

## A SURROGATE MODELING APPROACH FOR FATIGUE DAMAGE ASSESSMENT OF FLOATING WIND TURBINES

**Kolja Müller**  
University of Stuttgart  
Stuttgart, Germany

**Po Wen Cheng**  
University of Stuttgart  
Stuttgart, Germany

### ABSTRACT

Fatigue analysis for floating wind turbines poses a novel challenge to calculation workflows if a probabilistic load environment is to be considered. The increased complexity of the structure itself as well as its interaction with the environment require a coupled and more detailed analysis with respect to resolution of environmental conditions compared to fixed bottom systems.

Different approaches to address the computing challenge for floating turbines are possible to support engineering judgement and have been investigated in the past, with conservative binning on the one end of the accuracy scale and computation intensive Monte Carlo simulations on the other end. This study investigates the feasibility of regression based surrogate models based on radial basis functions. The investigation performed here is aligned with work performed in the H2020 project LIFES50+. Consequently, the considered system is the DTU 10MW Reference Wind Turbine installed on the LIFES50+ OO-Star Wind Floater Semi 10MW. The site under investigation is the LIFES50+ Site B (Gulf of Maine) medium severity representative site.

Results show a similar convergence of lifetime fatigue load prediction as with Monte Carlo simulations indicating that this technique may be an alternative if a response model of the considered system is of interest. This may be interesting if damage loading is to be calculated at a different site and if a classification of met-ocean conditions is available.

### INTRODUCTION

With the ongoing trends in offshore wind energy towards deeper waters and larger distances to the shore, floating offshore wind turbines have shown to be a possible addition to the offshore efforts to a sustainable energy mix. Industrialization efforts to

minimize system costs require a standardized, fast and accurate design process in order to check the feasibility of a given design for a specific site. The overall design process has increased in complexity moving from onshore to offshore to floating offshore systems, due to rising influence of the wave environment on overall system dynamics and loads. Especially the evaluation of the fatigue limit state, as described in the power production design load case 1.2 of common guidelines [1] requires a thorough understanding of the system response towards a large variation of environmental parameters. This is typically highly nonlinear as the wind turbine is considered to be in operational condition and thus, the complex interaction of the turbine with its environment becomes relevant. Fatigue load assessment is commonly established by comprehensive studies using integrated, time-domain simulations.

Fatigue damage assessment requires the (conservative) consideration of all possible events during the systems lifetime. This means that for relevant environmental conditions, the load response is calculated and weighted according to its occurrence probability. This is rather simple for an onshore turbine, as wind is the only relevant environmental condition. Because an onshore turbine foundation is typically circular, the directionality of the wind can be discarded. Then, only the wind speed and its variation, the turbulence intensity, are influencing the loading. For fixed-bottom offshore structures, ocean loading needs to be considered (wave height, wave period and current speed become relevant) and a symmetric substructure like a jacket may be added (wind, wave and current directionality become relevant). Going offshore thus significantly increases the number of environmental conditions to be considered, which leads to an increased modelling effort of the more complex system (i.e. substructure and ocean loads). The characteristics of fixed-bottom offshore structures however allow for simplified, conservative approaches: the use of damage equivalent

significant wave heights allows to link the wave height to the wind speed [3]. Stiff substructures result in natural periods well out of the wave period range, allowing the consideration of representative wave periods only (median or damage equivalent) [4] [5] [6]. Finally, it is expected that the largest loads are resulting from aligned wind, wave and current directions [7]. Thus, through simplifications, going offshore only increased the variable space by the wind direction. Still, the more complex structure and the added environmental variables significantly increase the simulation effort. Conservative assumptions help to keep the total effort on a feasible level. Further simplification may be possible due to the limited interaction of the turbine with the substructure by performing separate investigation of the rotor and substructure [9] [10] [11]. Conservative assumptions limit the environmental variable space and help to keep down the simulation effort. However, it was shown that fatigue loading on fixed-bottom wind turbines varies significantly due to changing environmental conditions [9] [12], indicating that a more accurate and less conservative design is possible.

For floating wind turbines, many of the simplifications used for fixed-bottom design may not be applied: Coupled models are required due to the close interaction of the rotor and the floating foundation [10]. The large natural periods of the floating platform result in largely increased sensitivity of the loading to the environment. Hydrodynamic response varies strongly with wave periods and at the same time with wave and wind inflow direction. Wind-wave-misalignment becomes relevant [13]. This means that the use of conservative assumptions may lead to overly conservative designs. On the other hand, considering a more detailed environment increases the simulation effort. Due to the large number of simulations to be performed, the simulation effort for fatigue analysis is substantial and poses a major challenge in the design and certification process of floating wind systems. Procedures exist which incorporate simplifications of the environment in order to limit the number of simulations, e.g. by using damage equivalent environmental conditions. Also, simple models may be used for initial design purposes and controller design [14]. However, given the strong variation in loading, these simplifications may lead to excessively conservative designs and/or may not consider important components of the turbine, such as rotor and mooring line loads [15].

Given the predicament of choosing between simulation effort and conservatism, it is regarded as an aspiring goal to find a methodology which allows the following: an accurate fatigue design considering the probability environment of the considered site and also keeping the simulation effort to a minimum. Problems of this kind are addressed by the general discipline of “design of simulation experiments” [16]. Here, the most baseline approach can be considered the simple simulation of random combinations of the input variables under consideration of their occurrence probability. This approach is commonly known as the Monte Carlo approach. The Monte Carlo approach allows a

discrete integration across a multidimensional design space following the equation

$$\int_{[0,1]^s} f(x) dx \approx \frac{1}{N} \sum_{i=1}^N f(\xi_i), \quad (1)$$

where  $\xi_i \in [0,1]^s$  are each of the  $N$  random, independent samples in the  $s$ -dimensional unit hypercube. It can be shown that the Monte Carlo approach converges to the correct solution, which is why this approach is of high value for use as a reference. However, Monte Carlo simulations generally show slow convergence and often efficient alternatives are available to simplify the calculation of the integration problem. In wind energy fatigue assessment, as described, feasible alternatives are available for onshore and offshore wind energy, but for floating systems, the problem has to be reviewed based on presented arguments.

One application of Monte Carlo integration for fatigue load assessment of floating wind turbines is presented in [17]. There, simple random sampling is applied considering the probability distribution of five different environmental parameters. It is mentioned that using Monte Carlo compared to a grid approach (an efficient alternative), the convergence rate is increased from  $h = \frac{1}{n^{1/d}}$  to  $h = \frac{1}{n^{1/2}}$ , with  $n$  simulations and  $d$  dimensions or independent variables. This results in an advantage for increased dimensionality. A possible limit with respect to increased nonlinearity (considered by increasing the fatigue slopes) is mentioned. A variation to standard random sampling is followed in [18], by which the convergence rate is further decreased by minimizing the discrepancy of the point set. The results presented therein also serve as baseline for the present study. This probability-based Monte Carlo approach produces quickly converging results for the lifetime damage estimate and may be of interest for a more accurate site-specific fatigue assessment. However, it is closely linked to the site under consideration which limits the reusability of the performed work (i.e. each combination of environmental conditions to be simulated are chosen according to their occurrence probability based on the probability model of the environment of the considered site).

