TU Delft

Delft University of Technology

Probabilistic data-driven surrogate modeling of wind turbine damage equivalent loads

Singh, D.

DOI
10.4233/uuid:15e9089b-ef07-472e-b1e5-b66746117620

Publication date
2026

Document Version
Final published version

Citation (APA)
Singh, D. (2026). *Probabilistic data-driven surrogate modeling of wind turbine damage equivalent loads*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:15e9089b-ef07-472e-b1e5-b66746117620

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Probabilistic Data-Driven **Surrogate** Modeling of Wind Turbine Damage Equivalent Loads

Deepali Singh

Probabilistic data-driven surrogate modeling of wind turbine damage equivalent loads

Probabilistic data-driven surrogate modeling of wind turbine damage equivalent loads

Dissertation

for the purpose of obtaining the degree of Doctor
at Delft University of Technology
by the authority of the Rector Magnificus, Prof. dr .ir. H. Bijl,
chair of the Board of Doctorates,
to be defended publicly on
Monday, 26 January 2026 at 17.30

by

Deepali SINGH

Master of Science in Aerospace Mechanics and Avionics,
Institut Supérieur de l'Aéronautique et de l'Espace, France,
born in Allahabad, Uttar Pradesh.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	Chairperson
Dr. R. P. Dwight,	Delft University of Technology
Prof. dr. ir. A. C. Viré,	Delft University of Technology
Independent members:	
Prof. dr. C. A. Crawford	University of Victoria
Prof. dr. E. Chatzi	ETH Zürich
Dr. N. Dimitrov	Technical University of Denmark
Dr. S. Barber	Eastern Switzerland University of Applied Sciences
Prof. dr. D. A. von Terzi	Delft University of Technology
Prof. dr. S. J. Watson	Delft University of Technology, reserve member

This research was a part of the European Industrial Doctorate programme, STEP4WIND, granted under the H2020 Marie-Curie Innovative Training Network initiative (H2020-MSCA-ITN-2019, grant agreement 860737).

Horizon 2020
European Union funding
for Research & Innovation

Keywords:	Machine learning, floating offshore wind turbines, mixture density network
Printed by:	Proefschriftspecialist
Cover:	Designed by Deepali Singh
Style:	TU Delft House Style, with modifications by Moritz Beller and Deepali Singh (https://github.com/Inventitech/phd-thesis-template)
Flipbook:	Designed by Dr. Erik Fritz, with modifications by Deepali Singh (https://github.com/erikfritz/FlipbookPageFooter)

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

To my family, and the friends who became family

Contents

Summary **xix**

Samenvatting **xxi**

1 Introduction **1**
1.1 Wind energy perspectives . . . 3
1.2 Floating wind technology . . . 4
1.3 The role of data . . . 7
1.4 Research motivation . . . 10
1.5 Research objective and research questions . . . 11
1.6 Thesis outline . . . 12

2 Background and motivation for data-driven site analysis **15**
2.1 Site analysis . . . 17
2.2 Fatigue calculations . . . 18
2.3 Probabilistic surrogates . . . 26
2.4 Advances in surrogate modeling of damage equivalent loads . . . 28

3 Methodology **31**
3.1 Background . . . 33
3.2 Methods for feature selection. . . . 33
3.2.1 Pearson's correlation coefficient . . . 34
3.2.2 Mutual information . . . 34
3.2.3 Conditional dependence coefficient. . . . 34
3.3 Regression . . . 36
3.3.1 Linear regression. . . . 37
3.3.2 Polynomial regression . . . 37
3.3.3 Gaussian process regression . . . 37
3.3.4 Chained Gaussian processes . . . 39
3.3.5 Artificial Neural Networks . . . 41
3.3.6 Mixture density networks . . . 43
3.3.7 Conditional generative adversarial networks . . . 45
3.4 Goodness of fit for probabilistic models . . . 47
3.4.1 Wasserstein distance . . . 47
3.4.2 Coefficient of determination . . . 48

4 Surrogate modeling for fixed bottom wind turbine loads **49**
4.1 Introduction and objective . . . 51
4.2 Numerical setup . . . 51
4.2.1 OpenFAST modeling approach . . . 51

4.2.2 Definition and sampling of the input features . . . 52
4.2.3 Responses . . . 53
4.2.4 Test datasets . . . 54
4.3 Data-driven modeling setup . . . 56
4.3.1 Probabilistic machine learning approach . . . 56
4.3.2 Surrogate modeling setup . . . 57
4.4 Results . . . 63
4.4.1 Convergence . . . 63
4.4.2 Load prediction . . . 68
4.5 Summary and discussion . . . 82

5 Data analysis of the TetraSpar demonstrator loads 85
5.1 Motivation and background . . . 87
5.2 Data processing . . . 89
5.2.1 Data description . . . 89
5.2.2 Preprocessing steps . . . 90
5.3 Simulation setup . . . 91
5.4 Results . . . 91
5.4.1 Data visualization and standard correlation metrics . . . 92
5.4.2 Feature ordering by conditional independence . . . 93
5.5 Conclusion . . . 99

6 Surrogate modeling for floating offshore wind turbine loads 101
6.1 Introduction . . . 103
6.1.1 Background . . . 103
6.1.2 Research objectives . . . 104
6.2 Numerical setup . . . 105
6.2.1 Definition of the floating wind turbine . . . 105
6.2.2 Numerical model . . . 105
6.3 Definition of relevant site features and responses . . . 109
6.3.1 Feature bounds . . . 110
6.4 Surrogate modeling setup . . . 114
6.4.1 Training and testing database generation . . . 114
6.4.2 Mixture density network hyperparameters . . . 115
6.5 Results . . . 116
6.5.1 Choice of hyperparameters . . . 116
6.5.2 Surrogate model validation . . . 121
6.5.3 Lifetime damage equivalent loads . . . 124
6.6 Discussion . . . 128
6.6.1 Summary . . . 128
6.6.2 Discussion and future work . . . 129

7 Conclusions and future work 131
7.1 Main outcomes . . . 133
7.2 Limitations and future work . . . 137

A Appendix A **139**
A.1 Scatter plots of the load channels on the TetraSpar demonstrator. 139
A.2 Feature importance calculated using a surrogate model 147

B Appendix B **151**
B.1 ERA5 locations used for defining feature ranges 151
B.2 Feature bounds. 152
B.2.1 Significant wave height 152
B.2.2 Peak spectral period 152

Bibliography **155**

List of Acronyms **174**

Acknowledgments **175**

Curriculum Vitæ **177**

List of Publications **179**

List of Figures

1.1 Left: Map of a portion of the North Sea showing the wind farms in all stages, from development zones to fully commissioned. Right: Norwegian wind farm projects. Image source: Hjelmeland et al. [6]. 3

1.2 The image shows the North Sea region with a water depth cut-off at 60 m, indicating large offshore areas where floating wind turbines could be installed. Image source: https://globalwindatlas.info/en. 4

1.3 Schematic of different types of wind turbine foundations based on water depth. For depths > 60 m, fixed bottom foundations are no longer economically viable. 5

1.4 Comparison of some of the main floater design categories. 6

1.5 Images of commercial floating wind turbine projects in the EU. 7

1.6 Schematic of the surrogate modeling mapping. On the left, the various site conditions defining the input are listed. On the right, the response channels of interest are shown. 11

2.1 Sketch of a typical time-varying stress profile on a wind turbine. 18

2.2 Process of calculating fatigue damage on wind turbine parts. 19

2.3 Example of an hourly wind speed time series obtained from the ERA5 reanalysis database [32] for Sud de la Bretagne [92] is shown on the left. The image on the right shows the corresponding histogram with the wind speed bins. 20

2.4 Illustration of turbulence boxes and the advection direction. The left image shows the Kaimal approach with the two-dimensional turbulence grids. The image on the right represents a 3D turbulence grid from the Mann model. 21

2.5 Time history of the wave elevation (ζ) with the successive wave heights (h) highlighted. 22

2.6 The corresponding probability density function and cumulative density function of the wave height measurements from time series data. 22

2.7 Comparison of Pierson-Moskowitz, JONSWAP and Torsethaugen wave spectral models for two wave periods, $H_s = 2$m [101]. 23

2.8 Various modules that comprise the coupled multi-physics engineering tools used for wind turbine simulations. 24

2.9 Schematic representation of the multi-valued mapping of the input-output pairs used to train the surrogate. 27

3.1 The left image shows sample functions from the prior with an RBF kernel. On the right, the image with the training data points shows the posterior function draws in dashed lines, with the confidence interval of the prediction in gray. . . . 38
3.2 Figures showing the difference between the standard Gaussian process regression assuming a fixed noise term, versus the heteroscedastic approach with chained Gaussian process regression. . . . 40
3.3 A schematic of an ANN with one hidden layer. . . . 41
3.4 Schematic representation of MDNs. . . . 43
3.5 Schematic of a standard conditional GAN model architecture. . . . 46

4.1 Pairplot of the input features for CASE-OFFSHORE showing the training samples along with the test datasets. The marginal distribution for the training dataset is shown along the diagonal. . . . 56
4.2 Schematic of the machine learning approach. . . . 57
4.3 The machine learning approach for the chained Gaussian process regression model. . . . 58
4.4 CASE-OFFSHORE: Comparison of accuracy metrics for various regularization parameter values with MDN (30, 20) trained on 5000 samples, tower bottom fore-aft standard deviation. . . . 60
4.5 CASE-OFFSHORE: Comparison of accuracy metrics for various combinations of layer widths in a two-layer network with MDN trained on 5000 samples, tower bottom fore-aft standard deviation, $L1 = 0.001$, $L2 = 0.001$. . 61
4.6 CASE-OFFSHORE: Comparison of accuracy metrics for a few combinations of layer widths in a 3-layer network with MDN trained on 5000 samples, tower bottom fore-aft standard deviation, $L1 = 0.001$, $L2 = 0.001$. . . . 62
4.7 Convergence of the R^2 value of the conditional mean and the d_{W2} of the conditional distribution of the predicted quantities as a function of the number of samples for CASE-ONSHORE. . . . 64
4.8 Convergence of the R^2 value of the conditional mean and the d_{W2} of the conditional distribution of the predicted quantities as a function of the number of samples for CASE-OFFSHORE. . . . 64
4.9 Calculated length scales per feature using ARD in the GPR model trained with 1500 training samples on tower bottom fore-aft moment standard deviation (CASE-OFFSHORE). . . . 66
4.10 Log marginal likelihood plotted as a function of the individual length scales for each input feature for the GPR model trained with 1500 training samples on tower bottom fore-aft moment standard deviation. . . . 66
4.11 Training and validation losses for MDN (10, 10) with 500 (a) and 4500 (b) training samples, CASE-ONSHORE. . . . 68
4.12 CASE-ONSHORE: Conditional mean (top left) and conditional standard deviation (top right) surrogate predictions vs. OpenFAST reference for blade root edgewise moment DEL. Zoomed-in image of the conditional standard deviation on the bottom. . . . 70

4.13 CASE-ONSHORE: Predicted and reference (OpenFAST) conditional pdf for the tower base fore-aft moment DEL. 72
4.14 CASE-ONSHORE: Predicted and reference (OpenFAST) conditional pdf for the tower top fore-aft moment DEL. 73
4.15 CASE-ONSHORE: Predicted and reference (OpenFAST) conditional pdf for the blade root flapwise moment DEL. 74
4.16 CASE-ONSHORE: Predicted and reference (OpenFAST) conditional pdf for the blade root edgewise moment DEL. 75
4.17 CASE-OFFSHORE: Predicted and reference (OpenFAST) conditional pdf for the tower base fore-aft moment DEL. 78
4.18 CASE-OFFSHORE: Predicted and reference (OpenFAST) conditional pdf for the tower top fore-aft moment DEL. 79
4.19 CASE-OFFSHORE: Predicted and reference (OpenFAST) conditional pdf for the blade root flapwise moment DEL. 80
4.20 CASE-OFFSHORE: Predicted and reference (OpenFAST) conditional pdf for the blade root edgewise moment DEL. 81

5.1 Image of the TetraSpar full-scale demonstrator moored at the test site off the coast of Norway [9]. 88
5.2 Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. standard deviation of wind speed, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality. 93
5.3 Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. significant wave height, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality. 94
5.4 MI and PCC values between the tower bottom resultant 10-minute DEL and the various environmental variables from the measurements and simulations. 95
5.5 MI and PCC values between the tower top resultant 10-minute DEL and the various environmental variables from the measurements and simulations. 95
5.6 MI and PCC values between the blade root edgewise moment 10-minute DEL and the various environmental variables from the measurements and simulations. 96
5.7 MI and PCC values between the blade root flapwise moment 10-minute DEL and the various environmental variables from the measurements and simulations. 96
5.8 Scatter plots from the TetraSpar measurements, showing the influence of significant wave height on the standard deviation of the platform pitch and blade root flapwise moment 10-min. DEL. The size of the symbols is based on the standard deviation of the wind speed. 97

6.1 Schematic of the surrogate modeling mapping for a floating offshore wind turbine. 104

6.2 Hywind Scotland spar buoy with crowfoot mooring line configuration [173]. 107

6.3 A simplified representation of the structural sub-elements used in the BHawC model [26]. 108

6.4 Schematic of the OrcaFlex simulation elements. 109

6.5 (a) Histograms of the shear exponent α for selected sites for ERA5 reanalysis data from 1990 to 2019. (b) Shear exponent shown as a function of wind speed, marked by a box denoting the selected sampling domain. 111

6.6 Examples of wind profiles on selected sites. 112

6.7 Chosen turbulence intensity range in dashed lines, along with the IEC class C turbulence profile for NTM, and measured turbulence at FINO 3 (German Bight) from the met-ocean analysis report on the Hywind Scotland project [173]. 113

6.8 Scatter plots of (a) the significant wave height vs. the wind speed at 100 m, and (b) peak spectral period vs. significant wave height for the selected sites (Table B.1) based on the ERA5 reanalysis data from 1990 to 2019. . . . 113

6.9 Paired scatter plots and marginal distributions of the training and testing datasets. 115

6.10 Study on the network architecture for two-layer networks. The top image shows the R^2 value for the mean of the conditional pdf of the tower bottom fore-aft DEL channel. The dashed line corresponds to an R^2 value of 0.99. The bottom image plots the d_{W2} values for the same channel with the dashed line corresponding to d_{W2} of 1. 117

6.11 Study on the network architecture for three-layer networks. The top image shows the R^2 value for the mean of the conditional pdf of the tower bottom fore-aft DEL channel. The dashed line corresponds to an R^2 value of 0.99. The bottom image plots the d_{W2} values for the same channel with the dashed line corresponding to d_{W2} of 1. 118

6.12 Study on the network architecture for four-layer networks. The top image shows the R^2 value for the mean of the conditional pdf of the tower bottom fore-aft DEL channel. The dashed line corresponds to an R^2 value of 0.99. The bottom image plots the d_{W2} values for the same channel with the dashed line corresponding to d_{W2} of 1. 119

6.13 Sensitivity of the MDN (30, 30, 50) surrogate to the number of mixture components. 120

6.14 Convergence plots for the tower bottom fore-aft DEL channel. (a) Shows the convergence of the predicted mean as a function of the number of training samples for three MDN architectures. (b) Shows the normalized Wasserstein distance as a function of the training samples. 121

6.15 Load predictions using MDN for the tower bottom fore-aft DEL (normalized). Figure (a) shows the surrogate predicted conditional mean at the test locations vs. the conditional mean calculated using BHawC. Figure (b) shows the predicted and reference standard deviations of the conditional pdf. Figures (c) and (d) compare the conditional pdf plots between the surrogate and the simulation at below-rated and near-rated conditions, respectively. The values in vector x denote: $[U_{\text{ref}}, \alpha, TI, H_s, T_p, W_{\text{dir}}, Yaw]$ with units specified in Table 6.4. 123

6.16 Load predictions using MDN for the tower top fore-aft DEL (normalized). Figure (a) shows the surrogate predicted conditional mean at the test locations vs. the conditional mean calculated using BHawC. Figure (b) shows the predicted and reference standard deviations of the conditional pdf. Figures (c) and (d) compare the conditional pdf plots between the surrogate and the simulation at below-rated and near-rated conditions, respectively. The values in vector x denote: $[U_{\text{ref}}, \alpha, TI, H_s, T_p, W_{\text{dir}}, Yaw]$ with units specified in Table 6.4. 124

6.17 Load predictions using MDN for the blade root flapwise DEL (normalized). Figure (a) shows the surrogate predicted conditional mean at the test locations vs. the conditional mean calculated using BHawC. Figure (b) shows the predicted and reference standard deviations of the conditional pdf. . . 125

6.18 Load predictions using MDN for the blade root edgewise DEL (normalized). Figure (a) shows the surrogate predicted conditional mean at the test locations vs. the conditional mean calculated using BHawC. Figure (b) shows the predicted and reference standard deviations of the conditional pdf. . . 125

6.19 Comparison of the site conditions at the four floating wind sites considered in this study (Table B.1). 126

6.20 The 25-year normalized M_{eq} calculated for four sites at the tower bottom fore-aft direction (top-left), tower top fore-aft (top-right), blade root flapwise (bottom-left), and blade root edgewise (bottom-right) channels. The mean M_{eq} obtained at the Hywind Scotland site is used as the reference to normalize the loads at the remaining locations. 127

A.1 Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. mean wind speed, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality. 140

A.2 Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. the wave peak spectral period, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality. 141

A.3 Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. power law exponent, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality. 142
A.4 Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. air density, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality. 143
A.5 Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. the wind direction, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality. 144
A.6 Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. current speed, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality. 145
A.7 Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. current direction, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality. 146
A.8 Tower bottom moment DEL list of feature importance as calculated by XGBoost. 148
A.9 Tower top moment DEL list of feature importance as calculated by XGBoost. 148
A.10 Edgewise blade moment DEL list of feature importance as calculated by XGBoost. 149
A.11 Flapwise blade moment DEL list of feature importance as calculated by XGBoost. 149

List of Tables

4.1 List of input variables and the corresponding variable bounds for CASE-ONSHORE (ON) and CASE-OFFSHORE (OFF). 53
4.2 List of response channels modeled by the data-driven surrogate models. . . 54
4.3 Variables and their corresponding values in TEST2 dataset. Δ indicates the discretization step. 55
4.4 The number of training samples and inducing points considered for the model's convergence study. 58
4.5 Summary of the ANN hyperparameters. 59
4.6 Summary of the MDN network hyperparameters. 59
4.7 Summary of the WCGAN network hyperparameters. 61
4.8 Description of the models used in the convergence study. 63
4.9 Comparison of the R^2_μ metric for various load channels' 10-minute damage equivalent loads in CASE-ONSHORE. 68
4.10 Comparison of the R^2_σ metric for various load channels' 10-minute standard deviation in CASE-ONSHORE. 69
4.11 Comparison of the d_{W2} metric for various load channels' 10-minute standard deviation in CASE-ONSHORE. 69
4.12 Values of the input features for the pdfs in Figure 4.13 to 4.16. 70
4.13 Comparison of the R^2_μ metric for various load channels' 10-minute damage equivalent loads in CASE-OFFSHORE. 76
4.14 Comparison of the R^2_σ metric for various load channels' 10-minute standard deviation in CASE-OFFSHORE. 76
4.15 Comparison of the d_{W2} metric for various load channels' 10-minute standard deviation in CASE-OFFSHORE. 76
4.16 Values of the input features for the pdfs in Figure 4.17 to 4.20. 77

5.1 List of variables characterizing the environmental conditions. 89
5.2 List of the load measurement locations and the corresponding description. 90
5.3 Feature ordering using FOCI on the TetraSpar measurement dataset. DEL refers to the 10-minute DEL. 97
5.4 Feature ordering using FOCI on the TetraSpar simulation dataset. DEL refers to the 10-minute DEL. 98

6.1 Parameters of 6MW Siemens Gamesa wind turbine. 106
6.2 Wind turbine tower and foundation properties [17, 168, 173]. 106
6.3 Catenary mooring line properties. 106
6.4 The list of input features provided to the surrogate model and their corresponding notation. 110

6.5 The list of output channels that the surrogate models are trained to predict. 110
6.6 Summary of the mixture density network hyperparameters. 116
6.7 The MDN architectures considered for the convergence study. 120
6.8 List of hyperparameters used for training the MDN model for the final load prediction. 122
6.9 Quantitative analysis of MDN model's predictions using d_{W2} and R^2 as evaluation metrics. 122

B.1 Description of the sites used for defining the feature ranges. 152
B.2 Tuning coefficients for defining the range functions for the spectral wave period. 153

Summary

As nations broaden their renewable energy portfolios, wind power is playing an increasingly important role in the energy landscape. Although onshore wind continues to dominate the wind energy mix in the EU, rising energy demand and ambitious climate targets have driven the development of offshore wind farms and spurred research into floating offshore wind turbines.

This thesis began with the goal of developing a data-driven modeling approach that maps site conditions to the load statistics on floating offshore wind turbines. The motivation for this research emerged from a need to accelerate the site-selection process, include more site variables, and achieve accurate load estimates while keeping the computational expense low.

Before installation, all modern wind turbines must be carefully assessed for structural loads to ensure safety and performance throughout their lifetimes. Floating turbines, in particular, introduce more complexity—more environmental variables to consider, and higher uncertainty in the dynamic response. This added complexity makes analyzing structural loads expensive and time-consuming, often more so than for fixed-bottom counterparts, and it makes understanding the behavior of floating wind turbines even more important. The process of fatigue damage calculation, in particular, is computationally intensive, often requiring thousands of costly simulations. Because it is not feasible to run simulations for every possible sea state, engineers typically reduce the problem by selecting a set of variables and binning or lumping sea states to limit the number of required simulations. However, for floating wind turbines, the choice of which variables to include and how to perform this binning remains an open question. This motivates the use of reliable data-driven surrogate models that can make quick estimates of loads on wind turbines while maintaining high accuracy.

Although most existing work relies on deterministic surrogate models, offshore wind environments exhibit strong stochasticity, making turbine loads inherently uncertain. This motivates the need for uncertainty quantification to characterize not only expected loads but also estimate their conditional variability. This dissertation, therefore, develops a probabilistic data-driven methodology that propagates environmental uncertainty to the 10-minute damage equivalent loads of onshore, offshore, and floating offshore wind turbines.

Several deterministic and probabilistic data-driven models are benchmarked on onshore and fixed-bottom offshore wind turbines. The evaluation consists not only of judging the accuracy of a model, but also of practical aspects such as the robustness of the model, sensitivity to hyperparameters, and ease of implementation. Compared to deterministic approaches, which require multiple seed repetitions prior to training, it is demonstrated that with probabilistic models, this step may not be necessary to achieve high-accuracy predictions ($R^2 > 0.95$), thereby saving precious computational resources needed to generate the training database. Widely used Gaussian process regression is shown to

accurately estimate the conditional mean of the response with a relatively small training dataset of $\mathcal{O}10^2$ - 10^3) samples. Wasserstein-conditional generative adversarial network is used as one of the probabilistic regression models. Despite learning the functional mapping with errors comparable to the best-performing models, it is found to be very complex to implement and requires extensive hyperparameter tuning. Simpler 4th-degree polynomials are shown to make good predictions in the onshore case, but are prone to overfitting and susceptible to the additional noise introduced by the hydrodynamic features. Overall, mixture density networks are shown to provide the best combination of consistency, accuracy, and practical ease of use.

Based on this analysis, the study extends the use of mixture density networks to a more sophisticated application of spar-type floating offshore wind turbine. It is shown to successfully capture the conditional response in terms of the normalized 2-Wasserstein distance despite the added complexity. The surrogate is further used to make probabilistic estimates of the lifetime damage equivalent loads on four potential floating wind turbine sites. Since the surrogate model is fast (order of milliseconds once trained), load predictions can be made over all sea states within seconds, without the need to lump or bin the sea states beforehand. The uncertainty in the aggregated lifetime fatigue loads due to stochastic inputs is extremely narrow, with variability on the order of only 0.1–0.5% of their mean values. This results from summing the 10-minute damage equivalent loads over a million occurrences, effectively nullifying the impact of the outliers. The use of a probabilistic surrogate that correctly captures the conditional distribution is still useful, as it minimizes the aggregation of error in the final response.

Through these analyses, it is demonstrated that surrogate models can be powerful tools for fatigue estimation in the site analysis process, especially for floating wind turbines, where the choice of variables and binning methods is still an open question. Additionally, using probabilistic surrogates like mixture density networks helps reduce bias in calculating the aggregate mean fatigue, as the conditional distributions are heteroscedastic and not always normally distributed.

Samenvatting

Nu landen hun portfolio's voor hernieuwbare energie verder uitbreiden, speelt windenergie een steeds belangrijkere rol in het energiesysteem. Hoewel windenergie op land nog steeds het grootste aandeel vormt in de Europese Unie, hebben de toenemende energievraag en ambitieuze klimaatdoelstellingen geleid tot de ontwikkeling van offshore windparken en tot intensief onderzoek naar drijvende offshore windturbines.

Dit proefschrift is gestart met als doel het ontwikkelen van een datagedreven modelleeraanpak die locatiecondities koppelt aan belastingsstatistieken van drijvende offshore windturbines. De motivatie voor dit onderzoek komt voort uit de behoefte om het locatiekeuzeproces te versnellen, meer locatievariabelen te kunnen meenemen en nauwkeurige belastingsschattingen te verkrijgen tegen een lage rekentijd.

Voorafgaand aan installatie moeten alle moderne windturbines zorgvuldig worden beoordeeld op structurele belastingen om veiligheid en prestaties gedurende de gehele levensduur te waarborgen. Drijvende windturbines brengen daarbij extra complexiteit met zich mee: er zijn meer omgevingsvariabelen van invloed en de dynamische respons kent een grotere onzekerheid. Deze toegenomen complexiteit maakt de analyse van structurele belastingen kostbaar en tijdrovend, vaak meer dan bij vaste funderingen, en onderstreept het belang van een goed begrip van het gedrag van drijvende windturbines. Met name de berekening van vermoeiingsschade is computationeel intensief en vereist vaak duizenden kostbare simulaties. Omdat het niet haalbaar is om simulaties uit te voeren voor elke mogelijke zeetoestand, reduceren ingenieurs het probleem doorgaans door een selectie van variabelen te maken en zeetoestanden te binnen of te clusteren om het aantal benodigde simulaties te beperken. Voor drijvende offshore windturbines is de keuze van relevante variabelen en de wijze van binning echter nog een open vraag. Dit motiveert het gebruik van betrouwbare datagedreven surrogaatmodellen die snelle belastingsschattingen kunnen leveren met behoud van hoge nauwkeurigheid.

Hoewel het merendeel van bestaand onderzoek gebruikmaakt van deterministische surrogaatmodellen, worden offshore windomgevingen gekenmerkt door sterke stochastische variatie, waardoor turbinebelastingen inherent onzeker zijn. Dit benadrukt de noodzaak van onzekerheidskwantificatie om niet alleen verwachte belastingen te beschrijven, maar ook hun conditionele variabiliteit te schatten. Dit proefschrift ontwikkelt daarom een probabilistische datagedreven methodologie die de onzekerheid in omgevingscondities doorvertaalt naar 10-minuten damage equivalent loads van onshore, offshore en drijvende offshore windturbines.

Verschillende deterministische en probabilistische datagedreven modellen worden vergeleken voor toepassingen op onshore en vaste offshore windturbines. De evaluatie richt zich niet alleen op nauwkeurigheid, maar ook op praktische aspecten zoals robuustheid, gevoeligheid voor hyperparameters en gebruiksgemak. In vergelijking met deterministische benaderingen, die meerdere herhalingen met verschillende willekeurige zaden vereisen voorafgaand aan training, wordt aangetoond dat probabilistische modellen deze stap niet

noodzakelijk maken om een hoge nauwkeurigheid te behalen, wat aanzienlijke besparingen oplevert in de benodigde rekeninspanning voor het genereren van trainingsdata. Veelgebruikte Gaussian process regression blijkt in staat om de conditionele verwachtingswaarde van de respons nauwkeurig te schatten met een relatief kleine trainingsdataset van orde $\mathcal{O}(10^2–10^3)$ samples. De Wasserstein-conditional generative adversarial network wordt toegepast als een van de probabilistische regressiemodellen. Hoewel dit model de functionele relatie leert met fouten vergelijkbaar met de best presterende modellen, blijkt het zeer complex in implementatie en vereist het uitgebreide afstemming van hyperparameters. Eenvoudigere vierdegraads polynomen leveren goede voorspellingen in het onshore geval, maar zijn gevoelig voor overfitting en voor de extra ruis die wordt geïntroduceerd door hydrodynamische effecten. Over het geheel genomen blijken mixture density networks de beste combinatie te bieden van consistentie, nauwkeurigheid en praktische toepasbaarheid.

Op basis van deze analyse wordt het gebruik van mixture density networks uitgebreid naar een meer geavanceerde toepassing voor een drijvende offshore windturbine van het spar-type. Aangetoond wordt dat dit model de conditionele respons succesvol vastlegt ondanks de toegenomen complexiteit. Het surrogaatmodel wordt vervolgens gebruikt om probabilistische schattingen te maken van de levensduur damage equivalent loads voor vier potentiële locaties voor drijvende windturbines. Omdat het surrogaatmodel zeer snel is (orde van milliseconden na training), kunnen belastingsvoorspellingen voor alle zeetoestanden binnen enkele seconden worden uitgevoerd, zonder voorafgaande binning of clustering van zeetoestanden. De onzekerheid in de geaggregeerde levensduurvermoeiingsbelastingen als gevolg van stochastische invoer blijkt zeer klein te zijn, met een variatie van slechts circa 0,1–0,5% van de gemiddelde waarde. Dit is het gevolg van het sommeren van 10-minuten damage equivalent loads over miljoenen voorkomens, waardoor de invloed van uitschieters grotendeels wordt uitgemiddeld. Het gebruik van een probabilistisch surrogaatmodel dat de conditionele verdeling correct weergeeft blijft desondanks waardevol, omdat het de accumulatie van fouten in de uiteindelijke respons minimaliseert.

Deze analyses tonen aan dat surrogaatmodellen krachtige hulpmiddelen kunnen zijn voor vermoeiingsschattingen in het locatieanalyseproces, met name voor drijvende offshore windturbines, waar de keuze van relevante variabelen en binningmethoden nog niet eenduidig is vastgesteld. Daarnaast draagt het gebruik van probabilistische surrogaten, zoals mixture density networks, bij aan het verminderen van bias in de berekening van gemiddelde levensduurvermoeiingsbelastingen, aangezien de conditionele verdelingen heteroscedastisch zijn en niet altijd een normale verdeling volgen.

1

Introduction

This thesis began with the goal of developing a data-driven modeling approach that maps site conditions to the load statistics on offshore wind turbines. The motivation for this research emerged from a need to accelerate the site-selection process, incorporate more site variables, and achieve accurate load estimates while keeping computational expense low. The introduction here provides a context for the remainder of the thesis. It begins with an overview of the significance of fixed and floating offshore wind in the current energy mix. It introduces various types of floating wind turbines and the elements characterizing floating wind technology. The concept of surrogate modeling is introduced, highlighting the role of data-driven methods in wind turbine design and operation. It establishes the background for presenting the research motivation and identifying the research gaps. Finally, the chapter outlines the objectives and structure of the thesis.

1.1 Wind energy perspectives

The global electricity demand in the year 2024-25 has been increasing at a rapid pace of four percent compared to the previous year, according to the electricity mid-year update by the International Energy Agency (IEA) [1]. This is partly due to the increasing electrification worldwide, along with global economic growth, as well as the intense heatwaves experienced in the first half of 2024. The increasing demand for electricity requires a corresponding increase in electricity production. In the year 2019, most of the energy needed for transportation, heating, cooling, industrial use, and electricity generation was still being produced from non-renewable sources [2]. However, due to positive policy changes, especially in the European Union (EU), electricity from renewable sources is on track to surpass coal for the first time in 2025 [3].

Wind resources form a big part of the renewable energy mix. About a fifth of all the electricity consumed in the EU in 2023 was covered from wind resources [4]. Europe now has more than 250 GW of installed wind capacity, the majority of which is located onshore. But to reach the climate targets set for 2030, the EU is still expected to install 260 GW of new wind capacity by the end of the decade, about 30GW per year. Globally, to stay within the 1.5 °C scenario, and to achieve net-zero by 2050, the International Renewable Energy Agency in the context of Sustainable Development Goal (7) envisions the total wind capacity to grow to 3500 GW from the current 1018 GW by 2030. Out of which, the global offshore wind capacity is expected to reach 500 GW. To put that into perspective, in the year 2023, the share of offshore wind was *only* 73 GW [5].

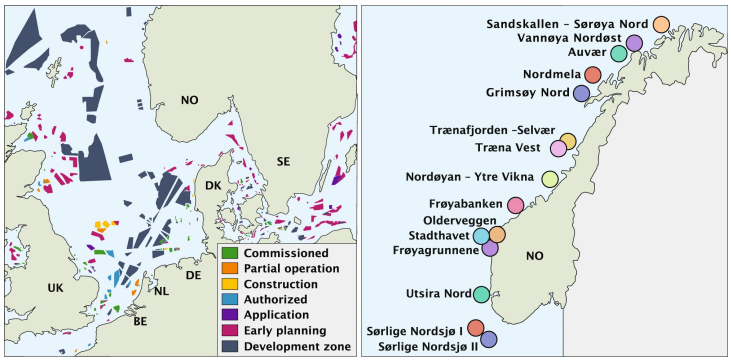


Figure 1.1: Left: Map of a portion of the North Sea showing the wind farms in all stages, from development zones to fully commissioned. Right: Norwegian wind farm projects. Image source: Hjelmeland et al. [6].

Evidently, the contribution of onshore wind, albeit significant, is not enough to keep up with the rising rate of energy consumption and ambitious energy targets. Tapping offshore wind resources is key to accelerating the growth of renewable energy. Offshore wind farms benefit from higher capacity factors, high-quality wind inflow, and relatively less opposition from the public, as the farms are located further away from residential areas. Rapid technological development for fixed-bottom wind turbines has shown a tremendous reduction of 59% in the weighted-average levelized cost of energy in the previous decade [7].

1

Figure 1.2: The image shows the North Sea region with a water depth cut-off at 60 m, indicating large offshore areas where floating wind turbines could be installed. Image source: https://globalwindatlas.info/en.

However, as shown in Figure 1.1, the shallow regions in the North Sea where fixed-bottom farms can be commissioned are increasingly overcrowded. The remaining development zones are in deeper waters (deeper than 60 m) beyond which installation complexity and costs increase sharply for fixed-bottom cases, and floating designs become much more competitive Figure 1.2. Floating technology is especially crucial for countries where the seabed drops very rapidly offshore, such as Norway. In addition, the floating wind resources across the world are at least four times larger, in theory, compared to fixed bottom regions [7].

1.2 Floating wind technology

Figure 1.3 shows a schematic of some of the common offshore wind turbine configurations and the nomenclature associated with floating wind turbine parts. The complex structure of Floating Offshore Wind Turbines (FOWTs) can be organized into three main categories. First, the wind turbine itself with the tower and the rotor-nacelle assembly (RNA), which is the same as fixed-bottom wind turbines. The scale of the wind turbines is typically also similar to that of the fixed-bottom offshore wind turbines. For example, the wind turbines in existing commercial floating projects, such as the Hywind Scotland (6 MW) or the Kincardine offshore wind models (10 MW) have rotor diameters of 154m and 164m respectively.

The second element is the floating foundation (also referred to as floating sub-structure, floating platform, or floater), as shown in Figure 1.3. The floating foundation holds the wind turbine upright and contributes to the stability of the ensemble. There are currently more than 100 unique floater concepts at various stages of development in the market [8]. They can be grouped roughly into three categories based on the stabilizing mechanism [8]. The first is gravity-stabilized, where the center of buoyancy is placed above and as far as possible from the center of gravity to provide a restoring force against the rotor thrust that is large enough to minimize power loss due to excessive pitching,

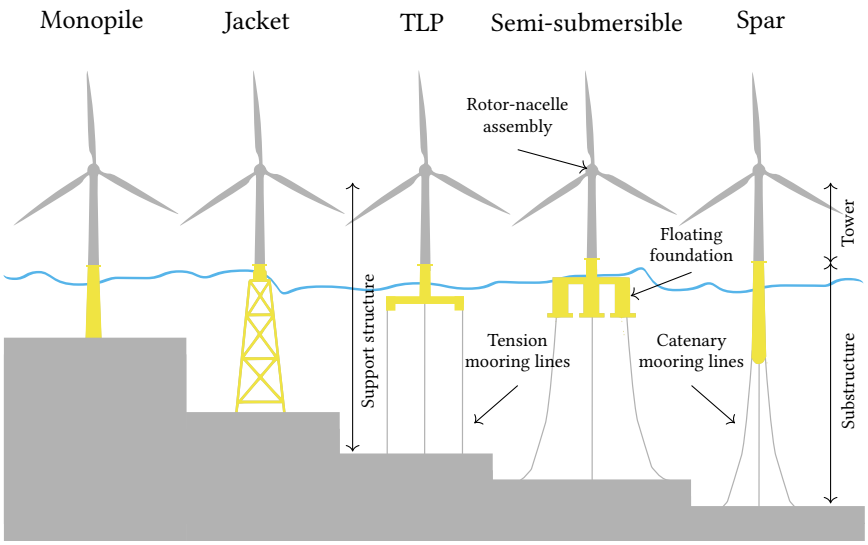


Figure 1.3: Schematic of different types of wind turbine foundations based on water depth. For depths > 60 m, fixed bottom foundations are no longer economically viable.

but not so large that the material costs explode. The spar buoy type floaters fall into this category. A spar is typically a single, ballasted cylinder made of concrete or steel. However, exceptions exist—such as the TetraSpar model [9]—which is also gravity-stabilized but comprises multiple cylindrical elements that form a suspended keel and a floating hull structure. The second mechanism is waterplane-stabilized. In this case, large columns piercing the waterplane provide stability in response to an external disturbance by moving the center of buoyancy towards the column that is pushed into the water. Semisubmersible substructures are waterplane stabilized. The floaters consist of three to five short, vertical columns connected with braces or pontoons. The columns may have stabilizing plates at the bottom that increase drag and added mass, thereby limiting heave motion. Finally, mooring-stabilized designs, such as Tension Leg Platforms (TLPs), use taut mooring lines or tendons to provide stiffness. TLP foundations are smaller than the other floater types, implying lower material costs in theory. They are, however, associated with the highest mooring line loads. The platform itself can be made very small in size, resulting in low deck space for maintenance activities. The TLP projects are currently at a much lower technological readiness level than their counterparts, and their installation and decommissioning processes are still very expensive.

The third design element is the mooring line system connecting the floating foundation to the seabed, coupled with the dynamic cables that connect the wind turbine to the offshore substation. The main purpose of the mooring line system is to restrict the motion of the wind turbine and provide a restoring force in response to wind and wave loads. They can be catenary, semi-taut, or taut. Spar buoys and semi-submersibles are often coupled with catenary mooring lines, whereas TLP designs use the taut configuration. They can be made of steel chains, wire rope, and synthetic material like Dyneema. The

1

mooring lines are designed to sustain the fatigue and ultimate loads resulting from snap events and the harsh marine environment. Snap events refer to the sudden tensile loading of a mooring line following a slack state [10]. The failure of a single line can be catastrophic for designs such as the TLP and can lead to the wind turbine capsizing. The mooring line is connected to the seabed via anchors. Taut mooring lines are typically coupled with suction anchors, whereas catenary mooring lines use either drag or suction anchors [11]. Most of the existing commercial projects, such as WindFloat, Hywind, and TetraSpar, use a semisubmersible or spar concept due to the ease of installation and design.

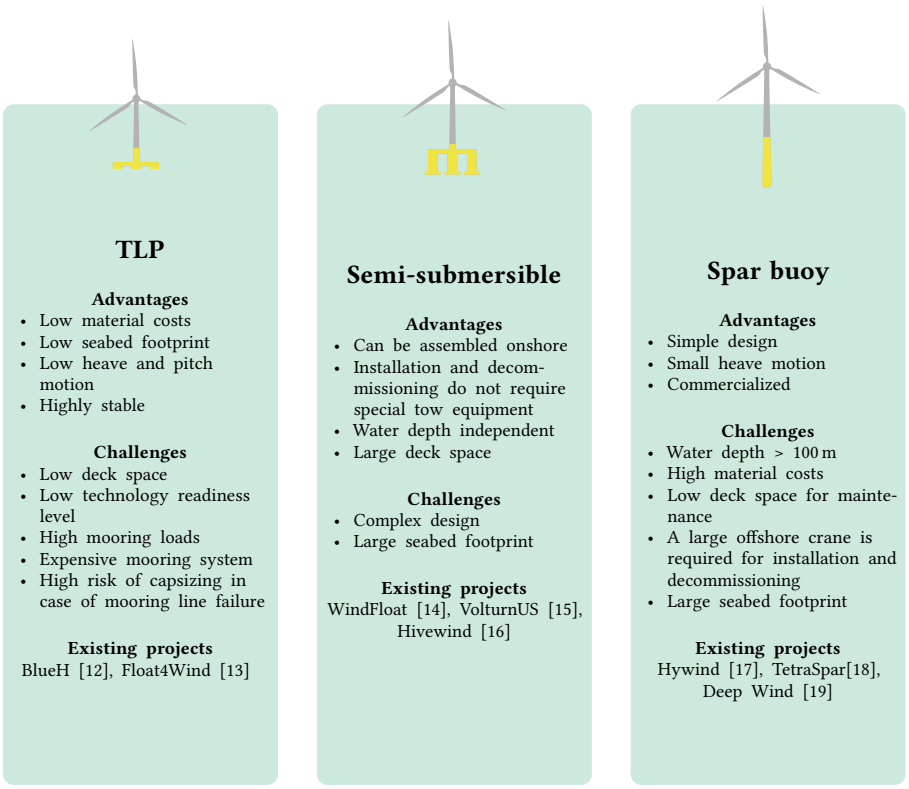


Figure 1.4: Comparison of some of the main floater design categories.

The design of such complex mega-structures comes with new technological, economic, and financial challenges. The turbine is subject to both irregular waves and turbulent wind loads, resulting in six degrees of freedom motion, unsteady aerodynamics, and sophisticated control systems. As such, wind turbine loads are a combined function of aerodynamic and hydrodynamic forces and the structural response of the material. Numerical simulations of FOWTs are complex and computationally expensive, and this added complexity makes understanding their behaviour even more important. In many cases, the total cost of computation exceeds practical limits. Could we find a simpler, faster, functional representation of the system using... *data*? Can data provide us with more insight into the system's behavior? Could it help resolve some of the bottlenecks in engineering workflows?

(a) WindFloat Kincardine semi-submersible platform design [20]

(b) Hywind Scotland spar buoy platform design[21]

(c) TetraSpar demonstrator [9]

Figure 1.5: Images of commercial floating wind turbine projects in the EU.

1.3 The role of data

Observations and data have always played a central role in the scientific understanding of our surroundings. In an engineering context, data broadly refers to information obtained through measurements, experiments, or numerical simulations. In the last decades, with the advent of accessible machine learning and statistical tools, data have become especially important. Even our actions, decisions, interactions, education, and research are significantly influenced by large data-driven models. The so-called data revolution has been so important that it is said to have given rise to the "fourth industrial revolution" [22].

With increasing access to cheaper, more advanced sensors, data are being constantly collected and analyzed from wind farms. Numerical models are utilized to generate valuable data essential for the design, validation, and monitoring of wind turbines. Some of the common data sources in the field of wind energy include:

- Time-domain numerical simulations: Multi-physics simulation tools such as OpenFAST [23, 24], HAWC2 [25] and BHawC [26, 27] provide a coupled aero-servo-hydro-elastic representation of the wind turbine ensemble. They are used to calculate time-domain responses such as power, torque, rotor speed, blade and tower loads, among other quantities. The engineering tools are used extensively at every step of the design, validation, and certification process. The wind turbine blades are represented by Euler-Bernoulli beam elements in simple cases, or fully nonlinear Finite Elements when large blade deflections are expected. The blades are subject to both gravitational loads and aerodynamic loads, which are calculated using approaches such as the Blade Element Momentum Theory (BEMT) and the Free Vortex Wake (FVW) model, among others. The sub-sea elements of offshore wind turbines are also subject to hydrodynamic loads from irregular waves. In the case of floating wind turbines, the mooring lines and the foundation elements also need to be treated appropriately.

- Frequency-domain numerical simulations: The wind turbine is represented by a linear, frequency-dependent system, which is used to solve the steady-state response at the excitation frequencies [28–30]. These numerical models are much faster than time-domain simulations for estimating the wind turbine's coupled dynamic response.

They are especially valuable for analyzing floating wind turbines, where the dynamic response of the floating structure can be computationally expensive to calculate using time-domain models.

- Computational Fluid Dynamics (CFD): With the growing interest in simulating entire wind farms to study the wakes and the interaction with the atmospheric boundary layer, CFD approaches have become central to wind turbine and farm design. Blade-resolved CFD simulations of wind turbines, let alone wind farms, are still computationally extremely expensive. In wind farm settings, wind turbines are often modeled as permeable discs in the flow field, known as Actuator Disc Models (ADM). A second approach is the Actuator Line Model (ALM), where the blades are represented as elements with certain airfoil properties that enforce body forces on the flow field. A third approach is the Actuator Surface Model (ASM), where the two-dimensional airfoil properties can be modeled as a function of the chord. CFD simulations are used to gain a deeper understanding of physical phenomena such as blade-vortex interaction, the unsteady aerodynamics and hydrodynamics of floating wind turbines in motion, and wake development. The high-fidelity data are also used to validate and improve lower-fidelity engineering models.

- Supervisory Control and Data Acquisition (SCADA) data: SCADA data are routinely collected on installed wind turbines to record quantities like the wind speed, strain, nacelle acceleration, wind direction, power, RPM, gearbox quantities, etc. SCADA data are generally stored as 10-minute statistics- mean, standard deviation, and range.

- High frequency sensors: Typically, on research turbines, higher frequency (e.g., 25Hz) measurements of various quantities on the turbine and of the environment are recorded. These are typically used for close condition monitoring of the wind turbine.

- Alarm logs: The SCADA system records alarms in case of incidents that induce an operational change. Alarms are activated, for instance, in cases of extreme environmental conditions, voltage dips in the grid, manual restrictions due to grid loads, maintenance activities, and component malfunction [31].

- Metocean data: Data from reanalysis models like the ERA5 reanalysis [32] at the wind farm site are combined with recorded data to form a complete picture of the operating conditions.

This diverse range of data provides valuable insights for a variety of applications, including predictive maintenance, improved energy forecasting, and enhancement of system reliability. The data-driven insights are used not only to drive operational improvements through digital twins but also during the early stages of conceptual design optimization, site selection, and developing our understanding of the dynamic behavior of the complex system. With the recent advancements in the field of machine learning, new approaches are developed to create a functional representation of complex systems in the form of data-driven surrogates. Some of the common applications of wind turbine data are elaborated as follows:

- Correlation studies: Analyzing data from simulations or sensors can be useful in establishing correlations between different quantities. Much research has already been performed on this topic for fixed-bottom wind turbines. [33–39]. For floating wind turbines, this is still an active area of research. Some recent studies have investigated the correlation between the environmental conditions and energy yield, as well as fatigue loads [40–44]. Identifying the combinations of conditions that exacerbate the fatigue loads on the structure can help identify important load cases during the design and certification process, design leaner structures, and identify safe installation sites.

- Site analysis: Before installing wind turbines and designing a site-specific tower, a thorough site analysis is conducted using available metocean data. This analysis helps estimate the turbine's performance and ensures a safe operational lifetime. Site-specific information is used to evaluate a series of Design Load Cases (DLCs) through advanced multi-physics numerical tools—this process can be computationally expensive. For situations where environmental variables are few, look-up tables provide a quick way to calculate loads. However, offshore wind turbines involve a larger set of variables, including the joint distributions of sea states, which complicates the analysis. In such cases, data-driven surrogates offer a significant advantage by mapping site conditions to wind turbine loads, accelerating the evaluation process [45–53].

- Condition Monitoring Systems: Continuously monitoring the state of the wind turbine components such as the gearbox, bearings, and the drivetrain, by analyzing the data collected from temperature and vibration sensors, can allow for condition-based maintenance in place of periodic maintenance activities [54–58]. This is especially useful for offshore wind turbines, where the operation and maintenance costs are generally high, up to about one-third of the total levelized cost of energy [59, 60].

- Structural Health Monitoring: Structural health monitoring techniques use data from a variety of online sensors such as accelerometers, strain gauges, tilt meters, etc., to extract damage-sensitive features and assess incremental structural changes over time [61–63].

- Virtual sensing: Virtual sensing entails mapping the observations from installed sensor locations to signals that cannot be measured on a real wind turbine, such as fatigue loads in the bearings and gear contacts. Data-driven models are trained to learn this mapping from data obtained from numerical simulations [59, 59, 64, 65].

- Fault detection and failure prediction: Digital twins relying on data are used for fault detection [64], investigation, and prediction of similar failures in the future using data-driven prediction strategies [66, 67].

- Wind turbine controller design: Modern techniques like model predictive control require the prediction of the future state of certain quantities to make control decisions. This entails having a representation of the nonlinear system to make fast

predictions of the future states. Data-driven models like reinforcement learning and long short-term memory have been explored in this context [68–72].

- Data-driven design optimization: Data-driven surrogates have been explored to accelerate the process of geometry optimization- including the floater [73], and blades [74–76]. This is because the numerical tools in the optimization loop are generally computationally expensive, and an iterative change in the geometry entails re-simulating the new designs several hundred times to find the optimal solution.
- Wake modeling: Data from CFD simulations has been used to enhance and validate the analytical or data-driven wake models used in wind farm layout optimization tools [77–81].
- Wake steering: Dynamic wake steering of individual wind turbines can result in improving the total annual energy production of the wind farm while reducing the fatigue loads. The calculation of wind turbine loads in the presence of wakes to solve such a multi-objective optimization problem is computationally expensive. Data-driven surrogate models trained on aeroelastic wind farm simulations enable the calculation of quick load estimates, which can be used to design control laws that determine the optimal yaw angle to achieve optimization objectives [82].
- Fatigue life extension: The true remaining fatigue life of the wind turbine can be determined by using the SCADA data as input to calculate the accumulated fatigue at any point in time. Data-driven surrogate models trained on numerical data are often used for this application, as it entails calculating loads on several thousand operating points [83, 84].

1.4 Research motivation

The previous section presented various scenarios in which different types of data can be leveraged to enhance, accelerate, and develop new industrial processes. However, in reality, engineering applications often have limited access to information, which is noisy due to stochasticity, missing information, and measurement inaccuracies. At the same time, the surrogate model is expected to have a high level of accuracy and not deviate from natural physical constraints. This motivates research into methods that can quantify this uncertainty in predictions. This is the primary focus of the thesis.

This research focuses on the need for accurate, data-driven surrogates that map site conditions to the corresponding loads on the wind turbine tower and blades, enabling quick site analysis. Data-driven surrogates for site analysis for fixed-bottom onshore wind turbines have been under investigation in recent years. However, the majority of the research has focused on complex uncertainty quantification approaches for simple onshore models or a deterministic mapping on fixed-bottom offshore wind turbines. Furthermore, there is a standard process for binning the site-specific wind speed distribution and scaling the short-term fatigue loads to the accumulated lifetime fatigue loads accordingly. However, for FOWTs, the number of relevant variables is larger, and properly accounting for the complex joint distribution while calculating the long-term fatigue loads remains a topic of

research. Figure 1.6 shows the interesting site conditions that would be provided to the surrogate as input to predict the 10-minute Damage Equivalent Loads (DELs) on the tower and blades. The motivation and novelty of this work lie in three aspects:

- Extending existing data-driven approaches for mapping metocean conditions to load statistics, which have largely been applied to simple turbine configurations, to complex FOWTs.
- Developing methods to provide uncertainty estimates in the loads of FOWTs to account for the stochasticity in the inflow conditions.
- Developing methods to account for the effect of the joint distribution of the input variables when calculating the long-term fatigue loads on FOWTs.

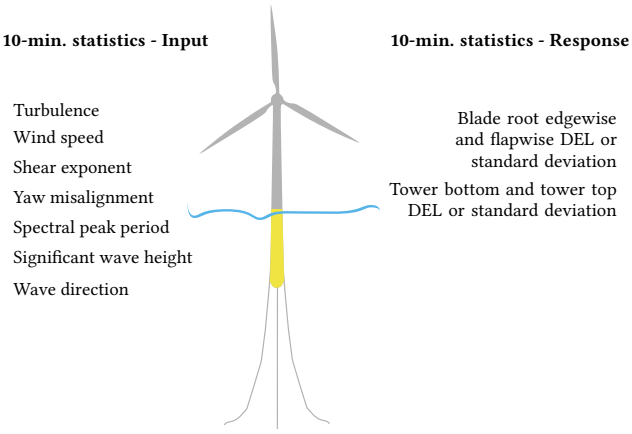


Figure 1.6: Schematic of the surrogate modeling mapping. On the left, the various site conditions defining the input are listed. On the right, the response channels of interest are shown.

1.5 Research objective and research questions

Given the identified research gap and motivation, the main research objective can be defined. As discussed in Section 1.3, data-driven surrogate models play a crucial role in wind energy applications, primarily by significantly reducing computational costs associated with complex processes.

The objective of this work is to develop an approach for training probabilistic surrogate models to quantify uncertainty in the 10-minute statistics of wind turbine load response, with site analysis as the primary application. As the number of dimensions, or features, increases, the total number of simulations performed to build a load response matrix increases exponentially. Surrogate models are a great tool to use sparse sampling to avoid grid-based sampling and reduce the simulation count. The approach presented in

1

the thesis for building probabilistic surrogates is very generic and can be applied to other areas, such as the estimation of remaining fatigue life and the prediction of extreme loads.

Research objective

RO To develop a workflow for training probabilistic surrogate models for quantifying the uncertainty in the 10-minute statistics of the load response of wind turbines for site analysis

In order to address the proposed research objective, it is necessary to investigate uncertainty in wind turbine load prediction from multiple perspectives. This includes identifying the primary sources of uncertainty inherent in load modeling, understanding the influence of environmental conditions on fatigue loading, and evaluating the performance of probabilistic surrogate models across different wind turbine configurations. Finally, the practical relevance of uncertainty quantification is addressed in the context of long-term fatigue load evaluations. The following research questions are therefore defined to structure the thesis.

Research questions

RQ1 What are the leading sources of uncertainty in wind turbine load prediction?

RQ2 What are the environmental conditions that most impact the fatigue loads in floating wind turbines?

