Surrogate Model Uncertainty in Wind Turbine Reliability Assessment

R M M Slot, J D Sørensen, B. Sudret, L Svenningsen, M L Thøgersen

ETH zürich

Surrogate model uncertainty in wind turbine reliability assessment

R.M.M. Slot, J. D. Sørensen, B. Sudret, L. Svenningsen, M.L. Thøgersen

Chair of Risk, Safety and Uncertainty Quantification
Stefano-Franscini-Platz 5
CH-8093 Zürich

Risk, Safety & Uncertainty Quantification

Data Sheet

Journal:

Report Ref.: RSUQ-2019-005

Arxiv Ref.:

DOI: -

Date submitted: June 25, 2019

Date accepted: -

Surrogate Model Uncertainty in Wind Turbine Reliability Assessment

René M. M. Slot[1,3], John D. Sørensen[1], Bruno Sudret[2], Lasse Svenningsen[3], and Morten L. Thøgersen[3]

[1]Department of Civil Engineering, University of Aalborg, Aalborg, 9220, Denmark
[2]Chair of Risk, Safety and Uncertainty quantification, ETH, Zürich, Switzerland
[3]EMD International A/S, Aalborg, 9220, Denmark

Correspondence to: René M. M. Slot (rmms@civil.aau.dk)

Abstract. Lowering the cost of wind energy entails an optimization of material consumption of wind turbine components without compromising structural safety. Typically, wind turbines are designed by the partial safety factor method which is calibrated by full probabilistic models and presented in the IEC 61400-1 design standard. This approach significantly reduces the amount of aero-elastic simulations required to assess the fatigue limit state of wind turbines, but it may lead to inconsistent reliability levels across wind farm projects. To avoid this, wind turbines may be designed by full probabilistic methods using surrogate models to approximate fatigue load effects. Doing so, it is important to quantify and model all relevant uncertainties including that of the surrogate model itself. Here we quantify this uncertainty according to Eurocode 1990 for polynomial chaos expansion (PCE) and Kriging using wind data from 99 real sites and the 5MW reference turbine by NREL. We investigate a wide range of simulation efforts used to train the surrogates and our results show that Kriging yields a higher accuracy per invested simulation compared to PCE. This improved understanding of using PCE and Kriging in fatigue reliability assessment may significantly benefit decision support in full probabilistic design of wind turbines.

Key words: Wind turbine, Fatigue loads, Structural reliability, Surrogate models, Model uncertainty

1. Introduction

To lower the cost of wind energy it is important to utilize wind turbines to their full load bearing capacity but without compromising structural safety. A typical design approach is to follow the partial safety factor method, calibrated by full probabilistic models and presented in standardized codes as the IEC 61400-1 design standard for wind turbines[1]. This semi-probabilistic approach accounts for variability and uncertainty in strength and load parameters by using characteristic values defined by quantiles. A final design equation is then adjusted by partial safety factors to meet a target structural reliability level which is defined with consideration of economic loss and risk of human lives to optimize material consumption from a societal point of view.[2] This simplified framework provides a direct advantage in computational requirements to assess whether a given wind turbine class is suited for a particular site and park layout; However, it may lead to inconsistent reliability levels as the simple characteristic input cannot fully explain the variation of the load response.[3,4] To cover

all relevant structural components partial safety factors are therefore typically calibrated based on conservative assumptions. As a result, it can be expected that wind turbines in general are over-designed thereby leading to a higher cost of wind energy than necessary. To avoid this excess use of materials site-specific assessment of wind turbines may be directly based on a full probabilistic analysis as described in the recent 4th edition of the IEC 61400-1 design standard[1].

The main challenge in full probabilistic analysis of wind turbines is the need for significantly more load evaluations than the safety factor approach. In particular, fatigue analysis during normal operation (Design load case 1.2[1]) involves an infeasible amount of load simulations as fatigue damage accumulates during the whole lifetime of the turbine. It is therefore necessary to assess the integrated fatigue load across the entire joint wind climate distribution. For onshore wind turbines this includes at least wind direction (θ), wind speed (U), turbulence (σ_U), vertical wind shear exponent (α), air density (ρ), and flow inclination (φ).[5] Consequently, the sheer amount of aero-elastic simulations that needs to be carried out to fully evaluate the lifetime fatigue load imposes a computational barrier to probabilistic design.[6] To circumvent this barrier a shortcut from wind climate to wind turbine fatigue loads is required.

Various methods have been proposed to simplify wind turbine fatigue load assessment using surrogate models, also referred to as meta-models, response surfaces or proxies. This motivated Dimitrov et al.[7] to benchmark the accuracy of several surrogate techniques against each other. Their emphasis was prediction of lifetime fatigue loads by importance sampling, quadratic regression, nearest-neighbour interpolation, polynomial chaos expansion (PCE) and Kriging. Overall, they found Kriging and PCE to be superior, with Kriging having the highest accuracy overall, but also an increased computational time compared to PCE when predicting new samples.

