment ( $Nm$ )
- $M_{tl}$  : bending moment in side-side direction ( $Nm$ )
- $M_{tn}$  : bending moment in fore-aft direction ( $Nm$ )
- $M_{tn,m}$  : measured bending moment in fore-aft direction ( $Nm$ )
- $N$  : number of data points
- $P$  : output power ( $W$ )
- $P_r$  : probability (-)
- $R$  : radius ( $m$ )
- $R_{rotor}$  : rotor radius ( $m$ )
- $R_i$  : inner radius (of cross section) ( $m$ )
- $R_o$  : outer radius (of cross section) ( $m$ )
- $R_0$  : gas constant of dry air =  $287,05 \frac{J}{kg \cdot K}$
- $T$  : air temperature ( $^{\circ}C$ )
- $t$  : wall thickness ( $mm$ )
- $V$  : (measured nacelle) wind speed ( $m/s$ )
- $V_n$  : air density corrected wind speed ( $m/s$ )
- $V_{rews}$  : rotor effective wind speed ( $m/s$ )
- $z$  : distance to hub ( $m$ )

## List of abbreviations

- 5PL-DE : 5 Parameter Logistic curve fitted by Differential Evolution
- CSF : Combined Safety Factor
- DEL : Damage Equivalent Load
- ERT : Extremely Randomized Trees
- EW : East-West
- FA : Fore-Aft
- FBGS : Fiber Bragg Grating Sensor
- FEM : Finite Element Model
- IQR : Inter-Quartile Range
- KNN : K-Nearest Neighbors regression
- LAT : Lowest Astronomical Tide
- LCoE : Levelized Cost of Energy
- MAE : Mean Absolute Error
- MDE : Modal Decomposition and Expansion
- MOB : Method Of Bins
- MRE : Mean (Absolute) Relative Error
- MP : Monopile
- MSF : Material Safety Factor
- NS : North-South
- O&M : Operation and Maintenance
- OHVS : Offshore High Voltage Station
- OWT : Offshore Wind Turbine
- OWF : Offshore Wind Farm
- PDF : Probability Density Functions
- RAE : Relative Absolute Error
- REWS : Rotor Effective Wind Speed
- RF : Random Forest regression

- RMSE : Root Mean Squared Error
- RPM : Rotations Per Minute
- RUL : Residual Useful Life
- SCADA : Supervisory Control And Data Acquisition System
- SCF : Stress Concentration Factor
- SE : Size Effect
- SEF : Static Extrapolation Factor
- SG : (resistive) Strain Gauge
- SGBRT : Stochastic Gradient Boosted Regression Trees
- SS : Side-Side or Sideways
- TI : Turbulence Intensity
- TP : Transition Piece
- TW : Tower

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Research context . . . . .	2
1.2	Introduction to the different topics . . . . .	5
1.2.1	Performance monitoring . . . . .	5
1.2.2	Fatigue assessment . . . . .	7
1.2.3	SCADA data . . . . .	9
1.3	Objectives and outline of the thesis . . . . .	9
1.4	Original contributions in this work . . . . .	11
<b>2</b>	<b>Technical background</b>	<b>13</b>
2.1	General concepts in (offshore) wind industry . . . . .	14
2.1.1	General working principle . . . . .	14
2.1.2	Availabilities . . . . .	18
2.1.3	Structural components of an offshore wind turbine . . . . .	18
2.1.4	Foundation concepts for offshore wind . . . . .	19
2.1.5	Turbulence and wake . . . . .	21
2.1.6	Loads . . . . .	21
2.1.7	Power coefficient . . . . .	23
2.2	Available data sources . . . . .	24
2.2.1	SCADA . . . . .	24
2.2.2	Data from a monitored turbine . . . . .	25
2.2.3	Meteorological data . . . . .	27
2.2.4	Overview of considered measurement campaigns . . . . .	27
<b>3</b>	<b>Performance monitoring</b>	<b>29</b>
3.1	Introduction . . . . .	30
3.2	Data . . . . .	30
3.2.1	Available datasets . . . . .	30
3.2.2	Data preprocessing . . . . .	31
3.3	Power curve modeling . . . . .	34
3.3.1	Method of bins . . . . .	35
3.3.2	Advanced techniques for power curve modeling . . . . .	36
3.4	Power curve monitoring . . . . .	39
3.4.1	Reference power curve . . . . .	39

3.4.2	Farm-wide power curve monitoring . . . . .	42
3.4.3	Power curve monitoring over time . . . . .	44
3.5	Rotor effective wind speed estimation . . . . .	49
3.6	Conclusions . . . . .	61
<b>4</b>	<b>Thrust modeling</b>	<b>63</b>
4.1	Introduction . . . . .	64
4.2	Available datasets . . . . .	64
4.2.1	Monitoring setup . . . . .	64
4.2.2	SCADA data . . . . .	65
4.2.3	Meteorological data . . . . .	67
4.3	Input parameter selection . . . . .	67
4.3.1	SCADA data . . . . .	67
4.3.2	Meteorological data . . . . .	79
4.4	Modeling technique . . . . .	79
4.4.1	Model based on 10min data . . . . .	81
4.4.2	Model based on 1s data . . . . .	85
4.4.3	Correlation between measured and modeled signal . . . . .	95
4.5	Cross validation . . . . .	96
4.6	Importance of data quality . . . . .	99
4.7	State-by-State modeling approach . . . . .	101
4.8	Conclusions . . . . .	104
<b>5</b>	<b>Full load estimation</b>	<b>107</b>
5.1	Introduction . . . . .	108
5.2	Monitoring campaign . . . . .	109
5.3	Modal decomposition and expansion . . . . .	110
5.3.1	Concept . . . . .	110
5.3.2	Results using strain gauges . . . . .	114
5.3.3	Results using a SCADA-based thrust model . . . . .	116
5.4	Kalman filter . . . . .	119
5.4.1	Concept . . . . .	119
5.4.2	Results . . . . .	121
5.5	Conclusion . . . . .	123
<b>6</b>	<b>Lifetime assessment of an instrumented turbine</b>	<b>125</b>
6.1	Introduction . . . . .	126
6.2	From stress history to fatigue damage . . . . .	126
6.2.1	Cycle counting . . . . .	126
6.2.2	Stress extrapolation within the structure . . . . .	128
6.2.3	Safety and correction factors . . . . .	129
6.2.4	Damage calculation . . . . .	133
6.3	From damage history to a fatigue life estimate . . . . .	135
6.3.1	Construction of the damage table . . . . .	136
6.3.2	Translating the damage table into RUL . . . . .	143
6.3.3	Reliability of lifetime calculation . . . . .	144

6.4	Case study . . . . .	145
6.4.1	Geometry, safety factors and applied S-N curves . . . . .	145
6.4.2	Measurements . . . . .	148
6.4.3	Results . . . . .	149
6.5	Conclusion . . . . .	167
<b>7</b>	<b>Fleet leader concept</b>	<b>169</b>
7.1	Introduction . . . . .	170
7.2	Data . . . . .	170
7.2.1	Strain data . . . . .	170
7.2.2	SCADA data . . . . .	171
7.2.3	Meteorological data . . . . .	171
7.3	Damage visualization . . . . .	171
7.4	Simple fleet leader . . . . .	175
7.4.1	Within one farm . . . . .	176
7.4.2	Across different wind farms . . . . .	180
7.5	Contribution of specific loads to damage . . . . .	183
7.6	Conclusion . . . . .	187
<b>8</b>	<b>Conclusions</b>	<b>189</b>
8.1	Conclusions . . . . .	189
8.2	Future work . . . . .	191
8.2.1	Developing a full monitoring approach . . . . .	191
8.2.2	Other future work . . . . .	193
<b>9</b>	<b>Bibliography</b>	<b>199</b>



# Chapter 1

## Introduction

*In this chapter the topic of the thesis is introduced. The general context of the research is given in Section 1.1, where the main topics are identified: performance monitoring and fatigue assessment. Both can be achieved (partly) using the existing SCADA data of the turbines. Those topics are discussed in more detail in Section 1.2, respectively in Sections 1.2.1, 1.2.2 and 1.2.3. The objectives set for this thesis are defined in Section 1.3. In the same section the outline of the thesis is given. To conclude, the main contributions in the thesis are summarized in Section 1.4.*

## 1.1 Research context

The importance of renewable energy and more specific wind energy has rapidly increased over the past decade, mainly in order to reach the Kyoto agreement. Figure 1.1 shows the evolution in time of the total installed capacity for several energy sources from 2005 until 2017 in the European Union. While the capacity of most traditional energy sources, like coal, fuel oil and nuclear energy, rarely changed or even decreased, most renewable energy sources, e.g. biomass, solar and wind energy, gained in capacity. The installed capacity by the wind industry in particular has quadrupled in a decade time, becoming the second largest form of power generation capacity in the European Union [1].

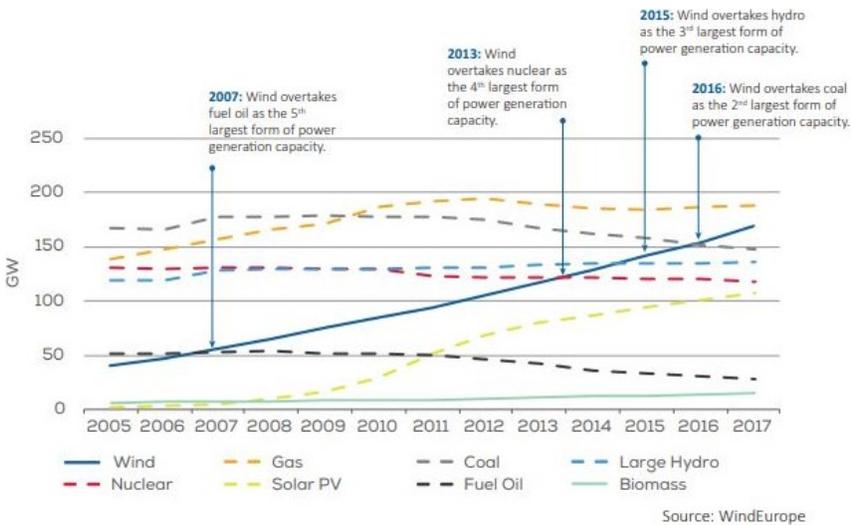


Figure 1.1: Total power generation capacity for all important energy sources separately in the European Union from 2005 until 2017 [1]

Looking into the wind industry more specifically, wind turbines can be installed on land, onshore (Figure 1.2a), or in shallow and deeper waters, offshore (Figure 1.2b). While installation onshore is usually easier and consequently cheaper, the average production of offshore wind turbines often surpasses that of onshore wind turbines. The annual report of 2017 of Wind Europe [1] showed the total installed capacity onshore exceeded the total installed capacity offshore in Belgium, while the percentage of the average annual electricity demand covered by wind can be equally divided among onshore and offshore. The higher production offshore is mainly caused by stronger and more consistent wind and less obstructions such as high buildings or trees. Installation on the other hand is more difficult and costly due to, among others, the additional construction of site-specific foundation structures, the installation on water using specialized vessels and the need for longer power cables to transport the produced energy to the consumers on land. Not only the installation is easier and cheaper onshore, the maintenance of offshore



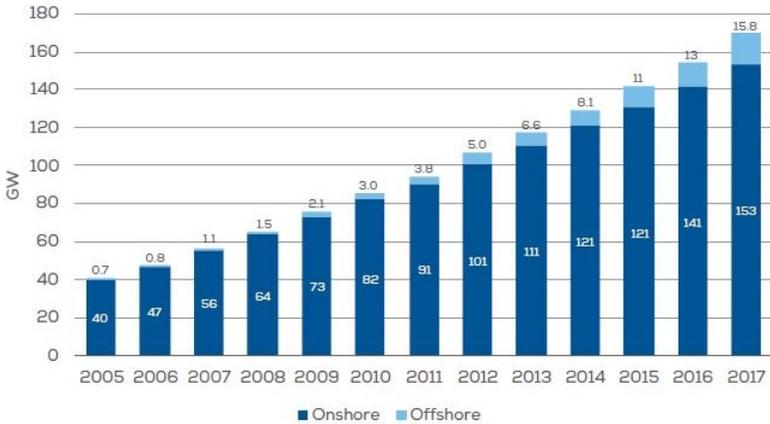
(a) Onshore [2]

(b) Offshore [3]

Figure 1.2: Wind farms are installed on land (onshore) or at sea (offshore)

wind turbines is more challenging as well. Maintenance costs are higher since offshore turbines are not as easily accessible as onshore wind turbines. In particular the specific conditions offshore, such as wave activity, complicate the maintenance schedule. E.g. if sea conditions block access to the turbines for boats, maintenance actions can only be performed using much more expensive alternatives such as helicopters.

Due to the more demanding challenges the offshore wind industry is facing, most of the wind turbines can be found onshore, as indicated by Figure 1.3. However a rapid increase in capacity installed offshore over the past years is clearly present.



Source: WindEurope

Figure 1.3: Total installed capacity of onshore and offshore wind in the European Union from 2005 until 2017 [1]

Looking at offshore wind farms in Belgium specifically, this trend could be seen as well in the past years. Moreover, the installed capacity offshore will continue to increase in the coming years with still two farms (partly) under construction and

three more farms planned. Figure 1.4 shows the currently operational wind farms (C-Power, Belwind, Northwind and Nobelwind), together with those (still partly) under construction (Rentel and Norther) and those expected (Seastar, Mermaid and Northwester 2).

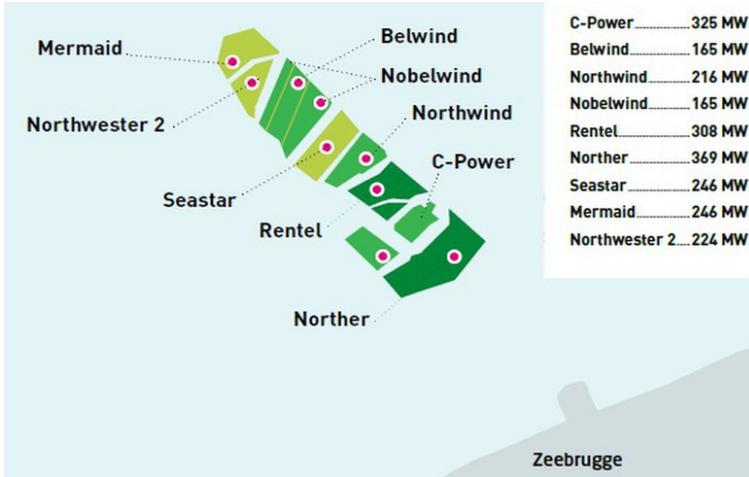


Figure 1.4: Belgian offshore wind farms [4]

The rapid growth in installed capacity offshore can be explained by, among others, the possibility to install more wind turbines in one farm and the increasing size of wind turbines installed on site. Recently it was announced the Belgian offshore wind farm Northwester 2 will install in 2019 wind turbines with a capacity of 9,5 MW, rotor diameters of 164 m and heights up to 220 m [5]. Installing bigger turbines in the farm has the advantage less turbines are needed to obtain the same total capacity. This reduces the installation costs, since also less, though bigger, foundations are required.

Despite the growth, it is still quite expensive to install and maintain offshore wind turbines. Wind farm operators are thus constantly looking to increase the energy production or to decrease the installation, operational and maintenance costs. All these gains are reflected in a reduction of the so-called Levelized Cost of Energy (LCoE), the actual cost of producing energy.

A reduction of the LCoE, can be obtained in multiple ways. First of all, the early detection of under performing wind turbines in the fleet is crucial. These turbines do not produce the intended amount of energy and should be maintained to optimize their production. Detecting these turbines in the wind farm can be achieved by performance monitoring.

An alternative approach to increase the total energy production of a single wind farm is by extending the lifetime of existing wind turbines and their foundation with a couple of years. In doing so the overall energy production is significantly enlarged with relatively low additional costs. For example a 5 year extension on a 20 year-old wind farm will imply a 25% increase in lifetime energy production, while only maintenance and operational costs continue. An accurate fatigue as-

assessment of each wind turbine structure individually, where the current strength or weakness of the metal is evaluated, can help in the decision for a possible lifetime extension. Moreover, such an assessment can help optimizing operational and maintenance and thus lower these O&M costs.

Such a fatigue assessment of existing wind turbines is also of great value to optimize the design of future wind farms as knowledge from the field flows back. With an optimized design, construction costs of new farms are reduced, ultimately leading to a lower LCoE.

In both performance monitoring and lifetime assessments the availability of the proper data is crucial. Installing additional sensors to obtain the proper data is often expensive or even unfeasible. However, all wind turbines installed are already equipped with a lot of sensors, generally referred to as the SCADA system of the turbine. Those measurements are mainly used by the control system of the turbine. Nonetheless, they are stored and usually available to operators as signals sampled at 1Hz and 10 minute statistics. While valuable in both performance and fatigue monitoring, the true potential of SCADA data remains often unused.

The next section will provide a general introduction to the topics discussed in this thesis. The existing challenges are briefly touched upon. More in depth analyses and discussions are provided throughout the thesis.

## 1.2 Introduction to the different topics

Section 1.2.1 gives an overview of the current practice of performance monitoring for existing wind farms. Section 1.2.2 introduces the importance of fatigue assessments for current and future wind farms. Finally Section 1.2.3 introduces the main data on which a lot of research is done for this thesis.

### 1.2.1 Performance monitoring

The primary purpose of wind turbines is to convert the wind into a renewable source of energy. The amount of energy generated by a turbine can be referred to as the production of the turbine. The instantaneous production depends mainly on the instantaneous wind conditions. By design wind turbines are optimized to do this conversion as efficient as possible for the site they are installed. However, over time several issues can arise that affect this efficiency and result in a loss of production. Typical examples of such issues are malfunctioning pitch motors, yaw-misalignment, ice accretion, blade erosion or any malfunction in the electrical drive of the wind turbine [6].

Production losses are minimized by early detection of under performing turbines, i.e. turbines whose production is not as high as one could expect for the current wind conditions. The current practice in industry to detect under performance is based on time-based and production-based availabilities [7],[8]. The time-based availability compares the actual time of operation to the ideal time of operation:

$$availability_{time} = \frac{operational\ time}{total\ time} \quad (1.1)$$

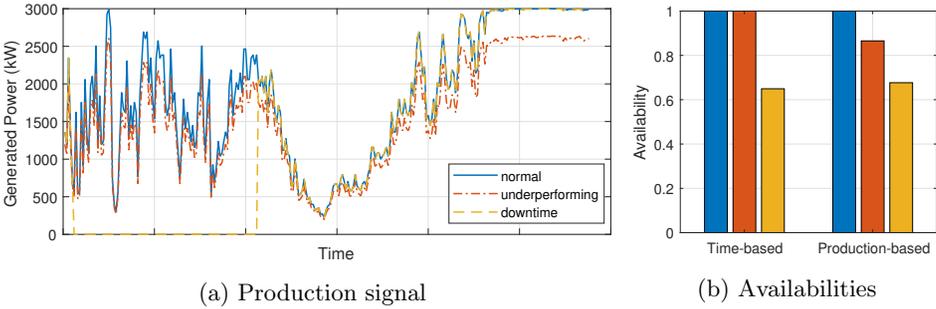


Figure 1.5: The production signal and corresponding availabilities of three fictional turbines. In blue, the production of a turbine performing ideally is shown, in red the production of a slightly under-performing turbine and in yellow the production of a turbine with some down-time.

The production-based availability compares the actual production to the ideal production:

$$availability_{production} = \frac{\text{realized production of energy}}{\text{theoretical amount of energy produced}} \quad (1.2)$$

Both indicators should be corrected for standstill conditions, when the turbine is not producing electricity while the wind conditions indicate production is possible. While relevant for contractual obligations, standstill conditions can obscure small but persistent changes in production, that in the long run are just as detrimental. This is illustrated by Figure 1.5, where the power generated by three different fictional turbines for two days and the corresponding availabilities are shown. One turbine produces energy as one can expect for the given wind conditions for that day, indicated by the blue full line in Figure 1.5a. The second turbine had to shut down for a moment for maintenance for example, indicated by the yellow dashed line. Low values barely exceeding 60% for both time and production-based availabilities are obtained for this turbine (Figure 1.5). The lost production during this downtime of less than a day would be equivalent to the yearly average energy consumption of 8 households. The third turbine seems to operate normally, but produces for some time instances and corresponding wind speeds less energy than expected. This is indicated by the red dashed-dotted line. This will lead to a perfect score of 100% for the time-based availability and an acceptable score of over 80% for the production-based availability. However, during those two days the same amount of energy was lost as three households would consume during one year (on average). In reality, a production based availability of 80% will be detected and looked into, but one can imagine a similar but less pronounced power loss can remain undetected for a long time (e.g. several months), leading to a lot more lost energy.

In this thesis, a performance monitoring strategy is presented to detect smaller but consistent production losses of individual wind turbines within a farm.

### 1.2.2 Fatigue assessment

A well-known mechanism to break a structure or object is to apply one single force, exceeding the ultimate breaking limit or the yield force of the material. However, multiple examples of failures due to much smaller forces can be found in history. For example, the Liberty ship S.S. Schenectady, built during the Second World War, split in two, even before its actual departure (Figure 1.6).

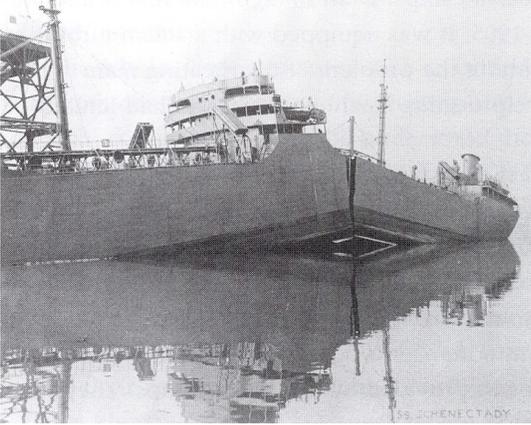


Figure 1.6: Failure of the Liberty ship S.S. Schenectady due to the formation of cracks and their propagation to critical lengths until sudden failure under relatively low but cyclic loads. [9]

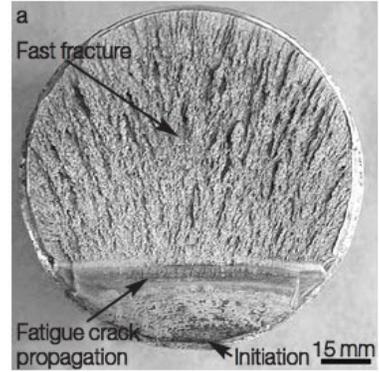


Figure 1.7: Typical fatigue fracture showing the crack initiation, the crack propagation and the sudden fracture. [10]

In this case the failure was caused by a relatively small force, but applied over and over again. By applying such a cyclic loading to a structure, one single force is applied and removed repeatedly. Such a cyclic loading causes microscopic cracks at the weakest spot in the original material, i.e. crack initiation. Such a microscopic crack will grow due to the applied force until a critical size is reached (crack propagation) and the structure suddenly fails (sudden or fast fracture). The cyclic loading thus causes the material to degrade over time. This mechanism is called fatigue.

A typical fatigue fracture surface is shown in Figure 1.7. Here the formation or initiation of a crack is shown. The propagation of the crack can be observed by the parallel lines in the material. Finally, the sudden, fast fracture is clearly visible.

In general, all forces exerted naturally to wind turbine structures, both onshore and offshore, are well below the ultimate force to break the material at once. However the total load applied on the wind turbine structure varies constantly over time and is thus assumed cyclic. Therefore, during design, the deterioration of the material caused by fatigue is taken into account. Offshore, fatigue is often the main driver for failure of the structure and more specifically the foundation. Given the fairly new industry, a lot of conservative choices were made during the design of the first wind farms. As the older wind farms now slowly reach their designed lifetime,

the question rises if the lifetime may be extended. Such a decision for lifetime extension is preferably supported by accurate fatigue and lifetime assessments of the entire structure of the wind turbine. This assessment boils down to interpreting the loads that the structure was subjected to and translating this in a residual lifetime. Moreover fatigue analysis of existing wind turbines and their foundation can optimize future design through a better understanding of loads and dynamics. The different parts of the substructure of an offshore wind turbine usually consist of multiple steel elements, welded together. This is shown by Figure 1.8. This means that, within the entire turbine structure, not all locations have the same material properties. Some locations, usually the welds between different elements, are more subjected to fatigue. This is usually caused by higher local stresses and more (small) defects in the material. These locations are thus more affected by loads and called fatigue hotspots. The fatigue induced by the load history at these hotspots is most valuable. Unfortunately for offshore wind turbines installed on monopiles, these hotspots are often located below water level or subsoil (indicated in red in Figure 1.9). For those locations it is unfeasible or impossible to mount sensors once the foundation has been installed.



Figure 1.8: Different elements are welded (white areas) together to form the monopile foundation of an offshore wind turbine. [11]

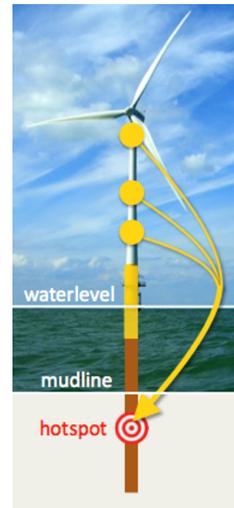


Figure 1.9: The fatigue hotspots of a wind turbine installed on a monopile are usually located around the mud-line. These locations are unreachable for existing wind turbines to mount sensors on. Therefore sensors will be installed at accessible locations within the tower and/or transition piece.

Currently, fatigue assessments of support structures, including the fatigue hotspots, are often based on measurements of the load history [12, 13, 14, 15]. Most of them imply continuous strain measurements at accessible locations. Pos-



Figure 1.10: The anemometer (encircled in the figure) used to measure wind speed and wind direction is usually located behind the rotor of the turbine [16]

sible accessible locations are indicated in yellow in Figure 1.9. However, for several reasons accelerometers are preferred over strain gauges, although this is not the case for measuring quasi-static loads. In the research presented in [13], the strain gauges are thus crucial to capture the quasi-static part of the loading.

In this thesis a model is presented to estimate the quasi-static load. This model is used as a substitute for the strain measurements during a fatigue assessment of an offshore wind turbine (OWT). Moreover a methodology to calculate the actual lifetime of an OWT based on strain measurements is given. Finally the possibility to extrapolate damage measurements within a wind farm is explored.

### 1.2.3 SCADA data

Every wind turbine is installed with a Supervisory Control And Data Acquisition (SCADA) system. The main purpose of the SCADA system is to monitor and control plants, for which reason it records data continuously. The main advantage of using SCADA data is its availability on all turbines in a fleet, providing relevant insight in the behavior of each of the turbines in the farm, without the need to instrument all of them.

However, the correct calibration and quality of the sensors contained in the SCADA is not guaranteed over the entire lifetime. A common example is the anemometer to measure wind speeds and wind directions. It is installed behind the rotor, as shown in Figure 1.10. The anemometer is known for its high uncertainties due to poor calibrations. Moreover, the quality and accuracy of the data can differ between the different wind turbine manufacturers. A proper preprocessing of the SCADA data and associated filtering process is advised.

## 1.3 Objectives and outline of the thesis

The main objective of this thesis is to optimize the use of the available SCADA data. This is done in two domains: performance monitoring and fatigue assess-



Figure 1.11: The different objectives set to achieve an accurate fatigue assessment of an offshore wind farm.

ment. In Chapter 3 the topic will be performance monitoring, while the following four chapters will handle (elements of) fatigue assessments.

To start Chapter 2 introduces some basic concepts in the wind industry. Moreover, an overview is given of the available datasets and the performed measurement campaigns at 11 offshore wind turbines, located across 5 different wind farms.

In terms of performance monitoring the objective set implies to detect under-performing wind turbines. Since SCADA data is already widely used for performance monitoring, the primary objective of the current research is not to detect severe defects and losses in production. Instead Chapter 3 introduces a technique based on power curves to detect small but persistent under-performances of turbines. In this chapter only 10min statistics of SCADA signals is used to monitor the entire wind farm.

For fatigue assessment the ultimate goal is to have an indication of consumed or expected lifetime for each turbine within a wind farm. To achieve this, multiple objectives were set, which should be achieved consecutively. These objectives are summarized by Figure 1.11 and elaborated upon in the following paragraphs.

An important input for any fatigue assessment is the load history as measured or expected at the OWT. Therefore the first objective regarding a fatigue assessment set in this thesis was to reconstruct as much as possible from the actual load history using SCADA data only. In Chapter 4 SCADA data, primarily 1s signals, is used to estimate one of the loads (the thrust load) acting on a wind turbine and its substructure. To do this, limited use of strain data is required. The focus in this chapter lies on only one turbine, but it is proven the concept is easily transferable to the entire fleet.

An accurate fatigue assessment is based on the combination of all loads acting on the wind turbine. For that reason, a second objective was set to estimate the stress history, at any location in the substructure of the OWT, caused by the full load acting on the OWT. Chapter 5 introduces two different techniques to combine the thrust load estimates with accelerometers to obtain a stress history at any location in the structure caused by all loads. In this chapter, the focus lies on only one turbine.

Once an accurate stress history is obtained for any location, including the fatigue hotspots, this should be translated into accumulated damage and finally into an expected lifetime for each of the locations of interest within the substructure. This was the next objective set. Chapter 6 describes the current practice for damage calculation. During this procedure a lot of choices have to be made. These are extensively discussed in this chapter. Results in this chapter focus on one turbine. However, the procedure can be easily transferred to any instrumented turbine.

By achieving the last objective, an accurate fatigue assessment can be performed for any instrumented turbine. However, not all turbines within the farm are instrumented. The last objective consists in the extrapolation of the results obtained from the instrumented turbines (fleet leaders) towards non-instrumented turbines within the farm. Essential analyses towards farm-wide fatigue assessment are presented in Chapter 7. In this chapter, only a couple of turbines are instrumented with additional strain sensors. For all other turbines in the farm 10 minute statistics of the common SCADA signals are available.

The datasets used in every chapter, the topic and the scope is summarized by Table 1.1.

Finally Chapter 8 concludes the thesis.

Table 1.1: Summary of the research topic, the data involved and the considered scope in every chapter

Chapter	Topic	Data involved	Scope
3	Performance monitoring	10min SCADA	farm
4	Thrust load estimation	10min SCADA 1s SCADA strain sensors (limited)	1 turbine, easily transferable to farm
5	Full load estimation	1s SCADA accelerometers strain sensors (limited)	1 turbine
6	Lifetime assessment	10min SCADA strain sensors	1 turbine, easily transferable to farm
7	Lifetime assessment	10min SCADA strain sensors	farm

## 1.4 Original contributions in this work

Results in this thesis are based on research presented at international conferences and published articles. Following, an overview is given of the most important contributions, together with the needed references.

1. The development of a performance monitoring strategy to detect slightly under performing turbines, or turbines of which the performance is decreasing over time, within a wind farm [17]. More information is given in Chapter 3.
2. The development of a model to estimate the acting thrust loads on a wind turbine, based on SCADA data only [18]. More information is given in Chapter 4.
3. Integration of thrust load estimates in the modal decomposition and expansion technique to assess full load stress signals at any location [19]. More

information is given in Section 5.3.

4. Two different techniques to reconstruct the full load stress signal at any location in the structure of the wind turbine were compared [20]. More information is given in Section 5.4.

On top of these contributions, some results presented herein are novel to this thesis, and have not been presented in current detail before:

1. The application of Rotor Effective Wind Speed on multiple turbines. To achieve this multiple functions to estimate the power coefficient were implemented, fitted and compared. Detailed information is given in Section 3.5.
2. A general framework for fatigue assessment of a single instrumented offshore wind turbine on a monopile foundation was implemented. More information is given in Chapter 6.
3. A first investigation into the feasibility of using the so-called fleet-leader concept for fatigue assessment of the foundations for offshore wind turbines using solely a limited number of instrumented turbines and the SCADA data. More information can be found in Chapter 7.

## Chapter 2

# Technical background

*In this chapter some basic principles and definitions about a wind turbine and the offshore wind industry are clarified in Section 2.1. In Section 2.2 information is given about the available data throughout this thesis. Additionally, the different measurement campaigns performed offshore are summarized.*

## 2.1 General concepts in (offshore) wind industry

### 2.1.1 General working principle

The purpose of a wind turbine is to generate electricity, using the potential energy of the wind passing by. As the wind passes the rotor blades, the rotor will rotate at a relatively low speed. Depending on the turbine type, this low-speed rotational movement is first converted into a high-speed rotational movement using a gearbox. The rotational movement is then transferred into electricity by a generator. The generator and, if present, the gearbox are located in the nacelle of the turbine, as shown in Figure 2.1.

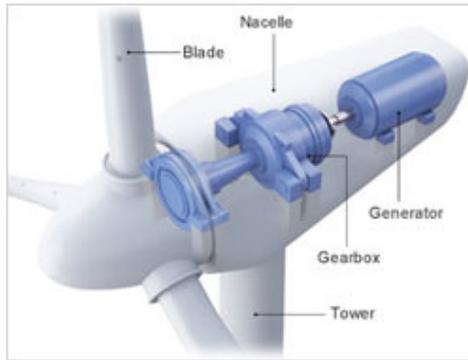


Figure 2.1: A schematic illustration of the nacelle of a wind turbine. The blades are pushed by the wind, causing the rotor to rotate. The gearbox transforms a low rotational speed into a high rotational speed. The torque is transformed into electric power by the generator. [21]

The higher the speed of the wind passing by, the faster the rotor will rotate (up to a certain speed) and the more electricity is generated by the generator. This is shown by the time signals of 10 minute averages of wind speed and generated power in Figure 2.2. Therefore, every wind speed can be coupled to a specific amount of power generation. This is already done during design, by the manufacturer, and resulted in a warranted power curve. Such a power curve is shown in Figure 2.3. As already mentioned, an increase in wind speed will cause an increase in rotor speed (up to a certain speed). Therefore, every wind speed can also be coupled to a specific value for rotor speed. This relation is illustrated in Figure 2.4.

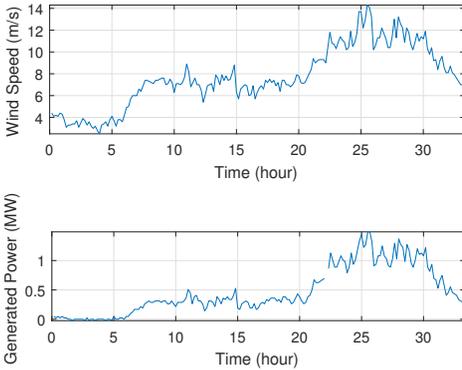


Figure 2.2: Time signals of 10 minute averages of two SCADA signals: wind speed and generated power

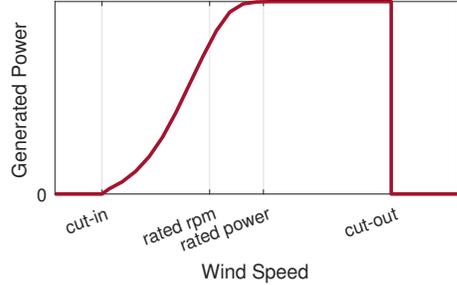


Figure 2.3: Illustration of a wind turbine's power curve

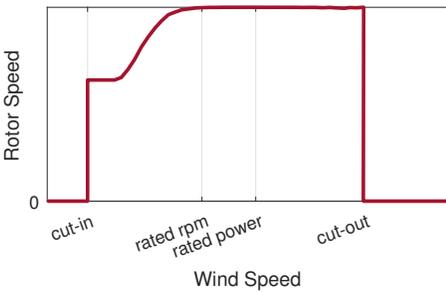


Figure 2.4: Rotor speed vs wind speed

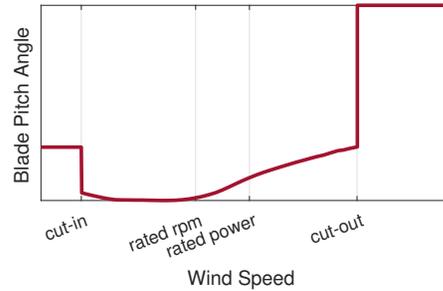


Figure 2.5: Blade pitch angle vs wind speed

As one can see, the turbine is not producing any electricity nor rotating under a certain wind speed due to too little available wind energy. The wind is not capable of exerting a high enough torque on the blades to make the rotor rotate. In this region, the turbine is so-called idling. The minimum wind speed for which the turbine starts generating electricity is called cut-in wind speed. The cut-in wind speed is typically around 3 or 4m/s.

From cut-in, the rotor speed increases as the wind speed increases until rated rotor rpm is reached. The rated rotor rpm is the maximum speed at which the rotor will rotate. Together with an increasing rotor speed, the electricity generation is increasing as well.

Once rated rotor rpm is reached, the blades will start to rotate around their longitudinal axis for increasing wind speeds. This rotation of the blades is called pitching the blades. By pitching the blades, only a part of the wind is caught by the rotor. The amount of pitch is determined by the controller of the turbine, based on measurements done by the SCADA system. Blades are pitched just as much as needed to keep the rotor rpm constant. Therefore, the higher the wind speed,

the more the blades will be rotated. The rotation of the blades is controlled by setting the ideal value of the blade pitch angle. The relation between blade pitch angle and wind speed is given by Figure 2.5. The blade pitch angle is illustrated in Figure 2.6.

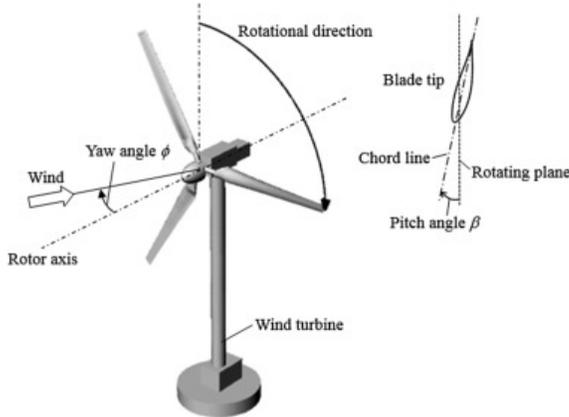


Figure 2.6: Illustration of yaw and blade pitch angle of a wind turbine, where the rotating plane is the plane the rotor blades rotate in [22]

As the wind speed still increases, the electricity generation rises as well. Once rated power is reached, the pitch angle will be adjusted in such a way the electricity generation never exceeds the rated power.

During storms, the risk of causing damage to the wind turbine's components by the high energy of the wind is too high if the rotor is still rotating. Therefore, the turbine is shut down once the wind speed reaches a certain value, referred to as cut-out wind speed. Cut-out wind speed is usually around  $25m/s$ , which can be related to 10 at the Beaufort scale and classified as a storm. This shut down is obtained by pitching out the blades, meaning rotate the blades in such a way all wind passes the rotor without being caught by the blades.

As the direction of the wind changes, the orientation of the wind turbine will change as well. This is controlled by the yaw angle of the turbine. The yaw angle is illustrated in Figure 2.6. A malfunctioning of the yaw motors can cause a difference between the orientation of the wind turbine and the wind direction. This is called yaw-wind misalignment and usually causes a decrease in production.

In theory, a wind turbine should always produce as much power as given by the manufacturer's power curve for a given wind speed. However, a variety of reasons can cause a deviation between the actual generated power and the ideal power as given by the power curve. During planned maintenance actions the turbine is not rotating regardless of the wind speed for instance. Whenever the turbine is not rotating nor producing electricity for any reason while the wind speed did exceed the cut-in wind speed, the turbine is so-called in standstill.

Another common deviation from the warranted power curve is de-rating or output-

limitation. In this case the turbine is still producing power, though not as much as expected. This is illustrated in Figure 2.7. This might be imposed by the wind farm operator to reduce the wake loads affecting the turbines downstream. Yet, de-rating can also be caused by a defective component of the turbine without the knowledge of the operator.

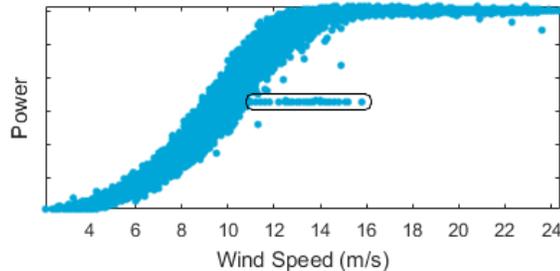


Figure 2.7: A measured power curve during which the turbine was output-limited. Data points measured during de-rating are framed.

Throughout this thesis, a distinction between operating states of the turbine is often made. To avoid confusion, an overview is given of the definitions used for the different operating states.

- **Idling:** the turbine is not rotating nor producing electricity because the actual wind speed is below the cut-in wind speed. If the wind suddenly increases however, the turbine will start rotating and generating power.
- **Generating:** the turbine is rotating and producing electricity. Here a distinction between under rated power and at rated power can be made.
- **Operational:** The turbine is operating. This means it produces energy when the wind speed is high enough and is idling when the wind speed is too low. This operational state is thus a merge between idling and generating and consequently excludes cut-out. The data corresponding to this operational state is selected based on the value of the blade pitch angle, which cannot exceed a predefined value.
- **Standstill, parked or non-operational:** the turbine is not rotating nor producing electricity although the wind speed is usually above cut-in wind speed. This operational state includes cut-out. For wind speeds below cut-in, the turbine is not ready to start rotating or generating power once the wind speed would increase. This state is generally characterized by a very high value for the blade pitch angle. This state is the opposite of operational.
- **Non-generating:** The turbine is not producing any electricity. This state is a merge between idling and standstill and is the opposite of generating.
- **Normal operating:** The turbine is operating as expected. This implies without any defect or fault, e.g. de-rating.

### 2.1.2 Availabilities

The state-of-the-industry to perform performance monitoring is based on the calculations of availabilities. Two main types can be distinguished: time-based and production-based availabilities.

The time-based availability is basically the percentage of the time the turbine was available to produce energy and calculated using Equation 2.1.

The production-based availability is rather a comparison between actual and ideal production and is calculated using Equation 2.2.

$$availability_{time} = \frac{available\ time}{available\ time + unavailable\ time} \quad (2.1)$$

$$availability_{production} = 1 - \frac{potential\ production - actual\ production}{potential\ production} \quad (2.2)$$

For both availabilities, the interpretation of available time or potential production can differ and is important to define. Therefore all available data is first divided in three groups. One group represents available time or actual production, a second group represents the unavailable time or potential production and the last group represents data which was not considered for the calculation. The decision on which data is divided in which group depends on the needed type of availability. Often a distinction is made between the operator's point of view and the manufacturer's. The first is called the operational availability, while the latter is called the technical availability. In general, the operator will consider less conditions in which the turbine is not producing power as available compared to the manufacturer, e.g. a technical standby or a requested shutdown. Moreover to calculate the technical availability, more categories are not considered in the calculation, e.g. scheduled maintenance. More information is given in [7] and [8].

### 2.1.3 Structural components of an offshore wind turbine

The main components of an offshore wind turbine can be seen in Figure 2.8. Starting from the top, one can see the nacelle, the tower, the transition piece and the foundation. As already shown in Figure 2.1, the nacelle contains the gearbox, if present, and the generator to produce electricity from the rotational energy of the rotor. The turbine blades are attached to the rotor hub, which is connected to the nacelle. The nacelle is placed on top of the turbine tower. The nacelle and the tower are the two components an offshore and an onshore wind turbine have in common.

The tower of an offshore wind turbine is installed on top of a transition piece (TP) between the tower and the foundation.

The substructure is defined as the merge of all components supporting the nacelle, being the tower, the TP and the foundation.

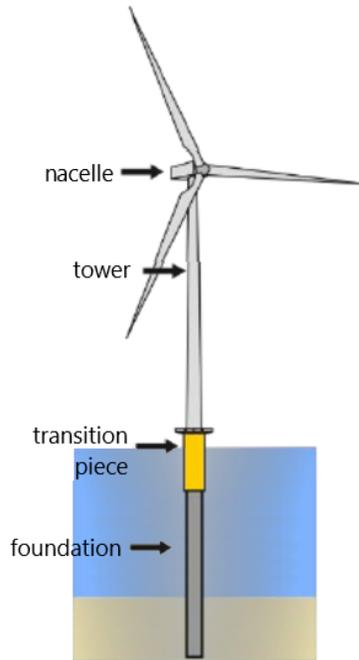


Figure 2.8: An offshore wind turbine and its structural components (modified) [23]

### 2.1.4 Foundation concepts for offshore wind

The design of the nacelle and turbine tower is done by the turbine manufacturer and is similar to the onshore turbines. Given the different site conditions between offshore wind farms, the foundation is designed site-specific. Multiple foundation concepts can be identified for offshore wind turbines. Several options are given in Figure 2.9: gravity-based, monopile, tripile, tripod, jacket or floating options.

Due to the simplicity and relatively low manufacturing costs, the monopile (MP) is often favored over the other foundation types for a wide range of water depths (up to ca. 50m). Figure 2.10 shows over 80% of the installed offshore wind turbines in Europe in 2016 have a monopile foundation. The dominance of monopiles in Europe can be explained by the relatively shallow waters. However at many other locations, in Europe and worldwide, waters are a lot deeper. For example in Scotland, Norway, the Atlantic coast of Portugal or the west coast of the United States of America. Here bottom-fixed substructures are not an option. Therefore floating wind turbines are considered.

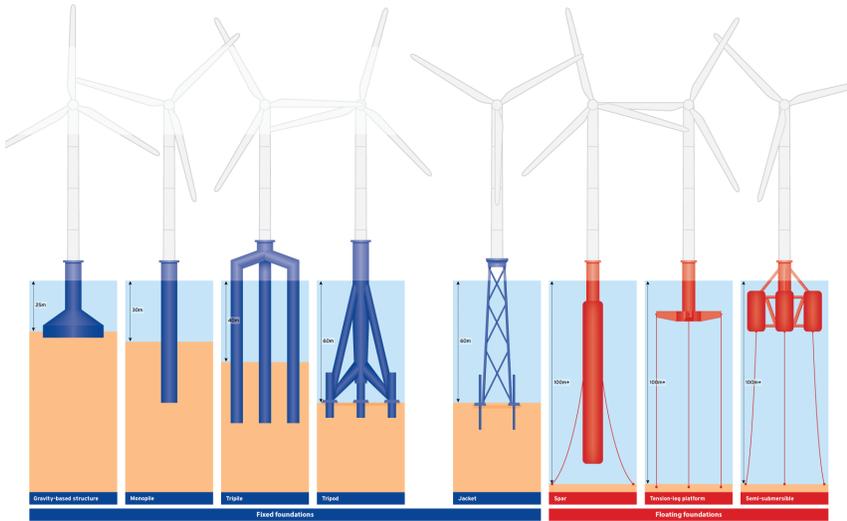


Figure 2.9: Multiple foundation concepts for offshore wind turbines. Bottom-fixed options are illustrated in blue, from left to right: gravity based, monopile, tripile, tripod, jacket. Floating concepts are illustrated in red. [24]

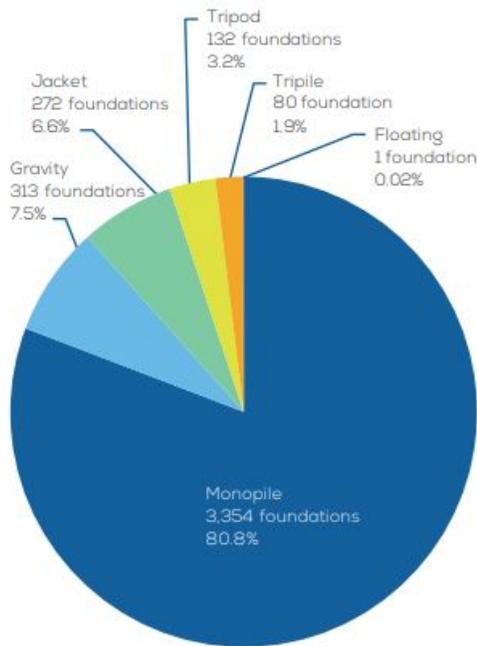


Figure 2.10: Installed offshore foundations in Europe in 2016 [25]

### 2.1.5 Turbulence and wake

As the wind passes the turbine rotor and energy is extracted from the air, the original air flow is changed. The wind after the turbine often has a decreased average speed. Moreover more chaotic and local changes in pressure and velocity are present. The latter is usually indicated as an increase in turbulence in the air flow. This phenomena of changing air flow is called the wake effect. Figure 2.11 visualizes the wake effect of the turbines within an offshore wind farm.



Figure 2.11: Visualization of wakes in an offshore wind farm [26]

Given the decrease in wind speed caused by wake, the performance of a downstream turbine is influenced by the wakes of upstream turbines. Moreover the more turbulent air in the wake of a turbine causes more damage to the downstream turbines than a wake-free air stream would cause.

### 2.1.6 Loads

Several different loads act on a wind turbine, both onshore and offshore, summarized by Figure 2.12.(a). In Figure 2.12.(b) the spectral composition of these loads is illustrated conceptually.

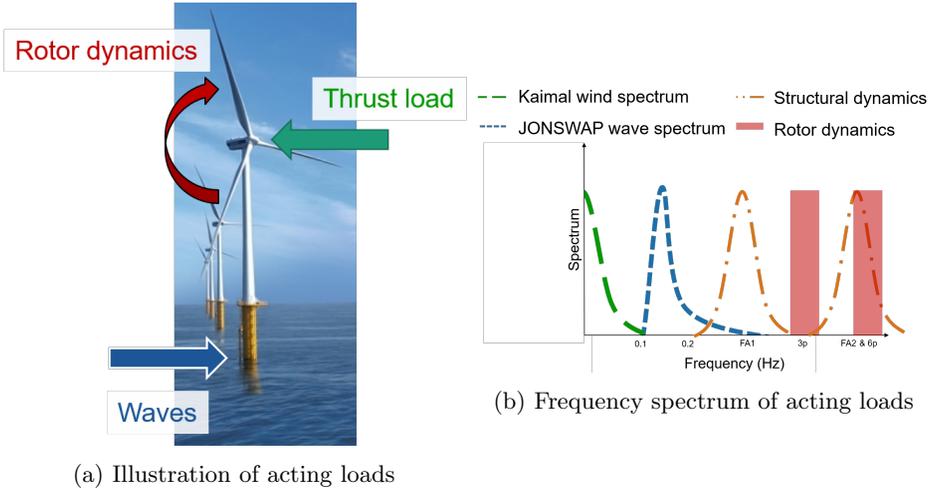


Figure 2.12: Overview of all loads acting on an offshore wind turbine.

First of all the wind caught by the rotor exerts a thrust load on the turbine structure. As the wind varies slowly, large, slow varying load cycles are induced on the turbine and the substructure. The variations in wind are referred to as turbulence. A typical representation of the frequency content of the atmospheric turbulence is the Kaimal wind turbulence spectrum [27, 28]. This Kaimal spectrum is indicated in green in Figure 2.12b. To quantify the thrust load  $F_T$  exerted by these wind variations, Equation 2.3 [29] is often used. Here, the air density  $\rho$ , the swept rotor area  $A_R$ , the wind speed  $V$  and the thrust coefficient  $c_T$  are required. Unfortunately, the thrust coefficient  $c_T$  is often only known to the manufacturer.

$$F_T = \frac{\rho}{2} c_T A_R V^2 \quad (2.3)$$

In case of an offshore wind turbine, an additional wave load is present. This wave load is caused by the variation in sea surface elevation or wave height. Usually the energy density spectrum of the surface elevation is represented by a wave spectrum. To represent the wave activity in the North Sea, the JONSWAP wave spectrum is often used [27, 30]. The frequency range corresponding to the wave spectrum is slightly higher than the wind turbulence spectrum, depending on site conditions. This is indicated in blue in Figure 2.12b.

Moreover, the passing of the blades introduces additional loads on the substructure of the turbine. These loads are called rotor harmonics and depend on the rotor speed. The exact frequency of the 3P harmonic can be calculated using the rotor speed and bearing in mind the passing of three blades every rotation. So for a rotor speed of 16RPM the 3P frequency occurs at 0,8Hz. Any multiple of this frequency is another rotor harmonic, being 6P, 9P, etc. Since the rotor speed varies, the associated frequencies change. Therefore a frequency zone is allocated to rotor dynamics in Figure 2.12b. In case a rotor imbalance is present, an additional 1P harmonic can be detected at a third of the frequency of the 3P harmonic.

These loads are amplified by the structural dynamics at the resonance frequencies. Every structure moves in a specific way when excited by an external force of a frequency close to its resonance frequency. Multiple resonance frequencies and corresponding mode shapes can be identified, of which the first two or three are the most important for a wind turbine. The first two mode shapes of an OWT installed on a monopile are shown in Figure 2.13, both in fore-aft and side-side direction. Each of these mode shapes correspond to a specific resonance or natural frequency, indicated by FA1 and FA2 for fore-aft direction in Figure 2.12b. The shape and natural frequency of a mode of a structure depend on the design of the structure. Therefore the careful positioning of the structural modes is part of design optimization. For example the first tower resonance frequency is targeted to remain below the 3P harmonic, but above the 1P harmonic, in what is called a soft-stiff design.

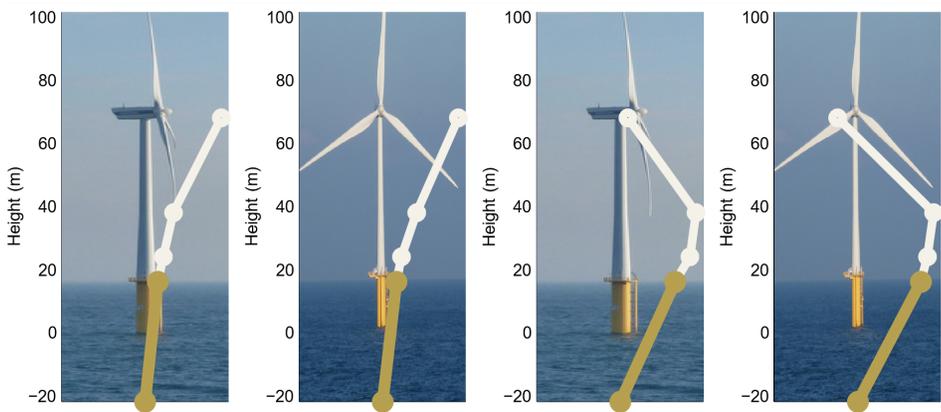


Figure 2.13: Visualization of first two mode shapes in fore-aft direction and side-side direction of an offshore wind turbine

### 2.1.7 Power coefficient

The power curve, as illustrated by Figure 2.3, is often represented by Equation 2.4 [31], where  $P$  the generated power,  $\rho$  the air density,  $c_P$  the power coefficient,  $A_R$  the swept rotor area and  $V$  the wind speed.

$$P = \frac{\rho}{2} c_P A_R V^3 \quad (2.4)$$

Just like the thrust coefficient  $c_T$ , the power coefficient  $c_P$  is often only known to the manufacturer. Although, based on the actuator disk model, both are related to each other by the axial induction factor  $a$  (Equations 2.5 and 2.6) [32].

$$c_P = 4a(1 - a)^2 \quad (2.5)$$

$$c_T = 4a(1 - a) \quad (2.6)$$

## 2.2 Available data sources

### 2.2.1 SCADA

Throughout the thesis, SCADA data is used consequently to assist in performance and lifetime monitoring. As already indicated, the SCADA system is installed on every turbine by default. Its purpose is mainly to control the plant, but the recorded data is stored as well. In industry, 10 minute statistics of SCADA signals, in short 10min SCADA data, are already widely used to perform performance monitoring for instance. Recently, also the SCADA signals sampled at 1 second have gained importance in industry. Both 10min as 1s SCADA data is used in this thesis.

#### Available parameters

The entire SCADA system consists of many sensors and measurement signals. However, for performance and lifetime monitoring only few of them are needed. Therefore, only a subset of the SCADA database was used in this thesis. The available parameters for this thesis were typically wind speed, generated power, blade pitch angle, rotor speed, yaw angle, wind direction and ambient temperature. In case of 10min SCADA data, usually only the mean value over 10 minutes was available. Although, in case of the wind speed often the standard deviation over 10 minutes was available as well. Sometimes the standard deviation, minimum and maximum over 10 minutes was available for all parameters.

#### Limitations

Despite the default existence of the SCADA system, there is no guarantee all sensors perform and are maintained properly over the entire lifetime of the wind turbine. Therefore the quality of the SCADA data can be questionable. As already mentioned in Section 1.2.3, a common example is the anemometer installed behind the rotor to measure wind speeds and wind directions. Often the quality of the SCADA data depends on the turbine manufacturer.

Moreover, the use of SCADA data is also limited depending on the application. When looking at load estimation, only one acting load can be estimated using SCADA data. As given in Section 2.1.6 and summarized by Figure 2.12 various loads act on an OWT. Part of these loads are high frequent, with a frequency exceeding 0,5Hz. However according to Nyquist-Shannon sampling theorem, only loads with a frequency up to 0,5Hz can be detected at their proper frequency (cfr. aliasing) using the 1s SCADA signals. On top of that, the SCADA system contains no signals related to wave activity. Therefore, the wave load cannot be reconstructed using SCADA data either.

#### Error codes

On top of the normal signals, the SCADA system also records error codes. Opposed to the typical SCADA signals, error codes are not stored continuously. Only

at timestamps where the error code has changed, data is recorded.

Error codes contain additional information about the state of the turbine. Therefore they are often used by the operators to explain the reason for down-time.

## 2.2.2 Data from a monitored turbine

On top of the available SCADA data sets, additional sensors were installed on a couple of wind turbines located in different offshore wind farms. This section will introduce the installed sensors and instrumented turbines.

### Typical instrumentation setup

An instrumented turbine is typically equipped with both strain gauges and accelerometers. All sensors are installed at easily accessible locations, like the existing platforms in the transition piece and tower. Figure 2.14 shows strain gauges are installed at the interface between tower and transition piece, indicated by the arrow left of the turbine. Moreover accelerometers were installed on multiple levels, among which the TP-TW interface, indicated by the arrows at the right of the turbine.

The specific number of sensors and heights of the platforms depend on the wind turbine. This is usually the same for all instrumented turbines within one farm.



Figure 2.14: Standard monitoring setup of an instrumented OWT

### Accelerometers

Two accelerometers are usually installed at typically 3 levels of the tower. These accelerometers capture the vibrations in two perpendicular radial directions. Sometimes two additional accelerometers are installed at the top level, to identify the torsional vibrations in the tower. Acceleration signals are useful to understand and capture the dynamics of the structure. For example, the estimation of the resonance frequencies and damping values of the structure can be done based on acceleration measurements [33].

### Strain gauges

Typically four or six strain gauges are installed in axial direction at the interface between tower and transition piece. Two types of strain gauges were installed during the different measurement campaigns, optical strain gauges (Fiber Bragg Grating Sensors, FBGS) and classical resistive strain gauges (SG). The first measures the change in the wavelength of reflected light, while the latter measures the electrical resistance. In general, installing a resistive strain gauge is cheaper

than installing an optical strain gauge. The only exception is when multiple strain gauges are needed along one single string of fiber (e.g. instrumenting the entire monopile). Another advantage of the optical strain gauge is the non-sensitivity to lightning. Both types of strain gauges are temperature-sensitive. Therefore the resulting measurements of each strain gauge should be compensated for the local temperature first.

The strain sensors are used to calculate the bending moment and resulting stresses acting on the structure. The measured (axial) strains  $\epsilon$  can be used to calculate the stress  $\sigma$ , using Young's modulus  $E$  and a simplified version of Hooke's law (Equation 2.7). The resulting stresses are still as observed by each sensor separately.

$$\sigma = E\epsilon \quad (2.7)$$

The general equation for normal stress  $\sigma_z$  induced by a normal force  $F_N$  and bending moments  $M_{NS}$  and  $M_{EW}$  in cylindrical coordinates is given by Equation 2.9, where  $A_{cs}$  represents the surface area of the cross section,  $R_o$  the outer radius,  $R_i$  the inner radius,  $\theta_h$  the heading and  $I_c$  the area moment of inertia. The area moment of inertia can be calculated using Equation 2.8. An illustration of a typical cross section is given in Figure 2.15.

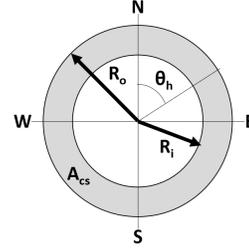


Figure 2.15: Illustration of a cross section of the tower, transition piece or monopile

Equation 2.9 can be used for every sensor installed on the turbine. However, at least three sensors are necessary to extract the normal load  $F_N$  and bending moments in both directions  $M_{NS}$  and  $M_{EW}$  (North - South and East - West) from the measurements. In general, one can find Equation 2.10, for a setup with three sensors.

$$I_c = \frac{\pi}{2}(R_o^4 - R_i^4) \quad (2.8)$$

$$\sigma_z = \frac{F_N}{A_{cs}} + \frac{R_i}{I_c}(M_{NS}\sin\theta_h - M_{EW}\cos\theta_h) \quad (2.9)$$

$$\begin{bmatrix} \sigma_{z1} \\ \sigma_{z2} \\ \sigma_{z3} \end{bmatrix} = \begin{bmatrix} \frac{1}{A_{cs}} & \frac{R_i}{I_c}\sin\theta_{h,1} & -\frac{R_i}{I_c}\cos\theta_{h,1} \\ \frac{1}{A_{cs}} & \frac{R_i}{I_c}\sin\theta_{h,2} & -\frac{R_i}{I_c}\cos\theta_{h,2} \\ \frac{1}{A_{cs}} & \frac{R_i}{I_c}\sin\theta_{h,3} & -\frac{R_i}{I_c}\cos\theta_{h,3} \end{bmatrix} \begin{bmatrix} F_N \\ M_{NS} \\ M_{EW} \end{bmatrix} \quad (2.10)$$

According to the standards for the wind industry [34] the bending moments acting on the tower are given in two directions, the fore-aft and the side-side. Basically, the fore-aft direction is the same as the wind direction, while the side-side is perpendicular to the fore-aft. This is illustrated by Figure 2.16.

The resulting bending moments in north-south and east-west directions are converted into bending moments in fore-aft  $M_{tn}$  and side-side direction  $M_{tl}$  by

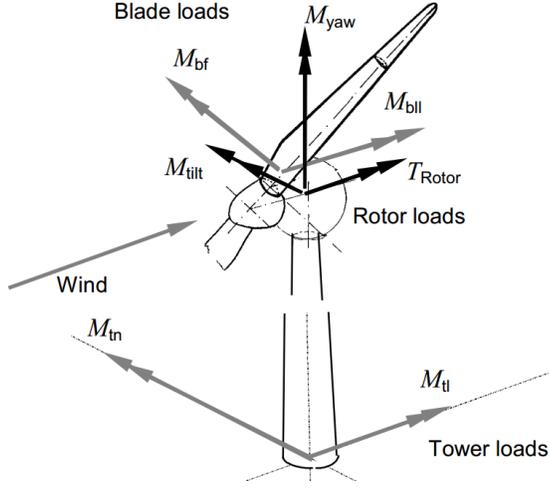


Figure 2.16: Load conventions according to IEC61400-13 [34]

multiplying by an additional rotation matrix  $\mathbf{R}$  (Equation 2.11) using the turbine yaw  $\psi$  in the SCADA.

$$\mathbf{R} = \begin{bmatrix} \cos(-\psi + 180^\circ) & \sin(-\psi + 180^\circ) \\ -\sin(-\psi + 180^\circ) & \cos(-\psi + 180^\circ) \end{bmatrix} \quad (2.11)$$

### 2.2.3 Meteorological data

On top of the SCADA dataset and the measurements performed during the several measurement campaigns, additional meteorological data is available from multiple sources as well.