RQ3 How does probabilistic model fitting compare in the case of onshore, fixed-bottom offshore, and floating wind turbines?

RQ4 Does uncertainty matter in long-term fatigue load evaluations?

1.6 Thesis outline

The structure of the thesis is outlined below.

Chapter 1 Introduction

A general introduction to the thesis covering the main research objectives and the chapter outline.

Chapter 2 Background and motivation for data-driven site analysis

Why investigate probabilistic load surrogates? This chapter provides a background, along with a detailed review of the existing literature, to highlight the research gaps.

Chapter 3 Methodology

This chapter presents the theoretical background of the various probabilistic data-driven surrogate models investigated in this part, along with the statistical correlation methods and the methods used to estimate the goodness of fit of a model.

Chapter 4 Load surrogates for fixed bottom wind turbines

The IEA 10-MW wind turbine's onshore configuration and offshore monopile are used as prototypes to develop surrogate models to map 10-minute statistics of site conditions to the corresponding load statistics. Several deterministic and probabilistic surrogate models are evaluated for predictive accuracy and ease of implementation. This chapter provides an overview of the simulation tools used to generate the data, set up the simulations, and analyze the results.

Chapter 5 Data analysis of the TetraSpar demonstrator measurements

A correlation study is conducted between the 10-minute DELs and the site parameters, using data obtained from the TetraSpar demonstrator. The analysis is repeated on one-to-one simulations performed with BHawC.

Chapter 6 Load surrogates for floating wind turbines

Mixture Density Network (MDN) is used to map 10-minute statistics of site conditions to the corresponding load statistics on a spar-buoy type floating wind turbine. The surrogate is further used to make probabilistic estimates of the lifetime fatigue damage. This chapter provides an overview of the simulation tools used for generating the data, setup, and analysis of the generated data before fitting the data-driven model.

Chapter 7 Conclusions

The conclusions, research experience, and lessons learned from the thesis are presented together in this chapter. It also presents the potential for future investigations and recommendations for the users of the surrogate models.

2

Background and motivation for data-driven site analysis

Selecting an appropriate site for a wind farm is crucial to ensuring the structure's safe and reliable operation throughout its lifespan. Site assessment typically involves detailed numerical analyses to evaluate the wind turbine's structural integrity under well-defined design load cases. However, as the number of site variables increases, the fatigue analysis, in particular, becomes computationally expensive. Data-driven models offer a solution by inferring fast-computing functions to map site conditions to fatigue loads, significantly reducing overall simulation time and expense. This chapter explores the motivation for developing such a framework, emphasizing the probabilistic nature of the problem. Additionally, it provides a comprehensive review of state-of-the-art data-driven load surrogate modeling.

This chapter is partly based on:
D. Singh et al. Probabilistic surrogate modeling of damage equivalent loads on onshore and offshore wind turbines using mixture density networks, Wind Energy Science [85].
D. Singh et al. Data-driven probabilistic surrogate model for floating wind turbine lifetime damage equivalent load prediction, Wind Energy Science Discussions [86].

2.1 Site analysis

To select a wind turbine for a specific site, the developers need assurance of its safe operation during its lifetime. The safety is compromised if any of the following limit states are encountered:

- Fatigue loads The wind turbine's structure, during its operational lifetime, is subject to constantly fluctuating loads due to complex and stochastic site conditions, leading to fatigue damage. The extent of this fatigue depends on various factors, including the type of foundation, material properties, control algorithms, and site-specific conditions.
- Ultimate loads Various scenarios, including extreme turbulence conditions and wave-induced loads, can induce loads that exceed the material's ultimate tensile strength, leading to breakage.
- Instability Resonance due to the excitation of the structure's natural frequency significantly increases the loads.
- Tower strike Modern wind turbines are designed with highly flexible blades that deflect and twist under dynamic aerodynamic loading. The wind turbine components must be designed so as to avoid the blade striking the tower under any circumstance.

It is up to the manufacturers to provide a detailed assessment of the structural response of the wind turbine for the site-specific conditions. Site analysis is crucial in estimating installation, maintenance, and operational costs, while ensuring that the fatigue and ultimate limit states are not exceeded during the wind turbine's lifetime.

The International Electrotechnical Commission (IEC) provides design standards [87, 88] that define DLCs for various loading, operational, and fault scenarios. These DLCs are simulated for site-specific analyses using a type-certified rotor-nacelle assembly with reference tower and foundation configurations. As projects progress, site-specific parameters, including foundation design, mooring systems (for FOWTs), and environmental conditions, are incorporated into detailed certification studies. The load simulations, performed using time-domain multi-physics engineering tools [23, 25, 26, 89, 90], calculate both fatigue and ultimate loads. However, the computational cost can be substantial, especially for fatigue calculations. In the case of offshore wind turbines, the computational cost of each 10-minute time-domain simulation can range from tens of minutes to a few CPU hours, depending on the complexity of the model, computing framework, and code efficiency. The total number of simulations that need to be evaluated at each site can be on the order of tens of thousands, significantly inflating the total computational overhead of a site assessment campaign.

Fatigue is a multi-scale phenomenon that depends on the material composition, composite structure, geometry, control operations, and inflow dynamics. The estimation of the fatigue damage for FOWTs, in particular, is computationally intense due to the more complex wave-structure coupling. The lifetime fatigue load assessment entails calculating the 10-minute DELs on multivariate bins of typical variables characterizing the site and scaling them to the observed probability of occurrence. Not all site variables

can be practically included in fatigue load analysis, as the required number of simulations increases exponentially with each additional variable. The choice of the variables in the offshore environment that have the most impact on FOWT fatigue is currently an active area of research [43, 91]. The total computational cost of the simulations also constrains the lower limit of the bin size. While industry-standard engineering tools are necessary for certification, the preliminary site analysis can benefit from data-driven surrogate models to provide quick load estimates. Data-driven surrogates can infer complex relationships from data observations alone, without requiring prior knowledge of the underlying physics.

2.2 Fatigue calculations

Structural fatigue

Fatigue damage is the initiation and progressive growth of microcracks in a structure from continuous stress fluctuations.

Fatigue damage results from continuous, cyclic loading on a structure over time. In wind turbines, the stochastic site conditions induce stresses on the structure that are irregular in nature, as illustrated in Figure 2.1. The stochastic loads are typically much lower than the yield stress; yet, over time, they lead to micro-cracks that eventually propagate and can cause severe damage. Fatigue assessment has gained increasing importance over the years as wind turbines have become larger and more economical, and are installed in sites with complex wind flow conditions and offshore locations, where they are subject to additional hydrodynamic loading.

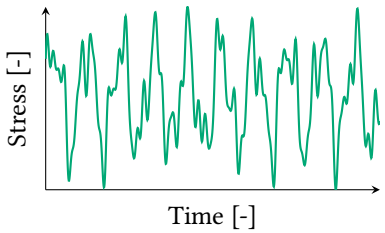


Figure 2.1: Sketch of a typical time-varying stress profile on a wind turbine.

Fatigue calculations typically follow the process shown in Figure 2.2. **Step 1** is to obtain historical data from the site. High-frequency (in seconds) information is useful in determining the turbulence characteristics of the wind profile, as well as extreme events such as gusts and jets. Low-frequency historical wind data (minutes to multi-hourly period) collected over several years is used to calculate the average wind histogram. As an example, in Figure 2.3, a time series of the reference wind speed is analyzed. The data are converted into a histogram with a bin size of 0.5ms^{-1}, allowing for the determination of wind speed frequency within specific ranges. From this, the probability of occurrence of each interval, P_{bin}, within the given time frame can be calculated. If additional variables are necessary to

characterize the site, their joint probability of occurrence is taken into account. In addition to the wind-related variables, FOWT fatigue calculations also require quantities such as significant wave height, wave direction, spectral wave period, wave spectrum, current speed, and current direction as a function of the water depth. The site data, in accordance with the certification requirements, is used to generate a table of simulations for fatigue calculations.

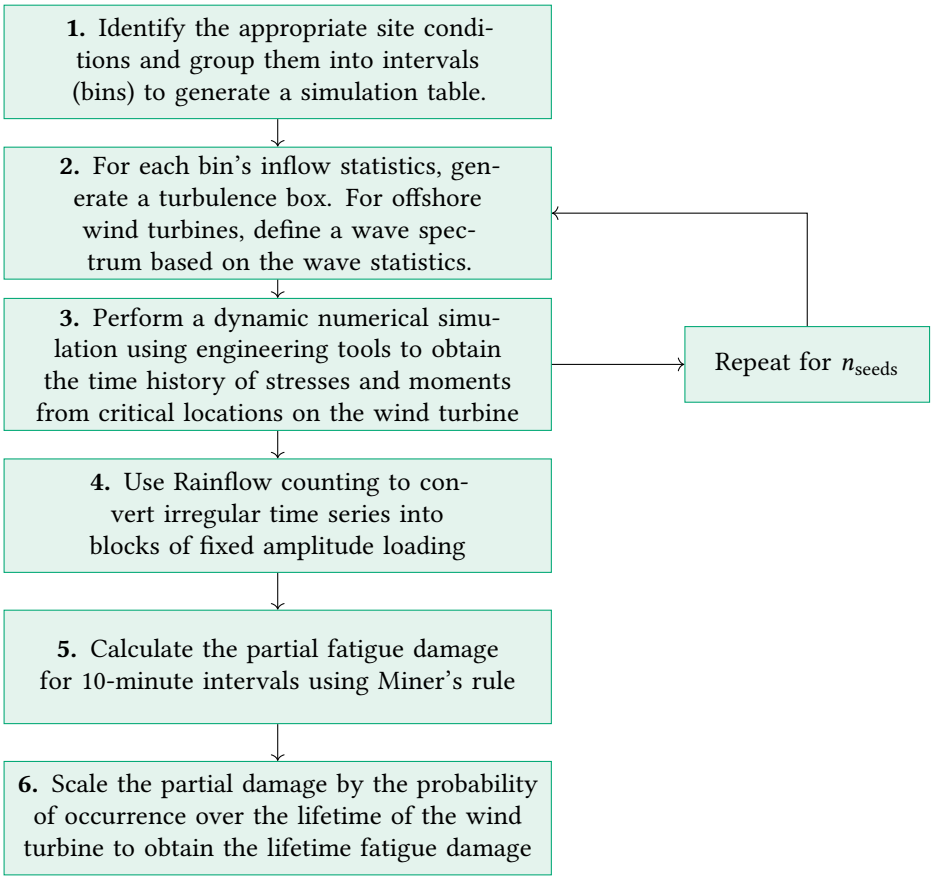


Figure 2.2: Process of calculating fatigue damage on wind turbine parts.

Step 2 involves generating a turbulent inflow field. Solving the Navier-Stokes equations to obtain the atmospheric turbulence field is computationally very expensive and, thus, infeasible. Instead, a number of stochastic methods have been proposed that model the statistics of the inflow rather than an accurate evolution of the eddies. Solvers such as TurbSim [93] and HAWC2 [25] use a spectral representation of turbulence based on the Kaimal [94] or the Mann model [95]. A turbulence spectrum describes how the energy of turbulent fluctuations is spread across different wave numbers (κ). κ, in turn, relates to the physical length scale (l) of the eddies. The turbulence generation modules of the time-domain engineering tools generate a turbulence field in the wave number space and

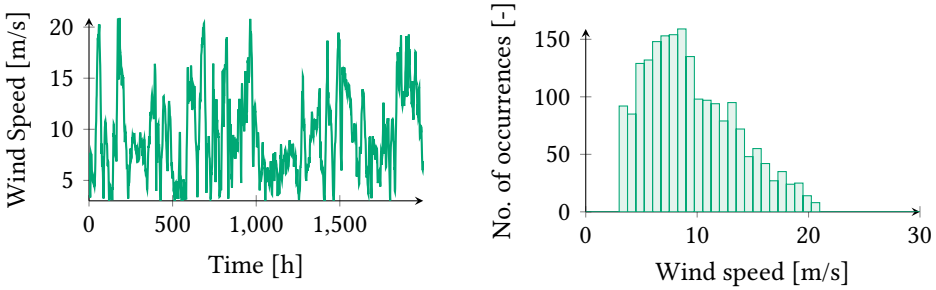


Figure 2.3: Example of an hourly wind speed time series obtained from the ERA5 reanalysis database [32] for Sud de la Bretagne [92] is shown on the left. The image on the right shows the corresponding histogram with the wind speed bins.

then transform it into physical space using the inverse Fourier transform. In the Kaimal approach, the turbulence box can be visualized as a series of two-dimensional grids of turbulent inflow, slightly larger than the rotor diameter. Subsequently, Taylor's frozen turbulence hypothesis is employed to propagate the two-dimensional grids downstream. An illustration is shown in Figure 2.4. The hypothesis assumes that the turbulent structures remain unchanged in the frame of reference of the mean wind speed. The Kaimal model accounts for spatial correlations between velocity components using empirical coherence functions, which approximate anisotropy but do not fully resolve it. Alternatively, the Mann or the uniform shear model uses rapid distortion theory to model turbulence affected by mean shear, combined with the concept of eddy lifetime, to produce a 3D spatial turbulence field. The advantage of this approach is that the three inflow components are generated simultaneously using an anisotropic spectral tensor formulation, which inherently captures the correlation between vertical, lateral, and longitudinal turbulence fluctuations. The stochastic methods can create any number of realizations of turbulence boxes by means of random seeds. The simulations are often repeated with multiple boxes to ensure seed independence.

Taylor's frozen turbulence hypothesis

Suppose the turbulent velocity fluctuations u' are much smaller than the mean velocity U, i.e., $\mid u' \mid \ll U$. In that case, Taylor's frozen turbulence hypothesis states that the turbulent eddies remain practically unchanged as they advect downstream with the speed U. This allows us to relate the spatial (x) and temporal (t) fields such that $t \approx \frac{x}{U}$.

For offshore wind turbines, it is crucial not only to define the turbulence model but also to characterize the wave energy spectra. This is essential for understanding the dynamic environmental conditions the turbines will face, as wave patterns significantly influence the instantaneous and fatigue loads on the system. There are various mechanisms of wave generation:

- The interaction between the wind and the surface of the ocean.

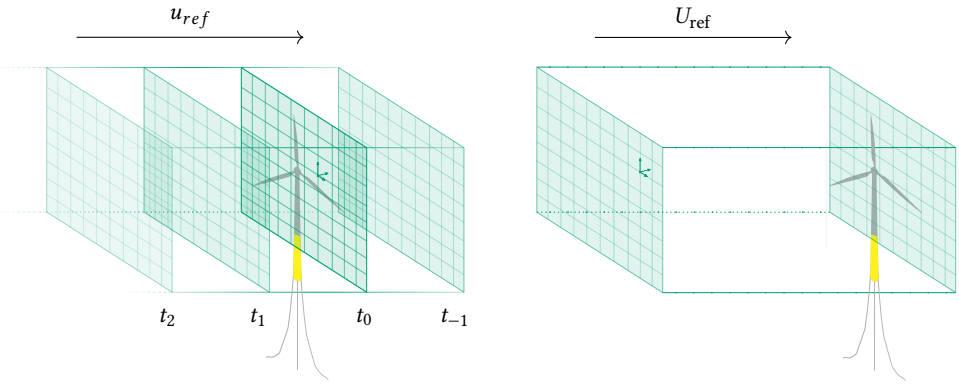


Figure 2.4: Illustration of turbulence boxes and the advection direction. The left image shows the Kaimal approach with the two-dimensional turbulence grids. The image on the right represents a 3D turbulence grid from the Mann model.

- Any large floating body that is moving forward or oscillating.
- The presence of a static large structure (relative to the wavelength) causes diffraction and alteration of the wave phase [96].
- Tidal movements due to astronomical forces.
- Movement in the ocean bottom due to earthquakes.
- The presence of internal waves as a result of temperature differences in the water body, causing density stratification.

Out of these, the predominant mechanism of wave generation is due to wind. The wind-generated waves are further characterized into two types. The first are called sea waves, which are local, irregular, and unpredictable perturbations of the ocean surface, generating short-crested waves. Swell is the other type of wave pattern that was wind-generated but has propagated several hundred kilometers away from the origin. These types of waves are more long-crested with smoother crest [97].

The wave elevation is generally measured at a point location using simple wave buoys to obtain a time history. Performing a statistical analysis of the time history yields the average wave period and wave height statistics. The mean period is simply the average period of the wave crests or troughs. Wave height statistics are more complex to compute.

As shown in Figure 2.5, the wave height (h) of successive zero-crossing peaks is calculated. The heights are binned into groups with fixed intervals. The number of occurrences is counted for each bin and divided by the total number of wave height counts to obtain the probability density function ($f(x)$) and the cumulative density function ($F(x)$) as shown in Figure 2.6. Significant wave height is defined in the time domain ($H_{1/3}$) as the average of the highest 1/3 of the wave heights,

$$H_{1/3} = \frac{3}{N} \sum_{j=1}^{N/3} h_j \tag{2.1}$$

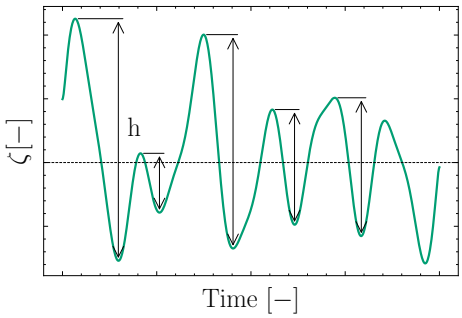


Figure 2.5: Time history of the wave elevation (ζ) with the successive wave heights (h) highlighted.

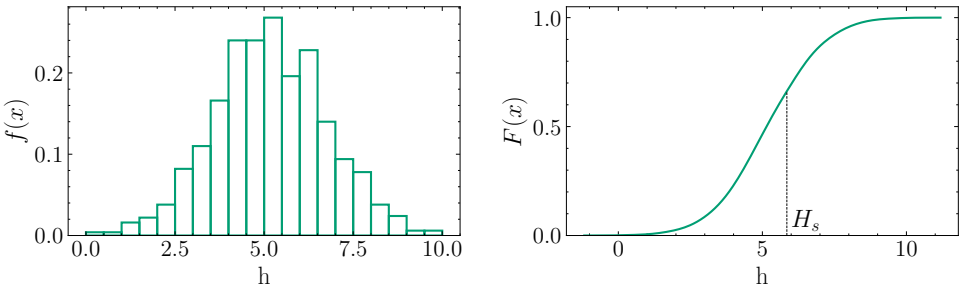


Figure 2.6: The corresponding probability density function and cumulative density function of the wave height measurements from time series data.

where j represents the rank and h_j are the individual wave heights.

More commonly, it is defined in the frequency domain (H_{m0}) as four times the standard deviation of the sea surface elevation σ_η. σ_η can be calculated by taking the square root of the zeroth-order moment of the wave spectrum (area under the spectral curve). Details of the frequency domain representation of waves are outlined in the following section. H_{m0} is expressed as,

$$H_{m0} = 4\sigma_\eta \tag{2.2}$$

H_{m0} and $H_{1/3}$ are used interchangeably as H_s in deep sea conditions. The spectral peak period or peak spectral wave period (T_p) is defined as the wave period with the highest energy peak in the wave spectrum.

Waves are stochastic in nature; therefore, they should be characterized in a probabilistic manner. Under the assumption of linearity, the irregular wave elevation profile can be decomposed into a set of regular sinusoidal waves with their own amplitude, phase, and frequency, which allows it to be transformed in the frequency domain. The energy content of the waves is defined using the wave energy spectra. The spectral models are often empirical, based on ocean data collected at several locations. Some of the simpler models, like the Pierson-Moskovitz spectrum [98], assume a single peak without a swell component for fully developed sea states. Fully developed sea state corresponds to conditions where

the wind-generated waves have reached their maximum potential/size for a given wind speed, wind duration and fetch. The Pierson-Moskowitz spectrum is described as,

$$S_{\mathrm{PM}}(\omega) = \frac{5}{16} H_s^2 \omega_p^4 \omega^{-5} \exp\left(-\frac{5}{4}\left(\frac{\omega}{\omega_p}\right)^{-4}\right) \tag{2.3}$$

where $\omega_p = \frac{2\pi}{T_p}$ is the angular spectral peak frequency. Hasselman et al. [99] later discovered through the Joint North Sea Wave Observation Project (JONSWAP) that the spectrum is never fully developed and presented an extension to the Pierson-Moskowitz spectra in the form of the JONSWAP model. It is described as,

$$S(\omega) = A_\gamma S_{\mathrm{PM}}(\omega) \gamma^{\exp(-0.5((\omega-\omega_p)/\sigma\omega_p)^2)} \tag{2.4}$$

where γ is the spectral peak parameter, which takes a value between one and seven. σ is the spectral width parameter, equal to 0.07 for $\omega \leq \omega_p$ and 0.09 otherwise. $A_\gamma = 1 - 0.287\log(\gamma)$. More recently, Torsethaugen et al. [100] published a two-peak spectral model suited for the North Sea conditions that also accounts for swell. It is a summation of two JONSWAP spectra fitted to measurements from the North Sea, with a number of empirical constants [100, 101]. Figure 2.7 compares the three spectral models for wind sea ($T_p = 6$s) and swell sea ($T_p = 12$s) states at $H_s = 2$m. In this work, the JONSWAP wave spectrum is used in Chapter 4 and Torsethaugen in Chapter 6.

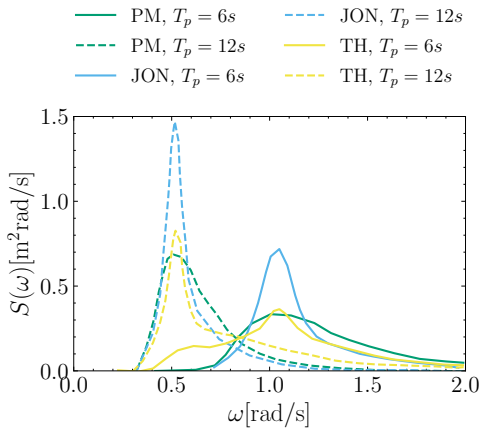


Figure 2.7: Comparison of Pierson-Moskowitz, JONSWAP and Torsethaugen wave spectral models for two wave periods, $H_s = 2$m [101].

Step 3 uses the inflow field and wave information to calculate the shear forces and moments on the blade and tower elements in response to the external aerodynamic and hydrodynamic forces. The Blade Element Momentum (BEM) model is used to calculate the forces on the rotor. BEM relies on several assumptions, such as incompressible, steady, inviscid flow, and blades acting as an actuator disk to facilitate practical computations of aerodynamic forces on the blades. Modern engineering tools combine sophisticated dynamic stall models, control laws, physical characteristics of the mechanical actuators,

hydrodynamic solvers, and structural models for blades, towers, foundation, and mooring lines to obtain loads on the complex wind turbine structure. Figure 2.8 highlights the various modules constituting a typical multi-physics engineering tool required for certification analyses. The length of the simulations per simulation is determined by the IEC standards [87]. For fatigue calculations on the tower and blades (DLC1.2), the length of the simulation is typically 10 minutes. The IEC also recommends that the simulations be statistically independent of the turbulence seed. Therefore, each simulation is repeated with multiple turbulence seeds, and the resulting short-term fatigue damage is averaged over the seeds. In FOWTs, the fatigue calculation of mooring lines requires longer simulations (several hours) as a result of their long natural periods.

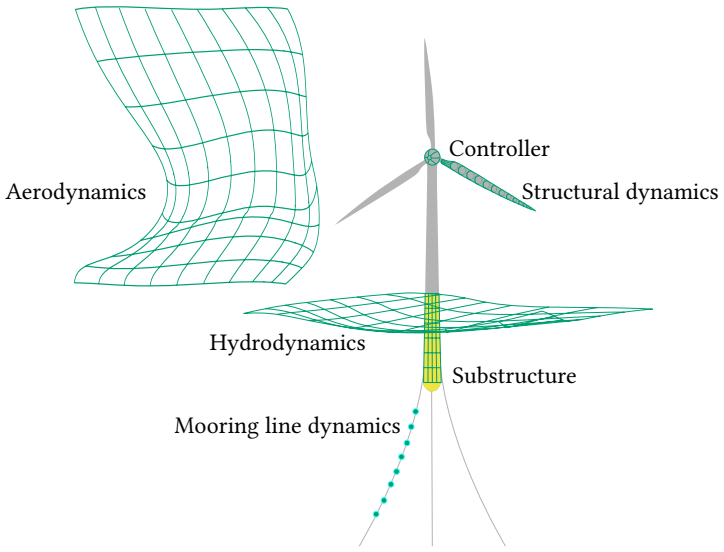


Figure 2.8: Various modules that comprise the coupled multi-physics engineering tools used for wind turbine simulations.

Rainflow counting

Method to count load range cycles from an irregular load time history. Rainflow counting decomposes the load time series into several sets, each consisting of a number of cycles per load range (or amplitude), in order to represent the fatigue load as a single value, termed equivalent load.

The irregular time history of stresses cannot directly be used to calculate the partial fatigue damage in the 10-minute window. In **step4**, it is first converted into blocks of constant amplitude loading using the Rainflow counting algorithm [102]. The algorithm identifies peaks and valleys in the load history, groups them into full and half cycles, and calculates their ranges and mean stress. As a result of Rainflow counting, the stress ranges (ΔS) and the corresponding number of occurrences (n) are obtained. In **step 5**, the S-N

curve is used to calculate the corresponding number of cycles to failure, N, if the material were subjected to a constant range cyclic loading with a stress range ΔS. If N cycles are not reached, only a fraction of the fatigue life is used up. This is termed partial damage. The partial damage for the i^{th} load range is simply $d_i = n_i/N_i$. The linear Palmgren-Miner accumulation law states that partial damages can be summed linearly to determine the total fatigue damage of the structure. Therefore, for the 10-minute stress signal, the total fatigue damage is given by $D_{10} = \sum d_i$.

Instead of dealing with many cycles of varying loading after Rainflow counting, an alternate approach is to calculate the 10-minute DELs. DEL is the constant range and frequency load that the structure would need to be subjected to in order to achieve the equivalent fatigue damage as the irregular stress signal. DEL is useful in comparing different designs and operating conditions, as the fatigue damage is reduced to a single representative value. To obtain 10-minute DEL, the equivalent number of load cycles, N_{ref}, is often chosen as 600, resulting in a frequency of 1Hz.

Damage equivalent load

DEL is the constant range and frequency load that the structure would need to be subjected to in order to achieve the equivalent fatigue damage as the irregular load signal. For example, a stress (or loads) time series can be decomposed into blocks of stress ranges S_i and the corresponding number of cycles n_i from the Rainflow counting algorithm. The DEL for the stress time series is given by,

$$\Delta S_{\text{eq}} := \left(\frac{\sum_{i=1}^{n} n_i \Delta S_i^m}{N_{\text{ref}}} \right)^{1/m} \tag{2.5}$$

where m is Wöhler's exponent. The value of m indicates how quickly the material accumulates damage as stress increases. N_{ref} is the reference equivalent number of cycles defined by the user.

To calculate the lifetime fatigue damage of a component, the probability of occurrence of each of the bins, P_{bin} obtained in the first step is required. In **step 6**, the probability is multiplied by the total number of 10-minute periods N_{10} in L years as,

$$N_{10} = P_{\text{bin}} \times 6 \times 24 \times 364.25 \times L. \tag{2.6}$$

The total damage in L years is a linear sum of 10-minute partial damages or of 10-minute DEL,

$$D = \sum N_{10i} \times D_{10i} \tag{2.7}$$

where i iterates over the bins. Finally, the safety factors for loads (γ_f), material (γ_m), and consequence of failure (γ_n) are taken into account to determine whether D is less than $1/(\gamma_f \gamma_m \gamma_n)$.

As seen here, the process of retrieving fatigue damage is extensive and requires many simulations. Additionally, there are several sources of uncertainty:

- The distributions of the site conditions derived from the historical data are not exact.

- The choice of the bin size is arbitrary.
- The turbulence and wave fluctuations are stochastic.
- The frequency of data acquisition (sample rate) can affect total fatigue.
- Engineering models and data-driven models have limitations and are dependent on parameters.

Predicting fatigue with fast, low-fidelity physics-based approaches is not straightforward, especially if the various uncertainties have to be accounted for, and can benefit from data-driven methods. Using surrogates that accurately estimate DEL can potentially eliminate the need for site data binning. This enables the direct use of multivariate joint distributions to represent the data more holistically. Surrogates can also be used to identify the most important features using global sensitivity analyses to ultimately reduce the dimensionality of the parameter space. Once trained, surrogate models can directly use all the available site information to estimate the site-specific DELs quickly. In addition, probabilistic surrogates can also propagate the statistical uncertainty from the stochastic input variables to the loads.

2.3 Probabilistic surrogates

Data-driven surrogates for wind turbine or wind farm-level loads are often designed with deterministic models. Given a training dataset with d-dimensional input parameters $\boldsymbol{x} \in \mathbb{R}^d$, the deterministic surrogate maps them to the corresponding output observations $y \in \mathbb{R}$. However, the assumption of a deterministic relationship between inputs and outputs does not hold in our case. For instance, keeping every other input constant, a single value of 10-minute mean wind speed can correspond to an infinite number of turbulent inflow patterns, resulting in an infinite number of DEL values with a certain probability distribution conditioned on that wind speed. Tools such as OpenFAST take as input a frozen turbulence field generated by TurbSim [93], consisting of stochastic turbulence patterns created using pseudo-random number generators initialized by random seeds. Thus, repeated simulations with the same input parameters but different seeds result in different values of output quantities, yielding a multi-valued mapping, such as that shown in Figure 2.9. On repeating the simulations with sufficient seeds, the load statistics converge towards a value of statistical moments that characterize a random variable, denoted $Y \mid \boldsymbol{X} = \boldsymbol{x}$, where $\boldsymbol{X}$ represents the joint Probability Density Function (pdf) of the input random variable and Y is the pdf of the corresponding response.

How is this uncertainty in DEL characterized? Uncertainty is categorized as either aleatoric or epistemic in nature. Aleatoric or irreducible uncertainty results from the intrinsic stochasticity or randomness of a process. The second type is epistemic uncertainty or reducible uncertainty, as it is a result of a lack of sufficient data or a subjective lack of knowledge. The combined effect of these can result in complex noise patterns in the data indicated by heteroscedasticity, non-Gaussianity, and multi-modality in the conditional response [103, 104]. The conditional target predicted by the surrogate combines the two sources of uncertainty.

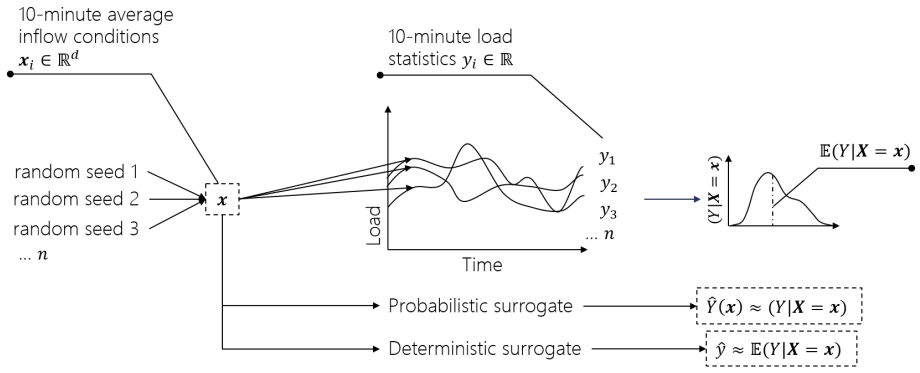
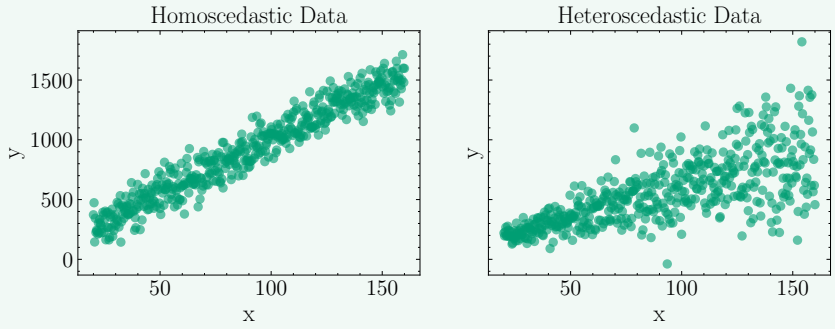


Figure 2.9: Schematic representation of the multi-valued mapping of the input-output pairs used to train the surrogate.

Heteroscedasticity and homoscedasticity

Homoscedasticity refers to the homogeneity of the conditional variance of a variable across all values of another set of variables. In contrast, heteroscedasticity occurs when the variance of a variable depends on the values of the independent variables. A simple example differentiating the two cases is shown in the figures below.



In regression problems, the homoscedasticity/ heteroscedasticity is talked about in terms of the nature of the errors or the residuals after fitting a model.

A simple regression model for a dataset with random variables $(\boldsymbol{X}, Y)$ can be expressed in the form

$$Y = f(\boldsymbol{X}) + \varepsilon, \tag{2.8}$$

where f is a deterministic function of the independent variable. Thus, $\mathbb{E}(Y) = f(\boldsymbol{X})$ is the deterministic component of the model. $\varepsilon \sim \mathcal{N}(0, \sigma)$ is the normally distributed noise component with a standard deviation of σ. A deterministic framework can only accommodate the average conditional of the target. Therefore, when deterministic models are used, the common practice is to convert the multi-valued problem to a single-valued

setting by averaging the response at each sample point $\boldsymbol{x}$ over many random seeds to approximate $\mathbb{E}(Y|\boldsymbol{X} = \boldsymbol{x})$. Some of the drawbacks of this approach include:

- A finite sampling of input loading due to stochastic representations of wind and waves in time-domain simulations introduces uncertainty in the turbine's load response. Although the IEC61400-3-1 [87] guidelines for DLC 1.2 do not explicitly state the number of seeds needed per case for statistical convergence, they do specify that the number and duration of simulations shall be sufficient to avoid seed dependence. Liew et al. [105] concluded that some load channels are more sensitive to the number of seeds than others. For training surrogates, it is common to run 60-minute-long simulations and obtain the average response or perform four to ten 10-minute simulations over different random seeds to obtain the average response before training the surrogate [45, 47, 53, 106]. The variability in the response with fewer seeds can be interpreted as noise, forcing the surrogate model to interpolate the noise in case of a small dataset or fit the mean of the response when the dataset is large. However, the mean may be wrongly inferred if the response is non-Gaussian. Seed repetitions add a significant computational cost to the data generation phase, especially when dealing with expensive simulations, as in the case of FOWTs.
- Modern wind turbines are equipped with sophisticated controllers that can lead to multi-modal responses in loads. Training the model on the expectation of a multi-modal distribution can misrepresent the actual load pattern, as the average of several correct target samples is not necessarily a meaningful target value.
- Most real-world learning tasks involve data sets with complex patterns of missing features that introduce aleatoric uncertainty. Unlike numerical simulations, seed repetitions cannot be performed on such datasets.

An alternate approach is to use probabilistic regression that models the targets as random variables with an unknown conditional distribution, $Y|\boldsymbol{X} = \boldsymbol{x}$. Probabilistic models provide a framework for informed decision-making by predicting a confidence interval in addition to the average response. For instance, the uncertainty in the 10-minute DELs can be propagated to the lifetime DELs while also considering the distribution of the wind speeds to design less conservative, site-specific structures. Probabilistic surrogates such as conditional generative adversarial networks and conditional variational autoencoders use latent variables to infer meaningful quantities from data with complex noise distributions [107–110]. Other probabilistic regression approaches like MDNs [111] and generalized lambda distributions [112] use maximum likelihood estimates to infer the conditional distributions in stochastic systems. Simulation repetitions before training are therefore unnecessary and significantly reduce training time.

2.4 Advances in surrogate modeling of damage equivalent loads

Wind turbine loads, for site-specific fatigue analysis and wind farm design, have commonly been approached with deterministic models like standard Artificial Neural

Networks (ANNs), [46, 47, 113]. ANNs are extremely powerful and emulate the loads well if the training data have been averaged over a set of inflow turbulence realizations. Shaler et al. [47] compare the performance of inverse distance weighting, ANNs, radial basis functions, Kriging with a partial least squares dimension reduction, and regularized minimal-energy tensor-product b-splines in a wind farm array. They observe the highest R^2 values for ANNs and the inverse distance weighting method. These approaches, however, do not aim to account for the variance of the load response.

The standard Gaussian Process Regression (GPR) [114] is a probabilistic surrogate that is capable of uncertainty quantification. However, in its standard form, it is restricted to normally distributed homoscedastic responses. Nevertheless, due to its flexibility and ease of implementation, it is widely used as a surrogate to estimate the fatigue load response in wind turbines [51, 115–117]. Gasparis et al. [118] compare GPR to other data-driven methods like linear regression and ANNs for modeling power and fatigue loads, showing superior performance by the GPR. Similarly, Dimitrov et al. [45] evaluate importance sampling, nearest-neighbor interpolation, Polynomial Chaos Expansion (PCE), GPR, and quadratic response surface (QRS), to conclude a better performance again by the GPR despite a computational penalty. Slot et al. [53] provide a thorough comparison of the performance of PCE and GPR for the uncertainty quantification of fatigue loads on NREL's 5MW reference onshore wind turbine. They conclude the need for a minimum of four random seeds per training sample in the case of GPR to make high-accuracy predictions. They also note that GPR performs better per invested training simulation than PCE. Haghi et al. [119] use temporal convolutional networks (TCNs) and acceleration-infused TCN to project synthetic unsteady wind time series obtained from TurbSim onto wind turbine DELs. The models take complex information from the wind field and make relatively good estimates of the load channels close to the rotor. The methods are, however, shown to be complex and lose accuracy in the tower bottom channels.

The dependence of short-term fatigue load variability on inflow conditions has been recognized for several decades, with early studies demonstrating statistical variation in fatigue loads due to wind speeds [120]. Therefore, interest in quantifying the uncertainty of the short-term fatigue loads as a function of the input parameters has initiated research into heteroscedastic surrogates. *Heteroscedasticity* refers to the heterogeneity in the response variance as a function of the inputs. The variance observed in DEL at the tower bottom, for instance, for very large values of significant wave height is generally larger than that in calm ocean conditions. It is, therefore, an important consideration when choosing the appropriate surrogate modeling approach for load uncertainty quantification. Murcia et al. [48] use 100 turbulent inflow realizations at each sample point to obtain the first two statistical moments of the fatigue response. Thereafter, they create two independent surrogates using PCE to model the mean and standard deviation of the fatigue loads on the DTU 10MW reference wind turbine separately. Even though they use only 140 training samples for their model, the replications scale the computational cost by 100, eventually leading to a costly training database. Another replication-based approach is taken by Zhu et al. [121] to model the load response using generalized lambda distributions. In this study, 50 inflow wind field realizations are used at each input sample to estimate the four lambda parameters. Four PCE surrogates are then used to model the parameters independently. The main drawback of replication-based methods is the cost of generating the training database,

which makes it challenging to apply them to computationally demanding applications such as FOWTs. Secondly, the goodness of fit relies heavily on estimating the statistical parameters in the first step.

Heteroscedasticity can also be modeled using statistical methods. Zhu et al. [112] extend the replication-based approach to derive a statistical method combining generalized least squares with maximum conditional likelihood to estimate the lambda parameters without replications. The main advantage of this method is that it does not assume a Gaussian distribution. However, it cannot handle multi-modality. Haghi et al. [122] use Approximate Gaussian process regression (AGPR) on hybrid simulation data from NREL's 5MW reference wind turbine [123] and SCADA data from a comparable, commercial turbine to make predictions of DELs on the tower and blades. AGPR captures the heteroscedasticity in the response. Abdallah et al. [124] utilize parametric hierarchical Kriging to predict the extreme loads on the blade root bending moment of a 2MW onshore wind turbine, which are heteroscedastic. Their approach combines low- and high-fidelity observations, where the low-fidelity model informs the high-fidelity GPR. They demonstrate that introducing hierarchy helps make the model selection process more robust than the manual tuning of GPR parameters. The application of probabilistic surrogates to FOWTs has only been studied to a limited extent in the literature. Li et al. [116, 117] model the uncertainty in the site conditions using a C-vine copula combined with ANN and GPR to make long-term probabilistic estimates of the DELs on a FOWT.

Evidently, only a few approaches attempt to model the uncertainty in the load response of the turbine and the tower. Of those that do, only some consider complex offshore fixed or floating systems, which is the primary focus of this thesis. In Chapter 4, a probabilistic machine learning framework to map site conditions to the 10-minute DELs is introduced. Several commonly used regression approaches on onshore and offshore fixed-bottom wind turbines are compared. These regression models range from the simplest deterministic polynomial fits to increasingly complex probabilistic models. Based on this analysis, in Chapter 6, the best regression model is selected to estimate the probabilistic fatigue response of a spar-type 6 MW FOWT. The 10-minute probabilistic estimates are further used to determine the aggregated probabilistic fatigue damage on the structure.

3

Methodology

This chapter presents various statistical approaches used subsequently in the thesis for feature analysis, surrogate modeling, and determining the goodness of fit.

This chapter is partly based on:
D. Singh et al. Probabilistic surrogate modeling of offshore wind-turbine loads with chained Gaussian processes, IOP Journal of Physics Conference Series [125].
D. Singh et al. Probabilistic surrogate modeling of damage equivalent loads on onshore and offshore wind turbines using mixture density networks, Wind Energy Science [85].
D. Singh et al. Data-driven probabilistic surrogate model for floating wind turbine lifetime damage equivalent load prediction, Wind Energy Science Discussions [86].
D. Singh et al. Data analysis of the TetraSpar demonstrator measurements [43].

3.1 Background

This chapter presents the methodology behind various statistical and machine learning techniques used in this thesis. First, the methods for feature selection applied in Chapter 5 are introduced. This is followed by the regression models used in Chapter 4 and Chapter 6. Finally, the commonly used metrics for measuring the goodness of fit of the regression model are presented.

3.2 Methods for feature selection

Feature selection is an essential pre-processing step to machine learning and data-driven modeling. A feature is any observed quantity or a derived value based on observations, serving as an input to the data-driven model [126, 127]. Training a surrogate with an extensive feature set or with redundant features can lead to increased training data needs, data storage, and computational and measurement costs. As not all available variables significantly impact the quantity of interest, they can often be merged together or eliminated through feature selection techniques. Feature selection has the following advantages:

- Reduction of the computational cost associated with generating training data or training the model.
- The system processes can be described and understood with a simpler model.
- The resulting surrogate is more generalizable.

Feature selection

One goal of feature selection is to identify the smallest number of features necessary to build an accurate predictor for estimating a quantity of interest.

Feature selection methods are grouped into three broad categories [127]. Filter methods use statistical correlation methods to find dependence and relationships between variables. Wrapper methods entail trying various combinations of features to repeatedly train the model and choose the ones that result in the best goodness-of-fit values. Embedded methods include specific tree-based methods, lasso regression, and regularization techniques that have feature selection embedded in the process. Thus, an explicit pre-processing step is not needed. Finally, an expert understanding of the underlying physical processes of the system can also help isolate its key features.

Filter methods are simple to use and model-independent. Therefore, they allow us to independently perform feature analysis on data before deciding which surrogate model fits best. One common approach in filtering is assessing correlation, which quantifies the degree and nature of statistical association between variables. Correlation-based feature selection generally involves the following steps. (i) A correlation matrix containing the correlation coefficients between each combination of input features is prepared. (ii) Highly

correlated features are identified. When two variables are strongly correlated, one may be redundant and can be excluded from model training. (iii) The correlation between the resulting features and the target is calculated to eliminate uncorrelated input quantities.

Various kinds of correlation coefficients exist to measure these linear or nonlinear relationships. This section presents the coefficients used in this study in more detail. Two random variables $A : \Omega \to \mathbb{R}$ and $B : \Omega \to \mathbb{R}$ are used as an example, with realizations a and b, respectively.

3.2.1 Pearson's correlation coefficient

Pearson's Correlation Coefficient (PCC) is a simple measure of linear correlation between two random variables. The values of PCC are bound by -1 and 1. 0 denotes a lack of linear correlation, whereas the negative and positive values are interpreted as inverse and direct linear relationships between the variables. Mathematically, it is expressed as

$$\rho(A,B) := \frac{\sum_{i=1}^{n}(a_i-\bar{a})(b_i-\bar{b})}{\sqrt{\sum_{i=1}^{n}(a_i-\bar{a})^2}\sqrt{\sum_{i=1}^{n}(b_i-\bar{b})^2}}. \tag{3.1}$$

It is calculated on a set of n discrete samples. $\bar{a}$ and $\bar{b}$ represent the means of the variable populations. PCC has a major drawback wherein the coefficient can be misleading if the variables are non-linearly correlated. Nevertheless, it has been widely used as a feature analysis tool for various wind energy surrogate modeling applications [59, 81, 128–130].

3.2.2 Mutual information

Mutual Information (MI) is a measure of the amount of information shared between two random variables. In other words, it measures the reduction in uncertainty of a random variable B, given the knowledge of another random variable A. The MI value ranges from 0 (complete independence) to a positive maximum (perfect correlation). The value is a function of the units of measurement of the quantities in question. For two discretely joint random variables A and B with a joint probability distribution $\mathbb{P}_{AB}(a,b)$, and marginal distributions $\mathbb{P}_A(a), \mathbb{P}_B(b)$, MI is defined as [131],

$$I(A;B) = \sum_{b\in B}\sum_{a\in A}\mathbb{P}_{A,B}(a,b)\log\left(\frac{\mathbb{P}_{A,B}(a,b)}{\mathbb{P}_A(a)\mathbb{P}_B(b)}\right). \tag{3.2}$$

MI is symmetric and invariant under invertible and differentiable transformations in the feature space. It can measure nonlinear correlations and is not sensitive to outliers, making it a robust metric.

3.2.3 Conditional dependence coefficient

Conditional Dependence Coefficient (CODEC) was introduced by Azadkia et al. [132]. As opposed to PCC and MI, CODEC is designed to account for the conditional dependence between variables. This is interesting because the correlation between two variables can change when conditioned on a third variable. This non-parametric approach can be used with randomly sampled, unstructured data, allowing us to use it on measurements directly. To define CODEC, firstly, a multivariate setting is established. Let $C : \Omega \to \mathbb{R}$ be a random variable. $\boldsymbol{A} : \Omega \to \mathbb{R}^{D_A}$ and $\boldsymbol{B} : \Omega \to \mathbb{R}^{D_B}$ be random vectors with dimensions

$D_B \geq 1$ and $D_A \geq 0$. CODEC defines the degree of conditional dependence between C and $\boldsymbol{B}$ given $\boldsymbol{A}$. The coefficient is defined as,

$$T(C, \boldsymbol{B} \mid \boldsymbol{A}) := \frac{\int_{-\infty}^{\infty} \mathbb{E}(\mathrm{Var}(\mathbb{P}(C \geq t \mid \boldsymbol{B}, \boldsymbol{A}) \mid \boldsymbol{A}))\mathrm{d}\mu(t)}{\int_{-\infty}^{\infty} \mathbb{E}(\mathrm{Var}(1_{\{C \geq t\}} \mid \boldsymbol{A}))\mathrm{d}\mu(t)}. \tag{3.3}$$

$\mathbb{P}(C \geq t \mid \boldsymbol{B}, \boldsymbol{A})$ is the conditional probability that the random variable C is greater than a threshold t, given $\boldsymbol{B}$ and $\boldsymbol{A}$. If the probability barely changes with $\boldsymbol{B}$, then its variance given $\boldsymbol{A}$ would be low. The numerator of $T(C, \boldsymbol{B} \mid \boldsymbol{A})$ integrates this variance over all possible thresholds t. The indicator function $1_{\{C \geq t\}}$ in the denominator is 1 if $C \geq t$, and 0 otherwise. The variance $\mathrm{Var}(1_{\{C \geq t\}} \mid \boldsymbol{A})$ captures the uncertainty in predicting whether C exceeds t based only on $\boldsymbol{A}$, ignoring $\boldsymbol{B}$. This is also integrated over t to account for the overall variability of C across different thresholds. In summary, $T(C, \boldsymbol{B} \mid \boldsymbol{A})$ provides a measure of how much of the variance in C is explained by $\boldsymbol{B}$ beyond what is already explained by $\boldsymbol{A}$.

The authors, Azadkia et al. [132], developed an estimator T_n to be able to approximate T on n discrete data samples. The main idea is to approximate the expectation and variance terms in Equation (3.3) using rank-based nearest neighbors. Let there be n independent and identically distributed samples from the joint distribution of $(C, \boldsymbol{A}, \boldsymbol{B})$ as $(C_1, \boldsymbol{A}_1, \boldsymbol{B}_1)...(C_n, \boldsymbol{A}_n, \boldsymbol{B}_n)$ with $n \geq 2$. Then, T_n is defined as,

$$T_n(C, \boldsymbol{B} \mid \boldsymbol{A}) := \frac{\sum_{i=1}^{n}(\min\{R_i, R_{M(i)}\} - \min\{R_i, R_{N(i)}\})}{\sum_{i=1}^{n}(R_i - \min\{R_i, R_{N(i)}\})}. \tag{3.4}$$

For the i^{th} data sample, $N(i)$ is the index j, such that $\boldsymbol{A}_j$ is the nearest neighbor of $\boldsymbol{A}_i$ in $\mathbb{R}^{D_A}$. Similarly, $M(i)$ is the index j such that $(\boldsymbol{A}_j, \boldsymbol{B}_j)$ is the nearest neighbor of $(\boldsymbol{A}_i, \boldsymbol{B}_i)$ in $\mathbb{R}^{D_A + D_B}$. R_i is the rank of C_i (number of j such that $C_j \leq C_i$). The equivalence of the estimator to Equation (3.3) is elaborated in the original publication [132]. The estimator is easy to implement as it only uses basic search algorithms to find the nearest neighbors and can be easily scaled to large datasets. However, it has an unstable behavior when used on a small dataset with repeating values.

The value of CODEC can be further used to order the features in terms of their importance on the response variable. This process is termed Feature Ordering by Conditional Independence (FOCI) [132]. The initial dataset comprises a response variable C and input variable $\boldsymbol{A}$. The FOCI algorithm is defined as follows:

1. Initialize:
   - Set the selected feature set $S = \emptyset$.
   - Set the original feature set $R = \{A_1, A_2, \ldots, A_{D_A}\}$.
2. Iterate until $T_n \leq 0$
   (a) For each feature $A_j \in R$, compute the conditional dependence coefficient
   $$T_n(C, A_j \mid S).$$
   (b) Select the feature A^* that maximizes the conditional dependence
   $$\arg\max_{A_j \in R} T_n(C, A_j \mid S).$$

(c) If $T_n(C, A^* \mid S) \geq 0$:
- Add A^* to S.
- Remove A^* from R.

(d) Else, exit the loop.

3. Return: The selected feature subset S.

3.3 Regression

Supervised regression is defined as the task to learn a mapping f from a set of d-dimensional input variables $\boldsymbol{x} \in \boldsymbol{X}$ to some observed quantity of interest $y \in Y$. The inputs are referred to as features or predictors, and the output is termed target or response. The regression training set is defined as a set of n_{train} pairs $\{\boldsymbol{x}_q, y_q\}$ with $q = 1...n$. $\boldsymbol{x}$ and y are real-valued and considered to be realizations of random variables $\boldsymbol{X} : \Omega \to \mathbb{R}^d$ and $Y : \Omega \to \mathbb{R}$, respectively. During the training process, the performance of the model with parameters $\boldsymbol{\theta}$ is measured using a loss function $\mathcal{L}(\boldsymbol{\theta})$. The training objective is to infer the parameters so that the loss is minimized.

Loss function

The loss function is a measure of the performance of the model during training. For deterministic approaches, a commonly used loss function is the Mean Squared Error (MSE), defined as:

$$MSE(\boldsymbol{\theta}) = \frac{1}{N_{\text{train}}} \sum_{n=1}^{N_{\text{train}}} (y_n - f(\boldsymbol{x}_n; \boldsymbol{\theta}))^2$$

Probabilistic models, on the other hand, commonly use the Negative Log-Likelihood (NLL) as the loss function. NLL is defined as,

$$NLL(\boldsymbol{\theta}) = -\frac{1}{N_{\text{train}}} \sum_{n=1}^{N_{\text{train}}} \log(p(y_n \mid f(\boldsymbol{x}_n; \boldsymbol{\theta})))$$

Minimizing the NLL allows us to compute the maximum likelihood estimate (MLE). The idea is to pick the parameters that assign the highest probability to the training data.

Prior to training, the data are always scaled or normalized, either with a minmax scaler or a standard scaler. For each vector dimension, the minmax scaler is defined as,

$$x_{i,\text{scaled}} = \frac{x_i - \min x}{\max x - \min x}. \tag{3.5}$$

Similarly, the standard scaler is defined as,

$$x_{i,\text{scaled}} = \frac{x_i - \bar{x}}{\sigma}, \tag{3.6}$$

where $\bar{x}$ is the mean, σ is the standard deviation, $\min x$ and $\max x$ are the minimum and maximum values of the dataset along each dimension. Feature scaling is an essential step to bringing the various physical quantities in different units to a single scale.

3.3.1 Linear regression

Linear regression is the simplest statistical approach to determining a linear relationship between variables. Because of its simplicity, it is easy to implement and interpretable. In linear regression, the predicted output is expressed as a linear combination of the input features:

$$f(\boldsymbol{x}, \boldsymbol{\theta}) = b + \boldsymbol{w}^\intercal \boldsymbol{x} \tag{3.7}$$

where $\boldsymbol{\theta} = (b, \boldsymbol{w})$ are the model's parameters consisting of the bias term b and the weights $\boldsymbol{w}$. The optimal parameters $\hat{\theta}$ are determined by minimizing the MSE [133].

3.3.2 Polynomial regression

Polynomial regression can be seen as an extension of the linear model. For a polynomial of degree P, a feature vector is defined as,

$$\phi(\boldsymbol{x}) = [1, \boldsymbol{x}, \boldsymbol{x}^2, \boldsymbol{x}^3, ..., \boldsymbol{x}^P]. \tag{3.8}$$

Thus, the model is defined as,

$$f(\boldsymbol{x}, \boldsymbol{\theta}) = \boldsymbol{w}^\intercal \phi(\boldsymbol{x}) \tag{3.9}$$

Similar to the linear model, f is a linear combination of $\phi(\boldsymbol{x})$, despite being a nonlinear function of $\boldsymbol{x}$. The optimal parameters are determined by minimizing the MSE.

