

# Gaussian process regression for fatigue reliability analysis of offshore wind turbines

D. Wilkie<sup>a</sup>, C. Galasso<sup>a</sup>

<sup>a</sup>*Department of Civil, Environmental and Geomatic Engineering, University College London, Gower Street, London, WC1E 6BT, United Kingdom*

---

## Abstract

The fatigue limit state (FLS) often drives the design of offshore wind turbine (OWT) substructures in European waters. Assessing fatigue damage over the intended design life of an OWT is computationally expensive, primarily as dynamic structural analyses have to be run for a large number of stochastic wind and wave loading conditions. This makes structural reliability assessment for the FLS a challenging task. In addition to evaluating load-induced fatigue damage, simulation-based structural reliability analysis also requires sampling of random variables that model uncertainties in the capacity of OWT structural components. To this aim, we develop and validate a computational framework for OWT fatigue reliability analysis that relies on Gaussian process (GP) regression to build surrogate models of load-induced fatigue damage. We demonstrate that the proposed approach can reduce the computational effort required to evaluate FLS reliability with high accuracy through application to three plausible offshore wind farm sites in Europe. The sensitivity of various goodness-of-fit metrics to different model assumptions is investigated to further reduce the computational effort required to perform GP regression/predictions. The results from this study can provide guidance for practical applications of the proposed framework in OWT projects.

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Fatigue assessment of OWTs</b>	<b>3</b>
2.1	Structural simulations	3
2.2	Fatigue damage estimation	4
<b>3</b>	<b>Problem statement</b>	<b>5</b>
3.1	Limit state definition	5
3.2	Dealing with the large number of simulations	6
<b>4</b>	<b>Methodology</b>	<b>7</b>
4.1	Considered environmental variables	7
4.2	Calculation framework	8
4.3	Considered GP surrogate model	8
4.4	Considered metrics for assessing the goodness of fit of the GP surrogate model	10

<b>5</b>	<b>Illustrative application</b>	<b>10</b>
5.1	Site and environmental conditions . . . . .	10
5.2	Fatigue damage calculation . . . . .	11
5.2.1	Structural model . . . . .	11
5.2.2	Load calculation . . . . .	13
5.2.3	Fatigue damage . . . . .	13
5.3	Training and validation data sets . . . . .	14
<b>6</b>	<b>Results and discussion</b>	<b>14</b>
6.1	Proposed surrogate model . . . . .	14
6.1.1	Different sites . . . . .	16
6.2	Fatigue reliability analysis . . . . .	17
<b>7</b>	<b>Conclusions</b>	<b>17</b>
<b>8</b>	<b>Acknowledgments</b>	<b>18</b>
<b>A</b>	<b>Gaussian process regression</b>	<b>21</b>

## 1. Introduction

Offshore wind energy is a rapidly growing form of electricity generation. This is especially true in Europe, where a large number of offshore wind farms (OWFs) are due to be installed in the North Sea basin [1]. Bidding for new OWFs is commonly driven by the levelized cost of energy, i.e., the amount of money per unit of electricity that must be earned over the design life of the assets to cover all the installation and operational costs. It is therefore important to ensure this value is minimized at the design stage while maintaining acceptable levels of structural safety, so as not to incur large repair or replacement costs over the intended design life of the farm.

The design requirements for offshore wind turbine (OWT) substructures are currently specified by prescriptive codes, e.g., the International Electrotechnical Commission (IEC) 61400-3 [2], or the Det Norske Veritas Germanischer Lloyd (DNVGL) DNVGL-ST-0126 [3]. These codes of practice implement the load-resistance factor design (LRFD) approach to attain a certain level of structural safety. This is ensured by prescribing various load cases encompassing the environmental and operating conditions expected to be encountered during the intended 20-25 year design life of the asset. The uncertainties affecting load-induced structural demands and structural capacities are accounted for by using partial safety factors that should be (in principle) calibrated to ensure a desired level of structural safety with consideration of the consequences of failure in terms of economic losses and risk of human lives [4]. For instance, the current partial safety factors are intended to provide a yearly probability of failure of  $5 \times 10^{-4}$  in IEC 61400-1 and an annual probability of failure of  $10^{-4}$  in DNVGL-ST-0126. However, the current set of partial safety factors used in the design of OWTs is often a combination of those used in the offshore oil and gas industry and those used to assess onshore wind turbines [5].

An essential step in understanding the suitability of the current code-based partial safety factors or to promote novel performance-based design and assessment approaches for OWTs (to decide, for instance, about lifetime extension of ageing OWFs) is to perform an explicit structural reliability assessment of OWT structural components. To date, most studies have focused on the OWT ultimate limit state (ULS) (e.g.,

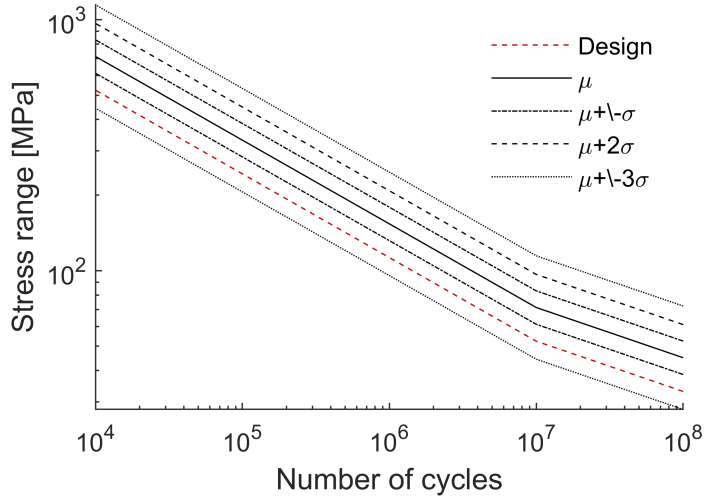
[6], among many others), including the development and use of statistical surrogate models to improve the computational efficiency of the assessment (e.g., [7]). These studies have utilized fully-coupled time-history simulations (i.e., modeling the rotor, structure, foundation and control system in a single simulation; e.g., [5]), which is considered the state-of-the-art approach in both design and assessment of OWTs [8]. However, there are very few studies assessing the fatigue limit state (FLS) reliability of OWTs, such as Gintautas and Sørensen [9]; such FLS often drives the design of OWT substructures in European waters [10]. Such an assessment should include all the environmental conditions (in terms of acting wind and wave loading and its variability) that are expected to be encountered during an OWT operation, requiring in the order of >50,000 time-domain structural simulations [8] for a single OWT. Additionally, the IEC 61400-3 code requires performing long dynamic analyses to ensure a stable estimation of fatigue damage accumulating over the OWT intended design life. The large number of environmental conditions combined with the need for long time-history analyses makes the assessment of an OWT at the FLS computationally expensive and impedes a full simulation-based structural reliability assessment.

Motivated by the need to improve the computational efficiency in assessing FLS structural reliability of OWTs, one application of which could be the improved calibration of code-based partial safety factors, this paper introduces a computational framework for assessing OWT fatigue reliability utilizing a surrogate modeling approach to predict load-induced lifetime fatigue damage. Most installed OWTs across the world use monopile foundations, and these are the focus of this study. The basics of the fatigue assessment for OWTs are briefly reviewed in Section 2; a discussion on the challenges in assessing the FLS and the problem statement of this study are presented in Section 3. Then, the proposed computational framework is introduced in Section 4, including an introduction to Gaussian process (GP) regression. An application, combining illustrative environmental conditions and a computational model of an archetype OWT, is introduced in Section 5. The accuracy of different GP modeling assumptions are discussed in Section 6 to guide practical application of the proposed framework. The number of samples necessary to estimate OWT reliability and the applicability of a surrogate model trained at one site to others with different environmental conditions are also evaluated and results of the analysis are critically discussed.

## 2. Fatigue assessment of OWTs

### 2.1. Structural simulations

During their design life, OWTs are exposed to highly-variable loading conditions from wind and waves. Structural assessment of OWTs generally utilizes time-history simulations to capture the fluctuating forces in a structural component resulting from different environmental conditions [8]. The variable wind and wave loading is modeled through a long-term and a short-term component [11]. The long-term component captures the site-specific climate conditions and includes mean wind speed (at a reference height), turbulence intensity, significant wave height, peak spectral period and wind-wave inflow angle. These variables are assumed as stationary over a given time window, typically 10-minutes for the wind conditions [12]. In state-of-the-art analysis approaches [8], short-term fluctuations are captured by generating wind-wave time series to be input in the time-history structural simulations. Specifically, the fluctuating environmental conditions are modeled through spectra with parameters depending on the long-term component of the environmental conditions. For instance, the Kaimal spectrum is frequently used to model wind turbulence [12] and depends on the mean wind speed and turbulence intensity. To accurately predict the total lifetime fatigue loading, time-history structural simulations need to be run at every foreseeable combination of environmental conditions. Additionally, in fully-coupled software environments, such as OpenFAST [13], the inverse Fourier transform is used to convert the environmental spectra into wind-wave time series. These are stochastic as the phase angle of the transformation is not known in advance and is sampled from an assumed probability distribution (e.g., uniform) [14]. Consequently, each analysis has to be repeated for every set of environmental conditions in order to produce a stable estimation of the total fatigue damage (where the number of analysis repetitions is referred to as “seeds”). The underlying process is stationary as a result of the inverse Fourier transform, and convergence of damage with the use of increasing numbers of seeds has been investigated by Zwick and Muskulus [15]. In particular, the provisions in IEC 61400-3 [2] require a large number of structural analyses to model different combinations of the environmental spectral parameters and



**Figure 1:** Scatter in the SN-curve when the full distribution of x-axis intersections is utilized. The 2.5th percentile curve used in structural design is highlighted as the mean curve minus two standard deviations [17].

for the structural analysis to be repeated at each, resulting in the >50,000 simulations mentioned previously.