- air pressure, 10 minute averages: measured at the Offshore High Voltage Station (OHVS) at Belwind
- wave height and wave period, 10 minute averages: measured at the Offshore High Voltage Station (OHVS) at C-Power and Belwind (only wave height)

### 2.2.4 Overview of considered measurement campaigns

Multiple measurement campaigns were performed in the framework of the OWI-lab. For this thesis a lot of data collected during those measurement campaigns was used. In this section a general overview of the different considered campaigns is provided. Due to confidentiality reasons, no specifics of the wind farms involved are given.

In total 11 turbines were instrumented across 5 different wind farms. Most of these turbines were installed on monopiles, some on jackets. Table 2.1 gives an

overview of the number of instrumented turbines in the different wind farms, together with the duration of the measurement campaigns. Some of these campaigns are still ongoing.

Table 2.1: Measurement campaigns at different offshore wind farms performed by OWI-lab

Wind farm	Number of turbines	Duration of measurement campaign (years)
WindFarm1	1	2
WindFarm2	2	3,25
WindFarm3	3	3
WindFarm4	2	2
WindFarm5	3	1

All 11 turbines were instrumented based on the typical instrumentation setup as explained in Section 2.2.2. This includes strain measurements at one level, typically close to the interface between tower and TP, and multiple levels of accelerometers in the tower. All the measurements done at all instrumented turbines combine into an available data set containing strain and acceleration measurements for a total period of almost 25 years.

## Chapter 3

# Performance monitoring

*In this chapter the performance of a Belgian offshore wind farm is monitored using only the 10min SCADA data set of the turbines. Monitoring is done based on the power curve, representing normal operation of the turbines. To make sure only normal operation of the turbine is accounted for, several filters are applied to the SCADA data.*

*In the wind energy industry, the power curve is usually obtained using the method of bins. This method is compared to other univariate and multivariate techniques, such as  $k$ -nearest neighbors regression or random forest regression. Univariate techniques give little improvements with respect to the standard method of bins, while multivariate techniques can improve the modeling significantly (up to 20%). Moreover, in this chapter modeled power curves are also used to monitor several turbines or an entire wind farm over time. The monitoring approach consists in calculating a power curve of normal operation regularly and tracking those calculated power curves using health indicators. A seasonal effect can be detected. Moreover under-performing turbines or degrading production can be easily discovered. However, an important dependency on measured wind speed and a lack of reliable wind speed measurements is found.*

*A solution for the unreliable wind speed measurements is suggested by calculating the rotor effective wind speed. To estimate the power coefficient needed, several functions from the literature are tested. It is shown the coefficient of these functions should be fit for the specific turbine type. Moreover the inclusion of a pitch offset might be advisable in some cases.*

## 3.1 Introduction

Wind farm operators are interested most in the performance of a wind turbine, and more specific in the possible production loss. A well-known parameter in the industry to detect performance issues is the availability (time-based [7] or production-based [8]), where actual and possible production (time) is compared. This is explained in more detail in Section 2.1.2

On the other hand, anomalies in power production are often detected using modeled power curves [35, 36, 37, 38, 39, 40], which represent the relation between wind speed at the turbine and power production of the turbine.

Both monitoring techniques are mainly used to detect down-time or relatively big power losses. But fail to detect small changes in production. If these small changes persist over time, the accumulated lost production can become significant. For this reason, a more sensitive monitoring approach is introduced to detect even the small changes over time.

The techniques proposed are applied on an operating wind farm. More information about the available data is given in Section 3.2.1. This approach is based on power curve modeling during normal operation. To be able to detect small changes, abnormalities in power production should be filtered out. Multiple filters are proposed in Section 3.2.2.

In general, the power curve modeling technique used in the industry, the method of bins, is utilized and explained in Section 3.3. Moreover, this standard technique is compared to other techniques, such as k-nearest neighbors regression or random forest regression, as well.

The monitoring approach consists in calculating the power curve under normal operation on a regular basis and tracking the resulting health indicators (Section 3.4). The proposed approach is highly dependent on a reliable wind speed measurement [40]. Unfortunately, the SCADA measurement is known to be unreliable [41, 42]. For that reason, Section 3.5 elaborates on the calculation of the rotor effective wind speed, using other SCADA parameters to estimate the wind speed. Finally, Section 3.6 concludes this chapter.

## 3.2 Data

### 3.2.1 Available datasets

For this chapter, a subset of SCADA data of the Belwind wind farm located in the Belgian North Sea was available. The Belwind wind farm consists of 55 V90 3MW Vestas turbines. The layout of the wind farm is shown in Figure 3.1.

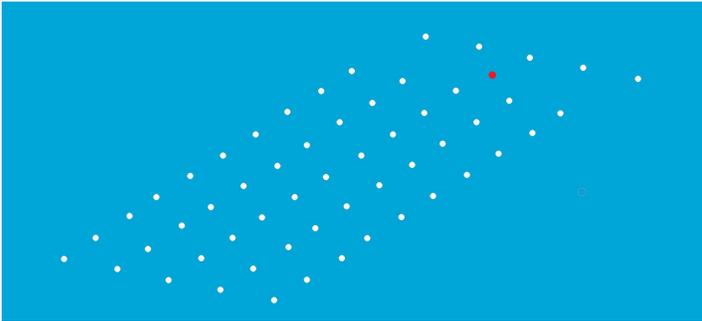


Figure 3.1: The layout of the Belgian Wind farm Belwind, consisting of 55 3MW V90 Vestas turbines (white dots) and one Offshore High Voltage Station (red dot)

The SCADA subset contained 10 minute statistics for a period of a little over 4 years for the following parameters: wind speed, produced power, rotor rpm, blade pitch angle, wind direction, yaw angle and ambient (air) temperature. For all parameters the averaged value was available, for wind speed, produced power, rotor rpm and blade pitch angle the minimum, the maximum and the standard deviation were also available. The wind speed is measured by an anemometer mounted on the nacelle, just behind the rotor of the turbine.

Although also error codes were logged during this period, it was chosen not to use those to obtain the final results. This decision was made because the added value of the error codes in this analysis was little due to the lack of understanding of their definition.

In addition, meteorological data is measured at the Offshore High Voltage Station (OHVS) located in the Belwind farm, as shown in Figure 3.1. From the available meteorological data at the OHVS only the 10 minute averages of air pressure were used.

### 3.2.2 Data preprocessing

Key in all data analyses is a proper preprocessing of the required data. As the saying goes, garbage in is garbage out. In order to obtain a valid result, invalid data points should be removed, i.e. filtered out. Additionally, where necessary, data signals should be corrected.

#### Filtering

Many possible filters are suggested in literature [43, 44, 45] to exclude all abnormal data from the dataset for performance monitoring. However, depending on the application not all filters seem advisable. Moreover some filters are impossible to implement due to missing or unavailable data. For example in [43] and [44] it is proposed to remove all data for which the external conditions other than wind

speed are out of the operating range of the wind turbine, however this information was not available in the considered research.

In order to be able to keep the flexibility to change the applied filters easily, a toolbox with several, adjustable filters was created. In Figure 3.2 an overview of these filters is given.

**Filters for SCADA-data**

Only datapoints during normal operation are included (based on errorcodes)

	Apply?	Minimum value	Maximum value
<input checked="" type="checkbox"/> Remove NaN-values	<input checked="" type="checkbox"/>	0	30
<input checked="" type="checkbox"/> Turbine offline (all zero)	<input checked="" type="checkbox"/>	0	360
<input checked="" type="checkbox"/> Blade icing	<input checked="" type="checkbox"/>	-500	4500
<input checked="" type="checkbox"/> Measurement Sector	<input checked="" type="checkbox"/>	-5	30
<input type="checkbox"/> Derating (based on errorcodes)	<input checked="" type="checkbox"/>	-5	100

Environmental conditions (except wind speed) are out of the operating range (based on errorcodes)

Filter manually based on plot

Severe filters: Timestamps before and after removed points will be removed as well

	Apply?	Period (hours)
Windspeed value is constant for x hours	<input type="checkbox"/>	2

Remove statistical outliers

Figure 3.2: Overview of available SCADA filters in the developed toolbox for performance monitoring

Here follows more information about all the implemented filters.

- According to [44], the wind turbine should be operating in normal operation during the measurement period. To ensure this, one could use the error codes if available. Depending on the wind turbine type, codes corresponding to abnormal operation should be identified. Then, the data points for which one or more of these error codes are recorded during the 10 minute interval are rejected.
- [44] also stated that all data points for which a required signal has been outside the signal range during the 10 minute interval should be excluded from the dataset. This can be done by setting a minimum and maximum allowed value for the mean, minimum or maximum recorded value during every 10 minute interval. This filter can also be used to exclude start/stop events from the dataset, as suggested in [45].
- **Remove NaN-values:** data points for which a required signal was unavailable or erroneous should be excluded [44]. If a signal was unavailable or not operating, a NaN-value is recorded. Hence a filter is included to exclude all data points for which a NaN-value is detected for one of the signals.

- Another possible filter consists in excluding all data points for which the external conditions (other than wind speed) were out of the operating range of the wind turbine [44]. This is another filter that can be included if error codes are available. In that case the error codes corresponding to abnormal external conditions should be identified first.
- **All parameter values are 0:** Data points for which the turbine was off-line can be excluded [44]. This usually corresponds to a value of exactly zero for all required signals.
- Similarly all data points for which the turbine is output-limited (de-rated) should be excluded [44]. While these points are clearly visible in the plots (for example Figure 2.7), it is not always easy to remove them automatically. It may be some error codes correspond to a de-rated condition of the turbine, but these error codes don't include all occasions of de-rating. Therefore they can be filtered out manually, based on the plot, as well if needed.
- Additionally, all data points for which the 10 minute averaged wind direction is outside the measurement sector should be removed according to [44]. The measurement sector represents all wind directions for which the air in front of the turbine is not disturbed by another turbine. This can be defined for every wind turbine based on the wind farm layout. However for the turbines in the middle of the farm the measurement sector is empty and no data will remain if this filter is applied.
- **Blade icing:** As stated in [44], all data points for which blade icing possibly occurs are removed. In practice this is done by removing all data points for which the measured air temperature at the turbine is below zero.
- **Severe filters:** If only 10 minute statistics should be available, the reason for which a data point is excluded could have started at the end of the previous timestamp or ended in the beginning of the next timestamp. Hence [45] suggests to exclude all timestamps just before and after already excluded data points.
- **Remove statistical outliers:** Finally a last filter was included in an attempt to exclude de-rated data points in an automatic manner. This filter consists in excluding the statistical outliers of the power residuals (compared to the warranty power curve). All data points within the interval  $[Q_1 - 3(Q_3 - Q_1); Q_3 + 3(Q_3 - Q_1)]$  remain, where  $Q_1$  indicates the first quartile and  $Q_3$  the third quartile of the residual signal.

### Data corrections

Once all required filters are set, some data corrections are proposed in [44]. The most important one is to correct the windspeed for variations in air density, related to variations in temperature and air pressure. The standard suggests to correct the data to two reference air densities, the sea level air density and the average of the measured air density data. Here, only the sea level air density

$(1,225 \frac{kg}{m^3})$  is chosen as reference air density  $\rho_0$ . The air density corrected wind-speed  $V_n$  is calculated using Equation 3.1, where  $V$  corresponds to the measured nacelle wind speed in  $m/s$  and  $\rho_{10min}$  corresponds to the averaged measured air density over 10 minutes.

$$V_n = V \left( \frac{\rho_{10min}}{\rho_0} \right)^{\frac{1}{3}} \quad (3.1)$$

The measured air density  $\rho_{10min}$  is calculated using Equation 3.2 [43], where the averaged measured air pressure over 10 minutes  $B_{10min}$ , the averaged measured air temperature  $T_{10min}$  and the gas constant of dry air  $R_0 = 287,05 \frac{J}{kg \cdot K}$  is utilized.

$$\rho_{10min} = \frac{B_{10min}}{R_0 T_{10min}} \quad (3.2)$$

Moreover, the standard [44] suggests to use a nacelle transfer function to correct for the distortion of the wind by the rotor and nacelle. Unfortunately, this was not possible to include here, since no measurement of the wind just in front of the rotor was available. On top of that, it is not clear whether the post processing of the raw SCADA data already includes such a nacelle transfer function.

### 3.3 Power curve modeling

Discussions on performance of wind turbines can be based on availabilities or power curves. While the first is more susceptible to down-time, the latter evaluates the performance during normal operation. A power curve essentially represents the relation between wind speed and produced power, under normal operating conditions. In literature, a power curve model is often used to detect and/or classify performance anomalies [35, 36, 37, 38, 39, 40]. By identifying changes in the power curve, a turbine's power production process can be optimized.

Many different data mining techniques are applied to model a power curve [46, 47]. In this section the one technique described in an industrial standard, the method of bins, is explained (Section 3.3.1) and compared to some other techniques (Section 3.3.2) such as k-nearest neighbors or random forest regression.

As mentioned before, the power curve gives an idea of power production under normal operating conditions. Therefore it is important to use only data representing normal operating conditions to model a power curve. Hence a set of filters is chosen to exclude all data corresponding to possible downtime or under-performance of the turbine. Table 3.1 shows the allowed minimum and maximum values for the required signals. On top of that filters on non-operation and blade icing are applied as well. Possible de-rating is filtered out manually.

Figure 3.3 shows the remaining (blue dots) and removed (red crosses) data points after applying these filters for one turbine over a period of 3 months.

Table 3.1: Interval in which the required signals are considered available and normally operating.

Wind Speed ( $m/s$ )	Power Production ( $kW$ )	Ambient Temperature ( $^{\circ}C$ )	Air Pressure ( $mbar$ )
[3; 25]	]0; 3300]	[-20; 40]	[950; 1050]

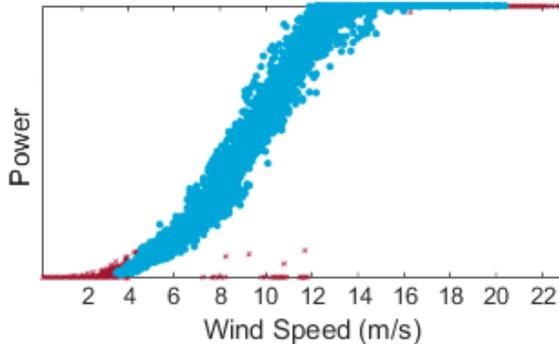


Figure 3.3: Measured power curve during 3 months for one turbine. The blue dots represent data points kept after applying the proposed filters, while the red crosses represent the rejected data.

### 3.3.1 Method of bins

To calculate a power curve, the method of bins (MOB) is usually used in industry. To apply the MOB, the data is divided into bins of  $0,5m/s$  (centered on multiples of  $0,5m/s$ ) based on the value of the air density corrected wind speed. Afterwards, the mean values of air density corrected wind speed  $V_{n,i}$  and the output power  $P_i$  are calculated for any bin  $i$  using Equations 3.3 and 3.4, in which  $V_{n,ij}$  and  $P_{ij}$  represent the air density corrected wind speed and the power output of data point  $j$  in bin  $i$  respectively and  $N_i$  the number of data points in bin  $i$ .

$$V_{n,i} = \frac{1}{N_i} \sum_{j=1}^{N_i} V_{n,ij} \quad (3.3)$$

$$P_i = \frac{1}{N_i} \sum_{j=1}^{N_i} P_{ij} \quad (3.4)$$

To make sure enough data was used to calculate the power curve, following checks are performed, as proposed by the standard [43]:

- The covered wind speed range extends from cut-in ( $3m/s$ ) to 1,5 times the wind speed at 85% of the rated power of the wind turbine (ca.  $17,5m/s$ )
- each bin contains at least 3 data points
- the total database contains at least 180 hours of sampled data

Figure 3.4 visualizes the concept of the method of bins. The grey lines indicate the bin borders, while the purple crosses indicate the resulting bin centers.

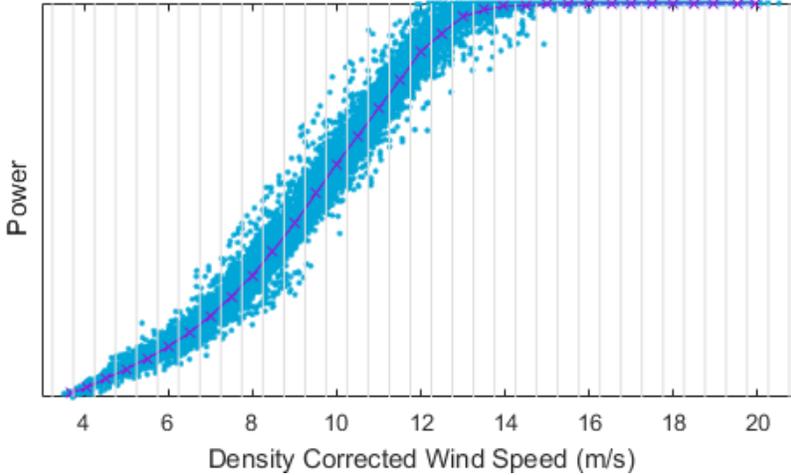
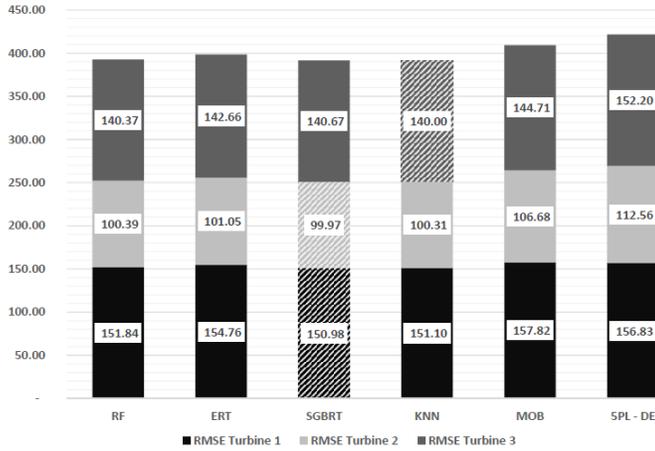


Figure 3.4: Visualisation of the method of bins. A measured power curve is divided into bins of  $0,5m/s$  (grey lines show the bin borders). For each bin the mean value for wind speed and power production is calculated (represented by the purple crosses).

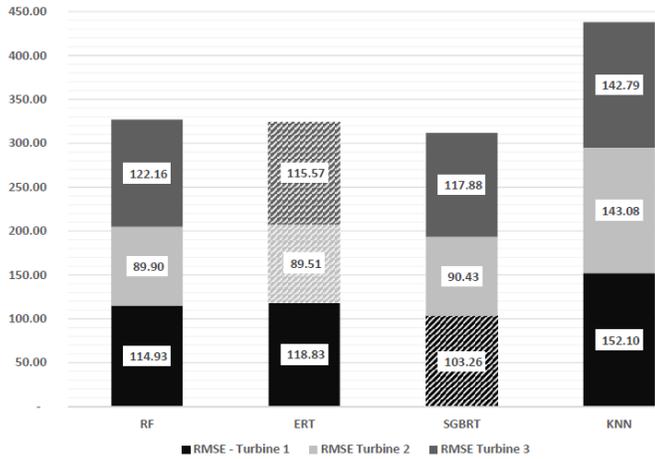
### 3.3.2 Advanced techniques for power curve modeling

The analyses described in this section were performed in collaboration with Olivier Janssens, Rik Van de Walle and Sofie Van Hoecke and were published in [48].

The method of bins (MOB), as described in Section 3.3.1, is a simple and fast method. However, its performance can be improved by other modeling techniques [48]. Using only wind speed as an input parameter (univariate models), [48] compared the method of bins to 5 other techniques: 5 parameter logistic curve fitted by differential evolution (5PL-DE), k-nearest neighbors regression (KNN), random forest regression (RF), extremely randomized trees (ERT) and stochastic gradient boosted regression trees (SGBRT). If possible, the techniques were applied on a combination of input parameters as well, so-called multivariate models. The considered input parameters were wind speed, yaw angle, blade pitch angle, rotor rpm and wind direction.



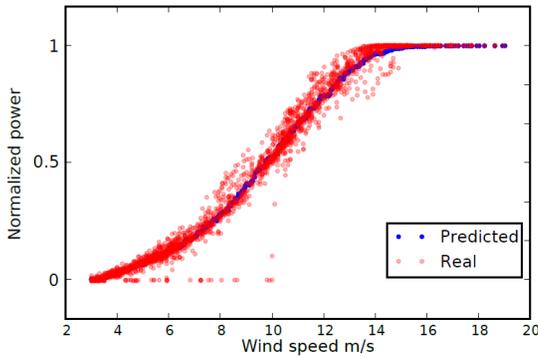
(a) RMSE of the univariate models



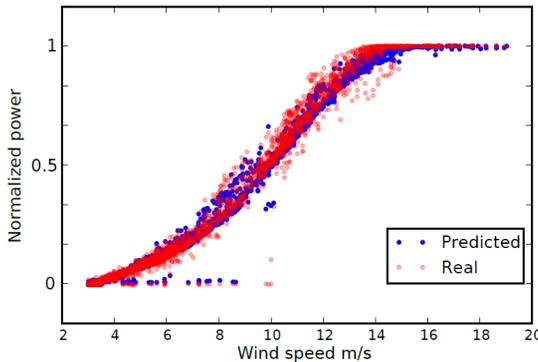
(b) RMSE of the multivariate models

Figure 3.5: Root Mean Squared Error between measured produced power and modeled produced power for 6 different modeling techniques and 3 different turbines. Both univariate (a) and multivariate (b) results are shown. Results are based on 17 days of measurements, training of the modeling was done using 75 days of measurements.

Different models were trained for 3 different turbines, based on 75 days of measurements. Figure 3.5 shows the resulting Root Mean Squared Errors (RMSE) between the measured and the modeled produced power during a validation period of 17 days. It can be seen that no significant improvement (maximal from  $144,71kW$  to  $140kW$  for the first turbine, from  $106,68kW$  to  $99,97kW$  for the second and from  $157,82kW$  to  $150,98kW$  for the third) is found when using a different univariate model compared to the MOB, while almost all multivariate models clearly improve the modeling performance. The best improvements obtained for multivariate models reduced the measured RMSE at the first turbine from  $144,71kW$  to  $115,57kW$ . For the second turbine a reduction from  $106,68kW$  to  $89,51kW$  was found and for the third from  $157,82kW$  to  $103,26$ . When using a univariate model, the modeled power curve basically boils down to one line. However, the measured power clearly shows some deviation from this line. If a multivariate model is used, this deviation from the line can be explained partially by other parameters. This explains the better results for multivariate models and is shown in Figure 3.6.



(a) The power curve of turbine three when only the wind speed as input is used



(b) Power curve when five input variables are used

Figure 3.6: Modeled and measured power curve for univariate tree (a) and multivariate tree (b).

## 3.4 Power curve monitoring

To detect smaller performance changes during normal operation, a monitoring approach based on power curve calculation is suggested here. A comparison between a power curve, calculated regularly with data during normal operation, and the warranted power curve given by the manufacturer is the basis of the approach.

### 3.4.1 Reference power curve

For every turbine type, the manufacturer provides a warranted power curve. This power curve gives the power production guaranteed by the manufacturer for the entire wind speed range. Figure 3.7 shows the warranted power curve for the Vestas 3MW V90 turbine, installed in the Belwind wind farm.

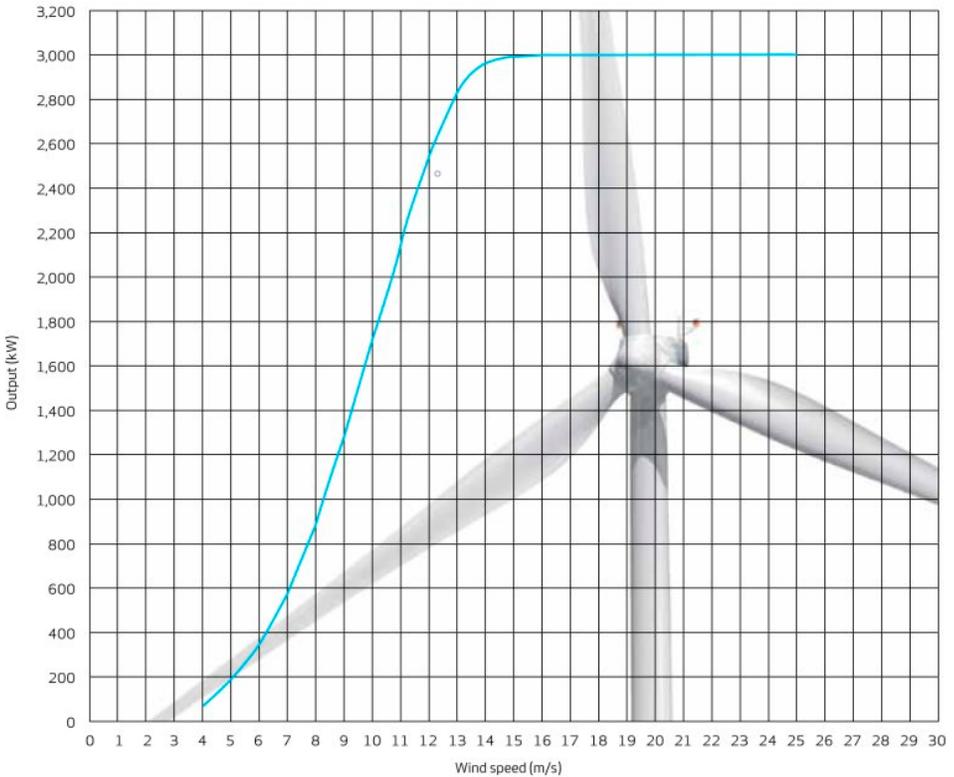


Figure 3.7: Warranted power curve of the turbine installed at Belwind, as given in the technical specifications of the manufacturer [49]

Multiple reasons can cause some deviation of the actual power curve with respect to the warranted power curve. Meteorological conditions can cause some deviation, such as wind shear, wind veer, turbulence or wake. While affecting the power production, these meteorological conditions are outside the control of

the operator. However, issues related to the turbine operation can decrease the power output as well, including the blade condition, turbine suitability, the control algorithm or maintenance [40]. Some of these issues can be resolved through maintenance, but only when the problem is detected. Therefore regularly calculating the actual power curve and comparing it to the warranted power curve can increase the total power production of a wind turbine by detecting possible faults early.

To compare the actual power curve to the warranted power curve, three different metrics are defined. One is based on the difference between measured and warranted power, the second is based on the area under the curves and the last is based on potential energy production. Conceptual illustrations of the two first metrics are provided in Figure 3.8.

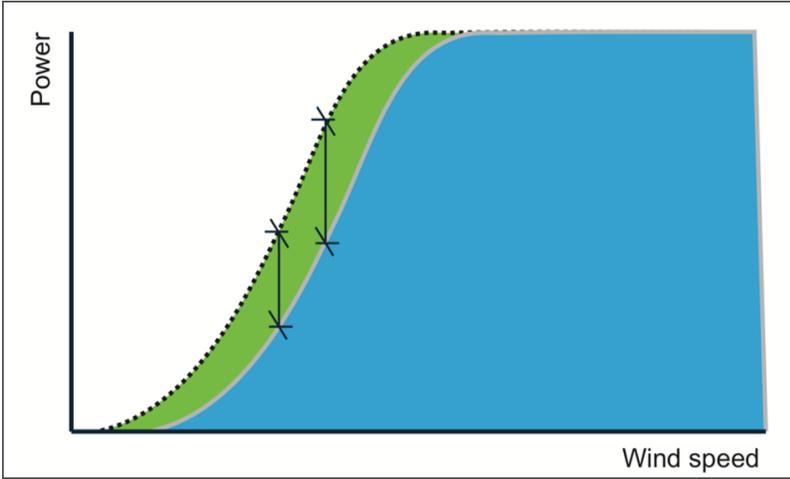


Figure 3.8: Health indications based on the actual power curve. The difference between the warranted curve (grey) and the actual power curve (dashed) can be defined by the difference between actual and warranted power (black) or by the ratio of the area under the actual power curve (green + blue) and the area under the warranted power curve (blue).

Assuming the warranted power curve is represented by an array of warranted power  $P_W$  with associated wind speed  $V_W$  and the actual power curve by  $P_A$  and  $V_A$ , the three different metrics can be defined as following.

The first one  $HI_{\Delta P}$  is based on the difference in power, as given by Equation 3.5.

$$HI_{\Delta P} = \sqrt{\sum_{i \in \text{range}_{meas}} \left( P_A(V_A(i)) - P_W(V_A(i)) \right)^2} \quad (3.5)$$

To calculate  $HI_{\Delta P}$ , first the warranted power associated with the actual averaged wind speed for each bin  $P_W(V_A(i))$  is calculated first based on linear interpolation between the given data points for the warranted power curve. This is represented by Equation 3.6.

$$\frac{P_W(V_A(i)) - P_W(V_W(i-1))}{P_W(V_W(i+1)) - P_W(V_W(i-1))} = \frac{V_A(i) - V_W(i-1)}{V_W(i+1) - V_W(i-1)} \quad (3.6)$$

Figure 3.9 illustrates the exact calculation, zooming in on one bin only.

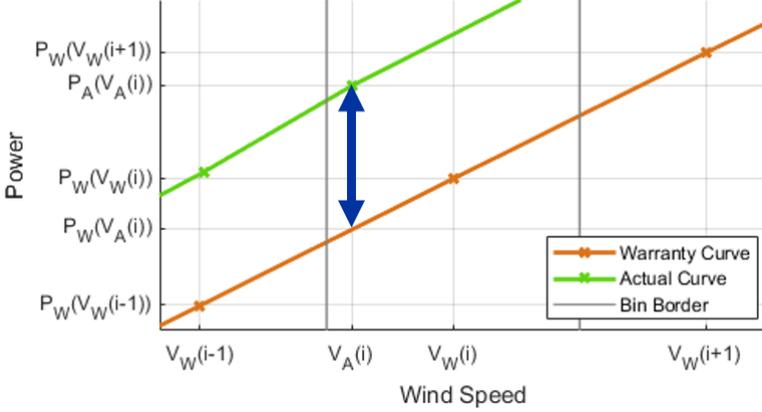


Figure 3.9: Calculation for health indication based on the difference in power (indicated by the blue arrow).

The second metric  $HI_{area}$  compares the area under the curves, as defined by Equation 3.7.

$$HI_{area} = \frac{\sum_{i \in range_{meas}} \frac{P_A(i) + P_A(i-1)}{2} \cdot (V_A(i) - V_A(i-1))}{\sum_{i \in range_{meas}} \frac{P_W(i) + P_W(i-1)}{2} \cdot (V_W(i) - V_W(i-1))} \quad (3.7)$$

For the last metric, the power curves are used to calculate the potential energy production based on the wind speed probability distribution  $P_r$  as expected during design. Therefore the potential energy production is calculated once based on the measured power curve and once based on the warranted power curve. The first one can represent the actual energy production, while the latter can represent the ideal energy production. The third metric  $HI_E$  gives the ratio between both, as given by Equation 3.8.

$$HI_E = \frac{\sum_{i \in range_{meas}} \frac{P_A(i) + P_A(i-1)}{2} \cdot P_r(V \in [V_A(i-1); V_A(i)])}{\sum_{i \in range_{meas}} \frac{P_W(i) + P_W(i-1)}{2} \cdot P_r(V \in [V_W(i-1); V_W(i)])} \quad (3.8)$$

The wind speed range for which both curves are compared is defined by the wind speed range measured during the analyzed period ( $range_{meas}$ ).

### 3.4.2 Farm-wide power curve monitoring

For this section the actual power curve of all turbines within the farm is calculated based on 2 years of data as an illustration. The actual power curve is calculated using the same filters as explained in Section 3.3. Once calculated the previously defined health indicators are applied to spot any variations in the farm. These results are provided in Figure 3.11.

As one can see in Figure 3.11, there is some variability in results. Based on the difference in power (Figure 3.11a) several turbines pop up with issues such as BBF05, BBA05, BBB02, BBA08, BBD06 and BBB10.

Thanks to  $HI_{area}$  (Figure 3.11b) turbines that perform better than expected can be distinguished from those performing below the expected level. If the calculated power curve exceeds the warranted power curve,  $HI_{area}$  will exceed 1 (indicated in green in the figure), while if the calculated power curve lies lower than the warranted power curve, and the turbine is thus underperforming,  $HI_{area}$  will be lower than 1 (in red in the figure). This is illustrated by Figure 3.10, where the power curve of BBD06 exceeds the warranted power curve, BBF05 produced less power than guaranteed and BBA01 performed as expected. One of the reasons for the difference between a calculated power curve and the warranted power curve can be a defective anemometry [40]. When the anemometer underestimates the actual wind speed, the power curve is shifted to the left and an above expected output is obtained. Oppositely when the wind speed is overestimated, the power curve is shifted to the right which will be perceived as below expected performance. Seeing the calculated power curve of turbine BBD06, a defective anemometer might cause the apparent excellent power production as there are no other mechanisms that would lead to a similar over-performance.

Looking at the results for  $HI_E$  (Figure 3.11c), the same observations as for the results of  $HI_{area}$  can be made. However, unlike  $HI_{area}$ , the resulting values give an indication of over- or underproduction in terms of energy. For example, BBB02 and BBD06 both seem to produce about 8% more energy than expected with the warranted power curve. On the other hand, BBA05 and BBF05 seem to produce about 8% less energy than expected.

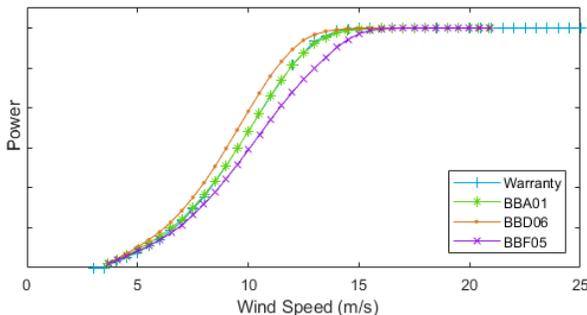
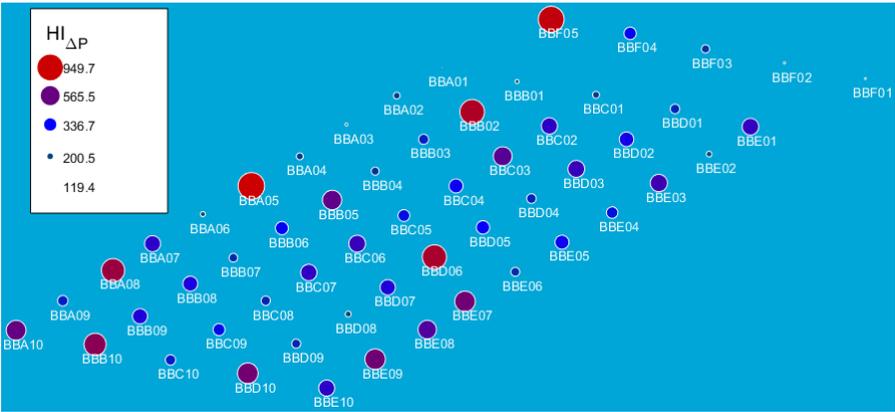
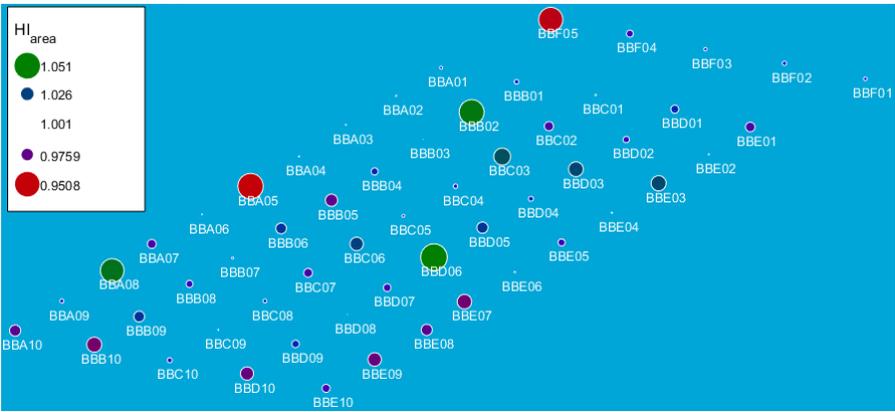


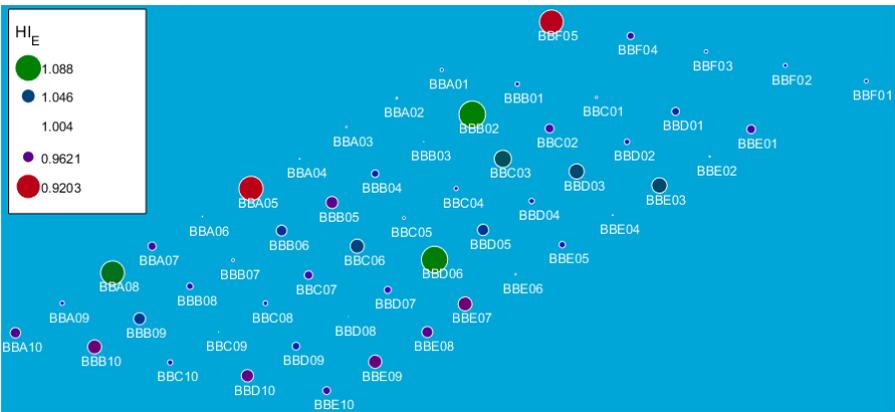
Figure 3.10: Example of different power curves leading to different values for the health indicators. One turbine with exceeding performance (BBD06), one with normal performance (BBA01) and one with low performance (BBF05) is chosen.



(a) Health indication based on the difference in power



(b) Health indication based on area under the curve



(c) Health indication based on potential energy

Figure 3.11: Resulting health indications for power curves calculated based on two years of data.

In some cases the health indicator based on the area  $HI_{area}$  and the energy  $HI_E$  show an excellent value of almost 1, while the health indicator based on the power difference  $HI_{\Delta P}$  is still a little higher than expected, for example for BBC01. In this case the shape of the power curve changed slightly (Figure 3.12): for lower wind speeds (under  $10\text{m/s}$ ) the calculated curve exceeds the warranted one, while for the higher wind speeds (above  $10\text{m/s}$ ) the inverse is true. Although the effect does not seem very severe, it can cause a lot of lost production over several months depending on the actual wind conditions. If the wind speeds for which the actual curve lies below the warranted curve often occur, a lot of production will be lost. However, again an issue with the anemometer, i.e. an erroneous sensitivity, could lead to this behavior.

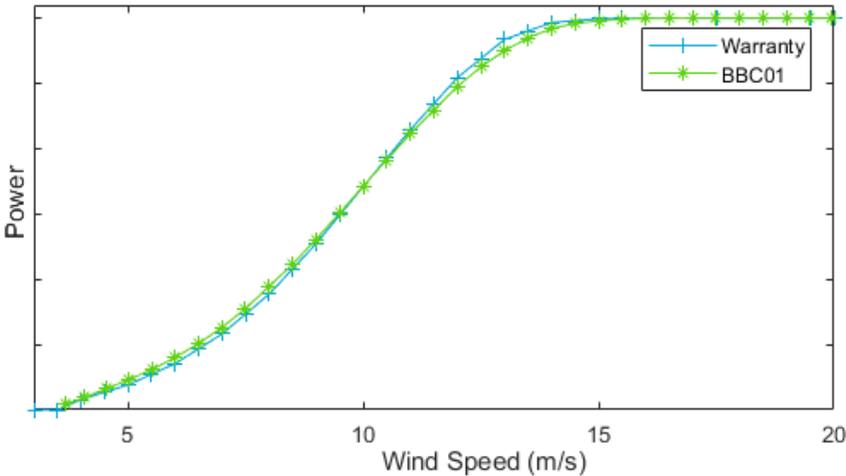


Figure 3.12: Example of almost perfect power curve according to the area-based health indicator, but not according to the health indicator based on the power difference

### 3.4.3 Power curve monitoring over time

To see long term changes in power production under normal operation, a monthly power curve can be calculated using data from the last month only. Again, the same filters are applied as in Section 3.3.

Figure 3.13 shows the results for the area-based health indicator  $HI_{area}$  for one turbine, BBF05, over a period of 4 years. One can observe a seasonal effect in the results: the area under the measured power curve seems to be lower in spring and summer, while it seems to increase during fall and early winter. Although a correction of wind speed is done for seasonal effects (through the air density correction explained in Section 3.2.2), this result suggests the correction suggested in [44] is not sufficient as the performance is better in the cooler months with denser air. However, it is shown atmospheric stability and turbulence influence

the power production as well [50]. Therefore, the observed seasonal change can also be the result of changing atmospheric conditions with different seasons.

Figure 3.13 also shows very low results for months of June and September in the final year. The calculated power curves for these months are shown in Figure 3.14, together with the calculated power curves of the months before, between and after. It can be seen the power curves of June and September are clearly lower than the power curves for the other months. No reason for this sudden reduction in performance was identified.

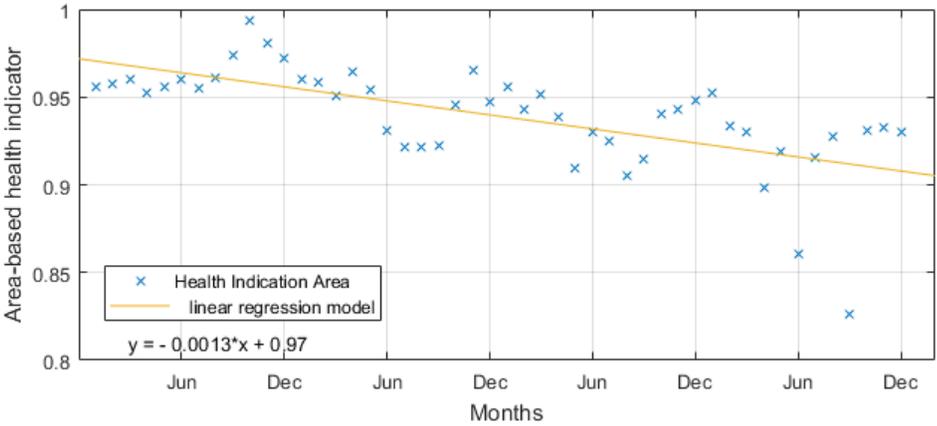


Figure 3.13: Results for area-based health indicator for power curves calculated each month during 4 years.

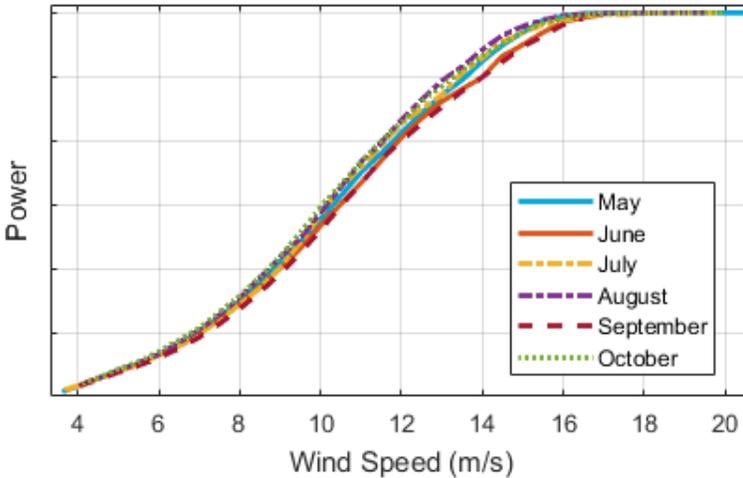


Figure 3.14: Calculated monthly power curves for BBF05, from the last May until October

Not only a seasonal effect is visible in Figure 3.13, a gradual decrease over time of the area under the measured power curve is also clearly present. This behavior hints at a gradual reduction of the turbine performance over time, e.g. as caused by blade erosion. To compare this decrease to the change over time for the other turbines within the farm, a linear regression model ( $HI_{area}(t) = S_{area} \cdot t + C$ ) is fit through the data. The resulting slope,  $S_{area}$  of each linear model is plotted for the whole farm in Figure 3.15.

The results in Figure 3.15 reveal that the gradual decrease over time was not unique to BBF05 as several turbines in the farm show similar or worse behavior.

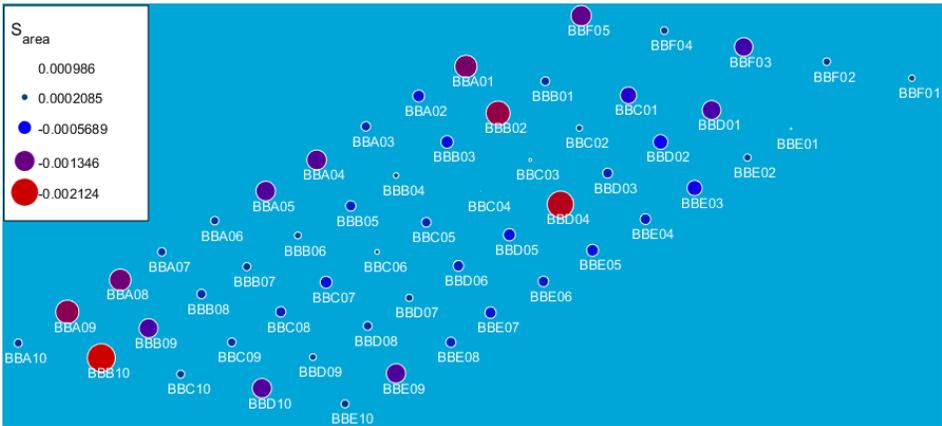


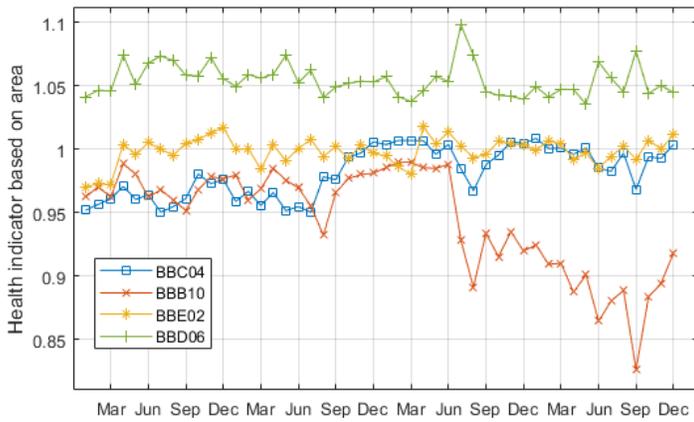
Figure 3.15: Resulting slope of linear regression model through area-based health indicator results of monthly power curves for all turbines within the farm.

Figure 3.16 shows the monthly evolution of the health indicators  $HI_{area}$ ,  $HI_{\Delta P}$  and  $HI_E$  for monthly curves over four years for several turbines. The turbines were selected based on their performance in Figure 3.15 ranging from the turbine with the highest increase, BBC04, the turbine with strongest decrease, BBB10, and two turbines that have intermediary results, BBE02 and BBD06.

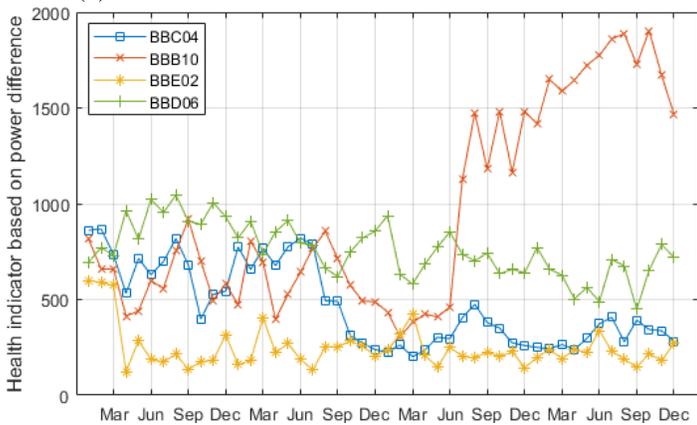
First of all, one can notice the seasonal effect is not clearly present for all turbines. For example turbine BBE02 has a steady behavior.

Furthermore, the low resulting slope of BBB10 can be explained by a sudden increase for  $HI_{\Delta P}$  (Figure 3.16b) and sudden drop for  $HI_{area}$  and  $HI_E$  (Figures 3.16a and 3.16c) in July of the third year. Such a sudden, albeit smaller, change can also be observed in April of the first year for BBE02. These changes are probably caused by a reposition of the anemometer that measures the wind speed. However, this is not confirmed with the maintenance records.

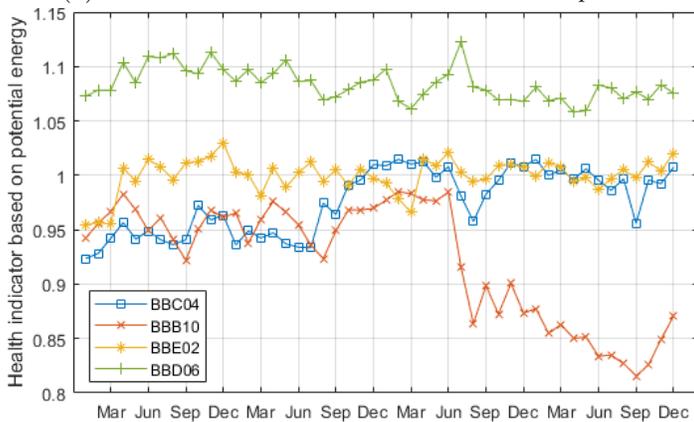
Finally, one can observe some sudden gains or drops for  $HI_{area}$ , and to a lesser extend for  $HI_E$  but no exceptional changes for  $HI_{\Delta P}$ , for example in July of the 3rd year for BBD06. Figure 3.17 shows that the power curve of July and August are very comparable. This explains the normal value for the health indicator based on power difference. However the power curve of July contains three bins less than the curve of August. Since these bins are all for wind speeds above rated power



(a) Health indication based on area under the curve



(b) Health indication based on the difference in power



(c) Health indication based on the potential energy

Figure 3.16: Variation of health indicators in time for 4 different turbines

and the turbine produces as expected from the warranty curve for those wind speeds, the results for the area-based health indicator differ. In August, the good behavior for lower wind speeds is compensated by those three bins leading to a lower value for the area-based health indicator. The same effect influences the results for the energy-based health indicator as well. However, to a lesser extent, since higher wind speeds are less expected.

In case the measured power curve is lower than the warranty curve, the same "lack" of presence of higher wind speeds can lead to lower values of the area-based health indicator. This is also partly the reason for the two very low results in Figure 3.13. In Figure 3.14, it can be seen that the highest wind speed measured in June and September is between  $16\text{m/s}$  and  $17\text{m/s}$ , while for the other months the wind speed easily exceeded  $19\text{m/s}$ .

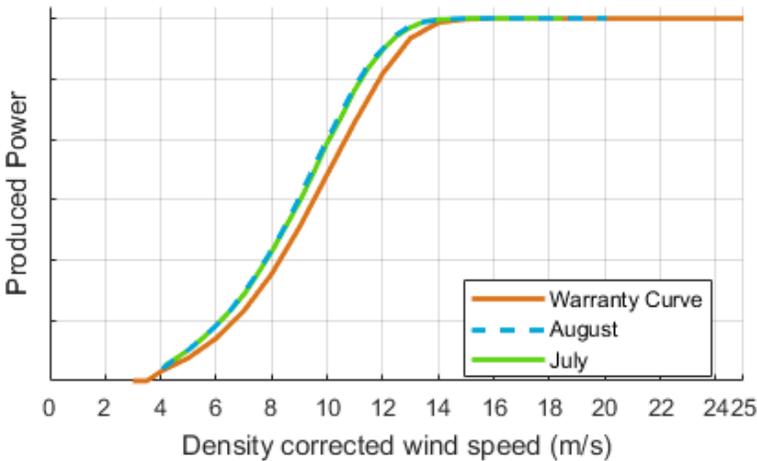


Figure 3.17: Monthly power curves of BBE06 for July and August in year 3. While the power curves in August and July are nearly identical the area based indicator was significantly larger for July due to less data in the upper wind speed ranges in July.

Unfortunately these artificial changes in  $HI_{area}$  will (wrongly) influence the results for  $S_{area}$  as well. One could consider to do a similar analysis for  $HI_{\Delta P}$ . However, for power curves very close to the warranted power curve, the resulting slope  $S_{\Delta P}$  will not indicate the actual change over time of the power curve. Moreover, depending on whether the turbine is over- or underproducing, an increasing power difference can be beneficial or disadvantageous. On top of that, similar artificial changes in  $S_{\Delta P}$  could occur.

A better solution might be to do the same analysis for  $HI_E$ . Although these artificial changes will still influence the resulting slope  $S_E$ , it will be less pronounced than for  $S_{area}$ .

### 3.5 Rotor effective wind speed estimation

As mentioned in Section 3.4.2 and [40], defective anemometry or an unreliable wind speed measurement can influence the power curves. [41, 42] point out the higher uncertainties of the SCADA wind speed measurement.

To illustrate this, a dataset containing two months of data is used. This data was filtered for four different sections of wind direction. For each wind sector, all turbines of one string in the farm were outside the wake of the other turbines. To make sure this was the case, each sector spanned 140 degrees. Moreover only data points were considered for which all turbines were operational.

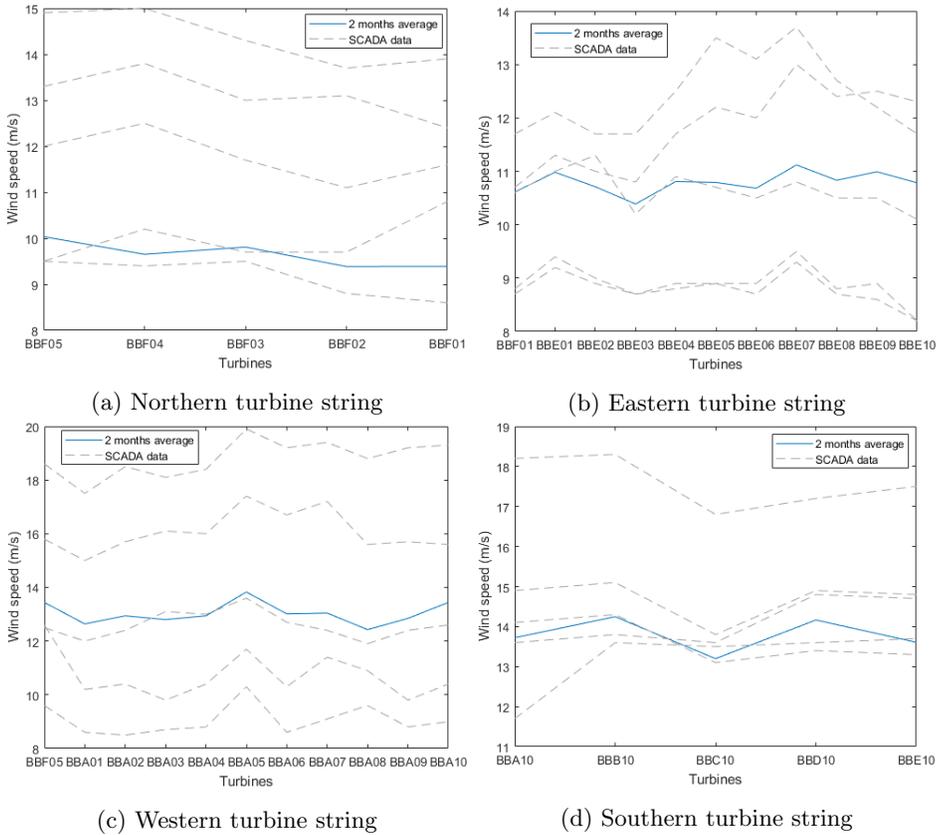


Figure 3.18: Wind speed measured by turbines in free wind, averaged over two months of measuring for the free wind sector (blue full line) together with some 10 minute averages of wind speed as captured by the SCADA system (gray dashed line). This is done for the northern (a), eastern (b), western (c) and southern (d) of free wind turbines.

Since all turbines of that one string are out the wake of the others, one assumes the 10 minute average SCADA value of wind speed for all time instances for those turbines should be very similar. However the results shown in Figure 3.18 indi-

cate differently. For each sector, 5 timestamps of ten minute averaged (free) wind speeds as given by the SCADA system are shown in gray. One can see differences up to more than 2 m/s. These differences are higher than expected for turbines standing in free wind.

When looking at the overall average value for wind speed of all remaining data points per turbine for each sector, it is expected the resulting averages are very similar to each other. This resulting averaged wind speed of each turbine in a free stream turbine string, based on the filtered data for each sector, is shown in blue in Figure 3.18. One can see differences among the free turbines of more than 1 m/s for the averages over two months, which is not considered close to each other. These results indicate the wind speed measurements by the anemometer of the wind turbine are not reliable for all turbines.

To bypass the erroneous wind speed measurement, [41] suggests to calculate the rotor effective wind speed, using Equation 3.9. This equation is based on the one for power production (Equation 2.4) as introduced in Section 2.1.7. In this equation the produced power  $P$  and the air density  $\rho$  are known from the SCADA dataset and meteorological dataset. The rotor radius  $R_{rotor}$  is known to be 45m for a Vestas V90 3MW turbine. To be able to calculate the rotor effective wind speed  $V_{rews}$ , the power coefficient  $c_P$  is needed as well. Basically, the power coefficient gives an idea on how efficient the energy conversion actually is.

$$P = \frac{1}{2} \rho c_P \pi R_{rotor}^2 V_{rews}^3 \quad (3.9)$$

One could choose to use the  $c_P$ -curve as given by the manufacturer (Figure 3.19). However, in abnormal operation, e.g. during down-regulation,  $c_P$  will not follow this ideal curve, as a result the wind speeds will be estimated wrongly in those cases. Therefore, numerous functions to estimate the  $c_P$  are found in literature [31, 52, 53, 54, 55, 56]. In this thesis three different functions are considered. For each function, different coefficients can be found. Two sets for each function are applied.

$$c_P(\lambda, \theta) = c_1 \left( \frac{c_2}{\lambda_i} - c_3 \theta - c_4 \theta^{c_5} - c_6 \right) e^{-\frac{c_7}{\lambda_i}} \quad (3.10)$$

$$\lambda_i = \left( \frac{1}{\lambda + c_8 \theta} - \frac{c_9}{\theta^3 + 1} \right)^{-1}$$

The first function (Equation 3.10), in the remainder of the section referred to as F1, is suggested by [31, 52], using the blade pitch angle  $\theta$  and the tip speed ratio  $\lambda$ :

$$\lambda = \frac{\Omega R_{rotor}}{V} \quad (3.11)$$

where  $\Omega$  is rotor speed in *radians/s*,  $R_{rotor}$  is the rotor radius in meter and  $V$  is the wind speed in *m/s*.

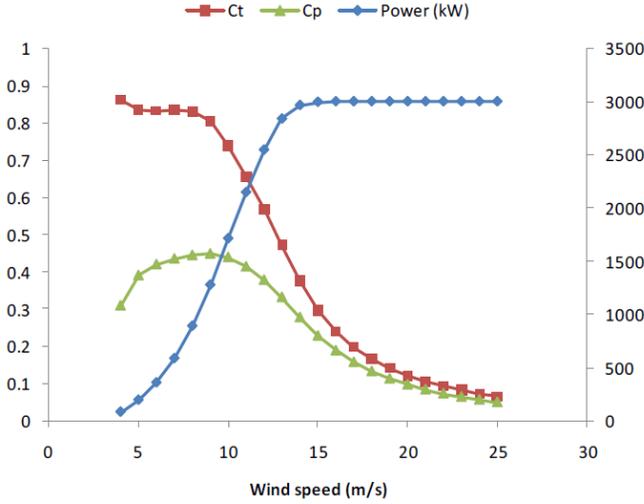


Figure 3.19: Produced power (kW), power coefficient  $c_P$  and thrust coefficient  $C_t$  versus wind speed for a Vestas V90 3MW turbine [51]

Two different sets of coefficients are found in the aforementioned papers and applied on the data. The coefficients used are given, along with their respective sources, in Table 3.2.

The second function (referred to as F2) slightly differs, as suggested by [53, 54, 55, 56]:

$$c_P(\lambda, \theta) = c_1 \left( \frac{c_2}{\lambda_i} - c_3\theta - c_4\theta^{c_5} - c_6 \right) e^{\frac{-c_7}{\lambda_i}} \quad (3.12)$$

$$\lambda_i = \left( \frac{1}{\lambda + c_8} - \frac{c_9}{\theta^3 + 1} \right)^{-1}$$

Again two different sets of values for the coefficients are found and applied (Table 3.2).

The last and third function (referred to as F3) implemented for this thesis was proposed by [57, 58, 59]:

$$c_P(\lambda, \theta) = [c_1 - c_2(\theta - c_3)] \sin \left( \frac{\pi(\lambda + c_4)}{c_5 - c_6(\theta - c_3)} \right) - c_7(\theta - c_3)(\lambda - c_8) \quad (3.13)$$

The two different sets of values for the coefficients applied here are again given in Table 3.2.

Table 3.2: Coefficient sets for multiple  $c_P$ -functions from literature as applied on available datasets.

	$F1_1$	$F1_2$	$F2_1$	$F2_2$	$F3_1$	$F3_2$
Equation	3.10	3.10	3.12	3.12	3.13	3.13
$c_1$	0,5	0,73	0,5	0,5	0,5	0,44
$c_2$	116	151	116	98	0,00167	0,0167
$c_3$	0,4	0,58	0,4	0,4	2	0
$c_4$	0	0,002	0	0	0,1	-3
$c_5$	1,5	2,14	1	1	12	15
$c_6$	5	13,2	5	5	0,3	0,3
$c_7$	21	18,4	21	16,5	0,00184	0,00184
$c_8$	0,08	0,02	0,088	0,089	3	3
$c_9$	0,035	0,003	0,035	0,035		
Reference	[31]	[52]	[53, 54]	[55, 56]	[57]	[58, 59]

Since the tip speed ratio and thus  $c_P$  depends on the wind speed, it is not as straightforward to solve Equation 3.9 for the wind speed  $V_{rews}$ . To solve this, the function given in Equation 3.14 is calculated for every timestamp  $i$ , where  $v$  represents an array of possible values for wind speed between  $0m/s$  and  $30m/s$ , with an accuracy of  $0,1m/s$ . The value for the rotor effective wind speed  $V_{rews}$  is taken as the value of  $v$  such as  $f(V_{rews})$  is zero. If multiple zeros occur, the closest to the measured wind speed is taken. If no zeros occur, the value of  $v$  is taken where the absolute value of  $f$  is minimum. This means that Equation 3.14 has to be solved for each timestamp separately.

