University of Strathclyde
Renewable Energy Marine Structures Centre for Doctoral Training
Department of Naval Architecture Ocean and Marine Engineering

Development of numerical and data models for the support of digital twins in offshore wind engineering

Mark Richmond

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Engineering
2021

Supervisors: Prof Athanasios Kolios
Prof Feargal Brennan

Declaration of authenticity and author's rights

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Mark Richmond

April 2021

Statement on previously published work

In the course of this thesis, a number of papers are published in scientific journals or still under review at the time of writing, as listed in Section 1.4. The author of this thesis was and is in all these publications - even if the paper is co-authored - the main responsible person. The author's contributions to the publications, thus, comprise conceiving the works, administering the studies, realizing the works, performing literature studies, developing the methodologies, performing the researches, developing and applying the approaches, working with and extending the software, curating the data, verifying and validating the methods and results, analysing and investigating the data and results, post-processing and visualizing the results and findings, writing the papers and preparing the original drafts, interacting with the journals' editors and reviewers, as well as reviewing and editing the papers for the final publications.

Mark Richmond

April 2021

Abstract

As offshore wind farms grow there is a continued demand for reduced costs. Maintenance costs and downtime can be reduced through greater information on the asset in relation to its operational loads and structural resistance to damage and so there is an increasing interest in digital twin technologies. Through digital twins, an operational asset can be replicated computationally, thus providing more information. Modelling these aspects requires a wide variety of models in different fields. To advance the feasibility of digital twin technology this thesis aims to develop the multi-disciplinary set of modelling domains which help form the basis of future digital twins. Throughout this work, results have been validated against operational data recorded from sensors on offshore structures. This has provided value and confidence to the results as it shows how well the mix of state-of-the art models compare to real world engineering systems.

This research presents a portfolio of five research areas which have been published in a mix of peer-reviewed journal articles and conference papers. These areas are: **1)** A computational fluid dynamics (CFD) model of an offshore wind farm conducted using a modified solver in the opensource software. This work implements actuator disk turbine models and uses Reynolds averaged Naiver Stokes approaches to represent the turbulence. This investigates the impact of modelling choices and demonstrates the impact of varied model parameters. The results are compared to operational site data and the modelling errors are quantified. There is good agreement between the models and site data. **2)** An expansion on traditional CFD approaches through incorporating machine learning (ML). These ML models are used to approximate the results of the CFD and thereby allow for further analysis which retains the fidelity of CFD at comparatively negligible computational cost. The results are compared to operational site data and the errors at each step are quantified for validation. **3)** A time-series forecasting of weather variables based on past measured data. A novel approach for forecasting time-series is developed and compared to two existing methods: Markov-Chains and Gradient Boosting. While this new method is more complex and requires more time to train, it has the desirable feature of incorporating seasonality at multiple timescales and thus providing a more representative time-series. **4)** An investigation of the change in modal parameters in an offshore wind jacket structure from damages or from changing operational conditions. In this work the detailed design model of the structure from Ramboll is used. This section relates the measurable modal parameters to the operational condition through a modelling approach. **5)** A study conducted using accelerometer data from an Offshore Substation located in a wind farm site. Operational data from 12 accelerometers is used to investigate the efficacy of several potential sensor layouts and therefore to quantify the consequence of placement decisions.

The results of these developments are an overall improvement in the modelling approaches necessary towards the realisation of digital twins as well as useful development in each of the component areas. Both areas related to wind loading as well as structural dynamics have been related to operational data. The validation of this link between the measured and the modelled domains facilitates operators and those in maintenance in gaining more information and greater insights into the conditions of their assets.

Contents

List of Figures

List of Tables

Preface/Acknowledgements

I would like to thank my supervisor, Professor Athanasios Kolios, for the support, resources and opportunities he has provided to me over the years.

I would like to thank my parents, Alan and Lucy Richmond, for their love and support. I would like to thank my brother.

I would like to thank Jenny Cheng for her love and support.

I would like to thank everyone on the Renewable Energy Marine Structures centre for doctoral training.

I would like to thank my colleagues and friends at Ramboll in Hamburg and Tufts University in Boston.

I would like to thank the countless individuals I met along the way, whether they became my friend, colleague or only a brief acquaintance, for the knowledge and thoughts shared.

And in the end, I would humbly like to give at least a little thanks to myself. Well done, me!

Chapter 1

1. Introduction

1.1. Background

Offshore wind is a rapidly growing industry and has established itself as a leading source of renewable energy [1]. Not only is this the case in Europe where the industry first emerged [2,3], but also now in other markets including the United States [4,5] as well as Asian markets such as Taiwan [6,7], Japan [8] and others [9]. Current projections show global offshore wind capacity to reach 130-140 GW by 2030 and as much as 1500 GW by 2050 [10]. The offshore wind industry is rapidly growing, but to achieve the upper limits of these projections it is vital to further reduce the cost so that it becomes more competitive in both current and developing markets.

The levelized cost of energy (LCOE) arises from a combination of capital expense to build and install the energy assets (CAPEX) as well as operational expense in running and maintaining the assets (OPEX). The balance of these costs varies between projects and depends as much on conscious choices from the designers as it does on circumstances of the project. A prevalent idea in offshore wind is that of 'conservativism' where designs are made stronger, or inspection is conducted more often than it must be. Conservativism is important for ensuring the asset achieves its intended design life, but it increases both CAPEX and OPEX. Conservativism, and hence LCOE, can be reduced by improving confidence in assessments of the wind farm. Improved confidence can come from either confidence in the design and its future loads or from the current state of the wind farm and this can come from improved models and data.

There has typically been a strong motivation towards measuring structures as well as their environments. These data are vital for either operational monitoring or for building and validating models and analytical approaches. Many organisations, research groups or classification societies, now give recommendations for what sensors to use and how to use them [11,12]. This data can be an incredibly valuable resource and as more data is collected and stored, more ways need to be researched for how to get even more value from it.

The concept of a digital twin is to replicate the operational asset as much as possible within a digital environment. This can be accomplished through a combination of data acquisition from the operational asset as well as a modelling framework. Models are constructed of the asset and its operational environment, which are then updated through sensor data, such as accelerometers, inclinometers or strain gauges [13]. Provided that the models behave sufficiently closely to the real asset, they can then give a valuable insight into its current and future state.

1.2. Problem statement

The offshore wind-energy research and industry space is filled with models designed for a plethora of applications. These models, which relate to the structural health, include models of wind, wave and other loading, models of the structure itself and aero-elastic models. Using these models can lead to computational requirement challenges. Generally, as model fidelity is increased, the computational requirements rapidly increase. Additionally, offshore wind farm analysis requires the consideration of multiple interdependent physical phenomena which are challenging to model – particularly in modelling wind farm level aerodynamic wake interactions. Hopes for developing digital twin approaches can be greatly improved by reducing computational demand for models used. Additionally, behaviour of the structural response due to changing parameters is important for damage detection and assessing asset life. This requires a detailed model approach of the support structure itself.

Many analytical offshore wind energy purposes should not be considered as deterministic but rather as subject to significant uncertainty [14]. Any analysis of the loads should consider that uncertainty, which is present in both the loading on the structure as well as the structural resistance. From a reliability perspective these two items are key in that the probability of failure of the structure can be considered as the overlap between the distributions of loading and resistance where the loading is the greater of the two. Therefore, research which works towards reducing the cost of offshore wind energy through lowering the uncertainty in the analysis should consider the modelling of both the load variation as well as structural variation, and should do so in such a way as to allow for uncertainty.

1.3. Aim and objectives

This research aims to develop modelling approaches related to offshore wind loading as well as to bridge the gap between structural sensor data and structural models by modelling structural response and studying accelerometer capabilities.

The contribution to knowledge is investigating the effectiveness of computational fluid dynamics (CFD) models for modelling offshore wind farms in comparison to operational data. This is developed further through a framework for extending CFD models using machine learning approaches. A new weather time-series model is introduced which outperforms existing models in some ways. The contribution from the structural part is a sensitivity study of an offshore jacket structure which uncovers new considerations and challenges to vibration-based damage detection. Additionally, a study of accelerometer placement, using measured data, gives new insights into the impact of sensor layout on real measurements.

The thesis aims to deliver the following objectives:

1. Conduct a comprehensive literature review of offshore wind farm modelling approaches in order to map current state of the art as well as provide a reference. This includes wind farm wake models, hydrodynamic models, geotechnical models and structural modelling which relate to this thesis's aim.

2. Model a reference offshore wind farm using CFD to evaluate the wake flow within the wind farm and the impact of changing parameters.
3. Expand the CFD model using machine learning so that further results can be gained rapidly without significant loss of accuracy. Compare several models and evaluate performance compared to measured data.
4. Investigate the effectiveness of weather time-series forecasting methods to determine the most suitable model for wind farm load estimation.
5. Conduct sensitivity study of a reference offshore wind jacket structure to determine the change in dynamics as a result of changed parameters.
6. Investigate sensor layout impact on the ability to extract modal parameters from accelerometers.

1.4. Structure

Due to the importance of modelling both wind farm load as well as structural resistance, this thesis approaches both challenges and does so structured as two parts. The first part comprises the wind load variation, modelling wake flow through a wind farm using computational dynamics, which is then approximated using machine learning to allow for rapid results while retaining fidelity. Additionally, weather time-series modelling is covered. The second part comprises the structural modelling aspect of the thesis. This includes sensitivity of the structural dynamics in response to changes in parameters and damage cases. There are also studies on accelerometer sensor layout using measured data. Finally, the project presents a coupled dynamic modelling analysis of the structure for crack growth estimation.

This thesis is composed of portfolio of work which has been published in peer-reviewed journals. Each of the two parts in this thesis are composed of the following publications:

Literature review – offshore wind farm modelling approaches

1. M. Richmond, D. Cevasco, A. Kolios “A review of modelling methods for arrays of offshore wind turbines” – **Submitted**

- **Part 1:**
  - **Offshore wind turbine wind wake modelling**

2. M. Richmond, A. Kolios, V. S. Pillai, T. Nishino, and L. Wang, “Development of a stochastic computational fluid dynamics approach for offshore wind farms” **J. Phys. Conf. Ser., 2018**. [15]
3. M. Richmond, A. Antoniadis, L. Wang, A. Kolios, S. Al-Sanad, J. Parol, “Evaluation of an Offshore Wind Farm Computational Fluid Dynamics Model Against Operational Site Data” Ocean Engineering 2019;193. https://doi.org/10.1016/j.oceaneng.2019.106579 [16]

  - **Machine learning expansion of wind turbine wakes**

4. M. Richmond, R. Pandit, A. Sobey, A. Kolios, “Stochastic assessment of aerodynamics within offshore wind farms based on machine-learning” – **Renewable Energy, 2020;161:650-61. https://doi.org/10.1016/j.renene.2020.07.083** [17]
    - **Time series weather forecasting**
5. M. Richmond, R. Pandit, S. Koukoura, A. Kolios “Effect of weather forecast modelling uncertainty to the availability assessment of offshore wind farms” – **In revision pending final decision**

- **Part 2**
    - **Offshore jacket foundation structural variation**
6. M. Richmond, U. Smolka, A. Kolios, “Feasibility for Damage Identification in Offshore Wind Jacket Structures through monitoring of global structural dynamics” – **energies, 2020;13:5791. https://doi.org/10.3390/en13215791** [18]
    - **Sensor layout study**
7. M. Richmond, S. Siedler, M. Häckell, U. Smolka, A. Kolios, “19052 Impact of Accelerometer Placement on Modal Extraction of Offshore Wind Structures” – **Ocean Mar. Arct. Eng. 2020, 2020.** [19]

Chapter 2

2. Literature review

Paper:

1. M. Richmond, D. Cevasco, A. Kolios “A review of modelling methods for arrays of offshore wind turbines” – In progress

2.1. Introduction

The European Union is well on its way to achieving its objective of 20% total energy consumption from renewable energy by the year 2020 [20] and a revised target of 32% by 2030 [21]. Two important aspects to be improved upon in achieving these targets from wind energy are reducing the cost of operational expenditure (OPEX), which can range from 25% – 30% of overall energy generation cost [22]; and increasing availability, which is presently around 97% [23,24]. To achieve higher performance, the research effort to numerically model offshore wind farms has been performed for two primary reasons: 1) increasing energy revenue at the design stage via optimizing the layout of wind turbine generators (WTG) throughout the site [25] and 2) determining load and resistance variation for individual WTG support structures throughout the wind farm [26][27]. These two points can also be used together for enhancing the control strategies and finding the best compromise between the farm power output and the loading on the downwind structures [28].

The grouping of turbines into an array leads to a reduction in the power production and increased dynamic loading of the downstream turbines, due to the introduction of velocity deficits and higher turbulence levels in the wake aerodynamics [29]. Additionally, uncertainties on the estimation of soil characteristics at the farm level, and the hydrodynamic loads on individual structures [30], can have a considerable impact on structural fatigue life estimation. Therefore, when aiming to forecast numerically the performance and loads on the turbines in a farm the key areas to be modelled are:

1) the physical and stochastic aspects of the environmental conditions;
2) the structure-environment interaction, in terms of the aerodynamic loads – for estimation of both performance and excitation on the structures, hydrodynamic loads – for wave and current excitations, and the soil-structure interactions.

The target audience of this chapter is wind farm developers or wind energy consultants that are looking for which model to use for a future offshore wind farm. The purpose of such a broad approach, rather than a more targeted chapter on one of the topics covered, is to show the interrelation and the commonality between the modelling domains. This chapter aims to review

the modelling approaches which can be applied to offshore wind farms at an array level as well as give a list of numerical tools which employ these models. By ‘array level’ it is meant that the variation between multiple turbines is modelled, for example the wake within a wind farm or the geotechnical variation within the farm site. Models related to individual wind turbines or foundations are only covered in how they form part of ‘array level’ models. This chapter can help guide wind farm modelling activities at either a design stage or an operation and maintenance stage. It is found that many of the areas covered are well developed and have a broad range of available numerical tools to implement them. Wind farm aerodynamics can be modelled in a wide variety of ways which have been tested and compared to each other. However, as computational capabilities improve, the balance between resources and requirements can lead to the use of more high-fidelity models. There are numerous well-established methods for wind, wave and current models, mostly based on statistical approaches. However, machine learning is an active area of research yielding new approaches which can be used here. Wave loading can be modelled analytically through linear or non-linear wave theory as well as computationally, however there is very little research on the difference in wave loads within wind farms for different support structures.

Statistical weather forecasting methods are presented because they can be used as direct parameter inputs to the wind farm variation models. More advanced mesoscale resource models are not presented because this would comprise a very large section without adding to the focus of the chapter. Weather forecasting models which can be used to predict wind and wave conditions are presented in section 2.2 beginning with statistical methods followed by machine learning approaches for time series forecasting. The wind and wave conditions vary in time and between the sites. These variations can be monitored by buoys and met masts by taking measurements at predetermined intervals. Once these data are gathered, they can be converted into models such as probability density functions which can be used as inputs for further analysis. Some commonly used functions in offshore wind are Weibull [31], Rayleigh [32] and lognormal [33], among others [34]. Spatial variation models are presented at the end of this section, discussing how surrogate modelling approaches can be used for predicting soil parameters.

Section 2.3 presents the loads and soil structure interaction in the offshore environment. This section starts presenting the commonly used wave theories, for estimating the wave patterns of a specific site and on a time-variant basis. Then the main approaches to calculate the hydrodynamics loads on the wind turbine structure are briefly introduced. While models exist for determining the wave field and load variation in arrays of wave energy converters and tidal turbines, there appears to be a gap in research on wave load within the wind farm; the wave load approaches tend to assume a single structure with no array level effects. The soil dynamics are also presented in this section where the modelling of soil-structure interaction is discussed, as well the phenomenon of scour.

The aerodynamic principles and models are presented in section 2.4, starting with models to determine the impact of the turbine on the flow field. These models include generalised actuator methods, blade element momentum theory, and direct computational fluid dynamics (CFD) modelling. This is followed by methods for evaluating the variation of the wind field within a wind farm through modelling the downstream flow field.

A review of performance requirements of an individual turbine which may support the Digital Twin Model is given in section 2.5, followed by a review of damage detection approaches, given in section 2.6

A comprehensive list of numerical tools which use the modelling methods discussed in this chapter are listed in section 2.7. This list includes applicable machine learning models as well as a list of wind flow modelling numerical tools.

The application and importance of these methods are discussed in section 2.8 and a conclusion is given in section 2.9.

2.2. Data-driven forecasting of wind and wave parameters

When modelling the loading on turbines in a farm, the variations in the wind, wave and soil parameters may be considered temporally and spatially. For instance, the wind loads can vary temporally due to seasonal and diurnal phenomena, and in the short term for the presence of turbulence and gusts, while spatial variations include the shear effect of the boundary layer and layout-dependent effects of the wake generated from the other turbines in the farm. In this section, an overview of traditional statistical methods as well as alternative methods for the time and spatial forecast of the environmental conditions is given for the inflow wind, wave field and soil characteristics.

2.2.1. Wind-wave-current data

The characteristic environmental conditions for wind, wave, and current are typically defined in terms of:

- Mean wind speed, wind direction and turbulence intensity, as well as a spectrum or turbulence modelling parameters, for the wind flow.
- Significant wave height (H_s) – as the average of the highest one-third among the waves, peak period time (T_p), direction, depth averaged current velocity and other spectrum parameters, for describing the sea state.
- Velocity, direction and depth profile, for representing the ocean currents.
- Joint correlations between separate, but interdependent, variables.

The wind and wave spectra provide an expression of their dynamic properties and fluctuations within a short-term stationary state. In other words, a spectrum describes the statistical properties of wind and waves over a time period (typically, of about an hour for wind fluctuation and of 3 hours for the sea state), covering all the energy distributions along a frequency range. [35].

2.2.1.1. Traditional statistical methods

Some common probability distribution functions used for modelling wind velocity at the hub height include: Weibull [31], Rayleigh [32], lognormal [33] and Poisson [34] distributions, however several others have been suggested [32]. The Weibull distribution, shown in equation (1) where

k is the shape parameter, c is the scale parameter and v is the velocity [31], is the most prominent distribution for wind energy use because it fits most wind regimes quite accurately [36]. This distribution is recommended by International standards, IEC 61400-12 [37] as well as the American Bureau of Shipping [38]. It is however not well suited to regimes with significant null wind periods [32]. The 2 parameter Weibull probability density function is as follows:

$$f(v) = \frac{k}{c} * \left(\frac{v}{c}\right)^{k-1} * e^{-(\frac{v}{c})^k} \quad (1)$$

The Rayleigh distribution, where the *k* parameter is set equal to 2 [39], is a special case of the two-parameter Weibull and it gives a good general estimate where there is a lack of data available [32]. Despite the widespread use of the Weibull distribution, sometimes it is outperformed, for example, by the Lognormal and Gamma functions [40], so it is worth investigating different distributions for each site. The time series of turbulent wind can then be generated either statistically, via the use of stochastic inflow turbulence codes such as the one of TurbSim [41] or by physics-based models, such as the Mann approach [42]. The implementation of the first became of common use in some of the most well-known aero-elastic solvers for wind turbine aerodynamics (i.e. in the AeroDyn of FAST [43] and the aerodynamics module of BLADED [44]), while the second is developed as a spectral tensor turbulence model and it is used in the generation of the inflow turbulence for wake modelling tools of the Dynamic Wake Meandering (DWM) model and EllipSys3D (later discussed in Section 2.7.2) [29].

In the design phase, the wave parameters of H_s and T_z are commonly modelled with the following distributions: Lognormal [45,46], Weibull [47,48], Gumbel [49,50] among others. For the description of the short-term waves, if the spectrum data cannot be obtained from measurements, numerical approximations are used as common practise, such as the Pierson-Moskowitz and JONSWAP spectra for wind sea, or the Torsethaugen and Gaussian swell spectra suitable for swell sea and implemented in [51]. The spectrum needs to be carefully chosen to correctly represent the local sea conditions and the frequency distribution depending on the wave spreading. The realization of irregular waves in the time domain is then obtained using inverse fast Fourier transforms based on the spectrum.

The long-term environmental conditions can be formulated as a site-specific joint probability distribution of the given parameters (as waves are generally a wind driven phenomenon) [33], or can be modelled independently. It is vital to account for the spatial heterogeneity within a wind farm as use of only a single wind speed distribution for an entire wind farm site can lead to significant error. Yang et. al. found that spatial heterogeneity can account for 24.3% error between predicted and measured power for total installed capacity [52].

As regards to the ocean currents, driven by wind, tides and geostrophic phenomena, they do not only contribute to the static loads on the substructure of the offshore wind turbine, but can also be responsible to modification in the amplitude and frequency of the wave loads (see Section 2.3.1) and/or affect the scour formation (see Section 2.3.2.2). According to standards, the velocity and directional characteristics of the sea current should be evaluated together and the extreme sea surface current velocities assessed via site-specific assessment and/or metocean database. Nonetheless, no requirements are given for determination of the current velocity with depth, thus

the standard profiles can be applied for evaluating the effect of sub- and near-surface and breaking wave induced currents [53].

Extreme values can be estimated from these statistical models and these extreme values can be an important part of offshore wind structural design. Examples of these extreme values are 50-year or 100-year wave, which is the largest wave likely to be encountered once within a given return period. A recommendation is given in IEC-61400-3 [54]. This commonly accepted design consideration has been used in many offshore wind design studies [55–57].

2.2.1.2. Regression methods for forecasting and simulation

In addition to traditional statistical methods, regression models exist which can either forecast or simulate the wind and wave time-series. Regression models fundamentally make predictions by modelling the relation between the independent and the dependant variables. Although the approaches can have a statistical basis, the application is done algorithmically. In the case of time series, the prediction is a combination of trend, seasonality, and noise. These can be modelled as dependant on time and so can be reproduced into the future.

A clear distinction needs to be made between time-series forecasting, where there is some expectation of relatively high accuracy, and time-series simulation, where the result behaves in a representative manner. In highly stochastic variables such as wind and waves, it can be difficult to accurately predict individual observations, particularly over longer time spans. In this regard, the time spans can be divided into very short term of a few seconds to 30 minutes, to long-term of between 1 day and 1 week [58]. For other applications such as micro-siting and lifelong O&M models, longer time spans of simulation are required, sometimes years ahead [59]. Very long-term models need to only replicate the mean and level of deviation, but can have high errors for individual observations, and can generally be extended for a much longer period. This distinction is important because it has a high influence on both the choice of model and the training approach. As an example, in forecasting one might take only the newest data, expecting it to have a greater correlation to near-future values. For many approaches, the training data needs to be made stationary, meaning that the trend and seasonality need to be removed, which can then be reintroduced into the prediction.

In terms of time series forecasting methods, recurrent neural networks are a popular method, particularly Long-Short Term Memory LSTM approaches [60,61], these methods are computationally heavy and can make predictions within a future timeframe of a few hours. ARIMA (Auto-Regressive Integrated Moving Average) [62] models are highly effective for short term forecasting but are limited in long-term predictions due to the tendency for the autoregressive part to converge towards the mean. Ensemble methods can be used as well, which combine a series of weak learners to form a strong learner. Methods such as random forest [63,64], which works by dividing the training data through hyperplanes and then giving predictions based on which section a new observation occurs. Another example is gradient boosting [65–67] which iteratively adds new learners to correct the residual from previous weak learners. These methods can be very fast to train relative to the others, making them efficient with large data sets.

For time series simulation, Markov-chains are a common method and have been employed successfully for wind modelling [68–71]. Markov-chains work well for long term simulation and can be constructed with a separate model trained on each month to capture annual seasonality. A hybrid statistical regression method for wind and wave time-series simulation is proposed by Richmond et. al [59]. In this approach a series of regressive models are used to predict time-dependant statistical parameters including mean, standard deviation and covariance matrixes. These statistical parameters are then used to randomly sample from a distribution, giving a likely simulated value. For simulating both wind and wave values together, a multi-variate distribution is used.

2.2.2. Soil characteristics

The way in which the soil supports the structure is not only vital for all life-cycle phases of a wind farm but also subject to considerable uncertainty. This uncertainty can result from many factors, including limited measurement data or from time dependant phenomena. Some of the modelling approaches aimed at addressing this uncertainty are now presented. First, stochastic spatial models are presented, which address improving site knowledge based on limited data.

Typically, there are in-situ soil tests conducted at the locations of each turbine before the turbines are installed [72]. However, this might not always be the case, particularly in the site layout planning stage before the layout is decided, so stochastic spatial variation models are required to model how the soil is likely to vary, based on known data points. A widely used stochastic spatial soil surrogate model is ‘Kriging’, which was first proposed by Krige [73] in 1951 for use in mining. Kriging is a geo-statistics method belonging to the ‘maximum likelihood’ family where there is assumed to be a spatial correlation between data points [74–76]. In the Kriging model, the output function is derived from a polynomial global approximation, vector of regression coefficient and a random process function [77]. A number of algorithms can be used to implement Kriging, including: turning bands method, fast Fourier transforms and discrete Fourier transforms [79, 80]. Kriging not only produces a response surface, which crosses through the data points, but also provides some estimate for the uncertainty [75].

However, throughout the wind farm there is some lack of certainty in the soil constitution and stratification. Vahdatirad et al. [72] investigated random field theory to create a stochastic model of the undrained strength and Mohr-Coulomb model for the constitutive behaviour then applied this to a limit state model to assess uncertainty in structural resistance. Carswell et al. [80] discussed how probability distribution functions such as Gaussian, log-normal and Beta distributions can be used to model uncertainty in soil variability. Andersen et al. [81] proposed a method for combining the random models of the soil strength with deterministic structural models and a Monte-Carlo approach to determine natural frequencies of a monopile foundation. The geotechnical parameters play a key role in modelling the dynamics of offshore wind turbine support structures and their uncertainty can account for a significant deviation between a modelled structure and measured dynamics [82].

2.3. Loads in the offshore environment

The hydrodynamic loads result from the waves and current. The importance of wave loads tends to increase as water depth increases, and the modelling of non-linearities could be essential for certain types of structures and/or site conditions. The modelling of soil-structure interaction is another driving criterion in both the design and operating phases as this can be a high-sensitivity factor in the structural response.

2.3.1. Hydrodynamic loads

The selection of the most suitable numerical model for the representation of the wave pattern and the calculation of the hydrodynamic loads on an offshore structure depends on its geometries (shape and size as well as fixed or floating body) and the site characteristics (i.e. sea level and wave conditions) [83]. The hydrodynamic pressure loads on slender structures – which can be assumed to have a little effect on the wave field – can be estimated by solving the undisturbed potential flow. If the body has a larger volume, the diffraction effect should be considered, and thus the solution of the velocity potential should account for the presence of the body. Finally, considering the motion of the structure, radiation loads should be considered by including the contribution of the added mass and linear damping.

2.3.1.1. Wave loads modelling

One of the most employed and well-known methods for the calculation of the loads on the structure is the Morison's equation [84,85]. This semi-empirical approach assumes hydrodynamic transparency of the structure, and it is thus suitable when the body diameter (D), for a vertical cylinder body, is small with respect to the wavelength (L). Morison's equation considers the contribution of two forces (integrated along strips of the length of the structure): the inertial force and nonlinear drag force. For non-slender bodies (for $D < 0.2\,L$), the MacCamy and Fuchs equation [86], which reformulates the inertia term of Morison's equation to consider diffraction effects, can be used in the case of simple shapes [83].

However, neither of these equations are suitable for accurately capturing higher-order effects, both neglecting the loads from nonlinear potentials. The consideration of nonlinear effects due to diffraction and nonlinear waves might be important in certain cases. For instance, Peeringa [37] and Henderson et al. [55, 56], among other authors [87], argued that linear wave theory is inadequate for wind turbines that are in shallow water and non-linear wave models should be used. Nonlinear wave models range from higher order Stokes, cnoidal wave, solitary wave and stream function wave theories [88]. Additionally, the high frequency loads related to steep, nonlinear waves might be better captured by the employment of the Rainey equation [89,90] and the FNV (Faltinsen, Newman, and Vinje [91,92]) theory. The comparison of these methods for the calculation of the loads on a monopile structure for steep nonlinear wave conditions was proposed by Ristić's thesis [93]. Chen [94] investigated the pros and cons of the application of Rainey equation and FNV method for the loads on a spar-type floating structure. Regarding the loads from breaking waves, the modelling of this complicated physics is an active area of several researchers [95–97], and has been summarised for offshore wind structure in the review of Chella et al [98].

Table 1. Hydrodynamic model overview, adapted from [83]

			Morrison equation		Diffraction theory	CFD
			Time domain	Freq. domain	Freq. domain	Time domain
Force	Transverse	inertia	✓	✓	✓	✓
		drag	✓	✓[3]	✗[1]	✓
	Lateral (drag)		✓	✓[3]	✗[1]	✓
	Pressure		✗[1]	✗[1]	✓	✓
Geometry	Diffraction		✗[2]	✗[2]	✓	✓
	Surface effect	1D	✓	✓	✗	✓
		3D	✗	✗	✗	✓
	Big structure		✗	✗	✓	✓
Wave model	Linear - Stochastic		✓	✓	✓	✗[4]
	Nonlinear - Wave extrapol.		✓	✗	✗[4]	✓
Applicability poor * okay ** good ***	Ease of use		***	**	**	*
	Calculation speed		**	***	***	*
	Commercial availability		***	***	***	**

[1] can be modelled by adding an extra term

[2] can be modelled by MacCamy-Fuchs correction

[3] linearised

[4] high computation power

In Table 1 an overview of the main approaches for the derivation of wave loads is given. This is adapted from the review work of [83]. More recently, and extensively, the OC5 (offshore code comparison collaboration continuation with correlation) research project investigated and compared the load prediction from several open access and commercial tools. For tank testing and load estimation on a scaled model of the OC3 monopile structure [99], most of the OC5 participants employed the Morison equation in its relative-motion form. A few others opted for a potential flow theory solved via a panel method approach and only one applied a CFD approach [100]. One of the major drawbacks of these latter approaches is the impossibility to and/or the high computation effort required in accounting for the deflection of the structure. A CFD approach is thus generally pursued if the interest is the detailed and local modelling of loads on secondary structures (e.g. turbine access ladder) [101], and for the study of floating structures, where capturing the fully coupled fluid-structure interaction is of greater concern than modelling the structure flexibility [102].

There are then a number of other approaches for modelling the hydrodynamic loads and behaviour of floating offshore wind turbines [103]. For instance, Zhang [104] developed a numerical model to implement a 'Fluid Impulse Theory', calculating, in the time domain, the nonlinear hydrodynamic wave loads and the resulting nonlinear responses of a floating wind turbine for various wave conditions. In [105] Leble et al, instead, calculated loads using smoothed particle hydrodynamics (SPH) method where particles represent a fluid volume and its motion and is calculated by solving Lagrangian Navier-Stokes equations. They then applied the SPH in a coupled floating offshore wind turbine model, noting the advantage of such a model in that the free-surface requires no special treatment [106].

Finally, to summarise the result from the OC5, during the first phase of the project (tank testing of pile structure), it was concluded that:

- All the models used for analysis were good at capturing the first order component of the force and sufficient for deep water application. In contrast, in shallower water the higher order components of the force are important to be captured to accurately describe the overall force on the structure and the potential increase in fatigue damage.
- In shallow water, the approaches which were able to model the seabed slope and nonlinear wave transformation were more accurate in representing the distribution of forces and acceleration. These are the ones deriving the wave kinematics from either CFD or the fully nonlinear potential-flow solvers, as opposed to standard wave theories, including those using direct measurement of the wave elevation. Nonetheless, it is also acknowledged that such approaches are computationally impractical for the load analysis of a wind turbine and should rather be employed for an in-depth analysis of particular conditions.

Although the highly nonlinear nature of the waves for shallow water conditions can be well approximated, additional effort and research is necessary to integrate the contribution from wave breaking loads.

2.3.1.2. Impact of the turbine array on the wave field

It also must be mentioned that the presence of marine and offshore structures can have an impact on the surrounding wave field. This aspect has mainly been of concern to researchers investigating wave energy converters. Although not designed for this purpose, large offshore jacket structures and turbine piles can affect the propagation of the wave fields with diffraction, reflection and blocking effects. These phenomena can influence the hydrodynamic loads for the array of structures in the farm, as well can have an impact on sediment transport [107] (discussed further in section 2.3.2.2), on the shoreline [108] and on the marine ecosystem [109].

Spectral models, such as the SWAN wave model developed at TU Delft [110], were employed to investigate the transformation of irregular waves inside an offshore wind farm. These models aim to predict the far field effect and were used by [109] to derive the wave field characteristics downwind of an offshore wind farm. Nonetheless, they cannot cope with the hydrodynamically accurate, and computational expensive solution, of the structure-fluid interactions, and thus have been mainly employed to quantify the potential environmental impact of offshore wind farm [111,112].

As regards to more advanced models for the derivation of the structure-fluid interaction, Sismani et al. [113] applied the NEMOH, an open-source boundary-element-method-based numerical model [114], for estimating the impact of an array of offshore monopiles, with a kilometre spacing. They noticed that an increased hydrodynamic interaction and loads on the structures are obtained for increasing wave period and specific incident wave direction, especially affecting the turbines positioned in the middle rows, due to the scattering effect in the offshore wind farm. From literature it was then found that several authors are currently applying and developing new numerical models for the solution of the water wave interaction with multiple-cylinder-array structures. In [115], Jia Fu et al. offer a good review of these models, suggesting on the other hand a quicker meshless approach as an alternative to the more popular finite element [116] and boundary element [117] numerical methods. Nonetheless, no further application and validation studies for offshore wind farms were found in the literature.

2.3.2. Bottom-fixed foundation loadings

This section gives an overview of soil-structure interaction (SSI) modelling approaches which relate directly to the modelling of offshore wind farms. Although this is used for modelling individual structures, it is useful to discuss this topic as these models are affected by spatial and temporal phenomena discussed in other sections. Therefore, when conducting spatial soil parameter modelling, it can be useful to bear in mind the individual structural models which might use those values. A good review of SSI analysis related to offshore wind support structures can be found at [118].

2.3.2.1. Models for soil interaction

Generally, the strength of the soil-structure interaction can be modelled to a high fidelity with finite element analysis (FEA) models [119]. However, quicker ‘engineering’ models have been developed and employed in offshore wind [120]. Among these are the widely used ‘*p-y’* models,

based on Winkler foundation theory [80]. A quasi-static *p-y* degradation model can be used which accounts for the change in parameters resulting from cyclic loads [121].

As the monopiles have much larger diameter to length ratios than the *p-y* method was intended for, the PISA project (Pile Soil Analysis) was conducted at Oxford University - applying largescale, onshore, cyclic, lateral load tests. These tests resulted in a greatly improved *p-y* design method which would allow for building in deeper waters and lower conservativism [122–124].

Constitutive models are needed to model the stress-strain behaviour of soils and to translate the soil parameters to values used in SSI modelling methods [125]. There are a wide range of constitutive models and no one model fits all soils. As an example Whyte et al. developed a constitutive model for suction bucket designs which can be integrated into an FEA model without excessive computational demand [126].

2.3.2.2. Sediment transport and scour development modelling

Scour is a time varying phenomenon which can have a significant impact on the structural health of a wind turbine. The scour and its development need to be monitored and/or predicted throughout the life of the wind farm. There are two types of scour: local scour, which is a conical pit around the monopile created by the erosion of the soil; and global scour which is the erosion of the entire top layer of soil [127]. Scour directly impacts the structural dynamics and has been the subject of much research within offshore wind structures [128–130].

Whitehouse et al. [131],[132] developed a conceptual scour model which illustrates how scour depth varies with different sediment types as well as whether scour tends to increase or decrease in extreme wave conditions for that sediment type. Further work from Harris et al. [133] developed the 'Scour Time Evolution Predictor' (STEP) model, an engineering model where scour development is modelled as an exponential function. Due to the fact that scour is significant from a design perspective, there are a number of numerical models which predict the evolution of scour throughout a wind farm [134]. Numerical scour models tend to have three parts: a hydrodynamic part; a wave part; and a sediment transport part. Some examples of this are work from Van den Eynde et al. [135,136] where they used the hydrodynamic models MU-BCZ, based on shallow water equation and TELEMAC-2D. They used the MU-SEDIM, which uses the Nikuradse bottom roughness model, and MU-STM for soil transport models, modelling advection and diffusion, and they used the MU-WAVE model. The models of MU-BCZ, MU-WAVE and MU-STM were also used by Fettweis et al. [137]. Developing these models together can be useful as the hydrodynamic and sediment transport models can share the same grid [135]. Vera et al [138] used a suite of numerical models to model the scour in the Hinder Bank region, they used OPTOS-BCZ, HYPAS, MU-SEDIM and MU-STM.

2.4. Aerodynamic models for wind farm simulations

Often to model the wind farm, the wind turbine and wind farm are modelled separately, feeding values such as velocity, turbulence, energy, and so on, between models. In this section, an overview on the approaches and level of fidelity for the modelling of wind turbines and farm aerodynamics is given. First, an overview of the physics involved and the level of fidelity of the

models is discussed. Leading on from this, the turbulence models and governing equations are presented. This gives the link between the physics being modelled and the numerical approaches used in the models. Wind turbine models are presented in section 4.3; it is useful to understand what these are as well as their influence on the larger flow stream. The wide range of array models are presented in section 4.4, from very fundamental models such as roughness models, up to high-fidelity computational fluid dynamics.

CFD modelling of wind turbines and wind-farm levels are discussed, and finally some existing CFD numerical tools used for offshore wind farm applications are presented. A list of specific numerical tools is presented in section 2.5.

2.4.1. Physics and level of fidelity of the models

A fundamental understanding of how wind turbines affect wind flow can come from one-dimensional actuator disk theory. In this theory, the turbine disk acts as an infinitely thin momentum sink, causing a discontinuity in flow energy and momentum, as well as a jump in pressure, as flow passes the disk [139,140]. Given the assumptions that there is a well-defined stream tube and that flow is entirely axial, homogeneous and steady, then the maximum power efficiency, the Betz limit, can be determined as 16/27 [141]. Additionally to the drop in momentum across the turbine, there is also vorticity introduced to the flow, which occurs from blade circulation and is shed at the blade tips [142]. According to the Kutta-Joukowski law, the lift force on the blades can be attributed to this vorticity [139]. The shed vortices then form a cylinder in the wake of the turbine separating the ambient flow from the wake in the near region. The flow is considered to be incompressible, given a Mach value below 0.3, and turbulent, given a high Reynolds number [143].

The wake behind the turbine can be classified as being near or far wake, with an intermediate phase between the two [139,144]. The near wake is directly affected by the geometry of the rotor and the tower and is typically considered to extend 1 – 2 rotor diameter distance downstream [140], but depending on definition, can be between 2 – 5 diameters [142]. It has been shown that the length is a function of thrust, tip speed ratio and turbulence intensity [145]. The far wake region is defined as when the profile of the wake has fully developed into a roughly Gaussian profile, in this region the geometry of the rotor has only an indirect impact [140]. The far wake is assumed to be roughly axisymmetric and self-similar [146].

Wind farm flow is of high Reynolds number, dominated by the kinetic, rather than viscous forces [16], and therefore has significant turbulence. Within the far wake, turbulence is the result of atmospheric turbulence, mechanical turbulence from blades and tower, and wake turbulence from the shed vortexes [140]. Turbulence is important in the far wake because it determines the level of mixing and hence the wake recovery [140]. Due to the complexity and importance of turbulence, it has been the subject of much research on approaches to model it. Models used resolve turbulence in a varying degree, from Reynolds averaging to direct numerical simulation [143].

From this discussion it can be seen that the real situation of wind turbine flow is very complex. In wind turbine aerodynamics there are multiple phenomena occurring together and

separate regions of airflow which behave quite differently from each other. In deciding which models to use for a wind farm, it is useful to consider what the aim is, what data is available and what physical behaviours should be included. For example, the self-similarity of the far wake is the basis for kinematic models [140,142]. The near wake is directly impacted by turbine geometry, and so data on that geometry is required, such as needing air foil data for blade-element-momentum models [147]. If one is interested in modelling the wind farm as a whole, with an interest more in the far wake, then a simpler actuator-disk model can be used [16]. Models typically employed for wind farm aerodynamics are presented in the following sections.

2.4.2. Turbulence models and governing equations

The turbulence characteristics in a wind farm are part of estimating the power production and the loads on the wind turbines. As the wind farms increase in size, which is happening particularly offshore, a good description of the wakes becomes crucial for an accurate performance prediction of the wind farm. The atmospheric turbulence enhances the wake recovery while together with the wake induced turbulence, they are the major source of fatigue loading on the wind turbines [1]. Both the atmospheric and the wake added turbulence are parametrised in various wake models in terms of Turbulence Intensity (TI) which is defined by $Ti = \sigma_u/U$ where σ_u is the standard deviation of the velocity and U is mean wind speed [148]. Modelling small turbulence length scales is prohibitively computationally expensive and so turbulence models are used [139]. The two most widely used types of models in wind energy are Reynolds Averaged Navier-Stokes (RANS) models and Large Eddy Simulation (LES) models. Detached Eddy simulation (DES) models are sometimes used, however they are in their infancy [143]. RANS and LES models are now discussed in sections 2.4.2.1 and 2.4.2.2 respectively.

2.4.2.1. Reynolds Averaged Navier-Stokes (RANS)

Reynolds Averaged Navier-Stokes (RANS) models are among the most widely used models for turbulence in wind energy. RANS models present a time-averaged solution to the Navier-Stokes equations to average the turbulent fluctuations [149][139].

The most commonly used RANS-based models, which contain the Reynolds stress terms, are *k-ε* and *k-ω* models (RSM) [143]. This method was used by Porté-Agel et al to validate LES, compare turbine models and evaluate the effects of layout [150] and by Kalvig et al [151] in comparing actuator disk, actuator line and full rotor simulation against wind tunnel measurements.

k-ε, proposed by Launder and Spalding [152], is one of the most widely used turbulence models [153]. In this two-equation model the turbulent kinetic energy, *k*, and the dissipation rate, *ε*, are calculated in order to determine the turbulent viscosity [154]. The method employs the Boussinesq approximation which assumes that the turbulence is isotropic, which is inaccurate for large-scale, atmospheric turbulence [139][153]. However, there have been a number of modifications to the standard *k-ε* model which have shown to improve results, the most common example is the Renormalized Group Method (RNG) *k-ε* model [143]. Based on a number of studies [155],[156], the RNG model has shown significant improvements over the standard k-ε, particularly for near wake. The standard *k-ε* model is given by the following equations:

$$\frac{\partial k}{\partial t} + \bar{u}_j \frac{\partial k}{\partial x_j} = \frac{\partial}{\partial x_j}\left[(\nu + \frac{\nu_t}{\sigma_\epsilon})\frac{\partial k}{\partial x_j}\right] + \nu_t \frac{\partial \bar{u}_i}{\partial x_j}\left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i}\right) - \varepsilon \quad (2)$$

$$\frac{\partial \varepsilon}{\partial t} + \bar{u}_j \frac{\partial \varepsilon}{\partial x_j} = \frac{\partial}{\partial x_j}\left[\left(\nu + \frac{\nu_t}{\sigma_\epsilon}\right)\frac{\partial \epsilon}{\partial x_i}\right] + C_1 \frac{\epsilon}{k} \nu_t \frac{\partial \bar{u}_i}{\partial x_j}\left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i}\right) - C_2 \frac{\varepsilon^2}{k} \quad (3)$$

Where $x_{i,j}$ are Cartesian space coordinates, $u_{i,j,k}$ are mean stream-wise velocity in the denoted Cartesian coordinate, ν is the kinematic viscosity and C_μ, C_1, C_2, σ_k and σ_ϵ are constants.

The k-ω method is another widely used eddy viscosity method which uses the specific dissipation rate, ω. This method uses a wall function close to the surface of objects [157]. Like with *k-ε*, the *k-ω* model has had several modifications proposed. It has been found that the *k-ω* Shear Stress Transport (SST) model is the most successful for 2D and 3D modelling [149] and combines aspect of both *k-ω* and *k-ε* [143]. The standard *k-ω* model is given by the following equations:

$$\frac{\partial k}{\partial t} + \bar{u}_j \frac{\partial k}{\partial x_j} = P_k - \beta^* \omega k + \frac{\partial}{\partial x_j}\left[(\nu + \sigma_k \nu_t)\frac{\partial k}{\partial x_j}\right] \quad (4)$$

$$\frac{\partial \omega}{\partial t} + \bar{u}_j \frac{\partial \omega}{\partial x_i} = \alpha S^2 - \beta \omega^2 + \frac{\partial}{\partial x_j}\left[(\nu + \sigma_\omega \nu_t)\frac{\partial \omega}{\partial x_j}\right] + 2(1 - F_1)\sigma_{\omega 2}\frac{1}{\omega}\frac{\partial k}{\partial x_i}\frac{\partial \omega}{\partial x_i} \quad (5)$$

Where x_j is Cartesian space coordinate, ν is kinematic viscosity, F_1, is a blending function, P_k is a production limiter, and $\sigma_{k,\omega,\omega 2}$, β, and α are closure coefficients.

The Reynolds Stress Model (RSM) is a second order closure model for resolving turbulence which is rarely used in wind energy [139]. This method works by introducing 7 new equations to the transport equations for all components of the stress tensor [143][154].

2.4.2.2. Large Eddy Simulation (LES)

Wind flow within an offshore wind farm contains turbulence at multiple scales, there is smaller scale turbulence resulting from interaction with the structures, but also large-scale turbulence associated with atmospheric fluctuations. Large Eddy Simulations (LES) differ from RANS models in that they directly resolve the unsteady flow properties above a filtered scale, turbulence below this level is parameterized through a sub-grid-scale (SGS) model [158],[150]. Turbulence can be divided into three main regions based on the scale: large scale, which contains most of the energy and is unaffected by viscosity; intermediate scale, which is unaffected by viscosity; and small scale which is viscosity dependant. LES, directly calculates large scale turbulence and averages the small scale turbulence [159]. The differentiation is based on specific requirements and computational capabilities and is applied through a filter. The small-scale turbulence is modelled with a sub-grid scale model, usually based on eddy viscosity [143]. LES was used by Churchfield et al [160] to model the 48 wind turbine array, Lillgrund in OpenFOAM. LES has been used with

an actuator line representation of turbines in several studies [158,161,162], and actuator disk representation by Yang & Sotiropoulos [163]. The method is reviewed by Mehta et al [159].

In modelling of offshore wind farms using an LES approach, typically a set of filtered, incompressible Naiver-Stokes equations are solved [161]:

$$\frac{\partial \widetilde{u_i}}{\partial x_i} = 0 \tag{6}$$

$$\frac{\partial \widetilde{u_i}}{\partial t} + \widetilde{u_j}\left(\frac{\partial \widetilde{u_i}}{\partial x_j} - \frac{\partial \widetilde{u_j}}{\partial x_i}\right) = -\frac{\partial \widetilde{p^*}}{\partial x_i} - \frac{\partial \tau_{ij}}{\partial x_j} + \nu\frac{\partial^2 \widetilde{u_i}}{\partial {x_i}^2} - \frac{f_i}{\rho} + \delta_{i1}F_P \tag{7}$$

Where the filtered velocity, $\widetilde{u_i}$, is in the i direction, $\widetilde{p^*}$ is the modified kinematic pressure, is the τ_{ij}, SGS stress tensor. ν is the kinematic pressure, f_i is the body forcing from the wind turbines, ρ is the air density, and F_P is the imposed pressure gradient.

2.4.2.3. The Dynamic Wake Meandering (DWM)

Dynamic Wake Meandering (DWM) is a wake model where the wake is modelled as a passive tracer driven by large-scale turbulence structures [164]. The velocity profile behind the turbine takes the assumption of being Gaussian, axisymmetric and self-similar, with lateral and vertical displacement dependant only on large-scale turbulence. The model is calculated as three parts [29]: 1) Velocity deficit, solved through parabolised Naiver-stokes equations, 2) Wake meandering of the horizontal, v, and vertical, w, velocities and 3) rotor added turbulence, which is assumed to be rotationally symmetric.

This model has been validated by a number of authors in the offshore wind environment who have found it to have very good agreement with measurement [165,166]. One limitation that was found was that model error could occur as a result of not predicting the non-linear interaction due to controller strategy [165]. The model has been compared to higher fidelity models and found to have better agreement with each of these models at different parts of the downstream wake, which is due to the different physical reason for wake recovery [166]. However, the model was assessed from a certification perspective by DNV-GL who identify some basic requirements for the model's use in site suitability studies [167].

The model is often coupled with fluid-structure-interaction models. For example, HAWC2 couples the model with the HAWC environment a single wind turbine. The simulated data were shown in good agreement with load measurements [165].

2.4.3. Wind turbine models

The following sections present methods used to represent wind turbines. These are models which have been used to influence the wake effects within a wind farm and so can be used to model the entire wind farm. While direct numerical simulation of the wind farm is possible, and this is presented in section 2.4.3.3, this approach becomes almost impossible with current hardware for

wind farm simulations due to computational cost. Therefore, simplifications are typically used for this approach. These simplified models include actuator disk, actuator line and surface methods. The more advanced methods are used when a more high-fidelity view of wake effects is required, such as dynamic modelling of turbulent fluctuations.

2.4.3.1. Actuator disk, line and surface models

The most common methods for modelling the turbine is the actuator disk (AD)[141]. AD is a low to medium fidelity approach where the turbine is modelled as a permeable disk whose loading influences the flow field [168]. The load can be evaluated in a number of ways, either through looking up aerodynamic values for the blade or through solving Euler or Naiver-Stokes equations over a discretised field [149]. The advantage of AD is its relative simplicity and reduced computational cost due to not having to resolve the boundary layer around the blade [150]. However, due to the inviscid assumption of the model it is inaccurate in predicting downstream wake because it under-predicts the turbulence required for wake dissipation [168,169].

A distinction of different turbine representations within CFD models is given in Table 2. This table gives the model names as well as the distinction between them and the data required to parameterise the models.

Table 2. Turbine numerical model representations for computational fluid dynamics.

Model	**Fully resolved**	**Actuator disk** [141]	**Actuator line** [170]	**Actuator surface** [171]
Rotor model description	Blades are modelled as a boundary within the CFD domain.	Swept area of the blades are modelled. Uses fundamental momentum theory.	Turbine blades are simplified to a line with the aerodynamic forces along it.	Turbine blade modelled as a surface with the aerodynamic forces imposed on the surface.
Data required	Surface geometry	Coefficient of thrust, C_T, for the whole turbine.	Coefficients of lift and drag, c_l and c_d, for blade sections.	Coefficients of pressure and force, C_p and C_f, for the blade.

The actuator disk model was used by Annoni et al. [168] for the purpose of controlling two turbines, by Castellania et al. [172] in investigating a method which reduces computational effort and Cabezón et al. [155] in evaluating a semi-parabolic wind farm model. Additionally, the method was compared to the methods of actuator line and full CFD. While Stergiannis et al. [169] found the method underpredicted wake effects downstream and was more sensitive to turbulence models than CFD, Kalvig et al. [151] agreed that CFD was the most accurate method but concluded also that AD performs well on far wake and gave magnitude values that were in-line

with measurements. It was also noted that AD has a much lower computational time than full CFD.

To conclude, a hierarchy of actuator models has emerged, from 'simple' actuator disks to more advanced actuator line and surface models. Gaining accuracy is accompanied by higher computational cost and the need for more detailed airfoil data: from an overall coefficient of thrust (uniform actuator disk) to section wise coefficients of lift and drag (non-uniform actuator disk, actuator line) to coefficients of pressure and force (actuator surface). The unsteady nature of actuator line and surface methods makes them most suitable for LES simulations and that the steady nature of actuator disk methods limits their application to mostly RANS simulations. However, the computational cost limits the use of the actuator line technique most often to single-wake investigations, and most LES simulations of wind farms are being performed with actuator disks.

2.4.3.2. Blade Element Momentum

A very common way of applying the AD model [150] is through Blade Element Momentum model (BEM), which is used to parameterise the turbine loading in most aeroelastic models [149]. Although this can be considered as a form of AD, its implementation is more complicated and requires more knowledge about the turbine blades profile geometries and performances.

To apply this model, which was first proposed by Glauert [173], the blade is first discretised into a series of elements along the blade's length. The local induction factor, combined with knowledge of the local velocity and pitch angle, is used to determine the local angle of attack. The lift and drag coefficients can be evaluated from tabulated values and combined with the local velocity and geometric values to determine lift and drag which can be transposed into a thrust value parallel to the free stream. Momentum theory is then used to determine a new estimate for radial and axial induction factors. This process is iterated until the induction factor values converge within an accepted residual [139].

The model assumes the flow is inviscid and so does not need to resolve the boundary layer near the blade [150]. There are a number of limitations to the model, it ignores vortex shedding [149], and the model fails after an axial induction factor larger than 0.4, however there are various corrections proposed for these [139],[149].

BEM was used by Wang [149] to develop an aeroelastic model for large wind turbine blades and was used by Eliassen [174] to investigate time-dependant aerodynamic loads on wind turbines in comparison with numerical methods. Engelen [175] used the method for frequency-domain calculation for offshore wind turbines in the TURBU offshore code. Mahmuddin used the method with the aim of optimizing wind turbine blade designs and validated the code developed against Qblade [176].

2.4.3.3. Fully resolved rotating blade

Computational fluid dynamics (CFD) approaches based on solving discretised Navier-Stokes (NS) or Euler equations are rarely used in commercial software for parameterising the individual

turbines, however research shows that the method is much more accurate than other turbine parameterisation approaches [151,169]. In this approach, the area around the turbine is discretised using a finite volume or finite element method and the fluid flow equations are solved across this domain.

Despite the rotation of the rotor, the flow can be solved in a steady-state manner, simplifying the model, if the reference frame is rotating, fixed to the rotor. Stergiannis [169] et al. used a multiple reference frame approach, using a rotating inner, ADM, frame and a stationary outer, CFD, frame to solve for two generators with the geometry of their blades. Dehrt et al. [177] presented their approach for using a rotating reference frame and its mathematical formulation.

Réthoré et al. [178] compared a direct CFD approach with an actuator disk approach and found that the increased computational expense did not justify the increased accuracy while Stergiannis et al [169] found in their comparison between CFD and ADM that ADM under-predicted wake effects and recommended using more advanced CFD methods for wind farm modelling. The results from Kalvig et al. [151] agrees with both of these findings, concluding that a fully resolved model provided 'superior' velocity results but at a computational expense. A direct CFD model can be used to support a comparatively simplified model, for example Lanzafame et. al. used a fully resolved CFD model to support a BEM model and compared results with experiments, finding the approach to give very accurate results [179].

Fully resolved, CFD solvers for wind turbines usually consider the air incompressible as the velocity generally has a low Mach number [139]. These codes solve the Navier-Stokes equations of fluid motion, incorporating the conservation of mass, momentum and energy, across a discretised domain [143]. These methods can be used to handle turbulence in a variety of ways including Reynolds Averaged Navier-Stokes and Large Eddy Viscosity, which was discussed further in section 2.4.2.

CFD was used in recent literature to evaluate the turbine in a paper by Stergiannis et al [169] to compare this method to an actuator disk. Similar comparisons between CFD and other methods were also made in a number of other papers [151,178]. This method was also used by L. Wang et all [180] for use in fluid structure interaction.

2.4.4. Methods of representing fluid flow in the domain

There are a wide range of wind farm wake models which can be used to evaluate site layouts depending on power production and/or loads on the structures. As anticipated in section 2.4.1, these models range in fidelity and in complexity, dealing with simple and two-dimensional problems, to complex three-dimensional models. Following an introduction of some of the most basic models including roughness, kinematicand progressively more advanced models is presented.

2.4.4.1. Roughness models

Wake models discussed by Crespo et al. [142] and Sanderse [139] consider the wind shear profile along the path of the wind, where the ground roughness influences the profile and the turbines

themselves act as added roughness elements. The wind shear is modelled as a logarithmic profile up to the wind turbine and then a second logarithmic profile is modelled above the turbine.

2.4.4.2. Kinematic models

Kinematic models are more complex than roughness models as they consider each turbine individually. An overview of kinematic modelling is presented by L. Simeng [181] and B. Sanderse [139]. Kinematic models consider the wake as a velocity profile and wake growth as a function of distance from the turbine, with wake interaction modelled through superposition. These models generally need to include a turbulence model (see Section 2.4.2) to accurately model wake closure.

As described in the literature [181,182], the Jensen wake model assumes linear wake expansion downstream and that when wakes overlap, their total kinetic energy deficit is equal to the sum of energy deficits. The Larsen model is based on turbulent boundary layer and a self-similar velocity profile. The Fuga model is a linearized CFD model for turbine wakes [29].

Gaumond et al. [182] compare the Jensen, Larsen and Fuga [183] models in evaluating the Horns Rev and Lillgrund offshore wind farms. In a study by Borràs et al. [184], validating three wind farms against data from Horns Rev, It was found that all three models, Jensen, Larsen and Ishihara, provide accurate predictions with the error in predicted power production around 1.5% over a large variation in incident wind direction of 30°. It was found that the models, particularly Larsen, performed well, especially when using a slightly higher wake decay constant, but that CFD models provide more accurate results. Johnson et al. [185] developed a model based on a modification of the analytical expressions used in the PARK wake model for wind farm control. Their stated motivation for this was feedback from industry suggesting that even CFD models didn't always accurately predict performance so a quicker model was desired.

2.4.4.3. Large wind farm models

An emerging field of research is large scale wind farms, which is becoming significant as wind farms grow. It has been found that conventional models underestimate the levels of drag in large scale wind farms and so overestimate power production, consequentially authors have revisited the Betz limit regarding turbines in these farms [186]. To this end, Ma et. al. proposed a two-scale coupled momentum balance method which better predicts large wind farm aerodynamics. A principle component of this extension is the incorporation of a support-structure drag term. This was validated using Wall-Modelled Large-Eddy Simulation in a periodic boundary condition [187,188].

2.4.4.4. Field models

Field models have been developed with the aim of efficiently estimating the effect of wake decay while also providing physical insight into the behaviour. These linearized models treat the wind turbine wake as only a small perturbance in the flow field and so in solving the Navier-Stokes equations of fluid motion for wind turbine wakes, several assumptions can be made to simplify the process. These assumptions are a steady state model where the wake is axisymmetric with no radial velocity, pressure gradients outside the wake are considered negligible. An early model

proposed by Sforza et. al. used only a simplified momentum equation assuming constant viscosity and compared it to experimentation finding good agreement given the simplicity [189]. Ainslie then expanded on this by including an eddy viscosity equation [190] – this model described eddy viscosity differently for near and far wake, using a filter function in near wake. Presenting a matrix inversion which was still quick to solve while providing more accurate estimates. Ainslie stated that such a model could be conducted 'speedily' on a personal computer.

Field models have seen some use in research, for example Kim et. al. [191] used an Ainslie model in comparison to wind tunnel experiments, scaled to have tip-speed ratios comparable to modern megawatt turbines finding that the Ainslie model was among the most accurate of similar models tested. Field models are also used for wake meandering, for example Larsen et. al. [192] used a formulation heavily influenced by Ainslie to implement wake meandering, finding satisfactory agreement in both statistics of the flow field and the response of the coupled structural model.

2.4.4.5. Computational Fluid Dynamics models

CFD models of wind farms require the fluid flow to be discretised into a series of two or three-dimensional sections across which Navier-Stokes equations can be solved [143]. Generally, the wind turbines themselves are parameterised using a representative model [139]. These models, as discussed in section 2.4.2, can either be another CFD model or, more commonly, can be an actuator disk or blade element momentum model. Additionally, to model wake dissipation, these models also require a turbulence model, turbulence models are discussed in section 2.4.2.

Castellani et al. [172] investigated a CFD model using an actuator disk model and RANS. Porté-Agel et al. [150] investigated a CFD model where the turbines were parameterised using both an actuator disk and actuator line model with turbulence modelled with large eddy simulation (LES). The open source solver OpenFOAM was used by Cabezón et al. [155] where the domain is split into sub-domains of actuator disk models and Reynolds Stress model; these sub-domains allowed for parallel processing.

There is considerable variation between models based on how the turbines are parameterised, discussed in section 2.4.2, as well as a significant difference based on the turbulence model employed. There are many existing models developed by different projects and organisations – these models are presented in section 2.7.

2.5. Foundation performance requirements for Digital Twins

The purpose of a digital twin in offshore wind is to replicate the condition of an asset to better facilitate risk-based operation and maintenance strategies and potentially exploit the potential for lifetime extension. With this goal in mind, it can be helpful to examine the performance requirements for individual offshore wind turbines and contextualise the work in this thesis. The key limit states can be taken as those given in DNVGL-ST-0126 [193] and summarized below:

- ULS – Ultimate Limit State, which results in the loss of the structure of component due to reaching the maximum load capacity.

- FLS – Fatigue Limit State, which corresponds to the cumulative damage of the structure over time resulting from dynamic loading.
- ALS – Accidental Limit State, which results from rare, accidental loading situations.
- SLS – Serviceability limit state, which is the tolerance criteria of the structure or component in normal use.

All these limit states must be considered within the digital twin modelling framework. Both ULS and FLS can be important for lifetime extension and ALS can be an important consideration to assess the structural capacity after an accident, however it is SLS which can present a more immediate concern. For fixed bottom foundations the primary limit can be quantified as the allowable rotation at the mudline as well as tilt and deflection [194]. For floating foundations, the limits can be much more complex, as is outlined by Bhattacharya et al. [195]. Serviceability limit states are often approached by the structure gradually throughout its life in iterative load cycles, so it is important that they are defined clearly and that an appropriate method is used to assess whether the limit is reached [196].

With limit states established, the reliability of the structure is then a combination of both the structural resistance as well as the loads applied. In conducting this analysis, methods such as FORM (first order reliability method) or SORM (second order reliability method) have been employed, depending on whether the response is linear or of higher order [197].

2.6. Damage detection approaches

The idea of modal analysis is that a damage or some other change to the structure causes a measurable change in the modal parameters. An example of where it can be particularly important to detect damage is after a seismic event, which can cause significant change in the structure's dynamic response. This is a problem facing offshore wind as it expands to locations with seismic activity [198,199]. While a single earthquake can cause notable damage, there is often repeated, sequential incidents which are even more impactful [200]. The action of earthquakes on offshore structures is also very complex, involving seismic waves and non-linear structural responses [201], so using measured data can provide valuable insights into these situations [202]. Damage detection of structures can either be in an experimental modal analysis, where a load is applied, or operational modal analysis (OMA) where operational loads provide excitation to the structure [203]. In OMA, the excitation energy comes from operational loads. The operational excitation should be uniform in terms of frequency domain, and therefore any non-uniformity in the measured response is due to the structure itself. This assumption of frequency domain uniformity is generally not consistent with wind energy but can be approximately satisfied enough to discern structural modes [204].

Vibration based damage detection has been developed since the 1970s and a broad range of detection methods have been derived, however, choosing a method for offshore wind structures should be done with the limitations in mind. With many newer constructions, 'virgin state' modal data is collected, so there is an 'intact' condition to compare to. This comparison can be done either on derived modal parameters [205–207] or directly on the measured signal [208,209],

although the latter is challenging due to the signal's direct dependence on environmental conditions.

A limitation however is that, due to the expense and challenge of recording data from offshore wind turbines, usually only a handful of sensors are placed. For vibration approaches these sensors are usually accelerometers [210,211] but can sometimes be Fibre-Bragg Grating [212,213]. This small number of sensors makes methods such as curvature-based methods [214], or the Rayleigh quotient-based methods [215] impractical, despite their promise. Thus, any method used must work with only a few sensors. There have been studies aimed at using a combination of data sources to assist in damage detection; an example is Cevasco et.al. investigated the potential for damage detection using typical SCADA data and found that the approach showed promise given the right machine learning models were used [216]. Additionally a data fusion approach was proposed in another paper which combined accelerometer with inclinometers and found that doing so greatly improved the effectiveness in damage detection [217].

However, if only a few accelerometer signals are recorded then damage detection can be dependent on the ability to discern certain modes of the structure and hence calculate mode-based metrics. Some suitable approaches include eigenfrequencies, modal assurance criterion (MAC) and modal flexibility. Some methods, such as multiple damage location assurance criteria (MDLAC) [218] compare numerically calculated modes to measured modes, to identify a damage, one would then take the case which maximized the MDLAC value. Following such an approach, of comparing to a numerical model, necessitates a good numerical model. The modal parameters can be attained with solving the eigenvalue problem.