3.3.3 Gaussian process regression

GPR is a non-parametric, flexible, Bayesian machine learning framework that models distributions over functions. A Gaussian process ($\mathcal{GP}$) is a generalization of the Gaussian distribution. Just as distributions are associated with random variables, processes describe the properties of functions. From a set of infinite possible functions, an assumption about the properties and distribution of the functions, before accessing any data, is encoded in a Gaussian prior. Figure 3.1 shows $\mathcal{GP}$ proper draws on the left, generated using a radial basis function (RBF) kernel. On the right, the posterior functions are updated in the presence of new information, that is, the data samples. A confidence interval of 2σ is indicated in gray. It collapses where the training data samples are present. The confidence interval grows smaller as more training samples are introduced.

Gaussian process

A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution.

C. E. Rasmussen & C. K. I. Williams,
Gaussian Processes for Machine Learning, 2006

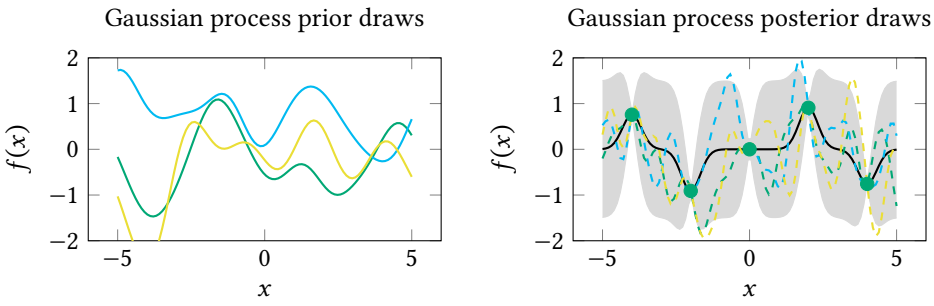


Figure 3.1: The left image shows sample functions from the prior with an RBF kernel. On the right, the image with the training data points shows the posterior function draws in dashed lines, with the confidence interval of the prediction in gray.

Model definition

The regression problem is defined as,

$$y = f(\boldsymbol{x}) + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2). \tag{3.10}$$

The standard GPR models the noise, ε, as a normally distributed quantity with a variance of σ^2. The Gaussian process is specified as,

$$f(\boldsymbol{x}) \sim \mathcal{GP}(\mu(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}')) \tag{3.11}$$

where the mean $\mu(\boldsymbol{x})$ and the covariance functions $k(\boldsymbol{x}, \boldsymbol{x}')$ are defined as,

$$\mu(\boldsymbol{x}) = \mathbb{E}[f(\boldsymbol{x})] \tag{3.12}$$

$$k(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[(f(\boldsymbol{x}) - \mu(\boldsymbol{x}))(f(\boldsymbol{x}') - \mu(\boldsymbol{x}'))]. \tag{3.13}$$

The covariance function dictates the smoothness of the function. In this work, the squared exponential kernel is used. It is also known as the Radial Basis Function kernel, or the Gaussian kernel. It is widely used for GPR as it is smooth, infinitely differentiable, and is characterized by $d+1$ parameters. It is expressed as

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma_h^2 \exp\left(-\frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}'\|_{\boldsymbol{l}}^2\right), \quad \|\boldsymbol{x} - \boldsymbol{x}'\|_{\boldsymbol{l}}^2 := \sum_{j=1}^{d} \frac{|x^{(j)} - x'^{(j)}|^2}{l^{(j)}}. \tag{3.14}$$

$x^{(j)}$ is the j-th component of $\boldsymbol{x}$. The kernel function $k(\boldsymbol{x}, \boldsymbol{x}')$ is used to compute the covariance matrix K_{XX}. The characteristic length-scales $\boldsymbol{l} \in \mathbb{R}^d$ are defined per input parameter, and these and the variance σ_h^2 are hyperparameters that are tuned based on the training data. The length scales, for instance, determine the degree of correlation the function value at $\boldsymbol{x}$ has with the function value at a neighboring point $\boldsymbol{x}'$. The smaller the length scale, the closer the adjacent function values become uncorrelated.

Loss function

The hyperparameters σ_h and $\boldsymbol{l}$ may be fixed by the user, but an optimal value is often inferred from the data using type-II maximum likelihood [114], wherein the negative

log marginal likelihood is minimized with respect to the hyperparameters. The negative log marginal likelihood is defined as,

$$-\log p(\boldsymbol{y}|\mathrm{X},\sigma_h,\boldsymbol{I}) = \frac{1}{2}(\boldsymbol{y}-\mu(\mathrm{X}))^\top (K_{\mathrm{XX}}+\sigma^2 I)^{-1}(\boldsymbol{y}-\mu(\mathrm{X})) + \frac{1}{2}\log|K_{\mathrm{XX}}+\sigma^2 I| + \frac{n_{train}}{2}\log 2\pi \tag{3.15}$$

The parameters of the model are determined using gradient-based optimization functions. In this application, Limited memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm [134] with box constraints (L-BFGS-B) has been used.

Prediction

The aim is to make predictions $y^\star$ on an unseen data point $\boldsymbol{x}^\star$. The observations $\boldsymbol{y}$ and prediction $y^\star$ are jointly Gaussian, as shown in Equation (3.16).

$$\begin{bmatrix}\boldsymbol{y}\\ y^\star\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\mu(\mathrm{X})\\ \mu(\boldsymbol{x}^\star)\end{bmatrix}, \begin{bmatrix}K_{\mathrm{XX}}+\sigma^2 I & K_{\mathrm{X}\boldsymbol{x}^\star}\\ K_{\boldsymbol{x}^\star\mathrm{X}} & K_{\boldsymbol{x}^\star\boldsymbol{x}^\star}+\sigma^2 I\end{bmatrix}\right) \tag{3.16}$$

The joint distribution is conditioned on the observed values to get the predictive distribution corresponding to a new input $\boldsymbol{x}^\star$ as,

$$y^\star \mid \boldsymbol{y},\mathrm{X},\boldsymbol{x}^\star \sim \mathcal{N}(\hat{\mu}(x^\star),\hat{\Sigma}(x^\star)) \tag{3.17}$$

$$\hat{\mu}(\boldsymbol{x}^\star) = \mu(\boldsymbol{x}^\star) + K_{\boldsymbol{x}^\star\mathrm{X}}(K_{\mathrm{XX}}+\sigma^2 I)^{-1}(\boldsymbol{y}-\mu(\mathrm{X})) \tag{3.18}$$

$$\hat{\Sigma}(\boldsymbol{x}^\star) = K_{\boldsymbol{x}^\star\boldsymbol{x}^\star} - K_{\boldsymbol{x}^\star\mathrm{X}}(K_{\mathrm{XX}}+\sigma^2 I)^{-1}K_{\mathrm{X}\boldsymbol{x}^\star}+\sigma^2 I \tag{3.19}$$

3.3.4 Chained Gaussian processes

In the standard GPR approach, the mean of the likelihood is an $\boldsymbol{x}$ dependent latent function while the additive noise element remains constant. For many applications, this assumption does not hold as the noise function itself can be dependent on $\boldsymbol{x}$. Such processes are termed heteroscedastic. Figure 3.2 shows a schematic of GPR and Chained Gaussian Process (CGP) on a heteroscedastic dataset, highlighting the benefit of including heteroscedasticity in the framework. For such data, the confidence intervals are much more reliable in the CGP case, as the assumption of heteroscedasticity is accounted for in the model definition.

Model definition

A logical generalization of Equation (3.10) is to write ε as a function of the input variables [135],

$$y = f(\boldsymbol{x}) + \varepsilon(\boldsymbol{x}), \tag{3.20}$$

$$\varepsilon \sim \mathcal{N}(0,\sigma^2(\boldsymbol{x})). \tag{3.21}$$

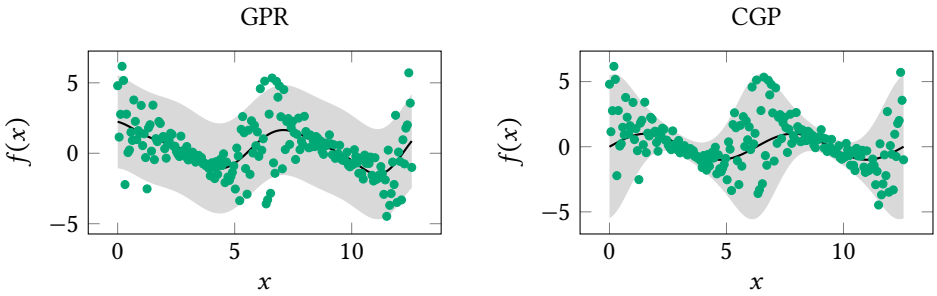


Figure 3.2: Figures showing the difference between the standard Gaussian process regression assuming a fixed noise term, versus the heteroscedastic approach with chained Gaussian process regression.

A new latent function $g(\boldsymbol{x})$ is introduced to describe the variance of the noise. Specifically, to guarantee positive values, the variance is assigned a log-GP prior,

$$y \sim \mathcal{GP}(f(\boldsymbol{x}), e^{g(\boldsymbol{x})}) \tag{3.22}$$

$$f(\boldsymbol{x}) \sim \mathcal{GP}(\boldsymbol{\mu}_f, k_f(\boldsymbol{x}, \boldsymbol{x}')) \tag{3.23}$$

$$g(\boldsymbol{x}) \sim \mathcal{GP}(\boldsymbol{\mu}_g, k_g(\boldsymbol{x}, \boldsymbol{x}')). \tag{3.24}$$

$f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are independent latent processes with individual means and covariances. Contrary to the standard GPR, the posterior is no longer analytically tractable and must be approximated. One of the emerging approximation methods in the field of statistical inference is variational inference (VI) [107]. In this study, CGPs [135] that use VI have been used to solve Equation (3.20). CGPs build on the idea of generalized linear models (GLMs), which use link functions to relate the mean function of a GP to the mean of any other distribution from the exponential family of distributions. A major drawback of GLMs is that the link function must be invertible, which renders the latent functions additive Gaussian. CGPs propose a generalized framework that allows for nonlinear combinations of any number of latent processes.

In Gaussian processes with n_{train} data points, the cost of inference is $\mathcal{O}(n_{train}^3)$, which severely limits its scalability. The CGP implementation used also includes a sparse scheme, which reduces the computational cost to $\mathcal{O}(n_{train} I^2)$ for I inducing points. The inducing points are used for parametrizing the covariance matrix, and their locations are inferred variationally. The posterior distribution to be inferred is defined as,

$$p(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{u}_f, \boldsymbol{u}_g | \boldsymbol{y}) = \frac{p(\boldsymbol{y} | \boldsymbol{f}, \boldsymbol{g}, \boldsymbol{u}_f, \boldsymbol{u}_g) p(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{u}_f, \boldsymbol{u}_g)}{p(\boldsymbol{y})}, \tag{3.25}$$

where the functions $\boldsymbol{f} = f(\boldsymbol{x})$ and $\boldsymbol{g} = g(\boldsymbol{x})$ are valued $\boldsymbol{u}_f$ and $\boldsymbol{u}_g$ at inducing point location $\boldsymbol{z}$. The issue is that the marginal distribution $p(\boldsymbol{y})$ is intractable; therefore, the exact posterior distribution cannot be determined. It is approximated by a variational distribution $q \in \mathcal{Q}$ such that,

$$p(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{u}_f, \boldsymbol{u}_g | \boldsymbol{y}) \approx p(\boldsymbol{f} | \boldsymbol{u}_f) p(\boldsymbol{f} | \boldsymbol{u}_f) q(\boldsymbol{u}_f) q(\boldsymbol{u}_g).$$

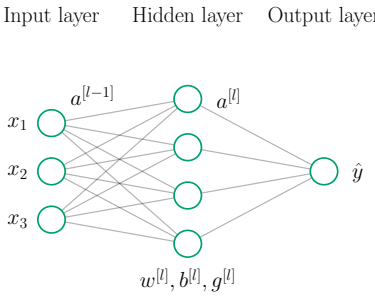


Figure 3.3: A schematic of an ANN with one hidden layer.

Loss function

The optimization problem reduces to finding the distributions $q(\boldsymbol{u}_f)$ and $q(\boldsymbol{u}_g)$ so that the Kullback-Leibler (KL) divergence between the approximate and the true distribution can be minimized. The KL divergence cannot be minimized directly, but one can minimize a function that is equal to it up to a constant using a variational bound for the log marginal likelihood,

$$\log p(\boldsymbol{y}) \geq \int q(\boldsymbol{f})q(\boldsymbol{g}) \log p(\boldsymbol{y}|\boldsymbol{f},\boldsymbol{g}) - KL(q(\boldsymbol{u}_f)||p(\boldsymbol{u}_f)) - KL(q(\boldsymbol{u}_g)||p(\boldsymbol{u}_g)), \tag{3.26}$$

where $q(\boldsymbol{f}) = \int p(\boldsymbol{f}|\boldsymbol{u}_f)q(\boldsymbol{u}_f)\mathrm{d}\boldsymbol{u}_f$ and $q(\boldsymbol{g}) = \int p(\boldsymbol{g}|\boldsymbol{u}_g)q(\boldsymbol{u}_g)\mathrm{d}\boldsymbol{u}_g$. Saul et al. [135] parametrize $q(\boldsymbol{u}_f)$ and $q(\boldsymbol{u}_g)$ as Gaussian distributions with variational parameters $\boldsymbol{\mu}_f, \boldsymbol{S}_f$ and $\boldsymbol{\mu}_g, \boldsymbol{S}_g$, respectively. A Gaussian prior is placed over $\boldsymbol{f}$ and $\boldsymbol{g}$ allowing $q(\boldsymbol{f})$ and $q(\boldsymbol{g})$ to be computed analytically as a function of the variational parameters, which are learned through optimization. The integral in Equation (3.26) factorizes over the data points and is solved analytically for a Gaussian likelihood.

3.3.5 Artificial Neural Networks

An ANN is a commonly used term for a supervised learning algorithm to model a complex nonlinear function between input parameter observations and the corresponding set of output quantities. Neural networks have become the bedrock of progress that machine learning has seen in the last few decades. They consist of input and output layers that sandwich the so-called hidden layers. Each layer is made up of nodes, and each node is connected to several other nodes in the adjacent layers. The nodes consist of weights and biases, which are the unknown parameters of the function. Figure 3.3 shows a schematic of a typical neural network architecture with one hidden layer. The nodes are shown in circles, and the connections between them are highlighted. The first is the input layer $\boldsymbol{x} \in \mathbb{R}^3$ in this example. The activation functions are denoted by g, the weights by w, and biases by b. The output layer is simply $\hat{y} \in \mathbb{R}$ in this case, but can be multi-dimensional.

Model definition

A multi-layer feed-forward neural network utilizing backpropagation is employed. It is a fully connected architecture where the nodes in the hidden layers are connected to all

the nodes in the adjacent layers, but not to other nodes in the same layer. At each hidden layer, a weighted summation is performed on the incoming values from the previous layer and passed through a nonlinear activation function. A commonly used activation function is the rectified linear unit (ReLU), which is defined as,

$$\text{ReLU}(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{otherwise.} \end{cases} \tag{3.27}$$

In a general form, the calculations performed at every node in a forward pass in the l^{th} layer, j^{th} node are:

$$\begin{aligned} z_j^{[l](p)} &= \sum_{i=1}^{m^{[l-1]}} w_{ji}^{[l]} a_i^{[l-1](p)} + b_j^{[l]} \\ a_j^{[l](p)} &= g^{[l]}\left(z_j^{[l](p)}\right) \end{aligned} \tag{3.28}$$

where $a_i^{[l-1](p)}$ is the input from the i^{th} node of the previous layer $(l-1)$ for the p^{th} training sample. $w_{ji}^{[l]}$ is the weight assigned to the connection between the i^{th} node of the previous layer and the j^{th} node of the current layer. b_j is the bias term and $g^{[l]}$ is the activation function. The neural network weights and biases are always assigned a random value to initialize the network, usually between -1 and 1. In matrix form, the same equation is stated as,

$$\begin{aligned} Z^{[l]} &= W^{[l]} A^{[l-1]} + B^{[l]} \\ A^{[l]} &= g^{[l]}(Z^{[l]}) \end{aligned} \tag{3.29}$$

Loss function and network training

The neural network training consists of adjusting the weights and biases such that the error in the prediction of the network output ($\hat{y}$) with respect to the "real" value (y) is minimized. This process, known as training, is essentially an optimization task. It aims to minimize a cost function, which quantifies the error between the predictions and the actual values as

$$C = \frac{1}{n_{train}} \sum_{i=1}^{n_{train}} \mathcal{L}(y^{(i)}, \hat{y}^{(i)}). \tag{3.30}$$

The parameters are adjusted based on the gradient of the cost function with respect to the weights ($\mathrm{d}W$) and biases ($\mathrm{d}B$) using gradient descent or stochastic gradient descent. The updated parameters are given as,

$$W^{[l]} \leftarrow W^{[l]} - \alpha \mathrm{d}W^{[l]} \tag{3.31}$$

$$B^{[l]} \leftarrow B^{[l]} - \alpha \mathrm{d}B^{[l]} \tag{3.32}$$

where α is the learning rate, which is a tuning parameter that affects the rate of convergence of the network. The gradients are derived using the backpropagation algorithm [136]. Several optimization algorithms are available in the literature, including Adam [137], momentum

3.3.6 Mixture density networks

MDN is a probabilistic regression method that combines Gaussian mixture models with ANNs [138].

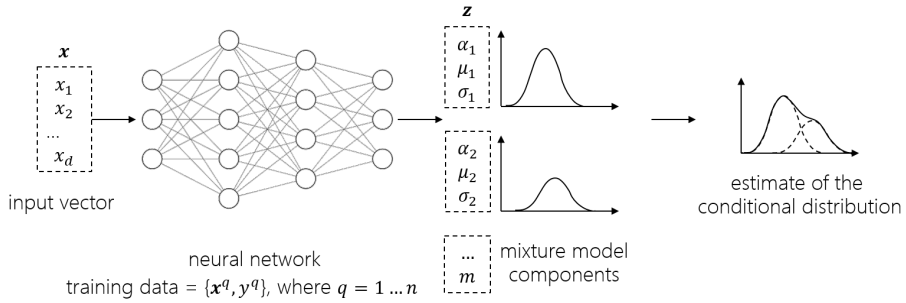


Figure 3.4: Schematic representation of MDNs.

Model definition

The conditional distribution of the target is represented by a mixture of Gaussian distributions,

$$p(y \mid \boldsymbol{x}) = \sum_{i=1}^{m} \alpha_i(\boldsymbol{x}) \mathcal{N}(y \mid \boldsymbol{\mu}_i(\boldsymbol{x}), \sigma_i^2(\boldsymbol{x})), \tag{3.33}$$

where $\alpha_i(\boldsymbol{x})$ are the weights or coefficients assigned to the i^{th} mixture component, and $\mathcal{N}(y \mid \boldsymbol{\mu}_i(\boldsymbol{x}), \sigma_i^2(\boldsymbol{x}))$ is a Gaussian kernel representing the conditional density of the i^{th} component of the target distribution, with parameters $\boldsymbol{\mu}_i(\boldsymbol{x})$ and $\sigma_i(\boldsymbol{x})$.

Instead of mapping the inflow features $\boldsymbol{x}$ to the load statistics y directly, the neural network is trained to predict the parameter vector, $\boldsymbol{z} \in \mathbb{R}$ consisting of α_j, μ_j and σ_j for $j = 1...m$.

The mixing coefficients $\alpha_i(\boldsymbol{x})$ must sum up to exactly 1. A *softmax* function is used to handle this constraint,

$$\alpha_i = \frac{\exp z_i^{\alpha}}{\sum_{j=1}^{m} \exp z_j^{\alpha}}, \tag{3.34}$$

where z_i^{α} are the network outputs predicting the mixture coefficients. Similarly, positive values of the standard deviation are ensured by representing them as exponential functions of the corresponding network outputs, z_i^{σ},

$$\sigma_i = \exp(z_i^{\sigma}). \tag{3.35}$$

Loss function

The likelihood $\mathcal{L}$ of the dataset is given by,

$$\mathcal{L} = \prod_{q=1}^{n} p(y^q \mid \boldsymbol{x}^q) p(\boldsymbol{x}^q) \tag{3.36}$$

A very commonly used error function for probabilistic models is the negative log of the likelihood. From (3.33) and (3.36), the error can be written as,

$$E^q = -\ln\{\sum_{i=1}^{m} \alpha_i(\boldsymbol{x}^q)\mathcal{N}(y \mid \boldsymbol{\mu}_i(\boldsymbol{x}^q), \sigma_i^2(\boldsymbol{x}^q))\}, \tag{3.37}$$

where $p(\boldsymbol{x}^q)$ is not included as it is constant with respect to the parameters or weights. The derivative of the error function is calculated at the output layer and is back-propagated to get its gradient with respect to the network weights. The hidden layers in our network use the rectified linear unit (ReLU):

$$ReLU(x) = \begin{cases} x & \text{for } x > 0 \\ 0 & \text{for } x \leq 0. \end{cases} \tag{3.38}$$

The output layer of the network does not have an activation function; therefore, the outputs are just linear combinations of the inputs from the previous layer. Finally, everything needed to minimize the error function using gradient descent optimization is in place. In this study, the Adam optimizer [137] is used to perform stochastic gradient descent. A 10-fold cross-validation set is performed over 600 samples at every training iteration.

Activation functions

The hidden layers in our network use the rectified linear unit (ReLU), defined as,

$$ReLU(x) = \begin{cases} x & \text{for } x > 0 \\ 0 & \text{for } x \leq 0 \end{cases} \tag{3.39}$$

The output layer of the network does not have an activation function; therefore, the outputs are just linear combinations of the inputs from the previous layer.

Regularization

Minimizing the error function is an ill-posed problem, as it involves a conflict between learning a function that fits the data perfectly and remaining robust under varying training data sets. As the network size grows, the function space increases, and the tendency of the neural network is to overfit. Among several ways to avoid overfitting [139], in this study, a combination of *early-stopping* [140] and *L1* and *L2 regularization* [141] is implemented.

Early stopping

The error function measured on the cross-validation dataset first decreases, then starts increasing as the network begins over-fitting the training data. This can be avoided by applying an early stopping mechanism that stops training as the validation loss stops decreasing over a certain number of iterations. After experimenting with a range of early stopping iterations, 100 was found to be sufficient for the negative log-likelihood on the validation dataset to converge without overfitting. That is, if the validation loss did not show any improvement after 100 iterations, the model training was stopped.

L1 and L2 regularization

L1-regularization penalizes the error function with the sum of the magnitude of the weights,

$$E_R^q = E^q + \lambda \sum |w_i| \tag{3.40}$$

It pushes the coefficients of uninformative features towards zero, effectively pruning the feature space.

Weight-decay or L2-regularization, on the other hand, penalizes the error function with a fraction of the squared magnitude of the weights,

$$E_R^q = E^q + \lambda \sum w_i^2 \tag{3.41}$$

L2-regularization encourages the weights to be small. In both approaches, λ is the *regularization parameter*. It is a hyperparameter that controls the model's complexity, and the optimal value can be determined using a search algorithm. It was found that heavy regularization with $l1$ and $l2$ values of 0.1 deteriorated the results by over-smoothing the conditional response. On the other hand, no $l2$ regularization also resulted in relatively smaller R^2 values for the standard deviation prediction, likely due to some degree of overfitting. On the basis of this hyperparameter study on one channel, a conservative value of $1e-3$ for both $l1$ and $l2$ was chosen.

3.3.7 Conditional generative adversarial networks

Generative Adversarial Networks (GANs) [142] and Conditional Generative Adversarial Networks (CGANs) [143] are types of deep generative models that have gained immense success in the realm of generation of high-dimensional data, including synthetic audio, images, and videos [144–146]. The fundamental assumption in GANs is that the output is a stochastic function of the input. The 'fake' samples generated by the GAN model are realizations of the random variable predicted by the underlying stochastic function. This framework can, in theory, also be applied to probabilistic regression problems, where the CGAN can be trained to predict samples from the conditional response $p(Y|X = x)$.

Model definition

A typical GAN framework consists of two players - generator (G) and discriminator (D)- that are pitted against each other, Figure 3.5. G and D are independent differentiable functions of their inputs and parameters with cost functions $J_G(\theta_G, \theta_D)$ and $J_D(\theta_G, \theta_D)$, respectively. The generator generates new samples using random samples from a noise distribution $\boldsymbol{z} \sim p_z(\boldsymbol{z})$ that's passed through the learned function G. In the conditional variant, the generator and discriminator are also given the feature vector $\boldsymbol{x} \in \mathbb{R}^d$ as input. The goal of the generator is to fool the discriminator into *thinking* that the synthetic samples it generates from $p_g(y \mid \boldsymbol{x})$ are coming from the real data distribution $p_d(y \mid \boldsymbol{x})$. The discriminator, on the other hand, needs to learn to differentiate between the samples from the model distribution and those from the data distribution.

Loss function

Training a GAN model consists of finding the functions G and D through a minmax problem [142]. The minmax formulation for an unconditional GAN is given as,

$$V(G,D) = \mathbb{E}_{y\sim p_d(y)}[\log(D(y))] + \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(\boldsymbol{z})))], \tag{3.42}$$

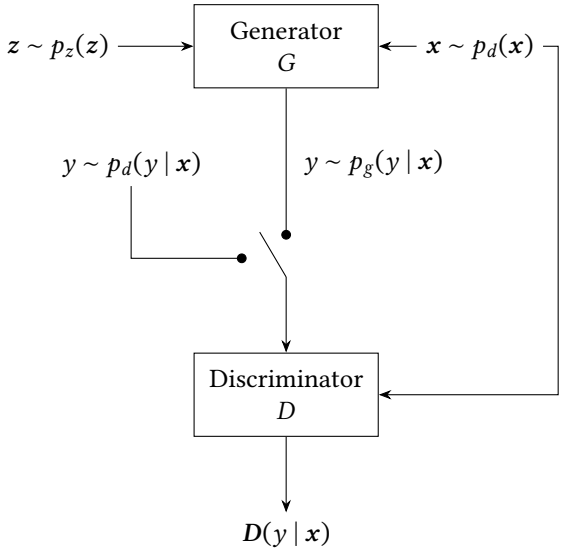


Figure 3.5: Schematic of a standard conditional GAN model architecture.

where V is the objective function. Maximizing the objective with respect to the discriminator means that the discriminator should assign high probabilities to the real samples, y, in the first term of Equation (3.42), and low probabilities to the generator samples, $G(z)$, in the second. Minimizing the objective with respect to the generator, on the other hand, means that the probability assigned by the discriminator to the generated samples in the second term must be as close to 1 as possible. The conditional variant for Equation (3.42) is given as,

$$\min_G \max_D V_c(D, G) \tag{3.43}$$

$$V_c(D, G) = \mathbb{E}_{x,y\sim p_d(x,y)}[\log(D(x, y))] + \mathbb{E}_{z\sim p_z(z),x\sim p_d(x)}[\log(1 - D(G(z, x)))]. \tag{3.44}$$

In practice, the training objective is to minimize the distance between the p_d and p_g such that, after the training is complete, the samples produced by the generator are indistinguishable from the real samples. In the case of low-dimensional probabilistic regression, this translates into inferring the underlying conditional distribution of the response. In the standard formulation, the optimal discriminator for a fixed generator results in minimizing the Jensen-Shannon (JS) divergence between p_g and p_d [142, 147]. Despite having a theoretical optimum for the training objective, determining the actual JS divergence is not possible as it requires knowledge of the probability density functions that are not explicitly available. Additionally, the methods used to approximate the JS divergence are costly and do not provide useful gradients to train the generator [148]. A promising alternative metric, called the Earth Mover's Distance, or Wasserstein-1 distance, has been investigated in the literature to calculate the distance between two probability distributions in the context of training GANs [147–149]. Arjovsky et al. [148] present a

way to approximate the Wasserstein-1 distance using the Kantorovich-Rubenstein duality [150] as,

$$W_1(p_d, p_g) = \sup_{\|f\|_L \leq 1} \mathbb{E}_{y \sim p_d(y)}[f(y)] - \mathbb{E}_{y \sim p_g(y)}[f(y)], \tag{3.45}$$

where f is a set of K-Lipshitz functions that can be approximated using neural networks (G and D). The modified objective for Wasserstein GANs is, thus, given as,

$$\min_G \max_{D \in \boldsymbol{D}} V_{c,W}(D, G) \tag{3.46}$$

$$V_{c,W}(D, G) = \mathbb{E}_{\boldsymbol{x} \sim p_d(\boldsymbol{x})}[D(y \mid \boldsymbol{x})] - \mathbb{E}_{\boldsymbol{z} \sim p_z(z)}[D(G(\boldsymbol{z} \mid \boldsymbol{x}) \mid \boldsymbol{x})]. \tag{3.47}$$

The Lipshitz constraint on D is applied in this model by adding a gradient penalty (GP) to the training objective to constrain the gradient of the discriminator's output with respect to its input, as proposed by Gulrajani et al. [149]. The penalty term is written as,

$$\lambda \mathbb{E}_{\hat{y} \sim p(\hat{y}|\boldsymbol{x})}[(\|\nabla_{\hat{y}} D(\hat{y} \mid \boldsymbol{x})\|_2 - 1)^2]. \tag{3.48}$$

$\hat{y}$ is a linear combination of real and generated samples , $\hat{y} = ty + (1-t)G(\boldsymbol{z}), t \sim \mathcal{U}(0,1)$. λ is another hyperparameter that must be tuned according to the dataset. This form of GANs has been termed WCGAN-GP. Petzka et al. [151] extended the framework from simple weight-clipping to penalizing the violation of the Lipshitz constraint using Lipshitz penalty (LP). The regularization term for WCGAN-LP is given as,

$$\lambda \mathbb{E}_{\hat{y} \sim p(\hat{y}|\boldsymbol{x})} \left[(\max\{0, \|\nabla_{\hat{y}} D(\hat{y} \mid \boldsymbol{x})\|_2) - 1\})^2 \right]. \tag{3.49}$$

The penalty term is simply added to Equation (3.47) as an additional regularization term. In this study, the WCAGN-LP formulation is implemented. The value of the penalty coefficient, λ, is set to 0.002 following the detailed hyperparameter studies on this dataset by Yap [152]. Once the CGAN model is trained, the predictor function is the generator function $G(z|\boldsymbol{x})$.

3.4 Goodness of fit for probabilistic models

3.4.1 Wasserstein distance

The Wasserstein metric is a distance function that compares the difference between the pdfs of two random variables. It is symmetric, non-negative, and satisfies the triangle inequality, making it a proper distance metric. In the case of 1-D distributions, the Wasserstein-2 distance between a reference empirical measure Y and predicted measure $\hat{Y}$, is defined as [153–155],

$$W_2(Y, \hat{Y}) = \left(\int_0^1 |F^{-1}(t) - G^{-1}(t)|^2 dt \right)^{1/2} \tag{3.50}$$

where F^{-1} and G^{-1} are the quantile functions of Y and $\hat{Y}$ respectively. The individual quantile functions are obtained from the samples of the empirical distributions and then integrated.

The distance metric is normalized by the standard deviation of the reference conditional distribution, Y. Therefore, a value of $W_2/\sigma(Y) = 1$ is the distance between a distribution with mean $\mu(Y)$, scale $\sigma(Y)$, and a degenerate distribution with the same mean. The global performance of the model is estimated by averaging the normalized Wasserstein distance over N_{test} test samples as,

$$d_{W2} = \mathbb{E}_{N_{test}}\left(\frac{W_2}{\sigma(Y)}\right) \tag{3.51}$$

3.4.2 Coefficient of determination

The coefficient of determination, also known as the R^2, is a common measure of the goodness of fit of a model. It is defined as,

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}, \tag{3.52}$$

where $\hat{y}_i$ is the predicted output, y_i is the observed value, and $\bar{y}$ is the mean of the observed values. R^2 measures the fraction of the variance in the observed data that is explained by the model predictions. When the dataset contains a significant amount of stochastic variability, R^2 cannot be reliably used to assess or compare point predictions, as it becomes strongly influenced by irreducible noise rather than model quality. In such cases, even a surrogate model that perfectly captures the conditional distribution of the response will not achieve $R^2 = 1$, since the reference data represent random realizations drawn from that distribution. For this reason, in the following chapters, R^2 is not applied to individual realizations of the response, but instead used to compare the predicted and reference conditional mean (R^2_μ) and conditional standard deviation (R^2_σ).

Let μ_i and σ_i denote the reference conditional mean and standard deviation of the response for input x_i, estimated from multiple realizations, and let $\hat{\mu}_i$ and $\hat{\sigma}_i$ denote the corresponding quantities predicted by the surrogate model. The coefficient of determination for the conditional mean, R^2_μ, and the conditional standard deviation R^2_σ are defined as

$$R^2_\mu = 1 - \frac{\sum_{i=1}^{N_{test}}(\mu_i - \hat{\mu}_i)^2}{\sum_{i=1}^{N_{test}}(\mu_i - \bar{\mu})^2}, \qquad \bar{\mu} = \frac{1}{N_{test}}\sum_{i=1}^{N_{test}}\mu_i, \tag{3.53}$$

$$R^2_\sigma = 1 - \frac{\sum_{i=1}^{N_{test}}(\sigma_i - \hat{\sigma}_i)^2}{\sum_{i=1}^{N_{test}}(\sigma_i - \bar{\sigma})^2}, \qquad \bar{\sigma} = \frac{1}{N_{test}}\sum_{i=1}^{N_{test}}\sigma_i. \tag{3.54}$$

where $\bar{\mu} = \frac{1}{N_{test}}\sum_{i=1}^{N_{test}}\mu_i$ is the mean of the reference conditional means and $\bar{\sigma} = \frac{1}{N_{test}}\sum_{i=1}^{N_{test}}\sigma_i$ is the mean of the reference conditional standard deviations.

4

Surrogate modeling for fixed bottom wind turbine loads

In this chapter, novel probabilistic methods, including Chained Gaussian processes, mixture density networks, and conditional generative adversarial networks, are proposed for mapping 10-minute average site conditions to corresponding load statistics on fixed-bottom wind turbines. Widely used deterministic models, such as polynomial regression, artificial neural networks, and Gaussian process regression, are also compared. The probabilistic approach enables the modeling of uncertainty in loads as a response to stochastic inflow conditions. The data-driven model is trained on the OpenFAST simulations of the IEA-10MW-RWT. The various surrogates are compared in terms of accuracy, as well as practical aspects such as ease of use, model sensitivity to changes in the dataset, and sensitivity to the hyperparameters. It is shown that mixture density networks are reliable and robust options for probabilistic surrogate modeling of loads. They can recover the accurate conditional mean response in all load channels with values for the coefficient of determination (R^2) greater than 0.95 on the test dataset.

This chapter is partly based on:

D. Singh et al. Probabilistic surrogate modeling of offshore wind-turbine loads with chained Gaussian processes, IOP Journal of Physics Conference Series [125].

D. Singh et al. Probabilistic surrogate modeling of damage equivalent loads on onshore and offshore wind turbines using mixture density networks, Wind Energy Science [85].

4.1 Introduction and objective

The chapter introduces a probabilistic surrogate modeling approach to map 10-minute site condition statistics to the corresponding 10-minute load statistics on the IEA-10MW-RWT reference wind turbine [156]. The process of generating training and validation data differs from the standard surrogate modeling approaches. To capture variation in the loads, the training data were constructed without any seed repetitions. The validation dataset, on the other hand, involves seed repetitions in order to characterize the conditional pdf of the load at any point. The chapter presents information on the OpenFAST setup used for generating the data, the parameters used in the surrogate models, and the predictions.

4.2 Numerical setup

4.2.1 OpenFAST modeling approach

The surrogate is modeled on the responses of an aero-hydro-servo-elastic code, OpenFAST, which serves as the ground truth in this study. It is a state-of-the-art, multi-physics numerical tool for modeling wind turbines. It combines analytical and empirical formulations with conservative assumptions to simplify the code and limit the computational cost. It can model environmental conditions like stochastic waves, currents, and a frozen wind turbulence field with randomized coherent turbulent structures superimposed on the random, homogeneous background turbulence.

OpenFAST is used to set up a numerical model of the real-world environment and system dynamics to produce training data for the surrogate. All simulations are performed on the IEA-10MW-RWT [156] offshore reference wind turbine. This wind turbine was chosen as it was one of the latest reference wind turbines available for OpenFAST. Additionally, WindEurope reported the average size of offshore wind turbines to be 8.2MW in 2020, with newer turbines averaging 8.7MW [157]. 10 MW was thought to be representative of the near-future trends in offshore wind turbines.

TurbSim [93] simulations for inflow turbulence generation are performed with the Kaimal turbulence spectrum[94], a grid resolution of 40 points in a $1.16D \times 1.16D$ domain, D being the rotor diameter. The spatial coherence is dictated by the IEC coherence function in the streamwise direction and no coherence in the plane orthogonal to the streamwise direction. The gradient Richardson number is set to 0. The total simulation duration is 900s, out of which the first 300s are discarded to exclude the initial transient. Ten-minute simulations on their own are insufficient for estimating fatigue loads. However, the statistical variations in loads that one would expect over longer periods or multiple seed repetitions can be potentially inferred indirectly via probabilistic surrogates based on the variation in the quantities of interest at neighboring training samples. Therefore, 10-minute statistics are sufficient for a complete description of the load response as long as unique turbulence and wave seeds are used for each training simulation.

The loads are calculated with the ElastoDyn module in OpenFAST, which uses the Euler-Bernoulli beam theory to calculate the bending moments by assuming the structure

is straight and isotropic. The ServoDyn module is used to control the wind turbine. The controller settings differ from those used in the DTU Wind Energy controller for the HAWC2 simulations of the IEA-10MW-RWT reference document [156]. The main difference appears at low wind speeds, where the rotor RPM is not restricted to 6, and the collective blade pitch remains zero until the rated wind speed is reached. The OpenFAST simulation output with these controller settings at low wind speeds may interfere with the tower's natural frequencies; however, modification of the controller is beyond the scope of this study. The simulations are performed with single precision to limit file size and simulation times without any significant impact on the accuracy of the loads.

The implementation of the IEA-10MW-RWT in OpenFAST is relatively new. As such, there are ongoing updates to the public model based on user feedback. This study uses the IEA-10MW-RWT as the ground truth for the surrogate modeling study. For this purpose, it does not need to be perfectly accurate but should be representative of the expected load response class. The machine learning methodology is expected to be easily transferable to different wind turbine types.

4.2.2 Definition and sampling of the input features

The IEA-10MW-RWT is designed for offshore conditions. However, in this study, it is simulated on both onshore (CASE-ONSHORE) and offshore (CASE-OFFSHORE) sites to evaluate the additional training requirements in the offshore case against an onshore reference.

CASE-ONSHORE

Aero-servo-elastic: The wind turbine is subjected only to aerodynamic loading. Wind speed, power-law exponent, and turbulence intensity are selected as the input parameters for the aerodynamic simulations as they have been shown to have the highest impact on the load response in previous studies [45]. The variable bounds are also the same as the ones defined in [45], listed in Table 4.1. The power-law exponent and turbulence intensity are functions of the wind speed. R and z are the rotor radius and the hub height, respectively. The samples are drawn from a three-dimensional Sobol sequence [158] to ensure an even spread of points in the sample space. The random seed for turbulence generation is not included as a training variable. Each sample is, therefore, associated with a unique random seed, there are no repetitions. The simulation time step used in OpenFAST is 0.01s.

CASE-OFFSHORE

Aero-servo-hydro-elastic: The offshore wind turbine is placed on a monopile foundation at 30m water depth and is subject to both aerodynamic and hydrodynamic loading. Along with the aerodynamic parameters of CASE-ONSHORE, there are additional wave parameters in this case as listed in Table 4.1. For designing load surrogates suitable for multiple sites, ideally, the $H_s - T_p$ diagrams from several sites should be combined to define conservative ranges for the two variables. In this study, for example, the values from a joint $H_s - T_p$ kernel from a representative distribution are sampled. Additionally, the minimum and maximum range of H_s may also be defined as a function of wind speed in order to sample from the joint $u - H_s - T_p$ distribution. The first-order waves are modeled using the JONSWAP spectrum in HydroDyn. The values of the aerodynamic variables are

the same as in CASE-ONSHORE. In particular, the Turbsim solution files are therefore shared between CASE-ONSHORE and CASE-OFFSHORE. Similar to CASE-ONSHORE, the wave and turbulence seeds are not included in training the model. Each sample is, therefore, associated with a unique set of random seeds; there are no repetitions. The simulation time step used in OpenFAST is 0.002s. This is much smaller than the time step needed to capture the relevant frequencies. However, larger time steps resulted in numerical instability in certain cases.

Table 4.1: List of input variables and the corresponding variable bounds for ON and OFF.

Case	Variable	Lower Bound	Upper Bound	Sampling
ON, OFF	Wind Speed (u) [ms^{-1}]	4	25	Uniform, Sobol
ON, OFF	Turbulence Intensity (ti) [%]	2.5	$\frac{18}{u}(6.8+0.75u+3(\frac{10}{u})^2)$	Uniform, Sobol
ON, OFF	Power Law Exponent (α) [−]	$0.15-0.23(\frac{u_{max}}{u})(1-(0.4\log\frac{R}{z})^2)$	$0.22+0.4(\frac{R}{z})(\frac{u_{max}}{u})$	Uniform, Sobol
OFF	Significant Wave Height (H_s) [m]	0	6	Kernel density estimate, pseudo-random
OFF	Spectral Peak Period (T_p) [s]	1	21	Kernel density estimate, pseudo-random
OFF	Wave Direction (W_{dir}) [deg]	-180	180	Uniform, pseudo-random
ON, OFF	Turbulence Random Seed [−]	−50000	50000	Uniform, pseudo-random
OFF	Wave Random Seed 1 [−]	−50000	50000	Uniform, pseudo-random
OFF	Wave Random Seed 2 [−]	−50000	50000	Uniform, pseudo-random

4.2.3 Responses

The quantities of interest are the 10-minute load statistics at the tower top and tower bottom in the fore-aft direction and at the blade root in the edgewise and flapwise directions. The fore-aft bending moment of the tower is typically the primary contributor to tower fatigue and overall support-structure dynamics. Although the side-side moment is not aerodynamically damped, it remains smaller in magnitude than the fore-aft direction. The methodology presented here can be easily extended to other tower load channels. Blade-root flapwise and edgewise moments are the standard channels used to characterize blade fatigue. The statistics include the 10-minute standard deviation of the moment's signal and the 10-minute damage equivalent load, M_{eq}. The modeled load channels are

listed with their notation in Table 4.2. The model weights for each channel are calibrated separately.

Table 4.2: List of response channels modeled by the data-driven surrogate models.

Response channel	10-minute statistics	Notation
Tower bottom fore-aft moment	stddev, M_{eq}	TwrBsMyt
Tower top fore-aft moment	stddev, M_{eq}	YawBrMyn
Blade root edgewise moment	stddev, M_{eq}	RootMxb1
Blade root flapwise moment	stddev, M_{eq}	RootMyb1

Since the direction of the incoming flow is always aligned with the rotor, the fore-aft direction at the tower base is defined in the local coordinate system of the inflow wind. The time series are saved at 100Hz and 500Hz for onshore and offshore cases, respectively. All the statistics except for M_{eq} are calculated using the original high-frequency signal. M_{eq}, however, is computed using a down-sampled load signal at 1Hz. Given the large scale of the database, this choice was made to save time, memory, and space during preprocessing. While this lower sampling frequency captures the dominant contributors to the fatigue, including the 1P, 3P, and lower-frequency hydrodynamic oscillations, it limits the contribution of the higher-frequency harmonics, resulting in overall lower M_{eq} values compared to those obtained from higher-frequency load signals. Expert feedback suggests that a sampling frequency in the range of 10–40 Hz would be more suitable for capturing the full spectrum of fatigue-relevant dynamics. This frequency range is, therefore, recommended for future studies while training the model. Nonetheless, further analysis confirms that the surrogate model's performance is not affected by the minor discrepancies in the M_{eq} values arising from downsampling. The process of calculating M_{eq} is presented in Equation (2.5), where the stress ranges are replaced by bending moment ranges calculated using the Rainflow counting algorithm. The Wöhler coefficient takes the values 3.5 for the tower, 10 for blade flapwise, and 8 for blade edgewise moments.

4.2.4 Test datasets

The prediction accuracy of the conditional pdf by MDNs is tested on two independently sampled test datasets, which have not been used in the training procedure, referred to as TEST1 and TEST2. TEST1 consists of 50 pseudo-randomly-selected points spanning the entire sampling domain, with parameter bounds the same as in Table 4.1. At each test point, engineering simulations with OpenFAST are repeated 300 times to get a reference pdf by keeping the input features constant but changing the turbulence and wave random seeds, resulting in a total of 15000 TurbSim and OpenFAST simulations. In TEST2, only the wind speed and turbulence intensity are altered, while keeping the other inflow parameters constant. TEST2 consists of 24 cases, each of which is repeated similarly with 300 turbulence and wave seeds. The purpose of building a separate TEST2 dataset was to visualize the model's performance through two-dimensional heat maps and analyze the integrated quantities derived from TEST1. The values of the parameters are listed in Table 4.3.

The test and training points for CASE-OFFSHORE are shown in Figure 4.1. The marginal distributions are shown for the training dataset. The test points for CASE-

Table 4.3: Variables and their corresponding values in TEST2 dataset. Δ indicates the discretization step.

Variable parameters	[Min : Max : Δ]
Wind Speed (u) [ms^{-1}]	[6 : 21 : 3]
Turbulence Intensity (ti) [%]	$\begin{cases} [10 : 40 : 10] & \text{for } u = 6 \\ [6 : 24 : 6] & \text{for } u = 9 \\ [4 : 16 : 4] & \text{for } u > 10 \end{cases}$
Fixed parameters	**Value**
Power Law Exponent (α) [−]	0.08
Significant Wave Height (H_s) [m]	1
Spectral Peak Period (T_p) [s]	7
Wave Direction (W_{dir}) [deg]	0
Random seeds	**Value**
Turbulence Random Seed [−]	$\mathcal{U}(-50000, 50000)$
Wave Random Seed 1 [−]	$\mathcal{U}(-50000, 50000)$
Wave Random Seed 2 [−]	$\mathcal{U}(-50000, 50000)$

ONSHORE are identical, but only for the turbulence inflow features, namely, u, ti, and α.

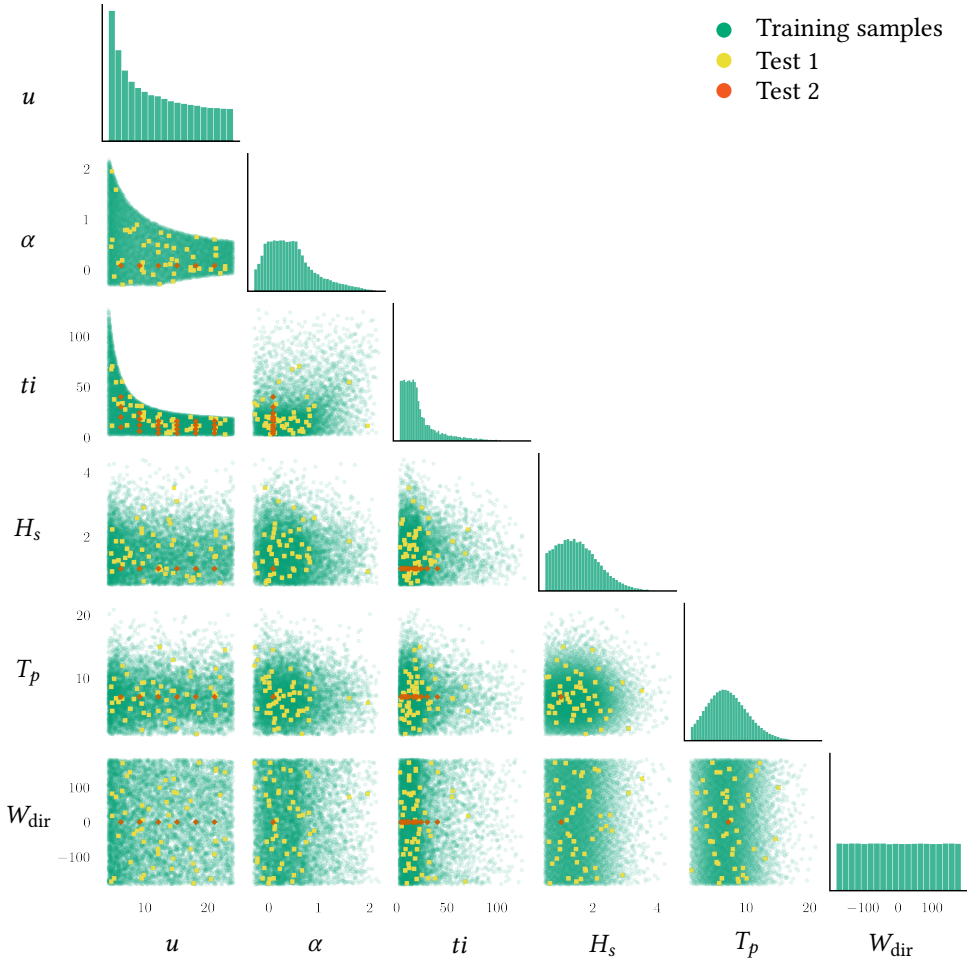


Figure 4.1: Pairplot of the input features for CASE-OFFSHORE showing the training samples along with the test datasets. The marginal distribution for the training dataset is shown along the diagonal.

4.3 Data-driven modeling setup

4.3.1 Probabilistic machine learning approach

Figure 4.2 shows the basic workflow used for model calibration and load estimation. All data generated from OpenFAST is included in the training base without any repetition or a pre-filtering step involved.

The validation approach in probabilistic modeling differs from the standard models, as the reference conditional distribution needs to be obtained from the simulations. This requires repeating the simulations with multiple seeds. Here, 300 random seeds are chosen at each validation point, which should be sufficient to capture the standard deviation and

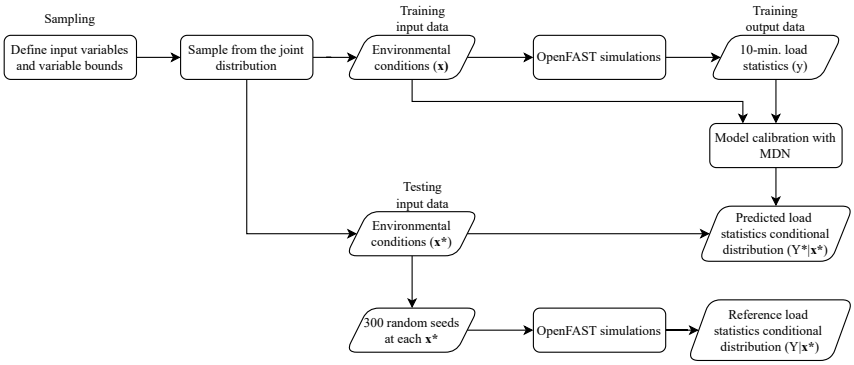


Figure 4.2: Schematic of the machine learning approach.

the tails of the distribution.

4.3.2 Surrogate modeling setup

The theoretical background of the surrogate models is presented in Section 3.3. Here, the specific details of the hyperparameters are presented. The data are always scaled before training using the standard scaler.

Polynomial regression

second and fourth-degree polynomial regression is tested on the dataset. Even though polynomial regression is a simplistic approach, it provides a reference for the performance of the rest, more complex models. The models are trained on 250, 500, 1500, and 2500 samples in the convergence study.

Gaussian process regression

The GPR model is trained on 250, 500, 1500, and 2500 samples to assess the convergence of the conditional mean. Training the GPR model is fairly straightforward, as there are only a limited number of hyperparameters to tune, namely, the initial length scale ($\boldsymbol{l}$). For all cases, the length is scaled per feature using automatic relevance determination [159]. The initial value is assigned to be 0.8 for every feature.

Chained Gaussian process regression

To ensure that positive quantities like the standard deviation of the loads or the fatigue loads always return a positive prediction from the surrogate model by assuming them to be log-normally distributed. This also required the data to be free from outliers, for which a Z-score filter with a threshold of 3σ is used for outlier detection. The CGP model has an additional hyperparameter known as the number of inducing points. In Section 4.4.1, the model is trained with the combinations of training samples and inducing points listed in Table 4.4. The inducing points are calculated as centers of K-means clusters of the training subset. The additional steps required for the CGP model are shown in the flowchart in Figure 4.3.

Table 4.4: The number of training samples and inducing points considered for the model's convergence study.

Training samples	Inducing points
250	250
500	250
1500	250, 1000
2500	250, 1000
3500	250, 1000

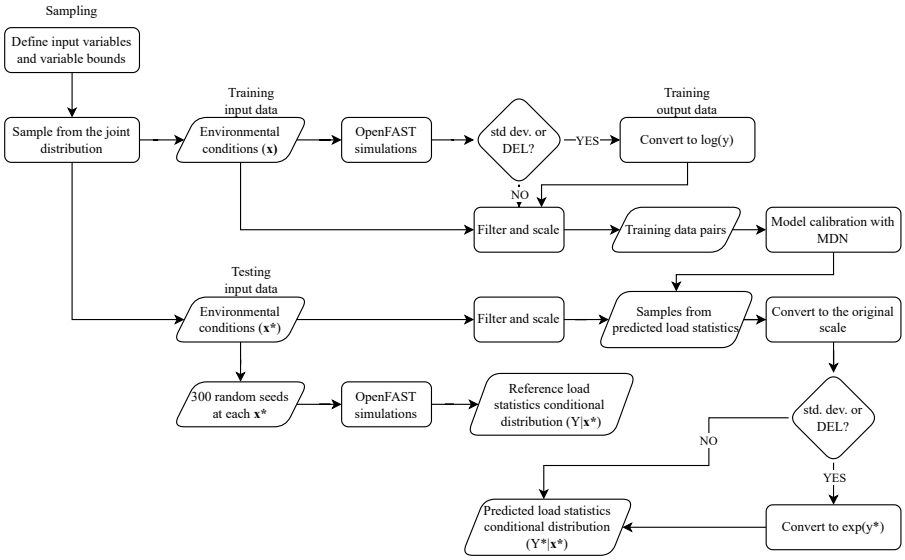


Figure 4.3: The machine learning approach for the chained Gaussian process regression model.

The kernels are separately and independently assigned to the latent functions $f(x)$ and $g(x)$. For both, the squared exponential kernel described in Equation (3.14) is used. The initial length scales are assigned to be 1 for each feature. Automatic relevance determination [159] is used to calculate the length scales per feature independently.

Artificial neural network

The hyperparameters of the ANN models introduced in Chapter 3 are listed in Table 4.5. CASE-ONSHORE is trained using a two-hidden-layer ANN, whereas CASE-OFFSHORE employs three hidden layers, as this configuration was found to yield improved predictive performance.