## 2.2. Fatigue damage estimation

The output from a time-history structural simulation is used to predict fatigue-damage accumulation by first extracting stress cycles, for instance by using the rainflow counting algorithm [14]. Subsequently, the fatigue damage caused by each stress cycle can be estimated using the stress-life approach, where the number of load cycles to fatigue failure ( $N$ ) resulting from a constant-amplitude load is a function of the applied stress range ( $\sigma_r$ ); this relationship is described by an SN-curve [16], shown in Figure 1 and described by the following equation:

$$N = K_\mu \sigma_r^{-m}. \quad (1)$$

In Eq. (1), the  $m$ -parameter is the slope of the SN-curve and  $K_\mu$  is a constant determining the location of the  $x$ -axis intersection. The ratio between the number of applied cycles ( $n$ ) and the tolerable number of cycles ( $N$ ) is assumed “to consume” a fixed proportion of the fatigue life. In the case of stochastic environmental inputs to each time-history analysis, the loading amplitude varies with time. The composite fatigue damage can be predicted using Miner’s rule [16], which assumes that the damage caused by individual stress cycles can be summed, resulting in the following equation:

$$D_{kj} = \sum_{i=1}^{N_{cyc}} \frac{n_{kji}}{N_{kji}} = \frac{1}{K_\mu} \sum_{i=1}^{N_{cyc}} \frac{n_{kji}}{\sigma_{r,kji}^{-m}}. \quad (2)$$

In Eq. (2),  $D_{kj}$  is the total fatigue damage calculated over all  $N_{cyc}$  stress ranges (each indexed by  $i$ ) occurring during a time-history analysis conducted at a set of environmental parameters (indexed by  $j$ ) and different seeds (indexed by  $k$ ). The fatigue damage over the design life ( $D_{kj,life}$ ) can be calculated by linearly scaling  $D_j$  from the time-history structural simulation length ( $T_{analysis}$ ) to the intended (i.e., assumed) life of the OWT ( $T_{life}$ ):

$$D_{kj,life} = \frac{T_{life}}{T_{analysis}} D_{kj}. \quad (3)$$

Fatigue damage is also commonly represented through the damage equivalent load (DEL), i.e., the constant-amplitude stress causing the same fatigue damage as the variable amplitude loading at a reference number of tolerable cycles ( $N_{eq}$ ) [11]. This number is typically easier to communicate, as fatigue loading calculated in terms of total damage can assume very small values.

$$DEL_{kj,life} = \left( \frac{K_\mu}{N_{eq}} D_{kj,life} \right)^{1/m}. \quad (4)$$

The total fatigue damage, calculated across a set of structural analyses, is:

$$D_{total} = \frac{1}{N_{seed}} \sum_{j=1}^{N_{sim}} \sum_{k=1}^{N_{seed}} D_{kj,life} Pr_j \approx \frac{1}{N_{sim} N_{seed}} \sum_{j=1}^{N_{sim}} \sum_{k=1}^{N_{seed}} D_{kj,life}. \quad (5)$$

In Eq. (5),  $N_{sim}$  is the number of fatigue environmental states (i.e., the number of simulated wind and wave conditions), and  $N_{seed}$  is the considered number of seeds. The composite damage is weighted by the occurrence probability of each fatigue sea state ( $Pr_j$ ). This can be approximated using Monte Carlo sampling, provided that  $D_{kj,life}$  is evaluated at samples drawn from the joint-probability distribution of wind and wave conditions.

A given structural component is assumed to fail when the fatigue damage reaches a pre-specified threshold, typically assumed equal to one. Equivalently, failure in terms of DEL occurs when the equivalent constant amplitude stress is larger than the SN-curve stress for the selected number of equivalent cycles. For structures assessed using the DNVGL fatigue code [17], a predicted fatigue failure corresponds to a crack growing through the thickness of the structures cross-section.

### 3. Problem statement

#### 3.1. Limit state definition

Assessing whether or not a structural component of an OWT is expected to fail due to fatigue during its operating life can be achieved by introducing a limit state (or performance) function [18]. In particular, a fatigue limit state function,  $G(\mathbf{X})$ , can be defined as in Eq. (6), based on a vector of random variables  $\mathbf{X}$ , capturing the uncertainties in load-induced structural demand and structural capacity. Specifically, the demand term is the fatigue damage incurred over the component intended design life ( $D_{total}$ ; i.e., Eq. (5)) and the capacity term is the fatigue damage resistance of the component ( $X_\delta$ ):

$$G(\mathbf{X}) = X_\delta - X_{SN}(D_{total,SN1}(X_{Str}X_{SCF}X_{GP})^{m_1} + D_{total,SN2}(X_{Str}X_{SCF}X_{GP})^{m_2}). \quad (6)$$

Eq. (6) accounts for the fact that a typical SN curve for steel consists of two slopes ( $m_1$  and  $m_2$ ), from which a damage term, Eq.(2), is calculated for each time history.

Significant uncertainty exists also in the prediction of a structural component fatigue (material) properties, and these are captured as random variables, i.e., the  $X_{SN}$  and  $X_\delta$  terms in Eq. (6). These random variables are defined following the recommendations of DNVGL [17] and are presented in Table 1. The SN-curve  $x$ -axis intersect ( $X_K$ ) variable is modeled following the guidance in DNVGL [17], which utilizes a single standard deviation for both slopes of the SN curve. These curves are assumed to be fully correlated following guidance in the JCSS Probabilistic Model Code [19]. A different modeling choice is to consider larger uncertainty for the second slope, i.e., using a larger standard deviation for the  $X_K$  variable corresponding to the second slope [20]. However, the DNVGL approach is utilized here because it is widely used in the design/assessment of offshore structures. It should be noted that Eq. (6) was simplified by including the re-weighting of the SN-curve  $x$ -axis intersect ( $X_K$ ) inside the random variable  $X_{SN}$  by (1) calculating the reciprocal of  $X_K$ ; and (2) evaluating the parameters of this distribution multiplied by a constant ( $K_\mu$ , for each SN-curve slope). Additionally, the modeling uncertainty introduced as a result of limitations in the structural analysis are taken from DNVGL guidance on probabilistic fatigue inspection [21]. In particular,  $X_{SCF}$  is the uncertainty

**Table 1:** Statistical properties of the random variables modeling the fatigue material behavior. \* = the parameter varies with the SN-curve intersection. \*\* Quantified during model training, as described in Section 6.1.

Variable	Reference	Distribution	Mean	Standard deviation
Tolerable damage ( $\ln(X_\delta)$ )	[17]	Normal	1.05	0.32
SN intersection ( $\log_{10}(X_K)$ )	[17]	Normal	$f(K_\mu)^*$	0.2
SN uncertainty ( $\ln(X_{SN})$ )	N/A	Normal	0.45	0.22
GP uncertainty ( $X_{GP}$ ) **	N/A	Normal	0.998	0.01
Stress concentration factor uncertainty ( $X_{SCF}$ )	[21]	Normal	1	0.12
Mechanical model uncertainty ( $X_{Str}$ )	[21]	Normal	1	0.05
SN slope 1 ( $m_1$ )	[17]	Deterministic	3	N/A
SN slope 2 ( $m_2$ )	[17]	Deterministic	5	N/A
SN slope 1 intersection ( $K_{\mu 1}$ )	[17]	Deterministic	$10^{11.61}$	N/A
SN slope 2 intersection ( $K_{\mu 2}$ )	[17]	Deterministic	$10^{15.35}$	N/A

due to the stress calculation and  $X_{Str}$  models uncertainty resulting from the structural analysis. The modeling uncertainty due to the use of a surrogate model for computing fatigue damage is captured by the  $X_{GP}$  variable, which is described in Section 6.1. These are multiplicative factors that act on the stress calculation and therefore are raised to the power of the relevant SN-curve slope. Other work, such Velarde et al. [22], include additional random variables such as wave loading uncertainty, which could be modeled as an additional multiplicative factor in the limit state equation introduced above.

The main challenge in assessing fatigue reliability for OWTs is the computation of the  $D_{total}$  terms (for each SN-curve slope) in the fatigue limit state function. As discussed in Section 2.1, a large number of structural simulations are typically required to accurately assess fatigue damage and this is combined with the computational expense of fully-coupled time-history simulations. Even when the large number of simulations required by IEC 61400-3 is performed, there are still various simplifying assumptions contained within the specification of load cases. For instance, the IEC code allows one to consider discrete 2m/s-wide bins for the mean wind speed and use of a single combination of wave parameters for each bin, which is a strong simplification as the environmental parameters vary continuously within bins [10].

### 3.2. Dealing with the large number of simulations

In practice, the large computational burden required to assess the FLS can be reduced by either making the analysis more computationally efficient, or by reducing the number of simulations to be run, using:

- Load-case reduction;
- Simplified mechanical models, based on frequency-domain simulations;
- Surrogate models, also commonly referred to as metamodels.

Load-case reduction mitigates the computational burden of evaluating the FLS by combining those load cases which produce similar outputs. This approach was first developed for application to OWTs by Kühn [23] and extended by Damgaard et al. [24] to encompass multi-directional environmental conditions. A set of 2,000 load cases were reduced by a factor of 10, with a reported error of less than 1% in terms of DEL [24]. Häfele et al. [25] found that similar reductions were possible for an OWT located at a different North Sea site, suggesting that the number of load cases can be reliably reduced. However, the computational burden for FLS assessment remains large as the full set of load cases need evaluated to calibrate the reduced set.

Replacing an expensive aero-elastic dynamic model with a simplified mechanical model results in simulations running much faster, enabling the evaluation of a large number of environmental load cases. These techniques have been applied to OWTs using frequency-domain calculations, which are summarized comprehensively by Baltrop [14] and Tempel [26]. Kühn [23] and Tempel [26] developed early approaches, where analysis of the fatigue damage caused by wind and waves was carried out separately. Similarly Yeters et al. [27] and

Ziegler et al. [28] developed variants based on decoupled frequency-domain models. This approach has been applied in FLS reliability assessment by Marquez-Dominguez and Sørensen [29], Dong et al. [30] and Oest et al. [31] for different OWT components. Muskulus and Schafhirt [32] developed a fatigue reliability model for OWTs on jackets substructures. Their study included uncertainty in both the environmental loading (e.g., wind and waves) and material properties. However none of these approaches can include a fully-coupled dynamic simulation, which is considered the state-of-the-art technique in OWT project certification [8].

Surrogate models approximate complex systems attempting to mimic their behavior yet remaining computationally cheap to evaluate. In other words, a surrogate model is nothing but a model of a model: it constructs a relationship between the design variables and corresponding responses using relatively few, intelligently-chosen sampling points. The original complex model is effectively replaced by a transfer function. Different surrogate models have been used to represent the outputs from complex structural analysis packages for OWTs. This includes using linear surrogate models to assess onshore wind turbines, using piece-wise fitting to represent nonlinearities in the response. Both Toft et al. [33] and Zwick and Muskulus [34] developed surrogate models using this approach, where a separate model was fit for each mean wind speed bin. However, the additional non-linearity in an OWTs response due to wave loading [8] means that fitting separate surrogate models may not be feasible. Complex linear surrogate models can capture a larger range of behaviors and may be able to better represent complex simulators, such as the output from a time-history OWT simulation. One highly-flexible type of surrogate model is a regression built on the use of a GP, which provides an adaptable and localized surrogate modeling approach. GP regression has been proven capable of addressing reliability calculations in problems with nonlinear limit state functions and large number of model parameters. For instance, in the case of OWTs, Huchet et al. [35] found that a GP model could estimate fatigue damage using four environmental parameters (mean wind speed, wind-wave direction, significant wave height and peak spectral period), and reported a low modeling error of around 4.4%. Similarly, Teixeira et al. found that a GP was suitable for modeling fatigue damage, represented by the mean wind speed and turbulence intensity [36]. A different approach was applied by Häfele et al. [37] who used a GP regression to represent post-processed fatigue damage against different geometrical parameters in an optimization procedure for OWTs on jacket substructures. Hence, GP regression represents a feasible approach for modeling fatigue loads on fixed OWT and to be used in FLS reliability assessment of OWTs, as proposed in this study. The basics theory supporting GP regression is provided in Appendix A; interested readers should refer to Rasmussen et al. [38] for a broader view on GP theory and discussions on underlying hypotheses.