A more flexible usage of the simulation results may be obtained by application of surrogate models. These can be seen as multidimensional regression surfaces representing the load response as a function of the (assumed independent) environmental variables. If a surrogate is available, it may simply be combined with the site-specific occurrence probability of environmental conditions in order to obtain the fatigue loads over the systems’ lifetime. Considering the possible introduction of substructure classes similar to wind turbine classes [2], the same design will be evaluated in different environments. Surrogate models may be site-independent and hence allow the creation of a database which can be evaluated for different environments without any simulation effort (assuming the site-independency of the mooring line system). Initial surrogate

models may be obtained from the first load evaluation and new information may be added later on to improve accuracy.

The determination of load response surrogates is seen in the same problem specification of curve fitting procedures applied in analysis of experimental data [32]: there is an unknown target function is to be constructed by a simpler known function (the simplest version is certainly a linear curve fitting procedure based on polynomial function of e.g. first- or second degree). First attempts related to load response modeling of floating wind turbines have been performed in the past (e.g. [30]). With the development and employment of artificial intelligence and statistical learning more complex alternatives as function approximators are available, which may also be applied in a curve fitting environment, such as neural networks. These methods are more complex and due to the high dimensionality of the problem, many supporting points are typically required. Thus, these approaches are especially interesting in the field of load monitoring where a large amount of data points can be assumed to be available [33]. As computational power is increasing the application of these methods for simulation based load evaluation may become interesting. In [19], a neural network was used based on probabilistic environmental conditions for the estimation of lifetime fatigue damage of floating wind turbines.

The present study is building on the past experiences by presenting the following:

- Use independent environmental variables in order to achieve more flexibility in the usage of the results.
- Use quasi-random sampling for design point selection. This procedure shows a better distribution of design points compared to other random sampling techniques such as e.g. latin-hypercube sampling [20] and thus shows better convergence qualities. Also, convergence analyses are facilitated, as newly added points follow the scheme of optimal distribution within the design space.
- Use radial basis function networks for surrogate modeling. This is considered as a simple alternative to neural networks representing the application of approximation functions in general [21].

The present paper starts with an introduction of the considered system and environment, as well as the selection of the baseline design points which are later used for the setup of the surrogate model. Afterwards, the models active in the simulation and post-processing routines are presented. Next, the setup and evaluation of the radial basis function surrogate model is described. Finally, the fatigue damage prediction obtained using the surrogate model is discussed.

## CONSIDERED SYSTEM

As this study is closely aligned with work performed in the H2020 project LIFES50+, the analysis is performed using an

early version of the public DTU 10MW reference turbine [22] installed on the LIFES50+ [23] OO-Star Wind Floater Semi 10MW floating platform. The tower of the DTU 10MW turbine was redesigned to a stiff-stiff design in order to mitigate rotor-excitation. A system-specific controller is applied taking into consideration the global system dynamics and avoiding the floating-wind specific negative damping problem. The mooring line system applies added weights in order to stiffen the global system motions. A public FAST model of the used system is provided in [24]. In this study, a preliminary version of the same is used, including minor variations of e.g. mooring line design and hydrodynamic properties as can be found in the final version of the public model. A sketch of the substructure is shown in FIGURE 1. Key information of the system is provided in TABLE 1.

TABLE 1: KEY PARAMETERS OF LIFES50+ OO-STAR WIND FLOATER SEMI 10MW FLOATING PLATFORM WITH DTU 10MW REFERENCE TURBINE. SYSTEM FREQUENCIES CONSIDER A STIFF SUBSTRUCTURE FROM MSL DOWNWARDS.

Property	Value
Cut-in wind speed	4 m/s
Rated wind speed	11.4 m/s
Cut-out wind speed	25 m/s
Number of mooring lines	3
Mass of clumped weight	50000 kg
Substructure mass	2.3618E+07 kg
System natural frequency surge	0.0055 Hz
System natural frequency heave	0.049 Hz
System natural frequency pitch	0.032 Hz
System natural frequency yaw	0.0086 Hz
System natural frequency tower	0.786 Hz

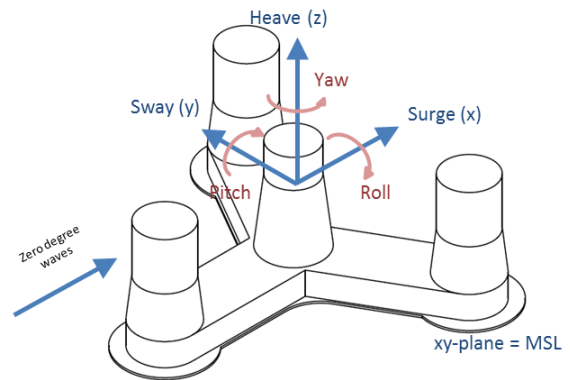


FIGURE 1: LIFES50+ OO-STAR WIND FLOATER SEMI 10MW FLOATING PLATFORM

## CONSIDERED ENVIRONMENT AND DESIGN POINT SELECTION

Environmental parameters are determined based on LIFES50+ representative site B (medium severity environment, reference

site: Gulf of Maine, USA) as provided by the design basis of LIFES50+ [25]. Three environmental conditions are considered in this study: wind speed, wave height and wave period. Wind and wave data are taken from measurement data from the NOAA buoy data network, which was also used by [26]. Hub height wind speeds are calculated using the power law for the wind shear [27]. Wind and wave directions are assumed co-aligned in this study. Turbulence intensity is considered according to IEC turbulence class C [2]. Hourly measurements were considered from 2003 to 2015, resulting in an overall database of 103,282 useable measurement points.

Based on the available data, a Nataf joint probability model of the considered environmental parameters was established based on available measurement data [18]. The model is used for weighting of environmental conditions for the determination of the lifetime damage loading. The use of a surrogate model for the environment is considered of higher value for this study compared to using raw data, as this way gaps in information due to sensor resolution are mitigated. As the model accuracy of the environmental model is not part of the investigation, the Nataf model is assumed to be a “perfect” description of the environment. The same Nataf model was used for all the simulations referred to in this study, so errors in the modeling do not influence the results of this study. Using this model, continuous information of the environment is available which facilitates probabilistic evaluation of the fatigue loading. The quality and performance of the environmental model is not part of this study.