Surrogate models make it feasible to carry out full probabilistic design of wind turbines. In Toft et al.[8] a quadratic response surface was used for reliability analysis of onshore wind turbines with focus on modelling wind climate uncertainties. Morató et al.[9] established a Kriging surrogate model to capture Von Mises stresses in a reliability analysis of offshore wind turbines in the ultimate limit state. In addition, they investigated the influence of the number of samples and seeds that were used to calibrate the model. With focus on offshore wind turbine fatigue loads Teixeira et al.[10] used a Kriging model to analyse the importance of different wind and wave climate parameters. In Murcia et al.[11] the uncertainty propagation properties of PCE was used to analyse the sensitivity of the wind climate on the power output and structural response of an onshore turbine. Focussing on blade design Hu et al.[12] proposed a reliability based design optimization which relied on multiple Kriging surrogate models to predict fatigue loads in critical structural hotspots. They included wind climate uncertainty, both spatial and temporal, while also considering manufacturing uncertainties of the composite laminate.

A common goal of the previous literature on wind turbine fatigue reliability is to establish novel reliability models and quantify the long chain of uncertainties from wind climate to wind turbine load effects. In this context, an important uncertainty is still missing in the literature; namely that of using a surrogate model to approximate fatigue loads instead of performing direct aero-elastic simulations. The scope of this paper is to study this uncertainty using wind measurements from 99 real wind turbine sites. This provides a solid basis to quantify a general uncertainty model for future applications. In this work PCE and Kriging are considered due to

their very promising capability in terms of capturing fatigue loads, propagating uncertainty and carrying out sensitivity analyses. The two techniques also represent two main approaches to predict a model output, namely regression (PCE) and interpolation (Kriging).

2. Wind Measurements

High quality measurements of wind direction, windspeed, turbulence, and wind shear from 99 real wind turbine sites are used in this study. All measurements are from meteorological masts or from wind power projects and represent a wide spread in geographical location and terrain complexity with and without nearby forestry. The measurements were also used by Slot et. al.[3] where a detailed description of the data may be found.

2.1. Ambient joint wind climate

The site-specific joint distribution of wind direction, windspeed, turbulence, and wind shear is described in terms of conditional distributions as summarized in Table 1. Measurements leading to air density and flow inclination were unavailable at most sites. Instead air density time series have been estimated from a meso-scale model while flow inclination is modelled as a fixed value dependent on the site-specific orography, see Table 2. These simplifications are not expected to have a significant effect given the low importance of both air density and flow inclination on site-specific fatigue loads compared to wind speed, turbulence and wind shear.[3,4,13]

Table 1: Joint wind climate distribution

Wind climate parameter	Notation	Description
Wind direction	$P_\theta(\theta)$	Discrete distribution[5]
Wind speed	$f_U(U\|\theta)$	Weibull distribution[5]
Turbulence[a]	$f_{\sigma_U}(\sigma_U\|U,\theta)$	Lognormal distribution[5]
Wind shear	$f_\alpha(\alpha\|\sigma_U,U,\theta)$	Normal distribution[4,13]
Air density	$f_\rho(\rho)$	Normal distribution[4]
Flow inclination	φ	Fixed value

To define the conditional distributions of wind speed the wind direction is binned into 12 sectors covering 30° each as recommended in the IEC 61400-1 standard[5]. The conditional distributions of turbulence are determined by further binning wind speed by 1 m/s. To model wind shear dependent on direction, wind speed and turbulence the procedure described in Dimitrov et al.[13]

[a] A Weibull distribution may also be considered as recommended in the IEC 61400-1 ed. 4 design standard.

is adopted. In each wind speed and direction bin all turbulence samples are ranked and divided into five equally sized bins, each assumed to represent a width of 0.2 of the turbulence cumulative distribution function. Subsequently a wind shear distribution is fitted to the shear values in each of the turbulence bins.