Mixture Density Networks

The primary hyperparameters used in this study to train the models to obtain the results are summarized in Table 4.6. In Section 4.4.1, the performance of the MDN model is evaluated with two two-layer network architectures.

Table 4.5: Summary of the ANN hyperparameters.

Network hyperparameter	Value
Hidden layers	2 (ON) and 3 (OFF)
Network width {layer1, layer2, layer 3}	{20, 20, ∅ } and {20, 20, 10}
Activation function (hidden layers)	ReLU
Activation function (output layer)	None
Learning rate	0.001
Maximum epochs	1000
Mini-batch size	200
Optimizer	Adam
Regularization	
λ for $L2-$regularization	10^{-3}
Early-stopping	
Early-stopping tolerance	10^{-4}

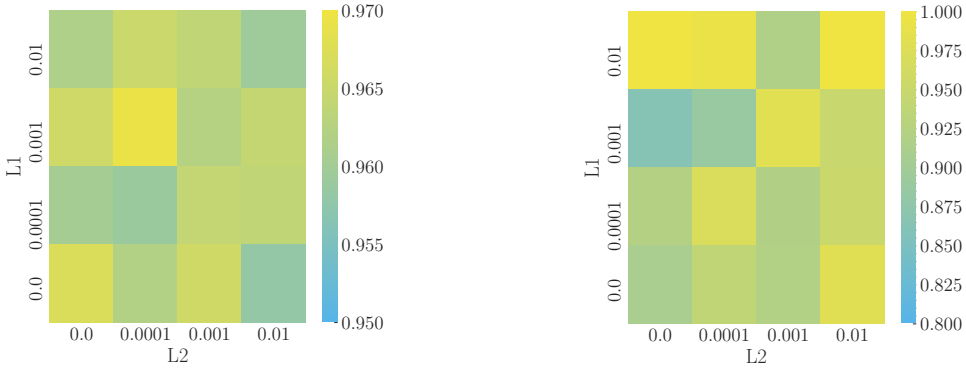
Table 4.6: Summary of the MDN network hyperparameters.

Network hyperparameter	Value
Number of mixture components	4
Hidden layers	2
Network width {layer1, layer2}	{10, 10} and {30, 30}
Activation function (hidden layers)	ReLU
Activation function (output layer)	None
Learning rate	0.005
Maximum epochs	1000
Mini-batch size	100
Optimizer	Adam
Regularization	
λ for $L1-$regularization	10^{-3}
λ for $L2-$regularization	10^{-3}
Early-stopping	
Early-stopping patience	100
Early-stopping monitor	validation loss
Number of early-stopping validation samples	600

A sensitivity analysis on the regularization parameters is shown in Figure 4.4. The conditional mean is easier to learn and less sensitive to regularization. However, the accuracy of the conditional distribution fit reduces if the $L1$ parameter is increased to $1e-2$. $L1$ regularization encourages the weights of the network to go to zero. A lower accuracy with a higher $L1$ means that the model is underfitting the data. A balance is found with $L1$

and $L2$ values of $1e-3$, which is used for the rest of the study.

A similar study is performed on a two-layer network with varying widths of the hidden layers in Figure 4.5. The scale of the plots shown in Figure 4.5 is quite narrow, supporting the robustness of the model. Predictions with large networks (>50) and small ones (<10) show a slight reduction in the prediction accuracy of the mean and the conditional distribution for a training sample size of 5000 points. The larger networks, containing more than 3000 parameters, would perform better with a bigger training dataset. The smaller networks are not sufficiently complex to capture the right behavior. The study is repeated on a 3-layer network with similar conclusions, showing that the model predictions are not sensitive to the depth, indicating that the two-layer network is already sufficient to capture the complexity.



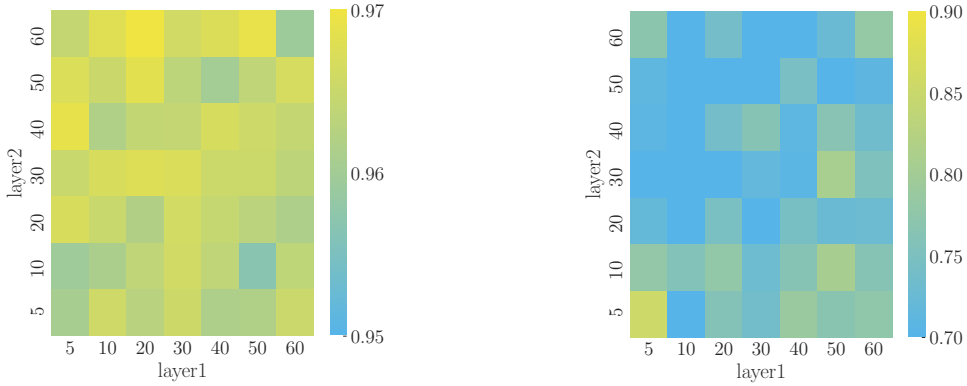
(a) R^2_μ results

(b) d_{W2} results

Figure 4.4: CASE-OFFSHORE: Comparison of accuracy metrics for various regularization parameter values with MDN (30, 20) trained on 5000 samples, tower bottom fore-aft standard deviation.

CGAN training

An extensive study on this dataset was performed by Yap et al. [152]. The Wasserstein-CGAN model (WCGAN) is used in Section 4.4 with the hyperparameters listed in Table 4.7.



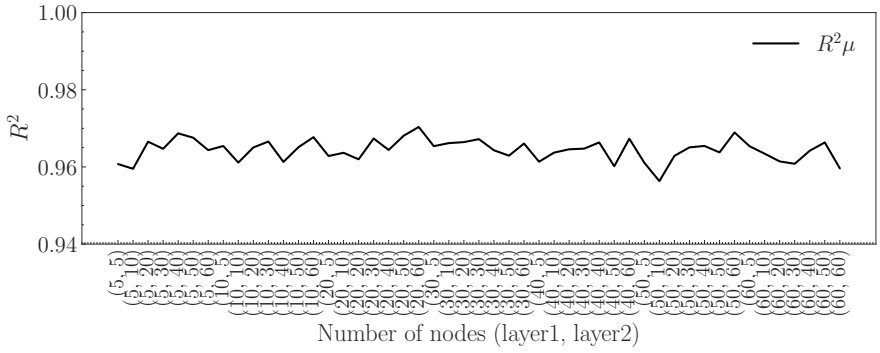
(a) R^2_μ results

(b) d_{W2} results

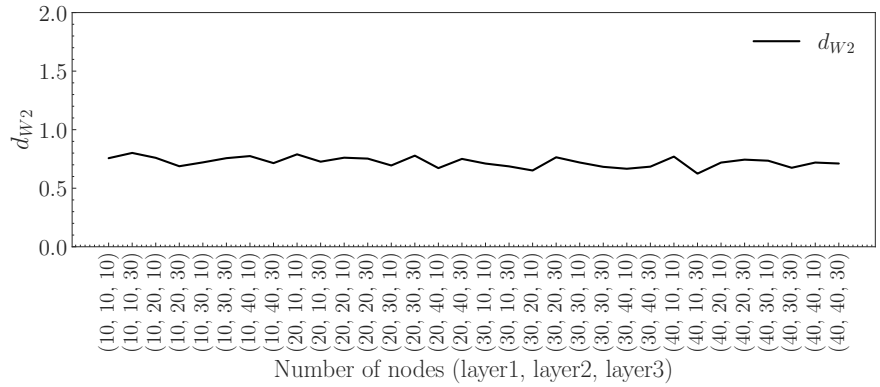
Figure 4.5: CASE-OFFSHORE: Comparison of accuracy metrics for various combinations of layer widths in a two-layer network with MDN trained on 5000 samples, tower bottom fore-aft standard deviation, $L1 = 0.001$, $L2 = 0.001$.

Table 4.7: Summary of the WCGAN network hyperparameters.

Network hyperparameter	Value
Epochs	30000
Critic learning rate	0.0001
Generator learning rate	0.0002
Batch size	200
Generator optimiser	RMSProp
Critic optimizer	RMSProp
Activation function (hidden layers)	ReLU
Gradient penalty coefficient	0.02



(a) R^2_μ results



(b) d_{W2} results

Figure 4.6: CASE-OFFSHORE: Comparison of accuracy metrics for a few combinations of layer widths in a 3-layer network with MDN trained on 5000 samples, tower bottom fore-aft standard deviation, $L1 = 0.001$, $L2 = 0.001$.

4.4 Results

In this section, the accuracy of the surrogate models in predicting the conditional load distribution on the TEST1 and TEST2 datasets, as mentioned in Section 4.2.4, is assessed. The first part focuses on convergence studies, specifically examining the impact of training data size on predicting the average 10-minute standard deviation of the tower bottom fore-aft moment. The goal is to measure how the accuracy of the predictions varies based on the hyperparameters and initialization of the optimization algorithm. Once the network architecture and training sample size are fixed, a rigorous analysis of the model's performance is performed.

4.4.1 Convergence

Generally speaking, more data translates to better accuracy. However, an increase in data after a certain point gives diminishing returns in accuracy. Given the computational cost of generating the training database, ensuring a good model fit with as little training data as possible is crucial. The models listed in Table 4.8 are trained on a range of training samples to determine the rate of increase of the accuracy of the model as a function of the number of training samples. For both CASE-ONSHORE and CASE-OFFSHORE, the convergence study shown here is performed on the tower bottom fore-aft moment standard deviation (TwrBsMyt [kN-m] stddev). In principle, the convergence study could be performed on any or all of the load channels mentioned in Section 4.2.3. However, TwrBsMyt [kN-m] stddev was chosen because it is found to have relatively the highest average d_{W2} value amongst all channels. Models are data-dependent, and each channel would ideally need a separate sensitivity study in terms of the number of training samples needed for convergence, the model architecture, and the model hyperparameters. However, for practical purposes, it is assumed that if a model can perform well on the channel with the poorest metrics, then it should also perform well, if not optimally, on the other load channels with the same network architecture. Care is taken to prevent overfitting by comparing the evolution of the negative log-likelihood on the training and test datasets during training.

Table 4.8: Description of the models used in the convergence study.

Name	Description
MDN(10,10)	MDN with two hidden layers of 10 nodes each
MDN(30,30)	MDN with two hidden layers of 30 nodes each
CGP M=250	CGP with 250 inducing points
CGP M=1000	CGP with 1000 inducing points
Polynomial degree 2	Second-degree polynomial fit
Polynomial degree 4	Fourth-degree polynomial fit
ANN (20,20)	ANN with two hidden layers of 20 nodes each
ANN (20,20,10)	ANN with three hidden layers of 20, 20 and 10 nodes

The convergence of the models with respect to the number of training samples is investigated. For the models MDN(10, 10), MDN(30, 30), an error bar is also plotted to

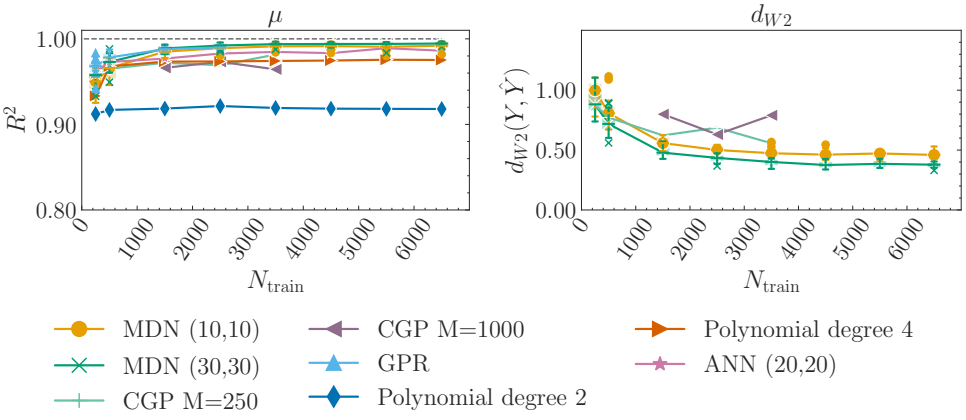


Figure 4.7: Convergence of the R^2 value of the conditional mean and the d_{W2} of the conditional distribution of the predicted quantities as a function of the number of samples for CASE-ONSHORE.

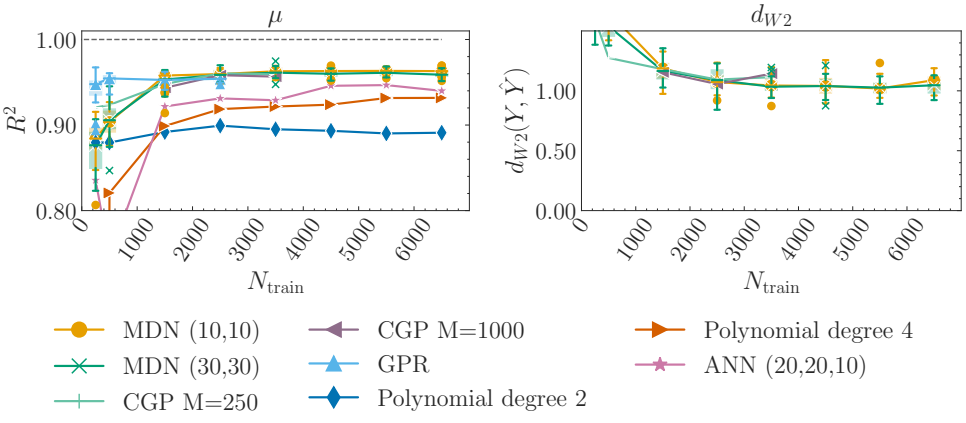


Figure 4.8: Convergence of the R^2 value of the conditional mean and the d_{W2} of the conditional distribution of the predicted quantities as a function of the number of samples for CASE-OFFSHORE.

highlight the deviation of the predictions as a function of the choice of the training data subset. At every N_{train}, the model training is repeated on 25 uniquely sampled subsets of the data with 10-fold cross-validation. The GPR model predictions are similarly repeated with 25 different samples of the training data pool at every N_{train} value. The variance in the predictions of the rest of the models due to differences in the choice of training dataset is not included for the sake of clarity. The plots Figure 4.7 and 4.8 show the convergence of the model in terms of predicting the R^2 value of the conditional mean and the normalized 2-Wasserstein distance. The metrics are averaged over the validation dataset, which is formed by combining TEST1 and TEST2. The boxes in Figure 4.7 and Figure 4.8 extend between the first ($Q1$) and third ($Q3$) quartiles of the data and the horizontal line across the box indicates the median. The difference between $Q1$ and $Q3$ defines the interquartile range (IQR). The upper whisker extends to the largest data values that are within 1.5IQR above $Q3$. The lower whisker, similarly, extends to the lowest data point within 1.5IQR below $Q1$. Outliers are visible as dots beyond the whisker boundaries. The MDN and GPR models are trained on 25 different subsets at each N_{train}. Some of the key observations of the convergence study are listed as follows:

- Comparison of the datasets: The performance of all evaluated models was consistently better on the onshore wind turbine dataset compared to the offshore counterpart. This difference is mainly due to the reduced complexity of the onshore dataset—fewer influencing features result in simpler underlying functions, thereby improving model predictability. In terms of the convergence, CASE-OFFSHORE benefits from 1000+ training samples, whereas in CASE-ONSHORE, the model predictions are approaching near-perfect estimates of R^2 with as few as 500 samples. GPR shows marginal but continuous improvements in the R^2 value with the number of samples in CASE-OFFSHORE, contrary to CASE-ONSHORE. One common reason for such behavior is the presence of redundant features combined with more noise, which ultimately limits training. It can be observed in the values of the optimized values of length scales derived from ARD, as seen in Figure 4.9. The extremely large length scales of the hydrodynamic parameters indicate that they are not relevant to the model training. Additionally, as shown in Figure 4.10, the log marginal likelihood function is not well-defined in the hydrodynamic parameters and plateaus at relatively large length scales. These figures were generated by varying the length scale of one dimension at a time while keeping the others constant. This investigation suggests that, in the case of fixed-bottom offshore wind turbines, fatigue even at the tower bottom is largely driven by the wind speed-related statistics.

- Comparison of the models: One of the key advantages of GPR is its ability to achieve the same level of accuracy as MDN while requiring fewer training samples. This is noticeable with both datasets in the R^2 values at 500 samples. The main drawback of GPR is that it has a matrix computation cost of $\mathcal{O}(N_{\text{train}}^3)$ [114], making it difficult to scale to larger datasets, if needed. To address scalability, a sparse implementation using CGP was employed, which reduces the complexity to $\mathcal{O}(N_{\text{train}}M^2)$ [160]. Still, convergence studies for CGP were also limited due to computational costs.

  The number of induction points in CGP does not seem to improve the prediction accuracy. In fact, in the case of CGP M=1000, performance even deteriorates with

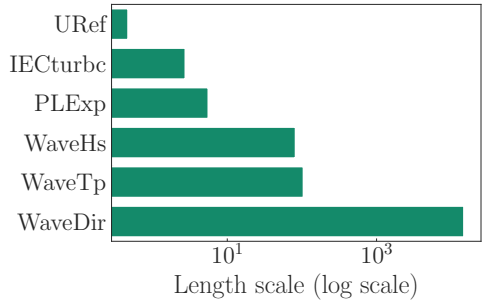


Figure 4.9: Calculated length scales per feature using ARD in the GPR model trained with 1500 training samples on tower bottom fore-aft moment standard deviation (CASE-OFFSHORE).

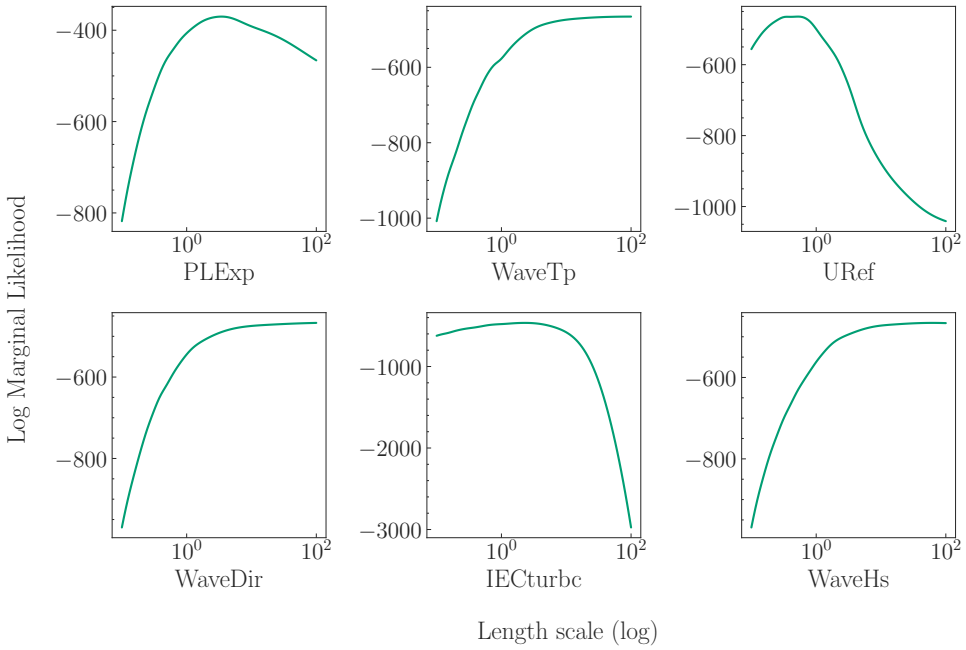


Figure 4.10: Log marginal likelihood plotted as a function of the individual length scales for each input feature for the GPR model trained with 1500 training samples on tower bottom fore-aft moment standard deviation.

increasing values of N_{train}. This behavior could be a result of the saturation of the accuracy already with 250 induction samples. Another possible hypothesis is that the k-means clustering algorithm is not the best choice for choosing the induction points. An alternative could be to randomly sample a subset or use the low-discrepancy sequences already generated for training to select a quasi-random subset. Finally, the solution may be converging to a local minimum due to the sensitivity to the initialization of the training or the learning rate.

The fourth-degree polynomial function, although simple, is able to capture the con-

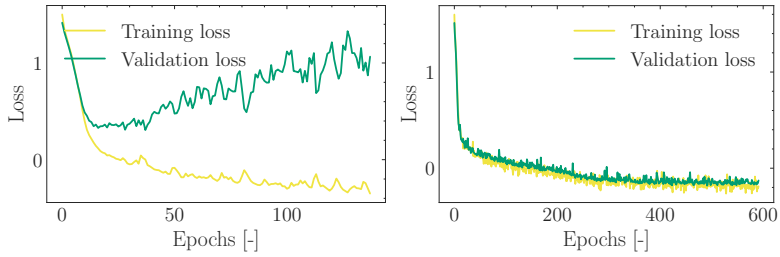
ditional mean with good accuracy in the two cases. The disadvantage of polynomial functions is that they are parametric. However, since they are much quicker to train than most other approaches tested here, they are easy to experiment with and should certainly be considered for simple setups where only the prediction of the conditional mean is interesting.

Detailed hyperparameter studies were not performed on the ANN model. Additional regularization and choice of a different architecture may converge to better estimates. Still, the ANN model predicts the conditional mean on average with R^2 values greater than 0.95 with a sufficiently large training dataset, indicating a good fit.

In both CASE-ONSHORE and CASE-OFFSHORE, the difference between the predictions of the two MDN architectures $(10, 10)$ and $(30, 30)$ is mainly seen in the d_{W2} estimates of the onshore turbine. An improvement of roughly 15% in terms of d_{W2} is seen in CASE-ONSHORE, whereas a negligible difference is observed in CASE-OFFSHORE. As the data pool is finite, the uncertainty bounds associated with the choice of training samples progressively decrease as the total available samples are approached. At smaller N_{train} values, the uncertainty is also driven by the missing information in the training data and the choice of the initial conditions used by the stochastic gradient descent optimizer. Significantly better GPR estimates of the response mean can be achieved for less than 1500 training samples, which can be attributed to the Bayesian formulation. Beyond that point, however, MDNs and GPR are comparable, with $R^2 > 0.99$ in CASE-ONSHORE and $R^2 > 0.95$ in CASE-OFFSHORE.

- Conditional distribution: The convergence of conditional distributions was evaluated using the Wasserstein distance metric (d_{W2}) for heteroscedastic models, including MDN and CGP. Results showed that d_{W2} values stabilized after approximately 3500 training samples. Beyond this point, additional training data did not significantly improve the quality of the predicted distributions. d_{W2} is 50% smaller in CASE-ONSHORE than in CASE-OFFSHORE, signifying, overall, a much better inference of the latent pdf in the onshore conditions than offshore. This is expected, as the larger the degrees of freedom, the larger the space between the neighboring points. The goodness of fit of the model in terms of d_{W2} depends on the rate of change of the statistics of the underlying distributions. The gradient of the standard deviation, in particular, is much harder to infer from a cloud of points than the gradient of the mean. The convergence plots for CASE-OFFSHORE could indicate either that this is the best estimate of d_{W2} achievable, given the information the model has, or that the number of samples needed to capture the variation in the standard deviation is even larger than the largest training sample size that was tested.

Figure 4.11 shows the training and validation losses plotted against the number of epochs for CASE-ONSHORE. MDN overfits the data at $N_{\text{train}} = 500$ because, as the training loss decreases, the validation loss increases, indicating that the model cannot handle previously unseen data. Figure 4.11b is well-fitted as the training and validation losses decrease at the same rate. The plot also shows the auto-stop algorithm at work, which halts training after 100 epochs of approximately zero-gradient loss to avoid overfitting.



(a) $N_{\text{train}} = 500$ (b) $N_{\text{train}} = 4500$

Figure 4.11: Training and validation losses for MDN (10, 10) with 500 (a) and 4500 (b) training samples, CASE-ONSHORE.

For the remainder of this study, each model is trained on the number of samples that best offered a subjective trade-off between the accuracy and computational effort of training and generating the dataset. The final N_{train} values are listed in Table 4.9.

4.4.2 Load prediction

In this section, the prediction of the 10-minute damage equivalent loads on the wind turbine for CASE-ONSHORE and CASE-OFFSHORE is presented. R^2 values reflect the proportion of variance in the output that is predictable from the surrogate models. The absolute magnitude of R^2 is sensitive to the optimization initialization, choice of the test samples, and the choice of the training subset, as seen in Figure 4.8. It is, therefore, important to note that the absolute R^2 values do not carry much objective meaning on their own. They are only used here for comparing the performance of models relative to one another.

CASE-ONSHORE

First, the performance of the models on CASE-ONSHORE is discussed. The different surrogate models are tested, and their performance is compared in Table 4.9, 4.10, and 4.11.

Table 4.9: Comparison of the R^2_μ metric for various load channels' 10-minute damage equivalent loads in CASE-ONSHORE.

Model	N_{train}	TwrBsMyt	YawBrMyn	RootMyb1	RootMxb1
MDN(10, 10)	4500	**0.99**	**0.99**	**0.99**	0.91
GPR	500	0.94	0.92	**0.99**	0.86
CGP M=250	3500	**0.99**	0.98	**0.99**	0.67
Polynomial 2	2500	0.96	0.97	0.88	0.7
Polynomial 4	2500	0.97	**0.99**	0.98	0.9
ANN (20,20)	4500	0.98	0.98	0.98	0.93
CGAN	5969	**0.99**	**0.99**	**0.99**	**0.96**

Discussion on the comparison of the model performance on CASE-ONSHORE:

Table 4.10: Comparison of the R^2_σ metric for various load channels' 10-minute standard deviation in CASE-ONSHORE.

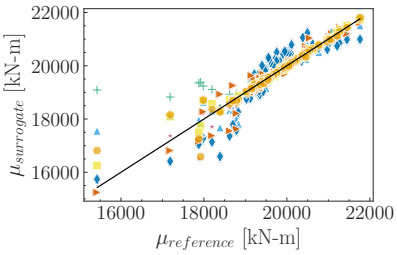
Model	N_{train}	TwrBsMyt	YawBrMyn	RootMyb1	RootMxb1
MDN(10, 10)	4500	**0.94**	**0.96**	**0.91**	0.86
CGP M=250	3500	0.7	0.13	0.35	-
CGAN	5969	0.9	0.94	**0.91**	**0.91**

Table 4.11: Comparison of the d_{W2} metric for various load channels' 10-minute standard deviation in CASE-ONSHORE.

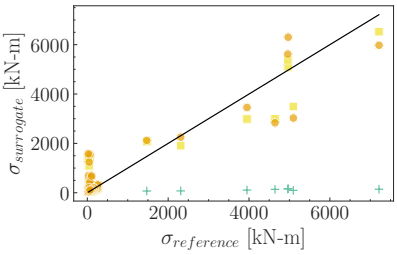
Model	N_{train}	TwrBsMyt	YawBrMyn	RootMyb1	RootMxb1
MDN(10, 10)	4500	**0.44**	0.65	0.76	2.8
CGP M=250	3500	0.62	0.88	0.9	**0.5**
CGAN	5969	0.51	**0.45**	**0.51**	2.4

- CGAN and MDN performance: CGAN and MDN consistently outperform the other surrogates, in terms of conditional mean, standard deviation, and the normalized Wasserstein distance values. The primary drawback of CGANs is the extensive hyperparameter tuning requirements and complex model architecture..
- CGP performance: Although CGP was trained using a different process, as explained in Section 4.3.2, all error metrics are calculated on the same test dataset as for the other models. Given this setup difference, larger deviations from the OpenFAST reference in CGP predictions are expected.
- Polynomial performance: The simpler, fourth-degree polynomial model outperforms some of the more complex surrogates in predicting the conditional mean. This suggests that the underlying dataset may not be as complex as initially assumed.
- GPR robustness: GPR is able to make good predictions of the conditional mean with a small training set. For this kind of application, it can be regarded as a robust and reliable approach when the dataset is limited or training must be performed on a limited computational expense.
- Blade root edgewise moment DEL: Across all models, the blade root edgewise moment shows notably poorer predictive performance. It is primarily a result of the presence of on/off multiple modes in the response at high wind speeds. It is further investigated in the following sections by analyzing the individual conditional distributions in Figure 4.16c. Figure 4.12 shows the predicted conditional mean and standard deviation plots for the blade root edgewise moment DEL, against the conditional reference from OpenFAST. The points where a disagreement exists between the various surrogates in Figure 4.12a correspond to the points where the multi-modal response exists. Figure 4.12b shows the $\sigma_{\text{surrogate}}$ values. It only shows a small number of points because the multi-model response results in a very large standard deviation of the loads compared to normal operation. On zooming in to the normal range of loads, the performance of the other models can be compared. CGP was not trained on

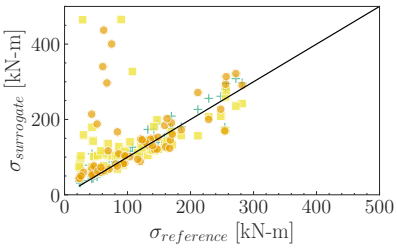
the 'outliers', or the points where the wind turbine was curtailed. Therefore, it is unable to correctly extrapolate the standard deviation in Figure 4.12b, but can do it accurately in the region where it was trained, Figure 4.12c.



(a) Conditional mean



(b) Conditional standard deviation



(c) Conditional standard deviation (zoomed in)

Figure 4.12: CASE-ONSHORE: Conditional mean (top left) and conditional standard deviation (top right) surrogate predictions vs. OpenFAST reference for blade root edgewise moment DEL. Zoomed-in image of the conditional standard deviation on the bottom.

The conditional DELs for the tower base fore-aft moment (Figure 4.13), tower top fore-aft moment (Figure 4.14), blade root flapwise moment (Figure 4.15), and blade root edgewise moment (Figure 4.16) are plotted at three operational conditions falling in low, medium, and high wind speed blocks. The values of the input features are noted in Table 4.12.

Table 4.12: Values of the input features for the pdfs in Figure 4.13 to 4.16.

Wind condition	u [ms^{-1}]	ti [-]	α [-]
Low wind speed	6	40	0.08
Medium wind speed	12	16	0.08
High wind speed	22.70	17.28	0.07

The conditional pdfs provide a closer look at the nuances between the various surrogate model performances. There are a few interesting things to note here:

- High wind speed: At high wind speeds, there were some test and training simulations when the wind turbine was partly curtailed due to high fluctuations in wind speeds, resulting in a bimodal response. This can be seen in Figure 4.13c, 4.14c, 4.15c, and 4.16c. The bimodality in the load response is primarily attributed to the behavior of the wind turbine controller rather than to inflow-related flow physics. Modern wind turbine controllers enforce hysteresis to ensure a smooth transition between operating regimes [161, 162]. This is to avoid continuous switching between wind turbine states when the wind speed is at the border of two operating regions (for instance, below cut-out and above cut-out wind speeds). However, the transition in this controller appears to be more abrupt. Studies in the future should investigate the exact cause of this behavior and determine whether the controller needs to be updated. However, it is beyond the scope of this study. It should further be noted that while mean loads during state switching scale approximately linearly with the fraction of time the turbine is operating or parked, the corresponding DELs exhibit strongly nonlinear behavior. The magnitude of the DEL is dominated by the largest load ranges. Hence, transient events during state changes govern the fatigue response. Consequently, the sequence of operating states is relevant, and it matters whether the turbine undergoes a restart or a shutdown within the averaging window. All surrogate models in this study, except MDN and CGAN, don't assume a multi-modal response. Hence, the predicted mean is shifted with respect to the OpenFAST reference. An alternative approach is to preprocess the training data to actively remove outliers before fitting the surrogates. Otherwise, the dominant mode of the bimodal response can be considered for lifetime fatigue calculations.

- Low wind speed: The probabilistic and deterministic surrogates agree well on the conditional mean across all channels in this wind speed range. The underlying distribution appears to be Gaussian or lognormal.

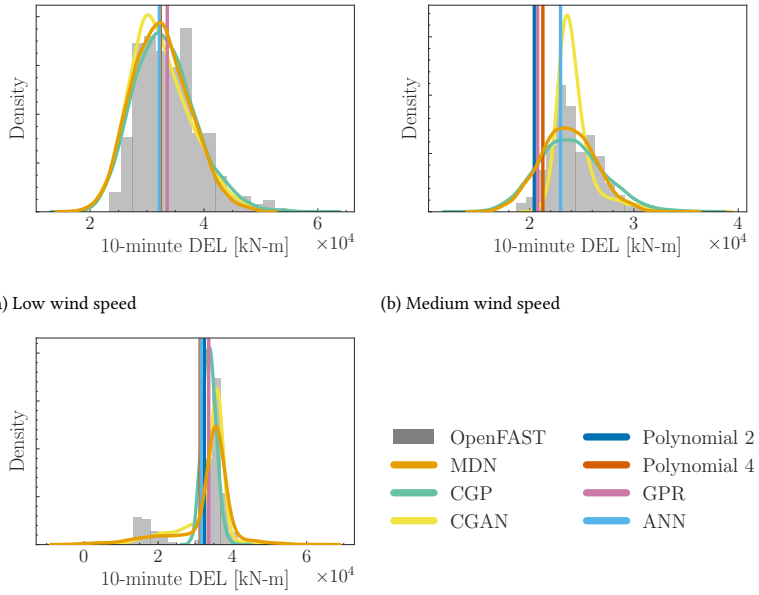
- Medium wind speed: This regime consists of more complex, nonlinear behavior in control systems and structural dynamics. The blades are highly loaded with pitch changes to regulate the power production. The deterministic model predictions are generally in good agreement with the reference, but compared to the low wind speed region, there is a wider spread of predictions. The probabilistic models capture the underlying distribution closely.

- CGAN performance: When examining the individual pdf, CGAN demonstrates qualitatively strong agreement with the reference distribution.

- Non-physical response: The probabilistic surrogates can predict distributions with tails that extend to negative values. DEL is a strictly positive quantity, but there is no mechanism in the training process to prevent the models from predicting negative samples. Limiting the realizations to positive values could potentially be a solution.

- Blade root edgewise moment DEL: The edgewise blade root moment has the most drastic difference in the load values at high wind speeds as a result of intermittent curtailment. Therefore, there is a larger disagreement with the predictions from the reference.

- Low wind speed resonance: It was later discovered that the controller on the reference wind turbine introduced a resonance in the structure at below-rated speeds. Therefore, the magnitude of DEL is at times larger than the other regions. This has been reported, and the controller has since been updated in the public repository of the 10-MW RWT.

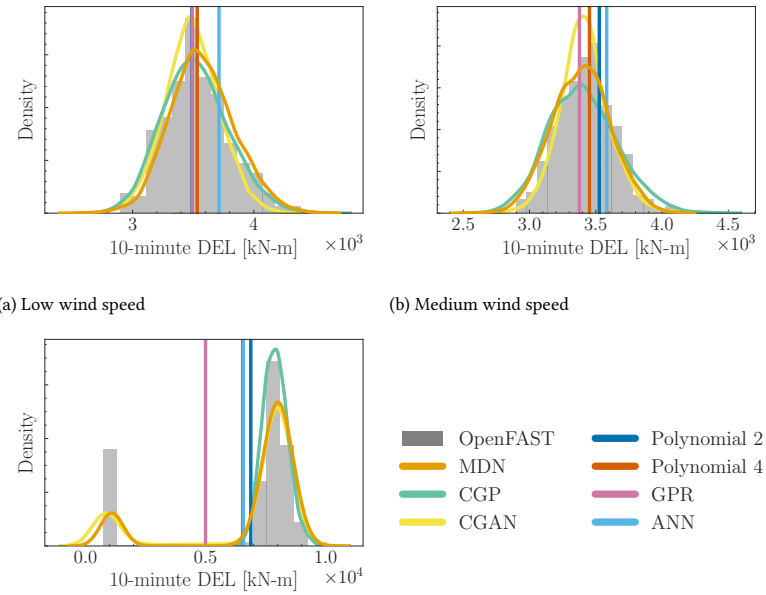


(a) Low wind speed

(b) Medium wind speed

(c) High wind speed

Figure 4.13: CASE-ONSHORE: Predicted and reference (OpenFAST) conditional pdf for the tower base fore-aft moment DEL.

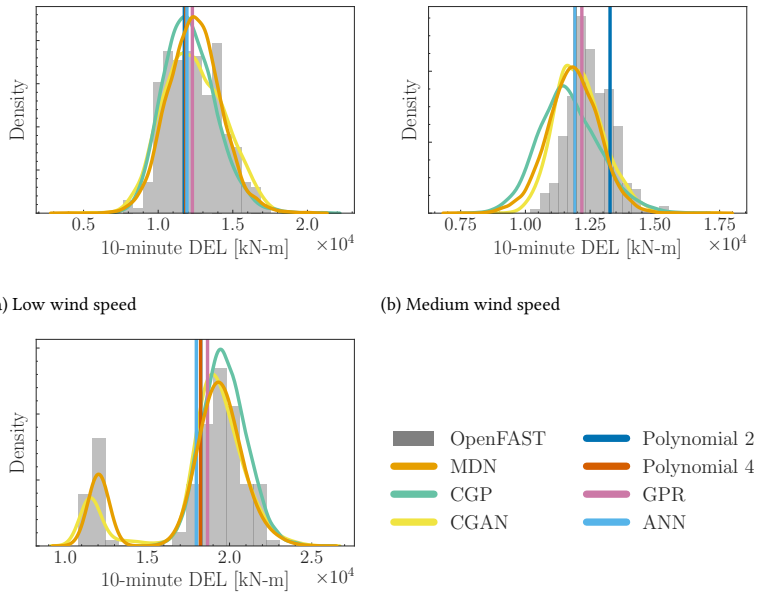


(a) Low wind speed

(b) Medium wind speed

(c) High wind speed

Figure 4.14: CASE-ONSHORE: Predicted and reference (OpenFAST) conditional pdf for the tower top fore-aft moment DEL.

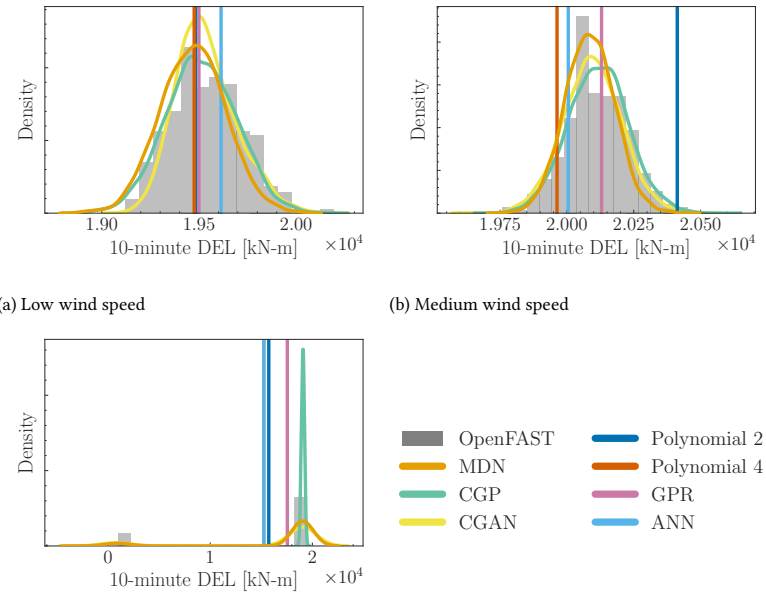


(a) Low wind speed

(b) Medium wind speed

(c) High wind speed

Figure 4.15: CASE-ONSHORE: Predicted and reference (OpenFAST) conditional pdf for the blade root flapwise moment DEL.



(a) Low wind speed

(b) Medium wind speed

(c) High wind speed

Figure 4.16: CASE-ONSHORE: Predicted and reference (OpenFAST) conditional pdf for the blade root edgewise moment DEL.

CASE-OFFSHORE

The different surrogate models are compared similarly for CASE-OFFSHORE in this section. The differences in training and prediction in this case, compared to the onshore counterpart, are simultaneously discussed. Firstly, the performance metrics of various models are compared in Table 4.13, 4.14, and 4.15. The standard deviation and Wasserstein distances are only measured for the probabilistic models. The R^2 values that are below zero are marked with a dash (-).

Table 4.13: Comparison of the R^2_μ metric for various load channels' 10-minute damage equivalent loads in CASE-OFFSHORE.

Model	N_{train}	TwrBsMyt	YawBrMyn	RootMyb1	RootMxb1
MDN(10, 10)	4500	**0.95**	**0.94**	**0.87**	0.91
GPR	500	0.94	0.89	0.86	0.48
CGP M=250	3500	0.92	0.90	0.85	**0.94**
Polynomial 2	2500	0.89	0.92	0.79	0.55
Polynomial 4	2500	0.84	0.84	0.8	-
ANN (20,20,10)	4500	0.89	0.9	0.86	0.6
CGAN	5969	0.9	0.91	**0.87**	0.93

Table 4.14: Comparison of the R^2_σ metric for various load channels' 10-minute standard deviation in CASE-OFFSHORE.

Model	N_{train}	TwrBsMyt	YawBrMyn	RootMyb1	RootMxb1
MDN(10, 10)	4500	**0.78**	**0.94**	**0.79**	0.85
CGP M=250	3500	0.2	-	0.35	0.64
CGAN	5969	0.7	0.83	**0.79**	**0.91**

Table 4.15: Comparison of the d_{W2} metric for various load channels' 10-minute standard deviation in CASE-OFFSHORE.

Model	N_{train}	TwrBsMyt	YawBrMyn	RootMyb1	RootMxb1
MDN(10, 10)	4500	**1.06**	**0.56**	**1.4**	2.6
CGP M=250	3500	1.29	0.7	2	**0.8**
CGAN	5969	1.43	0.6	**1.4**	2

The main observations are summarized here:

- Conditional mean compared to CASE-ONSHORE: Compared to CASE-ONSHORE, the conditional mean metrics are lower across all models. It has been shown in Section 4.4.1 for the tower bottom fore-aft moment that adding training samples only marginally improves the accuracy of the model. It is still possible that manual tuning per channel can improve the overall performance. However, time is often a limitation in an industrial context, and the time needed to perform even automated sensitivity studies can override the benefit of using a surrogate. The result of the lower R^2_μ values can be explained by the presence of higher dimensions and more

noise in the training data due to the presence of a new stochastic source (waves), as discussed in Section 4.4.1.

- MDN and CGAN performance: MDN is consistent in predicting the conditional mean with the overall highest accuracy. Neither MDN nor CGANs were specifically tuned for the offshore case, but as shown here, the performance of CGAN drops more drastically as the dataset is changed for the same architecture. MDN also predicts the conditional standard deviation with a relatively good accuracy as seen in Table 4.14 and 4.15.

- Polynomial predictions: Interestingly, the fourth-degree polynomial is found to be very sensitive to the number of training samples in the case of tower base and top fore-aft moment DEL. Its prediction is even worse than the second-degree polynomial case. This indicates overfitting to the noise and poor generalization.

- GPR predictions: Similar to CASE-ONSHORE, the predictions of the conditional mean with the GPR model on this dataset, even with a small training sample, are comparable to MDN.

- Blade root edgewise moment DEL: Similar to CASE-ONSHORE, the bad performance of the models on this channel is attributed to the presence of multiple modes in the conditional response during training and testing. This has been discussed in the previous section. The discrepancy in the $R^2\sigma$ and d_{W2} performance is because the Wasserstein distance integrates over the whole probability distribution and inherently weights differences by the associated probability mass of each value, which is not the case for standard deviation calculation.

The individual conditional distributions are shown in Figure 4.17, 4.18, 4.19, and 4.20 for selected points in the low, medium, and high wind speed regimes. The input vectors for the three cases are listed in Table 4.16.

Table 4.16: Values of the input features for the pdfs in Figure 4.17 to 4.20.

Wind condition	u [ms^{-1}]	ti [-]	α [-]	H_s [m]	T_p [s]	$wdir$ [deg]
Low wind speed	6	40	0.08	1.0	7.0	0
Medium wind speed	12	16	0.08	1.0	7.0	0
High wind speed	22.70	17.28	0.07	1.9	14.47	-48.57

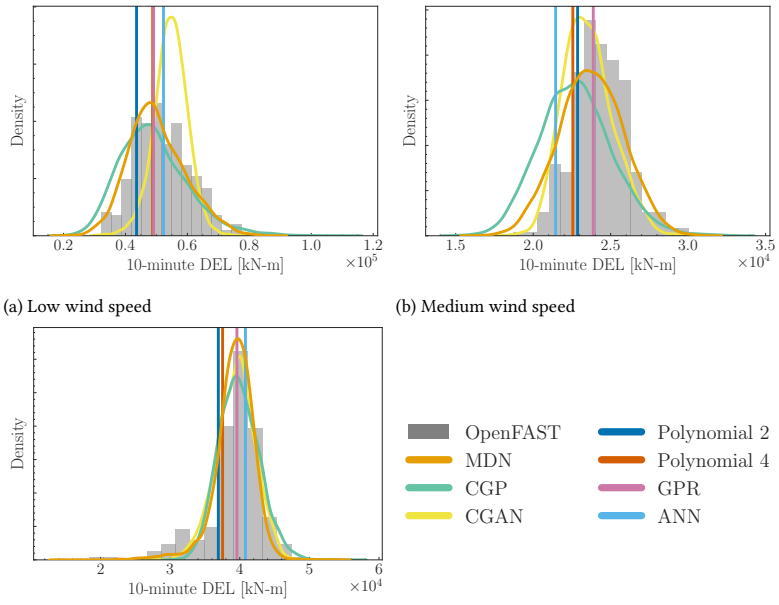
Some of the key observations on the conditional pdfs in CASE-OFFSHORE are listed below:

- Disagreement between the models Compared to CASE-ONSHORE, there is a lot more disagreement between the deterministic models in predicting the conditional mean across all channels. As a result of added features and noise due to some of the features being non-informative, there is a tendency for the models to overfit or underfit the data. Andrianakis and Challenor [163] describe two types of solutions in Gaussian

process modeling, in particular. However, the observation is applicable to any highly nonlinear surrogate. The first solution is where the observation noise is assumed to be very low, causing the surrogate to fit through all the training data. This results in overfitting as the model tries to capture all observed variation, including the noise. The second type of solution is when the model assumes high noise and is unable to model the true underlying function. It treats the actual patterns or structure in the data as if they were just random noise.

- CGAN mode separation inconsistency: It can be seen in Figure 4.19c, that if the two modes are close to each other, CGAN does not assign a zero probability to the space between them. One hypothesis for this is that because Wasserstein distance measures the cost of moving probability mass from the generator distribution to the real distribution, it doesn't directly force zero probability in low-density regions. It only tries to minimize the total transport cost. The failure to detect a separation between the modes could also be a result of not reaching a convergence before the training was stopped.

  That means the generator can still put some probability mass between two modes if doing so slightly reduces transport cost, even if the true distribution has none there

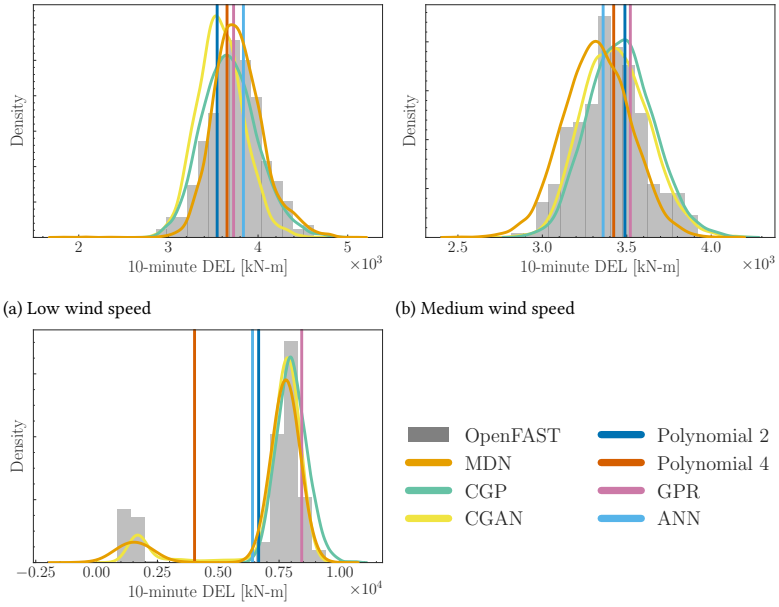


(a) Low wind speed

(b) Medium wind speed

(c) High wind speed

Figure 4.17: CASE-OFFSHORE: Predicted and reference (OpenFAST) conditional pdf for the tower base fore-aft moment DEL.

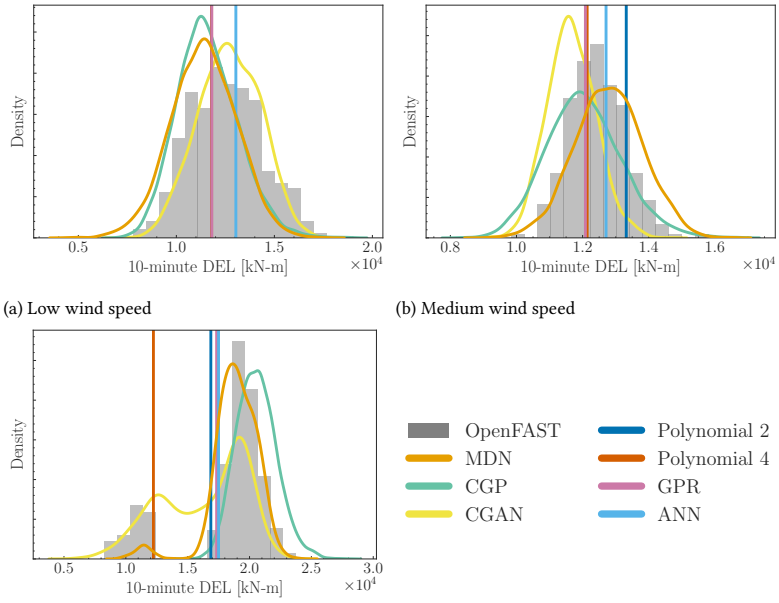


(a) Low wind speed

(b) Medium wind speed

(c) High wind speed

Figure 4.18: CASE-OFFSHORE: Predicted and reference (OpenFAST) conditional pdf for the tower top fore-aft moment DEL.

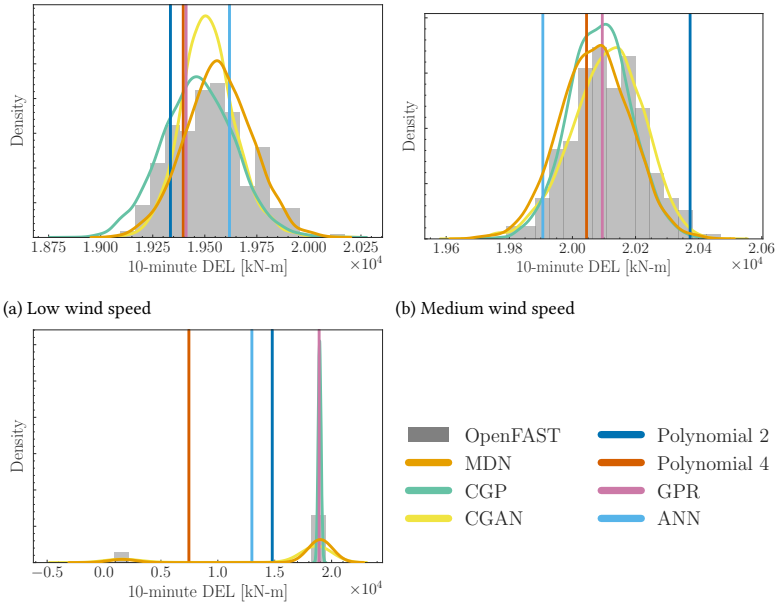


(a) Low wind speed

(b) Medium wind speed

(c) High wind speed

Figure 4.19: CASE-OFFSHORE: Predicted and reference (OpenFAST) conditional pdf for the blade root flapwise moment DEL.



(a) Low wind speed

(b) Medium wind speed

(c) High wind speed

Figure 4.20: CASE-OFFSHORE: Predicted and reference (OpenFAST) conditional pdf for the blade root edgewise moment DEL.

4.5 Summary and discussion

This chapter presents a comparison of several widely used surrogate models, including polynomials, ANN, and GPR. Additionally, it introduces a probabilistic approach based on MDNs, CGP, and conditional Wasserstein GANs to make efficient and flexible load surrogates for offshore siting. The inputs constitute the variables characterizing the site, and the output is the 10-minute damage equivalent load on the wind turbine blades and tower. The data-driven surrogate uses aero-servo-hydro-elastic OpenFAST simulations of the 10-MW reference wind turbine for training. The goal was to explore the hypothesis of whether including information about uncertainty in site conditions improves the surrogate predictions.

The surrogates are trained on two cases. The first is a wind turbine subject to aerodynamic (CASE-ONSHORE) loads onshore, and the second is its offshore counterpart with aero-hydrodynamic (CASE-OFFSHORE) loading. In order to reduce training costs, no seed repetitions were performed during training. This means that the variation in the load response due to the stochastic inflow conditions is treated as aleatoric uncertainty. The noise, if not properly accounted for while describing the model, can result in model sensitivity to the number of training samples and model complexity. The presence of noise forces highly nonlinear models into two types of solutions [163]. The first is when the noise is treated as data (overfitting), and the second is when the noise masks the underlying function (underfitting). The probabilistic models produce resulting conditional load distributions that reflect this aleatoric uncertainty, combined with the epistemic uncertainty arising due to the choice of model parameters.

In this work, the performance of the deterministic model is evaluated based on the quality of its prediction of the conditional mean. The probabilistic surrogates are evaluated on their predictions of the conditional standard deviation and the 2-Wasserstein distance in addition to the mean. The reference conditional distributions are generated in OpenFAST by varying the wind and wave random seeds 300 times per test sample.

Convergence studies show that in CASE-ONSHORE, several surrogate models achieve $R^2_\mu \geq 0.99$ for the conditional mean of the tower bottom fore–aft moment with approximately 500–1000 training samples, whereas in CASE-OFFSHORE, R^2_μ increases only marginally beyond approximately 1000 training samples. For probabilistic metrics, the normalized 2-Wasserstein distance d_{W2} stabilizes after approximately 3000–3500 training samples in both cases, with no significant improvement observed beyond this range.

The performance of the surrogate models is consistently better in the onshore case. The additional hydrodynamic parameters in the offshore case were found not to be as informative in predicting the fatigue loads, resulting in more noise in the data. Combined with a larger number of features, the surrogate model performance drops in comparison. Across load channels and surrogate models, the conditional mean coefficient of determination R^2_μ is reduced by approximately 0.05–0.10 compared to CASE-ONSHORE, while the normalized Wasserstein distance d_{W2} increases by approximately 50%–100%, depending on the channel. These results indicate that increasing the number of training samples alone does not fully compensate for the additional noise introduced by hydrodynamic loading.