Although real data can be used for lifetime extension [219], there is very limited data on already failed structures. The lifetime of the structure is dependent on complex, stochastic loading [220] and there is also significant interdependence between failure modes, for example corrosion, in particular pitting corrosion, plays a role in both the initiation and propagation of cracks [221]. To overcome the lack of existing cases of identified structural damage, in exploring damage detection approaches, researchers have used numerical models of offshore jacket structures with a simulated damage and calculated damage metrics. These structures have predominantly been oil and gas platforms, but some include wind energy structures. The aims have typically been evaluation of the metric and its suitability for a generic structure. Some researchers did so with the aim of evaluating effectiveness given measurement noise; for example, Malekzehtab et. al. used mode shapes and natural frequency in an objective function, combined with a penalty term to avoid false positives [222]. Liu et. al found, in a Monte-Carlo simulation, that using modal flexibility in a damage detection problem could still be successful with 3% randomness added to the samples [223]. Other researchers aimed for a method robust to environmental factors. Wang et. al. used Modal Strain Energy Decomposition to detect damage with temperature variations [224] while others used a residual strain energy-based approach [206], [207]. Neural networks have also been employed for damage detection, for example [209,225], both of which found this to achieve even higher accuracy than other methods.

This chapter can be of significant value for those looking to conduct modal-analysis based work on offshore wind jacket structures as it gives a detailed view of the response of these structures

to several potential damages and changes in parameters. By having several damage effects presented in this chapter, this can act as a reference for either mode-based damage detection or further study. This work differs from other studies in this area in four ways. 1) This chapter uses a more detailed model to investigate damages which can not only capture their effects in a more realistic way but also allows for the investigation of other aspects. 2) Other works have typically not included the mass and moment of inertia for the Rotor Nacelle Assembly (RNA), however this aspect is important for wind turbines as it makes the modal properties not rotationally symmetric around the tower and makes a substantial difference to the mode shapes – it is found that for jacket structure, the orientation of RNA must be considered. 3) A broader range of 'damages' are simulated including global scour, corrosion and marine growth, as well as a sensitivity study to various soil parameters. By having this broad range applied to the same structure and in one chapter allows for a clear reference. 4) Mode switching is a known and troublesome phenomenon in mode-based damage detection of pseudo-symmetrical structures, this chapter explores ways of overcoming it.

In terms of damage identification, there are four levels: damage detection, damage localization, damage severity assessment and damage consequence/ progression [226]. Simply knowing that a damage has occurred, damage detection, can be of great benefit for an offshore wind jacket structure as inspection teams can be sent to investigate. Further benefit can be gained through knowing where damage has occurred, as time and resources can be saved in planning the repair of the damage. Damage quantification can provide more benefit still, allowing judgements to be made about the criticality of the damage and how it affects future operation. To achieve all four levels, it is required to have both direct sensing and dedicated models [227], this can be very expensive. A more cost-effective approach is to use virtual sensing, along with pre-existing detailed design models. Operational modal analysis (OMA) can provide information about the structure, and numerically solving the eigenvalue problem can provide information on how the structure responds to damage cases.

2.7. Modelling tools

2.7.1. Machine learning modelling tools

A wide range of tools and packages exist to implement machine learning time-series prediction, and they can be applied for estimation of wind and wave conditions. These provide application programming interfaces (APIs) for a range of programming languages. There are a wide range of machine learning packages which implement ML models, and it is a rapidly changing field, so it would be inappropriate to compile an 'exhaustive' list here. However, a list of packages and APIs particularly suitable to wind and wave time-series forecasting and simulation are given in Table 3.

Table 3. Machine learning packages which can be used for wind farm wind and wave forecasting and simulation

Package name	Base model	Details	API programming language

TensorFlow	LSTM and other RNNs	TensorFlow [228] is a neural network package from Google which can be used to implement a range of models. This is typically applied though packages such as Keras.	Python
MATLAB toolbox	LSTM and other RNNs, tree-based approaches and more.	MATLAB offers packages for applying machine learning approaches including LSTM. MATLAB is widely used in forecasting research [68,229].	MATLAB
LightGBM	Gradient boosting trees.	LightGBM [230,231] employs a tree-based, gradient boosting approach. It can train on hundreds of thousands of samples in less than a minute.	Python, C and R
AWS Deep AR	RNN	Amazon Web Services. A cloud-based implementation of RNN.	Python
Forecasting package by Hyndman	ARIMA	Hyndman provides an entire forecasting package as well as a free online text-book on forecasting [232]. This package includes a range of forecasting methods, including ARIMA.	R
Python statsmodels	ARIMA	Statsmodels [233] is a python library which has several statistical models, including ARIMA.	Python

2.7.2. Turbine array modelling tools

A comprehensive list of all notable offshore wind-farm numerical modelling tools is presented in this section. The list includes only models which have some utility in offshore applications, and which can model at a farm level. These are tools used in both industry and academia. Some additional tools which model individual turbines are also included at the end because they can play an important part if used in conjunction with farm level models.

Table 4 presents tools which use or are derived from the OpenFOAM computational fluid dynamics solver. OpenFOAM is a versatile and powerful opensource solver with a considerable community of contributors and so is an obvious choice for basing a wind farm tool. A benefit for

users, but a limitation for commercial use, is that OpenFOAM operates under a GNU General Public Licence, GPL v3. This means that not only is OpenFOAM free and open, but if a derivative of the code is shared, then the derived code must be freely available. ‘In-house’ means that the software is used by the developing organisation but not necessarily publicly available.

Table 4. Wind Farm models based on, or derived from, the OpenFOAM fluid solver.

Solver	Model Name	Model details	Availability
OpenFOAM	OpenFOAM	Generic CFD solver, based on C++. OpenFOAM is widely used in offshore wind research to implement Navier–Stokes based models. [16,160,169,234,235]	Free and open source
	FlapFOAM	FlapFOAM is an in-house extension of OpenFOAM from Fraunhofer IWES. It uses CFD simulations, with an AD turbine, to define a numerical wake model. [236–239]	In-house
	Semi-Parabollic FOAM	This semi-parabolic model was developed by Cabezón et. al. and uses multiple sub-domains, modelling the turbines with an AD approach within a parabolic model. [155]	In-house
	VestasFOAM	VestasFOAM was developed by Vestas in order to automate the CFD analysis of wind farms, such that a non-expert can do it. The model includes RANS, DES and LES turbulence models and models the turbines as actuator disks or actuator lines. [240,241]	In-house
	SOWFA	Developed by NREL, SOWFA (Simulation fOr Wind Farm Applications) is a package which allows for wind plant investigation [242]. The software can be coupled with the FSI software FAST. The package is well supported by NREL and a considerable amount of research has been conducted using this tool [243–246].	Free and open source
	SimScale	SimScale is a cloud-based computational modelling platform which integrates opensource software. It can be used in much the same way as one would use OpenFOAM, however SimScale provide the computational resources [247]. Although it appears to be used in many applications in industry, it does not appear to have been used for much wind turbine academic research. [248,249]	Commercial

Both PARK, an analytical module in WindPro, and UPMWAKE, a parabolised solver, are two of the most widely used wind farm wake modelling tools. These models, as well as their derived tools, are presented in Table 5.

Table 5. Tools which use or are based on the PARK/ UPMWAKE model

Wind flow component	Model Name	Model details	Availability
PARK	PARK	PARK in a module in WindPro which implements the Jensen model [250] and Katíc [251].	The code can easily be written [252]
	Modified PARK	Modified PARK is a MATLAB implementation which takes into account the wakes of upstream turbines. [253]	In-house
UPMWAKE/ UPMPARK	UPMWAKE	UPMWAKE is a 3D model which uses a parabolisation of the Navier-Stokes equations and uses a standard k-ε turbulence model. [254,255]	
	WAKEFARM	Developed from a modified UPMWAKE model. Fortan code from ECN. [256]	In-house
	FarmFlow	A software tool developed by ECN. [257]	In-house

Table 6 lists other solvers which have been used as the basis for wind farm tools. In general, although certainly not always, the fact that further models are derived from these solvers typically indicates that they have a level of industry acceptance and/or technical versatility greater than other models.

Table 6. Other solvers used as the basis for wind farm tools and the derived tools

Solver	Model Name	Model details	Availability
Ansys Fluent	Ansys workbench	Generic, commercial solver in Ansys workbench. Both Ansys Fluent [258] and CFX [259] are commonly used for wind turbine CFD because they are very capable solvers, universities often already have Ansys licenses and Ansys is comparatively easy to use. They also allow extra flexibility through user-defined functions. [143,151,180]	Commercial
Ansys CFX			Commercial
	WindModeller	WindModeller is an Ansys package which steamlines CFX for wind use. This is sometimes used in both academia and industry [260–262]	Commercial
Fuga	WAsP	Fuga is the turbine wake model employed in WAsP, but WAsP is a separate package which combines meteorological data to calculate power predictions. [263,264]	Commercial
WindFarmer	WindFarmer	WindFarmer is a software package for resource assessment and farm layout optimisation developed by DNV-GL. [265]	Commercial
PHOENICS	PHOENICS	PHOENICS is a generic CFD solver which aims at being reliable and cost-effective. [266] Some wind energy research has been conducted with PHOENICS [267,268], but researchers are more likely to use WindSim when applying PHOENICS.	Commercial
	WindSim	WindSim uses the RANS turbulence model in PHOENICS and is tailored to wind energy [269]. It does not seem to be as popular as OpenFOAM in terms of academic research, but some researchers have used it. [270–272]	Commercial

Other models which can be used for wind farms, and employ their own solver, are listed in Table 7.

Table 7. Other models which can be used for wind farm level CFD

Model Name	Model details	Availability
CRES-flowNS	CRES-flowNS is an NS solver which uses the k-ω model and is 3D, although a 2d version exists [273]. It is designed for wind turbines and uses an elliptic approach. [274–276]	In-house
PALM	PALM (PArallelized large-eddy simulation model) is a meteorological model which employs an LES turbulence approach. Turbines can be modelled in this software as actuator disks or AD with rotation. [277–279]	Open source
OpenWind	Designed for wind farm layout optimization by AWS, this software uses a Deep Array Wake Model (DAWM). [280–282]	Commercial
FITNAH-3D	FITNAH is a non-hydrostatic mesoscale model. [283]	
EPFL-WIRE LES	This is an in-house, LES solver from École Polytechnique Fédérale de Lausanne (EPFL) which solves a filtered NS equation. [284]	In-house
ORFEUS	ORFEUS solves linearized NS equations for wind turbine farm wakes. [285,286]	In-house
EllipSys3D	EllipSys3D has been used to implement both actuator disk and actuator line methods and the turbulence models used include RANS models such as k-ε, k-ω, LES and DES models. [235,287–289]	In-house
RIAM - COMPACT	Developed at Kyushu University, RIAM-COMPACT is an LES, code which implements turbines in an actuator line approach. [290,291]	commercial
WFSim	A wind farm model designed to be used in closed loop wind farm control. [292]	Free and open source

There are tools which, rather than having a single solver, use a suite of methods to model wind farm aerodynamics. These are typically for wind farm layout and/or control. These use less computationally intensive methods of wake models. These modelling suites are presented in Table 8.

Table 8. Modelling suites for farm layout and control

Wind flow component	Model Name	Model details	Availability
PARK, WAsP	WindPRO	WindPRO is a wind farm design tool which utilizes a suite of tools such as PARK and WAsP to make predictions. [293,294]	Commercial
Dynamic Wake Meandering	FAST.Farm	FAST.Farm [295] is an aero-hydro-servo-elastic dynamics tool developed by NREL to determine turbine loads within a wind farm. [296,297]	Free and open source (yet to be released)
	Wind Farm Modeling Program (WFMP)	Precursor of FAST.Farm, it was developed by Hao [298] based on the work of [164,299]. The module runs the FAST model of the wind turbine serially by interacting with the input file to the AeroDyn [43] module of the downwind turbines.	Free and open source
	Bladed	Bladed is a multi-body dynamics software from DNV-GL which utilizes DWM as part of its solver [44]. and application by [300]	Commercial
FLaP	FLaP	Wake modelling utility which uses a 'field model' eddy-viscosity approach. [301]	n/a
Jensen, Jimenez and others	Floris	Developed by NREL and Delft University for wind farm control, Floris utilizes a number of wake models including Jensen, Jimenez and others. [246,302,303]	Free and open source
WakeBlaster	WakeBlaster	WakeBlaster is a purpose-built solver which models the windflow over the entire wind farm and then inserts the turbines as momentum sinks. Very low computational power requirement. [304]	Commercial
Aeolus	Aeolus	Aeolus is a dynamic wind farm simulation toolbox from Aalborg University. It uses a combination of ambient field model with a turbine wake model. [305]	Free (requires MATLAB and Simulink)
UPMORO, modified UPMPARK	Furow	Furow is a commercial wind farm siting tool which includes 8 wake models, one of which is a modified UPMPARK wake model called Furowake [306] but also includes a CFD model based on OpenFOAM [307].	Commercial

2.8. Discussion

The chapter has presented an overview of a broad range of topics regarding modelling approaches for offshore wind farms. Additionally, a list of modelling tools used in industry and academia was presented. In this section these topics are discussed critically to give the perspective modeller some guidance on which models and features to choose for a given task as well as gaps for future research.

Some of the gaps in research are as follows. Aerodynamic modelling is computationally a very intensive and there is a compromise drawn between requirements and resources, however as computational capabilities increase this line is shifting. In terms of weather modelling, new approaches which can be used here are being developed from machine learning, and specifically regression modelling. There is very little research on wave load variation within wind farms for different support structures.

For modelling wind flow within a wind farm, there are several key questions to ask which guide model selection. The first two questions are 1) What is the level of fidelity and accuracy required? 2) How quickly does the solution need to be found and what hardware will be used? Upon answering these questions, one can immediately narrow the choices from a very broad range to a much narrower subset. Further details about the model can be decided by the following questions: 3) Does the model need to be transient or steady state? 4) Is the interest in the near wake or the far wake? 5) How complex should the turbine model be and how much information required for modelling is available? Answering these final three questions should narrow down the selection to something specific. This process is shown diagrammatically in Figure 1.

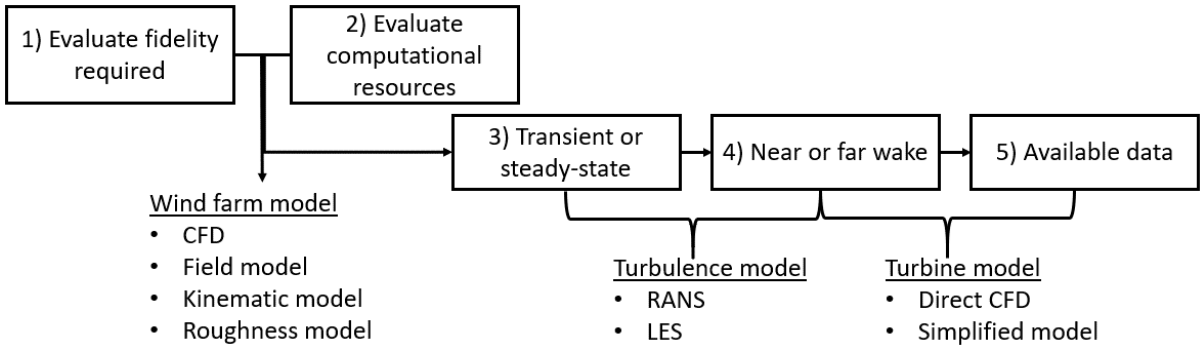


Figure 1. Diagram of model selection process

Answering questions 1 and 2 leads to a choice of whether a full CFD domain should be used or if a more simplified model with assumptions should be used such as parabolic models, field models or analytical models. If a high-fidelity approach is required and the choice is a CFD domain, this can be conducted using a generic Navier-Stokes solver such as OpenFOAM, Ansys Fluent or any of the other CFD modelling packages. A CFD model of an entire wind farm typically requires considerable processing power. If results need to be obtained rapidly, for example if many results need to be obtained, and processing power is limited, then field models or analytical

models are a better choice. In which case tools such as PARK or WAsP or others may be more appropriate.

The question of needing the model to be transient or steady state is also a consideration for CPU capabilities as a transient model will generally require more power than for a comparable steady-state model. Questions 3 and 4 can also influence the type of turbulence model required as in transient analysis LES may be preferable to RANS since RANS uses a turbulence averaging approach while LES can capture some of the stochastic nature of fluid flow. Questions 4 and 5 are both important to the choice of turbine model as a more detailed turbine model can predict near wakes better but more knowledge of the turbine performance is required for a more detailed turbine model.

Calculating hydrodynamics load can be conducted through understanding the behaviour of waves and how the structure interacts with the waves. The waves can be modelled using one of several wave theories such as Airy linear waves or Stokes non-linear wave theories. The load on the structure are generally calculated by using some implementation of Morrison's equation, which relatively easy and quick to implement. If a higher-fidelity approach is required, even a Navier-Stokes computational fluid dynamics approach can be used to model the waves interaction with particular details of the structure.

Modelling variation across a wind farm, did not raise much research interest in literature yet. However, some studies on water wave interaction with multiple-cylinder-array structures have already shown the potential high impact of the array organisation on the load distribution, due to the diffraction phenomenon. Thus, this effect could become of concern of designers for the next generation of offshore wind turbines, with bigger support structures.

Geotechnical variation within wind farms can be looked at in either a time variant or spatially variant perspective, and both ultimately impact soil-structure interaction. From a time-varying approach scour should be modelled which includes both local scour, removal of soil around the turbine, and global scour where soil is removed from the entire area. Scour models be based on conceptual studies resulting in engineering models or they can be higher fidelity where the water flow which drives the scour is modelled as well as the soil's structural resistance to scour.

As regards spatial variations, if in-situ soil tests are not present for each turbine location, such as in early planning phases, then stochastic spatial variation models are needed to extrapolate knowledge between sample locations. Kriging is a popular method for this, as well as other surrogate models such as radial basis function. The model needs to be able to make use of a limited number of data samples. These geotechnical variation models for the wind farm can then be used to inform structural models, which can be high-fidelity finite element models, or simpler engineering or analytical models such as the *p-y* design model.

Modelling of stochastic wind and wave variation requires statistical models and specifically statistical distributions and their fitting methods. For wind modelling, Weibull, Rayleigh, Lognormal and Poisson are commonly used while for waves Lognormal, Weibull, and Gumbel are used. The findings from research are that the best choice of distribution is dependent on the site and so it is important to conduct fit testing.

Machine learning approaches can also be used to model wind and wave time-series data. These models take statistical ideas and apply them in an algorithmic way. Regarding model choice and training approach, the time-horizon desired from the prediction should be considered. These approaches can either result in short-term forecasting or long-term simulation. Markov-chains are the most prominent method, however recurrent neural networks, auto-regression and tree-based approaches can also be suitable, depending on the problem.

2.9 Conclusion

In this chapter, modelling approaches for offshore wind arrays in the area of Aerodynamics, hydrodynamics, geotechnical structure and stochastic models have been presented. Although these topics are very expansive in their own rights, this chapter has attempted to concisely summarize them into pertinent findings. The main conclusions from each section are summarized as follows:

- Aerodynamics: A significant number of codes have been developed to explore the aerodynamics. These codes can follow one of essentially three approaches: 1) employ CFD solvers using either a RANS or LES turbulence method and parameterise turbines using an AD approach, although other methods exist. 2) Make assumptions specific to wind turbine aerodynamics leading to parabolised and other models, using field or wake models or 3) Base the solution on analytical expressions.
- Hydrodynamics: Wave theories and hydrodynamic models are the general approaches adopted for evaluating the hydrodynamic loads on and individual structure of an offshore wind turbine. Variation in loading from one turbine to the next can potentially have an impact on the structures fatigue life. However, it is typically not implemented in numerical models for the wind farms, as the modelling of this variation requires the use of high computation cost models difficult then to couple with the other important turbine dynamics.
- Geotechnical: Variation in soil structure can be modelled as a function of space in spatial variation using surrogate models such as Kriging or they can be a function of time modelling phenomenon of scour. Scour, being a physical process, can either be modelled through theoretical models or through modelling the physical processes themselves.
- Stochastic wind and wave models: Wind and wave loading are modelled stochastically, usually with probability distribution functions. They can either be modelled independently or correlation can be modelled but the ideal distribution is site dependant. Machine learning based approaches can also be used for time series prediction either for short-term forecasting or long-term simulation.

This chapter can be meaningful to anyone aiming to model some aspect of a wind farm at an array level. This covers not only the main modelling concepts but also relevant turbine specific modelling approaches which can lead to a more useful model. Most of this chapter covers aerodynamic modelling since this is where the majority of research interest in literature has been focused. Wind farm level models can be very complicated and consist of multiple smaller parts; this chapter discusses these smaller parts and how they fit together as well as consolidating a list of the numerical tools which implement these models.

Part A: Load-variation at an array level.

Papers:

2. M. Richmond, A. Kolios, V. S. Pillai, T. Nishino, and L. Wang, “Development of a stochastic computational fluid dynamics approach for offshore wind farms” **J. Phys. Conf. Ser., 2018**.
3. M. Richmond, A. Antoniadis, L. Wang, A. Kolios, S. Al-Sanad, J. Parol, “Evaluation of an Offshore Wind Farm Computational Fluid Dynamics Model Against Operational Site Data” Ocean Engineering – **Ocean Engineering** – **2019**
4. M. Richmond, R. Pandit, A. Sobey, A. Kolios, “Stochastic assessment of aerodynamics within offshore wind farms based on machine-learning” – **Renewable Energy**, **2020**
5. M. Richmond, R. Pandit1, S. Koukoura1, A. Kolios “Effect of weather forecast modelling uncertainty to the availability assessment of offshore wind farms” – **In revision pending final decision**

Chapter 3

3. Computational Fluid Dynamics model of a reference offshore wind farm

3.1. Introduction

Offshore wind power is growing rapidly, not only within the UK and EU but is beginning to gain traction around the world in other markets including Taiwan [6], China [308,309] and the United States [4,310–313]. To capture the maximum energy from an offshore wind farm, the most reliable assessment needs to be made of the potential wind farm before it is constructed. The models need to be able to assess the wind behaviour for the entire wind farm, rather than for a single turbine, to be able to get the most accurate picture of how the wakes interact within the wind farm and conduct further analysis such as structural modelling [314–319].

A variety of methods exist and have been used in literature for modelling the wake of turbines within a wind farm for horizontal or vertical axis turbines. These methods include Navier-Stokes Computational Fluid Dynamics (CFD) solvers where the turbine is modelled as an Actuator Disk (AD) [15,150,169] to more computationally intensive models of fully resolved wind turbines [25,169,316,320]. At the most computationally intensive end of the methods are fully resolved wind turbine. These can provide accurate answers and predict the flow realistically relative to less detailed methods but are limited to modelling only short time intervals, relative to the lifetime of the wind farm, of a few seconds due to their computationally intensive nature. There are a wide range of Navier-Stokes (NS) based approaches to choose from, actuator line is a common approach for representing the turbines which takes the idea of an actuator disk but applies it to a rotating line in a transient analysis [150,151,159]. The simplest NS based models are porous disks or momentum sinks which can provide good estimates but lack detail in their assessments [143,153]. AD is a low to medium fidelity approach where the turbine is modelled as a permeable disk whose loading influences the flow field [168]. The load can be evaluated either through looking up aerodynamic values for the blade or through solving Euler or Navier-Stokes equations over a discretised field [149]. The advantage of AD is its relative simplicity and reduced computational cost due to not having to resolve the boundary layer around the blade [150]. However, due to the inviscid assumption of the model it is inaccurate in predicting downstream wake because it under-predicts the turbulence required for wake dissipation [168,169]. For evaluating wind load variation within a wind farm, AD appears to be a good compromise method between engineering, or analytical, models and fully resolved rotors. The method is mature, at least 30 years old, with a variety of implementations. Researchers can either be interested in near wake or in the wake deficit and studies have shown that AD produces good agreement with measurements regarding wake deficit [143,151].

The turbulence model used can have a significant impact on both the accuracy and stability of the solution and can be either steady-state or unsteady. Comparisons of these have been made by other researchers [151,154,172]. Different levels of turbulence lengths scales might be modelled, for example many researchers have used Large Eddy Viscosity (LES) to model turbulence which models large turbulence length scales while averaging smaller length scales [150,159,289]. However, a less computationally expensive option is to model the turbulence using Reynolds averaged Navier-Stokes (RANS) approaches based on eddy viscosity turbulence models including different implementations of k-ε [153,321,322] and k-ω[143,323] turbulence models. As these researchers have found, using the right turbulence model can have a large impact and it's not always clear from the start which model is a better choice.

Comparisons have been conducted in wind tunnel conditions [151,169,324],[162] because those are easier to measure and compare to than full wind farms, however real wind farms have much more complex flow conditions. Some authors have modelled full wind farms but they've done so only at a very limited range of wind directions and speeds if there is any range at all – for example [289] was conducted at 7 directions but appears to have only been modelled at one far-field velocity while [160] was only conducted at one direction. To accurately assess the wind resource, it is not enough to simply assess wind flow at one input condition but rather over a range of conditions. These conditions can include inflow velocity and direction, because even a small change in either of these can have large implications for the wind farm. Some researchers have conducted wind farm level research at multiple inflow directions [156]. However, when a turbulence model is assessed it is usually assessed only at a limited range of inputs and not over a broad range [151]. Very few papers combine a range of directions necessary to evaluate a wind farm model and use wind turbine SCADA data, for example in [325] they did this but used engineering wake models and not CFD-AD models. The work presented here uses an NS based, giving the reader more knowledge of the performance of such models in this case, so that the reader can decide which type of model to use for their own purposes as well as better understanding what impact modelling choices make in this application.

The main contribution of this chapter is that a CFD simulation of an entire wind farm is conducted at multiple free-stream directions and speeds which is then compared to SCADA and met-mast data held for that wind farm, thus evaluating the approach in a real case. These simulations are conducted for two different steady-state RANS turbulence models, allowing a direct comparison. An additional contribution of this chapter is that it demonstrates a modification to an open source actuator disk code in OpenFOAM [326] which expands the code's capabilities by incorporating individual turbine performance. Methods for evaluating the model against real site data, by approximating the measurement data with regression models, is presented. The work not only demonstrates best practice which has been shown in other research, but also builds on it.

This work expands on a previous conference paper [15] with a strong focus on the CFD aspect. Specifically, much more is done to evaluate the results against the wind farm data, the construction of the model is much more detailed and twice as many CFD cases are used.

In section 2, the reference offshore wind farm and the reference data available are presented. In section 3 the CFD model is discussed in detail including the boundary conditions,

meshing strategy and the turbulence models. Section 4 presents the actuator disk model used and the modifications to it, section 5 outlines the validation approach including modelling of SCADA data for comparison, section 6 presents the results in comparison to the real data from the wind farm and discusses the findings. Finally, a conclusion is given in section 7.

3.2. Reference offshore wind farm

In this work, an offshore wind farm containing 25 turbines is modelled using actuator disks in a CFD domain. The comparisons are made to data obtained from a meteorological mast (met-mast, MM). This mast is permanently positioned on site roughly between two turbines on an outer corner. The MM has sensors which can record wind speed and direction at five elevations up to a maximum of the hub height of the turbines. These data are recorded as averages in 10-minute intervals. A map of the reference wind farm is shown in Figure 2, in this figure 'turbine 16' is circled in red. The figure shows that the reference zero direction for the CFD model is normal to the turbine rows. This orientation puts the MM in the wake of two turbines at the reference zero, but also shows some interesting behaviour at the two extremes because the met mast starts being in the wake of the turbines next to it which causes a large velocity deficit which allows for investigation of the edge of the wake. There are 25 turbines in the wind farm, but Figure 2 shows 30 to protect confidentiality of the company who provided data.

Although the orientation of the direction values was not given, it is possible to determine this out from comparison of the MM data to the map. The typical convention is wind coming from the north, going south, is zero and then the value goes up clockwise [327]. However, that assumption doesn't fit the MM data. The correct orientation for this wind farm is rotated 25 degrees counter-clockwise from a north zero. This type of sensor calibration issue is not uncommon [12].

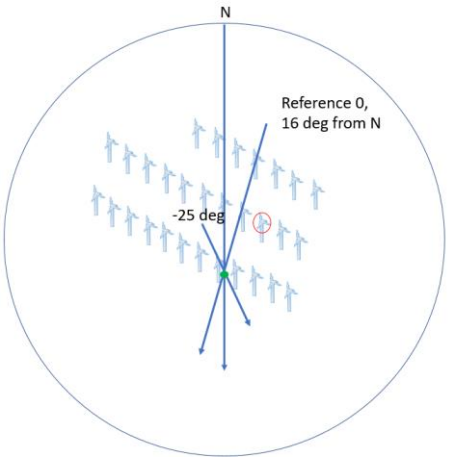


Figure 2. A map of the wind farm, turbines in red and met mast in green, 'turbine 16' circled in red, blue arrow show wind directions for reference.

3.3. Methodology

3.3.1. Governing equations and numerical discretisation.

Within this section the CFD framework is detailed. The governing equations, physics models and numerical schemes are presented. A computational aerodynamic database is constructed within a parameterization process.

Within the CFD simulation framework the governing equations of fluid dynamics are numerically solved. The Navier-Stokes equations for incompressible fluids encompass the principles of continuum mechanics and conservation laws which, along with more details on the solver used, can be found in [328] and are also given here:

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{8}$$

$$\frac{\partial u_i}{\partial t} + \frac{\partial u_j u_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j}\left(\mu \frac{\partial u_i}{\partial x_i}\right) \tag{9}$$

Where u_i is the cartesian velocity vector with components in three spatial dimensions $\vec{u} = [u_{1,}u_{2,}u_{3,}]^T$, x_i the spatial coordinate direction vector, respectively; where ρ is the fluid's density, p the pressure and μ the dynamic viscosity.

As the flow of wind turbines is turbulent of high-Reynolds number, Reynolds averaging is applied thus the final equations solved are the Reynolds Averaged Navier-Stokes (RANS). Reynolds averaging decomposes the mean flow variables to mean and fluctuating terms, where temporal and spatial ensemble averaging ensures steady state solutions are obtained. To close the set of equations Boussinesq approximation of the eddy-viscosity hypothesis is set where turbulence modelling mathematically describes the eddy viscosity μ_t.

The governing equations are formulated within the finite volume numerical framework and discretized in three space dimensions, where the SIMPLE algorithm is employed to obtain numerical solutions [329,330]. This is a widely used solver algorithm which has been used by other researchers in modelling wind turbines, including in OpenFOAM's simpleFoam [331] solver, which is for incompressible, steady flow without changes in temperature [169,332,333]. The SIMPLE algorithm solves the governing equations in a sequential approach, iteratively until convergence is reached.

The governing equations are discretised in time and the time derivatives are set to zero; the gradient scheme is the standard Gaussian finite volume discretisation, where the scalar gradient is computed using Green-Gauss theorem [334] and the values at the cell faces are linearly interpolated from the cell centre values. The divergence schemes are the second order Gauss for the velocity field and bounded Gauss for k, ε and ω. More on the OpenFOAM numerical schemes can be found on the OpenFOAM user guide [335].

The convergence criteria were set to 10^{-8} for difference in turbulence kinetic energy, k, and each velocity component and was set to 10^{-7} for the difference in pressure. These convergence criteria were reached after around 5000 iterations, with some variation of +/- 1000 for all case set-ups.

Tracking values for velocity and turbulence kinetic energy at the met-mast during solving, the values used were appropriate as the physical values converged in around as many iterations.

3.3.2. Turbulence models

Several published works compare turbulence models used with AD models for wind turbines wake models, for example some researchers have found that standard *k-*ε under predicts wind speed deficit [151,154]. Some researchers, comparing *k-*ε and *k-*ω models recommend *k-*ε because the high level of turbulence in *k-*ω leads to quick wake recovery and hence less accurate predictions in mid and far wakes regions [332]. Although modified versions of the *k-*ε and *k-*ω models can be less robust, they generally perform better than the standard models [154,169]. Therefore, to use the most suitable model for this work, a comparison of several RANS turbulence models was conducted. The three turbulence models compared are standard *k-*ε, Realizable *k-*ε and *k-*ω SST.

Standard k-ε *turbulence model*

The standard *k-*ε model is a widely used turbulence model developed by Launder et al. [152]. This model calculates a turbulence kinematic viscosity, ν_t, based on the turbulence kinetic energy, *k*, and dissipation rate, ε. The standard transport equations are:

$$\frac{\partial k}{\partial t} + \bar{u}_j \frac{\partial k}{\partial x_j} = \frac{\partial}{\partial x_j}\left[(\nu + \frac{\nu_t}{\sigma_\epsilon})\frac{\partial k}{\partial x_j}\right] + \nu_t \frac{\partial \bar{u}_i}{\partial x_j}\left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i}\right) - \varepsilon \tag{10}$$

$$\frac{\partial \varepsilon}{\partial t} + \bar{u}_j \frac{\partial \varepsilon}{\partial x_j} = \frac{\partial}{\partial x_j}\left[\left(\nu + \frac{\nu_t}{\sigma_\epsilon}\right)\frac{\partial \epsilon}{\partial x_i}\right] + C_1 \frac{\epsilon}{k} \nu_t \frac{\partial \bar{u}_i}{\partial x_j}\left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i}\right) - C_2 \frac{\varepsilon^2}{k} \tag{11}$$

Where $x_{i,j}$ are Cartesian space coordinates, $u_{i,j,k}$ are mean stream-wise velocity in the denoted Cartesian coordinate. ν is the kinematic viscosity and C_μ, C_1, C_2, σ_k and σ_ϵ are standard constants given in the original publication from Launder et al. [152] these values have been used by other authors in modelling wind turbines [151,169], [332], however some authors have shown that different choices of constant values are more appropriate for atmospheric flow [336].

From which the turbulence kinematic viscosity can be calculated using:

$$\nu_t = C_\mu \frac{k^2}{\varepsilon} \tag{12}$$

Realizable k-ε *turbulence model*

The realizable eddy viscosity model was developed by Shih et al. [337] and implements a realizable eddy viscosity formulation through modified dissipation rate equation. This modified dissipation rate equation is [338]:

$$\frac{\partial \varepsilon}{\partial t} + \bar{u}_j \frac{\partial \varepsilon}{\partial x_i} = \frac{\partial}{\partial x_j}\left[\left(\nu + \frac{\nu_t}{\sigma_\epsilon}\right)\frac{\partial \varepsilon}{\partial x_i}\right] + C_1 S\varepsilon - C_2 \frac{\varepsilon^2}{k + \sqrt{\nu\varepsilon}} + S_\varepsilon \quad (13)$$

Where $S\varepsilon$ is the mean strain rate and v is the kinematic viscosity. The difference being that the Reynolds stresses do not appear in this equation. The standard model constants are given in the original publication [337].

K-ω shear stress transport

The *k-ω* shear stress transport (SST) model was introduced by Mentor el al [339]. This model uses *k-ω* from Wilcox [340] to calculate the inner boundary layer flow and a *k-*ε model in the free-shear flow.

The *k-ω* equations for turbulent kinetic energy, *k* and specific rate of dissipation, *ω* are:

$$\frac{\partial k}{\partial t} + \bar{u}_j \frac{\partial k}{\partial x_j} = P_k - \beta^* \omega k + \frac{\partial}{\partial x_j}\left[(\nu + \sigma_k \nu_t)\frac{\partial k}{\partial x_j}\right] \quad (14)$$

$$\frac{\partial \omega}{\partial t} + \bar{u}_j \frac{\partial \omega}{\partial x_i} = \alpha S^2 - \beta \omega^2 + \frac{\partial}{\partial x_j}\left[(\nu + \sigma_\omega \nu_t)\frac{\partial \omega}{\partial x_j}\right] + 2(1 - F_1)\sigma_{\omega 2}\frac{1}{\omega}\frac{\partial k}{\partial x_i}\frac{\partial \omega}{\partial x_i} \quad (15)$$

Where x_j is Cartesian space coordinate, ν is kinematic viscosity, F_1, is a blending function, P_k is a production limiter, and $\sigma_{k,\omega,\omega 2}$, β, and α are closure coefficients given in the original *k-ω* formulation from Wilcox [340].

3.3.3. Model set-up

The velocity inlet uses a fully developed wind shear profile. The logarithmic profile was fitted to an averaged set of met mast wind velocity samples, which records velocities at different heights. Only samples for which the met mast was in the free stream, and not in the wake of the wind farm, were used. There are some potential limitations to this, for example that this is the opposite to the direction at which the CFD cases used to build training data were conducted at; however, it gives the most accurate assessment of the local wind shear profile possible. Figure 3 gives an example of what the samples are like, each dot represents an individual value and the five lines each link points from the same sample for the sake of representation. The velocity at 58 m was limited to 7.5 m/s +/- 0.1 m/s in this plot.

The velocity profile used is of the form in (16) which comes from the OpenFOAM API guide [341].

$$U = \frac{u_*}{K}\ln\left(\frac{z - z_g + z_0}{z_0}\right) \quad (16)$$

Where *K* is the von Karmen constant, z_g is the ground surface height, z_0 is the roughness height and u_* is given by (17). From this study, it was found that the most suitable value for roughness height for this location was z_0 = 0.015. Although since waves are predominantly wind driven, z_0 would change for a given wind speed.

$$u_* = K \frac{U_{ref}}{\ln\left(\frac{z_{ref} + z_0}{z_0}\right)} \tag{17}$$

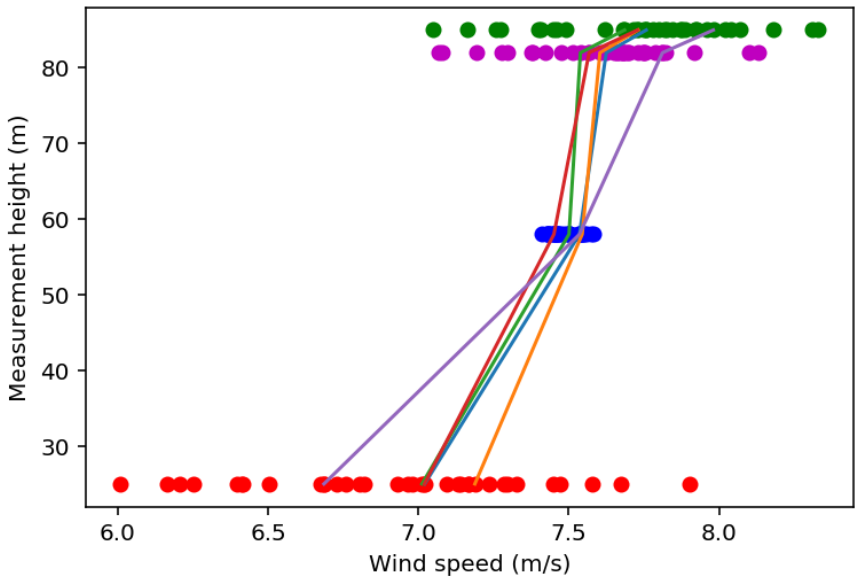


Figure 3. Met-Mast samples of wind velocity (m/s) against measurement height (m). Lines are drawn between points which come from the same sample for 5 randomly chosen samples. Green, purple, blue and red dots indicate samples from different heights; the colors of the lines are only to differentiate the lines.

The atmospheric boundary layer model used is the standard atmospheric boundary layer in OpenFOAM which is the model presented by Richards et. al [342]. In all of turbulence models a turbulence kinetic energy value of 1.3 was used. Wall functions were used for the sea surface using the inlet values. The top wall uses a slip condition; however, it is 2 km above the turbine hubs so should have no noticeable impact on the wind turbines.

3.3.4. Meshing strategy

The grid used is a structured grid with over 63 million cells. To construct such a large grid, OpenFOAM's blockMesh utility is used on an academic high-performance computer cluster which has 118 compute nodes, each node having 16 CPU cores with a base frequency of 2.10 GHz and 128 GB of shared memory. The domain is much larger than the area of the wind farm to avoid the influence of the walls on the turbines; the domain is 16 km wide, 6.8 km long and 2 km tall. The grid is refined in such a way to have smaller cells around where the turbines are located and larger cells further away from the turbines.

Despite being such a large grid with over 63 million cells, there are roughly 12 cells across the diameter of the turbines. This is a consequence of having a structured grid with such a large domain. The advantage of a structured mesh however is that it ensures a high-quality mesh.

A grid sensitivity study is conducted to determine suitable cell sizes with a base set of values used being taken from a paper on validating OpenFOAM's actuator disk model [156], though a more fine, structured grid is used in the end. The domain used for the work is shown in Figure 4 the grid used is demonstrated in Figure 5.

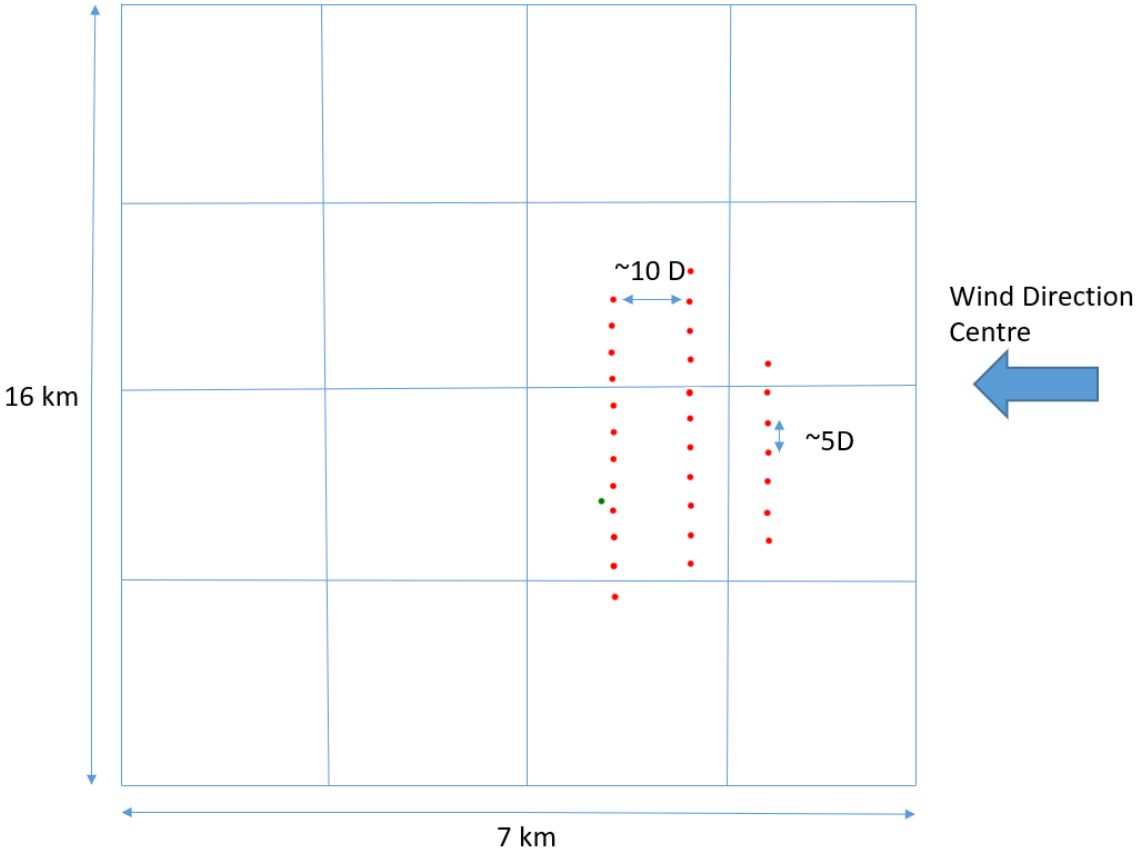


Figure 4. Domain map. Red dots represent turbine locations, green dot represents MM location. D is the turbine diameter. Blue lines represent domain partitions. The domain was also partitioned horizontally for a total of 32 partitions.

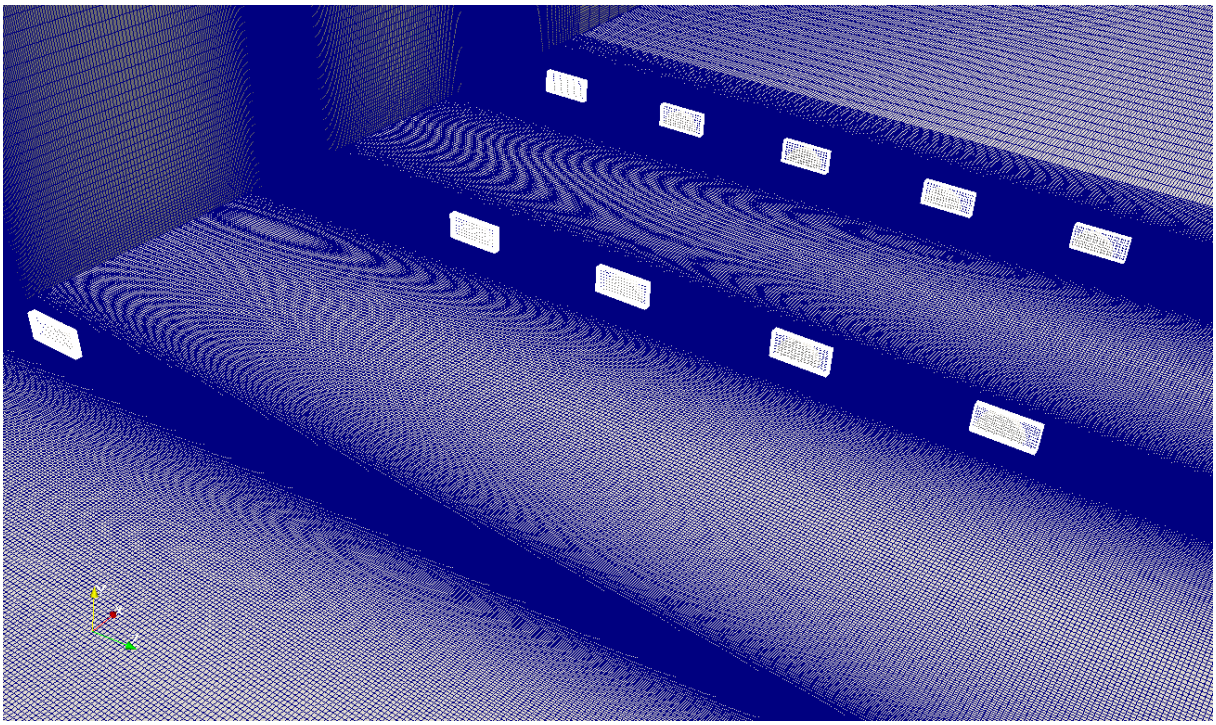


Figure 5. Demonstration of grid cut away vertically through the centre and horizontally at 80 m. Turbines represented as cylinders and a wireframe of the surrounding grid shown in white.

3.4. Actuator disk model

3.4.1. Original actuator disk code

The simpleFoam code is modified to implement an actuator disk model, with tangential force momentum source, to represent the turbines. The code used for the study is based upon Svenning's actuator disk model [326] which represents the turbine not as an infinitely thin disk but as a three dimensional region where cell centres within the region are assigned new momentum terms. This gives more flexibility in changing the direction of the turbines.

3.4.2. Modifications to actuator disk code

To implement turbine performance it is necessary to determine thrust and torque curves for a given wind speed, such as the curves which can be found for the NREL 5-MW turbine as shown in [343]. This was done by first filtering out data not representative of general turbine performance, e.g. speed below cut-in and pitch higher than usual values. Then generator power was converted to rotor power through an efficiency curve. Rotor speed was used to convert rotor power to rotor torque and finally velocity at the disk was used to convert rotor power to rotor thrust. This is resulted in clear performance curves which are very similar to the NREL curves in general shape. Scatter plots of the turbine performance are shown in Figure 6.

To automate selection performance for each turbine, that is the perpendicular and tangential components of velocity, this performance needed to be modelled and the curve-fitted approximation model needed to be integrated into the AD code. Polynomial functions were then

fitted to this data using the method of least squares, two polynomials for each value with the change occurring at rated power. From cut-in speed to 10.85 m/s a fourth order polynomial was used for both thrust and torque, after 10.85 m/s, a third order polynomial was used for thrust and a straight line for torque. These shapes were chosen after comparison to other shapes, they fit this particular data set the best. The use of a polynomial model was chosen because it fit the data very well, as can be seen in Figure 6, and also it could easily and efficiently be integrated into the OpenFOAM code.

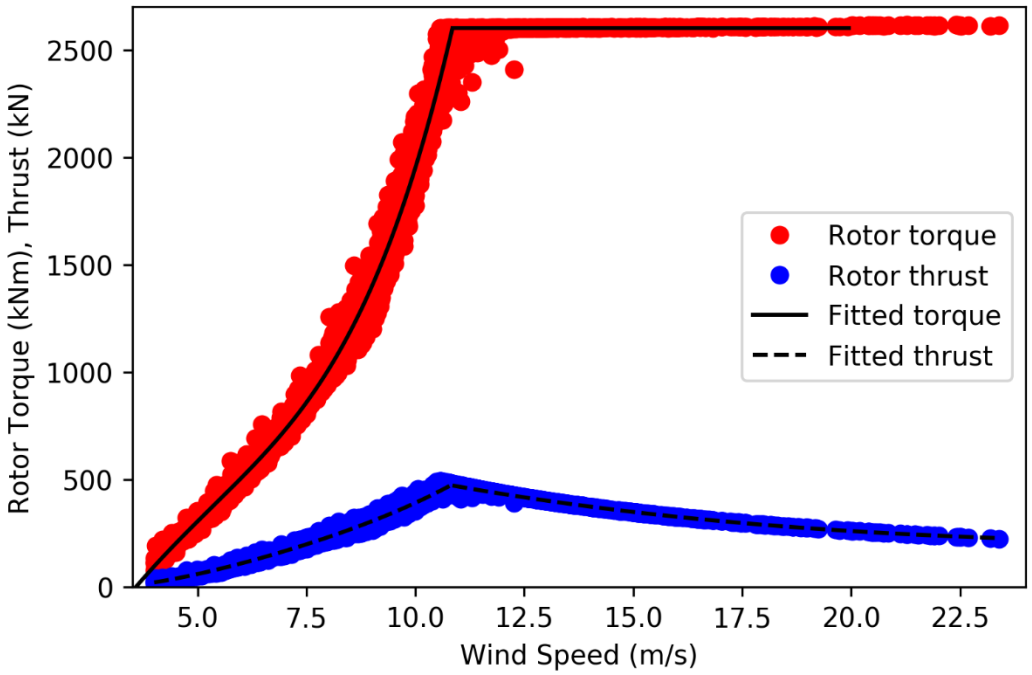


Figure 6. Plot showing rotor thrust (blue scatter) and torque (red scatter) as well as fitted torque (black solid line) and fitted thrust (black dashed line)

These performance curves were based on inflow velocity, however this isn't what is observed by the turbines within the wind farm, so instead it was necessary to determine a theoretical upstream velocity based on the velocity at the disk and AD theory. The theoretical upstream velocity was used instead by the code to determine the thrust and torque values for each given turbine. Fundamental AD equations were used for this which come from [344]. These equations are:

$$U_D = U_\infty(1-a) \tag{18}$$

$$C_T = \frac{Thrust}{\frac{1}{2}\rho {U_\infty}^2 A_D} \tag{19}$$

$$C_T = 4a(1-a) \tag{20}$$

Where U_D is the velocity at the disk, in the CFD code this is velocity averaged over the disk volume and in the SCADA data this is velocity measured by the anemometer on the nacelle, this is a source of some minor discrepancy. U_∞ is the far-field velocity, C_T is the coefficient of thrust, A_D is the area of the disk and a is the axial induction factor. Equations (18 - (20 were used to derive the following equation for axial induction factor in terms of thrust and velocity at the disk:

$$a = \frac{B}{4 + B} \tag{21}$$

Where:

$$B = \frac{Thrust}{\frac{1}{2}\rho {U_D}^2 A_D} \tag{22}$$

This equation is solved iteratively within the AD code, for each turbine at each solver iteration, as thrust is dependent on axial induction. Finally, the original code used two point values (*x,y,z*) to define the 'start' and 'end' points which define the yaw of the turbine; all turbines set to a single direction. The same direction assumption is not necessarily true and the difference in mean direction between two turbines is an average of around 9% for this SCADA data, however this assumption was made for simplification. The process followed by the code is shown in Figure 7.

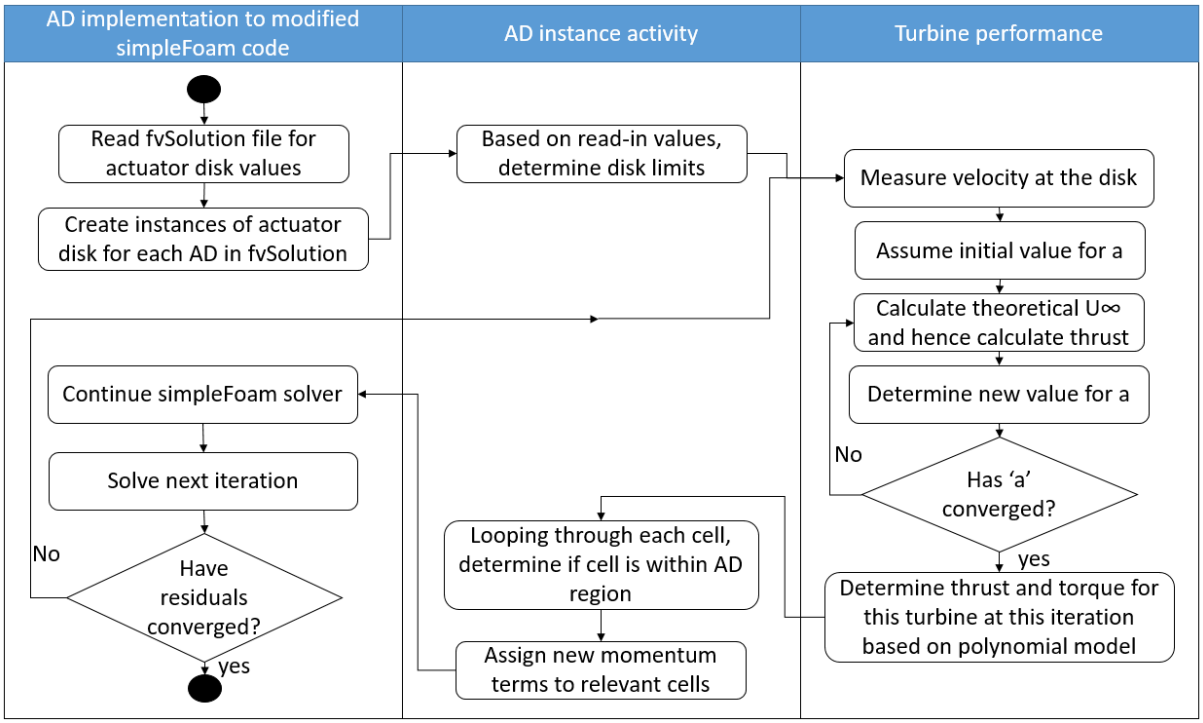


Figure 7. Flowchart showing modified AD code process

3.5. Validation approach

Samples in the MM data occur at random input conditions. To overcome this, samples close to the CFD input conditions were taken, put into bins and averaged. The bins are made as small as possible to only take representative samples but still be able to determine an average, their size were chosen in order to have between 8 and 10 data points in each. Additionally, there are very few samples around 15 m/s inflow velocity but a very large amount around 7.5 m/s. Even with thousands of samples, some of the limits had to be large to ensure there were enough samples in each bin. Different limits were chosen for each inflow velocity with the aim of having good representation for each point, between roughly 3 to 10 samples each. The bin sizes as well as the MAE for each inflow velocity is presented in Table 9.

Table 9. Bin sizes for each inflow velocity

Parameter	5.0 m/s	7.5 m/s	10.0 m/s	12.5 m/s	15.0 m/s
Velocity limit +/-(m/s)	0.7	0.2	0.2	0.6	1
Direction limit +/- (degrees)	0.5	1.0	3.0	8.0	15
MAE (%)	11.77	9.06	7.03	10.27	7.98

The values compared for the CFD are Root Mean Squared Error (RMSE) and Mean Absolute Error both as a value and a percentage of the value being predicted and are calculated using standard equations. The error is determined in comparing the CFD results to the real data from the wind farm.

With so much scatter in the met-mast and SCADA data, it is difficult to discern the general behaviour in the data from a scatter or other types of plots.

Although putting the data into averaged bins is a commonly used and straight forward approach, the resulting bins can be relatively large in some places, as shown previously in Table 9. Additionally, the observations within each bin are not weighted, therefore observations far from the bin centre have as much impact on the average as those far away. Regression models have the potential to be more representative of the data. Therefore, to determine the behaviour of the data as well as for comparison to the CFD values, regression models were applied to the SCADA/MM data. There are two independent variables, inflow velocity and inflow direction, which were used to predict the one independent variable, met-mast velocity. After restricting the direction to +/- 35 degrees from the zero-reference direction, data set was reduced to 5,046 samples.

For the regression modelling, the ‘regression learner’ app in Matlab’s ‘Statistics and machine learning toolbox’ was used which has a wide selection of regression methods. All of the models available in the app were tested and the ones which fit the data best, based on their RMSE values, were Gaussian Process Regression using an exponential kernel function and an ensemble of trees using bagged trees with a minimum leaf size of 25 and 2000 learners.

Gaussian process regression employs a Bayesian approach where a prior, which is a Gaussian distribution based on a 'noise' parameter for each point, is combined with observations to form a posterior [345–347]. In Gaussian process regression the prediction is given by[348]:

$$P(f|X) \sim N(f|0, K(X,X)) \tag{23}$$

Where *f|X* is the mean function and *K(X,X)* is the covariance matrix. *K(X,X)* is determined using the exponential kernel which defines it as [349]:

$$k(x_i, x_j|\theta) = \sigma^2{}_f \exp\left(-\frac{r}{\sigma_l}\right) \tag{24}$$

Where σ_f and σ_l are signal standard deviation and characteristic length scale respectively. The Euclidian distance between points x_i and x_j is defined as:

$$r = \sqrt{(x_i - x_j)^T (x_i - x_j)} \tag{25}$$

An ensemble of trees approach combines the predictions of multiple decision tree models, with their segregation of the variable space randomly determined. The use of multiple trees can reduce the level of overfitting to the data compared to a single tree [63,64]. For a regression problem, the prediction value is the weighted average using selected trees [350].

$$\hat{y}_{bag} = \frac{1}{\sum_{t=1}^{T} \alpha_t I(t \in S)} \sum_{t=1}^{T} \alpha_t \hat{y}_t I(t \in S) \tag{26}$$

Where α_t is the weight of the tree, $\hat{y}_t$ is the prediction from tree *t*, and *S* is the set of trees used in the prediction.

The goodness of fit parameters for these two models in predicting the data are given in Table 10. These values are limited by the scatter in the data. These two independent variables can predict a significant amount of the variation in the results, but do not completely account for the variation. However, for determining the general trend in the data and in comparison to the CFD results, which used the same independent variables, the results from these models are suitable based on their goodness-of-fit parameters.

For calculating the comparison values of RMSE and MAE, only the results from the ensemble method are used, not the GPR method. This is because ensemble is a simpler but more robust method, as the ensemble method takes averages from within multiple partitions, the result is more constrained to where the data points are, rather than GPR which is also somewhat constrained to forming a continuous function, which is not necessarily representative of the data. However, GPR is used for representation as it provides a smooth function which shows the behaviour well.

Table 10. Goodness-of-fit values for regression models to site data

Model	RMSE	R^2	MSE	MAE
GPR	1.06	0.89	1.11	0.74
Ensemble	1.09	0.89	1.18	0.76

3.6. Results and discussion

3.6.1. Grid sensitivity study

The discretisation error of the spatial numerical scheme can be assessed and evaluated with a grid sensitivity analysis. The grid sensitivity analysis entails the consistent generation of consecutively refined grids, at least three, simulations with the same numerical setup and evaluation of representative quantities [351]. Four grids are constructed, the flow conditions are set as follows: velocity magnitude measured at 80 meters above sea-level was used as a comparative value since much of the future comparisons are made from this. The simulations are performed at 0 degree inflow direction at 10 m/s. Grids specifications are shown in Table 11 and solver times are using 2 nodes, with 16 cores each on the high-performance computer described in section 3.4.

Table 11. Grid statistics and solver time for grid sensitivity study including for each grid the number of cells, points and time taken for the solver to finish.

Grid reference	1	2	3	4
Cells (e+8)	17.49	36.48	63.30	124.17
Points (e+8)	18.07	37.30	64.39	125.87
Time (hours)	7.15	12.73	21.47	48.03

For comparison, the velocity at the met-mast at 80 m was used. The results from this are shown in Figure 8.

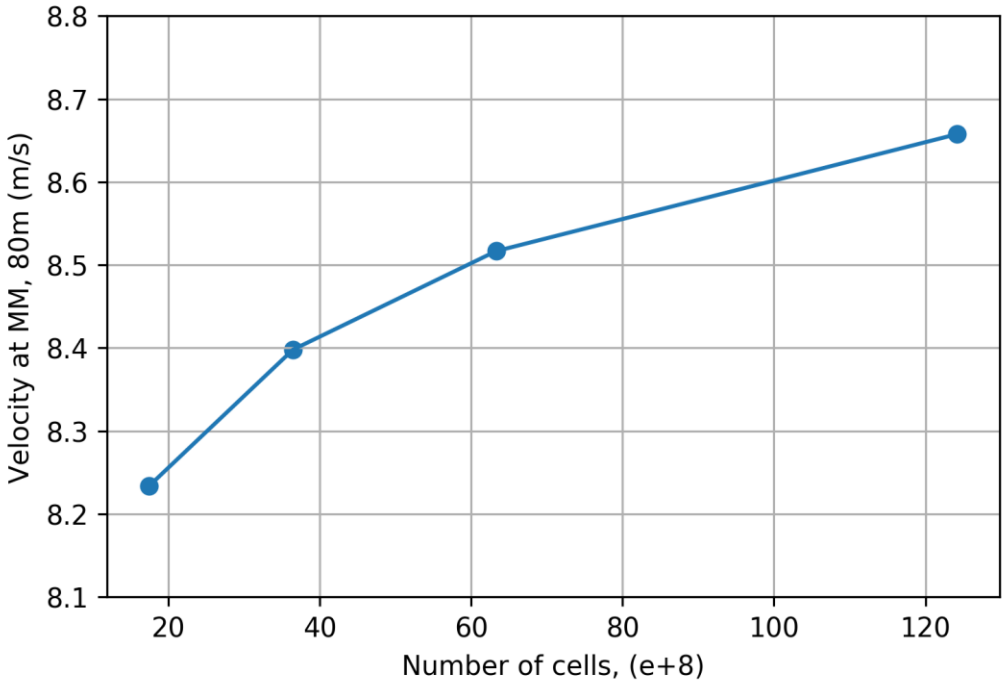


Figure 8. Results from grid convergence analysis. Velocity magnitude at the met-mast, 80 m above sea level along the y-axis, against the number of cells in millions.

The grids were increased by increasing the number of cells along all directions by a certain, consistent amount and resulted in a grid refinement ratio of around 2 (+/-0.1). Given that a second order scheme is used, the Richardson extrapolation can be used to determine an estimate for the continuum value [352].

$$f_{h=0} \cong f_1 + \frac{f_1 - f_2}{r^p - 1} \qquad (27)$$

Where $f_{h=0}$ is the estimated value, f_1 and f_2 are the results from the first and second finest grids respectively, r is the refinement ratio and p is the model order.

Using the results obtained gives an estimated continuum value of 8.70 m/s, this is 0.5% higher than the value in grid 4 and 2.2% higher than the value for grid 3. The grid 3 level of refinement was used for the studies, this is due to resource constraints, given the number of cases to run it wasn't possible to use the highest refinement.

3.6.2. Turbulence model sensitivity study

The initial turbulence comparison was done at 0 degrees free stream velocity (normal to the rows along which the turbines are aligned) and at 15 degrees clockwise from this. The data compared is the velocity at the met mast.