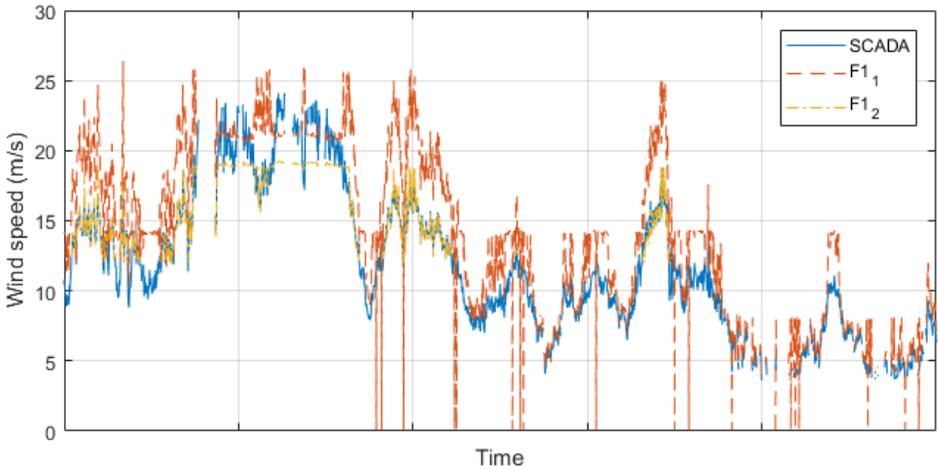
$$f(v) = P(i) - \frac{1}{2} \rho c_P(v, \Omega(i), \theta(i)) \pi R_{rotor}^2 v^3 \quad (3.14)$$

Since the calculation of rotor effective wind speed can only be accurate if the turbine is producing power, it is only calculated for those data points for which the measured produced power is greater than 0 kW during the full 10 minutes of each data point. To illustrate, the proposed strategy is applied on a small period for one turbine for all functions and coefficient sets.

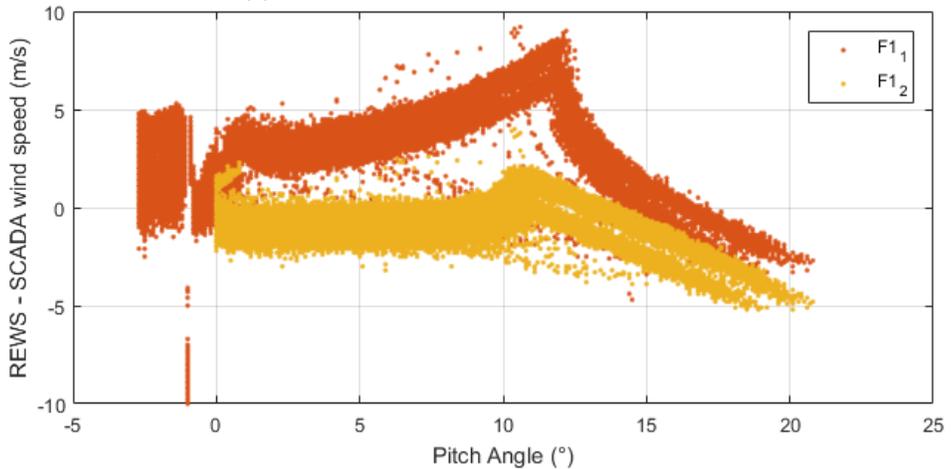
Figure 3.20 shows the resulting rotor effective wind speed calculated using Equation 3.10 for both sets of coefficients,  $F1_1$  and  $F1_2$  respectively. In Figure 3.20a one can see that the coefficients proposed in [31],  $F1_1$  lead to an overestimation of the wind speed fairly often and that for some periods the estimated wind speed stays almost constant. Figure 3.20b shows the overestimate by  $F1_1$  is mainly for pitch angles around  $12^\circ$ . For larger pitch angles, the difference between the estimated and the measured wind speed decreases fairly quickly. Moreover, instability can be observed for pitch angles around  $-1$ , caused by values close to zero in the denominator of Equation 3.10.

The coefficients proposed in [52],  $F1_2$ , never return a value for wind speed higher than ca.  $19m/s$  (Figure 3.20a). This is also represented by the decrease in differ-

ence with the SCADA wind speed for pitch angles higher than  $10^\circ$  (Figure 3.20b). Moreover, no valid estimations are found for pitch angles lower than 0, because of the impossibility to calculate a fractional exponent of a negative base. Besides those flaws the match seems rather nice.



(a) SCADA wind speed and REWS vs time

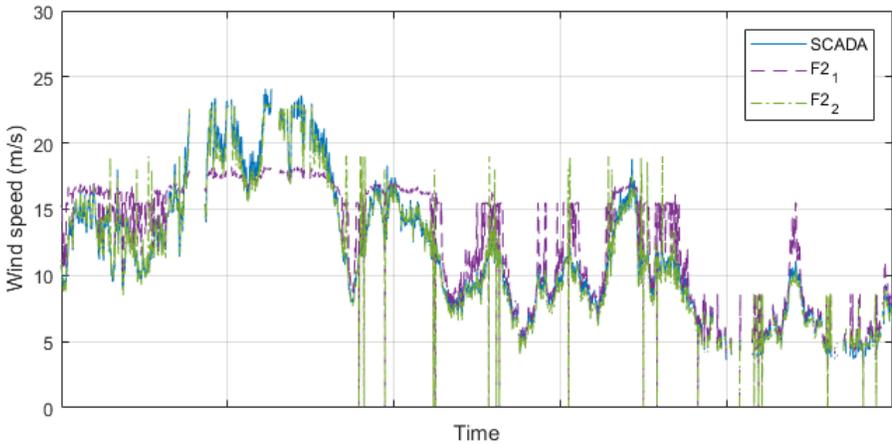


(b) Difference between REWS and SCADA wind speed vs pitch angle

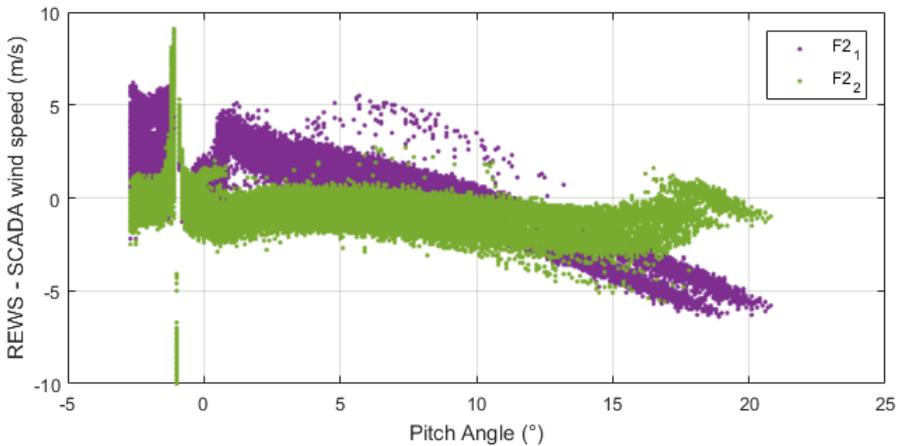
Figure 3.20: Rotor effective wind speed using Equation 3.10 compared to measured wind speed by SCADA system

Figure 3.21 shows the resulting rotor effective wind speed calculated using Equation 3.12 for both sets of coefficients,  $F2_1$  and  $F2_2$  respectively. In Figure 3.21a one can see the use of the coefficients proposed in [53],  $F2_1$  never gives a value for wind speed higher than ca.  $18\text{m/s}$ . Looking at Figure 3.21b, the difference with the measured wind speed is decreasing starting from a pitch angle of  $0^\circ$ . Again, instability can be observed around a pitch of  $-1^\circ$ , caused by values close to zero in the denominator of Equation 3.12.

Based on Figure 3.21 the agreement between the measured wind speed and the rotor effective wind speed calculated using Equation 3.12 and the coefficients proposed in [55],  $F2_2$ , seems to be fairly good with differences barely exceeding  $3\text{m/s}$ . However, instability is observed here as well for pitch angles around  $-1^\circ$ .



(a) SCADA wind speed and REWS vs time

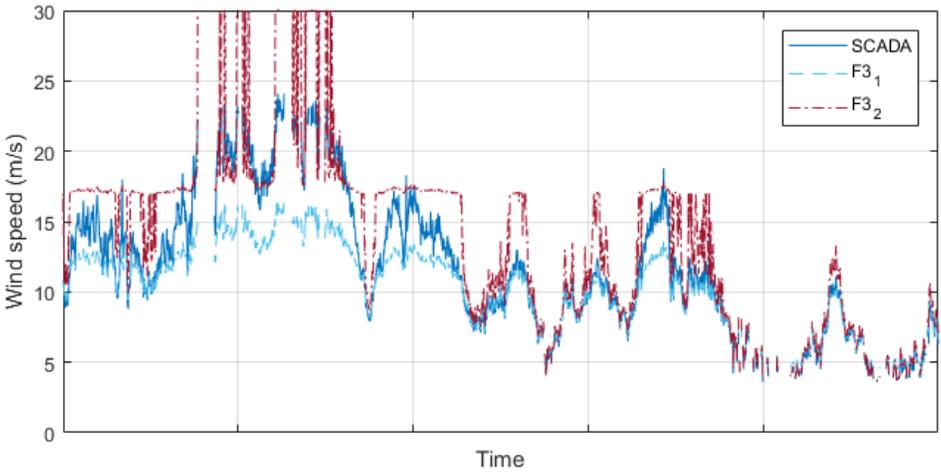


(b) Difference between REWS and SCADA wind speed vs pitch angle

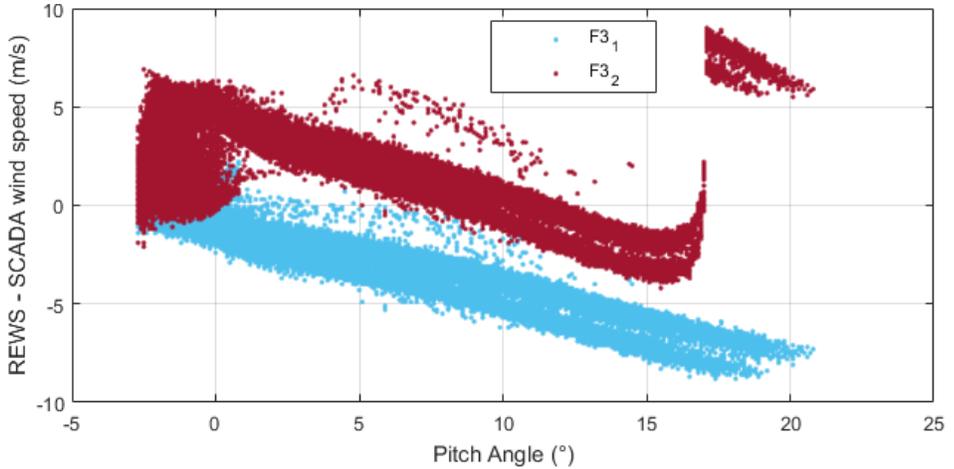
Figure 3.21: Rotor effective wind speed using Equation 3.12 compared to measured wind speed by SCADA system

Finally, Figure 3.22 shows the results for the third function (Equation 3.13),  $F3$ . For the coefficients suggested in [57],  $F3_1$ , a scaling problem seems to be present (Figure 3.22a), which is represented by the decreasing line for the difference to the SCADA wind speed as well (Figure 3.22b).

The coefficients proposed by [58],  $F3_2$ , on the other hand, seem to have a discontinuity around a pitch angle of  $17^\circ$ . For greater pitch angles, which occur typically for higher wind speeds, the estimated wind speed is well above the wind speed recorded in the SCADA, leading to high differences as well.



(a) SCADA wind speed and REWS vs time



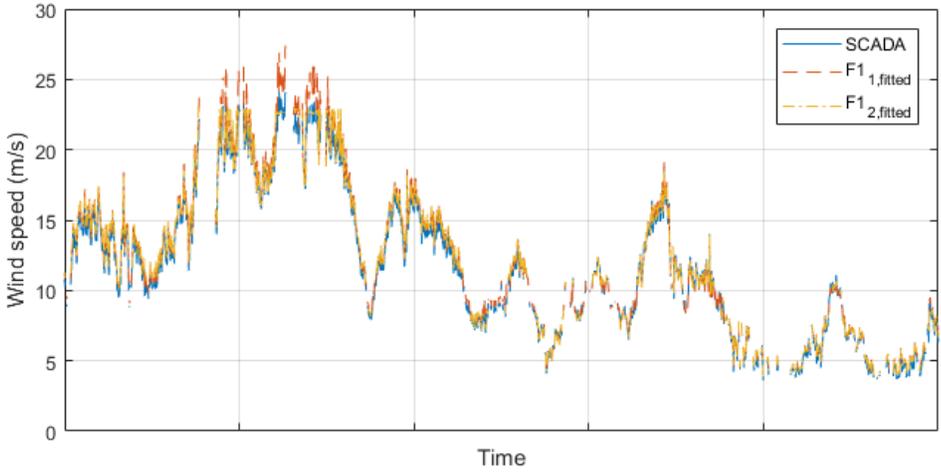
(b) Difference between REWS and SCADA wind speed vs pitch angle

Figure 3.22: Rotor effective wind speed using Equation 3.13 compared to measured wind speed by SCADA system

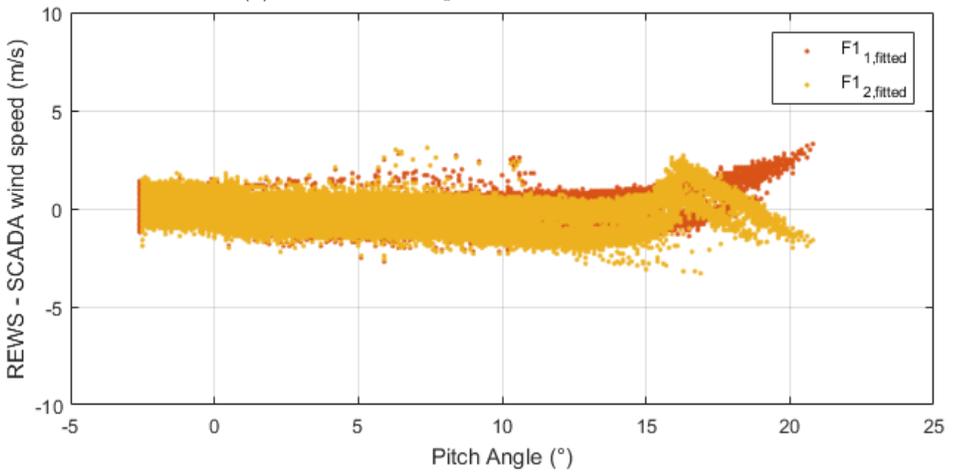
From the results above it is clear that none of the proposed solutions in literature are adequate for the intended purpose. Rather than using pre-defined coefficients, [41] proposes to obtain the actual coefficients by fitting the functions for  $c_P$  to the curve given in Figure 3.19. Unfortunately this is only possible when the  $c_P$ -curve is publicly available. The fitting is done using the standard multidimensional unconstrained nonlinear minimization method of MATLAB, Nelder-Mead, between the  $c_P$  curve and the  $c_P$  functions. To obtain the needed rotor speed and pitch angle values for the  $c_P$  functions, a rotor speed curve and a pitch angle curve is calculated in the same way as the power curve in Section 3.3.1 using the method of bins. These curves are shown in Figures 2.4 and 2.5. This standard fitting method of MATLAB requires start values for the parameters to be optimized. In this case, these parameters are the coefficients of the  $c_P$  function. As start values for the coefficients, the values from the literature given in Table 3.2 are used. This means each function is optimized twice, each time with a different set of start values.

Moreover, to tackle the problems caused by negative pitch angles for the functions given by Equations 3.10 and 3.12, a pitch offset with a starting value of 3 degrees is introduced and fitted together with the coefficients.

Figure 3.23 shows the updated results using the fitted coefficients for the  $c_P$  function expressed by Equation 3.10 with an included offset for the pitch angle. The errors between the predicted wind speed and those found in the SCADA shown in Figure 3.23b are for both coefficient sets clearly lower than in Figure 3.20b. For the first coefficient set (started from the values of [31])  $F1_{1,fitted}$ , the high wind speeds are a little bit overestimated, indicated by the increasing difference for high pitch angles (Figure 3.23b) and seen in Figure 3.23a. In case of the second coefficient set (started from the values given by [52])  $F1_{2,fitted}$ , a clear decreasing line for higher pitch angles is still present, indicating a topped estimation of wind speed for higher wind speeds (as can be seen in Figure 3.23a for wind speeds above ca.  $22m/s$ ).



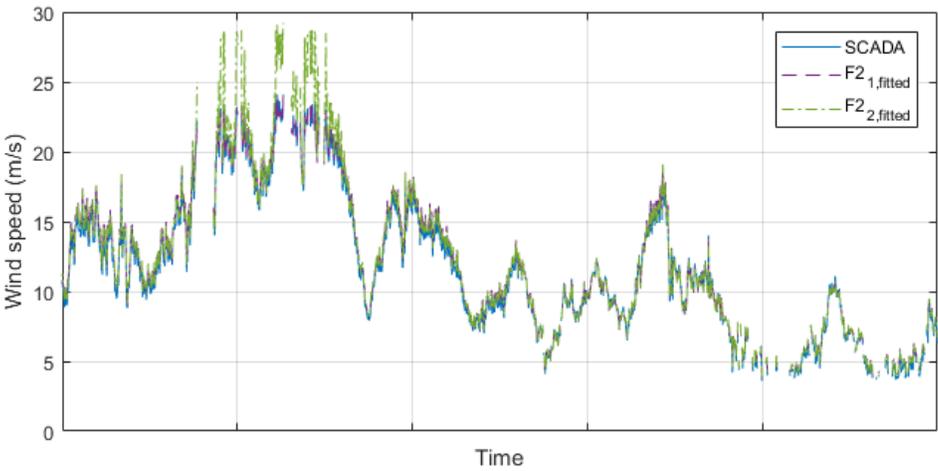
(a) SCADA wind speed and REWS vs time



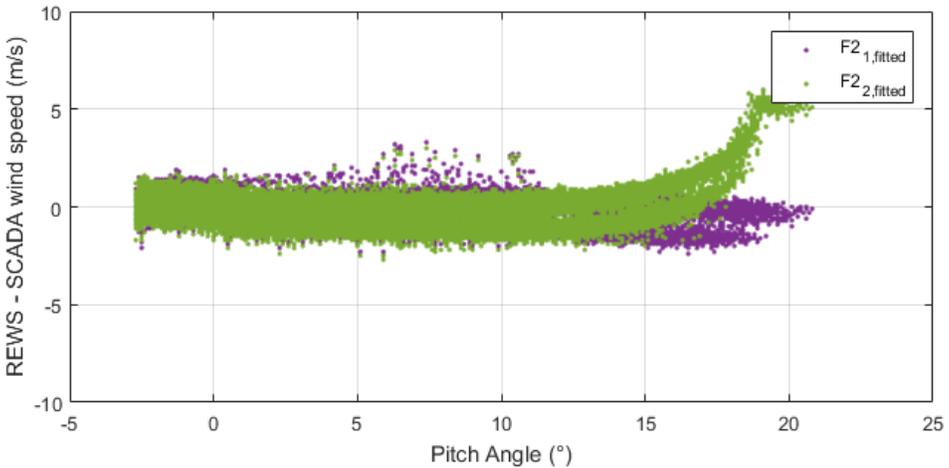
(b) Difference between REWS and SCADA wind speed vs pitch angle

Figure 3.23: Rotor effective wind speed using Equation 3.10, fitted coefficients and an included pitch offset compared to measured wind speed by SCADA system

Figure 3.24 shows the updated results using the fitted coefficients for the  $c_P$  function expressed by Equation 3.12 with an included offset for the pitch angle. A very good agreement, with differences barely exceeding  $2\text{ m/s}$ , is found between the SCADA wind speed and the wind speed estimation using Equation 3.12 with the first fitted coefficient set (started from values given by [53, 54])  $F2_{1,\text{fitted}}$ , both in time series (Figure 3.24a) as in difference to SCADA (Figure 3.24b). For the second fitted coefficient set (started from values of [55, 56])  $F2_{2,\text{fitted}}$ , a good agreement is found until pitch angles of ca. 14 degrees. For higher pitch angles (and wind speeds) an overestimation is visible.



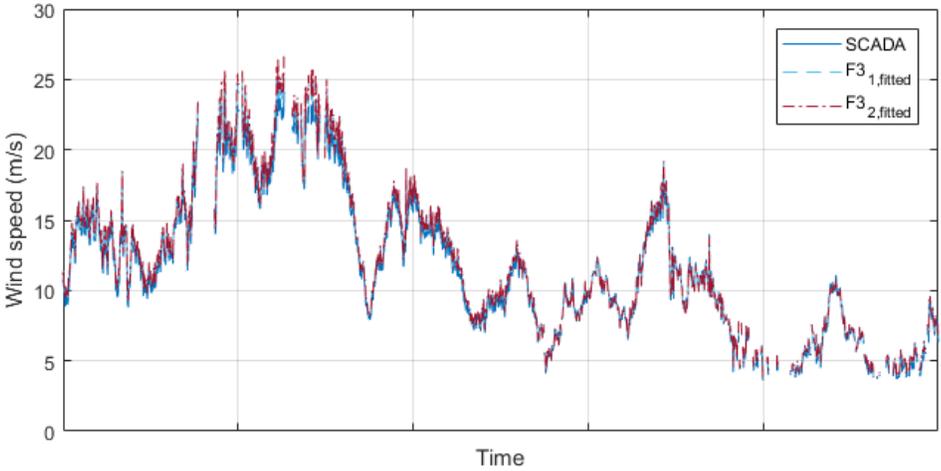
(a) SCADA wind speed and REWS vs time



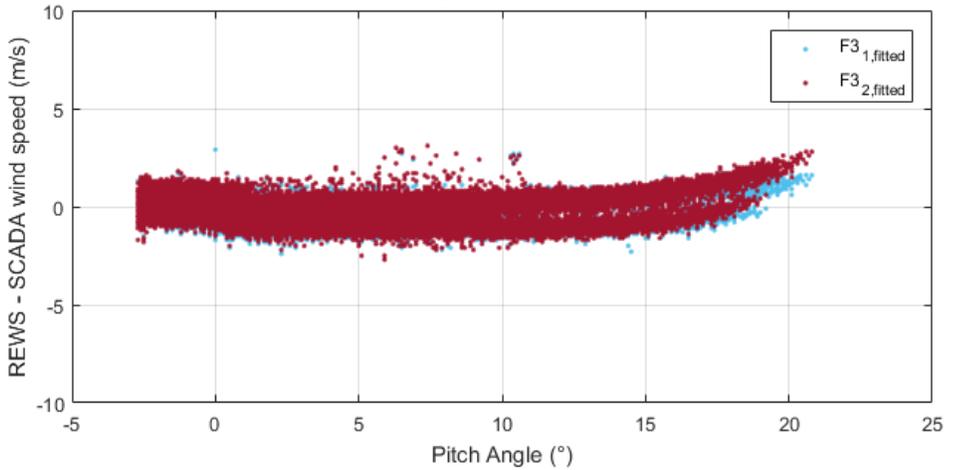
(b) Difference between REWS and SCADA wind speed vs pitch angle

Figure 3.24: Rotor effective wind speed using Equation 3.12, fitted coefficients and an included pitch offset compared to measured wind speed by SCADA system

Figure 3.25 shows the updated results using the fitted coefficients for the  $c_p$  function expressed by Equation 3.13. The estimated wind speeds clearly match better to the measured SCADA wind speed with respect to Figure 3.22. For both coefficient sets, an overestimation is seen for higher pitch angles, coinciding with high wind speeds.



(a) SCADA wind speed and REWS vs time



(b) Difference between REWS and SCADA wind speed vs pitch angle

Figure 3.25: Rotor effective wind speed using Equation 3.13 and fitted coefficients compared to measured wind speed by SCADA system

The resulting values for the rotor effective wind speed, for each  $c_P$  function and each coefficient set, are now compared for free stream measurements only. Figure 3.26 shows the calculated rotor effective wind speed for the same free turbine strings and periods from Figure 3.18. One can see that the differences between the rotor effective wind speeds are much lower than those between the measured SCADA wind speeds. These results suggest the rotor effective wind speed can definitely be a better estimation for the actual wind speed seen by the turbine. However a comparison to an independent measurement such as a met mast is still needed as confirmation. Unfortunately no such data is available for the considered Belwind site.

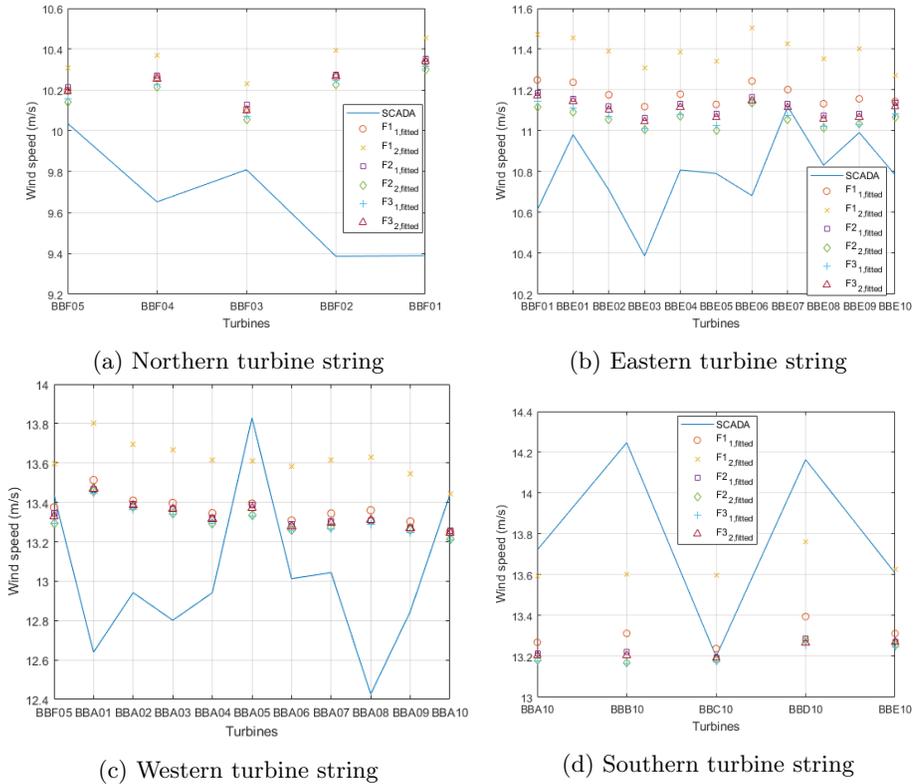


Figure 3.26: Wind speed measured by turbines in free wind, averaged over two months of measuring for the free wind sector (blue full line) together with the calculated rotor effective wind speed, using three different functions with 2 fitted coefficient sets each.

Multiple possibilities for a  $c_P$  function can work out, the preference after the obtained results in this thesis would be the function expressed by Equation 3.12 with the first set of fitted coefficients (started from the values given by [53, 54]),  $F_{2,1,fitted}$ , or the function expressed by Equation 3.13 with any of the two sets of

fitted coefficients.  $F3_{1,fitted}$  and  $F3_{2,fitted}$ .

Given the impossibility to validate the results with an independent wind speed measurement, it is chosen not to use the rotor effective wind speed in the remainder of this thesis.

## 3.6 Conclusions

In this chapter 10min SCADA data was used for performance monitoring. Instead of detecting instant and large anomalies in the performance, the aim was to detect small performance changes over time. For that reason, power curves to represent the normal operation of a turbine were calculated on a regular basis, using the method of bins. This method was also compared to other univariate and multivariate techniques. Although multivariate techniques shown an improvement of ca. 20% with respect to the method of bins, the latter was used for the performance monitoring, as the benefits of simplicity and the low computational effort outweighed the better results of multivariate techniques.

To monitor the performance, the calculated power curves are compared to the warranted power curve based on three health indicators. By tracking the value of those over time, seasonal effects and the general decline in performance of a turbine could be detected. Moreover turbines within a farm can be compared to each other based on those health indicators. In that way, under-performing turbines can be identified and investigated. The general decline or increase in performance over time is also compared on farm level.

An important dependency of the results on the quality of the wind speed measurement is observed. Since the SCADA measurements tend to be unreliable, a solution is suggested. The rotor effective wind speed is calculated using different functions for the power coefficient. Results show the coefficients of these  $c_P$  functions should be fit to the  $c_P$  curve given by the manufacturer in order to obtain acceptable results. Moreover, for the functions defined in Equations 3.10 and 3.12 a pitch offset needs to be inserted to deal with mathematical instabilities. Final results show potentially a more reliable wind speed estimation than the measurements available in the SCADA system by using the concept of the rotor effective wind speed. However, without an independent measurement of the free wind speed, it is impossible to conclude this.



# Chapter 4

## Thrust modeling

The majority of the work presented in this chapter has been published in [18].

*A reliable load history is crucial for a fatigue assessment of wind turbines. However, installing strain sensors on every wind turbine is economically not feasible. In this chapter, a technique is proposed to reconstruct the thrust load history of a wind turbine based on high frequent SCADA data. Strain measurements recorded during a short period of time are used to train a neural network. The selection of appropriate input parameters is done based on Pearson correlation and mutual information. Once the training is done, the model can be used to predict the thrust load based on SCADA data only. Multiple alternatives of the same technique are clarified and validated using 10 minute averages and mainly 1 second SCADA data from both a real world turbine and simulation data (FAST). In general, the relative error between simulated/measured and predicted thrust load barely exceeds 20% under normal operating conditions. Furthermore, the model obtained for one turbine is also applied on a different turbine of the same type to validate the transferability of the model to the rest of the wind farm. The cross validation resulted in comparable results and is considered successful. Finally, some drawbacks of the technique were illustrated by performing the technique on data of lesser quality.*

## 4.1 Introduction

Load estimation is an important part of a fatigue assessment. To have an idea about the loads acting on an OWT, installing additional sensors is a common solution. However, not all additional sensors stay reliable over time. For example, strain sensors tend to drift in time. Moreover some data, e.g. SCADA data, is already available and still underutilized. The research presented in this chapter aims to replace the use of strain gauges for the estimation of quasi-static loads. Existing approaches to estimate thrust loads are based on simulations and additional design information (e.g. thrust coefficient) or acceleration measurements [29, 60].

Although SCADA data is available for every wind turbine by default, its possibilities for load monitoring are still underutilized. Several authors [61, 62] have suggested to use 10min SCADA statistics to estimate the loads on the blades. If the estimated model uses SCADA data solely, it can be translated to every turbine in the farm without the need of installing additional sensors. Recently, also the use of 1s SCADA signals is becoming common practice in industry. Therefore the use of (1s) SCADA data to estimate the thrust load, acting on the wind turbine and its substructure, is proposed.

A correct estimation of the acting thrust load is not only of value in terms of fatigue assessments. It is also possible to associate the thrust with properties of wake flows. Therefore, an accurate estimation of thrust also proved important in estimating wake wind speeds and turbulences [63].

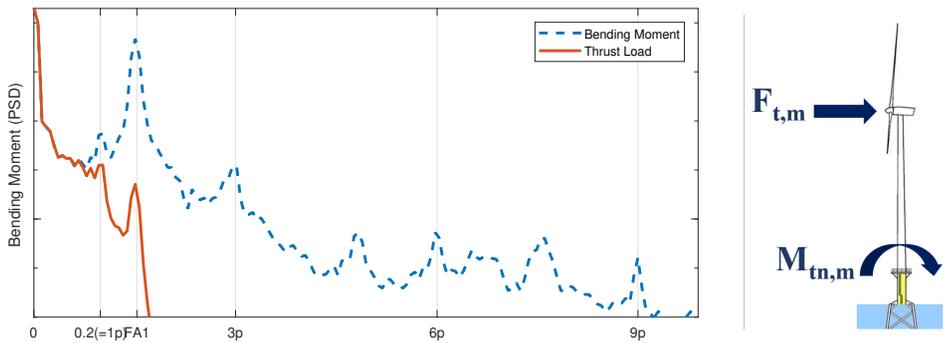
Moreover, since the thrust coefficient and the power coefficient can be linked to each other (as mentioned in Section 2.1.7), an accurate thrust estimation can also help in improving, or facilitating if the  $c_P$ -curve is not available, the calculation of rotor effective wind speed (Section 3.5).

## 4.2 Available datasets

### 4.2.1 Monitoring setup

In this chapter, results are shown using measurements taken at an offshore wind turbine. The monitored turbine is installed on a jacket and instrumented with strain gauges at the interface between transition piece and tower (Figure 4.1b).

Although this chapter will focus on the estimation of the thrust load, additional loads with higher frequencies are measured by the strain gauges as well. These additional loads were introduced in Section 2.1.6. Figure 4.1ba shows the frequency spectrum of measured bending moments, as induced by the thrust load and the additional loads. The quasi-static contribution of the thrust load to the measured bending moment  $M_{tn,m}$  is obtained by using a Butterworth filter of 4th order on the recorded bending moments in a frequency range from  $0Hz$  to  $0,2Hz$ . This frequency band is defined in a way the filtered signal is not influenced by the first natural frequency ( $0,31Hz$ ), since this is unrelated to any SCADA signal. The final filtered signal is shown by the red solid line in Figure 4.1a. The filtered signal represents the quasi static load and does no longer contain the effects of any



(a) Frequency spectrum of measured tower bending moment in fore-aft direction during 10 minutes (blue dashed line). The quasi-static part of the bending moment is filtered out (red solid line). The targeted quasi static load (filtered) does no longer contain the effects of rotor harmonics.

(b) The strain gauges are located at the interface between tower and transition piece.

Figure 4.1: The measured thrust load  $F_{t,m}$  is obtained using the bending moment  $M_{tn,m}$  measured by strain gauges located at the interface between tower and transition piece.

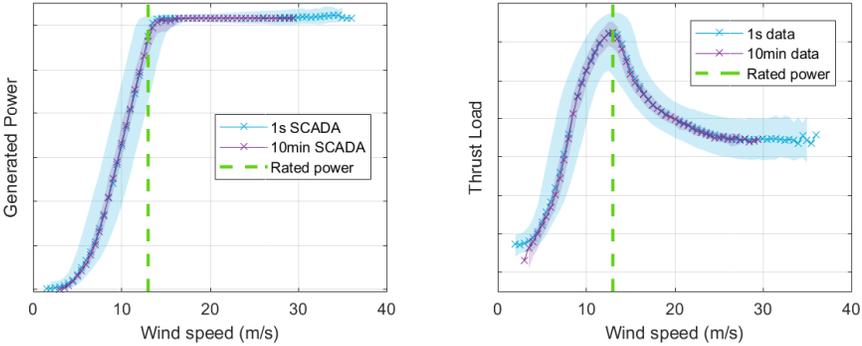
additional load.

The resulting signal is then transformed into thrust load  $F_{T,m}$ , using the distance between the sensors (= location of the measured bending moment) and the hub (= location of acting thrust force) [63]. To match the time-steps of the SCADA data, the obtained thrust load is down-sampled using an anti-aliasing filter to a time frame of 1 second and additionally averaged over 10 minutes. As the turbine is installed on a jacket, the role of wave loading in the quasi static bending moment is assumed to be negligible.

### 4.2.2 SCADA data

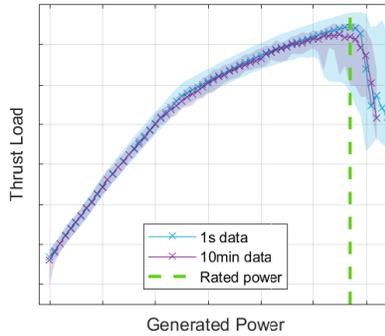
For this research, SCADA data is utilized, both the 10min averages as the 1s data, i.e. data sampled at  $1Hz$ . In this case, the preprocessing and filtering process for the SCADA data consisted in exclusion of improbable and unrealistic values for wind speed (outside interval  $[0; 50]m/s$ ) and for generated power (outside interval  $[-0, 1; 1, 25] \cdot P_{rated}$ ) and periods of constant wind speed from the dataset. In total, only 0,02% of the total dataset is removed.

For this research a subset of one year and three months of both 10min statistics and  $1Hz$  signals of SCADA data was available. The subset consisted in both cases of measurements for wind speed, rotor speed, generated power, blade pitch angle and yaw angle. The subset of 10min statistics additionally contained measurements of wind direction and ambient (air) temperature.



(a) Produced power vs wind speed

(b) Measured thrust load vs wind speed



(c) Measured thrust load vs produced power

Figure 4.2: Characteristic curves for power output and load obtained using SCADA data in combination with averages of thrust load measurements. Operational data for a period of 2,5 months is shown. Data in both 10 minute time frame (purple) and 1 second time frame (blue) is shown. The line indicates the median value, calculated per bin of  $0,5m/s$  or  $100kW$ , whereas the surface spans the 5th to 95th percentile of the data. Rated power is reached for wind speeds of approximately  $13m/s$  (indicated by the green dashed line).

Figure 4.2a shows the power curves obtained with both the 1s and 10min SCADA, respectively in blue and purple, for a period of 2,5 months. The lines indicate the median value of the dataset, while the surface spans the 5th to 95th percentile of the data. The median values of both datasets usually coincide quite nicely. A higher difference in spread of the data however is observed, since the blue surface is much bigger than the purple surface. The power curve generated using the 1s SCADA thus shows a much larger variability in power compared to the 10min averaged SCADA. The same difference in variability can be observed in Figure 4.2b and 4.2c, where 1s and 10min averages of measured thrust load are plotted versus 1s and 10min SCADA wind speed and generated power respectively.

The present variability in 1s data is not only the result of noise, but is mainly due to the inertias within the controlling system and the wind turbine. For example, when the wind speed increases, the power output increases only a few seconds after (Figure 4.3). These inertias result in time delays up to several seconds between e.g. the wind speed and the generated power. These delays are not considered constant over time and will differ for every SCADA parameter. Moreover, they last for only a couple of seconds and in consequence they can not be observed within 10 minute averages. Although the variability obtained using 10min data is far less compared to 1s data, it is still existent. This variability in the data was the reason for which better results were obtained for multivariate power curve modeling techniques compared to univariate modeling techniques in Section 3.3.2.

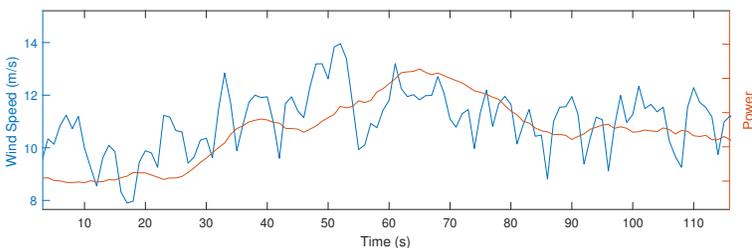


Figure 4.3: Wind speed signal and power signal as captured by the SCADA system. An inertia of the power signal with respect to the wind speed signal is visible, as power output only peaks at around 65s while the wind speed had peaked about 15s earlier.

### 4.2.3 Meteorological data

Measurements of air pressure are available from a nearby met mast (15km). Using the ambient temperature, contained in the considered SCADA dataset, the air density is calculated using Equation 3.2.

## 4.3 Input parameter selection

### 4.3.1 SCADA data

According to [29], it is possible to calculate the thrust load  $F_T$  with Equation 4.1, where  $\rho$  is the air density,  $V$  the wind speed,  $R_{rotor}$  the rotor radius,  $c_T$  the thrust coefficient,  $\theta$  the blade pitch angle and  $\Omega$  the rotor speed.

$$F_T = \frac{1}{2} \rho V^2 \pi R_{rotor}^2 c_T(\theta, V, \Omega) \quad (4.1)$$

In this equation, all parameters are measured by the SCADA system except for the air density and the thrust coefficient  $c_T$ . The influence of air density is commented on later in Section 4.3.2. The thrust coefficient depends on the design of the turbine

and is considered to be dependent on the blade pitch angle and the tip speed ratio [29]. Since the tip speed ratio can be calculated using the wind speed and the rotor speed, the thrust coefficient should be dependent on SCADA parameters only: blade pitch angle, wind speed and rotor speed. Hence a model to estimate the thrust load can be created based on SCADA data. In this chapter, a data-driven model will be created, which will approximate the model given by Equation (4.1). Once such a model is found, it should be valid for every turbine of the same type and hopefully can be transferred to all wind turbines in the farm.

Figure 4.4 shows a couple of time series of measured thrust loads together with the same 1s time series of a SCADA parameter: wind speed (4.4a), generated power (4.4b) and blade pitch angle, the additive inversed, (4.4c). These time series do not represent the same operational conditions, but are chosen in such a way the different dependencies of thrust load can be visualized. A clear match between the signals can be seen. This intuitively confirms the idea to model the thrust load based on SCADA data.

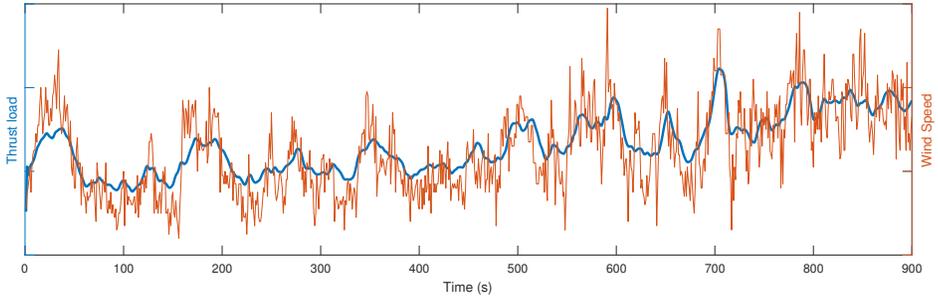
A crucial part in the model creation is the parameter selection. Input parameters are chosen from the SCADA dataset based on their Pearson correlation and mutual information to the thrust load.

### Pearson correlation

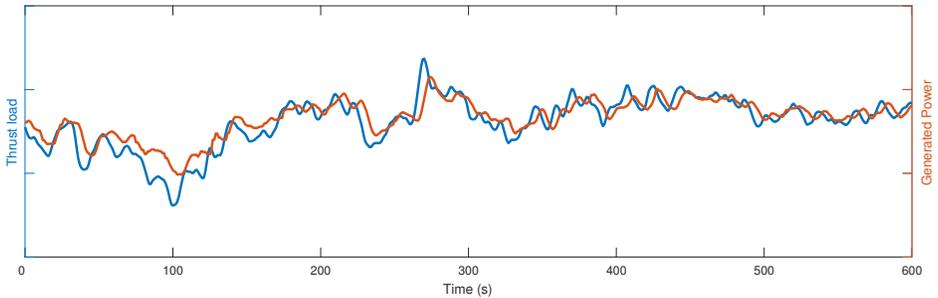
A common practice to select input variables is by calculating the Pearson correlation  $R(X, Y)$  between the possible input parameters and the given output parameter. The Pearson correlation between two signals  $X$  and  $Y$ , both with  $n$  time instances, is calculated using Equation 4.2, in which  $\bar{X}$  is the mean value of the signal  $X$  and similar for  $\bar{Y}$  [64].

$$R(X, Y) = \frac{\sum_{i=1}^n (X(i) - \bar{X}) (Y(i) - \bar{Y})}{\sqrt{\sum_{i=1}^n (X(i) - \bar{X})^2 \sum_{i=1}^n (Y(i) - \bar{Y})^2}} \quad (4.2)$$

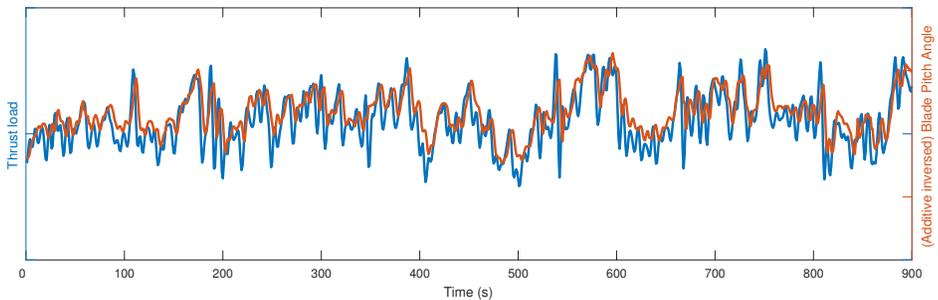
The calculation of Pearson correlation is done for data during different operating states. Both 1s data and 10min data for a period of 2,5 months is used. A first dataset contained all data points for which the turbine was generating. During this period the full wind speed range is covered, as shown by Figure 4.2. Additionally, the generating data sets are divided into two subsets: data when the turbine was generating below rated power (64 % of 10min and 62 % of 1s operational data) and at rated power (36 % of 10min and 38 % of 1s operational data). The distinction between below and at rated power is made based on the value for power (below or over 95% of rated power as given by the manufacturer). Moreover, the correlation is calculated for a fourth dataset, where the turbine was not generating.



(a) Thrust load time series compared to 1s SCADA wind speed time series



(b) Thrust load time series compared to 1s SCADA power time series



(c) Thrust load time series compared to 1s SCADA pitch angle time series (additive inverted)

Figure 4.4: A good match is found between time series of measured thrust loads and time series of SCADA parameters (1s). The shown time series all represent different time instances and do not necessarily have the same lengths.

Results for all different datasets and all parameters of interest from the SCADA are given in Figure 4.5. For the wind direction, no 1s dataset is available. Therefore only results for 10min data are shown.

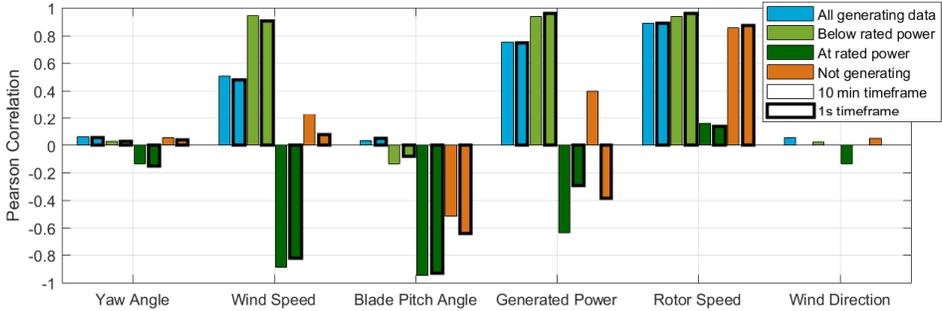


Figure 4.5: Pearson Correlation between measured thrust load and several SCADA, calculated for 10min and 1s dataset.

**Behavior during generating** Focusing first on the results for the total generating data set, a high Pearson correlation can be found for rotor speed (0,8871 and 0,8935 for 10min and 1s data respectively), generated power (0,7513 and 0,7497) and to a lesser extent wind speed (0,5082 and 0,4829).

In case of the generating data below rated power, the resulting values are even higher for rotor speed (0,9383 and 0,9622), power (0,9396 and 0,9603) and wind speed (0,9417 and 0,9086).

In contrast, generating data at rated power reveals a high Pearson correlation of the blade pitch angle (0,9499 and 0,9298) and wind speed (0,8898 and 0,8194). This difference in behavior is explained as follows. Once the turbine reached its rated power value, the only parameter acting to varying wind speed and thrust load will be the blade pitch angle. Hence a significant lower correlation for the rotor speed (0,1562 and 0,1385) is found. However generated power is still correlated to thrust load with a significant value (0,6354 and 0,2915). Figure 4.2c reveals a very steep curve between thrust and generated power once rated power is reached.

In the results of Pearson correlation (Figure 4.5) negative values are the result of an additive inversed relationship between the depicted parameter and the thrust load. For a turbine generating below rated power, a higher wind speed results in a slightly lower blade pitch angle and an increased thrust load. Therefore a decreasing blade pitch angle (due to an increase in wind speed) leads to a higher thrust load. Hence, a negative value for Pearson correlation between pitch angle and thrust load when the turbine is generating below rated power is expected. Once rated power is reached, increasing wind speeds result in higher blade pitch angles, slightly increasing generated power and decreasing thrust loads (Figure 4.2b). Accordingly an increase in blade pitch angle and generated power (thanks to an increase in wind speed) enforces a decrease in thrust load. And thus, a resulting

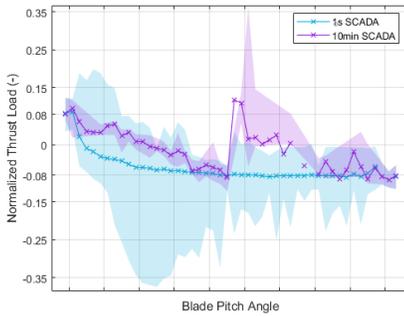
negative Pearson correlation between thrust load and wind speed, generated power and pitch angle for generating data at rated power is consistent.

It is obvious the turbine reacts differently to varying wind speeds depending on the operating state. Once rated power is reached, the relation of the thrust load to the depicted SCADA parameters often differs. This leads to lower correlation values for the total generating dataset in comparison to the generating states separately. In case of the pitch angle, the correlation even is non-existing when looking at the total operational dataset.

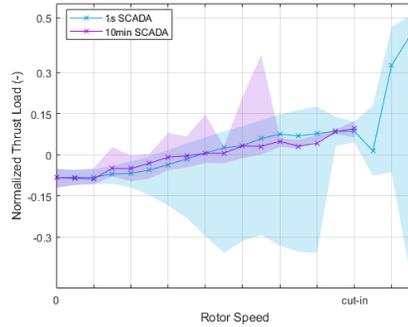
**Behavior in non generating state** When looking at the results for the non-generating dataset, a high correlation is found for the rotor speed (0,8534 and 0,8773 for 10min data and 1s data respectively). This can be explained using Figure 4.6b, where one can see a clear (linear) correlation between thrust load and low rotor speeds up to 8rpm. Striking is the existence of data points for which the rotor speed is not equal to zero, although the data set was filtered to contain only data for which the turbine was not generating. As long as the turbine is slowing down or speeding up from/to 8 rpm for any reason (e.g. rotor stop, too low wind, cut-out) the turbine is not producing power. It starts producing power once the rotor speed reached 8 rpm. This change in rotor speed is actually triggered by a changing blade pitch angle, as shown in Figure 4.6d. This also explains the significant correlation between the measured thrust load and the blade pitch angle (0,52 and 0,6399 for 10min and 1s data respectively), as can be seen in Figure 4.6a too. It can be seen that, the more the turbine is slowing down, the greater the pitch angle becomes, therefore the lower the thrust load is.

The moment the turbine stops producing energy, but is still rotating, it consumes the highest amount of energy (Figure 4.6e) and thus the most negative values for generated power are measured. The more the turbine slows down, the lower the thrust load is and the less energy it consumes (thus the generated power increases but stays negative). This relation between thrust load and generated power can be seen in Figure 4.6c and in the significant value for Pearson correlation (0,3953 and 0,3827 for 10min and 1s data respectively) as well.

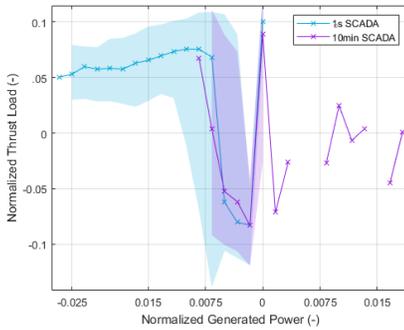
Results in Figure 4.6 show that for pitch angles greater than 40 degrees, the 10 minute data actually contains the transition from non-generating to generating or vice versa. This means that the turbine shut down or started during the 10 minutes of which the averages values are considered as not operational. Therefore, the correlation results based on 10min averages are less reliable, for these particular points.



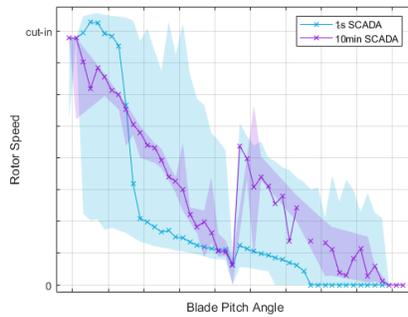
(a) Measured thrust load vs blade pitch angle



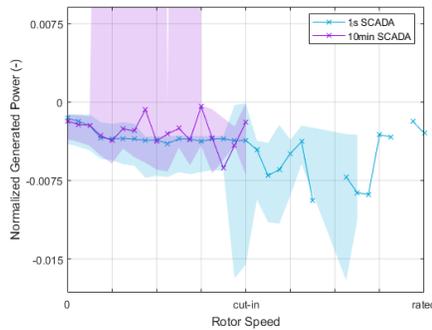
(b) Measured thrust load vs rotor speed



(c) Measured thrust load vs generated power



(d) Rotor speed vs blade pitch angle



(e) Generated power vs rotor speed

Figure 4.6: Characteristic curves obtained using SCADA data in combination with averages of thrust load measurements. Only **non-generating data**, i.e. without power production, during a period of 2,5 months is shown. Data in both 10 minute time frame (blue) and 1 second time frame (purple) is shown. The line indicates the median value, calculated per bin, whereas the surface spans from the 5th to the 95th percentile of the data.

### Mutual information

The Pearson correlation studied in the previous section evaluates the existence of a linear relationship between two parameters. However, the problem we are facing is not necessarily linear. As such an analysis to identify and quantify possible chaotic or non-linear dependence is performed as well. A relevant metric is mutual information, a measure of dependence based on information theory and the notion of entropy. The mutual information  $I(X; Y)$  between two signals  $X$  and  $Y$  is determined by equation 4.3[65], or discretized by equation 4.4[66].

$$I(X; Y) = \int \int \mu_{X,Y}(x, y) \log \frac{\mu_{X,Y}(x, y)}{\mu_X(x)\mu_Y(y)} dx dy \quad (4.3)$$

$$I(X; Y) = \sum_{x \in X} \sum_{y \in Y} h(x, y) \log \frac{h(x, y)}{f(x)g(y)} \quad (4.4)$$

To calculate the mutual information, the continuous probability density functions (pdf)  $\mu_X(x)$ ,  $\mu_Y(y)$  and  $\mu_{X,Y}(x, y)$  or the discretized functions  $f(x)$ ,  $g(y)$  and  $h(x, y)$  are used. To obtain these pdf estimations, both kernel or histogram based estimations are commonly used. Unfortunately, the kernel based calculation of mutual information for continuous variables requires too much computational effort for the considered amounts of data. Therefore only the histogram based estimation for discretized functions is implemented.

To obtain the discretized pdf estimations, a histogram based estimation as explained by [67], is implemented. To estimate the joint density  $h(x, y)$ , a bi-dimensional histogram  $\hat{h}_{k,j}$  ( $1 \leq k \leq m, 1 \leq j \leq m$ ) is created. The idea is to build a set of rectangular tiles spanning the area created by the ranges of both signals. After which the number of pairs  $(x, y)$  falling into each tile (Equation 4.5) are counted.

$$\hat{h}_{k,j} = \#\{(x, y) \mid a_k \leq x \leq a_{k+1} \text{ and } b_j \leq y \leq b_{j+1}\} \quad (4.5)$$

The size of the tiles is defined by using Equation 4.6, with the interquartile range of the data IQR and the number of data points  $N$ , as suggested by [68].

$$TileSize = 2 \cdot IQR \cdot N^{-\frac{1}{3}} \quad (4.6)$$

The resulting histogram  $\hat{h}_{k,j}$  can be used to estimate  $h(x, y)$ . Moreover  $f(x)$  and  $g(y)$  can be estimated from the bi-dimensional histogram as well:

$$\hat{f}_k = \sum_j \hat{h}_{k,j} \quad (4.7)$$

$$\hat{g}_j = \sum_k \hat{h}_{k,j} \quad (4.8)$$

Eventually, the mutual information can be calculated from:

$$\hat{I}(X;Y) = \sum_{k,j} \hat{h}_{k,j} \cdot \log \frac{\hat{h}_{k,j}}{\hat{f}_k \hat{g}_j} \quad (4.9)$$

Opposed to Pearson correlation coefficients, mutual information doesn't have a general maximum value indicating perfect dependence between two signals. Therefore the resulting mutual information should be normalized first. Multiple possibilities are found in literature. Four of them are applied and compared here. For these normalization, often the entropy of a signal is used. Therefore, Equation 4.10 defines entropy  $H(X)$  for a signal  $X$  and Equation 4.11 gives the joint entropy  $H(X, Y)$  for two signals  $X$  and  $Y$ .

$$H(X) = - \sum_{x \in X} f(x) \log f(x) \quad (4.10)$$

$$H(X, Y) = - \sum_{x \in X} \sum_{y \in Y} h(x, y) \log h(x, y) \quad (4.11)$$

The first one utilizes the entropy of the signals separately (Equation 4.12) [69].

$$I_{normalized}(X;Y) = \frac{2I(X, Y)}{H(X) + H(Y)} \quad (4.12)$$

For the second one, normalization is done by dividing by the joint entropy  $H(X, Y)$  of the two signals [70], as indicated by Equation 4.13.

$$I_{normalized}(X;Y) = \frac{I(X, Y)}{H(X, Y)} \quad (4.13)$$

A third method to normalize takes the maximum of the entropies of both signals as a reference (Equation 4.14)[71].

$$I_{normalized}(X;Y) = \frac{I(X, Y)}{\max(H(X), H(Y))} \quad (4.14)$$

A last suggestion for normalization doesn't need the mutual information calculation at all, as seen in Equation 4.15[72].

$$I_{normalized}(X;Y) = \frac{H(X) + H(Y)}{H(X, Y)} \quad (4.15)$$

Mutual information is calculated for the same period and datasets as Pearson correlation. The resulting mutual information between the measured thrust load and several SCADA parameters for all 10 minute data sets is depicted in Figure 4.7 for all normalization methods. Similar results are found for all normalization methods, although for the fourth (Equation 4.15) the reference value for no correlation is 1 instead of 0 as is the case for the other normalization techniques.

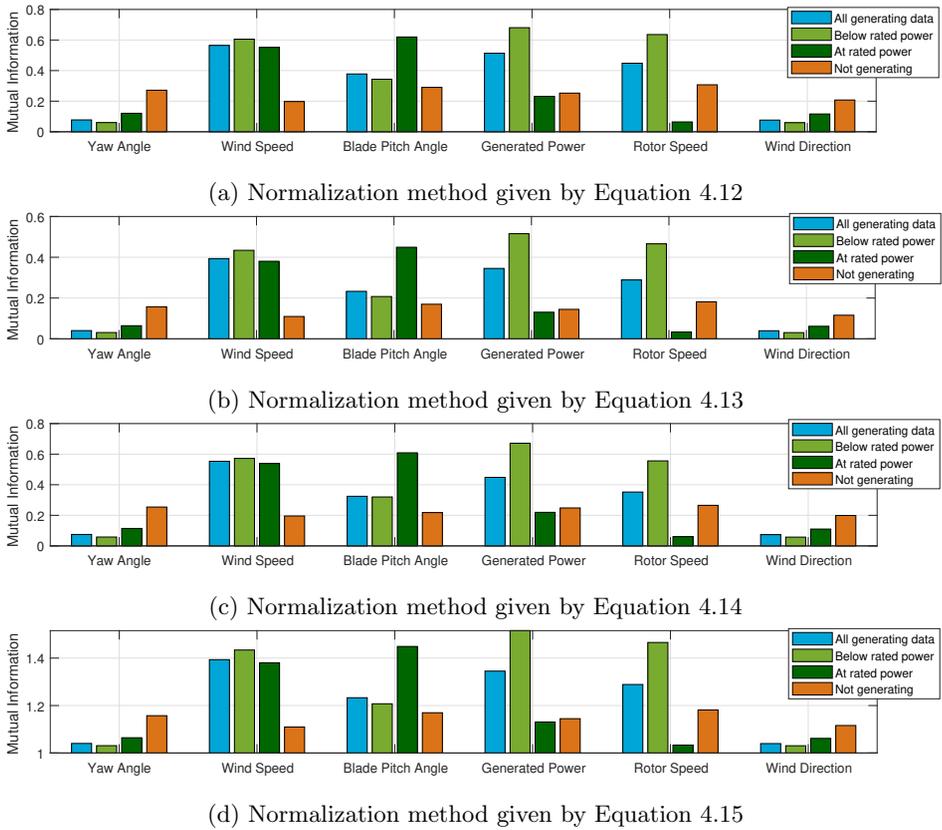


Figure 4.7: The mutual information is calculated between thrust load averages and several SCADA parameters for four different datasets. Four different normalization methods are compared based on results for 10min averages only.

Because little difference is found among these normalization methods only one normalization method is applied on the 1s dataset. Only the normalization given by Equation 4.13 was performed. The results for both 1s and 10min data is given in Figure 4.8.

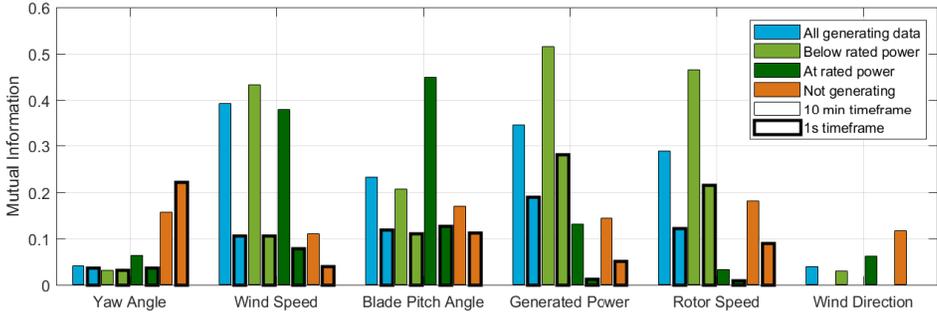


Figure 4.8: Mutual information between measured thrust load and several SCADA, calculated for 10min and 1s dataset.

**Behavior in generating conditions** When looking at the results of mutual information for the total generating data sets, the highest dependencies are again found for wind speed (0,3934 and 0,1068 for 10min and 1s data respectively), generated power (0,3453 and 0,1904) and rotor speed (0,2888 and 0,1217). Interestingly, also a relatively high value is found for the blade pitch angle (0,233 and 0,1186). When taking into account the existing non linearities, as done by mutual information, the blade pitch angle is clearly correlated with the thrust load based on the total generating dataset as well.

In case of the generating data below rated power, the values for mutual information are even higher for power (0,5158 and 0,2811), rotor speed (0,466 and 0,2161) and wind speed (0,4342 and 0,1057). Here, the values for mutual information of the pitch angle decreased to 0,2075 and 0,1103 for the 10min and 1s dataset respectively.