To define the turbulence and wind shear distributions the first and second moments of the available data are required. To obtain robust estimates only direction and wind speed bins with 50 or more samples are considered. For bins with less than 50 samples the distribution parameters are extrapolated in order to get a full description of the joint wind climate. Following the IEC 61400-1 standard[1] the mean value (μ_{σ_U}) and standard deviation (σ_{σ_U}) of the turbulence distribution are extrapolated by linear models as described by Eqs. (1) and (2).

$$\mu_{\sigma_U}(U|\theta) = a_{\mu_\sigma}(\theta) \cdot U + b_{\mu_\sigma}(\theta) \quad (1)$$

$$\sigma_{\sigma_U}(U|\theta) = a_{\sigma_\sigma}(\theta) \cdot U + b_{\sigma_\sigma}(\theta) \quad (2)$$

The mean value of the wind shear distribution (μ_α) is approximated as the median of μ_α at the three highest accepted wind speed bins given direction and turbulence. This imitates that atmospheric stability typically turn towards neutral conditions at medium to high wind speeds where the mean wind shear becomes constant.[14] Finally, the standard deviation of the wind shear distribution is extrapolated inversely proportional to wind speed as shown in Eq. (3).[15] All model parameters a, b, and c in are fitted based on the available data in accepted bins.

$$\sigma_\alpha(U|\theta, \sigma_U) = \frac{c_\alpha(\theta, \sigma_U)}{U} \quad (3)$$

Table 2: Flow inclination model

Site-specific orography	Flat terrain	Hilly terrain	Steep terrain
Fixed flow inclination	0°	6°	12°
Number of sites	62	27	10

2.2. Wake added turbulence

All the included measurements represent ambient climates. This does not reflect the reality of most turbines, where wakes are present in some directions. Wake added turbulence is therefore considered by assuming a rectangular grid layout where a neighbouring turbine is placed 5 rotor diameters (RD) up- and downwind in the main wind direction, and 3 RD perpendicular to that. The wake added turbulence ($\sigma_{U,wake}$) is modelled according to the IEC 61400-1 standard[1] as outlined in Eq. (4), where C_T is the thrust coefficient.

$$\sigma_{U,wake}(U,\theta) = \frac{U}{1.5 + \frac{0.8\,RD(\theta)}{\sqrt{C_T(U)}}} \tag{4}$$

To significantly limit the required computations in this paper without loss of generality the wakes are assumed to perfectly align with the defined sectors regardless of the distance between the two turbines. This is a simplification but the main purpose of introducing the wakes is to reflect a real case representative range of turbulence values. The ambient and wake added turbulence may be combined to the total turbulence that is experienced by the turbine ($\sigma_{U,T}$) by Eq. (5).

$$\sigma_{U,T}(U,\theta) = \sqrt{\sigma_U(U,\theta)^2 + \sigma_{U,wake}(U,\theta)^2} \tag{5}$$

3. Probabilistic Model for Fatigue Failure

The main objective of a probabilistic design approach is to assess the structural reliability of a given failure mode and check if it meets the target reliability level. This requires a representative limit state equation (LSE) to calculate the failure probability which cover relevant uncertainties on strengths and loads. In this section a simple yet representative LSE for wind turbine fatigue failure is presented.

3.1. Wind turbine simulation and fatigue loads

To define a representative LSE it is relevant to discuss how fatigue loads are calculated. In this paper the framework of "damage equivalent loads" (DEL) is adopted. It is therefore implicitly assumed that fatigue strength of materials is modelled by a linear SN-curve and that Miner's rule[16] may be used to accumulate fatigue damage from varying load effect amplitudes.

All fatigue loads are based on 10 min effective[b] simulations of the 5MW reference wind turbine by NREL[17] using its baseline controller. The turbine is simulated in the aero-servo-elastic software FAST[18] and each realized wind field is computed in TurbSIM[19] using the Kaimal spectrum[20]. The output of the simulations are timeseries of load effects for various sensors on the main components of the turbine, which are reduced to a spectrum of load effect amplitudes (ΔF_i) and a corresponding number of cycles (n_i) by Rainflow counting.[21] This is then further condensed to a single scalar, the DEL, which represents the load effect range that produces the same fatigue damage as the entire spectrum. The DEL is outlined in Eq. (6) where the wind climate parameters except turbulence are gathered in the vector $\boldsymbol{C} = [U, \alpha, \rho, \varphi]$ to ease notation. The equivalent number of cycles, N_{eq}, is used as a reference value and may be selected arbitrarily if it is kept the same when comparing different DELs.

[b] The term "effective" is used to indicate that transient start-up behaviour is removed.

$$DEL(\boldsymbol{C}, \sigma_{U,T}) = \left(\frac{1}{N_{eq}} \sum_i n_i \Delta F_i^m \right)^{1/m} \quad (6)$$

The combined site-specific equivalent fatigue load (F_{eq}) with a one year reference period may be assessed by Eq. (7) where f_{Site} models the joint wind climate distribution and $c_{sim} = \frac{1year}{T_{sim}}$ is a correction factor accounting for the simulation time (T_{sim}).