The results from this comparison are in line with literature which states that the *k-ε* model is more accurate for this case and that the *k-ω* SST model over predicts the intensity of the wake [151]. The results from this comparison are shown in Figure 9 along with predictions from the two regression models which were trained on the SCADA data. As can be seen in the figures, the *k-ω* SST model in this case consistently shows much lower values than is observed by the real met-mast, while k-*ε* generally performs more in line with the real data, as is consistent with the findings of other researchers [332]. At higher velocities, around 15 m/s, and at 0 degrees, all models predict velocities that are too low. This under prediction by the models is partially due to having so few data points at higher velocities for comparison, and so these points come from slightly further away from the centre, where the wake is less prominent.

It is still possible that *k-ω* as implemented performs better than *k-ε* for some combinations of initial conditions, as is observed in both 0 and 15 degrees for 5 m/s, therefore the 35 cases are conducted with both a *k-ω* model and a *k-ε* model. The *k-ε* model chosen for this is standard *k-ε*. This is because standard *k-ε* is shown to predict closer to the MM data.

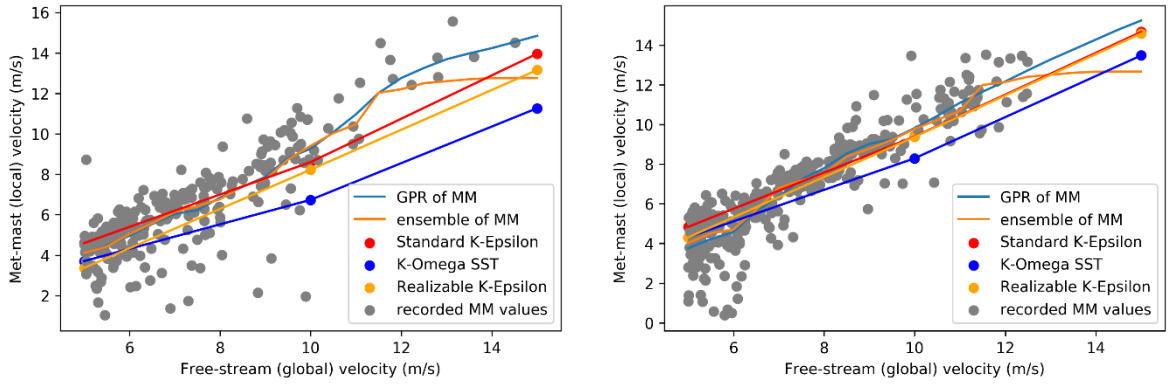


Figure 9. Met mast velocity comparison for the turbulence models investigated at inflow directions (a) 0 degrees and (b) 15 degrees.

3.6.3. Case set-up

A wide range of cases were conducted to evaluate the model over a spectrum of conditions. The wind farm is modelled at a range of incoming wind angles between 35 degrees and -35 degrees and the freestream velocity is modelled between 5 m/s and 15 m/s. All combinations of direction including 0 degrees as well as plus and minus 10, 20 and 35 degrees are modelled with freestream velocities of 5, 7.5, 10, 12.5 and 15 m/s, resulting in 35 cases. The limit of incoming wind direction is chosen to capture a board range of inflow directions while still capturing a relatively high resolution in terms of angular variation.

3.6.4. *Comparison to site data approximation*

The numerical solutions are compared with real data to identify potential discrepancies and evaluate the levels of computational uncertainty. Therefore, data was extracted along a vertical line where the met mast is located to make predictions of local wind velocities.

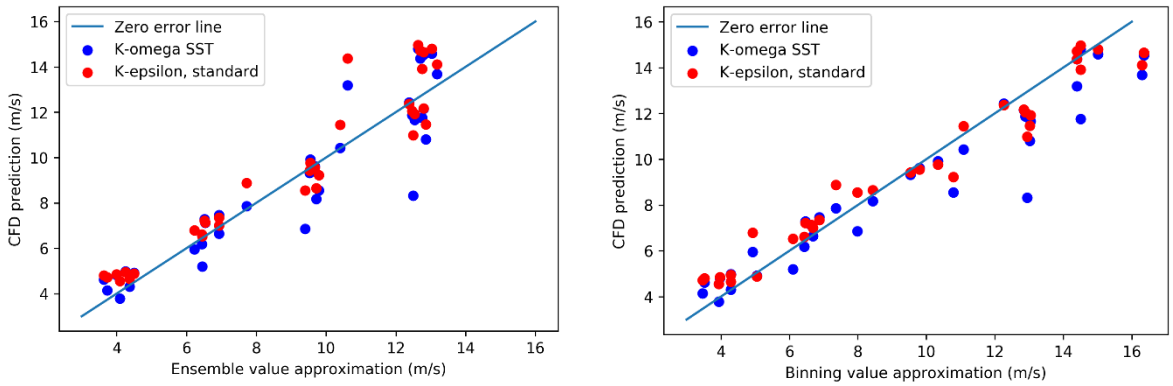
Once the extracted data are generated from the CFD model, they are compared to the wind farm data samples to determine their accuracy. From each of the 35 CFD cases, local velocity values at the met-mast were sampled at hub height, the same as one of the sensors on the met-mast. The results from this comparison are shown in Table 12. The mean absolute error for the k-ε is 0.761 m/s, approximately 8 % relative to the comparison values and k-ω SST is 0.932 m/s approximately 10.5%, using the binning approach. This is roughly in line with the error values achieved in CFD-AD by other authors comparing to wind tunnel measurements [151,169] as well as for wind far measurements [156].

Table 12. Metrics for computational fluid dynamics results

	Binning		Ensemble	
Error Measure	K-Epsilon	K-Omega	K-Epsilon	K-Omega
RMSE	0.96	1.34	1.17	1.32
MAE (m/s)	0.76	0.93	0.89	0.96
MAE (%)	8.06	9.87	9.88	10.69

The results from the CFD analysis are compared to the regression model and binning approaches for each individual data point in Figure 10, which is a prediction against true plot where a 45-degree line indicates zero error between the two values. For both the binning and ensemble approaches there is good agreement to the CFD results, however the values are generally closer for the binning approach. This increased accuracy of the binning approach is partially due to an averaging of the result; the ensemble approach can show sharp fluctuations as the data fluctuates which would be averaged out in the binning approach. This is can be a desirable behaviour to some extent, but if there is too much averaging than it ceases to capture the behaviour of the real situation.

Figure 10 also shows larger deviations from the zero-error line in both approaches at higher wind speeds. This is at least partially due to a lower amount of data present at these higher values and so both approaches become less reliable. Additionally, this could be due to higher unsteadiness at higher velocities, if standard deviation values were present in the SCADA and met-mast data, this could be investigated.



(a) *(b)*

Figure 10. Approximation method to measured data against prediction plots comparing the MM velocity values for CFD results to (a) ensemble regression (b) binning

Figure 11 shows the CFD predictions compared to the measured data from the SCADA/MM for 7.5 m/s free-stream condition. This figure shows the level of scatter present in the measured data and necessitates the use of the regression models. This also shows that a significant portion of the difference between CFD prediction and data comes from the difficulty in getting a reliable representation of the measurement data. As can be seen from other researchers [156,353] measured data is typically not as smooth and determinate as steady-state CFD results.

Figure 12 shows the results for the CFD model and the regression models, normalised by inflow velocity and averaged for inflow direction. At 0 degrees the MM is in the direct wake of two turbines, as seen in Figure 13, which causes a minor dip in the wind flow velocity; this dip is much higher for the *k-ω* SST model than for the *k-ε* model. There are numerous things which can affect turbulence kinetic energy including calibration of constants, initialisation, boundary conditions and so on. In this case, the *k-ω* SST model predicts lower turbulence viscosity than the *k-ε*, and researchers have found that a lower turbulence viscosity results in less wake recovery [169].

Towards -35-degree inflow direction there is a significant drop in the MM velocity shown by the regression models which is captured more closely by this k-ω SST turbulence model. The results from this direction are shown in Figure 14. Although the met mast is roughly equidistant between the nearest two turbines, this significant drop is not observed in the +35-degree direction. The drop at -35 degrees appears to be an interaction of both the near wind turbine and two upstream turbines.

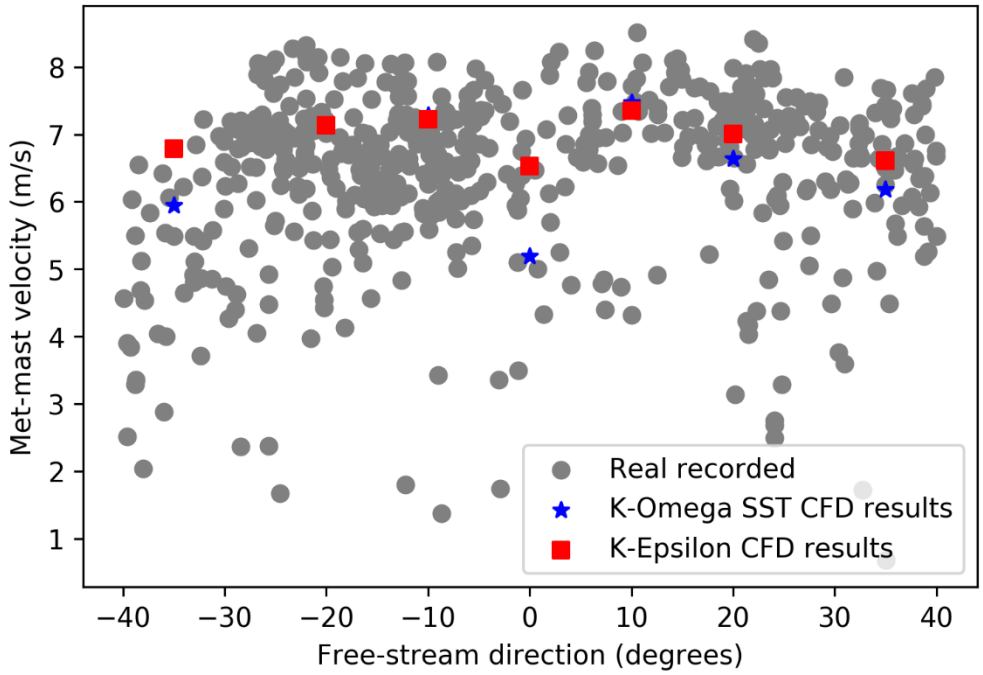


Figure 11. CFD predictions compared to measured data values by direction for 7.5 m/s free-stream condition.

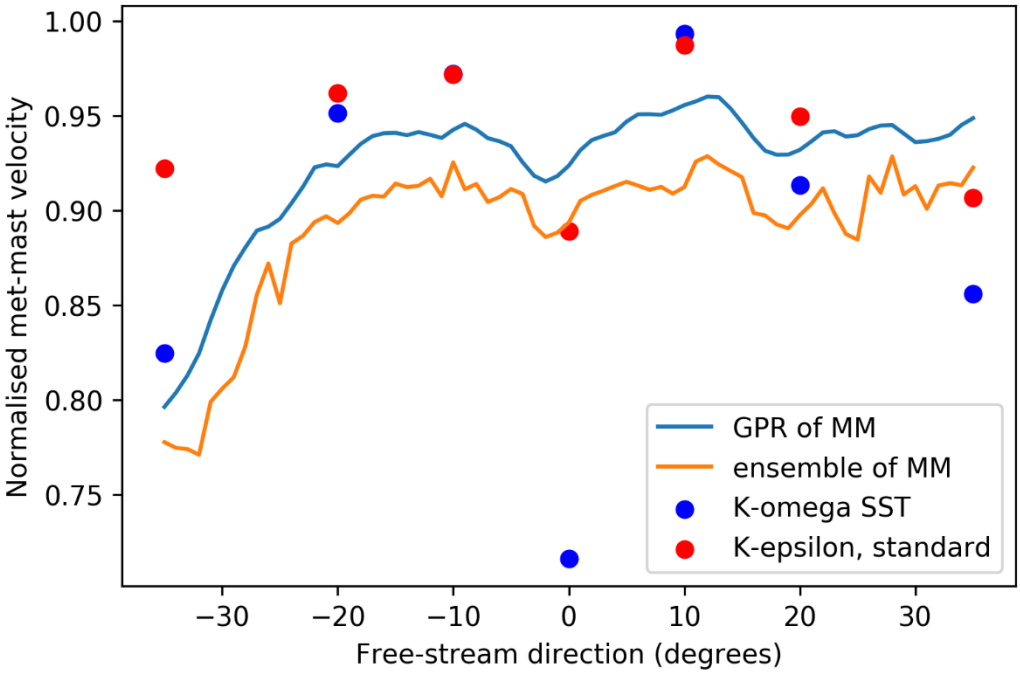


Figure 12. Averaged by wind direction and normalised MM velocity values against inflow direction. Showing the CFD compared to both of the regression methods.

A typical result from the simulation in shown in Figure 13 and Figure 14 which show a two-dimensional cross section of the flow field through the turbine centres. This image shows the

velocity magnitude field. The case is set up as described in the methodology section and the inputs for this particular case are inflow velocity of 10 m/s and an inflow direction of 0 degrees (normal to the three rows). Several observations can be made from this plot. This plot shows how, at this direction, the met-mast is in the wake of two turbines, and hence there is a velocity reduction at the MM. Another observation is that there is more of a velocity deficit on the right of the wake, when viewed in the direction of wind flow, than on the left. This is due to how the turbines interact with the wakes from upstream turbines. The third observation is that there appears to be a global reduction in velocity – even where there is no immediate wake, the velocity is nearly one m/s lower on the left side of the figure compared to the right of the figure.

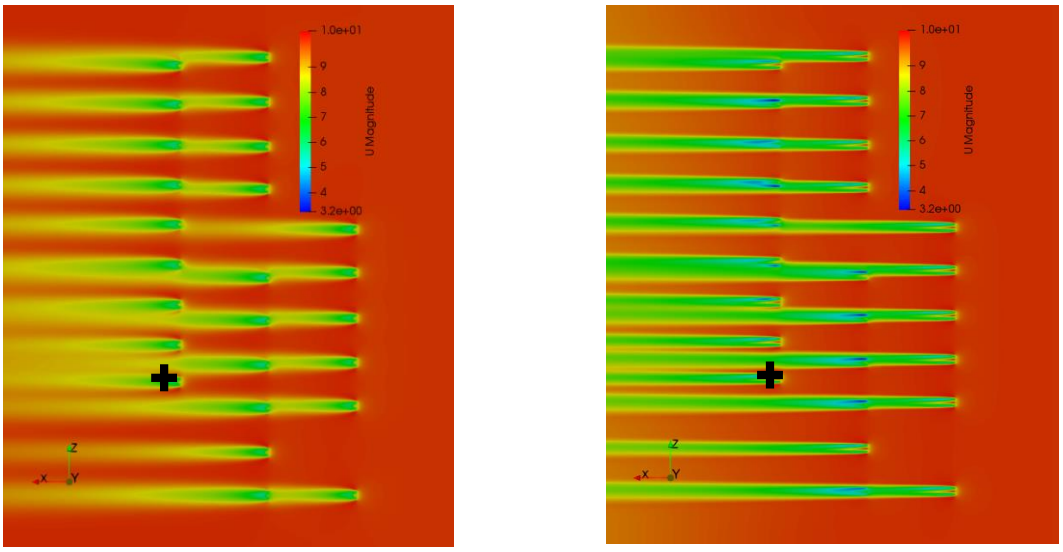


Figure 13. Velocity-magnitude field from steady state actuator disk model at hub height. k-ε model left and k-ω SST right. Met-mast location denoted by a black plus. 0-degree inflow, 10 m/s.

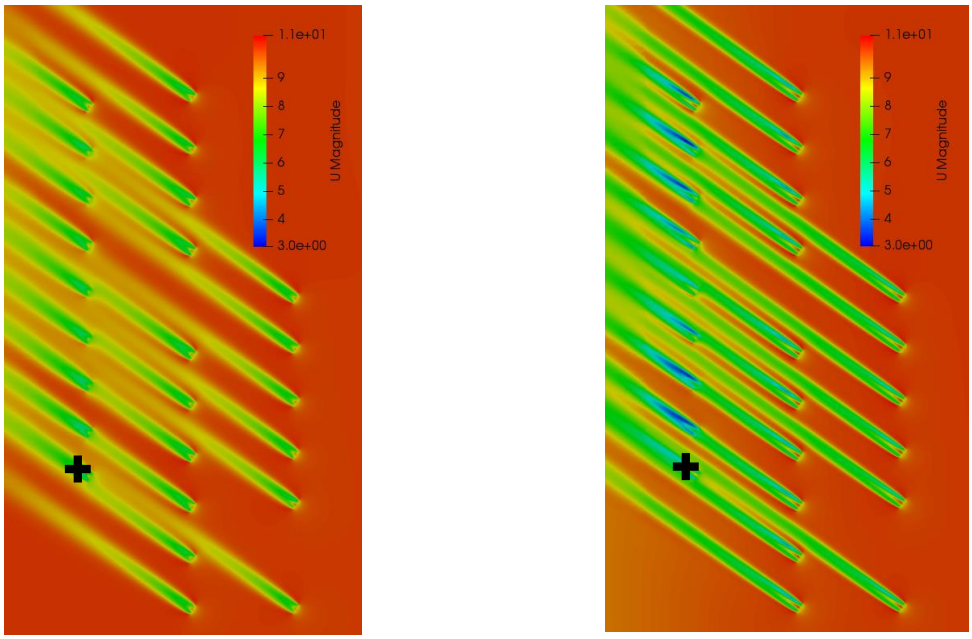


Figure 14. Velocity-magnitude field from steady state actuator disk model at hub height. k-ε model left and k-ω SST right. Met-mast location denoted by a black plus. 35-degree inflow at 10 m/s.

The turbulence viscosity for *k*-ε, where the freestream direction is 35 degrees with a free-stream velocity of 10 m/s, is shown in Figure 15. Throughout the chapter, the *k*-ε model is mainly assessed due to it performing better for this model set up. This shows that the air downstream of the wind farm is still turbulent for a large distance. Additionally, the turbulence is not so high until after the third row. This figure shows that the wake of the turbine is much broader than it appears in the velocity plot. Just by looking at velocity it might appear there are times when the MM is not in a turbine wake, but from the turbulence plot one would conclude that this is never the case for the directions modelled.

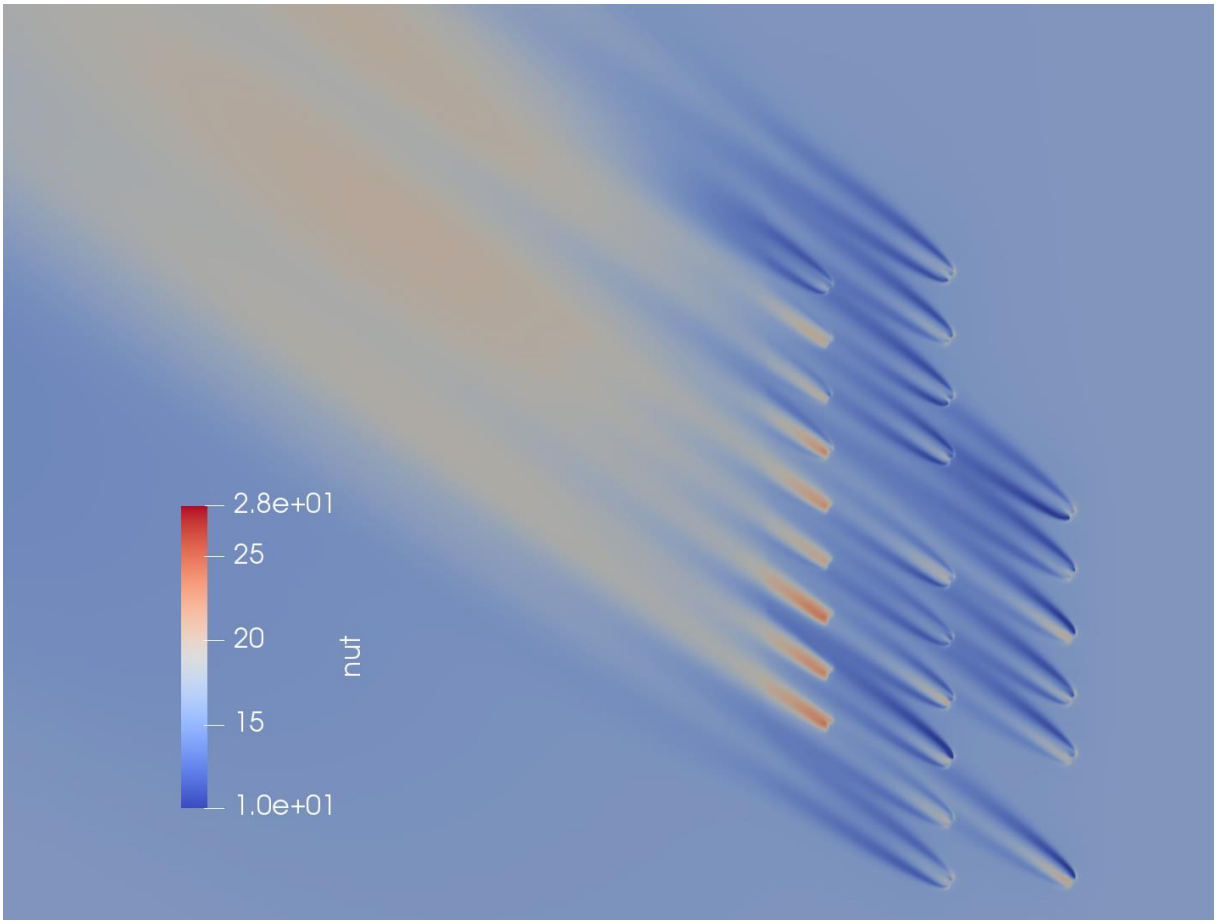


Figure 15. Turbulence viscosity for k-ε where the freestream direction is 35 degrees and the freestream velocity is 10 m/s.

Table 12 shows the metrics for comparison of the two turbulence models against the binning and regression approaches. This shows that for conducting all cases *k*-ε outperforms *k*-ω for how these models were set up, however the difference is not as significant as it might be for a more limited range of inflow directions.

3.7. Conclusions

In this study a computational fluid dynamics analysis of an existing wind farm is conducted using the actuator disk method. Not only are the cases conducted at multiple wind directions and velocities, but the results are also compared to real recorded data from the wind farm. The opensource code OpenFOAM is used and modifications to the code are presented so that the process followed is reproducible. Additionally, methods for processing the wind farm data are presented.

The two turbulence models of standard k-ε and k-ω SST were compared and used for all simulations. It was found that both perform generally well but for this model setup k-ε performs better, with a mean absolute error of 8% for all cases compared to 10% for k-ω SST when binning is used to compare to the met-mast data. In the case where the met-mast is in the direct wake of two turbines up stream, k-ε shows good agreement with the recorded data. However, at 35 degrees, when the met-mast is on the edge of a very close turbine, k-ε underpredicts wake deficit. Therefore, an overall error value calculated for one direction only quantifies the error for that direction and not others.

New methods for the use of SCADA data in both the construction and evaluation of the wind farm models are shown. Two approaches of approximating the SCADA data to more readily compare it to the CFD predictions are shown and compared. These approaches are binning the SCADA data and constructing regression models from the data. To compare the deterministic CFD results to the scattered measurement data, regression methods are used to approximate the MM data. The methods chosen are Gaussian Process Regression and Ensemble bagging, with mean absolute errors of 0.74 and 0.76 respectively. Both provide a suitable means of comparison however it is not determined which one is better.

The current practices for actuator disk method are well suited to modelling wake deficit for an entire wind farm at a range of inflow conditions. However, some care must be taken when inflow direction is varied as the error is not consistent.

In terms of potential future work, with the growth in computational capabilities, CFD can more be applied in a way that expands the compromise between capturing flow complexity and computational time. By also combining machine learning approaches, using the CFD results as training data, this opens the possibility to use CFD for evaluating the wind farm in ways that were previously not possible.

Chapter 4

4. A stochastic expansion of CFD results

4.1. Requirement for stochastic analysis

To continue its growth and expansion into other markets the cost needs to be reduced, particularly operational expense which can range from 13.1% to 56.5% of the levelized cost of the electricity [354]. Offshore wind turbines have grown significantly in size over the past few decades, while the average offshore wind turbine installed in 2017 was 5.9 MW there are currently concepts from manufacturers that will generate around 12 MW. Additionally, the size of the offshore wind farms are growing, with more turbines on new sites and expansions of existing sites. This brings new challenges, including in assessing the loads on turbines; but it also brings new opportunities to reduce costs through effective strategies. To aid in this area, new methods need to be developed to more effectively predict the loads on turbines not only at discrete points in time, but over the entire 25-30 year life span of a project.

A wide range of methods exist for modelling the wake of turbines within a wind farm, from simple engineering approximations to full computational fluid dynamics (CFD). The Navier-Stokes CFD solvers range from models where the turbine is treated simply as an actuator disk (AD) [150,169], to more computationally intensive models of fully resolved wind turbines [169,323]. The fully resolved methods can provide accurate answers and predict the flow to a relatively high accuracy but are limited to modelling only relatively short time intervals due to their computationally intensive nature [317]. For evaluating wind load variation within a wind farm, the actuator disk approach is a good compromise between engineering models and fully resolved rotors. The method is mature, at least 30 years old, with a variety of implementations. Studies have shown that actuator disk produces good agreement with measurements regarding wake deficits, [143,151] and that turbulence approaches, such as Reynolds Averaging, perform well at reproducing experimental results [355–357]. However, it is hard to cover the wide range of variables experienced by a wind farm through CFD; the wind speed can range from a turbine's cut-in speed of around 3-4 m/s to its cut-out speed at around 24-25 m/s and can flow from all directions. The pitching strategy employed results in a non-linear response and there are a number of parameters that must be accounted for as they affect turbulence intensity, yaw misalignment, blade rotational speeds, wind shear and temperature gradients, which creates a stochastic environment.

Monte-Carlo simulations can be used to account for this stochastic environment to generate a frequency-based probabilistic view [358][359]. However, these simulations require rapid modelling techniques such as the engineering wake models [182] but these are limited and cannot be used to evaluate time-variant phenomena and are inaccurate when modelling wake redirecting or meandering. Therefore, a more adaptable stochastic method is needed which can not only predict simple cases but also which can be adapted to transient and less conventional problems.

This chapter develops a method which combines CFD model results with approximation models in order to build a computationally efficient model of wind velocity values at turbine locations [15]. The purpose of this is that the model could then be used with stochastic inputs to build a probabilistic model of the wind farm. To accomplish this, several CFD cases are run across the range of variables of interest, information of interest is extracted from these cases and used to train supervised machine learning models. From the resulting machine learning models, any combination of the previously varied input variables can be explored, with accurate replication of the CFD model but with significantly lower computational cost; facilitating probabilistic methods. The challenge in applying the machine learning model is to build a suitable approximation model for interpolating between training data points, which are individual numerical values in this case. There are a wide range of potentially suitable models for this task which can be used in a regression or surrogate approach; five state-of-the-art methods are chosen for this study: Artificial Neural Networks (ANN), Gaussian Process (GP) with a Radial Basis Function kernel, Radial Basis Function (RBF), Random Forest (RF) and Support Vector Regression (SVR). These methods are compared in both their ability to replicate CFD data after training, as well as the overall accuracy in predicting the wind flow values within a reference wind farm. In addition, several different methods of generating the training data from the CFD data are also compared. This results in a novel method for stochastic analysis of wind turbine loading within a wind farm, as it is now possible to produce results which account for three-dimensional effects in a computationally rapid manner. This approach combines the ability to capture the physics of the problem to a suitable level using CFD analysis, combined with the computational efficiency of machine learning methods to be able to conduct a stochastic analysis.

In section 2, the reference offshore wind farm and the reference data available are presented. In section 3 the CFD model, previously verified in [16], is summarised. Section 4 presents the machine learning methods that are benchmarked. Results are presented in sections 5. A discussion of the findings is given in section 6 and finally, conclusions are given in section 7.

4.2. Description of the reference offshore wind farm

A series of CFD analyses are conducted for a discrete set of variables, in this case different wind directions and freestream velocities. The results from these analyses are then used to train machine learning models, from which predictions for other values can be generated. The trained approximation models can then be used to replicate the CFD, especially useful in computationally expensive simulation methods like Monte Carlo Simulations, Particle Swarm Optimisation and Genetic Algorithms.

A real offshore wind farm is simulated using actuator disks in a CFD domain based on weather data collected from the farm and the SCADA data collected at four turbine locations. The turbines have a hub height of 80 m. A meteorological mast (met mast) is permanently positioned on the edge of the site, shown on the reference map of the wind farm in Figure 1. The met mast has various sensors which can record wind speed and directions at heights of 25, 30, 58, 72 and 82 m. There are also sensors that record barometric pressure and temperature at 20 m as well as relative humidity and temperature at 82 m. These data are recorded as averages over 10-minute intervals, along with a time-stamp for each sample.

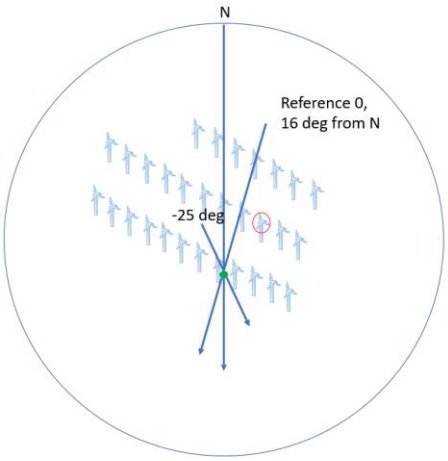


Figure 16. Map of the modelled wind farm, turbine 16 is circled in red, velocities at Turbine 16 are predicted in later sections. Blue arrow shows wind directions at 16 degrees and -25 degrees for reference.

4.3. CFD simulation of the wind farm

The purpose of the CFD model is to generate training data for the machine learning techniques, replicating the real turbines on the farm. The results of the predictions from the CFD-ML approach are compared to real data to evaluate the predictive capabilities.

The CFD model used is a steady-state, Reynolds Averaged Navier-Stokes (RANS) model in the opensource solver OpenFOAM using the SIMPLE solver algorithm. In this approach the three-dimensional wind farm domain is divided into roughly 60 million cells and discretised form of the Navier-Stokes equations are solved iteratively across the domain. The turbines are modelled using actuator-disks which impart a momentum loss, both axially and tangentially, to the fluid flow. The performance of each turbine is controlled individually within the solution process. The CFD results used in this chapter were from the k-ε turbulence model. The k-ε model averages the turbulence fluctuations and solves the quantities of k, turbulence kinetic energy, and ε, turbulence dissipation rate. Full details about the CFD model, how it is constructed and an accuracy comparison to real data, can be found at [16].

The wind farm is modelled at a range of incoming wind angles between 35° and -35°, shown in Table 13**,** and the freestream velocity is modelled between 5 m/s and 15 m/s in steps of 2.5 m/s; all combinations of these velocities and angles are modelled with the CFD analysis, resulting in 65 cases. These 65 cases comprise the full dataset. A reduced dataset of only 35 samples was also used to investigate the suitable number of samples. 13 additional points are also included at zero wind speed and power, to reduce the region of extrapolation for the models and ensure a good fit at lower wind speeds. There are 13 points as they are one for each wind direction.

The values for the incoming wind angle are chosen as the accuracy is compared with values from the met mast and this range includes several different wake behaviours: including the met

mast being directly behind two turbines and being partially in the wake of close neighbouring turbines. These differences should result in variability in the real values and in the model's ability to predict them.

Table 13. Independent variable values from which all CFD cases were based.

List of Directions (degrees)	-35	-27.5	-20	-15	-10	-5	0	5	10	15	20	27.5	35

The training samples from the CFD model are compared to the wind farm data samples to determine their accuracy by comparing predicted velocity values to those at the met mast, shown in Figure 2. The values compared, due to the availability of the data, are the mean velocity deficit value at 10 minute intervals. The mean absolute error is 0.85 m/s, or 8.94% relative to the comparison values. This accuracy should be kept in mind when evaluating the accuracy of the ML predictions. Conducting the CFD analysis is also computationally demanding, even for a steady state analysis. Each of these cases required around 22 hours using 32 CPU cores at a time. On this system, 128 cores can be used simultaneously, so 4 cases can be solved at once. This CFD-ML approach is then only feasibly if there is access to large computational capabilities.

4.4 Stochastic Inputs

As an additional study, a stochastic analysis was performed on turbine 16 to predict the probability of the power production being at the rated power of 3.6 MW. Turbine 16 was chosen because it is inside the wind farm, not at the freestream side, and SCADA data is present for this turbine. Based on the SCADA data, this turbine is at rated power 10.1 % of the time.

For comparison, 10000 predictions are made based on random sampling of the independent variables for the wind speed at turbine 16. The distributions from which samples were taken are a triangular distribution for direction and a Weibull distribution for freestream velocity. These distributions were fitted to filtered data using Palisade @RISK software [32]. The values for the distributions are given in Table 14. The assumption is made that if the velocity is at or above nominal velocity then the turbine is at rated power. The nominal velocity is between 13-14 m/s so 13.5 m/s was used. An axial induction factor of 0.3 is used to convert turbine velocity to upstream velocity.

Table 14. *Distributions used for the Monte-Carlo analysis*

	Minimum	*Most likely*	*Maximum*
Direction (Triangular distribution)	-25.104	296.8	368.04
	Shape α	*Scale β*	*Shift*
Freestream Velocity (Weibull distribution)	1.486	5.8776	1.4761

4.4. Machine learning methods

Several approximation methods are tested to evaluate their applicability as rapid surrogates for the CFD. The effectiveness of these predictions is compared with additional CFD test cases. The models compared are: Artificial Neural Networks (ANN), Gaussian Process (GP), Radial Basis Function (RBF), Random Forest Regression (RF) and Support Vector Regression (SVR). These models were selected because they are in common use within wind energy [227,348].

Due to the CFD cases requiring so much computational resource, only a few cases can be run and so only a small number of samples can be generated. The model is steady state and the site data available is mean velocity values at 10-minute intervals, and so each case relates to one value. The number of samples used to train the approximation method is an important consideration and so it should be determined how much data is required. While regression methods typically perform better with hundreds or even thousands of data samples [219], surrogate models can perform well with far fewer, but tend to overfit with too many sample points. For example, Queipo et al. [360] performed a case study using 54 designs to optimise 4 parameters using a surrogate model based on radial basis function (RBF). In another example using RBF [361], 200 samples were used to estimate soil parameters in a 400 km by 700 km area, finding this number of training samples suitable. For the response surface method, for a full factorial design of experiments, it is recommended to have a minimum of either 2^N samples or 3^N samples including center points, where N is the number of variables [362], this is more of an absolute minimum. Realistically, this depends on the situation and often 10^N is more suitable. The following study uses 65 training sample points to model the 2 variables.

4.4.1. Artificial Neural Networks

Artificial Neural Networks (ANN) were first coded by Rosenblatt as a ‘perceptron’ in 1957 [363], although the idea had been considered in physical models earlier than this [364]. Neural networks are widely used in wide range of applications including wind energy in both condition monitoring, [365,366], for damage detection [367–369], load forecasting [370], to infer loads from standard signals including yaw misalignment, generator speed and electrical power [371] and in power production and resource estimation [58,372–374]. Due to the power and versatility of ANNs, they have found wide spread use [227].

Multi-Layer Perceptron Neural Networks employ several layers of neurons: an input layer, one or more hidden layers and an output layer, as shown in Figure 17. The neurons have activation functions to relate the input they receive to the output they send to the next layer. In addition, there are weights between different neurons and biases which are trained iteratively, traditionally using reinforcement learning through back-propagation where the weights are adjusted between each neuron based on how accurate the prediction is to the required result. The model can be used in an approximation problem by using an appropriate output layer function and minimizing a cost function.

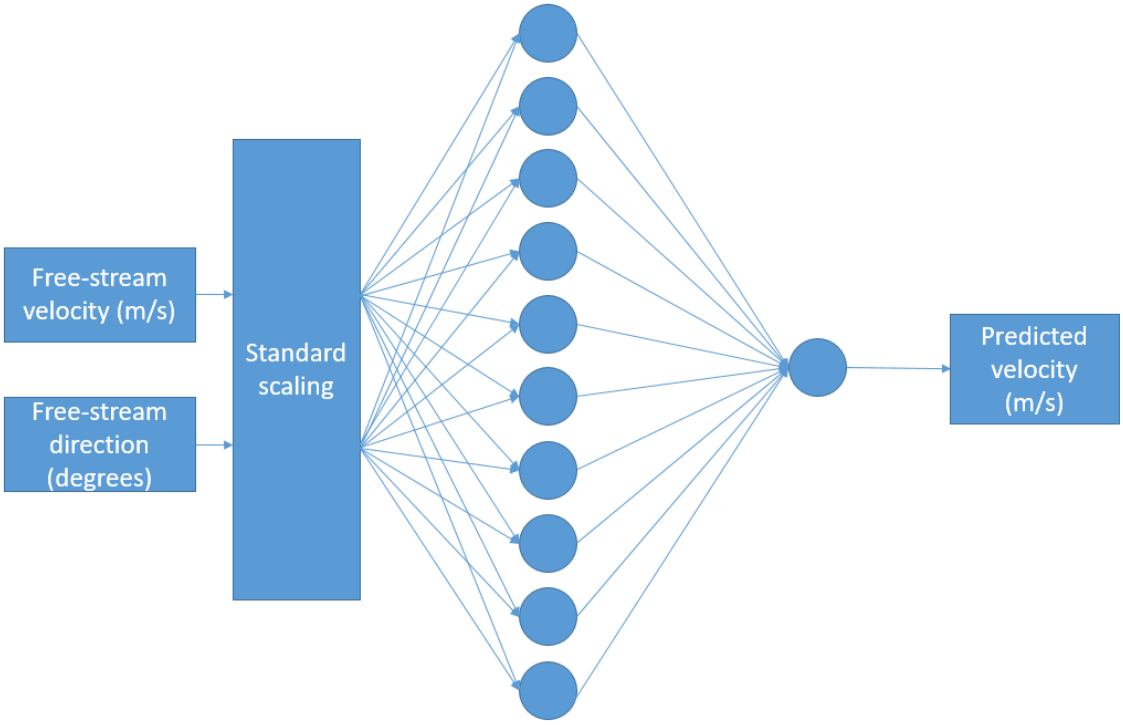


Figure 17. Architecture of the implemented ANN

Results were obtained using Keras with a TensorFlow backend. This model requires a large number of epochs to acheice suitable convergence, and hence a long training time. A minimum of 25,000 epochs are required with the best results at 60,000. After hyperparameter tuning the best results were obtained using: 10 nodes for the hidden layers, 'elu' activation for the hidden layer and 'linear' activation for the output node. Only one hidden layer was used as more were not necessary to capture the features. The system was trained using the 'adam' optimizer and the 'mean_square_error' loss parameter.

4.4.2. Gaussian Process Regression

Gaussian process regression, or kriging, is a stochastic process where a distribution is defined over a basis function which can take any form, in this case a squared exponent function. One of the earliest instances of a Gaussian process being used for curve fitting is from O'Hagan [375]. The methodology is as follows: after obtaining a predictive distribution, the regression is then applied over a basis function which projects the input onto the feature space, using a Bayesian approach to take a prior distribution and updating it to form a posterior distribution [17-19]. The output of the Gaussian Process is defined as the mean and covariance matrix, the covariance matrix defining the smoothness and can be represented by the kernel function. Given a set of independent variables, x, Gaussian process can be fully defined with a mean, $m(x)$, and a covariance, $k(x, x')$, as given in equation (28) [347];

$$\mathrm{F}(x) \sim GP(m(x), k(x, x')). \tag{28}$$

Gaussian Process has been used in wind energy for wind power forecasts [62,376–378] as well as determining wind turbine performance and efficiency [379] and condition monitoring through SCADA signal assessment [348,380–382] where the ability to determine confidence bounds in the prediction has been useful.

In the results presented below the algorithm is implemented in python using the Scikit learn library, *sklearn.gaussian_process.GaussianProcessRegression* using a radial basis function kernel [383]. This implements a GP model using algorithm 2.1 by Rasmussen and Williams [347]. A useful feature here is that it allows a prediction without requiring a fit for the prior. In this study, a squared exponential kernel is used which is parameterised by a length scale, l, which in this case was set at 0.1. This kernel is given in equation (29) [384];

$$k(x_i, x_j) = \exp\left(-\frac{1}{2} d\left(\frac{x_i}{l}, \frac{x_j}{l}\right)^2\right). \tag{29}$$

4.4.3. Radial Basis Function

A typical Radial Basis Function (RBF) model, originally derived by Broomhead and Lowe [385], is a form of feed-forward neural network composed of single neurons using radial basis function transfer functions [386–388]. The result of this approach is that a radial basis function fits a surface through the measured sample points. The values between the sample points are determined from functions based on the radial distance from the original point. The method has been used in soil strength parameter estimation [22, 23] and recently in wind energy to create a wind turbine controller which can deal with system nonlinearities [391] as well as in a LIDAR (Light Detection and Ranging) based controllers [392].

The equation for a multi-quadratic basis function is given in (30) [393],

$$\phi(r) = \sqrt{1 + (\varepsilon r)^2} \tag{30}$$

where ε is the shape parameter and r is the radius. The output from the entire model can then be given in equation 5, where the basis-function is here shown as a function of the Euclidian norm which is effectively the radius shown previously.

$$s(x) = \sum_{i=1}^{n} \lambda_i \phi(\|x - x_i\|) \tag{31}$$

where *s(x)* is the output of the model and λ_i is is the weight of the *i*-th node.

RBF nodes can be applied in a variety of models, however they are traditionally applied in a single-layer model with a single node for each sample point in the training set, the structure of which is shown diagrammatically in Figure 18.

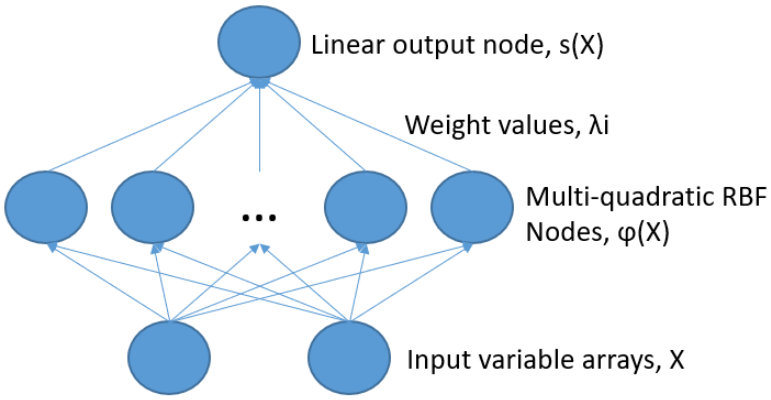


Figure 18. RBF schematic diagram

The code was developed in MATLAB to apply a multi-quadratic radial basis function. The code iteratively optimises the shape parameter, ε. As the RBF method forces the interpolation at the sample points, a least-squares method is not applicable and therefore LOO cross-validation is used to assess accuracy.

4.4.4. Random Forest Regression

This method is more recent and was first proposed by Ho [63] in 2001 and improved by Breiman [64]. Random forest regression is an ensemble method based on Decision Tree Analysis where the results of multiple Decision Tree models are averaged to produce a prediction. The key idea is that for multiple noisy, but mostly unbiased models, accuracy can be improved by reducing variance. The forest of trees is as accurate as the trees from which it is made, but the combination can reduce variance. To best accomplish this, trees must be generated in such a way as to minimise the correlation between trees while maintaining low variance of individual trees. The variance of the forest of *B* identically distributed trees can be shown in equation (32) [394],

$$\rho\sigma^2 + \frac{1-\rho}{B}\sigma^2, \qquad (32)$$

where σ^2 is the variance of the trees and ρ is the correlation between individual trees. To accomplish this low correlation and low variance of trees, the random forest process randomly selects variables, or sets of variables, to use to form the splits at nodes. The distinction between RF classification and RF regression is simply that for classification the trees cast votes for the prediction while in regression an average is taken.

The method is widely used in some software applications because it is fast and robust, for example it was used by Microsoft for pose recognition in their Xbox Connect [395]. However the

method has seen only limited use in wind energy; RF classifiers have been used to predict blade icing [396] and wind turbine stoppages [397] but so far have received limited attention in wind energy applications.

The algorithm is implemented in python using the Scikit learn library, *sklearn.ensemble.RandomForestRegression* [398]. Implementing this algorithm, the parameter changed is the number of trees, which in this case was set at 10,000. The criteria for the quality of split was mean-squared error. There are other parameters which can be set in order to avoid 'fully-grown' and 'unpruned' trees to save on memory, however with the size of the dataset this wasn't an issue and these parameters were set at their default as defined in the documentation.

4.4.5. Support Vector Regression

Support vector machines (SVM) can be used for regression and when doing so are called Support Vector Regression (SVR). SVMs, first identified by Vapnik [399], work through linear domain division where the division is made to be as large as possible. This can also be extended to higher order domains and be used for regression through the use of kernels [19, 20]. In this method, a support vector is drawn such that the error is minimized by selecting a hyperplane which maximises the margin. For linear SVR, an approximation can be derived from equation (33) [402];

$$y = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*) * \langle x_i, x\rangle + b, \quad (33)$$

where α_i and α_i^* are Lagrange multipliers, there are N training variables and b is a real number constant. For non-linear SVR, a kernel is applied to $\langle x_i, x\rangle$. For a Gaussian radial basis function kernel this becomes equation (34), where σ is a free parameter,

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right). \quad (34)$$

Support vector machines can be applied in either regression or classification form and have been used in power forecasting for wind energy [403–405], where it was found that accurate predictions can be made and that hybrid methods can improve predictions. SVM has also been used in fault-detection [406,407] with good performance, but also relatively quick training times compared to other methods tested due to their simplicity.

The algorithm is implemented in python using the Scikit learn library, *sklearn.svm.SVR* using a radial basis function kernel [408]. An RBF kernel was used in this case. The two tuneable parameters for this SVR implementation with RBF kernel are the penalty parameter, *C*, and epsilon-tube where no penalty is associated, ε, which were optimized for this data with values of 7 and 0.05 respectively.

4.5. Results

The predictions from the approximation models, trained to the CFD data, are compared against measured wind speed values from the real wind farm. A total number of 2016 data samples recorded from the met mast recorded data are taken and predicted by the CFD-approximation process. These samples are 10-minute average velocity values. The met mast values are measured at 82 m above sea level and the CFD results for comparison are extracted at this same coordinate in the model.

For each of the 2016 met mast recorded samples, there are corresponding free-stream conditions which were recorded at the same time; these free-stream conditions are used as independent variables in the approximation models to predict the values at the met mast. However, there is a slight mismatch between the free-stream conditions recorded at the time and the values recorded at the met mast as they are from different sources and matched by their timestamps. In addition, large scale turbulence makes the data stochastic, resulting in variations of the flow. Therefore, a second comparison is made between the already trained approximation model and the CFD results at the met mast for seven new free-stream conditions. These new CFD cases are modelled with free-stream conditions shown in Table 15 and were chosen with a strategy to position them roughly midway between free-stream conditions previously used to generate training data. The 'free-stream' values are measured at 80 m and the met mast values are measured at 82 m, however this difference in height is consistent for both the CFD model as well as the measured data and so should not lead to a difference in this comparison.

Table 15. Independent variables for the three CFD comparison cases

	1	2	3	4	5	6	7
Wind Direction (degrees)	0.00	12.50	12.50	7.00	18.75	6.25	18.75
Freestream Velocity (m/s)	8.25	10.00	8.25	6.50	11.75	13.50	6.50

4.5.1. Artificial Neural Networks

Figure 19 shows a *Y-Y* plot of recorded met mast velocities against predictions from the combined CFD-ANN approach, for 100 randomly selected cases to predict, with the metrics calculated for 2016 total number of samples. The figure shows points for when both the 35 training samples and 65 training samples are used as well as when a 0 free-stream velocity case is used to reduce extrapolation. Additionally, the mean absolute errors are shown in the legend. The 45° line represents the points where the x-axis values are equal to the y-axis values, that the error is 0.

Including the zero reduces the deviation from the real values at lower velocities slightly. This is reflected in the mean absolute error (MAE) values which are slightly lower for the 'with 0' cases than without. Although this changes the mean bias error (MBE), it's not consistently one way or the other. Using 65 training samples rather than 35 also improves the MAE, although this can't

be seen in the scatter plot because the difference is minimal. The CFD predictions are, erroneously, a little lower than the met mast values at 12.5 m/s. These MAE values are consistently the second lowest of the approximation methods, very closely behind SVR.

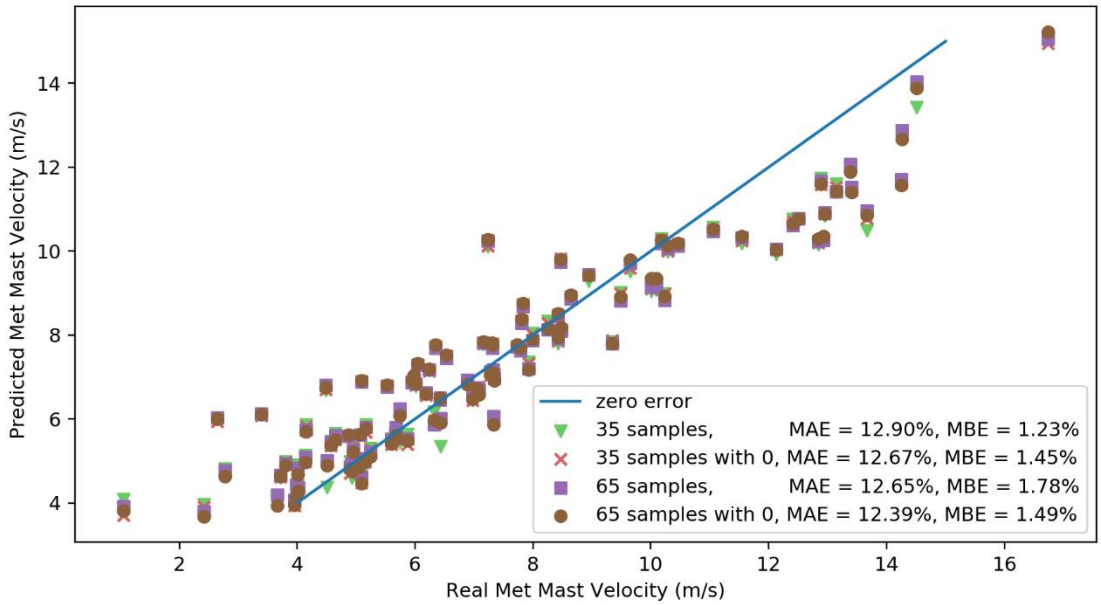


Figure 19. ANN, plotted 100 predicted cases, real velocity (m/s) against predicted velocity (m/s). The legend shows the training dataset used for the predictions as well as error metric calculated using 2016 total samples, including the mean absolute error, MAE, and mean bias error, MBE, achieved by that model.

4.5.2. Gaussian Process Regression

Figure 20 shows a *Y-Y* plot of recorded met mast velocities against predictions from the combined CFD-GP approach, for 100 randomly selected cases to predict, with the metrics calculated for 2016 total number of samples. The figure shows points for when both the 35 training samples and 65 training samples are used as well as when a 0 free-stream velocity case is used to reduce extrapolation. Additionally, the mean absolute errors are shown in the legend. The 45° line represents the points where the x-axis values are equal to the y-axis values, that the error is 0.

At the lower values, where the met mast value is less than 4 m/s, the predicted values are all too high, despite otherwise being close to the 45° line. When training sets which include 0 values are used, then this behaviour is less of an issue, however it is still present. This is reflected in the MBE values which reduce when training sets including the 0 values are used. Surrogate models predict values which pass through the training points and so they are effectively pinned at those points. If values being predicted are outside of the training set, then they are not bound by training values and are free to fluctuate. By adding the 0 values into the training set, this helps pin the surrogate model, extending its applicability.

This model makes predictions which pass through the training set; however, it is free to fluctuate away from the training data. Above 5 m/s the training data is spaced by 2.5 m/s intervals, however there is a 5 m/s gap to 0 m/s, this is apparently too large.

The decrease in MAE from the 35 samples training set to the 65 samples training set is the largest of any model. This is hard to tell from the scatter plot, there's nowhere that the larger training set results in significantly better results, so the improvement is throughout a lot of prediction points. The added 0 values make a large improvement in the MAE for reasons already discussed.

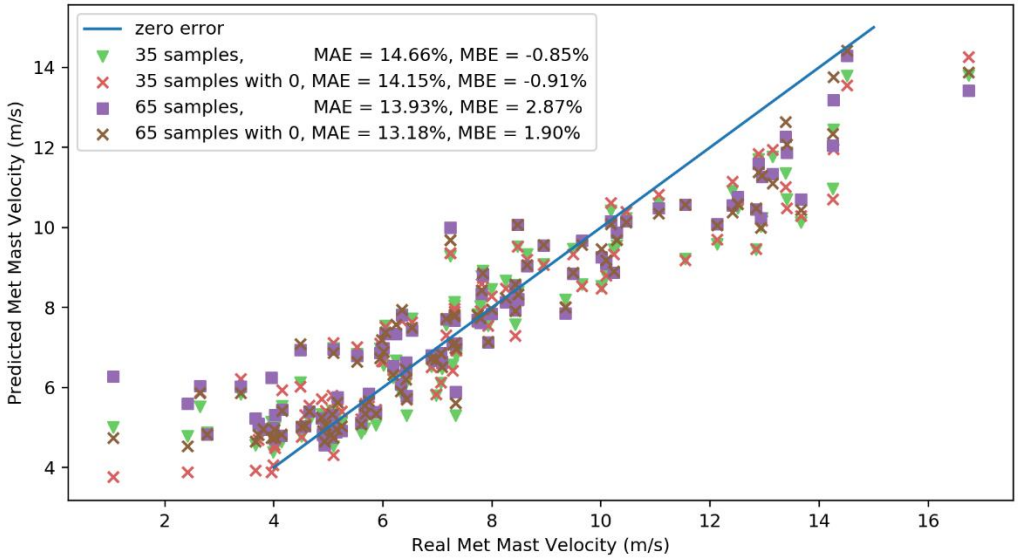


Figure 20. GP 100 plotted cases, real velocity (m/s) against predicted velocity (m/s). The legend shows the training dataset used for the predictions as well as error metric calculated using 2016 total samples, including the mean absolute error, MAE, and mean bias error, MBE, achieved by that model.

4.5.3. Radial Basis Function

Figure 21 Shows a *Y-Y* plot of recorded met mast velocities against predictions from the combined CFD-GP approach, for 100 randomly selected cases to predict, with the metrics calculated for 2016 total number of samples. The figure shows points for when both the 35 training samples and 65 training samples are used as well as when a 0 free-stream velocity case is used to reduce extrapolation. Additionally, the mean absolute errors are shown in the legend. The 45° line represents the points where the x-axis values are equal to the y-axis values, that the error is 0.

Although there are a few more points for the 35 training samples model which appear far away from the line, there isn't any significant difference in model accuracy with different training sets, the MAE is consistently close to 13%. As will be shown in section 4.5.7, there is a large difference in accuracy compared to further CFD cases when the RBF model is trained with more training data, which is not observed in comparison to the met mast. There appears to be a lower limit of MAE which is limited by the accuracy of the CFD to the met mast, rather than by the approximation method to the CFD, and that is just under 12%.

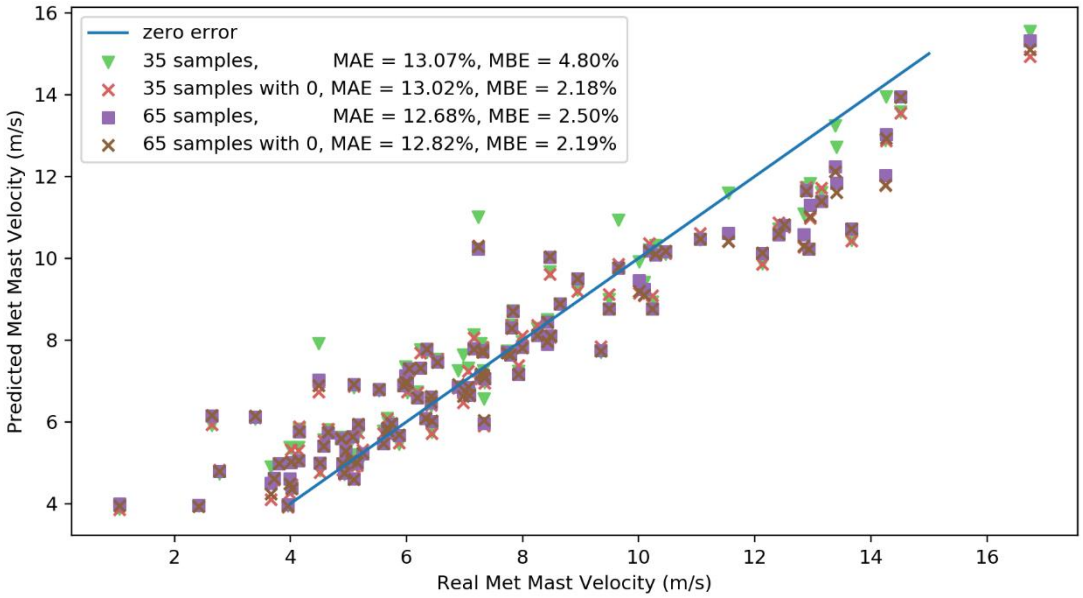


Figure 21. RBF 100 plotted cases, real velocity (m/s) against predicted velocity (m/s). The legend shows the training dataset used for the predictions as well as error metric calculated using 2016 total samples, including the mean absolute error, MAE, and mean bias error, MBE, achieved by that model.

4.5.4. Random Forest Regression

Figure 22 shows a *Y-Y* plot of recorded met mast velocities against predictions from the combined CFD-RF approach, for 100 randomly selected cases to predict, with the metrics calculated for 2016 total number of samples. The figure shows points for when both the 35 training samples and 65 training samples are used as well as when a 0 free-stream velocity case is used to reduce extrapolation. Additionally, the mean absolute errors are shown in the legend. The 45° line represents the points where the x-axis values are equal to the y-axis values, that the error is 0.

The MAE values for the RF model is high compared to other models, at around 14.7%, and this is due to the overfitting displayed as clusters are formed around the training data. Because of this overfitting, there is almost no difference between the main four training sets. By using the 'artificial samples' training set the MAE is reduced by nearly two percentage points and becomes in line with the other models. From the scatter plot the overfitting is resolved with the 'artificial samples' training set as the points now no longer form discrete clusters. The mean bias error of RF is the lowest of any of the models.

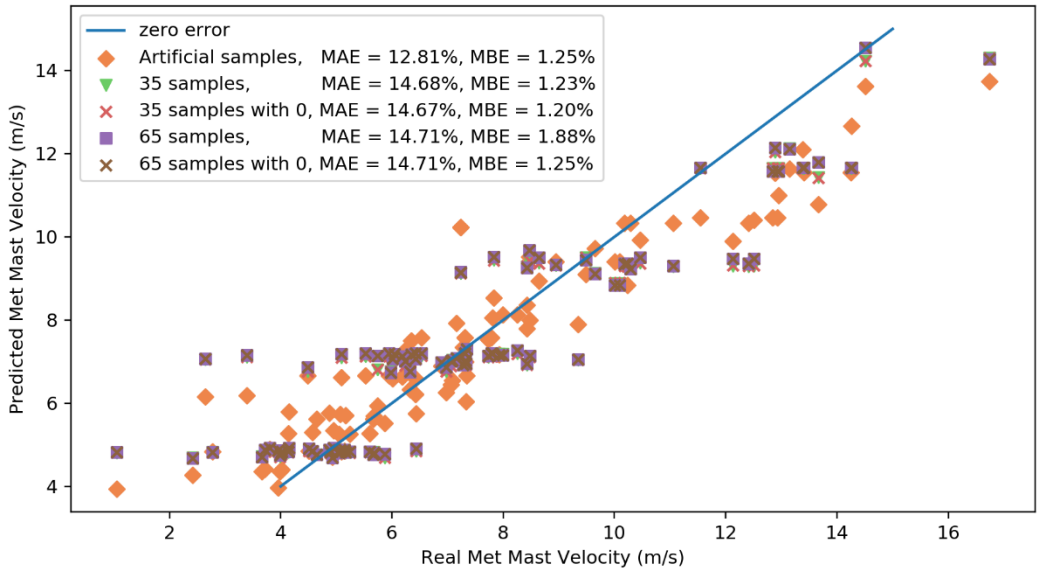


Figure 22. RF 100 plotted cases, real velocity (m/s) against predicted velocity (m/s). The legend shows the training dataset used for the predictions as well as error metric calculated using 2016 total samples, including the mean absolute error, MAE, and mean bias error, MBE, achieved by that model.

4.5.5. Support Vector Regression

Figure 23 shows a *Y-Y* plot of recorded met mast velocities against predictions from the combined CFD-SVR approach, for 100 randomly selected cases to predict, with the metrics calculated for 2016 total number of samples. The figure shows points for when both the 35 training samples and 65 training samples are used as well as when a 0 free-stream velocity case is used to reduce extrapolation. Additionally, the mean absolute errors are shown in the legend. The 45° line represents the points where the x-axis values are equal to the y-axis values, that the error is 0.

As can be seen from the MAE values, there's not much difference between the training data sets, except that the training data which included the 0 values are 0.5% points lower than without. The reason can be seen at the lower end of the plot where the points marked with an '*x*' are closer to the 45° line than the other points. Without the 0 points the model is extrapolating. Between the 35 training data points to 65 training points, there is only a very minor improvement in MAE. There's so little difference in the 35 training and 65 training models that the predictions are on top of each other and the green triangles are not visible.

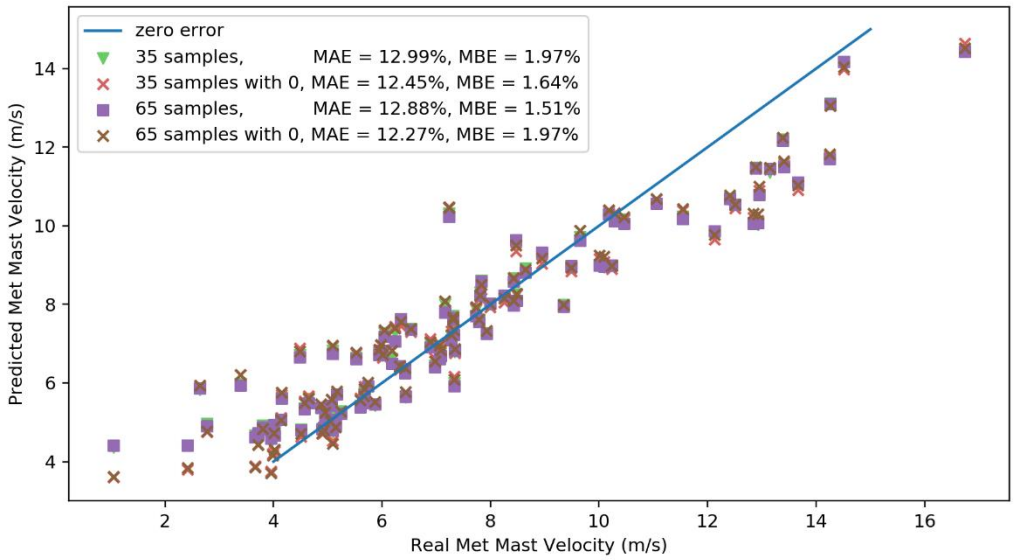


Figure 23. SVR 100 plotted cases, real velocity (m/s) against predicted velocity (m/s). The legend shows the training dataset used for the predictions as well as error metric calculated using 2016 total samples, including the mean absolute error, MAE, and mean bias error, MBE, achieved by that model.

4.5.6. Model results with direction

Figure 24 and Figure 25 show the model behaviour along the range of directions for a given free-stream velocity for the 35 training sample dataset and the 65 training sample dataset respectively. In the physical system being modelled, there is a relation between wake deficit at the met mast and the inflow direction because the wakes of the turbine follow wind direction and so the MM will move in and out of turbine wakes. This physical behaviour is captured in the CFD and hence the blue dots in the following figures, indicating CFD result values, are not all at the same value. The purpose of the machine learning approaches trained on the CFD results is to approximate this behaviour in values between the CFD values.