On the contrary, generating data at rated power reveals a high value for mutual information for the blade pitch angle (0,4492 and 0,1268) and for the wind speed (0,3804 and 0,07928). Again, this is because of the pitch action of the controller once rated power is reached. Also, a significant lower value for mutual information for the rotor speed (0,0333 and 0,0098) is found. Such as seen for the Pearson correlation, the value for mutual information of generated power in case of 10min data is significantly higher as well (0,1308).

**Behavior in non-generating conditions** Looking at the non-generating dataset, values for mutual information are significantly lower than for the other datasets. Again, rotor speed (0,1814 and 0,08987 for 10min and 1s data respectively), pitch angle (0,1697 and 0,1119) and generated power (0,1447 and 0,05072) can be considered as influencing parameters. In contrast to the results of Pearson correlation, the yaw angle has a high value for mutual information (0,1573 and 0,2219).

Figure 4.9 shows a clear non-linear relation between the median value of measured thrust and the measured yaw angle. This can be explained by the transformation of strain measurements to fore-aft and side-side bending moments. When the transformation done here is not according to the actual wind direction, a depen-

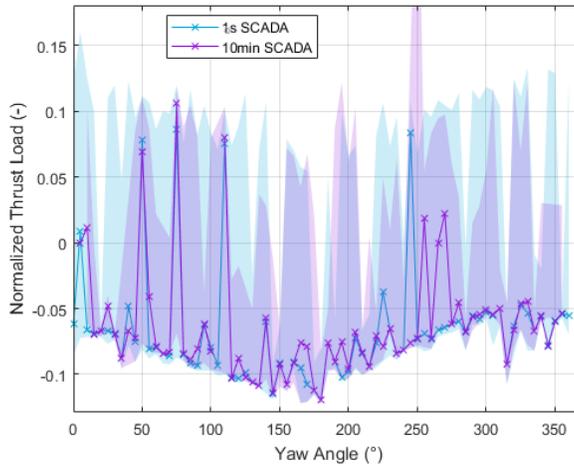


Figure 4.9: Measured thrust load vs SCADA yaw angle for non-generating 10min and 1s data

dency on the yaw angle can be found. When the turbine is not operating, it is often not rotated in the actual wind direction. The transformation is based on the measured yaw angle. Therefore the transformation can be inaccurate in case the turbine is not operating.

For the continuation of this research, all SCADA parameters except for the yaw angle will be considered as an input parameter. The yaw angle is not selected due to its small correlation for the total generating dataset (0,0598 and 0,0596 for 10min and 1s data respectively, Figure 4.5) and mutual information (0,0403 and 0,0367 for 10min and 1s dataset respectively, Figure 4.8) with the thrust load.

### Analysis of time lags between 1s SCADA parameters

For all four considered subsets in Figure 4.8 the values for mutual information are less for 1s SCADA over the results for the 10 minutes averages. This can be explained by the time delays of several seconds between different parameters, as a result of the inertias present within the system. In this section an additional analysis is done to see the effect of including a time lag between signals.

A time lag between the different SCADA signals and the thrust load could already be observed in Figure 4.4. This time lag is not constant over time and differs for every SCADA parameter, this behavior excludes a fixed time delay between the load-monitoring system and the SCADA system. A more likely cause for the observed time delays are the inertias within the controlling system and the wind turbine. The existence of these inertias is confirmed in [60]. Due to these inertias, the Pearson correlation and mutual information based on the 1s SCADA data can be recalculated, including different values for a time lag.

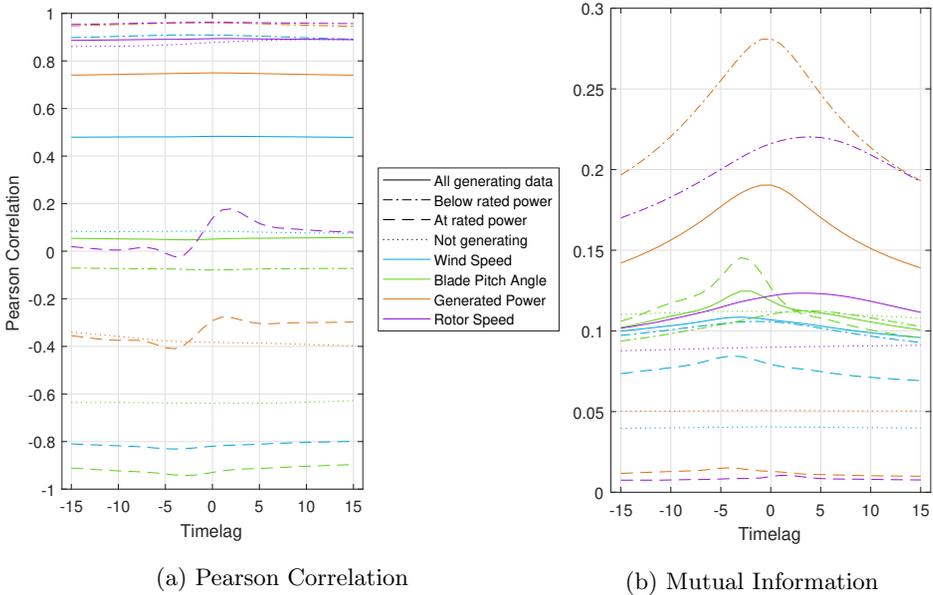


Figure 4.10: Correlation analyses for different SCADA signals and different datasets (based on operational status of the turbine) including different time lags

Figure 4.10 shows the resulting values for Pearson Correlation and mutual information for time lags from -15 to 15 seconds. Negative time lags indicate that the thrust signal is lagging behind the SCADA signal, positive time lags the other way around. In general, a bigger influence of a possible time lag is found for mutual information than for Pearson correlation. In Figure 4.10a very little difference is found for almost all signals and all different datasets, except for the rotor speed and generated power at rated power. Here, the SCADA signals are not dominated by big, very low frequent oscillations. Therefore the correlation is based on the existing small but relatively high frequent oscillations with a period of a couple of seconds and thus more influenced by the introduction of a time lag.

Table 4.1 shows the time lag for which the correlation or mutual information was highest for every parameter and every dataset. When a value is missing, no maximum was found. In general, the thrust signal is not lagging more than 5 seconds behind any SCADA parameter. These results indicate the thrust load a turbine is subjected to at a certain timestamp can be induced by a SCADA parameter as measured a few seconds (up to 5) ago. Therefore it is concluded that if the thrust load is estimated based on 1s data, not only the instantaneous values for all SCADA parameters will be considered but also those up to 5s before.

Table 4.1: Ideal time lag, in seconds, found for different SCADA signals and different datasets

(a) Pearson Correlation

	$V$	$\theta$	$P$	$\Omega$
All generating	1	/	0	1
Below rated power	-2	0	-1	0
At rated power	-4	-3	-4	2
Not generating	-1	1	/	12

(b) Mutual Information

	$V$	$\theta$	$P$	$\Omega$
All generating	-3	-3	0	3
Below rated power	-1	4	0	4
At rated power	-4	-3	-4	1
Not generating	0	-3	0	/

### 4.3.2 Meteorological data

According to [29], thrust loads are greatly influenced by air density. This comes to no surprise as results in Section 3.2.2 already introduced a density-correct wind speed to account for the strong relation between air density and the power production. However, while changes in the depicted SCADA variables happen within seconds, air density changes on a different time scale of several hours. Because of this *slow* behavior including air density into the set of input parameters for training the model implies that long-term data is needed to train the 1s SCADA model. Instead the effect of air density is accounted for as a correction of the modeled thrust load  $\hat{F}_T$ . Equation 4.1 and [29] shows a linear correlation between air density and the thrust load. Therefore Equation 4.16 is proposed to perform an air density correction on the modeled thrust load.

$$\hat{F}_{T,corr} = \rho \hat{F}_T \quad (4.16)$$

## 4.4 Modeling technique

Seeing the relation between thrust load and the depicted SCADA parameters is non-linear, a model will be created using a neural network and the Neural Network toolbox of MATLAB. A neural network is capable of finding and characterizing non-linear dependencies within datasets. Therefore it can handle the inverted relations between thrust load and the considered SCADA parameters once rated power is reached. The neural network used in this section has 3 hidden layers with 4 neurons each, as shown in Figure 4.11.

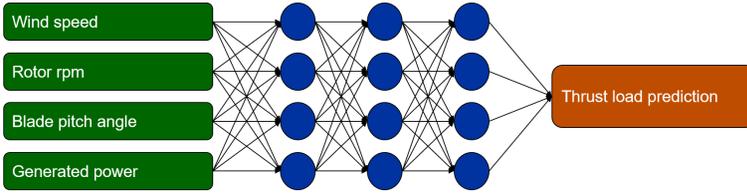


Figure 4.11: Schematic visualization of chosen neural network topography.

The network is trained using all data remaining after the preprocessing and filtering process explained in Section 4.2.2. The training data consisted of 1s or 10min SCADA data and the corresponding 1s or 10min averages of thrust load measurements ( $F_{T,m,training}$ ). As explained in Section 4.3.2, the effect of air density is accounted for by applying a correction on the model results. To make sure the effect of air density is not present in the training data, the inversed correction is applied on the measured thrust loads of the training dataset:  $F_{T,m,training} = \frac{F_{T,m}}{\rho}$

From the conclusions of Section 4.3 the following input parameters were chosen for the model: wind speed, blade pitch angle, rotor speed and generated power.

The modeling technique will be performed using 10min averages and 1s data. To account for the inertias in the system visible in 1s data, not only instantaneous SCADA values, but also the values of 5 previous seconds are given as input for a model trained based on 1s data.

Since the model will be used in every operating state of the wind turbine, it is important the full operational wind speed band is covered in the training data set. For every operating state, e.g. during a down-rating or curtailment, that is not represented in the training data, the model will probably not be able to predict the thrust load correctly.

To train the neural network, the Neural Network toolbox of MATLAB is used with the default settings [73]. This means the preprocessing is done by a min-max mapping function, tan-sigmoid transfer functions are used for hidden layers and a linear transfer function is used for the output layer. Furthermore the data chosen to train the model is randomly divided into 70 % of training data, 15 % of validation data and 15 % of test data. This so-called hold-out method is preferred over cross-validation to reduce the computational load, since large datasets are used. Training is done using the training data and the Levenberg-Marquardt algorithm. Training of the network is stopped when the error on the validation data failed to decrease for 6 iterations or a maximum number of 1000 iterations is reached. The test data is used as an independent dataset of the network training, to calculate the final model error.

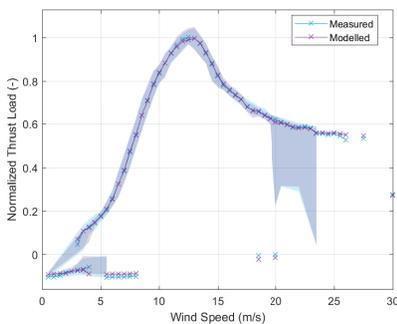
For all models created and validated in the remaining of this section, the measured thrust load  $F_{T,m}$  is compared to the modeled thrust load  $\hat{F}_T$ . This is done in absolute terms, but also by using the relative error between both, obtained by Equation 4.17.

$$\Delta\epsilon = \frac{abs(F_{T,m} - \hat{F}_T)}{F_{T,m}} \quad (4.17)$$

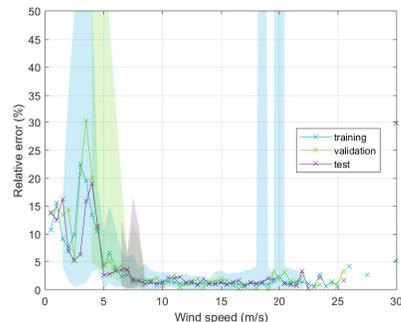
### 4.4.1 Model based on 10min data

Firstly, the thrust load is modeled using the 10 minute average SCADA data and 10 minute average load measurements. The 10min averages could be preferred over 1s data because less data is needed to cover the entire wind speed and a wide meteorological range. This reduced amount of data inherently reduces the time to train the model significantly.

The training of the model is done based on a dataset of 2 weeks. This training dataset was chosen carefully to make sure the entire wind speed range was covered, as is shown in Figure 4.12a. Moreover, it was chosen to keep the training period as short as possible for multiple reasons. First of all, the computational time for a shorter period is significantly lower. Furthermore, the shorter the needed training period, the faster the application of the model can be used for future applications. However it is very important the full wind speed range is covered during the training period. This means for future applications, that might be the main driver to define the training period.



(a) Modeled and measured thrust loads for the total dataset of 2 weeks during the training phase, including training, validation and test set



(b) The absolute relative error for the different datasets during the training phase

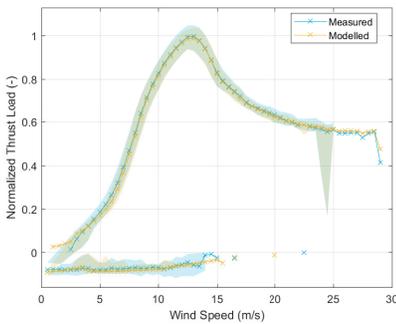
Figure 4.12: Results based on 10 minute averages of SCADA parameters and measured strains for the dataset used during the training phase. The lines indicate median values for every wind speed bin of  $0, 5\text{ m/s}$ ; the surface spans from the 5th to the 95th percentile of the data. For the thrust curve, two datasets are considered for each wind speed bin. One for high thrust loads and one for low thrust load (the limit is ca 5% of the maximum thrust).

Figure 4.12 shows the results of the training phase. Figure 4.12a shows the measured and modeled thrust load for the 10 minute averages. A good agreement is found over the entire wind speed range, for both operational and not operational data. Figure 4.12b shows the relative absolute error for the training, validation and test set during the training phase. Median errors of over 20% are found for low wind speed, mainly caused by the low absolute value of the measured thrust load. During operating conditions for wind speeds higher than 7 m/s, the relative error

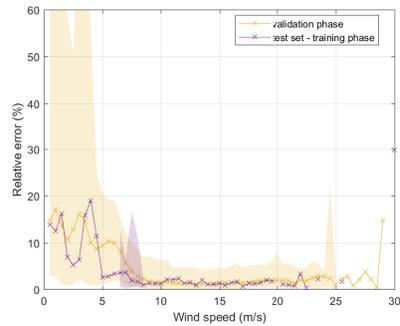
stays under 5%. A similar behavior for the three datasets is found under operating conditions, indicating the training set was representative for the validation and test set.

### Applied on 10min data

This found model is extensively validated on a dataset containing one year of 10 minute averages. The training data of 2 weeks is not included in this validation data. Figure 4.13 shows the results of this validation. Figure 4.13a shows a good agreement between measured and modeled thrust load over the entire wind speed range. Figure 4.13b shows a similar relative error under operating conditions for wind speeds higher than  $8\text{m/s}$ . For low wind speeds, relative errors are higher and more inconsistent due to the low value of the measured thrust load. Less importance is given to the low values of thrust load, because other loads acting on the structure are dominant in that case.



(a) Modeled and measured thrust loads for the total validation dataset of one year

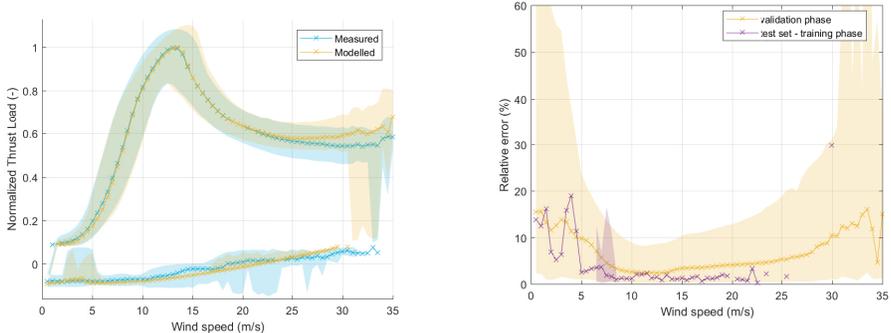


(b) The absolute relative error for the validation dataset of one year and the test set during the training phase

Figure 4.13: Results based on 10 minute averages of SCADA parameters and measured strains for the validation set of one year. The lines indicate median values for every wind speed bin of  $0,5\text{m/s}$ ; the surface spans from the 5th to the 95th percentile of the data. For the thrust curve, two datasets are considered for each wind speed bin. One for high thrust loads and one for low thrust load (the limit is ca 5% of the maximum thrust).

### Applied on 1s data

To cover all relevant fatigue cycles over the lifetime of the wind turbine the load modeling will have to be performed at a higher rate than once every ten minutes. At this point it is still uncertain whether the ten-minute model can be combined with the 1s SCADA to obtain a sufficiently accurate model of the thrust loading. Therefore 1s data is used as an input for the model that was trained with 10 minute



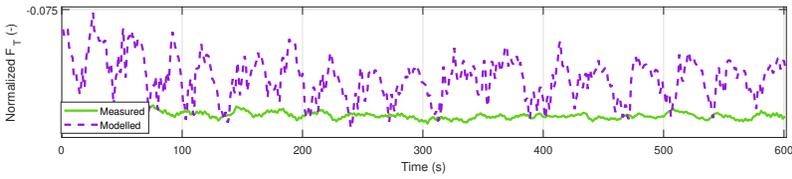
(a) Modeled and measured thrust loads for the total validation dataset of one year

(b) The absolute relative error for the validation dataset of one year and the test set during the training phase

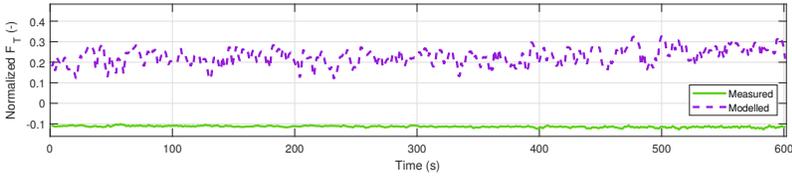
Figure 4.14: Results based on using 1s data of SCADA parameters as the model-input and measured strains for one year. The model was trained based on 10 minute averages of SCADA data and measured strains of 2 weeks. The lines indicate median values for every wind speed bin of  $0,5m/s$ ; the surface spans from the 5th to the 95th percentile of the data. For the thrust curve, two datasets are considered for each wind speed bin. One for high thrust loads and one for low thrust load (the limit is ca 5% of the maximum thrust).

averages of data. Again, a validation period of one year is chosen. Figure 4.14 shows the results.

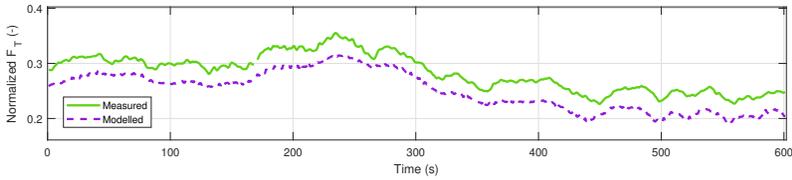
A good agreement in the general behavior between modeled and measured thrust load is found for operating conditions up to wind speeds of  $20m/s$  (Figure 4.14a). Also for parked conditions a fairly good agreement is found. However close to cut-out wind speed (around  $25m/s$ ), an increasing difference between measured and modeled thrust load is observed for operating conditions. This is also represented by the relative error (Figure 4.14b). The median error and the surface clearly increase for increasing wind speeds from  $15m/s$ . Moreover, higher errors are again observed for very low wind speeds, due to the low absolute value of thrust loads. Finally, one can see that the surface spanning the 5th to the 95th percentile of the measured data is bigger in general compared to the results of 10min data (Figure 4.13b). This can be easily explained by the difference in variability in both datasets, as was already explained in Section 4.2.2 and by Figure 4.2.



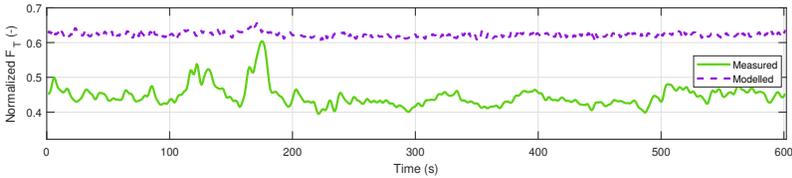
(a) Not generating. MRE = 5,36%



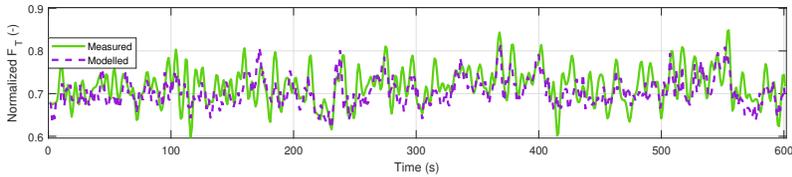
(b) Not generating. MRE = 268%



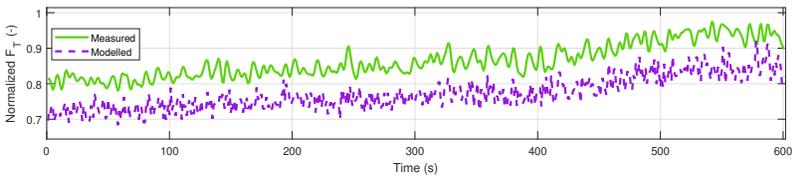
(c) Below rated power. MRE = 11,7%



(d) Below rated power. MRE = 38,3%



(e) At rated power. MRE = 4,23%



(f) At rated power. MRE = 10,1%

Figure 4.15: Time series (spanning 10 minutes) of modeled and measured thrust loads for not generating (a,b), below rated power (c,d) and at rated power (e,f). For each operational state, the time series with the highest averaged absolute error is shown (b,d,f). MRE shows the averaged absolute relative error over 10 minutes.

Six time series spanning 10 minutes during the validation phase are shown in Figure 4.15, two when the turbine is not generating (Figures 4.15a and 4.15b), two while operating below rated power (Figures 4.15c and 4.15d) and two while operating at rated power (Figures 4.15e and 4.15f). The time series of 10 minutes with the highest averaged absolute error between measured and modeled thrust loads while not generating, operating below rated power and operating at rated power are shown in Figure 4.15b, Figure 4.15d and Figure 4.15f respectively. Almost all time series show an offset, being a significant difference between the mean value of both signals. The modeled signal in Figures 4.15a, 4.15b and 4.15f shows more variations than the measured signal. These are probably caused by variations in wind speed. This indicates the model is too influenced by the variations in SCADA signals such as wind speed. On the other hand, the model clearly missed some variations in Figure 4.15d. It was observed the variation in measured thrust load was mainly induced by a varying pitch angle. Thus, one could conclude the model is too much influenced by variations in wind speed and too little by variations in pitch angle. This over- or under-sensitivity to specific SCADA parameters can be caused by the use of 10 minute averages to train and 1s data as input to the model. For example, little variations in pitch angle can cause little variations in thrust load. However, when looking at the 10 minute averages of the pitch angle, these little variations are non-existent.

Finally, in case of rated power conditions, the model is not always capable of fully capturing the amplitudes of all cycles, as can be seen in Figure 4.15e.

For these reasons it is not recommended to train a thrust load model, intended for fatigue life estimation, based on 10 minute averages. An alternative is to train the model with 1s data instead of 10 minute averages.

#### 4.4.2 Model based on 1s data

The modeling method proposed in Section 4.4 is also validated when trained using 1s data. For this purpose two different datasets of 1s data are considered. The first data is obtained by simulation in FAST in order to obtain a fully controlled and reproducible dataset. The second dataset is obtained from a measurement campaign performed at an offshore wind turbine (as explained in Section 4.2).

##### Simulated data

The simulated data is obtained by using the software FAST v8 [74], offered by NREL. The chosen simulated turbine is the NREL 5,0 MW Baseline Wind Turbine, installed on an OC3 Monopile RF configuration. All simulation specifications are kept as proposed by the software [74] for use of this turbine type. This means that turbulence and irregular waves are also accounted for. To make sure the full wind speed range is sufficiently covered in the simulation data, several input wind files with varying average wind speed between  $3m/s$  and  $25m/s$  are generated using TurbSim. Each wind speed is accounted for equally. In essence, the wind speed distribution is thus considered as uniform.

The output parameters of interest for this research are specified in Table 4.2.

As the results obtained using simulated data will serve to validate a real-world methodology, only parameters for the SCADA data and the measured bending moment whom are available in the real-world are considered.

During these simulations the air density was kept constant. Therefore the applied corrections for air density did not influence the results.

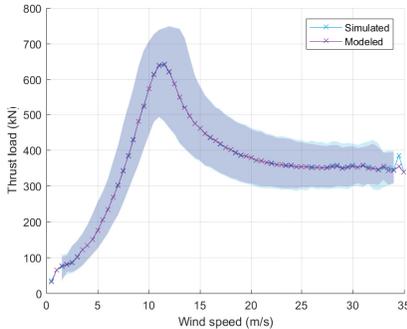
Table 4.2: The selected output variables for FAST simulations

Parameter	Category	Description	Unit
Wind1VelX	InflowWind	Nominally downwind component of the hub-height wind velocity	m/s
BldPitch1	ElastoDyn - Blade Pitch Motions	Blade pitch angle (position)	deg
LSSGagVxa	ElastoDyn - Shaft Motions	Low-speed shaft strain gage angular speed (on the gearbox side of the low-speed shaft)	rpm
YawPzn	ElastoDyn - Nacelle Yaw Motions	Nacelle yaw angle (position)	deg
TwrBsMyt	ElastoDyn - Tower Base Loads	Tower base pitching (or fore-aft) moment (i.e., the moment caused by fore-aft forces)	kNm
GenPwr	ServoDyn - Generator and Torque Control	Electrical generator power	kW

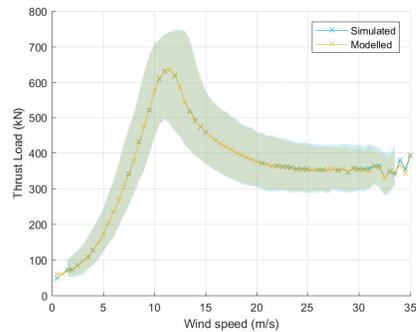
To train the model a dataset with a total simulated time of approximately 2,5 days is used. Additionally, the model is validated on an additional simulated dataset with a total simulated time of approximately 1,5 days. This data is not used to train the model and can thus be used as a fully independent validation set.

Results are shown in Figure 4.16. A good match between modeled thrust load  $\hat{F}_T$  and simulated thrust load  $F_{T,s}$  can be found during both training and validation phase (Figures 4.16a and 4.16b). Figure 4.16c shows the relative error  $\Delta\epsilon$  between simulated and modeled thrust load ( $\Delta\epsilon = \frac{abs(F_{T,s} - \hat{F}_T)}{F_{T,s}}$ ) versus wind speed for the test set during the training phase and the total dataset during the validation phase. The line indicates the median value of the relative error, calculated for each wind speed bin of  $0,5m/s$ . The surface spans from the 5th to the 95th percentile of the data. In general, the 95th percentile of the relative error of both datasets barely exceeds 10%, except for very low wind speeds. Here, a higher relative error is found due to the lower absolute values of the thrust load. For higher wind speeds, errors are increasing. Starting from  $12m/s$  an increasing variability in relative error can be observed for increasing wind speed. A similar behavior was found for the training and the validation set during the training phase. This indicates

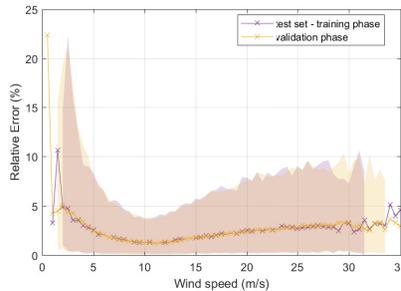
the training set was representative for the validation set, the test set and the total dataset during the validation phase.



(a) Modeled and simulated thrust loads for the total dataset of ca 2,5 days during the training phase, including training, validation and test set



(b) Modeled and simulated thrust loads for the total dataset of ca 1,5 days during the validation phase

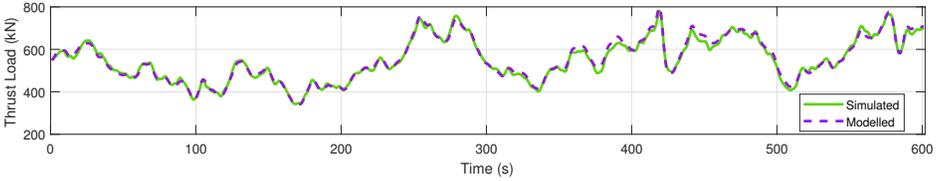


(c) The relative error for the test dataset during the training phase and the additional dataset during the validation phase

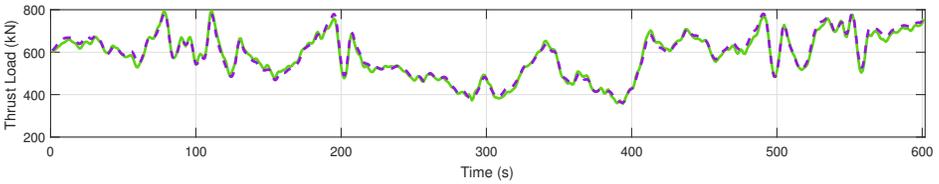
Figure 4.16: Results based on simulation data obtained with FAST. The lines indicate median values for every wind speed bin of  $0,5\text{m/s}$ ; the surface spans from the 5th to the 95th percentile of the data.

For further illustration, four time series spanning 10 minutes during the validation phase are shown in Figure 4.17, two while operating below rated power (Figures 4.17a and 4.17b) and two while operating at rated power (Figures 4.17c and 4.17d). The time series of 10 minutes with the highest averaged absolute error between simulated and modeled thrust loads below rated and at rated power are shown in Figure 4.17b and Figure 4.17d respectively. In all cases a very good match between the model and the simulated data is found. This is represented by a low value for the averaged absolute relative error for those time series (MRE < 4,5%).

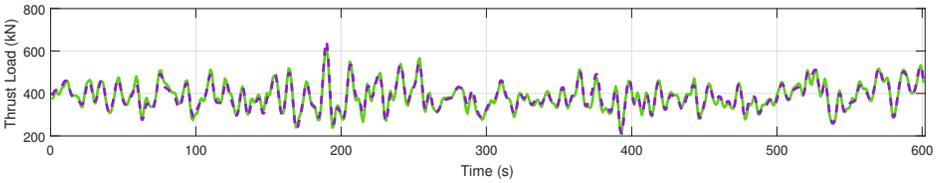
The presented results confirm the feasibility of the proposed methodology to reproduce the thrust loading purely on SCADA data.



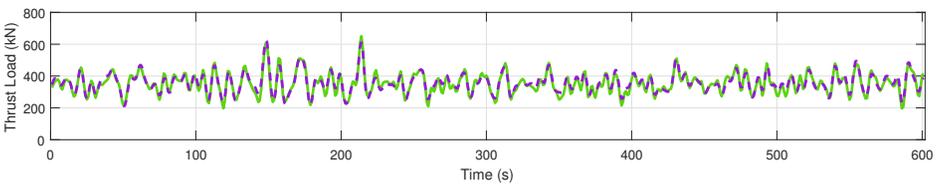
(a) Below rated power. MRE = 1,6042%



(b) Below rated power. MRE = 1,8039%



(c) At rated power. MRE = 2,7471%

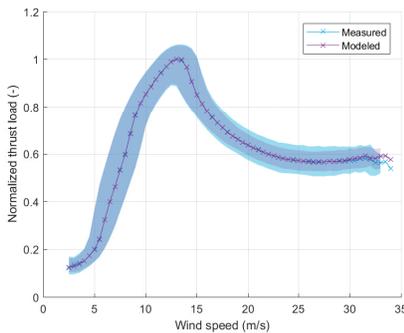


(d) At rated power. MRE = 4,4626%

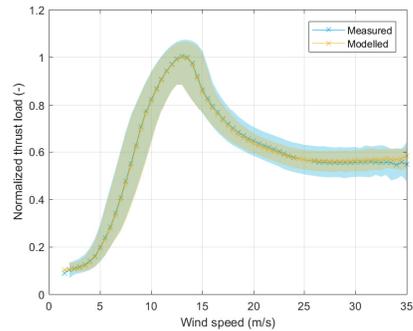
Figure 4.17: Time series (spanning 10 minutes) of modeled and simulated thrust loads below rated (a,b) and at rated power (c,d). For each operational state, the time series with the highest averaged absolute error is shown (b,d). MRE shows the averaged absolute relative error over 10 minutes.

### Application to real world: offshore wind turbine

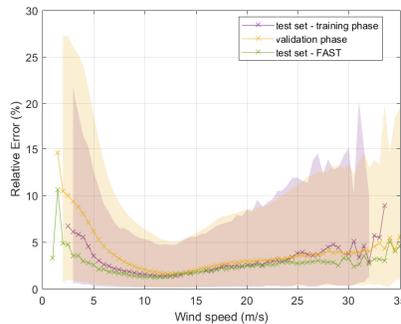
The proposed modeling method is also tested on an operating wind turbine. First, results from [18] are shown and discussed. In the current analysis, the model is trained using 2 weeks of data and validated using one year of data. The validation period includes the training period. A different period from the one described in Section 4.4.1 is used. The model was trained and validated using generating conditions only.



(a) Modeled and measured thrust loads for the total dataset of 2 weeks during the training phase, including training, validation and test set



(b) Modeled and measured thrust loads for the total dataset of one year during the validation phase



(c) The relative error for the test dataset during the training phase and the additional dataset during the validation phase

Figure 4.18: Results based on measurement data obtained from an operating offshore wind turbine. Only data during generating conditions is considered. The lines indicate median values for every wind speed bin of  $0,5\text{m/s}$ ; the surface spans from the 5th to the 95th percentile of the data.

Results are shown in Figure 4.18. A good match between measured thrust loads  $F_{T,m}$  and modeled thrust loads  $\hat{F}_T$  can be found during both training and

validation phase. Although above roughly  $18m/s$ , the modeled thrust curve shows less variability than the measured curve, meaning the difference between the 5th and the 95th percentile of modeled thrust is lower than the difference between the 5th and the 95th percentile of measured thrust (Figure 4.18a,b). Figure 4.18c shows the relative error ( $\Delta\epsilon = \frac{abs(F_{T,m} - \hat{F}_T)}{F_{T,m}}$ ) of the test set during the training phase, being 15 % of the total training set of 2 weeks, and the total dataset during the validation phase of one year of operation. Again the line indicates the median value, calculated for every wind speed bin of  $0,5m/s$ . The surface spans from the 5th percentile to the 95th percentile of the data. A similar behavior among the training, validation and test set during the training phase was obtained. In general, the relative error does not exceed 15 %. Moreover, with a median value barely exceeding 5 %, results are promising.

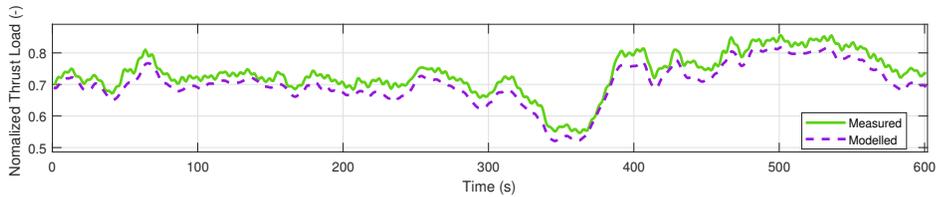
In general, the errors have increased with respect to the results using FAST (as shown in Figure 4.18c).

The errors obtained for lower wind speeds up to  $10m/s$  have increased due to a difference between the averaged measured and the averaged modeled value during a certain period, called an offset. These offsets have even increased for the long term validation set, with respect to the test set during the training phase. As a result, the errors have increased as well.

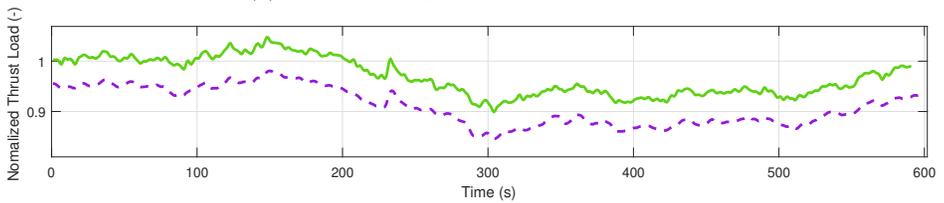
Furthermore, the errors obtained for wind speeds higher than ca.  $18m/s$  are slightly higher due to the loss of variability in the tail of the thrust curve. Again, an increasing variability in relative errors can be observed for increasing wind speeds, starting from ca.  $12m/s$ .

To further illustrate these observations, four time series of 10 minutes are shown in Figure 4.19. Two of them show operation below rated power (Figures 4.19a and 4.19b), while the other two show operation at rated power (Figures 4.19c and 4.19d). The time series depicted in Figure 4.19b and Figure 4.19d show 10 minutes with the highest averaged absolute error when operating below or at rated power respectively. An offset can be observed in Figures 4.19a and 4.19b, while the loss of variability can be observed in Figure 4.19d. The values of the averaged absolute relative error indicate the match is still acceptable (MRE smaller than 6,5 %). As explained, the resulting errors are influenced a lot by present offsets between the measured and the modeled thrust load signal. However, these offsets won't influence a fatigue assessment performed according to common practice in industry. This practice consists in cycle counting of the stress signals and transforming the cycle counts into damage using the Miner's rule. As this approach only considers the size of the cycles, the following fatigue assessment is not influenced by the mean value of the cycles.

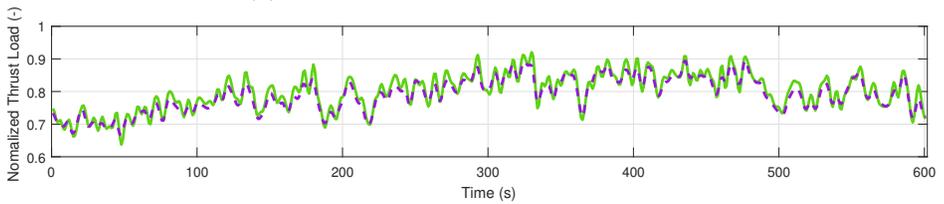
However, when looking at a longer period of time than 10 minutes, it can be seen these offsets are not constant over time. This means that they will influence a fatigue assessment if the assessment is performed based on signals consisting much longer than 10 minutes.



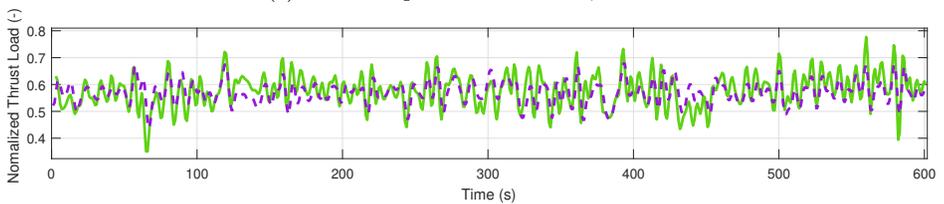
(a) Below rated power. MRE = 4,3505%



(b) Below rated power. MRE = 6,0214%



(c) At rated power. MRE = 1,6433%



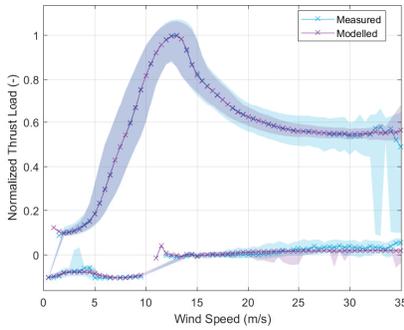
(d) At rated power. MRE = 6,4893%

Figure 4.19: Time series (spanning 10 minutes) of modeled and measured thrust loads below rated (a,c) and at rated power (b,d). For each operational state, the time series with the highest averaged absolute error is shown (c,d). MRE shows the averaged absolute relative error over 10 minutes.

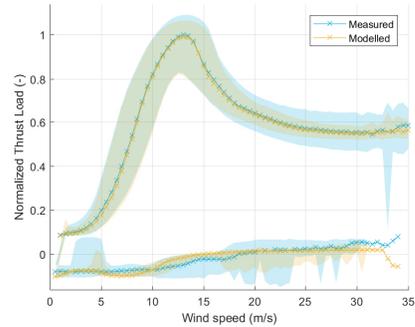
The same procedure was repeated using all data, including operational and parked conditions. Here, the same training and validation period is used as explained in Section 4.4.1. Results are shown in Figure 4.20. Again, a good match between measured thrust loads  $F_{T,m}$  and modeled thrust loads  $\hat{F}_T$  can be found during both training and validation phase, for both operational and parked conditions. Although above  $18m/s$ , the modeled thrust curve shows less variability than the measured curve (Figure 4.20a and Figure 4.20b). This was also observed when only generating data was used to train and validate (Figure 4.18). However the difference in variability is even bigger when parked and idling conditions are included in the training data (Figure 4.20). In Figures 4.20a and 4.20b a distinction is made between very low values for thrust and higher values. The normal thrust curve, for which thrust loads go up to their maximum value, primarily represents the generating cases. The very low values for thrust load (below zero), represented by the lower part of the figure, are mainly caused by non-generating cases. When looking at the non-generating cases in Figure 4.20b, an acceptable match between measured and modeled thrust loads is observed as well. This is even the case for the wind speed range where the training set lacked a bit of data (between  $10m/s$  and  $15m/s$ ). Figure 4.20c shows the absolute relative error of the test set during the training phase, being 15 % of the total training set of 2 weeks, and the total dataset during the validation phase of one year of operation. Again the line indicates the median value, calculated for every wind speed bin of  $0,5m/s$ . The surface spans from the 5th percentile to the 95th percentile of the data. A similar behavior among the training, validation and test set during the training phase was obtained.

Generally, the results of the relative absolute error are promising. The median value stays far below 10%, except for very low and very high wind speeds. Striking are the very high errors for high wind speeds in the training test set, with respect to the validation set. Most likely, the training test set has a higher percentage of data points when the turbine was not generating power than the validation dataset. During non-generating conditions, the absolute value of thrust load is very low. This low value can lead to a very high relative error, even though the error is not that big in absolute terms.

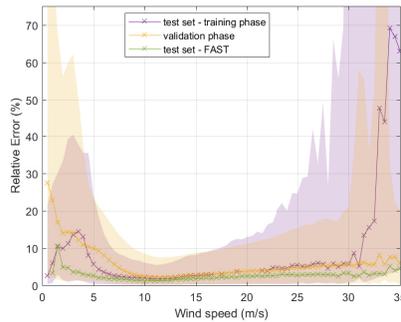
When comparing these results to the ones obtained when only generating conditions were considered (Figure 4.18), an increase in error can be observed. Again, this can be explained by the existence of data points when the turbine is not generating and their higher relative errors. Since non-generating conditions usually occur for very low wind speeds below cut-in and very high wind speeds above cut-out, these regions show a higher relative error.



(a) Modeled and measured thrust loads for the total dataset of 2 weeks during the training phase, including training, validation and test set

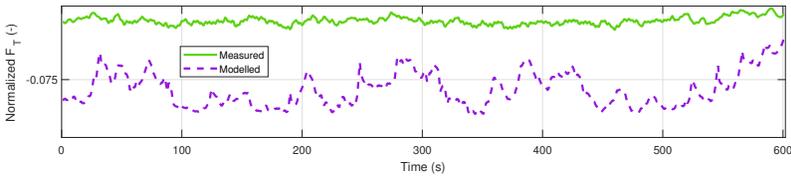


(b) Modeled and measured thrust loads for the total dataset of one year during the validation phase

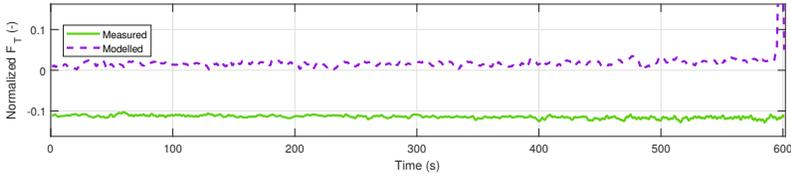


(c) The relative error for the test dataset during the training phase and the additional dataset of one year during the validation phase

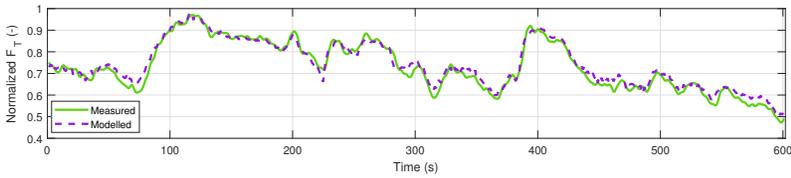
Figure 4.20: Results based on measurement data obtained from an operating offshore wind turbine. The lines indicate median values for every wind speed bin of  $0,5\text{m/s}$ ; the surface spans from the 5th to the 95th percentile of the data. For the thrust curve, two datasets are considered for each wind speed bin. One for high thrust loads and one for low thrust load (the limit is ca. 5% of the maximum thrust).



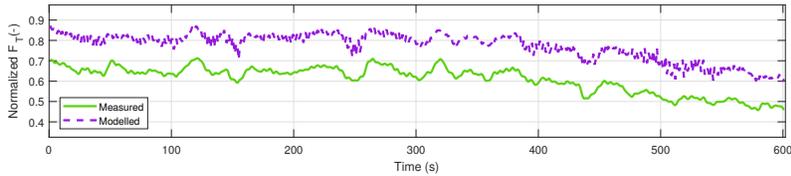
(a) Not generating. MRE = 30,4%



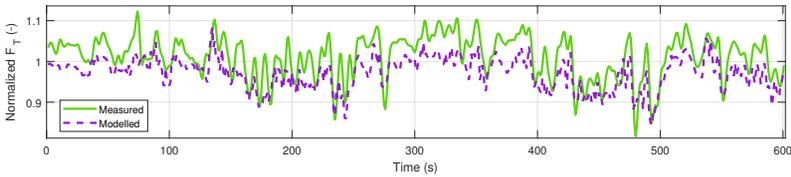
(b) Not generating. MRE = 113%



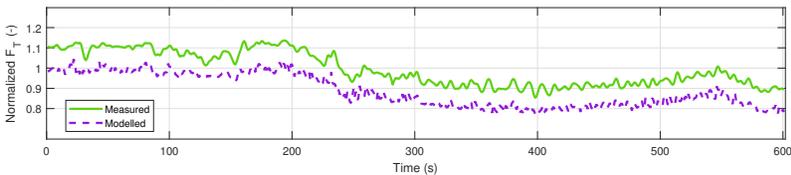
(c) Below rated power. MRE = 2,45%



(d) Below rated power. MRE = 25,4%



(e) At rated power. MRE = 4,48%



(f) At rated power. MRE = 10,5%

Figure 4.21: Time series (spanning 10 minutes) of modeled and measured thrust loads for not generating(a,b), below rated power (c,d) and at rated power (e,f). For each operational state, the time series with the highest averaged absolute error is shown (b,d,f). MRE shows the averaged absolute relative error over 10 minutes.

Again, six time series of 10 minutes are shown in Figure 4.21. Two of them show non-generation (4.21a and 4.21b), two show operation below rated power (4.21c and 4.21d) and the remaining two show operation at rated power (4.21e and 4.21f). The time series depicted in Figure 4.21b, Figure 4.21d and Figure 4.21f show 10 minutes with the highest averaged absolute error when not generating, operating below or at rated power respectively. Again, an offset can be observed for almost all examples and the amplitude of the cycles at rated power is not always modeled correctly (Figure 4.21e). For non-generating conditions, one can observe high offsets and the oversensitivity of the model to SCADA parameters such as wind speed (Figure 4.21a) or blade pitch angle (Figure 4.21b). The values of the averaged absolute relative error indicate the match is usually still acceptable, but in case of bigger offsets or non-generating the MRE values get quite high.

In conclusion, one could decide to split up the model based on operating conditions. However, a prior more profound analysis of the resulting errors between measured and modeled thrust load is advisable. This analysis should split up the resulting errors properly based on operating conditions. In that case it would be possible to conclude whether or not the inclusion of all data points, with respect to only generating data points, influences the performance of the neural network in case of power generation.

Another possibility is to only include generating conditions in the training set, since the thrust load during non-generating conditions is typically negligible compared to the other loads acting on the wind turbine.

### 4.4.3 Correlation between measured and modeled signal

In Section 4.3, the Pearson correlation and the mutual information were calculated between the measured thrust signal and several SCADA signal to decide which parameters would be the most valuable. A similar analysis can be done between the measured and the modeled (1Hz) thrust signal. This is done for two types of models, one is trained with 10 minute averages (Section 4.4.1), while the second model was trained with 1s data (Section 4.4.2). For both models the same 2,5 month period was considered

Figure 4.22 shows that for all generating data, the Pearson correlation (Figure 4.22a) and the mutual information (Figure 4.22b) for both modeled thrust loads have greatly improved over the individual SCADA parameters. These increased values indicate a lot of the present variability in the thrust load can be explained by combining several SCADA parameters even on a 10min basis.

When looking at the results for non-generating data, the performance for the modeled thrust loads are not always higher than for some of the SCADA parameters, such as the blade pitch angle and rotor speed. This also indicates improvement can still be achieved in modeling the thrust when the turbine is not generating power.

To compare the results of both models, the absolute values are given in Table 4.3. One can see that in almost all cases the model trained using 1s data obtained

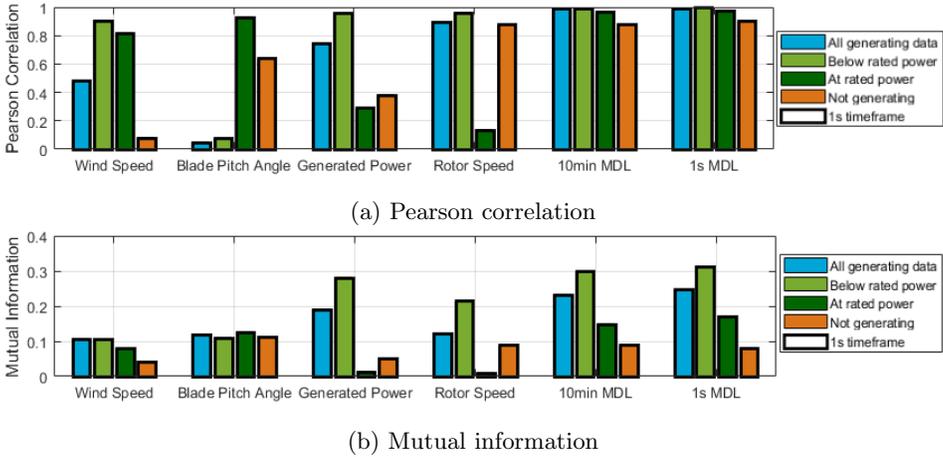


Figure 4.22: Pearson correlation and mutual information between measured thrust and the input parameters and the two types modeled thrust load for 1s data. One modeled thrust load is obtained using a model trained with 10min data (Section 4.4.1). The second modeled thrust load is obtained using a model trained with 1s data, including non-generating conditions (Section 4.4.2). For the Pearson correlation of the input parameters, the absolute value is shown.

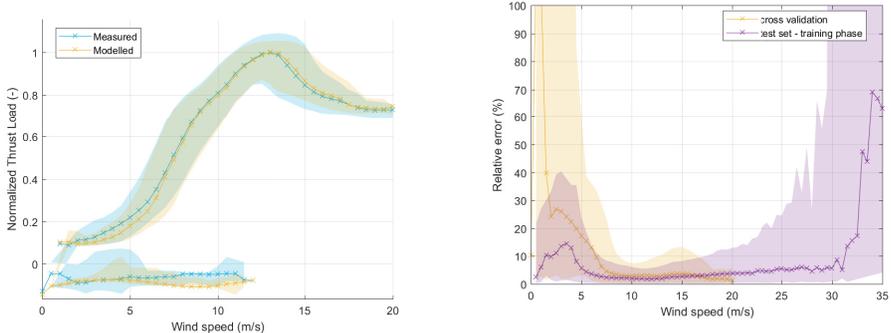
higher (linear and non-linear) correlation values than the model trained using 10 minute averages. The slight raise of the 1s model over the 10min model indicates allowing some latency, or time-lag, between the different signals improved the prediction even more. This confirms the previous analyses, where the model trained with 1s data also showed better results.

Table 4.3: The resulting values for Pearson correlation and mutual information between measured and modeled thrust load. Two types of models are used.

	Pearson Correlation		Mutual Information	
	10min model	1s model	10min model	1s model
All generating data	0,9927	0,9941	0,2331	0,2472
Below rated power	0,9959	0,9965	0,2991	0,3113
At rated power	0,9672	0,9746	0,1498	0,1695
Not generating	0,8791	0,9040	0,0903	0,0820

## 4.5 Cross validation

In theory one thrust model is valid for every turbine of the same type, as indicated in Section 4.2.2. To validate this statement, the model obtained using all 1s data from one turbine (as shown in Section 4.4.2) was applied on a different turbine



(a) Modeled and measured thrust loads for the total validation dataset of two months

(b) The absolute relative error for the validation dataset of two months and the test set during the training phase

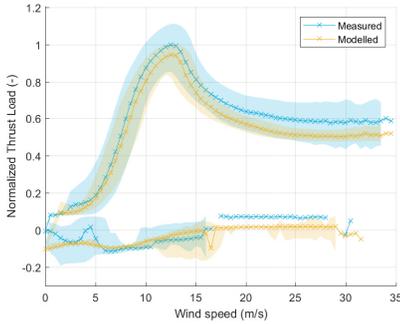
Figure 4.23: Cross-validation results where the model was trained on a different turbine in the farm. Model outcome using 1s data of SCADA parameters and measured strains for two months. The model was trained based on 1s SCADA data and measured strains of 2 weeks. The lines indicate median values for every wind speed bin of  $0,5\text{m/s}$ ; the surface spans from the 5th to the 95th percentile of the data. For the thrust curve, two datasets are considered for each wind speed bin. One for high thrust loads and one for low thrust load (the limit is ca 5% of the maximum thrust).

of the same type that was also instrumented with a load monitoring setup. This cross validation was done for a shorter period of 2 months, as availability of data on the second turbine was less.

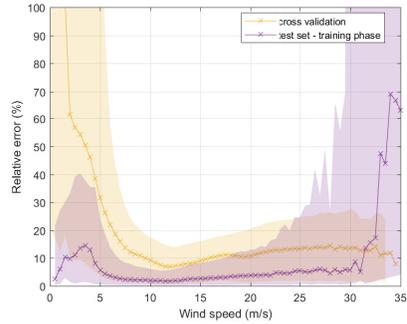
Figure 4.23a shows the measured and resulting modeled thrust load for this turbine and Figure 4.23b shows the relative absolute error between both.

A good agreement between modeled and measured thrust load is found (Figure 4.23a). Given the smaller dataset, no measurements were obtained for high wind speeds during these two months. This is also represented by the relative error (Figure 4.23b). Again, the biggest errors are found for very low wind speeds, due to the low absolute value of thrust loads. In general, the errors obtained for the cross validation are slightly higher than those of the test set during the training phase. The match is still acceptable and the cross validation can be considered a success.

The trained model was cross validated a second time, on a turbine of the same type but at which a rotor imbalance caused an additional cyclic loading at the rotational (1P) frequency. Since this frequency is quite low, it is still within the frequency band of the thrust load ( $[0, 0, 2]Hz$ ). However, SCADA data is believed not to be influenced by such a 1P harmonic. Figure 4.24 shows the results for this cross validation. Figure 4.24a shows a clear difference between measured and modeled thrust load. First of all, the modeled thrust load seems to be underestimated by the model. Moreover, the surface spanning from the 5th



(a) Modeled and measured thrust loads for the total validation dataset of two months



(b) The absolute relative error for the validation dataset of two months and the test set during the training phase

Figure 4.24: Cross-validation results between two turbines in a single farm, however for the considered period the validation turbine had a significant load-imbalance.

to the 95th percentile is clearly smaller in case of the modeled thrust load with respect to the measured thrust load. This was also the case for the validation on the same turbine (Figure 4.20). But it was not observed for the first cross validation (Figure 4.23), although it is usually mainly present for very high wind speeds over  $18m/s$ . When looking at the absolute relative error (Figure 4.24b), again an significant increase in errors can be observed over the entire wind speed range.

The explanation of this increase in errors can be illustrated using a time series (Figure 4.25). Looking at the time series, the overall trend of both signals is similar. However, the measured signal shows a lot more cycles at a specific frequency. These are the result of rotor imbalance, causing a 1P variation. It is clear from these results the effects of the additional load due to a rotor imbalance are not reflected by the model.

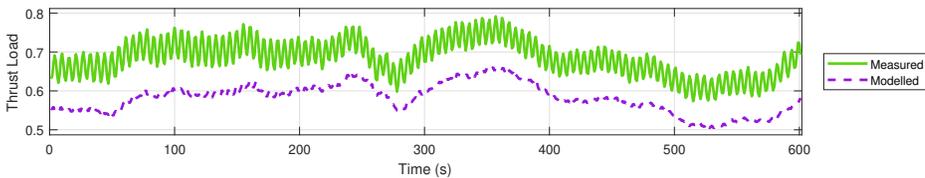


Figure 4.25: Time series (spanning 10 minutes) of modeled and measured thrust loads for operation below rated power. The turbine is influenced by a rotor imbalance causing a 1P variation in the measured signal.

## 4.6 Importance of data quality

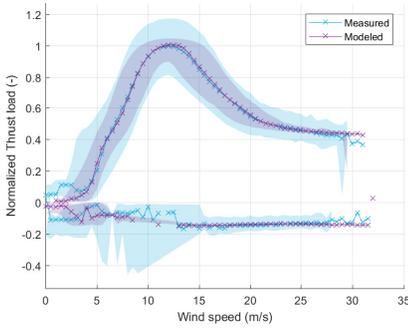
To illustrate the importance of the quality of the SCADA data used to train and apply a thrust model, the same analysis is done for a different turbine (of a different type). This turbine was installed on a monopile instead of a jacket. First strong data filters were applied to remove all unrealistic data. The filters consisted in removing all data points for which the wind speed or the generated power had a value outside the interval  $[0; 50]m/s$  or  $[-0, 1; 1, 25] \cdot P_{rated}$  respectively. Moreover a comparison is made between the wind speed value in the 10min SCADA dataset and the average value over 10 minutes based on 1s SCADA data. If the difference is bigger  $0, 1m/s$ , indicating an inconsistency in the SCADA data, the entire time series was rejected. Finally, any other data points for which the wind speed value is exactly the same for two or more consecutive data points, are rejected as well. Most of these inconsistencies are believed to be the result of a storing issue of the data.

After these filters, only 54,57% of the total original dataset remained. This implies that when a 1sec model would be used, it can only be applied 55% of the time, as the 1s SCADA and consequently the model outcome is otherwise unreliable.

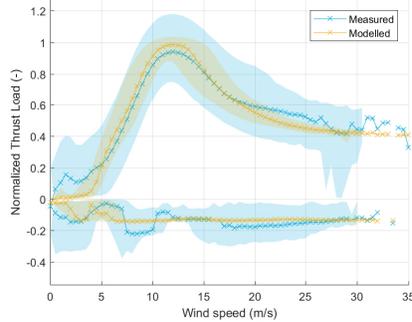
An additional difference in data quality was the accuracy of the data. Instead of having two decimal places for all measurements, the SCADA data of this turbine only had one decimal place. This affects the observability of small changes in the SCADA signals.

Two weeks of the remaining 1s data was used to train a thrust model. Afterwards the model was validated on 1s data of 6 months, again after applying the same strong data filters. Results are given by Figure 4.26. For both the training data and the validation data a fairly good match can be found for the median thrust curve (Figures 4.26a and 4.26b). Although, a clear difference in results can be observed between the validation and the training set. When looking at the surface spanning the 5th to the 95th percentile, the model is not capable of capturing the entire surface, especially for maximum thrust and high wind speeds. This is the case for both training data and validation data. This is translated into a much higher absolute relative error for both datasets (Figure 4.26c), compared to previous models (for example Figure 4.20c). It is believed this is due to the poor quality of the SCADA data. Since the quality is less, lower variation in the specific parameters can be observed and used by the neural network.

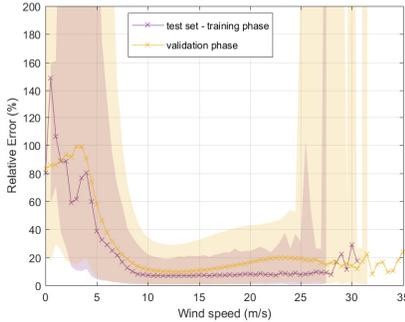
It should be noted, given the different foundation, the role of waves can play a role as well. As monopiles are more prone to wave loads than jackets, it might be some wave loading is still present in the measured thrust signal. However, this wave loading cannot be predicted by the thrust model based on SCADA data. More in depth analysis to distinguish the role of the quality of the data and the role of the waves might be advisable. Nonetheless the perfect results obtained using the simulation data of FAST in Section 4.4.2 and depicted in Figure 4.16, were based on a OWT installed on a monopile subjected to wave loading. This could indicate the role of wave loading in this approach is minor.



(a) Modeled and measured thrust loads for the total dataset of 2 weeks during the training phase, including training, validation and test set



(b) Modeled and measured thrust loads for the total dataset of six months during the validation phase



(c) The relative error for the test dataset during the training phase and the additional dataset of six months during the validation phase

Figure 4.26: Results based on measurement data obtained from an operating offshore wind turbine with data of less quality. The lines indicate median values for every wind speed bin of  $0,5m/s$ ; the surface spans from the 5th to the 95th percentile of the data. For the thrust curve, two datasets are considered for each wind speed bin. One for high thrust loads and one for low thrust load (the limit is at zero).

## 4.7 State-by-State modeling approach

An alternative modeling approach has been published in [19]. Instead of a single model to cover all states of the turbine, the presented approach trained different neural network models for different operational states of the turbine. The considered states were:

- not generating power
- generating below rated power
- generating at rated power

The decision of which input parameters should be used to create the model is based on the results for the Pearson correlation and mutual information for each operational state (Section 4.3). When the turbine is not producing, the highest correlations are found for blade pitch angle, generated power and rotor speed. For operation below rated power, rotor speed, wind speed and generated power are selected. Finally, when operating at rated power: blade pitch angle, generated power and wind speed are the considered input parameters. The aforementioned is summarized in Table 4.4.

Table 4.4: For every regime, a neural network is trained with a different input parameter set

	Non-producing	Below rated power	At rated power
blade pitch angle	x		□
rotor speed	x	△	
wind speed		△	□
produced power	x	△	□

For each of the three regimes, a neural network is trained using the default MATLAB functions. This means the training dataset is randomly split up into training (70%), validation (15%) and test data (15%). A feed-forward neural network with one hidden layer with four neurons is trained with the remaining training data, based on the Levenberg-Marquardt algorithm. The transfer function for the hidden layer is a tan-sigmoid transfer function; for the output layer it is a linear transfer function. The training stops when the validation error failed to decrease for six iterations. Finally, the resulting model to estimate the thrust load combines the three neural networks, where for each data point first is decided which neural network will be used to estimate the thrust load, based on the SCADA values.