## 4. Methodology

### 4.1. Considered environmental variables

A GP surrogate model can be used to estimate the fatigue damage ( $D_{kj}$ ; Eq. (2)) at a set of environmental conditions; estimates of  $D_{kj}$  can then be used to rapidly compute the total lifetime fatigue damage ( $D_{total}$ ; Eq. (5)). This allows one to evaluate efficiently the limit state function of Eq. (6), without requiring a time-history analysis for each sample of the environmental conditions. However, the GP surrogate model must be trained against all the environmental conditions that are important for predicting the OWT response. Hübler et al. [39] applied a four stage sensitivity analysis starting with 100 variables (covering wind, waves, soil and geometrical properties) and ranked the 14 which had the greatest impact on the FLS response. For an OWT on a monopile substructure, mean wind speed ( $V_w$ ), peak spectral period ( $T_p$ ), significant wave height ( $H_s$ ), wind direction ( $\theta_{wind}$ ) and wave direction ( $\theta_{waves}$ ) were the main variables identified by the study, and were not ranked in order of importance. Then, soil properties, marine growth thickness, pile dimensions and water depth were found to be important. Turbulence intensity ( $T_i$ ) was not on the final list of important variables, but has been consistently highlighted as a critical environmental quantity in studies of onshore wind turbines [33, 40]. Brandt et al. [41] conducted a global sensitivity analysis for an OWT on a jacket substructure and identified similar main environmental variables to Hübler et al. [39], but reported that the FLS was insensitive to  $\theta_{wind}$ ,  $\theta_{waves}$  and yaw error. It should be noted that the ranking of variables in a global sensitivity analysis may not provide the same results as a sensitivity analysis for fatigue damage based on structural reliability estimates; however, a global sensitivity analysis does indicate



which parameters are crucial for the prediction of fatigue damage and therefore should be included in the surrogate model definition.

This study models the environmental variables that have been consistently identified as important by various studies; in particular,  $V_w$ ,  $H_s$  and  $T_p$  together with  $T_i$ ,  $\theta_{wind}$  and  $\theta_{waves}$  are included. They are distributed according to site-based measurements, see Section /refsec:SiteEnv for details.

In addition, the fatigue damage is calculated by using six random seeds which are selected using a random number generator for each sample of the environmental conditions. This approach follows the guidance in IEC 61400-3 [2], which recommends 60-minute simulations (as each analysis is 10 minutes in length).

#### 4.2. Calculation framework

The FLS reliability can be computed by using the computational framework proposed in Figure 2. The first step of the framework involves fitting a GP surrogate model to realizations of fatigue damage computed using a set of time-history analyses. To this aim, a random design of experiments (DoE) is used to select combinations of the environmental variables and seeds at which to evaluate fatigue damage. After generating the training sample (for which a sensitivity study is presented in Section 6.1), structural simulations are run on a high-performance computing cluster (HPCC), and fatigue damage is estimated using the calculation described in Section 2.2. The GP surrogate model is developed by firstly specifying a kernel, and then by fitting its parameters to the training data. The kernel defines the spatial relationship between fatigue damage at different input values (i.e., environmental conditions). This finally results in a predictive distribution used to estimate the output (i.e., fatigue damage) for any arbitrarily-selected input vector not included in the training dataset for which the output is wanted. The predictive distribution depends on the parameters of the kernel (referred to as hyperparameters in the remainder of this text) and a noise term ( $\sigma_n$ ) which is used to model seed variability. In fact, OpenFAST produces different prediction of fatigue damage at the same input variable location because of the stochastic loads produced from using different random seeds (described in Section 2.1), and this introduces noise into the GP model. These parameters are selected to minimize the log marginal likelihood [38] using a quasi-Newton optimizer, which iterates from a set of assumed, initial conditions to the optimal combination of parameter values. The overall goodness of fit can be assessed using the metrics described in Section 4.4. Some background theory on GP regression is provided in Appendix A and different kernel options adopted in this study are described in Section 4.3.

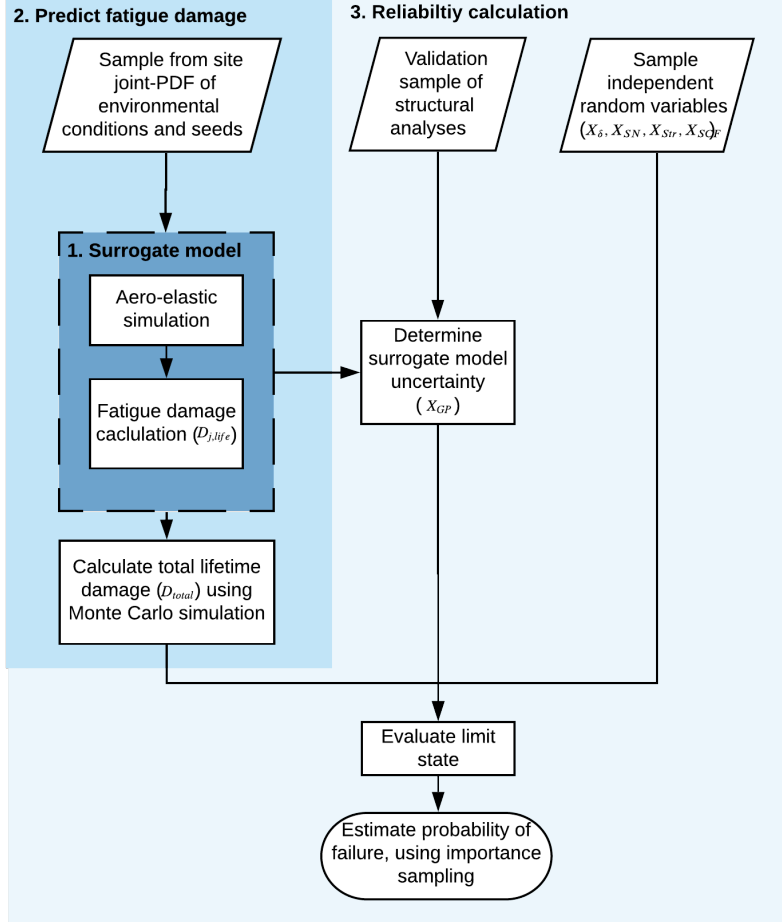
The total fatigue damage, Eq. (5), can be evaluated numerically by using plain Monte Carlo sampling [42] using the surrogate model predictions and the probability distributions describing the environmental conditions at a given site. The limit state function of Eq.(6) includes both segments of the SN curve and a surrogate model can be fit to each independently. In practice, if low stress ranges are observed, the term related to the first segment of the SN curve may be neglected.

Finally, FLS reliability can be assessed using the limit state function (Eq.(6)) and any structural reliability technique [18]. Specifically, importance sampling [42] is employed in this study due to the low probabilities of failure involved in the considered problem. In particular, the tolerable damage, the SN-curve intersection and the modeling uncertainties (i.e., the stress concentration factor uncertainty, the uncertainty in the mechanical model and the uncertainty due to the surrogate modeling of fatigue damage) are all modeled as random variables, as discussed above. The other structural parameters, such as geometry, are considered as deterministic as their variability is assumed small in comparison to the material properties [19]. The implementation of these steps is described in greater detail with reference to an illustrative OWT structure in Section 5.

#### 4.3. Considered GP surrogate model

Different assumptions concerning the form of the GP surrogate model can be tested, as described by Rasmussen and Williams [38], including: (1) the Kernel type, for which popular choices include squared exponential (SE), general exponential, Matern (including the 3/2, 5/2, and generalized variants), and the rational quadratic; and (2) basis function on which the GP model is fit (e.g., none, linear function or quadratic function).





**Figure 2:** Proposed computational framework for FLS reliability assessment of OWTs. The variables have been defined in Section 2.2.

The kernel function defines the relationship between points in the input space ( $\mathbf{x}$ ; i.e., the considered environmental variables) of the GP model. The kernel parameters are referred to as hyperparameters because they apply equally to all entries in the covariance matrix and can be estimated from the training data using maximum likelihood estimation [38]. For instance, the SE kernel assumes that response estimates vary smoothly with changing inputs; however, it has been noted that this may not be suitable for representing complex physical processes and real-world applications [43]. The Matern kernel can model functions that are not infinitely smooth; it is defined as:

$$k_{MT}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left( \frac{\sqrt{2\nu(\mathbf{x} - \mathbf{x}')(\mathbf{x} - \mathbf{x}')^T}}{l} \right)^\nu K_\nu \left( \frac{\sqrt{2\nu(\mathbf{x} - \mathbf{x}')(\mathbf{x} - \mathbf{x}')^T}}{l} \right), \quad (7)$$

where  $\Gamma(\cdot)$  is the gamma function,  $K_\nu$  is the Bessel function of the second kind, and  $\sigma_f^2$ ,  $l$  and  $\nu$  are the kernel hyperparameters. The  $l$  parameter defines the length scale over which the GP-model output varies. The  $\nu$  parameter determines the mean-squared differentiability of the GP predictions. Two commonly used

variants are the 5/2 ( $k_{MT,5/2}$ ) and 3/2 ( $k_{MT,3/2}$ ), the latter being defined as:

$$k_{MT,3/2}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \left( 1 + \frac{\sqrt{3(\mathbf{x} - \mathbf{x}')(\mathbf{x} - \mathbf{x}')^T}}{l} \right) \exp\left( - \frac{\sqrt{3(\mathbf{x} - \mathbf{x}')(\mathbf{x} - \mathbf{x}')^T}}{l} \right) \quad (8)$$

The kernel functions considered in the study are of automatic relevance detection type [38], meaning that they allow different length scales for each input dimension. The parameter values are calibrated to the training data using the maximum likelihood method.