The predictions for Radial Basis Function and Gaussian Process, which uses a Radial Basis Function kernel, both pass through all the CFD sample points as they utilise these points within their prediction. Support vector regression complies loosely with the training sample points, although the prediction dips slightly around 0 degrees, it doesn't dip as much as low as the CFD prediction. The Artificial Neural Network model performs very well based on these figures, the predictions pass closely to the CFD training points and does not show any strange or excessive behaviour. Of the regression models, RF appears to perform the best; RF captures the general behaviour of the direction change relatively well however it does not reach the lowest point at the 0-degree direction. An advantage of the RF method is that it is unlikely to produce extreme, unrealistic results as it takes the average of decision tree predictions, while some of the other methods can produce extreme, unrealistic results. In that regard, RF is the most robust of the methods tested, but not necessarily the most accurate.

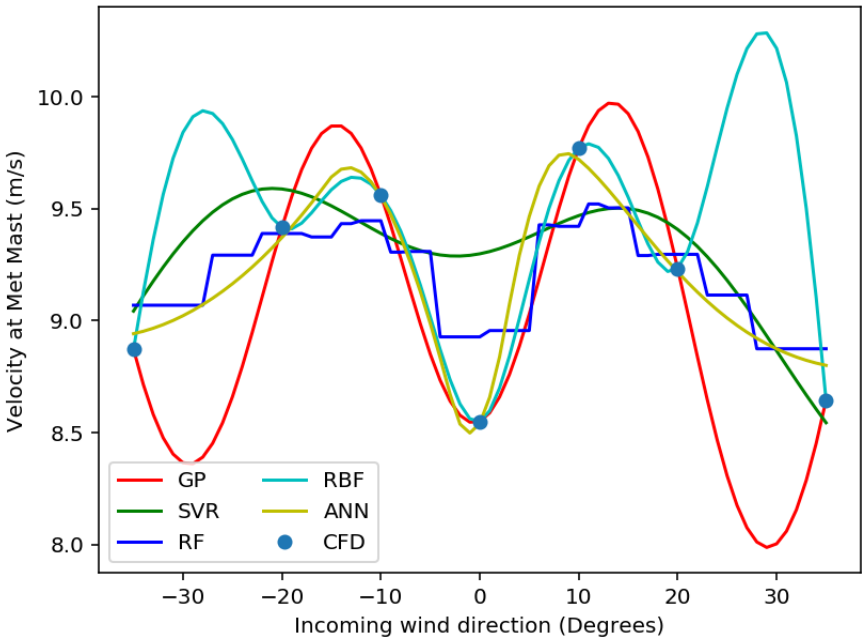


Figure 24. Reduced CFD dataset of 35 cases, *Model prediction for the met mast velocity at different free-stream velocities, k-ε, 15 m/s,*

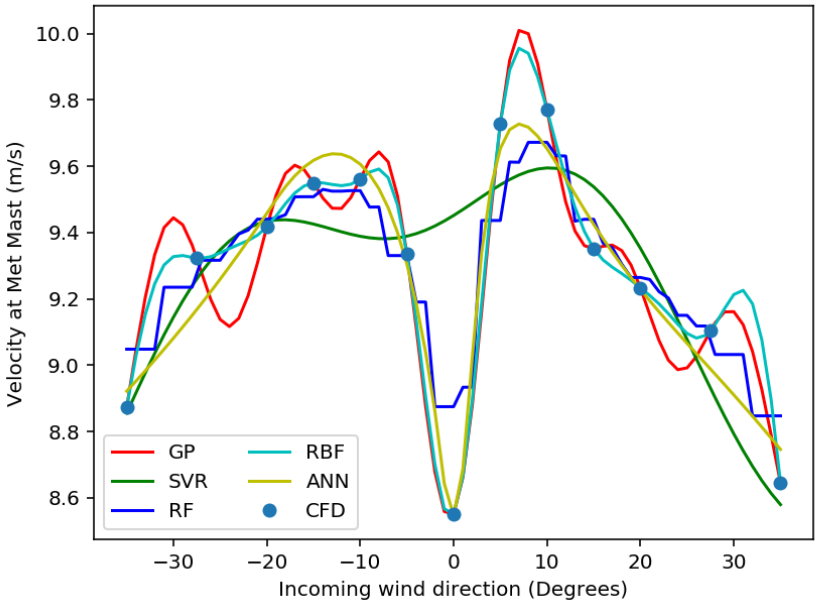


Figure 25. Full CFD dataset of 65 cases, plots of model prediction for met mast velocity against direction for a given free-stream velocity, k-ε, 15 m/s,

4.5.7. Comparison to CFD model

The following charts show a comparison of the predictions from the approximation models with the CFD model results. In this case, seven further CFD cases were conducted and the approximation models were used to predict the results in order to get an estimate for what level of error comes from the ML approximation of CFD. The MAE as a percentage of the mean CFD value are shown in Figure 26 and the maximum errors are shown in Figure 27.

Additionally, a sensitivity study was conducted where the number of training samples was increased from the minimum of 35 cases to the maximum of 65. These two additional steps were using 55 and 45 training samples. To arrive at 55 samples, the results at the direction of +/- 27.5º were removed and for 45 samples, results with direction +/- 15º were removed from that.

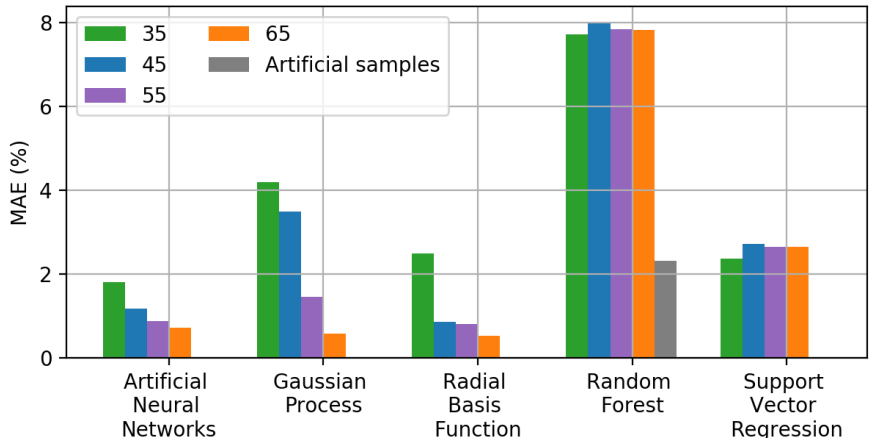


Figure 26. MAE as a percentage of the mean, for each approximation method compared to CFD model

The MAE values range from 0.54% for RBF with the 65-training point data set to 7.72% for RF with the 35-training points data set.

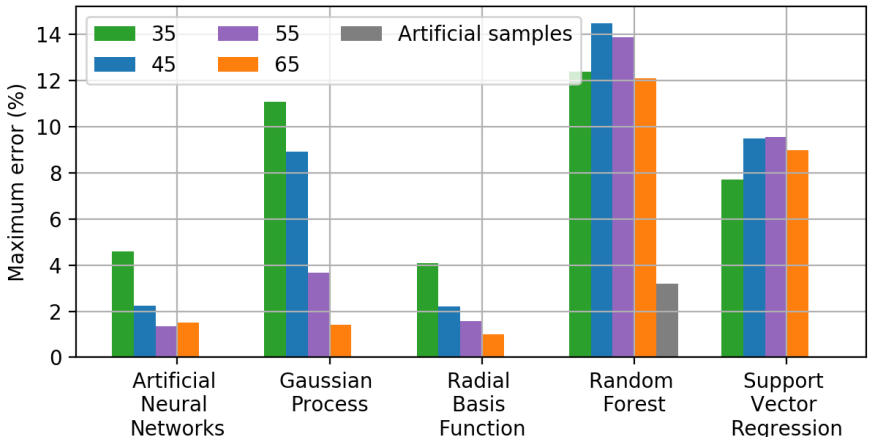


Figure 27. Maximum error percentage, for each approximation method compared to CFD model

The maximum error values range from 1.0% for RBF with the 65-training point data set to 12.4% for RF with the 35-training points data set.

In these results the maximum error is typically twice the MAE with the exceptions of the 'artificial samples' training set for RF where the maximum is slightly less, and for SVR where the maximum is roughly three times the MAE.

4.5.8. Comparison to established method Jensen

It can be useful to compare these results to other methods used in literature; however, this comparison can be challenging. In a wind tunnel experiment, analytical models can give very close results to reality, for example Garcia et. al. achieved a 2% error relative to a single turbine in wind tunnel tests [146]. For wind farm measurements, where parameters are far less controlled and there are more complications, the error can be much larger. For example, comparing wind farm efficiency using 4 analytical park models, Shao et. al found a MAE at Lillgrund of between 6.48% and 17.06% and at Horns Rev an error of between 5.81% and 10.73% [409]. Gaumond et. al. found that error level was dependant on an accurate knowledge of wind direction and found errors of 20.9%, 20.9%, 21.7% for Jensen, Larsen and Fuga respectively [410]. It can be difficult to compare these numbers directly, not only because they often are errors for different things but also because the error can be dependent on the wind farm itself as well as factors specific to the data recording and collection. Conducting the CFD for this study it was found that a large contribution of error was from the level of scatter in the data and the CFD model achieved an MAE of around 11% relative to measurement data [16].

For a direct comparison, the Jensen wake model was used to evaluate the same wind farm. This is an analytical model which evaluates the wake deficit and propagation as a function of the distance from the turbine. Wake interaction is accounted for by superposition of the wakes. [29,411] The same 2016 data samples from the met mast are evaluated and the error metrics achieved (Mean Absolute Error, MAE, and Mean Bias Error, MBE) in this are given in Figure 28. The bias from Jensen is slightly lower than the ML-CFD approach using ANN, around 0.7%, however the MAE is 4% higher at around 16%. It should be noted that the Jensen model predicts velocity at hub height, which is 80 m, and the met mast records values at 82 m.

The effectiveness of this approach can be visualised by plotting the 10-minute averaged met mast values sequentially as they were observed along with model predictions, forming a time-series plot, as in Figure 28. The model used in this case is ANN with 65 training samples and the added 0 points. The behaviours discussed earlier in the chapter can be observed in this plot. Typically, the predictions track closely with the real values. There is deviation between 10 and 13 m/s met mast velocities, as mentioned before this is from the CFD model under predicting in this area. There is also some over-prediction at very low values.

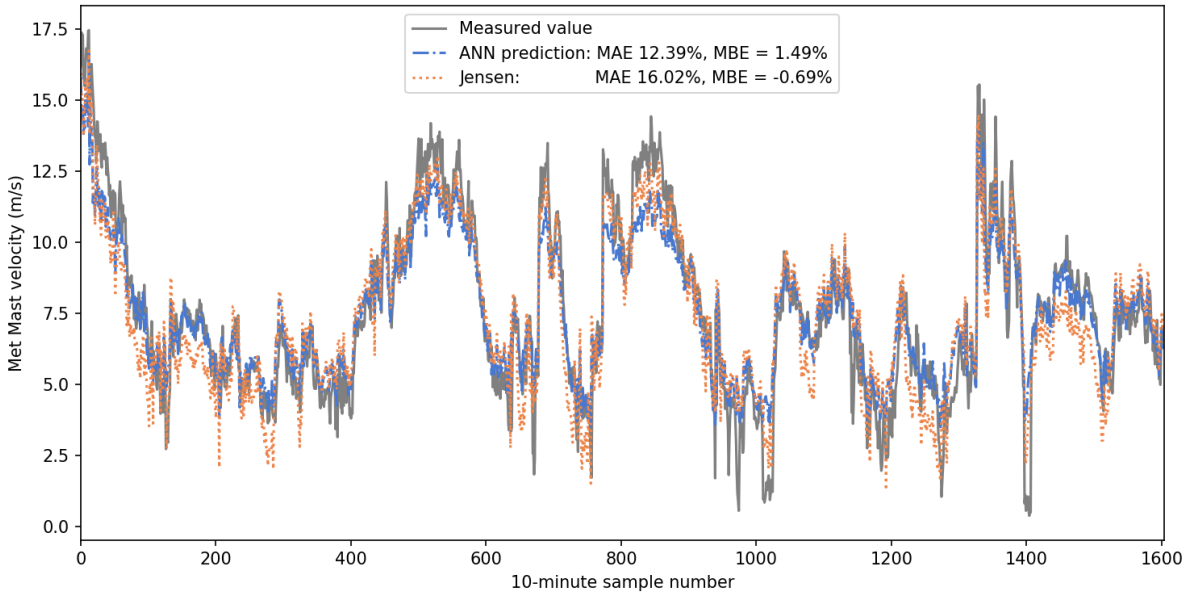


Figure 28. Observations plotted sequentially comparing the met mast values to the predictions from the best method, ANN with 65 training samples and the added 0 points, compared to Jensen wake model.

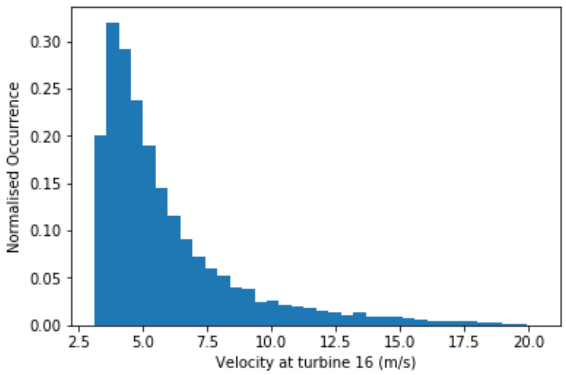
4.5.9 Stochastic Analysis

Each model was used to conduct a stochastic analysis to determine the percentage of time at which turbine 16 was at or above rated velocity. The purpose of this is to test the utility of the method in a stochastic analysis. The real value for time at rated power from the data for this particular turbine was 10.1% of the time. The method, outlined in section 2.4, was to randomly sample free-stream values from fitted distributions to produce independent variables for each approximation model which was fitted to the CFD results for the original data set. The results from this are presented in **Table 16** which shows the predicted percentage of time at rated power and also the difference from the actual percentage of time. The distributions from which the free-stream conditions were samples are presented in **Table 14**.

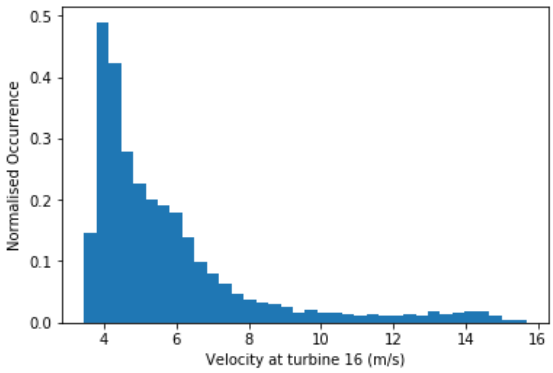
Table 16. *Predictions and errors for percentage of time turbine 16 is at rated power*

Method	ANN	GP	RBF	RF	SVR
Percentage at rated power	10.5	8.5	13.0	11.2	12.0
Difference from actual	+0.4	-1.6	+2.9	+1.1	+1.9

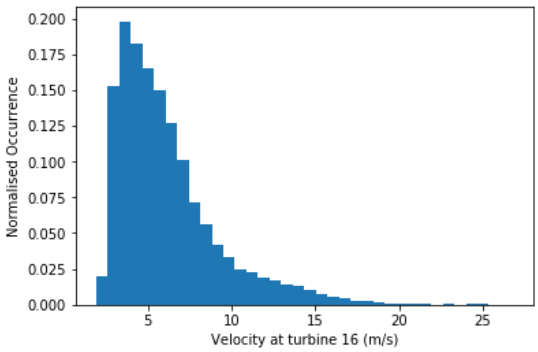
As can be seen in **Table 16**, ANN and RF produced the closest prediction with a difference in values of +0.4 and +1.1 respectively. SVR was the least accurate, over predicting the percentage of time at rated power by a difference of 1.9. The distributions are shown in **Figure 29**.


Normalised Occurrence
0.30
0.25
0.20
0.15
0.10
0.05
0.00
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Velocity at turbine 16 (m/s)

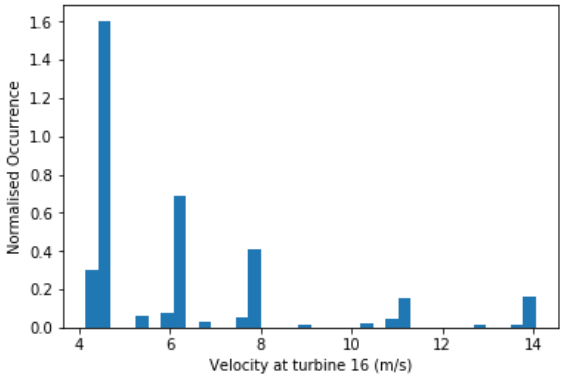
(a)


Normalised Occurrence
0.5
0.4
0.3
0.2
0.1
0.0
4
6
8
10
12
14
16
Velocity at turbine 16 (m/s)

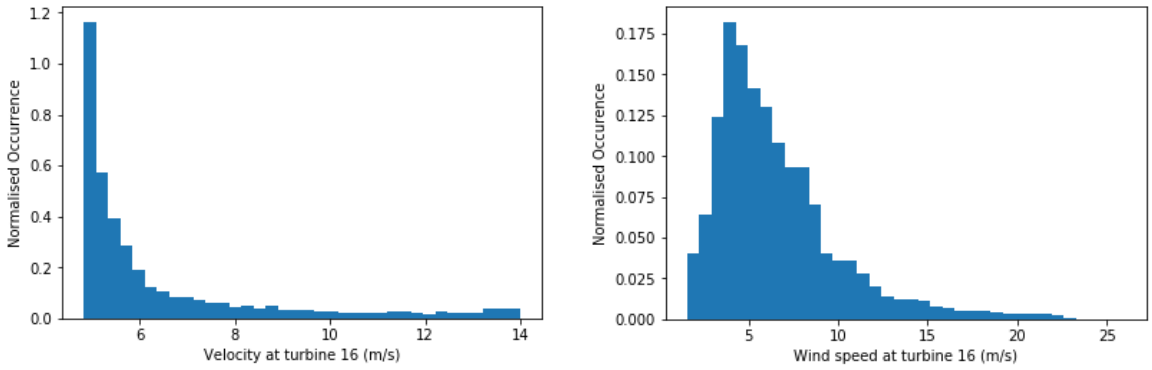
(b)


Normalised Occurrence
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
5
10
15
20
25
Velocity at turbine 16 (m/s)

(c)


Normalised Occurrence
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
4
6
8
10
12
14
Velocity at turbine 16 (m/s)

(d)



(e) (f)

Figure 29. *Probability distribution functions for the velocity at turbine 16 for (a) ANN, (b) GP, (c) RBF, (d) RF, (e) SVR, (f) Real data for comparison.*

The prediction from GP is the only model below the expected value, GP has a larger amount of dispersion in the predictions as seen in section 3.1.1. The other methods generally over predict time at rated power. This can partially be explained by large scale, atmospheric turbulence which is not accounted for in the CFD model.

Comparing the distributions produced from each of these models shown in **Figure 29** shows a clear distinction between the models. ANN, RBF and GP all produce distributions roughly similar in shape to the distribution from the turbine data shown in **Figure 29** (f). The most similar distribution is produced by RBF which has a head at the closest height, roughly 0.2 compared to the actual value of roughly 0.175, and a tail reaching a similar maximum of roughly 25 m/s. The two which are the least good are SVR, which has a higher skewness and a sharp drop just below the mean, which is itself below the actual value.

The RF model shows considerable overfitting, only producing discrete peaks and not a smooth curve. This behaviour was addressed by the method outlined in section 2.3.3 where artificial samples were produced and used to train the RF model. The result of this approach to the resulting distribution is shown in **Figure 30**. This distribution reaches a maximum normalised occurance only slightly higher than the RBF model and produces a similair general shape to the actual distribution. The main discrepancy is the large value of ‘Velocity at turbine 16’ at 14 m/s, which occurs because the model was not trained on any values above this.

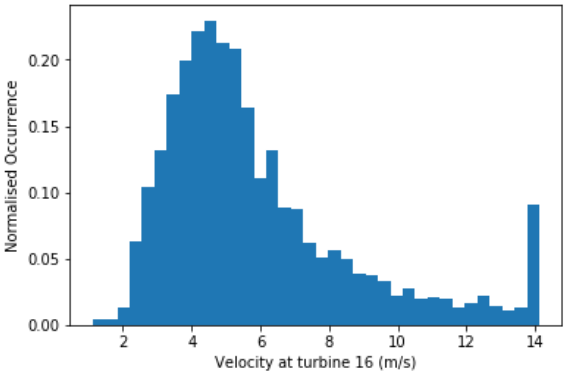


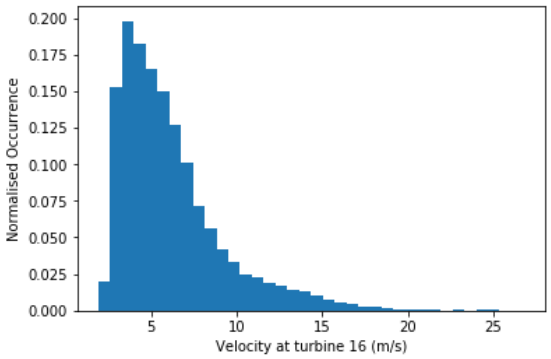
Figure 30. *Random Forest with artificial samples*

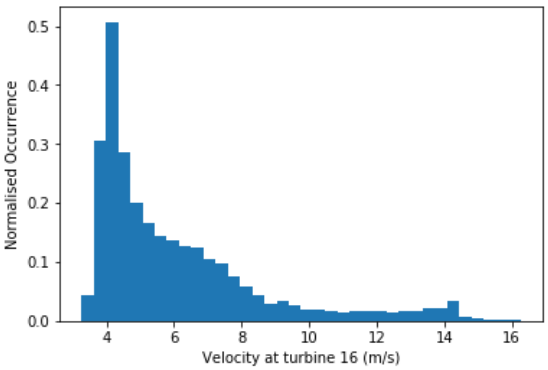
The models trained using the expanded data set were then used to conduct a Monte-Carlo analysis and produce joint probability distributions. However, as discussed, the increase in training data does not necessarily improve the accuracy of the predictions since many of the real data points are lower than the model predictions, possibly due to large scale turbulence. Similarly, some of the predictions for time at rated power have gotten worse, but others have improved. The values for this are shown in **Table 17**.

Table 17. *Stochastic prediction for each model trained using the expanded data set*

Method	**ANN**	**GP**	**RBF**	**RF**	**SVR**
Percentage at rated power	13.0	10.1	11.7	11.4	11.8
Difference from actual	+2.9	0	+1.6	+1.3	+1.7

The prediction for GP is the best in this case, now with a difference from the actual value of 0, although this is likely due to chance as the model is not including noise at each data point, although the model can do this. The only other two improvements are RBF, going from +2.9 difference to +1.6, and SVR, going from +1.9 to +1.7.





(a)

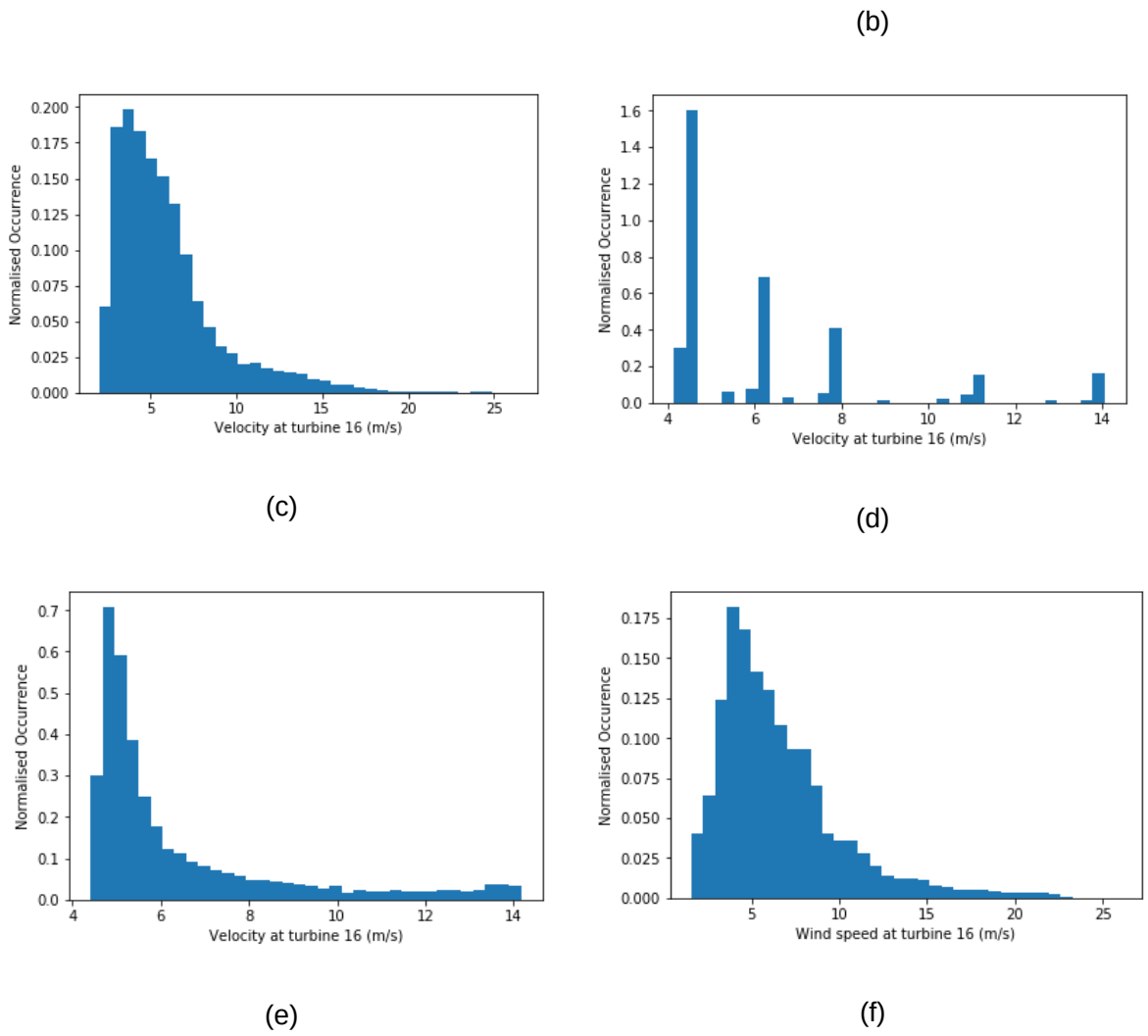


Figure 31. *Probability distribution functions for the velocity at turbine 16 for (a) ANN, (b) GP, (c) RBF, (d) RF, (e) SVR, (f) Real data for comparison.*

Figure 31 (a)-(e) show the probability distributions produced by the models and (f) shows again the actual distribution from the data. There is a slight improvement in the distributions, particularly in ANN, GP and SVR where the normalised occurrence of the mode has come down and the general shape is closer to (f) in those cases. The distribution from ANN is now very similar to the distribution from RBF. The distribution from RF is almost identical to how it was with the original data set.

4.6. Discussion

The approximation models were trained on CFD generated training data and then used to make predictions which were compared to both further CFD cases as well as against met mast recordings. Against further CFD cases, all methods were able to perform well, with a MAE under 3% for the best training data. However, the methods which clearly performed best against further CFD cases were ANN, RBF and GP as their MAE were all below 1% with the largest training set. In terms of approximating velocity deficit at a given point, these approximation methods can

successfully provide similar results as conducting more CFD cases. It should be noted that this approach models the input/output relationship between free-stream values and measured values within the windfarm wake. The approach predicts are individual 10-minute values. If forecasts were to be made, then an appropriate method for statistically sampling free-stream values would need to be used which includes seasonality and cycles.

The goal of the combined CFD-ML approach is to predict the physical system and the accuracy to which this is accomplished is measured by the comparison to met mast results. The CFD model predicts a range of parameters including velocity deficit at various locations as well as turbulence. The ML portion of the approach can then approximate these results to sets of independent variables not used in the CFD models, in this case velocity deficit. The accuracy is compared to measured values recorded on the real wind farm as well as accuracy of the ML models relative to the CFD model. In this study, the lowest MAE was achieved with SVR using the 65 CFD training sample data set with the additional 0 values, with a MAE of only 11.8%. This is despite still having a relatively large error compared to further CFD cases of 2.65% for the 65-point training set. This is because SVR case provides a response which doesn't change as much with direction, while the CFD, based on this, changes too much with direction. It appears from this that the stochastic nature of the real measurement data smears out sharp changes in the result, and so SVR, which approximates without the sudden dip at 0 degrees, achieves a lower error than the other methods against measurement data. Applying this in the future however, it would be good practice to either reduce the error from the CFD or to deliberately account for the stochastic nature of the real data. Otherwise, the best method after SVR is ANN with a lowest MAE of 12.39%. This agrees with the other results, that it is one of the best approximations of the CFD predictions.

Using the larger training set improves the results for the comparison to further CFD cases with all models except for SVR which becomes slightly worse. The improvement from training on the larger data is substantial for comparison to further CFD, reducing MAE to one seventh compared to the smaller training set in one case. The improvement compared to the met mast measurements is not so significant and doesn't even reduce for some models. This is because the error of the approximation models to the CFD is much less than the error of the CFD to the met mast.

Adding the zero values improved the accuracy of almost all cases and can therefore be recommended to be included in future such studies. This provides a noticeable improvement at no computational cost.

The use of artificial samples, effectively using a polynomial regression for one variable before then using the RF, had a big impact in all error metrics. With this training set, the RF model performed as well as the other models. The reason for this is clear in the plot of the randomly chosen cases as the points no longer form discrete clusters due to overfitting.

From a practical standpoint, some of the methods approximate the CFD results to a small error of only around 1% maximum error, but the error to the met mast is much larger. With the use of these approximation methods, similar results to conducting more CFD cases can be obtained at vastly lower computational cost. The overall computational cost of this approach is still high due

to requiring so many CFD cases as training data, which in this case required roughly a month on a high-performance computer cluster.

This ML-CFD approach can be useful in cases when measured data isn't available but when a layout is already established. It is more computationally intensive than the Jensen wake model but much more accurate and inherently accounts for physical behaviour of the wake. Therefore, it is useful in cases such as: determining values at points which aren't measured, evaluating different wind farm control strategies or predicting future wind farm performance.

4.7. Conclusions

In this study, a method is developed for conducting stochastic analysis of an offshore wind farm. This method takes results from CFD analyses at a set of input values, builds an approximation model of the system from the CFD data and produces predictions between the modelled inputs. CFD cases were conducted at a range of free-stream speeds and directions in order to build a suitable training set based on the 65 CFD cases. Velocity deficit at the met mast was predicted by the trained models and was compared to measurement data from the real met mast.

Five Approximation methods were compared: Artificial Neural Networks, Gaussian Process with a Radial Basis Function kernel, Radial Basis Function and Support Vector Regression. These are state-of-the-art methods which all work in very different ways. The suitability and accuracy of each for this task is assessed.

On balance, both Artificial Neural Network and Radial Basis Function were the best method with the most accurate prediction of CFD results, achieving a mean absolute error of 0.73% and 0.54% respectively. Increasing the number of data samples for training had a clear improvement on the replication of CFD results by the approximation methods with the surrogate models, Gaussian Process and Radial Basis Function, showing the largest improvements in replicating the CFD results with an increase in the training data set size. For support vector regression however, there was no improvement with the larger training set.

In the case of comparing against the measured site data, the best method was SVR with a MAE of 12.27%, followed by ANN and RBF which had MAEs of 12.39% and 12.68% respectively. Good practice in future work would be to use either ANN or RBF, which predict CFD results well, and then account for the stochastic nature of the real measurements.

Future work could build on this by incorporating the stochastic capabilities of some approximation models in order to replicate the stochastic nature of the wakes. Additionally, other variables which can be calculated could be approximated such as turbulence values, or a time-series prediction can be generated, which can be useful for estimating fatigue.

Chapter 5

5. Wind and wave time series forecasting

5.1. Literature Review

5.1.1. Markov Chains

Markov chains (MC) is a widely accepted statistical tool [412] for modelling a variety of natural phenomena, including wind speed values. Authors have used MCs for both on and offshore wind [69,413] and discussed the uses in reliability forecasting [414], [415]. There are some limitations to using Markov Chains for wind speed prediction; it has been found that the autocorrelation plots for short intervals are often inaccurate due to lacking the persistence found in real data. As a result, they do not perform well in ranges from 15 to 40 minutes, it was found that the synthetic wind data fluctuates more rapidly in first order MCs than real data for these time scales [416]. However, authors have devised modifications for Markov chains which improve upon this aspect and others. For example, the aforementioned limitation was addressed by incorporating a second lag as well as a running-average filter [417]. Other improvements include nested Markov Chains which are better able to capture the temporal self-dependence in wind speed [413]. More improvements exist, such as accounting for the uncertainty in the transition matrix through Bayesian inference, which in a case study was found to outperform traditional Markov Chains in a credible interval criterion. Another interesting idea is to first cluster the wind data as a method of discretisation. This was done on two dimensional anemometer data and found to show 'characteristic' wind behaviour of the site [71].

An important aspect for the present chapter is seasonality. There is seasonality in the real wind data and this needs to be represented in the model. One author trained separate models for summer and winter in order to incorporate seasonality [415]. Another author, with only one year of training data, trained a separate Markov Chain for each month. They not only found that this captured the seasonality present in the wind speed data but also found that one month of data per model was sufficient to capture the statistical properties [70].

5.1.2. Gradient boosting and trees

Gradient boosting is often used in stochastic value prediction [418]. The method has been used to win two Global Energy Forecasting competitions [419,420] where it was required not only to forecast into the future but also to back-cast missing data. The method works well at forecasting at different time scales – some authors have used it to forecast 1 to 6 hours ahead [421], others up to one day [66], and others have used it for even longer time scales [67]; all found it to perform very well compared to other methods tested in those studies. Boosted regression trees have also been used for prediction of maximum wind speed based on geographical considerations, though this is not forecasting, it shows the diversity of the approach [422].

Some authors have found ways to improve upon the gradient based boosting method for certain problems. For example, by incorporating instance-based transfer learning, Cai et. al. were able to incorporate related data sources into their training and improve against a benchmark approach [423]. Chen et. al. improved on Gradient boosting by incorporating Markov Chain mixing rate to derive upper bounds in the loss function, which improves convergence of the problem [65].

5.1.3. Probabilistic methods

For estimating sea-states and wind speeds, statistical distributions are normally used, and this practice is also recommended in engineering standards. For example, prediction of annual energy production is recommended by International standards IEC 61400-12 to be conducted with a two-parameter Weibull distribution [424]. Similarly, the American Bureau of shipping recommends a two-parameter Weibull distribution for mean wind speed estimation, but finds that sometimes a Rayleigh distribution is appropriate [38]. Researchers have used a range of distributions to model both significant wave height and mean wind speed, particularly for investigations at the design phase. Authors have used Weibull [47,48], mentioned before, as well as Gumbel [49,50], Lognormal [45,46] and many other distributions.

Seasonality has been incorporated into the use of statistical distributions for wind speed. Jaramillo et. al. fitted a distribution for winter, spring, autumn and fall for a site in Mexico [37]. This is the same approach used by a number of other authors who also characterized wind speed at sites using Weibull distributions [425,426]. A challenge in addressing seasonality was discusses by Erikson et. al. who found that at different times of the year, much of the wind speed can exhibit non-Weibull behaviour [427]. To address this, Drobinski et. el. proposed an alternative which fits both the diurnal and seasonal variability better [428].

5.2. Weather forecasting methods

5.2.1. Markov chains

A Markov Chain is a stochastic process which defines a set of probabilities for the next possible set of states, given the current state. The set of probabilities for moving from one state to the next are called transition probabilities and are dependant only on the state being moved from. In a regression problem, each possible discretised numeric value can be defined as a state. In Markov chains, the probability of any state at time t_n for a state, i, within a countable set of states, S, is independent of all previous states except for the last one, as shown [429]:

$$\Pr\{X(t_n) = i_n | X(t_1) = i_1, \dots X(t_{n-1}) = i_{n-1}\} = \Pr\{X(t_n) = i_n | X(t_{n-1}) = i_{n-1}\} \quad (35)$$

Forecasting into the future is accomplished as a series of steps and any state in the future is dependent on all of the probabilities in between those states. With a given start time i, the probability of a certain state at time j which is r steps from the value, is given in the Chapman-Kolmogorov equation [412]:

$$P_{ij} = \sum_{k=1}^{r} P_{ik} P_{kj} \quad (36)$$

Markov chains and how they work are described in detail in references [412,429,430].

The Markov Chain model used in this chapter is a first-order, observation driven model which generates a probability vector of wave height given the previous (t_{n-1}) wave height, with a separate matrix for each month. A probability matrix of wind speed, given wave height, is constructed as well as a matrix of wind direction given wave height. These are constructed by counting the number of occurrences within each matrix position and then normalising them by the total number. To determine each next iteration, the model samples from the probability matrixes given the current state.

5.2.2. Gradient boosting

In this chapter gradient boosting is applied using LightGBM which is an API for Python [230]. Gradient boosting is an ensemble of weak learners where new learners are added sequentially in such a way as to minimize the gradients of the loss function. At each iteration, a new model is added which aims to correct the error of the previous model. Each new term fits to the residual of the previous model, R=$y - F_m(x)$. By identifying this as the gradient of the squared error loss function, the method can be generalized with other loss functions. Each new iterated model, $F_m(x)$, is defined from the previous model, $F_{m-1}(x)$, in the following equations [431]:

$$F_m(x) = F_{m-1}(x) + \rho_m h(x; a_m) \quad (37)$$

Where $h(x; a_m)$ is the constrained negative gradient. The 'line search', ρ_m, along the direction of the gradient is defined as:

$$\rho_m = \arg min_\rho \sum_{i=1}^{N} L(y_i, F_{m-1}(x_i) + \rho h(x_i; a_m)) \quad (38)$$

$L()$ is the loss function between the value y_i and the previous model plus the new term.

In applying this method, the time variable, *t*, needs to be converted to lag variables. A starting point is variables of 'month', 'day' and 'hour'. However, this was found to not be enough to capture the time-series behaviour. In the same concept, generic lag variables were also created where a repeated duration is subdivided into a set of intervals. These durations were chosen from the optimized time-lengths found through the hybrid statistical method discussed in the next section.

There are many parameters which can be set when implementing a LightGBM model [231]. A Gradient Boosting Decision Tree is used for training with the loss metric of mean absolute percentage error.

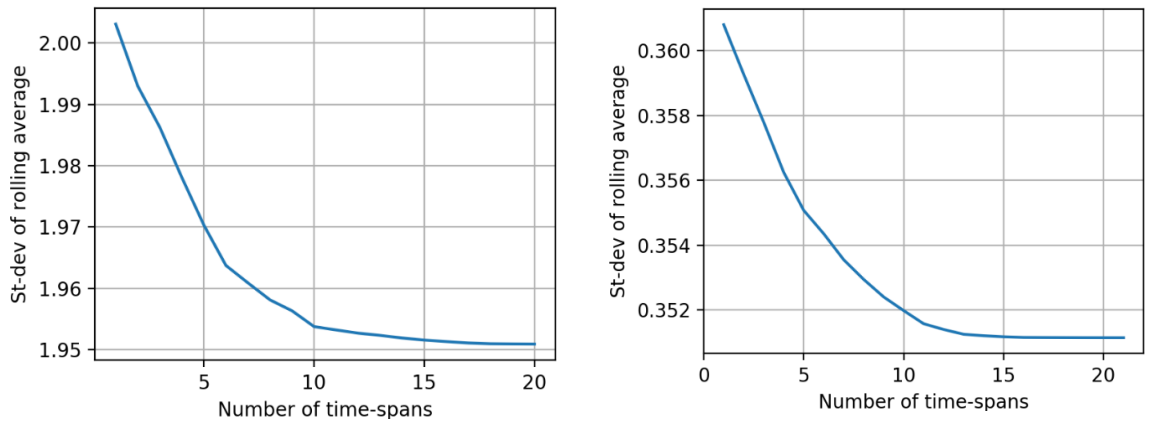
5.2.3. Hybrid probabilistic method

The key premise to the proposed concept, in terms of time-series forecasting, is that any time-series is composed of three components: trend, seasonality and randomness. Trend is the gradual change over time, seasonality is repeated patterns which are expected to continue and randomness is behaviour which cannot be fully explained based on past patterns. To an extent, trend and seasonality can be removed from a time-series data by various approaches, leaving only the randomness of the data. This randomness can be parameterised and replicated through distribution fitting and sampling. If one is interested in predicting accurate values for any given observation, then this is not enough. However, when the objective is to simulate, rather than predict, then this can be appropriate, as this chapter will show.

The hybrid statistical/ regression method proposed in this chapter uses a combination of sampling from statistical distributions with regression of parameters. The distribution parameters are fitted to the time series data through repeated polynomial regression. In the case of a univariate model these parameters are the mean and standard deviation as they vary with time, and are used to fit a normal distribution. In the case of a multivariate model this is the mean and the covariance between the training data arrays, and this is used to fit a multivariate normal distribution. Both of these distributions are widely used in literature [432]. For wind speed prediction, authors tend not to use normal distributions opting for others [31,32,34]. However, with the seasonality removed, the data is much closer to a normal distribution than before and the small number of parameters in a normal distribution is appealing for applying this method.

The seasonality and trend are first removed from the data by iterative fitting of polynomial regression curves where the curve is subtracted from the training data before the next iteration is conducted. The first curve is a first order polynomial to remove trend and the following sets of polynomials are all fourth order as this appears to fit the trends well. A least-squares based fitting approach is used.

The regression fitted to seasonality should cover each time span for which there is a significant repeated pattern. This may initially be taken as each year, each month, each day and so on, but this is insufficient. The process for selecting the seasonality time spans is automated in this approach. At each iteration, a curve for all time lengths below a set value is fitted to, and then subtracted from, the data. The time span leading to the lowest standard deviation in rolling average for the new, de-trended data is then used and another iteration is conducted. The new maximum time length for the next iteration is a value just under the previous selected time length. The reduction in the error metric is shown in Figure 32. These parameters can be visualised over the measurement values, as in Figure 33.



(a) *(b)*

Figure 32. Reduction in error metric, standard deviation of rolling mean, for the de-seasoned data at each iteration. (a) u10, (b) Hs

A curve for the standard deviation (for a univariate model) or the covariance (for a multivariate model), is then fitted to the de-trended data using only an annual time scale. These fitted curves can then be extended into the future. This whole process is shown diagrammatically in Figure 34.

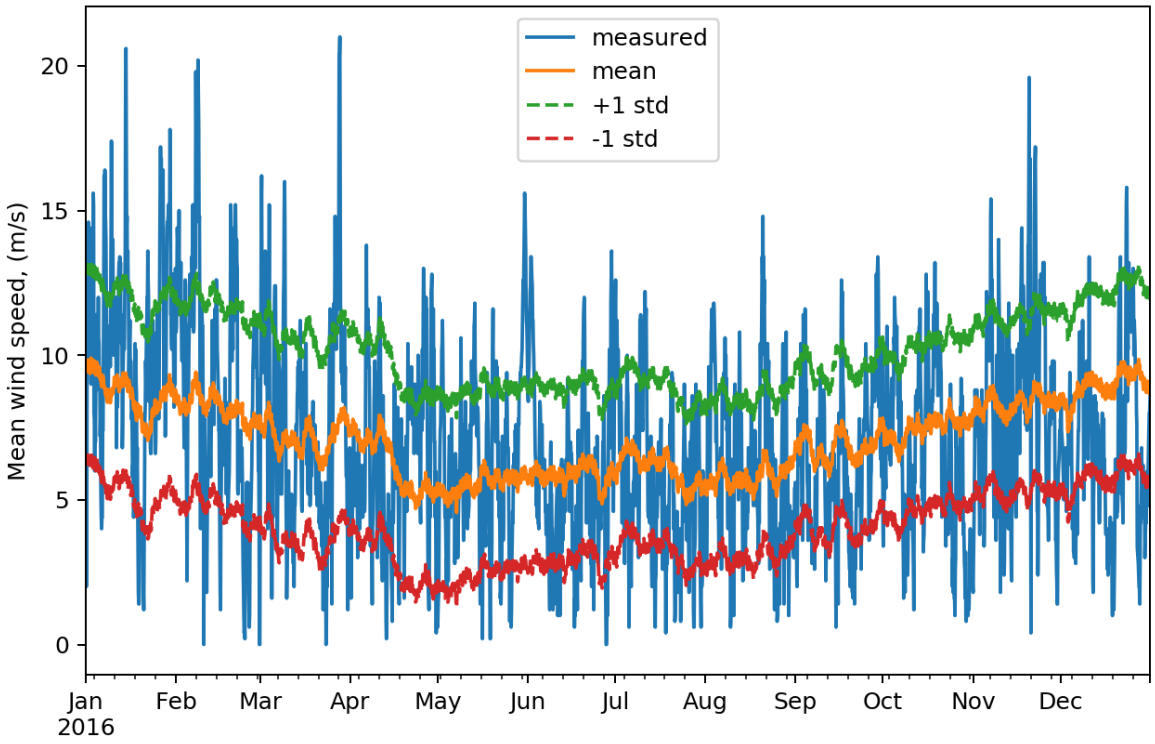


Figure 33. Mean curve as well as the mean +/-1 standard deviation over the measured values for wind speed for the year 2016.

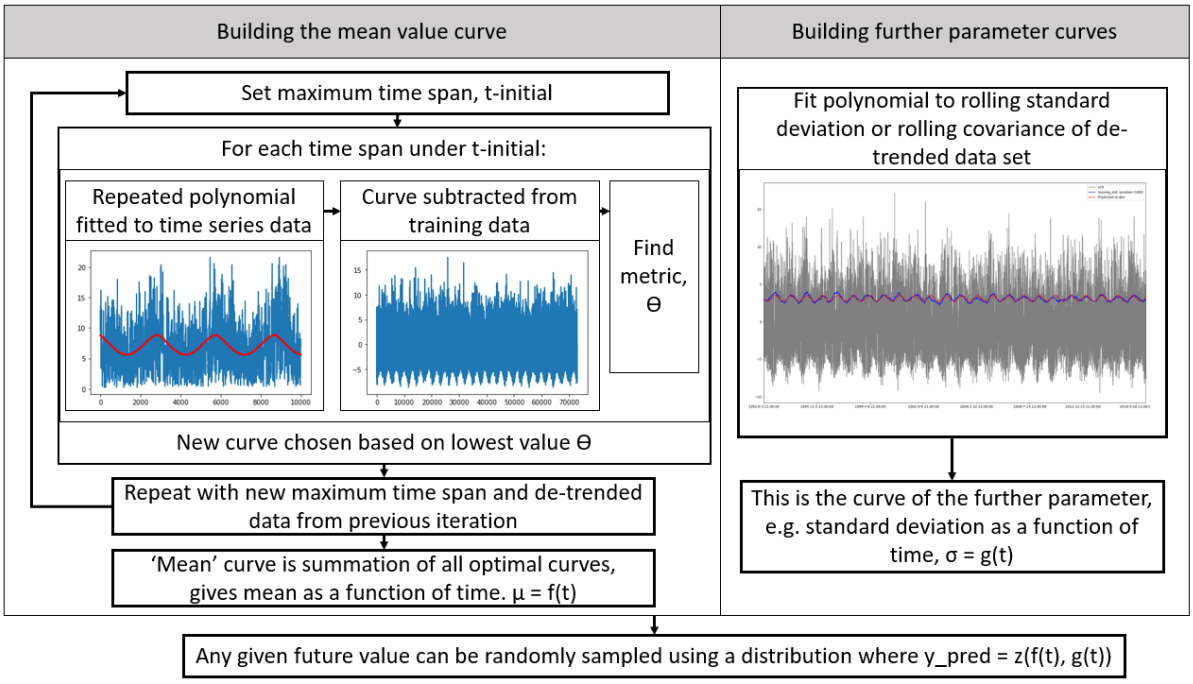


Figure 34. Hybrid method process diagram showing the approach used to fit curves for distribution parameters from time series data.

The model can be expressed in the pdf of equation 5 (univariate) or equation 6 (multivariate).

$$f(t) = \frac{1}{\sigma(t)\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{z-\mu(t)}{\sigma(t)}\right)^2} \tag{39}$$

$$f(t) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma(t)|^{\frac{1}{2}}} e^{-\frac{1}{2}(z-\mu(t))^T\Sigma^{-1}(z-\mu(t))} \tag{40}$$

Where the parameters are time dependant and are calculated in this model as:

$$\mu(t) = \sum_{n=0}^{N}\sum_{k=0}^{K} a_k t^k \tag{41}$$

$$\sigma(t) \; or \; \Sigma_{i,j}(t) = \sum_{k=0}^{K} a_k t^k \tag{42}$$

With *N* as the number of optimized timescales, and hence the number of curves, and *K* is the polynomial order of each curve. The items *i* and *j* are indices in the covariance matrix. To match

the rate of change of the real measurements, samples are taken every 5 intervals and linearly interpolated between these.

5.2.4. Strengths and weaknesses of methods

The following table presents the strengths and weaknesses of the methods presented in section 5.2.

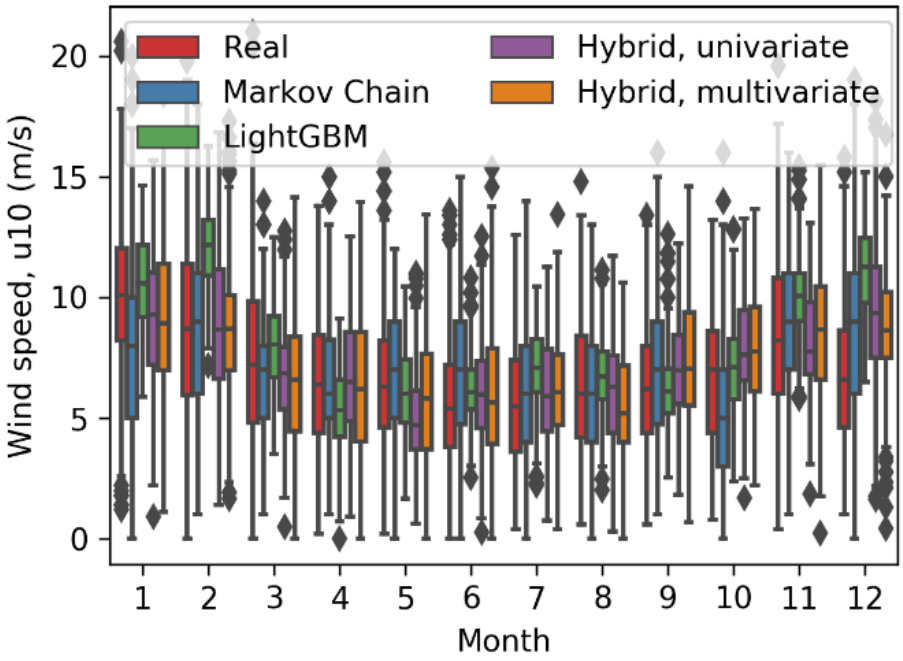
Table 18. Strengths and weaknesses of the methods used in this chapter. Markov Chains, Gradient Boosting and the hybrid regression method.

	Strengths	Weaknesses
Markov Chains	• Accurately replicates the probability distribution [416]. • Can be generated with limited data (e.g. one year) [70] • Well developed with improved implementations. [413]	• Seasonality accomplished by having separate models, not unified in one, therefore limited. [70] • Persistence dependant on prediction rate and limited to certain timescales. [416]
Gradient Boosting	• Very low computational requirement and quick to train. • Can be useful at a range of timescales. [66,67,421] • Flexible and suitable for a variety of problems [419,420].	• Not a statistically based method, so it won't inherently capture statistical distribution. • Deterministic method and so cannot be used as a stochastic input without modification.
Hybrid Regression	• Captures trends at all significant timescales. • Level of persistence is a tuneable parameter. • Uses statistical distributions and is therefore closer to recommendations from standards. [38,53]	• Slow to train in its current form as the code is not optimized. • Depends on linear interpolation between sample points. • New idea and so not developed yet by other authors.

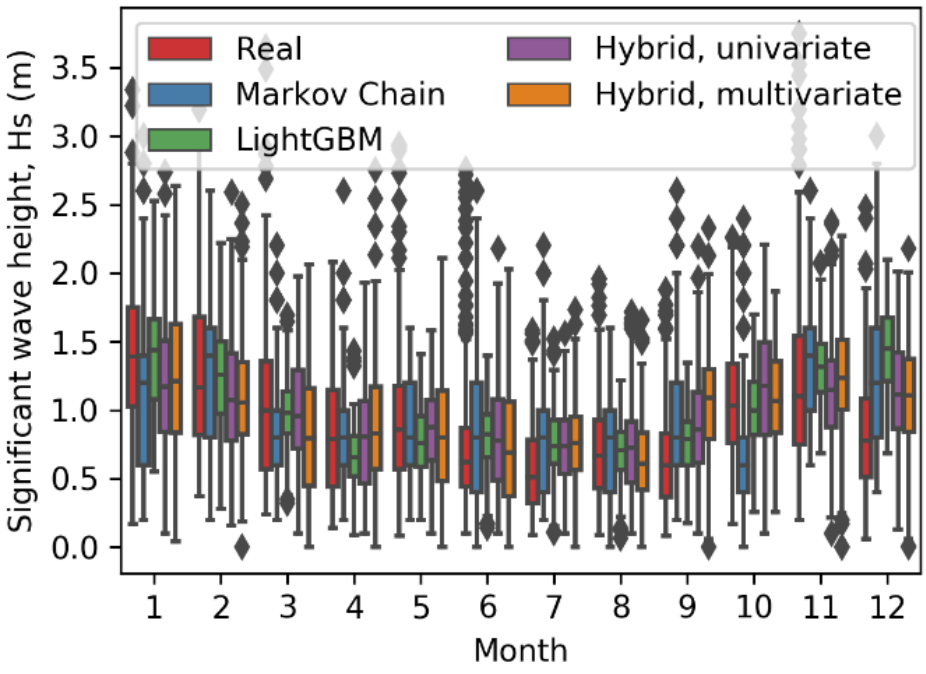
5.3. Results and discussion

The time series forecasting methods are now compared in terms of their numerical similarity to the real time series data. This is to show which is a closer approximation of the real data as well as to help explain differences which will be seen in the operational KPIs. The methods are compared for the year 2016, which is the last year that real data is available. There is a large amount of scatter in real wind and wave values, so it is unreasonable to compare individual 3-hour observations. Instead, the results are compared on a monthly basis.

Figure 35 shows box plots of the real as well as predicted values for the year 2016. This shows that the general results of all models are similar and roughly comparable with the real values. However, there are months where the mean values of predictions as well as the level of deviation do not match as closely with the real data, particularly in December where all methods overestimate the mean value. To investigate this further, the difference in mean and standard deviation are calculated.



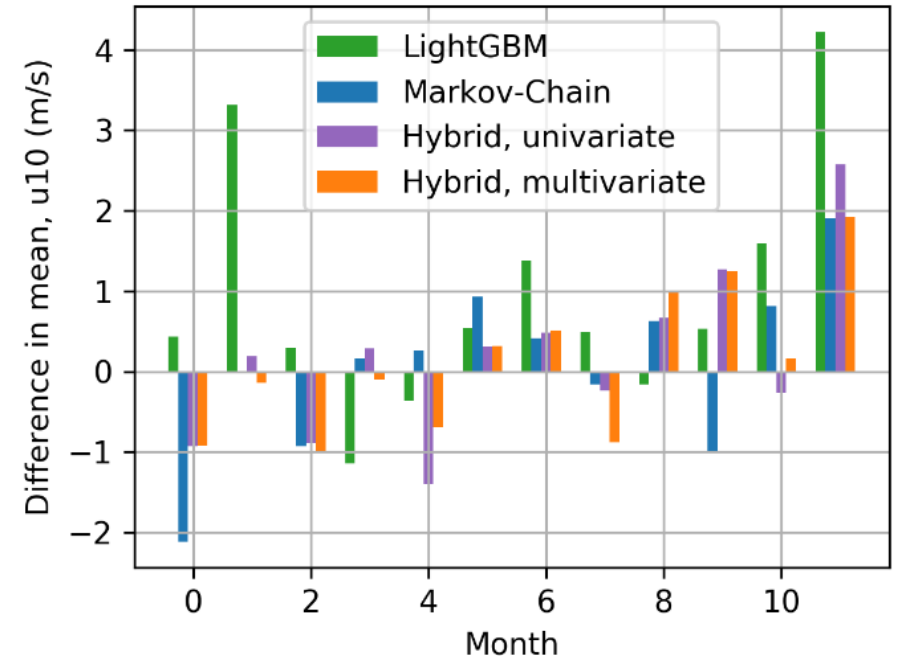
(a)



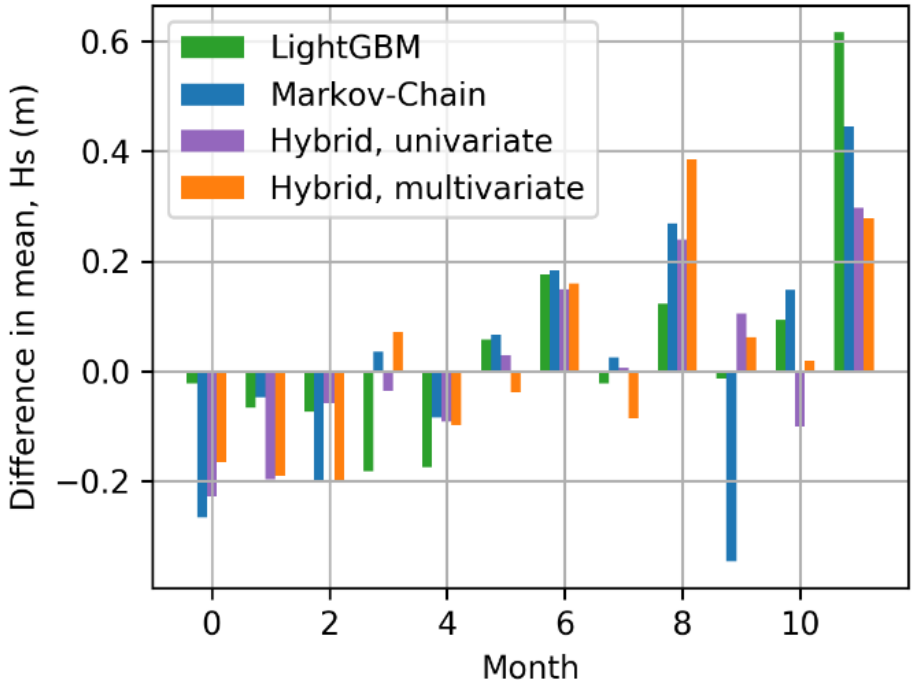
(b)

Figure 35. Box and whisker plot showing real, monthly values (left, red) compared to predicted monthly values for wind speed (a) and significant wave height (b) for 2016.

Figure 36 shows this difference, from the real measurements, in monthly mean values for wind speed and significant wave height for 2016. In both *u10* and *Hs* the methods all under predict the early months and over predict the later months. There are occasional months where all but one method shows a low difference, for example in October Markov Chains under predicts the mean values while all other methods slightly over predict. These discrepancies are typically quite small, roughly 10% of the mean values for the month. These differences are summarised in the following table.



(a)



(b)

Figure 36. Difference in mean, for each month, compared to the real data for 2016. (a) Wind speed, u10 and (b) Significant wave height, Hs. Showing deviation from real average.

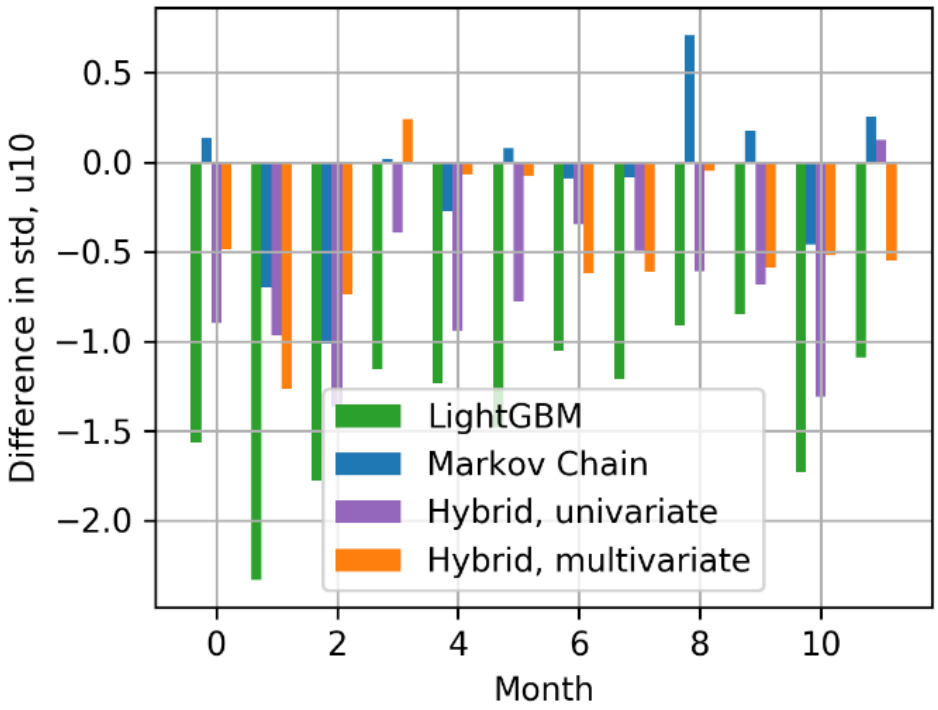
The conclusion from the values shown in Table 19 is that the tree-based approach in LightGBM gives predictions with mean values furthest from the real mean, and also gives comparatively large maximum differences. Markov Chains, as well as the two hybrid methods,

give the closest monthly mean values with the hybrid multivariate approach showing slightly better results. However, it is not close enough to say definitively that the hybrid methods outperform Markov Chains in this, only that they are all similarly good.

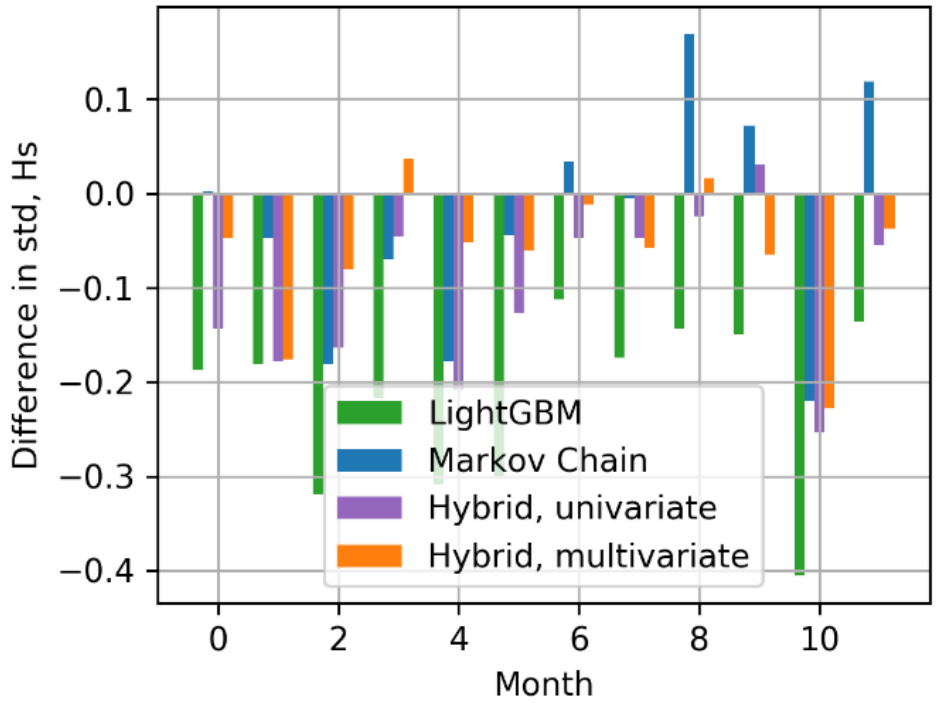
Table 19. Summary of difference between mean of real and forecast monthly values for 2016.