The evaluation of model also provides the limits of the design space to be considered. This is done by considering three load ranges (LR) which divide the wind speed into three regions: below rated (LR1),  $v_{cut-in} < v \leq v_{rated} - 0.15 \cdot v_{rated}$ , transition (LR2),  $0.15 \cdot v_{rated} \leq v < v_{rated} + 0.15 \cdot v_{rated}$  and above rated (LR3),  $0.15 \cdot v_{rated} \leq v < v_{cut-out}$ . Within the load ranges, the 1<sup>st</sup> and 99<sup>th</sup> percentiles of samples extracted from the environmental model of wave height and -period provide the limits of the design space. This way, a more efficient consideration of the environment is given by roughly taking into account the dependence of wave heights on wind speed. While this approach may be of advantage for the considered environment in this study, it may be a disadvantage when the resulting surrogate model is to be transferred to another site. Then, new simulations may be necessary in order to mitigate extrapolation errors. A more general approach by defining different classes (in the form of design space border values) of marine environmental conditions linked to the wind environment would help to limit the amount of simulations required for a predefined range of environmental conditions (i.e. all environments covered by a certain class can be considered with a given surrogate model).

The limits of the design space for the considered environmental conditions are given in Table 2. Figure 2 shows a scatterplot of all combinations of environmental conditions (design points) simulated in this study. For each load range, 345 design points were selected using the quasi-random sampling based on Sobol’ sequences as described in [28] and [29]. It can be seen that the

density of design points is smaller for wave height in LR3, which is due to the larger range of wave heights that need to be considered in this region. The total number of data points used in this study is 1035 (345x3) for which 3105 simulations were performed (1035 x 3 seeds).

TABLE 2: ENVIRONMENTAL CONDITIONS FOR DIFFERENT LOAD RANGES (LR)

Load range	variable	Design space limits [min, max]
LR1	Wave height [m]	0.3, 3.2
LR1	Wave period [s]	1.7, 13.3
LR2	Wave height [m]	0.5, 5.0
LR2	Wave period [s]	1.2, 12.3
LR3	Wave height [m]	0.7, 7.0
LR3	Wave period [s]	0.9, 11.9

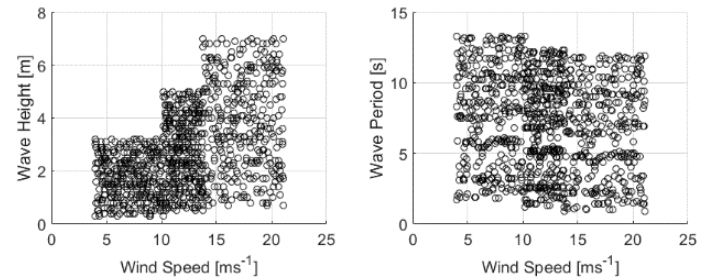


FIGURE 2 : SCATTERPLOT OF DESIGN POINTS CONSIDERED IN STUDY

## SIMULATION

FAST v8.16.00a-bjj was used to perform coupled simulations. Overall simulation time was 60 minutes with an additional 10 minutes that were neglected in the post-processing as run-in time to allow initial transients to fade out. Turbulent wind based on the Kaimal model and 10 minutes periodic wind fields were applied for the simulation. The simulation time for floating wind turbines has been investigated in [34] with the result that 10 min simulations are sufficient for a 5MW Spar buoy. A detailed investigation for the system and environment used in this study is presented in [35]. From the results presented therein, it can be assumed that the load uncertainty is below 10% (ca. 7/6/3 % for blades/tower/fairlead, respectively) for the applied setup in this study. Note that a SN-curve slope of  $m = 4$  is applied for all sensors, which means that all results for the blade loading are only indicative (see also further below). For aerodynamic modeling only simple blade element momentum theory was taken into account, i.e. no corrections for dynamic stall or dynamic inflow were considered. The platform hydrodynamics are calculated considering both Morison equations as well as potential flow theory. The potential-flow model was established in a pre-processing step using the panel-code Ansys AQWA. Mooring line dynamics are determined based on a dynamic model using NRELs lumped mass mooring line modeler

MoorDyn. The Jonswap model and a superposition of airy waves was applied for modeling the wave environment.

Rainflow counting was applied to obtain the distribution of the load amplitudes  $\Delta L$  for each time series and the Palmgren-Miner linear damage accumulation law was used to calculate damage equivalent load amplitudes  $\Delta L_{DE}$  (commonly known as damage equivalent loads or DEL) of the obtained 1 hour time series:

$$\Delta L_{DE,Simulation} = \left( \sum \frac{\Delta L_i^m \cdot n_i}{N_{ref,Simulation}} \right)^{\frac{1}{m}}, \quad (2)$$

where  $\Delta L_i$  are the load amplitudes from the rainflow counting,  $n_i$  are the number of occurrences of the detected load cycles,  $N_{ref,Simulation}$  is the reference cycle number applied for each simulation (set to  $2e6$  in this study) and  $m$  is the slope of the SN-curve. In this study,  $m = 4$  was assumed for all evaluated positions. This may not be adequate for all positions (in particular for composite materials typically  $m = 10$  is used), but is regarded as sufficient for the demonstration purpose of the method.

The results for the 3 seeds considered for each design point were averaged. Exemplary results of the tower base fore-aft bending moment DELs are shown for different load ranges (columns) and environmental conditions (rows) in FIGURE 3. The large influence of wind speed in the first load range is visible as well as the increased impact of the wave height with increasing magnitude. The DEL response also shows a distinctive peak at a wave period of about 7.2s. This peak is somewhat shifted to the pitch response from a potential-flow response magnitude evaluation. This may be linked to the additional dynamics of the mooring lines and wind turbine, which was not considered in the potential-flow analysis.

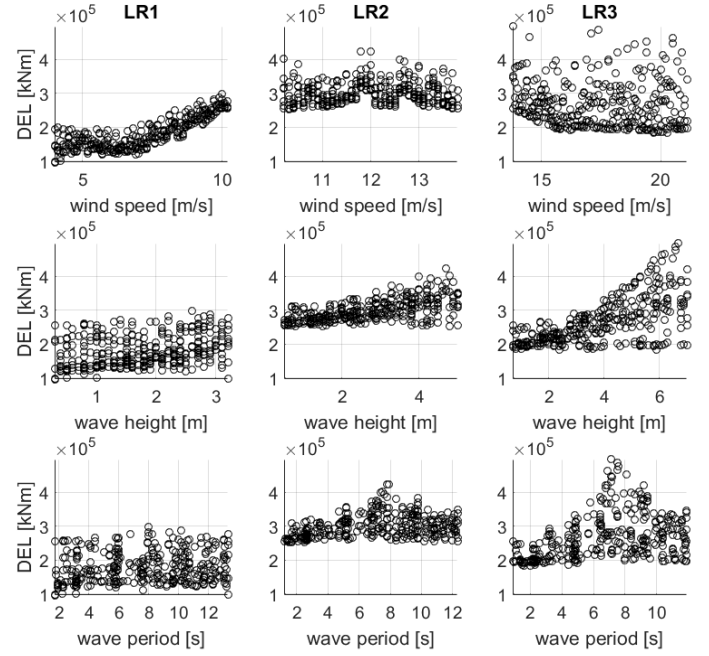


FIGURE 3: SCATTERPLOT TOWER BASE FORE-AFT BENDING MOMENT DELS FOR DIFFERENT LOAD RANGES (COLUMNS) AND ENVIRONMENTAL CONDITIONS (ROWS).