#### 4.4. Considered metrics for assessing the goodness of fit of the GP surrogate model

The accuracy of different GP surrogate models is quantified and compared using various metrics:

- The normalized root mean squared error ( $NRMSE$ ), which is the root mean squared error normalized by the mean of the test observations ( $\mathbb{E}(y)$ ); i.e.,  $NRMSE = \sqrt{\mathbb{E}(y - \hat{y})^2} / \mathbb{E}(y)$ , where  $\hat{y}$  is the the GP surrogate model prediction for  $D_{kj}$  and  $y$  is the observed  $D_{kj}$  value at same location, computed using structural analysis.
- The coefficient of determination ( $R^2$ ) which estimates the overall goodness of fit [44].
- The coverage probability of the 50%, 90% and 95% prediction intervals (hereafter simply called coverage), which is the proportion of test samples falling within a considered set of confidence intervals. This metric can identify whether the GP model introduces large modeling uncertainty, as, if this is the case, the standard deviation in the GP prediction would become much larger than that of the validation data, resulting in a large coverage.
- The ratio between the lifetime fatigue damage ( $D_{total}$ ) predicted using the GP surrogate model and that computed for the validation set observations - therefore quantifying the overall error in the estimation of the total fatigue damage due to the surrogate modeling of  $D_{kj}$ .

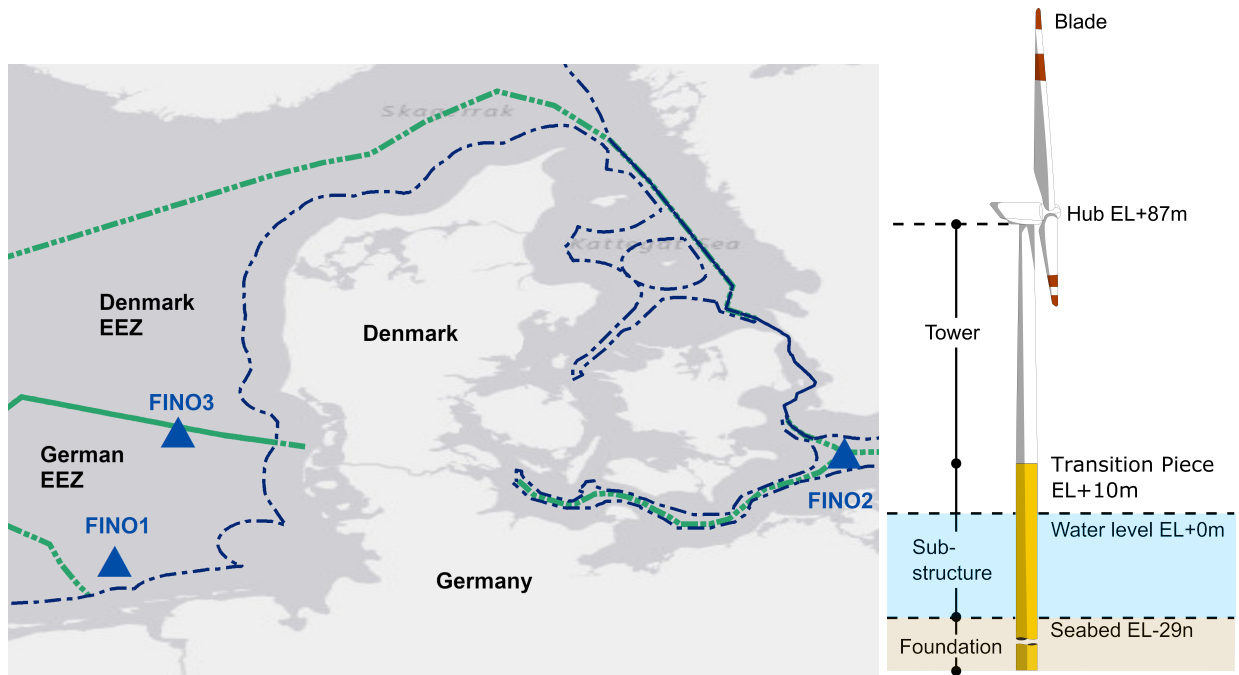
These metrics are all assessed using an independent validation set of analyses comprised of 1,000 samples, as discussed in the following sub-sections. The accuracy of the proposed GP surrogate model was also confirmed by assessing the  $NRMSE$  and coverage metrics using a leave-one-out-cross-validation approach, where the error is sequentially evaluated for a GP model fit by using each sample in the training set except one. Results from this second validation approach are not presented for the sake of brevity.

## 5. Illustrative application

### 5.1. Site and environmental conditions

This study uses environmental data measured by three met masts which are part of the FINO project to record environmental conditions in the North and Baltic Seas. This raw data was processed into probability distributions by Hubler et al. [45] who discretized the data into bins and fit a series of conditional distributions to each bin. The resulting distributions are used in this study. The FINO masts are located in the German North Sea (FINO1 and FINO3) and the Danish sector of the Baltic Sea (FINO2) as indicated in Figure 3. Each site has between 7 and 13 years worth of continuously recorded environmental data.

The FINO3 met mast is the main focus of this study; it has a water depth of 22m (the site depth varies between 22m and 30m). The dependencies between the considered environmental parameters are visible by plotting the conditional probability distributions for the FINO3 site. Figure 4 and 5 show the discrete bins into which the continuous environmental variables were segmented; the distribution types are summarized in Table 2. Similar distributions were provided for the FINO1 and FINO2 sites [45], but are not shown here to save space. The distribution of wave angle is not plotted as it is dependent on three variables.



**Figure 3:** Location of the three FINO met-mast sites (left). Territorial waters are shown with a dashed line and exclusive economic zones (EEZ) using a double-dashed line. A image of the NREL 5MW OWT with the main elevations (right).

## 5.2. Fatigue damage calculation

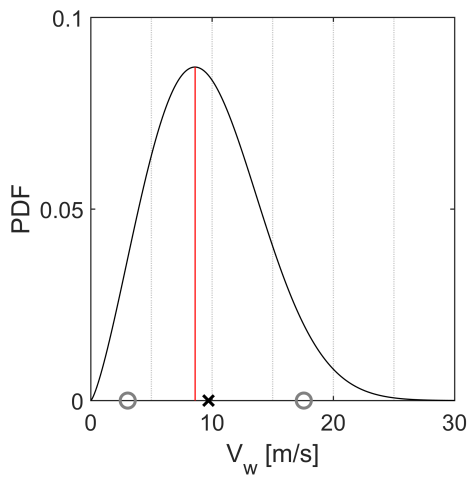
### 5.2.1. Structural model

The 3-bladed National Renewable Energy Laboratory (NREL) 5MW OWT on monopile substructure [46] is used as the reference structure in this study, with the main elevations shown on Figure 3 (right). This turbine has a 126m diameter rotor and the monopile has a 6m diameter with a constant 0.06m wall thickness, while the tower has a tapered section with the diameter decreasing linearly from 6m at the bottom to 3.87m at the top. The turbine is operational between mean wind speeds 3m/s to 25m/s and the rated rotor speed is 12.1rpm.

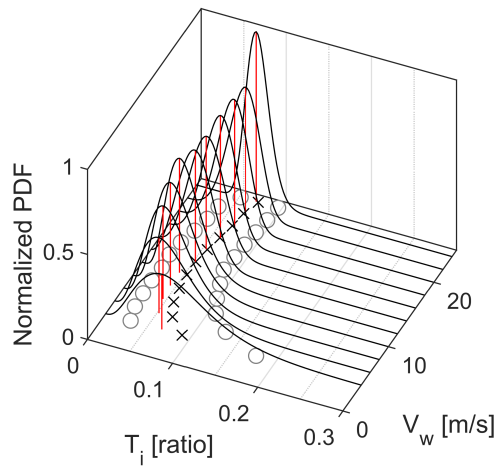
The steel used for the monopile has an elastic modulus of 210GPa, a Poisson ratio of 0.3 and a density of 7850kg/m<sup>3</sup>. The tower has the same material properties except that the steel density was increased (to 8500kg/m<sup>3</sup>) to incorporate the additional weight of secondary steel. The full set of dimensions and material properties used to model this OWT are detailed by Jonkman et al. [46].

**Table 2:** Probability distributions used to model environmental conditions at the FINO3 site showing the conditional dependencies, the distribution parameters are defined by Hubler et al. [45].

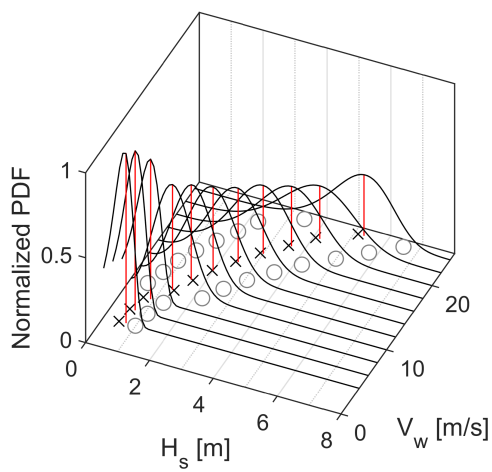
Parameter	Variable	Distribution	Dependencies
Mean wind speed	$V_w$	Weibull	N/A
Turbulence intensity	$T_i$	Weibull; Gamma	Mean wind speed
Significant wave height	$H_s$	Gumbel; Weibull	Mean wind speed
Peak spectral period	$T_p$	Bimodal Gumbel	Significant wave height
Wind direction	$\theta_{wind}$	Non-parametric KDE	Mean wind speed
Wave direction	$\theta_{wave}$	Non-parametric KDE	Significant wave height; Wind direction



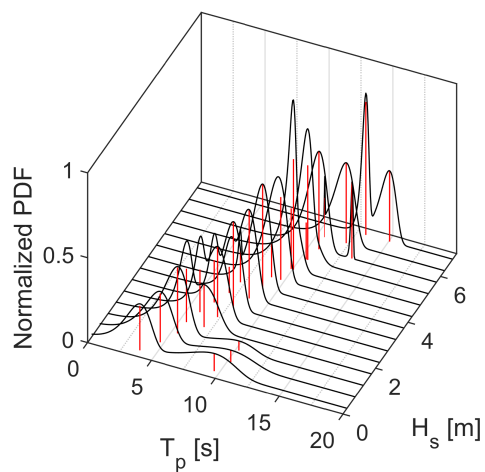
(a) Mean wind speed.



(b) Turbulence intensity dependent on mean wind speed.

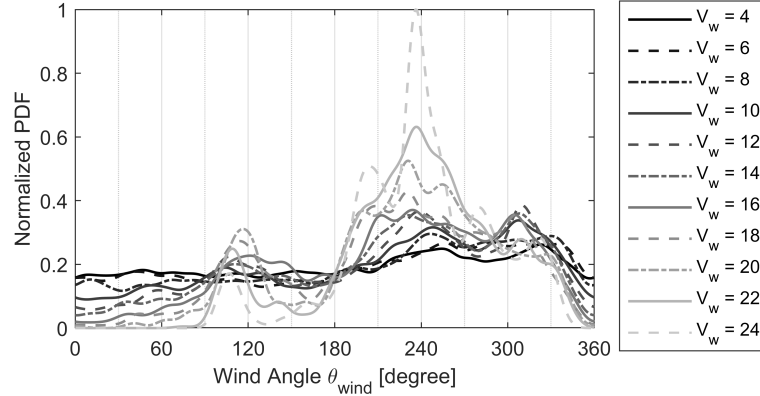


(c) Significant wave height dependent on mean wind speed.