The main advantage of this modeling technique is the increased flexibility with respect to input parameter selection. Different parameters can be chosen for different regimes. This might increase the understanding of the system and facilitate a transition to a regression technique. [60] mentions the necessity to establish separate regression models over the entire operating range.

On the other hand, this technique requires an increased foreknowledge about the different operating states of the turbine. Moreover, by splitting the model, chances

are bigger to have worse predictions in transition areas.

The alternative technique is applied on the 1s measurement dataset used in Section 4.4.2. Figure 4.27 shows the measured thrust load for the same training dataset as in Section 4.4.2, divided into 3 groups based on the operating state of the turbine. De-rated data is classified in the category 'below rated power'.

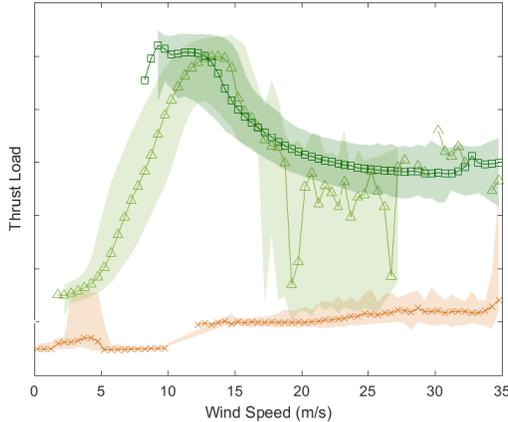


Figure 4.27: Measured thrust load, calculated from measured strains, vs wind speed. The data points are divided into 3 regimes: non-producing, producing below rated power and producing at rated power.

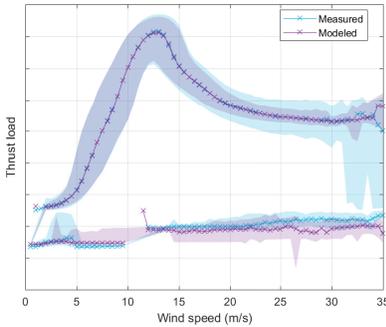
Table 4.5 gives the resulting mean and median absolute relative error for the test set of each regime. The mean value for the operating state below rated power is rather high, possibly caused by some points of de-rating.

Table 4.5: The mean and median relative absolute error (RAE) of each neural network on the test data set

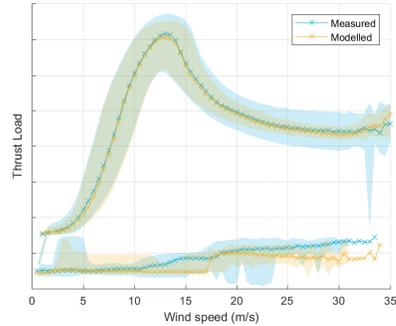
	Non-producing	Below rated power	At rated power
Mean RAE	0,4281	10,6370	0,0425
Median RAE	0,1438	0,0269	0,0336

Figure 4.28a shows the resulting modeled thrust load, together with the measured thrust load for the training set. Figure 4.28b does the same for the validation data set of one year utilized in Section 4.4.2. Figure 4.28c shows the relative error between measured and modeled thrust load for the test set during the training phase and the validation dataset of one year.

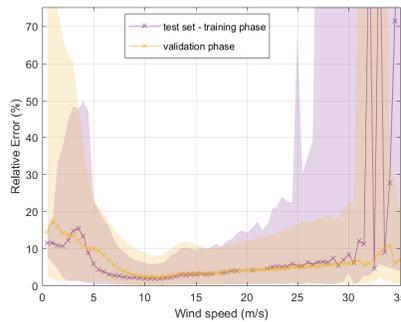
Figures 4.28a and 4.28b show a good agreement between measured and modeled, except for the non-producing data points for wind speeds higher than 10m/s. Moreover the modeled thrust load has clearly much less variability for producing data points for wind speeds higher than 20m/s, just as the model trained by all



(a) Modeled and measured thrust loads for the total dataset of 2 weeks during the training phase, including training, validation and test set



(b) Modeled and measured thrust loads for the total dataset of one year during the validation phase



(c) The relative error for the test dataset during the training phase and the additional dataset during the validation phase

Figure 4.28: Results based on measurement data of an offshore wind farm. The model utilized is based on a split according to the operational state of the turbine. The lines indicate median values for every wind speed bin of  $0, 5m/s$ ; the surface spans from the 5th to the 95th percentile of the data. For the thrust curve, two datasets are considered for each wind speed bin. One for high thrust loads and one for low thrust load (the limit is ca 5% of the maximum thrust).

data at once. Relative errors are comparable to those obtained using one neural network for all data points (Figure 4.20c).

With some additional effort and enough foreknowledge about the turbine, this approach might result in an improved model with respect to the single model, e.g. in non-generating conditions. However, the results obtained with the models as proposed in this section did not show a great improvement. Therefore, if the obtained results with one single model are already satisfying, there is no need to combine different models based on the operating state of the turbine.

## 4.8 Conclusions

An approach to estimate thrust load signals based on SCADA data is explained and validated in this section. Wind speed, rotor speed, blade pitch angle and generated power are selected as input parameters based on both linear and non-linear correlation analyses, Pearson correlation and mutual information respectively. Strain sensors are used to measure the acting thrust load. This thrust load signal is combined with SCADA signals to train a neural network. This was done for both 10min and 1s data.

For a model trained with 10min data, good results were obtained for the application of the model on 10min data. However application of such a 10min model on 1s signal is not advised. Overall, the 10min model can predict the larger cycles in the 1s signal. However, the equally important smaller variations can be missed or over-predicted.

Therefore most thrust load models in this chapter were trained using 1s data. All training was done based on only two weeks of data. This training period seemed to be sufficient for the prediction. However it is important to make sure the full wind speed range occurred during this training period. If this is not the case, a different or longer training period has to be selected.

The validation is done using FAST simulation data and data measured at an offshore wind turbine during one year. Time series show a good match between modeled and measured or simulated thrust signals. In general, the absolute relative error barely exceeds 20 % and the median value per wind speed barely 5% under normal operating conditions. Results obtained using FAST data are slightly better than those of the real world offshore wind turbine.

When only considering generating conditions, results are even better. An improvement for parked and idling conditions could be suggested, but in reality the thrust load during those conditions is not of high value. During non-generating conditions, other loads are more dominant than the thrust load.

The use of 1s SCADA data can be considered as the main advantage of this approach. A cross validation was successfully performed, validating one of the trained models on a different turbine of the same type. Thus the model proves to be transferable among turbines of the same type. Therefore this approach can be applied on any (non-instrumented) wind turbine within a wind farm, if the turbine type matches.

In theory, any instrumented turbine can be used to train the thrust model. However, it was shown that an additional 1P variation caused by a rotor imbalance was not filtered out of the measured thrust load signal. However, such an additional loading is not captured by the SCADA data. Therefore, the model should be trained on a turbine without a rotor imbalance.

Essential in this approach is the preprocessing and the quality of the SCADA data. It was shown less accurate SCADA data resulted in higher relative errors between measured and modeled thrust load. At least two decimal places would be advisable. Moreover, the more SCADA data has to be removed, the less modeled thrust load values can be obtained. Keeping that in mind, a good storing procedure of 1s SCADA values is advised as well.

Further improvement of the proposed technique is definitely still possible. First of all, more foreknowledge about the turbine and its working principles can help improving the model by, for example, a better selection of input data or a possible split in model based on the operating state. Moreover, as the results obtained in this section were based on the default settings of the neural network toolbox in MATLAB, improvement of the model settings of the neural network can still be done. Neither was the predefined topology of the neural network changed during the analyses. This means (slight) improvements might be possible here as well, both in terms of computational time and resulting errors.

Additionally, the robustness of the model training should be checked for future applications.



# Chapter 5

## Full load estimation

The majority of the work presented in this chapter has been published in [19, 20].

*An accurate stress or strain history at fatigue critical locations is a vital input for a fatigue assessment for offshore wind turbines. Unfortunately, it is not always feasible to install strain gauges at these fatigue hot spots, especially on existing wind turbines. This chapter compares two techniques to obtain a reliable stress history at any location of the turbine structure using the combination of a SCADA-based thrust load and additional accelerometers. This simplifies the existing techniques for virtual sensing [13] as it eliminates the need to use strain gauges, which are typically unreliable and costly to install, by using a data-source that is available on every turbine.*

*The first technique in this chapter is based on modal decomposition and expansion, the second is based on a Kalman filter. Both techniques are validated and compared using data from an offshore wind turbine monitored by OWI-lab. The monitored turbine is instrumented with strain gauges at the interface between transition piece and tower and accelerometers at multiple levels. The installed strain gauges allow to validate the proposed techniques with respect to the reality.*

## 5.1 Introduction

For offshore wind turbines, both quasi-static wind/thrust loads and dynamic loads, as induced by turbulence, waves and the turbine's dynamics, contribute to its fatigue life progression. To estimate the remaining useful life of an offshore wind turbine, the stresses acting on the fatigue critical locations within the structure are an important input. Therefore they should be monitored continuously. Unfortunately, in case of the most common monopile foundations these locations are often situated below sea-level and near the mud line and thus difficult or even impossible to access for existing turbines. To resolve this limitation OWI-lab has worked on the concept of virtual sensing [75, 76, 13]. Virtual sensing allows to estimate stresses at hotspots using measurements taken at more accessible locations within the structure such as on the tower structure. In particular accelerometers are favored as sensors, due to the ease of installation and their reliability on the long term. However, low frequency thrust loads cannot be adequately captured using accelerometers. Up to this point strain gauges have been part of the virtual sensing strategy in order to quantify the low-frequency contributions. However, strain gauges are not favorable for application offshore. While the sensor itself is cheap, the installation is time-consuming and prone to error even for skilled technicians. Moreover it is very likely the strain gauges will fail over time.

This research investigates to what extent SCADA data can be used, as explained in Chapter 4, to replace the role of the strain gauges once these have failed or on different sites in the wind-farm where no strain gauges are installed. In this chapter two techniques are introduced to combine both quasi-static (obtained from the SCADA model) and dynamic loads (derived from the accelerometer) to estimate the full band strain history at any location on the foundation induced by the loads acting on the entire structure.

The focus of this chapter is to validate the possibility to merge a thrust load signal sampled at 1Hz, as would be the outcome of the SCADA-model of Chapter 4, with the outputs of a high-frequency dynamic prediction. To avoid any errors originating from the SCADA model, the thrust load signal used for this validation is obtained from the measured strain signal. This is done in the same way to obtain the measured thrust signal to which modeled strain were compared in previous chapter (Chapter 4) and is explained in Section 4.2.1. As such a *perfect* model could be assumed.

To show the full potential, one of the techniques is also applied using the SCADA-based thrust load estimations as obtained using a model as explained in Chapter 4.

The proposed techniques are validated using strain measurements above sea level, at the interface between tower and TP. However, the final goal of virtual sensing is to estimate the stress signals at fatigue hotspots, often near the mud line. Therefore the work presented in this chapter is only the first step towards validation of virtual sensing. [77] shows the continuation of the validation at multiple locations in the monopile under water and subsoil.

## 5.2 Monitoring campaign

The techniques proposed in this contribution will be validated using measurements taken at an offshore wind turbine located in the middle of the Belgian offshore wind farm Belwind, 46 km off the Belgian coast. This Vestas 3MW V90 turbine is installed on a monopile foundation and was instrumented with additional accelerometers and strain gauges during multiple measurement campaigns.

The turbine was instrumented with accelerometers in the beginning of 2012 and with strain gauges in September 2014. Acceleration measurements are taken at 4 levels using a total of 10 accelerometers. Eight accelerometers (two per level) capture the vibrations in the X-Y direction and the two additional accelerometers at the highest level (tower top) are utilized to identify torsional vibrations in the tower. The locations are chosen based on the convenience of sensor mounting, such as the vicinity of platforms. Moreover four fiber Bragg grating (FBG) sensors at the Tower/Transition Piece interface have been installed. Figure 5.1 gives an overview of the instrumented OWT at the Belwind farm.

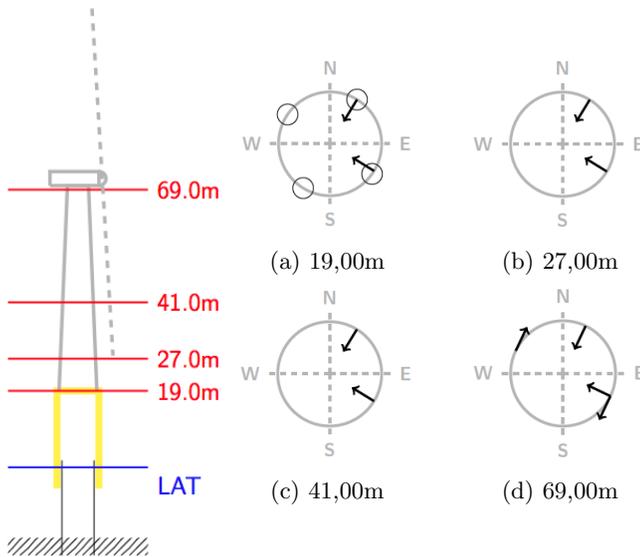


Figure 5.1: Instrumentation of BBC01 wind turbine at the Belwind farm during the first monitoring campaign. The circles in (a) indicate the presence of a Fiber Bragg Grating strain sensor; the arrows in (a-d) indicate the presence of an accelerometer.

In 2016 these sensors were replaced by a total of eight accelerometers and six strain sensors (SG). The accelerometers were installed at four different levels in the turbine tower (two per level) in order to capture vibrations in both X and Y direction. The strain gauges were installed on the lowest level, being the interface between tower and transition piece. An overview of the location of the sensors during this campaign is given in Figure 5.2.

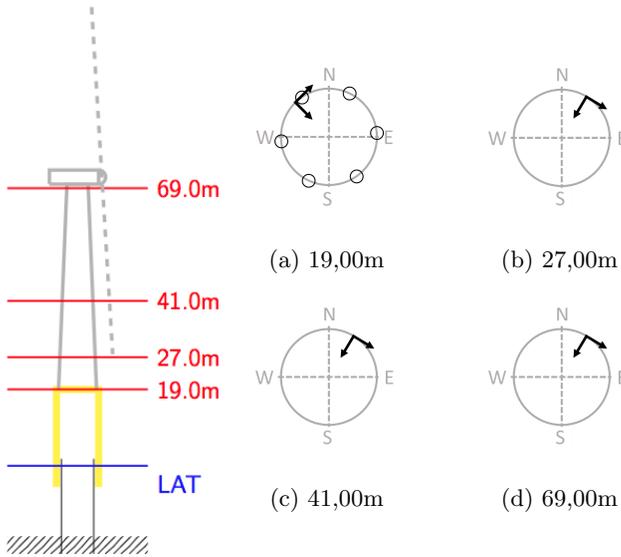


Figure 5.2: Locations of the additional sensors at of BBC01 wind turbine at the Belwind farm during the second measurement campaign. The circles in (a) represent a strain gauge; the arrows in (a-d) represent an accelerometer.

The accelerations measured at all levels together with the thrust load will be used to predict the strains measured at the tower-TP interface and finally to predict the strains at hotspots located below the water level.

### 5.3 Modal decomposition and expansion

The work explained in this section is a collaboration with Alexandros Iliopoulos and is published in two conference papers ([19] and [20]).

#### 5.3.1 Concept

The main goal of virtual sensing is to estimate the stresses at the fatigue-sensitive hotspots without the need of mounting sensors at these exact locations. This section will summarize the multi-band virtual sensing technique based on modal decomposition and expansion. This technique combines measurements with a tuned finite element model (FEM) of the structure to predict stresses at any location of the structure. The FEM is created using pipe elements with the as designed dimensions of the turbine's substructure. This FEM is tuned to match the modal properties (the natural frequencies, the mode shapes and the damping ratios of the turbine) obtained experimentally using the continuous data acquisition and by applying state of the art operational modal analysis techniques that have been fully automated [33, 78, 79]. More information about the finite element model and

the tuning can be found in [75]. Figures 5.3 and 5.4 show the resulting structural and strain mode shapes respectively.

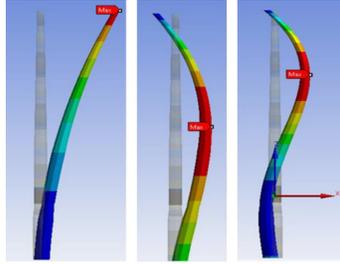


Figure 5.3: First three structural mode shapes of an offshore wind turbine. The first mode is captured best using the top accelerometer, while the second and third mode are captured better using the lower sensors.

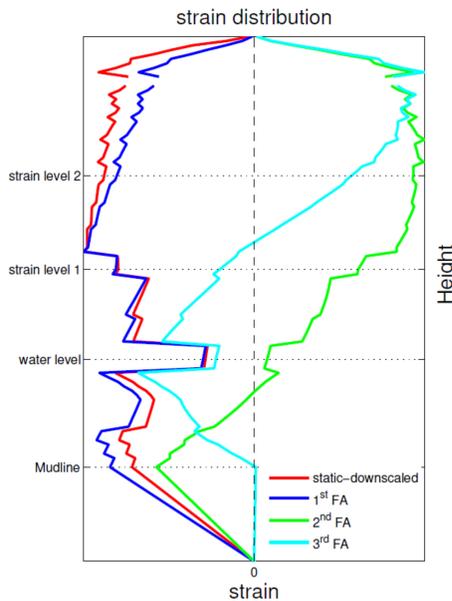


Figure 5.4: The static strain mode shape together with the strain mode shapes of the first three modes from mudline to tower top level.

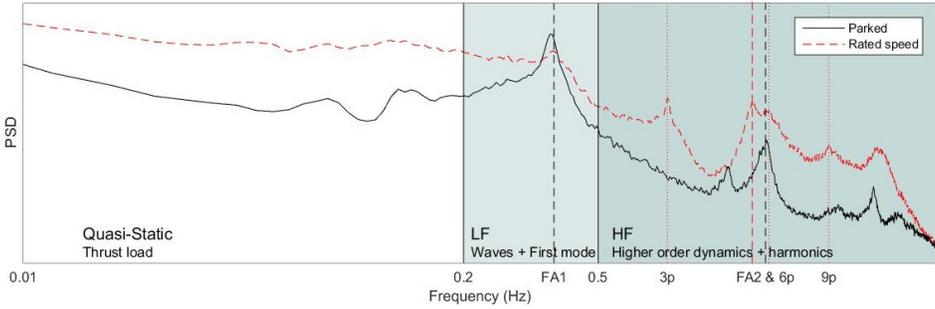


Figure 5.5: Frequency spectrum of the different loads acting on an offshore wind turbine in parked conditions and at rated rotor speed. The quasi-static load is dominated by the thrust load induced by variations in wind. Dynamic variations are induced by waves, structural dynamics and rotor harmonics.

As indicated in Section 2.1.6, the stresses observed at any location in the structure are induced by a variety of loads, each acting in a different frequency range. This can be seen in Figure 5.5. Since the structure responds differently to quasi-static loads than to dynamic loads, both types of loads are treated differently. This distinction is made based on the frequency spectrum of the loads acting on a turbine. The quasi-static thrust load can be estimated using 1s SCADA data. Unfortunately, with a 1 second sampling rate only loads up to 0,5Hz can be captured. Loads with higher frequencies such as the rotor dynamic and turbulence induced loads are thus not captured by SCADA parameters. Moreover, wave and current loads have no relation to SCADA and can thus also not be captured using a SCADA model. Both wave and high frequent loads are combined as dynamic loads. To capture these dynamic loads accelerometers are installed at easily accessible locations.

At very low frequencies situated well below the first eigenfrequency and below the site-specific wave peak frequency, a.k.a. quasi-static region of frequencies, the induced strains are caused by thrust loading. Under thrust load the strain distribution of the turbine differs from the strain distribution of the lowest structural mode (Figure 5.4). This implies the need for different strain mode shape components to represent the thrust load induced strains. Equation (5.1) is used to predict the quasi-static strains:

$$\varepsilon_{\mathbf{p}}^{\text{QS}}(t) = \phi_{\varepsilon_{\mathbf{p}}}^{\text{QS}} \frac{F_T(t)}{F_{T,ref}} \quad \forall t \quad (5.1)$$

where  $\phi_{\varepsilon_{\mathbf{p}}}^{\text{QS}} \in \mathbb{R}^{n_p \times 1}$  is the quasi-static strain distribution at the  $n_p$  DOFs which correspond to the virtual sensor locations  $p$ ,  $F_T(t)$  is the acting thrust load for each time instance  $t$  and  $F_{T,ref}$  is the reference thrust load which is exerted as a static load at the tower top of the tuned FEM in order to obtain the quasi-static strain distribution.

The modal decomposition and expansion (MDE) based virtual sensing technique [80, 81] is used for dynamic strain prediction. The theoretical background of this approach has been extensively presented in [76, 82, 75, 83, 84] and is repeated briefly hereafter.

The basic concept behind MDE is that each deflection of the structure is considered a combination of deflection caused by the excitation of different modes. Based on the acceleration measurements  $\ddot{x}(t)$  and the acceleration mode shapes  $\Phi_{acc}$  the measurements are decomposed in the contributions of the different considered modes given as *acceleration* modal coordinates  $\mathbf{q}(t)$ . This is visualized by Figure 5.6 and expressed in Equation 5.2.

$$\mathbf{q}(t) = \Phi_{acc}^{-1} \ddot{\mathbf{x}}(t) \tag{5.2}$$

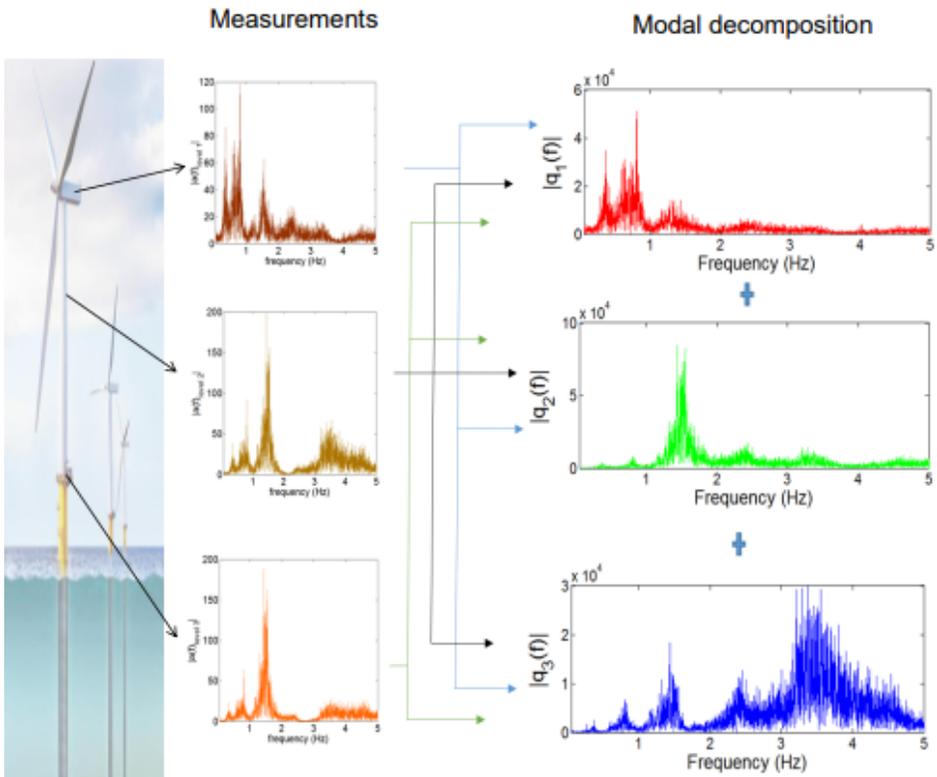


Figure 5.6: Acceleration modal coordinates are obtained by decomposing accelerations measured at various locations in the tower using the acceleration mode shapes.

Note that this operation requires that the number of modes considered is restricted by the number of available acceleration measurements. In practice this implies, that without prior knowledge, the number of considered modes is limited

by the number of accelerometers installed.

The modal decomposition results in the estimation of the *acceleration* modal coordinates  $\mathbf{q}(t)$ . Since strain is related to modal displacements, the found *acceleration* modal coordinates are integrated twice. In this work, a double integration in the Laplace domain was used, i.e. a multiplication of the Laplace spectrum with  $\frac{1}{s^2}$ . After which, a change from Laplace domain to the time domain for continuous strain prediction is achieved using the inverse Laplace transformation,  $\mathcal{L}^{-1}\{\bullet\}$ . Finally, the estimated time domain displacement modal coordinates are multiplied with the corresponding strain mode shape components  $\Phi_{\varepsilon p}$  derived numerically from a tuned finite element model. Doing so, the prediction of dynamic strain in any virtual location is established. Equation (5.3) summarizes the discussed process.

$$\varepsilon_{\mathbf{p}}^{\mathbf{D}}(t) = \Phi_{\varepsilon p} \mathcal{L}^{-1} \left\{ \frac{1}{s^2} \mathcal{L} \{ \mathbf{q}(t) \} \right\} \quad (5.3)$$

where  $\Phi_{\varepsilon p} \in \mathbb{R}^{n_p \times n}$  are strain mode shapes of the  $n$  considered modes at the  $n_p$  DOFs which correspond to the virtual sensor locations  $p$ ,  $\mathbf{q}(t) = \{q_1(t), q_2(t), \dots, q_n(t)\}^T$  are the modal coordinates that quantify the participation of each mode and  $\mathcal{L}\{\bullet\}$ ,  $\mathcal{L}^{-1}\{\bullet\}$  are the Laplace and inverse Laplace operations respectively.

The dynamic frequency band is subdivided in two parts as seen in Figure 5.5. The first part captures the lower frequency turbine dynamics including waves and the first structural mode ( $0, 2Hz$  up to  $0, 5Hz$ ) and the second part (higher than  $0, 5Hz$ ) captures all the remaining dynamics and modal behavior of the structure. The reason for this subdivision is the benefit from the optimal use of the best performing sensors in each dynamic frequency band. For example, sensors near the top of the turbine are very valuable for assessing the first order motion, while sensors closer to the bottom are nearly insensitive to this first order motion and they barely measure above the noise floor. For higher modes in the high-frequent dynamic band multiple modes and multiple sensors are necessary in order to optimally capture the dynamics.

Both the quasi-static and the dynamic strain contribute to the fatigue of an offshore wind turbine. Therefore a superposition of both quasi-static and dynamic contributions is made, leading to a prediction of the entire strain time history. This is called multi-band virtual sensing and is summarized by Equation (5.4).

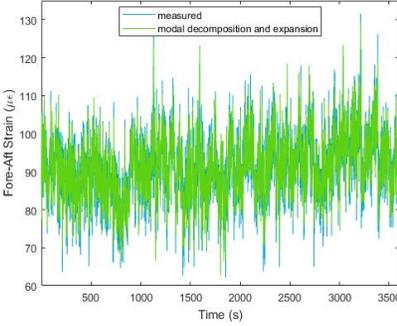
$$\varepsilon_{\mathbf{p}}(t) = \varepsilon_{\mathbf{p}}^{\mathbf{QS}}(t) + \varepsilon_{\mathbf{p}}^{\mathbf{D}}(t) \quad (5.4)$$

One of the drawbacks of this approach is the discontinuity in the frequency spectrum, since the signal is split in three frequency bands.

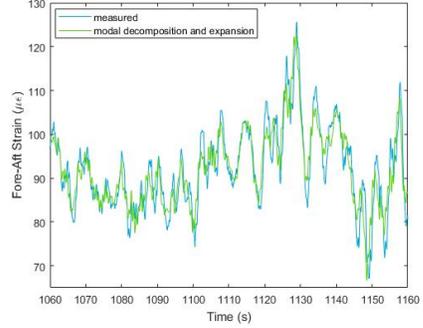
### 5.3.2 Results using strain gauges

The modal decomposition and expansion technique has been applied on measurements done during the second measurement campaign. Figure 5.7 shows the measured and predicted strain signal for one hour of measurements. In general,

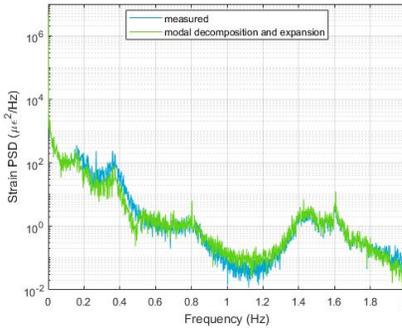
the predicted signals by both techniques match the measurements good in time domain. This is also represented by a fairly low value of mean absolute error between measurement and prediction of  $2,73\mu\epsilon$ .



(a) Time series of 1 hour. A mean absolute error between measured and predicted signal of  $2,73\mu\epsilon$  is found.



(b) Zoom of 100 s



(c) Frequency spectrum

Figure 5.7: Multi-band virtual sensing based on modal decomposition and expansion validated for a period of 1 hour using strain data as input. The blue line represents the actual measured signal and the green line the predicted signal.

In the frequency domain the technique can estimate quasi-static strain signals up to  $0,2Hz$  almost perfectly, since the thrust signal used was calculated from the measured strain signal. For higher frequencies the differences are clearly higher but still the signals seem to match pretty well, although the first mode is slightly underestimated.

Moreover, a drop can be observed around  $0,2Hz$  and more clearly around  $0,5Hz$ . These frequencies coincide with the limits at which the signal is divided to determine the different contributions to the resulting strain signal (quasi-static, low frequent and high frequent). This means the accuracy of the predicted signal reduces at the chosen limits in frequency for the different frequency bands. At those frequencies not only one of the possible loads dominates the response but

an interaction between multiple loads results in the actual response. However, the prediction only captures one load, the one presumably dominating the response.

### 5.3.3 Results using a SCADA-based thrust model

The technique was also applied on data from the first measurement campaign, in combination with a preliminary thrust model for a site with relatively poor quality SCADA data. Here, a thrust model as explained in Section 4.7 is trained using 10 minute averages and applied on 1s SCADA data. Table 5.1 shows the input parameters selected for every operational state.

Table 5.1: The different input parameter sets used for the thrust model as trained to proof the concept of Virtual Sensing with a SCADA-based thrust model and modal decomposition and expansion.

	Non-producing	Below rated power	Rated power
blade pitch angle	x	△	□
rotor speed	x	△	□
wind speed	x		
produced power		△	

In Figure 5.8(a-b) indicative results of SCADA-driven quasi-static strain response  $\varepsilon_{\mathbf{p}}^{\text{QS}}$  and dynamic strain response  $\varepsilon_{\mathbf{p}}^{\text{D}}$  are shown. Moreover, a superposition of these contributions ( $\varepsilon_{\mathbf{p}}$ ) is also given in Figure 5.8(c). These results are obtained for an operating offshore wind turbine at rated speed. It is nicely observed that the predicted strains adequately match the measured strains both in the sub-bands and in the entire band of interest.

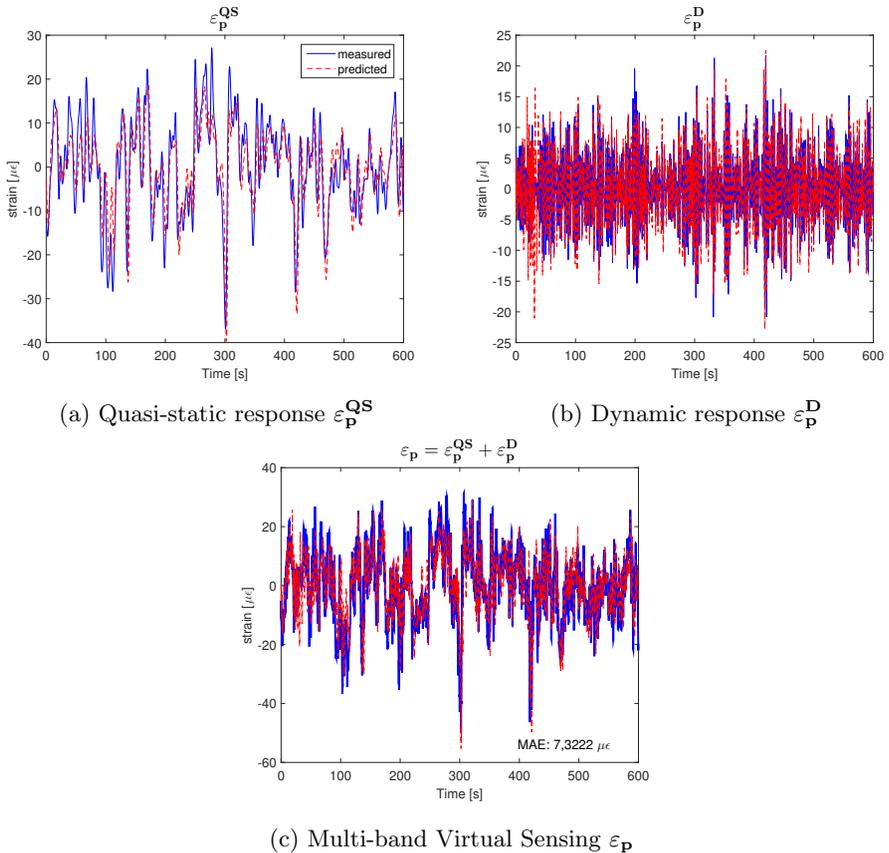


Figure 5.8: SCADA-driven quasi-static strain response and acceleration-driven dynamic strain response superimposed resulting in the so-called multi-band virtual sensing. The mean absolute error obtained for this example is shown in (c). The blue full line is the actual measured signal and the red dashed line is the predicted signal with the proposed technique. The example dataset corresponds to normal operating condition at rated rotor speed of the offshore wind turbine.

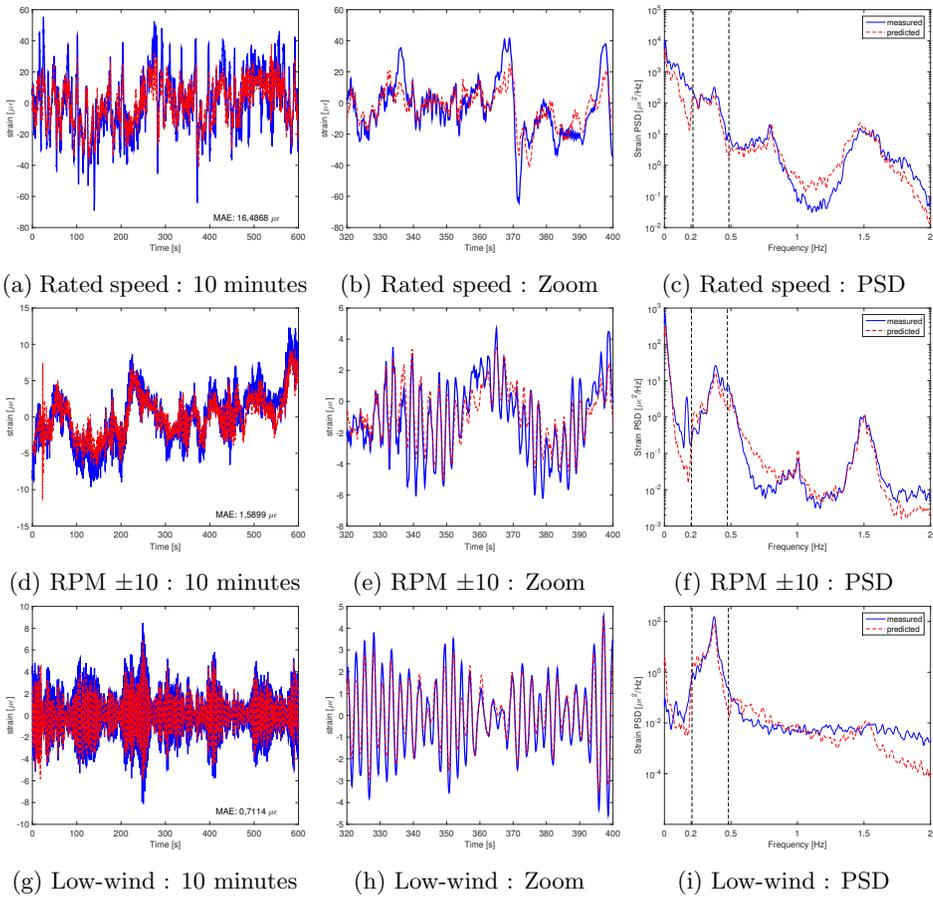


Figure 5.9: Multi-band virtual sensing validated for a variety of operational cases. The blue full line represents the actual measured signal and the red dashed line the predicted signal with the proposed technique. The mean absolute error for each example is shown in (a,d,g).

The performance of the technique is also demonstrated by a couple of examples in Figure 5.9, where various operational states are represented. A good match seems to be found between the predicted and the measured strains, both in terms of amplitude and in terms of temporal evolution. This is reflected in both time and frequency domain. In case of the two examples shown first (a-c) and (d-f), an important contribution of the quasi-static thrust load can be observed. This is also reflected in the frequency domain, as the spectral density for frequencies lower than  $0,2Hz$  is comparable to the spectral density of the first mode. However for the third example (g-i), the thrust load contributes very little. This indicates that for low winds, the dynamics are dominant over quasi static load, whereas for higher winds the thrust load will have a contribution of equal importance. The same observation can be made in Figure 5.5. Moreover, the frequency spectrum for frequencies higher than  $0,2Hz$  (i.e. the dynamic part) indicates a dominance of the first mode. This justifies the use of only one mode to estimate the dynamic behavior of the turbine.

The mean absolute error (MAE) is calculated for the given examples. The values for MAE are promising for the second and third example (d,g),  $1,5899\mu\epsilon$  and  $0,7114\mu\epsilon$  respectively, but an improvement is needed based on the values obtained for the first example (a) and the example shown in Figure 5.8,  $16,4868\mu\epsilon$  and  $7,3222\mu\epsilon$  respectively, both representing production at rated speed. The frequency spectra indicate the importance of the quasi-static part is higher in this case than for non-producing or producing below rated speed. Therefore it's most likely an improvement of the thrust model, as explained in Section 4.4 and trained with 1s SCADA, will influence the result considerably.

## 5.4 Kalman filter

The work presented in this section is a collaboration with Konstantinos Tatsis and Eleni Chatzi and is published in a conference paper ([20]).

An alternative course for tackling the problem of fatigue estimation on the basis of a limited number of vibration sensors consists in fusing the available measurements with a Kalman filter for extrapolating the response at unmeasured locations. The main advantage of this approach is the continuity in the frequency domain. There's no need to split the signals into frequency ranges. Another potential advantage in using the Kalman filter is the fact the filter will attribute uncertainty to the measurements. As such the filter will ignore values of the modeled thrust load, caused by bad SCADA data, and the acceleration data that are unrealistic given the underlying physical model of the turbine. Such a behavior does not exist in the MDE approach as both contributions are directly combined (Equation. 5.4).

### 5.4.1 Concept

The starting point for implementing a Kalman filter towards state estimation on wind turbines, is the continuous-time linear system of dynamic equations of motion, which is transformed to a discrete-time modally reduced state-space system.

Such an approach is widely reported in the literature and herein it is only briefly elaborated since the entire formulation and notation is adopted from [85]. In this sense, the system and measurement equations may be written as

$$\zeta_{k+1} = \mathbf{A}\zeta_k + \mathbf{B}\mathbf{p}_k + \mathbf{w}_k \quad (5.5)$$

$$\mathbf{y}_k = \mathbf{G}\zeta_k + \mathbf{J}\mathbf{p}_k + \mathbf{v}_k \quad (5.6)$$

where  $\zeta_k \in \mathbb{R}^{n_s}$  is the state vector containing modal displacements and velocities,  $\mathbf{y}_k \in \mathbb{R}^{n_y}$  is the output vector,  $\mathbf{p}_k \in \mathbb{R}^{n_p}$  is the input force vector and  $\mathbf{w}_k \in \mathbb{R}^{n_s}$  along with  $\mathbf{v}_k \in \mathbb{R}^{n_p}$  are zero-mean white processes, with covariance matrices  $\mathbf{Q} \in \mathbb{R}^{n_s \times n_s}$  and  $\mathbf{R} \in \mathbb{R}^{n_y \times n_y}$ , representing the system and measurement noise, respectively. Finally,  $\mathbf{A} \in \mathbb{R}^{n_s \times n_s}$  and  $\mathbf{B} \in \mathbb{R}^{n_s \times n_p}$  are the system matrices while  $\mathbf{G} \in \mathbb{R}^{n_y \times n_s}$  and  $\mathbf{J} \in \mathbb{R}^{n_y \times n_p}$  are the output and feedthrough matrices. These matrices are based on a modal model, composed of the first six, three fore-aft and three side-to-side, vibration modes. These modes are obtained from a simple Finite Element (FE) model which is tuned in order to be in accordance with the identified modal properties, i.e. frequencies, damping ratios and mode shapes.

In the absence of information with respect to the driving forces, the state of the system may be augmented with the input vector, so that  $\zeta_k^a = \text{vec}([\zeta_k \ \mathbf{p}_k]) \in \mathbb{R}^{n_s+n_p}$ , in order to form the so-called augmented state-space model

$$\zeta_{k+1}^a = \mathbf{A}^a \zeta_k^a + \mathbf{w}_k^a \quad (5.7)$$

$$\mathbf{y}_k = \mathbf{G}^a \zeta_k^a + \mathbf{v}_k \quad (5.8)$$

where superscript  $a$  designates the augmented quantities. By doing so, the evolution of input is dictated by the augmented system matrix  $\mathbf{A}^a \in \mathbb{R}^{(n_s+n_p) \times (n_s+n_p)}$ , whereby it is postulated that the input can be captured by a random-walk process.

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \boldsymbol{\eta}_k \quad (5.9)$$

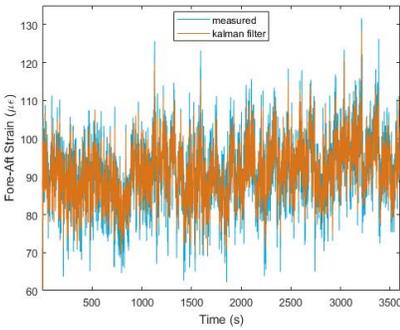
with  $\boldsymbol{\eta}_k$  being a zero-mean white Gaussian process with covariance matrix  $\mathbf{S} \in \mathbb{R}^{n_p \times n_p}$ . Within this context, both input and state may be estimated recursively through the standard Kalman filter operating on the augmented state-space model.

The measured response quantities of the considered turbine comprise accelerations at four different elevations as well as the thrust force. The accelerations are measured with a sampling rate of  $20Hz$  and lack information on the quasi-static loads. This lack of information is captured by the thrust force, measured with a sampling rate of  $1Hz$ . Due to this difference in the sampling rates, the above-described filter is employed in a multi-rate fashion which is materialized with the use of time-varying measurement noise. Considering that the lack of measurements is equivalent to optimal filtering with large measurement errors [86] and hence zero gain, the measurement noise corresponding to the thrust is set to an arbitrarily large value when thrust measurements are not available and it is reset to the tuned value as soon as thrust is measured again.

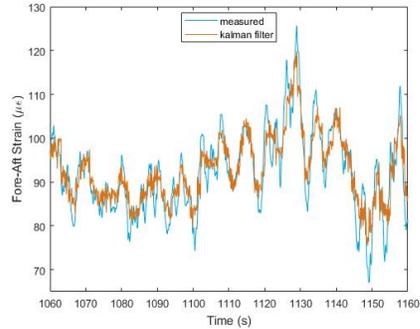
The estimation is subsequently performed assuming that the dynamics of the turbine are driven by the thrust force, applied on the tower top, and an equivalent wave force exerted at the hydrodynamic center. In contrast with the modal decomposition and expansion technique, the estimation using the Kalman filter does not assume any explicit distinction between quasi-static, low frequency and high frequency regimes. Instead, the quasi-static part of the response is captured by the thrust force, sampled at a rate of 1Hz, and the higher frequency dynamics are dictated by the acceleration measurements.

### 5.4.2 Results

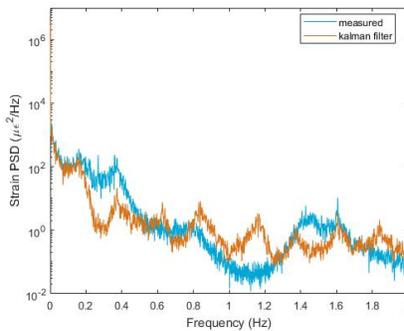
Figure 5.7 shows the measured and predicted strain signal for the technique based on the Kalman filter.



(a) Time series of 1 hour. A mean absolute error between measured and predicted signal of  $2,89\mu\epsilon$  is found.



(b) Zoom of 100 s



(c) Frequency spectrum

Figure 5.10: Multi-band virtual sensing based on a Kalman filter validated for a period of 1 hour. The blue line represents the actual measured signal and the red line the predicted signal.

As seen in Figure 5.10, the measurement setup based on the Kalman filter

provides sufficiently good response estimation, which additionally does not suffer from instability issues. In terms of time domain, a similar error is found for the Kalman-based method as for the MDE-based method (shown in Figures 5.10a and 5.7a respectively). However, in terms of the frequency domain of the strain signal the match is not as good as for the results obtained using modal decomposition and expansion (Figures 5.10c and 5.7c respectively).

It should be underlined that despite the straight-forward way of fusing different types of measurements, tuning of the covariance matrices is always the key feature in obtaining an optimal estimation using Kalman-type filters. This is herein achieved by firstly adjusting the system and measurement covariances, so that the predicted accelerations show good agreement with the measured ones. This is highlighted in Figure 5.11, through the frequency domain representation of the two signals, which are well matching. Finally, once the two covariance matrices are adjusted, the input process is tuned using the L-curve.

Although the frequency spectrum of measured and predicted acceleration signal is well matching (Figure 5.11), this didn't result in a similar match between both strain signals. The predicted strain signal by the Kalman filter is based on acceleration measurements at the lower three levels. The top level acceleration signal is dominated by the first mode and therefore out of phase to the other acceleration signals. Such an effect is heuristically seen to work at the expense of the stress estimates and as result the top acceleration was not used in the filter.

Another possible explanation for the mismatch in frequency domain is the lower measurement frequency of the thrust signal. This causes the prediction to be updated at some intermediate timestamps and results in big jumps in the predicted signal, as shown in Figure 5.12.

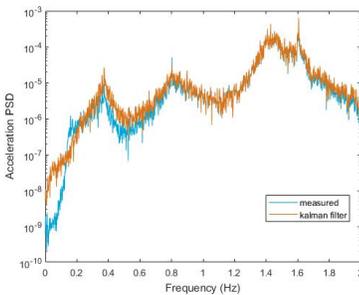


Figure 5.11: Frequency spectrum of the predicted and measured accelerations at the level of 19,0m using the Kalman filter.

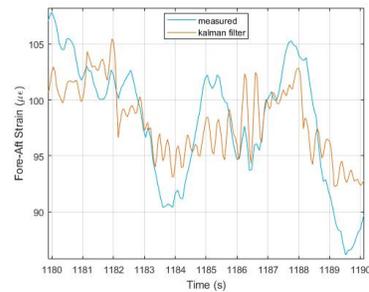


Figure 5.12: A drop in predicted strain is observed around 1182 s due to the discontinuous updating of the thrust force signal.

## 5.5 Conclusion

In this chapter two possible techniques to reconstruct a strain history at fatigue critical locations of monopile foundations of offshore wind turbines were compared. Both techniques combine measured accelerations and a measured or modeled thrust load signal. The first technique is based on modal decomposition and expansion, the second is based on a Kalman filter. For both techniques a good match is obtained in time domain. However in frequency domain, the technique based on modal decomposition and expansion showed better results than the Kalman filter. Although the technique based on the Kalman filter has some advantages over the technique based on modal decomposition and expansion, improvements are still needed to obtain similar results.



## Chapter 6

# Lifetime assessment of an instrumented turbine

*Performing a lifetime assessment for a turbine basically means giving an accurate estimate of when the turbine's substructure likely will fail. In this chapter, a methodology is presented to translate a measured stress signal at an accessible location in the substructure of the OWT into damage and eventually into an expected lifetime for multiple potentially fatigue hotspots of the structure. This methodology consists of multiple steps. In a first step, the stress signal is transformed to a fatigue damage. This is explained in Section 6.2. The actions needed for this first step can be divided into three parts: rainflow counting of the stress signal, the introduction of several stress correction factors and the application of the Miner's rule for damage calculation.*

*Section 6.3 elaborates on the second step in the process, being how to calculate the expected lifetime from damage measurements. For this second step, a lot of choices have to be made. The possibilities and their consequences are explained.*

*Moreover, the proposed methodology was applied on a fictional OWT using realistic measured stress signals. This case study is extensively explained in Section 6.4.*

*Finally Section 6.5 concludes the chapter.*

## 6.1 Introduction

Being able to reconstruct the stress history at any location in the structure of an offshore wind turbine is of big importance for a fatigue analysis. But the only number a wind farm operator is truly interested in, is the (remaining) lifetime of all (critical) locations in the structure. In this chapter the necessary steps to translate a stress history, measured directly or reconstructed using e.g. the techniques disclosed in Chapter 5, and the available SCADA data into a realistic life time assessment are summarized.

The chapter will touch upon following topics:

- From stress history to fatigue damage
  - Fatigue spectra
  - The different stress concentration factors
  - S-N curves and Miner's rule
- From damage history to a fatigue life time estimate

The proposed strategies are applied to a case study of a fictional offshore wind farm.

## 6.2 From stress history to fatigue damage

The first step to a full fatigue assessment of an offshore wind farm is the translation of a stress history into a damage estimate. Either a deterministic or a probabilistic approach can be followed here. A deterministic damage estimate quantifies how distant a fatigue failure is. The found fatigue damage is a number ranging from 0 to 1, where zero implies no fatigue damage and 1 implies failure and consequently end of life. A probabilistic damage estimate returns a likelihood the structure failed due to fatigue. End of life is reached when the likelihood of failure has reached a certain upper limit, e.g. 2,3% as the design S-N curves suggest [87].

In the current research a deterministic strategy is followed, as this is still the most common practice in industry and design. The steps required to translate a stress history into a life time estimate are summarized in Figure 6.1. The first steps from strain measurement to stresses in FA and SS direction are discussed in Section 2.2.2. In this case the measured stresses in compass directions are solely converted to stresses in FA and SS direction, independent of the exact headings. However, in theory a fatigue assessment should be performed for each heading (sector) individually. It is believed by reducing this to only FA and SS stress signals and taking the most damaging of both, a conservative approach is followed. Each step illustrated in Figure 6.1 starting from the stresses in Fore Aft and Side-Side direction will be discussed in the following sections.

### 6.2.1 Cycle counting

The state-of-the-industry technique to calculate damage, discussed later in Section 6.2.4, is based on the number of cycles combined with the stress range (or

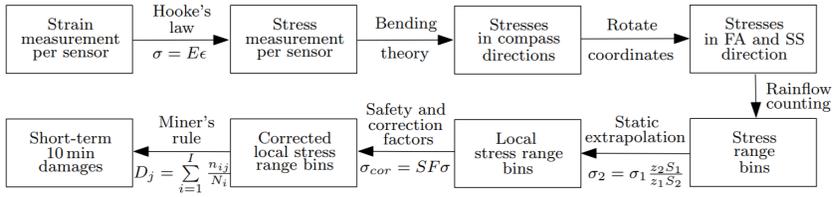


Figure 6.1: A simplified overview of the steps required to go from stress histories to a fatigue damage estimate.[88]

amplitude) of these cycles. Therefore, the measured (or extrapolated) stress signal should be reduced to a histogram, representing the number of cycles in predefined stress ranges. A well-known method for cycle counting in literature is so-called rainflow counting.

The principle is explained in [89] and illustrated by Figure 6.2. When plotting the stress or strain signal on the x-axis and the time on the y-axis downwards, an imaginary raindrop starts at the beginning of the signal and at the inside of every peak and can fall down the signal to lower roofs. Half a cycle is counted every time one of the following conditions is met:

- the raindrop comes opposite a maximum (minimum) higher (lower) than where it started from. In Figure 6.2 the raindrop started at point 1, stops at the opposite of point 5, since the strain value at point 5 is lower than the value at point 1.
- the raindrop meets another raindrop from the roof above. In Figure 6.2 the raindrop started at point 3 and stops at point 2', since at that point the raindrop that started at point 1 fell down.

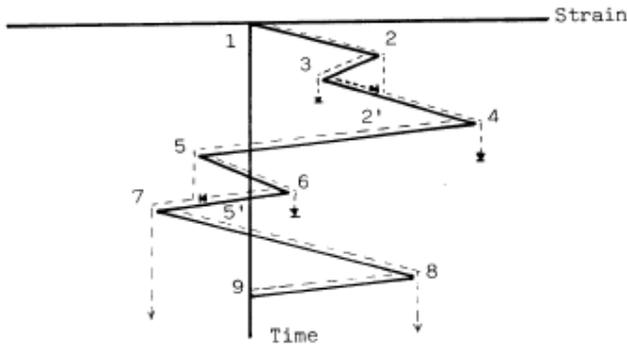


Figure 6.2: An example to illustrate the concept of rainflow counting [89]

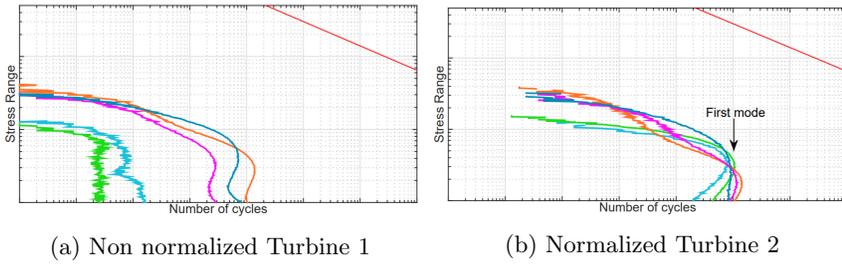


Figure 6.3: Exemplary fatigue spectra obtained after cycle counting the stress histories of strain measurements at offshore wind turbines. The different colors represent the different operational cases of the turbines. The normalization in (b) provides the equivalent fatigue spectra when each case is present the same amount of time. [91]

For this thesis, an existing toolbox WAFO [90], consisting of MATLAB code, was used to perform the rainflow counting. The rainflow counting is performed on the stress signal, both for fore-aft direction and side-side direction, based on the turning points in the signal.

All measured stress cycles are distributed over 500 predefined bins, where the bin centers were logarithmically spaced between 10 kPa and 1 GPa. The resulting histogram is also referred to as a fatigue spectrum, as it represents the different contributors to the final fatigue life.

Unique to the application in this research is that the measured fatigue spectra are stored for every ten-minute timestamp. As such it is possible to link each individual fatigue spectrum to the current condition of the turbine as is reflected by the SCADA data. In Figure 6.3 two examples of the resulting fatigue spectra are plotted with different colors indicating the different operational cases of the turbine. The results reveal a considerable difference in the behavior of the turbine, and thus the way it accumulates fatigue. E.g. in non-operational conditions (blue and green) the number of very large cycles is very low compared to operational cases (all other colors) as the thrust loading, responsible for the slow but large cycles (see also Chapter 4), is negligible in non-operational conditions.

### 6.2.2 Stress extrapolation within the structure

A fatigue assessment of a wind turbine should include multiple locations within the structure. Most valuable are the specific locations which are, according to design, most fatigue sensitive. However, other locations could be more affected by fatigue during the lifetime than expected. Unfortunately, mounting sensors at all these locations is not cost-effective and often physically impossible or unfeasible. Therefore, the first step in a full lifetime assessment for a wind turbine is the extrapolation of measurements done at easily accessible locations in the structure. This extrapolation can be done using the techniques described in Chapter 5. However, as the research in this chapter ran in parallel to development of virtual

sensing techniques a simplified but less accurate method was used in this chapter. The stress signal observed at the sensor level is assumed to be the result of one static load applied on the nacelle of the wind turbine ( $F_{hub}$ ). In this case, the bending moment ( $M_{meas}$ ) observed at the sensor location with a distance  $z_{sensor}$  from the hub can be extrapolated along the length of the structure using static extrapolation as given by Equation 6.1, where  $M_{ext}$  represents the bending moment at the location of interest with distance  $z_{ext}$  to the hub.

$$F_{hub} = \frac{M_{meas}}{z_{sensor}} = \frac{M_{ext}}{z_{ext}} \quad (6.1)$$

The bending moment and stress are related to each other by Equation 6.2. Here, the area moment of inertia  $I_c$  and the radius of where the stress in the section is needed  $R$  are used. Knowing this relation, a factor (Stress Extrapolation Factor) can be defined by which the measured stress signal should be multiplied to obtain the extrapolated stress signal. This is given in Equation 6.3.

$$M = \sigma \frac{I_c}{R} \quad (6.2)$$

$$\begin{aligned} SEF &= \frac{\sigma_{ext}}{\sigma_{meas}} \\ &= \frac{M_{ext}}{M_{meas}} \frac{R_{ext}}{I_{c,ext}} \frac{I_{c,sensor}}{R_{i,sensor}} \\ &= \frac{z_{ext}}{z_{sensor}} \frac{R_{ext}}{R_{i,sensor}} \frac{I_{c,sensor}}{I_{c,ext}} \end{aligned} \quad (6.3)$$

If available, one could choose to use an extrapolation factor based on the as designed distribution of the bending moment along the structure instead of assuming a static extrapolation.

### 6.2.3 Safety and correction factors

The entire substructure of a wind turbine consists of a large number of specific weld details that influence the local stress behavior and ultimately the fatigue life. Rather than an in-depth assessment of each detail the current practice is to use correction factors to account for the particularities of each weld detail. In the following sections an overview of several relevant parameters for offshore substructures on monopiles is provided. As well as a short discussion on the developed framework to automatically generate the factors for any given geometry of monopile.

The final outcome of all factors is the so-called Combined Safety factor (CSF). As the safety factors differ for each weld detail the fatigue life of each detail is greatly influenced by these CSF. Therefore the stress range associated to each detail is multiplied with the corresponding CSF.

### Size Effect (SE)

The first safety factor serves to account for the difference in thickness of plates between the test specimen used in the material fatigue testing and the actual component. According to [92] the measured stress should be multiplied by SE given in Equation 6.4, if the thickness of the actual component is larger than the reference thickness  $t_{ref}$  of 25mm. In this equation the SE is provided from the actual thickness  $t$ , through which a crack most likely will grow, and the thickness exponent on fatigue strength  $k$ , given by the type of S-N curve.

$$SE = \left( \frac{t}{t_{ref}} \right)^k \quad (6.4)$$

In [93] an addition was made. In case of butt welds, often used in the manufacturing of the substructure of offshore wind turbines, the actual thickness  $t$  can be replaced by an effective thickness  $t_{eff}$ . The effective thickness can be calculated using the weld width  $L_t$  and the minimum wall thickness of both adjacent sections  $t_{min}$ . Both of them should be expressed in mm. The exact values for the weld width are specified for every weld in the design documents.

$$t_{eff} = \min(14 + 0, 66L_t, t) \quad (6.5)$$

For applications in offshore wind SE is typically larger than 1. As the wall thicknesses used in offshore wind turbines are typically well above 25mm. The SE larger than one reflects that due to more likely inclusions of defects in the thicker walls fatigue is more likely to progress in thicker walls. The introduction of the effective thickness, Equation. 6.5, resulted in a reduction of the SE.

In general the specific values of SE taken for the different welds of an OWT substructure are specified in design documents. In case of a weld between two sections with a different wall thickness, the smallest wall thickness is chosen to calculate SE.

### Stress Concentration Factor (SCF)

Changes in geometry or welded attachments can cause local stress hot spots in the structure. Often fatigue cracks can initiate or propagate more easily due to these local stress augmentation. To account for this additional local stress, the measured (nominal) stress is multiplied with a stress concentration factor (SCF). Depending on type of transition in the structure, the calculation of SCF can be done analytically or obtained from a finite element analysis. For a thickness transition or a conical section, equations are suggested by [92]. The values for SCF applied for every detail during design, obtained analytically or using a finite element model, are typically also specified in the design documents. When available the values in the design document are preferred for a fatigue assessment.

#### SCF Thickness transition

When the wall thickness of two adjacent tubular sections is not the same, a thick-

ness transition is needed. This is illustrated by Figure 6.4. In general the transition is made at the inside of the monopile, as opposed to the illustration.

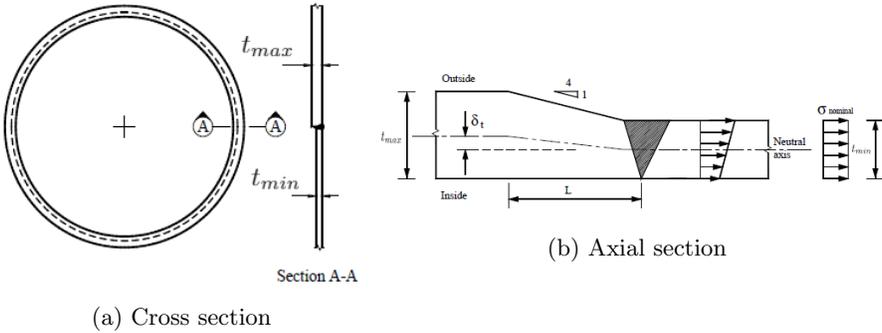


Figure 6.4: Illustration of a thickness transition between two tubular sections with a different thickness (modified) [92]

In this case, Equation 6.6 can be used to calculate the resulting stress concentration factor. In this equation  $t_{min}$  and  $t_{max}$  represent respectively the smallest and highest wall thickness,  $R_o$  the outer radius of the section,  $\delta_t$  half of the difference in thickness ( $\delta_t = \frac{t_{max} - t_{min}}{2}$ ),  $\delta_m$  a misalignment difference,  $\delta_0$  the misalignment inherent in the S-N data and  $L$  the length over which the transition takes place. Often  $\delta_m$  is taken equal to  $\delta_0$ , being 10% of  $t$ . However, newer design take into account a misalignment difference of  $3mm$ . This is specified in design documents. Moreover, a thickness transition is usually performed at a ratio of 1:4. This means  $L = 4 \cdot (t_{max} - t_{min})$ . This ratio is also specified in design documents. In practice, sometimes Equation 6.6 is simplified by taking 0 for  $\alpha$  and 1,5 for  $\beta$ . Both simplifications usually result in a higher value of SCF and thus lead to additional conservatism in the results.

$$SCF = 1 + \frac{6(\delta_t + \delta_m - \delta_0)}{t} \frac{1}{1 + \left(\frac{T}{t}\right)^\beta} e^{-\alpha}$$

$$\alpha = \frac{1,82L}{\sqrt{2}R_o t} \frac{1}{1 + \left(\frac{T}{t}\right)^\beta}$$

$$\beta = 1,5 - \frac{1}{\log\left(\frac{2R_o}{t}\right)} + \frac{3}{\left(\log\left(\frac{2R_o}{t}\right)\right)^2}$$

(6.6)

### SCF Conical section

In a conical section (Figure 6.5) additional stresses are induced at any transition between two sections with a different angle. These additional stresses are caused by a higher concentration of stress flow lines at the inner side of a bend in the material. This means, for a conical section, the additional stresses are at either the

inner wall or the outer wall, depending on the upper or lower part of the conical section. In Figure 6.5, the stress for the upper bend is concentrated at the outside wall (where the inner side of the bend is located), while the stress for the lower bend is concentrated at the inside wall.