Several deterministic and probabilistic models were evaluated. The simplest

polynomial surrogate of the fourth-degree showed good agreement with the reference in CASE-ONSHORE, achieving R^2_μ values above 0.97 for most load channels. In CASE-OFFSHORE, the polynomial fit was susceptible to overfitting, with R^2_μ values dropping below 0.90. GPR showed consistently good performance even with very limited training data, achieving $R^2_\mu \approx 0.94$–0.99 with approximately 500 training samples. The main drawback of GPR is poor scalability; however, since performance improvements beyond approximately 1000 training samples are marginal, as observed in the convergence plots, the need for scaling is limited in this application. ANNs demonstrate good agreement with the reference, typically achieving $R^2_\mu > 0.95$ with sufficiently large training datasets, but are prone to over- and underfitting and require hyperparameter tuning in each case. The process of training the CGP involved removing outliers and transforming the training data from a lognormal to a normal distribution, which improved model accuracy and enabled the capture of dominant modes in the conditional distribution, with d_{W2} values of approximately 0.6–1.3, depending on the load channel. MDNs are robust to changes in training sample size and model parameters, maintaining $R^2_\mu > 0.99$ in CASE-ONSHORE and $R^2_\mu \approx 0.90$–0.95 in CASE-OFFSHORE, with corresponding d_{W2} values of approximately 0.4–0.8 and 1.0–1.4, respectively. Conditional GANs are found to be the most complex models to train for regression in this study. Nevertheless, the results are comparable to MDN in CASE-ONSHORE, and the conditional response is well captured.

In conclusion, incorporating uncertainty in the surrogate modeling process is effective in improving the quality of the results. More so in complex cases, as it helps mitigate the adverse effect of noise in the dataset. With the added benefit of not requiring seed repetitions prior to training, it is shown that all probabilistic approaches cut down significantly on the computational cost associated with generating the training database without compromising prediction accuracy. Further modeling of floating wind turbines in Chapter 6 is done using MDNs.

5

Data analysis of the TetraSpar demonstrator loads

The complex interaction between the structural behavior of the floating offshore wind turbine and the stochastic site conditions is an active area of research. Characterizing the relationship between the environmental conditions and loads may help design reduced-order models, surrogate models, and physics-based engineering models for floating wind turbines. This study uses data from the TetraSpar Demonstrator, which is equipped with a 3.6 MW Siemens Gamesa wind turbine. One-to-one simulations performed using aero-servo-hydro-elastic software are included for comparison. Various tools, including linear correlation, MI, and feature ordering using conditional independence, are used for the assessment. This study is also helpful in validating the engineering model for future global sensitivity analysis using elementary effects or Sobol indices that require a rigid sampling structure of features and can, therefore, only be calculated with simulation tools. A good agreement is found between the measurements and the simulations. The 10-minute damage equivalent loads on the tower show a correlation, particularly with the wind speed statistics and the significant wave height.

This chapter is partly based on:
D. Singh et al. Data analysis of the TetraSpar demonstrator measurements [43].

5.1 Motivation and background

FOWTs offer vast potential to extract energy from deep offshore regions with stronger and more consistent wind resources. However, the interaction between the floating system and the stochastic site conditions, which ultimately results in uncertainty in the structural response, still needs to be better understood. The outputs from modeling tools depend on a large number of input parameters, such as environmental conditions, controller design, and the design parameters of the structure. However, not all inputs influence FOWT loads equally, and the response of each part to changes in the environment is unique. Moreover, in the case of floating wind turbines, the response is also a function of the substructure type [8]. This means that the sensitivity analysis for semi-submersibles may not be entirely relevant for the spar buoys.

It is essential to identify the correlations between the loads and site conditions, as this information can be used to develop more effective engineering, reduced-order, and data-driven surrogate models. Several studies in the past have focused on addressing the sensitivity of wind turbine loads to environmental conditions [34–36, 164, 165]. The combined effect of environmental and structural parameters has been analyzed on fixed-bottom [38, 39] and floating wind turbines with semi-submersible platforms [41, 42]. This study was motivated by the need to (i) validate the simulation tools against measurements for a unique floating wind turbine substructure by performing one-to-one simulations, (ii) identify correlations between the site conditions and the damage equivalent loads on the wind turbine tower and blades for both simulation data and measurements. One-to-one comparison refers to simulations where inputs are derived from time-synchronized measurements of a wind turbine prototype. Each simulation corresponds to a single measured realization and is used to validate the aeroelastic model's ability to reproduce observed load time histories.

The measurement data are obtained from the measurement campaign of the TetraSpar demonstrator [9, 166]. It is equipped with a 3.6MW Siemens Gamesa Renewable Energy wind turbine. An image of the platform with the wind turbine is shown in Figure 5.1. It is a unique spar-type configuration consisting of a slender tubular steel structure with a suspended keel. It is a promising design candidate for floating wind turbine foundations, as it is made up of simple parts and minimizes material costs. Due to its unique configuration, it presents an interesting test case to validate the accuracy and limits of the state-of-the-art engineering tools used in industry. The TetraSpar Demonstrator has been installed at the Marine Energy Test Centre, 10 km off Karmøy in Norway, with a water depth of 200 m and has been operating since July 2021 [9]. The data used for the analysis consists of 10-minute statistics of environmental conditions and the corresponding 10-minute DELs obtained from the measurement campaign.

The first section of the chapter provides information about the data used in the study, lists the various pre-processing steps and provides some background on the simulation setup used for the numerical analysis. The results are subdivided into two sections. The first part provides a qualitative comparison using scatter plots of the measured and simulated damage equivalent loads as a function of the environmental features. A visual comparison is useful in identifying key differences in the simulations and in highlighting

the similarities and regions for improvement in the numerical setup. The second part of the result section uses statistical approaches to investigate the quantitative correlation between the loads and the site conditions.

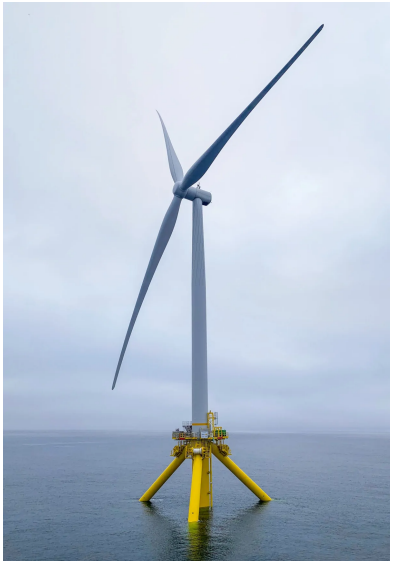


Figure 5.1: Image of the TetraSpar full-scale demonstrator moored at the test site off the coast of Norway [9].

For the correlation study, CODEC and a corresponding variable selection algorithm, called FOCI [132], are used. Other standard correlation metrics, including MI and PCC, are also evaluated. A detailed description of these methods can be found in Chapter 3.

An extension of this study would be to conduct a global sensitivity analysis (GSA). In the context of wind turbines, GSA is typically performed using either derivative-based methods, such as elementary effects (EE), or variance-based methods, such as Sobol's indices. These approaches rely on structured sampling strategies and require an explicit function to evaluate the input-output relationship, often an expensive engineering tool. Consequently, they cannot be directly applied to randomly sampled field measurement data. To enable GSA on field data, one approach involves fitting a surrogate model to the observations and then sampling the responses accordingly. Alternatively, GSA can be performed using simulations. However, a critical question arises: do the simulation results align with observed field data? To assess the reliability of state-of-the-art engineering tools used in the industry for such analyses, one-to-one simulations are conducted using the in-house multi-physics tool, BHawC [167, 168], coupled with OrcaFlex [169]. An open-source tool, such as OpenFAST, was not used in this project, as the numerical models were set up and simulated by the commercial partners. The correlation studies are then repeated on both measured and simulated data to evaluate consistency and validate the applicability of these tools.

5.2 Data processing

5.2.1 Data description

The Tetraspar Demonstrator is heavily instrumented, with the sensors onboard providing data to a central data acquisition system, which stores data at a frequency of 25 Hz. The turbulence intensity, wind speed at hub height, and shear profile are processed from the nacelle-based LiDAR measurements. The LiDAR is configured only to output the longitudinal turbulence component. The transverse and vertical turbulence components are assumed to be 0.7 and 0.5 times the longitudinal turbulence component. The free-stream wind conditions are measured 364 m (2.8 rotor diameter) upstream of the rotor plane. The onboard wave radar measures the wave elevation and direction, which are then processed to obtain values of significant wave height and peak spectral frequency. The wave radar is located at the railings of the foundation, targeting the area directly below the wind turbine. It is, therefore, subject to additional disturbances due to the surrounding floating structure compared to the open ocean. A detailed study on the validity of the wave radar measurements is performed by McCauley et al. [170]. Ocean current characteristics in the form of current speed and direction are not measured directly at the location and are obtained from the Norwegian Meteorological Institute's reanalysis data. Air density is derived from temperature and pressure measurements.

The responses are the damage equivalent loads measured at the blade root, tower top, and tower bottom. They are listed in Table 5.2. For the tower top and bottom, the results for the resultant loads are shown, and not specifically the fore-aft or side-side moments. The TetraSpar, unlike a traditional spar buoy, has a hydrodynamic asymmetry, which means its dynamic response can be directional. Therefore, it is easier to interpret the sensitivities for the resultant loads on the tower base and top. The loads are normalized using the minmax scaler. The input parameters considered for this study are listed in Table 5.1. The periods of normal operation are selected and processed to remove any faulty sensor measurements between 3 March 2022 and 31 October 2022.

Table 5.1: List of variables characterizing the environmental conditions.

Variable	Source
Turbulence intensity, hub height mean wind speed, shear profile power law exponent	Nacelle-based LiDAR
Air density	Derived from temperature and pressure
Platform pitch angle	Inclinometer (tower top)
Significant wave height, wave elevation peak frequency, wave direction	Wave Radar
Current speed, Current direction	Meteorologisk Institut (met.no)

Table 5.2: List of the load measurement locations and the corresponding description.

Variable [Unit]	Description
Tower bottom res [-]	Tower bottom resultant moment (10-minute DEL)
Tower top res [-]	Tower top resultant moment (10-minute DEL)
Edgewise blade root [-]	Edgewise blade root moment (10-minute DEL)
Flapwise blade root [-]	Flapwise blade root moment (10-minute DEL)

5.2.2 Preprocessing steps

The filtered 10-minute measurements and simulation output are further processed to render them suitable for our statistical analysis tools.

- **Missing values**: For any missing measurement channel in the dataset, all measurements from that timestamp are discarded.

- **Outlier detection**: CODEC and FOCI analysis have an additional outlier detection step. A simple metric called the *z-score* is used to mark any sample more than 2.5 standard deviations away from the mean as an outlier. For the i^{th} data sample with value x_i, the z-score value is defined as

$$z_i = \frac{(x_i - \bar{x})}{\sigma} \tag{5.1}$$

  where $\bar{x}$ and σ are the mean and standard deviation of the feature x.

- **Acquisition frequency**: Current speed, current direction, and density are recorded every hour, and therefore, their values repeat over a period of time until the next acquisition point. This does not introduce an issue with MI and PCC. However, the repeated values can introduce an artificial bias in the data for CODEC, and FOCI. Therefore, the CODEC and FOCI analyses are performed on a much smaller dataset where the repeating values of density, current speed, and current direction are discarded.

After outlier detection, roughly 9489 10-minute samples remain to perform MI and PCC. The number of samples was further reduced to 1145 for CODEC and FOCI as the samples with repeating feature values were discarded. CODEC uses nearest neighbors to approximate the conditional dependence between variables. In this study, 10 stochastic variables are analyzed using a total of 1145 samples, which is considered relatively limited given the dimensionality. Because the variables are stochastic, the measurements exhibit noise or variation—stemming from either measurement tolerances or the inherent randomness of the variables. Given the limited amount of noisy data to perform the analysis, some uncertainty in the metrics is expected.

DEL (ΔM_{eq}) is calculated following the procedure outlined in Chapter 2, Equation (2.5). The n_{ref} is chosen as 600 for the 10-minute signal. m is the Wöhler exponent with values 3.5 for the tower, 10 for blade flapwise, and 8 for blade edgewise moments.

5.3 Simulation setup

BHawC is used to perform the numerical simulations. The aeroelastic code is dynamically coupled with the hydrodynamic solver, OrcaFlex, which models the platform hydrodynamic loads, mooring loads, and response. At each coupling step, BHawC calculates the aero-servo-elastic response, providing OrcaFlex with the displacement, velocity, and acceleration of the interface between the tower and the foundation. With these inputs, OrcaFlex calculates the hydro-elastic response of the full structure and condenses it at the interface location. The suspended keel and tubular structures are modeled as stacked spar-buoy cylindrical elements under the category of six-dimensional buoys in OrcaFlex. The mooring line dynamics are solved using a Finite Element model, where the line is discretized into a series of lumped mass nodes connected by straight, massless segments with axial and torsional spring dampers at each end. The total simulation duration per simulation is set to 600 s, to match the 10-minute statistics from the TetraSpar Demonstrator measurements. The simulations have a mean sample period of 0.01 s, but the output is logged at every fourth time step, resulting in a frequency of 25 Hz to match the sample rate of the measurements.

The Mann model [95] is used to generate spatially coherent turbulent flow structures in the inflow. It uses a frozen-Taylor assumption (see Chapter 2) to represent a wind time series as a spatial wind field. The turbulence box is scaled based on the wind speed, turbulence intensity, and the shear profile values, which are obtained from the measurement campaign. The statistical variation in the turbulent field is reproduced by using different random seeds (ranging from 1 to 45) for the simulations.

The double-peaked Torsethaugen wave spectrum [171] is chosen to generate stochastic waves in OrcaFlex. It represents sea states that include both swell and wind-generated waves. An interpolated current profile is specified, based on the reanalysis data from the Norwegian Meteorological Institute.

5.4 Results

This section is divided into two parts. First, the simulation predictions are compared with the measurements obtained from the TetraSpar Demonstrator using scatter plots, MI, and PCC. The scatter plots provide a qualitative comparison of the simulated and measured loads as a function of each environmental feature. MI and PCC provide a quantitative assessment of the degree of correlation between the fatigue loads as a function of the environmental conditions. Note that the variations in the response are a cumulative effect of changes in all the parameters. However, the correlation metrics disregard any conditional dependence between the input variables.

In the second part, the results obtained using FOCI are looked at, which takes into account the conditional dependence between the input variables when calculating the score. The conditional dependence coefficient is helpful in determining a subset of interesting features that most explain the behavior of the loads.

5.4.1 Data visualization and standard correlation metrics

Figure 5.2 shows the scatter plots of the 10-minute damage equivalent loads at tower bottom, tower top, and blade root from measurements and one-to-one numerical simulations against the standard deviation of wind speed at hub height. Firstly, the simulations and measurements have a good qualitative agreement, and the statistical trends are similar. Regarding absolute values, the blade root edgewise DELs and tower DELs in simulations show a larger variance than the measurements. Subsequent investigation revealed a limitation in the structural modeling of the suspension lines connecting the hull to the keel. In simulations, these suspension lines do not remain as taut as the TetraSpar Demonstrator, primarily due to smaller structural damping. In certain cases, it results in artificially increased fatigue, especially in the non-damped side-side tower moment and the edgewise blade loads.

The tower bottom moment DEL plotted against the significant wave height in Figure 5.3 shows a slightly higher variance and, thereby, higher load magnitudes at lower significant wave heights in the measurements. Whereas in the simulations, the correlation is relatively linear. This could indicate that the measured loads are more wind-driven than the simulations. Conversely, the measured tower top moment DEL shows much less variance with respect to the site variables. Generally, the tower top loads are damped by the aerodynamic forces and control operations, and this effect is more pronounced in the measurements.

The plots highlight the importance of validating the simulated data against real measurements for complex designs. Minor differences in the setup can result in unexpected differences in behavior, particularly for quantities like fatigue, which is sensitive to dynamic effects like resonance and platform motion.

Figure 5.4 shows the MI and PCC scores for the tower bottom moment DEL. MI scores quantify what was qualitatively visible in the scatter plots. The measurements show a higher correlation with wind speed statistics, whereas the simulations show a higher correlation with the significant wave height. PCC scores follow the same pattern for simulations as MI, since the scatter plots appear to be more linear. However, the PCC scores for the measurements contradict MI, highlighting the limitation of using linear correlation metrics. The measured tower top loads in Figure 5.5 display a stronger correlation with the significant wave height as compared to BHawC simulations. Pitch and surge motion of the structure can induce tilting and displacement of the tower top even when the wind speed is constant. PCC and MI show a similar trend, which means that the correlations are linear in nature. Fatigue is tied to dynamic fluctuations. The correlation with mean wind speed seen here is a reflection of the changes in the operational states, control actions, and rotor speed.

The blade loads show a good agreement between the simulations and the measurements in terms of the correlation scores. The blade loads are mainly correlated to the wind statistics, especially the flapwise fatigue, which is turbulence-driven, as expected. Further investigation was conducted into how the flapwise fatigue correlates with the significant wave height.

Figure 5.8a shows the standard deviation of the measured platform pitch angle (measured at the tower top) plotted against the significant wave height. The color of each point corresponds to the mean wind speed, and its size reflects the wind speed standard

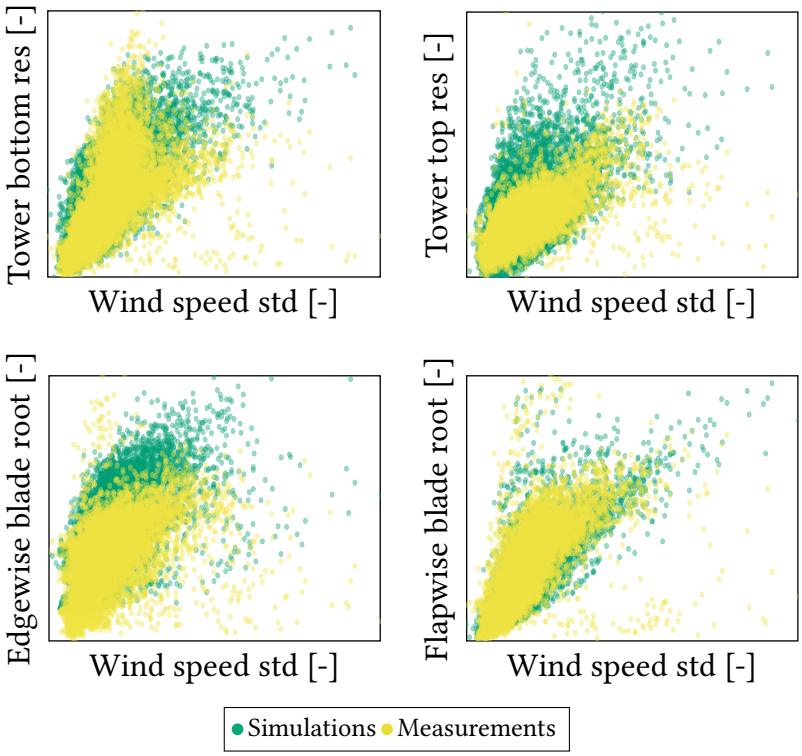


Figure 5.2: Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. standard deviation of wind speed, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality.

deviation. Then, in Figure 5.8b, the blade root flapwise 10-minute DEL against the platform pitch standard deviation is shown. Here, point color represents significant wave height, and size reflects the wind speed standard deviation. Figure 5.8a shows that higher blade root flapwise fatigue tends to occur when the significant wave height is high, and the standard deviation of hub height wind speed is large. The connection between significant wave height and flapwise fatigue might be influenced by the relationship between the platform pitch standard deviation and significant wave height, and how the platform pitch standard deviation relates to flapwise fatigue. This illustrates that it is challenging to account for the conditional dependence between all variables and trace a unique cause of fatigue damage.

5.4.2 Feature ordering by conditional independence

Feature ordering is an essential pre-processing step in developing data-driven models. Table 5.3 shows the features chosen by this algorithm to contain the most information about the loads. The values of the conditional dependence coefficient calculated with all the features versus those calculated with just the important features are also shown in the table. Note that the values of the coefficient range from zero to one; therefore, the closer

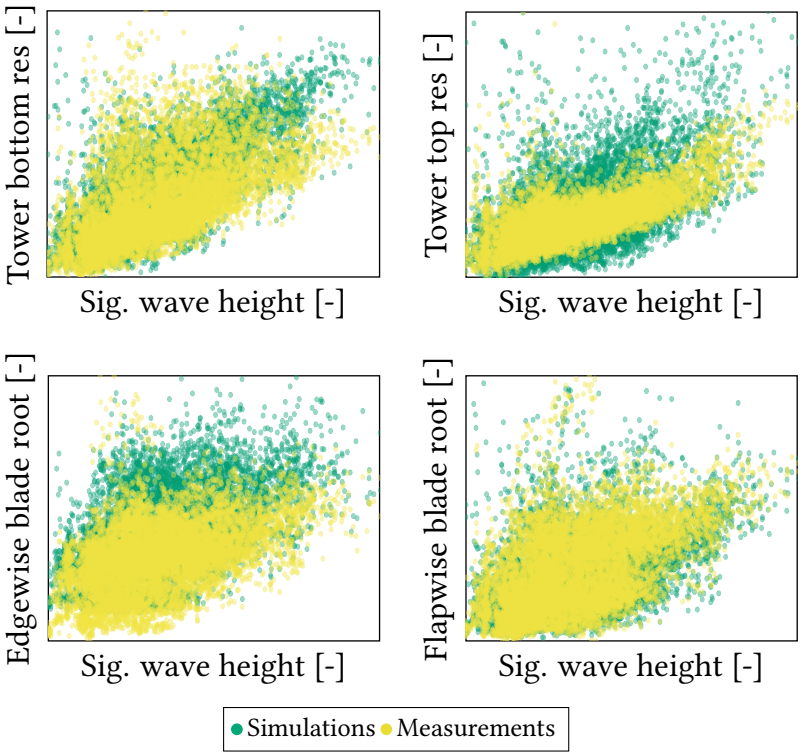


Figure 5.3: Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. significant wave height, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality.

the values are to unity, the higher the chances of being able to predict the loads using the information from the values of the important parameters. The most relevant features align with the trends seen in Section 5.4.1. Additionally, based on the values of the conditional dependence coefficient, C, it is observed that including non-important features in the dataset can lead to a reduction in the quality of information contained in it. Comparing the $C(Y|X(FOCI))$ values for this sample set, one would expect to have more success in training a data-driven model for predicting the blade root edgewise fatigue than the tower top resultant fatigue, as the latter is largely unexplained by the ten environmental variables considered in this study. It is also worth mentioning that FOCI employs the k-nearest neighbor algorithm to rank the distances between samples in practice. If two features are closely related in relevance, they may switch their order of importance with changes in the dataset.

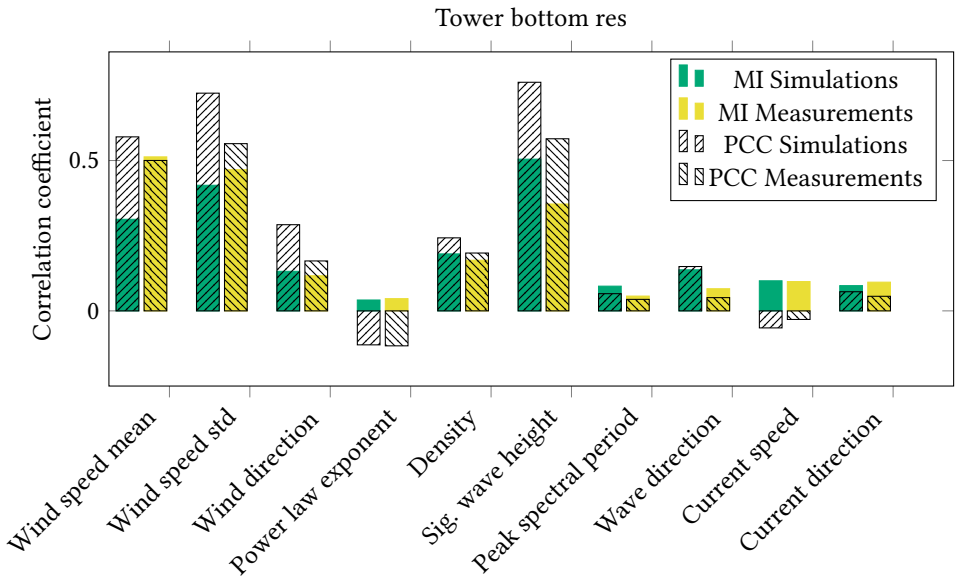


Figure 5.4: MI and PCC values between the tower bottom resultant 10-minute DEL and the various environmental variables from the measurements and simulations.

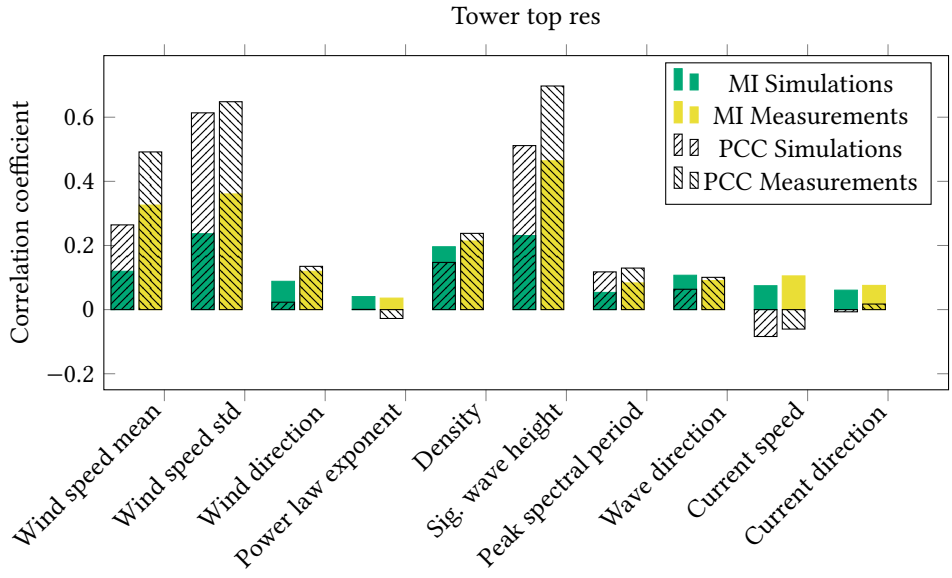


Figure 5.5: MI and PCC values between the tower top resultant 10-minute DEL and the various environmental variables from the measurements and simulations.

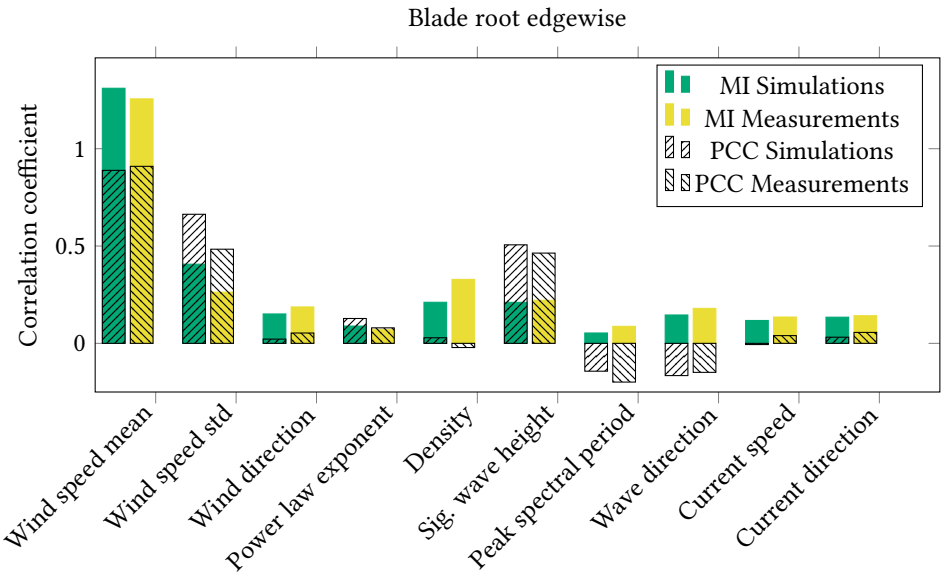


Figure 5.6: MI and PCC values between the blade root edgewise moment 10-minute DEL and the various environmental variables from the measurements and simulations.

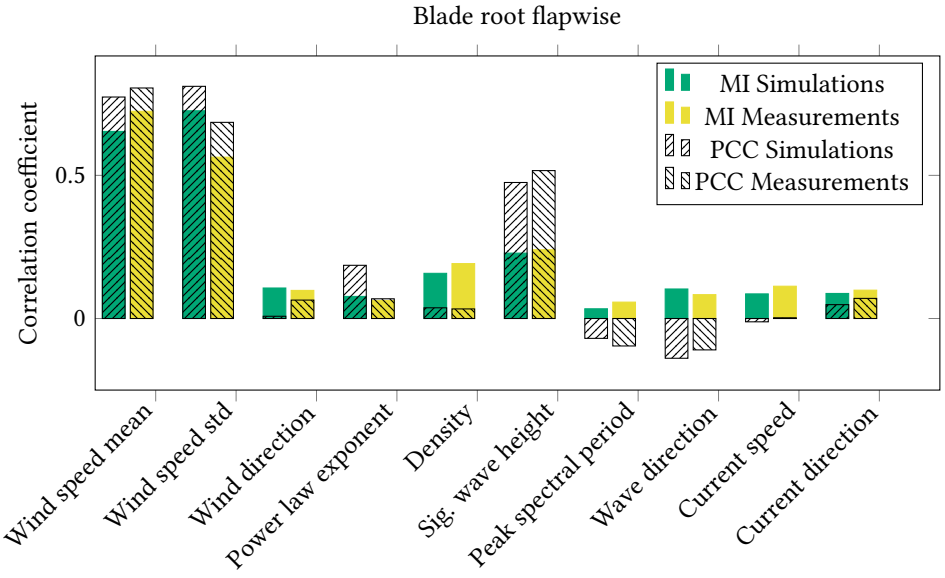
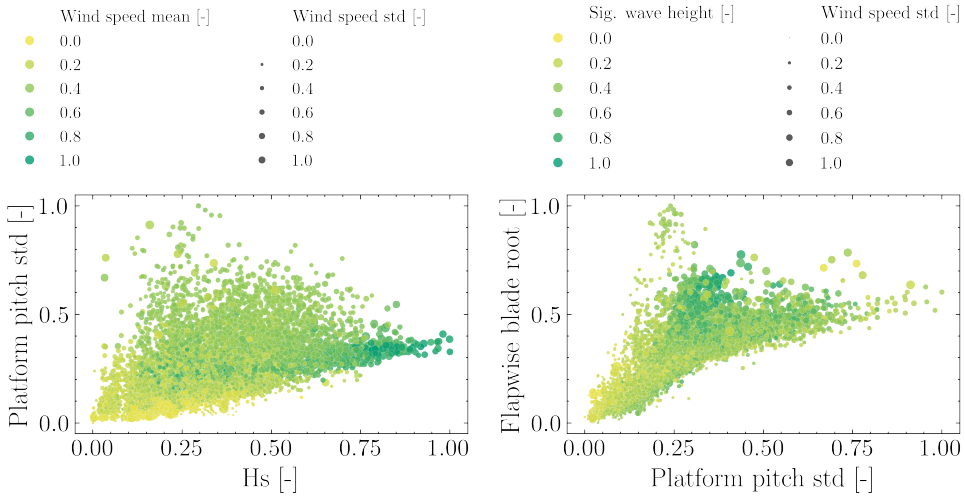


Figure 5.7: MI and PCC values between the blade root flapwise moment 10-minute DEL and the various environmental variables from the measurements and simulations.



(a) The standard deviation of the platform pitch angle vs. the sig. wave height, shaded by the mean wind speed

(b) Blade root flapwise 10-min. DEL vs. the standard deviation of the platform pitch angle, shaded by the sig. wave height

Figure 5.8: Scatter plots from the TetraSpar measurements, showing the influence of significant wave height on the standard deviation of the platform pitch and blade root flapwise moment 10-min. DEL. The size of the symbols is based on the standard deviation of the wind speed.

Table 5.3: Feature ordering using FOCI on the TetraSpar measurement dataset. DEL refers to the 10-minute DEL.

Feature importance	Tower bottom res. DEL	Tower top res. DEL	Blade root edgewise DEL	Blade root flapwise DEL
1	Wind speed mean	Sig. wave ht.	Wind speed mean	Wind speed mean
2	Wind speed std.	Wind speed mean	Density	Wind speed std.
3	Sig. wave ht.	Wind speed std.	Sig. wave ht.	-
4	-	Wave direction	-	-
$C(Y\|X)$	0.65	0.52	0.69	0.61
$C(Y\|X(FOCI))$	0.7	0.57	0.81	0.7

Table 5.4: Feature ordering using FOCI on the TetraSpar simulation dataset. DEL refers to the 10-minute DEL.

Feature importance	Tower bottom res. DEL	Tower top res. DEL	Blade root edgewise DEL	Blade root flap-wise DEL
1	Sig. wave ht.	Wind speed std.	Wind speed mean	Wind speed std
2	Wind speed std.	Wind direction	Wind speed std.	Wind speed mean
3	Wind speed mean	Wave direction		
4	Wind direction	Wind speed mean	-	-
$C(Y\|X)$	0.64	0.44	0.73	0.64
$C(Y\|X(FOCI))$	0.66	0.46	0.81	0.75

5.5 Conclusion

This study investigated the relationship between environmental conditions and 10-minute fatigue loads on the TetraSpar floating demonstrator equipped with a 3.6 MW Siemens Gamesa wind turbine, using measured data and numerical simulations. By comparing one-to-one simulations with full-scale measurements, the accuracy of engineering tools in capturing real-world behavior and identifying the environmental variables that affect turbine loads is assessed.

The analysis with MI and PCC showed that wind speed statistics and significant wave height are consistently the primary factors influencing fatigue loads on both the tower and blades. For tower bottom moments, measurements showed a stronger correlation with wind speed mean and standard deviation (MI=0.5) than with the significant wave height (MI=0.35). On the other hand, the simulations reflected a larger influence from the significant wave height (MI=0.5) than the wind speed mean (MI=0.27) and wind speed standard deviation (MI=0.4). This difference suggests the need to refine hydrodynamic modeling in simulation tools, particularly regarding damping effects in platform components, such as the suspended keel.

Blade root edgewise loads were mainly governed by mean wind speed, linked to rotor rotational speed. Flapwise loads were influenced by both wind turbulence and platform pitch motion. Current speed and direction show a negligible correlation with the fatigue loads for this specific configuration and site. In the future, the impact of modeling the current as measured by the wave buoy should be investigated, versus using the remodeled current, which introduces additional unknowns and uncertainty in the analysis. The use of multiple correlation metrics— PCC, MI, and conditional dependence analysis through FOCI—helped confirm these findings and provide useful information for feature selection in future surrogate model development.

It is evident from this analysis that it is very difficult to account for the conditional dependence between all variables and trace a unique cause of fatigue damage, particularly when the data are in a non-standard format. It is also important to validate the simulation setup before performing a correlation analysis or a sensitivity study, as the differences in the load response can result in differences in feature ordering. For future studies, global sensitivity analysis tools like Elementary Effects or Sobol sensitivity indices can be explored with simulation data.

6

Surrogate modeling for floating offshore wind turbine loads

Floating offshore wind turbines experience complex hydrodynamic and aerodynamic loading, influenced by the type of substructure and stochastic environmental conditions. Accurately estimating lifetime fatigue loads requires analyzing thousands of operational scenarios, resulting in high computational costs. Moreover, choosing the right input features that drive fatigue in floating wind systems and appropriately binning them remains an open question. A fast probabilistic surrogate that maps site conditions to loads on the wind turbine is presented. The probabilistic aspect allows the propagation and quantification of statistical uncertainties from the stochastic input quantities to the resulting loads. A fast surrogate eliminates the need to fit a distribution to the site conditions or bin the input data. Rather, all available met-ocean data can be directly used as input, which automatically accounts for the joint distribution in the calculations. The surrogate model in this study uses mixture density networks to predict the conditional distribution of the 10-minute damage equivalent loads for a 6 MW spar-type floating wind turbine. The mixture density network model achieves high accuracy ($R^2 > 0.99$) in capturing the conditional mean of the damage equivalent loads while efficiently propagating the statistical uncertainties. Furthermore, the surrogate enables quick estimation of 25-year lifetime fatigue damage across a range of potential floating wind farm sites, demonstrating its capability to facilitate rapid decision-making during preliminary site analysis.

This chapter is partly based on:
D. Singh et al. Data-driven probabilistic surrogate model for floating wind turbine lifetime damage equivalent load prediction [86].

6.1 Introduction

6.1.1 Background

As explained in Chapter 1, FOWT technology has witnessed a surge in research interest in recent years following the rapidly increasing demand for renewable power production. The structural response of an FOWT is a crucial indicator of its performance, safety, and reliability. During its operational lifetime, a FOWT accumulates fatigue damage as it undergoes time-variable loading in response to the complex and stochastic offshore environment. The nature, magnitude, and extent of fatigue are unique to the type of floating foundation, mooring line configuration, wind turbine material, geometry, mass, control algorithms, and site conditions.

The first step in the design load case analysis involves simulations of a type-certified rotor-nacelle assembly with a reference tower and floating foundation. More detailed information about the site is included later as the project progresses. Subsequently, a site-specific tower, foundation, and mooring line configuration is defined, and a site-specific certification study is performed.

The lifetime fatigue load assessment entails calculating the 10-minute damage equivalent loads on multi-variate bins of typical variables characterizing the site and scaling them to the observed probability of occurrence. Not all site variables can be practically included in fatigue load analysis, as the required number of simulations increases significantly with each additional variable. The choice of the variables in the offshore environment that have the most impact on FOWT fatigue is currently an active area of research [91]. The total computational cost of the simulations also constrains the lower limit of the bin size. While industry-standard engineering tools are necessary for certification, the preliminary site analysis can benefit from *data-driven surrogate* models to provide quick load estimates. Data-driven surrogates can infer complex relationships from data observations alone, without requiring prior knowledge of the underlying physics. Fatigue, which is difficult to model using lower-fidelity physics-based approaches, can benefit from such data-driven methods. Using surrogates that can accurately predict DELs can potentially eliminate the need to bin the site data, fit a multivariate joint distribution to it, or limit the total number of parameters. Once trained, surrogate models can utilize all the available site information to quickly estimate the site-specific loads. Additionally, probabilistic surrogates can also propagate statistical uncertainty from the stochastic input variables to the loads.

This study focuses on using a probabilistic data-driven surrogate to map 10-minute statistics of the environmental conditions to the corresponding conditional probability distribution of the DEL on the floating spar buoy with a 6 MW Siemens Gamesa wind turbine. The DEL values are calculated using the Siemens Gamesa in-house tool that couples the aeroelastic code, BHawC [26], with the hydrodynamic solver, OrcaFlex [172]. A schematic of this mapping is shown in Figure 6.1. A highly flexible probabilistic machine learning approach for the surrogate, MDN [138] is used in this study. The probabilistic estimates of DELs are used to subsequently calculate the lifetime fatigue loads at various potential floating wind sites.

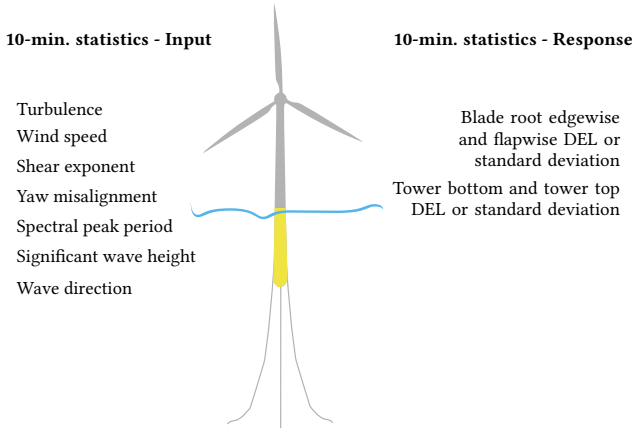


Figure 6.1: Schematic of the surrogate modeling mapping for a floating offshore wind turbine.

6.1.2 Research objectives

The objectives of this work are threefold:

- To apply and evaluate a probabilistic machine learning model for predicting short-term fatigue loads in floating offshore wind turbines:
  Not only are FOWT loads affected by the nonlinear dynamic behavior of the floating wind turbine, but also the underlying simulations are much more complex, the stochastic hydrodynamic loading has a bigger impact on the tower, the samples are more 'noisy' due to the additional stochastic variables, the parameter space is larger and, thereby, the number of simulations in certain design load cases is larger. This necessitates investigations into robust surrogates. Probabilistic data-driven models are often validated on simpler cases. Complex cases are modeled using deterministic approaches, or with the assumption that the response is Gaussian or homoscedastic. Using a mixture model effectively solves many of the gaps in the literature by propagating uncertainty, including heteroscedasticity, and modeling higher-order moments.

- To demonstrate that such a model can provide insights into probabilistic long-term fatigue estimation, especially in high-dimensional spaces where traditional binning approaches become computationally restrictive:
  The lifetime fatigue estimation for FOWTs involves additional complexity, including the joint distribution of metocean conditions and sensitivity to bin sizing. These issues are exacerbated by the growth in the number of simulations with each additional dimension. A surrogate model enables the direct use of historical site condition data, bypassing the need for parametric joint distribution assumptions.

- To investigate the impact of stochastic variability in site conditions on fatigue prediction over long time spans:
  Four potential floating wind sites with similar water depths that use the surrogate

model to estimate probabilistic lifetime fatigue loads are selected. The effect of accounting for the stochastic variability in the 10-minute loads relative to the variation in site conditions is investigated. It is demonstrated that due to the law of large numbers, the influence of random seeds on long-term tower load estimation diminishes significantly over a 25-year period, even with 500 seeds (from the probabilistic surrogate). This indicates that stochastic inflow conditions contribute minimally to uncertainty in long-term fatigue predictions.

By addressing these points, this work contributes a validated, flexible surrogate modeling approach that accounts for the complexities of floating wind systems, advancing the state of probabilistic load modeling in floating offshore wind research.

The rest of the chapter is structured as follows. In Section 6.2, the floating wind turbine model is introduced and the simulation tool described. The responses and the relevant site features are identified and the bounds defined in Section 6.3. The details of generating the database are described in Section 6.4. The results, presented in Section 6.5, are divided into three sub-sections. First, the model convergence across a range of training samples is investigated. Next, the selected MDN model's 10-minute conditional distribution estimates under randomly chosen environmental conditions are validated and compared with those generated by BHawC. Finally, the probabilistic 10-minute estimates are used to propagate the statistical uncertainty to the lifetime fatigue damage.

6.2 Numerical setup

6.2.1 Definition of the floating wind turbine

The floating wind turbine in this study is based on a modified geometry of the Hywind Scotland spar buoy foundation [173]. It comprises a 6MW Siemens Gamesa Renewable Energy direct drive wind turbine assembly, SWT-6.0-154, mounted on a spar buoy. The characteristic wind turbine parameters are listed in Table 6.1. The simulations use a tower with a larger diameter than the tower designed for the Hywind Scotland site. It is, therefore, stiffer and has a higher natural frequency than its installed counterpart. The geometry details of the tower and the floating platform used in the simulation are provided in Table 6.2. The floating substructure is attached to the ocean floor using catenary mooring lines, equally spaced at 120° in crowfoot configuration using bridle lines as shown in Figure 6.2 [173]. The structural properties of the main mooring lines and the bridle lines are listed in Table 6.3.

6.2.2 Numerical model

The damage equivalent loads are obtained through time-domain hydro-servo-aeroelastic simulations performed using BHawC-OrcaFlex coupled implementation. BHawC has been used for several years at Siemens Gamesa for wind turbine load calculations and is continuously validated against turbine prototypes and the entire operational fleet. A similar analysis may be performed with OpenFAST [23, 24] coupled with OrcaFlex [174] via FASTLink for reproducibility. Arramount et al. [172] present the mathematical background for the software coupling. In short, the tower, the rotor nacelle assembly, and

Table 6.1: Parameters of 6MW Siemens Gamesa wind turbine.

Parameter	Property
Rated power	6000 kW
Power control	Pitch
Drivetrain	Direct drive
Rotor diameter	154 m
Hub height	96 m
Rated wind speed	12 m s^{-1}
Rated tip speed	89 m s^{-1}
Nacelle mass	360 t

Table 6.2: Wind turbine tower and foundation properties [17, 168, 173].

Parameter	Value
Tower bottom outer diameter	9.45 m
Tower bottom thickness	0.08 m
Tower top outer diameter	4.89 m
Tower top thickness	0.029 m
Tower bottom elevation above SWL	13 m
Draft	78 m
Platform top geometry – length	12 m
Platform top geometry – diameter	9.4 m
Platform taper length	15 m
Platform bottom geometry – length	58 m
Platform bottom geometry – diameter	14.4 m

Table 6.3: Catenary mooring line properties.

Parameter	Value
Number of mooring lines	3
Angle between mooring lines	120°
Mooring bridle line length	50 m
Mooring bridle line mass per unit length	0.348 t m^{-1}
Mooring main line length	610 m
Mooring main line mass per unit length	0.4322 t m^{-1}
Mooring line anchor radius	640 m

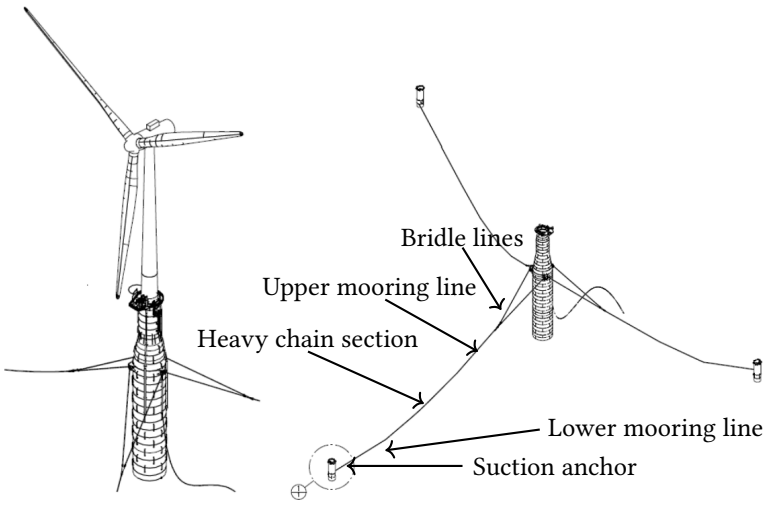


Figure 6.2: Hywind Scotland spar buoy with crowfoot mooring line configuration [173].

the blade elements are dynamically modeled in BHawC. The BHawCLink module serves as a communication channel between the dynamic link library and the OrcaFlex API. OrcaFlex simulates the hydrodynamic response of the floater element. Time integration is performed individually on both elements while accounting for the response of the other structure per iteration [172].

The inflow turbulence is modeled using a spatially varying frozen wind field based on the Mann model [95]. The tangential and axial induced velocities are calculated on several aerodynamic nodes on the blades using the blade element momentum theory coupled with Prandtl's tip loss correction and thrust correction at high induction values. Skewed and unsteady inflow is modeled using the method introduced by Björck et al. [175].

The structural elements are modeled using the co-rotational formulation, providing geometric nonlinearity [176]. The tower, shaft, and blade substructures are modeled using beam elements [89].

The Torsethaugen two-peak wave spectrum generates swell and local wind-driven waves [100]. The various elements of the OrcaFlex model are shown in Figure 6.4. A 6-DOF rigid buoy in OrcaFlex represents the floating substructure. The mooring lines are modeled in OrcaFlex. Each line is divided into several massless spring segments, joined by elements with lumped properties such as mass, damping, added mass, buoyancy, and material properties.

The simulations are initialized in BHawC with a quasi-static approach where the environmental loads (wind) and inertial loads (gravity and buoyancy) are ramped up in small steps. For every load step, OrcaFlex determines the mooring line static equilibrium based on the floater position determined by BHawC. BHawC calculates the global equilibrium position based on the stiffness matrices and interface loads provided by OrcaFlex. Once the global equilibrium is calculated, the next load step is applied. The dynamic part of the simulations consists of an initialization phase of 300 s to eliminate any initial transients

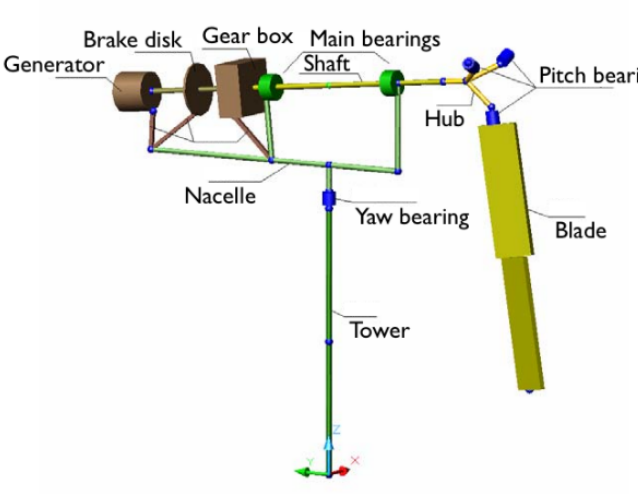


Figure 6.3: A simplified representation of the structural sub-elements used in the BHawC model [26].

as the wave dynamics, turbulence, and substructure motion build up, and the artificial structural damping is slowly ramped down. The final post-processing is performed on 600 s dynamic simulations that follow the initialization phase. The simulations for training the surrogate may be performed for a longer duration if necessary. For instance, to estimate mooring line fatigue correctly due to its long natural period. The effect on the tower and blade fatigue is shown to not change significantly with larger simulation windows, but rather with the fatigue calculation algorithm used to account for the unclosed cycles [177].

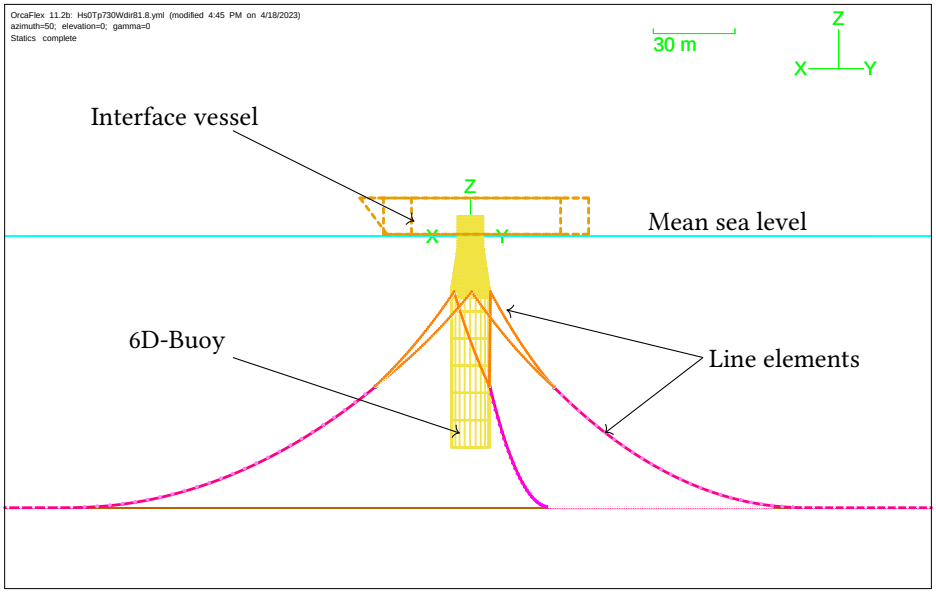


Figure 6.4: Schematic of the OrcaFlex simulation elements.

6.3 Definition of relevant site features and responses

Having a large feature space can lead to a very expensive surrogate training process, as the number of training samples required grows rapidly with the number of input variables due to the *curse of dimensionality*. It is, therefore, important to identify which variables have the most significant impact on fatigue to reduce the input space effectively.

Several studies in the past have focused on addressing the sensitivity of wind turbine loads to environmental conditions [34–36, 43, 164]. The combined effect of environmental and structural parameters has been analyzed on fixed-bottom [38, 39, 45] and floating wind turbines[41–44, 178].

Wiley et al. [41] demonstrate that for the OC4-DeepCwind semi-submersible platform, the standard deviation of wind speed in the inflow is the most influential parameter affecting the fatigue and ultimate loads on the tower and blades. Secondary drivers of fatigue on the tower bottom moment include turbulence coherence parameters as well as wave characteristics, such as significant wave height and peak spectral period. For blade

pitching fatigue, secondary factors include the yaw misalignment angle and geometric features like the blade twist angle. Whereas the wind-wave misalignment and the current speed and direction seem to have a secondary effect on the blade root bending moment fatigue. Reddy et al. [44] perform elementary effects analysis to determine the most significant parameters affecting tower bottom fatigue on the OC3-Hywind Spar platform and the OC4-DeepCwind semi-submersible design. In both cases, the significant wave height is found to be the primary driver. Current-related parameters are shown to have a strong effect, mainly on the mooring line fatigue. Singh et al. [43] use measurement data from the TetraSpar demonstrator to find that the tower and blade fatigue are most highly correlated with the wind speed mean and standard deviation, and significant wave height values.

As observed, although current has a big impact on the mooring line loads, its effect is found not to be as significant on the tower and blades in the literature. Therefore, it is not included as a variable feature in the training of the surrogate. A variation in the water depth was also not considered because the mooring line system must be redesigned for any new value of water depth. Automating this process is non-trivial. Furthermore, for slack mooring lines, Lin et al. [178] show a negligible impact on the tower fatigue with an increase in water depth.

Table 6.4: The list of input features provided to the surrogate model and their corresponding notation.

Feature	Label
Wind speed [ms^{-1}]	U_{ref}
Shear exponent [−]	α
Turbulence Intensity [%]	TI
Significant wave height [m]	H_s
Spectral peak period [s]	T_p
Wave direction [°]	W_{dir}
Yaw error/ misalignment [°]	Yaw

Table 6.5: The list of output channels that the surrogate models are trained to predict.

Response
Tower bottom fore-aft DEL [−]
Tower top fore-aft DEL [−]
Blade root edgewise DEL [−]
Blade root flapwise DEL [−]

The targets the surrogates are trained on are listed in Table 6.5. Each target is trained with a separate surrogate model. This study only calculates the short-term DELs in the local coordinate system of the component. DELs result from the conversion of the irregular load time series to a constant amplitude and frequency signal that produces an equivalent fatigue damage. Rainflow counting [102] algorithm is used to obtain the load ranges S_i and the number of load cycles n_i needed to calculate the DEL as shown in Equation (2.5). n_{ref} is 600 for 1 Hz DELs over 10 minutes. m is the Wöhler coefficient with values 3.5 for the tower, 10 for blade flapwise, and 8 for blade edgewise moments. Finally, Z-score normalization is used to scale the DEL to its dimensionless form.