(d) Peak spectral period dependent on significant wave height

**Figure 4:** Probability distributions of environmental conditions at the FINO3 met mast. The mean of each distribution is identified by an 'x', the 0.05 and 0.95 quantiles by a circle, and the mode by a red line. The normalized values are calculated such that the highest value mode occurs at a value of 1.



**Figure 5:** Density functions of wind angle dependent on the mean wind speed, increasing wind speed is indicated by lighter colored line. Normalized so that the maximum mode has a value of 1.

### 5.2.2. Load calculation

The fatigue loading on the illustrative OWT structure due to different environmental conditions is calculated in the time domain using the aero-hydro-elastic wind turbine simulation package OpenFAST [13]. The mean wind speed and the turbulence intensity are used in the Kaimal spectrum, which defines the frequency content of turbulent wind [47]. The spectrum is subsequently converted into stochastic time series (using the inverse Fourier transform). The mean and turbulent wind speed components are summed and used to calculate wind loading on the OWT blades using the blade element momentum theorem [11]. Similarly, the significant wave height and peak spectral period are used in the JONSWAP wave spectrum [47], which defines the energy of different wave amplitudes within a stochastic sea state. Individual wave amplitudes are converted into wave kinematics (water particle velocities and accelerations) using Airy wave theory and are summed to evaluate the entire sea-state. Structural wave loads are evaluated using Morrison’s equation [47]. The wind and wave direction are combined into a single parameter, the misalignment ( $\theta_{mis}$ ), by taking advantage of the symmetry of the monopile sub-structure. Only the misalignment between wind and wave loading matters for calculating the fatigue damage, as the loads can be rotated around the monopile cross section in a post-processing phase to account for the absolute angle of loading. The below mudline foundation significantly influences the structural response in the dynamic analysis of OWTs [48]. In this paper, the foundation is modeled using the apparent fixity (AF) method [49], with the fixity length selected to match the natural frequency between the AF supported OWT and an OWT with a linear-elastic foundation model. The soil properties were taken from Damgaard et al. [24]. Individual analyses of 11-minute length are conducted and the first minute is deleted to remove transient effects [15].

In a post-processing step, the mudline bending moment time-series is extracted and converted into stresses. Assuming that the stresses in the wall of the monopile behave in-plane (with no bending moments), a membrane stress calculation can be used [50]. This avoids the need for complex finite-element analysis and can be implemented using a simple set of membrane stress equations.

### 5.2.3. Fatigue damage

Fatigue damage is calculated for a single weld located at the mudline. The weld is a transverse butt weld with no weld toe grinding, and assessed using a DNV class E [17] SN curve which assumes the weld is protected from corrosion by cathodic protection. The DNV SN curves [17] are bi-linear in the log scale. Fatigue damage is estimated at 12 locations equally distributed around the mudline cross-section, with the largest value extracted for use in the fatigue life calculation. The results are presented in terms of DEL calculated using  $10^7$  reference cycles. It should be noted that the proposed GP models are assessed in terms of DEL only for numerical convenience, and this is converted back into damage for evaluating the limit state of Eq. (6).

**Table 3:** Input range of each environmental variable in the GP model.

Variable	Lower bound	Upper bound	Criteria
$V_w$	4	24	Range of $V_w$ where the OWT is operational
$T_i$	0.01	0.18	The upper limit is two standard deviations above the mean $T_i$ of the lower $V_w$ bin; the lower limit is selected above 0 (for turbulence spectrum equation).
$H_s$	0.01	7	The upper limit is two standard deviations above the mean $H_s$ of the largest $V_w$ bin; the lower limit is above 0 (for wave spectrum equation).
$T_p$	1	15	The upper limit is the mode of 2nd peak in the largest $H_s$ bin.
$\theta_{mis}$	0	180	Full range of misalignments, given the symmetry of the OWT.

The stress time history is copied 35 times, generating a six-hour long time series. This mitigates against observed increases in predicted fatigue damage with increased analysis length which are due to the effect of open cycles in the rainflow counting algorithm [45], as the ratio of open to closed cycles is reduced by duplicating the stress time history. The fatigue damage is calculated assuming a design life of 20-years, a typical figure for OWTs [8].

### 5.3. Training and validation data sets

A training set is used to fit the surrogate model and a validation set is used to assess the overall goodness of fit of the proposed surrogate model for the FINO3 site. Additionally, for the FINO1 and FINO2 sites, validation sets are also generated. The number of dimensions in the input vector is five and matches the number of parameters in the environmental model subtracting the simplification that the wind and wave angle can be combined into a single variable,  $\theta_{mis}$ , because of the axisymmetric cross section of the monopile.

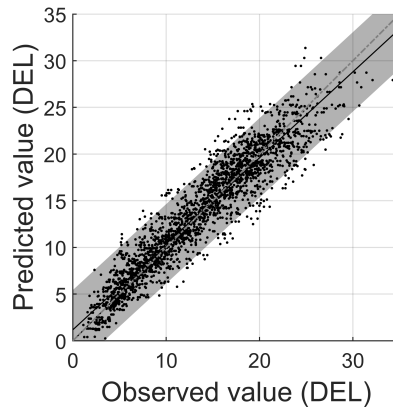
The single training and three validation data sets are each made of 1,000 realizations of the environmental conditions drawn randomly from the relevant site-specific probability distributions of the environmental conditions (as discussed above). Then the fatigue damage is calculated at each using six random seeds which are selected using a random number generator for each sample of the environmental conditions. This approach follows the guidance in IEC 61400-3 [2], which recommends 60-minute simulations (as each analysis is 10 minutes in length) and results in 6,000 structural simulations.

In a pre-processing stage, the site-specific conditional distributions are truncated to prevent unrealistic environmental conditions. This is achieved by modifying the limits of the uniformly distributed random numbers to produce acceptable ranges of the environmental conditions values. The mean wind speed is truncated to values occurring between 3m/s and 25m/s, conditions in which the turbine is operational [46]. This means that only operating wind and wave conditions are captured in this illustrative application, and that extreme conditions, where the OWT is in its shut-down state, are neglected. Additionally, the significant wave height (at low mean wind speeds) and peak spectral period are both represented by extreme value distributions which support negative values. However, numbers less than zero are not permissible as inputs into the spectral equations which are used to generate generating environmental conditions, described in Section 5.2.2.

## 6. Results and discussion

### 6.1. Proposed surrogate model

This section starts by assessing the accuracy of the proposed GP surrogate model. In particular, the Matern 3/2 kernel is selected in this study as it represents a popular and practical choice for applications related to both onshore and offshore wind turbine. Its flexibility stems from the variable smoothness of the resulting GP predictions, as described in Section 4.3. This allows one to better capture the fact that most real-world problems are not infinitely smooth (as assumed for instance by the squared-exponential kernel). In addition, both the considered Matern 3/2 kernel and the more complex generalized Matern kernel lead to very similar



**Figure 6:** GP-predicted vs observed DEL values based on the GP surrogate model with the Matern 3/2 kernel kernel. The line shows the result of a simple linear regression fit to the data and the shaded areas show the 95% confidence intervals.

model performance, i.e., a  $NRMSE$  of about 0.15 and  $R^2$  of about 0.94, in both cases. The Matern 3/2 kernel has also the advantage of being practically implemented in many commonly used software packages such as Matlab. A plot of predicted vs observed DEL values based on the GP surrogate model with the Matern 3/2 kernel kernel is shown in Figure 6 (note that only half of data points are plotted to improve the clarity of the figure and they were selected randomly). The corresponding kernel parameters and goodness-of-fit metrics are listed in Table 4. The lifetime damage values show very high accuracy compared to the scatter in Figure 6; this is partly due to an averaging effect when considering all the environmental conditions (in fact, as 47.6% of the samples are above the 1:1 trendline and 52.4% are below it). These findings are in very good agreement with those of Slot et al. [51], who used a Matern 3/2 kernel to represent fatigue loads on an onshore wind turbine, and Häfele et al. [25] who found a Matern 5/2 kernel to be best for representing joint fatigue loads for a OWT on jacket substructure.

The computational burden required to run 6,000 dynamic structural simulations is large. In practice, this sample size may be infeasible on a standard workstation. Moreover, GP surrogate models need to take into account the whole set of training data each time they are used for predictions. This means that the computational cost of each prediction scales with the number of training samples. Generally, it is not clear in advance whether reducing the number of samples or seeds will have a greater impact on the accuracy of the GP surrogate model. The number of seeds will affect the number of analyses at a specific set of input environmental conditions and may affect the coverage metrics (as the scatter at each input point will be defined better). Reducing the number of samples will affect the sampling quality over the considered input conditions, which should, in turn, effect the  $NRMSE$ .

Different numbers of random samples and seeds are tested by drawing 1,000 sets (an appropriate sample for evaluating scatter in the sample mean [52]) of bootstrapped samples from the full set of validation results ( $1,000 \times 6$ ). The resulting goodness-of-fit metrics are assessed for these reduced analysis sets and the results, presented in Figure 7, suggest that both the number of samples and seeds can be reduced without a large impact on the  $NRMSE$  and the coverage. A combination of 300 samples and two seeds ( $300 \times 2$ ) will be used in the rest of this study because this provides a good trade off in terms of  $NRMSE$  and lifetime fatigue damage accuracy.

Additionally, the accuracy of the proposed GP surrogate model was further confirmed using the bootstrap samples in Figure 7. The mean and standard deviation of the ratio between the observed on predicted damage values were evaluated to be 0.998 and 0.01 respectively. These parameters were used to define the  $X_{GP}$  in Eq.(6). A Normal probability model is assumed for  $X_{GP}$ .

Finally, the lifetime damage predictions in terms of DEL and  $D_{total}$  are also compared against the obser-



**Table 4:** Kernel parameters and goodness of fit metrics for different kernel functions at the FINO3 site. The  $\sigma_n$  parameter refers to the noise added to the GP that accounts for seed-to-seed variability (see Eq. (11)).

No Samples	1000*6	300*2
Gaussian Process type	Matern 3/2 kernel	Matern 3/2 kernel
$\sigma_f$	9.681	15.261
Length scale ( $V_w$ )	1.394	1.492
Length scale ( $T_i$ )	6.694	20.966
Length scale ( $H_s$ )	2.266	2.347
Length scale ( $T_p$ )	1.195	2.752
Length scale ( $\theta_{Mis}$ )	4.446	8.478
$\sigma_n$	2.092	2.110
$NRMSE$	0.154	0.163
$R^2$	0.942	0.935
Coverage:		
95%	0.923	0.943
90%	0.890	0.911
50%	0.572	0.584

**Table 5:** Comparison between the lifetime fatigue damage computed using the validation sets across the different FINO sites and the predictions from the GP model fit to the FINO3 site with a Matern 3/2 kernel. The errors are computed by dividing the validation set value (observation) by the relevant GP value (prediction).