In this case the stress concentration factor can be calculated using Equation 6.7, where  $t$  represents the wall thickness and  $R_o$  the outer of the tubular adjacent tubular section. Here, the wall thickness is assumed uniform over the conical section and adjacent tubular sections, as it is usually the case for conical sections in offshore wind turbine substructures. Using the simple rules of trigonometry, the height of the section  $H$  and the difference in radius  $\Delta R_o$ ,  $\tan\alpha$  is calculated (Equation 6.8).

$$SCF = 1 + \frac{0,6t\sqrt{4R_o}}{t^2}\tan\alpha \quad (6.7)$$

$$\tan\alpha = \frac{\Delta R_o}{H} \quad (6.8)$$

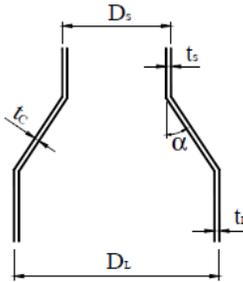


Figure 6.5: Illustration of a conical section in a monopile [92]

### Material Safety Factor (MSF)

[94] suggests to use a material factor  $\gamma_m$  or material safety factor to account for possible unfavorable deviations, uncertainties or differences between the actual structure and the test specimen with respect to the resistance of the material. In practice the value of MSF is determined during design based on the chosen inspection strategy. When a value of 1 is chosen, checks for fatigue cracks are required every 7 years for a designed fatigue life of 20 years. For a value of 1,15 and a designed fatigue life of 20 years, checks are needed every 13 years. In general, a value of 1,25 is chosen, corresponding to no checks at all during the designed lifetime of 20 years. The exact value of the material safety factor is specified for every detail in the design documents of the monopile and transition piece.

### Combined Safety Factor (CSF)

In the end the stress bin centers of the stress histogram are multiplied by the combined safety factor. The combined safety factor consists of the multiplications of all safety factors specified before. Equation 6.9 summarizes this.

$$CSF = SE \cdot SCF \cdot MSF \quad (6.9)$$

If the stress extrapolation within the structure is done using a factor, this factor can be included in the combined safety factor as well.

### 6.2.4 Damage calculation

The state-of-the-industry damage calculation, as suggested by [92], is based on the Palmgren-Miner rule, assuming linear cumulative damage. In this section we will explain the Palmgren-Miner rule.

#### Using Palmgren-Miner Rule

From the Palmgren-Miner rule the accumulated fatigue damage  $D$  is obtained as the sum of the ratios of occurred cycles  $n_i$  over cycles to failure  $N_i$  for all occurred stress ranges  $\Delta\sigma_i$ . This is given by Equation 6.10.

$$D = \sum_{i=1}^k \frac{n_i}{N_i} \quad (6.10)$$

The number of cycles to failure for a given stress range is defined by a S-N curve. A S-N curve represents how many cycles  $N$  of a constant stress range  $\Delta\sigma$  a specimen can hold before failure. Examples of design S-N curves are given in Figure 6.6

The choice for S-N curve is done based on, among others, the considered material, the fabrication method and the expected inspection of the detail. Moreover a distinction is made for the environment around each weld detail. Examples of such environments are by air or water, with or without cathodic protection. Because of the dependency to the environment, the specific S-N curve can differ from one detail to another, even though they are geometrically identical. All the specific S-N curves are specified in the design documents for every detail of the structure. And in a correct fatigue assessment the lifetime of each weld needs to be calculated using the corresponding S-N curve.

A basic S-N curve can be represented by Equation 6.11, where  $m$  represents the negative inverse slope of the S-N curve in logarithmic axes and  $\bar{a}$  the intercept of the design S-N curve with the N axis.

$$\log N = \log \bar{a} - m \log \Delta\sigma \quad (6.11)$$

For offshore applications bilinear S-N curves, as shown in Figure 6.6, are often used. Here, the slope of the curve changes starting from a specified number of

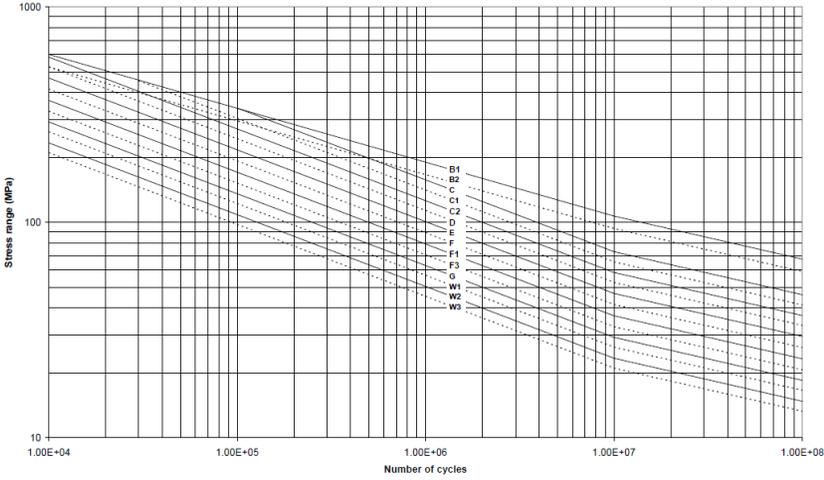


Figure 6.6: Illustration of several design S-N curves in air [87]

cycles, e.g.  $10^7$ , as represented by Equation 6.12. Here,  $\bar{a}_1$  and  $m_1$  are always defined, together with either  $\bar{a}_2$  or  $m_2$  or both.

$$\log N = \begin{cases} \log \bar{a}_1 - m_1 \log \Delta \sigma & \text{if } N \leq 10^7 \\ \log \bar{a}_2 - m_2 \log \Delta \sigma & \text{if } N > 10^7 \end{cases} \quad (6.12)$$

In case of bilinear S-N curves, the stress range associated with  $N = 10^7$  is calculated first. If necessary, the value for  $\bar{a}_2$  or  $m_2$  is calculated from the other known parameters.

Afterwards, the number of critical cycles  $N$  is calculated for every bin of stress cycles using Equation 6.12 together with the known parameters and the value for the stress range, being the stress bin center.

From the selected S-N curve the damage estimate  $D$  is calculated using Equation 6.10. In a deterministic fatigue assessment it is considered that a turbine has reached end-of-life when  $D$  equals one. This has led to the common misconception that  $D$  equal to one implies the structure will collapse under fatigue. However, the correct interpretation is, given the way S-N curves are established, that the turbine has a 97,7% probability to survive as long as  $D$  does not exceed one. This nuance makes it therefore perfectly possible to accumulate to a  $D$  well above 1 before actual failure. In a probabilistic fatigue assessment the probabilistic nature of the S-N curve is considered.

S-N curves are defined based on fatigue tests on small-scale specimens in laboratories. And the application of these S-N curves to structures the size of offshore wind turbines has led to the introduction of the correction factors discussed in Section. 6.2.3. However, a growing community is accepting that the original S-N curves are outdated for both the scale of offshore wind and state-of-the-art fabrica-

tion methodologies that are currently being used. Given the importance of fatigue progression in the final cost of the foundation, recent research projects have started to re-evaluate the original S-N curves for monopiles (e.g. SLIC project [95]) and the more complex jacket nodes (e.g. JaCo project [96]).

An alternative strategy is to step away entirely from the S-N curves and the Palmgren-Miner rule and evaluate fatigue life using crack propagation models and the maximum allowable crack size (cfr. the British Standard BS6835).

## 6.3 From damage history to a fatigue life estimate

The deterministic approach to translate these damage estimates into a (residual) lifetime (RUL), is to determine at what point in time the accumulated damage  $D$  reaches the critical value of 1. Therefore the easiest way to perform a lifetime calculation using an actual measurement campaign is simply by extrapolating the measured damage  $D_m$  over time as done in Equation 6.13.

$$RUL = \frac{1}{D_m} \cdot Period_{measured} \quad (6.13)$$

In essence, if the accumulated damage  $D_m$  over a measurement period of one year equals 0,05, then the resulting RUL is 20 years. The inherent assumption in this extrapolation is the continuity of the environmental and operational conditions. Or in other words, by performing the extrapolation as in Equation 6.13 two main assumptions are made. Firstly, one assumes the environmental conditions as seen during the measurement campaign are fully representative for the entire expected lifetime of the wind turbine. The other assumption is basically the same but for the operational conditions of the turbine. This means the percentage of time the turbine was parked or malfunctioning, and the corresponding environmental conditions, during the measurement campaign is the same as during the entire lifetime. A similar assumption is made for start and stop events or de-rating periods. Unfortunately, this assumption only holds if end of life is near and thus if the measurement period is very long, e.g. close to 20 years. In reality, such a long measurement period is not available for offshore wind turbines. Therefore more advanced extrapolation methods are needed.

A possible extrapolation method is proposed in Figure 6.7. This method consists of two main parts: the damage binning and the damage extrapolation, respectively elaborated upon in Section 6.3.1 and Section 6.3.2. For both stages, some choices have to be made. These choices are indicated in diamonds in Figure 6.7. The required data for the next step is indicated in circles and the needed actions are indicated in squares. All steps are explained in more detail in the following two sections.

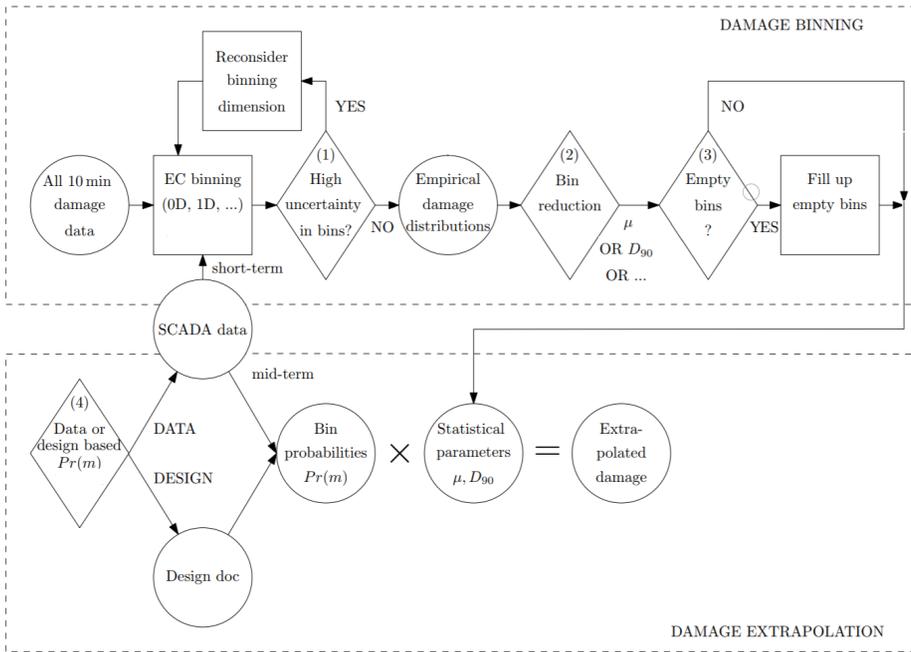


Figure 6.7: A simplified overview to extrapolate fatigue data into a fatigue life estimate, with indication of several choices that are part of the procedure, modified from [88].

### 6.3.1 Construction of the damage table

To be able to perform an extrapolation in time considering possible different conditions, the measured damage is linked with specific conditions. To do this, the damage is calculated as discussed in Section 6.2, for each time frame of 10 minutes. The damage obtained for every 10 minute interval can be linked to environmental conditions, such as (a combination of) wind speed, turbulence intensity or wave period. This is done by binning of the damage data based on the required environmental parameters. Which environmental parameters are of interest is the first choice that has to be made during the process. The final binning of the damage data results in a damage table. However, to calculate a lifetime one value for every bin is required instead of all damages measured during the requested conditions. Here, the appropriate metric has to be chosen to reduce the damage table. The next step consists in filling the empty bins of the damage table. This is the third choice to be made. Once the empty bins are filled, damage measurements related to environmental conditions are obtained. Finally, the behavior of the turbine differs significantly for different operational states. This difference in behavior results in a different fatigue spectrum (cfr. Figure 6.3) and a different damage accumulation as well. Therefore periods of standstill, de-rating or intervals during which a rotor start or stop occurred, can be treated separately.

More information about all required choices is given in the next paragraphs.

### Parameters in the damage table

Several environmental parameters influence the damage accumulation of the substructure of a wind turbine. In general, the obtained values for damage have a wide range. Therefore damage is often hard to interpret on a linear scale. As a workaround, often the Damage Equivalent Load (DEL) is used instead of damage. In essence, the DEL is obtained by assuming a hypothetical load signal with a fixed number of cycles and a fixed amplitude. When applying this load signal on the structure, it would cause exactly the same damage as the measured (more complex) stress or load history. The DEL represents the value of the fixed amplitude needed to obtain the same damage. More information can be found in [97] and [98]. To visualize the correlation between damage and environmental parameters, the DEL might be favored over the measured damage. However, since the resulting damages were often only an intermediate result, the damage is usually plotted on a logarithmic scale instead of the DEL on a linear scale in the remainder of this thesis.

As one can imagine, the substructure of an OWT will fatigue more when wind speeds are higher. This is shown in Figure 6.8 for two different wind turbines.

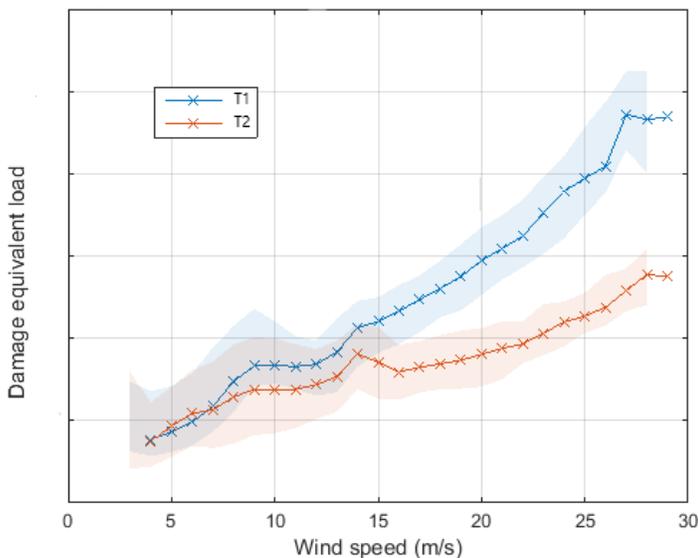


Figure 6.8: The normalized Damage Equivalent Load measured at two OWTs, shown versus wind speed. The DEL clearly increases for an increasing wind speed. The crosses and full line show the median values for different wind speed bins, while the blue and orange zones show the 10th until 90th percentile of the data within the different bins.

In [98], the influence of several parameters on the damage accumulation of sub-structures of OWTs is discussed. Among these parameters, some environmental parameters are considered as well. Here, it is shown the wind direction influences the DEL (Figure 6.9). The wind directions for which a higher damage was obtained can be linked with wind directions for which the considered wind turbine is standing in the wakes of other wind turbines in the farm. The wakes of the other turbines will cause a higher turbulence intensity at the considered turbine. Therefore, another possibility to take wakes into account is to consider the turbulence intensity measured at the turbine instead of the wind direction. By using turbulence intensity instead of wind direction, the damage in and out of wake at one turbine can be compared and extrapolated easier to the damage at another turbine.

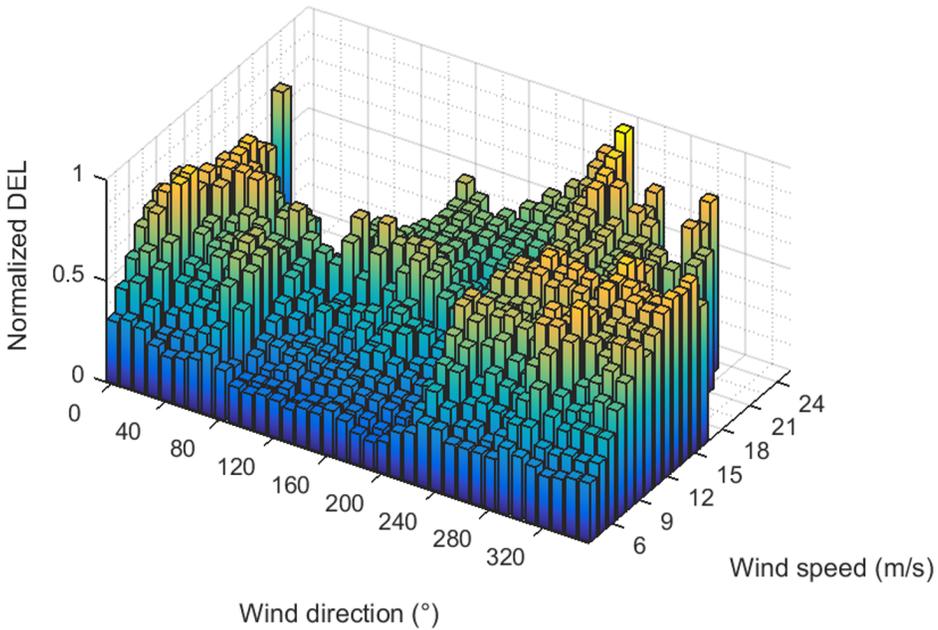


Figure 6.9: The normalized Damage Equivalent Load measured at an OWT, shown versus wind speed and wind direction. For the wind directions between 80 and 200 degrees, no other turbines are causing wake [98].

Moreover, a relation between damage and wave height is shown in Figure 6.10 [98]. In the figure, only data in a small interval of wind speeds is shown, since wave height and wind speed are strongly correlated to each other.

For the existing offshore wind farms, wind speed is often highly correlated to damage. For that reason, it will always be one of the environmental parameters considered in this thesis for any extrapolation based on environmental conditions. As suggested by [34], a standard bin size of  $1\text{ m/s}$  is taken. Only for wind speeds below  $1,5\text{ m/s}$  a bin size of  $1,5\text{ m/s}$  is taken and for wind speeds between  $22,5\text{ m/s}$

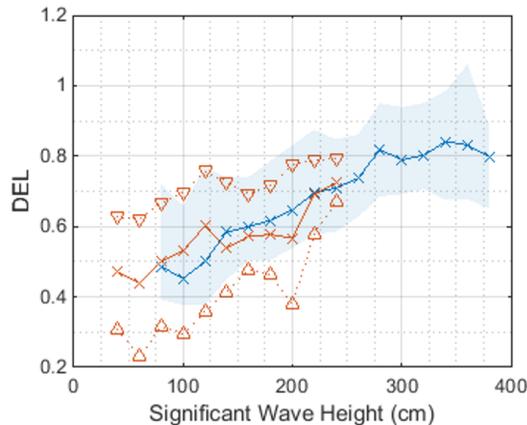


Figure 6.10: The normalized Damage Equivalent Load measured at an OWT, shown versus the significant wave height. The crosses and full line show the median values for different wave height bins, while the blue zone and the orange triangles show the 10th until 90th percentile of the data. Only data for a small interval of wind speed is shown [98].

and  $26,5\text{ m/s}$  the bin size is increased to  $2\text{ m/s}$ . Moreover the highest bin contains all data points for which the wind speed exceeded  $26,5\text{ m/s}$ .

Other parameters of interest are turbulence intensity and wave height, which might be combined with wind speed. These parameters often depend on wind speed themselves, as illustrated in Figure 6.11 by the blue dots. Because of this correlation, it is chosen to define the bin borders for these parameters separately for each wind speed bin. For both parameters, 4 different bins are created for each wind speed bin based on the 25th, 50th and 75th percentile. The lowest bin border is chosen well below the minimum value to include all data points. The highest bin border is chosen well above the maximum value for the same reason. The values of 0 and Inf can be chosen respectively. The bin borders are shown by the green lines in Figure 6.11.

In case of turbulence intensity, not all data points are shown in the figure because of the very high value obtained for low wind speeds. It was chosen to focus on the data points with a higher value since these are far more important in terms of damage and fatigue.

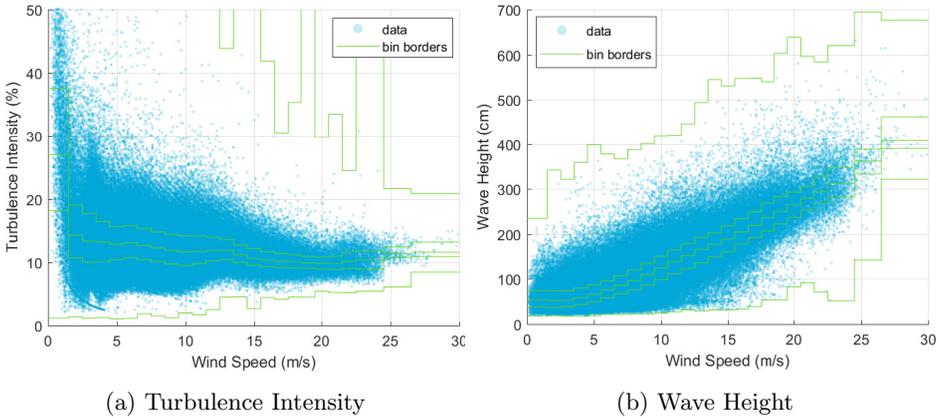


Figure 6.11: Some environmental parameters versus wind speed. The bin borders for these environmental parameters depend on the value of the wind speed, as shown by the green lines.

An illustration of a resulting damage table will be given and discussed upon in Section 6.4.3.

### On reducing the damage table

Up to now, each bin contains all damage values measured during the environmental conditions, as specified by the bin borders. To calculate the lifetime in a deterministic way, a damage table is needed where one damage value  $D_i$  is given for each possible combination of environmental conditions. This damage represents the damage the turbine would accumulate if it would operate an entire year in these environmental conditions. An obvious choice might be to take the mean value of damage measurements in one bin. However, any statistical metric can be used to obtain one value, e.g. the 75th or 90th percentile. These three possibilities are shown in Figure 6.12 by the full yellow, dashed purple and dotted line red line respectively.

It is shown in Figure 6.13 and [88] that for long measurement periods, the results using the 90th percentile are very conservative. Therefore, depending on how conservative the results should be and the length of the measurement period, the appropriate metric is chosen.

For the remainder of this thesis, the mean value was chosen to reduce the damage table.

Should one be interested in a probabilistic fatigue assessment one could also determine the damage distribution per bin and calculate damage accumulation using these distributions instead [99].

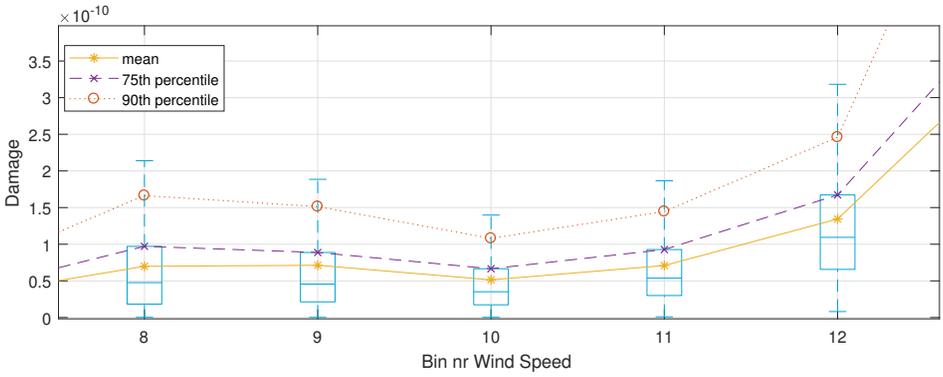


Figure 6.12: Damage distributions in each bin of damage table, composed based on wind speed and turbulence intensity. The results for several wind speed bins (number 8 to 12) but only one turbulence intensity bin are shown. The distributions are indicated using box plots. The mean value, the 75th percentile and the 90th percentile of each bin are indicated by respectively the full yellow, dashed purple and dotted red line.

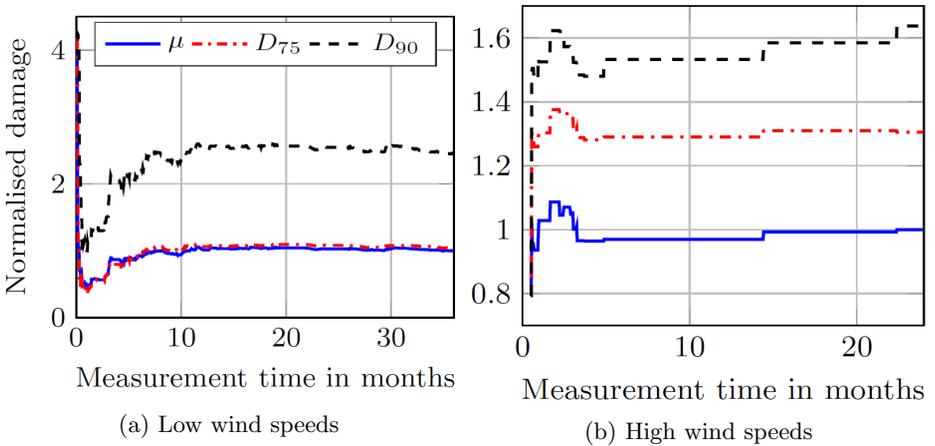


Figure 6.13: Convergence of normalized damage calculated using the mean value ( $\mu$ ), the 75th percentile ( $D_{75}$ ) and the 90th percentile ( $D_{90}$ ) [88]

### Filling empty bins

As explained in previous sections, the damage table can be composed based on several (combinations of) environmental parameters. Ideally the environmental parameters and their bin sizes and limits are chosen in such a way each bin contains enough damage data. A guideline for the minimum amount of damage data needed in each bin to obtain a representative damage table is given by [34]. In reality however it is possible and even very likely some combinations of environmental conditions did not occur during the measurement period. In that case empty bins should be filled with a well considered value. Depending on the availability of the needed information, the filling can be based on design documents or data-driven. If design documents contain information about the load case tables obtained by simulations, those values can be used to fill up the empty bins in the damage table. On the other hand, the empty bins can be filled based on measured data as well. It is the latter option that is executed for this thesis. Each empty bin is filled with the maximum value found in the neighboring bins. This is shown in Figure 6.14, where the two transparent bins initially were empty. In both cases, the neighboring bins are indicated with purple or orange lines and the highest damage of these is copied to the empty bin.

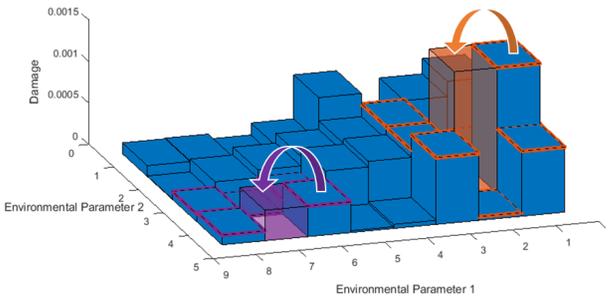


Figure 6.14: Conceptual illustration on how empty bins were filled. The maximum value found in the surrounding bins of the damage table is used.

By filling up the damage table in such a way, a kind of artificial damage data is added to the damage table. This can lead to a bias in resulting lifetime, as illustrated in Figure 6.15. This figure shows the error between the 1st or 5th percentile of the resulting lifetime distribution and the mean value of the lifetime distribution, obtained by extrapolation based on wind speed only. For small measurement periods (shown on the x-axis), a higher error is obtained in case of 2d extrapolation (i.e. based on two environmental parameters) with respect to 1d extrapolation (i.e. based on one environmental parameter). This is explained by the increased number of empty bins for 2d extrapolation compared to 1d extrapolation.

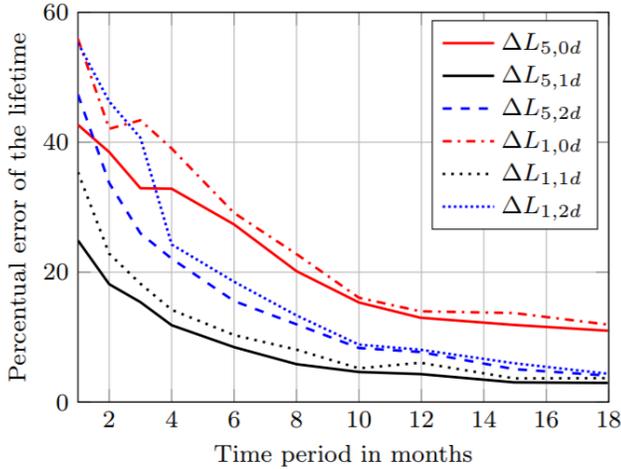


Figure 6.15: Error of the 1st or 5th percentile of the lifetime distributions calculated using different dimensionality of damage tables, 0d, 1d or 2d, with respect to the mean value of the lifetime distribution using wind speed only.[88]

To conclude, if a damage table with a lot of empty bins is obtained, filling up empty bins can not always be done in a realistic way. Therefore one can reconsider the selected binning table. This can be done by reducing the dimensionality (e.g. the number of selected environmental parameters) or adjusting the bin borders.

### Operational conditions

As shown by the fatigue spectra in Figure 6.3, the measured stress cycles differ significantly between standstill and operating conditions. In this case, also the accumulated damage will differ quite a lot. The same is true for events such as a rotor stop.

Due to these differences, multiple damage tables will be composed in this thesis: one for operational conditions, one for standstill conditions and one for rotor stops.

### 6.3.2 Translating the damage table into RUL

Once the damage table is composed, this damage table should be multiplied by the needed probability distributions  $P_r$  associated with each of the bin properties. The probability distribution can be the one as used in design or a histogram composed using measurements from several years. If the latter is used, one should take care the probability function is based on a much longer measurement period than the one used to calculate the damage table. If not, the resulting lifetime will approximate the 0d damage extrapolation as given by Equation 6.13. On the other hand, the as designed probability distribution might not be correct either, since a validation is not existent. The probability distribution in design might also

be conservative, since it usually takes into account some rare storm conditions as well. Unfortunately, this cannot be demonstrated with real-life data due to confidentiality reasons.

By summing up the result,  $D_{extrapolated}$  is obtained. This is summarized by Equation 6.14.

$$D_{extrapolated} = \sum_{i \in ECbins} D_i \cdot P_{r,i} \tag{6.14}$$

To account for the difference in damage accumulation during operating and standstill conditions, the same procedure can be repeated twice, once with operating data and once with non-operating data. A weighted average (based on the expected ratio of operation - standstill) of the resulting yearly damages gives an indication of yearly damage under the required environmental and operational conditions.

Finally, Equation 6.13 can be used again to calculate the lifetime, by replacing  $D_{measured}$  by  $D_{extrapolated}$  and taking  $Period_{measured}$  as one year.

### 6.3.3 Reliability of lifetime calculation

For the last part of this analysis, a look is taken at the uncertainty of the lifetime calculation. To do this, the concept of sample-based bootstrapping [100] is applied on the calculations. Therefore, the lifetime calculation is repeated multiple times to give an estimate of the variation of the lifetime. Here, not all available measurements  $N_i$  in a bin  $i$  are used to calculate the yearly damage for that bin. Instead the same number of measurements,  $N_i$ , is picked randomly out of the available measurements, where the same measurement can be picked more than once. This random combination of measurements is then used to calculate the yearly damage for that bin. This is illustrated in Figure 6.16.

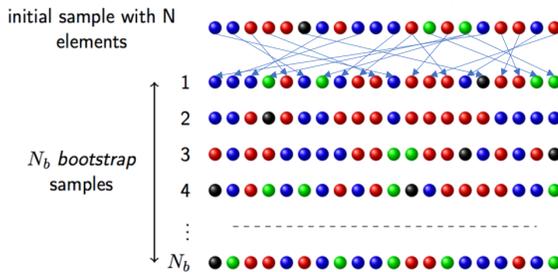


Figure 6.16: Conceptual illustration of sample-based bootstrapping (modified from [101])

In this analysis, this procedure is repeated  $N_b$  times for every bin, resulting in  $N_b$  damage tables and finally in  $N_b$  different values for lifetime. These resulting distributions can then be plotted for each detail in the structure. As such, an uncertainty on the final resulting lifetime is obtained.

## 6.4 Case study

In this final section we will use the geometry of a fictional offshore wind turbine on a monopile foundation and normalized measurements from one of the monitoring campaigns to give an example of the full work flow for fatigue life estimation on a single turbine. The fictional OWT was based on the geometry of the OWT given in [102]. Results will be analyzed to discuss some of the elements raised in earlier sections.

### 6.4.1 Geometry, safety factors and applied S-N curves

During this research a framework was developed that takes the geometry of an offshore wind turbine as an input and automatically generates a table of expected safety and correction factors. Additionally it automatically proposes a specific S-N curve for each weld based on the height of the weld. This table can then later be updated for any particular detail, not clear from the original geometry, such as secondary attachments like boat-landings and ladders. Moreover, the proposed values for safety factors can be adjusted if a difference with the design documents is found. The same is true for the proposed S-N curves. It might be some welds were treated by grinding or shot peening during construction, to reduce the number of defects in the material. In that case, typically a different S-N curve is used to account for the effect on the fatigue of the weld by the weld treatment.

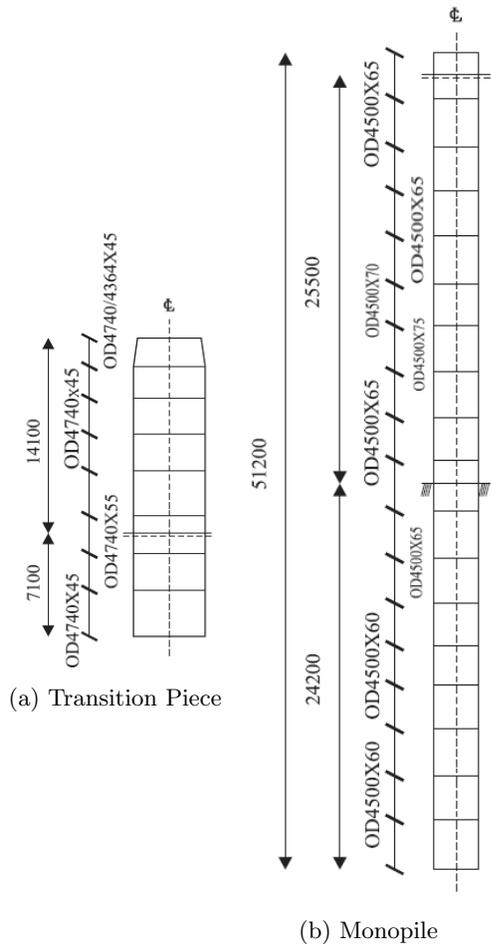


Figure 6.18 and Table 6.1 show the geometry as taken from [102]. In Tables 6.2 and 6.3 the output of the developed script is shown.

Figure 6.18: Illustration of the geometry of a fictional OWT (modified) [102]

Table 6.1: The geometry details of a fictional OWT, inspired by [102]

Height	Section Height	Outer Diameter	WT
(m), w.r.t. LAT	(mm)	(mm)	(mm)
<b>Transition Piece</b>			
14,1	0	4364	45
12,15	1950	4740	45
9,35	2800	4740	45
7,3	2050	4740	45
4,7	2600	4740	45
1,9	2800	4740	55
-0,75	2650	4740	55
-3,75	3000	4740	55
-7,1	3350	4740	45
<b>Monopile</b>			
1,5	0	4500	65
-1,5	3000	4500	65
-4,5	3000	4500	65
-7	2500	4500	65
-10	3000	4500	65
-13	3000	4500	65
-15,6	2600	4500	70
-18,6	3000	4500	75
-21,6	3000	4500	65
-24,6	3000	4500	65
-27	2400	4500	65
-30	3000	4500	65
-33	3000	4500	65
-36	3000	4500	60
-38,3	2300	4500	60
-41,3	3000	4500	60
-44,3	3000	4500	60
-46,7	2400	4500	60
-49,7	3000	4500	60

Table 6.2: Fatigue details of the monopile of a fictional OWT, as calculated by the developed framework during this thesis.

Height	Label	Structure	SNCurve	SCF	SE	MSF
<b>Monopile</b>						
-1,5	CW02-O	MP	DNV-D-W	1	1,211	1,25
-1,5	CW02-I	MP	DNV-D-W	1	1,211	1,25
-4,5	CW04-O	MP	DNV-D-W	1	1,211	1,25
-4,5	CW04-I	MP	DNV-D-W	1	1,211	1,25
-7	CW06-O	MP	DNV-D-W	1	1,211	1,25
-7	CW06-I	MP	DNV-D-W	1	1,211	1,25
-10	CW08-O	MP	DNV-D-W	1	1,211	1,25
-10	CW08-I	MP	DNV-D-W	1	1,211	1,25
-13	CW10-O	MP	DNV-D-W	1	1,211	1,25
-13	CW10-I	MP	DNV-D-W	1	1,211	1,25
-15,6	CW12-O	MP	DNV-D-W	1	1,229	1,25
-15,6	CW12-I	MP	DNV-D-W	1	1,229	1,25
-18,6	CW14-O	MP	DNV-D-W	1	1,211	1,25
-18,6	CW14-I	MP	DNV-D-W	1,059	1,211	1,25
-21,6	CW16-O	MP	DNV-D-W	1	1,211	1,25
-21,6	CW16-I	MP	DNV-D-W	1	1,211	1,25
-24,6	CW18-O	MP	DNV-D-W	1	1,211	1,25
-24,6	CW18-I	MP	DNV-D-W	1	1,211	1,25
-27	CW20-O	MP	DNV-D-W	1	1,211	1,25
-27	CW20-I	MP	DNV-D-W	1	1,211	1,25
-30	CW22-O	MP	DNV-D-W	1	1,211	1,25
-30	CW22-I	MP	DNV-D-W	1	1,211	1,25
-33	CW24-O	MP	DNV-D-W	1	1,191	1,25
-33	CW24-I	MP	DNV-D-W	1	1,191	1,25
-36	CW26-O	MP	DNV-D-W	1	1,191	1,25
-36	CW26-I	MP	DNV-D-W	1	1,191	1,25
-38,3	CW28-O	MP	DNV-D-W	1	1,191	1,25
-38,3	CW28-I	MP	DNV-D-W	1	1,191	1,25
-41,3	CW30-O	MP	DNV-D-W	1	1,191	1,25
-41,3	CW30-I	MP	DNV-D-W	1	1,191	1,25
-44,3	CW32-O	MP	DNV-D-W	1	1,191	1,25
-44,3	CW32-I	MP	DNV-D-W	1	1,191	1,25
-46,7	CW34-O	MP	DNV-D-W	1	1,191	1,25
-46,7	CW34-I	MP	DNV-D-W	1	1,191	1,25

Table 6.3: Fatigue details of the transition piece of a fictional OWT, as calculated by the developed framework during this thesis.

Height	Label	Structure	SNCurve	SCF	SE	MSF
<b>Transition Piece</b>						
14,1	CW01-O	TP	DNV-D-A	1	1,125	1,25
14,1	CW01-I	TP	DNV-D-A	1	1,125	1,25
12,15	CW03-O	TP	DNV-D-A	1	1,125	1,25
12,15	CW03-I	TP	DNV-D-A	1,84	1,125	1,25
9,35	CW05-O	TP	DNV-D-A	1	1,125	1,25
9,35	CW05-I	TP	DNV-D-A	1	1,125	1,25
7,3	CW07-O	TP	DNV-D-A	1	1,125	1,25
7,3	CW07-I	TP	DNV-D-A	1	1,125	1,25
4,7	CW09-O	TP	DNV-D-W	1	1,125	1,25
4,7	CW09-I	TP	DNV-D-W	1,187	1,125	1,25
3,3	BL01-O	TP	DNV-D-W	2	1,2	1,25
1,9	CW11-O	TP	DNV-D-W	1	1,171	1,25
1,9	CW11-I	TP	DNV-D-W	1	1,171	1,25
-0,75	CW13-O	TP	DNV-D-W	1	1,171	1,25
-0,75	CW13-I	TP	DNV-D-W	1	1,171	1,25
-3,75	CW15-O	TP	DNV-D-W	1	1,125	1,25
-3,75	CW15-I	TP	DNV-D-W	1,187	1,125	1,25

With the developed framework all relevant parameters are stored and can later be used to provide a correct estimate of fatigue life, respecting all structural details of the considered substructure.

## 6.4.2 Measurements

To illustrate the given approach for a lifetime assessment, it is applied using data from an offshore wind turbine, installed on a monopile. It is shown by [88] a minimum period of 9 months is needed to obtain reliable lifetime results. Therefore a measurement period of one year is taken. During this period, three types of data are available. First of all strain sensors at the interface between tower and transition piece are used to calculate the fore-aft and side-side stress signals. The calculations needed for this transition are explained in Section 2.2.2. In practice, cycle counting is performed on both FA and SS stress signal. Afterwards, damage is calculated using the maximum number of cycles counted among both FA and SS for each stress range bin. This strategy introduces an additional conservatism since it assumes all damage is accumulated in the same direction and thus at the same location along the circumference of the weld.

Moreover a subset of 10 minute statistics of SCADA parameters and meteorological data representing wave conditions is available for slightly less than seven years. The subset contains wind speed (mean value and standard deviation) and wind direction. Based on the mean value and the standard deviation of the wind speed, the turbulence intensity can be calculated using Equation 6.15.

$$TI = \frac{sV}{\bar{V}} \quad (6.15)$$

All available SCADA was filtered first to remove erroneous data. All values lower than a predefined minimum value or higher than a predefined maximum value were excluded. The chosen set of criteria is given in Table 6.4 and reflect improbable or impossible values for SCADA. Moreover to exclude one time outliers, all values  $x_i$  for which all expressions in Equations 6.16, where  $p$  is a predefined percentage and  $T$  a predefined threshold value, are fulfilled are excluded too. The values chosen for  $p$  and  $T$  are given for each parameter in Table 6.4.

$$\begin{aligned} \left| x_i - \frac{x_{i-1} + x_{i-2}}{2} \right| &> p \cdot x_i \\ \left| x_i - \frac{x_{i+1} + x_{i+2}}{2} \right| &> p \cdot x_i \\ \left| x_i - \frac{x_{i-1} + x_{i-2}}{2} \right| &> T \\ \left| x_i - \frac{x_{i+1} + x_{i+2}}{2} \right| &> T \end{aligned} \quad (6.16)$$

Table 6.4: Predefined values for different SCADA parameters during filtering

	wind speed (m/s)	TI (%)	wind direction (°)
absolute minimum	0	2	-360
absolute maximum	50	999	360
p (%)	100	100	0
T	5	20	999

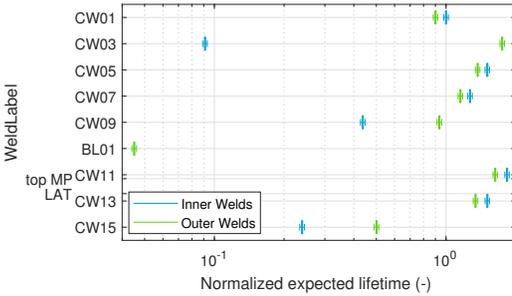
### 6.4.3 Results

While performing a lifetime assessment, a lot of choices have to be made. For some of them, one can rely on the design documents. For example the values for several safety factors or the choice of S-N curves. To illustrate the effect of different S-N curves and a different value for the combined safety factor (CSF), the lifetime is calculated for all welds of the substructure of the fictional OWT. The calculation is done using a damage table composed with operational data only and binned based on wind speed only. All factors are applied as specified in Section 6.4.1, together with a static extrapolation factor to account for the difference in bending moment according to the exact location of the weld. Moreover, the S-N curve as specified in Section 6.4.1 is used for every detail in the damage calculation. The extrapolation in time is performed based on a wind speed distribution as measured over almost 7 years.

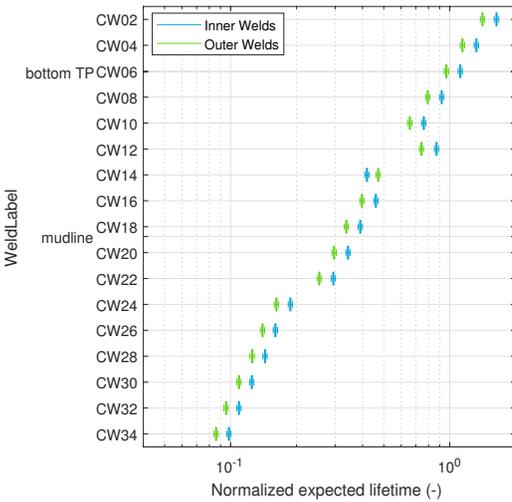
Sine the resulting estimates of lifetime are purely for illustration, they are normalized with respect to the resulting lifetime estimate of the weld closest to the

sensor location (the inside weld of CW01).

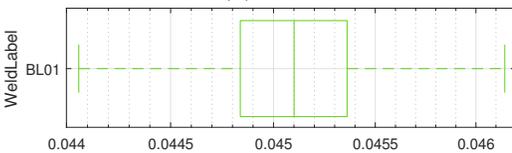
By only using measurements under normal operating conditions, the assumption is made the turbine will operate its entire lifetime under normal operating conditions, without one rotor stop or one second of standstill. Off course, this is not true in reality. The method can be easily extended to multiple operating conditions, as discussed later in this section.



(a) TP



(b) MP



(c) Most critical weld

In general, it can be observed that the expected lifetime of the outer welds (illustrated in green in Figure 6.20) is slightly lower than the expected lifetime of the inner welds (illustrated in blue in Figure 6.20). This is because of the slightly higher value for the static extrapolation factor of the stress. Since the radius of the location of interest  $R_{ext}$  in case of outer welds is a little higher than for inner welds, the SEF is slightly increased (Equation 6.3). An increased stress range causes an increased damage and a lower expected lifetime.

All exceptions for which the expected lifetime of the inner weld is lower than for the outer weld can be explained by a significant thickness transition at the inside of the structure (CW09, CW15 and CW14) or a conical section (CW03).

Overall, the effect of a difference in distance to the hub on the SEF and on the lifetime can be easily observed. An increase in SEF due to an increasing  $z_{ext}$  (Equation 6.3), resulting in an increased stress range and thus a higher damage and lower lifetime expectation can be observed for all consecutive welds with the same cross section and same CSF. Examples of these are CW02 to CW10, CW16 to CW22 and CW24 to CW34.

Figure 6.20: Lifetime assessment for all weld details on a fictional OWT, based on an extrapolation using wind speed only

Some salient differences in expected lifetime can be observed. The expected lifetime for the top weld (CW01) is different because of the smaller diameter with respect to the other welds of the transition piece. This difference in diameter will lead to a smaller  $I_{c,ext}$  (Equation 2.8) and thus a higher SEF (Equation 6.3) and lower lifetime expectation.

Usually a change in wall thickness can explain some jumps in expected lifetime as well. For example CW11, CW13, CW12 and CW22 have higher wall thicknesses than neighboring welds. These higher wall thicknesses will lead to a higher value for size effect (Equation 6.4), which would lead to lower lifetime expectations. However, a higher wall thickness also causes a higher  $I_{c,ext}$  (Equation 2.8) and thus a decrease for the SEF (Equation 6.3). The decrease in SEF is usually bigger than the increase in SE, which eventually will lead to a higher expected lifetime. The lowest estimate for expected lifetime can be found for BL01. This is explained by the highest value for CSF for this boat landing stub weld.

When looking at the difference between the fatigue characteristics of the outer welds of CW07 and CW09, the main difference would be a different S-N curve (i.e. DNV-D-A and DNV-D-W respectively) except for the difference in SEF. Results suggest the role of the SEF is more dominant than the change in S-N curve. This was confirmed when the expected lifetime for the outer weld of CW09 was calculated with the same values for safety factors and extrapolation factor but by using DNV-D-A as S-N curve instead of the proposed S-N curve, DNV-D-W. The resulting lifetime distribution for the different S-N curve was almost exactly the same. Less than 1 % difference in median expected lifetime was obtained.

Other choices in a lifetime assessment often depend on the availability of the needed information. For example, if no load measurements are available, one could use the as designed load tables, giving the expected load for each specific operational state. However, the as designed load tables are usually not publicly available. Often they're even not available for the operators or not sufficiently detailed.

The most important choice in this section is how the extrapolation over time will be performed. And more particularly, based on which parameters and operational conditions. In general, the more complex the choices and combinations, the closer to reality the results should be in theory if the measurement period is long enough. However, more problems and limitations pop up in practice.

The required choices were already introduced in Section 6.3. The remainder of this section will focus on four choices in particular. The first one deals with the dimensionality of the damage table. The second one compares multiple environmental parameters to base the probability distribution on. The third compares the division in multiple operational conditions. And the last one is whether to use the as designed probabilities or data-based probabilities. For the simplicity of the plots, it is chosen to only plot the median (indicated by the vertical line in the middle of the box of the boxplots) for expected lifetime of the outer welds in most of the following figures.

## On the dimensionality of the damage table

The choices made for the extrapolation decide the complexity of the load case table. In other words, the more environmental parameters chosen for the probability distribution, the more bins the measured damages should be divided in. And thus the more likely some of these bins will not have sufficient data points in them. In the case a bin remains empty, a (conservative) approach is followed by taking the highest damage in the neighboring bins (Figure 6.14), which is not necessarily close to the actual damage.

Moreover, the more bins, the more the damage calculation will be prone for outliers. Since less data points are available in each bin, outliers will have a higher influence on the calculated mean in the bins.

In this section, results for lifetime based on multiple damage tables with different dimensions are compared. To obtain these results, all available data was used, including parked conditions and rotor stops. In total 4 different approaches are compared. Two of them are only extrapolated in time, without any binning based on environmental conditions, both called 0d extrapolation. The third, called 1d extrapolation, uses one environmental parameter, wind speed, to extrapolate in time. Finally the last, called 2d extrapolation, uses two environmental parameters, wind speed and turbulence intensity.

For the first 0d extrapolation, the damage accumulated during one year is calculated based on the entire stress signal of that year. The resulting damage is then transformed in lifetime using Equation 6.13. This means no bootstrapping was performed. Moreover, by using this approach it is impossible to split up different operating conditions, such as operational, parked or rotor stop.

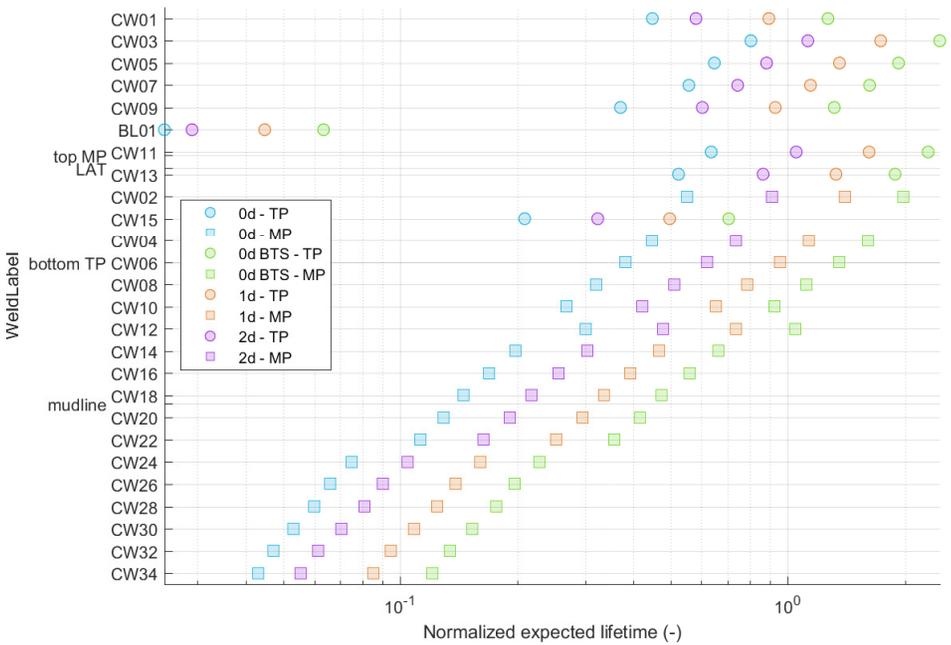
The second 0d extrapolation however was performed based on accumulated damage calculated each 10 minutes. Afterwards, only bootstrapping was performed on all data without binning it first. The main advantage of this approach is that the data could be split up first based on operating conditions. This approach is called 0d BTS in the remainder of the section. If for the construction of the damage table and for the calculation of the environmental distribution exactly the same data were chosen, the resulting expected lifetime would essentially be the same as this 0d BTS extrapolation.

The 1d and 2d extrapolation is obtained by binning all damage data first based on wind speed only or wind speed and turbulence intensity respectively. The extrapolation in time is based on environmental distributions as measured over almost 7 years.

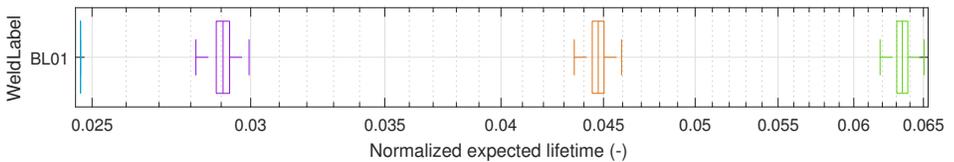
Figure 6.21 shows the results for both 0d extrapolations, 0d and 0d BTS, in blue and green respectively. The results for the 1d extrapolation are shown in orange. Finally, the results for the 2d extrapolation are shown in purple.

A rather big difference can be observed between the two 0d extrapolations. Two differences between the two approaches can be identified: signal splitting and bootstrapping. During bootstrapping, the random choice of data points will influence the resulting lifetime for that particular bootstrap. The environmental and operational distribution of chosen data points will probably differ from the general

distribution of the entire dataset. Moreover, it is possible a higher percentage data points with a high damage is chosen. The same is true for data points with a low damage. It is assumed however that the median value of the resulting lifetime distribution represents a lifetime estimate based on the same environmental and operational distribution as the entire data set. Moreover no over- or underestimation of more or less damaging data points is assumed to be present in the median value of the lifetime distribution. Therefore it is believed the distributions used for both 0d extrapolations are exactly the same.



(a) Median value of the resulting lifetime distribution (in logarithmic scale) for all outer welds of the substructure (TP in circles and MP in squares)



(b) Zoom on distributions of critical weld

Figure 6.21: Expected lifetime for outer welds of the substructure, extrapolated using a data-based environmental distributions if required. The lifetime was obtained using different damage tables with different dimensions: 0d in blue, 0d BTS in green, 1d in orange and 2d in purple.

This means the main difference between the two extrapolations is the division of the entire signal into smaller signals of 10 minutes each. Splitting the signal

causes a lot of damage to be missed if the obtained damage for each smaller signal is just added. The very slow cycles, with periods over 5 minutes, are missed. However, these cycles are usually very big and still induce a lot of damage. This explains the difference in lifetime estimates. Although this is a big drawback of the binning methodology of damages, the advantages of linking damage to operating and environmental conditions still outweigh this loss of information. However, a compensation for this loss of damage should be included in the future. In case one is only interested in the accumulated damage over the past time, the 1s SCADA data can be a solution. The slow cycles are induced by variations in wind, which are captured by the SCADA data. Another workaround could be to store the minimum and maximum stress recorded during each 10min interval and perform an additional cycle count on the concatenation of these.

To compare the impact of including environmental parameters in the lifetime extrapolation, the comparison will be made with respect to 0d BTS extrapolation. Figure 6.21 shows the 0d BTS extrapolation result in lower accumulated damage and higher expected lifetimes than the extrapolations based on environmental conditions. This is because the environmental conditions seen during the measurement period were less damaging than expected for the entire lifetime. This difference is illustrated by Figure 6.22.

Here, the extrapolated damage is shown for 0d BTS and 1d extrapolation as a stacked bar. Each colored bar in the stacked bars represent the contribution to the extrapolated damage of one wind speed bin. This contribution is the product of the averaged measured damage for that wind speed bin and the probability of occurrence of that wind speed bin. The average measured damage used is calculated at the sensor location, using the DNV-D-A S-N curve, the value for SE as calculated by Equation 6.4, 1,25 as MSF and 1 for SCF. The probability of occurrence for the 0d BTS extrapolation is the measured probability during the measurement period. As both damage extrapolations are based on a damage table composed for exactly the same period, both damage tables are exactly the same. The only difference between both extrapolations is thus the environmental probability.

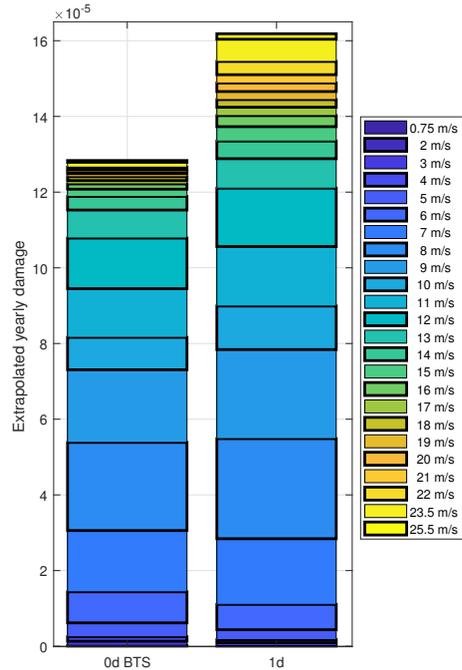


Figure 6.22: Extrapolated damage during one year for 0d BTS and 1d extrapolation. The contribution of each wind speed bin is shown by a different color.

It can be seen that for wind speeds lower than  $4m/s$ , more extrapolated damage is accumulated for the 0d BTS extrapolation, whereas for all higher wind speeds the extrapolated damage of the 1d extrapolation is clearly higher. This is caused by a higher probability of higher wind speeds during the extrapolation period than the measurement period. Or in other words, because the measurement period was a gentle period with respect to the expectation for the entire lifetime.

The 2d extrapolation results in even lower lifetimes than the 1d extrapolation. The reason for this can be two-folded. Firstly, again the difference between seen and expected environmental conditions will influence this extrapolation.

This effect is illustrated by Figure 6.23, where the extrapolated damage for each turbulence intensity bin within one wind speed bin is shown for both 1d and 2d extrapolation. As for the 1d extrapolation, no distinction is made for expected turbulence intensity during the turbine's lifetime, the environmental probability utilized to obtain the extrapolated damage is the as measured during the measurement period. Since again the damage table used for both extrapolations is exactly the same, the only difference between both extrapolations is the environmental probabilities. It is clear that during the measurement period the environmental conditions were dominated by low turbulence intensity. However, given the clear difference in accumulated damage for the 2d extrapolation, this is not what is expected for the entire lifetime of the turbine. Although this is only demonstrated for one wind speed bin, it is valid for most wind speed bins higher than  $5m/s$ .

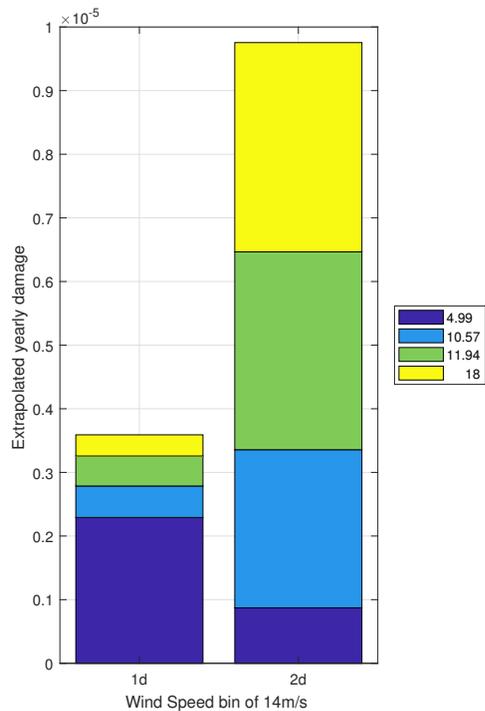


Figure 6.23: Extrapolated damage during one year for 1d and 2d extrapolation, for one specific wind speed bin. The contribution of each turbulence intensity bin is shown by a different color.

Moreover, filling up empty bins and the higher dependence of average damage bin values on outliers can cause additional conservatism in the extrapolation leading to lower expected lifetimes.

### Parameters in the damage table

Several environmental parameters influence the measured damage, as shown in Section 6.3.1. This can already be demonstrated by looking at the obtained damage distributions in each bin of multiple damage tables. Again, damages shown in the distribution are calculated at the sensor location, using the DNV-D-A S-N curve, the value for SE as calculated by Equation 6.4, 1,25 as MSF and 1 for SCF. Figure 6.24 shows an example of a damage table, composed based on wind speed only and using only operational data. One can observe an increased damage around 12 m/s due to the maximum thrust loading. For very high wind speeds, the damage is again increased. This increase can be explained in multiple ways. Higher dynamic loads caused by higher wave heights usually occur only for high wind speeds. Moreover higher wind speeds also induce higher frequent wind turbulence, which can excite the first resonance frequency more.

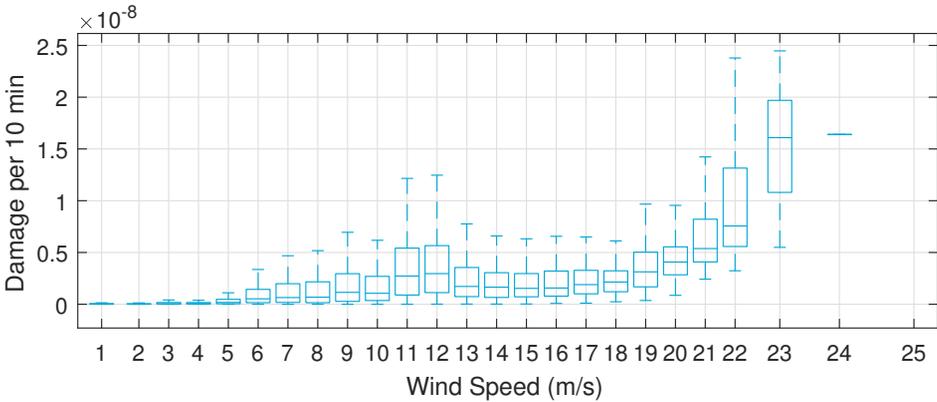
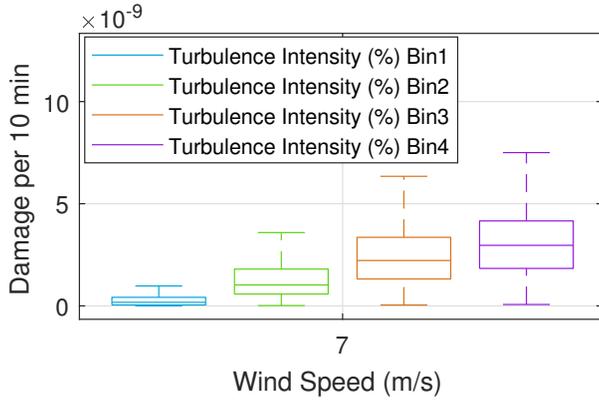
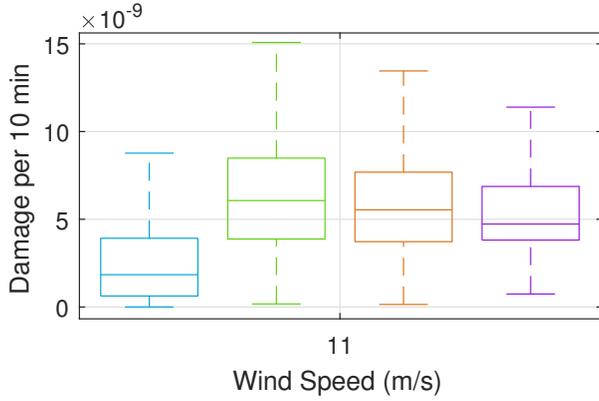


Figure 6.24: Damage distributions in each bin of the damage table, composed based on wind speed only. The distributions are indicated using box plots.