## SURROGATE MODEL

Based on the simulation results, a surrogate model can be established for the DEL response of different load sensors as a function of the environmental conditions. The derivation of a surrogate model for the fatigue load response of floating wind turbines can essentially be viewed as a curve-fitting operation in multidimensional environmental variable space. Taking into account the complexity of the problem, a simple curve fitting based on polynomial functions is not feasible, as was shown in [30]. For highly nonlinear, multidimensional curve-fitting (also: function approximation, pattern-recognition) problems, machine-learning algorithms are a common approach in the field of data analysis.

Classical regression is typically performed in the sense of multiple linear regression [36], which means linear in a statistical sense (i.e. with respect to the regression coefficients, so the base functions of the regression may be non-linear):

$$Y = X\beta + \varepsilon. \quad (3)$$

Here,  $Y \in \mathbb{R}^{R \times 1}$  is the system response over  $R$  evaluations of varying combinations of the  $p$  independent variables  $X \in \mathbb{R}^{R \times p+1}$  combined by regression coefficients  $\beta \in \mathbb{R}^{1 \times p+1}$  and errors between observed and predicted values  $\varepsilon \in \mathbb{R}^{R \times 1}$ . The regression coefficients of a predefined set of linear combinations

of functions may be tuned such that the sum-of-squares of the errors is minimized, which is a simple procedure.

An alternative to linear regression is nonlinear regression, which offers a much broader range for the selection and combination of baseline functions, however the error-minimization then needs to be performed iteratively, embedded into a non-linear optimization framework. The general problem of regression with analytical functions, however, is that they have to be defined beforehand which becomes increasingly difficult with increasing number of dimensions, especially when the regression is nonlinear. This is where neural networks are of interest as they may provide a cost-effective, general, parametrized, non-linear response approximation procedure given the existence of a set of input-output combinations [37]. They are cost-effective from a computational point of view in the sense that they superpose simplified functions (neurons) depending on single variables (weights) rather than taking into account many variables at the same time. This way, they mitigate the “curse of dimensionality” faced by conventional (non-linear) series expansion. However, still, as neural networks are nonlinear in the parameter space, learning must be based on nonlinear optimization techniques which may cause new challenges when trying to mitigate results in local minima [38].

Radial basis networks (RBN) may be viewed as a simple representation of artificial neural networks (ANN) due to their simple definition and fast and efficient learning procedures [38]. They also show good performance for interpolation of noisy data [39]. Implementation of RBNs is done as two layer feed forward networks (i.e. one hidden and one output layer; information flows only from the input over the hidden to the output layer). The hidden layer employs radial basis functions at fixed points (neurons), whose values depend only on the distance towards the center point. Due to fixing the radial functions to predefined points (and hence the nonlinearities) in the hidden layer, a fixed nonlinear transformation with no adjustable parameters of the input vector is performed. The results of this nonlinear transformation is then weighted in a second linear output layer whose weights resemble the only adjustable parameters of the model. Due to the setup, it is expected that RBNs show similar performance as other two-layer ANNs, hence presenting a good alternative to classical ANNs for this work.

The network output for a given input vector is calculated as follows:

- (1) In the radial basis layer, the distance between the input vector and the weight matrix (containing the weight vectors) is calculated.
- (2) The obtained distances are multiplied elementwise by a constant bias vector, which tailors the sensitivity of the different radial basis neurons.
- (3) The biased distances are then inserted into the radial basis function. In this work, radial basis functions of the type  $\phi(n) = e^{-n^2}$  are used, with distance  $n$  between function center and evaluation point. The results of the

hidden layer are then forwarded to the output layer. A spread parameter may be defined, which defines the range of influence of individual neurons. In this study, two different spreads were used such that the medium response is achieved after 10% and 20% of the design space. A smaller spread can be expected to require more neurons to cover the overall design space but allows more complex response surfaces (as the overlap between neurons is less).

- (4) In the output layer, the results of the different neurons are added and a constant bias value is added to the results from the radial basis layer.

The response mapping in  $k$ -dimensional space  $f: \mathbb{R}^k \rightarrow \mathbb{R}^1$  performed by the network can be summarized in the formula:

$$f(\mathbf{x}) = \lambda_0 + \sum_{i=1}^{k_r} \phi(\lambda_i \cdot \|\mathbf{x} - \mathbf{w}_i\|) \quad (4)$$

With the input vector  $\mathbf{x}$ , bias or weight values  $\lambda_i$ , weight vector  $\mathbf{w}$  and the number of neurons  $k_r$ . The curve fitting may then be performed by first defining  $\mathbf{w}$  and using a least-square method to find optimal  $\lambda_i$ . An alternative approach is the orthogonal least squares learning algorithm as proposed in [38] and applied in this work. Therein, radial basis neurons are sequentially added at points where the greatest error of the network is detected so that the error at these points is minimized. This procedure is repeated until the mean squared error (MSE) of the network reaches a value below the predefined threshold ( $\epsilon_{goal} = 10^{-3}$  and  $10^{-4}$  in this study). A more detailed description of the algorithm used in this study can be found in [31] and [38].

More advanced neural networks may employ sigmoid transfer functions in the hidden layer, which may provide more accurate results. Also more research may investigate more closely the number of hidden layers and feedback loops. Also, algorithms to avoid subjective parameters such as smooth factors or center locations as is required in the procedure used here may provide additional advantages [39]. Finally other options for surrogate models have been used elsewhere, and may also be promising for the modeling of the loads of floating offshore wind turbines, such as Kriging, multi-layer neural networks, support vector machines or polynomial chaos. It remains to be found the most feasible (i.e. simple, robust, computationally efficient and exact) solution for the presented curve fitting problem. This is not part of this work, as only the general applicability of complex function approximation techniques is to be shown for the derivation of surrogates for the load response of floating wind turbines.

For derivation in this study, the available data set of 1035 points is subdivided into sets for training and validation. The number of data sets for training the network ranges between 50 and 1000, as indicated in Fig. 6 (x-axis values). For each considered training size, random points from the database are chosen



repetitively (1000 times for each training size) and a surrogate model is created.

## MODEL EVALUATION

The evaluation of any individual surrogate model has to be performed for any surrogate model which is to be applied. A feasible approach is presented in this section and exemplary applied to the tower fore-aft bending moment loads and a surrogate model using spreads of 10% and a target MSE of  $10^{-4}$  (considering the training set).

The model evaluation typically involves the comparison of response predictions from the newly created model to some reference data (i.e. validation data set). In the present work, simulation results are available from a related study based on direct Monte-Carlo simulation (probability based) using 16,200 simulations to determine the fatigue load response at 5400 design points. From the results of the study, the lifetime DEL can be predicted with an accuracy of  $\pm 5\%$  after 300 simulations for the blade, tower base and fairlead positions. Thus, using all of the 5400 DEL results from that study is considered valid baseline data for this study. The study assumed the same Nataf model used in this study and the same system and environmental conditions were used. It is presented in detail in [18].