		FINO1	FINO2	FINO3
$DEL_{total}$	GP 300x2	12.625	12.181	13.723
	Validation	12.458	11.995	13.745
$D_{total}$	GP 300x2	7.95E-04	6.64E-04	1.21E-03
	Validation	7.44E-04	6.15E-04	1.22E-03
Ratio	$DEL_{total}$	0.987	0.985	1.002
	$D_{total}$	0.936	0.926	1.008

vations for independent validation sets, allowing a direct comparison between the surrogate model and the structural analysis results for different sites. The results for the FINO3 site, shown in Table 5, demonstrate that the proposed surrogate model produces predictions of lifetime fatigue damage similar to the observations computed using structural simulations. The GP model error is lower than that reported by Huchet et al. [35] who found a 4.43% error between an adaptive Kriging algorithm and aero-elastic analyses in terms of damage. Similarly, small model error (using the EN 1990 method) was found by Slot et al. [51] in the case of onshore wind turbines, using aero-elastic simulations with 4 seeds and 100 environmental samples (although it should be noted that different environmental conditions were simulated).

### 6.1.1. Different sites

The GP surrogate model predictions using the FINO3 training set are compared with the observations from the validation sets from the FINO1 and FINO2 sites in Table 5. This indicates that, while the accuracy of the prediction decreases when the GP model trained at FINO3 is validated against fatigue damage computed at other sites, the observed errors are not large. This is confirmed in Table 5 where it can be seen, for instance, that the error is below 2% in terms of DEL for the FINO1 site. As the environmental probability distributions change between the sites, a reduced accuracy of the GP surrogate model is observed when considering the FINO2 data. This is visible by comparing the  $T_p$  parameter from all sites; the FINO1 and FINO3 distributions are broadly similar while the FINO2 distribution is more peaked, i.e., Figure 8. However,

it should be noted that if this assumption was to be tested in practice, the site water depth should also be explicitly modeled as an additional variable within the proposed surrogate model.

### 6.2. Fatigue reliability analysis

FLS reliability is finally assessed using the importance sampling approach [42], in combination with the limit state function defined in Eq. (6). The GP surrogate model with Matern 3/2 kernel trained using  $300 \times 2$  samples is used to evaluate damage - as it combined one of the most accurate and practical kernel functions with an appropriate number of samples/seeds, as discussed previously. It is worth noting that the computed damage values by OpenFAST were all found to lie within the second segment of the SN curve; consequently, the fatigue damage prediction is only based on a GP model considering the second slope of the SN curve. In Eq. (6), the 20-year damage estimates are computed by sampling from the site joint distributions of the environmental conditions and evaluating numerically Eq. (5) using a Monte Carlo integration with  $1 \times 10^6$  samples, using the upper and lower boundaries in Table 3 (i.e., same intervals as used in the DoE sample).

The finite sample of the GP is associated with some error which is assessed by rerunning the integration 1,000 times for each site using the mean function of the GP conditioned on the OpenFAST observations with  $1M$  samples. The full GP distribution was also sampled at the FINO3 site. The resulting distributions of fatigue damage are shown in Figure 9. A normal probability distribution is fit to the empirical distributions and this assumption is confirmed using the Anderson-Darling normality test at a 5% significance level. The resulting distribution of the lifetime fatigue damage are comparable as shown in Figure 9 and the coefficient of variation in the empirical distribution increases from 3.37% in the mean function of the GP to 3.95% in the full GP-distribution case. In the subsequent analysis fatigue damage is predicted using the normal distribution of fatigue damage predictions.

The mean values of lifetime damage calculated for the respective sites are  $6.70 \times 10^{-3}$  (FINO1),  $8.30 \times 10^{-3}$  (FINO2) and  $8.69 \times 10^{-3}$  (FINO3). The similarity between fatigue damage using the FINO1 and FINO3 met masts is consistent with the findings in Section 6.1 due to similar environmental conditions between the two sites. However, the damage predicted at the FINO2 site is lower than that at the others site primarily as a result of lower significant wave heights [45].

The probability of failure is finally evaluated using importance sampling. Limit state samples calculated over the life of the structure are plotted in Figure 10, which also shows the best fit Gumbel distribution.

The resulting probability of failure over a 20-year design life is  $3.940 \times 10^{-6}$ , which corresponds to an annual probability of failure of  $9.005 \times 10^{-7}$ , using the relationship defined by Velarde et al. [22]. The relationship between these values over time is shown in Figure 11. Comparing this value to code target reliability levels seem to suggest that the NREL 5MW located at the FINO3 site has a highly conservative level of safety. IEC 61400-1:2019 recommends a probability of failure of  $10^{-4}$ . However, the NREL 5MW has not been specifically designed for the FINO site-conditions and therefore a direct comparison between calculated and code-based failure rates is only illustrative.

The Gumbel distribution has a CoV of 0.319, indicating the high uncertainty introduced by the random variables in the limit state. In particular, within the limit state equation, the two random variables  $X_\delta$  and  $X_{SN}$  act against each other. Modeling the SN curve as random increases capacity (as the design curve is conservative), whereas modeling tolerable damage as random reduces capacity (because the median of random variable is below one).

## 7. Conclusions

This paper presented a novel application of Gaussian process (GP) regression to assess fatigue reliability of offshore wind turbines (OWTs) using site-specific environmental conditions and advanced dynamic structural analyses. This is especially important as fatigue is often the limit state driving structural design of OWTs in European waters. The approach can also be used to calibrate code-based partial safety factors to be used in the design of OWTs. The proposed approach allows the rapid evaluation of monopile fatigue damage without requiring computationally-expensive time-history analyses to be run for each sample of the site-specific environmental conditions. This allows, in turn, a fast yet accurate estimation of fatigue

failure/reliability of OWT structural components.

The sensitivity of various goodness-of-fit metrics to the number of structural analysis samples has also been investigated to further reduce the computational effort required to perform GP regression/predictions in practical applications. The use of the efficient GP surrogate model for fatigue damage prediction allowed a comprehensive assessment of FLS structural reliability for an archetype OWT at three sites in Europe, considering all the relevant random variables modeling uncertainties in load-induced fatigue demands and structural capacity of the OWT substructure. In particular, it was shown that the considered random variables, particularly those related to modeling uncertainties, can cause large scatter in the fatigue limit state function, introducing a coefficient of variation of about 30%.

The annual probability of failure for the illustrative OWT structure was found to be  $\approx 1 \times 10^{-6}$ . Other OWTs with different nominal properties could be similarly assessed as the approach taken in this paper is flexible, potentially reducing the conservatism in OWT design and assessment. Situations where the turbine is not operational, such as wind speeds above and below the operating range, can be assessed by fitting a similar GP regressions to models representing those conditions. For instance, assessing OWTs on jacket substructures would require multiple welds to be assessed in order to evaluate system reliability of a structure with redundancy. Additionally, higher values of the probability of failure are expected if different random-variable models/statistics are used for each slope of the SN curve in Eq.(6), e.g., [20]. However, for illustrative purposes, it was decided to use the single random variable approach to be consistent with the DNVGL code [17].

Extending the fatigue calculation to other sites in European waters demonstrated that a surrogate model developed for a single site may be usable at other locations characterized by similar environmental conditions. This suggests that further computational savings would be possible if a representative GP surrogate model is developed for a OWT model which is used across multiple nearby locations.

It is finally worth noting that the accuracy of/computational effort required by plain Monte Carlo sampling and importance sampling to generate the random DoE used in this study and to solve the limit state equation was found acceptable (also given the relatively small number of random variables involved in the analysis and their statistics). However, the use of adaptive sampling in the surrogate model development, leveraging on adaptive training input to focus on the failure region of the limit state function, could further increase the computational efficiency of the proposed framework. Results from this study can provide guidance for practical applications of the proposed framework in OWT projects.

## 8. Acknowledgments

This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC), DTP grant EP/M507970/1 for University College London. The authors acknowledge the use of the UCL Myriad High Performance Computing Facility (Myriad@UCL), and associated support services, in the completion of this work. We sincerely thank Prof Bruce Ellingwood and four anonymous reviewers for very helpful and insightful comments that substantially improved the quality of this study.

## References

- [1] A. Mbistrova, [Financing and investment trends; the european wind industry in 2016](#), Tech. rep., WindEurope (2017). URL <https://windeurope.org/about-wind/reports/financing-investment-trends-2016/>
- [2] International Electrotechnical Commission, IEC 61400-3:2019 Wind energy generation systems. Design requirements for fixed offshore wind turbines (2019).
- [3] DNV GL, DNVGL-ST-0126. Support structures for wind turbines (2018).
- [4] K. Fischer, C. Viljoen, J. Kohler, M. H.Faber, Optimal and acceptable reliabilities for structural design, *Structural Safety* 76 (2019) 149–161. doi:10.1016/j.strusafe.2018.09.002.
- [5] A. Morató, S. Sriramula, N. Krishnan, Kriging models for aero-elastic simulations and reliability analysis of offshore wind turbine support structures, *Ships and Offshore Structures* (2018) 1–14doi:10.1080/17445302.2018.1522738.
- [6] D. Wilkie, C. Galasso, A probabilistic framework for offshore wind turbine loss assessment, *Renewable Energy* 147 (2020) 1772–1783. doi:10.1016/j.renene.2019.09.043.
- [7] A. A. Taffanidis, E. Loukogeorgaki, D. C. Angelides, Offshore wind turbine risk quantification/evaluation under extreme environmental conditions, *Reliability Engineering & System Safety* 115 (2013) 19–32. doi:10.1016/j.ress.2013.02.003.