	U10 average	U10 maximum	Hs average	Hs maximum
LightGBM	1.206	4.221	0.135	0.616
Markov Chains	0.776	2.119	0.176	0.445
Hybrid, univariate	0.793	2.579	0.128	0.297
Hybrid, multivariate	0.741	1.919	0.146	0.385

With a few exceptions, all methods under predict the level of deviation in each month, as shown in Figure 37. LightGBM particularly under predicts the level of deviation, which is to be expected as tree-based approaches use an average of a leaf and so some level of scatter in the measurements is averaged out. All methods are closest to the real value of standard deviation during the mid-months such as 4 and 6; this can simply be due to the fact that there is much less deviation in the real values for these months. These differences in standard deviation are summarised in the following table.



(a)



(b)

Figure 37. Difference in standard deviation for each month compared to the real measurement values. (a) wind speed u10 and (b) significant wave height, Hs

The conclusion from Table 20 is that LightGBM shows the largest difference from the real standard deviation while Markov chains shows the smallest with the Hybrid multivariate values

very close to the Markov Chain values. The values for Markov Chains and the two hybrid methods are very close and so it can only tentatively be concluded that Markov Chains performs slightly better in this.

Table 20. Summary of difference between standard deviation of real and forecast monthly values for 2016.

	U_{10} average	U_{10} maximum	H_s average	H_s maximum
LightGBM	1.367	2.334	0.220	0.406
Markov Chains	0.332	1.015	0.095	0.220
Hybrid, univariate	0.743	1.368	0.110	0.254
Hybrid, multivariate	0.485	1.267	0.072	0.228

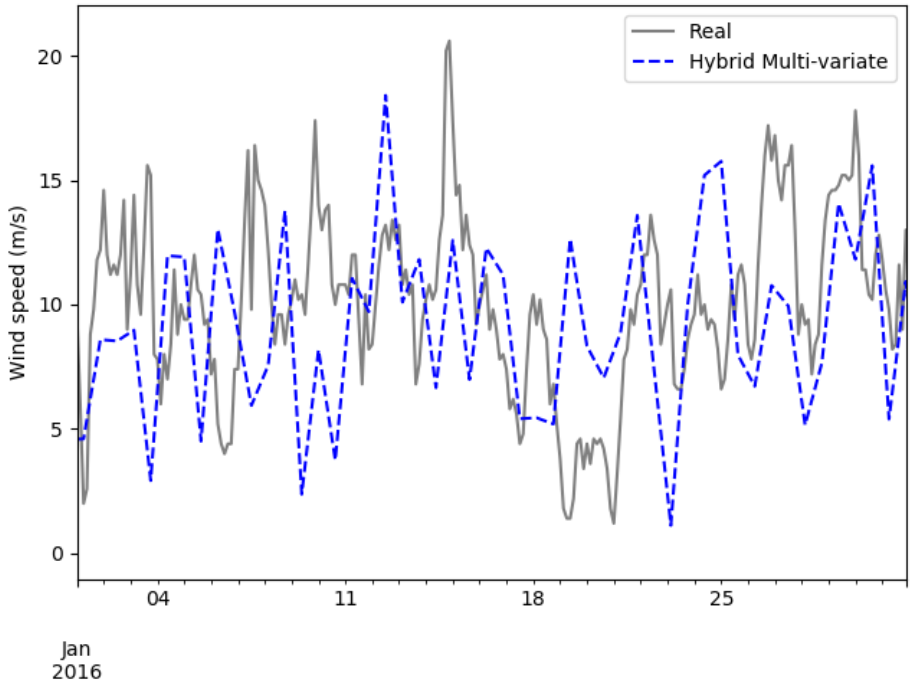
The final value which is important in the replication of the forecasts compared to the real values is the level of correlation between the wind speed and the significant wave heights. This level of correlation is important for O&M modelling as some activities are dependent on both values. The Pearson correlation coefficients are given in the following table. This coefficient shows the level of linear correlation between two sets of variables in both magnitude and direction. A value of +1 indicates perfect positive correlation and a value of -1 indicates perfect negative correlation. 0 indicates no relationship between the two sets of variables.

Table 21. Pearson correlation coefficients between wind speed, u10, and significant wave height, Hs, for the real measurement observations and the forecasts.

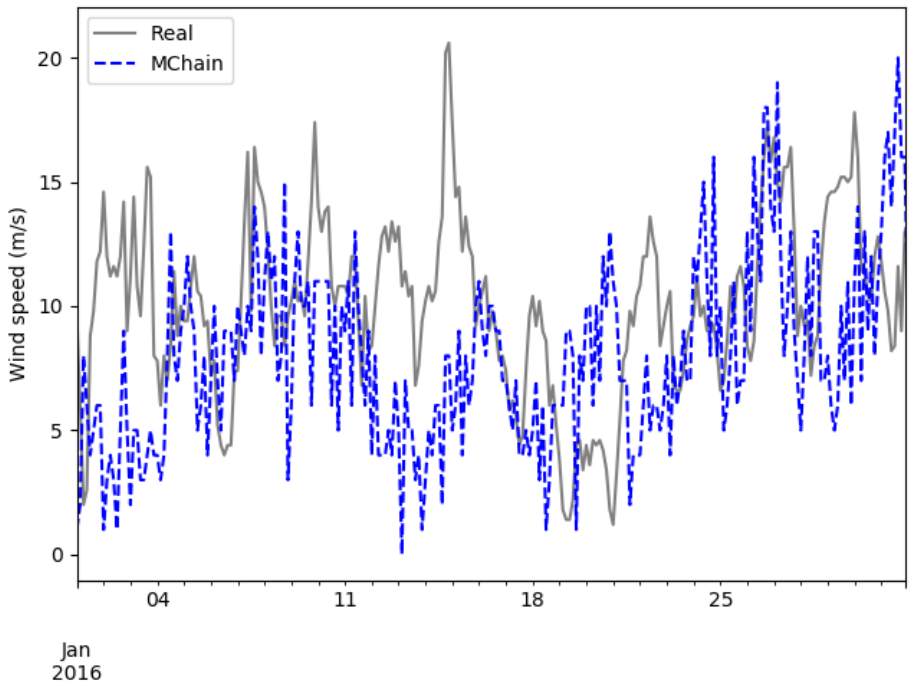
Real	LightGBM	Markov Chains	Hybrid, univariate	Hybrid, multivariate
0.829	0.579	0.807	0.120	0.810

The level of correlation in the real observations between u10 and Hs is roughly 0.8, which is also accomplished in Markov Chains and the hybrid multivariate approach. The tree based LightGBM shows some level of correlation and the hybrid univariate shows no correlation.

Based on the previous metrics, Markov chains and the hybrid multi-variate approach appear to produce comparable estimates. However, looking within one month shows this to not be the case. There are several distinctions which can be seen in plots for January 2016, as an example, shown in Figure 38.



(a)



(b)

Figure 38. Time series plots within the month of January 2016 comparing real values to predictions. Recorded wind speed time-series is plotted in grey. (a) - prediction from the hybrid multivariate approach for that month, (b) - prediction from Markov Chain.

The Markov Chain appears to show a higher persistence, it remains around a value for longer, than the hybrid method. The persistence in Markov Chains is a fundamental quality of the model and is dependent on the training data. For the hybrid method persistence is determined by how frequently predictions are made, and so is a changeable parameter.

The Markov Chain captures seasonality within a year by training a separate model for each month. However, any trend which occurs within a month is not captured using this approach. Conversely, the hybrid approach fits the parameters to trends at multiple timescales and so shows trends within a month. This point about shorter trends is clear in Figure 38 where the general movement of the hybrid method follows the real data in a more deliberate way then the Markov Chain.

5.4 Conclusion

In this chapter a comparison of numerical weather forecasting methods and their results were presented. Three numerical methods were used including Markov Chains, gradient boosting and a proposed hybrid regression/ statistical method. The methods are compared in terms of their numerical similarity to the measurement data.

The results show that, numerically, both Markov Chains and the hybrid method perform similarly to the measurement data. Markov Chains and the hybrid method are roughly equal in terms of prediction of monthly mean and deviation values as well as correlation between wind and wave vectors. The largest difference resulted from the gradient boosting method. The difference in Markov Chains and the hybrid method comes down to the level of persistence in the two methods as well as capturing trends at multiple timescales, where the hybrid method excels.

This work is novel because it evaluates the results of numerical weather model choice in comparison to site data. This chapter is also novel because it presents a hybrid forecasting approach which combines existing ideas in a way which is tailored for this problem and results in forecasts which are numerically similar to measurement data. Future work will focus on developing the numerical forecasting models further. This work will be of value to practitioners who are developing maintenance strategies for offshore wind farms, as contractual relationships between developers/operators and the grid often depend on ensuring adequate levels of availability; accurate prediction of such value is critical towards reducing risks of such agreements. Further, the study can be of value to researchers and developers of numerical tools which utilise forecasting algorithms as it shows how different algorithms can be modelled, and at the same time discusses performance characteristics for each method applied.

Part B: Structural variation and data

Papers:

6. M. Richmond, U. Smolka, A. Kolios, “Feasibility for Damage Identification in Offshore Wind Jacket Structures through monitoring of global structural dynamics” – **energies, 2020**
7. M. Richmond, S. Siedler, M. Häckell, U. Smolka, A. Kolios, “Impact of Accelerometer Placement on Modal Extraction of Offshore Wind Structures” – **OMAE, 2020**

Chapter 6

6. Damage identification through monitoring of global structural dynamics

6.1. Introduction

Operation and maintenance activities make up a large portion of the levelized cost of energy for offshore wind power. Operators spend this money because the cost of failure is larger and comprises cost of repairs, loss of productivity and loss of remaining fatigue life caused in the time the asset still operates while the structure is in a condition of exceeded design parameters. These exceeded design parameters could either be caused by damage, such as a crack, or environmental parameters such as excessive scour. Detecting the exceedance of design parameters quickly can save a lot of money, identifying the type of damage and assessing its severity can provide considerable further benefits. For example, it was found in a cost-benefit analysis that implementing structural health monitoring on offshore wind support structures significantly reduced the operational expense [433,434].

Data on failed structures is typically not available, being able to identify failure when it occurs is of significant economic benefit. To overcome this lack of data, researchers have modeled structures to simulate damage. Some researchers aimed to detect damage in their model in the presence of measurement noise; for example, Malekzehtab et. al. used mode shapes and natural frequency in an objective function, combined with a penalty term to avoid false positives [222]. Liu et. al found, even with 3% noise, modal flexibility could still effectively detect damage [223]. Wang et. al. investigated their approach in the presence of temperature variation and used Modal Strain Energy Decomposition to detect damage [224] while Xu et. al. used a residual strain energy-based approach [206].

This work differs from others in that a detailed design model is used which includes more complex aspects of the structure and allows for a more detailed investigation. This work includes the mass and moment of inertia for the nacelle and blades, an aspect which is not present in many other studies, but which is important in the modal properties of a wind turbine structure. Both element damage and scour are simulated allowing for direct comparison of their impacts on the same structure.

There are four levels of damage identification: damage detection, damage localization, damage severity assessment and damage consequence/ progression [226]. Detecting a damage is beneficial because while the turbine continues to operate, the remaining fatigue life is used much

more rapidly. By detecting the damage and parking the turbine, the excess consumption of fatigue life can be reduced. By localizing the damage, time and resources can be saved in its repair. With the quantification and assessment of damage, decisions regarding the growth and operational impact can be made. A great deal of knowledge is required to achieve all four levels. The use of virtual sensing, along with pre-existing detailed design models, can a cost-effective way of improving the detection and assessment of damage. Operational modal analysis (OMA) can provide information about the structure, and numerically solving the eigenvalue problem can provide information on how the structure responds to damage cases. The aim of this chapter is to show what information can be gained from a numerical, structural model and what lessons can be learned to aid those conducting OMA in the hopes to detect, identify and quantify damage as it occurs.

The approach used in this chapter is shown diagrammatically in Figure 39. The process followed is to alter a detailed design model with simulated damages, solve the eigenvalue problem for that model and compare the modal properties to the unaltered model. 'Damages' include element damage, global scour, and corrosion. The operational condition altered is the nacelle direction.

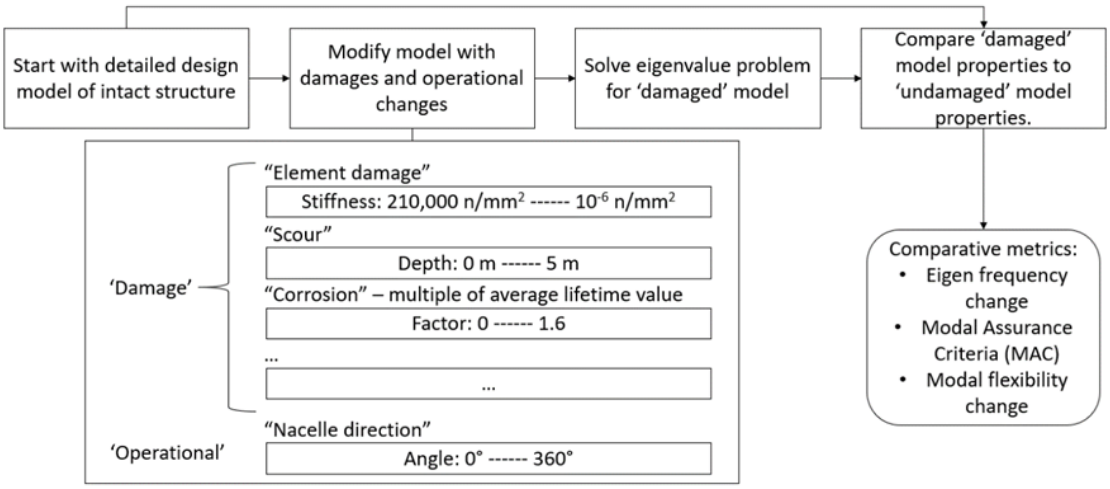


Figure 39. Diagram of damage process used to map damages implemented in a structural model with structural response caused by a change in the structure.

A limitation of this chapter is that there is no laboratory tests or operational data for comparison. This is unfortunate, but unavoidable. The real structure from which the model was based is not damaged, so no damaged operational data is available. Additionally, because so many parameters are investigated in this chapter, and the model incorporates so many details, a laboratory model of all of this would be unfeasible. There are other published publications in this field with the same limitation, for a few examples: [206,210,435].

The chapter begins with the damage quantification approach in section 2 which presents the metrics used and the equations for calculating them. In section 3 the structure evaluated in this case study is presented along with information about the numerical model used. The types of damage replicated in the numerical model and the method for implementing them are given in section 4. Results are presented in sections 5 and 6; the results in 5 are presented separately

because they are useful for understanding 6, which is the outcomes from each damage case. Finally, a discussion is given in section 7 and a conclusion in section 8.

6.2. Damage Identification approach

6.2.1. Metrics and the Modal Assurance Criterion equation

Two of the fundamental modal parameters of the structure are the natural frequency and the mode shape. The natural frequency is the frequency at which each mode shape occurs, and the mode shape is the vector in which the structure oscillates. The natural frequencies can easily be compared as they are scalar quantities; however, the mode shapes are vectors and so require an additional step for comparison.

MAC is commonly used to compare mode shapes; not only in numerical modelling but also in OMA. The equation for calculating the MAC, for two vectors $\{\varphi_A\}$ and $\{\varphi_X\}$, is given in equation (43) [436]. This gives a value ranging from 1 to 0 where a value of 1 indicates that the mode shapes are fully consistent, while a value of zero indicates that they are totally inconsistent.

$$MAC(A,X) = \frac{\{\varphi_A\}^T\{\varphi_X\}^2}{(\{\varphi_A\}^T\{\varphi_A\})(\{\varphi_X\}^T\{\varphi_X\})} \quad (43)$$

6.2.2. Algorithms, MAC vs frequency

When solving the eigenvalue problem, a set of natural frequencies and their corresponding mode shapes are determined. For a set of mode shapes, the convention is to assign labels to these modes based on the order of low to high natural frequency. It is observed that damage can cause the dynamics of a system to change, resulting in a shift of both natural frequency and mode shapes. As each 'damaged' case is distinct and not tied through some time history to an 'undamaged' case (and the same for two observations in an OMA situation) making comparisons based only on mode number is not always appropriate, as will be shown in the results of this chapter. This mode switching, which is an issue for pseudo-symmetrical structures, raises the question of which comparison should be made for a reliable identification of damage. This is important because in this study, the wrong pairing could show that a type of damage is easily detected, when it is not or that changes in the wrong modes indicate a certain type of damage. If the wrong pairing is chosen in an OMA situation then it could lead to a false alarm.

An approach which is used in this chapter is to select mode pairings such that the total sum of MAC values is maximized, thus assuming the minimum change in the system. This assumption is not always valid so care is taken when applying it – in this chapter it is found to be not appropriate when there is a very large change in the mode shape. To maximize the MAC value selection, a Kuhn-Munkres algorithm, sometimes called the 'Hungarian algorithm', is used [437]. This method reliably selects the highest total MAC value. Both this method and the method of selecting the diagonal elements are used, but often only one is shown.

6.2.3. Modal Flexibility based damage metric

The approach of using modal flexibility for damage detection was presented by Liu et. al. [223]. Modal flexibility presents an attractive means of deriving a damage index, as it is calculated as a combination of mode shapes and frequencies, there is no need for matching of mode shapes between damaged and intact. Additionally, as will be shown in the results, it is also relatively unaffected by changes to the turbine direction compared to the MAC values. One downside is that modal flexibility is affected more by lower modes than higher ones; in this case, changes only observed in higher modes cause less of a change in the modal flexibility. With virtual sensing the global modes 1-5 are targeted to reduce the number of sensors required in a potentially real case. In this study, the damage index used is the Frobenius norm of the modal flexibility residuals. This is calculated using the following equations. First, the modal flexibility is:

$$F = \sum_{i=1}^{n} \frac{\varphi_i \varphi_i{}^T}{\omega_i{}^2} \tag{44}$$

Where φ_i is the mode vector and ω_i the natural frequency for each i mode. A modal flexibility residual is then calculated as the difference between the intact and damaged cases.

$$F_{Residual} = F_{Healthy} - F_{Damaged} \tag{45}$$

A single damage index is then derived using the Frobenius norm of this residual.

$$\|F\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{m} F_{Residual}{}^2} \tag{46}$$

This approach is useful to indicate the level of damage but since the damage index is only a single value, it does not give enough information to diagnose the damage. Used in combination with MAC and change in natural frequency, more information about the effect of the damage can be ascertained.

6.3. Case study structure and model

The structure being evaluated is a four-legged jacket structure from the Wikinger wind farm, which is owned by Iberdrola. These jackets are situated in the Baltic Sea at a water depth of between 39-42 meters. Their footprint is 23 m square, they are 62 meters tall and support 5 MW, Adwen AD 5-135 turbines [438]. The structure is fixed with 40 m long piles at each leg [439]. The 67 m long blades are not modelled; however, the structure is modelled with the mass and moment of inertia of the blades and nacelle.

Dynamic analysis is conducted using Ramboll's proprietary, in-house software, Ramboll Offshore Structural Analysis Package (ROSAP), which Ramboll have developed over the past 30+ years. In this case, version 53, the latest at the time, is used [440]. The ROSAP package, and the module

ROSA, has been used for the design of the majority of offshore wind turbine foundations worldwide [441].

ROSA is a beam element-based design software used for frames, truss structures and piping systems. The transition piece is modelled as a 'super element' where the complex shape is analysed in ANSYS and the results are used as an element in this model. The soil structure interaction is conducted considering each soil layer and in accordance with API standards [442].

Only the first five mode shapes and corresponding eigenfrequencies are investigated in this chapter and the reasons for this come from the rationalization that the aim is to reflect practical measurement of operational global dynamics. While a coupled dynamics analysis will result in modes from sources such as the blades [99], this eigen analysis only gives structural modes and so the first five modes are major structural modes. The reason for not using higher modes is that in practice, in an OMA situation for offshore wind structures, lower modes can more reliably be detected and the most significant are first and second order fore-aft and side-side modes [211]. In this structure, the frequency of the 6th mode is much higher than the 5th, and so even with significant damage the frequencies do not change enough for a mode 6 or higher to be mistaken for mode 5 or lower. An example demonstrating the difference in eigenfrequencies can be seen in Figure 40 which shows the eigenfrequencies of the first 10 modes as a leg element is damaged. This demonstration shows that even after a large change in the structure, the 6th eigenfrequency is still much higher than the 5th.

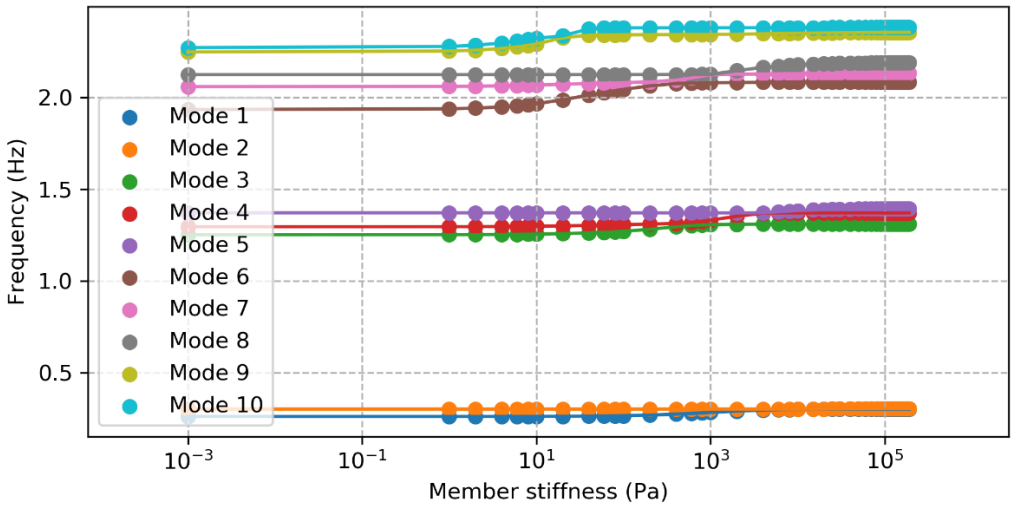


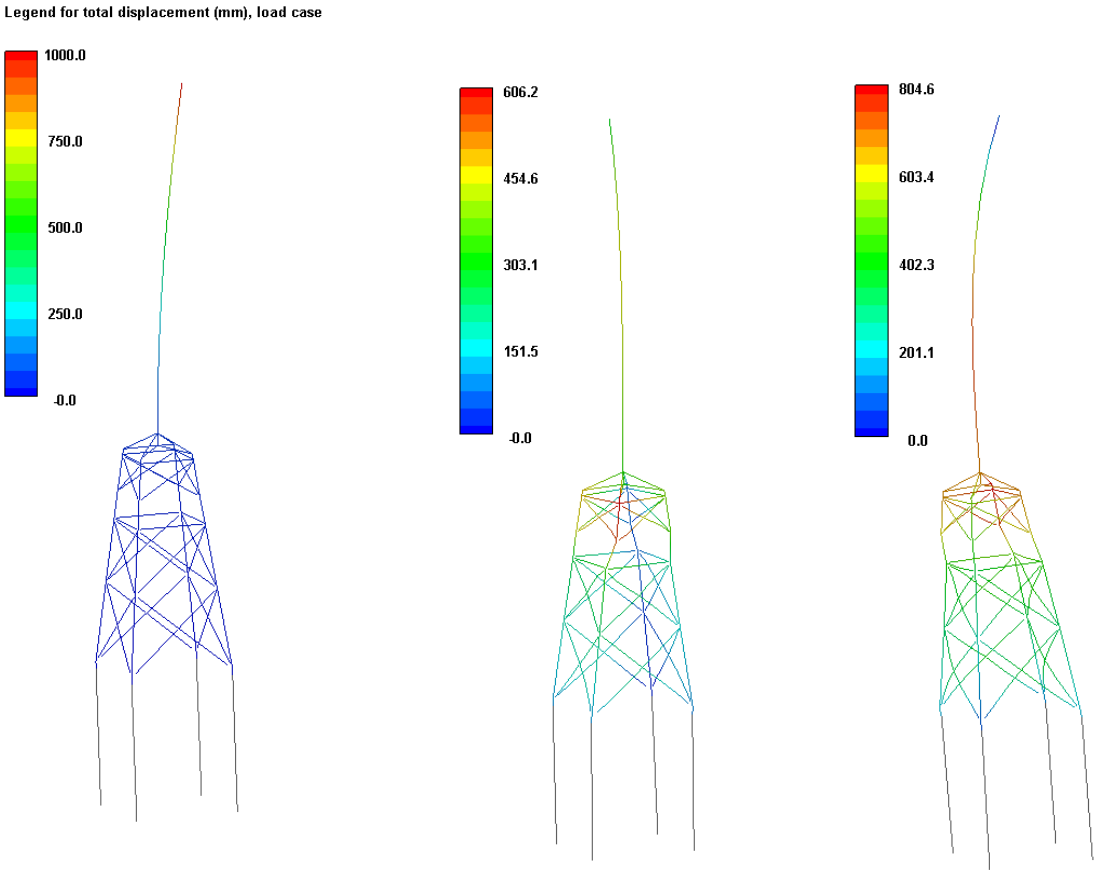
Figure 40. Example case showing the eigenfrequencies of the first 10 modes as the stiffness of a leg element is reduced from design stiffness (right) to complete loss of stiffness (left).

Of the five modes calculated, the first two are both tower modes which are tower fore-aft and side-side. The third mode is a torsional mode while the fourth and fifth are both tower-swaying modes. Plots of these modes along with their respective frequencies, are shown in Figure 41, with modes 1 and 2 as well as modes 4 and 5 shown together. A representation of the structural model is shown in Figure 42.

Fundamentally, ROSA works in the same way as any beam-element based structural model which solves the eigenvalue problem and so these results can be reproduced with other, similar software.

A wide range of relevant parameters are included in ROSA and can be adjusted by the user in ‘cards’ within the text-based input file. Some of these parameters include element and joint stiffness; corrosion; scour, both global and local in addition to a range of other items. The parameters changed are discussed in more detail in the following section.

There are several failure mechanisms which are time dependent and expected to occur, such as marine growth, scour, corrosion and some others [443,444]. The model includes these at the level which they are generally expected, or an average amount over the lifetime of the structure. In cases when one of these parameters is varied, they might be given as a multiple of this expected value or as a measurable value, as appropriate. For example, corrosion is defined as a profile with respect to depth and so this is varied as multiplied by a factor, however, global scour can be quantified with just scour depth and so this value is given.



(a)

Mode 1: 0.3017 (1/s)

(b)

Mode 3: 1.3107 (1/s)

(c)

Mode 4: 1.3712 (1/s)

Mode 2: 0.3042 (1/s) Mode 5: 1.3949 (1/s)

Figure 41. Shapes of first five modes along with their respective frequencies. Coloured by total displacement. (a) tower modes 1 and 2 (b) torsional mode 3 (c) swaying modes 4 and 5

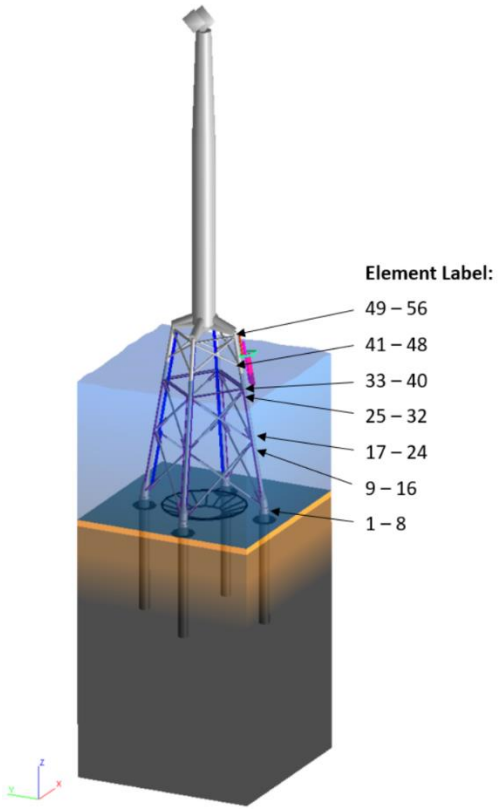


Figure 42. Representation of the structure in ROSA, including appurtenance, sea and soil. Labels of joints presented.

6.4. Structural model and damage modelling

6.4.1. General approach

Each case is an eigen value solution from the ROSAP model. This has both advantages and disadvantages and it's important to be aware of them. The disadvantage is that environmental impacts such as damping from wind or waves is not considered as they would be in a coupled dynamics simulation. The advantages are that it is possible to consider a wide range of damage cases and focus on the direct impact to the structural dynamics. Environmental impacts are not considered here because they can be measured in a real structure without causing any damage and so real measurements can be referenced for this [210]. Some examples of these measurements in literature, which include temperature variation, can be found at these references [211,445]. Frequency based measurements are temperature dependent and so when results of frequency changes are presented in this chapter, the reader should bear in mind this change.

As each damage case is implemented, only one damage is implemented at a given time. There is an average, expected state, which has some level of scour and corrosion at values which are expected, and everything is changed relative to this.

6.4.2. Joint damage

Researchers who have developed damage detectability models by implementing some damage into an Finite Element Analysis (FEA) modelled structure have typically limited their investigation to joint damage [222]. The way that joint damage has typically been implemented is through the reduction of either joint or element stiffness. Within many equations used to derive the equations of motion for the beam members, the effect from a change in stiffness and from area are proportional, as shown in the Lagrange equation given for reference in equation (47).

$$U = \frac{1}{2} EA \int_0^L \left(\frac{\partial u}{\partial x}\right)^2 dx \tag{47}$$

Where *U* is the strain energy, *x* is the length along the element and *u* is the displacement, while *E* is modulus of elasticity and *A* is cross-sectional area. The assumption that a change in stiffness is proportional to a change in area due to the growth of a crack is not always strictly realistic, but for investigating the first five natural frequencies this assumption has been made by other authors [224,446].

Two types of joint damage are implemented: complete loss of the joint and gradual loss. The jacket structure is composed of four legs and the legs are joined by cross members. The cross members are less significant to the modal properties than the legs and so all of them are tested in a complete loss. The legs are integral to the structural integrity of the jacket and their complete loss is typically not survivable. Instead of modelling the complete loss of the legs, the gradual reduction in stiffness of two leg elements is modelled. This is done with the idea of determining whether the reduced stiffness impacts the modal properties in a noticeable enough way that the impending loss can be detected through OMA before complete loss occurs. This is modelled in multiple stages where, at each stage, the stiffness of the joint is altered and then the modal properties are determined.

Based on the previous discussion in this section, the physical meaningfulness of these changes in Young's Modulus is representative of a crack at two general values: 1) relatively low reductions in *E* where the effect is roughly proportional to reduction in area due to the crack and 2) reduction of *E* to 0, where the joint no longer supports load. The range of *E* where it is significantly reduced but not yet 0 is too large to draw a direct reference to a crack growth.

6.4.3. Pile interaction

Multiple parameters are modelled and can be changed in ROSA. The soil-structure interaction is modelled using the approach recommended by the API recommended practice RP 2A [442] and dynamic loading. The soil is modelled as springs with stiffness dependent on displacement factors

which can be applied to these curves to either compress or elongate them. The soil itself is modelled in layers, each with a characteristic strength. Both local and global scour are modelled.

The parameters investigated are scour, *P-Y/Z/W* displacement factors and characteristic soil strength. Scour is varied from 0 to 5.5 meters, which in this case is removing the top layers, exposing pile length. The displacement curves are multiplied by a factor which either compresses or elongates the curves. This is not representative of any physical process but rather investigates the possibility of uncertainty in this parameter. All characteristic strength values are multiplied by a factor from their original, 'expected' value – this is also not representative of a process but an investigation of variation.

6.4.4. Bolted connection

The tower is connected to the transition piece through a bolted flange connection, where tightness of the bolted connection must always be ensured. Loosening can be detected with sensors on the bolts or with an inspection, but this study investigates if this can be detected in modal parameters.

The individual bolts are not modelled in ROSA but rather the interface between the tower and transition piece is modelled as a node with a stiffness value. To model the loosening of the bolts, the stiffness of this node is gradually reduced from its original values down to 0. The physical meaningfulness of this is not a reduced joint stiffness with the members in contact, but rather points at which separation of the clamped members occurs resulting in a lower stiffness. The equation for the stiffness, k_{fr}, of an individual bolted connection (frustum) is given in equation (48) and the total joint stiffness, k_{grip}, is given in equation (49) [447]. Where d is the inner diameter of the frustum, *D* is the smallest value of the frustum outer diameter, *t* is the frustum thickness and α is the pressure angle of the cone. The tension in the bolt does not come into these equations and so, based on this, as the tension is lost, there would be no change in joint stiffness. However, when significant load is applied, there is a point of separation between the two bolted members and all the load is then carried in the bolts, which have a lower stiffness than the members. It is this stiffness after separation which might be observed by the change in modal parameters.

$$k_{fr} = \frac{\pi E d tan(\alpha)}{ln\left[\frac{(2t * \tan(\alpha) + D - d)(D + d)}{(2t * \tan(\alpha) + D - d)(D + d)}\right]} \quad (48)$$

$$k_{grip} = \frac{1}{\frac{1}{k_{fr1}} + \frac{1}{k_{fr2}} + \cdots + \frac{1}{k_{frN}}} \quad (49)$$

6.4.5. Corrosion and marine growth

In ROSA, corrosion is modelled as a reduction in material thickness with both internal and external corrosion, but with the mass of the corroded material still included in calculations. Equivalent

outside diameter is unaffected by this. The amount of corrosion is set based on the elevation, since the level of corrosion is driven by variables that change with depth. A certain amount of expected corrosion is included in the model, the values used in the model are a multiple of this expected level of corrosion by a factor, assuming a linear growth. The expected values for corrosion are both internal and external and range from around 1 mm to around 12 mm depending on height on the structure and exposure zone. One of the main concerns of marine corrosion is that it can exacerbate the initiation and growth of cracks and so detecting it is important [448].

Marine growth affects the mass as well as the surface roughness height, thereby affecting structural mass as well as damping due to changed fluid dynamics [449–451]. As with corrosion, the values are set based on depth. The values are plotted as a multiple of an expected value.

6.5. Detectability considerations

6.5.1. Dependence on nacelle direction of modal properties in intact condition

The turbine blades and nacelle have a large mass and moment of inertia and as a result their position relative to the rest of the structure has a significant impact on the modal parameters of the structure, which is particularly apparent due to the structure's lack of rotational symmetry. Therefore, it is necessary to investigate the effect from changing nacelle yaw direction.

The nacelle and blades are modelled here as point masses with pre-assigned moment of inertia values. To model the nacelle yawing from the original reference direction, the point masses are moved around the tower by 1 degree each iteration and the moment of inertia components are transformed correspondingly. This process of changing the direction is then repeated for a 'damaged' case, with joint loss at the middle of a leg element.

To determine the response of the structure with the nacelle yawed away from the original direction and compare it back to the original reference direction, the mode shapes are then transformed back to the original reference direction and compared to the 'intact' case at the original direction; the transformation is conducted using elementary coordinate rotation as shown in equation (50) [452] which is applied for both translation and rotation values. An alternative approach is to compare 'damaged' at θ degrees to 'intact' at θ degrees where θ is each change from the reference. The first method is used here rather than the second as this would make future comparison to operational data more convenient since the directional values in operational data are effectively continuous. This method also allows them to use more of the data rather than only comparing subsets.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} cos\,\theta & sin\,\theta & 0 \\ -sin\,\theta & cos\,\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (50)$$

As the turbine direction yaws to make a complete rotation, the natural frequency does not change by an amount that exceeds a 4% threshold where it would be considered detectible. The maximum change in frequency calculated using the ROSA model, normalized by the original frequencies, for the first five modes are: 0.002, 0.009, 0.002, 0.022, 0.024 respectively. However,

the mode shapes, as quantified by the MAC values, change considerably. Figure 43 shows the MAC values for the first five mode shapes, as calculated after the shapes have been transformed back to the original reference. MAC is calculated in comparison to 0 degrees case.

The first two modes are the tower fore-aft and side-side modes. In the first two modes the tower always oscillates roughly in the direction of the jacket legs. There is an edge-case exception to this which is when the nacelle is pointed equally between two jacket legs, which is the case for 45, 135, 225 and 315 degrees. In cases when the nacelle is pointed equally between jacket legs, there appears to be a transition as the direction of the first two modes swap with each other. The MAC value for 45, 135, 225 and 315 degrees is 1 because of the transformation approach. Without the mode shape being transformed, the MAC values in the leg directions are 0.5 as the MAC changes between 1 and 0.

Mode 3 is a torsional mode with the otherwise vertical axis of oscillation offset slightly in the direction of the nacelle. As the nacelle turns, the axis about which mode 3 oscillates also moves.

Modes 4 and 5 are tower swaying modes and their direction rotates at the rate of the nacelle yaw, but in the opposite direction (in a sense of clockwise/counterclockwise about the tower centerline). At 0 degrees of yaw, mode 5 is in line with the nacelle and 4 is perpendicular. At 45 degrees of yaw, mode 4 is in line with the nacelle yaw and mode 5 is perpendicular.

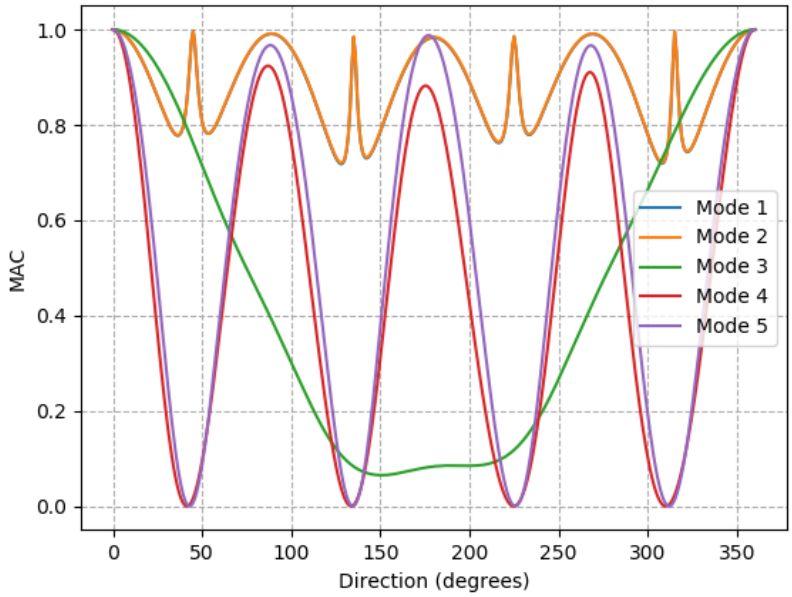


Figure 43. MAC values for the undamaged case as nacelle direction is rotated in comparison to 0 degrees. Comparison is made to the original direction at 0 degrees. Mode shapes are transformed to the original reference.

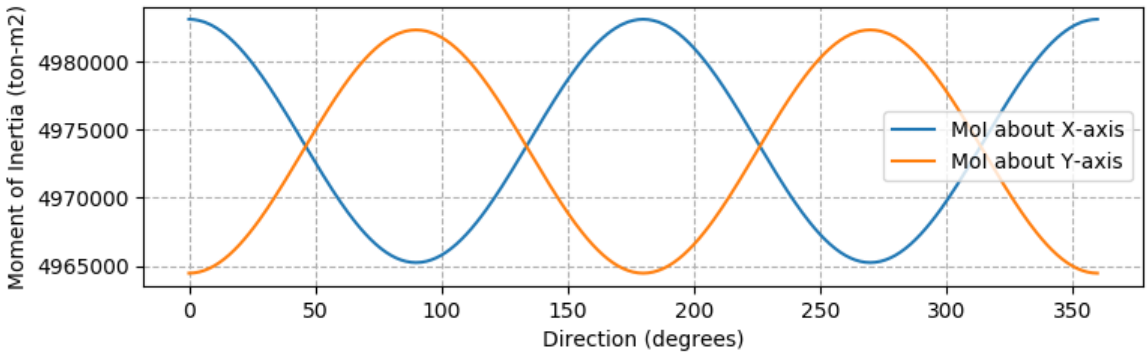


Figure 44. Total Structural Dynamic Moment of Inertia (MoI) with changing yaw direction

The change in mode vectors resulting from RNA yaw is important to be aware of because without accounting for it, even an intact structure would appear damaged using MAC. However, being aware of this, MAC can be used effectively to detect and identify damage.

Figure 44 shows the dynamic moment of inertia of the structure calculated by the model. This shows a smooth change in moment of inertia with nacelle direction which gives some indication that the model is performing as expected. We also see that the spikes in the 1st and 2nd modes seen in Figure 43 correspond to directions when the x-axis and Y-axis Moment of Inertia are equal; this is where the change in oscillating direction occurs.

The ‘damage index’ value calculated using modal flexibility varies less with direction, except for the four peaks. This damage index is not quantifying a damage in this case, there is no damage, it is showing the effect on this metric of an operational behaviour - it is a difference in the flexibility of the structure due to this operational behaviour. Besides the peaks, the values are all under 0.9, which by itself does not mean much but can be used as a reference when quantifying damage cases. The seemingly anomalous results in this plot are the peaks at 45, 135, 225 and 315 degrees. The modal flexibility, and hence damage index seen in Figure 45, is more affected by lower modes and these peaks correspond with the rapid change in MAC seen in Figure 43, which are changes in mode shape between the first and second modes. It is possible that these four peaks are an artefact of the numerical solution process and might not be seen in measurements of a real structure, so this aspect should be taken with caution.

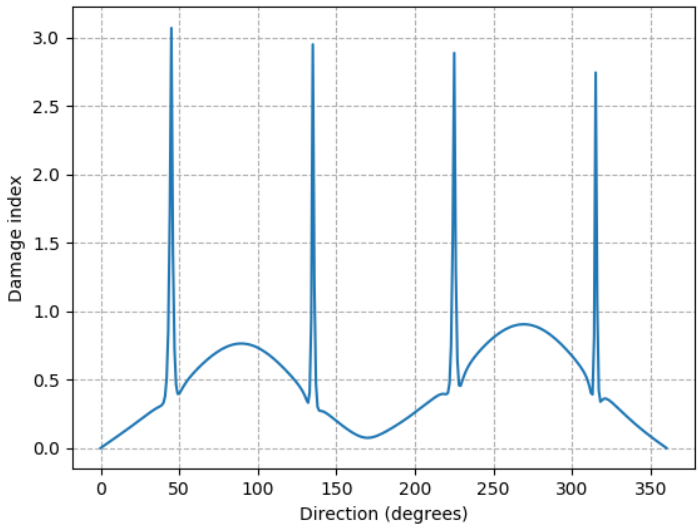


Figure 45. 'Damage index' as a result of changing nacelle direction, with no damage implamented. This comparative value is calculated in comparison to the nacelle at 0 degree case.

6.5.2. Choice of damage metric

The MAC is a comparative value between two observations. If an observation is made at two points in time, there is not necessarily any time history linking these two observations, instead there are a set of modes to be compared with another set of modes. This raises the question of how that pairing should be done to best represent the change. This is particularly important in the case of mode switching where the labelling of modes can change due to changes in frequency. This is important for the task of damage identification as well as potentially in sensor placement when determining which mode shapes need to be identified. Although each numerical model represents a discreet state, gradually induced change in the model can help give an indication of whether a comparison makes logical sense.

To illustrate the problem, as well as a potential solution to the problem, a leg element is gradually damaged, meaning the stiffness at a joint is gradually reduced. This is also useful from the perspective of damage identification to determine when a damage becomes detectible. In this case the leg element is halfway up the structure. Figure 46 (a) and (c) show the mode frequencies as simulated damage is gradually induced. (a) shows comparisons made where the order is based on frequency order, as is the established approach. Choosing the modes to compare based on the order of frequencies is not always the correct approach. Some of the modes have a gradual reduction in frequency, but the reduction stops when it becomes the same as the next mode down, but then the frequency of the next mode down continues the downwards trend – when one frequency passes another, the comparison changes. (c) Shows comparisons made where the sum MAC value is maximized using the Hungarian algorithm. With this approach, one frequency can go below the other without the comparison changing. The implication when comparing

frequency is that rather than determining the frequency for mode 1 changes by 14% from start to finish, instead mode 2 changes 14% and mode 5 changes 10%.

The effect of this change is even more apparent looking at MAC values. Figure 46 (b) and (d) show the MAC values calculated with these two approaches. If the comparison is made based on frequency order, then all modes have points where they become completely inconsistent with the undamaged mode. If the comparison is made based on maximizing the sum of MAC, then these drops happen only briefly before the MAC returns to a relatively high value. The drops occur when two frequencies approach equality. If there is a mode shape which is so close to being consistent with an intact mode shape, such as a MAC of greater than 0.9, then that is a strong argument that this pairing is what should be considered. This also gives more information for determining which damage has occurred, since if all MAC values are roughly 0 then this may also be the case for many types of damage.

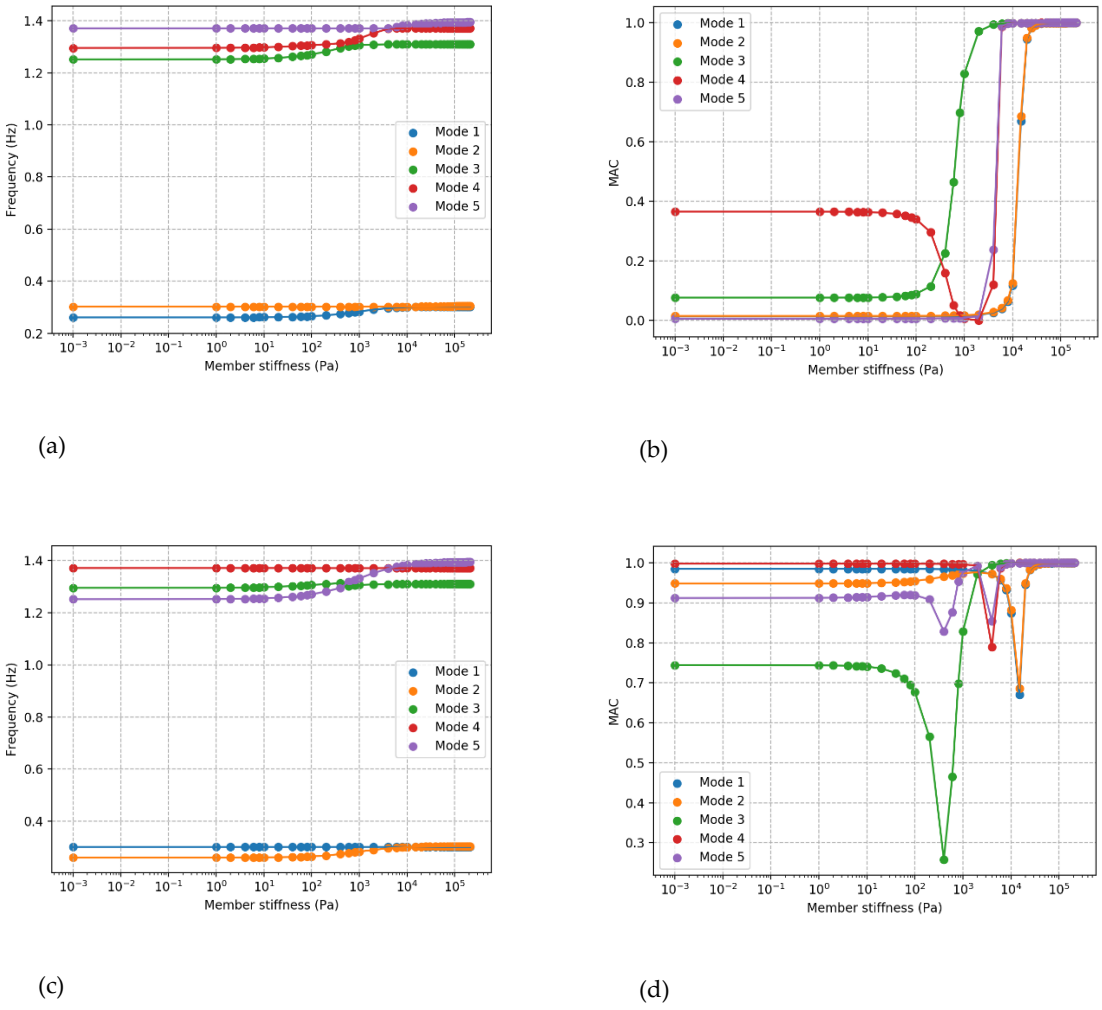


Figure 46. MAC values for the first five modes as simulated damage to a leg element is implemented. (a) Mode frequency value (b) MAC value, both with comparison based on mode frequency. (c) Mode frequency value (d) MAC value, both with comparison based on maximized MAC values

The damage index calculated through the modal flexibility does not have this same question of which modes to compare as all modes are combined before damaged and intact are compared. The damage index calculated for this gradual implementation of damage is shown in Figure 47. This figure shows a peak at the same time modes 1 and 2 have a sharp reduction at 2E4, but not any of the other similar sharp reductions. While the MAC returned to a relatively high value, the damage index continued to increase as more damage was implemented, therefore the damage index is a better choice for quantifying the level of damage.

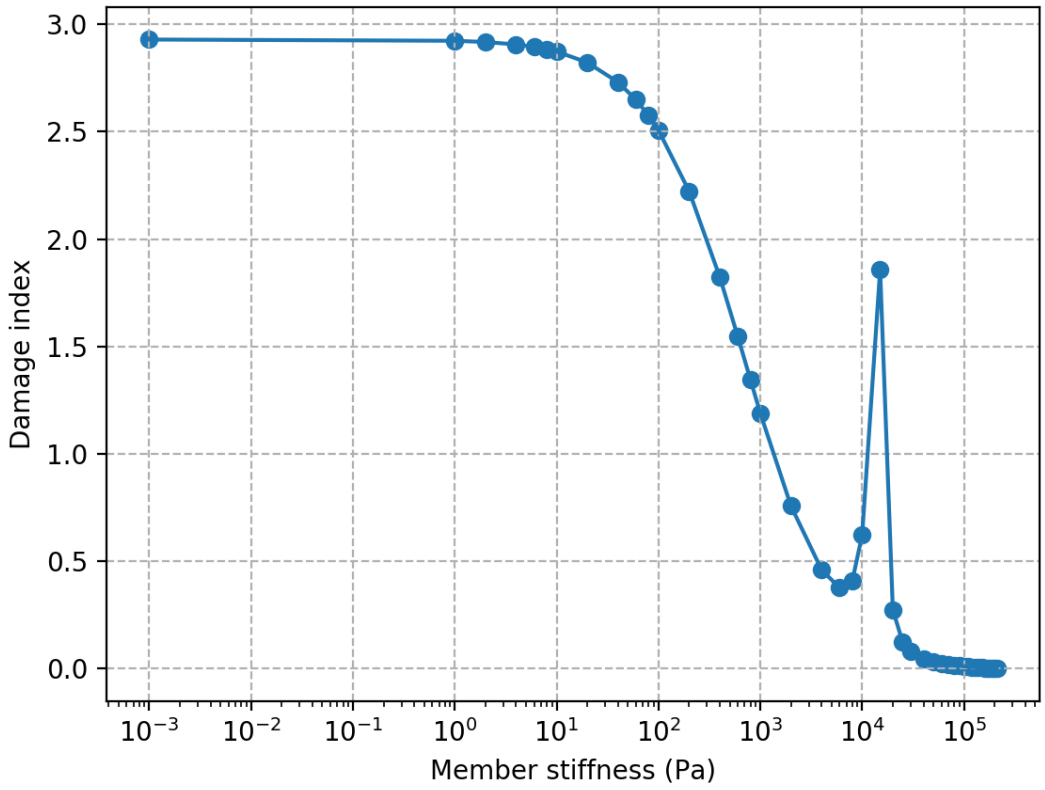


Figure 47. Damage index as simulated leg damage is implemented.

6.6. Results – Damage types

6.6.1. Joint Damage

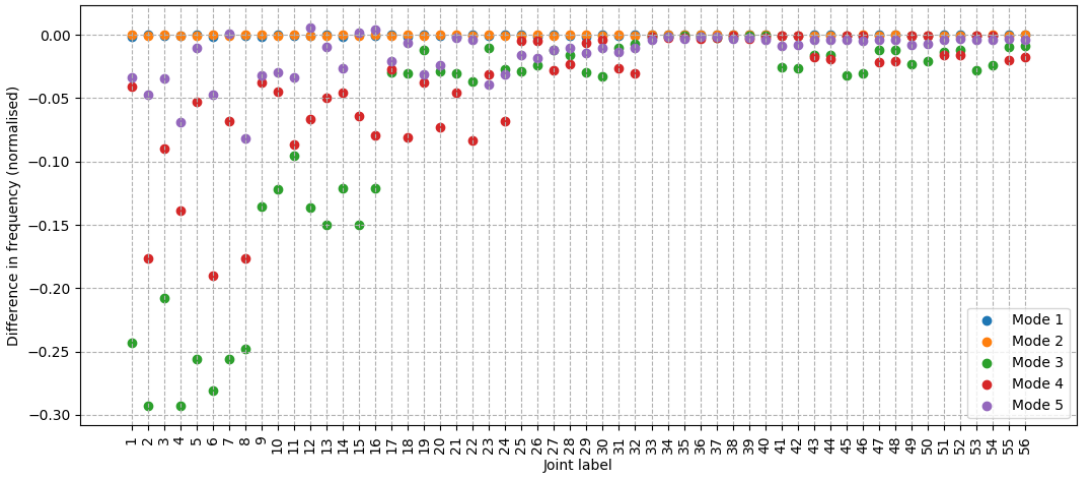
6.6.1.1. joint loss

This section gives the results for complete loss of joints (Young's Modulus equal to zero). In the first section this is shown for all joints in the structure with no change in nacelle yaw angle. In the second section, the focus is on the impact of nacelle yaw angle to damage detectability. Figure 48a shows the change in natural frequency of the first five modes, normalized by the intact natural frequency, for all cross-member joints. Figure 48b shows the MAC values for the structure with the loss of each joint, calculated relative to the intact structure. These results are all relative to

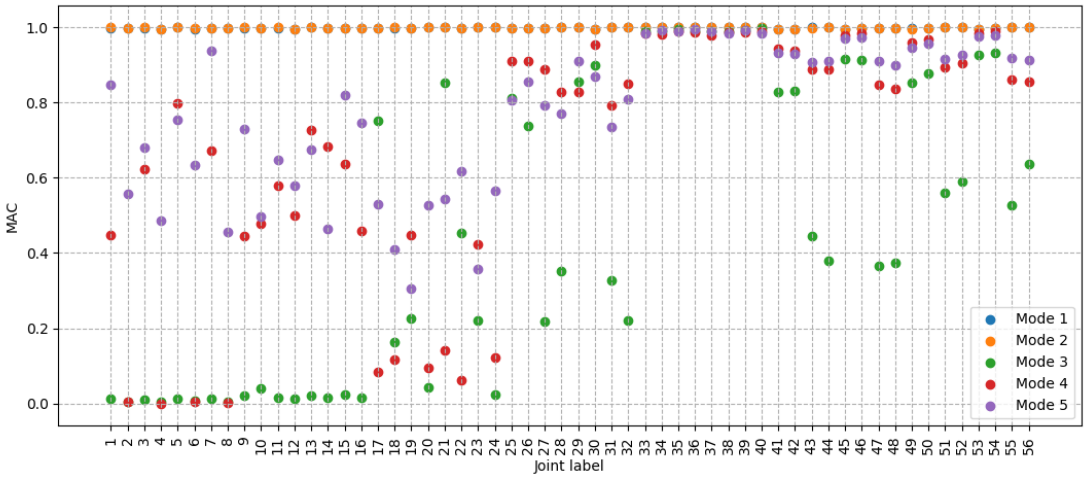
the intact, 0-degree nacelle yaw case and does not include any leg member joints, only cross-member joints, leg member joints are addressed later. Only a single damage is implemented at a time. Normalization is the change from the intact frequency, which is typically negative, divided by the intact frequency.

Frequencies for modes 1 and 2, the tower side-side and fore-aft shapes, are not significantly affected for any cross-member joint. However, the other three modes are changed noticeably. Based on field experience, frequency changes larger than 4% and a MAC value less than 0.8 are detectable, in that it can be discerned clearly from the standard variation caused by environmental and other factors. However, more study is needed to clearly establish detectability bounds as this depends on sensor quality, placement and site characteristics, among other factors. The joints lower down in the structure have generally more of an impact than joints higher up, as has been found in other studies [224]. Based on these graphs, no damage on cross members above 25m is detectible through frequency, and joints above 50m, elements 40 upwards, are only detectible through MAC if rotational modes are captured. There is almost no measurable effect from damage to the horizontal member joints, 33 – 40, since they are a relatively small part of this structure.

Using the ‘damage index’ as a metric results in the same conclusions as for the frequency and MAC plots. The one exception is that the values for the horizontal member joints are higher, between 0.2 and 0.3, which is as high as for elements which are higher than the horizontal ones. This would still be difficult, but easier to detect than through frequency or MAC.



(a)



(b)

Figure 48. Change in mode frequency (a) and MAC value (b) calculated relative to undamaged case.

6.6.1.2. Damage at varied nacelle direction

It can be observed from the previous section that, for a given nacelle yaw angle, the impact of loss from a given joint is not the same as a joint on the other side of the structure at otherwise the same location. Due to the mass and momentum of the nacelle and blades, the structure is not rotationally symmetric and so joint losses on different sides have different impacts. This lack of rotational symmetry can be best demonstrated by looking at the effect of damage to a leg element. Figure 49 shows the MAC for the case where a joint halfway up the leg is damaged; the nacelle and blade point-masses are turned a full rotation with their mode shapes transformed back to the original reference. The undamaged plot which this can best be compared to is Figure 43b, which is the undamaged case with transformation. The key point from this figure is that the MAC values

now exhibit a sinusoidal behaviour. If only the first and second modes are detected, then there are directions where the damage is not detectible by MAC as it is above the threshold, but also nacelle directions where it is much easier to detect the damage as the MAC reaches 0. Additionally, to identify which damage has occurred then the MAC value should also be assessed knowing what direction the nacelle was pointing when the mode shapes were recorded. If a leg next to this leg, at 90 degrees, is damaged at the same height, then the plot of the MAC value for different directions looks the same except that the phase is shifted by 90 degrees.

The frequency of the modes varies more significantly by direction than in the undamaged case. The change in frequency of mode 5 ranges from -0.04 to 0 and mode 3 varies from -0.08 to -0.04. The other three, modes 1, 2 and 4, stay roughly around their 0-degree change values of -0.14, 0, and -0.04 respectively.

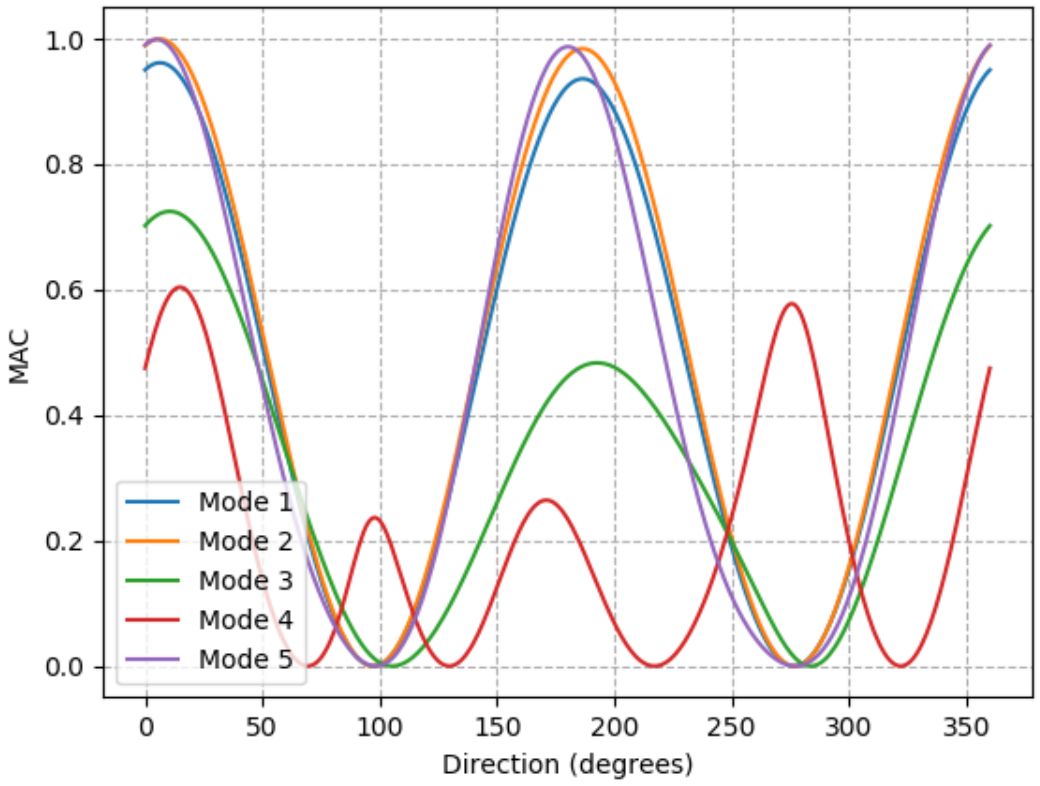


Figure 49. MAC values for the Leg damaged case at varying nacelle directions compared to undamaged case at 0 degrees. Mode shapes have been transformed back to original reference.

From these results, if the higher modes cannot be discerned in measurements, then at some directions of the nacelle the damage might not be detectible with the MAC values. In this case, the alternative approach of the modal flexibility can be used. The Frobenius norm of the flexibility residual stays between 0.65 and 0.25 for all nacelle directions and there are none of the peaks seen in the undamaged case. This indicates that the damage can be identified with modal

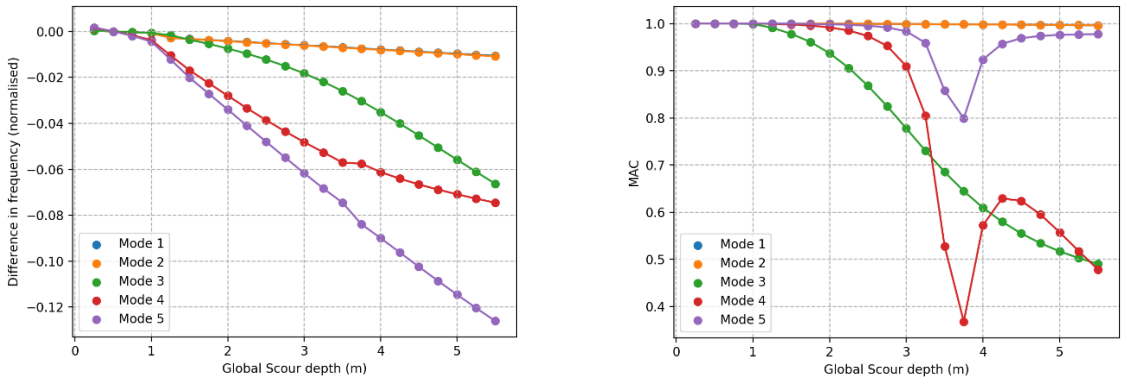
flexibility regardless of nacelle direction. In the damaged case there are no peaks as direction is varied, unlike in the undamaged case. The peaks have disappeared because there are no longer any sudden changes in mode shape, the transitions are now smoother.

6.6.2. Soil

The soil parameters investigated were global scour, variation in displacement curve scaling and variation in characteristic strength values. Global scour is a natural phenomenon which is expected to change over time, the other two parameters are not expected to change, but are investigated here as a sensitivity study of their impact on the modal properties.