Using the abovementioned MC baseline data, the following procedure for evaluation of surrogate models is set up including the following items:

- **Scatterplots** over all dimensions provide a first intuitive overview of the accuracy of the model
- **Error analysis** of the error of predicted DEL is performed by plotting the
  - **Variation of error** depending on the predicted response and the independent variables. This allows insuring that errors are randomly distributed across the design and response space. This is complemented by
  - **Error histograms** to ensure normalized distribution of errors and plotting
  - **Predicted over observed values** (from baseline data).
  - **2D error maps** are used additionally to identify more clearly the relation of errors to the location in the design space.

The procedure is independent on the modeling approach presented in this work. Hence, the procedure and the related plots may be of use for similar studies which apply surrogate modelling.

Figure 4 shows the scatterplot evaluation of a surrogate model based on 800 simulations. It is visible how the main characteristics of the load response are well captured (load peak around rated wind speed, increased loading with increasing wave height and small impact of wave period with some increased

loading around 8s periods). Some outliers are also visible which apparently lead to smaller predicted loads.

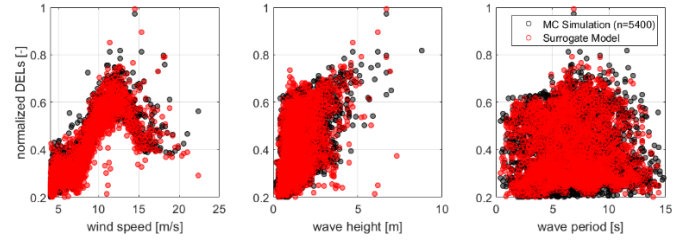


FIGURE 4: SURROGATE MODEL VERIFICATION USING SCATTERPLOTS. RESULTS SHOWN TOWER BASE FORE-AFT BENDING MOMENT.

For the error evaluation, the normalized error is calculated as follows:

$$\epsilon_{DEL} = \frac{\Delta L_{DE,surrogate} - \Delta L_{DE,baseline}}{E(\Delta L_{DE,baseline})}, \quad (5)$$

with  $\Delta L_{DE,surrogate} \in \mathbb{R}^{1 \times 5400}$  the predicted results from the surrogate model,  $\Delta L_{DE,baseline} \in \mathbb{R}^{1 \times 5400}$  the observed results from the baseline simulation study, and expected value  $E$ .

In curve fitting problems, it is to be ensured that errors are distributed randomly and do not show dependence on the independent or the predicted variable. This analysis may be done using a plot set as given in Figure 5, which is based on the same data as used in Figure 4. It shows a constant random distribution of errors across both predicted and independent variables. The histogram of the errors shows some increased kurtosis compared to the reference standard normal distribution (scattered line). Some more weight of the probability distribution is seen for positive errors. Overall, the errors are well within the 0.5 margin. However, some outliers are visible, in particular for large wave heights and small DEL values.

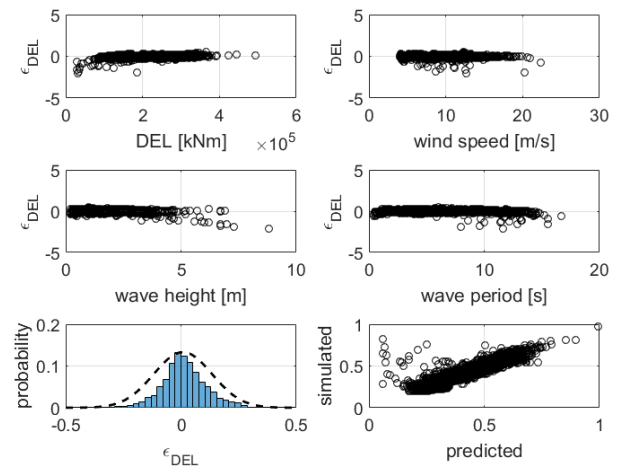


FIGURE 5: ERROR EVALUATION. RESULTS SHOWN FOR TOWER BASE FORE-AFT BENDING MOMENT.

Figure 6 shows the 2D error maps for the tower base fore-aft bending moment DEL response. These plots indicate the borders of the design space used for the setup of the surrogate model (red boxes) and the points evaluated for the error analysis (and simulated in the baseline MC study). With this plot it becomes clearly visible how extrapolation beyond the design space leads to the previously observed large errors. As the design space borders are defined according to the 1<sup>st</sup> and 99<sup>th</sup> percentile values, it is clear that in a model evaluation of 5400 design points, some values will fall out of the design space. Because the radial basis network positions radial basis functions across the design space, an underestimation of the loads outside of the design space is to be expected. The problem of extrapolation can only be mitigated by further increasing the design space, which does not seem feasible due to the lower density of design points. An alternative is to simply neglect points outside of the design space, the error of which would have to be assessed. Also, less points could be sufficient to calculate lifetime damage which means that the problem does not occur for smaller samples. For standardized application, it should be clear how many design points are required for the calculation of lifetime damage. Environmental contours are seen as a promising solution to give a limit for consideration of sampling points. These could then also give an indication of the applicability of a surrogate model for a specific site. In [18] it was shown that singular events do not have a large influence on the lifetime DEL, hence the impact of these extrapolation is not found critical in the present study.

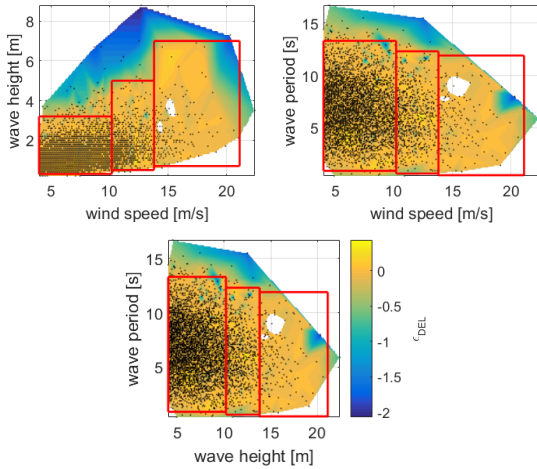


FIGURE 6: 2D ERROR MAPS. SHOWN FOR TOWER BASE FORE-AFT BENDING MOMENT.