- [8] F. Vorpahl, H. Schwarze, T. Fischer, M. Seidel, J. Jonkman, Offshore wind turbine environment, loads, simulation, and design, *Wiley Interdisciplinary Reviews: Energy and Environment* 2 (5) (2013) 548–570. doi:10.1002/wene.52.
- [9] T. Gintautas, J. Sørensen, B. Sudret, L. Sørensen, M. Thøgersen, Reliability-based inspection planning of 20MW offshore wind turbine jacket, *International Society of Offshore and Polar Engineers* (2018).
- [10] C. Hübler, C. G. Gebhardt, R. Rolfes, Methodologies for fatigue assessment of offshore wind turbines considering scattering environmental conditions and the uncertainty due to finite sampling, *Wind Energy* 21 (11) (2018) 1092–1105. doi:10.1002/we.2216.
- [11] T. Burton, N. Jenkins, D. Sharpe, E. Bossanyi, *Wind Energy Handbook*, 2nd Edition, John Wiley & Sons, 2011. doi:10.1002/9781119992714.  
URL <http://doi.wiley.com/10.1002/9781119992714>
- [12] DNV GL, DNVGL-RP-C205. Environmental conditions and environmental loads (2017).
- [13] B. Jonkman, J. Jonkman, FAST Readme v8.12.00a-bjj, Tech. rep., NREL, Denver (2015).
- [14] N. Barltrop, A. Adams, Dynamics of fixed marine structures, 3rd Edition, Elsevier, 1991. doi:10.1016/c2013-0-04571-9.
- [15] D. Zwick, M. Muskulus, The simulation error caused by input loading variability in offshore wind turbine structural analysis, *Wind Energy* 18 (8) (2014) 1421–1432. arXiv:arXiv:1006.4405v1, doi:10.1002/we.1767.
- [16] J. Schijve, Fatigue of structures and materials, 2nd Edition, Springer Science + Business Media, Amsterdam, 2009.
- [17] DNV GL, DNVGL-RP-C203 Fatigue design of offshore steel structures, Tech. rep., DNV GL (2016).
- [18] R. E. Melchers, A. T. Beck, Structural reliability analysis and prediction, 3rd Edition, John Wiley & Sons Ltd, 2017. doi:10.1002/9781119266105.
- [19] Joint Committee Structural Safety, *JCSS Probabilistic Model Code*, Tech. rep., Technical University of Denmark (2013).  
URL <http://www.jcss.byg.dtu.dk/Publications/Probabilistic{ }Model{ }Code.aspx>
- [20] J. Sørensen, S. Frandsen, N. Tarp-Johansen, Effective turbulence models and fatigue reliability in wind farms, *Probabilistic Engineering Mechanics* 23 (2008) 531–538. doi:10.1016/j.probengmech.2008.01.009.
- [21] DNV GL, DNVGL-RP-C210 Probabilistic methods for planning inspections fatigue cracks offshore, Tech. rep., DNV GL (2015).
- [22] J. Velarde, E. Bachynski, *Design and fatigue analysis of monopile foundations to support the DTU 10 MW offshore wind turbine*, *Energy Procedia* 137 (2017) 3–13. doi:10.1016/j.egypro.2017.10.330.  
URL <https://doi.org/10.1016/j.egypro.2017.10.330>
- [23] M. Kuhn, Dynamics and design optimization of offshore wind conversion systems, Ph.D. thesis, Delft University of Technology (2001).
- [24] M. Damgaard, L. Andersen, L. Ibsen, H. Toft, J. Sørensen, A probabilistic analysis of the dynamic response of monopile foundations: Soil variability and its consequences, *Probabilistic Engineering Mechanics* 41 (2015) 46–59. doi:10.1016/j.probengmech.2015.05.001.
- [25] J. Häfele, C. G. Gebhardt, R. Rolfes, A comparison study on jacket substructures for offshore wind turbines based on optimization, *Wind Energy Science* (2018). doi:10.5194/wes-2018-58.
- [26] J. V. D. Tempel, Design of support structures for offshore wind turbines, Ph.D. thesis, Technical University of Delft (2006).
- [27] B. Yeter, Y. Garbatov, C. Soares, Fatigue reliability assessment of an offshore supporting structure, in: *Maritime Technology and Engineering*, CRC Press, 2014, pp. 671–680. doi:10.1201/b17494-89.
- [28] L. Ziegler, S. Voormeeren, S. Schafhirt, M. Muskulus, Design clustering of offshore wind turbines using probabilistic fatigue load estimation, *Renewable Energy* 91 (2016) 425–433. doi:10.1016/j.renene.2016.01.033.
- [29] S. Márquez-Domínguez, J. D. Sørensen, Fatigue reliability and calibration of fatigue design factors for offshore wind turbines, *Energies* 5 (6) (2012) 1816–1834. doi:10.3390/en5061816.
- [30] W. Dong, T. Moan, Z. Gao, Fatigue reliability analysis of the jacket support structure for offshore wind turbine considering the effect of corrosion and inspection, *Reliability Engineering & System Safety* 106 (2012) 11–27. doi:10.1016/j.ress.2012.06.011.
- [31] J. Oest, R. Sørensen, L. C. T. Overgaard, E. Lund, Structural optimization with fatigue and ultimate limit constraints of jacket structures for large offshore wind turbines, *Structural and Multidisciplinary Optimization* 55 (3) (2016) 779–793. doi:10.1007/s00158-016-1527-x.
- [32] M Muskulus, S Schafhirt, Reliability-based design of wind turbine support structures, *Proc SRES 2015* 1 (October 2015) (2015) 1–13. doi:10.13140/rg.2.1.5125.5766.
- [33] H. S. Toft, L. Sørensen, W. Moser, J. D. Sørensen, M. L. Thøgersen, Assessment of wind turbine structural integrity using response surface methodology, *Engineering Structures* 106 (2016) 471–483. doi:10.1016/j.engstruct.2015.10.043.
- [34] D. Zwick, M. Muskulus, Simplified fatigue load assessment in offshore wind turbine structural analysis, *Wind Energy* 19 (2016) 265–278. arXiv:arXiv:1006.4405v1, doi:10.1002/we.1831.
- [35] Q. Huchet, C. Mattrand, P. Beaurepaire, N. Relun, N. Gayton, *Ak-da: An efficient method for the fatigue assessment of wind turbine structures*, *Wind Energy* 22 (5) (2019) 638–652. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/we.2312, doi:10.1002/we.2312.  
URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/we.2312>
- [36] R. Teixeira, M. Nogal, A. O’Connor, J. Nichols, A. Dumas, Stress-cycle fatigue design with kriging applied to offshore wind turbines, *International Journal of Fatigue* 125 (2019) 454–467. doi:10.1016/j.ijfatigue.2019.04.012.
- [37] J. Häfele, R. Damiani, R. King, C. G. Gebhardt, R. Rolfes, A systematic approach to offshore wind turbine jacket pre-design and optimization: Geometry, fatigue, cost, and surrogate structural code check models, *Wind Energy Science* (May) (2018) 1–29. doi:10.5194/wes-2018-39.
- [38] C. Rasmussen, C. Williams, *Gaussian processes for machine learning*, MIT Press, Cambridge, 2006.  
URL <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6279024>

- [39] C. Hübler, C. G. Gebhardt, R. Rolfes, Hierarchical four-step global sensitivity analysis of offshore wind turbines based on aeroelastic time domain simulations, *Renewable Energy* 111 (2017) 878–891. doi:10.1016/j.renene.2017.05.013.
- [40] H. S. Toft, L. Svenningsen, J. D. Sørensen, W. Moser, M. L. Thøgersen, Uncertainty in wind climate parameters and their influence on wind turbine fatigue loads, *Renewable Energy* 90 (2016) 352–361. doi:10.1016/j.renene.2016.01.010.
- [41] S. Brandt, M. Broggi, J. Häfele, C. G. Gebhardt, R. Rolfes, M. Beer, Meta-models for fatigue damage estimation of offshore wind turbines jacket substructures, *Procedia Engineering* 199 (2017) 1158–1163. doi:10.1016/j.proeng.2017.09.292.
- [42] J. L. Beck, K. M. Zuev, *Rare-event simulation*, Springer International Publishing, 2015. doi:10.1007/978-3-319-11259-6\_24-1.
- [43] M. Stein, *Interpolation of spatial data: some theory for kriging*, Springer, New York, 1999.
- [44] A. M. Mood, F. A. Graybill, D. C. Boes, *Introduction to the theory of statistics*, 3rd Edition, McGraw-Hill, Singapore, 1974.
- [45] C. Hübler, C. G. Gebhardt, R. Rolfes, Development of a comprehensive database of scattering environmental conditions and simulation constraints for offshore wind turbines, *Wind Energy Science* 2 (2) (2017) 491–505. doi:10.5194/wes-2-491-2017.
- [46] J. Jonkman, S. Butterfield, W. Musial, G. Scott, *Definition of a 5-MW reference wind turbine for offshore system development*, Tech. Rep. February (2009).  
URL [http://tethys-development.pnnl.gov/sites/default/files/publications/Jonkman\\_et\\_al\\_2009.pdf](http://tethys-development.pnnl.gov/sites/default/files/publications/Jonkman_et_al_2009.pdf)
- [47] DNV GL, DNVGL-RP-C205. Environmental conditions and environmental loads (2017).
- [48] E. Bush, L. Manuel, The influence of foundation modeling assumptions on long-term load prediction for offshore wind turbines, in: *Proceedings of the ASME 28th International Conference on Ocean, Offshore and Arctic Engineering*, ASME, Honolulu, 2009, pp. 1–9.
- [49] M. Damgaard, V. Zania, L. Andersen, L. Ibsen, Effects of soil–structure interaction on real time dynamic response of offshore wind turbines on monopiles, *Engineering Structures* 75 (2014) 388–401. doi:10.1016/j.engstruct.2014.06.006.
- [50] DNV GL, DNV-RP-C202. Buckling strength of shells (2013).
- [51] R. Slot, J. Sørensen, B. Sudret, L. Svenningsen, M. Thøgersen, Surrogate model uncertainty in wind turbine reliability assessment, *Renewable energy* (11 2019). doi:10.1016/j.renene.2019.11.101.
- [52] B. Efron, R. Tibshirani, *An introduction to the bootstrap*, Chapman & Hall, New York, 1993.

### A. Gaussian process regression

From a general, high-level perspective, the complex aero-hydro-elastic simulation to derive OWT structural response (e.g., in terms of component stress) can be considered as a “black box” (or a “simulator”) which takes a vector of unique inputs ( $\mathbf{x}$ ) and produces an output/target ( $y$ ) described by the relationship  $y = f(\mathbf{x})$ . This simple equation form means that the simulator can encompass a wide range of specific analysis procedures and assumptions. Metamodels, and specifically GP regression in this study, aim to mimic the behaviour of the complex system (the simulator) by predicting the value of the target variable(s) given a combination of the input variables. In other words, a statistical model is adopted to surrogate the relation  $f(\mathbf{x})$  between inputs and output such that it is possible to make predictions for any input vector while remaining computationally cheap to evaluate. In a GP regression,  $f(\mathbf{x})$  is regarded as a realized value of a gaussian process, describing the distribution of functions  $f(\mathbf{x})$ , rather than scalars or vectors.