6.3.1 Feature bounds

The feature bounds are defined based on the observations of data on sites where floating wind farms could potentially exist [179] and where data were readily available. Table B.1 in the Appendix lists the selected sites with their location and water depth values. The ERA5 reanalysis data, produced by the European Center for Medium-Range Weather Forecasts (ECMWF) on behalf of the European Union's Copernicus Climate Change Service

(C3S), is used for the analysis in this section.

Average wind speed at hub height

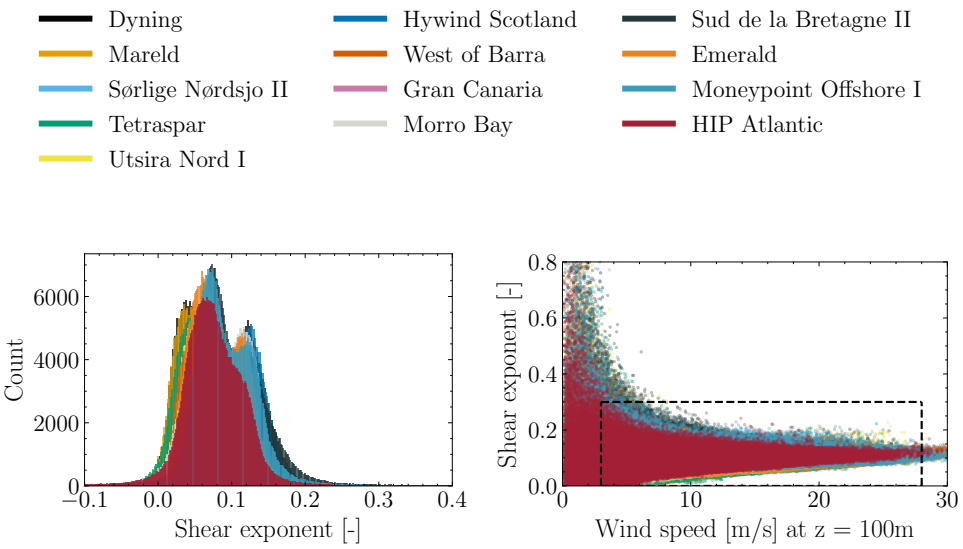
The wind speed at hub height (U_{ref}) is uniformly sampled between 3 and 28 m s^{-1}, which is the operational range of the wind turbine investigated in this study.

Shear exponent

The shear exponent α is defined according to the wind profile power law as,

$$\frac{U}{U_{\text{ref}}} = \left(\frac{z}{z_{\text{ref}}}\right)^{\alpha}, \tag{6.1}$$

where, U is the wind speed at height z, and U_{ref} is the known wind speed at height z_{ref}. The ERA5 reanalysis data are used to obtain wind speed values at 10m and 100m for the sites listed in Table B.1. Assuming the wind profile follows the power law, the shear exponent is calculated using Equation (6.1). The distribution of the shear exponent is shown in Figure 6.5a, with values primarily ranging between 0 and 0.2. It is also plotted against wind speed in Figure 6.5b. In our database, the shear exponent is uniformly distributed in the region corresponding to the dashed box in Figure 6.5b.



(a) Shear exponent distribution

(b) Shear exponent vs wind speed

Figure 6.5: (a) Histograms of the shear exponent α for selected sites for ERA5 reanalysis data from 1990 to 2019. (b) Shear exponent shown as a function of wind speed, marked by a box denoting the selected sampling domain.

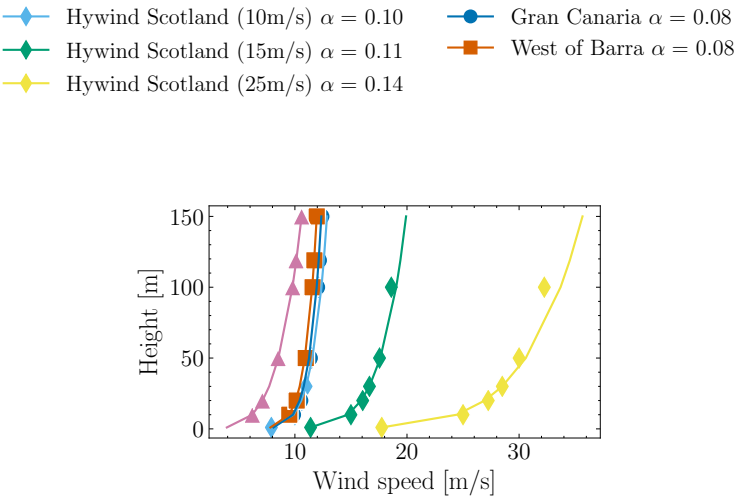


Figure 6.6: Examples of wind profiles on selected sites.

Turbulence intensity

The lower range of turbulence intensity is 0.1%, and the upper limit is designed to be 20% greater than the prescribed IEC Class C standard [87] for the Normal Turbulence Model (NTM). The function for class C turbulence intensity (in percentage) is given by,

$$TI = 100 \times \frac{I_{\text{ref}}(0.75 V_{\text{hub}} + 5.6)}{V_{\text{hub}}} \tag{6.2}$$

where, $I_{\text{ref}} = 0.12$ is the expected value of turbulence intensity at 15 ms^{-1} [87]. The upper limit of turbulence intensity for sampling is, therefore, $1.2 \times TI$. Figure 6.7 shows the chosen range, with the IEC class C turbulence and the measured turbulence at the FINO 3 metmast, which is referenced in the Buchanan deep met-ocean report [173].

Significant wave height

Waves in deep water primarily originate from two sources: wind-induced waves and swell waves. It is useful to consider the correlation between wind speed and significant wave height while training the surrogate to avoid including non-physical wind-wave combinations. Figure 6.8a illustrates a scatter plot of significant wave height (H_s) versus wind speed, based on ERA5 reanalysis data for the selected sites. The sampling domain is also a function of wind speed, highlighted with the dashed lines.

The upper and lower ranges of sampling for the significant wave height are defined empirically based on these observations. In this case, the functions are rather conservative and subject to modification based on the kind of sites the user would want to use the surrogate model on. The equations for the upper and lower limits for H_s are listed in

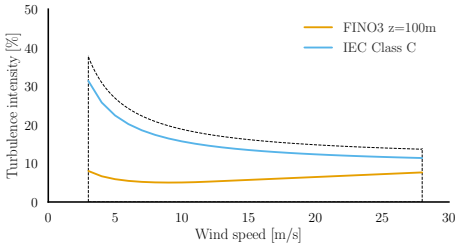
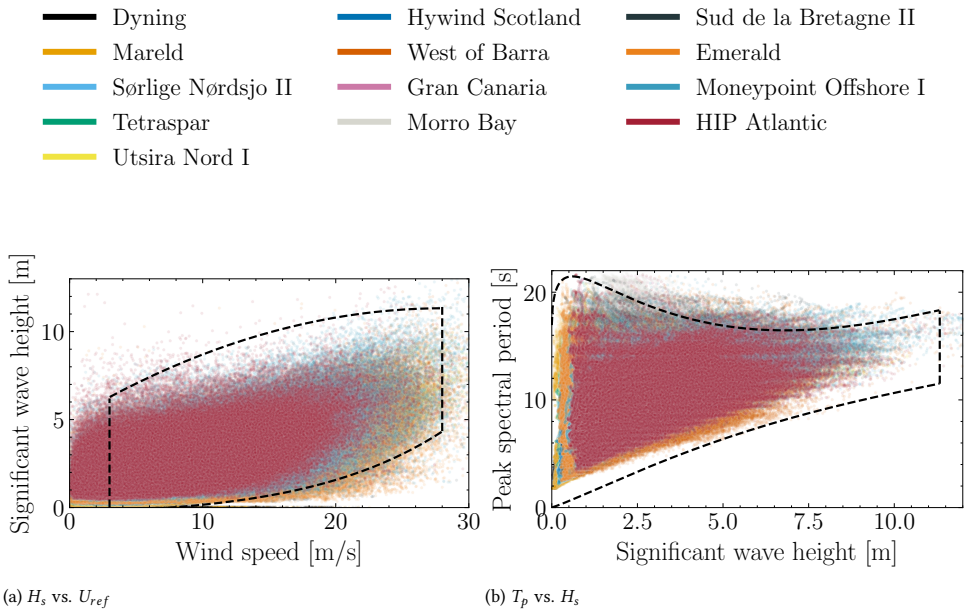


Figure 6.7: Chosen turbulence intensity range in dashed lines, along with the IEC class C turbulence profile for NTM, and measured turbulence at FINO 3 (German Bight) from the met-ocean analysis report on the Hywind Scotland project [173].



(a) H_s vs. U_{ref} (b) T_p vs. H_s

Figure 6.8: Scatter plots of (a) the significant wave height vs. the wind speed at 100 m, and (b) peak spectral period vs. significant wave height for the selected sites (Table B.1) based on the ERA5 reanalysis data from 1990 to 2019.

Section B.2.1. Recent studies have benchmarked various methods for generating environmental contours [180]. The readers are encouraged to create contours using automated approaches such as Kernel Density Estimate [181], IFORM [182], ISORM [183], etc., to avoid manually defining the empirical functions.

Peak spectral period

The empirical functions are defined for the spectral period range based on the significant wave height. Figure 6.8b illustrates the sampling domain with dashed lines,

overlaid on observational data from the ERA5 reanalysis. This plot also includes $H_s - T_p$ values corresponding to wind speeds below the cut-in speed and above the cut-out speed. The functions defining this range are detailed in Section B.2.2. As with the significant wave height, these bounding functions can be adjusted based on the region of primary interest to the user.

Wave direction

The wind turbine is always assumed to face the inflowing wind. Therefore, only the wave direction is varied to introduce wind-wave misalignment. Wave direction is considered to be an independent variable and sampled uniformly between 0° and 360°. For asymmetric floating foundations, however, wind directions would also need to be considered as an independent parameter.

Initial yaw misalignment

The effect of the initial yaw misalignment is chosen to be evaluated at −5.6°, 0°, and 5.6° while performing fatigue calculations. Therefore, the sampling bounds between $1.1 \times -5.6°$ and $1.1 \times 5.6°$ are selected.

6.4 Surrogate modeling setup

6.4.1 Training and testing database generation

Training database

Sobol sampling [158] is used to jointly sample in seven dimensions to generate the training dataset. The sampling is uniform in all dimensions. The samples lying outside the aforementioned feature bounds are discarded, resulting in a total of 9041 training samples. Alternatively, a multivariate distribution fitting the available data can be used to define the sampling space bounds. Each sample corresponds to a unique wave seed in OrcaFlex and a single inflow turbulence seed in BHawC. This approach is designed to emulate the inherent stochasticity of real-world inflow variables. Note that the statistical variation in the flow field is constrained by the BHawC implementation to only 45 turbulence seeds. Consequently, these seeds had to be reused, and the inflow turbulence box could not be uniquely defined for every case.

Testing database

The values of the shear exponent, turbulence intensity, and yaw misalignment are not randomly assigned to the test cases. Instead, they take the values commonly used while performing fatigue design load case evaluations. The shear exponent was fixed at 0.08, with yaw misalignment values of −5.6°, 0°, and 5.6°, and turbulence intensity corresponding to IEC Class C values. H_s, T_p, TI, and U_{ref} were jointly sampled in a random manner, without being tied to any specific location, but constrained within the defined feature bounds. For future studies, jointly sampling the test points across all variables is recommended for a fairer evaluation of the surrogate's performance.

In total, $n_{\text{test}} = 47$ test samples were used in this study. Each test sample simulation was repeated with $n_{\text{seeds}} = 44$ random seeds for turbulence and waves to capture the

statistical variation in the DEL values from the variation in the wind and wave fields. The seed repetition establishes a *reference* conditional distribution for each sample, which is used to compare against the probabilistic predictions of the surrogate model in Section 6.5.2. The samples used for training and testing the surrogate models are shown in Figure 6.9.

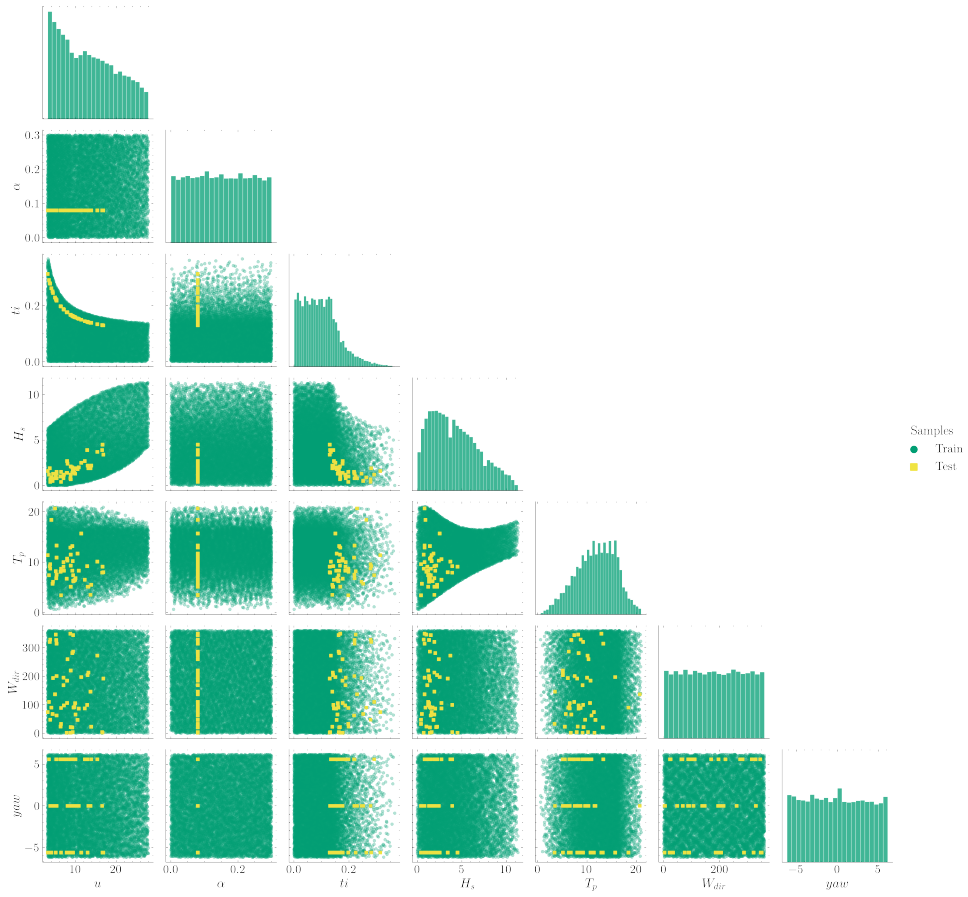


Figure 6.9: Paired scatter plots and marginal distributions of the training and testing datasets.

6.4.2 Mixture density network hyperparameters

The MDN model is described in Section 3.3.6. During training, the MDN model seemed susceptible to overfitting. In order to avoid that, a combination of *early-stopping* [140] and *L1* and *L2 regularization* [141] is used in this study.

The main hyperparameters used in this study to train the models to obtain the results are summarized in Table 6.6. In subsequent sections, the performance of the MDN model with various architectures is evaluated. The features and targets are scaled with the standard scaler before training.

Once the weights are calibrated, the model can output the parameters of the

Table 6.6: Summary of the mixture density network hyperparameters.

Network hyperparameter	Value
Number of mixture components	4
Activation function (hidden layers)	ReLU
Activation function (output layer)	None
Learning rate	$5e^{-3}$
Maximum epochs	5000
Mini-batch size	100
Optimizer	Adam
Regularization	
λ for $L1-$regularization	$1e^{-3}$
λ for $L2-$regularization	$1e^{-3}$
Early-stopping	
Early-stopping patience	100
Early-stopping monitor	validation loss
Number of early-stopping validation samples	600
Steps per epoch	number of training samples / batch size

mixture model for any new site condition $\boldsymbol{x}^{\star}$. To obtain the predicted pdf, 5000 samples from the mixture of Gaussians are generated.

6.5 Results

This section is divided into three parts. The first part presents a convergence study on the number of training samples, highlighting the model's robustness and demonstrating a clear trade-off between the computational cost of data generation and the resulting accuracy. A related hyperparameter study to determine the network architecture is presented in Subsection 6.5.1. The second part validates the performance of the surrogate on the test dataset. The validated model is used to make lifetime fatigue damage estimates on the wind turbine components in response to different site conditions in the third part.

In this section, the Wasserstein distance is calculated between the conditional distribution predicted for each sample ($\hat{Y}$) and the conditional distribution obtained as a reference through seed repetitions in BHawC (Y). $\hat{Y}$ consists of 5000 samples, and Y is obtained from 44 turbulence seed repetitions. Similarly, R^2 is calculated for the reference and predicted conditional pdf's mean and standard deviation values on the test dataset.

6.5.1 Choice of hyperparameters

Number of layers and nodes

Large networks are better at capturing complex expressions in data, but are susceptible to overfitting with a small training set. The objective of this study is first to observe the robustness of the model relative to the number of network parameters for a

particular training data size. And second, to choose a network architecture suitable for the rest of the study.

A sensitivity study on the number of nodes and layers is performed in this section for a training dataset of 8250 samples. The number of mixture parameters is four in all cases, and the rest of the network hyperparameters are fixed to the values listed in Table 6.6. Networks with two (Figure 6.10), three (Figure 6.11), and four (Figure 6.12) layers with various widths are tested. The x-axis in the figures lists the combinations of widths per layer evaluated in this study. The tower bottom fore-aft DEL channel is chosen for this study. The figures also highlight a dashed line with a value of 0.99 for R^2 and d_{W2} of 1, indicating a threshold for goodness of fit.

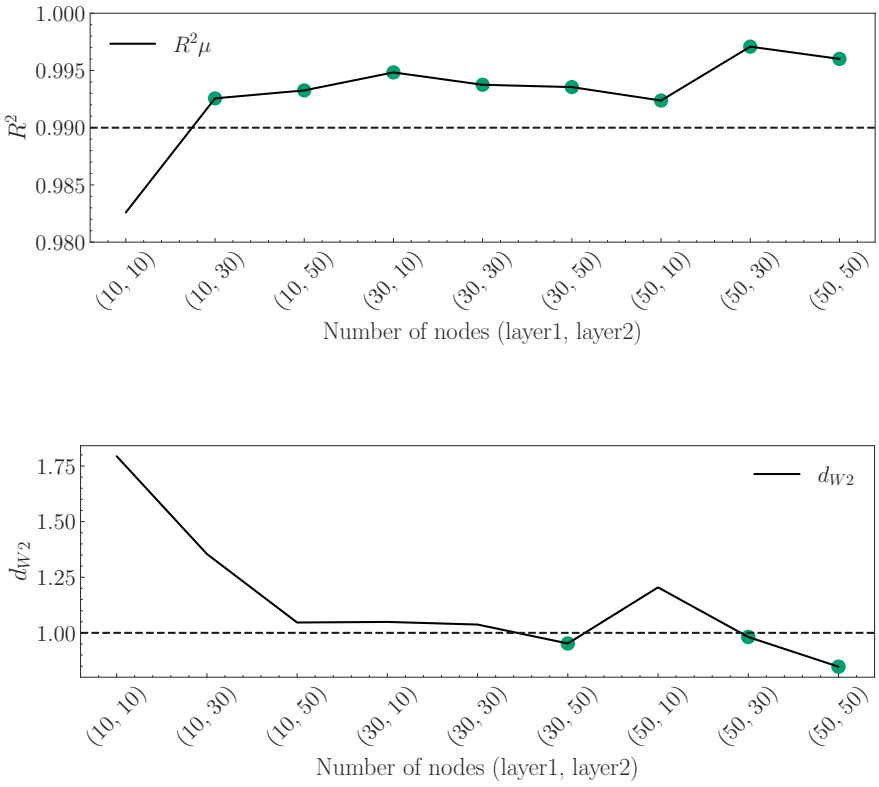


Figure 6.10: Study on the network architecture for two-layer networks. The top image shows the R^2 value for the mean of the conditional pdf of the tower bottom fore-aft DEL channel. The dashed line corresponds to an R^2 value of 0.99. The bottom image plots the d_{W2} values for the same channel with the dashed line corresponding to d_{W2} of 1.

The R^2 values of the DEL are notably good for the architectures tested, indicating good model robustness. The main differences observed are in d_{W2}, where the large three

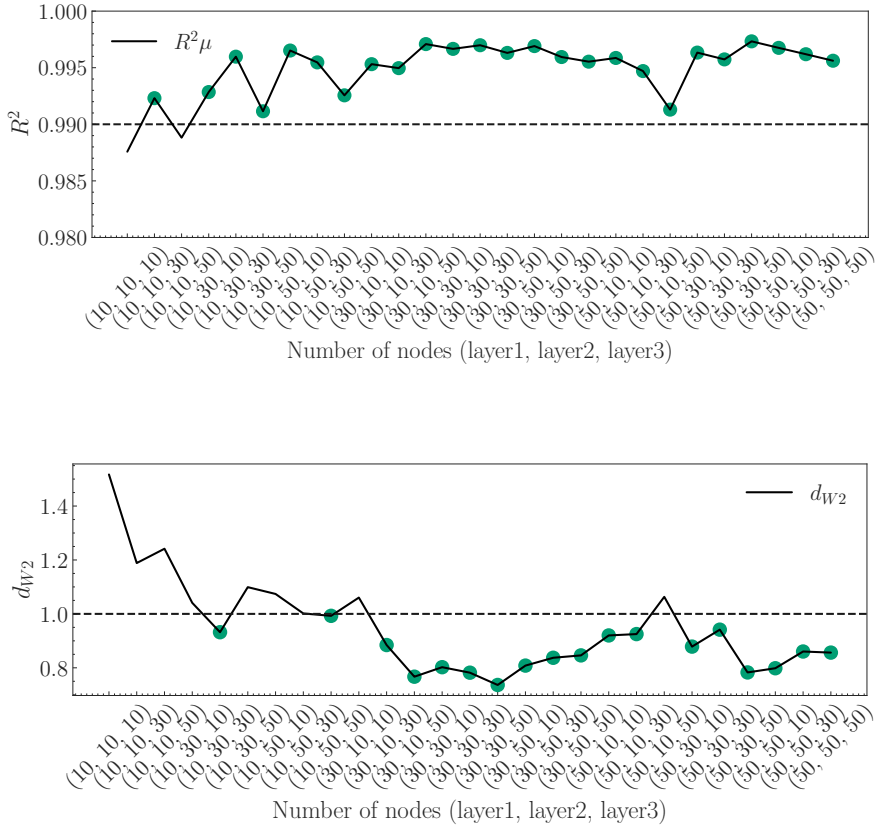


Figure 6.11: Study on the network architecture for three-layer networks. The top image shows the R^2 value for the mean of the conditional pdf of the tower bottom fore-aft DEL channel. The dashed line corresponds to an R^2 value of 0.99. The bottom image plots the d_{W2} values for the same channel with the dashed line corresponding to d_{W2} of 1.

or four-layer networks are generally better at capturing the complete conditional pdf. For the remainder of this study, a three-layer network with 30 nodes in the first layer, 30 nodes in the second layer, and 50 nodes in the third layer is chosen, in combination with 8250 training samples.

Number of mixture elements

The number of Gaussian distributions in the mixture controls the complexity of the predicted conditional pdf. However, a large number of unnecessary mixture elements adds redundancy and increases the computational complexity of the surrogate. In this section, 6250 and 8250 training samples are used with a three-layer architecture, width = (30, 30, 50). The performance of 4, 12, and 20 mixture elements on the tower bottom

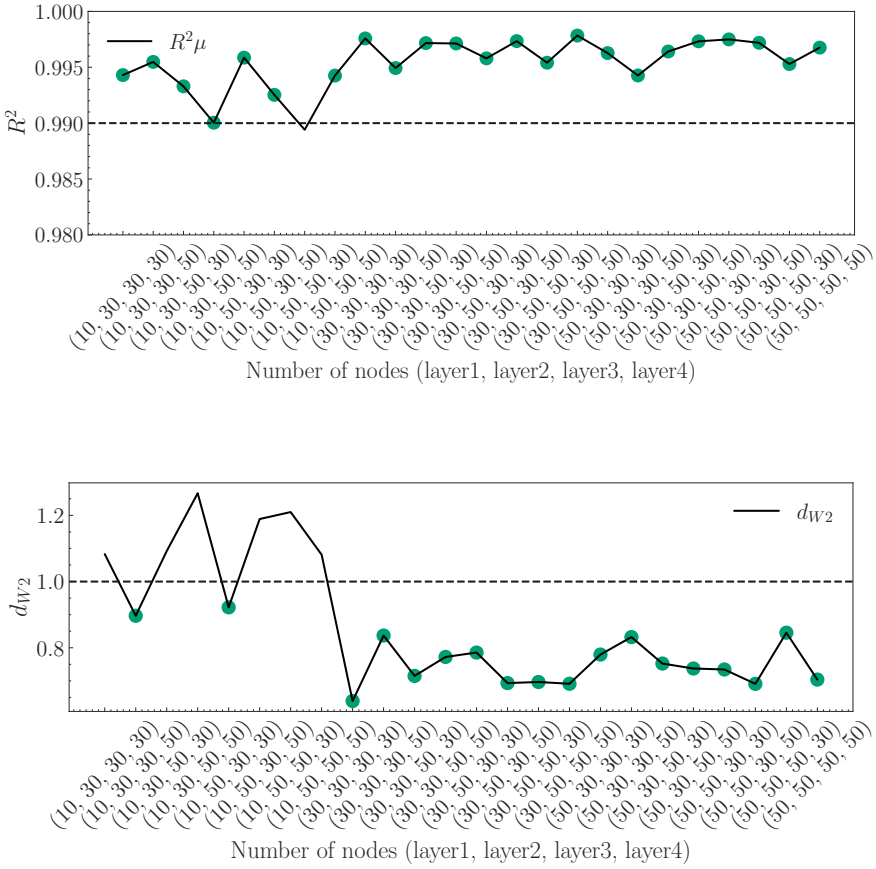


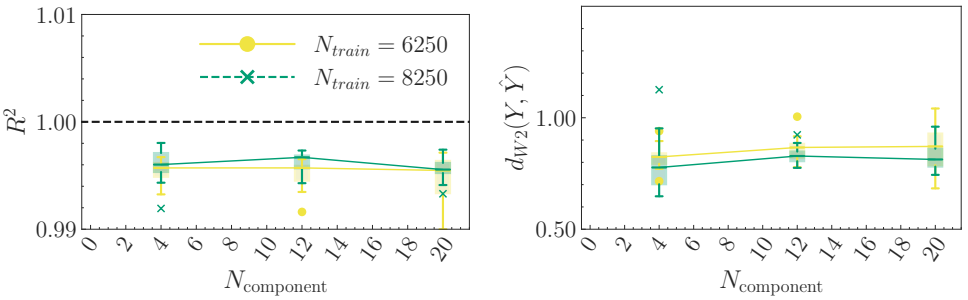
Figure 6.12: Study on the network architecture for four-layer networks. The top image shows the R^2 value for the mean of the conditional pdf of the tower bottom fore-aft DEL channel. The dashed line corresponds to an R^2 value of 0.99. The bottom image plots the d_{W2} values for the same channel with the dashed line corresponding to d_{W2} of 1.

fore-aft DEL channel is tested.

The number of components does not affect the estimation of the tower bottom fore-aft DEL mean. A slight improvement can be seen in Figure 6.13b with four components. The model, therefore, appears to be robust regarding the choice of the number of mixture components. In other words, it does not necessarily benefit from a large set of mixture components. MDN models in the remainder of the study are trained with four kernels.

Choice of training data size

This section shows the convergence study with respect to the number of training samples for the tower bottom fore-aft DEL. It is assumed that the same architecture can be



(a) R^2 values for the predicted mean

(b) d_{W2} for the predicted conditional pdf

Figure 6.13: Sensitivity of the MDN (30, 30, 50) surrogate to the number of mixture components.

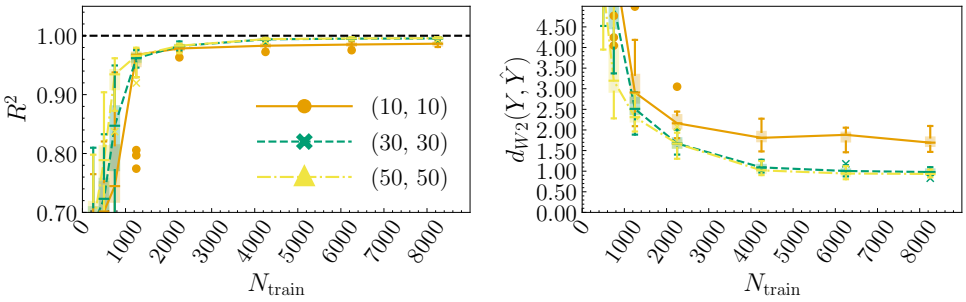
used to predict the remaining channels. Three networks for the MDNs are compared to test the robustness of the approach, as listed in Table 6.7. The MDNs contain four mixture elements. The rest of the hyperparameters are as specified in Table 6.6. In Figure 6.14, at each N_{train} value, the models are trained on 25 different subsets of the total training data space to capture the sensitivity of the model's fit to the choice of the training samples. The boxes reflect the variation in the R^2 values as a result of the choice of training data points. The boxes extend between the data's first ($Q1$) and third ($Q3$) quartile, and the horizontal line across the box indicates the median. The difference between $Q1$ and $Q3$ defines the interquartile range (IQR). The upper whisker extends to the largest data values within $1.5IQR$ above $Q3$. The lower whisker, similarly, extends to the lowest data point within $1.5IQR$ below $Q1$. Outliers are visible as dots beyond the whisker boundaries. Figure 6.14a is the R^2 value obtained from predicting the mean of the conditional pdf of the tower bottom fore-aft DEL, averaged over the test dataset. The mean in the BHawC reference is calculated using 44 realizations of the wind and wave fields. The diminishing size of the IQR as the number of samples grows is a combination of the increasing robustness of the model and the smaller variability in the test samples as fewer untrained samples remain in the dataset.

Notation	Number of layers	Number of nodes per layer
(10, 10)	2	10
(30, 30)	2	30
(50, 50)	2	50

Table 6.7: The MDN architectures considered for the convergence study.

MDN uses a neural network framework capable of inferring extremely complex underlying functions given sufficient data. In this case, MDN converges to consistent values of both R^2 of the conditional mean and d_{W2} above 4250 samples. It can be seen that MDN estimates are closer to the ground truth with larger networks of 30 or 50 nodes per layer. The performance of MDN(30, 30) and MDN(50, 50) is almost identical in this region, indicating good model robustness with respect to the width of the layer.

Figure 6.14b shows the normalized 2-Wasserstein distance between the predicted



(a) R^2 values for the predicted mean

(b) d_{W2} for the predicted conditional pdf

Figure 6.14: Convergence plots for the tower bottom fore-aft DEL channel. (a) Shows the convergence of the predicted mean as a function of the number of training samples for three MDN architectures. (b) Shows the normalized Wasserstein distance as a function of the training samples.

and reference pdf. The Wasserstein distance quantifies the similarities between the predictions and the reference. It is, thus, a good indicator of whether or not the surrogate can correctly estimate the variation in the target resulting from a combination of epistemic and aleatoric sources. The predicted pdf is based on 5000 realizations from the estimated Gaussian mixture in MDN. The reference is based on $n_{\text{seeds}} = 44$ BHawC/OrcaFlex realizations. Since 44 samples are insufficient to characterize the reference pdf fully, there is certainly an error associated with the d_{W2} values; therefore, d_{W2} cannot be expected to be zero in practice. Beyond 4250 samples, there is a small but marginal improvement in the d_{W2} values from MDN(30, 30) and MDN(50, 50).

In conclusion, the two-layered MDN surrogates – MDN(30, 30) and MDN(50, 50) – reach a degree of convergence in terms of d_{W2} already at 4250 samples. However, for subsequent sections, the MDN models will be trained with a dataset of 8250 points, as this provides slightly better predictions with only a slight increase in model fitting cost.

The choice of the number of layers and nodes in the neural network, as well as the number of mixture components, is based on a hyperparameter study, presented in Section 6.5.1.

6.5.2 Surrogate model validation

In this section, the performance of the MDN model is evaluated for the load channels listed in Table 6.5, on the selected test dataset presented in Section 6.4.1. Based on studies in Section 6.5.1, the values of hyperparameters used in training the models, in addition to Table 6.6, are listed in Table 6.8.

Table 6.9 provides a quantitative analysis of the model performance in terms of the average R^2 and d_{W2} values. The conditional mean is accurately captured by the MDN model with an R^2 exceeding 0.99 on the test dataset. The goodness of fit on the conditional distribution is evaluated using d_{W2}. Lower d_{W2} values indicate a smaller difference between the predicted and reference conditional distributions across the test database. As the d_{W2} values are normalized by the local reference standard deviation, the performance of the models across different load channels can be compared. MDN's performance remains consistently good on the tower top and blade targets. The tower bottom channel shows a

Table 6.8: List of hyperparameters used for training the MDN model for the final load prediction.

MDN hyperparameter	Value
Number of hidden layers	3
Width layer 1	30
Width layer 2	30
Width layer 3	50
Number of mixture components	4
Number of training samples	8250

larger deviation in the d_{W2} values, which is investigated further in Figure 6.15.

Table 6.9: Quantitative analysis of MDN model's predictions using d_{W2} and R^2 as evaluation metrics.

Model	Tower bottom FA	Tower top FA	Blade root edgewise	Blade root flapwise
d_{W2}	0.86	0.35	0.36	0.36
R^2_μ	0.99	0.99	0.99	0.99

Figure 6.15 shows the statistics of the conditional distribution of the DEL variation at the tower bottom fore-aft direction Since 44 seeds is a relatively small sample size to determine the true mean and standard deviation of the population, a gray area is highlighted in Figure 6.15a and Figure 6.15b to reflect the uncertainty in the reference values. For the mean, the 95% confidence interval (CI_t) is calculated with the t-distribution [184], assuming the response is normal. It is defined as,

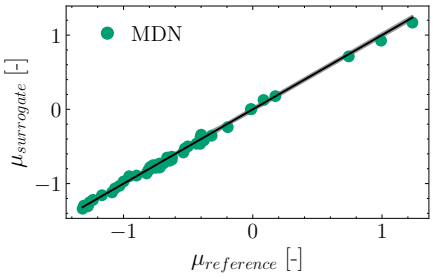
$$CI_t = \mu_{\text{reference}} \pm t \frac{\sigma_{\text{reference}}}{\sqrt{n_{\text{seeds}}}}, \tag{6.3}$$

where $\mu_{\text{reference}}$ is the mean and $\sigma_{\text{reference}}$ is the standard deviation calculated from the simulation samples. $n_{\text{seeds}} = 44$ is the number of seeds with which the simulations were repeated. t is the *t-score* for 95% confidence, given n_{seeds} samples from a normally distributed population. The bounds are similarly calculated using the χ^2 distribution for the standard deviation. The bounds are asymmetric as the χ^2 distribution is skewed. χ_L and χ_R are based on 5% and 95% tails of the χ^2 distribution. The true standard deviation, σ, is expected to lie between the bounds,

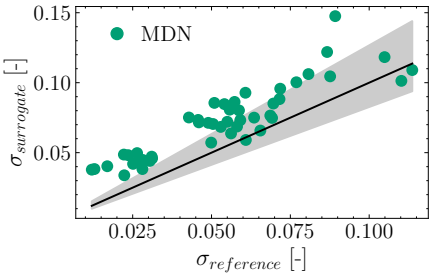
$$\sqrt{\frac{(n_{\text{seeds}}-1)}{\chi^2_L}}\sigma_{\text{reference}} \leq \sigma \leq \sqrt{\frac{(n_{\text{seeds}}-1)}{\chi^2_R}}\sigma_{\text{reference}}. \tag{6.4}$$

The individual pdfs are shown in Figure 6.15 for two site conditions. The reference BHawC realizations are plotted as histograms overlayed with kernel density estimate (KDE) plots generated from 5000 samples from the conditional pdf predicted by the surrogate model. The estimates include two sources of uncertainty. The first is the epistemic uncertainty of inferring a function from limited data. The second is due to the irreducible noise term (aleatoric), which is a part of the observed stochastic process. The subsequent plots assume the standard deviation of the combined uncertainty.

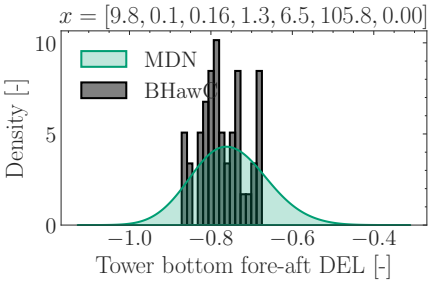
Figure 6.15a shows the predicted conditional mean ($\mu_{\text{surrogate}}$) of the normalized tower bottom fore-aft DEL as a function of the reference conditional mean ($\mu_{reference}$) derived from BHawC/Orcaflex simulations. As already indicated in Table 6.9, the R^2 values are greater than 0.99 for MDN, indicating an excellent fit. Similarly, the standard deviation derived from the surrogates ($\sigma_{\text{surrogate}}$) is plotted against the ground truth reference ($\sigma_{reference}$) in Figure 6.15b. Despite the slight overprediction of the standard deviation, MDN is able to capture the heteroscedastic trend in the data. Figure 6.15c corresponds to a below-rated velocity of 9.8 ms^{-1} and a wind-wave misalignment of 105°. Under these conditions, the BHawC reference is a short-tailed conditional pdf. Since MDN assumes a medium-tailed Gaussian mixture conditional, there is a tendency for the surrogate model to overestimate the standard deviation (Figure 6.15b). The reason for the tower bottom fore-aft DELs to be restricted between a very small range, resulting in such a short-tailed distribution, is not obvious and demands a deeper investigation into the behavior of the tower structure and control laws, which is beyond the scope of this study. A similar pattern is not observed in the other three load channels. Figure 6.15d corresponds to an example of a near-rated wind speed case, where the MDN predictions show a closer match to the reference conditional distribution.



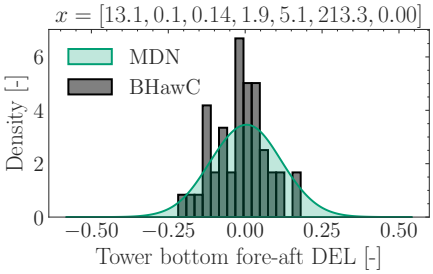
(a) Prediction of the mean DEL



(b) Prediction of the standard deviation of the DEL



(c) Conditional pdf below rated conditions

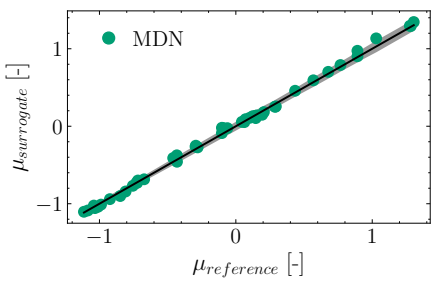


(d) Conditional pdf near rated conditions

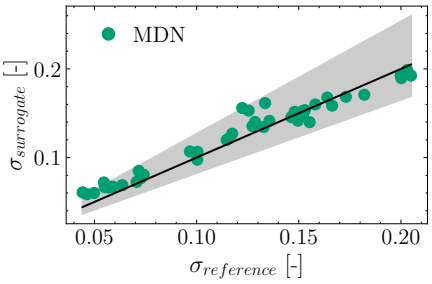
Figure 6.15: Load predictions using MDN for the tower bottom fore-aft DEL (normalized). Figure (a) shows the surrogate predicted conditional mean at the test locations vs. the conditional mean calculated using BHawC. Figure (b) shows the predicted and reference standard deviations of the conditional pdf. Figures (c) and (d) compare the conditional pdf plots between the surrogate and the simulation at below-rated and near-rated conditions, respectively. The values in vector x denote: $[U_{\text{ref}}, \alpha, TI, H_s, T_p, W_{\text{dir}}, Yaw]$ with units specified in Table 6.4.

A similar analysis is performed for the tower top fore-aft DEL channel in Figure 6.16. The conditional standard deviation estimated from the MDN surrogate are within

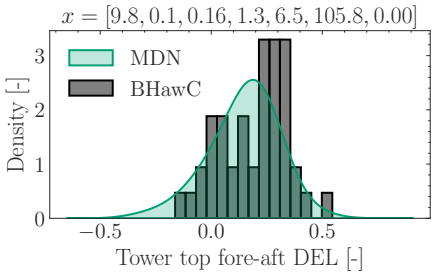
the error bounds of the small population assumption, indicating a very good fit.



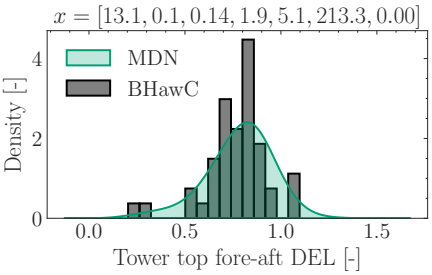
(a) Prediction of the mean DEL



(b) Prediction of the standard deviation of the DEL



(c) Conditional pdf below rated conditions



(d) Conditional pdf near rated conditions

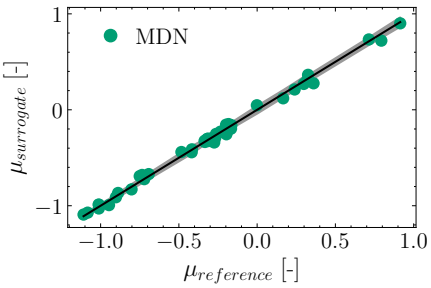
Figure 6.16: Load predictions using MDN for the tower top fore-aft DEL (normalized). Figure (a) shows the surrogate predicted conditional mean at the test locations vs. the conditional mean calculated using BHawC. Figure (b) shows the predicted and reference standard deviations of the conditional pdf. Figures (c) and (d) compare the conditional pdf plots between the surrogate and the simulation at below-rated and near-rated conditions, respectively. The values in vector x denote: $[U_{\text{ref}}, \alpha, TI, H_s, T_p, W_{\text{dir}}, Yaw]$ with units specified in Table 6.4.

Figure 6.17 and Figure 6.18 show the surrogate models' performance on the blade root flapwise and edgewise DEL, respectively. Similar to the tower top, the standard deviation and mean estimates from the MDN surrogate agree very well with the BHawC reference in both blade channels.

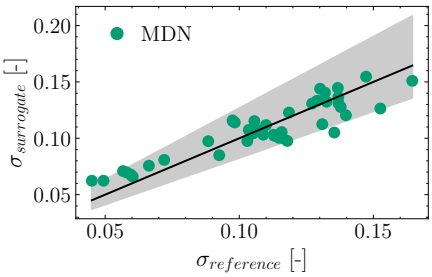
These results indicate that the surrogate model demonstrates a good level of reliability in accurately predicting the DELs with respect to the BHawC/OrcaFlex reference. Consequently, it is assumed that the model can be extended to other operating conditions without necessitating further verification.

6.5.3 Lifetime damage equivalent loads

The calculation of aggregated fatigue loads in onshore wind cases consists of binning the wind speed and scaling the loads at each bin by the probability of occurrence of the wind speed during the operating lifetime of the wind turbine. Floating wind turbine fatigue evaluations are more complex because, firstly, many more environmental parameters must be considered to characterize the site. Secondly, the bins need to be defined on a joint probability space. The process of selecting the right variables for the fatigue analysis and the size of the bins is not yet standardized and is a topic of ongoing research [91]. With a fast surrogate model, however, it is possible to account for every single observation in the

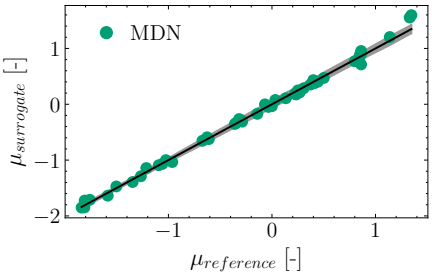


(a) Prediction of the mean DEL

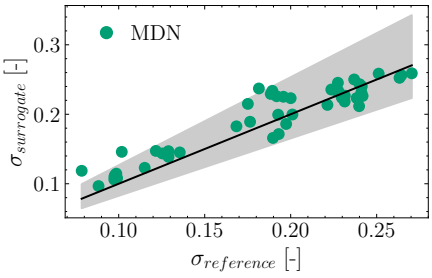


(b) Prediction of the standard deviation of the DEL

Figure 6.17: Load predictions using MDN for the blade root flapwise DEL (normalized). Figure (a) shows the surrogate predicted conditional mean at the test locations vs. the conditional mean calculated using BHawC. Figure (b) shows the predicted and reference standard deviations of the conditional pdf.



(a) Prediction of the mean DEL



(b) Prediction of the standard deviation of the DEL

Figure 6.18: Load predictions using MDN for the blade root edgewise DEL (normalized). Figure (a) shows the surrogate predicted conditional mean at the test locations vs. the conditional mean calculated using BHawC. Figure (b) shows the predicted and reference standard deviations of the conditional pdf.

previous years, without the need to lump probabilities or limit the number of variables. The joint probabilities of the sea states are, therefore, automatically accounted for.

In this section, the validated surrogate model from Section 6.5.2 is used to make probabilistic estimates of the equivalent loads (M_{eq}) for 10 million reference load cycles on the floating wind turbine structure. The site data are obtained from the ERA5 database for four sites with an approximate water depth of 100 m, namely Sud de la Bretagne II, Emerald, Hywind Scotland, and HIP Atlantic (Table B.1). Figure 6.19 provides an overview of the site conditions observed at the four selected sites. For simplicity, the foundation and mooring line design are assumed to be the same across the four sites. It is assumed that the difference in load distributions between the design in use and the site-optimized foundation will not be significant. The ERA5 hourly conditions are converted to 10-minute inputs by repeating each set of values six times. An alternative approach could be to draw the 10-minute values from a normal distribution with the hourly values as the mean and an assumed standard deviation. The observations below the cut-in and above the cut-out wind speed are excluded from the calculations. The site data consists of the average wind speed at 100m, significant wave height, peak spectral period, wind-wave misalignment

(converted to wave direction in OrcaFlex coordinates) and the shear exponent. The yaw misalignment values are sampled from a normal distribution with zero mean and a standard deviation of 2°. The turbulence intensity is calculated for each case based on the wind speed, assuming the IEC 61400-1 turbulence class C classification.

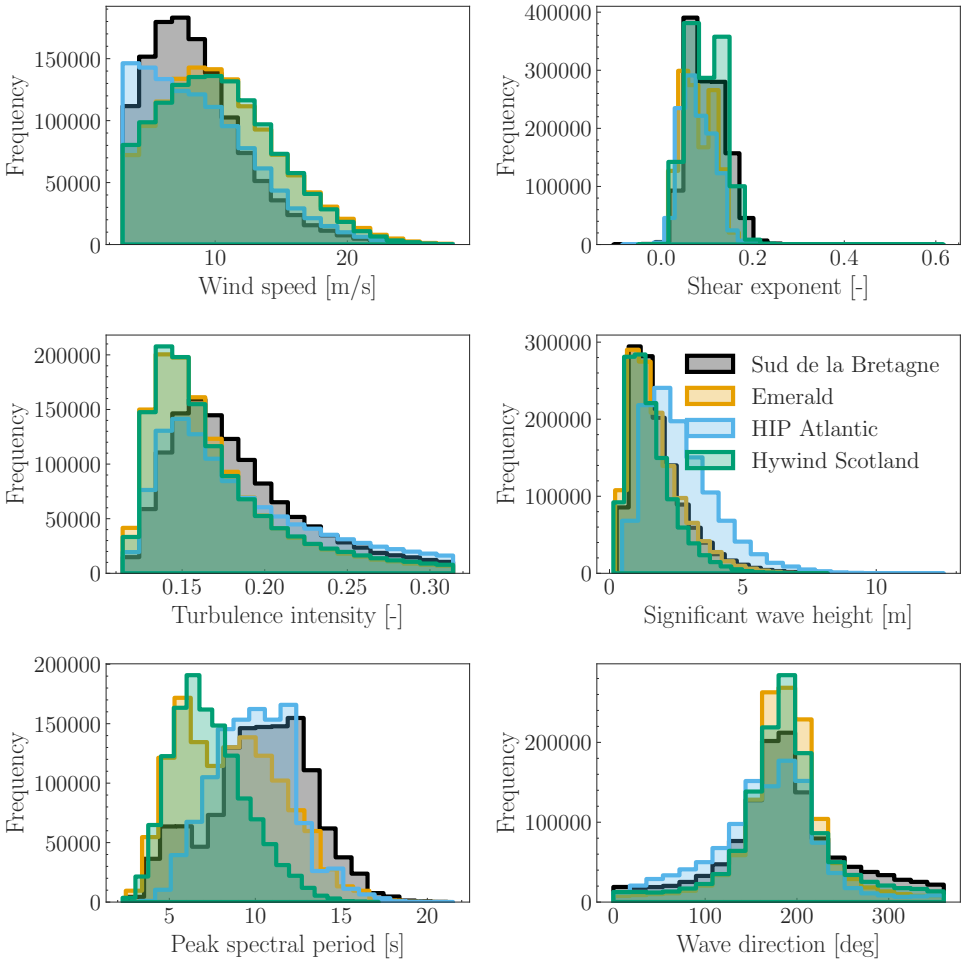


Figure 6.19: Comparison of the site conditions at the four floating wind sites considered in this study (Table B.1).

The value M_{eq}, represents the cyclic load amplitude which produces the equivalent lifetime damage given n_{eq} cycles of oscillation over $L = 25$ years. In Equation (6.5), M_i is the DEL for the i^{th} 10-minutes of operation, n_{ref} is the reference number of cycles per 10-minutes, set to 600. m is the Wöhler exponent, with part-specific values listed in Section 6.3. n_L is the number of 10-minute periods in L years. The loads do not have to be scaled as the probability of occurrence of each condition is equal. Since the surrogate has been validated in previous sections, it is assumed here that its predictions are accurate, and each M_i can be treated as a probabilistic output from the MDN model. From each M_i pdf, 500 samples

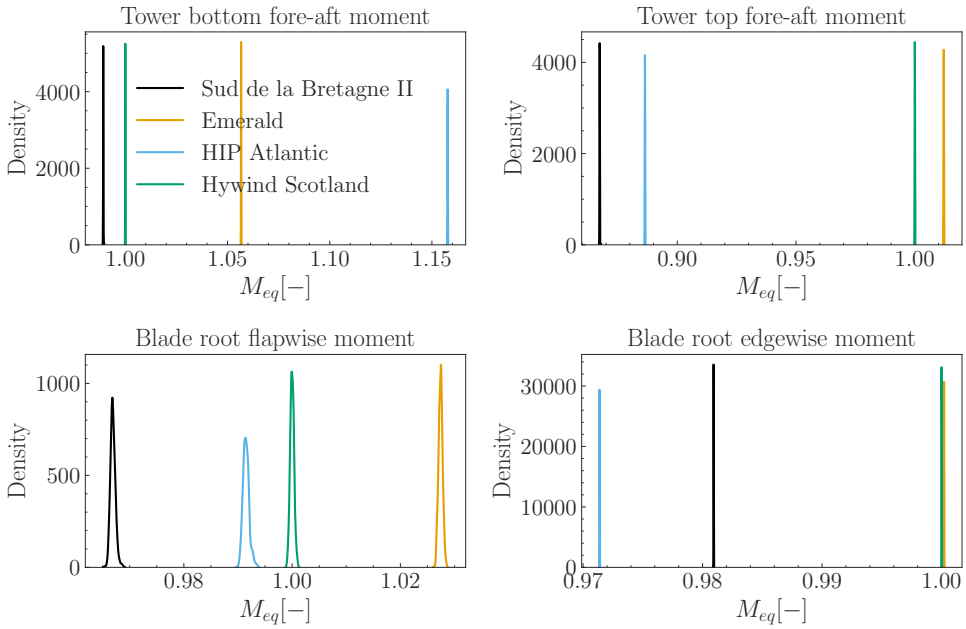


Figure 6.20: The 25-year normalized M_{eq} calculated for four sites at the tower bottom fore-aft direction (top-left), tower top fore-aft (top-right), blade root flapwise (bottom-left), and blade root edgewise (bottom-right) channels. The mean M_{eq} obtained at the Hywind Scotland site is used as the reference to normalize the loads at the remaining locations.

are drawn, resulting in a probabilistic estimation of M_{eq}. M_{eq} is defined as,

$$M_{\text{eq}} = \left(\frac{n_{\text{ref}}}{n_{\text{eq}}} \sum^{n_L} M_i^m \right)^{1/m} \tag{6.5}$$

where n_{eq} is 10^6 and n_{ref} is fixed to 600 oscillations per 10 minute period. The probabilistic M_{eq} value can be further used to calculate the stress reserve factor when re-designing the tower, or to calculate the fatigue damage during the structure's operating lifetime.

Figure 6.20 shows the kernel density estimate of the normalized M_{eq} values from the surrogate for the four selected sites. M_{eq} has been normalized by the average of the predicted M_{eq} values at the Hywind Scotland site for every channel. Firstly, it is interesting to note that the uncertainty in M_{eq} at each site is very small compared to the mean. This aligns with the law of large numbers, which states that for M_i with a mean μ and variance σ^2, the standard deviation of the average of the distribution of $(\sum M_i)$ decreases as $\sigma / \sqrt{n_L}$. Since n_L is in the order of 10^6, the standard deviation becomes extremely small as we get closer to the true mean. Even though the effect of M_i being raised to the power of m means that any variability in the sum is amplified, subsequently taking the m^{th} root has the opposite, damping effect. Therefore, the effect of the outliers is essentially nullified due to the averaging. It is important to note that this study considers only the statistical uncertainty arising from stochastic input sources. In practice, other sources of uncertainty

may contribute to the analysis [185]. For example, uncertainties related to the underlying joint distribution of site conditions represent another significant source of variability. Including these additional sources of uncertainty may introduce bias in the long-term mean, which is reflected as an uncertainty in the aggregated statistics of the outputs.

Secondly, the loads on different channels do not scale uniformly across sites. At the HIP Atlantic site, for instance, the cumulative tower bottom fore-aft moment is the highest, as shown in Figure 6.20. This is primarily due to the influence of the significant wave height, which is expected to have a larger impact on the tower bottom fatigue [8, 41, 43]. The marginal distribution of significant wave height at this site indicates a higher probability of larger waves compared to other locations, which supports the observed increase in tower bottom loads.