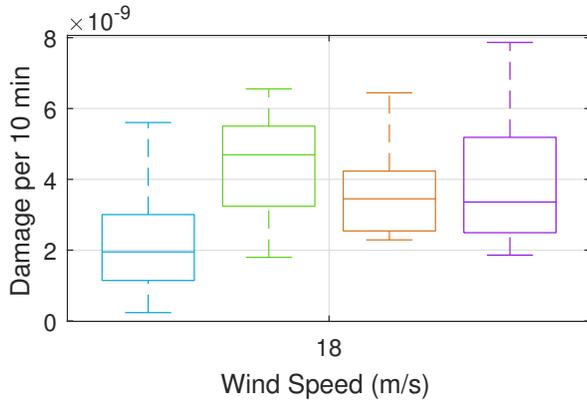
Apart from the wind speed, also the waves and the wake flows in the farm have an influence. Therefore, composing a damage table based on wind speed and turbulence intensity can have an added value. This is illustrated by Figure 6.25, where the damage distribution is shown for all turbulence intensity bins of a few wind speed bins. Usually an increased damage and increased damage range can be observed for increasing turbulence intensity in one wind speed bin for low wind speeds (up to 9 m/s). This is shown in Figure 6.25a. For higher wind speeds, starting from wind speeds around maximum thrust loading (11 m/s) however, the turbulence intensity seems to be less critical. Both in terms of absolute damage values and damage ranges, the higher the turbulence doesn't necessarily mean the higher the damage (range). This can be seen in Figures 6.25b and 6.25c. Although, for the lowest turbulence intensity, the lowest damages are found in every wind speed bin.



(a) Wind speed around 7m/s



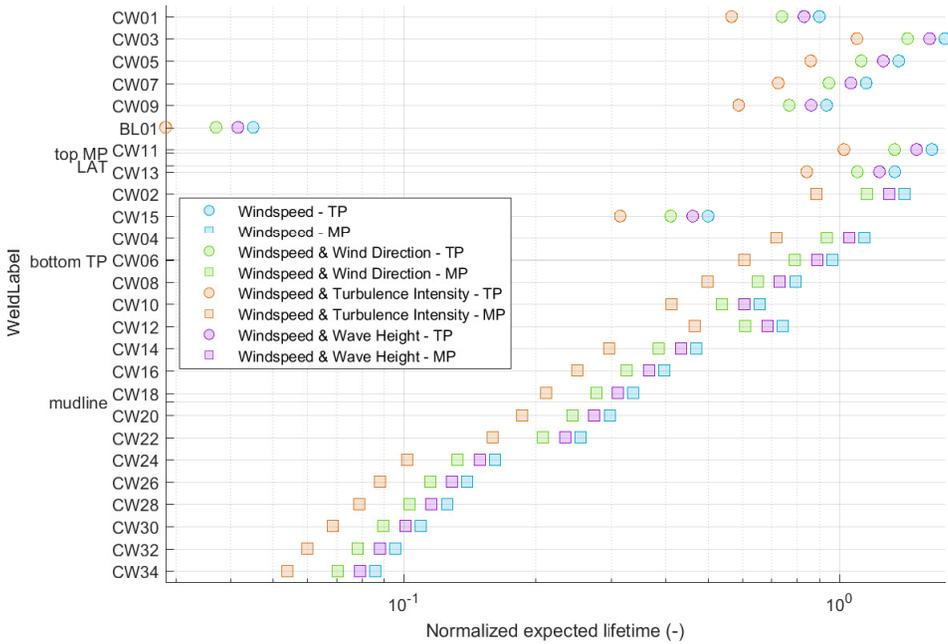
(b) Wind speed around 11m/s



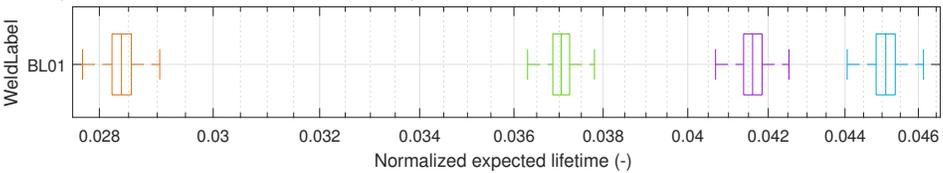
(c) Wind speed around 18m/s

Figure 6.25: Damage distributions in all turbulence intensity bins of the damage table, for a specific range of wind speeds. The distributions are indicated using box plots.

The differences in damage tables can be translated to lifetime distributions as well. Figure 6.26 shows the resulting median expected lifetime for all outer welds for four combinations of environmental parameters. The extrapolation in time is based on environmental distributions as measured over almost 7 years. The first one is simply wind speed only, shown in blue. The second (wind speed and wind direction) and third (wind speed and turbulence intensity) take into account the possible wake flows in the farm, shown in green and orange respectively. The last (wind speed and wave height) considers the wave activity as well, shown in purple.



(a) Median value of the resulting lifetime distribution for all outer welds of the substructure (TP in circles and MP in squares)



(b) Zoom on distributions of critical weld

Figure 6.26: Expected lifetime for outer welds of the substructure, calculated using different damage tables. The damage tables were composed using several combinations of environmental parameters: wind speed only in blue, wind speed and wind direction in green, wind speed and turbulence intensity in orange and wind speed and wave height in purple.

To explain these results, it is important to know for which bins the highest

average damage was measured. Heat maps were created to visualize the damage table of each extrapolation option. In these heat maps, high damage is represented by dark red, while low damage is represented by dark blue. These are depicted in Figure 6.27. Bins that needed to be filled are shown with a larger border. It can be seen highest damage is measured for high wind speed. If the damage is binned based on two parameters, usually one of the bins show less damage than the others. For high wind speeds, usually some bins needed filling up, often with a high damage value.

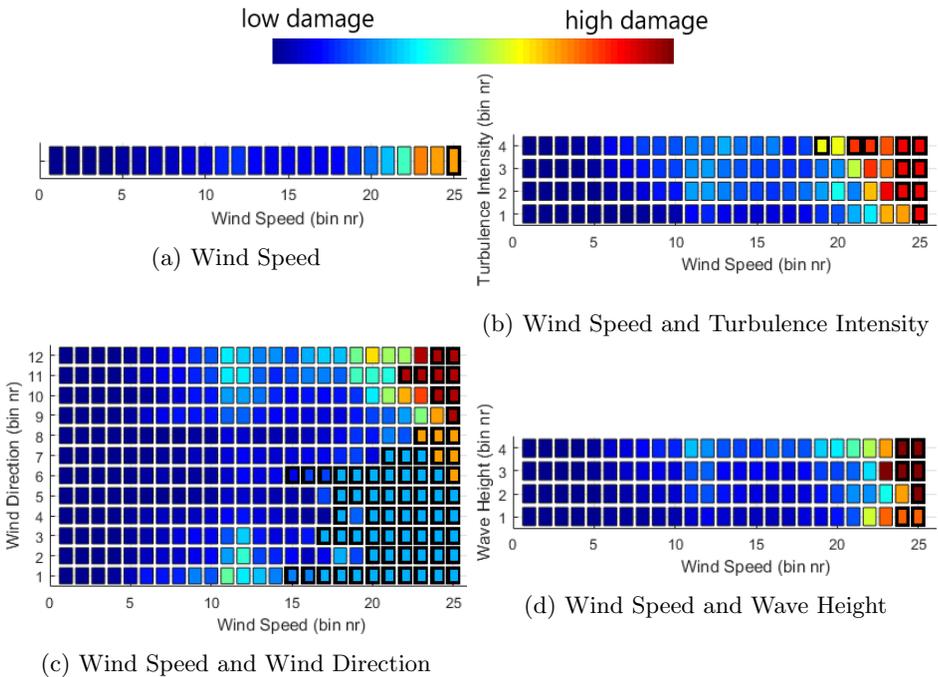


Figure 6.27: Damage tables for several lifetime extrapolations based on different combinations of environmental parameters

Moreover the difference in occurrence of these specific bins during the damage measurement period, the measurement period, and the period on which the extrapolation was based upon, the extrapolation period, is important as well. To quantify the occurrence of each specific environmental bin relative to the total period, the relative occurrence of all bins is calculated for both periods. The relative occurrence is defined as the number of data points in the specific bin  $N_{bin}$ , divided by the total number of data points  $N_{total}$  ( $P = \frac{N_{bin}}{N_{total}}$ ). To visualize this difference, heat maps were created, where the colors indicate the difference in relative occurrence between the extrapolation period and the measurement period ( $P_{extrapolation} - P_{measurement}$ ). Red indicates a higher relative occurrence during the extrapolation period and blue a higher relative occurrence during the measurement period. These heat maps are shown in Figure 6.28. In general, it can be

seen the measurement period had more gentle average environmental conditions than the extrapolation period. That is to say, bins represented by blue to green colors in Figure 6.28 occur all for low wind speeds, low turbulence intensity or low wave height. This difference is even more visible for 2d tables (Figures 6.28b, 6.28d and 6.28c). This observation was already made in previous section as well.

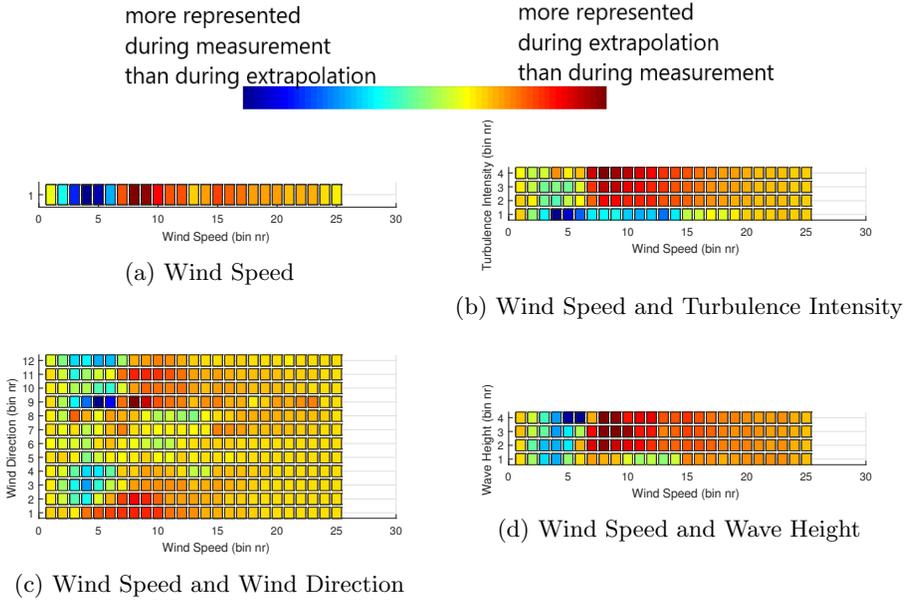


Figure 6.28: Differences between relative occurrence of the specific environmental conditions during the measurement period with respect to the extrapolation period, based on several combinations of environmental parameters

All these figures show that conditions for which the average damage is high (red to green in Figure 6.27) occurred more often during the extrapolation period than during the measurement period (yellow in Figure 6.28). Again, this indicates the measurement period was rather gentle in comparison to the expectation for the entire lifetime. This was also already shown in previous section.

Results in Figure 6.26 show the extrapolation based on wind speed only gives the highest lifetimes for all welds. This was also observed in previous section when 1d and 2d extrapolation was compared. This results in Figure 6.26 indicate this is not only true for turbulence intensity but also for the other parameters. The explanation for this difference remains the same. That is to say, the environmental conditions during the measurement period were quite gentle with respect to the extrapolation period. This difference in environmental conditions is more pronounced when binning is done based on two parameters. For all wind speeds also less damaging conditions, e.g. low turbulence or wave heights, are distinguish from more damaging conditions, e.g. high turbulence or wave heights. In case

of extrapolation based on wind speed only, the damage used for a specific wind speed is averaged over the different damaging conditions as they occurred during the measurement period. This can be observed in Figure 6.27. Here, the highest damage observed for the wind speed bins is clearly lower than the highest damage observed in the damage table for any other extrapolation. Since the measurement period was quite gentle, more damaging bins are given more importance during extrapolation. This is true for the extrapolation based on wind speed only, but even more so for the other extrapolations. One can see that for the other extrapolations, the highest differences in relative occurrence between measurement and extrapolation periods are found for bins around wind speed bin number 10 (bins 2 to 4 in Figures 6.28b and 6.28d and bins 1, 2 and 11 in Figure 6.28c). These bins coincide with an increased damage compared to the averaged damage for that wind speed (Figures 6.27b, 6.27d and 6.27c respectively). Therefore more influence of the higher damaging bins is found during all 2d-extrapolations. This leads to a higher extrapolated damage and lower estimates for expected lifetime.

Looking at the two options taking wake effects into account, the extrapolation using the turbulence intensity results consistently into the lowest lifetime estimate. This is explained by the erroneous filling of empty bins. Looking at the lay-out of the farm, the turbine is standing in the wake of other turbines for wind directions in bins number 1 to 3 in Figure 6.27c (among others). However, these bins (for higher wind speeds) are filled with too low average damage values in the damage table based on wind direction (Figure 6.27c). This leads to a lower extrapolated damage and thus higher lifetimes. By using turbulence intensity instead, wind directions for which the turbine stand in the wake of others but did not occur during the measurement period will be extrapolated correctly. Moreover, less bins needed filling (Figure 6.27b).

Finally, a difference in extrapolation based on turbulence intensity and extrapolation based on wave height can be observed. This can mainly be explained looking at the most dominant wind speed bins, being numbers 8 to 13 approximately. For these bins, more bins based on turbulence intensity seem to have higher average damages than the bins based on wave height. This can explain the higher extrapolated damage and thus lower expected lifetimes when considering turbulence intensity instead of wave height.

## Operating conditions

The load acting on an OWT differs from one operating condition to another. For example, during standstill conditions loads are for the current offshore wind turbines on relatively small monopiles usually much lower than during normal operating conditions of the wind turbine. Moreover, the correlation between the acting loads and the different environmental conditions can be different for another operational state. In this respect, it might be useful to compose multiple load case tables for multiple operational states. Figure 6.29 shows the resulting damage table for all operational data and for all non-operational data. Again, the damage is calculated at the sensor location. The damages are distributed based on the SCADA values for wind speed and turbulence intensity. For both operational

conditions, the same bin borders are chosen.

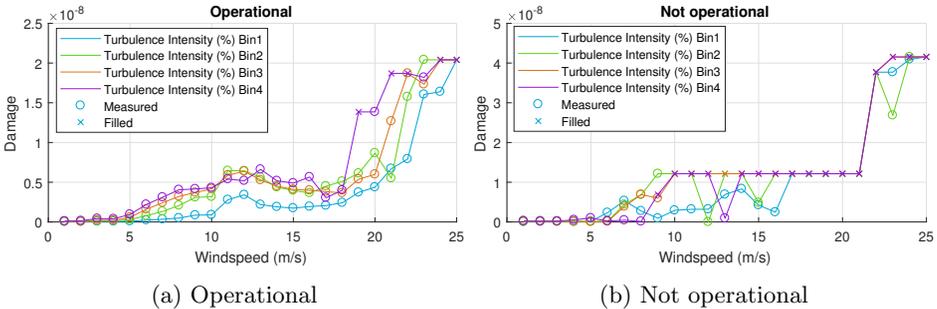


Figure 6.29: Damage tables for operational and non-operational data, based on wind speed and turbulence intensity bins

Both figures show circles when the resulting damage was calculated based on measured data in the same bin. If crosses are shown, the damage bin was empty initially and needed filling up. Here, it can be seen a lot of bins needed to be filled for non-operational conditions. Since filling up that many bins can cause a serious bias in the results, it is advised to reduce the dimensionality of the table, to increase the bin size or to collect more data. In this case, the dimensionality is reduced. Another possibility is to use damage values based on the load case tables as calculated during design, if these are available.

For this analysis, the lifetime was calculated using three different datasets. Firstly, a dataset was used where the turbine was constantly operating. In reality, this is not feasible. A second dataset contained 20% of non-operational data. This non-operational data consisted of a combination of standstill conditions and a few rotor stops. A third dataset also contained 20% of non-operational data. Again the non-operational data consisted of a combination of standstill conditions and rotor stops. For this dataset however a lot more rotor stops were recorded. Moreover standstill conditions occurred less. In case of the two last datasets, two damage tables were composed. One based on operational data only, using wind speed and turbulence intensity to bin. A second based on non-operational data only, using only wind speed to bin. The two resulting damage tables based on non-operational data are shown in Figure 6.30.

It can be seen the mean damage recorded for the dataset with more rotor stops (shown in orange) is often higher than for the dataset dominated by standstill (shown in green). However, still a lot of empty bins were detected and needed filling. For the dataset dominated by standstill, the damage clearly increases for increasing wind speed starting at  $18\text{m/s}$ . This might be because these bins can include cut-out events where the rotor stops rotating due to too high wind speeds.

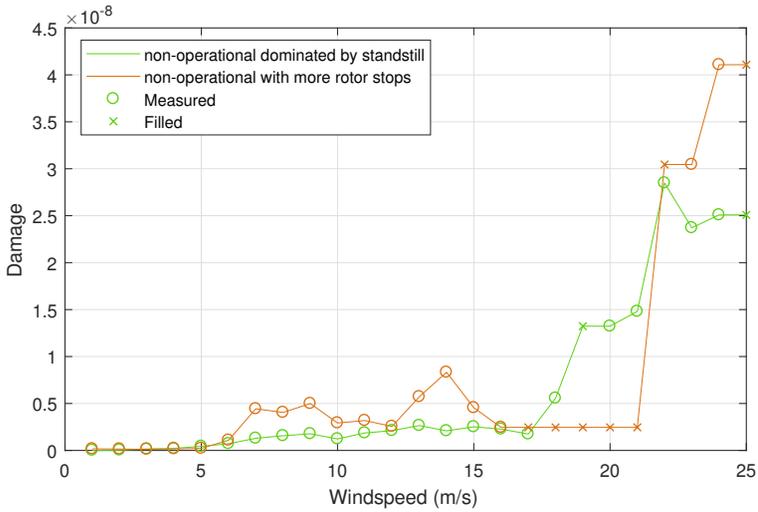
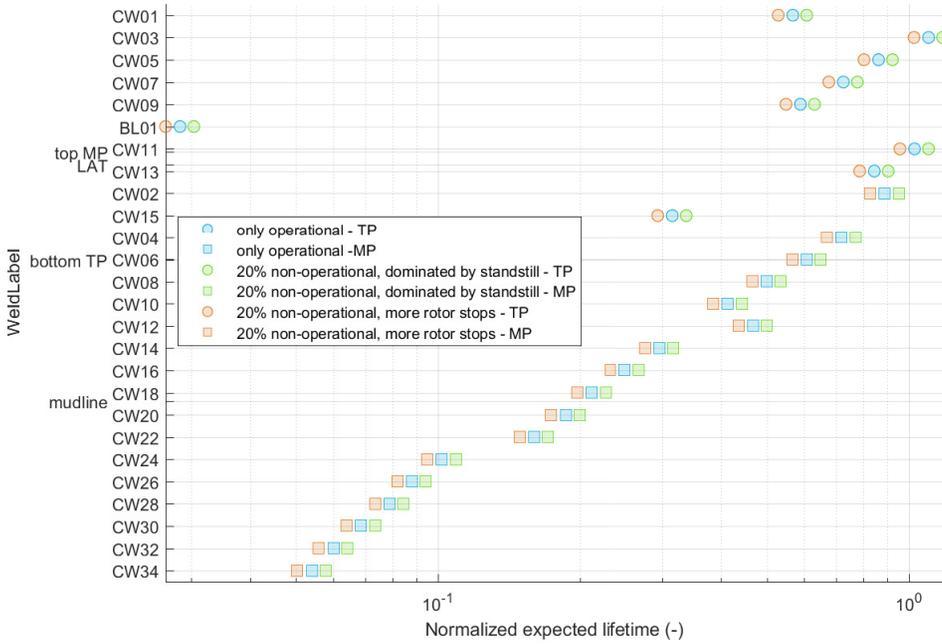


Figure 6.30: Damage tables based on two sets of non-operational data. Set 1 was dominated by standstill conditions, while set 2 contained more rotor stops.

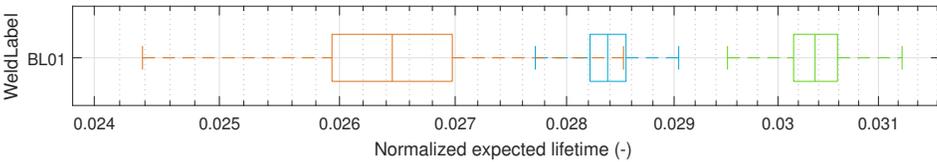
The resulting expected lifetimes are shown in Figure 6.31.

The results show a consistent lower estimate for lifetime for the dataset containing more rotor stops. This is explained by the more damaging character of a rotor stop with respect to operating conditions or standstill. Moreover, in terms of damage accumulation non-operation can be advantageous. It can be seen that the results for the dataset including non-operational data dominated by standstill show a higher expected lifetime than the dataset with only operational data. This confirms the statement that standstill conditions are less damaging than operating conditions, for this type of turbine and foundation. To achieve a reliable expected lifetime, it is thus important a distinction between operating conditions, standstill and rotor stops is made. It is not difficult to distinguish operational data from non-operational data based on standard SCADA parameters. However, when only 10 minute statistics of SCADA data are available, the distinction between standstill and rotor stops is not as easily made.

When performing a lifetime assessment, creating different damage tables based on different operational conditions might be very useful. If necessary, also the environmental probability distributions used for each operational state can differ. Eventually, a weighted combination of the resulting extrapolated damages can be used to calculate the lifetime.



(a) Median value of the resulting lifetime distribution (in logarithmic scale) for all outer welds of the substructure (TP in circles and MP in squares)

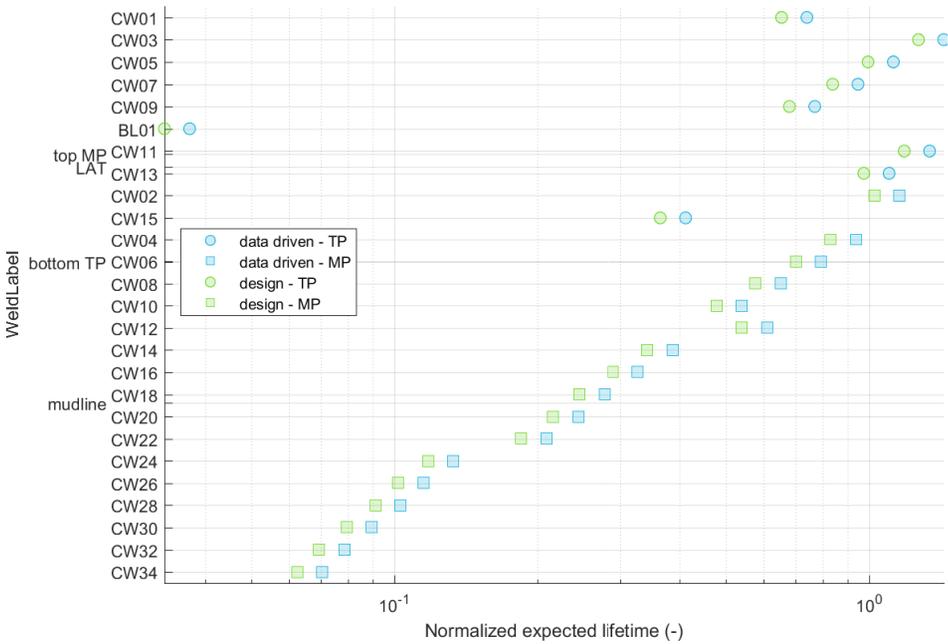


(b) Zoom on distributions of critical weld

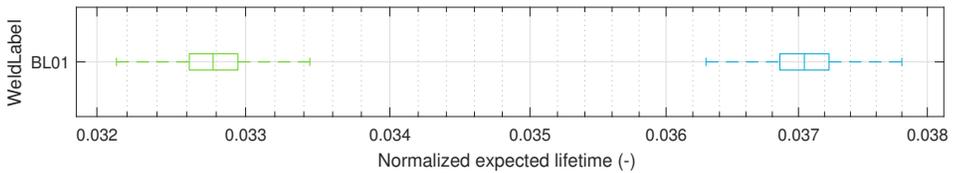
Figure 6.31: Expected lifetime for outer welds of the substructure, based on three different scenarios. The first assumes only operating conditions over the entire lifetime, shown by the blue markers. The second includes non-operational conditions, dominated by standstill, shown by the green markers. The third also includes non-operational conditions but including more rotor stops, shown by the orange markers.

### Design vs measurements

In design documents, one can find the distribution used to represent the environmental conditions over the entire lifetime of the wind turbine. This distribution is believed conservative, as it also includes some storm conditions that only rarely occur. Moreover, it is not sure if the conditions meteorologists expect are the same as those really occurring.



(a) Median value of the resulting lifetime distribution (in logarithmic scale) for all outer welds of the substructure (TP in circles and MP in squares)



(b) Zoom on distributions of critical weld

Figure 6.32: Expected lifetime for outer welds of the substructure, based on operational data only. Extrapolation is performed based on wind speed and wind direction distributions, both from design (green markers) and data-based (blue markers).

The alternative is to calculate a distribution based on measurements performed during a longer period of time. This distribution is closer to the truth for the past period. However, it is not sure if this distribution is representative for the coming years.

Figure 6.32 shows the difference in expected lifetime when an as designed distribution or a data-based distribution is used. In both cases the distribution was made based on wind speed and wind direction and only data under operating conditions was used. Results based on as designed probabilities show a lower expected lifetime for all welds, compared to results based on data driven probabilities. These results thus confirm design takes into account a higher probability of high damaging environmental conditions.

### **Lifetime assessments for decision support**

In the case study, a lot of confidence is given to the values as reported or suggested by design documents and available standards. However during the lifetime of a wind turbine, the values for some safety factors or the choice for S-N curve could change. A possible reason for this is corrosion. In that case, a different S-N curve is applied to account for a faster fatiguing of the material. Moreover, corrosion pits can cause additional local stresses leading to a necessary increase in SCF. Or as material corrodes, the radius of the structure might change.

Another possibility is the adjustment of standard procedures as a result of new research. For example recently additional research was done to update the S-N curves as used for offshore wind industry. Or, as indicated in Section 6.2.3, an adjustment of the standard formulas to calculate SE or SCF.

The developed framework facilitates the change in such parameters. Therefore dedicated analyses to see the effect of such changes on the lifetime could be done easily.

### **Impact of additional inspections on lifetime assessments**

For some of the safety factors as explained in Section 6.2.3 the value is chosen based on the required inspection plan over the lifetime of the wind turbine. But even if these values are chosen high enough and no inspections are strictly required, the resulting information coming from such an inspection can improve the results of a lifetime assessment. By conducting inspections, information is available of the actual cracks in the structure. With this additional information, a more accurate lifetime assessment can be performed using a fracture mechanics model, e.g. based on Paris' Law. This model requires measurements of the initial crack size and the fatigue load acting on the structure [103, 104].

On the other hand, the information obtained from a structural health monitoring system can influence the required inspections as well. In general, an inspection plan is made in such a way an inspection is scheduled when the expected accumulated damage, including the uncertainty often given by CoV, has increased significantly with respect to the previous inspection. However, by using the information obtained from a SHM system, this uncertainty CoV can be decreased significantly. A reduction of 70% is shown for marine structures in [105]. Therefore each required inspection can be postponed and the total amount of required inspections over the lifetime of a turbine can be reduced. As a result, the total expected maintenance costs can be reduced if a monitoring system is present, as

indicated in [106].

## 6.5 Conclusion

In this chapter the state-of-the-art approach for damage calculation was explained. This approach consists in performing a cycle counting algorithm on the stress signal, in this case rainflow counting. Afterwards damages are calculated using the Miner's rule and the appropriate S-N curves. In industry a lot of additional safety factors are applied to consider for the difference in dimension between the actual components and the test specimen which are used to generate the S-N curves. More information about the purpose and the calculation is given for several safety factors, such as stress concentration factor, size effect and material safety factor. The damage calculation was performed for multiple welds in the structure. Here, the specific values for the safety factors as proposed by the newest standard [93] were applied. Moreover a static extrapolation of the stress signal was performed to account for the difference in bending moment along the structure.

The general method to calculate the expected lifetime from measured damages was explained. This method is based on some decisions made for environmental and operational conditions over the entire lifetime of the turbine. Moreover, the uncertainty of the resulting lifetime was obtained using sample based bootstrapping.

The damage and lifetime calculations are demonstrated using data from an operating OWT and a geometry of a fictional OWT. The influence of the possible decisions regarding environmental and operational conditions is illustrated. To start with the effect of including one or two environmental parameters in the extrapolation over time is shown. In general, two environmental parameters would be preferred IF enough damage data is available. If insufficient damage data is available, bins need to be filled up. This filling can lead to a bias and thus unrealistic results. Moreover, if only few data points are present in one bin, the influence of outliers is high.

The choice for environmental parameters in the extrapolation can depend on application. For wake-dominated wind farms, the choice for wind direction or turbulence intensity seems more logical. Of those two, turbulence intensity is preferred over wind direction to reduce the probability for empty bins. If the damage of the turbines would be wave-driven, the choice for a wave-related parameter is more advisable.

Furthermore, it is shown the expected lifetime of a turbine is heavily influenced by the occurrence of parked conditions and rotor stops during its lifetime. It is advisable to calculate multiple damage tables for different operating conditions, i.e. operational, standstill and rotor stops. Moreover a realistic probability of occurrence for all operating conditions is important for the lifetime extrapolation. Moreover, if the environmental probability distribution from design is used to perform a lifetime assessment, a more conservative result is obtained with respect to probability distribution based on a long period of data.



# Chapter 7

## Fleet leader concept

*At this point, the available SCADA data does not contain enough information and the right measurements to perform turbine-specific lifetime assessments. Therefore the installation of additional sensors on the OWT's substructure is necessary. The current practice in industry is to follow up only a few OWTs in the farm by installing additional sensors, e.g. strain gauges, on those. This chapter elaborates on the possibility to extrapolate the measured damages at those instrumented turbines, the fleet leaders, to the other wind turbines in the farm. This extrapolation is mainly based on environmental conditions. After giving an overview of the available data in Section 7.2, first the damage of all available fleet leaders is visualized with respect to the environmental conditions in Section 7.3. In Section 7.4 the concept of extrapolating damage based on environmental conditions is demonstrated on multiple wind farms. However, differences in structural properties proved to influence the measured damage at OWTs greatly. To gain better understanding of these influences, the contribution of different frequency ranges of the stress signal to the accumulated damage is investigated in Section 7.5. Finally Section 7.6 concludes this chapter.*

## 7.1 Introduction

As (offshore) wind farms are growing older, the question about possible lifetime extension is getting more important. Currently, most lifetime assessments require additional instrumentation of the OWT's substructure, such as strain gauges [15]. However, installing additional sensors, especially strain gauges, is still considered quite expensive for the operators. The current practice in industry is to instrument about 10% of the OWTs in a wind farm with additional sensors. This requires the extrapolation of the measurements from those turbines to predict the fatigue life of the other wind turbines. This concept was introduced for blade loading by [107] as the "Fleet Leader Concept" and will be applied on turbine foundations of multiple wind farms in this chapter. First the differences in damage between several turbines are visualized and commented upon (Section 7.3). Afterwards the fleet leader concept is applied based on environmental conditions only (Section 7.4). Finally the differences in damage accumulation are discussed (Section 7.5).

## 7.2 Data

For the analyses performed in this chapter, data of different wind turbines located in different wind farms is used. For each of these wind turbines, strain data and 10 minute statistics of SCADA data was available. Moreover, one set of meteorological data was available. More information about all data can be found below.

For all turbines, a period of 9 months is used to visualize the accumulated damages, while a period of one year is used to validate the fleet leader concept. Only data for which the turbine was operational was used.

### 7.2.1 Strain data

Each of the wind turbines included in this analysis, was instrumented with strain sensors at the interface between tower and transition piece. In all cases, the measured strain signals are translated into a stress signal for fore-aft movement and a stress signal for side-side movement, using the orientation of the wind turbine as given by the SCADA yaw angle. This was explained in Section 2.2.2. The measurement setup was similar for most of the instrumented wind turbines and consisted of four or six classical strain gauges.

The instrumented turbines cover different foundation types and turbine dimensions. Moreover, often the design of the substructure is site specific. Therefore the height of interface between tower and transition piece or water depth can differ as well. This often leads to differences in structural behavior, represented by differences in resonance frequencies, damping values and mode shapes. The difference in design of the substructure might also lead to differences in values for the combined safety factor for the fatigue analysis.

The instrumented turbines are distributed over several offshore wind farms. Usually at least two different turbines in one farm are instrumented. In general, a lot of specifications are identical or very similar among the turbines within one farm.

For example the turbine type and size, the foundation type, the height of TP-TW interface and values for safety factors.

### 7.2.2 SCADA data

For all turbines a subset of SCADA data was available. The subset contained 10 minute averages and standard deviations of wind speed. The turbulence intensity is calculated as well. The available SCADA data was filtered first. The same filters from Section 6.4.2 are applied.

### 7.2.3 Meteorological data

Additionally, 10 minute averages of wave height measurements, taken in one of the wind farms, were available. Given the slow variations in wave conditions and the fact all farms are located fairly close to each other, the wave measurements are assumed to be applicable for all wind turbines.

## 7.3 Damage visualization

Different turbines, with different structural properties, react differently to changes in environmental parameters. In order to visualize these differences, the damage for every 10 minutes is calculated. The same procedures as explained in Chapter 6 are applied. In this case, the same S-N curve (DNV-D-A) and the same value for stress concentration factor, size effect and material safety factor (all 1) is used to calculate the damage. No extrapolation within the structure is performed. Therefore, the damages shown are at the TW-TP interface.

The resulting damages are divided in bins based on the environmental conditions.

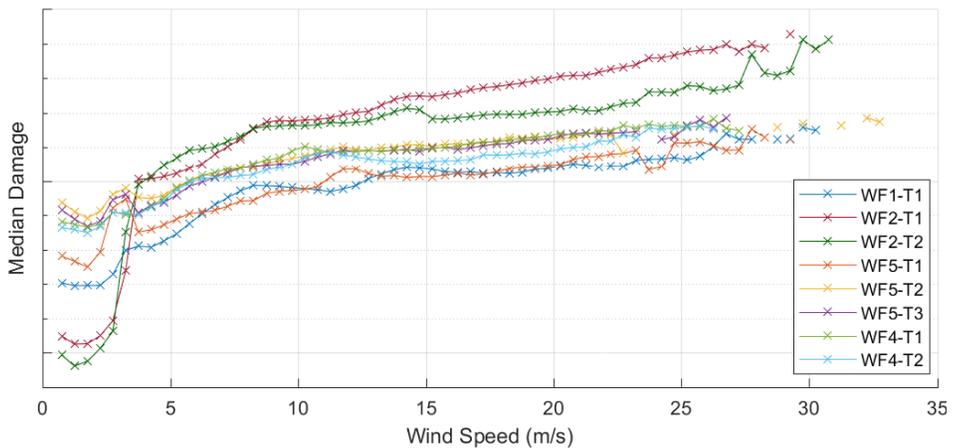


Figure 7.1: Median value of damage vs wind speed for multiple wind turbines across 4 different farms.

Figure 7.1 shows the resulting median damage for each wind speed bin for eight different wind turbines. The curves of the turbines of WindFarm2 (represented by the red and green line) are clearly different from those of the other wind farms. First of all, the median damage in each bin above cut-in wind speed is clearly higher compared to the other turbines. This can be explained by the increased size of the turbine blades and the resulting higher thrust load with respect to the other turbines. However, below cut-in wind speed the damage of the turbines of WindFarm2 is clearly lower. The explanation for this would lie in the foundation type. The turbines of WindFarm2 are installed on jacket substructures, which are less susceptible for waves. Although the wave load for low wind speeds is usually very low, it is still present. Since monopiles are more affected by them, higher damages can be observed.

Since one of the turbines of WindFarm2, WF2-T1, has an additional 1P variation, the median damage is increased with respect to the other wind turbine of WindFarm2, WF2-T2. This increase however is only visible for wind speeds higher than ca.  $9m/s$ .

The turbines installed on monopiles show a similar behavior. Also here an increase in damage around cut-in wind speed can be observed, although much smaller.

Two turbines show a lower damage: WF1-T1 (the dark blue line) and WF5-T1 (the orange line). These are two turbines installed in shallower waters. Moreover, the turbine size of WF1-T1 is slightly smaller.

Furthermore, a difference in damage can be observed between WF4-T1 and WF4-T2. Here WF4-T2 seems to accumulate less damage for wind speeds higher than ca.  $11m/s$ . Although these turbines are installed in the same farm and thus have some specifications, such as turbine type, in common, the differences in structural properties are relatively high between both.

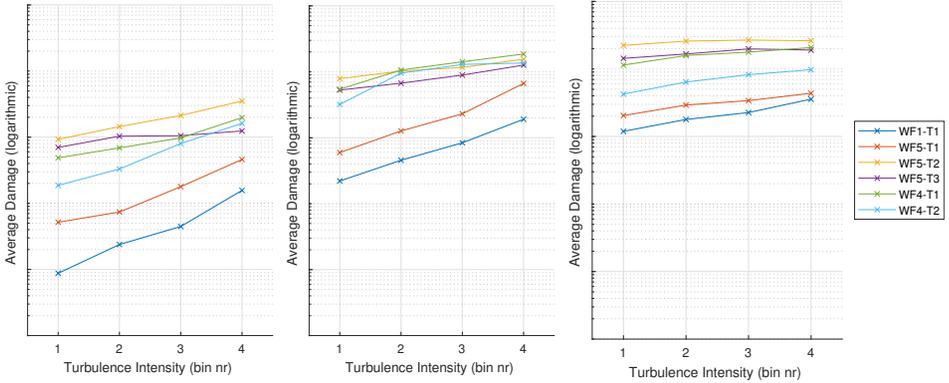
Finally the damage of three turbines seem to match very nicely: WF5-T2, WF5-T3 and WF4-T1. This is despite the known differences in structural properties.

Given the clear difference in damage behavior and accumulation with respect to wind speed, the two turbines of WindFarm2 will be considered separately from the others in the remainder of this section.

Figures 7.2 and 7.3 show the dependency of damage on turbulence intensity for turbines installed on monopiles and jackets respectively. Each sub-figure shows the median damage of the bins, versus the turbulence intensity (indicated as bin number). The bin borders of the turbulence intensity bins depend on the wind speed and are calculated for each turbine individually as explained in Section 6.3.1 using the first, second and third quartile. All axis limits of the sub-figures are taken exactly the same for each figure individually.

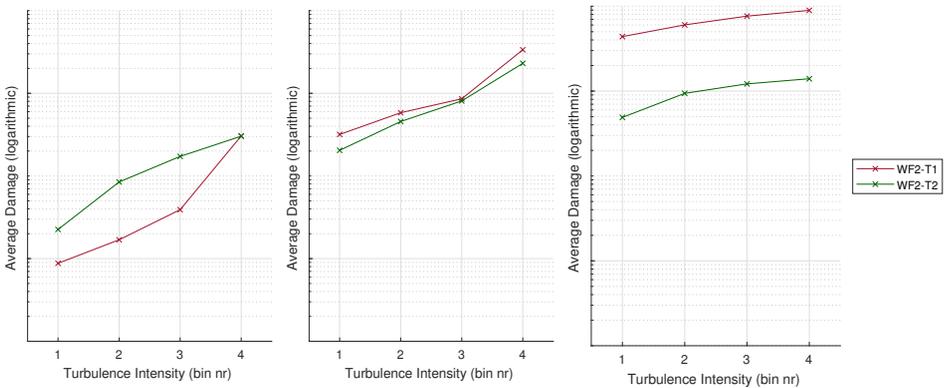
In general the damage increases for increasing turbulence intensity. Only for the higher wind speed of  $18m/s$  approximately, the damage seems almost constant for WF5-T2, WF5-T3 and WF4-T1. In general the same observations as in Figure 7.1 can be made. Moreover most curves seem comparable to each other in the same wind speed bin. However, small differences can be observed. In Figure 7.2 the damage of WF1-T1, WF5-T1 and WF4-T2 seem to differ more for different

turbulence intensity bins and thus to be more influenced by the turbulence intensity than the damage of the other turbines. In Figure 7.3a the shape of the curves of both figures differs significantly. However, for the other wind speed bins, the curves are very comparable.



(a) Wind speed between 4, 5 and 5, 5m/s      (b) Wind speed between 10, 5 and 11, 5m/s      (c) Wind speed between 17, 5 and 18, 5m/s

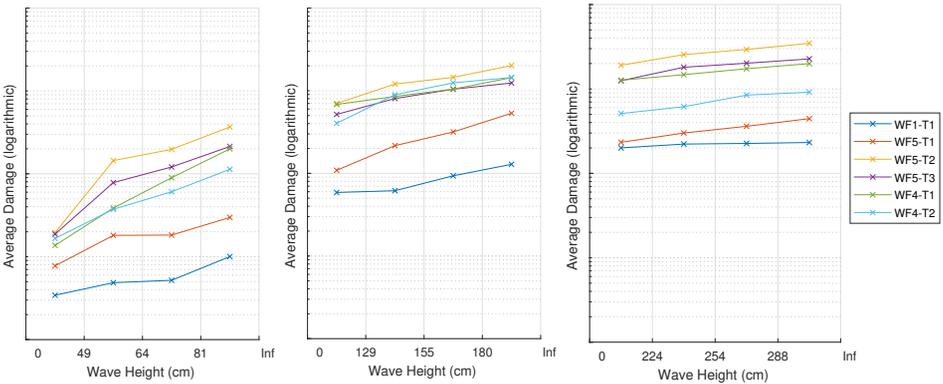
Figure 7.2: Average value of damage vs turbulence intensity for multiple wind turbines, all installed on monopiles. Data shown corresponds to one wind speed bin each figure



(a) Wind speed between 4, 5 and 5, 5m/s      (b) Wind speed between 10, 5 and 11, 5m/s      (c) Wind speed between 17, 5 and 18, 5m/s

Figure 7.3: Average value of damage vs turbulence intensity for multiple wind turbines, all installed on jackets. Data shown corresponds to one wind speed bin each figure

Finally also the dependency of the damage on wave height is shown for the different turbines. Figures 7.4 and 7.5 show the median damage of the specific bins for turbines installed on monopiles and jackets respectively. Each sub-figure focuses on only one wind speed bin. The bin borders again depend on the wind speed. However, since the (significant) wave height measurement is the same for all turbines, the bin borders are defined based on the wind speed bins of one turbine (WF1-T1). Therefore the same bin limits for wave height are used for all turbines within one wind speed bin. These are also indicated in the figures. Moreover, all axis limits of the sub-figures are again taken exactly the same for each figure individually.

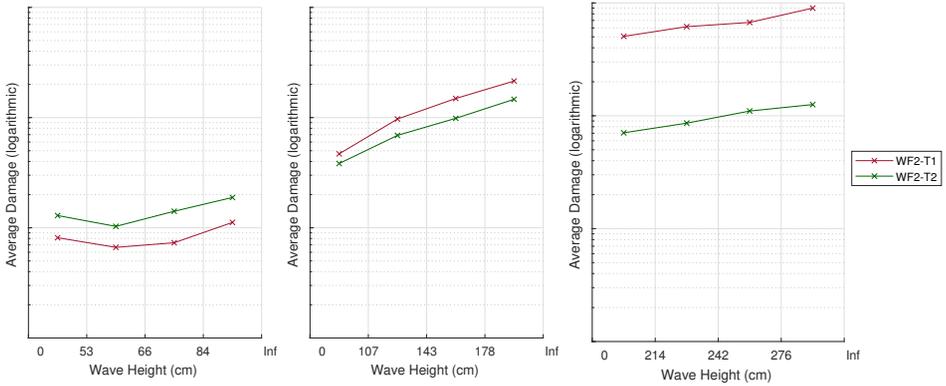


(a) Wind speed between 4,5 and 5,5m/s (b) Wind speed between 10,5 and 11,5m/s (c) Wind speed between 17,5 and 18,5m/s

Figure 7.4: Average value of damage vs wave height for multiple wind turbines, all installed on monopiles. Data shown corresponds to one wind speed bin each figure

A first observation can be made for all turbines. Usually the damage increases with increasing wave heights. In Figure 7.4, the damage of WF1-T1 and WF5-T1 seem to be influenced less by the wave height than the damage of the other turbines. These two turbines are the shallowest, which can explain this difference. For higher wind speeds (above 10m/s), WF5-T1 seems to be influenced a bit more than WF1-T1. This is probably because of the slightly higher turbine size (and thus bigger monopile) of WF5-T1.

For all other turbines the curves seem very comparable to each other. This is also true for the turbines installed on jackets (Figure 7.5).



(a) Wind speed between 4,5 and 5,5m/s (b) Wind speed between 10,5 and 11,5m/s (c) Wind speed between 17,5 and 18,5m/s

Figure 7.5: Median value of damage vs wave height for multiple wind turbines, all installed on jackets. Data shown corresponds to one wind speed bin each figure

## 7.4 Simple fleet leader

In Section 7.3 one could see the damage is influenced by, among others, the measured wind speed, wave height and turbulence intensity. Given the similarities seen between the damage of turbines, the question rises whether one could extrapolate measured damage at one turbine (the fleet leader) to another turbine solely based on environmental conditions. In this section, it is checked if the damage of one turbine can be estimated using another turbine. Thus one turbine is chosen as a fleet leader and the extrapolated damages of the other turbines are compared to the measured damage. The extrapolation is done twice for each turbine, once based on wind speed only and a second time based on wind speed in combination with turbulence intensity.

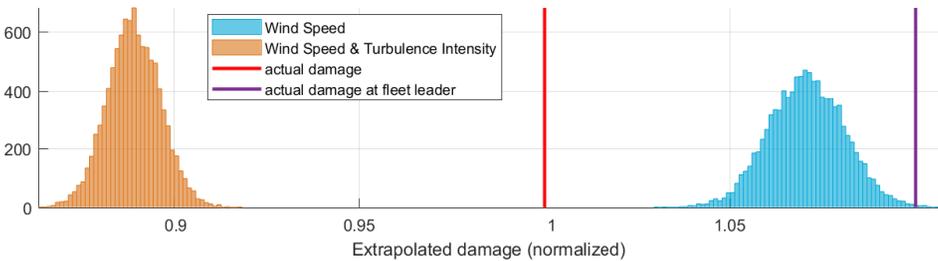
To simplify the analysis, only data points for which the turbine is operating are considered. Moreover, the damage is calculated as if the turbine would have been operating a full year in operating conditions in the same environmental conditions as during the remaining data points.

Basically, the method used is similar to the one used in Chapter 6 to calculate the expected lifetime of an instrumented wind turbine. Only in this case the environmental distribution used to extrapolate is the one as measured at the desired wind turbine and no lifetime is calculated using the extrapolated damage. Moreover the damage is only calculated for the sensor location, including using the specific safety factors as given by the individual design documents.

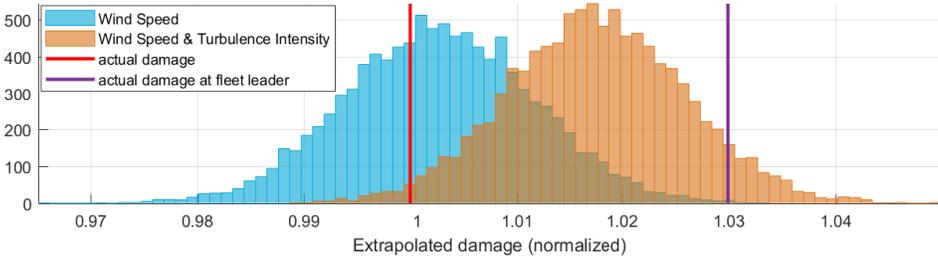
### 7.4.1 Within one farm

#### WindFarm3

First, the extrapolation is performed in an older wind farm, WindFarm3. Here all monopiles installed in the wind farm have the same design. Therefore, structural properties of all turbines are very comparable. Figure 7.6 shows the results where the damages measured at the fleet leader, WF3-T1, were extrapolated on two other wind turbines in the farm, WF3-T2 and WF3-T3. The resulting extrapolated damages are normalized with respect to the actual measured damage at WF3-T2 and WF3-T3 respectively. In blue the extrapolated damage based on the wind speed distribution measured at WF3-T2 or WF3-T3 is shown. The red vertical line shows the amount of damage actually measured at WF3-T2 or WF3-T3, whereas the purple vertical line indicates the amount of damage measured at the fleet leader WF3-T1.



(a) Extrapolated and measured damage at WF3-T2



(b) Extrapolated and measured damage at WF3-T3

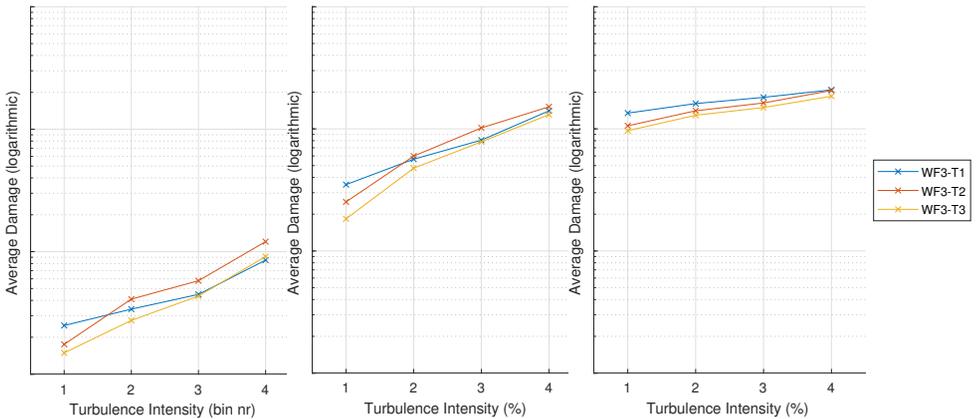
Figure 7.6: Damage extrapolation of measured damage at WF3-T1 to two other wind turbines, based on wind speed only (blue) and a combination of wind speed and turbulence intensity (orange). Results are normalized with respect to the actual damage measured at the extrapolation turbine, which is shown in red. The actual damage measured at the fleet leader is shown in purple.

In general, differences up to only 10% of the measured damage are found between the extrapolated damage and the actual damage. This is an improvement over taking the measured damage at the fleet leader as is for every turbine in the farm. This illustrates the possible gain in damage extrapolation using the fleet leader concept for turbines with comparable structural dynamics.

Taking a closer look at the results, one can observe an almost perfect match between the center of the blue distribution for WF3-T3 and the actual damage (Figure 7.6b).

However, such a perfect result was only obtained for an extrapolation based on wind speed exclusively. In orange the result for the extrapolation based on the combination of wind speed and turbulence intensity is given. An overestimation of only ca 2% can be observed based on the median of the extrapolated damage distribution with respect to the actual measured damage. This is caused by small differences in the measured damage tables.

The difference in damage table for all turbines of WindFarm3 is illustrated by Figure 7.7. Here, the damage tables of all turbines were composed based on 2d binning using wind speed and turbulence intensity. The values of the bin borders for turbulence intensity bins are different for every wind speed bin. These values were established based on the environmental data of the fleet leader and used for the other turbines as well. The resulting damage values are shown versus the turbulence intensity, for a few wind speed bins. It can be seen the damage accumulated by WF3-T1 is often higher than the damage accumulated by WF3-T3 although the environmental conditions were exactly the same, especially for low turbulence. This causes a (slight) overestimate of damage when WF3-T1 is taken as the fleet leader to estimate the damage at WF3-T3.



(a) Wind speed between 4,5 and 5,5 m/s  
 (b) Wind speed between 10,5 and 11,5 m/s  
 (c) Wind speed between 17,5 and 18,5 m/s

Figure 7.7: Average value of damage vs turbulence intensity for all turbines of WindFarm3. The data shown in one sub-figure corresponds to one wind speed bin.

When looking at the results of the extrapolation of WF3-T1 on WF3-T2, differences between measured and extrapolated damage are clearly higher. Here, the difference between predicted damage and the actual damage goes up to ca. 13%. However, this is still reasonable.

For this extrapolation, a higher difference between the two types of extrapolation

can be observed. This can be explained by a difference observed in environmental conditions at the turbines. When looking at the probability of turbulence intensity (Figure 7.8), WF3-T2 sees a little bit less turbulence than the other wind turbines. This is caused by its location in the farm, being at the edge. Due to this lower turbulence, during the extrapolation more weight will be given to the lower turbulence bins and thus often lower damage.

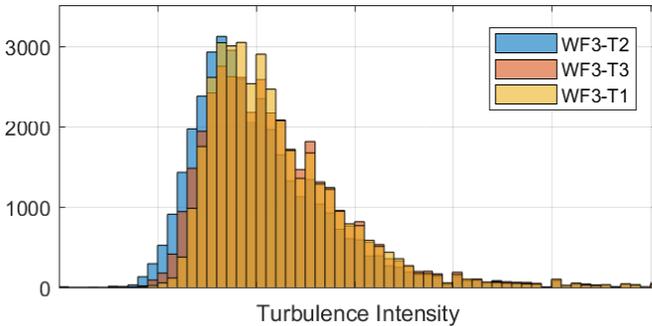


Figure 7.8: The occurrence of turbulence intensity at WF3-T2 in blue, WF3-T3 in red and WF3-T1 in yellow.

## WindFarm2

The same concept is applied on a second wind farm, where the turbines are installed on jackets. Looking at the results given in Figure 7.9, none of both extrapolations is successful. The estimated damage is only 20 to 38% of the actual measured damage. This is explained by the existence of a rotor imbalance at WF2-T1. Therefore the turbine accumulates more damage due to an additional cyclic loading and the extrapolated damage based on the damage measurements at WF2-T2 is underestimating the actual damage.

Moreover, a big difference between both extrapolations (about 15% of the actual damage) can be observed. This can be explained by the difference in position in the farm. The fleet leader, WF2-T2, is surrounded by other turbines in all wind directions and thus experiencing a lot of turbulence. The extrapolated turbine however has free wind flow for the majority of the wind directions. Therefore the measured turbulence intensity at this turbine will often be a lot lower. When extrapolating based on wind speed only, this distinction in turbulence intensity is not made. Although, when extrapolating based on wind speed and turbulence intensity, the extrapolated damage will be dominated by the lower turbulence intensity bins with lower average damages. Therefore the extrapolated damage will be a lot lower compared to the extrapolated damage based on wind speed only.

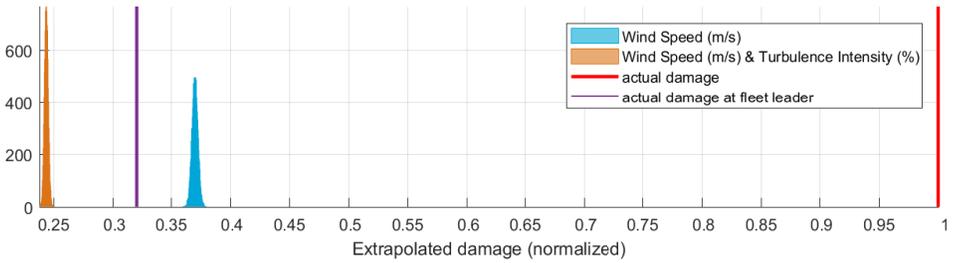


Figure 7.9: Extrapolated damage distribution of WF2-T1 based on measured damage at WF2-T2, compared to actual damage. Extrapolation is done based on wind speed only (blue) and a combination of wind speed and turbulence intensity (orange)

### WindFarm4

The next wind farm at which the concept was applied has two instrumented turbines. Here the structural properties of the two turbines are not comparable to each other. Looking at the results given in Figure 7.10, none of both extrapolations is successful. The damage was overestimated by 60 to 70% of the actual measured damage. This difference is not only explicable by small differences in environmental conditions. Nor suffered one of the turbines from an additional 1P oscillation. The main reason for this difference in damage can be explained by a difference in structural properties.

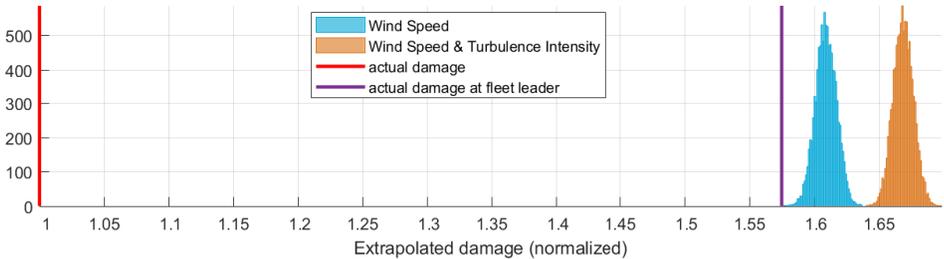
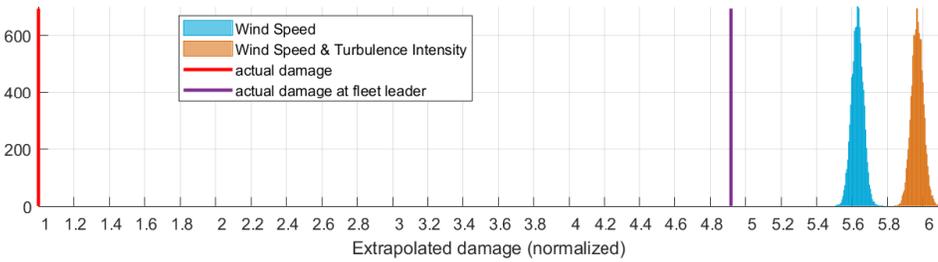


Figure 7.10: Extrapolated damage distribution of WF4-T2 based on measured damage at WF4-T1, compared to actual damage. Extrapolation is done based on wind speed only (blue) and a combination of wind speed and turbulence intensity (orange)

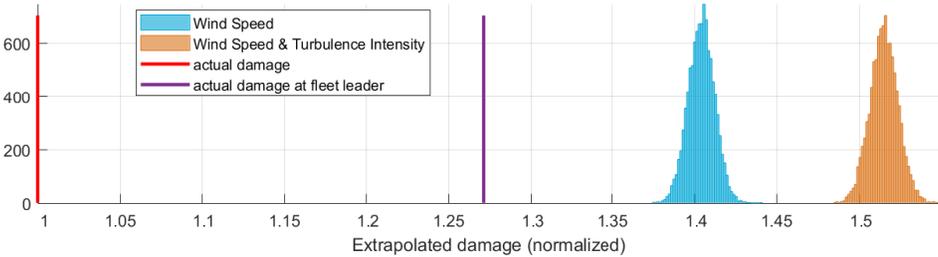
### WindFarm5

The last wind farm where the concept was tested upon was WindFarm5. Here WF5-T2 was chosen as a fleet leader. The extrapolated damage of WF5-T2 on WF5-T1 and WF5-T3 are shown in Figure 7.11. Results show a big difference among the two turbines. The extrapolated damage based on WF5-T2 on WF5-T1

is a little less than 6 times bigger than the measured damage at WF5-T1. This is caused by the big difference in accumulated damage between both turbines, as was already established in previous section based on Figure 7.1. This difference in accumulated damage is because the structural properties are very different for both turbines. The results for WF5-T3 on the other hand are much closer to the actual damage. Here, the difference is only 40 to 55% of the actual measured damage at WF5-T3. This might indicate the structural properties of WF5-T2 and WF5-T3 are more comparable. This could already be expected, since comparable damages were measured for both turbines (Figure 7.1). However, compared to WindFarm3, the difference is still quite big and is probably caused by (little) differences in structural dynamics.



(a) Extrapolated and measured damage at WF5-T1

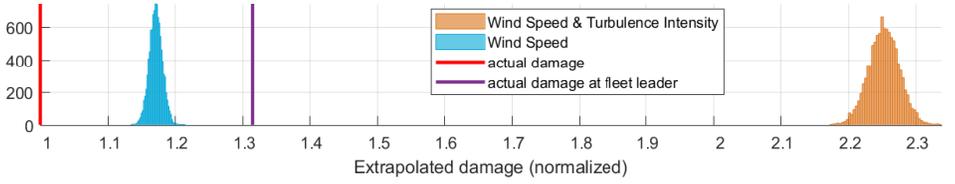


(b) Extrapolated and measured damage at WF5-T3

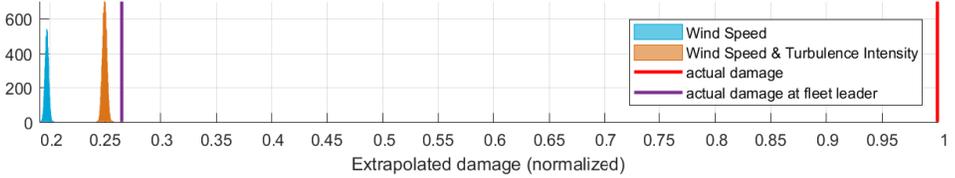
Figure 7.11: Damage extrapolation of measured damage at WF5-T2 to two other wind turbines, based on wind speed only (blue) and a combination of wind speed and turbulence intensity (orange)

### 7.4.2 Across different wind farms

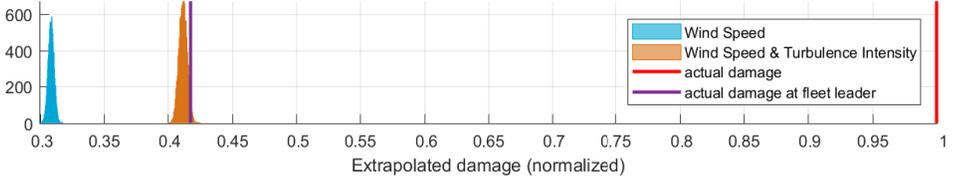
Taking it one step further, one could try to extrapolate the damage at one turbine in one farm to another turbine in another farm. In reality, this might be useful for wind farms where no turbine is instrumented, but where instrumented turbines are present in a neighboring farm. Figure 7.12 shows the results for WF5-T1 as a fleet leader.