The marginalization property of a GP allows a practical, finite-dimensional, model to be extracted from the infinite-dimensional process. When it is conditioned on training data, i.e., pairs of input variables  $\mathbf{x}$  and function values  $y$  (collected in the matrix  $\mathbf{X}$  and the vector  $\mathbf{y}$ , respectively), the GP reduces to a multivariate Gaussian distribution. Consequently, estimates of the functions values ( $\hat{y}$ ) corresponding to a new set of input variables ( $\hat{\mathbf{x}}$ ) can be obtained by drawing samples from the obtained conditional distribution [38]. This multivariate conditional Gaussian distribution is fully defined by a mean vector ( $\mu_{\hat{\mathbf{x}}}$ ) and a covariance matrix ( $\Sigma_{\hat{\mathbf{x}}}$ ):

$$Pr(\hat{y}|\hat{\mathbf{x}}, \mathbf{X}, \mathbf{y}) \sim \mathcal{N}(\mu_{\hat{\mathbf{x}}}, \Sigma_{\hat{\mathbf{x}}}) \quad (9)$$

$$\mu_{\hat{\mathbf{x}}} = \mathbf{k}_{\hat{\mathbf{x}}, \mathbf{X}} (\mathbf{k}_{\mathbf{X}, \mathbf{X}} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y} \quad (10)$$

$$\Sigma_{\hat{\mathbf{x}}} = \mathbf{k}_{\hat{\mathbf{x}}, \hat{\mathbf{x}}} - \mathbf{X}_{\hat{\mathbf{x}}, \mathbf{X}} (\mathbf{k}_{\mathbf{X}, \mathbf{X}} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{k}_{\mathbf{X}, \hat{\mathbf{x}}} \quad (11)$$

where  $\sigma_n^2$  models the variability in the function value at an input location and the identity matrix  $\mathbf{I}$  assigns noise to the diagonal terms in the covariance matrix. The full covariance matrix  $\mathbf{k}$  is partitioned:

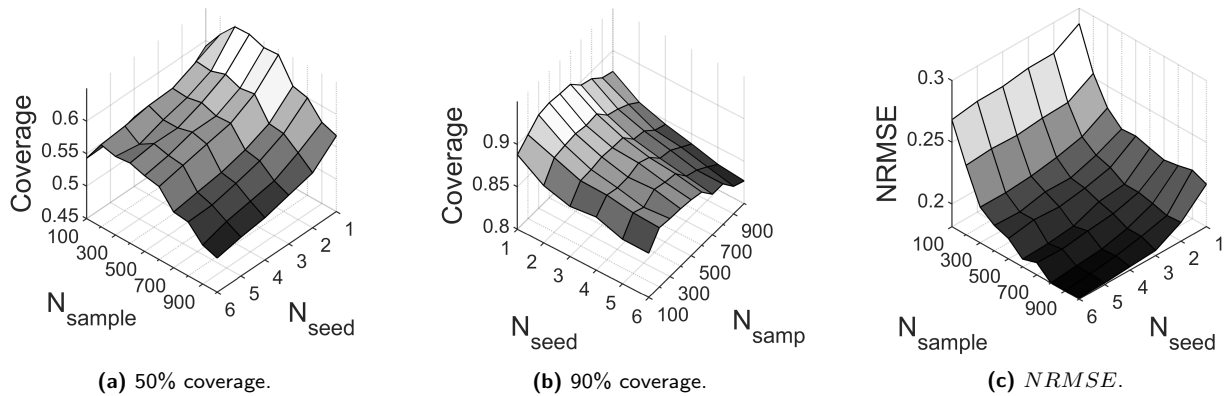
$$\mathbf{k} = \begin{bmatrix} \mathbf{k}_{\mathbf{X}, \mathbf{X}} + \sigma_n^2 \mathbf{I} & \mathbf{k}_{\mathbf{X}, \hat{\mathbf{x}}} \\ \mathbf{k}_{\hat{\mathbf{x}}, \mathbf{X}} & \mathbf{k}_{\hat{\mathbf{x}}, \hat{\mathbf{x}}} \end{bmatrix}. \quad (12)$$

Each entry is defined by evaluating a kernel function at the corresponding training or test input. The form of the kernel function is variable and encodes assumptions about the spatial correlation between the response at different input locations.

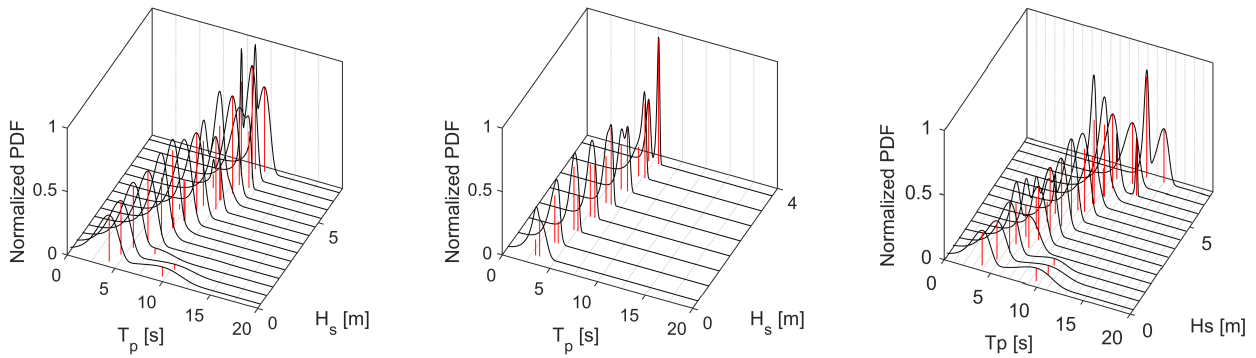
The form of the conditional Gaussian distribution given in Eq. (9) assumes the unconditioned GP has a constant mean equal to zero. If prior information is known, this assumption can be altered, e.g., fitting the GP to an underlying linear function to capture trends in the data [38].

GP regression is also applied under the name kriging in the field of geostatistics [43]. Two types of GP model are relevant for this work: prediction using the conditional mean function (simple kriging) and using the full GP to model uncertainty in the predictions (standard or universal kriging).

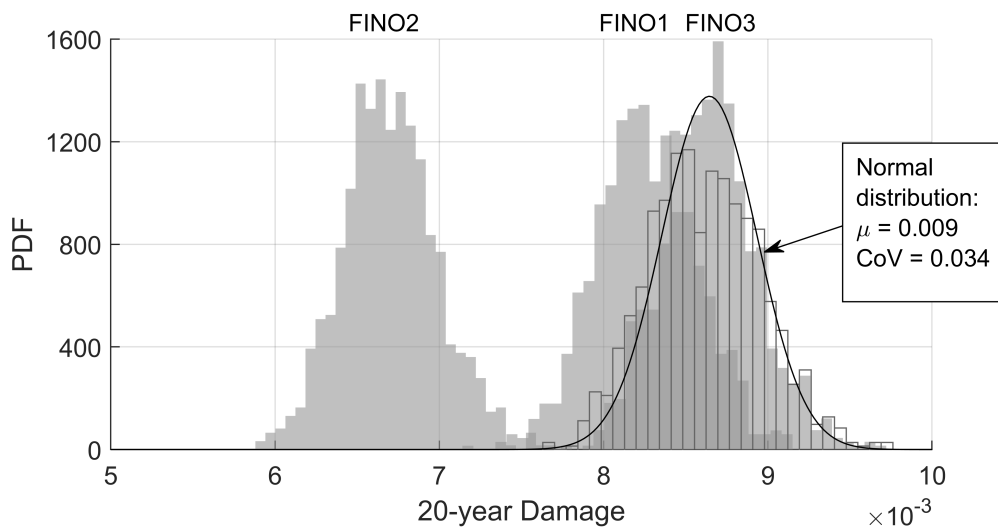




**Figure 7:** Impact of different combinations of samples and seed on the GP goodness-of-fit metrics; these estimates were generated by taking the mean of 1,000 bootstrap samples drawn from the full set ( $1,000 \times 6$ ) of random samples.

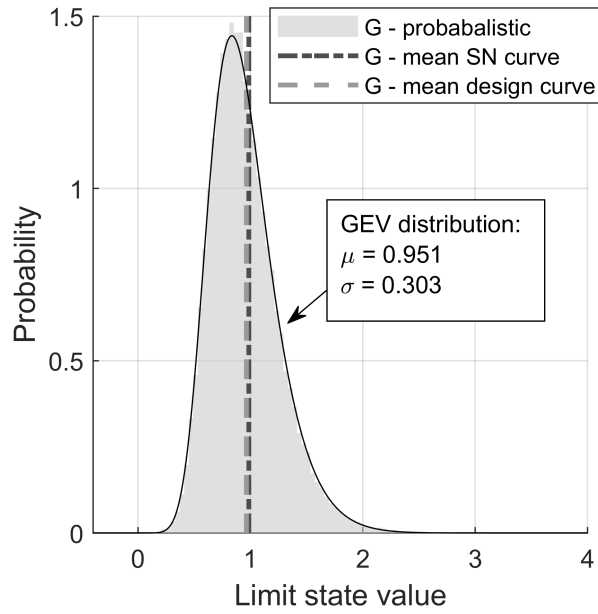


**Figure 8:** Probability distribution of  $T_p$  parameter across the FINO1 (left), FINO2 (middle) and FINO3 (right) sites.

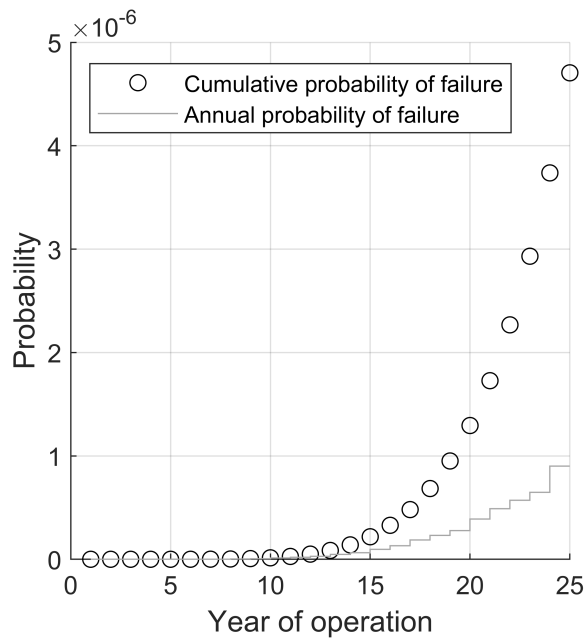


**Figure 9:** Empirical distributions of fatigue damage calculated for the three FINO sites (in block gray). The FINO3 site calculation was also run using the fully probabilistic GP with results shown in gray lines and the best-fit normal distribution in black.





**Figure 10:** Empirical distribution of the Monte Carlo limit state evaluations with a Gumbel distribution fit using MLE. The limit state evaluated using the mean SN curve and design SN curve with a damage threshold of 1 are also shown.



**Figure 11:** Probability of failure increasing with time.