6.6.2.1. Global scour

Jacket structures and monopiles respond differently to scour; others have found that the first mode frequency changes significantly with scour, around 5% for a scour/diameter ratio of 1 [453] for a monopile, however, the results presented here in Figure 50 show that for this jacket structure the frequency of the first mode is almost unaffected and the MAC does not go below 0.996. Modes 3, 4 and 5 are more affected by this change. At just under 4m scour depth, the frequencies of the 4th and 5th modes cross each other, resulting in a dip in the MAC values for these modes. The damage index calculated for this case is almost completely linear, fitting a linear trendline gives an increase in damage index by 0.22 per meter. A certain level of scour is expected, so over time from installation, the natural frequencies will reduce. [120]



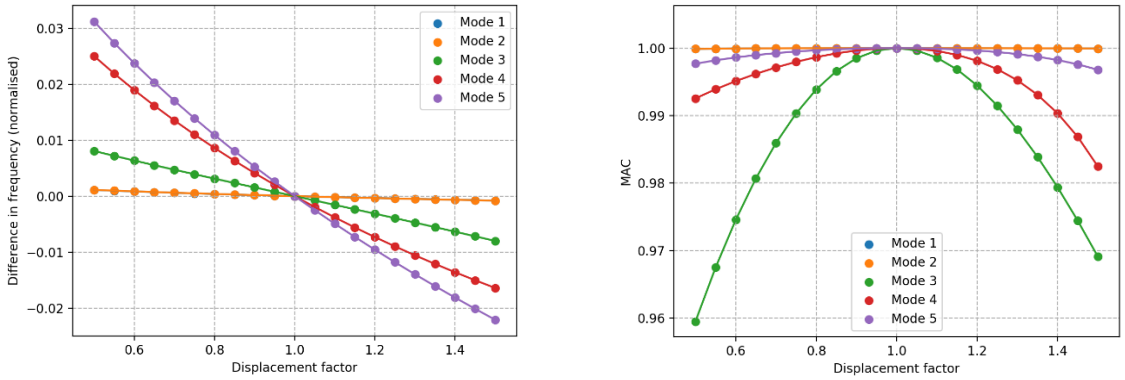
(a) (b)

Figure 50. Effect of gradually increased depth of global scour, compared to 0 scour depth. Modes compared based on maximized MAC (a) Difference in natural frequency, (b) MAC value. Scour depth in meters for the 23m x 23m jacket

6.6.2.2. Displacement factor

Values of the three displacement – stiffness curves are multiplied by the displacement factor; this gives an investigation of sensitivity of modal parameters to soil stiffness. A displacement factor of

1 indicates the expected curve and other values indicate either a compressed or elongated curve. The graphs for change in frequency against factor and MAC against factor are similar between the *P-Y*, *T-Z* and *Q-W* curves, although the amount the response is affected by the factor varies slightly. Figure 51 shows the change in frequency and the MAC values calculated as a result of changing the displacement factor for the *P-Y* curves, which was found to be the most significant of the three. The largest change in natural frequency is just over 0.03 and the smallest MAC is 0.96, these values are small, and the parameter is not time dependent. This is not an observable change but rather might act to make the measured parameters different from the 'intact' case.

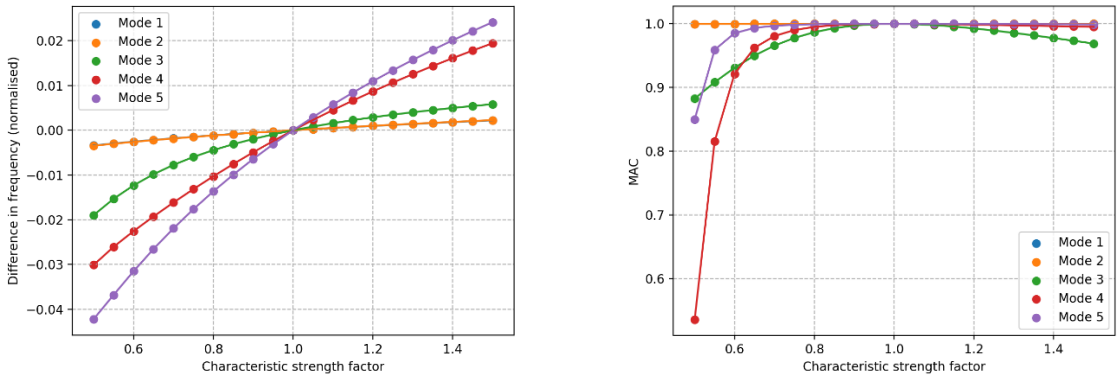


(a) (b)

Figure 51. Effect of displacement factor compared to a factor of 1 (a) Difference in frequency (b) MAC value

6.6.2.3. Characteristic strength

Variation in the characteristic strength of the soil has only a minor impact for this jacket structure. If the strength value is not below 0.6 of what is expected, then there is not much change to either the natural frequencies or the MAC values. Below a factor of 0.6 the MAC values drop quickly, however, even at this low value modes 1 and 2 show negligible change. This is shown in Figure 52, where the change in natural frequency is shown in (a) and MAC in (b). The effects are very similar to what is observed with a change in displacement factor.



(a) (b)

Figure 52. Characteristic soil strength sensitivity. (a) Difference in frequency (b) MAC value

6.6.3. Bolted connection

The bolts are not specifically modelled, so instead, the stiffness of the node at the location of the bolted joint is reduced. In this model, stiffness is reduced from its original value down to nearly 0, though this range is just for the sake of exploration and not necessarily realistic, as described in section 6.4.4. The change in frequency for the modes is shown in Figure 53. The largest change is in the frequencies for modes 1 and 2 - these are both tower modes. This change in stiffness has almost no change on the mode shapes, the MAC values for modes 1 and 2 only reduce to a minimum of 0.999. The damage index is similarly not much changed, only going as high as 0.11 for the minimum stiffness. This is investigated further by taking the lowest stiffness case and applying a change in nacelle direction, as discussed in section 4.4. The results from changing direction are very much like the undamaged case, except that the first and second mode frequencies are shifted down by the amount shown in Figure 53. As discussed in section 6.4.4 the effect of loosened bolts would not be apparent as a gradual change but rather appearing to suddenly change between ‘intact’ and ‘damaged’ depending on the load at the time and if it exceeds the clamping strength.

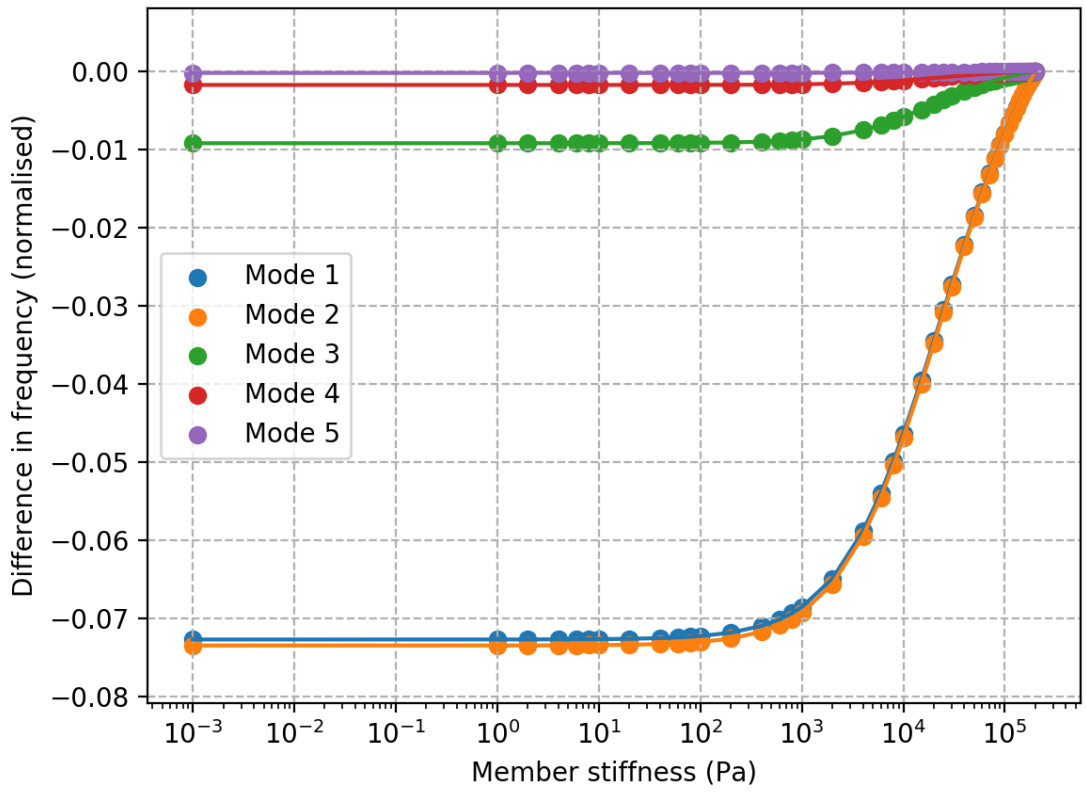


Figure 53. Change in frequency from reduced node stiffness at tower base.

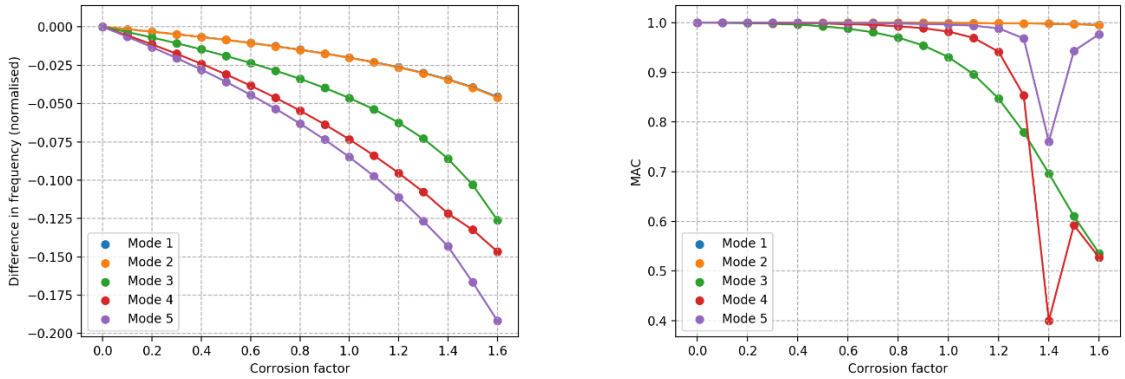
6.6.4. Corrosion and marine growth

6.6.4.1. Corrosion

Corrosion is modelled in this case as a wall thickness reduction with the mass of the corroded material retained. This has the effect of lowering the stiffness of the structure but not the mass. There is an expected corrosion profile given in section 6.4.5 and this is multiplied by the factor 'corrosion factor'. The effect from corrosion strongly resembles the effects from scour. This similarity to the effects from scour is potentially useful, both behaviors are expected to some extent and so the similarity of their effect potentially makes them easier to differentiate from other factors which are not expected. The normalized change in natural frequency and the MAC values are shown in Figure 53. The frequencies of the 4th and 5th mode become the same at around a corrosion factor of 1.4, causing the dip in MAC.

The main difference between how corrosion affects modal parameters, in this model, and how scour affects it is that the frequency changes more with corrosion; at the dip in MAC for corrosion, the frequency for the 4th and 5th modes reduced more than 12%, whereas with scour the frequencies of these same modes changed only 6% and 8% respectively. It would be difficult to distinguish between these two without a good understanding of the specific structure. There is a much higher damage index for corrosion due to the larger change in modal frequency, around 1.3

for the dip while it is only around 0.7 at the dip from scour. However, this is subjective to how this is implemented and how it is modelled.

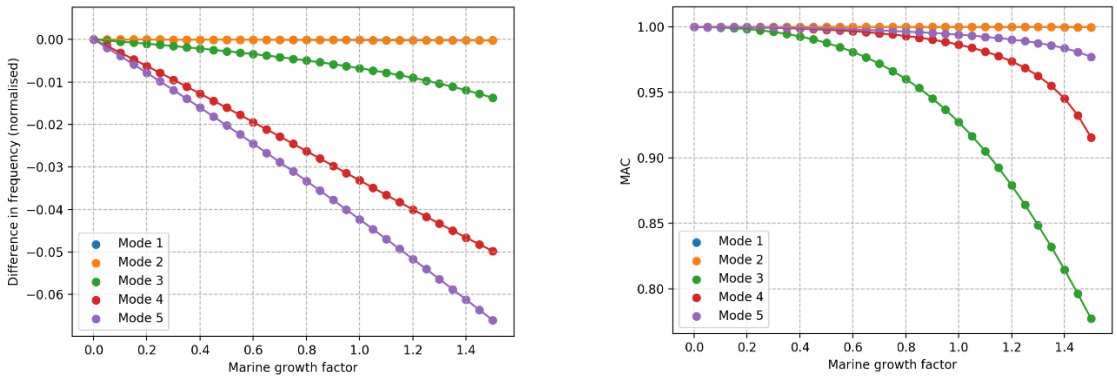


(a) (b)

Figure 54. Effect of corrosion relative to the uncorroded structure. (a) Change in frequency (b) MAC values

6.6.4.2. Marine growth

The marine growth increases the mass of the structure and increases surface roughness, which affects hydrodynamic properties and both factors are accounted for in this model. The effect of this is to decrease the eigen frequencies almost linearly with an increase in marine growth. There is no dip in the MAC values observed with other cases as the frequencies of the modes never pass each other; this observation seems to be consistent when the nacelle is rotated.



(a) (b)

Figure 55. Effect of marine growth relative to the unfouled structure. (a) Change in frequency (b) MAC values

6.7. Discussion

The results of this study reveal several findings about the behaviour of this jacket structure and about the potential for damage detection based on these measurement values. Some consideration should be made as to the applicability of these results to other cases. In terms of similar models of similar jacket structures, it should be expected that comparable results would be achieved. It should also be expected that this is a good approximation of the real system as this study was conducted using the design model to which the system was built. However, if a more complex model is used, such as one which explicitly models the blades, the results may differ due to the more complex dynamic behaviour. Measurement results from a real wind turbine may be different still, and so it would be important to understand this difference for any case of interest.

There is a consistent order in normalized change in frequency and MAC values for effects which are global on the structure, that is all factors investigated that are not a damage to a specific structural element. The order in terms of largest change to smallest is consistently 5^{th}, 4^{th}, 3^{rd}, 2^{nd}, 1^{st} for normalized change in frequency and for MAC, within moderate changes, the order is 3^{rd}, 4^{th}, 5^{th}, with negligible change in 1^{st} and 2^{nd}. This order is of course dependent on how the comparison is chosen, but the conclusion is that based on the order of frequency change or MAC values, these effects cannot be distinguished. To distinguish between these effects, this can potentially be done based on the quantity of the changes with respect to each other as well as using time history observing the development of the change.

Damages to the structural member joints have a nonsymmetric effect and therefore, the observed change in modal parameters is closely related to the nacelle direction. Because of this directionality, the side of the structure on which the damage has occurred is identifiable. The level of effect the joint damage has can also be characterized by the height of the damage and therefore it should be possible to locate joint damage with reasonable accuracy.

It is useful to calculate several metrics when trying to detect and identify damage. This gives more information about what change has occurred and can make the situation easier to understand. In calculations which require comparison of specific modes, it is useful to make this selection both with the order of the frequency and by maximizing the total MAC value, as it appears that neither approach is always the better choice. The damage metric calculated with the modal flexibility can be hard to interpret on its own as the value indicating ‘damage’ changes between damage types. For example, the damage index is 1.2 with 6 m of scour, 0.35 with a marine growth factor of 1.45 and 2 at a corrosion factor of 1.6, while just changing the direction of the nacelle the value can go up to 0.3. By itself, it can only show if there is a change in the structure, but it can be useful if combined with other information.

When the frequencies of two modes intersect, there is a sharp change in the shape of these two modes which is observed in the MAC values. When this happens with lower modes, the change is also observed in the modal flexibility.

6.8. Conclusion

This chapter presented the simulated dynamic response of an offshore wind turbine jacket structure and how it changes in response to a range of situations. These situations include: 1) operational changes, the turbine nacelle yaw 2) damage including cracks on structural joints and 3) exceeding design parameters for corrosion, scour, marine growth and tower bolt tension. This can serve as a useful reference for future work as these situations were investigated on a single structure using a highly detailed design model. The feasibility of detecting a range of damages, based on simulated damages to a numerical model, was investigated. An eigenvalue solution using a detailed design model was conducted for the sake of calculating damage metrics at these conditions and sensitivity studies to a range of parameter changes was conducted. The model is more detailed than other models typically used in similar studies, such as including the mass and moment of inertia of the nacelle and blades. The effects simulated included not only cracks on structural elements but also other types of damage such as scour, marine growth and corrosion. The metrics calculated are global metrics to reflect real-world limitations, these metrics are natural frequency changes, MAC values and a value derived from modal flexibility.

It was found the nacelle direction substantially alters the mode shapes and hence the MAC values, although the frequencies are not significantly altered. Therefore, any damage detection algorithm using mode shapes should include knowledge of the nacelle yaw direction. This also impacts how cracks in joint elements appear in terms of MAC value; the result is that at some nacelle yaw directions it may be hard to detect damage, but this potentially aids in locating the damage.

If there is a significant enough change in the modal properties caused by the damage, then the frequency of one mode can move below the frequency of another, this mode-switching is an issue encountered for semi-symmetric structures. Two methods for dealing with this were used and assessed, which includes using the Hungarian algorithm to maximize the sum of MAC values as well as using a modal stiffness-based damage metric. Both approaches were shown to be effective, although the Hungarian algorithm was only effective up to a limited level of change in some parameters.

There are several observations which would aid in the localization of cracks through changes in global dynamics. One of these observations confirms what is already known, that for cracks on elements, the severity on the structural dynamics is a function of height in the structure; lower elements change mode shape and frequency more than higher ones. This combined with the directionality of the nacelle can make localization of the damaged member joint possible using global metrics.

For non-joint effects which are symmetric on the structure such as global scour or corrosion, the change in frequencies and the MAC values follow a consistent order, given moderate levels of damage. This can make distinguishing these types of damage difficult.

Chapter 7

7. Sensor placement impact on modal extraction

7.1. Introduction

The way in which a structure vibrates, in the eigenvectors, frequency and other parameters, can give a great deal of information about the structure. The source loading to excite the structure can be either deliberately applied, as in experimental modal analysis, or in the case of operational modal analysis, the result of natural loading such as wind and wave [204].

By measuring the modal parameters after a structure is built, models of the structure can be updated so that they more closely resemble the real structure. This updated model can then be used for purposes such as to estimate fatigue life. Measurement of the modal parameters can also be useful in offshore wind for the purpose of damage detection, where a change in parameters indicates a change in the structure [226,227].

Accelerometer measurements can be processed with a variety of methods to determine modal parameters from simple approaches such as peak picking, where a significant peak in the frequency-response-function corresponds to one mode, to more advanced methods such as Stochastic Subspace Identification [454,455]. Once extracted, the structure's modal parameters can be evaluated based on Eigenfrequency, mode shape and damping. The mode shape can be represented as a matrix of node displacements and so a comparison of mode shapes requires use of a statistical parameter. A widely used tool is the modal-assurance criterion (MAC) [436,456]. This parameter gives the level of consistency between two vectors. This is not always ideally suited to a problem and so there are other, similar parameters, such as sensor elimination algorithm using modal assurance criterion (SEAMAC) and others.

For any of these aforementioned purposes for using the modal parameters, the parameters need to accurately and reliably be measured. However, measurement and the process for determining parameters from the measurement is not straight forward when it comes to the practical implementation. The sensors need to be placed at locations where the natural vibration of the structure leads to measurable displacements, but there are practical limitations to sensor placement. Installing sensors offshore and maintaining them can be expensive, so it is desired to not only limit the number of sensors but also install them at locations which are easily accessible.

When determining sensor placement, an optimal sensor placement study can be conducted where the sensors are placed at locations where they can optimally capture the

structure's natural mode shapes. This is a numerical study which can employ an optimization algorithm to best place sensors. The optimized layout is where a limited number of sensors are placed in a way that the diagonal elements of a MAC matrix are maximum, without having large off diagonal.

This chapter is unique in that accelerometer data is analysed on a full scale and operational, offshore substation which allows for the comparison of multiple optimized sensor layouts as well as non-optimized layouts on the same structure during the same time period. This allows for extraction of modal parameters for each of the sensor layouts which can then be compared to the modal parameters predicted by the FEA model from which the layouts were optimized. Additionally, the observed modal parameters can be clustered and analysed in a theoretical real-world case where the consequence of choosing each layout can be shown.

This chapter begins in section 1 by showing the structure analysed as well as the first five mode shapes predicted from the finite element analysis (FEA). Background to the optimal sensor placement is then given in section 2, followed by the resulting sensor layouts. The methodology and data processing are given in section 3. The results are given in section 4, first the results of clustering, followed by comparison to the FEA model and then to the cluster averages. A discussion on the practical implications is given in section 5 and finally a conclusion is given.

7.2. Reference structure

The structure measured in this study is an offshore sub-station (OSS) located in the Wikinger wind farm, owned by Iberdrola. The main purpose of wind farm OSS structures is to collect the low voltage power from the turbines and transform it to a higher voltage to be exported back to land. The OSS at the Wikinger site is built on a jacket structure which is fixed to the sea-bed through six pile foundations. This is a large structure, relative to the turbines, and is equipped with a crane, a helicopter landing-pad and a considerable amount of electrical equipment such as transformer, switchgear and shunt reactor.

The main source of excitation of the structure is wave loading. Although the structure doesn't have a wind turbine, there is still wind loading on the structure. Throughout this study it is assumed that the structure is stationary, it is assumed that the mass and stiffness matrix are not time-variant, however it is possible that this is not completely true as components, related to power conversion, may have been added or removed in the course of the study.

The FEA model of the structure is illustrated in Figure 56 and the first five mode shapes from the FEA model are illustrated in Figure 57.

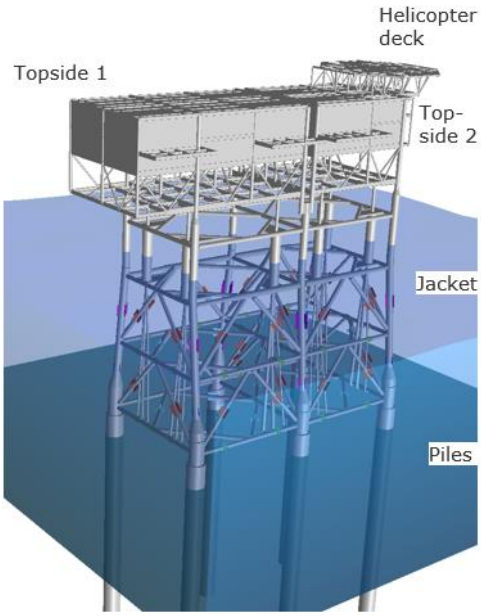
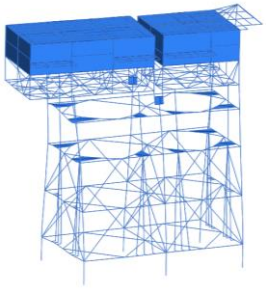
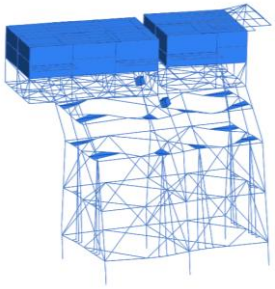


Figure 56. FEA model of the Wikinger OSS

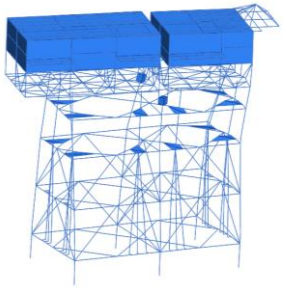
A natural frequency analysis of the OSS structure is performed, and the first five natural frequencies and mode shapes are shown in Figure 57. These will be compared to the operational modal parameters of modal vectors and eigenfrequencies.



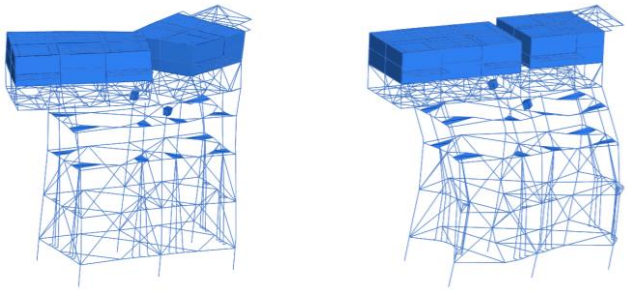
1st mode – 0.434 Hz



2nd mode – 0.492 Hz



3rd mode – 0.508 Hz



4th mode – 1.784 Hz 5th mode – 1.802 Hz

Figure 57. FEA model mode shape illustration

7.3. Sensor placement

7.3.1. Optimal sensor placement study

The objective of the optimal sensor placement study is to assess the optimal number and locations for mounting accelerometers on the structure. It is based on a large choice of criteria. The aim is for a sensor layout with the capability of extracting linearly independent mode shapes up to a defined number of harmonics. This should be achieved by utilizing minimum amount of hardware (sensors, cabling and data acquisition system) to save costs while keeping high-quality data analysis levels and considering sensor redundancy. Mounting locations must be regarded as acceptable in terms of installation risks, maintainability and exposure to potentially harming environment.

For the OSS structure a SEAMAC algorithm – sensor elimination algorithm using modal assurance criterion – is chosen. It tests the removal of DoFs for each candidate sensor location to find the setup with minimum off-diagonal terms in the MAC matrix. This ensures the highest linear independence of the specified mode shapes. The sensor locations have been limited to the topside of the OSS only as sensors in the splash zone and below water are considered unfeasible due to maintainability.

7.3.2. Sensor layouts

Three sensor layouts are developed according to the criteria listed in Table 22. They aim to extract the first five mode shapes of the OSS. For comparability two non-optimized sensor layouts are considered and outlined in

Table 23.

The number of sensor types, total sensors, maximum MAC off-diagonal value and total MAC off-diagonal values are summarized in Table 24. The off-diagonal MAC properties give an interpretation of how linearly independent mode shapes are extracted. Low values are favourable. The Bundesamt für Seeschifffahrt und Hydrographie (BSH) govern offshore activity for Germany [433] and their relevant standards are given in [457].

Table 22. Optimized sensor layout definitions.

Layout	Definition
1a	Aims to balance number of sensors and small off-diagonal MAC values. Utilizes sensors from the permanent sensor setup. Additional sensor locations on the jacket legs above the splash zone are allowed.
1b	Aims to balance number of sensors and small off-diagonal MAC values. Utilizes sensors from the permanent sensor setup.
2a	Aims to reduce the number of sensors to a minimum while keeping off-diagonal MAC values around 25%. Sensor locations on the jacket legs above the splash zone are allowed.
2b	Aims to reduce the number of sensors to a minimum while keeping off-diagonal MAC values around 25%.

Table 23. Non-optimized sensor definitions.

Layout	Definition
Complete	Combination of permanent sensors and temporary sensors.
Permanent	Utilizes the sensor setup that was established in accordance with the BSH monitoring standards.

Table 24. Number of sensors and MAC properties per layout.

Sensor DoF	Permanent	Complete	Setup 1a	Setup 1b	Setup 2a	Setup 2b
3D	6	12	1	0	3	1
2D	0	0	2	3	0	0
1D	0	0	4	4	0	4
Total	6	12	7	7	3	5
Max. off-diagonal	70%	62%	11%	18%	27%	33%

Total off-diagonal	184%	153%	72%	111%	146%	166%

The sensor locations of all layouts on the OSS topside are shown in Figure 58.

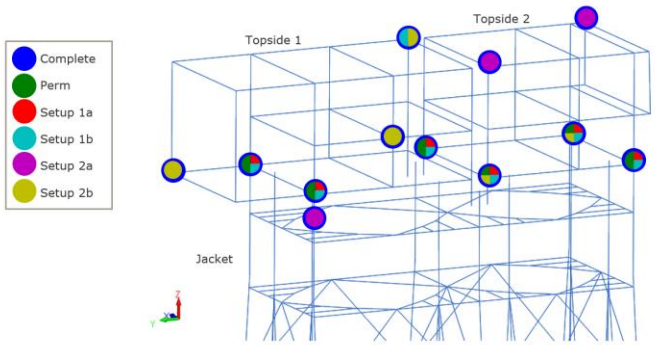


Figure 58. Sensor locations according to specified sensor layouts.

7.4. Methodology

7.4.1. Initial data collection and processing

The data present for this work consists of a permanent and a temporary measurement campaign. Between these two sets of sensors, there is one month of overlap running from the middle of November to the middle of December 2018. Multiple types of sensors are placed at the nodes measured. The sensors used for this study are 3-axis accelerometers which record accelerations in the *X*, *Y* and *Z* axis over time.

These time series signals were divided into 30-minute intervals. Prior experience with these structures has shown that 30 minutes is the ideal amount of time to extract modal parameters. The two configurations, the temporary sensor configuration and the permanent configuration, are then combined at this stage. An ARTeMIS batch process was then used to process all the resulting samples. ARTeMIS is a commercial modal analysis tool. Each of the sensor layouts shown in Figure 58 can be constructed be omitting sensors from the ARTeMIS analysis. Within ARTeMIS the SSI-UPC (Stochastic Subspace Identification – Unweighted Principle Component) was used to extract the modal parameters. The result of this is that for each of the 30-minute samples, and for each of the possible sensor layouts, the modal parameters are extracted. For each mode captured, the eigenfrequency, damping and modal displacements are determined.

7.4.2. Data processing

The main purpose of this study is to determine the ability to detect the modal parameters between optimized and non-optimized sensor layouts. To this end, it is necessary to determine the level of consistency between the mode displacement vectors observed in a given sample and some

reference mode vectors. The statistical tool used for comparing mode vectors in this work is the Modal Assurance Criterion (MAC) [458], this is a scalar constant which is a measure of the level of consistency between two vectors. A MAC value of 0 indicates that the modes are inconsistent and a value of 1 indicates that they are totally consistent [456]. The equation for comparing vector *c* to the reference vector *d* is given, for degree of freedom *q* and mode *r*, as:

$$\mathrm{MAC} = \frac{\left|\sum_{q=1}^{N_0} \Psi_{cqr} \Psi^*_{dqr}\right|^2}{\sum_{q=1}^{N_0} \Psi_{cqr} \Psi^*_{cqr} \sum_{q=1}^{N_0} \Psi_{dqr} \Psi^*_{dqr}} \quad (51)$$

Where the * indicates the complex conjugate of the vector Ψ with modal coefficient reference *p*, degree of freedom *q*, mode *r*. N_0 is the number of outputs. This relatively simple statistical tool is highly used in operational modal analysis for both model updating and damage detection [445,459]. However, there are limitations to this tool, the two to be most aware of for this research are: 1) Due to the least squares fitting approach used, the value is highly influenced by outlier values. 2) The prediction becomes more affected by slight errors if there are only a few degrees of freedom used (between 2 and 5) [456]. These will be discussed more in the results section. In all calculations of MAC, if a degree of freedom is not present then it is not part of the MAC calculation to avoid zero padding.

7.4.3. Clustering

For the sake of comparison, a reference set of averaged modes needs to be determined. Over 1400 30-minute samples were evaluated for the 6 layouts, each of which can have more than one mode extracted. This provides a significant number of observations. Based on the FEA model of the structure we obtain that there are a certain number of distinct modes, each of which are at a frequency which is noticeably different from each other. This is also observed in the operational data, that the eigenfrequencies of modes are far enough apart from each other that, despite scatter, they form distinct, observable clusters. The MAC value is used to determine if a mode vector is consistent with another, however, there needs to be some reference vector with which the observed vector's consistency can be judged. The approach used for clustering can be assessed given the hypothetical scenario where no FEA model is available and we believe the structure has not fundamentally changed in the observation period. Then to determine how the mode vector changes relative to some reference vector derived from those observations, it can be enough to use a clustering algorithm based on eigenfrequency alone. In this way the value being investigated, variation in mode vector, is not included as part of creating the reference, but the reference is still meaningful. Mode vector could be included in the clustering approach, but that is made challenging by the fact that some degrees of freedom are not available in some observations.

In this chapter, a k-means clustering approach is used, implamenting the scikit-learn Python library [460]. The limitation to this is that the number of clusters needs to be manually set. In this chapter, the predominant interest is the first three modes which have eigenfrequencies bellow 0.7, this is seen both in the reference FEA model as well as the operational data. To set the number of clusters, the minimum number to form three clusters below 0.7 is used. The

frequencies of these three modes are close together compared to the higher modes and so this approach also distinguishes higher modes with good fidelity.

7.4.4. Comparison approaches

The two key questions this chapter aims to answer are 1) how do operational measurements vary relative to the FEA model from which optimal sensor placement was conducted? 2) Given a set of sensor layouts, how does the choice of layout influence the consistency of modal parameter extraction?

Towards answering the first question, each operational mode is compared to the FEA model. The MAC value is used to identify which FEA mode each observed mode most resembles.

To answer the second question, a baseline set of operational modes are established. All the observations are first clustered based on their eigenfrequencies and then the eigenvectors within a cluster are averaged together to create an average mode shape for that cluster. Each operational observation is then compared to this 'baseline' mode to investigate level of variation resulting from the layout choice.

A factor influencing the ability to extract modes appears to be sensor availability and number of sensors. The permanent campaign sensors are reliable with no loss of DoFs in the observed period, however the temporary campaign is not as reliable and there is frequent loss of DoFs. This can influence the MAC value as fewer DoFs lead to the result being more influenced by outliers.

7.5. Results

7.5.1. Resulting eigenfrequencies of observed modes

Figure 59 and Figure 60 show scatter plots of the observed eigenfrequencies for each sample for all layouts. Each layout is represented by a different shape and colour. The first three mode frequencies are relatively close together and so the second figure shows the same as the first, except the y-axis is limited to between 0.5 and 0.7 Hz.

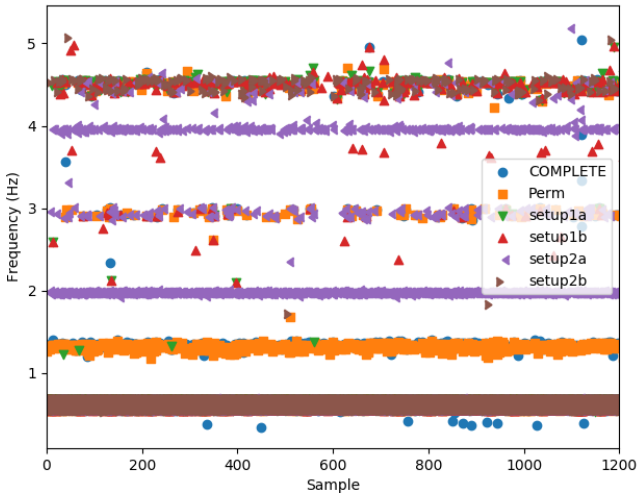


Figure 59. Scatter plot of observed operational eigenfrequencies for each observation sample and all layouts.

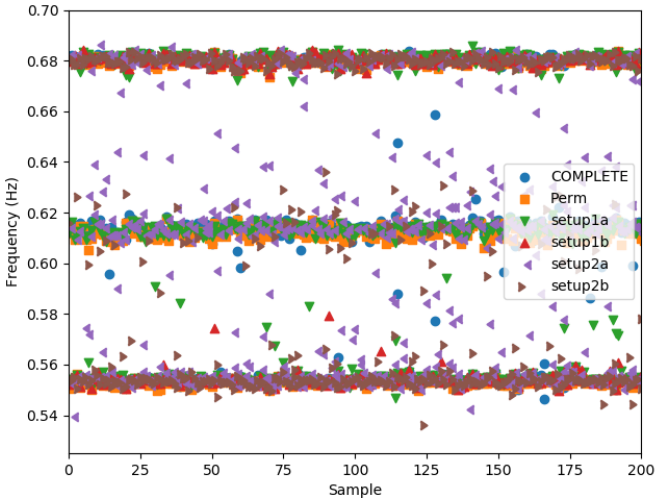


Figure 60. Scatter plot of all observed operational eigenfrequencies with the y-axis limited to between 0.5 and 0.7 Hz and x-axis limited to the first 200 samples.

There are two points to these figures, the first is that the scatter plots show distinct clusters by frequency, justifying the clustering approach presented in section 3.3. The second observation is that even only in terms of eigenfrequencies, the choice of sensor layout is important in order to obtain consistent values – although it will be shown that this is influenced by sensor availability. Additionally, and this will be discussed further in subsequent sections, the sensor layout also affects whether an eigenmode is identified or not.

Based on the clustering approach, the mean and standard deviations of the operational clusters are presented in Table 25. For reference, the FEA eigenfrequencies are: 0.434, 0.492, 0.508, 1.784 and 1.802 Hz.

Table 25. Mean, μ, and standard deviation, σ, of mode observation clusters each of the sensor layouts.

Complete		**Permanent**		**Setup 1a**	
μ (Hz)	***σ (Hz)***	***μ (Hz)***	***σ (Hz)***	***μ (Hz)***	***σ (Hz)***
0.553	0.0018	0.553	0.0015	0.555	0.0031
0.614	0.0041	0.612	0.0016	0.613	0.0043
0.681	0.0153	0.680	0.0024	0.683	0.0357
1.337	0.0288	1.318	0.0369	-	-
2.952	0.0937	2.935	0.0499	2.747	0.0722
-	-	-	-	-	-
4.500	0.1281	4.499	0.0627	4.504	0.3370
Setup 1b		**Setup 2a**		**Setup 2b**	
μ (Hz)	***σ (Hz)***	***μ (Hz)***	***σ (Hz)***	***μ (Hz)***	***σ (Hz)***
0.554	0.0016	0.557	0.0066	0.554	0.0018
-	-	0.617	0.0071	0.613	0.0041
0.681	0.0027	0.679	0.0098	0.680	0.0108
-	-	1.976	0.0147	1.773	0.0316
3.026	0.1072	2.943	0.0262	-	-
-	-	3.961	0.0510	4.424	0.0548
4.503	0.4229	4.476	0.1275	4.562	0.071

For all layouts, the first three eigenfrequencies are 0.55, 0.61 and 0.68 Hz, except for ‘setup 1b’ which does not find the 0.61 Hz mode. ‘Setup 2a’ locates more modes, specifically at frequencies 1.98 and 3.96 Hz, which are not captured in any other layout but is similar in terms of frequency and MAC to the 4th and 5th FEA modes. The ‘Complete’ and ‘Permanent’ layouts locate a mode at 1.34 Hz and 2.95 Hz. For reference, the FEA model Eigenfrequencies were 0.434 Hz, 0.492 Hz, 0.508 Hz, 1.784 Hz and 1.08 Hz.

7.5.2. Results compared to structural model

The purpose of comparing the operational observations to the FEA model is to investigate the influence of the off-diagonal elements in the MAC matrix, which was generated from optimal sensor placement study, on the ability to extract modes. To accomplish this, it is first necessary to relate the observed modes to the model modes. This is so that the high MAC values in the optimal sensor study can be tied to specific observations.

Comparison of the observations to a specific model mode can be done through Eigenfrequency of eigenvectors, the consistency of which is determined through the MAC. To get an initial overview, these two quantities are shown in box and whisker plots in Figure 61 and Figure 62, where the circles are individual observations. Typically, the lower the mode number, the higher the MAC number. This is reasonable since the lower modes are first-order modes and are less influenced by the complexity of the top-side structure. An issue of comparing only to the first five FEA modes is that higher order modes are sometimes observed in the sensor data and therefore are erroneously compared to lower modes. This also shows that, in terms of eigenvector, the mode found by setup 2a at roughly 2 Hz most closely resembles the FEA, torsional modes 4 and 5, which are also around that frequency.

As shown in Figure 62, the more optimized layouts, setup 2a and 2b, have fewer low MAC values for the first three modes, than the permanent layout. This is because the sensors were placed with the objective of best capturing mode displacements which the model shows as indicative of that mode, these are high displacement DoFs from the model. Therefore, the MAC value calculation for these layouts will only use DoFs which were determined to be important in the model, and so will be more significant in the MAC calculation.

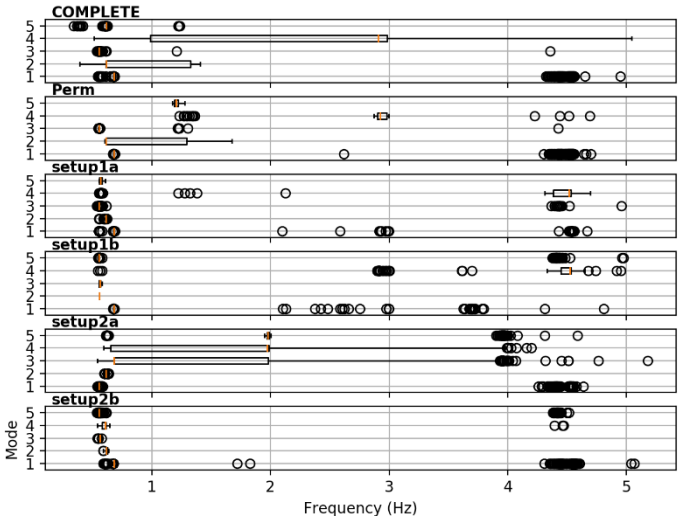


Figure 61. Box and whisker plot of measured eigenfrequency of each observed mode, for each sensor layout, grouped by which FEA mode it most closely matches (based on highest MAC).

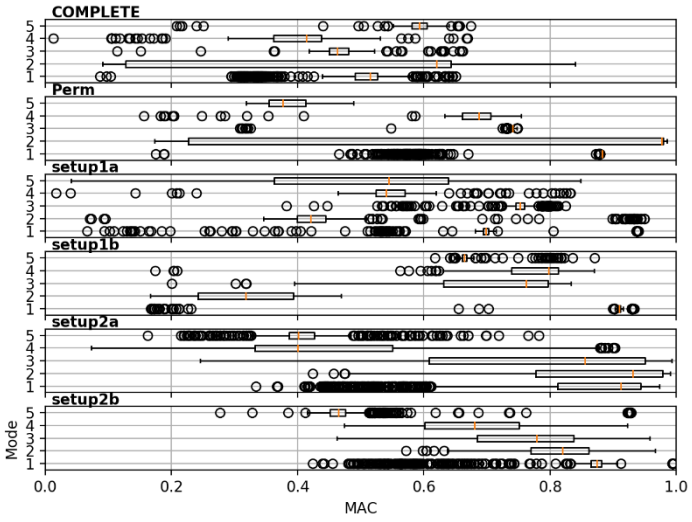


Figure 62. Box and whisker plot of highest MAC value of each observed mode, for each sensor layout, grouped by which FEA mode it most closely matches (based on highest MAC).

To show this in more detail, Figure 63 shows histograms number of observations within each measurement cluster for each sensor layout. The histogram is coloured by which FEA model eigenvectors the observation is most consistent with (highest MAC). Each subplot has a title showing which layout as well as the frequency mean values of each cluster. The frequency of each numbered mode is given under the layout name and the number corresponds to the clusters given in Table 25.

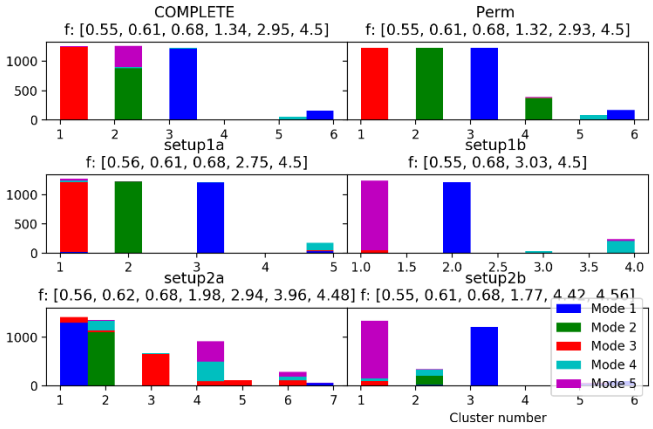


Figure 63. Histogram of number of observations in a cluster for each layout, coloured by the FEA mode the observation is most consistent.

From these histograms, it can be seen that typically the order of observed mode to model mode is reversed for the first three modes, which is that mode 1 of the model matches mode 3 of the observations. The mode which is captured by setup 2a at roughly 2 Hz matches with the fourth and fifth model modes.

7.5.3. Results compared to cluster averages

For each cluster of modes for a given sensor layout, an averaged mode vector is calculated, and this acts as a reference mode vector to which each observation is then compared.

Figure 64 and Figure 65 show the frequency and MAC value respectively of all observations and layouts compared to the cluster averaged reference vector with which it has the highest match. This shows that the MAC values are typically much higher for each observation than compared to the FEA model, although this is to be expected.

For some 30-minute samples, the SSI process detects two modes which are at very similar Eigenfrequencies. It is assumed that one of these is anomalous and would not be useful to include in damage detection or model updating. To deal with this, if two similar modes are detected in the same sample, only the mode with the Eigenfrequency closest to the cluster average is used to form part of the cluster eigenvectors average. This criterion is also used when calculating the MAC value of each observation relative to the cluster average. This results in a reduction in the amount of low MAC values which otherwise indicate a greater level of scatter than is present.

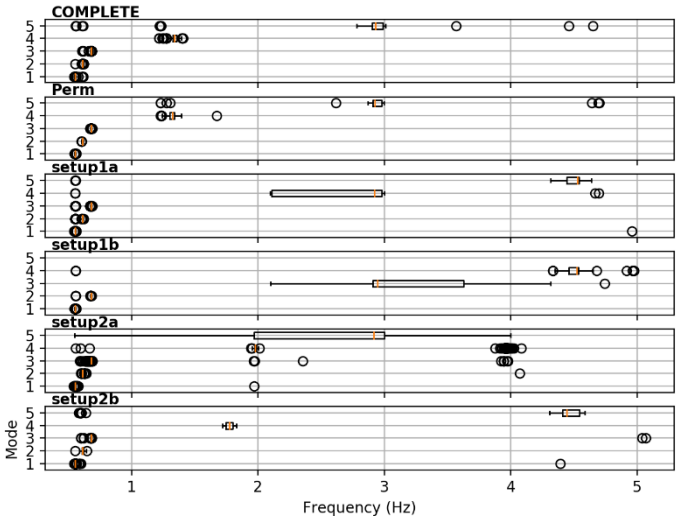


Figure 64. Box and whisker plot of measured eigenfrequency of each observed mode, for each sensor layout, grouped by which cluster mode it most closely matches (based on highest MAC).

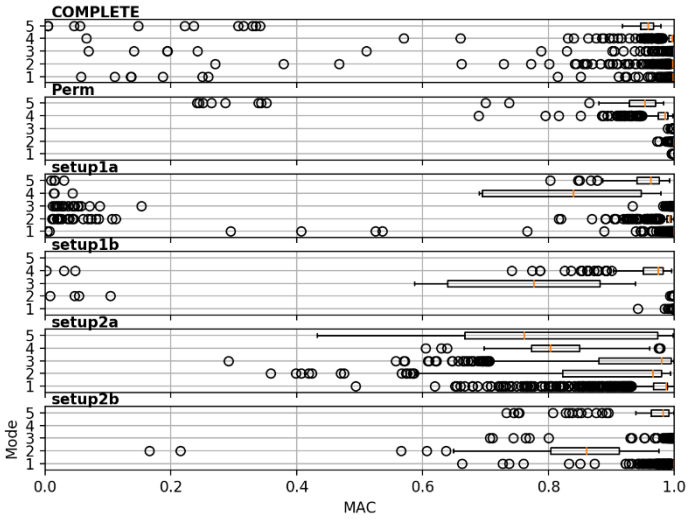


Figure 65. Box and whisker plot of highest MAC value of each observed mode, for each sensor layout, grouped by which cluster mode it most closely matches (based on highest MAC).

The decision process for if a mode should be compared to a given reference mode is based on a thought process which might be followed in an operational situation and so these are treated as new observations. If these were new observations and not observations already in a cluster, then deciding which cluster it belongs to should be done first by which cluster the new observation has the highest MAC value with and second by whether the frequency is relatively close. That is the process followed here in order to replicate an operational condition.

Figure 65 shows the consistency of the mode vector for each layout. For the permanent layout the MAC is always very near one, meaning the mode shape derived is always very consistent with the reference. This is good for using the operational data. For model updating, only a few observations are needed and for damage detection there is little scatter in the healthy state. Setup 2a however shows recordings which are typically inconsistent with the

reference. Even though setup 2a can detect modes which the permanent layout cannot detect. The permanent sensor layout and setup 2a do not share any sensors.

Setup 2a shows relatively low availability. This layout consists of only 3 sensors, so there are only 9 degrees of freedom if everything is working correctly. To show the level of availability of each layout, the number of degrees of freedom in each observation is shown per layout. Despite this, setup 2a is still able to distinguish the largest number of modes by cluster.

Table 26. Maximum, minimum and mode of number of degrees of freedom for the observations for each sensor layout.

Sensor layout	Maximum	Minimum	Mode
Complete	33	27	31
Permanent	18	18	18
Setup 1a	11	8	11
Setup 1b	10	9	10
Setup 2a	9	3	7
Setup 2b	6	3	6

Table 26 shows the statistics for number of degrees of freedom available for each 30-minute observation over for each sensor layout. The permanent sensor layout has 100% availability with 18 DoFs throughout the entire time period. The temporary campaign however is less reliable with the occasional loss of a DoF as well as a sensor (with 3 DoFs) becoming unavailable after the 21st of November. One might conclude from this that the MAC value is dependent on the number of sensor DoFs available. However, plotting MAC value against number of degrees of freedom shows this to not be the case. An implication of this when designing a sensor layout is that more degrees of freedom from more sensors doesn't necessarily lead to a more consistent extraction of modal parameters.

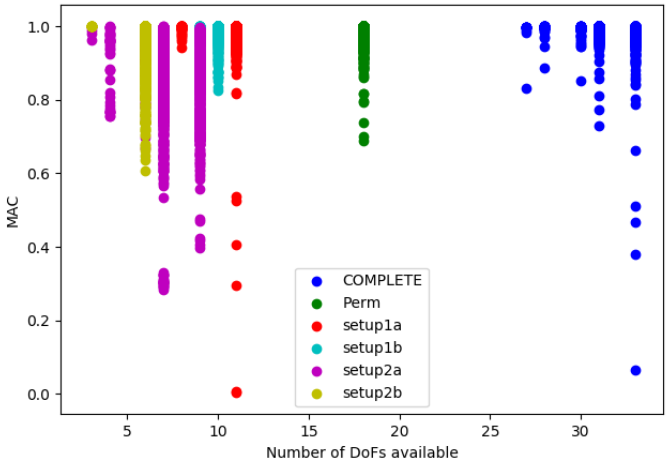


Figure 66. Scatter plot showing MAC value against number of DoFs available. Coloured by sensor layout.

7.6. Discussion

Although the permanent sensor layout, constructed through engineering judgement, resulted in very reliable modal extraction, there is potential to improve upon this. There are clear practical considerations for these results on why an optimal placement sensor study should be conducted. For a single structure like this OSS, sensors are relatively inexpensive. However, for the case of wind turbine structures, where there can be tens or even hundreds in a wind farm, sensors for all structures are expensive to procure, install and maintain. More sensors also lead to more complexity in signal processing and storage. If fewer sensors can give equal or better results, as this chapter has shown, then that saves money. The results can be better, as shown in this chapter, the most optimized layout was able to detect more modes than the non-optimal setup. Damage sensitive modes can also be targeted through conducting an optimal sensor study. Furthermore, redundancy can be designed deliberately through an optimal layout if this is required. Knowledge of installation requirements cannot be overlooked as the sensors need to be installed in accessible and solidly mounted locations. There may be resistance to conduct an optimization as it is a technical challenge which requires a detailed FEA model, however it requires only the application of an algorithm.

7.7. Conclusion

In this chapter the influence of accelerometer sensor placement on the ability to extract modal parameters was investigated. Operational readings covering 1 month from 12 accelerometers placed on an OSS platform from the Wikinger wind farm (owned by Iberdrola) were used. The SSI-UPC method in ARTeMIS modal was used to extract modal parameters. Observed modes are clustered using a k-means approach and then compared to both the original FEA model as well as a cluster averaged eigenvector.

These results show that while relatively high MAC values between the observation and the FEA model can be obtained, there is still a significant enough difference between the FEA model and the observed modal parameters to draw any conclusions on the impact of off diagonal elements in the MAC matrix from optimal sensor placement for the first 5 modes.

Comparing the modal parameters of each observation to a cluster-averaged vector shows level of fluctuation in the observed mode for each sensor layout. While the most optimized sensor layout (setup 2a) is able to detect the most modes, it shows the greatest level of fluctuation in both eigenfrequency and mode vector. A contributing factor to this appears to be sensor availability which is less high for the temporary campaign (which includes setup 2a) than the permanent.

The recommendation from this is that sensors should be placed with knowledge of the mode vectors (optimal sensor placement) and sensor reliability is critical to obtaining consistent results.

Chapter 8

8. Discussion

The goal of this thesis is to advance areas required for digital twin technology. This is inherently a multidisciplinary task which this thesis has approached in a two-part approach viewing both load and resistance. These two parts are: Part 1) Load variation at an array level, Part 2) structural variation and data. The parts within each section all show important findings and considerations.

8.1 Literature

The literature review conducted as part of this research highlighted models that can be used and highlighted gaps in the research which can be filled. It provided an overview of topics in modelling wind, wave and geotechnical variation within a wind farm, as well as consolidating a list of applicable models which can be used.

It was identified that there are a wide range of computational fluid dynamics models and parts of models which could be used to model the wave within a wind farm. These models all provide different levels of fidelity and capture behaviour in different areas. As computational capabilities that researchers and engineers have access to advances, more complicated models become suitable for tasks which they previously were not suitable for. Additionally, in some cases the simplifications which had been made to the models for performance reasons can be removed. The literature review also highlighted the growing role of machine learning in engineering applications, particularly in predicting wind and wave time-series applications.

The literature review highlighted that there is not much existing research on load variation within wind farms resulting from the impact of other structures on the waves. However, there is a plethora of research on wave and current models which are applicable, this can be particularly useful in sites with significant depth variations.

When conducting wind farm aerodynamic wake models, it is important to answer 5 key questions to identify the models most suitable for the task: 1) What level of fidelity is required? 2) How quickly are the results needed? This should be considered in conjunction with available computational capability. 3) Are time domain results required or are steady state, time averaged results sufficient? 4) What is the area of interest? Is it near or far wake? 5) What level of complexity is required for the turbine model? This question can also be combined with what information on the turbine is available since many require detailed aerodynamic information, but some do not. Answering all these questions is a vital first step to any turbine wake modelling.

Waves and wave loading in normal conditions can be modelled suitably with existing models. There exist many wave models of various orders to represent sea level as well as statistical models for sea state which are suitable for different regions. Loading can be modelled through

analytical means such as Morrison's equation or more complex loading can be modelled numerically.

Geotechnical variation within a wind farm results from two things: either there is spatial variation between turbines, for which a ground model can be built through consolidating numerous data sources. The second source of variation is time-varying depth caused by scour, which can be either a global reduction or local around the structure. Scour can be roughly predicted analytically based on soil type but can be modelled much more accurately by considering fluid flow erosion as well as soil transport and other factors.

Part 1

8.2 Computational fluid dynamics

The computational fluid dynamics chapter in this thesis forms the foundation of the load variation investigation in this thesis. It is used for further investigation in other chapters but also provides novel and useful findings.

An important part of modelling is the quantification of error, and that was done at multiple stages in conducting the computational fluid dynamics model. A grid sensitivity study was conducted, and it was found that the results of the grid used are roughly 2.2% from the continuum value. This grid was used, rather than being more refined, due to computational limitations.

Turbulence models were investigated, comparing some of the most commonly used Reynolds Averaged Navier Stokes models. It was found that in this application k-ε gives the closest results to the operational values with a mean absolute error of around 8%, compared to 10% for k-ω SST. This is in general agreement with literature which shows k-ε to be better for this application.

A range of inflow directions and speeds were considered to consider multiple different behaviours which can be observed in turbine wakes. While k-ε typically performs better, there are some cases when the lower velocity in k-ω SST, resulting from slower wake recovery, results in closer predictions.

There is a large amount of scatter in the observed measurements and so comparing to individual observations was unsuitable. Instead, two approaches were compared which allow comparison to an averaged observation. These two approaches are binning and regression. In the binning approach the CFD results are compared to averages within a range while in the regression approach the data is approximated and the CFD is compared to the approximation. While binning generally results in a lower error observed, the regression approach gives more information about the response within the wind farm and avoids the potential of being near the edge of a bin. Results ate higher speed show higher errors, but this region in the data is also less reliable as there are fewer observations at higher speeds.

There is a reduction in velocity everywhere in the wind farm, even when the met mast is not directly in line with a turbine wake. This is observed in the operational data as well as in the CFD results. The size of the wake itself is more apparent in the turbulence viscosity plots

which show the wake to be broad, at the turbine the turbulence change is wider than the turbine itself.

These results were interesting as they show in detail where sources of error occur and since the model is compared to operational data, which is rarely available to researchers. This also forms the training data set for further study in this thesis.

8.3 Machine learning expansion

Several machine learning approaches were applied to the CFD data with the intention of expanding on the applicability of that CFD data. This includes a range of models which work in fundamentally different ways to identify behaviours which are suitable or not suitable for this task. These models were then compared to further CFD results as well as the operational data to validate the approach.

Compared to further CFD cases, the machine learning models were all within 3% mean absolute error, with the smallest error of 1% resulting from artificial neural networks. This validates this idea of the approach in that the machine learning models can be used instead of conducting further CFD analysis. The approach could then either be used with existing data to predict the loads each turbine has experienced or with a time-series forecasting approach to predict future loads.

While support vector machines provided poor predictions relative to further CFD cases, it provided the best results when compared to met-mast observations. This was surprising, but it was because the scatter in the observed site data causes a smearing of the result and a reduction in the sharpness of dips due to direct wakes. Future models should account for this stochastic nature deliberately rather than with a model with a more general fit. Otherwise, ANN provided the best results compared to MM data.

Random forest showed an inability to generate smooth predictions between training data which are located relatively far from each other, as was the case for wind speed. The use of ‘artificial samples’ where polynomial models were used in lower dimensional training sets to fill the gap resulted in a significant improvement in the RF model predictions such that their error metrics were in-line with other approaches.

A small ‘trick’ was found to improve the performance of all models – this was to include the 0 m/s inflow wind speed values in the training data. These are trivial cases when CFD would not be conducted since the wind speed at the met mast is simply 0, but including them greatly improves the machine learning performance since low wind speed values (less than 5 m/s) are no longer outside of the range of modelled values.

From a practical perspective this approach can result in comparable results to conducting more CFD cases without the corresponding computational expense. This allows for analysis involving the lifetime of the structure or Monte-Carlo analysis. However, it does not totally avoid significant computational expense since the training data needs to be generated, which in this case required one month on a high-performance computer cluster. This approach can provide values in cases where the layout is already established but modelling would provide some value. For example, this could be useful when assessing different control strategies or in repowering.

8.4 Wind and wave time series forecasting

Three different time series forecasting methods were compared to a measured 3-hour data set and were evaluated for their numerical similarity to the observed values. These three methods include Markov-Chains, Gradient Boosting and a novel method involving a combination of deterministic fitting of seasonality with random sampling, referred to as the hybrid method.

It is unreasonable to compare to individual values as there is so much randomness in each observation, instead comparison is made to metrics for monthly values. In general, all three models produce usable predictions, but some are better suited to the task. Comparing difference from the mean, Markov-Chains and the hybrid approach are both very close, while Gradient Boosting has the largest error from the means. In terms of the size of standard deviation of the predictions within a month, again, both Markov-Chains and the hybrid approach perform well while the gradient boosting approach shows the least amount of deviation. It is desirable to have a have larger deviation since that is what is observed in the real values, but GB has too much averaging and no inherent randomness in the model. It is very important to model correlation between wind and wave values because, for a lot of application, both values are important. For example, correlation is needed for vessel weather windows or predicting contemporaneous loads. Both Markov-Chains and the hybrid method (when used with multivariate distributions) gave levels of correlation comparable to the observed value (0.8), but gradient boosting did not show as much correlation (0.6).

Based on monthly metrics, both Markov-Chains and the hybrid approach provide similar values, however there are still important differences between them in terms of the forecasts that they produce. They are different in terms of the level of persistence in the forecast (how long high wind or waves stay high), which is changeable in the hybrid approach but is inherent and unchangeable in first-order Markov-Chains.

The most significant difference between Markov-Chains and the hybrid approach is that of seasonality. Markov-Chains model seasonality by having a separate model for each month. Within a month, there is no seasonality and the trend of the forecast is relatively free to wander up and down. In the hybrid approach, seasonality is modelled at several levels by the underlying series of polynomial models. This results in a forecast which follows trends shorter than a month and so follows the observed data in a more deliberate way.

This can be used to provide more accurate inputs for calculation of loading for structural models, or to better predict the ability to conduct maintenance on the structures.

8.5 Part 1 discussion

The first part of the thesis investigated aspects related to loading within a wind farm. Although no load values were calculated, the research conducted incorporates the stochastic nature of how wind behaves in a wind farm as well as the forecasting of time series to input into the models. This forms the basis for the different loading within an offshore wind farm and so allows for a more accurate, turbine specific load analysis. A potential use case could be to 1) first model the wind farm through computational fluid dynamics, then 2) train a machine learning model on the results 3) to forecast freestream wind speeds using the hybrid approach and 4) use the machine learning model to predict wind speeds at each turbine based on the

forecast freestream values over the life of the structure. For making predictions with a digital twin model, this approach would provide very accurate estimates for the wind values to use in a load analysis for each structure.

Part 2

8.6 Damage detection

An investigation of damage detection through global modal analysis was conducted using a detailed design model of the Wikinger wind turbine structure, which uses a four-legged jacket. This model is a very close analogue of the real structure as it is very detailed and is the model that the real structure was designed from. There is some limitation that the model provides an eigenvalue solution and does not model the time series with blade rotation, but for the sake of the current work the model is sufficient.

There were some consistent patterns observed in the rate at which different modes change frequency. For effects which are global and symmetric, such as scour or corrosion, the order from most to least change is 5^{th}, 4^{th}, 3^{rd}, 2^{nd}, 1^{st}. When these are normalised by the starting frequency of each mode, the order becomes 3^{rd}, 4^{th}, 5^{th} with the 1^{st} and 2^{nd} having negligible change.

Changes which are non-symmetric, such as damage to a single element, have unique behaviour. The side of a structure that a damaged element is located on can be identified when the nacelle changes direction and the identification should be made with knowledge of nacelle direction. The height of the damage can be identified by the amount of change in metrics.

In this analysis it was found that more metrics provided more information and so ideally several metrics should be used. When determining the pairing for damaged against undamaged frequency and MAC values, using a pairing which maximises the MAC value can provide the most appropriate and consistent pairing. Modal flexibility can be a useful metric to use because it is calculated without a pairing of individual modes between damaged and undamaged. However, modal flexibility does not provide much information or give have a consistent value for ‘damaged’ and ‘undamaged’ so it is best to use it in conjunction with other metrics.

8.7 Sensor placement

In this chapter, an optimal sensor placement study had been conducted on a structure resulting in 6 possible sensor configurations using the sensors which were installed on an offshore substation that had been operating in the Baltic sea. These sensor placements were composed of two campaigns which had 1.5 months of overlap.

The permanent sensor campaign resulted in the more consistent modal extraction and had no cases of missing data, while the temporary campaign sensors occasionally had missing data. While the permanent layout had the most consistent model extraction, the optimized layouts were able to detect more mode shapes using fewer sensors. This shows that by placing the sensors with consideration for the structure’s mode vectors, fewer sensors can be used.

For a single structure, having fewer sensors might not save much money or time, but if multiple structures within a wind farm have fewer sensors then this can add up to a more significant reduction. However, sensors cannot be placed based on optimised locations alone, they should also be placed with consideration to access and other concerns.

8.8 Part 2 discussion

In part 2 the modal response of the structure in a range of circumstances was investigated and the ability for accelerometers on an operational structure to detect those mode properties was also investigated. This can provide value to an operator who can combine the data they receive from the operational structure with the knowledge about the structure from their digital twin in order to quickly and accurately assess the condition of the structure.

9. Conclusion

9.1 Summary and key findings

The main purpose of this research is to develop methodologies required for the assessment of offshore wind turbine assets through digital twins. Digital twins are inherently multidisciplinary, with both load and resistance aspects to consider as well as the requirement to incorporate data and stochasticity into the models. The components of this research include fluid dynamics, structural dynamics, weather time series, accelerometer data processing and machine learning. The outcome of this research is to support digital twins in offshore wind by developing, verifying and validating modelling approaches required to simulate operational offshore wind assets. The conclusions of this thesis are summarised in the following sub-sections.