The distributions of wind speed, turbulence intensity, and significant wave height at the Emerald and Hywind Scotland sites (Figure 6.19) are nearly identical. This results in comparable tower top fore-aft and blade root edgewise damage. However, there remains a significant difference in blade root flapwise fatigue accumulation. This result is surprising, given that the loads at this location are primarily wind-driven. Nevertheless, it underscores the complexity of fatigue damage accumulation, which can yield different outcomes with minor variations in site conditions, even with respect to non-dominant variables.

6.6 Discussion

6.6.1 Summary

This chapter presents a methodology to develop probabilistic surrogate models for predicting floating offshore wind turbine fatigue loads for site analysis. The surrogate maps the environmental conditions from potential farm sites to the 10-minute damage equivalent loads experienced by a spar-type floating wind turbine. The main advantage of using probabilistic surrogates for this application is the ability to estimate conditional statistics with high accuracy to account for the statistical uncertainty resulting from the stochastic site conditions while minimizing the computational cost of training by avoiding seed repetitions. Based on the reanalysis data from the ERA5 database for several comparable floating sites, the surrogate model is used to propagate the statistical uncertainties to the 25-year fatigue loads on the wind turbine.

In this study, the analysis is performed on a spar-buoy floating foundation based on a modified Hywind-Scotland 6MW wind turbine. The damage equivalent loads are considered on critical locations on the tower and blades and are calculated using a coupled implementation of BHawC/OrcaFlex for training and validating the surrogate. The features characterizing a floating farm site and the appropriate ranges are defined. The probabilistic model considered in this study is the MDN, as it is flexible, robust, and interpretable, and has performed well for fixed bottom load emulation in the literature.

Since MDN is based on a neural network parametrization, several hyperparameters require tuning prior to training. Therefore, a hyperparameter study is conducted to determine the optimal neural network layout and the minimum number of training samples required to reach a high accuracy in terms of R^2_μ and d_{W2}. The conditional distribution predicted by the chosen model is validated on a set of 47 operating conditions, each

simulated with 44 random seeds in BHawC/Orcaflex to obtain a reference conditional distribution for each test case. The R^2 value for estimating the conditional mean is > 0.99 on all channels with the surrogate, indicating an excellent fit. The standard deviation of the conditional distribution is over-predicted by the model in the case of the tower bottom fore-aft moment, but within the range of uncertainty bounds for the tower top and blade root channels.

Finally, the validated surrogate model is used to make probabilistic estimates of the 25-year equivalent damage on the tower and blades for four different sites. Since the surrogate model is fast, load predictions can be made quickly on all observed site conditions without lumping or binning the sea states a priori. It is demonstrated that surrogate models can be powerful tools for site analysis, particularly for floating wind turbines, where the choice of variables and binning methods remains an open question. Additionally, using probabilistic surrogates like MDNs helps reduce bias in calculating the aggregate mean fatigue, as the conditional distributions are not always normally distributed.

6.6.2 Discussion and future work

This section provides a critical discussion of the study's results, along with practical considerations and limitations associated with the use of MDNs.

10-minute conditional DEL prediction

Given the stochastic nature of the site conditions, it is natural to model the 10-minute DEL response within a probabilistic framework. MDN is demonstrated in this study to be a reliable tool for modeling the conditional distribution of 10-minute DELs on the spar buoy floating wind turbine's tower and blades. MDN predictions are shown to remain robust across different network architectures and numbers of mixture components. The conditional means of the DELs are predicted with high accuracy, achieving an $R^2 = 0.99$. Additionally, the Wasserstein distance between the predicted and reference conditional distributions shows a strong match at the blade roots and tower top. However, at the tower bottom, the conditional standard deviation of the 10-minute fore-aft DEL is consistently over-predicted. It is corroborated by the relatively larger normalized Wasserstein distance value, indicating a bigger difference between the reference and predicted pdfs. Two main factors contribute to this: (i) the reference BHawC distributions are not converged at all simulated test locations with 44 random seeds. The tails of some distributions are not developed, resulting in short-tailed distributions that the MDN cannot easily capture; and (ii) the tower bottom fatigue is shown in the literature to have a stronger correlation to the hydrodynamic parameters, leading to higher noise in the data. As MDN is trained to minimize the negative log-likelihood, it is rewarded for predicting higher variance when there is less confidence.

Probabilistic lifetime DEL aggregation

The uncertainty in the aggregated lifetime fatigue loads due to stochastic inputs is found to be much smaller in scale compared to the mean. This results from summing the 10-minute DELs over a million occurrences, effectively nullifying the impact of the outliers. The use of a probabilistic surrogate that correctly captures the conditional distribution is still useful, as it minimizes the aggregation of error in the final response.

Notes on MDNs

MDNs, due to their flexibility in modeling the conditional response, are well-suited for the problem of probabilistic load estimation. One big advantage of the method is the ease of implementation and robustness, as demonstrated in this work. Compared to deterministic models that often assume a Gaussian response to determine the conditional mean, mixture models can account for skewness and multimodality, and improve the mean estimates. This is especially important for quantities like DELs, which may have non-Gaussian, heteroscedastic variations. MDNs scale well and are cost-effective to train compared to models that use Bayesian inference or variational inference [107].

MDNs without regularization can result in overfitting. Therefore, in this study, both L1 and L2 regularization are implemented. Secondly, MDNs rely on a stochastic optimizer that is sensitive to the initialization of the model parameters. Hence, a 10-fold cross-validation is recommended to ensure the optimizer is not stuck on a false minimum. As seen in the tower-bottom fore-aft channel, minimizing the negative log-likelihood can result in the over-prediction of the standard deviation of the conditional response when the underlying distribution is short-tailed. MDNs here are not restricted to strictly positive values; in some cases, the tails may also extend to negative values. A potential solution is to assume a lognormal distribution for the output. This can be done by directly predicting the parameters of a lognormal distribution during training or by transforming the output to a normal distribution before training.

Future work

Future studies could use such surrogates to identify optimal methods for grouping sea states in order to reduce the number of physics-based simulations required to achieve the same lifetime fatigue loads as using all observed site data. This type of analysis would be computationally impractical with an engineering tool, as it would require performing millions of simulations to establish a baseline reference. Surrogates offer an alternative for reducing the computational demands while maintaining accuracy. Surrogate models can also be used in this context to isolate combinations of sea states that produce the highest fatigue on the wind turbine structure. Furthermore, it is interesting to include other sources of uncertainty in the analysis of loads, such as introducing an uncertainty in the parameters defining the joint distribution of the site conditions. Once trained, probabilistic surrogate models can be used to propagate the different uncertainty sources to the loads to study the combined effect without additional costs. This approach opens new opportunities for integrating reliability-based decision-making into the design process.

7

Conclusions and future work

The chapter reflects on the principal findings, discussing their implications and summarizing the research contributions. It revisits the main research objectives and summarizes the limitations and assumptions of the various approaches. The recommendations for future work are presented.

7.1 Main outcomes

The research objectives introduced in Chapter 1 are revisited in the following section. The main research objective of the thesis is:

Research objective

RO To develop a workflow for training probabilistic surrogate models for quantifying the uncertainty in the 10-minute statistics of the load response of wind turbines for site analysis

These surrogates aim to reduce the computational costs associated with engineering simulations. This process allows for the quick assessment of the suitability of a new turbine type for a candidate offshore site, based on the environmental conditions at the site. The RNA suitability is justified if the site-specific emulated loads are below the certified design loads for the turbine type. In the early stages of the project, the site-specific emulated tower loads can serve as initial loads for the initial tower design. Although it is possible to apply scaling laws from site to site, the uncertainty may be large due to non-linearities. Hence, the surrogate models serve as alternatives. The problem with most interpolator objects used for siting purposes in offshore environments is the large amount of parameters which increases the dimensionality and explodes the amount of simulations. With surrogates, the solution space is generally sparse and broad; therefore, it can cover all possible environmental inputs more effectively.

In this work, rather than using deterministic surrogates that map point-to-point predictions, probabilistic surrogates are investigated. The motivation is to directly account for the stochastic nature of the inflow to quantify model uncertainty when predicting structural loads. Chapter 4 focuses on fixed-bottom wind turbine data and compares the performance of several surrogate models of varying complexity. Chapter 6 extends this methodology to floating wind turbines, where fatigue calculations typically require several thousand simulations, resulting in more than 5000 CPU hours per site. In the fixed-bottom case, Gaussian processes demonstrate strong capability, accurately capturing the underlying mappings even with a relatively small training set of roughly 500 quasi-random samples and making predictions in a matter of seconds. This reduces the total computational cost of training and prediction to only about 250 CPU hours. The floating wind turbine loads are mapped using MDNs trained on approximately 8000 simulations, which translates to 4000 CPU hours for generating the training data for a model that can be used on any site. Thereby, massively reducing the computational costs without impacting the accuracy.

The subsequent research questions are addressed in the following section.

RQ1

What are the leading sources of uncertainty in wind turbine load prediction?

Modeling complex structures like floating wind turbines numerically involves various sources of uncertainty, some of which are discussed in detail in this work.

- Stochastic environmental conditions, characterized by turbulent wind speed fluctuations and wave-structure interactions, subject wind turbine components to variable-amplitude cyclic loading, thereby inducing fatigue. In practice, load statistics are calculated in 10-minute data blocks, where detailed inflow variations are reduced to aggregate measures such as mean and standard deviation, introducing aleatoric uncertainty in the loads due to the loss of high-frequency information. Chapter 2.

- Engineering models are designed with many assumptions to simplify numerical analysis, introducing epistemic uncertainty in load calculations. Chapter 5 shows that performing one-to-one simulation validation can be critical when dealing with non-standard, complex wind turbine models.

- The information about certain site-specific variables, such as wind veer, detailed wind profiles, soil stiffness, and ocean currents, is often unavailable. Data-driven models approximate these quantities using assumed functional forms, introducing epistemic uncertainty.

- Assumptions about controller performance and tuning can significantly affect the conditional distributions of the predicted loads. For example, a bimodal distribution in the 10-minuteDEL is observed for the fixed-bottom case in Chapter 4. This behavior is attributed to the controller design, in which the wind turbine transitions between operating states at high wind speeds. In these simulations, although the mean wind speed provided to the simulation tool as input is within the operational range, instantaneous wind speeds may exceed the cut-out threshold for extended periods to trigger a change in turbine state. Such state transitions introduce additional uncertainty in the resulting fatigue loads. Fatigue magnitude is highly nonlinear and is governed primarily by the largest load ranges occurring within a given time interval. Consequently, the resulting DEL depends not only on the fraction of the time the wind turbine is operating but also on the temporal sequence of events, specifically whether the turbine undergoes a restart or a shutdown within the 10-minute window. Modern wind turbine controllers introduce hysteresis to encourage smooth switching between modes to limit loads and maximize the power output.

- Fatigue and extreme loads are often estimated through extrapolation of short-term simulations, which introduces inherent statistical (aleatoric) uncertainty, as noted in Chapter 6. In fatigue calculations, additional epistemic uncertainty arises from assumptions regarding the parameters of the joint distribution describing site conditions, as these parameters are themselves uncertain.

- The uncertainty in the parameters of the surrogate model due to insufficient data, stochastic optimization, and hyperparameter tuning is an important source of epistemic uncertainty.

When designing data-driven models, it is important to recognize the various sources of uncertainty, as they directly affect the reliability and accuracy of load predictions (Chapter 4, Chapter 6).

RQ2

What are the environmental conditions that most impact the fatigue loads in offshore wind turbines?

Feature selection is an essential preprocessing step for machine learning and data-driven modeling. A feature is any observed quantity or a derived value based on observations, serving as an input to the data-driven model. Training a surrogate with an extensive feature set or with redundant features can lead to increased training data needs, data storage, and computational and measurement costs. Environmental conditions such as wind speed, turbulence intensity, wind shear, wave height, spectral peak period, and wind–wave misalignment have an influence on fatigue.

In Chapter 4, the offshore wind turbine loads were trained with the mean wind speed, turbulence intensity, power law exponent, significant wave height, spectral peak period, and wave direction. During training, it was found that the wind parameters most affected the wind turbine DEL in this dataset. The hydrodynamic parameters for the analyzed dataset were not highly informative for fatigue load estimation in fixed-bottom offshore wind turbines. However, in sites with more prominent wave activity, these parameters remain relevant. The correlation study in Chapter 5, based on measurements from the TetraSpar demonstrator, highlighted wind speed standard deviation and significant wave height as key drivers of 10-minute DELs.

Furthermore, it is demonstrated that fatigue in each component is affected differently by changes in site conditions, with some components showing greater sensitivity than others. To accurately quantify these effects and account for correlations between site variables, a deeper investigation using global sensitivity analysis is warranted.

RQ3

How do probabilistic models compare in the case of onshore, fixed-bottom offshore, and floating wind turbines?

In Chapter 4, deterministic and probabilistic surrogate modeling methods are compared for predicting 10-minute DELs on fixed-bottom offshore wind turbines. The results show clear differences in predictive capability, robustness, and practical utility between the two classes of models.

Deterministic models such as polynomial regression, ANN, and GPR performed well in predicting the conditional mean of DELs, particularly in the simpler onshore case. In this case, high R^2_μ values were achieved with relatively small training datasets. However, their performance degraded in CASE-OFFSHORE, where additional hydrodynamic parameters increased data complexity and noise. Probabilistic models, in particular, MDNs, consistently matched or exceeded the conditional mean accuracy of deterministic models while also accurately predicting the conditional distributions.

The deterministic models inherently assume a normally distributed response with a deterministic relationship between site conditions and loads. This limitation arises when the conditional load distribution is heteroscedastic and non-Gaussian. Assuming a

normal distribution introduces a bias in the predictions that can aggregate as the loads are extrapolated to multiple years.

Probabilistic surrogate models allowed for training without seed repetitions, filtering, or data transformation. This contrasts with deterministic approaches, which are typically trained on averages over multiple seeds to account for variability in inflow conditions. By learning conditional distributions directly, probabilistic models reduce computational costs associated with dataset generation.

While GPR showed strong performance with small datasets, its computational scalability is limited due to matrix inversion costs. MDNs offered a balance between accuracy, scalability, and practical implementation effort, with relatively straightforward hyperparameter tuning compared to CGANs and CGPs.

In Chapter 6, MDNs were used to model a 6MW floating wind turbine. MDN predictions of the conditional mean were extremely accurate with $R^2 = 0.99$. This was an improvement on the fixed-bottom offshore case due to a combination of a larger network architecture, a bigger training dataset, and an advanced controller that minimized structural resonance. The conditional distributions exhibited a good agreement with the BHawC reference.

For offshore wind turbine site assessment and load modeling tasks where uncertainty quantification is important, probabilistic surrogates provide a more comprehensive and efficient alternative to deterministic methods. MDNs, in particular, demonstrated reliability and ease of use across onshore, offshore, and floating offshore cases, supporting their recommendation for future surrogate modeling tasks in offshore wind applications.

RQ4

Does uncertainty matter in long-term fatigue load evaluations?

In Chapter 6, the validated surrogate model is used to make probabilistic estimates of the 25-year equivalent damage on the tower and blades for four different sites. Since the surrogate model is fast, load predictions can be made quickly on all observed site conditions without lumping or binning the sea states a priori. It is found that the uncertainty in the aggregated lifetime fatigue loads due to stochastic inputs is much smaller in scale compared to the mean. This results from summing the 10-minute DELs over a million occurrences, effectively nullifying the impact of the outliers. Does this mean that probabilistic surrogates are not useful? It is demonstrated in Chapter 4 that using probabilistic surrogates provides robustness against the presence of noise in the data and a smaller computational cost for training. A probabilistic surrogate that accurately captures the conditional distribution remains useful, as it minimizes the aggregation of errors in the long-term load calculations. This is especially true when the loads are not always homoscedastic, normally distributed, as several widely used deterministic models assume a normal distribution of noise and model the approximate conditional mean of the response. Additionally, for floating wind turbines, where the choice of variables and binning methods is still an open question, surrogates provide a way to bypass the process from historical site conditions to load predictions without having to make assumptions on the bin size or joint site distribution parameters. Probabilistic surrogates can also help improve the process of lumping sea

states in a way that minimizes the errors in the fatigue calculations when simulations are later performed with numerical tools.

Quantifiable differences relative to standard approaches for calculating lifetime fatigue loads have not yet been assessed. As a next step, a DLC 1.2 simulation set should be generated, with loads scaled using available site data following conventional practice. The resulting lifetime fatigue loads can then be compared against those predicted by the surrogate model using the methodology presented in this thesis. Additionally, evaluating the errors introduced (separately) when only the conditional mean is considered, as well as the impact of varying the number of samples per case, and binning strategies, would enable a more systematic quantification of these effects.

7.2 Limitations and future work

The probabilistic surrogate model introduced in this work currently requires updates for every new wind turbine or for structural modifications on the same wind turbine. In other words, a new training dataset must be created for the model to be trained when predictions need to be made on a new wind turbine. Future steps would include generalizing the approach by including additional variables during training that define the structure. For instance, the geometric parameters of the tower, of the floating or fixed foundation, and the soil properties for fixed-bottom wind turbines. For floating wind turbines, the water depth and current information may be included in the future.

The surrogates in this work focus solely on the tower fore-aft and blade loads. Even though the tower side-side moment is not the biggest contributor to tower fatigue, it should be modeled in the future, as it is not damped by the aerodynamic forces and is more susceptible to changes in the hydrodynamic parameters.

This study considers only the statistical uncertainty arising from stochastic input sources. In practice, other sources of uncertainty may contribute to the analysis. For example, uncertainties related to the underlying joint distribution of site conditions represent another significant source of variability. Including these additional sources of uncertainty may introduce bias in the long-term mean, which is reflected as an uncertainty in the aggregated statistics of the outputs.

Future studies should account for the effects of upstream wakes in the inflow. Wind turbines modeled with dynamic wake inflow exhibit more realistic load predictions and may experience higher cumulative damage. As wake control strategies such as wake steering [186] and Helix [187] are being deployed in real wind farms, understanding their impact on structural loads has become an important research question. Incorporating wake information in the inflow would also enable the application of such surrogates in a wind farm context.

MDNs do not differentiate between epistemic and aleatoric uncertainty. In the next steps, models that differentiate between the two should be investigated in order to isolate the errors introduced by the surrogate. Combining the MDN output vector with Bayesian neural networks could enable the decomposition of predictive uncertainty into its constituent elements, as Bayesian neural networks treat the network weights as random variables with (typically) Gaussian priors. A second drawback of MDNs, also shown in

Chapter 4, is that in cases where the loads are very close to zero, the tails of the predicted conditional distributions by MDN can extend to values below zero. In the future, MDNs with the parameters of a lognormal distribution should be investigated, or conditions on the realizations should be introduced.

If operational data are available, similar surrogate models can be trained on SCADA input that maps to the loads. This process allows for 10-minute DEL calculations in the event of sensor failure based on lean measurements. Additionally, the models can be used to obtain DELs at critical and inaccessible locations where direct sensor placement is not possible. Both offer long-term fatigue monitoring opportunities, justifying the feasibility of upgrades or lifetime extension.

Instead of predicting load statistics, an extension of this study would include modeling the real-time variation of loads as a function of environmental input, as it provides a more comprehensive picture of the wind turbine's response. Purely data-driven methods generally fail in time series settings or are extremely difficult to train; therefore, a physics-based time series approach with a data-driven correction to introduce nonlinearity should be investigated. The time series model can be further extended to inverse problems, such as utilizing wind turbines as sensors.

A

Appendix A

A.1 Scatter plots of the load channels on the TetraSpar demonstrator

A

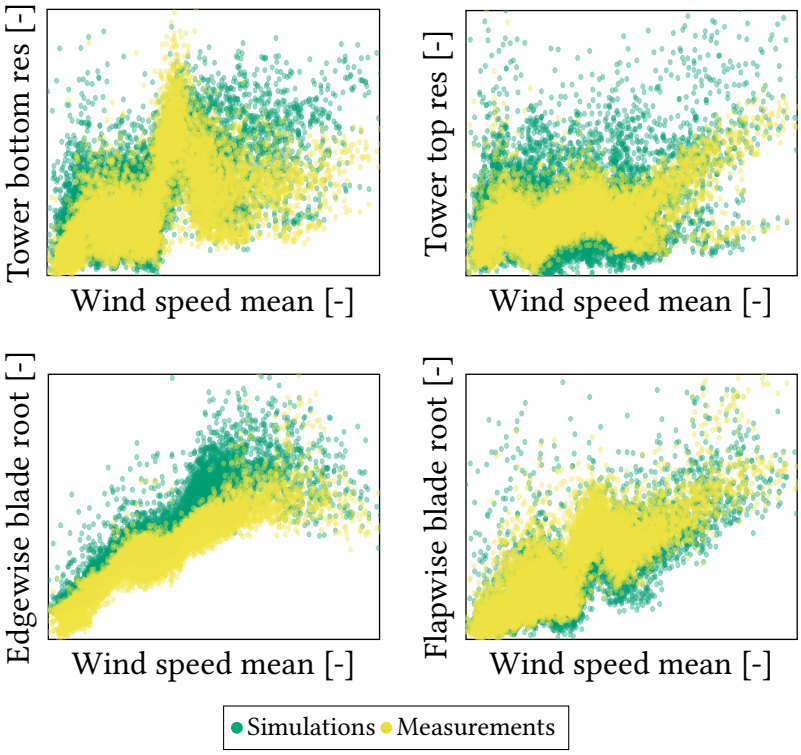


Figure A.1: Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. mean wind speed, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality.

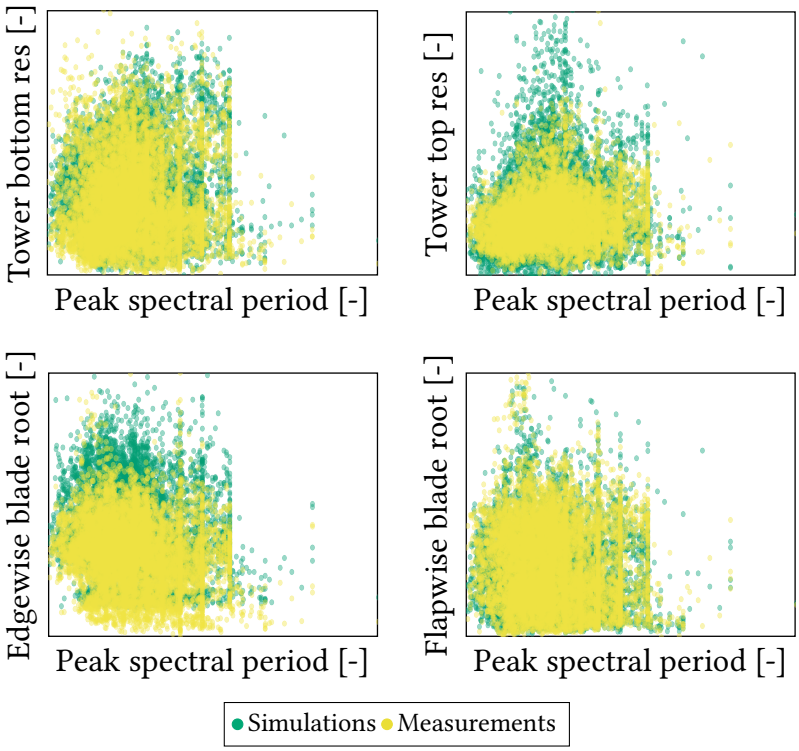


Figure A.2: Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. the wave peak spectral period, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality.

A

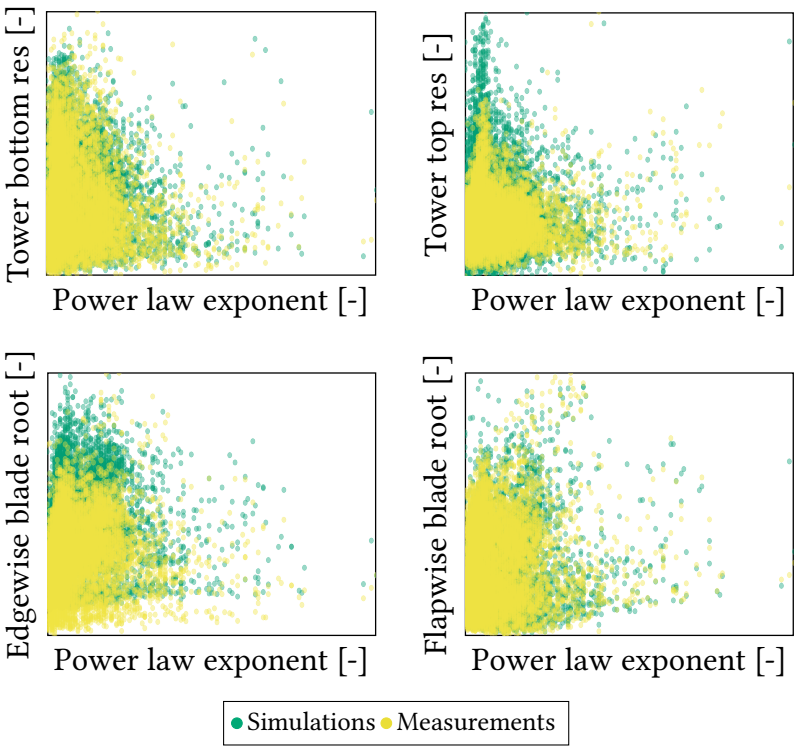


Figure A.3: Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. power law exponent, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality.

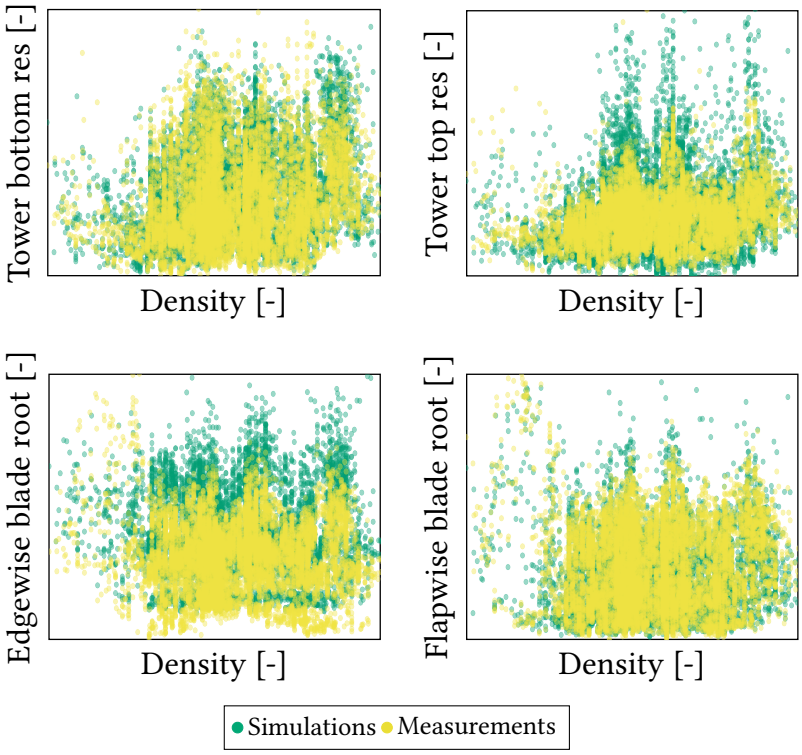


Figure A.4: Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. air density, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality.

A

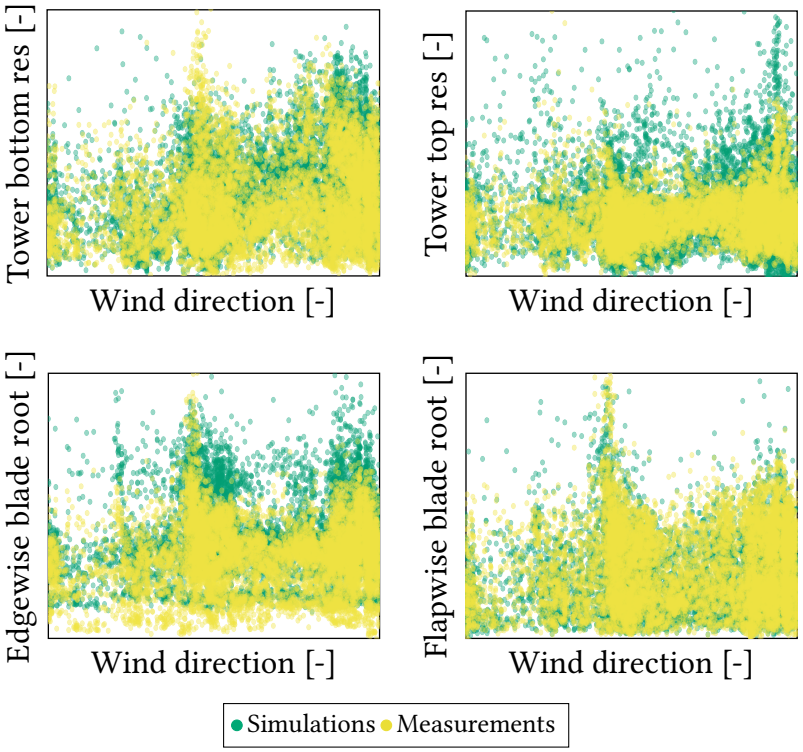


Figure A.5: Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. the wind direction, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality.

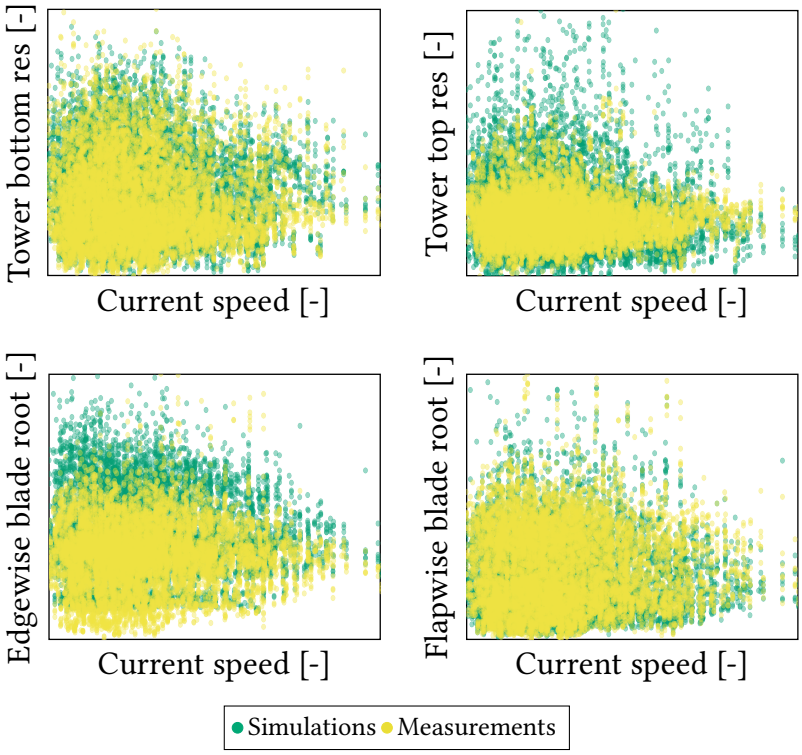


Figure A.6: Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. current speed, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality.

A

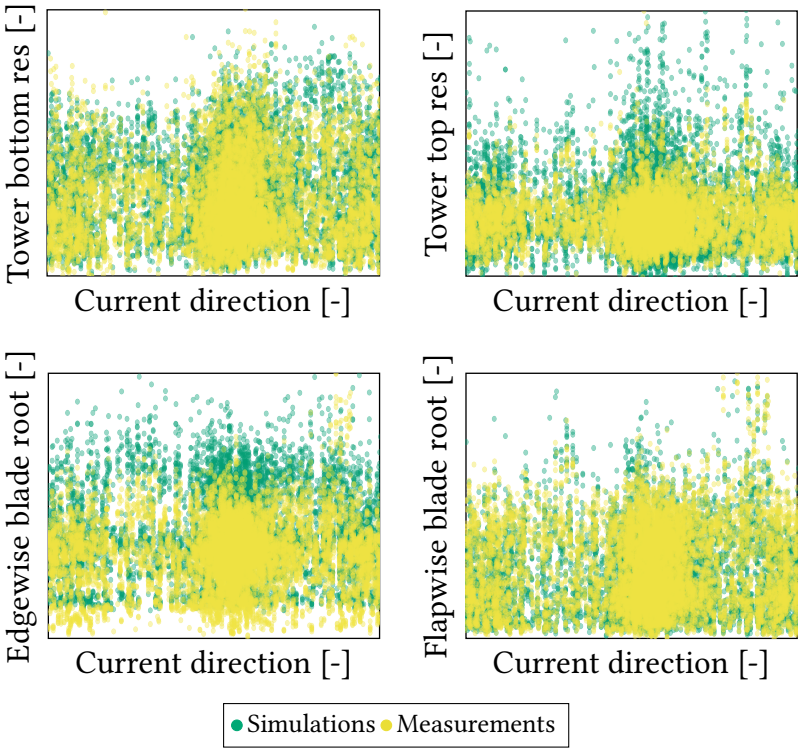


Figure A.7: Scatter plots of tower bottom moment DEL (top-left), tower top moment DEL (top right), edgewise blade root moment DEL (bottom left), and flapwise blade root moment DEL (bottom right) vs. current direction, comparing the measurements with the one-to-one simulations. The scales of the axes have been removed to preserve confidentiality.

A.2 Feature importance calculated using a surrogate model

In this section, a surrogate-based feature analysis method is presented. The data-driven surrogate uses XGBoost [188]. XGBoost calculates feature importance based on how much each feature contributes to the decision-making process of the model. The feature importance shown here is based on 'gain', which is the average improvement in loss each time a feature is used in a split. Gain is an indicator of how much better or worse the predictions get when a split is made, and is defined as [189],

$$Gain = \frac{1}{2}\left[\frac{G_L^2}{H_L+\lambda} + \frac{G_R^2}{H_R+\lambda} - \frac{(G_L+G_R)^2}{H_L+H_R+\lambda}\right] - \gamma \tag{A.1}$$

where γ is the regularization term for the number of leaves and λ is the regularization term for leaf weights. G and H are the first and second derivatives of the loss, respectively. The first term is the "score" on the new left leaf, the second term is the score on the new right leaf, and the third term is the score on the original leaf. After the trees are built, the gain for each feature is summed up across all trees. The highest gain, therefore, translates to the highest predictive value the feature has.

Based on this understanding, the feature importance calculated by XGBoost on the simulation and measurement values of loads is shown in the Figure A.10, Figure A.11, Figure A.8, and Figure A.9. The outcome aligns well with what has been observed in Section 5.4.1.

A

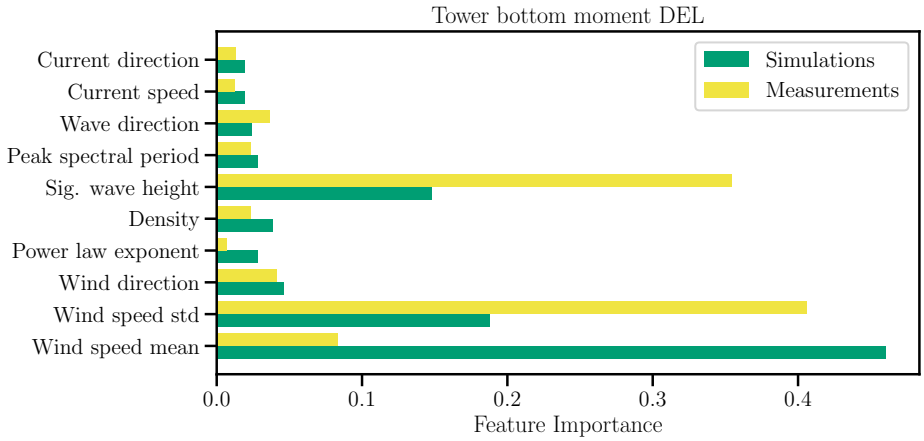


Figure A.8: Tower bottom moment DEL list of feature importance as calculated by XGBoost.

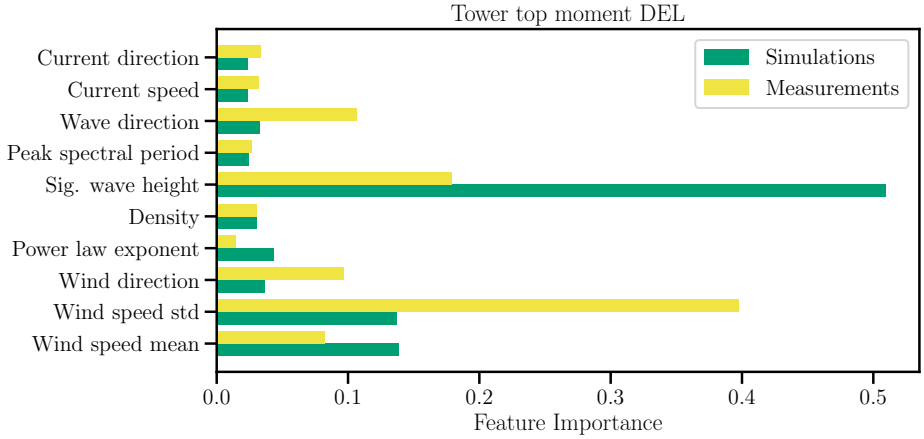


Figure A.9: Tower top moment DEL list of feature importance as calculated by XGBoost.

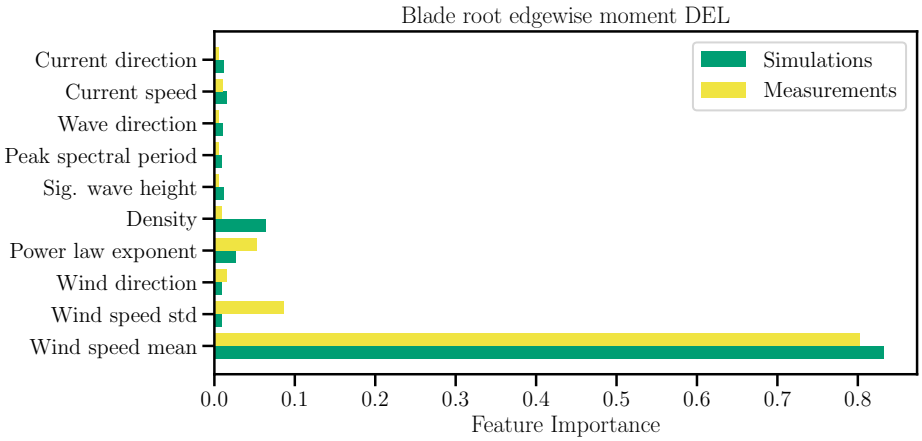


Figure A.10: Edgewise blade moment DEL list of feature importance as calculated by XGBoost.

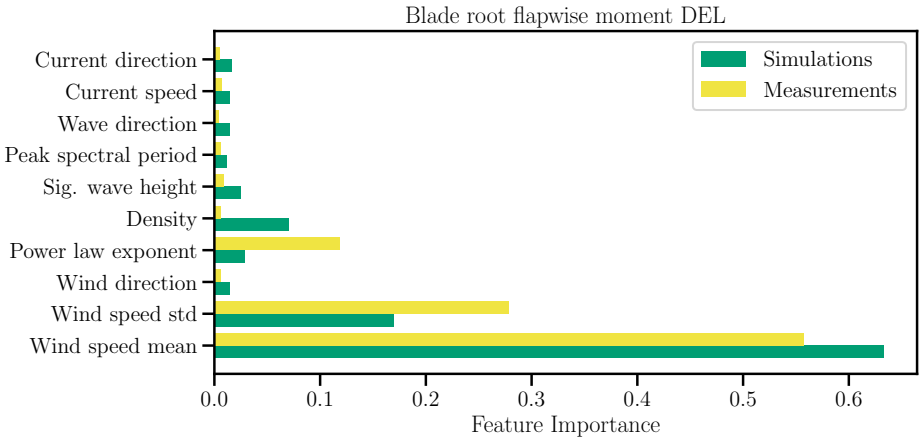


Figure A.11: Flapwise blade moment DEL list of feature importance as calculated by XGBoost.

B

Appendix B

B.1 ERA5 locations used for defining feature ranges

Table B.1 lists the locations used for defining the feature ranges in Section 6.3. The data are downloaded from the years 1979 to 2020. The database consists of the hourly average wind speeds at 10 m and 100 m, the significant wave height, spectral peak period, and wave direction. The shear law exponent is derived from the wind speed values, assuming a power law profile for the atmospheric boundary layer.

Table B.1: Description of the sites used for defining the feature ranges.

Site	Location [°]		ERA5 approx. location [°]		
	Latitude	Longitude	Latitude	Longitude	Depth [m]
Dyning [179]	58.218	17.860	58.00	17.75	141
Mareld [179]	58.161	10.575	58.25	10.50	233
Sørlige Nørdsjo Phase II [179]	56.783	4.918	56.75	5.00	60
Tetraspar	59.15	5.013	59.00	5.00	200
Utsira Nord Phase I [179]	59.276	4.540	59.00	4.50	273
Buchanan Deep [190]	57.45	−1.31	57.50	−1.25	100
West of Barra [19]	56.885	−7.947	57.00	−7.75	100
Gran Canaria [19]	27.75	−15.33	27.75	−15.00	200
Morro Bay [19]	35.083	−121.5	35.5	−121.75	870
Sud de la Bretagne II [179]	47.3247	-3.6594	47	-3.7	94
Emerald [179, 191]	51.3565	-8.0761	51.5	-8	90
Moneypoint Offshore I [179, 192]	52.519	-10.276	52.5	-10.5	102
HIP Atlantic [179]	63.6325	-16.3756	63.5	-16.5	98

B.2 Feature bounds

B.2.1 Significant wave height

The upper and lower limits for the significant wave height are defined as functions of the wind speed at hub height U_{ref}.

The upper limit is a quadratic function of the form:

$$H_{s_U} = -0.008U_{ref}^2 + 0.45U_{ref} + 5 \tag{B.1}$$

The lower limit is defined as:

$$H_{s_L} = 0.719e^{(0.0832U_{ref})} - e^{(0.04U_{ref})} \tag{B.2}$$

B.2.2 Peak spectral period

The peak spectral period range is designed to be a function of the significant wave height (which is in turn, a function of the wind speed at hub height). Scaling functions A, B and C are defined as,

$$A = a_1 + a_2 H_s^{a_3} \tag{B.3}$$

$$B = b_1 + b_2 e^{-b_3 H_s} \tag{B.4}$$

$$C = c_1 + c_2 e^{-c_3 H_s} \tag{B.5}$$

The scaling functions are used to define the upper bound T_{p_U} and lower bound T_{p_L} as,

$$T_{p_\mu} = e^{(A+0.5B)} \tag{B.6}$$

$$T_{p_L} = T_{p_\mu}(1-3\times\sqrt{(e^B-1)}) \tag{B.7}$$

$$T_{p_U} = T_{p_\mu}(1+3\times\sqrt{(e^C-1)}) \tag{B.8}$$

The coefficients used to fit the curve in this study are listed in Table B.2.

a_1	a_2	a_3	b_1	b_2	b_3	c_1	c_2	c_3
1.3	0.57	0.37	0.005	0.1	0.43	0.005	0.75	0.6

Table B.2: Tuning coefficients for defining the range functions for the spectral wave period.

Bibliography

References

[1] International Energy Agency. Electricity mid-year update. Technical report, International Energy Agency, 2024. URL https://www.iea.org/reports/electricity-mid-year-update-july-2024. [Accessed: 13-12-2025].

[2] International Energy Agency. World energy balances: Overview. Technical report, International Energy Agency, 2020. URL https://www.iea.org/reports/world-energy-balances-overview. [Accessed: 13-12-2025].

[3] International Energy Agency. Electricity 2024: Analysis and forecast to 2026. Technical report, International Energy Agency, 2024. URL https://www.iea.org/reports/electricity-2024. [Accessed: 13-12-2025].

[4] WindEurope. Wind energy in europe: 2023 statistics and the outlook for 2024–2030. Technical report, WindEurope, 2024. URL https://windeurope.org/data/products/wind-energy-in-europe-2023-statistics-and-the-outlook-for-2024-2030/. [Accessed: 13-12-2025].

[5] United Nations Statistics Division, World Health Organization, International Energy Agency, International Renewable Energy Agency, and World Bank. Tracking SDG 7: The Energy Progress Report 2024. Technical report, World Bank, Washington, DC, 2024. URL https://www.irena.org/Publications/2024/Jun/Tracking-SDG-7-The-Energy-Progress-Report-2024. [Accessed: 13-12-2025].

[6] Martin Hjelmeland and Jonas Kristiansen Nøland. Correlation challenges for North Sea offshore wind power: a Norwegian case study. *Scientific Reports*, 13(1):1–17, 2023. ISSN 20452322. doi: 10.1038/s41598-023-45829-2.

[7] International Renewable Energy Agency. Floating offshore wind outlook. Technical report, International Renewable Energy Agency, Abu Dhabi, 2024. URL https://www.irena.org/Publications/2024/Jul/Floating-offshore-wind-outlook. [Accessed: 13-12-2025].

[8] Emma C. Edwards, Anna Holcombe, Scott Brown, Edward Ransley, Martyn Hann, and Deborah Greaves. Evolution of floating offshore wind platforms: A review of at-sea devices. *Renewable and Sustainable Energy Reviews*, 183(June), 2023. ISSN 18790690. doi: 10.1016/j.rser.2023.113416.

[9] Stiesdal Offshore Technologies. The tetraspar full-scale demonstration project, 2024. URL https://www.stiesdal.com/offshore/the-tetraspar-full-scale-demonstration-project/. [Accessed: 03-09-2024].

[10] Wei Ting Hsu, Krish P. Thiagarajan, and Lance Manuel. Snap load criteria for mooring lines of a floating offshore wind turbine. In *ASME 2018 1st International Offshore Wind Technical Conference (IOWTC 2018)*, November 2018. ISBN 9780791851975. doi: 10.1115/IOWTC2018-1047.

[11] Shengjie Rui, Zefeng Zhou, Zhen Gao, Hans Petter Jostad, Lizhong Wang, Hang Xu, and Zhen Guo. A review on mooring lines and anchors of floating marine structures. *Renewable and Sustainable Energy Reviews*, 199(12):114547, 2024. ISSN 18790690. doi: 10.1016/j.rser.2024.114547.

[12] Blue H Engineering. Technology, 2025. URL https://www.bluehengineering.com/technology.html. [Accessed: 13-12-2025].

[13] SBM Offshore. Sbm offshore launches float4wind, 2024. URL https://www.sbmoffshore.com/newsroom/news-events/sbm-offshore-launches-float4windtm. [Accessed: 03-09-2024].

[14] Principle Power. The windfloat, 2024. URL https://www.principlepower.com/windfloat. [Accessed: 03-09-2024].

[15] Christopher Allen, Anthony Viselli, Habib Dagher, Andrew Goupee, Evan Gaertner, Nikhar Abbas, Matthew Hall, and Garrett Barter. Definition of the UMaine VolturnUS-S Reference Platform Developed for the IEA Wind 15- Megawatt Offshore Reference Wind Turbine. Technical report, National Renewable Energy Laboratory, Golden, CO, 2020. URL https://www.nrel.gov/docs/fy20osti/76773.pdf. [Accessed: 13-12-2025].

[16] HiveWind: semi-submersible floating platform for offshore wind, 2024. URL https://www.group.sener/en/project/hivewind-floating-platform-offshore-wind/. [Accessed 03-09-2024].

[17] Equinor ASA. Hywind floating offshore wind brochure, 2025. URL https://www.equinor.com/content/dam/statoil/documents/newsroom-additional-documents/news-attachments/brochure-hywind-a4.pdf. [Accessed: 13-12-2025].

[18] Jonas Bjerg Thomsen, Roger Bergua, Jason Jonkman, Amy Robertson, Nicole Mendoza, Cameron Brown, Christos Galinos, and Henrik Stiesdal. Modeling the tetraspar floating offshore wind turbine foundation as a flexible structure in OrcaFlex and OpenFAST. *Energies*, 14(23):1–14, 2021. ISSN 19961073. doi: 10.3390/en14237866.

[19] Fernando Vigara, Lara Cerdán, Rubén Durán, Sara Muñoz, Mattias Lynch, Siobhan Doole, Climent Molins, Pau Trubat, and Raúl Guanche. Corewind d1.2 design basis. Technical report, COREWIND Project, 2019. URL https://corewind.eu/wp-content/uploads/files/publications/COREWIND-D1.2-03-Design-Basis.pdf. [Accessed: 13-12-2025].

[20] Principle Power, Inc. Projects: Kincardine offshore wind farm - principle power, 2024. URL https://www.principlepower.com/projects/kincardine-offshore-wind-farm. [Accessed: 03-09-2024].

[21] Equinor ASA. World's first floating wind farm has started production, 2024. URL https://www.equinor.com/news/archive/worlds-first-floating-wind-farm-started-production. [Accessed: 03-09-2024].

[22] Klaus Schwab. *The Fourth Industrial Revolution*. Crown Publishing Group, New York, 2017. ISBN 9781524758868.

[23] Jason Jonkman. The new modularization framework for the FAST wind turbine CAE tool. In *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, Grapevine, 09–11 May 2013. AIAA. doi: 10.2514/6.2013-202.

[24] NREL. Openfast/openfast: Main repository for the nrel-supported openfast whole-turbine and fast.farm wind farm simulation codes., 2022. URL https://github.com/OpenFAST/openfast. [Accessed: 29-09-2022].

[25] Larsen, Torben J. Hansen, and Anders Melchior. *How 2 HAWC2*. Number 1597(ver. 3-1)(EN). 2007. ISBN 978-87-550-3583-6.

[26] Skjoldan, Peter Fisker. *Aeroelastic modal dynamics of wind turbines including anisotropic effects*. PhD thesis, DTU, 2011.

[27] J. Thirstrup Petersen, K. Thomsen, and Helge Aagaard Madsen. Local blade whirl and global rotor whirl interaction. Technical Report Risø Report 1067(EN), Forskningscenter Risø, 1998.

[28] Matthew Hall, Stein Housner, Daniel Zalkind, Pietro Bortolotti, David Ogden, and Garrett Barter. An open-source frequency-domain model for floating wind turbine design optimization. *Journal of Physics: Conference Series*, 2265(4):042020, 2022. ISSN 1742-6588. doi: 10.1088/1742-6596/2265/4/042020.

[29] Antonio Pegalajar-Jurado, Michael Borg, and Henrik Bredmose. An efficient frequency-domain model for quick load analysis of floating offshore wind turbines. *Wind Energy Science*, 3(2):693–712, 2018. ISSN 23667451. doi: 10.5194/wes-3-693-2018.

[30] Richard C. Lupton and Robin S. Langley. Improved linearised models of wind turbine aerodynamics and control system dynamics using harmonic linearisation. *Renewable Energy*, 135:148–162, 2019. ISSN 18790682. doi: 10.1016/j.renene.2018.11.067.

[31] E. Gonzalez, M. Reder, and J. J. Melero. SCADA alarms processing for wind turbine component failure detection. *Journal of Physics: Conference Series*, 753(7), 2016. ISSN 17426596. doi: 10.1088/1742-6596/753/7/072019.

[32] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, Saleh Abdalla, Xavier Abellan, Gianpaolo Balsamo, Peter Bechtold, Gionata Biavati, Jean Bidlot, Massimo Bonavita, Giovanna De Chiara, Per Dahlgren, Dick Dee, Michail Diamantakis, Rossana Dragani, Johannes Flemming, Richard Forbes, Manuel Fuentes, Alan Geer, Leo Haimberger, Sean Healy, Robin J.

Hogan, Elías Hólm, Marta Janisková, Sarah Keeley, Patrick Laloyaux, Philippe Lopez, Cristina Lupu, Gabor Radnoti, Patricia de Rosnay, Iryna Rozum, Freja Vamborg, Sebastien Villaume, and Jean Noël Thépaut. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020. ISSN 1477870X. doi: 10.1002/qj.3803.

[33] Fernando Echeverría, Fermín Mallor, and Unai San Miguel. Global sensitivity analysis of the blade geometry variables on the wind turbine performance. *Wind Energy*, 20 (9):1601–1616, 2017. doi: https://doi.org/10.1002/we.2111.

[34] Rui Teixeira, Alan O'Connor, and Maria Nogal. Probabilistic sensitivity analysis of offshore wind turbines using a transformed Kullback-Leibler divergence. *Structural Safety*, 81(6):101860, 2019. ISSN 01674730. doi: 10.1016/j.strusafe.2019.03.007.

[35] Kelsey Shaler, Amy Robertson, and Jason Jonkman. Sensitivity Analysis of Turbine Fatigue and Ultimate Loads to Wind and Wake Characteristics in a Small Wind Farm. *Wind Energy Science*, 8(12):25–40, 2023. doi: 10.5194/wes-8-25-2023.

[36] Amy N. Robertson, Kelsey Shaler, Latha Sethuraman, and Jason Jonkman. Sensitivity analysis of the effect of wind characteristics and turbine properties on wind turbine loads. *Wind Energy Science*, 4(3):479–513, 2019. ISSN 23667451. doi: 10.5194/wes-4-479-2019.

[37] Amy N. Robertson, Latha Sethuraman, Jason Jonkman, and Julian Quick. Assessment of wind parameter sensitivity on ultimate and fatigue wind turbine loads. In *Wind Energy Symposium, 2018*, Kissimmee, Florida, 2018. ISBN 9781624105227. doi: 10.2514/6.2018-1497. URL https://arc.aiaa.org/doi/abs/10.2514/6.2018-1728.

[38] Joey Velarde, Claus Kramhøft, and John Dalsgaard Sørensen. Global sensitivity analysis of offshore wind turbine foundation fatigue loads. *Renewable Energy*, 140: 177–189, 2019. ISSN 18790682. doi: 10.1016/j.renene.2019.03.055.

[39] Clemens Hübler, Cristian Guillermo Gebhardt, and Raimund Rolfes. Hierarchical four-step global sensitivity analysis of offshore wind turbines based on aeroelastic time domain simulations. *Renewable Energy*, 111:878–891, 2017. ISSN 18790682. doi: 10.1016/j.renene.2017.05.013.

[40] Markus Lerch, Mikel De-Prada-Gil, and Climent Molins. The influence of different wind and wave conditions on the energy yield and downtime of a Spar-buoy floating wind turbine. *Renewable Energy*, 136:1–14, 2019. ISSN 18790682. doi: 10.1016/j.renene.2018.12.096.

[41] Will Wiley, Jason Jonkman, Amy Robertson, and Kelsey Shaler. Sensitivity analysis of numerical modeling input parameters on floating offshore wind turbine loads. *Wind Energy Science*, 8(10):1575–1595, 2023. doi: 10.5194/wes-8-1575-2023.