## LIFETIME DAMAGE EQUIVALENT FATIGUE LOADS

Using the surrogate model, a continuous description of 1hour DEL response across the environmental design space is available. This may now be combined with the continuous model of the occurrence probability of the environmental parameters in order to arrive at a lifetime damage response. This is done in the present study using Monte-Carlo integration. As described above, the Nataf joint probability model of the environment is

determined based on marginal probability distributions and correlations of environmental variables obtained from measurements. The model allows a mapping of random points in a unit hypersquare into the variable space, taking into account the occurrence probability of different environmental conditions. With the environmental model  $10^6$  environmental events are selected taking into account the probability of occurrence. The DELs for the resulting set of environmental conditions are then obtained from the DEL surrogate model. Finally, the lifetime damage equivalent load may be calculated by Monte Carlo integration using:

$$\Delta L_{DE,Lifetime} = \left( \sum_{i=1}^{n_{samples}} \Delta L_{DE,i}^m \frac{N_{ref,Lifetime,i}}{N_{ref,Simulation}} \right)^{\frac{1}{m}}, \quad (6)$$

where  $N_{ref,Lifetime,i}$  is the reference cycle number for the full life time of the system.  $N_{ref,Lifetime}$  is calculated by weighting each evaluation point according to their relative occurrence probability over the entire life time:  $N_{ref,Lifetime} = w_{sim} \cdot N_{ref,Simulation}$ . In the case of  $10^6$  environmental conditions, the weighting is calculated according to  $w_{sim} = \frac{10^6}{20 \cdot 8760 \cdot 3600s}$ .

## CONVERGENCE STUDY

The convergence of the overall procedure is investigated with respect to increasing the number of simulations used for the derivation of the surrogate model. An uncertainty is implemented in the surrogate modeling due the choice of training samples for the setup of the surrogates. This is included in the convergence analysis by repeating the regression procedure for each number of considered simulations  $10^3$  times. The resulting lifetime DEL are normalized with respect to the lifetime DEL obtained from the Monte Carlo baseline. Two different types of radial basis networks are investigated here in order to explain some fundamental characteristics specific to radial basis networks. As described before, the specific characteristics of radial basis functions are the defined spread of the radial basis neurons and the target accuracy. In this comparison the first setup (RBN1) is a 20% spread value of neurons and a target error of  $10^{-3}$ . The second setup (RBN2) has a spread of 10% and a target error of  $10^{-4}$  and hence is more accurate, but computationally expensive and may be prone to overfitting.

Figure 7 shows the results of the convergence analysis which indicates that already based on 100 design points (evaluated using 3 seeds each), the error of the prediction lifetime DEL is likely to be below 5% for RBN1. For RBN2 an underestimation is visible if only a limited number of simulations are available for training. This indicates that if the spread of neurons is small, a larger number of training points is required in order to reach sufficient coverage of the design space.



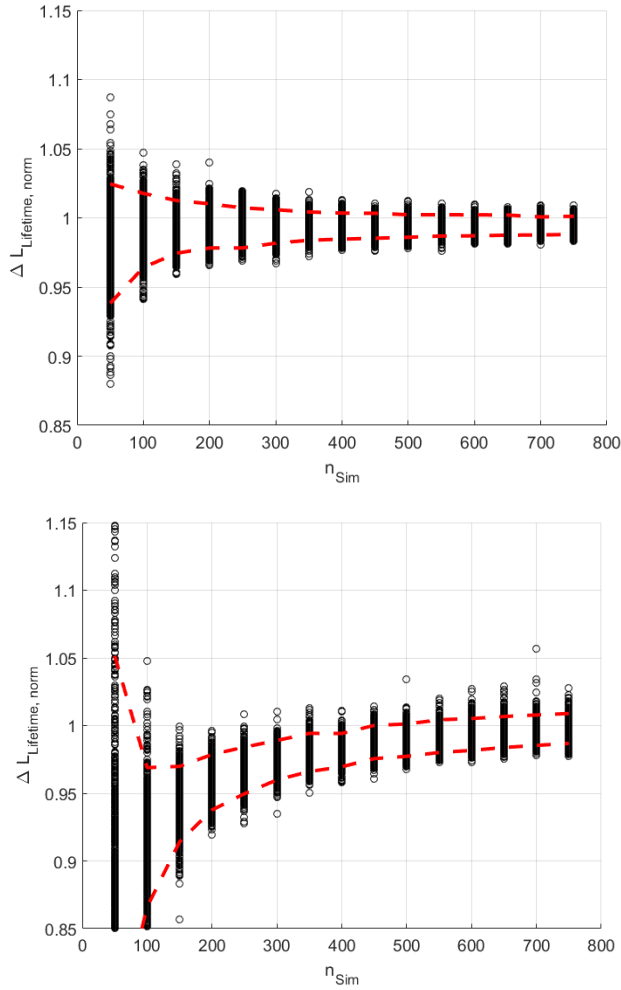


FIGURE 7: CONVERGENCE OF ERROR OF LIFETIME DEL DETERMINED FROM SURROGATE MODEL WITH INCREASING NUMBER OF SIMULATIONS. RED DOTTED LINES INDICATING 5TH AND 95TH PERCENTILES. RESULTS SHOWN FOR TOWER BASE FORE-AFT BENDING MOMENT. TOP: RBN1, BOTTOM: RBN2

The prediction error of the lifetime DEL may be weighted again with the SN-slope in order to show convergence of the damage prediction error:

$$\frac{D_1}{D_2} = \left( \frac{\Delta L_{DE,1}}{\Delta L_{DE,2}} \right)^m. \quad (7)$$

Following Eq. (7), the normalized damage error is equal to the normalized error in DEL calculation to the power of  $m$ .

Figure 8 shows the convergence of damage adjusted error for all sensors investigated as part of this study.

For RBN1, it indicates a bias of up to around 5% compared to the Monte Carlo reference study for the lower load locations (i.e. for the tower base and fairlead the center of the plotted percentile

curve-pairs stays below 1 even for a large number of simulations). The bias is linked to the lower target accuracy of RBN1. This leads to consistent underprediction of the load response (may be found through error analysis). Apart from this bias, the convergence behavior is similar for all load locations. This being said, sensors which are located lower, show a slower speed of convergence. This is due to the increased impact of the wave environment in the lower parts of the system and hence the increased complexity of the response function.

For RBN2, the underprediction for the fairlead tension persists even after the before mentioned underprediction. A clear reason for this bias could not be found. It is however likely that an underprediction is done in a region with high occurrence probability and fatigue loads (i.e. around rated wind speed). There, even small errors will be penalized strongly. It could be a possibility for improvement of surrogates to take this occurrence probability into account in the curve fitting procedure (e.g. by adding occurrence probability and load based weights to the error calculation).

Overall, the damage prediction for two different approaches results in a damage model accuracy of around  $\pm 10\%$  using below 1000 design points. In this way, when comparing the results with the Monte Carlo reference study performed in [18], it can be seen that the convergence behavior using surrogate models may be comparable to the direct Monte Carlo approach (there, 200-500 design points were found to be sufficient for fatigue design). Due to the flexible application of surrogates to other environments, it would thus be recommended to choose a surrogate modeling over direct Monte Carlo simulation: The break-even point would be reached with the second or third load evaluation. It should also be possible to add design points sequentially, increasing the design space and hence the region of model applicability (adding dimensions, however is easily implemented). However, it is highlighted that care is to be taken in the model derivation and a thorough error analysis and quality evaluation needs to be performed in order to ensure that the results are trustworthy. In this way future research on surrogate modeling on loads of floating wind turbines should be followed in order to identify standardized procedures for modeling and model evaluation. Some suggestions have been presented here, but will need extension.