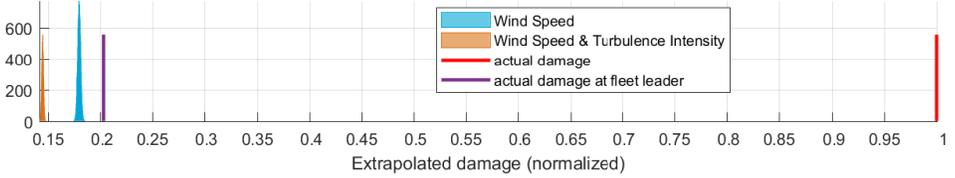
(a) Extrapolated and measured damage at WF1-T1



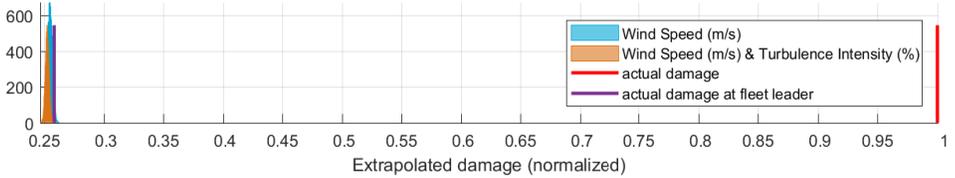
(b) Extrapolated and measured damage at WF4-T1



(c) Extrapolated and measured damage at WF4-T2



(d) Extrapolated and measured damage at WF5-T2



(e) Extrapolated and measured damage at WF5-T3

Figure 7.12: Extrapolated damage distribution of turbines comparable in size installed on monopiles based on measured damage at WF5-T1, compared to actual damage. Extrapolation is done based on wind speed only (blue) and a combination of wind speed and turbulence intensity (orange)

It can be seen, the damage at most turbines is underestimated except for WF1-T1. It was already seen in Section 7.3 that the damages of WF1-T1 and WF5-T1 were comparable, despite their location in different farms and different turbine types. This is now confirmed with a difference of only 20% in case of extrapolation based on wind speed only. As could be seen in Figures 7.1 and 7.2, the damages of both turbines were not for all wind speed and turbulence intensity bins comparable. This probably explains the difference in both extrapolation methods. If more bins with comparable damage get higher weights by the 1d extrapolation compared to the 2d extrapolation, a better result is obtained.

For all other turbines, the extrapolated damage is maximum a little over 40% of the actual measured damage. This could be expected given the high difference in damage between WF5-T1 and the other turbines, as shown in Section 7.3. Again, the differences can be explained by different structural dynamics.

Additionally, the same exercise is repeated for different fleet leaders. Results for the extrapolation based on wind speed only are given in Table 7.1. The values given in the table indicate the median of the extrapolated damage distribution, normalized with respect to the actual measured damage in %. Only for two combinations acceptable results are obtained, surprisingly enough none within the same wind farm. The first couple of turbines is WF1-T1 and WF5-T1. For these turbines, comparable damages were shown in Section 7.3 indeed. The second couple of turbines is WF4-T1 and WF5-T2.

Table 7.1: The median of extrapolated damage distribution, normalized with respect to the measured damage at the extrapolated turbine, for different fleet leaders and different extrapolated turbines

		$\frac{med(D_{extr})}{D_{meas}}$ (%)	Extrapolated Turbine				
		WF1-T1	WF4-T1	WF4-T2	WF5-T1	WF5-T2	WF5-T3
Fleet Leader	WF1-T1	100,00	16,67	26,85	<b>88,45</b>	15,73	21,83
	WF4-T1	598,28	99,99	160,92	519,22	<b>92,23</b>	130,55
	WF4-T2	375,90	64,32	100,00	313,06	57,49	80,34
	WF5-T1	<b>117,16</b>	19,76	30,86	99,98	17,99	25,49
	WF5-T2	642,96	<b>109,09</b>	177,12	563,31	100,00	140,44
	WF5-T3	455,98	76,58	124,57	401,09	70,73	99,99

In general, the fleet leader concept with extrapolation based on environmental conditions only does not perform accurately. Clearly not only different environmental conditions cause difference in damage. Also differences in structural properties are important. For example differences in resonance frequencies can cause differences in accumulated damage. The closer the wave frequency gets to a resonance frequency of the structure, the more damage will be accumulated [108]. The same is true if the frequency of variations in wind or the frequency of rotor dynamics get close to a resonance frequency. Moreover the damping values of each structure can be important as well, just like the mode shapes.

Additionally, differences in as designed values for safety and correction factors or S-N curves should be taken into account.

This simple application of the fleet leader concept does however work for similar turbines, of the same type, same foundation and similar dynamic and structural behavior. Unfortunately, designs of foundations in newer wind farms are more and more specific to site and location and thus rarely comparable to all other foundations in the farm. Identifying and quantifying these differences between the different structures and their effect on damage accumulation is key to ensure a correct damage extrapolation. This will get even more important for the bigger wind turbines installed on bigger monopiles. Due to the increase in height and width of the monopile and the resulting decrease in resonance frequencies, wave loading will get more and more important for a fatigue assessment.

## 7.5 Contribution of specific loads to damage

In Section 7.4 it is shown that the damage of one turbine can only be extrapolated to another turbine based on differences in environmental conditions exclusively, if the structural properties of the turbines are almost identical. For turbines where the structural properties are different, a (significant) mismatch is observed. In this section, the contribution to the total damage of different loads acting on the structure is examined. Since several loads act within different frequency ranges, the analysis is based on filtering the original stress signal for different frequency bands. For all bands the lowest admissible frequency is set to 0, the highest admissible frequency is varied from 0,02Hz to 5Hz in a step of 0,02Hz. Every filtered stress signal is cycle counted and translated into a damage value. The same S-N curve is used for all stress signals and no correction factors were applied. The resulting array of damages represent the accumulated damage for an increasing frequency band.

The analysis is done for multiple turbines and for three different days. The days are chosen in such a way each of them shows a different operation state. During one of them only operating conditions occur. For another day very high wind speeds were measured and all turbines stopped producing power at a certain point due to cut-out. During the last day, the wind speed increased from very low to normal. Therefore that day represents cut-in conditions

In all figures in this section, a frequency band  $i$  is represented by the maximum value of the band, indicated as cut off frequency  $f_i$  (Hz). Moreover, instead of showing the accumulated damage, the damage contribution for each cut off frequency is shown. This is calculated by subtracting the damage for the previous (slightly smaller) frequency band  $[0, f_{i-1}]$  from the damage obtained for the new frequency band  $[0, f_i]$ . This is summarized by Equation 7.1.

$$D_{contr, f_i} = D_{total, [0, f_i]} - D_{total, [0, f_{i-1}]} \quad (7.1)$$

Figure 7.13 shows the damage contributions (on a logarithmic scale) for all turbines for a normal day. It can be seen the damage of all turbines is dominated by frequencies up to 0,5 Hz, or in case of WindFarm2up to 0,7 Hz. The loads acting in this frequency range are the quasi-static thrust load, the cyclic loading as induced by wave loads, the first mode of the structure and possibly some rotor dynamics (1P and 3P).

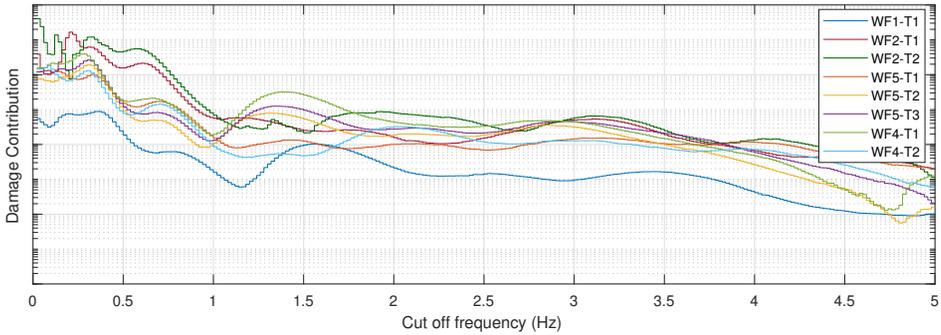


Figure 7.13: The contribution to the total damage by each specific frequency band during a normal operational day for multiple turbines

It can be seen the turbines of WindFarm2 not only have higher damage contributions, but also the general shape of the curves differ. The higher damage can be explained by the bigger turbine, the difference in shape is mainly caused by the different foundation.

Clear differences can be distinguished among the turbines of WindFarm2 as well. This is better illustrated by Figure 7.14, where a detail is taken out of Figure 7.13.

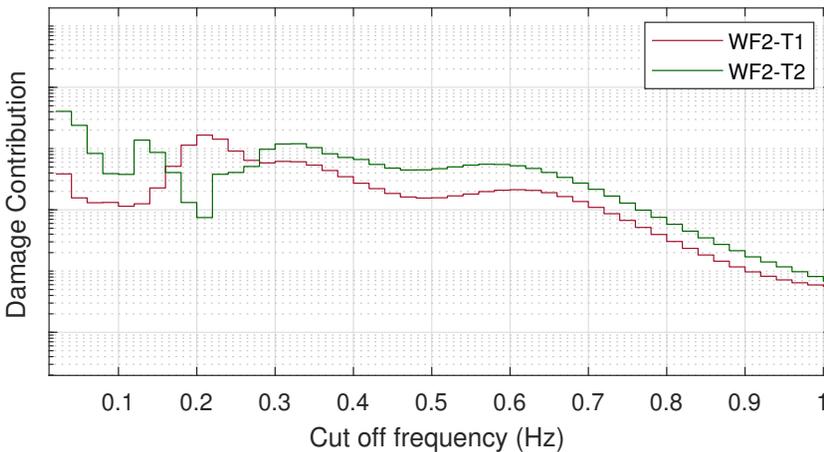


Figure 7.14: The contribution to the total damage by each specific frequency band during a normal operational day for the turbines of WindFarm2

The contribution of the 3P movement (around 0,6 Hz), caused by the periodic passing of the blades, is significant for both turbines of WindFarm2. This is probably caused by the bigger size of the turbine. In general it can be seen WF2-T2 has higher damage contributions than WF2-T1, except for the region around 0,2Hz. As already mentioned, WF2-T1 has an additional cyclic loading in this frequency band at 1P, induced by a rotor imbalance.

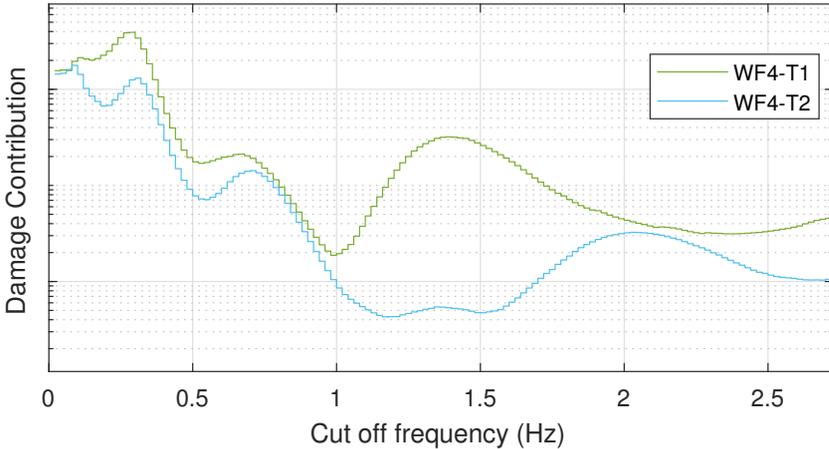


Figure 7.15: The contribution to the total damage by each specific frequency band during a normal operational day for the turbines of WindFarm4

A closer look is taken to the turbines of WindFarm4. In previous Sections 7.3 and 7.4 it was clear the difference in damage between both turbines was relatively high. In Figure 7.15 the difference in damage contribution can be seen. It is clear WF4-T1 accumulates a lot more damage than WF4-T2. The first resonance frequency is a bit lower for WF4-T1 than for WF4-T2, both around 0,3 Hz. Although it seems a small difference, the slightly lower resonance frequency of WF4-T1 is closer to the wave frequency as seen on the site that day. Therefore the first mode of WF4-T1 is more excited by the wave frequency than the first mode of WF4-T2. The other big difference between both turbines is the contribution of the second resonance frequency. Here, the difference is over 0,5 Hz: for WF4-T1 the second natural frequency is slightly lower than 1,5 Hz, while the one for WF4-T2 is around 2 Hz. Due to this difference, the second mode of WF4-T1 is far more excited by the 6p rotor harmonic than the second mode of WF4-T2. Again, this causes more damage at WF4-T1 than at WF4-T2.

Figure 7.16 shows the results for a very windy day, with wind speeds above cut-out. The most important observation is the importance of the first mode compared to the quasi-static contribution. For all turbines (except WF2-T1), the damage is dominated by the first order movement of the tower (and substructure). In general, turbines installed on monopiles seem to be more sensitive for damage at high winds than turbines installed on jackets. It is to say, for both turbines of WindFarm2 the first order movement does not seem to dominate entirely, since the 3P rotor harmonics contributes still quite a lot. For all other turbines, it is definitely the first mode that dominates the damage accumulation.

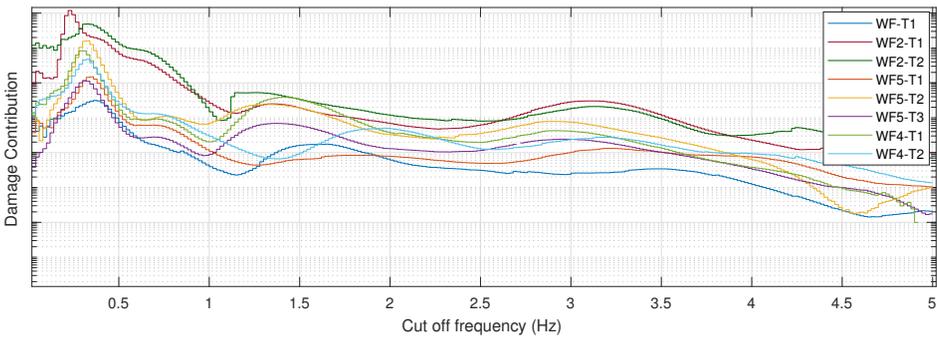


Figure 7.16: The contribution to the total damage by each specific frequency band during a very windy day, up to cut-out, for multiple turbines

Figure 7.17 shows the results for a day which started with very low wind speeds and ended with wind speeds around rated wind speed. Here the damage is clearly dominated by the thrust load and the first mode. For frequencies higher than 0,5 Hz, the contribution become very low and even negative (hence missing data points on the figure). The dominance of the thrust load in this case might not surprise since the variation in wind speed this day was very high. Big, though slow, stress cycles were introduced due to the wind varying between cut-in and rated.

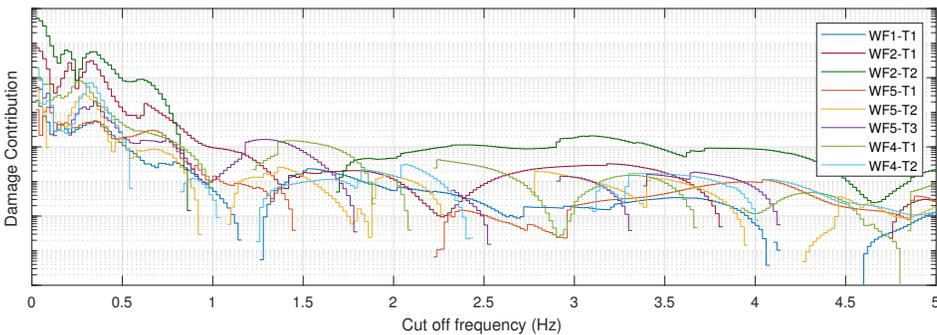


Figure 7.17: The contribution to the total damage by each specific frequency band during a gentle day, starting from cut-in, for multiple turbines

## 7.6 Conclusion

In this chapter the accumulated damage of multiple instrumented turbines was compared to each other based on environmental conditions. Although differences in measured damage could be observed among the turbines, mainly due to different structural properties, their relation to environmental conditions seemed very similar. The difference in structural properties among the turbines can be explained by a different foundation, a difference in water depth, a different design, a different turbine, ...

Given the similar influence of environmental conditions on damage of multiple turbines, the fleet leader concept is believed to be valid for those wind farms. Therefore, one of the instrumented turbines was chosen as a fleet leader for the farm. The measured damage at the fleet leader was extrapolated towards the other instrumented wind turbines and compared to the actual measured damage. Results of an older wind farm consisting of OWTs with very similar structural properties proved the concept, with errors of maximum ca. 15%. However, it was also shown complicating the extrapolation by including the turbulence intensity is not always beneficial for the results.

Results of a different wind farm showed the concept is not valid if one of the turbines has a defect, such as a rotor imbalance.

The same concept was applied on two other wind farms as well, where the difference in structural properties between the OWTs are bigger. Also for these wind farms no successful extrapolation was obtained.

A last application of the fleet leader was across different wind farms. Although the turbine type usually differed from one farm to another, some OWTs seemed surprisingly comparable. This indicates the structural properties of the OWT have a bigger influence on the accumulated damage than the exact turbine type, if the turbines are still comparable in terms of rated power and size. Moreover, the differences in as designed fatigue values, such as safety factors, correction factors or S-N curves, should be taken into account.

As the fleet leader concept based on environmental conditions was not successful for most turbines, the reason for differences in damage was looked for as well. In the last section of this chapter, a look was taken at the damage contribution of different frequency bands.

With these insights the fleet leader concept could be updated by applying it separately for different frequency bands. A correction factor specified for each frequency band based on the structural properties of the OWT can then be applied. However the interaction between several frequency bands will be lost in this approach. Moreover, to quantify the needed correction factors, a very large number of instrumented turbines would be needed.

An alternative for damage extrapolation within the farm is installing an additional accelerometer on every turbine to capture the dynamic loads. In combination with 1s SCADA a stress history can be reconstructed for every turbine, as explained in Chapter 5. This stress history can then be used to perform a lifetime assessment on each turbine individually as explained in Chapter 6.

A last alternative given in this thesis relies on the successful results among OWTs

with similar structural properties. Usually the OWTs in an offshore wind farm can be divided into clusters. Within one cluster, all turbines have a similar design for the foundation. Therefore the structural properties like mode shapes, resonance frequencies and damping are comparable as well. This would mean the fleet leader concept could be applied within one cluster and thus at least one turbine in each cluster should be instrumented.

# Chapter 8

## Conclusions

### 8.1 Conclusions

For this PhD thesis, SCADA was extensively utilized, both 10min statistics and 1s data. Multiple applications were elaborated upon. First of all, a couple of techniques based on 10min SCADA data were proposed for performance monitoring in Chapter 3. All these techniques relied on the calculation of the actual power curve of every turbine. Therefore, different techniques to calculate the power curve were compared. The techniques for performance monitoring proved ability to detect under performing turbines within a farm, as well as turbines whose production is rapidly decreasing. However, the quality of the SCADA data proved to be of big importance for the reliability of the results. More specifically the quality of the wind speed measurement seemed to be insufficient for some turbines. The reason would be the location of the anemometer, just behind the rotor. A workaround was given by estimating the rotor effective wind speed. Unfortunately since no data from an independent met mast was available, the method of rotor effective wind speed could not be validated. The existing issues in the SCADA data however, seem to have improved by using the REWS.

In Chapter 4, the transition towards fatigue assessment and high frequent SCADA data was made. An important element in a fatigue assessment is a reliable representation of the actual loads acting on the wind turbine. One of these loads can be estimated using 1s SCADA data, the thrust load. The estimation of the thrust load is done using a neural network. A correlation analysis (both linear and nonlinear) was performed to decide on the required input parameters. Wind speed, rotor speed, blade pitch angle and generated power proved to be essential. The suggested method was applied and validated on a variety of data sets. Using a dataset obtained by simulations, the concept was proven in a controllable environment.

Moreover the technique was validated on an offshore wind turbine with SCADA data of good quality. Very good results were obtained when the model was trained to estimate only under generating conditions. However, once standstill and idling were included in training and validation dataset, results showed that some im-

provement might be useful for non-generating conditions.

The resulting neural network was also validated on a different turbine of the same type. Good results were obtained for this cross validation. This confirmed the transferability of one thrust model among all turbines of the same type.

However, the technique proved to be very sensitive to the quality of the SCADA data. When applying exactly the same technique to a different wind turbine with SCADA data of less quality, strong filters were needed to avoid unrealistic results. For the remaining data, results were acceptable but inferior to previous results. It is believed these inferior results were caused by less accurate SCADA data.

The accurate estimation of only one load acting on the wind turbine is not enough to perform an accurate and thorough fatigue assessment. Since all other loads are not correlated to SCADA data or they are acting in a frequency range too high to estimate with 1 Hz data, an alternative approach for full load estimation is given in Chapter 5. Two different approaches are compared to estimate the stress history at any location in the structure. Both approaches combine 1s thrust estimates with additional accelerometers installed in the tower. One of the approaches uses a modal decomposition and expansion technique, where the signal is divided into different parts based on the frequency spectrum. Each part is then processed differently. The second technique utilizes a Kalman filter, where no division is necessary. Both techniques are validated on an operating offshore wind turbine. Despite the drawback of inevitable division, the technique of modal decomposition and expansion still outperforms the Kalman filter. Thus, the technique based on the Kalman filter needs more improvement.

Moreover a lifetime assessment, using state-of-the-art techniques, is performed for every weld of the substructure of an fictional offshore wind turbine in Chapter 6. The approach consisted in the cycle counting based on rainflow counting of the stress signal and a damage calculation based on the Palmgren-Miner rule and a predefined S-N curve. The measured stress signal was extrapolated to all specific locations of the welds. In this case a simplified extrapolation was applied, being a simple static extrapolation. Based on the guidelines as used in industry the stress signal is multiplied by several safety factors compensating the difference in dimensions between the actual components and the test specimens used to compose the S-N curve. The value of these safety factors and the choice for S-N curves is based on recommendations given in industrial standards as well.

Damages are calculated every 10 minutes and linked to environmental and operating conditions based on SCADA and other meteorological data. Based on these conditions, a damage table is composed and the measured damages are divided in the appropriate bins. Afterwards the resulting damage table is multiplied with the expected environmental probabilities to obtain the expected accumulated damage for each considered operating state. Finally all resulting elements are superposed to obtain an expected yearly damage and thus expected lifetime. The uncertainty of the lifetime results is obtained by sample-based bootstrapping.

During the described procedure, a lot of options and choices are possible. If possible, the same values were chosen as during design. However, the possibilities for environmental and operational probabilities were investigated a little deeper. Comparison is made between the probabilities used in design and probabilities

obtained from a long period of data. Moreover, the choice of (number of) environmental parameters on which the probability histogram depends is given some more thought. Finally, also the importance of including non-operational data, representing standstill or rotor stops, is illustrated.

In the last Chapter 7 the damages measured at multiple turbines, each with its own specifications, was compared. The ultimate goal was to validate whether the damage measured at one turbine could be extrapolated accurately to other turbines in the farms. To obtain this, the dependence of damage on environmental conditions was compared. Significant differences were observed between turbines with a different foundation and significant difference in size.

The fleet leader concept was applied on the turbines based on environmental conditions only. The concept was proven using a wind farm consisting of OWTs with very similar substructures and structural properties. However, some limitations were shown as well. Acceptable differences in damage tables can be enlarged by the extrapolation, especially when multiple environmental parameters are utilized. Moreover, if one of the turbines suffers from an additional load compared to the others, e.g. a rotor imbalance, the extrapolation results won't hold.

The same concept was tested on other offshore wind farms as well. Unfortunately, the damages of most turbines proved to be too different and results differed significantly from the measured damage. These increased differences are mainly explained by differences in structural properties.

Given the big differences in damages, a final analysis was performed to find out in which frequency bands the most damage was accumulated for each turbine individually. This was very different for all turbines and differed significantly for different operating conditions. The difference in structural dynamics, indicated by resonance frequencies, mode shapes and damping values, between the turbines is considered the main driver for the different results.

In conclusion, the analyses done in this thesis can significantly increase the use of the available SCADA data for both performance monitoring and lifetime assessment. However, results of multiple analyses showed an important dependency on the quality of the SCADA data. Therefore proper preprocessing and filtering of SCADA data is crucial.

## 8.2 Future work

The possible improvements for the work presented in this thesis can be divided in two main categories. First of all, the different steps taken towards a full farm fatigue assessment should be combined. This is explained in more detail in Section 8.2.1. On the other hand, improvements specific to the different parts are proposed as well in Section 8.2.2.

### 8.2.1 Developing a full monitoring approach

A lifetime assessment where most of considered techniques are combined should be done. As explained in Section 1.3, the ultimate goal of the lifetime assessment approach as proposed in this thesis was to have an accurate indication of the

consumed and remaining lifetime of all critical locations for all wind turbines within one farm.

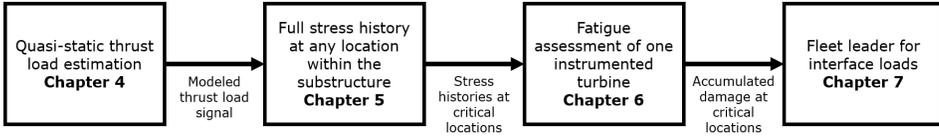


Figure 8.1: The obtained objectives (given in the different boxes) and the missing pieces (indicated by the connecting arrows) to perform an accurate farm-wide fatigue assessment of an offshore wind farm

Figure 8.1 shows the different steps taken to achieve this objective. Although the main methodology was implemented and tested successfully for each of these steps, the feed-through of information between the different steps is often still missing. In order to obtain this, additional validation or an improvement of the proposed techniques is needed.

First of all, the thrust load model based on SCADA data is used to estimate the quasi-static load. This load signal should then be combined with acceleration measurements using MDE or a Kalman filter, in order to obtain a reconstruction of the stress history at multiple locations, preferably weld details, in the substructure of the OWT. This should be validated using reliable SCADA data and for multiple locations within the structure, preferably under water and even subsoil.

Furthermore a lifetime assessment is performed for each of these locations, as explained in Chapter 6. However, instead of using a static extrapolation factor, every reconstructed stress signal is cycle counted individually. For this lifetime assessment, it is important the reconstructed stress signal represents a period of at least 9 months and preferably over one year. Moreover the environmental parameters based on which the extrapolation will be done, should be chosen carefully. It is advisable all resulting bins include enough data, to prevent the need for filling up empty bins and to reduce the effect of outliers. Moreover it is strongly advised to extrapolate the damage separately for different operating conditions (at least operating, standstill and rotor stops), given the high differences in damage. To do this, a proper methodology is still needed to separate data points during standstill from data points representing a rotor stop.

Finally a farm-wide fatigue assessment is required, for multiple fatigue-critical locations in the substructure. Here, several approaches are possible. A first one is to develop the fleet leader concept as explained in Section 7.4 for multiple damage contributions. However, by splitting up in damage contributions, the interaction between different loads might be lost. Moreover, it will be extremely difficult to quantify the effect of slightly different structural parameters, such as resonance frequency, damping value or modeshapes, on the damage accumulation.

Another possibility is to install at least one accelerometer in each turbine of the farm. With this accelerometer of good quality and sufficiently high sample frequency and 1s SCADA data it would be possible to reconstruct the stress signal at any location in structure using the techniques explained in Chapter 5 and thus

perform a turbine-specific lifetime assessment based on the reconstructed stress signals. With installation of only one accelerometer, only the first mode can be taken into account during the reconstruction of the stress signal.

A third option can be to fully instrument at least one turbine within each cluster of comparable turbines and perform the fleet leader concept for each cluster.

For each of these possibilities, an analysis comparing the accumulated damage at different locations within the substructure is required. Moreover, the influence of wave loads should be analyzed and taken into account as well. To do this, accurate measurements regarding wave activity is needed.

Finally, a thorough analysis is needed about the effect of SCADA data of less quality on the final results for accumulated damage and lifetime.

### 8.2.2 Other future work

As already indicated throughout this thesis, some topics still need improvement, proper validation or simply more in-depth analyses:

- A proper validation of the rotor effective wind speed, as calculated in Section 3.5 is necessary. Once this alternative for wind speed as measured by the SCADA system is validated, the performance monitoring tools (and if applicable any other methodologies presented in this thesis) can be updated with this improved wind speed signal.
- An improvement of the thrust model, as explained in Section 4.4.2 is still possible, especially for non-generating conditions. Although it is possible, it is not advisable at this point. This is because the thrust load in non-generating conditions has only very little influence on the total load and can thus be ignored. Moreover, the impact of the data quality on the model results is substantial. Therefore, such an improvement has no influence if the SCADA data is of bad quality to begin with.
- A more useful improvement relates to the combination of the modeled thrust load and acceleration measurements. Essentially, for both presented methodologies improvements are possible. In case of the technique using modal decomposition and expansion, a workaround to deal with the transition between different frequency bands can improve the results significantly. In case of the Kalman filter, improvement is needed to obtain better results, especially in frequency domain. In both cases a thorough validation is needed of the combination with an improved 1s SCADA thrust model. This validation is not only required for the TP-TW interface but, even more important, also for different locations under water. Here, the effect of wave activity should be checked as well.
- In terms of lifetime assessment, a better understanding of the effect of safety factors and S-N curves on the damage and eventual lifetime can be beneficial to account for possible future changes (e.g. due to corrosion). The developed framework enables these analyses.

- A more in-depth analysis of how the damage is accumulated at different OWTs and its main drivers is still needed to enable extrapolation within the entire wind farm.
- The comparison between damages accumulated at different wind turbines should be repeated for several heights in the structure. The effect of the different loads can be different above or under sea level.

# List of contributions

## Conference paper

- 1. Effective virtual sensing scheme for fatigue assessment of monopile offshore wind turbines**  
Iliopoulos, A., Noppe, N., Weijtjens, W., Van Hemelrijck, D. & Devriendt, C., 2016, *Life-Cycle of Engineering Systems: Emphasis on Sustainable Civil Infrastructure: Proceedings of IALCCE 2016*. Taylor & Francis, 6 p. (PROCEEDINGS AND MONOGRAPHS IN ENGINEERING, WATER AND EARTH SCIENCES).
- 2. Fatigue life assessment of existing offshore wind turbines**  
Weijtjens, W., Noppe, N., Verbelen, T. & Devriendt, C., 2016, *2nd International Conference on Offshore Renewable Energy, Official Conference Proceedings*. ASRANet Ltd, p. 62-69
- 3. Fatigue life assessment of three offshore wind turbines**  
Weijtjens, W., Noppe, N., Iliopoulos, A., Verbelen, T. & Devriendt, C., 2016, *Life-Cycle of Engineering Systems: Emphasis on Sustainable Civil Infrastructure: Proceedings of IALCCE 2016*. Taylor & Francis, p. 742-747 (PROCEEDINGS AND MONOGRAPHS IN ENGINEERING, WATER AND EARTH SCIENCES).
- 4. Fatigue stress estimation of offshore wind turbine using a Kalman filter in combination with accelerometers**  
Noppe, N., Tatsis, K., Chatzi, E., Devriendt, C. & Weijtjens, W., 2018, *ISMA: International Conference on Noise and Vibration Engineering*.
- 5. Fleet-wise structural health monitoring of (offshore) wind turbine foundations**  
Weijtjens, W., Noppe, N., Verbelen, T. & Devriendt, C., 2016, *Proceedings of EWSHM 2016*. 8 ed. NDT.net, Vol. 21. 8 p. 292. (E-JOURNAL OF NON-DESTRUCTIVE TESTING).
- 6. Full load estimation of an offshore wind turbine based on SCADA and accelerometer data**  
Noppe, N., Iliopoulos, A., Weijtjens, W. & Devriendt, C., Oct 2016, *The*

*Science of Making Torque from Wind (TORQUE 2016)*. p. 1-11 11 p.  
(Journal of Physics: Conference Series; vol. 753).

7. **Monitoring the Consumed Fatigue Life on Three Offshore Wind Turbines**  
Weijtjens, W., Noppe, N., Iliopoulos, A. & Devriendt, C., 1 Sep 2015, *Proceedings of the 10th International Workshop on Structural Health Monitoring 2015: System Reliability for Verification and Implementation*. p. 2809-2816 8 p.
8. **Offshore wind turbine foundation monitoring, extrapolating fatigue measurements from fleetleaders to the entire wind farm**  
Weijtjens, W., Noppe, N., Verbelen, T., Iliopoulos, A. & Devriendt, C., Oct 2016, *The Science of Making Torque from Wind (TORQUE 2016)*. IOP Publishing, p. 1-10 (Journal of Physics: Conference Series; vol. 753).
9. **SCADA-based thrust load estimation of an offshore wind turbine**  
Noppe, N., Weijtjens, W., Devriendt, C. & Iliopoulos, A., 2016, *2nd International Conference on Offshore Renewable Energy, Official Conference Proceedings*. ASRANet Ltd, p. 375-381 7 p.
10. **Subsoil strain measurements on an operational wind turbine for design validation and fatigue assessment**  
Henkel, M., Noppe, N., Weijtjens, W. & Devriendt, C., 19 Jun 2018, *Journal of Physics: Conference Series*. IOP Publishing Ltd., Vol. 1037. 10 p.

## Article

1. **Data-driven multivariate power curve modeling of offshore wind turbines**  
Janssens, O., Noppe, N., Devriendt, C., Van de Walle, R. & Van Hoecke, S., Oct 2016, In : *Engineering Applications of Artificial Intelligence*. 55, p. 331-338 8 p.
2. **Modeling of quasi-static thrust load of wind turbines based on 1 s SCADA data**  
Noppe, N., Weijtjens, W. & Devriendt, C., 22 Mar 2018, In : *Wind Energy Science*. 3, p. 139-147

## Unpublished paper

1. **Subsoil stress reconstruction for fatigue monitoring of offshore wind turbine using accelerometers on the tower**  
Henkel, M., Noppe, N. (ed.), Weijtjens, W. (ed.) & Devriendt, C. (ed.), 4 May 2018. 12 p.
2. **Thrust load estimation based on high frequent SCADA data**  
Noppe, N., Weijtjens, W. & Devriendt, C., 2017.

## Poster

1. **Determining the remaining useful life of offshore wind farms**  
Noppe, N., Iliopoulos, A., Weijtjens, W. & Devriendt, C., 2016.
2. **Development of a SCADA-driven model to quantify wake effects within a windfarm**  
Noppe, N., Weijtjens, W. & Devriendt, C., 2015.
3. **Empirical analysis of wake effects in an operating wind farm**  
Noppe, N., Weijtjens, W. & Devriendt, C., 2015.
4. **Estimation of the quasi-static thrust force based on 1s SCADA data**  
Noppe, N., Weijtjens, W. & Devriendt, C., 2016.
5. **Fatigue Assessment of an operational wind farm**  
Weijtjens, W., Noppe, N. & Devriendt, C., 2015.
6. **Integrated turbine condition and health monitoring**  
Helsen, J., Weijtjens, W., Noppe, N., De Sitter, G. L. & Devriendt, C., 2015.
7. **Reliable empirical analysis of effects of turbulent air in an operating wind farm based on unreliable SCADA-data**  
Noppe, N., Weijtjens, W. & Devriendt, C., 2015.
8. **SCADA-based estimation of the quasi-static thrust load**  
Noppe, N., Weijtjens, W. & Devriendt, C., 4 Dec 2017.
9. **Subsoil strain measurements on an operational wind turbine for design validation and fatigue assessment**  
Henkel, M., Noppe, N. (ed.), Weijtjens, W. (ed.) & Devriendt, C. (ed.), 20 Jun 2018.
10. **Towards Farm-wide Fatigue Monitoring: Extrapolating fatigue measurements at a limited number of turbines to the entire farm**  
Weijtjens, W., Noppe, N. & Devriendt, C., 2015.



# Chapter 9

## Bibliography

- [1] WindEurope. Wind in power 2017 - annual combined onshore and offshore wind energy statistics, 2018. <https://windeurope.org/wp-content/uploads/files/about-wind/statistics/WindEurope-Annual-Statistics-2017.pdf>.
- [2] <http://corporate.engie-electrabel.be/>. Accessed on 2018-12-13.
- [3] <http://www.flanderstoday.eu/belgiums-largest-windfarm-breaks-ground-north-sea>. Accessed on 2018-12-13.
- [4] <https://renewablesnow.com/news/belgian-govt-reaches-deal-on-support-for-716-mw-of-offshore-wind-projects-588871/>. Accessed on 2019-02-18.
- [5] [https://www.vrt.be/vrtnws/en/2018/10/05/world\\_s-biggest-wind-turbines-for-new-belgian-wind-farm/](https://www.vrt.be/vrtnws/en/2018/10/05/world_s-biggest-wind-turbines-for-new-belgian-wind-farm/). Accessed on 2018-11-07.
- [6] Michael C Brower. Wind turbine performance: Issues and evidence, 2012. Presentation at EWEA Technology workshop "Analysis of operating wind farms: Learning the lessons from operational sites".
- [7] Francisco Javier Corcuera Infante and Niels Raben. WIND TURBINES - Part 26-1: Time based availability for wind turbines. International Electrotechnical Commission (IEC), 2010. Committee Draft (CD).
- [8] Robert Sherwin and Niels Raben. WIND TURBINES - Part 26-2: Production based availability for wind turbine generating systems. International Electrotechnical Commission (IEC), 2011. Committee Draft (CD).
- [9] <https://metallurgyandmaterials.wordpress.com/2015/12/25/liberty-ship-failures/>. Accessed on 2019-02-04.
- [10] <https://www.totalmateria.com/page.aspx?ID=CheckArticle&site=kts&NM=299>. Accessed on 2019-02-04.

- [11] <http://www.clocss.nl/portfolio/parkwind-nv-nobelwind-offshore-windfarm/>. Accessed on 2019-02-18.
- [12] Christophe Loraux and Eugen Brühwiler. The use of long term monitoring data for the extension of the service duration of existing wind turbine support structures. In *Journal of Physics: Conference Series*, volume 753, page 072023. IOP Publishing, 2016. doi: 10.1088/1742-6596/753/7/072023.
- [13] Alexandros Iliopoulos, Wout Weijtjens, Danny Van Hemelrijck, and Christof Devriendt. Fatigue assessment of offshore wind turbines on monopile foundations using multi-band modal expansion. *Wind Energy*, 20(8):1463–1479, 2017. doi: 10.1002/we.2104.
- [14] Marcel Schedat, Torsten Faber, and Abiman Sivanesan. Structural health monitoring concept to predict the remaining lifetime of the wind turbine structure. In *Domestic Use of Energy (DUE), 2016 International Conference on the*, pages 1–5. IEEE, 2016. doi: 10.1109/DUE.2016.7466723.
- [15] Lisa Ziegler, Ursula Smolka, Nicolai Cosack, and Michael Muskulus. Brief communication: Structural monitoring for lifetime extension of offshore wind monopiles: can strain measurements at one level tell us everything? *Wind Energy Science*, 2:469–476, 2017. doi: 10.5194/wes-2-469-2017.
- [16] <https://www.windpowerengineering.com/construction/installation/siemens-surges-but-vestas-wind-systems-remains-top-wind-turbine-installer-in-2014/>. Accessed on 2018-12-13.
- [17] Nymfa Noppe, Wout Weijtjens, and Christof Devriendt. Performance monitoring by tracking estimated power curves on a wind farm level. *WindFarms 2015, International colloquium on "Large Wind-Power Plants: Interaction, Control and Integration"*, 7 2015.
- [18] Nymfa Noppe, Wout Weijtjens, and Christof Devriendt. Modeling of quasi-static thrust load of wind turbines based on 1 s scada data. *Wind Energy Science*, 3(1):139–147, 2018. doi: 10.5194/wes-3-139-2018.
- [19] N Noppe, A Iliopoulos, W Weijtjens, and C Devriendt. Full load estimation of an offshore wind turbine based on scada and accelerometer data. In *Journal of Physics: Conference Series*, volume 753, page 072025. IOP Publishing, 2016. doi: 10.1088/1742-6596/753/7/072025.
- [20] N Noppe, K Tatsis, E Chatzi, C Devriendt, and W Weijtjens. Fatigue stress estimation of offshore wind turbine using a kalman filter in combination with accelerometers. In *Proceedings of ISMA2018 International Conference on Noise and Vibration Engineering*, 2018.
- [21] <https://www.olympus-ims.com/it/applications/rvi-wind-turbine/>. Accessed on 2018-12-13.
- [22] <https://www.sciencedirect.com/science/article/pii/S0360544216306958#fig5>. Accessed on 2018-12-13.

- [23] [https://www.researchgate.net/figure/Breakdown-of-Soil-Structure-Interaction-of-Offshore-Wind-Turbines-into-two-types-of\\_fig3\\_289585642](https://www.researchgate.net/figure/Breakdown-of-Soil-Structure-Interaction-of-Offshore-Wind-Turbines-into-two-types-of_fig3_289585642). Accessed on 2018-12-13.
- [24] <https://www.windpoweroffshore.com/article/1210054/foundations-types-depth-limits---alternative-solutions>. Accessed on 2018-12-13.
- [25] WindEurope. The european offshore wind industry. key trends and statistics 2016, 2017. <https://windeurope.org/wp-content/uploads/files/about-wind/statistics/WindEurope-Annual-Offshore-Statistics-2016.pdf>.
- [26] <https://www.sciencedaily.com/releases/2011/04/110426151040.htm>. Accessed on 2018-12-13.
- [27] Jan Van Der Tempel. Design of support structures for offshore wind turbines. PhD thesis, TU Delft, 2006.
- [28] Tony Burton, Nick Jenkins, David Sharpe, and Ervin Bossanyi. Wind energy handbook. John Wiley & Sons, 2011. doi: 10.1002/9781119992714.
- [29] Rico Baudisch. Structural health monitoring of offshore wind turbines. Master's thesis, Danmarks Tekniske Universitet, 2012.
- [30] Klaus Hasselmann, TP Barnett, E Bouws, H Carlson, DE Cartwright, K Enke, JA Ewing, H Gienapp, DE Hasselmann, P Kruseman, et al. Measurements of wind-wave growth and swell decay during the joint north sea wave project (jonswap). Ergänzungsheft 8-12, 1973.
- [31] Siegfried Heier. Grid integration of wind energy: onshore and offshore conversion systems. John Wiley & Sons, 2014. doi: 10.1002/9781118703274.
- [32] Emrah Kulunk. Aerodynamics of wind turbines. In Fundamental and Advanced Topics in Wind Power. InTech, 2011. doi: 10.5772/17854.
- [33] Christof Devriendt, Filipe Magalhães, Wout Weijtjens, Gert De Sitter, Álvaro Cunha, and Patrick Guillaume. Structural health monitoring of offshore wind turbines using automated operational modal analysis. Structural Health Monitoring, 13(6):644–659, 2014. doi: 10.1177/1475921714556568.
- [34] TS IEC. 61400-13: Measurement of mechanical loads. International Electrotechnical Commission, 2001.
- [35] Andrew Kusiak, Haiyang Zheng, and Zhe Song. On-line monitoring of power curves. Renewable Energy, 34(6):1487–1493, 2009. doi: 10.1016/j.renene.2008.10.022.
- [36] Joon-Young Park, Jae-Kyung Lee, Ki-Yong Oh, and Jun-Shin Lee. Development of a novel power curve monitoring method for wind turbines and its field tests. IEEE Transactions on Energy Conversion, 29(1):119–128, 2014. doi: 10.1109/TEC.2013.2294893.

- [37] Staffan Lindahl and K Harman. Analytical techniques for performance monitoring of modern wind turbines. In proc. of the European Wind Energy Association Conference, 2012.
- [38] Onder Uluycul and Girija Parthasarathy. Multi-turbine associative model for wind turbine performance monitoring. In Proceedings of the Annual Conference of the Prognostics and Health Management Society, 2012.
- [39] Edzel Lapira, Dustin Brisset, Hossein Davari Ardakani, David Siegel, and Jay Lee. Wind turbine performance assessment using multi-regime modeling approach. Renewable Energy, 45:86–95, 2012. doi: 10.1016/j.renene.2012.02.018.
- [40] David McLaughlin, Peter Clive, and Joanna McKenzie. Staying ahead of the wind power curve. Renewable Energy World Magazine, 2010.
- [41] Tuhfe Göçmen Bozkurt, Gregor Giebel, Niels Kjølstad Poulsen, and Mahmood Mirzaei. Wind speed estimation and parametrization of wake models for downregulated offshore wind farms within the scope of posspow project. In Journal of Physics: Conference Series, volume 524, page 012156. IOP Publishing, 2014. doi: 10.1088/1742-6596/524/1/012156.
- [42] J Beltrán, JJ Guerrero, JJ Melero, and A Llombart. Detection of nacelle anemometer faults in a wind farm minimizing the uncertainty. Wind Energy, 16(6):939–952, 2013. doi: 10.1002/we.1535.
- [43] International Electrotechnical Commission (IEC). WIND TURBINES - Part 12-1: Power performance of electricity producing wind turbines, 2005. First edition.
- [44] International Electrotechnical Commission (IEC). WIND TURBINES - Part 12-2: Power performance of electricity producing wind turbines based on nacelle anemometry, 2008. Committee Draft (CD).
- [45] T Burchhart. Power curve online monitoring - an alternative approach to execute the power curve warranty content, 2013. Presentation at Wind Farm Data Management and Analysis Conference Hamburg.
- [46] Meik Schlechtingen, Ilmar Ferreira Santos, and Sofiane Achiche. Using data-mining approaches for wind turbine power curve monitoring: a comparative study. IEEE Transactions on Sustainable Energy, 4(3):671–679, 2013. doi: 10.1109/TSTE.2013.2241797.
- [47] M Lydia, S Suresh Kumar, A Immanuel Selvakumar, and G Edwin Prem Kumar. A comprehensive review on wind turbine power curve modeling techniques. Renewable and Sustainable Energy Reviews, 30:452–460, 2014. doi: 10.1016/j.rser.2013.10.030.
- [48] Olivier Janssens, Nymfa Noppe, Christof Devriendt, Rik Van de Walle, and Sofie Van Hoecke. Data-driven multivariate power curve modeling of offshore

- wind turbines. Engineering Applications of Artificial Intelligence, 55:331–338, 2016. doi: 10.1016/j.engappai.2016.08.003.
- [49] Vestas. Vestas v90 3.0 mw. Catalogue, 2011.
- [50] Clara M St Martin, Julie K Lundquist, Andrew Clifton, Gregory S Poulos, and Scott J Schreck. Wind turbine power production and annual energy production depend on atmospheric stability and turbulence. Wind Energy Science (Online), 1:221–236, 2016. doi: 10.5194/wes-1-221-2016.
- [51] Amy Stidworthy, David Carruthers, and Julian Hunt. Cerc activities under the topfarm project: Wind turbine wake modelling using adms, 2011.
- [52] JG Slootweg, SWH De Haan, H Polinder, and WL Kling. General model for representing variable speed wind turbines in power system dynamics simulations. IEEE Transactions on power systems, 18(1):144–151, 2003. doi: 10.1109/TPWRS.2002.807113.
- [53] Sachin Khajuria and Jaspreet Kaur. Implementation of pitch control of wind turbine using simulink (matlab). International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 1(4):pp-196, 2012.
- [54] Thomas Ackermann. Wind power in power systems. John Wiley & Sons, 2005. doi: 10.1002/0470012684.
- [55] Andrés Bravo Cuesta, Francisco Javier Gomez-Gil, Juan Vicente Martín Fraile, Jesús Ausín Rodríguez, Justo Ruiz Calvo, and Jesús Peláez Vara. Feasibility of a simple small wind turbine with variable-speed regulation made of commercial components. Energies, 6(7):3373–3391, 2013. doi: 10.3390/en6073373.
- [56] K Raiambal and C Chellamuthu. Modeling and simulation of grid connected wind electric generating system. In TENCON’02. Proceedings. 2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering, volume 3, pages 1847–1852. IEEE, 2002.
- [57] Sanjiba Kumar Bisoyi, RK Jarial, and RA Gupta. Modeling and analysis of variable speed wind turbine equipped with pmsg. International Journal of Current Engineering and Technology, Special, (2):421–426, 2014. doi: 10.14741/ijcet/sp1.2.2014.78.
- [58] Ezzeldin S Abdin and Wilson Xu. Control design and dynamic performance analysis of a wind turbine-induction generator unit. In Power System Technology, 1998. Proceedings. POWERCON’98. 1998 International Conference on, volume 2, pages 1198–1202. IEEE, 1998. doi: 10.1109/ICPST.1998.729275.
- [59] A Murdoch, JR Winkelman, SH Javid, and RS Barton. Control design and performance analysis of a 6 mw wind turbine-generator. IEEE

- transactions on power apparatus and systems, (5):1340–1347, 1983. doi: 10.1109/TPAS.1983.318083.
- [60] Nicolai Cosack. Fatigue load monitoring with standard wind turbine signals. PhD thesis, Universitt Stuttgart, 2010.
- [61] Claudia Hofemann, Gerard van Bussel, and Herman Veldkamp. Forecasting of wind turbine loads based on SCADA data. 6th PhD Seminar Wind Energy, pages 17–22, 2010.
- [62] Luis Vera-Tudela and Martin Kühn. Analysing wind turbine fatigue load prediction: The impact of wind farm flow conditions. Renewable Energy, 107:352–360, 2017. doi: 10.1016/j.renene.2017.01.065.
- [63] Pierre-Elouan Réthoré. Thrust and wake of a wind turbine: Relationship and measurements. Master’s thesis, Technical University of Denmark, 2006.
- [64] Robert May, Graeme Dandy, and Holger Maier. Review of input variable selection methods for artificial neural networks. In Artificial neural networks-methodological advances and biomedical applications, pages 19–44. InTech, 2011. doi: 10.5772/16004.
- [65] Fabrice Rossi, Amaury Lendasse, Damien François, Vincent Wertz, and Michel Verleysen. Mutual information for the selection of relevant variables in spectrometric nonlinear modelling. Chemometrics and intelligent laboratory systems, 80(2):215–226, 2006. doi: 10.1016/j.chemolab.2005.06.010.
- [66] Brian V Bonnländer and Andreas S Weigend. Selecting input variables using mutual information and nonparametric density estimation. In Proceedings of the 1994 International Symposium on Artificial Neural Networks (ISANN94), pages 42–50, 1994.
- [67] Nabil Benoudjit, Damien François, M Meurens, and Michel Verleysen. Spectrophotometric variable selection by mutual information. Chemometrics and Intelligent Laboratory Systems, 74(2):243–251, 2004. doi: 10.1016/j.chemolab.2004.04.015.
- [68] David Freedman and Persi Diaconis. On the maximum deviation between the histogram and the underlying density. Probability Theory and Related Fields, 58(2):139–167, 1981. doi: 10.1007/BF00531558.
- [69] Josien PW Pluim, JB Antoine Maintz, and Max A Viergever. Mutual information matching in multiresolution contexts. Image and Vision Computing, 19(1-2):45–52, 2001. doi: 10.1016/S0262-8856(00)00054-8.
- [70] Gerlof Bouma. Normalized (pointwise) mutual information in collocation extraction. Proceedings of GSCL, pages 31–40, 2009.

- [71] Aaron F McDaid, Derek Greene, and Neil Hurley. Normalized mutual information to evaluate overlapping community finding algorithms. arXiv preprint arXiv:1110.2515, 2011.
- [72] Zeger F Knops, JB Antoine Maintz, Max A Viergever, and Josien PW Pluim. Normalized mutual information based registration using k-means clustering and shading correction. Medical image analysis, 10(3):432–439, 2006. doi: 10.1016/j.media.2005.03.009.
- [73] Matlab Users Guide. Neural network toolbox. The MathWorks, 2002.
- [74] J Jonkman and B Jonkman. Nwtc information portal (fast v8). last modified 27-July-2016.
- [75] Alexandros Iliopoulos, Rasoul Shirzadeh, Wout Weijtjens, Patrick Guillaume, Danny Van Hemelrijck, and Christof Devriendt. A modal decomposition and expansion approach for prediction of dynamic responses on a monopile offshore wind turbine using a limited number of vibration sensors. Mechanical Systems and Signal Processing, 68:84–104, 2016. doi: 10.1016/j.ymsp.2015.07.016.
- [76] Alexandros Iliopoulos, Wout Weijtjens, Danny Van Hemelrijck, and Christof Devriendt. Full-field strain prediction applied to an offshore wind turbine. In Model Validation and Uncertainty Quantification, Volume 3, pages 349–357. Springer, 2016. doi: 10.1007/978-3-319-29754-5\_34.
- [77] Maximilian Henkel, Nymfa Noppe, Wout Weijtjens, and Christof Devriendt. Subsoil stress reconstruction for fatigue monitoring of offshore wind turbine using accelerometers on the tower. In 9th European Workshop On Structural Health Monitoring (EWSHM 2018), 2018.
- [78] Wout Weijtjens, Gert De Sitter, Christof Devriendt, and Patrick Guillaume. Automated transmissibility based operational modal analysis for continuous monitoring in the presence of harmonics. In Proceedings of the 9th International Conference on Structural Dynamics (EURODYN), Porto, Portugal, pages 2231–2238, 2014.
- [79] Mahmoud El-Kafafy, Christof Devriendt, Wout Weijtjens, Gert De Sitter, and Patrick Guillaume. Evaluating different automated operational modal analysis techniques for the continuous monitoring of offshore wind turbines. In Dynamics of Civil Structures, Volume 4, pages 313–329. Springer, 2014. doi: 10.1007/978-3-319-04546-7\_35.
- [80] P. Avitabile and P. Pingle. Prediction of full field dynamic strain from limited sets of measured data. Shock and Vibration, 19(5):765 – 785, 2012. doi: 10.3233/SAV-2012-0686.
- [81] Javad Baqersad, Christopher Niezrecki, and Peter Avitabile. Full-field dynamic strain prediction on a wind turbine using displacements of optical targets measured by stereophotogrammetry. Mechanical Systems and Signal Processing, 6263:284 – 295, 2015. doi: 10.1016/j.ymsp.2015.03.021.

- [82] K. Maes, A. Iliopoulos, W. Weijtjens, C. Devriendt, and G. Lombaert. Dynamic strain estimation for fatigue assessment of an offshore monopile wind turbine using filtering and modal expansion algorithms. Mechanical Systems and Signal Processing, 76 77:592 – 611, 2016. doi: 10.1016/j.ymssp.2016.01.004.
- [83] A N Iliopoulos, W Weijtjens, D Van Hemelrijck, and C Devriendt. Prediction of dynamic strains on a monopile offshore wind turbine using virtual sensors. Journal of Physics: Conference Series, 628(1):012108, 2015. doi: 10.1088/1742-6596/628/1/012108.
- [84] A. Iliopoulos, W. Weijtjens, D. Van Hemelrijck, and C. Devriendt. Long-term prediction of dynamic responses on an offshore wind turbine using a virtual sensor approach. In Proceedings of the 10th International Workshop on Structural Health Monitoring 2015: System Reliability for Verification and Implementation, p. 2809-2816, Stanford, CA-USA, 2015. doi: 10.12783/SHM2015/346.
- [85] Konstantinos Tatsis and E Lourens. A comparison of two kalman-type filters for robust extrapolation of offshore wind turbine support structure response. In Life-Cycle of Engineering Systems: Emphasis on Sustainable Civil Infrastructure: Proceedings of the Fifth International Symposium on Life-Cycle Civil Engineering (IALCCE 2016), 16-19 October 2016, Delft, The Netherlands, page 209. CRC Press, 2016. doi: 10.1201/9781315375175-25.
- [86] Arthur Gelb. Applied optimal estimation. MIT press, 1974.
- [87] Det Norske Veritas. Fatigue design of offshore steel structures. DNV Recommended Practice DNV-RP-C203, 2011.
- [88] Clemens Hübler, Wout Weijtjens, Raimund Rolfes, and Christof Devriendt. Reliability analysis of fatigue damage extrapolations of wind turbines using offshore strain measurements. In Journal of Physics: Conference Series, volume 1037, page 032035. IOP Publishing, 2018. doi: 10.1088/1742-6596/1037/3/032035.
- [89] Turan Dirlik. Application of computers in fatigue analysis. PhD thesis, University of Warwick, 1985.
- [90] WAFO-group. WAFO - A Matlab Toolbox for Analysis of Random Waves and Loads - A Tutorial. Math. Stat., Center for Math. Sci., Lund Univ., Lund, Sweden, 2000.
- [91] Wout Weijtjens, Alexandros Iliopoulos, Jan Helsen, and Christof Devriendt. Monitoring the consumed fatigue life of wind turbines on monopile foundations. EWEA Offshore, 2015. doi: 10.13140/RG.2.1.1614.6162.
- [92] Det Norske Veritas. Fatigue design of offshore steel structures. DNV Recommended Practice DNV-RP-C203, 2010.

- [93] Det Norske Veritas. Fatigue design of offshore steel structures. DNV Recommended Practice DNV-RP-C203, 2016.
- [94] Det Norske Veritas DNV et al. Design of offshore wind turbine structures. Standard DNV-OSJ101, Det Norske Veritas AS (DNV), 2014.
- [95] Isaac Tavares and Feargal Brennan. The slic project: Advancing fatigue design guidance & standards for offshore wind, 2015. Poster presentation at EWEA Offshore 2015, Copenhagen.
- [96] <https://www.ocas.be/cases/ocas-uses-unique-technique-challenge-fatigue-standards-convinces-carbon-trust/>. Accessed on 2018-12-13.
- [97] HB Hendriks and BH Bulder. Fatigue equivalent load cycle method. ECN, The Netherlands, Tech. Rep. ECN-C-95-074, 1995.
- [98] Wout Weijtjens, Nymfa Noppe, Tim Verbelen, Alexandros Iliopoulos, and Christof Devriendt. Offshore wind turbine foundation monitoring, extrapolating fatigue measurements from fleet leaders to the entire wind farm. In Journal of Physics: Conference Series, volume 753, page 092018. IOP Publishing, 2016. doi: 10.1088/1742-6596/753/9/092018.
- [99] Anh Quang Mai. Updating failure probability of a welded joint considering monitoring and inspection for offshore wind turbine structures. PhD thesis, University of Liege, 2018.
- [100] Bradley Efron. Bootstrap methods: another look at the jackknife. In Breakthroughs in statistics, pages 569–593. Springer, 1992.
- [101] <https://yashueth.blog/2017/12/02/bootstrapping-a-resampling-method-in-statistics/>. Accessed on 2019-02-20.
- [102] Mads Damgaard, Mehdi Bayat, Lars Vabbersgaard Andersen, and Lars Bo Ibsen. Assessment of the dynamic behaviour of saturated soil subjected to cyclic loading from offshore monopile wind turbine foundations. Computers and Geotechnics, 61:116–126, 2014. doi: 10.1016/j.compgeo.2014.05.008.
- [103] John Dalsgaard Sørensen. Framework for risk-based planning of operation and maintenance for offshore wind turbines. Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology, 12(5):493–506, 2009. doi: 10.1002/we.344.
- [104] Wenbin Dong, Torgeir Moan, and Zhen Gao. Fatigue reliability analysis of the jacket support structure for offshore wind turbine considering the effect of corrosion and inspection. Reliability Engineering & System Safety, 106:11–27, 2012. doi: 10.1016/j.ress.2012.06.011.
- [105] Ulf T Tygesen, Michael S Jepsen, Jonas Vestermark, Niels Dollerup, and Anne Pedersen. The true digital twin concept for fatigue re-assessment of marine structures. In ASME 2018 37th International Conference on Ocean,

- Offshore and Arctic Engineering, pages V001T01A021–V001T01A021. American Society of Mechanical Engineers, 2018. doi: 10.1115/OMAE2018-77915.
- [106] Pablo G Morato, Jannie Sønderkær Nielsen, Anh Quang Mai, and Philippe Rigo. Pomdp based maintenance optimization of offshore wind substructures including monitoring. In 13th International Conference on Applications of Statistics and Probability in Civil Engineering, 2019.
- [107] H Braam, TS Obdam, and TW Verbruggen. Low cost load monitoring for offshore wind farms. In International Conference on Noise and Vibration Engineering 2012, ISMA 2012, including USD 2012: International Conference on Uncertainty in Structure Dynamics, 2012.
- [108] Marc Seidel. Wave induced fatigue loads: Insights from frequency domain calculations. Stahlbau, 83(8):535–541, 2014. doi: 10.1002/stab.201410184.

As the existing wind farms are growing older, fatigue assessments are gaining importance to the wind energy industry. In this thesis the possibility to optimize the use of the available SCADA data (1s and 10 minute statistics) is explored for fatigue and lifetime assessment and performance monitoring.

It is shown that under-performing turbines in a wind farm can be detected by calculating and comparing power curves. Moreover, changes in performance over time can be detected as well. However, results are very dependent on the quality of the wind speed measurement. To overcome this problem, the concept of Rotor Effective Wind Speed is implemented.

For the remainder of the thesis the transition towards fatigue is made. An important input for any fatigue assessment is the estimation of acting loads on a structure. In the particular case of a wind turbine, various loads can be detected. The first load that is estimated in this thesis is the thrust load. It is proposed to use a neural network, trained with a combination of SCADA data and measured thrust loads (obtained using strain sensors installed on the substructure). The method was validated on multiple datasets of different turbines at different wind farms, including the cross validation of a trained model on different wind turbines of the same type in a single offshore wind farm.

For a fatigue assessment, the stress history caused by all loads at fatigue critical locations in the structure is highly valuable. Two different techniques to reconstruct the stress history caused by all acting loads at any location in the structure, including the fatigue hotspots, are presented and compared. Both techniques use a combination of low frequent thrust load estimations and accelerometers installed at accessible locations in the tower. One of the techniques is based on modal decomposition and expansion, while the other makes use of a Kalman filter.

The next step in this thesis is to perform a lifetime assessment of multiple fatigue critical locations in the substructure of the offshore wind turbine, including the monopile and the transition piece. Therefore the state-of-the-industry techniques are implemented for cycle counting and damage calculation. Moreover safety factors and specific S-N curves are decided based on industry standards. To extend the calculated damage towards a lifetime, probability distributions regarding environmental conditions are needed. This thesis proposes a strategy for life time assessment and elaborates on several choices that can be made to perform such a life time assessment. The influence of these choices is illustrated based on a fictional offshore wind turbine.

Lastly, the correlation between measured damage and environmental conditions is shown by comparing the results of multiple instrumented turbines. The possibility of extrapolating measured damages at one turbine to other turbines within the same farm using environmental probability distributions is explored and commented upon. In the final chapter the main findings of this research are concluded and future perspectives are discussed.

ISBN

[www.CRAZYCOPY.BE](http://www.CRAZYCOPY.BE)