[42] Yipin Wang, Qi Lu, Tiancheng Yao, Taofeng Yin, Yongsheng Zhao, Zhaolong Han, Yuwang Xu, and Zhiyu Jiang. Global sensitivity analysis of a semi-submersible floating wind turbine using a neural network fitting method. *Ocean Engineering*, 285 (August):1–14, 2023. ISSN 00298018. doi: 10.1016/j.oceaneng.2023.115351.

[43] Deepali Singh, Erik Haugen, Kasper Laugesen, Ayush Chauhan, and Axelle Viré. Data analysis of the TetraSpar demonstrator measurements. *IOP Journal of Physics: Conference Series*, 2767(062025), 2024. doi: 10.1088/1742-6596/2767/6/062025.

[44] Likhitha Ramesh Reddy, Dimitra Karystinou, Daniel Milano, John Walker, and Axelle Viré. Influence of wind-wave characteristics on floating wind turbine loads: A sensitivity analysis across different floating concepts. *Journal of Physics: Conference Series*, 2767(6):062030, 2024. doi: 10.1088/1742-6596/2767/6/062030.

[45] Nikolay Dimitrov, Mark C Kelly, Andrea Vignaroli, and Jacob Berg. From wind to loads: Wind turbine site-specific load estimation with surrogate models trained on high-fidelity load databases. *Wind Energy Science*, 3(2):767–790, 2018. doi: 10.5194/wes-3-767-2018.

[46] Nikolay Dimitrov. Surrogate models for parameterized representation of wake-induced loads in wind farms. *Wind Energy*, 22(10):1371–1389, 10 2019. ISSN 1095-4244. doi: 10.1002/we.2362.

[47] Kelsey Shaler, John Jasa, and Garrett E. Barter. Efficient Loads Surrogates for Waked Turbines in an Array. *Journal of Physics: Conference Series*, 2265(3), 2022. ISSN 17426596. doi: 10.1088/1742-6596/2265/3/032095.

[48] Juan Pablo Murcia, Pierre Elouan Réthoré, Nikolay Dimitrov, Anand Natarajan, John Dalsgaard Sørensen, Peter Graf, and Taeseong Kim. Uncertainty propagation through an aeroelastic wind turbine model using polynomial surrogates. *Renewable Energy*, 119:910–922, 2018. ISSN 18790682. doi: 10.1016/j.renene.2017.07.070.

[49] Laura Schröder, Nikolay Krasimirov Dimitrov, and David Robert Verelst. A surrogate model approach for associating wind farm load variations with turbine failures. *Wind Energy Science*, 5(3):1007–1022, 8 2020. ISSN 2366-7451. doi: 10.5194/wes-5-1007-2020.

[50] Nikolay Dimitrov and Anand Natarajan. From SCADA to lifetime assessment and performance optimization: how to use models and machine learning to extract useful insights from limited data. *Journal of Physics: Conference Series*, 1222:012032, 5 2019. ISSN 1742-6588. doi: 10.1088/1742-6596/1222/1/012032.

[51] Rui Teixeira, Alan O'Connor, Maria Nogal, Nandakumar Krishnan, and James Nichols. Analysis of the design of experiments of offshore wind turbine fatigue reliability design with Kriging surfaces. *Procedia Structural Integrity*, 5:951–958, 2017. ISSN 24523216. doi: 10.1016/j.prostr.2017.07.132.

[52] Manuel A. Vega and Michael D. Todd. A variational Bayesian neural network for structural health monitoring and cost-informed decision-making in miter gates. *Structural Health Monitoring*, 21(1):4–18, 2022. ISSN 17413168. doi: 10.1177/1475921720904543.

[53] René M.M. Slot, John D. Sørensen, Bruno Sudret, Lasse Svenningsen, and Morten L. Thøgersen. Surrogate model uncertainty in wind turbine reliability assessment. *Renewable Energy*, 151:1150–1162, 2020. ISSN 18790682. doi: 10.1016/j.renene.2019.11.101.

[54] J. McMorland, M. Collu, D. McMillan, and J. Carroll. Operation and maintenance for floating wind turbines: A review. *Renewable and Sustainable Energy Reviews*, 163 (May):112499, 2022. ISSN 18790690. doi: 10.1016/j.rser.2022.112499.

[55] Obafemi O. Olatunji, Paul A. Adedeji, Nkosinathi Madushele, and Tien Chien Jen. Overview of Digital Twin Technology in Wind Turbine Fault Diagnosis and Condition Monitoring. *Proceedings of 2021 IEEE 12th International Conference on Mechanical and Intelligent Manufacturing Technologies, ICMIMT 2021*, pages 201–207, 2021. doi: 10.1109/ICMIMT52186.2021.9476186.

[56] Jake Walker, Andrea Coraddu, Maurizio Collu, and Luca Oneto. Digital twins of the mooring line tension for floating offshore wind turbines to improve monitoring, lifespan, and safety. *Journal of Ocean Engineering and Marine Energy*, 8(1):1–16, 2022. ISSN 21986452. doi: 10.1007/s40722-021-00213-y.

[57] Farid K. Moghadam and Amir R. Nejad. Online condition monitoring of floating wind turbines drivetrain by means of digital twin. *Mechanical Systems and Signal Processing*, 162(December 2020):108087, 2022. ISSN 10961216. doi: 10.1016/j.ymssp.2021.108087.

[58] Hergen Pargmann, Dorthe Euhausen, and Robin Faber. Intelligent big data processing for wind farm monitoring and analysis based on cloud-Technologies and digital twins: A quantitative approach. *2018 3rd IEEE International Conference on Cloud Computing and Big Data Analysis, ICCCBDA 2018*, pages 233–237, 2018. doi: 10.1109/ICCCBDA.2018.8386518.

[59] Felix C. Mehlan, Jonathan Keller, and Amir R. Nejad. Virtual sensing of wind turbine hub loads and drivetrain fatigue damage. *Forschung im Ingenieurwesen / Engineering Research*, 87(1):207–218, 2023. doi: 10.1007/s10010-023-00627-0.

[60] Giovanni Rinaldi, Philipp R. Thies, and Lars Johanning. Current status and future trends in the operation and maintenance of offshore wind turbines: A review. *Energies*, 14(9), 2021. ISSN 19961073. doi: 10.3390/en14092484.

[61] Zhou, Xingguo, Tian, Yankang, Qin, Yi, Charitidis, Costas A., Milickovic, Tanja K., and Termine, Stefania. An advanced structural health monitoring iot platform for offshore wind turbine blades. *Manufacturing Rev.*, 12:12, 2025. doi: 10.1051/mfreview/2025008.

[62] Ao Zhang, Ming Li, and Lin Zhou. Structural health monitoring of offshore wind turbine based on online data-driven support vector machine. In *2018 IEEE 7th Data Driven Control and Learning Systems Conference (DDCLS)*, 2018. doi: 10.1109/DDCLS.2018.8516109.

[63] Agnes Marie Horn, Gudfinnur Sigurdsson, Ole Gabrielsen, Gaute Storhaug, Joar Håland, and George Jagite. Structural health monitoring of offshore structure by use of sensor data. Technical Report 2023-1241, Rev. 1, Petroleumstilsynet / DNV AS, Technology Centre Oslo Advisory, Stavanger, Norway, 2024.

[64] E. E. Ambarita, A. Karlsen, F. Scibilia, and A. Hasan. Industry 4.0 digital twins in offshore wind farms. *Wind Energy Science Discussions*, 2023:1–34, 2023. doi: 10.5194/wes-2023-108.

[65] Nikolay Dimitrov and Tuhfe Göçmen. Virtual sensors for wind turbines with machine learning-based time series models. *Wind Energy*, 25(9):1626–1645, 2022. doi: https://doi.org/10.1002/we.2762.

[66] Fuxing Li, Luxi Li, and You Peng. Research on Digital Twin and Collaborative Cloud and Edge Computing Applied in Operations and Maintenance in Wind Turbines of Wind Power Farm. *Advances in Transdisciplinary Engineering*, 17:80–92, 2021. ISSN 23527528. doi: 10.3233/ATDE210263.

[67] Kuan Cheng Lin, Jyh Yih Hsu, Hao Wei Wang, and Mu Yen Chen. Early fault prediction for wind turbines based on deep learning. *Sustainable Energy Technologies and Assessments*, 64(2):103684, 2024. ISSN 22131388. doi: 10.1016/j.seta.2024.103684.

[68] Peng Lu, Ning Zhang, Lin Ye, Ershun Du, and Chongqing Kang. Advances in model predictive control for large-scale wind power integration in power systems. *Advances in Applied Energy*, 14(4):100177, 2024. ISSN 26667924. doi: 10.1016/j.adapen.2024.100177.

[69] Jingjie Xie, Hongyang Dong, and Xiaowei Zhao. Data-driven torque and pitch control of wind turbines via reinforcement learning. *Renewable Energy*, 215(June):118893, 2023. ISSN 18790682. doi: 10.1016/j.renene.2023.06.014.

[70] Xiuxing Yin and Xiaowei Zhao. Data driven learning model predictive control of offshore wind farms. *International Journal of Electrical Power and Energy Systems*, 127(10):106639, 2021. ISSN 01420615. doi: 10.1016/j.ijepes.2020.106639.

[71] Hafiz Ghazali Bin Muhammad Amri, Daniela Marramiero, Deepali Singh, Jan Willem Van Wingerden, and Axelle Viré. Data-driven time series forecasting of offshore wind turbine loads. *Journal of Physics: Conference Series*, 2767(5), 2024. ISSN 17426596. doi: 10.1088/1742-6596/2767/5/052060.

[72] Carlos E. García, David M. Prett, and Manfred Morari. Model Predictive Control: Theory and Practice a Survey. *Automatica*, 25(3):335–348, 1989. doi: https://doi.org/10.1016/0005-1098(89)90002-2.

[73] M. Baudino Bessone, D. Singh, T. Kalimeris, E. Bachynski-Polić, and A. Viré. Surrogate-assisted optimization of floating wind turbine substructure. *Journal of Physics: Conference Series*, 2767(6), 2024. ISSN 17426596. doi: 10.1088/1742-6596/2767/6/062032.

[74] Xiao Wang, Haibin Li, Hexi Baoyin, Shuqin Han, and Changchun Bao. Aerodynamic optimization of wind turbine blades via surrogate-assisted deep reinforcement learning. *Physics of Fluids*, 37(4):047141, 04 2025. ISSN 1070-6631. doi: 10.1063/5.0256389.

[75] Rui Meng, Xuhui Chen, Liangliang Chen, Nenggang Xie, Lu Wang, and Bofeng Xu. Aerodynamic and structural co-optimization of offshore wind turbine blades using a novel adaptive surrogate-based optimization method. *Ocean Engineering*, 340:122291, 2025. ISSN 0029-8018. doi: 10.1016/j.oceaneng.2025.122291.

[76] Thanasis Barlas, Néstor Ramos-García, Georg Raimund Pirrung, and Sergio González Horcas. Surrogate-based aeroelastic design optimization of tip extensions on a modern 10 MW wind turbine. *Wind Energy Science*, 6(2):491–504, 2021. doi: 10.5194/wes-6-491-2021.

[77] Guo Nai-Zhi, Zhang Ming-Ming, and Li Bo. A data-driven analytical model for wind turbine wakes using machine learning method. *Energy Conversion and Management*, 252(10), 2022. ISSN 01968904. doi: 10.1016/j.enconman.2021.115130.

[78] S. Ashwin Renganathan, Romit Maulik, Stefano Letizia, and Giacomo Valerio Iungo. Data-driven wind turbine wake modeling via probabilistic machine learning. *Neural Computing and Applications*, 34(8):6171–6186, 2022. ISSN 14333058. doi: 10.1007/s00521-021-06799-6. URL https://doi.org/10.1007/s00521-021-06799-6.

[79] Nikolaos Bempedelis, Filippo Gori, Andrew Wynn, Sylvain Laizet, and Luca Magri. Data-driven optimisation of wind farm layout and wake steering with large-eddy simulations. *Wind Energy Science*, 9(4):869–882, 2024. ISSN 23667451. doi: 10.5194/wes-9-869-2024.

[80] G. V. Iungo, C. Santoni-Ortiz, M. Abkar, F. Porté-Agel, M. A. Rotea, and S. Leonardi. Data-driven Reduced Order Model for prediction of wind turbine wakes. *Journal of Physics: Conference Series*, 625(1), 2015. ISSN 17426596. doi: 10.1088/1742-6596/625/1/012009.

[81] Balthazar Arnoldus Maria Sengers, Gerald Steinfeld, Paul Hulsman, and Martin Kühn. Validation of an interpretable data-driven wake model using lidar measurements from a field wake steering experiment. *Wind Energy Science*, 8(5):747–770, 2023. ISSN 23667451. doi: 10.5194/wes-8-747-2023.

[82] Jaime Liew, Riccardo Riva, Mikkel Friis-Møller, and Tuhfe Göçmen. Wind Farm Control Optimisation Under Load Constraints Via Surrogate Modelling. *Journal of Physics: Conference Series*, 2767(9), 2024. ISSN 17426596. doi: 10.1088/1742-6596/2767/9/092039.

[83] Jannie Sønderkær Nielsen, Lindsay Miller-Branovacki, and Rupp Carriveau. Probabilistic and risk-informed life extension assessment of wind turbine structural components. *Energies*, 14(4):1–16, 2021. ISSN 19961073. doi: 10.3390/en14040821.

[84] Ding Peng Liu, Giulio Ferri, Taemin Heo, Enzo Marino, and Lance Manuel. On long-term fatigue damage estimation for a floating offshore wind turbine using a surrogate model. *Renewable Energy*, 225(2):120238, 2024. ISSN 18790682. doi: 10.1016/j.renene.2024.120238.

[85] Deepali Singh, Richard P. Dwight, and Axelle Vire. Probabilistic surrogate modeling of damage equivalent loads on onshore and offshore wind turbines using mixture density networks. *Wind Energy Science*, 9(3):1–28, 2024.

[86] Deepali Singh, Erik Haugen, Kasper Laugesen, Richard P. Dwight, and Axelle Viré. Data-driven probabilistic surrogate model for floating wind turbine lifetime damage equivalent load prediction. *Wind Energy Science*, 10(12):2865–2888, 2025. ISSN 2366-7451. doi: 10.5194/wes-10-2865-2025.

[87] International Electrotechnical Commission. Wind energy generation systems – part 3-1: Design requirements for fixed offshore wind turbines. Technical Report IEC 61400-3-1:2019, International Electrotechnical Commission, Geneva, Switzerland, 2019.

[88] International Electrotechnical Commission. Wind energy generation systems – part 3-2: Design requirements for floating offshore wind turbines. Technical Report IEC TS 61400-3-2:2019, International Electrotechnical Commission, Geneva, Switzerland, 2019.

[89] Philippe J. Couturier and Peter F. Skjoldan. Implementation of an advanced beam model in BHawC. *Journal of Physics: Conference Series*, 1037(6):0–10, 2018. ISSN 17426596. doi: 10.1088/1742-6596/1037/6/062015.

[90] Tristan De Lataillade, Wenchao Yu, Maxime Pallud, and Matteo Capaldo. Seahowl: Partitioned multiphysics and multifidelity modelling of wind turbines with monolithically coupled elastodynamics. *Journal of Physics: Conference Series*, 2767(5):052051, 2024. doi: 10.1088/1742-6596/2767/5/052051.

[91] Francesco Papi and Alessandro Bianchini. A new perspective on offshore wind turbine certification using high performance computing. *Journal of Physics: Conference Series*, 2767(5):052008, 2024. doi: 10.1088/1742-6596/2767/5/052008.

[92] Bretagne Ocean Power. Floating wind — bretagne ocean power, 2025. URL https://bretagneoceanpower.fr/en/bretagne-mre/floating-wind/. [Accessed: 13-02-2025].

[93] National Renewable Energy Laboratory (NREL). Turbsim user's guide: Revised february 2007 for version 1.21. Technical report, National Renewable Energy Laboratory, 2007. [Accessed: 13-12-2025].

[94] C. Kaimal, J, Y. Izumi Wyngaard, and O. R. Coté. Spectral characteristics of surface layer turbulence over the sea. *Quarterly Journal of the Royal Meteorological Society*, 98(417):563–589, 1972. doi: https://doi.org/10.1002/qj.49709841707.

[95] Jakob Mann. Wind field simulation. *Probabilistic Engineering Mechanics*, 13(4): 269–282, 1998. ISSN 02668920. doi: 10.1016/s0266-8920(97)00036-2.

[96] R. C. MacCamy and R. A. Fuchs. Wave forces on piles: A diffraction theory. Technical memorandum no. 69, Beach Erosion Board, U.S. Army Corps of Engineers, Washington, D.C., USA, 1954.

[97] J.M.J. Journée and W.W. Massie. *Offshore Hydromechanics*. Delft University of Technology, Delft, The Netherlands, 2001. URL https://ocw.tudelft.nl/wp-content/uploads/OffshoreHydromechanics_Journee_Massie.pdf. [Accessed: 13-12-2025].

[98] Willard J. Pierson Jr. and Lionel Moskowitz. A proposed spectral form for fully developed wind seas based on the similarity theory of s. a. kitaigorodskii. *Journal of Geophysical Research (1896-1977)*, 69(24):5181–5190, 1964. doi: https://doi.org/10.1029/JZ069i024p05181.

[99] K. Hasselmann, T. P. Barnett, E. Bouws, H. Carlson, D. E. Cartwright, K. Eake, J. A. Euring, D. E. Hasselmann, P. Kruseman, A. Meerburg, P. Mullen, D. J. Olbers, K. Richter, W. Sell, and H. Walden. Measurements of wind-wave growth and swell decay during the joint north sea wave project (jonswap). Technical Report Supplement A(8), Deutsches Hydrographisches Institut, Hamburg, Germany, 1973.

[100] *Simplified Double Peak Spectral Model For Ocean Waves*, volume The Fourteenth International Offshore and Polar Engineering Conference of *International Ocean and Polar Engineering Conference*, 05 2004.

[101] Lin Li, Sverre Haver, and Nikita Berlin. Assessment of operational limits: Effects of uncertainties in sea state description. *Marine Structures*, 77(November 2020):102975, 2021. ISSN 09518339. doi: 10.1016/j.marstruc.2021.102975.

[102] Masanori Matsuishi and Tatsuo Endo. Fatigue of metals subjected to varying stress. In *Preliminary Proceedings of the Kyushu District Meeting, The Japan Society of Mechanical Engineers*, pages 37–40, Fukuoka, Japan, 1968.

[103] Hermann G. Matthies. Quantifying uncertainty: Modern computational representation of probability and application. In *Extreme Man-Made and Natural Hazards in Dynamics of Structures*, pages 105–135, Dordrecht, 2007. Springer Netherlands. ISBN 978-1-4020-5656-7.

[104] Armen Der Kiureghian and Ove Dalager Ditlevsen. Aleatoric or epistemic? does it matter? *Structural Safety*, 31(2):105–112, 2009. ISSN 0167-4730. doi: 10.1016/j.strusafe.2008.06.020.

[105] Jaime Liew and Gunner Chr. Larsen. How does the quantity, resolution, and scaling of turbulence boxes affect aeroelastic simulation convergence? *Journal of Physics: Conference Series*, 2265(3):032049, 2022. doi: 10.1088/1742-6596/2265/3/032049.

[106] Nikolay Dimitrov and Anand Natarajan. From SCADA to lifetime assessment and performance optimization: How to use models and machine learning to extract useful insights from limited data. *Journal of Physics: Conference Series*, 1222(1), 2019. ISSN 17426596. doi: 10.1088/1742-6596/1222/1/012032.

[107] David M. Blei, Alp Kucukelbir, and Jon D. McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, 2017. doi: 10.1080/01621459.2017.1285773.

[108] Yibo Yang and Paris Perdikaris. Conditional deep surrogate models for stochastic, high-dimensional, and multi-fidelity systems. *Computational Mechanics*, 64(2):417–434, 2019. ISSN 01787675. doi: 10.1007/s00466-019-01718-y.

[109] Thomas Kneib. Beyond mean regression. *Statistical Modelling*, 13(4):275–303, 2013. ISSN 1471082X. doi: 10.1177/1471082X13494159.

[110] Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. In *Proceedings of the 2nd International Conference on Learning Representations (ICLR)*, 2014.

[111] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer New York, 2006. ISBN 9780387310732. doi: 10.1007/978-3-030-57077-4_11.

[112] Xujia Zhu and Bruno Sudret. Emulation of stochastic simulators using generalized lambda models. *SIAM/ASA Journal on Uncertainty Quantification*, 9(4):1345–1380, 2021. doi: 10.1137/20M1337302.

[113] Laura Schröder, Nikolay Krasimirov Dimitrov, David Robert Verelst, and John Aasted Sørensen. Wind turbine site-specific load estimation using artificial neural networks calibrated by means of high-fidelity load simulations. *Journal of Physics: Conference Series*, 1037(6):062027, jun 2018. doi: 10.1088/1742-6596/1037/6/062027.

[114] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA, 2006. ISBN 978-0-262-18253-9.

[115] Luis David Avendaño-Valencia, Imad Abdallah, and Eleni Chatzi. Virtual fatigue diagnostics of wake-affected wind turbine via Gaussian process regression. *Renewable Energy*, 170:539–561, 2021. ISSN 18790682. doi: 10.1016/j.renene.2021.02.003.

[116] Xuan Li and Wei Zhang. Probabilistic fatigue evaluation of floating wind turbine using combination of surrogate model and copula model. *AIAA Scitech 2019 Forum*, 2019. doi: 10.2514/6.2019-0247.

[117] Xuan Li and Wei Zhang. Long-term fatigue damage assessment for a floating offshore wind turbine under realistic environmental conditions. *Renewable Energy*, 159:570–584, 2020. ISSN 0960-1481. doi: 10.1016/j.renene.2020.06.043.

[118] Georgios Gasparis, Wai Hou Lio, and Fanzhong Meng. Surrogate models for wind turbine electrical power and fatigue loads in wind farm. *Energies*, 13(23):6360, 2020. doi: 10.3390/en13236360.

[119] Rad Haghi and Curran Crawford. Data-driven surrogate model for wind turbine damage equivalent load. *Wind Energy Science*, 9(11):2039–2062, 2024. doi: 10.5194/wes-9-2039-2024.

[120] Kenneth Thomsen. The statistical variation of wind turbine fatigue loads. In B. Maribo Pedersen, editor, *IEA Joint Action on Wind Turbine Fatigue*, pages 1–15. Technical University of Denmark, Department of Fluid Mechanics, 2000. Proceedings of the 5th IEA Expert Meeting on Wind Turbine Fatigue, 25–26 October 1999.

[121] X. Zhu and B. Sudret. Replication-based emulation of the response distribution of stochastic simulators using generalized lambda distributions. *International Journal for Uncertainty Quantification*, 10(3):249–275, 2020. doi: 10.1615/Int.J.UncertaintyQuantification.2020033029.

[122] Rad Haghi, Cassidy Stagg, and Curran Crawford. Wind Turbine Damage Equivalent Load Assessment Using Gaussian Process Regression Combining Measurement and Synthetic Data. *Energies*, 17(2), 2024. ISSN 19961073. doi: 10.3390/en17020346.

[123] Jason M. Jonkman, Sandy Butterfield, Walt Musial, and George Scott. Definition of a 5-MW Reference Wind Turbine for Offshore System Development. Technical Report NREL/TP-500-38060, National Renewable Energy Laboratory (NREL), Golden, CO, 2009. URL https://www.nrel.gov/docs/fy09osti/38060.pdf.

[124] Imad Abdallah, Christos Lataniotis, and Bruno Sudret. Parametric hierarchical kriging for multi-fidelity aero-servo-elastic simulators —application to extreme loads on wind turbines. *Probabilistic Engineering Mechanics*, 55:67–77, 2019. ISSN 0266-8920. doi: https://doi.org/10.1016/j.probengmech.2018.10.001.

[125] Deepali Singh, Richard P. Dwight, Kasper Laugesen, Laurent Beaudet, and Axelle Viré. Probabilistic surrogate modeling of offshore wind-turbine loads with chained gaussian processes. *Journal of Physics: Conference Series*, 2265(3):032070, may 2022. doi: 10.1088/1742-6596/2265/3/032070.

[126] Jorge R. Vergara and Pablo A. Estévez. A review of feature selection methods based on mutual information. *Neural Computing and Applications*, 24(1):175–186, 2014. ISSN 09410643. doi: 10.1007/s00521-013-1368-0.

[127] Isabelle Guyon. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, 2003.

[128] Shahram Hanifi, Xiaolei Liu, Zi Lin, and Saeid Lotfian. A critical review of wind power forecasting methods: Past, present and future. *Energies*, 13(15):3764, 2020. doi: 10.3390/en13153764.

[129] Andrew Kusiak and Wenyan Li. Estimation of wind speed: A data-driven approach. *Journal of Wind Engineering and Industrial Aerodynamics*, 98(10-11):559–567, 2010. ISSN 01676105. doi: 10.1016/j.jweia.2010.04.010.

[130] Yuan Zeng, Yuzhuo Hu, Haosen Cao, Fei Wang, and Qichao Chen. Data-Driven Assessment of Wind Power Available Frequency Modulation Capacity. *IEEE Power and Energy Society General Meeting*, (2):1–5, 2024. ISSN 19449933. doi: 10.1109/PESGM51994.2024.10688720.

[131] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656, 1948.

[132] Mona Azadkia and Sourav Chatterjee. A simple measure of conditional dependence. *The Annals of Statistics*, 49(6):3070–3102, 2021. doi: 10.1214/21-AOS2073.

[133] Kevin P. Murphy. *Probabilistic Machine Learning: An Introduction*. MIT Press, Cambridge, MA, 2022. ISBN 978-0-262-54304-0.

[134] Ciyou Zhu, Richard H. Byrd, Peihuang Lu, and Jorge Nocedal. Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. *ACM Trans. Math. Softw.*, 23(4):550–560, dec 1997. ISSN 0098-3500. doi: 10.1145/279232.279236.

[135] Alan D. Saul, James Hensman, Aki Vehtari, and Neil D. Lawrence. Chained Gaussian Processes. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pages 1431–1440. PMLR, 2016.

[136] Ping Jiang, Qi Zhou, and Xinyu Shao. Surrogate-model-based design and optimization. In *Surrogate-Model-Based Design and Optimization*, Springer Tracts in Mechanical Engineering, pages 135–236. Springer, 2020. ISBN 978-981-15-0731-1. doi: 10.1007/978-981-15-0731-1_7.

[137] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015.

[138] Christopher M. Bishop. Mixture density networks. Technical Report NCRG/94/004, Neural Computing Research Group, Aston University, 1994.

[139] Grégoire Montavon, Geneviève B. Orr, and Klaus-Robert Müller, editors. *Neural networks: Tricks of the trade*. Springer Berlin, Heidelberg, 2 edition, 2012. ISBN 978-3-642-35288-1. doi: 10.1007/978-3-642-35289-8.

[140] Yuan Yao, Lorenzo Rosasco, and Andrea Caponnetto. On early stopping in gradient descent learning. *Constructive Approximation*, 26(2):289–315, Aug 2007. ISSN 1432-0940. doi: 10.1007/s00365-006-0663-2.

[141] Andrew Y. Ng. Feature selection, L1 vs. L2 regularization, and rotational invariance. In *Proceedings of the 21st International Conference on Machine Learning (ICML)*, ICML '04, pages 78–85, New York, NY, USA, 2004. ACM. doi: 10.1145/1015330.1015435.

[142] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, 2020. ISSN 15577317. doi: 10.1145/3422622.

[143] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *arXiv preprint*, 2014. URL https://arxiv.org/abs/1411.1784.

[144] Han Zhang, Tao Xu, and Hongsheng Li. StackGAN: Text to Photo-Realistic Image Synthesis with Stacked Generative Adversarial Networks. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 5908–5916. IEEE Computer Society, 2017. doi: 10.1109/ICCV.2017.629.

[145] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 2223–2232. IEEE, 2017. doi: 10.1109/ICCV.2017.244.

[146] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5967–5976. IEEE, 2017. doi: 10.1109/CVPR.2017.632.

[147] Cameron Fabbri et al. Conditional wasserstein generative adversarial networks, 2018. URL https://arxiv.org/abs/1802.02279.

[148] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein gan, 2017. URL https://arxiv.org/abs/1701.07875.

[149] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C. Courville. Improved Training of Wasserstein GANs. In *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)*, pages 5767–5777. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper/2017/file/892c3b1c6dccd52936e27cbd0ff683d6-Paper.pdf.

[150] Cédric Villani. *Optimal Transport: Old and New*. Springer, Berlin, Heidelberg, 2009. ISBN 978-3-540-71050-9. doi: 10.1007/978-3-540-71050-9.

[151] Henning Petzka, Asja Fischer, and Denis Lukovnikov. On the regularization of wasserstein GANs. In *International Conference on Learning Representations (ICLR)*, 2018. URL https://openreview.net/forum?id=Hk5p6k-RZ.

[152] Li-Toong Yap. Probabilistic regression of wind turbine loads using conditional generative adversarial networks. Master's thesis, Delft University of Technology, 2023. URL https://resolver.tudelft.nl/uuid:e86ab65e-3b80-4949-9666-a858ccecbb3e.

[153] Cédric Villani. The Wasserstein distances. In *Optimal Transport: Old and New*, pages 93–111. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. ISBN 978-3-540-71050-9. doi: 10.1007/978-3-540-71050-9_6.

[154] Gabriel Peyré and Marco Cuturi. Computational optimal transport: With applications to data science. *Foundations and Trends in Machine Learning*, 11(5-6):355–607, 2019. ISSN 1935-8237. doi: 10.1561/2200000073.

[155] Aaditya Ramdas, Nicolas Garcia, and Marco Cuturi. On wasserstein two-sample testing and related families of nonparametric tests, 2015. URL https://arxiv.org/abs/1509.02237.

[156] Pietro Bortolotti, Helena Canet Tarrés, Katherine Dykes, Karl Merz, Latha Sethuraman, David Verelst, and Frederik Zahle. Systems engineering in wind energy – wp2.1 reference wind turbines. Technical report, IEA Wind TCP Task 37, National Renewable Energy Laboratory (NREL), May 2019.

[157] WindEurope. Wind energy in Europe: 2020 Statistics and the outlook for 2021–2025. Industry report, WindEurope, 2021. URL https://windeurope.org/data/products/wind-energy-in-europe-2020-statistics-and-the-outlook-for-2021-2025/. [Accessed: 12-12-2025].

[158] Ilya M. Sobol. On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, 7(4):86–112, 1967. ISSN 0041-5553. doi: https://doi.org/10.1016/0041-5553(67)90144-9.

[159] William J. Welch, Robert J. Buck, Henry P. Wynn, Jerome Sacks, Toby J. Mitchell, and Max D. Morris. Screening, predicting, and computer experiments. *Technometrics*, 34(1):15–25, 1992. ISSN 15372723. doi: 10.2307/1269548.

[160] Michalis K. Titsias. Variational learning of inducing variables in sparse gaussian processes. In *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 5 of *Proceedings of Machine Learning Research*, pages 567–574, 2009.

[161] Ali Nouriani and Hamed Moradi. Smooth switching in power control of wind turbines using a combination strategy of hysteresis and modified middle regions. *Sustainable Energy Technologies and Assessments*, 37:100585, 2020. ISSN 2213-1388. doi: https://doi.org/10.1016/j.seta.2019.100585.

[162] Christian Jourdain. AEP Increase Solution, August 2017. URL https://drudu6g9smo13.cloudfront.net/wp-content/uploads/2017/09/OM-Workshop_SiemensGamesa_ChristianJourdain-TROCAR.pdf. Siemens Gamesa Renewable Energy, presentation at OM Workshop; accessed: 12 December 2025.

[163] Ioannis Andrianakis and Peter G. Challenor. The effect of the nugget on gaussian process emulators of computer models. *Computational Statistics & Data Analysis*, 56 (12):4215–4228, 2012. ISSN 0167-9473. doi: https://doi.org/10.1016/j.csda.2012.04.020.

[164] Amy Robertson, Latha Sethuraman, and Jason M. Jonkman. Assessment of wind parameter sensitivity on extreme and fatigue wind turbine loads. In *2018 Wind Energy Symposium*. American Institute of Aeronautics and Astronautics, 2018. doi: 10.2514/6.2018-1728.

[165] Stian H. Sørum, George Katsikogiannis, Erin E. Bachynski-Polić, Jørgen Amdahl, Ana M. Page, and Rasmus T. Klinkvort. Fatigue design sensitivities of large monopile offshore wind turbines. *Wind Energy*, 25(10):1684–1709, 2022. ISSN 10991824. doi: 10.1002/we.2755.

[166] Michael Borg, Morten Walkusch Jensen, Scott Urquhart, Morten Thøtt Andersen, Jonas Bjerg Thomsen, and Henrik Stiesdal. Technical definition of the tetraspar demonstrator floating wind turbine foundation. *Energies*, 13(18), 2020. ISSN 19961073. doi: 10.3390/en13184911.

[167] Srinivas Guntur, Jason Jonkman, Ryan Sievers, Michael A. Sprague, Scott Schreck, and Qi Wang. A validation and code-to-code verification of FAST for a megawatt-scale wind turbine with aeroelastically tailored blades. *Wind Energy Science*, 2(2): 443–468, 2017. ISSN 23667451. doi: 10.5194/wes-2-443-2017.

[168] Pieter Jaime Matthijs Bussemakers. Validation of aero-hydro-servo-elastic load and motion simulations in BHawC/OrcaFlex for the Hywind Scotland floating offshore wind farm. Master's thesis, Norwegian University of Science and Technology, 2020. URL https://hdl.handle.net/11250/2780185.

[169] Orcaflex help — orcina ltd., 2025. URL https://www.orcina.com/webhelp/OrcaFlex/Default.htm. Accessed: 12 December 2025.

[170] Guy McCauley, Hugh Wolgamot, Paul H. Taylor, and Jana Orszaghova. Validation of wave radar measurements from the TetraSpar Demonstrator floating wind turbine. *Ocean Engineering*, 302(March):117627, 2024. ISSN 00298018. doi: 10.1016/j.oceaneng.2024.117627.

[171] Knut Torsethaugen and Sverre Haver. Simplified double peak spectral model for ocean waves. *Proceedings of the International Offshore and Polar Engineering Conference*, pages 76–84, 2004. ISSN 10986189.

[172] V. Arramounet, C. E.De Winter, N. Maljaars, S. Girardin, and H. Robic. Development of coupling module between BHawC aeroelastic software and OrcaFlex for coupled dynamic analysis of floating wind turbines. *Journal of Physics: Conference Series*, 1356(1):1–15, 2019. ISSN 17426596. doi: 10.1088/1742-6596/1356/1/012007.

[173] Equinor ASA. HyWind Scotland Pilot Park Decommissioning Programme. Technical report, Equinor ASA, Scotland, 2022. URL https://marine.gov.scot/data/decommissioning-programme-hywind-scotland-pilot-park.

[174] Marco Masciola, Amy Robertson, Jason Jonkman, and Frederick Driscoll. Investigation of a fast–orcaflex coupling module for integrating turbine and mooring dynamics of offshore floating wind turbines. Technical Report NREL/CP-5000-52896, National Renewable Energy Laboratory, Golden, CO, United States, 2011. URL https://www.osti.gov/biblio/1029021.

[175] A. Björck. Aerforce: Subroutine package for unsteady blade-element/momentum calculations. Technical Report TN 7-5, Flygtekniska Försöksanstalten (FFA), 2000.

[176] Rune Rubak and J. T. Petersen. Monopile as part of aeroelastic wind turbine simulation code. In *Proceedings of the Copenhagen Offshore Wind Conference*, pages 1–10, 2005. URL http://wind.nrel.gov/public/SeaCon/Proceedings/Copenhagen.Offshore.Wind.2005/documents/papers/Design_calculations_and_risks/R.Rubak_Monopile_as_Part_of_Aeroelastic_WindTurbineSim.pdf.

[177] G. Stewart, M. Lackner, L. Haid, D. Matha, J. Jonkman, and A. Robertson. Assessing fatigue and ultimate load uncertainty in floating offshore wind turbines due to varying simulation length. *Safety, Reliability, Risk and Life-Cycle Performance of*

Structures and Infrastructures - Proceedings of the 11th International Conference on Structural Safety and Reliability, ICOSSAR 2013, pages 239–246, 2013. doi: 10.1201/b16387-33.

[178] Zi Lin, Xiaolei Liu, and Saeid Lotfian. Impacts of water depth increase on offshore floating wind turbine dynamics. *Ocean Engineering*, 224(September 2020):108697, 2021. ISSN 00298018. doi: 10.1016/j.oceaneng.2021.108697. URL https://doi.org/10.1016/j.oceaneng.2021.108697.

[179] Shauna Creane, Pedro Santos, Konstanze Kölle, Davide Airoldi, Mostafa Bakhoday-Paskyabi, Michael Biglu, William Brown, Etienne Cheynet, Laura Ecenarro Diaz-Tejeiro, Lars Frøyd, et al. Iea wind tcp task 49: Reference site conditions for floating wind arrays. Technical report, National Renewable Energy Laboratory (NREL), Golden, CO (United States), 09 2024.

[180] Andreas F. Haselsteiner, Ryan G. Coe, Lance Manuel, Wei Chai, Bernt Leira, Guilherme Clarindo, C. Guedes Soares, Ásta Hannesdóttir, Nikolay Dimitrov, Aljoscha Sander, Jan-Hendrik Ohlendorf, Klaus-Dieter Thoben, Guillaume de Hauteclocque, Ed Mackay, Philip Jonathan, Chi Qiao, Andrew Myers, Anna Rode, Arndt Hildebrandt, Boso Schmidt, Erik Vanem, and Arne Bang Huseby. A benchmarking exercise for environmental contours. *Ocean Engineering*, 236:109504, 2021. ISSN 0029-8018. doi: https://doi.org/10.1016/j.oceaneng.2021.109504.

[181] J.A. Ferreira and C. Guedes Soares. Modelling bivariate distributions of significant wave height and mean wave period. *Applied Ocean Research*, 24(1):31–45, 2002. ISSN 0141-1187. doi: https://doi.org/10.1016/S0141-1187(02)00006-8.

[182] Steven R Winterstein, Todd C Ude, C Allin Cornell, Peter Bjerager, and Sverre Haver. Environmental parameters for extreme response: Inverse form with omission factors. *Proceedings of the ICOSSAR-93, Innsbruck, Austria*, pages 551–557, 1993.

[183] Wei Chai and Bernt J. Leira. Environmental contours based on inverse sorm. *Marine Structures*, 60:34–51, 2018. ISSN 0951-8339. URL https://www.sciencedirect.com/science/article/pii/S0951833917304707.

[184] Mathieu Rouaud. *Probability, statistics and estimation. Propagation of Uncertainties*. Self-published, 2013. URL https://www.incertitudes.fr/book.pdf. [Accessed: 12-12-2025].

[185] International Electrotechnical Commission. IEC TS 61400-9 Ed.1: Wind Energy Generation Systems – Part 9: Probabilistic Design Measures for Wind Turbines. Technical Specification IEC TS 61400-9, International Electrotechnical Commission, Geneva, Switzerland, 2024.

[186] Michael F. Howland, Sanjiva K. Lele, and John O. Dabiri. Wind farm power optimization through wake steering. *Proceedings of the National Academy of Sciences*, 116(29):14495–14500, 2019. doi: 10.1073/pnas.1903680116.

[187] Joeri A. Frederik, Bart M. Doekemeijer, Sebastiaan P. Mulders, and Jan-Willem van Wingerden. The helix approach: Using dynamic individual pitch control to enhance wake mixing in wind farms. *Wind Energy*, 23(8):1739–1751, 2020. doi: https://doi.org/10.1002/we.2513.

[188] ZhenZhe Ying, Zhuoer Xu, Zhifeng Li, Weiqiang Wang, and Changhua Meng. Mt-gbm: A multi-task gradient boosting machine with shared decision trees, 2022. URL https://arxiv.org/abs/2201.06239.

[189] Introduction to Boosted Trees, 2014. URL https://xgboost.readthedocs.io/en/latest/tutorials/model.html. [Accessed 11-07-2025].

[190] Martin Mathiesen, Anja K. Meyer, and Børge Kvingendal. Hywind Buchan Deep Metocean Design Basis RE2014-002. Technical Report RE2014-002, Equinor ASA, 2014.

[191] Emerald Floating. Emerald floating. URL https://emeraldfloatingwind.com/. [Accessed: 2025-02-13].

[192] Moneypoint offshore wind – site investigation: Schedule of works submission to the department of housing, planning and local government. Technical Report Document No. QS-R-353, Electricity Supply Board (ESB), november 2021.

List of Acronyms

ANN Artificial Neural Network.

BEM Blade Element Momentum.

BEMT Blade Element Momentum Theory.

CFD Computational Fluid Dynamics.

CGAN Conditional Generative Adversarial Network.

CGP Chained Gaussian Process.

CODEC Conditional Dependence Coefficient.

DEL Damage Equivalent Load.

DLC Design Load Case.

ERA5 ECMWF Reanalysis v5.

EU European Union.

FOCI Feature Ordering by Conditional Independence.

FOWT Floating Offshore Wind Turbine.

GAN Generative Adversarial Network.

GPR Gaussian Process Regression.

IEA International Energy Agency.

IEC International Electrotechnical Commission.

MDN Mixture Density Network.

MI Mutual Information.

MSE Mean Squared Error.

NLL Negative Log-Likelihood.

OFF CASE-OFFSHORE.

ON CASE-ONSHORE.

PCC Pearson's Correlation Coefficient.

PCE Polynomial Chaos Expansion.

pdf Probability Density Function.

SCADA Supervisory Control and Data Acquisition.

TLP Tension Leg Platform.

Acknowledgments

I believe that who we are is shaped primarily by the people we meet during different phases of our lives. Over the last few years, I've had the good fortune to meet some of the best people, and it is not an overstatement to say that they have collectively changed my life. A few lines of acknowledgment are insufficient to express the deep and genuine gratitude I have for the people who have been a part of my journey here.

Firstly, the biggest thanks to my promotors, Axelle and Richard. I am grateful to have had the opportunity to work with you. I cannot thank you enough for all the care, trust, guidance, and support you have provided over the years, even if that meant many extra hours spent on paperwork. From you, Axelle, I learned what good leadership looks like. Richard, thank you for creating an open and fun environment for discussions. Having little knowledge of machine learning at the beginning of my PhD, I have no one else to thank more than you for showing me the way. Kasper, everything I know about floating wind I've learned from you. I truly appreciate all the extra hours you spent supporting this work by explaining, correcting papers, preparing simulation files, providing system access, and handling all the associated paperwork, despite your very packed schedule. Erik, the same goes for you; you went out of your way to help me whenever I needed it. I also really enjoyed the nice outings in Copenhagen. Thank you for being such a nice colleague and friend! Thanks to Laurent for the guidance and help to get me started with this project.

I don't think I have ever laughed as much as I have when I hung out with the STEP4WIND family. Omar, you are the light of my life. Every single moment with you is hilarious. I can talk to you about anything, anytime, and rely on your support, no matter what. You have made my life infinitely better. Ricardo, my Siemens comrade and soon-to-be finance boss! Shout-out to your ingenious *dadly* jokes that always manage to kill us. Likhitha, thank you for bringing sunshine wherever you go. Your laughter and energy always lift me up. Matteo, I am so grateful to have had the chance to work with you, but even more to call you my friend. You are genuine, honest, hardworking, and super funny. Top-tier office mate! Huzafia, it has been a real pleasure to see you grow from '*Does this PhD actually pay?*' to '*I took a taxi from Eemshaven to Schiphol, and it cost 600 Euros.*' Fell-ipe, so excited to see you tackle real floating wind projects as you become the boss of the boss! Rahul, it is always great to hang out with you and get inspired by your positive outlook. Omer, I know you don't enjoy big groups, but our big group always enjoyed your company. Alejandro, it's been a pleasure getting to know you through this project. Oh, and I am still waiting to be invited to that housewarming someday, hahh! Whenever I felt demotivated at work, I would think back at our group trips and be instantly filled with happy memories. I couldn't have finished this PhD without this wholesome group. I love us!

Erik, thank you so much for being the best office mate and hype man, for lending your ear, and especially for giving me that extra push every now and then. Forever grateful for our friendship. Rishi, your ability to make anyone laugh is amazing. Sharing the office

in the first year was a lot of fun! I like listening to your perspective and discussing life and philosophy with you. David, you are truly the funniest guy I know. It's been a pleasure organizing the cooking and craft evenings and attending many fun conferences together. And of course, Bivio! You are missed in Delft, but I am looking forward to future workations with our group. Jelle, I have learned so much about work-life balance from you. I enjoyed our philosophical discussions and your obsession with David Deutsch. Thanks for letting me know what it is like to use all forty days of vacation a year. Mihir, I admire your quiet confidence. You are such a smart person, and I am so happy your work is being recognized. Also enjoyed all the Wind Turbine Design TA and dance sessions together. Oriol, it is always amazing to see how good you are at everything- but especially sports and board games! Kiran, thanks for organizing cycling trips, Bombay Corner outings, badminton, joining the bouldering sessions, and especially for your kindness. Shantanu, thanks for always laughing at my jokes. I'm still sad that you stopped climbing, but so proud of your running progress. George, it's always nice to see you in the office because every time, I end up learning about a random fact that comes out of your machine-like memory. Abhyuday, macha full psych scene only you've put bro. Haoyuan and Guanqun, your savage and funny comments never fail to make me laugh. Jingna, thanks for having me as your paranymph! You have given me just as much inspiration and strength. Augustin, you are the Frenchest French I've met, but also one of the nicest. Thanks for the much-needed dark humor in the office! Chethan, thanks for making my Rouen experience somehow fun with your energy and bouldering sessions. Shyam, thanks for your wholesome and kind nature. It's always nice to hang out with you and hope you move back here!

Adhyanth, Mehtab, Simone, Flavio, Clem, Anwesha, Evita, Livia, Seba, Rushi, Carlos, Nirav, Nikita, Jiaxin, Dylan, Ali, Hafiz, Nico, Arjun, Tom, Raji, João, what a privilege it is to be surrounded by such nice people who make the office a cozy and inviting space. Jenna, it has been so great working with you. Thank you for your constant support and understanding. Pratik, I really enjoyed the cool collaboration! Thanks to Dr. Veldkamp and Dr. Bierbooms for sharing your knowledge on fatigue loads and probability. It significantly helped improve the quality of my thesis. Thanks to Sylvia and Oana for managing the administrative chaos, allowing our work to continue smoothly.

I am so glad to have Priyanka and Ewoud in Delft. They are my home away from home, especially when it comes to feeding me tasty home-cooked meals. Thank you to the Paris family for their support over many years. I would like to acknowledge my CrossFit journey, which has played an important role in keeping me sane over the years. For that, thanks to my coaches, Ron and Barry, who have both helped me grow and surpass what I thought would be my limits. Machiel, you know I can't thank you enough for your understanding, patience, and kindness. I am looking forward to the future and many adventures. I owe a lot to my parents, Sudha and Dinesh, for being so open-minded. I know it has been hard because my life's trajectory didn't fit the mainstream Indian stereotype, but you have been very supportive. Lastly, I want to thank Saket for being the best brother ever. I have watched you grow into someone kind, driven, and grounded, and I couldn't be prouder of you. I look forward to all the years of shared jokes and fun ahead.

Deepali
Delft, 2025

Curriculum Vitæ

Deepali Singh

28/06/1992 Born in Allahabad, Uttar Pradesh, India

Education

12/2020-01/2026	**PhD, Aerospace Engineering**
Delft, The Netherlands	Delft University of Technology
09/2014-09/2016	**M.Sc. Aerospace Mechanics and Avionics**
Toulouse, France	Institut supérieur de l'aéronautique et de l'espace
08/2010-08/2014	**B.E. Aeronautical Engineering**
Manipal, India	Manipal Institute of Technology
08/2008-07/2010	**High School**
Indore, India	Emerald Heights International School

Experience

10/2025-present	**Postdoctoral Research Candidate**
Delft, The Netherlands	Delft University of Technology
07/2022-11/2023	**Industrial Secondment**
Rouen, France	Siemens Gamesa Renewable Energy
07/2019-11/2020	**High-fidelity CFD Simulation Engineer**
Paris, France	Safran Tech
10/2016-06/2019	**Aerospace and Defense Solution Consultant**
Stuttgart, Germany	Dassault Systèmes

List of Publications

Journal articles

4. **Deepali Singh**, Erik Haugen, Kasper Laugesen, Richard P. Dwight, and Axelle Viré. Data-driven probabilistic surrogate model for floating wind turbine lifetime damage equivalent load prediction. *Wind Energy Science*, 10(12):2865–2888, 2025. doi: 10.5194/wes-10-2865-2025.

3. Pratik Rao, Richard P. Dwight, **Deepali Singh**, Jin Maruhashi, Irene Dedoussi, Volker Grewe, and Christine Frömming. The ozone radiative forcing of nitrogen oxide emissions from aviation can be estimated using a probabilistic approach. *Communications Earth & Environment*, 5(1):558, 2024. doi: 10.1038/s43247-024-01691-2.

2. **Deepali Singh**, Richard P. Dwight, and Axelle C. Viré. Probabilistic surrogate modeling of damage equivalent loads on onshore and offshore wind turbines using mixture density networks. *Wind Energy Science*, 9(10):1885–1904, 2024. doi: 10.5194/wes-9-1885-2024.

1. Nicolas Gourdain, **Deepali Singh**, Thierry Jardin, and Sébastien Prothin. Analysis of the turbulent wake generated by a micro air vehicle hovering near the ground with a lattice Boltzmann method. *Journal of the American Helicopter Society*, 62(4):042003, 2017. doi: 10.4050/JAHS.62.042003.

Conference proceedings

11. **Deepali Singh**, Erik Haugen, Kasper Laugesen, Ayush Chauhan, and Axelle C. Viré. Data analysis of the TetraSpar demonstrator measurements. *Journal of Physics: Conference Series*, 2767(6):062025, 2024. doi: 10.1088/1742-6596/2767/6/062025.

10. Hafiz Ghazali Bin Muhammad Amri, Daniela Marramiero, **Deepali Singh**, Jan-Willem van Wingerden, and Axelle C. Viré. Data-driven time series forecasting of offshore wind turbine loads. *Journal of Physics: Conference Series*, 2767(5):052060, 2024. doi: 10.1088/1742-6596/2767/5/052060.

9. Matteo Baudino Bessone, **Deepali Singh**, Theodoros Kalimeris, Erin Bachynski-Polić, and Axelle C. Viré. Surrogate-assisted optimization of floating wind turbine sub-

structure. *Journal of Physics: Conference Series*, 2767(6):062032, 2024.
doi: 10.1088/1742-6596/2767/6/062032.

8. **Deepali Singh**, Richard P. Dwight, Kasper Laugesen, Laurent Beaudet, and Axelle C. Viré. Probabilistic surrogate modeling of offshore wind-turbine loads with chained Gaussian processes. *Journal of Physics: Conference Series*, 2265(3):032070, 2022.
doi: 10.1088/1742-6596/2265/3/032070.

7. **Deepali Singh**, Benedikt König, Ehab Fares, Mitsuhiro Murayama, Yasushi Ito, Yuzuru Yokokawa, and Kazuomi Yamamoto. Lattice-Boltzmann simulations of the JAXA JSM high-lift configuration in a wind tunnel. *AIAA SciTech Forum Conference Proceedings*, 2019.
doi: 10.2514/6.2019-1333.

6. Benedikt König, **Deepali Singh**, and Ehab Fares. Lattice Boltzmann high-lift simulations — a step beyond classical CFD. *31st Congress of the International Council of the Aeronautical Sciences (ICCAS)*, 2018.

5. Benedikt König, **Deepali Singh**, Ehab Fares, and Martin Wright. Transonic Lattice Boltzmann Simulations of the NASA-CRM in the European Transonic Windtunnel. *AIAA Applied Aerodynamics Conference*, 2018.
doi: 10.2514/6.2018-3900.

4. Andre F. Ribeiro, Benedikt König, **Deepali Singh**, Ehab Fares, Raoyang Zhang, Pardeep Gopalakrishnan, Avinash Jammalamadaka, Yanbing Li, and Hudong Chen. Buffet simulations with a lattice-Boltzmann based transonic solver. *AIAA Aerospace Sciences Meeting*, 2017.
doi: 10.2514/6.2017-0514.

3. **Deepali Singh**, Andre F. Ribeiro, Benedikt König, and Ehab Fares. Lattice Boltzmann simulations of a supersonic cavity. *35th AIAA Applied Aerodynamics Conference*, 2017.
doi: 10.2514/6.2017-3515.

2. Andre F. Ribeiro, **Deepali Singh**, Benedikt König, and Ehab Fares. On the stall characteristics of iced wings. *AIAA Aerospace Sciences Meeting*, 2017.
doi: 10.2514/6.2017-1966.

1. Parag Deshpande, Roshan Antony, Prasobh Narayanan, Anirudh Rajashesan, **Deepali Singh**, and G. Ramesh. Visualization and PIV measurements of leading-edge vortex generated by rigid flapping wings. *32nd AIAA Applied Aerodynamics Conference*, 2014.
doi: 10.2514/6.2014-2151.

Talks

5. **Deepali Singh**, Erik Haugen, Kasper Laugesen, Richard P. Dwight, and Axelle Viré. Data-driven probabilistic prediction of floating wind turbine lifetime damage equivalent load. *Wind Energy Science Conference*, Nantes, 2025.

4. **Deepali Singh**, Richard P. Dwight, Li-Toong Yap, Erik Haugen, and Axelle C. Viré. Probabilistic data-driven load surrogate model for floating wind turbines. *Wind Energy Science Conference*, Glasgow, 2023.

3. **Deepali Singh**, Richard P. Dwight, and Axelle C. Viré. Site-specific load prediction using mixture density networks: a probabilistic approach. *NAWEA/WindTech*, Delaware, 2022.

2. **Deepali Singh**, Richard P. Dwight, Kasper Laugesen, Laurent Beaudet, and Axelle C. Viré. Data-driven reduced order model for offshore wind turbines. *Wind Energy Science Conference*, Hamburg, 2021.

1. **Deepali Singh** and Rafael Ihi. Aero-Structural Coupling Using Lattice Boltzmann Method. *GAMM CSE Workshop*, Aachen, 2017.

This work explores probabilistic machine learning methods, such as mixture density networks, for mapping site conditions to the short-term and lifetime damage equivalent loads of onshore, fixed-bottom offshore, and floating offshore wind turbines.