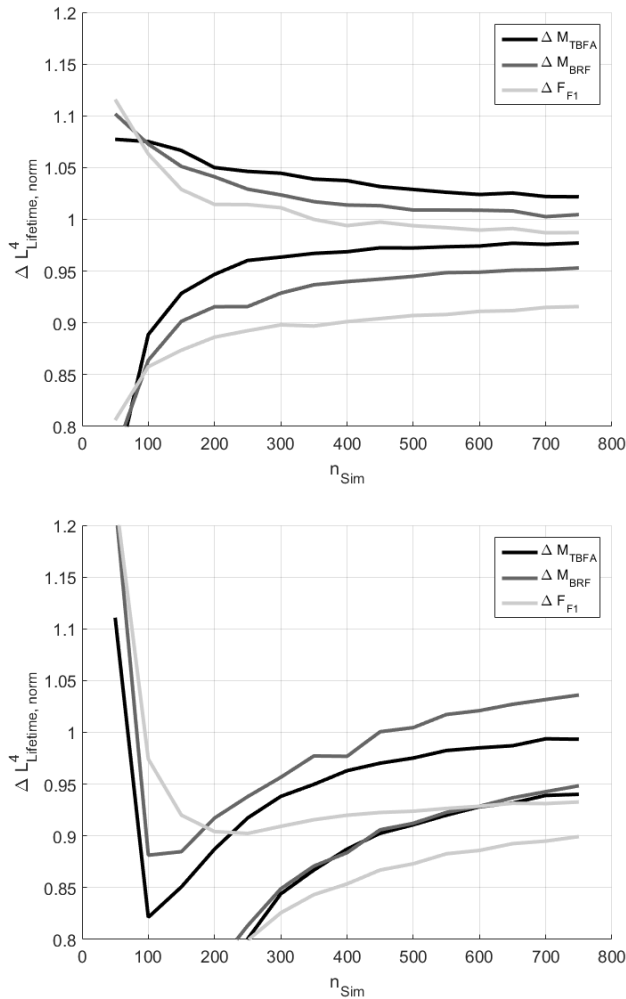


FIGURE 8: CONVERGENCE OF DAMAGE ADJUSTED ERROR OF LIFETIME DEL DETERMINED FROM SURROGATE MODEL WITH INCREASING NUMBER OF SIMULATIONS. SHOWING 5TH AND 95TH PERCENTILES FOR DIFFERENT SENSORS. TOP: RBNI, BOTTOM: RBN2

## SUMMARY, CONCLUSIONS AND OUTLOOK

A methodology for a probabilistic fatigue load assessment based on surrogate modeling was presented and evaluated. The chosen procedure here is based on the following steps (indicating the different choices made in the present work):

- 1) **Define load ranges linked to distinctive system characteristics**  
(e.g. below rated, transition, above rated)
- 2) **Set up joint probability environmental model**  
(e.g. Nataf transformation)
- 3) **Determine probabilistic design space limits for each load range**  
(e.g. 1<sup>st</sup> and 99<sup>th</sup> percentiles)

- 4) **Obtain design points based on quasi-random sampling procedures**  
(e.g. Sobol' sequences)
- 5) **Calculate fatigue load response at design points using time-domain aero-servo-hydro-elastic simulation tools**  
(e.g. FAST8.16)
- 6) **Set up surrogate model**  
(e.g. radial basis network)
- 7) **Combine probability and surrogate model to obtain lifetime fatigue response**  
(e.g. Monte Carlo simulation)

Several notes on the above points are given here for context and outlook:

(1) and (3) Load ranges are used at this stage to ensure accuracy for the different system behavior in different operation conditions. Also, the design space is more efficiently covered, partially taking into account the probabilistic nature of the environment (i.e. no large wave heights at very low wind speeds). It may well be that future applications will not need the load ranges. For now, it is considered in order to ensure robust results.

(2) Different environmental models may be used. In particular the Rosenblatt transformation may be of interest for higher accuracy. This introduces further modeling error (and need for convergence studies), due to the implied binning procedure and questionable interpolation of statistical properties. If the Nataf model is assumed as a correct representation of the environment, a Monte Carlo integration can be easily applied. In this way, Nataf modeling is preferred in an academic environment, as long as the environmental model is not under investigation.

(6) Alternatives may be more comprehensive artificial neural networks, support vector machines, kriging, polynomial chaos expansion or others.

It was shown that for the considered site, environmental conditions and system, a limited number of design points is sufficient to achieve a converged and accurate prediction of the fatigue damage of different relevant components of a floating wind turbine system. Detailed error analysis is a helpful tool in order to determine the quality and applicability of a surrogate model. Further study will be necessary to support the findings from this study, considering different environments, systems and additional loading positions and materials. Because the convergence characteristics are similar to the results from a direct Monte Carlo simulation study, the results of this work promise the general applicability of surrogate models. If this insight stands after further investigation, surrogate models combined with classified met-ocean conditions may largely facilitate the early design of floating wind turbine systems. Then, the fatigue response model of a system may be directly combined with any site within a given met-ocean class and an estimate for the lifetime fatigue damage can be obtained without any additional simulation effort.

## ACKNOWLEDGMENTS

The present work is carried out as part of the LIFES50+ project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640741. The funding and support is gratefully acknowledged. Also, we are grateful to Dr. techn. Olav Olsen AS for the permission and contribution to set up the public semi-submersible design based on their concept of the OO Star Wind Floater ([www.olavolsen.no](http://www.olavolsen.no)).