9.1.1 Computational Fluid Dynamics

Modelling wind turbine wake effects at a range of wind speeds and directions with actuator disk (AD) models can provide insight but also be challenging. With any model it is important to quantify the level of error, but this can also present a challenge when comparing a steady-state model to measurement data with scatter. This research models wind flow in a wind farm at a range of wind speeds and directions using an AD implementation. The results from these models are compared to data collected from the actual farm being modelled. An extensive comparison is conducted, constituted from 35 cases where two turbulence models, the standard k-ε and k-ω SST are evaluated. The steps taken in building the models as well as processes for comparing the AD computational fluid dynamics (CFD) results to real-world data using the regression models of ensemble bagging and Gaussian process are outlined. Turbine performance data and boundary conditions are determined using the site data. Modifications to an existing opensource AD code are shown so that the predetermined turbine performance can be implemented into the CFD model. Steady state solutions are obtained with the OpenFOAM CFD solver. Results are compared in terms of velocity deficit at the measurement locations. Using the standard k-ε model, a mean absolute error for all cases together of roughly 8% can be achieved, but this error changes for different directions and methods of evaluating it.

9.1.2 Stochastic expansion of Computational Fluid Dynamics model

Wind turbine flow field prediction is difficult as it requires computationally expensive computational fluid dynamics (CFD) models. The contribution of this research is to propose and develop a method for stochastic analysis of an offshore wind farm using CFD and a non-intrusive stochastic expansion. The approach is developed through testing a range of machine-learning methods, evaluating dataset requirements and comparing the accuracy against site measurement data. The approach used is detailed and the results are compared with real measurements obtained from the existing wind farm to quantify the accuracy of the predictions. An existing offshore wind farm is modelled using a steady-state CFD solver at several deterministic input ranges and an approximation model is trained on the CFD results. The approximation models compared are Artificial Neural Networks, Gaussian Process, Radial Basis Function, Random Forest and Support Vector Regression. RBF achieves a mean absolute error relative to the CFD model of only 0.54% and the error of the SVR predictions relative to the real data, with scatter, was 12%, compared to 16% from Jensen. This approach has the potential to be used in more complex situations where an existing analytical method is either insufficient or unable to make a good prediction.

9.1.3 Wind and wave time-series forecasting

With the growing demand for offshore wind energy and the continued drive for reduced levelised cost of energy, it is necessary to further to make operation and maintenance (O&M) activities more effective and reduce related costs. A key factor in achieving this aim is to more representatively model O&M activities and to do this, simulation models should include more accurate weather forecasting algorithms. In this research three weather forecast modelling methods are used to generate projections of wind and wave values which are then compared to measured, operational data. These methods include Markov Chains, gradient boosting and a novel hybrid regression/statistical approach which has been developed and is presented herein. It is shown that numerically the Markov Chain model and the hybrid model perform similarly, although the hybrid method has some additional desirable features.

9.1.4 Damage identification through global dynamics

The modal response of a four-legged jacket structure to damages are explored and resulting considerations for damage detection are discussed. A finite element model of the Wikinger jacket structure is used to investigate damage detection. Damages, such as cracks, scour, corrosion and more, are modelled in a simulation environment. The resulting modal parameters are calculated, these parameters are compared to those from an unaltered structure and metrics are calculated including frequency change, modal assurance criterion and modal flexibility. A highly detailed design-model is used to conduct a sensitivity study on modal parameters for a range of changes. By conducting this on the same structure, this acts as a useful reference for those interested in the dynamic response of offshore wind jacket structures. Additionally, this research addresses the issue of changes in mode parameters resulting from turbine yaw. This research also considers the challenge of mode-swapping in semi-symmetric structures and proposes several approaches for addressing this. Damage typically results in a reduction of frequency and change in mode shapes, but in ways which can be distinguished from other structural changes, given the extent of this model. These findings are important considerations for modal-based damage detection of offshore wind support structures.

9.1.5 Impact of accelerometer placement layout

The modal parameters extracted from a structure by accelerometers can be used for damage assessment as well as model updating. To extract modal parameters from a structure, it is important to place accelerometers at locations with high modal displacements. Sensor placement can be restricted by practical considerations, and installation might be conducted more based on engineering judgement rather than analysis. This leads to the question of how important the optimal sensor placement is, and if fewer sensors suffice to extract the modal parameters.

In this research, an offshore wind substation (OSS) from the Wikinger offshore wind farm (owned by Iberdrola) is instrumented with 12, 3-axis accelerometers. This sensor setup consists of 6 sensors in a permanent campaign where sensors were placed based purely on engineering judgement, as well as 6 sensors in a temporary campaign, placed based on a placement analysis. An optimal sensor placement study was conducted using a finite element model of the structure in the software package FEMtools, resulting in optimal layouts. The temporary campaign sensors were placed such that they, in combination with the permanent campaign, can be used to complete the proposed layouts.

Samples for each setup are processed using the software ARTeMIS modal to extract the mode shapes and natural frequencies through the Stochastic Subspace Identification (SSI) technique. The frequencies found by this approach are then clustered together using a k-means algorithm for a comparison within clusters.

The modal assurance criterion (MAC) values are calculated for each result and compared to the finite element model from which the optimal sensor placement study was conducted. This is to match mode shapes between the two and thus determine the importance of off diagonal MAC elements in the sensor optimization process. MAC values are also calculated relative to a cluster-averaged set of eigenvectors to determine how they vary over the 1.5 months.

The results show that for all sensor layouts, the three lower frequency modes are consistently identified. The most optimized sensor layout, consisting of only 3 sensors, was able to distinguish an additional, higher frequency mode which was never identified in the 6-sensor permanent layout. However, the reduced sensor layout shows slightly more scatter in the results than the 6-sensor layout. There is a higher signal to noise ratio in the temporary campaign which results in scatter. We conclude that with an optimized placement of accelerometers, more modes can be identified and distinguished. However, off diagonal elements in the original MAC matrix, as well as loss of sensor degrees of freedom, can result in additional scatter in the measurements. Some of these findings can be extended to other offshore jacket structures, such as those of wind turbines, in that it gives a better understanding of the consequence of an optimal sensor placement study.

9.2 Contribution to knowledge

This research contributes to knowledge in a way which is novel, scientifically sound and provides value to stakeholders. Three scientific journals have been successfully published as well as four peer-reviewed, scientific conference papers.

Section	Novelty	Value to stakeholders
Computational fluid dynamics	• Simulation of an entire wind farm at multiple free-stream wind speeds and directions - comparing turbulence models for this. Other research has either not shown this range of variables or has not done so at close enough resolution to observe the effects. • Demonstrates a novel modification of an existing actuator disk code.	Researchers and industry will benefit from the wind farm computational fluid dynamics findings as they give a better understanding for how wind flow behaves and how errors related to measurements vary with different wind speeds and directions. The actuator disk code developed can be used to model other offshore wind farms and achieve similar results. The methods shown for comparing accuracy of measured results to CFD results can be of value for others comparing their CFD results to measurements.
Stochastic expansion of CFD	• Develops an approach for expanding upon CFD results using machine learning. This provides a novel method for stochastic analysis of wind turbine loading within a wind farm.	The method used can be applied to gain more value from CFD results at either an initial planning phase of a wind farm or when considering different turbine control behaviour. The validation of results can give confidence to researchers and industry when applying machine learning techniques to similar problems as well as demonstrating which methods might be most suitable and why.
Wind and wave time-series forecasting	• Comparison of existing weather forecasting time-series methods, providing insights into the strengths and weaknesses as well as difference in error between the approaches. • Development of a novel time-series forecasting approach which combines deterministic and statistical models. This approach shows considerable promise in this field as it incorporates features	The new weather forecasting time series method is shown to be as accurate as existing methods while having features which are desirable for weather forecasting. When closer replication of real wind/wave time series behaviour is required in order to reduce conservativism in offshore wind load calculation, this new method can be used and reduce cost.

	which other methods do not.	
Damage identification through global dynamics	• Use of more detailed model than other works for an offshore structural damage detection study. • Includes mass/moment of inertia of rotor nacelle assembly which is found to be important for when observing relative changes in damage metrics. • A broad range of damage were simulated on a single structure which allows for direct comparison between damages. • Explores methods for overcoming the known issue of mode-switching.	The findings of the damage identification work are vital for developing an operational-modal-analysis based damage detection approach in industry. It shows how much the eigenfrequencies and vectors change as a result of various damages which informs practitioners how to identify damage in real structures. Additionally, researchers are often unsure how to approach the issue of mode-switching in semi-symmetric structures as a result of damage, this research shows the effectiveness of different approaches.
Impact of accelerometer placement layout	• Investigate measurements from multiple sensor layouts for a single operational structure. Shows direct implication of sensor layout optimization for operational scenarios.	The research on accelerometer placement impact is important for industry because accelerometer layouts which give better readings usually requires placing them at locations with more difficult access. This research shows the real impact of choosing better placements over accessibility.

Scientific soundness

The research is scientifically sound in that all the best practice procedures have been followed, results have been validated against operational site data and state-of-the-art methods have been used. Both through an extensive literature review, as well as working with experts in academia and industry, it has been assured that the research followed the highest and most appropriate standards. Throughout the work , models and results have been validated against operational site data which provides a metric for the efficacy of the methods used and gives confidence in their results and usefulness.

9.3

9.3 Future research

Regarding potential future work from a CFD perspective, with the growth in computational capabilities, CFD can more be applied in a way that expands the compromise between capturing flow complexity and computational time.

The machine learning based expansion of CFD can be developed further by incorporating the stochastic capabilities of some approximation models in order to replicate the stochastic nature of the wakes. Additionally, other variables which can be calculated could be approximated such as turbulence values, or a time-series prediction can be generated, which can be useful for estimating fatigue.

The time-series forecasting method developed could be developed further by applying the novel time-series method to forecasting other types of time series. This would test its ability under different conditions.

The damage identification model could be applied to other types of structures. Additionally, it could be valuable to apply damage detection algorithms to synthetic data generated from these model results.

References

[1] Enevoldsen P, Xydis G. Examining the trends of 35 years growth of key wind turbine components. Energy Sustain Dev 2019;50:18–26. https://doi.org/10.1016/j.esd.2019.02.003.

[2] Morthors P-E, Awerbuch S. The Economics of Wind Energy. Belgium: 2009. https://doi.org/10.1111/j.1745-6622.2009.00231.x.

[3] Gordon J. Mechanisms of Offshore wind growth : Feasibility of 2030 targets in the North Seas Energy Cooperation. Central European University, 2020.

[4] Mills AD, Millstein D, Jeong S, Lavin L, Wiser R, Bolinger M. Estimating the value of offshore wind along the United States' Eastern Coast. Environ Res Lett 2018;13:21. https://doi.org/10.1088/1748-9326/aada62.

[5] U.S. Department of Energy. 2018 Offshore Wind Technologies Market Report. Oak Ridge, TN: 2018.

[6] Yue C-D, Liu C-C, Tu C-C, Lin T-H. Prediction of Power Generation by Offshore Wind Farms Using Multiple Data Sources. Energies 2019;12:700. https://doi.org/10.3390/en12040700.

[7] Norwegian Energy Partners. Offshore wind in Taiwan. n.d.

[8] Li A, Xu Y. The governance for offshore wind in Japan. Energy Procedia 2019;158:297–301. https://doi.org/10.1016/j.egypro.2019.01.092.

[9] deCastro M, Salvador S, Gómez-Gesteira M, Costoya X, Carvalho D, Sanz-Larruga FJ, et al. Europe, China and the United States: Three different approaches to the development of offshore wind energy. Renew Sustain Energy Rev 2019;109:55–70. https://doi.org/10.1016/j.rser.2019.04.025.

[10] Blanch M, Davies C, Roberts A. Offshore wind energy. Elsevier Inc.; 2018. https://doi.org/10.1016/B978-0-12-818762-3.00007-8.

[11] NGI group. BSEE Offshore Wind Recommendations - Guidlines for Structural Health Monitoring for Offshore Wind Turbine Towers and Foundations. 2017.

[12] Catapult. Data Analytics in the Offshore Wind Industry Pilot Case Study Outcomes. 2018.

[13] Branlard E, Giardina D, S. D. Brown C. Augmented Kalman filter with a reduced mechanical model to estimate tower loads on a land-based wind turbine: A step towards digital-Twin simulations. Wind Energy Sci 2020;5:1155–67. https://doi.org/10.5194/wes-5-1155-2020.

[14] Shittu AA, Mehmanparast A, Wang L, Salonitis K, Kolios A. Comparative study of structural reliability assessment methods for offshore wind turbine jacket support structures. Appl Sci 2020;10. https://doi.org/10.3390/app10030860.

[15] Richmond M, Kolios A, Pillai VS, Nishino T, Wang L. Development of a stochastic computational fluid dynamics approach for offshore wind farms. J Phys Conf Ser 2018.

[16] Richmond M, Antoniadis A, Wang L, Kolios A, Al-Sanad S, Parol J. Evaluation of an offshore wind farm computational fluid dynamics model against operational site data. Ocean Eng 2019;193. https://doi.org/10.1016/j.oceaneng.2019.106579.

[17] Richmond M, Sobey A, Pandit R, Kolios A. Stochastic assessment of aerodynamics within offshore wind farms based on machine-learning. Renew Energy 2020;161:650–61. https://doi.org/10.1016/j.renene.2020.07.083.

[18] Richmond M, Smolka U, Kolios A. Feasibility for Damage Identification in Offshore Wind Jacket Structures through monitoring of global structural dynamics. Proc. 38th IMAC, A Conf. Expo. Struct. Dyn. 2020, 2020. https://doi.org/10.1007/978-3-030-47717-2.

[19] Richmond M, Siedler S, Häckell M, Smolka U, Kolios A. 19052 IMPACT OF ACCELEROMETER PLACEMENT ON MODAL EXTRACTION OF OFFSHORE WIND STRUCTURES. Ocean Mar. Arct. Eng. 2020, 2020.

[20] European Commission. EU on track to meeting 20% renewable energy target 2017. https://ec.europa.eu/energy/en/news/eu-track-meeting-20-renewable-energy-target (accessed April 29, 2017).

[21] Directive (EU) 2018/2001 of the European parliament and of the council of 11 December 2018 on the promotion of the use of energy from renewable sources. Off J Eur Union 2018.

[22] Sobral J, Soares C. Offshore Wind Farms Maintenance Strategy using the Analytic NetworN Process. Proc. 29th Eur. Saf. Reliab. Conf., 2019. https://doi.org/10.3850/978-981-11-2724-3.

[23] Bosch J, Staffell I, Hawkes AD. Global levelised cost of electricity from offshore wind. Energy 2019;189:116357. https://doi.org/10.1016/j.energy.2019.116357.

[24] Wang J, Zhao X, Guo X. Optimizing wind turbine's maintenance policies under performance-based contract. Renew Energy 2019;135:626–34. https://doi.org/10.1016/j.renene.2018.12.006.

[25] Rodrigues R, Lengsfeld C. Development of a Computational System to Improve Wind Farm Layout, Part I: Model Validation and Near Wake Analysis. Energies 2019;12:940. https://doi.org/10.3390/en12050940.

[26] Nelson B, Quéméner Y. Fatigue Life Analysis of Offshore Wind Turbine Support Structures in an Offshore Wind Farm. ASME 2018 1st Int. Offshore Wind Tech. Conf., American Society of Mechanical Engineers; 2018. https://doi.org/10.1115/IOWTC2018-1061.

[27] Tygesen UT, Jepsen MS, Vestermark J, Dollerup N, Pedersen A. The true digital twin concept for fatigue re-assessment of marine structures. Proc Int Conf Offshore Mech Arct Eng - OMAE 2018;1. https://doi.org/10.1115/OMAE2018-77915.

[28] Dimitrov N, Kelly MC, Vignaroli A, Berg J. From wind to loads: wind turbine site-specific load estimation with surrogate models trained on high-fidelity load databases. Wind Energy Sci 2018;3:767–90. https://doi.org/10.5194/wes-3-767-2018.

[29] Göçmen T, Laan P Van Der, Réthoré PE, Diaz AP, Larsen GC, Ott S. Wind turbine wake models developed at the technical university of Denmark: A review. Renew Sustain Energy Rev 2016;60:752–69. https://doi.org/10.1016/j.rser.2016.01.113.

[30] Popko W, Robertson A, Jonkman J, Wendt F, Thomas P, Uller KM, et al. Validation of numerical models of the offshore wind turbine from the alpha ventus wind farm against full-scale measurements within OC5 phase III. Proc Int Conf Offshore Mech Arct Eng - OMAE 2019;10. https://doi.org/10.1115/omae2019-95429.

[31] Carrillo C, Cidrás J, Díaz-Dorado E, Obando-Montaño AF. An approach to determine the weibull parameters for wind energy analysis: The case of Galicia (Spain). Energies 2014;7:2676–700. https://doi.org/10.3390/en7042676.

[32] Sohoni V. A comparative analysis of wind speed probability distributions for wind power assessment of four sites 2015:1–12. https://doi.org/10.3906/elk-1412-207.

[33] Ditlevsen O. Stochastic model for joint wave and wind loads on offshore structures. Struct Saf 2002;24:139–63. https://doi.org/10.1016/S0167-4730(02)00022-X.

[34] Cakmakyapan S, Kadilar G. The Poisson Gamma distribution for wind speed data. AIP Conf Proc 2016;1726. https://doi.org/10.1063/1.4945896.

[35] Merz K. Offshore Wind Energy Technology" - chapter 4 Author(s): Olimpo Anaya-Lara John O. Tande Kjetil Uhlen. 2018.

[36] Morgan EC, Lackner M, Vogel RM, Baise LG. Probability distributions for offshore wind speeds. Energy Convers Manag 2011;52:15–26. https://doi.org/10.1016/j.enconman.2010.06.015.

[37] Jaramillo OA, Borja MA. Bimodal versus Weibull wind speed distributions: An analysis of wind energy potential in La Venta, Mexico. Wind Eng 2004;28:225–34. https://doi.org/10.1260/0309524041211404.

[38] American Bureau of Shipping. Design Standards for Offshore Wind Farms. Houston Texas: 2011.

[39] Parajuli A. A Statistical Analysis of Wind Speed and Power Density Based on Weibull and Rayleigh Models of Jumla, Nepal. Nepal Energy Power Eng 2004;8:271–82. https://doi.org/10.4236/epe.2016.87026.

[40] Hou Y, Peng Y, Johnson a L, Shi J. Empirical Analysis of Wind Power Potential at Multiple Heights for North Dakota Wind Observation Sites 2012;4:1–9. https://doi.org/10.3968/j.est.1923847920120401.289.

[41] Kelley ND, Jonkman BJ. Overview of the TurbSim Stochastic Inflow Turbulence Simulator:

Version 1.21 (Revised February 1, 2001). NREL/TP-500-41137 2007:1–13.

[42] Mann J. Wind field simulation. Probabilistic Eng Mech 1998;13:269–82. https://doi.org/10.1016/s0266-8920(97)00036-2.

[43] NREL. AeroDyn Manual V15.04a. AeroDyn Man NREL Portal n.d. https://nwtc.nrel.gov/AeroDyn.

[44] DNV-GL. BLADED engineering feature summary. 2015.

[45] Teng C, Palao I. Wave Height and Period Distributions From Long-Term Wave Measurement. Coast Eng Proc 1996:368–79.

[46] Burrows R, Salih B a. Statistical Modelling of Long-Term Wave Climates. Coast Eng 1982:42–56.

[47] Prevosto M, Krogstad HE, Robin A. Probability distributions for maximum wave and crest heights. vol. 40. 2000. https://doi.org/10.1016/S0378-3839(00)00017-X.

[48] Satheesh SP, Praveen VK, Jagadish Kumar V, Muraleedharan G, Kurup PG. Weibull and Gamma distributions for wave parameter predictions. J Indian Geophys Union 2005;9:55–64.

[49] Persson K. Exponentiated Gumbel Distribution for Estimation. J Environ Stat 2010;1:1–12.

[50] van Gelder PHAJMHAJMVHAJM Van, Vrijling JKJK. On the distribution function of the maximum wave height in front of reflecting structures. Proc Coast Struct … 1999;1:37–46.

[51] Orcina Ltd. orcaflex Introduction 2020. https://orcina.com/webhelp/OrcaFlex/.

[52] Yang M, Zhang L, Cui Y, Yang Q, Huang B. The impact of wind field spatial heterogeneity and variability on short-term wind power forecast errors. J Renew Sustain Energy 2019;11. https://doi.org/10.1063/1.5064438.

[53] BSI British Standards Wind turbines. Part 3: Design requirements for offshore wind turbines. BS EN 61400-3:2009. 2009.

[54] International Electrotechnical Commissio. IEC 61400-3-1:2019 , Wind energy generation systems - Part 3-1: Design requirements for fixed offshore wind turbines. 2019.

[55] Jang H-K, Park S, Kim M-H, Kim K-H, Hong K. Effects of heave plates on the global performance of a multi-unit floating offshore wind turbine. Renew Energy 2019;134:526–37. https://doi.org/10.1016/j.renene.2018.11.033.

[56] Lee, Hua H, Hsieh H-Y. Dynamic Behaviour of a Mono-pile with Skirted Template Structure for Offshore Wind-Power System. J Phys Conf Ser 2019;1298:012028. https://doi.org/10.1088/1742-6596/1298/1/012028.

[57] Hsu W, Thiagarajan KP, MacNicoll M, Akers R. Prediction of Extreme Tensions in Mooring Lines

of a Floating Offshore Wind Turbine in a 100-Year Storm. Vol. 9 Ocean Renew. Energy, American Society of Mechanical Engineers; 2015. https://doi.org/10.1115/OMAE2015-42015.

[58] Foley AM, Leahy PG, Marvuglia A, McKeogh EJ. Current methods and advances in forecasting of wind power generation. Renew Energy 2012;37:1–8. https://doi.org/10.1016/j.renene.2011.05.033.

[59] Richmond M, Pandit R, Koukoura S, Kolios A. Effect of weather forecasting modelling uncertainty to the availability assessment of offshore wind farms. Submitt (Copy File with Author) 2020.

[60] Qin Q, Lai X, Zou J. Direct multistep wind speed forecasting using LSTM neural network combining EEMD and fuzzy entropy. Appl Sci 2019;9. https://doi.org/10.3390/app9010126.

[61] Gangwar S, Bali V, Kumar A. Comparative Analysis of Wind Speed Forecasting Using LSTM and SVM. ICST Trans Scalable Inf Syst 2018:159407. https://doi.org/10.4108/eai.13-7-2018.159407.

[62] Wang J, Hu J. A robust combination approach for short-term wind speed forecasting and analysis - Combination of the ARIMA (Autoregressive Integrated Moving Average), ELM (Extreme Learning Machine), SVM (Support Vector Machine) and LSSVM (Least Square SVM) forecasts usi. Energy 2015;93:41–56. https://doi.org/10.1016/j.energy.2015.08.045.

[63] Tin Kam Ho. Random decision forests. Proc. 3rd Int. Conf. Doc. Anal. Recognit., vol. 1, IEEE Comput. Soc. Press; 1995, p. 278–82. https://doi.org/10.1109/ICDAR.1995.598994.

[64] Breiman L. Random forests. Mach Learn 2001;45:5–32. https://doi.org/10.1023/A:1010933404324.

[65] Chen T, Singh S, Taskar B, Guestrin C. Effcient Second-Order Gradient Boosting for Conditional Random Fields. Proc. 18th Int. Conf. Arti cial Intell. Stat., vol. 38, 2015.

[66] Verbois H, Rusydi A, Thiery A. Probabilistic forecasting of day-ahead solar irradiance using quantile gradient boosting. Sol Energy 2018;173:313–27. https://doi.org/10.1016/j.solener.2018.07.071.

[67] Browell J, Gilbert C, McMillan D. Use of turbine-level data for improved wind power forecasting. 2017 IEEE Manchester PowerTech, Powertech 2017 2017. https://doi.org/10.1109/PTC.2017.7981134.

[68] Ciucu S, Rebenciuc M. The use of Markov chains in forecasting wind speed: Matlab source code and applied case study. Comput Methods Soc Sci 2016;4:44–53.

[69] Carpinone A, Giorgio M, Langella R, Testa A. Markov chain modeling for very-short-term wind power forecasting. Electr Power Syst Res 2015;122:152–8. https://doi.org/10.1016/j.epsr.2014.12.025.

[70] Karatepe S, Corscadden KW. Wind Speed Estimation: Incorporating Seasonal Data Using

Markov Chain Models. ISRN Renew Energy 2013;2013:1–9. https://doi.org/10.1155/2013/657437.

[71] Sánchez-Pérez PA, Robles M, Jaramillo OA. Real time Markov chains: Wind states in anemometric data. J Renew Sustain Energy 2016;8:1–22. https://doi.org/10.1063/1.4943120.

[72] Vahdatirad MJ, Andersen LV, Ibsen LB. Probabilistic Three-Dimensional Model of an Offshore Monopile Foundation : Reliability Based Approach FOUNDATION : RELIABILITY BASED APPROACH 2013.

[73] Krige DG. A statistical approach to some basic mine valuation problems on Witwatersrand. Chem Met Min Soc South Africa 1951;52:119–39.

[74] Morato A, Sriramula S, Krishnan N. Reliability analysis of offshore wind turbine support structures using Kriging models. Risk, Reliab. Saf. Innov. Theory Pract., CRC Press; 2017, p. 1425–31.

[75] Lukasik B, Goddard KF, Rotaru MD, Sykulski JK. Kriging assisted design of a synchronous superconducting generator with YBCO windings n.d.:310–1.

[76] Georgia Bayliss-Brown. Adding value to monitoring: The impact of offshore wind farm developments on benthic organisms and sediments in the Eastern Irish Sea. University of East Anglia, 2012. https://doi.org/10.13140/RG.2.1.2648.9369.

[77] Yang H, Zhu Y. Robust Requirements and Design Optimization for Supporting Structure of Offshore Wind Turbine Utilizing Approximate Model Technology 2014;3:250–9.

[78] Firouzianbandpey S. Reliability-Based Design of Wind Turbine Foundations: Geotechnical Site Assessment. Aalborg University, 2016.

[79] Fritz J, Neuweiler I, Nowak W. Application of FFT-based algorithms for large-scale universal kriging problems. Math Geosci 2009;41:509–33. https://doi.org/10.1007/s11004-009-9220-x.

[80] Carswell W, Arwade SR, DeGroot DJ, Lackner MA. Soil–structure reliability of offshore wind turbine monopile foundations. Wind Energy 2014. https://doi.org/10.1002/we.1710.

[81] Andersen L V., Vahdatirad MJ, Sichani MT, Sørensen JD. Natural frequencies of wind turbines on monopile foundations in clayey soils-A probabilistic approach. Comput Geotech 2012;43:1–11. https://doi.org/10.1016/j.compgeo.2012.01.010.

[82] Todo Bom L, Siedler S, Tautz-Weinert J. Validation of CPT-based initial soil stiffness in sand for offshore wind jacket piles. Mess. der Geotech. 2020, Braunschweig: 2020. https://doi.org/10.24355/dbbs.084-201912181435-0.

[83] Henderson AR, Zaaijer MB, Camp TR. Hydrodynamic Loading on Offshore Wind Turbines. Order A J Theory Ordered Sets Its Appl 2003;2:1–15.

[84] Heege A, Bonnet P, Bastard L, Horcas SG, Sanchez JL, Cucchini P, et al. Numerical simulation of

offshore wind turbines by a coupled aerodynamic , hydrodynamic and structural dynamic approach. DEWI Mag 2011:6–15.

[85] Peeringa JM. Wave loads on offshore wind turbines 2004.

[86] MacCamy RC, Fuchs RA. Wave Forces on Piles: A Diffraction Theory. 1954.

[87] Mockutė A, Marino E, Lugni C, Borri C. Comparison of nonlinear wave-loading models on rigid cylinders in regular waves. Energies 2019;12:1–22. https://doi.org/10.3390/en12214022.

[88] Copyright Orcina Ltd. Orcina: OrcaFlex user manual n.d.

[89] Rainey RCT. A new equation for calculating wave loads on offshore structures. J Fluid Mech 1989;204:295–324. https://doi.org/10.1017/S002211208900176X.

[90] Rainey RCT. Slender-body expression for the wave load on offshore structures. Proc. Math. Phys. Sci., vol. 450, No. 1, 1995, p. 391–416.

[91] Faltinsen OM, Newman JN, Vinje T. Nonlinear wave loads on a slender vertical cylinder. J Fluid Mech 1995;289:179–98. https://doi.org/10.1017/S0022112095001297.

[92] Kristiansen T, Faltinsen OM. Higher harmonic wave loads on a vertical cylinder in finite water depth. J Fluid Mech 2017;833:773–805. https://doi.org/10.1017/jfm.2017.702.

[93] Ristić M. Validation of Load Models and Calculations of Response for a Monopile in Steep Water Waves. 2018.

[94] Chen J. Nonlinear Wave Loads on Offshore Wind Support Structure. 2014.

[95] Wienke J, Oumeraci H. Breaking wave impact force on a vertical and inclined slender pile - Theoretical and large-scale model investigations. Coast Eng 2005;52:435–62. https://doi.org/10.1016/j.coastaleng.2004.12.008.

[96] De Ridder EJ, Aalberts P, Van Den Berg J, Buchner B, Peeringa J. The dynamic response of an offshore wind turbine with realistic flexibility to breaking wave impact. Proc Int Conf Offshore Mech Arct Eng - OMAE 2011;5:543–52. https://doi.org/10.1115/OMAE2011-49563.

[97] Bredmose H, Schløer S, Sahlberg-Nielsen L, Slabiak P, Larsen TJ, Kim T, et al. Loads and response from steep and breaking waves on monopiles. 2013.

[98] Chella MA, Tørum A, Myrhaug D. An overview of wave impact forces on offshore wind turbine substructures. Energy Procedia 2012;20:217–26. https://doi.org/10.1016/j.egypro.2012.03.022.

[99] Jonkman J, Musial W. Offshore code comparison collaboration (OC3) for IEA task 23 offshore wind technology and deployment. vol. 303. Golden, Colorado, USA: 2010. https://doi.org/NREL/TP-5000-48191.

[100] Robertson AN, Wendt F, Jonkman JM, Popko W, Borg M, Bredmose H, et al. OC5 Project Phase Ib: Validation of Hydrodynamic Loading on a Fixed, Flexible Cylinder for Offshore Wind Applications. Energy Procedia 2016;94:82–101. https://doi.org/10.1016/j.egypro.2016.09.201.

[101] DHI. Hydrodynamic loads on offshore wind turbines. n.d.

[102] Liu Y, Xiao Q, Incecik A, Peyrard C, Wan D. Establishing a fully coupled CFD analysis tool for floating offshore wind turbines. Renew Energy 2017;112:280–301. https://doi.org/10.1016/j.renene.2017.04.052.

[103] Borg M, Collu M. Offshore floating vertical axis wind turbines, dynamics modelling state of the art. Part III: Hydrodynamics and coupled modelling approaches. Renew Sustain Energy Rev 2015;46:296–310. https://doi.org/10.1016/j.rser.2014.10.100.

[104] Zhang Y. Wave Loads on Offshore Wind Turbines. MIT, 2015.

[105] Leble V, Barakos GN. Detailed Simulation of Offshore Wind Turbine n.d.:10–5.

[106] Leble V, Barakos GN. A Coupled Floating Offshore Wind Turbine Analysis with High-fidelity Methods. Energy Procedia 2016;94:523–30. https://doi.org/10.1016/j.egypro.2016.09.229.

[107] Besio G, Losada MA. Sediment transport patterns at Trafalgar offshore windfarm. Ocean Eng 2008;35:653–65. https://doi.org/10.1016/j.oceaneng.2008.01.002.

[108] Millar DL, Smith HCM, Reeve DE. Modelling analysis of the sensitivity of shoreline change to a wave farm. Ocean Eng 2007;34:884–901. https://doi.org/10.1016/j.oceaneng.2005.12.014.

[109] van der Molen J, Smith HCM, Lepper P, Limpenny S, Rees J. Predicting the large-scale consequences of offshore wind turbine array development on a North Sea ecosystem. Cont Shelf Res 2014;85:60–72. https://doi.org/10.1016/j.csr.2014.05.018.

[110] Delft T. SWAN User Manual 2020.

[111] Ponce de León S, Bettencourt JH, Kjerstad N. Simulation of irregular waves in an offshore wind farm with a spectral wave model. Cont Shelf Res 2011;31:1541–57. https://doi.org/10.1016/j.csr.2011.07.003.

[112] Christensen ED, Johnson M, Sørensen OR, Hasager CB, Badger M, Larsen SE. Transmission of wave energy through an offshore wind turbine farm. Coast Eng 2013;82:25–46. https://doi.org/10.1016/j.coastaleng.2013.08.004.

[113] Sismani G, Babarit A, Loukogeorgaki E. Impact of fixed bottom Offshore Wind farms on the surrounding wave field. Proc Int Offshore Polar Eng Conf 2016;2016-Janua:511–8.

[114] Babarit A, Delhommeau G. Theoretical and numerical aspects of the open source BEM solver NEMOH. Proc. 11th Eur. Wave Tidal Energy Conf., 2015, p. 1–12. https://doi.org/hal-01198800.

[115] Fu ZJ, Xie ZY, Ji SY, Tsai CC, Li AL. Meshless generalized finite difference method for water wave interactions with multiple-bottom-seated-cylinder-array structures. Ocean Eng 2020;195:106736. https://doi.org/10.1016/j.oceaneng.2019.106736.

[116] Bai W, Feng X, Taylor RE, Ang KK. Fully nonlinear analysis of near-trapping phenomenon around an array of cylinders. Appl Ocean Res 2014;44:71–81. https://doi.org/10.1016/j.apor.2013.11.003.

[117] Hsiao GC, Sánchez-Vizuet T, Sayas FJ. Boundary and coupled boundary-finite element methods for transient wave-structure interaction. IMA J Numer Anal 2017;37:237–65. https://doi.org/10.1093/imanum/drw009.

[118] Bhattacharya S, Nikitas G, Arany L, Nikitas N. Soil–Structure Interactions for Offshore Wind Turbines. Eng Technol Ref 2012;1. https://doi.org/10.1049/etr.2016.0019.

[119] W G Versteijlen, Oliveira JM De, Van BKN, Metrikine DAV. Method for extracting an equivalent Winkler model of the 3D dynamic soil-structure interaction of large-diameter offshore monopile foundations. Proc XLIII Int Summer Sch APM 2015 Identified 2015.

[120] Martinez-Luengo M, Kolios A, Wang L. Parametric FEA modelling of offshore wind turbine support structures: Towards scaling-up and CAPEX reduction. Int J Mar Energy 2017;19:16–31. https://doi.org/10.1016/j.ijome.2017.05.005.

[121] Rajashree SS, Sundaravadivelu R. Degradation model for one-way cyclic lateral load on piles in soft clay. Comput Geotech 1996;19:289–300. https://doi.org/10.1016/S0266-352X(96)00008-0.

[122] Byrne B, McAdam R, Burd H, Houlsby G, Martin C, c L, et al. New design methods for large diameter piles under lateral loading for offshore wind applications. Front Offshore Geotech III 2015:705–10. https://doi.org/10.1201/b18442-96.

[123] Oxford University. Pile Soil Analysis Project n.d. http://www.eng.ox.ac.uk/geotech/research/PISA (accessed January 23, 2018).

[124] Kallehave D, Byrne BW, LeBlanc Thilsted C, Mikkelsen KK. Optimization of monopiles for offshore wind turbines. Philos Trans R Soc A Math Phys Eng Sci 2015;373:20140100–20140100. https://doi.org/10.1098/rsta.2014.0100.

[125] Lade P V. Overview of Constitutive Models for Soils. Soil Const. Model., Reston, VA: American Society of Civil Engineers; 2005, p. 1–34. https://doi.org/10.1061/40771(169)1.

[126] Whyte S, Rattley M, Erbrich C, Burd H, Martin C. A Practical Constitutive Model for Offshore Foundation Problems Involving Dense Sands in Partially Drained to Undrained Conditions. Offshore Site Investig. Geotech. 8th Int. Conf. Proc., Society of Underwater Technology; n.d., p. 400–7. https://doi.org/10.3723/OSIG17.400.

[127] van der Tempel J, Zaaijer MB, Subroto H. The effects of Scour on the design of offshore wind turbines. Proc 3rd Int Conf Mar Renew Energy 2004:27–35.

[128] Mayall R, Byrne B, Burd H, McAdam R. Modelling of foundation response to scour and scour protection for offshore wind turbine structures. 2018 ICSE, 2018.

[129] Mayall R, Mcadam R, Byrne BW, Burd H, Sheil B, Cassie P, et al. Experimental modelling of the effects of scour on offshore wind turbine monopile foundations. 9th Int. Conf. Phys. Model. Geotech., 2018.

[130] Chavez CEA, Stratigaki V, Wu M, Troch P, Schendel A, Welzel M, et al. Large-Scale Experiments to Improve Monopile Scour Protection Design Adapted to Climate Change—The PROTEUS Project. Energies 2019;12.

[131] Whitehouse RJS, Harris JM, Sutherland J, Rees J. The nature of scour development and scour protection at offshore windfarm foundations. Mar Pollut Bull 2011;62:73–88. https://doi.org/10.1016/j.marpolbul.2010.09.007.

[132] Whitehouse RJS. Scour at Coastal Structures. 3rd Int Conf Scour Eros 2006.

[133] Harris Whitehouse, R.J., Benson, T JM. The time evolution of scour around offshore structures. Ice 2012;Maritime E:16. https://doi.org/10.1680/maen.2010.163.

[134] Sarah VE, Mieke M, Piet H, Michiel S, Marc S. Offshore wind farms : sand as indispensable as wind 2012;2:101–3.

[135] Van den Eynde D, Giardino A, Portilla J, Fettweis M, Francken F, Monbaliu J. Modelling the effects of sand extraction, on sediment transport due to tides, on the Kwinte Bank. J Coast Res 2010;SI No. 51:101–16. https://doi.org/10.2307/40928823.

[136] Van Den Eynde D. Interpretation of tracer experiments with fine-grained dredging material at the Belgian Continental Shelf by the use of numerical models. J Mar Syst 2004;48:171–89. https://doi.org/10.1016/j.jmarsys.2003.03.003.

[137] Fettweis M, Van Den Eynde D. The mud deposits and the high turbidity in the Belgian-Dutch coastal zone, southern bight of the North Sea. Cont Shelf Res 2003;23:669–91. https://doi.org/10.1016/S0278-4343(03)00027-X.

[138] Van Lancker Vera, Baeye Matthias, Evangelinos Dimitris, Francken Frederic V den E, Dries, De Mesel Ilse, Kerckhof Francis, Norro Alain V den BR. Integrated monitoring of sediment processes in an area of intensive aggregate extraction, Hinder Banks, Belgian part of the North Sea. Economie 2014:59–71.

[139] Sanderse B. Aerodynamics of wind turbine wakes: Literature review. Energy Res Cent Netherlands 2009:1–46. https://doi.org/10.1002/we.

[140] Sanderse B, Pijl SP van der, Koren B. Review of computational fluid dynamics for wind turbine wake aerodynamics. Wind Energy 2011;14:799–819. https://doi.org/0.1002/we.458.

[141] Betz A. Das Maximum der theoretisch möglichen Ausnutzung des Windes durch Windmotoren. Zeitschrift für das gesamte Turbinenwes. 26, 1920, p. 307–9.

[142] Crespo A, Hernandez J, Frandsen S. Survey of modelling methods for wind turbine wakes and wind farms. Wind Energy 1999;2:1–24. https://doi.org/10.1002/(SICI)1099-1824(199901/03)2:1<1::AID-WE16>3.0.CO;2-7.

[143] Johnson B. Computational Fluid Dynamics (CFD) modelling of renewable energy turbine wake interactions. University of Central Lancashire, 2015.

[144] Vermeer LJ, S JN, Crespo A. Wind turbine wake aerodynamics. Prog Aerosp Sci 2003;39:467–510. https://doi.org/10.1016/S0376-0421(03)00078-2.

[145] Sørensen JN, Mikkelsen R, Sarmast S, Ivanell S, Henningson D. Determination of wind turbine near-wake length based on stability analysis. J Phys Conf Ser 2014;524. https://doi.org/10.1088/1742-6596/524/1/012155.

[146] Garcia L, Vatn M, Mühle F, Sætran L. Experiments in the wind turbine far wake for the evaluation of an analytical wake model. J Phys Conf Ser 2017;854. https://doi.org/10.1088/1742-6596/854/1/012015.

[147] Hansen MOL, Aagaard Madsen H. Review paper on wind turbine aerodynamics. J Fluids Eng Trans ASME 2011;133:1–12. https://doi.org/10.1115/1.4005031.

[148] Clifton A, Wagner R. Accounting for the effect of turbulence on wind turbine power curves. J Phys Conf Ser 2014;524. https://doi.org/10.1088/1742-6596/524/1/012109.

[149] Wang L. Nonlinear aeroelastic modelling of large wind turbine composite blades. University of Central Lancashire, 2015.

[150] Porté-Agel F, Lu H, Wu Y-T. Interaction between Large Wind Farms and the Atmospheric Boundary Layer. Procedia IUTAM 2014;10:307–18. https://doi.org/10.1016/j.piutam.2014.01.026.

[151] Kalvig S, Manger E, Hjertager B. Comparing different CFD wind turbine modelling approaches with wind tunnel measurements. J Phys Conf Ser 2014;555:012056. https://doi.org/10.1088/1742-6596/555/1/012056.

[152] Spalding BE, Launder DB. Lectures in mathematical models of turbulence. London: Academic Press; 1972.

[153] Johnson PB, Jonsson C, Achilleos S, Eames I. On the spread and decay of wind turbine wakes in ambient turbulence. J Phys Conf Ser 2014;555:012055. https://doi.org/10.1088/1742-6596/555/1/012055.

[154] Cabezón D, Migoya E, Crespo A. Comparison of turbulence models for the computational fluid dynamics simulation of wind turbine wakes in the atmospheric boundary layer. Wind Energy 2011;14:909–21. https://doi.org/10.1002/we.516.

[155] Cabezón D, Migoya E, Crespo A. A semi-parabolic wake model for large offshore wind farms based on the open source CFD solver OpenFOAM. ITM Web Conf 2014;2:06002.

https://doi.org/10.1051/itmconf/20140206002.

[156] Javaheri A (DEWI), Cañadillas B (DEWI). Wake Modeling of an Offshore Wind Farm Using OpenFOAM. Dewi Mag 2013:15–22.

[157] Dhert T. High-fidelity aerodynamic shape optimization of wind turbine blades by 2015.

[158] Xie S, Archer C. Self-similarity and turbulence characteristics of wind turbine wakes via large-eddy simulation. Wind Energy 2015;18:1815–38. https://doi.org/10.1002/we.1792.

[159] Mehta D, van Zuijlen AH, Koren B, Holierhoek JG, Bijl H. Large Eddy Simulation of wind farm aerodynamics: A review. J Wind Eng Ind Aerodyn 2014;133:1–17. https://doi.org/10.1016/j.jweia.2014.07.002.

[160] Churchfield M, Lee S, Moriarty P, Martinez L, Leonardi S, Vijayakumar G, et al. A Large-Eddy Simulations of Wind-Plant Aerodynamics. 50th AIAA Aerosp. Sci. Meet. Incl. New Horizons Forum Aerosp. Expo., Reston, Virigina: American Institute of Aeronautics and Astronautics; 2012, p. AIAA 2012-0537. https://doi.org/10.2514/6.2012-537.

[161] Wu YT, Porté-Agel F. Large-Eddy Simulation of Wind-Turbine Wakes: Evaluation of Turbine Parametrisations. Boundary-Layer Meteorol 2011;138:345–66. https://doi.org/10.1007/s10546-010-9569-x.

[162] Stevens RJAM, Martínez-Tossas LA, Meneveau C. Comparison of wind farm large eddy simulations using actuator disk and actuator line models with wind tunnel experiments. Renew Energy 2018;116:470–8. https://doi.org/10.1016/j.renene.2017.08.072.

[163] Yang X, Sotiropoulos F. On the predictive capabilities of LES-actuator disk model in simulating turbulence past wind turbines and farms. Proc Am Control Conf 2013:2878–83. https://doi.org/10.1109/acc.2013.6580271.

[164] Larsen GC. Dynamic wake meandering modeling. vol. 1607. 2007.

[165] Larsen TJ, Larsen GC, Pedersen MM, Enevoldsen K, Madsen HA. Validation of the Dynamic Wake Meander model with focus on tower loads. J Phys Conf Ser 2017;854. https://doi.org/10.1088/1742-6596/854/1/012027.

[166] Ennis BL, Kelley CL, Maniaci DC. Dynamic wake meandering model comparison with varying fidelity models for wind turbine wake prediction. Annu Forum Proc - AHS Int 2015;4:2885–94.

[167] Bayo RT, Parro G. Site Suitability Assessment with Dynamic Wake Meandering Model. A Certification Point of View. Energy Procedia 2015;76:177–86. https://doi.org/10.1016/j.egypro.2015.07.893.

[168] Annoni J, Seiler P, Johnson K, Fleming P, Gebraad P. Evaluating wake models for wind farm control. Proc Am Control Conf 2014:2517–23. https://doi.org/10.1109/ACC.2014.6858970.

[169] Stergiannis N, Lacor C, Beeck J V., Donnelly R. CFD modelling approaches against single wind

turbine wake measurements using RANS. J Phys Conf Ser 2016;753:032062. https://doi.org/10.1088/1742-6596/753/3/032062.

[170] Sørensen JN, Shen WZ. Numerical Modeling of Wind Turbine Wakes. J Fluids Eng 2002;124:393–9. https://doi.org/10.1115/1.1471361.

[171] Dobrev I, Massouh F, Rapin M. Actuator surface hybrid model. J Phys Conf Ser 2007;75. https://doi.org/10.1088/1742-6596/75/1/012019.

[172] Castellani F, Gravdahl A, Crasto G, Piccioni E, Vignaroli A. A practical approach in the CFD simulation of off-shore wind farms through the actuator disc technique. Energy Procedia 2013;35:274–84. https://doi.org/10.1016/j.egypro.2013.07.180.

[173] Glauert H. Aeroplane Propellers. Aerodyn Theory 1935;4:169–360.

[174] Eliassen L. Aerodynamic loads on a wind turbine rotor in axial motion Lene Eliassen. University of Stavanger, 2015.

[175] Engelen TG Van. Frequency domain load calculation for offshore wind turbines (TURBU Offshore). Eur Wind Energy Conf 2004:22–5.

[176] Mahmuddin F. Rotor Blade Performance Analysis with Blade Element Momentum Theory. Energy Procedia 2017;105:1123–9. https://doi.org/10.1016/j.egypro.2017.03.477.

[177] Dhert T, Ashuri T, Chen S, Martins JRRA. Aerodynamic Shape Optimization of Wind Turbine Blades Using a Reynolds-Averaged Navier-Stokes Model and an Adjoint Method. Wind Energy 2015:1–26. https://doi.org/10.1002/we.2070.The.

[178] Réthoré P, Zahle F, Sørensen NN. Comparison of an actuator disc model with a full Rotor CFD model under uniform and shear inflow condition Comparison p of an Actuator Disc model with a Full rotor computation, 2008.

[179] Lanzafame R, Mauro S, Messina M. Wind turbine CFD modeling using a correlation-based transitional model. Renew Energy 2013;52:31–9. https://doi.org/10.1016/j.renene.2012.10.007.

[180] Wang L, Quant R, Kolios A. Fluid structure interaction modelling of horizontal-axis wind turbine blades based on CFD and FEA. J Wind Eng Ind Aerodyn 2016;158:11–25. https://doi.org/10.1016/j.jweia.2016.09.006.

[181] Li S. Wind Array Performance Evaluation Model for Large Wind 2014.

[182] Gaumond M, Réthoré P-E, Bechmann A, Ott S, Larson GC, Pena A, et al. Benchmarking of wind turbine wake models in large offshore wind farms. DTU Wind Energy 2012.

[183] Hansen KS, Réthoré PE, Palma J, Hevia BG, Prospathopoulos J, Peña A, et al. Simulation of wake effects between two wind farms. J Phys Conf Ser 2015;625. https://doi.org/10.1088/1742-6596/625/1/012008.

[184] Morales JB, Meissner C. Analytical wake models validation using WindSim. EWEA Resour Assess 2015 – Helsinki– 2-3 June 2015 2015:2015.

[185] Johnson KE, Thomas N. Wind farm control: addressing the aerodynamic interaction among wind turbines. Proc Am Control Conf 2009:2104–9. https://doi.org/10.1109/ACC.2009.5160152.

[186] West JR, Lele SK. Wind Turbine Performance in Very Large Wind Farms: Betz Analysis Revisited. Energies 2020;13:1078. https://doi.org/10.3390/en13051078.

[187] Ma L, Nishino T. Preliminary estimate of the impact of support structures on the aerodynamic performance of very large wind farms. J Phys Conf Ser 2018;1037. https://doi.org/10.1088/1742-6596/1037/7/072036.

[188] Antoniadis AF. Prediction of the impact of support structures on the aerodynamic performance of large wind farms. J Renew Sustain Energy 2019;11:063306. https://doi.org/10.1063/1.5120602.

[189] Sforza PM, Sheering P, Smorto M. Three-dimensional Wakes of Simulated Wind Turbines. AIAA J 1981;19:1101–7.

[190] Ainslie JF. Calculating the flowfield in the wake of wind turbines. J Wind Eng Ind Aerodyn 27 1988;27:213–24. https://doi.org/10.1016/0167-6105(88)90037-2.

[191] Kim H, Kim K, Bottasso CL, Campagnolo F, Paek I. Wind Turbine Wake Characterization for Improvement of the Ainslie Eddy Viscosity Wake Model. Energies 2018. https://doi.org/10.3390/en11102823.

[192] Larsen GC, Madsen HA, Torben J, Troldborg N, Troldborg N. Wake modeling and simulation 2008;1653.

[193] DNV. DNVGL-ST-0126, Support structures for wind turbines. 2018.

[194] Arany L, Bhattacharya S, Macdonald JHG, Hogan SJ. A Critical Review of Serviceability Limit State Requirements for Monopile Foundations of Offshore Wind Turbines. All Days, OTC; 2015. https://doi.org/10.4043/25874-MS.

[195] Bhattacharya S, Lombardi D, Amani S, Aleem M, Prakhya G, Adhikari S, et al. Physical modelling of offshore wind turbine foundations for trl (Technology readiness level) studies. J Mar Sci Eng 2021;9. https://doi.org/10.3390/jmse9060589.

[196] Page AM, Klinkvort RT, Bayton S, Zhang Y, Jostad HP. A procedure for predicting the permanent rotation of monopiles in sand supporting offshore wind turbines. Mar Struct 2021;75:102813. https://doi.org/10.1016/j.marstruc.2020.102813.

[197] Wang M, Wang C, Hnydiuk-Stefan A, Feng S, Atilla I, Li Z. Recent progress on reliability analysis of offshore wind turbine support structures considering digital twin solutions. Ocean Eng 2021;232:109168. https://doi.org/10.1016/j.oceaneng.2021.109168.

[198] Wang Y-K, Chai J-F, Chang Y-W, Huang T-Y, Kuo Y-S. Development of Seismic Demand for Chang-Bin Offshore Wind Farm in Taiwan Strait. Energies 2016;9:1036. https://doi.org/10.3390/en9121036.

[199] Ju S-H, Huang Y-C. Analyses of offshore wind turbine structures with soil-structure interaction under earthquakes. Ocean Eng 2019;187:106190. https://doi.org/10.1016/j.oceaneng.2019.106190.

[200] Konstandakopoulou F, Konstantinidou M, Pnevmatikos N, Hatzigeorgiou GD. Safety and performance of offshore platforms subjected to repeated earthquakes. Infrastructures 2020;5. https://doi.org/10.3390/infrastructures5040038.

[201] Konstandakopoulou FD, Evangelinos KI, Nikolaou IE, Papagiannopoulos GA, Pnevmatikos NG. Seismic analysis of offshore platforms subjected to pulse-type ground motions compatible with European Standards. Soil Dyn Earthq Eng 2020;129:105713. https://doi.org/10.1016/j.soildyn.2019.105713.

[202] Cevasco D, Koukoura S, Kolios AJ. Reliability, availability, maintainability data review for the identification of trends in offshore wind energy applications. Renew Sustain Energy Rev 2021;136:110414. https://doi.org/10.1016/j.rser.2020.110414.

[203] Luczak MM, Telega J, Zagato N, Mucchi E. On the damage detection of a laboratory scale model of a tripod supporting structure by vibration-based methods. Mar Struct 2019;64:146–60. https://doi.org/10.1016/j.marstruc.2018.11.002.

[204] Tcherniak D, Chauhan S, Hansen MH. Applicability limits of operational modal analysis to operational wind turbines. Conf. Proc. Soc. Exp. Mech. Ser., vol. 1, Jacksonville, Florida USA: 2011, p. 317–27.

[205] Liu G, Zhai Y, Leng D, Tian X, Mu W. Research on structural damage detection of offshore platforms based on grouping modal strain energy. Ocean Eng 2017;140:43–9. https://doi.org/10.1016/j.oceaneng.2017.05.021.

[206] Xu M, Wang S, Li H. A residual strain energy based damage localisation method for offshore platforms under environmental variations. Ships Offshore Struct 2018;14:747–54. https://doi.org/10.1080/17445302.2018.1558727.

[207] Li Y, Zhang M, Yang W. Numerical and experimental investigation of modal-energy-based damage localization for offshore wind turbine structures. Adv Struct Eng 2018;21:1510–25. https://doi.org/10.1177/1369433217750725.

[208] Fathi A, Esfandiari A, Fadavie M, Mojtahedi A. Damage detection in an offshore platform using incomplete noisy FRF data by a novel Bayesian model updating method. Ocean Eng 2020;217:108023. https://doi.org/10.1016/j.oceaneng.2020.108023.

[209] Mousavi Z, Varahram S, Ettefagh MM, Sadeghi MH, Razavi SN. Deep neural networks–based damage detection using vibration signals of finite element model and real intact state: An evaluation via a lab-scale offshore jacket structure. Struct Heal Monit 2020. https://doi.org/10.1177/1475921720932614.

[210] Oliveira G, Magalhães F, Cunha Á, Caetano E. Vibration-based damage detection in a wind turbine using 1 year of data. Struct Control Heal Monit Hoboken, New Jersey, USA 2018;25:e2238. https://doi.org/10.1002/stc.2238.

[211] Weijtjens W, Verbelen T, Capello E, Devriendt C. Vibration based structural health monitoring of the substructures of five offshore wind turbines. Procedia Eng 2017;199:2294–9. https://doi.org/10.1016/j.proeng.2017.09.187.

[212] Yi JH. Laboratory tests on local damage detection for jacket-type offshore structures using optical FBG sensors based on statistical approaches. Ocean Eng 2016;124:94–103. https://doi.org/10.1016/j.oceaneng.2016.07.060.

[213] Mieloszyk M, Luczak MM, Mucchi E, Telega J, Ostachowicz WM. Damage detection in laboratory scale model of the offshore support structure using two different measurement techniques. 8th Eur Work Struct Heal Monit EWSHM 2016 2016;4:2430–9.

[214] Stubbs N, Farrar C, Kim J. Field Verification of a Nondestructive Damage Localization and Severity Estimation Algorithm, Nashville, Tennessee, USA: 1995.

[215] Wandji WN. Rayleigh's quotient-based damage detection algorithm: Theoretical concepts, computational techniques, and field implementation strategies. Struct Heal Monit 2017. https://doi.org/10.1177/1475921717691018.

[216] Cevasco D, Tautz-Weinert J, Kolios AJ, Smolka U. Applicability of machine learning approaches for structural damage detection of offshore wind jacket structures based on low resolution data. J Phys Conf Ser 2020;1618:022063. https://doi.org/10.1088/1742-6596/1618/2/022063.

[217] Jeong S, Kim EJ, Shin DH, Park JW, Sim SH. Data fusion-based damage identification for a monopile offshore wind turbine structure using wireless smart sensors. Ocean Eng 2020;195:106728. https://doi.org/10.1016/j.oceaneng.2019.106728.

[218] Fan W, Qiao P. Vibration-based Damage Identification Methods: A review and Comparative Study. Struct Heal Monit 2010. https://doi.org/10.1177/1475921710365419.

[219] Ziegler L, Cosack N, Kolios A, Muskulus M. Structural monitoring for lifetime extension of offshore wind monopiles: Verification of strain-based load extrapolation algorithm. Mar Struct 2019;66:154–63. https://doi.org/10.1016/j.marstruc.2019.04.003.

[220] Ivanhoe RO, Wang L, Kolios A. Generic framework for reliability assessment of offshore wind turbine jacket support structures under stochastic and time dependent variables. Ocean Eng 2020;216:107691. https://doi.org/10.1016/j.oceaneng.2020.107691.

[221] Shittu AA, Mehmanparast A, Shafiee M, Kolios A, Hart P, Pilario K. Structural reliability assessment of offshore wind turbine support structures subjected to pitting corrosion-fatigue: A damage tolerance modelling approach. Wind Energy 2020;23:2004–26. https://doi.org/10.1002/we.2542.

[222] Malekzehtab H, Golafshani AA. Damage detection in an offshore jacket platform using genetic

algorithm based finite element model updating with noisy modal data. Procedia Eng 2013;54:480–90. https://doi.org/10.1016/j.proeng.2013.03.044.

[223] Liu K, Yan R-J, Guedes Soares C. Damage identification in offshore jacket structures based on modal flexibility. Ocean Eng 2018;170:171–85. https://doi.org/10.1016/j.oceaneng.2018.10.014.

[224] Wang S, Zhang M, Li H. Damage Localization of an Offshore Platform considering Temperature Variations. Math Probl Eng 2015;2015:1–10. https://doi.org/10.1155/2015/954926.

[225] Nguyen CU, Huynh TC, Kim JT. Vibration-based damage detection in wind turbine towers using artificial neural networks. Struct Monit Maint 2018;5:507–19. https://doi.org/10.12989/smm.2018.5.4.507.

[226] Rytter A. Vibrational Based Inspection of Civil Engineering Structures. Aalborg Universitet, 1993.

[227] Martinez-Luengo M, Kolios A, Wang L. Structural Health Monitoring of Offshore Wind Turbines: A review through the Statistical Pattern Recognition Paradigm. Renew Sustain Energy Rev 2016;64:91–105. https://doi.org/10.1016/j.rser.2016.05.085.

[228] Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems 2016.

[229] Buatois A, Gibescu M, Rawn B, van der Meijden M. Analysis of North Sea Offshore Wind Power Variability. Resources 2014;3:454–70. https://doi.org/10.3390/resources3020454.

[230] Microsoft Corporation. Welcome to LightGBM's Documentation 2020. https://lightgbm.readthedocs.io/en/latest/#.

[231] Microsoft Corporation. Parameters 2020. https://lightgbm.readthedocs.io/en/latest/Parameters.html.

[232] Hyndman R, Athanasopoulos G. Forecasting: Principles and Practice. Monash University, Australia: 2018.

[233] Perktold J, Seabold S, Taylor J. About statsmodels 2019. https://www.statsmodels.org/stable/about.html#module-statsmodels.

[234] OpenCFD Ltd. About OpenFOAM 2017. http://www.openfoam.com/ (accessed May 18, 2017).

[235] Cavar D, Réthoré P, Bechmann A, Sørensen NN, Martinez B, Zahle F, et al. Comparison of OpenFOAM and EllipSys3D for Neutral Atmospheric Flow over Complex Terrain. Presentation 2013:18. https://doi.org/10.5194/wes-2016-3.

[236] Schmidt J, Stoevesandt B. Wind farm layout optimization with wakes from fluid dynamics simulations. Eur Wind Energy Assoc Conf Exhib 2014, EWEA 2014 2014. https://doi.org/10.13140/2.1.2544.3847.

[237] Schmidt J, Chang CY, Dörenkämper M, Salimi M, Teichmann T, Stoevesandt B. The consideration of atmospheric stability within wind farm AEP calculations. J Phys Conf Ser 2016;749. https://doi.org/10.1088/1742-6596/749/1/012002.

[238] Peralta C, Nugusse H, Kokilavani SP, Schmidt J, Stoevesandt B. Validation of the simpleFoam (RANS) solver for the atmospheric boundary layer in complex terrain. ITM Web Conf 2014;2:01002. https://doi.org/10.1051/itmconf/20140201002.

[239] Schmidt J, Stoevesandt B. Wind farm layout optimization in complex terrain with CFD wakes. Eur Wind Energy Assoc Annu Conf Exhib 2015, EWEA 2015 - Sci Proc 2015:166–9.

[240] Mintgen F, Hristov Y V. Lessons learned from more than 10 years use of OpenFOAM at Vestas. OpenFOAM Conf., Berlin: 2019.

[241] Hahn S, Machefaux E, Hristov Y V., Albano M, Threadgill R. Estimation of annual energy production using dynamic wake meandering in combination with ambient CFD solutions. J Phys Conf Ser 2016;753. https://doi.org/10.1088/1742-6596/753/3/032043.

[242] NREL. SOWFA 2015. https://nwtc.nrel.gov/SOWFA.

[243] Barthelmie RJ, Churchfield MJ, Moriarty PJ, Lundquist JK, Oxley GS, Hahn S, et al. The role of atmospheric stability/turbulence on wakes at the Egmond aan Zee offshore wind farm. J Phys Conf Ser 2015;625. https://doi.org/10.1088/1742-6596/625/1/012002.

[244] Jonkman J, Doubrawa P, Hamilton N, Annoni J, Fleming P. Validation of FAST.Farm Against Large-Eddy Simulations. J Phys Conf Ser 2018;1037:062005. https://doi.org/10.1088/1742-6596/1037/6/062005.

[245] Fleming PA, Gebraad PMO, Lee S, van Wingerden JW, Johnson K, Churchfield M, et al. Evaluating techniques for redirecting turbine wakes using SOWFA. Renew Energy 2014;70:211–8. https://doi.org/10.1016/j.renene.2014.02.015.

[246] Fleming P, Annoni J, Shah JJ, Wang L, Ananthan S, Zhang Z, et al. Field test of wake steering at an offshore wind farm. Wind Energy Sci 2017;2:229–39. https://doi.org/10.5194/wes-2-229-2017.

[247] SimScale. SimScale - Computational Fluid Dynamics Software n.d. https://www.simscale.com/product/cfd/.

[248] Chase AV. Computational Fluid Dynamic Modelling of a Small-Scale Vertical Axis Wind Turbine University of Vermont Computational Fluid Dynamic Modelling of a Small-Scale Vertical Axis Wind Turbine. University of Vermont, 2017.

[249] Winter M. Benchmark and Validation of open source CFD codes, with focus on compressible and rotating capabilities. Chalmers University of Technology, 2013. https://doi.org/10.1007/BF03192151.

[250] Jensen NO. A note on wind generator interaction. Risø-M-2411 Risø Natl Lab Roskilde 1983:1–

16. https://doi.org/Riso-M-2411.

[251] Katic, I., Hojstrup, J., Jensen NO. A simple model for cluster efficiency. Ewea 1986 1986:407–10.

[252] Choi J, Shan M. Advancement of Jensen (PARK) wake model. Eur. Wind Energy Conf. Exhib., 2013, p. 1790–7.

[253] Peña A, Rathmann O. Atmospheric stability-dependent infinite wind-farm models and the wake-decay coefficient. Wind Energy 2014;17:1269–85. https://doi.org/10.1002/we.1632.

[254] Crespo A, Hernandez J, Fraga E, Andreu C. Experimental validation of the UPM computer code to calculate wind turbine wakes and comparison with other models, Netherlands: 1988.

[255] Zhang MH. Wind Resource Assessment and Micro-siting: Science and Engineering. Singapore: Wiley; 2015.

[256] Subramanian G. Improving the prediction of pressure gradient field in wind farm numerical simulation method, WAKEFARM. University of Twente, 2018.

[257] Bot ETG. FarmFlow validation against four full scale wind farms. ECN 2015.

[258] Ansys. ANSYS Fluent 2019. https://www.ansys.com/products/fluids/ansys-fluent.

[259] ANSYS. ANSYS CFX 2019. https://www.ansys.com/products/fluids/ansys-cfx.

[260] ANSYS and SSE. Simulation of Wind Turbine Sites Increases. 2016.

[261] Argyle P. Computational fluid dynamics modelling of wind turbine wake losses in large offshore wind farms , incorporating atmospheric stability. Loughborough University, 2014.

[262] Tumenbayar F, Son J, Ko K. Comparison of CFD Prediction and Actual Condition for Wake Effect on an Onshore Wind Farm. Int J Renew ENERGY Res 2018;8.