## REFERENCES

- [1] Det Norske Veritas AS, "Design of Floating Wind Turbine Structures," *DNV-OS-J103*, 2013.
- [2] International Electrotechnical Commission, "Wind Turbines - Part 1: Design Requirements," *61400-01*, 2005.
- [3] P. Passon and K. Branner, "Condensation of long-term wave climates for the fatigue design of hydrodynamically sensitive offshore wind turbine support structures," *SHIPS OFFSHORE STRUC*, vol. 11, no. 2, pp. 142-166, October 2016.
- [4] P. Passon, "Damage equivalent wind-wave correlations on basis of damage contour lines for the fatigue design of offshore wind turbines," *RENEW ENERG*, vol. 81, pp. 723-736, September 2015.
- [5] M. J. Kühn, "Dynamics and design optimisation of offshore wind energy conversion systems," Phd Thesis. Delft University Wind Energy Research Institute, 2001.
- [6] M. Seidel, "Wave induced fatigue loads on monopiles - New approaches for lumping of scatter tables and site specific interpolation of fatigue loads," in *Conference Proceedings IWE*, Hanover, 2014.
- [7] International Electrotechnical Commission, "Wind Turbines - Part 3: Design Requirements for Offshore Wind Turbines," *61400-03*, 2009.
- [8] D. Kaufer and P. W. Cheng, "Validation of an Integrated Simulation Method with High Resolution Load Measurements of the Offshore Wind Turbine REpower 5M at Alpha Ventus," *Journal of Ocean and Wind Energy*, 2014.
- [9] K. Müller, M. Reiber and P. W. Cheng, "Comparison of Measured and Simulated Structural Loads of an Offshore Wind Turbine at Alpha Ventus," *INT J OFFSHORE POLAR*, vol. 03, pp. 209 - 218, September 2016.
- [10] D. Matha, K. Müller, F. Lemmer and G. Perez Moran, "Comparative Analysis of Industrial Design Methodologies For Fixed-Bottom and Floating Wind Turbines," in *Proceedings of the 35th International Conference on Ocean, Offshore and Arctic Engineering*, Busan, South Korea, 2016.
- [11] P. Passon and K. Branner, "Load calculation methods for offshore wind turbine foundations," *Ships and Offshore Structures*, 2014.
- [12] J. Häfele, C. Hübler, C. Gebhardt and R. Rolfes, "Efficient Fatigue Limit State Design Load Sets for Jacket Substructures Considering Probability Distributions of Environmental States," in *Twenty-seventh International Ocean and Polar Engineering Conference*, San Francisco, 2017.
- [13] L. Barj, S. Stewart, G. Stewart, M. Lackner, J. Jonkman, A. Robertson and D. Matha, "Wind/Wave Misalignment in the Loads Analysis of a Floating Offshore Wind Turbine," in *32nd ASME Wind Energy Symposium, AIAA SciTech Forum*, (AIAA 2014-0363), 2014.
- [14] D. Matha, F. Sandner and D. Schlipf, "Efficient critical design load case identification for floating offshore wind turbines with a reduced nonlinear model," *Journal of Physics: Conference Series*, vol. 555, no. 1, 2012.
- [15] D. Matha, Impact of Aerodynamics and Mooring System on Dynamic Response of Floating Wind Turbines, 2017.
- [16] J. Kleijnen, Design and analysis of simulation experiments, New York: Springer, 2008.
- [17] P. A. Graf, G. Stewart, M. Lackner, D. Katherine and P. Veers, "High-throughput computation and the applicability of Monte Carlo integration in fatigue load estimation of floating offshore wind turbines," *Wind Energy*, vol. 19, no. 5, 2016.
- [18] K. Müller and P. W. Cheng, "Application of a Monte Carlo Procedure for Probabilistic Fatigue Design of Floating Offshore Wind Turbines (under review)," *Wind Energ. Sci. Discuss.*, 2017.
- [19] K. Müller, M. Dazer and P. W. Cheng, "Damage Assessment of Floating Wind Turbines Using Latin Hypercube Sampling," *Energy Procedia*, no. 137, pp. 119-133, October 2017.
- [20] D. Robinson and C. Atcitty, "Comparison of quasi-and pseudo-Monte Carlo sampling for reliability and uncertainty analysis," *40th Structures, Structural Dynamics, and Materials Conference and Exhibit*, 1999.
- [21] S. Chen, C. Cowan and P. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Transactions on Neural Networks*, vol. 2, no. 2, pp. 302 - 309, March 1991.
- [22] C. Bak, F. Zahle, R. Bitsche, K. Taeseong, A. Yde, L. Christian, H. Anand, N. Morten and H. Hansen, "Description of the DTU 10 MW Reference Wind Turbine," DTU Wind Energy, 2013.
- [23] L. project, "<http://www.lifes50plus.eu>," [Online].
- [24] A. Pegalajar-Jurado, H. Bredmose, M. Borg, H. Andersen, T. Landbo and J. G. S. Straume, "State-of-the-art model for the LIFES50+ OO-Star Wind Floater Semi 10MW floating

- wind turbine. Under review," in *DeepWind conference*, 2018.
- [25] G. Ramachandran, L. Vita, A. Krieger and K. Müller, "Design Basis for the Feasibility Evaluation of Four Different Floater Designs," *in review*, 2017.
- [26] G. M. Stewart, A. Robertson, J. Jonkman and M. A. Lackner, "The creation of a comprehensive metocean data set for offshore wind turbine simulations," *WIND ENERGY*, vol. 6, no. 19, pp. 1151-1159, 2016.
- [27] P. Gomez, G. Sanchez, A. Llana and G. Gonzales, "Oceanographic and meteorological conditions for the design," *LIFES50+ Deliverable 1.1*, 2015.
- [28] I. M. Sobol, "On the distribution of points in a cube and the approximate evaluation of integrals," *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, 1967.
- [29] P. Bratley and B. Fox, "Implementing Sobols Quasirandom Sequence Generator (Algorithm 659)," *ACM T MATH SOFTWARE*, vol. 14, no. 1, pp. 88-100, 2003.
- [30] G. M. Stewart, "Design Load Analysis of Two Floating Offshore Wind Turbine Concepts," PhD Thesis, University of Massachusetts, 2016.
- [31] M. Hudson Beale, M. T. Hagan and H. B. Demuth, *Neural Network Toolbox, User's Guide*, MathWorks, 2016.
- [32] Li, Michael M., and Brijesh Verma. "An improved RBF neural network approach to nonlinear curve fitting." *International Work-Conference on Artificial Neural Networks*. Springer, Cham, 2015.
- [33] Smolka, U., Kaufer, D., and Cheng, P. W. "Are sea state measurements required for fatigue load monitoring of offshore wind turbines?" *Journal of Physics: Conference Series*. Vol. 555. No. 1. IOP Publishing, 2014.
- [34] Haid, Lorentz, et al. "Simulation-Length Requirements in the Loads Analysis of Offshore Floating Wind Turbines." 2013. *ASME 2013 32nd International Conference on Ocean, Offshore and Arctic Engineering*. American Society of Mechanical Engineers.
- [35] Faerron Guzmán, R, Müller, K and Cheng, P. W., „Simulation requirements and relevant load conditions in the design of floating offshore wind turbines (in review),“ *Proceedings of the 37th International Conference on Ocean, Offshore & Arctic Engineering*, 06 2018.
- [36] Mata, J. "Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models" *Engineering Structures*. Vol. 33 No. 3. March 2011.
- [37] Bishop, C.M., Roach, C.M., "Fast curve fitting using neural networks" *Review of Scientific Instruments*. Vol. 63. No. 10. 1992.
- [38] Chen, S., Cowan, C.F.N., Grant, P.M., "Orthogonal least squares learning algorithm for radial basis function networks" *IEEE Transactions on Neural Networks*. Vol. 2. No. 2. March 1991.
- [39] Trejo-Caballero, G., H. Rostro-Gonzales, C.H. Garcia Capulin, O.G. Ibarra-Manzano, J.G. Avina-Cervantes, and C. Torres-Huitzil. "Automatic Curve Fitting Based on Radial Basis Functions and a Hierarchical Genetic Algorithm." *Mathematical Problems in Engineering*, December 2015.