[263] DTU Energy. Welcome to the world of WAsP by DTU Wind Energy n.d. http://www.wasp.dk/.

[264] DTU Wind Energy. Wake model for offshore wind farms n.d. https://www.wasp.dk/fuga.

[265] DNV-GL. Wind Resource Assessment software - WindFarmer: Analyst 2019. https://www.dnvgl.com/services/wind-resource-assessment-software-windfarmer-analyst-3766.

[266] CHAM. PHOENICS Overview n.d. http://www.cham.co.uk/phoenics/d_polis/d_docs/tr001/tr001.htm.

[267] Simisiroglou N, Karatsioris M, Nilsson K, Breton SP, Ivanell S. The Actuator Disc Concept in Phoenics. Energy Procedia 2016;94:269–77. https://doi.org/10.1016/j.egypro.2016.09.182.

[268] Irshad W. Wind Resource Assessment : Statistical and Computational Fluid-Dynamic Analysis. Edinburgh Napier University, 2012.

[269] Meissner C. WindSim: Getting started. Tønsberg: 2010. https://doi.org/10.1007/978-3-540-44397-1_3.

[270] Simisiroglou N. Wind power wake modelling. Uppsala Universitet, 2018.

[271] Dhunny AZ, Toja-Silva F, Peralta C, Lollchund MR, Rughooputh SDDV. Computational fluid dynamics simulation and full-scale experimental model inter-comparison of the wind flow around a university campus. Wind Eng 2017;41:43–54. https://doi.org/10.1177/0309524X16666460.

[272] Dhunny AZ, Lollchund MR, Rughooputh SDDV. Numerical analysis of wind flow patterns over complex hilly terrains: Comparison between two commonly used CFD software. Int J Glob Energy Issues 2016;39:181–203. https://doi.org/10.1504/IJGEI.2016.076339.

[273] Prospathopoulos JM, Papadakis G, Sieros G, Voutsinas SG, Chaviaropoulos TK, Diakakis K. Assessment of the aerodynamic characteristics of thick airfoils in high Reynolds and moderate Ma numbers using CFD modeling. J Phys Conf Ser 2014;524. https://doi.org/10.1088/1742-6596/524/1/012015.

[274] Barranger N, Ternisien T, Kallos G. An intercomparison study between RAMS and CRES-Flow-NS models and evaluation with wind tunnel experimental data: Toward improving atmospheric modeling for wind resource assessment. J Wind Eng Ind Aerodyn 2015;142:272–88. https://doi.org/10.1016/j.jweia.2015.04.003.

[275] Politis ES, Prospathopoulos J, Cabezon D, Hansen KS, Chaviaropoulos PK, Barthelmie RJ. Modeling wake effects in large wind farms in complex terrain: the problem, the methods and the issues. Wind Energy 2012;15:161–82. https://doi.org/10.1002/we.481.

[276] Prospathopoulos JM, Politis ES, Chaviaropoulos PK. Application of a 3D RANS solver on the complex hill of Bolund and assessment of the wind flow predictions. J Wind Eng Ind Aerodyn 2012;107–108:149–59. https://doi.org/10.1016/j.jweia.2012.04.011.

[277] Boersma S, Doekemeijer BM, Siniscalchi-Minna S, van Wingerden JW. A constrained wind farm controller providing secondary frequency regulation: An LES study. Renew Energy 2019;134:639–52. https://doi.org/10.1016/j.renene.2018.11.031.

[278] Boersma S, Rostampour V, van Wingerden BDJW, Keviczky T. A Model Predictive Wind Farm Controller with Linear Parameter-Varying Models∗. IFAC-PapersOnLine 2018;51:241–6. https://doi.org/10.1016/j.ifacol.2018.11.020.

[279] Witha B, Steinfeld G, Dörenkämper M, Heinemann D. Large-eddy simulation of multiple wakes in offshore wind farms. J Phys Conf Ser 2014;555. https://doi.org/10.1088/1742-6596/555/1/012108.

[280] Ramos DA, Guedes VG, Pereira RR, Valentim TA, Netto WA. Further considerations on WAsP,

OpenWind and WindSim comparison study: Atmospheric flow modelling over complex terrain and energy production estimate. Brazil Wind 2017. https://doi.org/10.1360/zd-2013-43-6-1064.

[281] Wagner M, Day J, Neumann F. A Fast and Effective Local Search Algorithm for Optimizing the Placement of Wind Turbines. BRAZIL Wind., 2012, p. 1–16.

[282] AWS Truepower. OpenWind - Theoretical Basis and Validation 2010:27.

[283] Gross G. The exploration of boundary layer phenomena using a nonhydrostatic mesoscale model. Meteorol Zeitschrift 2002;11:295–302. https://doi.org/10.1127/0941-2948/2002/0011-0295.

[284] Lin M, Porté-Agel F. Large-eddy simulation of yawedwind-turbine wakes: comparisons withwind tunnel measurements and analyticalwake models. Energies 2019;12:1–18. https://doi.org/10.3390/en12234574.

[285] Segalini A, Castellani F. Wind-farm simulation over moderately complex terrain. J Phys Conf Ser 2017;854:012042. https://doi.org/10.1088/1742-6596/854/1/012042.

[286] Segalini A, Dahlberg JA. Global Blockage Effects in Wind Farms. J Phys Conf Ser 2019;1256. https://doi.org/10.1088/1742-6596/1256/1/012021.

[287] Nathan J, Meyer Forsting AR, Troldborg N, Masson C. Comparison of OpenFOAM and EllipSys3D actuator line methods with (NEW) Mexico results. J Phys Conf Ser 2017;854. https://doi.org/10.1088/1742-6596/854/1/012033.

[288] Van Der Laan MP, Storey RC, Sørensen NN, Norris SE, Cater JE. A CFD code comparison of wind turbine wakes. J Phys Conf Ser 2014;524. https://doi.org/10.1088/1742-6596/524/1/012140.

[289] Ivanell S, Sørensen J, Henningson D. Numerical computations of wind turbine wakes. Tech Reports from KTH Mech 2007:1–91. https://doi.org/10.1007/978-3-540-33866-6_48.

[290] Uchida T, Ohya Y. Micro-siting technique for wind turbine generators by using large-eddy simulation. J Wind Eng Ind Aerodyn 2008;96:2121–38. https://doi.org/10.1016/j.jweia.2008.02.047.

[291] Uchida T, Ohya Y. Latest developments in numerical wind synopsis prediction using the RIAM-COMPACT® CFD model-design wind speed evaluation and wind risk (terrain-induced turbulence) diagnostics in Japan. Energies 2011;4:458–74. https://doi.org/10.3390/en4030458.

[292] Boersma S, Doekemeijer B, Vali M, Meyers J, van Wingerden J-W. A control-oriented dynamic wind farm model: WFSim. Wind Energy Sci 2018;3:75–95. https://doi.org/10.5194/wes-3-75-2018.

[293] Acker T, Chime AH. Wind Modeling using WindPro and WAsP Software. Flagstaff: 2011.

[294] Simisiroglou N. Wind Resource Assessment Comparison on a Complex Terrain Employing Windpro and Windsim 2012.

[295] Jonkman J. FAST.Farm n.d. https://nwtc.nrel.gov/FASTFarm.

[296] Jonkman J, Annoni J, Hayman G, Jonkman B, Purkayastha A. Development of FAST.Farm : A New Multiphysics Engineering Tool for Wind Farm Design and Analysis 2017:1–22.

[297] Jonkman J, Doubrawa P, Hamilton N, Annoni J, Fleming P. Validation of FAST.Farm Against Large-Eddy Simulations. J Phys Conf Ser 2018;1037:062005. https://doi.org/10.1088/1742-6596/1037/6/062005.

[298] Hao Y. Wind Farm Wake Modeling and Analysis of Wake Impacts in a Wind Farm. UMAss Amherst, 2016.

[299] Larsen GC, Madsen HA, Thomsen K, Larsen TJ. Wake meandering: a pragmatic approach. Wind Energy 2008;11:377–95. https://doi.org/10.1002/we.267.

[300] Schmidt B, Smolka U, Hartmann S. Validation of the Dynamic Wake Meander Model with AREVA M5000 Load Measurements at alpha ventus.pdf 2014:14.

[301] Lange B, Waldl HP, Guerrero AG, Heinemann D, Barthelmie RJ. Modelling of offshore wind turbine wakes with the wind farm program FLaP. Wind Energy 2003;6:87–104. https://doi.org/10.1002/we.84.

[302] King J, Fleming P. FLORIS Wake Modeling Utility 2019. https://floris.readthedocs.io/en/develop/.

[303] Simley E, Fleming P, King J. Design and Analysis of a Wake Steering Controller with Wind Direction Variability. Wind Energy Sci Discuss 2019:1–26. https://doi.org/10.5194/wes-2019-35.

[304] Schlez W. Virtual Wind Farm Simulation A Closer Look at the WakeBlaster Project. Wind. Int., 2017, p. 22–4.

[305] Grunnet JD, Soltani M, Knudsen T, Kragelund M, Bak T. Aeolus toolbox for dynamics wind farm model, simulation and control. Eur Wind Energy Conf Exhib 2010, EWEC 2010 2010;4:3119–29.

[306] SOLUTE. Furow - Micrositing 2015. https://furow.solute.es/en/furow/micrositing.

[307] Magdalena J. SOLUTE - FUROW Wind Energy Science Conference WESC 2017 Comparison of FUROW wind flow model with WAsP and OpenFOAM on Gaussian hills. WESC 2017, 2017.

[308] Poulsen T, Hasager CB. The (R)evolution of China: Offshore wind diffusion. Energies 2017;10:1–32. https://doi.org/10.3390/en10122153.

[309] Yang J, Liu Q, Li X, Cui X. Overview of wind power in China: Status and future. Sustain 2017;9:1–

12. https://doi.org/10.3390/su9081454.

[310] Chipindula J, Botlaguduru V, Du H, Kommalapati R, Huque Z. Life Cycle Environmental Impact of Onshore and Offshore Wind Farms in Texas. Sustainability 2018;10:2022. https://doi.org/10.3390/su10062022.

[311] Wiser R, Lantz E, Mai T, Zayas J, DeMeo E, Eugeni E, et al. Wind Vision: A New Era for Wind Power in the United States. Electr J 2015;28:120–32. https://doi.org/10.1016/j.tej.2015.09.016.

[312] Musial W. Offshore Wind Resource, Cost, and Economic Potential in the State of Maine. NREL 2015.

[313] Donovan A, Development G, Menard J, Development G. 2018 Massachusetts Offshore Wind Workforce Assessment. 2018.

[314] Gentils T, Wang L, Kolios A. Integrated structural optimisation of offshore wind turbine support structures based on finite element analysis and genetic algorithm. Appl Energy 2017;199:187–204. https://doi.org/10.1016/j.apenergy.2017.05.009.

[315] Wang L, Kolios A, Nishino T, Delafin PL, Bird T. Structural optimisation of vertical-axis wind turbine composite blades based on finite element analysis and genetic algorithm. Compos Struct 2016;153:123–38. https://doi.org/10.1016/j.compstruct.2016.06.003.

[316] Wang L, Quant R, Kolios A. Fluid structure interaction modelling of horizontal-axis wind turbine blades based on CFD and FEA. J Wind Eng Ind Aerodyn 2016;158:11–25. https://doi.org/10.1016/j.jweia.2016.09.006.

[317] Wang L, Liu X, Kolios A. State of the art in the aeroelasticity of wind turbine blades: Aeroelastic modelling. Renew Sustain Energy Rev 2016;64:195–210. https://doi.org/10.1016/j.rser.2016.06.007.

[318] Rasheed A, Holdahl R, Kvamsdal T, Åkervik E. A comprehensive simulation methodology for fluid-structure interaction of offshore wind turbines. Energy Procedia 2014;53:135–45. https://doi.org/10.1016/j.egypro.2014.07.222.

[319] Zhu R, Chen D, Wu S. Unsteady Flow and Vibration Analysis of the Horizontal-Axis Wind Turbine Blade under the Fluid-Structure Interaction 2019;2019.

[320] Zahle F, Martins JRRA, Section AD, Arbor A. Multipoint high-fidelity CFD-based aerodynamic shape optimization of a 10 MW wind turbine 2018:1–37.

[321] Réthoré P-E, Sørensen NN. Turbulence Closure Strategies for Modelling Wind Turbines Wake in Atmopsheric Turbulence. Euromech Colloq 508 Wind Turbine Wakes 2009:24.

[322] Réthoré P-E, Bechmann a, Sørensen NN, Frandsen ST, Mann J, Jørgensen HE, et al. A CFD model of the wake of an offshore wind farm: using a prescribed wake inflow. J Phys Conf Ser 2007;75:012047. https://doi.org/10.1088/1742-6596/75/1/012047.

[323] Wang L, Quant R, Kolios A. Fluid structure interaction modelling of horizontal-axis wind turbine blades based on CFD and FEA. J Wind Eng Ind Aerodyn 2016;158:11–25. https://doi.org/10.1016/j.jweia.2016.09.006.

[324] Réthoré P-E, Sørensen NN, Zahle F, Bechmann A, Madsen HA. CFD model of the MEXICO wind tunnel. Ewea 2011 2011:1–8.

[325] Creech A, Früh WG, Maguire AE. Simulations of an Offshore Wind Farm Using Large-Eddy Simulation and a Torque-Controlled Actuator Disc Model. Surv Geophys 2015;36:427–81. https://doi.org/10.1007/s10712-015-9313-7.

[326] Svenning E. Implementation of an actuator disk in OpenFOAM. 2010.

[327] Mittelmeier N, Blodau T, Steinfeld G, Rott A, Kühn M. An analysis of offshore wind farm SCADA measurements to identify key parameters influencing the magnitude of wake effects. J Phys Conf Ser 2016;753. https://doi.org/10.1088/1742-6596/753/3/032052.

[328] OpenFOAMWiki. SimpleFoam 2019. http://openfoamwiki.net/index.php/SimpleFoam.

[329] Mangani L, Bianchini C. Heat transfer application in turbomachinery. OpenFOAM Conf 2007:26–7.

[330] Ambatipudi V. SIMPLE Solver for Driven Cavity Flow Problem. Asme 2006.

[331] OpenFOAMWiki. OpenFOAM guide/The SIMPLE algorithm in OpenFOAM 2014. https://openfoamwiki.net/index.php/OpenFOAM_guide/The_SIMPLE_algorithm_in_OpenFOAM.

[332] Stergiannis N, van Beeck J, Runacres MC. Full HAWT rotor CFD simulations using different RANS turbulence models compared with actuator disk and experimental measurements. Wind Energy Sci Discuss 2017:1–20. https://doi.org/10.5194/wes-2017-6.

[333] Digraskar D a. a. Simulations of flow over wind turbines. University of Massachusetts Amherst, 2010.

[334] Federer H. THE GAUSS-GREEN THEOREM 1944;55:438–56.

[335] Greenshields C. OpenFOAM v6 User Guide: 4.4 Numerical schemes 2018. https://cfd.direct/openfoam/user-guide/v6-fvschemes/.

[336] El Kasmi A, Masson C. An extended k - ε{lunate} model for turbulent flow through horizontal-axis wind turbines. J Wind Eng Ind Aerodyn 2008;96:103–22. https://doi.org/10.1016/j.jweia.2007.03.007.

[337] Shih T-H, Liou WW, Shabbir A, Yang Z, Zhu J. A New k-\epsilon Eddy Viscosity For High Reynolds Number Turbulent Flows. Compurers Fluids 1995;24:227–38. https://doi.org/10.1063/1.4799179.

[338] Fluent inc. 12.4.3 Realizable k-e Model 2006. https://www.sharcnet.ca/Software/Fluent6/html/ug/node480.htm (accessed May 2, 2019).

[339] Menter FR. Two-equation eddy-viscosity turbulence models for engineering applications. AIAA J 1994;32:1598–605. https://doi.org/10.2514/3.12149.

[340] Wilcox DC. Formulation of the k-w Turbulence Model Revisited. AIAA J 2008;46:2823–38. https://doi.org/10.2514/1.36541.

[341] OpenFOAM. atmBoundaryLayer Class Reference n.d. https://www.openfoam.com/documentation/guides/latest/api/classFoam_1_1atmBoundaryLayer.html.

[342] Richards PJ, Hoxey RP. Appropriate boundary conditions for computational wind engineering models using the k-E turbulence model. J Wind Eng Ind Aerodyn 1993;46–47:145–53. https://doi.org/10.1016/0167-6105(93)90124-7.

[343] Jonkman J, Butterfield S, Musial W, Scott G. Definition of a 5-MW reference wind turbine for offshore system development. Contract 2009:1–75. https://doi.org/10.1002/ajmg.10175.

[344] Burton T, Jenkins N, Sharpe D, Bossanyi E. Wind Energy Handbook. 2011.

[345] Do CB, Lee H. Section notes 9 - Gaussian Processes. Lect Notes 2008:1–14.

[346] Davis RA. Gaussian Processes. Encycl Environmetrics Stoch Model Environ Chang 2006:6.

[347] Rasmussen CE, Williams CKI. Gaussian Process For Machine Learning. 2006. https://doi.org/10.1093/bioinformatics/btq657.

[348] Pandit RK, Infield D. Comparative analysis of Gaussian Process power curve models based on different stationary covariance functions for the purpose of improving model accuracy. Renew Energy 2019;140:190–202. https://doi.org/10.1016/j.renene.2019.03.047.

[349] MathWorks. Kernel (Covariance) Function Options 2019. https://www.mathworks.com/help/stats/kernel-covariance-function-options.html.

[350] MathWorks. MathWorks TreeBagger predict 2019. https://www.mathworks.com/help/stats/treebagger.predict.html.

[351] ASME. Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer 2009.

[352] Roache PJ. Verification and Validation in Computational Science and Engineering. Albuquerque, New Mexico: Hermosa Publishers; 1998.

[353] Avila M, Gargallo-Peiró A, Folch A. A CFD framework for offshore and onshore wind farm simulation. J Phys Conf Ser 2017;854. https://doi.org/10.1088/1742-6596/854/1/012002.

[354] Poulsen T, Hasager C, Jensen C. The Role of Logistics in Practical Levelized Cost of Energy Reduction Implementation and Government Sponsored Cost Reduction Studies: Day and Night in Offshore Wind Operations and Maintenance Logistics. Energies 2017;10:464. https://doi.org/10.3390/en10040464.

[355] Delafin PL, Nishino T, Kolios A, Wang L. Comparison of low-order aerodynamic models and RANS CFD for full scale 3D vertical axis wind turbines. Renew Energy 2017;109:564–75. https://doi.org/10.1016/j.renene.2017.03.065.

[356] Delafin P, Nishino T, Wang L, Kolios A. Effect of the number of blades and solidity on the performance of a vertical axis wind turbine. J Phys Conf Ser 2016;753:022033. https://doi.org/10.1088/1742-6596/753/2/022033.

[357] Delafin P, Nishino T, Wang L, Kolios A, Bird T. Comparison of RANS CFD and lower-order aerodynamic models for 3D Vertical Axis Wind Turbines. Eur Wind Energy Conf Exhib 2015:2–8.

[358] Richmond M, Balaam T, Causon P, Cevasco D, Leimeister M, Kolios A, et al. Multi-Criteria Decision Analysis for Benchmarking Human-Free Lifting Solutions in the Offshore Wind Energy Environment. Energies 2018;11:1175. https://doi.org/10.3390/en11051175.

[359] Leimeister M, Kolios A. A review of reliability-based methods for risk analysis and their application in the offshore wind industry. Renew Sustain Energy Rev 2018;91:1065–76. https://doi.org/10.1016/j.rser.2018.04.004.

[360] Queipo N V., Haftka RT, Shyy W, Goel T, Vaidyanathan R, Kevin Tucker P. Surrogate-based analysis and optimization. Prog Aerosp Sci 2005;41:1–28. https://doi.org/10.1016/j.paerosci.2005.02.001.

[361] Rusu C, Rusu V. Radial basis functions versus geostatistics in spatial interpolations. IFIP Int Fed Inf Process 2006;217:119–28. https://doi.org/10.1007/978-0-387-34747-9_13.

[362] Gunst RF. Response Surface Methodology: Process and Product Optimization Using Designed Experiments. Technometrics 1996;38:284–6. https://doi.org/10.1080/00401706.1996.10484509.

[363] Rosenblatt F. The Perceptron a Perceiving and Recognizing Automation. Buffalo, NY: 1957.

[364] McCulloch WS, Pitts WH. A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 1943;5:115–33.

[365] Schlechtingen M, Ferreira Santos I. Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. Mech Syst Signal Process 2011;25:1849–75. https://doi.org/10.1016/j.ymssp.2010.12.007.

[366] Lu B, Li Y, Wu X, Yang Z. A review of recent advances in wind turbine condition monitoring and fault diagnosis. Electron Mach Wind 2009:1–7. https://doi.org/10.1109/PEMWA.2009.5208325.

[367] Dervilis N, Choi M, Taylor SG, Barthorpe RJ, Park G, Farrar CR, et al. On damage diagnosis for a wind turbine blade using pattern recognition. J Sound Vib 2014;333:1833–50. https://doi.org/10.1016/j.jsv.2013.11.015.

[368] Gonzalez E, Reder M, Melero JJ. SCADA alarms processing for wind turbine component failure detection. J Phys Conf Ser 2016;753. https://doi.org/10.1088/1742-6596/753/7/072019.

[369] Helbing G, Ritter M. Deep Learning for fault detection in wind turbines. Renew Sustain Energy Rev 2018;98:189–98. https://doi.org/10.1016/j.rser.2018.09.012.

[370] Dervilis N, Worden K, Rogers T, Scott D, Simpson T. Machine Learning for Energy Load Forecasting. J Phys Conf Ser 2018;1106:012005. https://doi.org/10.1088/1742-6596/1106/1/012005.

[371] Smolka U, Kaufer D, Cheng PW. Are sea state measurements required for fatigue load monitoring of offshore wind turbines? J Phys Conf Ser 2014;555:0–10. https://doi.org/10.1088/1742-6596/555/1/012095.

[372] Li S, Wunsch DC, O'Hair E, Giesselmann MG. Comparative Analysis of Regression and Artificial Neural Network Models for Wind Turbine Power Curve Estimation. J Sol Energy Eng 2001;123:327. https://doi.org/10.1115/1.1413216.

[373] Ramasamy P, Chandel SS, Yadav AK. Wind speed prediction in the mountainous region of India using an artificial neural network model. Renew Energy 2015;80:338–47. https://doi.org/10.1016/j.renene.2015.02.034.

[374] Bechrakis DA, Sparis PD. Wind speed prediction using Artificial Neural Networks. Wind Eng 1998;22:287–95.

[375] O'Hagan A. Curve Fitting and Optimal Design for Prediction. J R Stat Soc Ser B 1978;40:1–24. https://doi.org/10.1111/j.2517-6161.1978.tb01643.x.

[376] Lee D, Baldick R. Short-term wind power ensemble prediction based on gaussian processes and Neural networks. IEEE Trans Smart Grid 2014;5:501–10. https://doi.org/10.1109/TSG.2013.2280649.

[377] Hu J, Wang J. Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression. Energy 2015;93:1456–66. https://doi.org/10.1016/j.energy.2015.10.041.

[378] Chen N, Qian Z, Nabney IT, Meng X. Wind power forecasts using gaussian processes and numerical weather prediction. IEEE Trans Power Syst 2014;29:656–65. https://doi.org/10.1109/TPWRS.2013.2282366.

[379] Hart E, Leithead WE, Feuchtwang J. Wind turbine C p , max and drivetrain-losses estimation using Gaussian process machine learning 2018;1037:1–8. https://doi.org/10.1088/1742-6596/1037/3/032024.

[380] Li Y, Liu S, Shu L. Wind turbine fault diagnosis based on Gaussian process classifiers applied to operational data. Renew Energy 2019;134:357–66. https://doi.org/10.1016/j.renene.2018.10.088.

[381] Pandit RK, Infield D. SCADA-based wind turbine anomaly detection using Gaussian process models for wind turbine condition monitoring purposes. IET Renew Power Gener 2018;12:1249–55. https://doi.org/10.1049/iet-rpg.2018.0156.

[382] Pandit RK, Infield D, Kolios A. Comparison of advanced non-parametric models for wind turbine power curves. IET Renew Power Gener 2019;13:1503–10. https://doi.org/10.1049/iet-rpg.2018.5728.

[383] scikit learn. sklearn.gaussian_process.GaussianProcessRegression n.d. http://scikit-learn.org/stable/modules/generated/sklearn.gaussian_process.GaussianProcessRegressor.html#sklearn.gaussian_process.GaussianProcessRegressor (accessed February 22, 2018).

[384] scikit learn. Radial-basis function (RBF) kernel n.d. https://scikit-learn.org/stable/modules/gaussian_process.html#radial-basis-function-rbf-kernel.

[385] Broomhead DS, Lowe D. Multivariable Functional Interpolation and Adaptive Networks. Complex Syst 1988;2:321–55. https://doi.org/10.1126/science.1179047.

[386] Tino P, Benuskova L, Sperduti A. Springer Handbook of Computational Intelligence. vol. 8. Berlin, Heidelberg: Springer Berlin Heidelberg; 2015. https://doi.org/10.1007/978-3-662-43505-2.

[387] Orr MJLM. Introduction to radial basis function networks. 1996. https://doi.org/10.1109/MED.2007.4433897.

[388] Schaback R. A Practical Guide to Radial Basis Functions. Sci Comput with Radial Basis Funct 2007:1–58. https://doi.org/https://num.math.uni-goettingen.de/schaback/teaching/sc.pdf.

[389] Zhu Q, Lin HS. Comparing ordinary kriging and regression kriging for soil properties in contrasting landscapes. Pedosphere 2010;20:594–606. https://doi.org/10.1016/S1002-0160(10)60049-5.

[390] Mondrago M. Probabilistic Modelling of Geotechnical Conditions for Offshore Wind Turbine Support Structures. Cranfield University, 2014.

[391] Roodsari BN, Macnab CJB, Nowicki EP. A novel adaptive controller using radial basis function neural network for the wind energy conversion system. Proc IEEE Int Conf Ind Technol 2017:715–20. https://doi.org/10.1109/ICIT.2017.7915447.

[392] Han B, Zhou L, Zhang Z. LIDAR-assisted radial basis function neural network optimization for wind turbines. IEEJ Trans Electr Electron Eng 2017;13:195–200. https://doi.org/10.1002/tee.22514.

[393] Fornberg B, Piret C. On choosing a radial basis function and a shape parameter when solving a

convective PDE on a sphere. J Comput Phys 2008;227:2758–80. https://doi.org/10.1016/j.jcp.2007.11.016.

[394] Hastie T, Tibshirani R, Friedman J. Random Forests. Elem. Stat. Learn. Data Mining, Inference, Predict., 2009, p. 587–605. https://doi.org/10.1007/b94608_15.

[395] Shotton J, Fitzgibbon A, Cook M, Sharp T, Finocchio M, Moore R, et al. Real-time human pose recognition in parts from single depth images. Cvpr 2011 2011:1297–304. https://doi.org/10.1109/CVPR.2011.5995316.

[396] Zhang L, Liu K, Wang Y, Omariba ZB. Ice detection model of wind turbine blades based on random forest classifier. Energies 2018;11. https://doi.org/10.3390/en11102548.

[397] Leahy K, Gallagher C, Bruton K, O'Donovan P, O'Dullivan DTJ. Automatically Identifying and Predicting Unplanned Wind Turbine Stoppages Using SCADA and Alarms System Data: Case Study and Results. J Phys Conf Ser 2017;926. https://doi.org/10.1088/1742-6596/926/1/012011.

[398] Scikit learn. 3.2.4.3.2. sklearn.ensemble.RandomForestRegression n.d. http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html (accessed February 22, 2018).

[399] Vapnik V. The Nature of Statistical Learning Theory. Springer; 1995.

[400] Chen P-H, Fan R-E, Lin C-J. A study on SMO-type decomposition methods for support vector machines. IEEE Trans Neural Netw 2006;17:893–908. https://doi.org/10.1109/TNN.2006.875973.

[401] Suthaharan S. Machine Learning Models and Algorithms for Big Data Classification. vol. 36. 2016. https://doi.org/10.1007/978-1-4899-7641-3.

[402] Smola AJ, Sch B. A tutorial on support vector regression. Canberra: 2003.

[403] Kramer O, Gieseke F. Short-term wind energy forecasting using support vector regression. Adv Intell Soft Comput 2011;87:271–80. https://doi.org/10.1007/978-3-642-19644-7_29.

[404] Zendehboudi A, Baseer MA, Saidur R. Application of support vector machine models for forecasting solar and wind energy resources: A review. J Clean Prod 2018;199:272–85. https://doi.org/10.1016/j.jclepro.2018.07.164.

[405] Zeng J, Qiao W. Support vector machine-based short-term wind power forecasting. 2011 IEEE/PES Power Syst Conf Expo PSCE 2011 2011;0511:1–8. https://doi.org/10.1109/PSCE.2011.5772573.

[406] Santos P, Villa LF, Reñones A, Bustillo A, Maudes J. An SVM-based solution for fault detection in wind turbines. Sensors (Switzerland) 2015;15:5627–48. https://doi.org/10.3390/s150305627.

[407] Laouti N, Sheibat-Othman N, Othman S. Support vector machines for fault detection in wind turbines. IFAC Proc Vol 2011;44:7067–72. https://doi.org/10.3182/20110828-6-IT-1002.02560.

[408] scikit learn. sklearn.svm.SVR n.d. http://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html (accessed February 22, 2018).

[409] Shao Z, Wu Y, Li L, Han S, Liu Y. Multiple Wind Turbine Wakes Modeling Considering the Faster Wake Recovery in Overlapped Wakes. Energies 2019;12:680. https://doi.org/10.3390/en12040680.

[410] Gaumond M, Réthoré P-E, Ott S, Peña A, Bechmann A, Hansen KS. Evaluation of the wind direction uncertainty and its impact on wake modeling at the Horns Rev offshore wind farm. Wind Energy 2014;17:1169–78. https://doi.org/10.1002/we.1625.

[411] Katic I, Højstrup J, Jensen NO. A Simple Model for Cluster Efficiency. EWEC'86. Proc., vol. 1, 1986.

[412] Grinstead C, Snell J. Introduction to Probability. 1988.

[413] Tagliaferri F, Hayes BP, Viola IM, Djokić SZ. Wind modelling with nested Markov chains. J Wind Eng Ind Aerodyn 2016;157:118–24. https://doi.org/10.1016/j.jweia.2016.08.009.

[414] Scheu MN, Kolios A, Fischer T, Brennan F. Influence of statistical uncertainty of component reliability estimations on offshore wind farm availability. Reliab Eng Syst Saf 2017;168:28–39. https://doi.org/10.1016/j.ress.2017.05.021.

[415] Chen P, Berthelsen KK, Bak-Jensen B, Chen Z. Markov model of wind power time series using Bayesian inference of transition matrix. IECON Proc (Industrial Electron Conf 2009:627–32. https://doi.org/10.1109/IECON.2009.5414993.

[416] Brokish K, Kirtley J. Pitfalls of modeling wind power using Markov chains. 2009 IEEE/PES Power Syst Conf Expo PSCE 2009 2009. https://doi.org/10.1109/PSCE.2009.4840265.

[417] Pesch T, Schröders S, Allelein HJ, Hake JF. A new Markov-chain-related statistical approach for modelling synthetic wind power time series. New J Phys 2015;17. https://doi.org/10.1088/1367-2630/17/5/055001.

[418] Esteoule T, Bernon C, Barthod M. Improving wind power forecasting through cooperation: A case-study on operating farms. Proc Int Jt Conf Auton Agents Multiagent Syst AAMAS 2019;4:1940–2.

[419] Hong T, Pinson P, Fan S. Global energy forecasting competition 2012. Int J Forecast 2014;30:357–63. https://doi.org/10.1016/j.ijforecast.2013.07.001.

[420] Hong T, Pinson P, Fan S, Zareipour H, Troccoli A, Hyndman RJ. Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond. Int J Forecast 2016;32:896–913. https://doi.org/10.1016/j.ijforecast.2016.02.001.

[421] Persson C, Bacher P, Shiga T, Madsen H. Multi-site solar power forecasting using gradient boosted regression trees. Sol Energy 2017;150:423–36. https://doi.org/10.1016/j.solener.2017.04.066.

[422] Fischer P, Etienne C, Tian J, Krauß T. Prediction of wind speeds based on digital elevation models using Boosted Regression Trees. Int Arch Photogramm Remote Sens Spat Inf Sci - ISPRS Arch 2015;40:197–202. https://doi.org/10.5194/isprsarchives-XL-1-W5-197-2015.

[423] Cai L, Gu J, Ma J, Jin Z. Probabilistic wind power forecasting approach via instance-based transfer learning embedded gradient boosting decision trees. Energies 2019;12. https://doi.org/10.3390/en12010159.

[424] International Energy Agency. Estimation of cost of energy from windenergy conversion system 1994.

[425] Bilir L, Imir M, Devrim Y, Albostan A. Seasonal and yearly wind speed distribution and wind power density analysis based on Weibull distribution function. Int J Hydrogen Energy 2015;40:15301–10. https://doi.org/10.1016/j.ijhydene.2015.04.140.

[426] Perea-Moreno AJ, Alcalá G, Hernandez-Escobedo Q. Seasonal wind energy characterization in the Gulf of Mexico. Energies 2019;13:1–21. https://doi.org/10.3390/en13010093.

[427] Erickson DJ, Taylor JA. Non-Weibull behavior observed in a model-generated global surface wind field frequency distribution. J Geophys Res 1989;94:12693. https://doi.org/10.1029/JC094iC09p12693.

[428] Drobinski P, Coulais C, Jourdier B. Surface Wind-Speed Statistics Modelling: Alternatives to the Weibull Distribution and Performance Evaluation. Boundary-Layer Meteorol 2015;157:97–123. https://doi.org/10.1007/s10546-015-0035-7.

[429] Tolver A. An Introduction to Markov Chains. An Introd. to Markov Chain., Copenhagen: 2016, p. 15.

[430] Serfozo R. Probability and its Applications. 2009. https://doi.org/10.1017/cbo9780511606656.010.

[431] Friedman JH. Greedy Function Approximation: A Gradient Boosting Machine. 1999.

[432] Gut A. An Intermediate Course in Probability. Springer; 2009.

[433] Martinez-Luengo M, Shafiee M. Guidelines and Cost-Benefit Analysis of the Structural Health Monitoring Implementation in Offshore Wind Turbine Support Structures. Energies 2019;12:1176. https://doi.org/10.3390/en12061176.

[434] Kolios AJ, Smolka U. Risk-based maintenance strategies for offshore wind energy assets. Proc - Annu Reliab Maintainab Symp 2020;2020-Janua. https://doi.org/10.1109/RAMS48030.2020.9153642.

[435] Song Y, Liang L, Du Y, Sun B. Railway polygonized wheel detection based on numerical time-frequency analysis of axle-box acceleration. Appl Sci 2020;10. https://doi.org/10.3390/app10051613.

[436] Pastor M, Binda M, Harčarik T. Modal Assurance Criterion. Procedia Eng 2012;48:543–8. https://doi.org/10.1016/j.proeng.2012.09.551.

[437] Kuhn HW. Variants of the hungarian method for assignment problems. Nav Res Logist Q 1956;3:253–8. https://doi.org/10.1002/nav.3800030404.

[438] BLADT Industries. Wikinger Offshore Wind Farm 2017. https://www.bladt.dk/UserFiles/05_Downloads/Foundations/2018_Update/Wikinger_Blaa_bjaelke_ver03_Low.pdf (accessed June 4, 2019).

[439] Iberdrola. Wikinger, the project that consolidates Germany as a strategic market n.d. https://www.iberdrola.com/about-us/lines-business/flagship-projects/wikinger-offshore-wind-farm.

[440] Ramboll. ROSA, Program ROSA Structural Analysis, User's guide. 2019.

[441] Passon P, Branner K, Larsen SE, Rasmussen HJ. Offshore Wind Turbine Foundation Design. Technical University of Denmark - DTU, 2015.

[442] American Petroleum Institute. Recommended Practice for Planning, Designing and Constructing Fixed Offshore Platforms, API RP2A-WSD, 21st edition. 2000.

[443] Scheu MN, Tremps L, Smolka U, Kolios A, Brennan F. A systematic Failure Mode Effects and Criticality Analysis for offshore wind turbine systems towards integrated condition based maintenance strategies. Ocean Eng 2019;176:118–33. https://doi.org/10.1016/j.oceaneng.2019.02.048.

[444] Luengo MM, Kolios A. Failure mode identification and end of life scenarios of offshore wind turbines: A review. Energies 2015;8:8339–54. https://doi.org/10.3390/en8088339.

[445] Weijtjens W, Verbelen T, De Sitter G, Devriendt C. Foundation structural health monitoring of an offshore wind turbine--a full-scale case study. Struct Heal Monit 2015:1–14. https://doi.org/10.1177/1475921715586624.

[446] Guo J, Wu J, Guo J, Jiang Z. A Damage Identification Approach for Offshore Jacket Platforms Using Partial Modal Results and Artificial Neural Networks. Appl Sci 2018. https://doi.org/10.3390/app8112173.

[447] Budyanas RG, Nisbett JK. Shigley's Mechanical Engineering Design. 9th ed. NY: McGraw-Hill Book Company; 2011.

[448] Adedipe O, Brennan F, Mehmanparast A, Kolios A, Tavares I. Corrosion fatigue crack growth mechanisms in offshore monopile steel weldments. Fatigue Fract Eng Mater Struct 2017;40:1868–81. https://doi.org/10.1111/ffe.12606.

[449] Martinez-Luengo M, Causon P, Gill AB, Kolios AJ. The effect of marine growth dynamics in offshore wind turbine support structures. Prog. Anal. Des. Mar. Struct. - Proc. 6th Int. Conf. Mar. Struct. MARSTRUCT, May 8 - 10, 2017, Lisbon, Portugal: n.d., p. 889–98.

[450] Jusoh I, Wolfram J. Effects of Marine Growth and Hydrodynamic Loading on Offshore Structures. J Mek 1996.

[451] Pascual G. INVESTIGATION ON THE INFLUENCE OF MARINE GROWTH ON THE DYNAMICS OF FOWT. Cranfield University, 2018.

[452] Anton H, Rorres C. Elementary Linear Algebra. vol. 3. 11th ed. Wiley; 2014.

[453] Weinert J, Smolka U, Schümann B, Cheng PW. Detecting Critical Scour Developments at Monopile Foundations Under Operating Conditions. Proc Eur Wind Energy Assoc Annu Event, EWEA 2015 2015:135–9.

[454] Häckell MW, Rolfes R. Monitoring a 5 MW offshore wind energy converter - Condition parameters and triangulation based extraction of modal parameters. Mech Syst Signal Process 2013;40:322–43. https://doi.org/10.1016/j.ymssp.2013.04.004.

[455] Dong X, Lian J, Yang M, Wang H. Operational modal identification of offshore wind turbine structure based on modified stochastic subspace identification method considering harmonic interference. J Renew Sustain Energy 2014;6. https://doi.org/10.1063/1.4881876.

[456] Allemang RJ. The Modal Assurance Criterion – Twenty Years of Use and Abuse. vol. 1, 2003, p. 14–21.

[457] Bundesamt für Seeschifffahrt und Hydrographie (BSH). Minimum Requirements Concerning the Constructive Design of Offshore Structures within the Exclusive Economic Zone. Hamburg, Germany: 2015.

[458] Allemang RJ, Brown DL. A correlation coefficient for modal vector analysis. Comput Sci 1982.

[459] Ciang CC, Lee JR, Bang HJ. Structural health monitoring for a wind turbine system: a review of damage detection methods. Meas Sci Technol 2008;19:122001. https://doi.org/10.1088/0957-0233/19/12/122001.

[460] scikit learn. Scikit learn KMeans 2019. https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html.

Appendix

A.1 Additional research, damage detectability in the presence of scatter

Additionally, there other research has been conducted but has not yet been completed and so has not been made part of this thesis.

A.1.1 Additional research objective

In chapter 6, the potential for damage detection through global dynamics was presented. While this research showed the change in modal dynamics as a result of changes in environmental conditions and damage situations, it did not show whether this damage is likely to be detectable in the presence of natural scatter observed in measured data. Further research was conducted to answer this question and provide an estimate for how large damage should be in order to be reliably detected.

A.1.2 Method

Damage to the structure, in this case a crack, is simulated in increments as it progresses from the design condition to complete loss of the joint. This gradually changes the modal parameters. The crack is simulated as a reduction in stiffness of the joint. The frequency is shown in the figure where the design value for stiffness is on the right and this reduces towards the left. The same modal parameters are used in this case as were used in the previous modal properties study.

Modal properties obtained through operational modal analysis of real structures contain scatter. For the level 3 damage severity approach it is necessary to replicate that scatter in the modelled, synthetic data. The NFA results are deterministic and do not have any scatter. So, the NFA results are combined with observations from real OMA results in order to fit statistical distributions which are then sampled in order to generate 'scattered' modal property observations.

The method followed is shown diagrammatically in . The first two steps are the same as were followed in chapter 6, where an intact structure is modelled, and the modal parameters are calculated through solving the eigenvalue problem. The modal parameters are also calculated for a simulated damage state. The new approach is to the simulate scatter in the 'damaged' modal parameters. The simulated, scattered data is created by fitting a normal distribution to the modal parameters, with the mean set as the deterministic value, the standard deviation is based on the level of deviation which was observed in chapter 7 where real accelerometer measurements from an offshore substation at the Wikinger site were investigated. The scattered values were sampled from these fitted distributions. Next, machine learning, classification, models were trained on the scattered data and the test ability of these models to predict the damage state was judged.

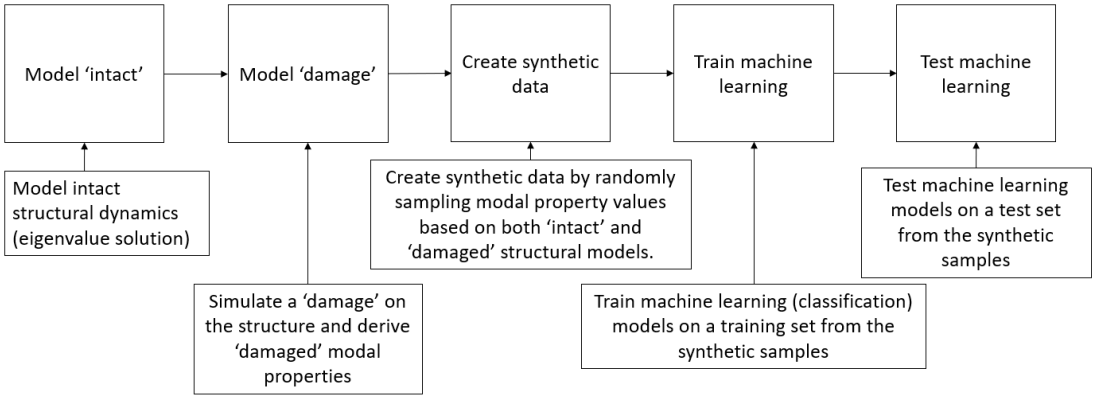


Figure 67. Method followed to investigate the potential for damage detection from modal parameters in the presence of scatter.

An example of the scattered sate is shown in . In chapter 7 it was observed that the level of scatter is higher for higher eigenfrequency modes and so this was replicated in the approach by setting the standard deviation as proportional to the mean value of frequency.

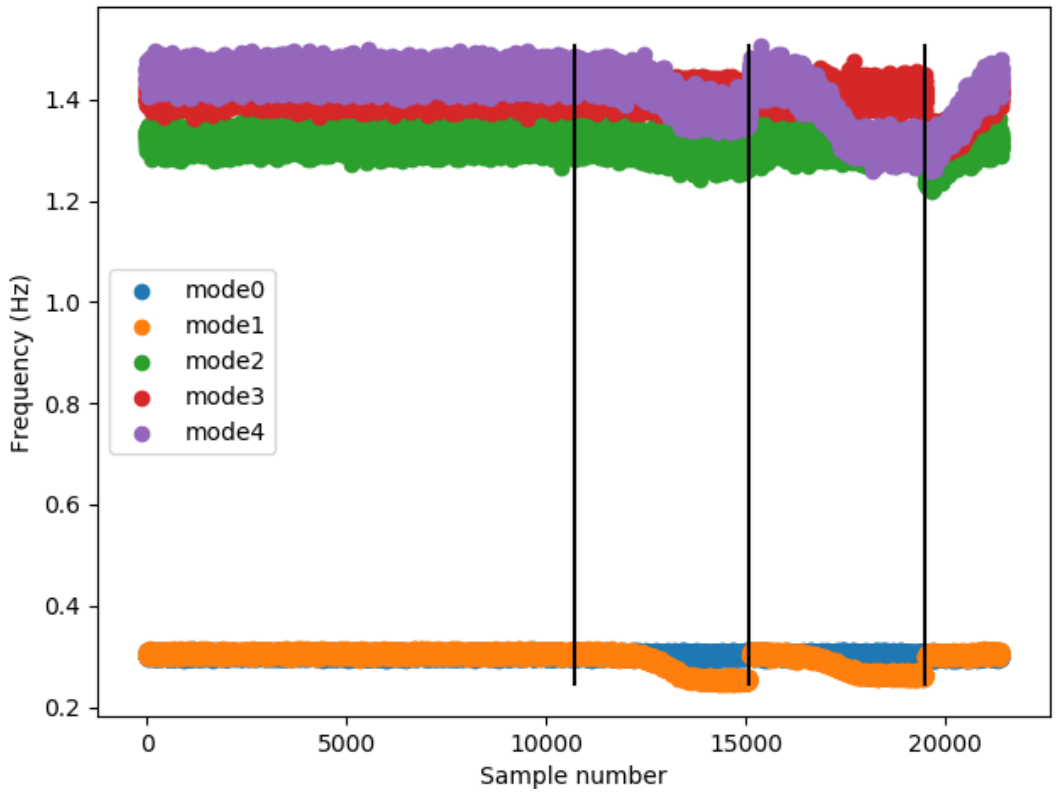


Figure 68. Scattered data generated in order to replicate real scatter in a modelled damage scenario. This shows the eigenfrequency for three damage cases as well as 'healthy' scattered data.

A.1.4 Machine learning models tested

The two machine learning approaches used for this study are 1) random forest classification and 2) logistic regression classification. The motivation for this choice is as follows:

1) Random Forest

This represents the ideal case as the model does not have a predefined shape it is able to fit closely to the specific data used. It should have a higher prediction accuracy but not be as representative of a real situation where the real damage response of the structure, rather than modelled, is unknown.

In summary:

- Fits well to the training data and accurately predicts the test data
- High bias, suited to this data but conclusions cannot be extended beyond this case
- Ideal case

2) **Logistic regression classification**

This is a simpler model where the decision boundary is defined linearly. This is useful because it is more representative of the performance which might be expected in a real situation, because it doesn't fit the data as closely. Additionally, this approach is interesting because the linear nature of the model means that several metrics and variable reduction approaches are valid. The variable reduction approach itself can provide useful insights into damage detection.

In summary:

- Simple model with linear decision boundaries, low variance, low bias
- Allows for methods to reduce variables (Information value and P-value)
- Gives probabilities and not just classification
- Findings are more generic

A.1.5 Results

A.1.5.1 Random Forest Classification results

This shows the confusion matrix from random forest where all damage levels are included in a multi-class classification. It shows that early on, there are a large number of false negative results where minor damage is classified as healthy. This is in the bottom right of the plot. As the damage becomes more severe it becomes more detectible and, beyond a point, the level of damage is very accurately predicted. The most important part of this graph is the bottom right, damage should be detected as soon as possible because there can be sudden failure if a crack grows too large. How realistic the middle section of the confusion matrix is subject to how well the behaviour of the model replicates the behaviour of the real structure and model updating is only conducted for the healthy case.

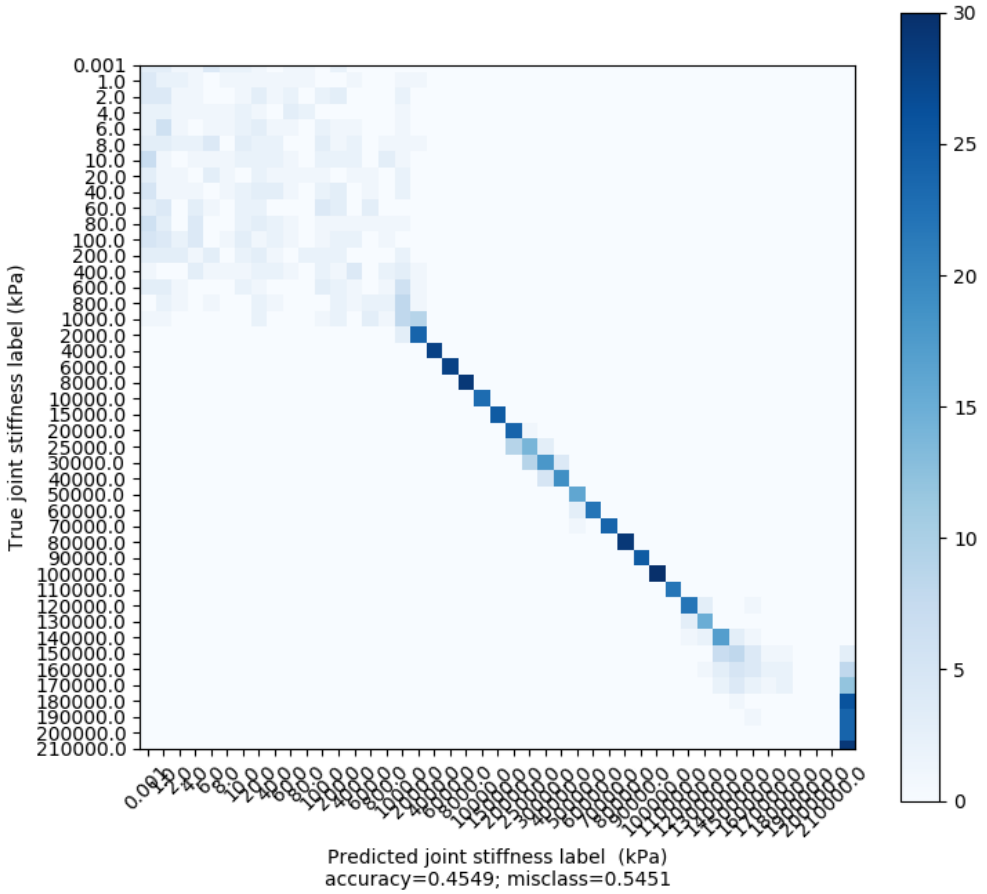


Figure 69. Confusion matric from random forest classification showing the true joint stiffness against the predicted joint stiffness.

shows the accuracy of classification between healthy and damaged, conducted in a sequential, binary approach rather than a multi-class approach. This shows more clearly when damage becomes detectible. When the stiffness value has reduced to 80 % of the design value it can be distinguished from the intact case. This shows that, in the ideal case where a random forest model is used, damage can be detected relatively early.

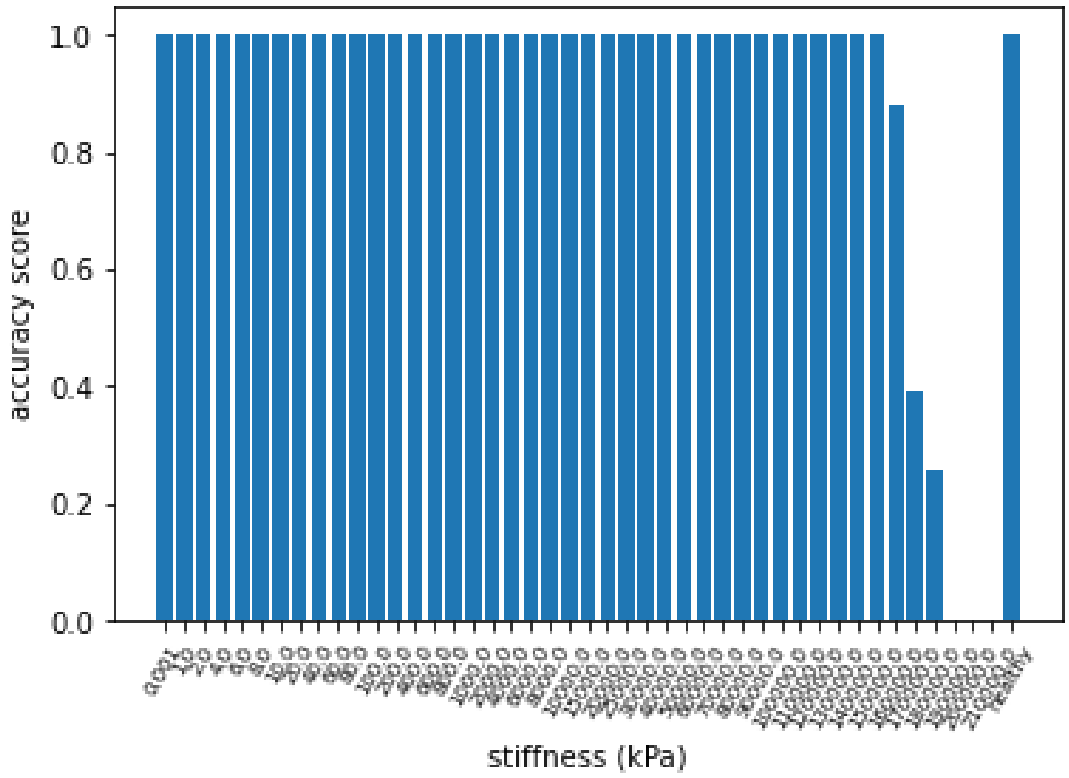


Figure 70. Accuracy score achieved by random forest classification for changing damage levels (represented in this case by member stiffness)

A.1.5.2 Logistic Regression Classification results

The second ML model used in the level 3 study is logistic regression. The relative simplicity of this linear model makes it more useful for some purposes. Statistical methods can be used to reduce the number of variables. While variable reduction is normally conducted to reduce numerical cost, in this case it can serve the very useful purpose of reducing sensors. To reduce variables, first the 'information value' is calculated for each variable and lower IV variables are removed, with the model retrained and tested. Information value is calculated according to the following equation:

$$IV = \sum (Event\% - Non\ event\%) * \ln\left(\frac{Event\%}{Non\ event\%}\right) \quad (52)$$

shows that the number of variables can be reduced significantly, from 185 to 29 variables, without much loss of accuracy. Logistic regression is a binary model and so this is predicting between 'damaged' and 'undamaged' labels.

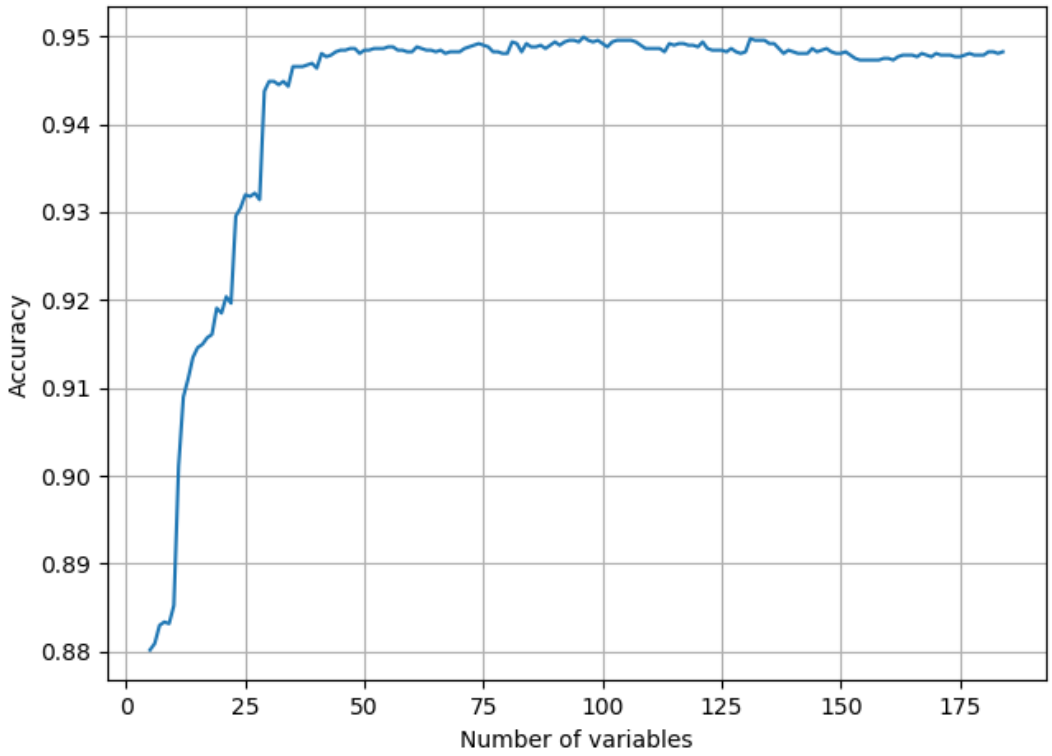


Figure 71. Accuracy of the logistic regression model to predict damaged/undamaged against the number of variables used as the number of variables is reduced based on information value

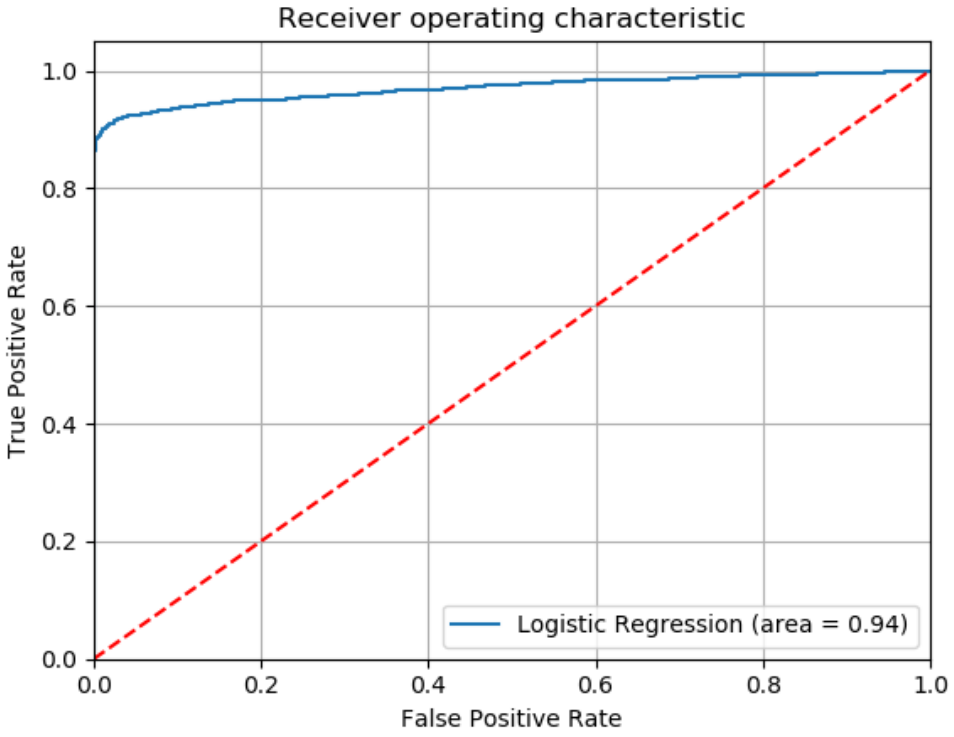
The number of variables can be further reduced by iteratively removing variables with p-values above a given threshold. Doing this reduces the number of variables further to only 23.

The final list of variables after the complete variable reduction approach is shown in . There is only one frequency value, the rest are node displacements. There are also no vertical displacements, only *dy* and *dx*, therefore, to detect this damage, only 2 dimensional accelerometers are needed. Some nodes only include either *dy* or *dx*, however this may change when nacelle yaw is changed. A potential use for this approach is for sensor reduction – if the aim of installing sensors is damage detection, then the variables can be reduced in this way while directly associating it with the loss in accuracy.

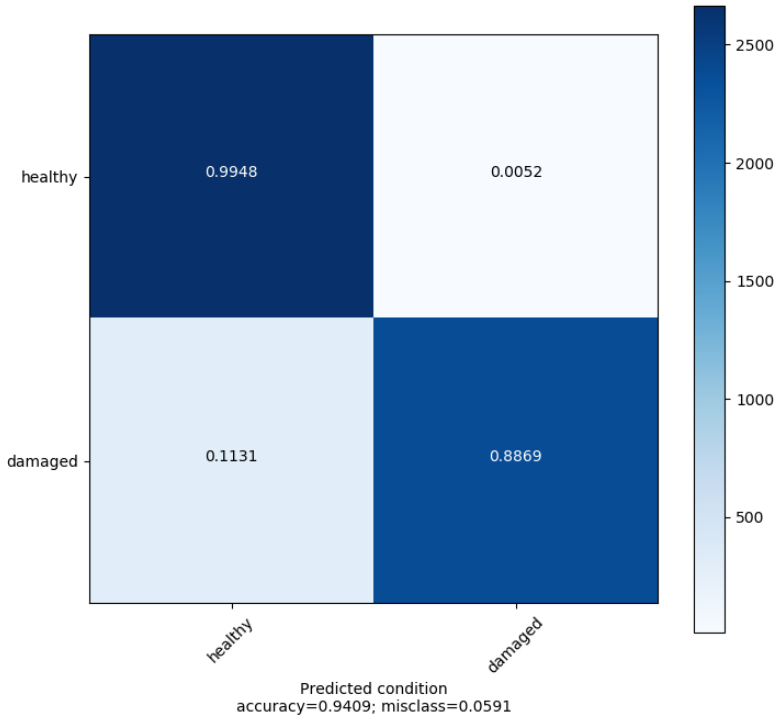
Table 27. List of variables after variable reduction for logistic regression

Number	Value type	Mode number	Node	axis
1	displacement	3	95A0P0	dy
2	displacement	3	95B0P0	dy
3	displacement	3	95B0Q0	dy
4	displacement	4	TOWN07	dy
5	displacement	3	99ABPQ	dy
6	displacement	1	TOWTOP	dx
7	displacement	3	95A0Q0	dy
8	displacement	1	TOWN07	dx
9	displacement	4	95A0P0	dy
10	displacement	3	TOWN07	dy
11	displacement	2	TOWTOP	dy
12	displacement	4	95A0Q0	dx
13	displacement	4	99ABPQ	dx
14	frequency	5	-	-
15	displacement	4	TOWN07	dx
16	displacement	4	95B0Q0	dx
17	displacement	5	TOWN07	dx
18	displacement	2	TOWN07	dy
19	displacement	4	95B0P0	dx
20	displacement	4	95A0P0	dx
21	displacement	4	95B0P0	dy
22	displacement	5	95A0Q0	dy
23	displacement	2	TOWTOP	dx

The accuracy of the model, using the reduced set of variables is shown in the following figure:


Receiver operating characteristic
True Positive Rate
False Positive Rate
Logistic Regression (area = 0.94)
0.0
0.2
0.4
0.6
0.8
1.0

(a)


healthy
damaged
0.9948
0.0052
0.1131
0.8869
healthy
damaged
Predicted condition
accuracy=0.9409; misclass=0.0591
2500
2000
1500
1000
500

(b)

Figure 72. Logistic regression model accuracy using reduced set of variables. a) Receiver operating characteristic showing the impact of the choice of decision boundary, b) confusion matrix showing the prediction results.

This shows the ROC curve and the subsequent confusion matrix of the final LR model. LR calculates probabilities of an observation belonging to a class, rather than simply the class, and so the ROC curve helps show where the boundary in the probability should be set. The confusion matrix shows that the number of false positives can be made smaller than the number of false positives, which reduces cases where the inspection team is sent unnecessarily.

A.1.6 Conclusion

The conclusions for level 3 tracking modal properties is that, given a good enough model, damage can begin to be detected relatively early. Once the damage is detected, the severity can be classified quite accurately as the rate of change in the independent variables becomes larger. With LR, variable reduction can be used to significantly reduce the number of variables. This has implications for sensor reduction for the purpose of damage detection.

There are several aspects to be further investigated in the future which can make the results more robust as well as giving new insights. The nature of the scatter can be investigated to more accurately replicate observed scatter. The applicability of using machine learning for sensor reduction can be further developed.