$$F_{eq,direct} = \left(c_{sim} \sum_\theta P_\theta(\theta) \int_{\sigma_U} \int_{\boldsymbol{C} \in \mathbb{R}^4} f_{Site}(\boldsymbol{C}, \sigma_U | \theta) DEL\left(\boldsymbol{C}, \sigma_{U,T}\right)^m d\boldsymbol{C} d\sigma_U \right)^{1/m} \quad (7)$$

To limit the amount of results that are shown in this paper only the six main sensors listed in Table 3 are considered. The sensors represent varying sensitivities to the wind climate parameters and controller actions and reflect the overall path of the wind loads from acting on the blades until being reacted by the soil. The fatigue strength of each component is modelled by typical Wöhler exponents used in the literature.[4,11,13,22]

Table 3: Wind turbine sensors

Component	Sensor description	Notation	Unit	Wöhler exponent
Blade	Blade root flap-wise bending	RootMyb1	kNm	10
Blade	Blade root edge-wise bending	RootMxb1	kNm	10
Main shaft	Low speed shaft torque	LSSGagMxa	kNm	6
Tower (top)	Yaw bearing tilt	YawBrMyp	kNm	4
Tower (top)	Yaw bearing yaw	YawBrMzp	kNm	4
Tower (bottom)	Tower bottom fore-aft bending	TwrBsMyt	kNm	4

3.2. Limit state equation

Following the described approach of evaluating fatigue loads directly by aero-elastic simulations a LSE for fatigue failure (g_{direct}) is defined as shown in Eq. (8).[22,23]

$$g_{direct}(z, t) = \Delta - \frac{N_{eq} t}{K} \left(X_{Load} X_{SCF} \frac{F_{eq,direct}}{z} \right)^m \quad (8)$$

Here z is the design parameter relating load effects to stresses and t is time in years. The model uncertainties Δ and K account for Miner's rule and the SN-approach respectively, and X_{Load}

and X_{SCF} models the uncertainty on wind load effects and stress concentration factors[c]. Typical examples of the model uncertainties are presented in Table 4.[8,22,23]

Table 4. Stochastic models

Variable	Distribution	Expected value	Standard deviation		
			m=4	m=6	m=10
Δ	Normal	1	0.30	0.40	0.50
X_{Load}	Lognormal	1	0.15	0.15	0.15
X_{SCF}	Lognormal	1	0.10	0.15	0.15
$\log K$	Normal	-	0.20	0.15	0.25

Based on the LSE the accumulated failure probability (P_f) and associated reliability index (β) of the considered component may be estimated by Eq. (9), where Φ is the cumulative standard normal distribution.

$$P_f = \Phi(-\beta) = P(\mathrm{g}_{direct}(z,t) \leq 0) \tag{9}$$

To evaluate the failure probability the accumulated fatigue damage across the entire joint wind climate has to be estimated. Direct aero-elastic simulation for this application is extremely computationally demanding, and in most cases unfeasible. This barrier may be overcome by accurate surrogate models to predict DELs as discussed in the introduction. Surrogates makes it possible to evaluate the site-specific fatigue loads but with a reduced accuracy compared to direct simulation ($F_{eq,proxy}$). This should be properly accounted for in the LSE by introducing an additional model uncertainty related to the surrogate model itself (X_{Proxy}). The model uncertainty is related to the load effect, similar to X_{Load} and X_{SCF}, and is applied alongside these as shown in Eq. (10).

$$\mathrm{g}_{Proxy}(\mathrm{z},\mathrm{t}) = \Delta - \frac{N_{\mathrm{eq}} t}{K}\left(X_{Load} X_{SCF} X_{Proxy} \frac{F_{eq,proxy}}{z}\right)^m \tag{10}$$

4. Surrogate Models for Fatigue Load Prediction

The surrogate model uncertainty will depend on the surrogate model type and how many simulations that have been invested in training it.[7] Two surrogate techniques are included in this work namely Kriging and PCE, and both are implemented in the general purpose uncertainty quantification framework UQLab[24]. This section outlines the input domain and the experimental design that is used for training the models followed by a brief summary of each surrogate technique with emphasis on how they are configured in this specific work. For a more general explanation of the details and theory behind Kriging we refer to Santner et al.[25] and for PCE we refer to Sudret[26].

[c] The linear relationship between load effects and stresses is based on simple beam theory. To account for non-linear effects a "stress concentration factor" is typically applied.

4.1. Input domain and experimental design

Regardless of the choice of surrogate model it is necessary to sample an experimental design. This requires an input domain that covers the joint wind climate distributions across all 99 sites as to avoid extrapolation by the surrogates which may lead to very unpredictable results. Meanwhile, the input domain should also encompass "physically realistic" wind climate combinations to ensure validity of the aero-elastic simulations used to estimate the output DELs. An input domain that meets these two objectives was defined by Dimitrov et al.[7] partly based on theoretical considerations of atmospheric stability. In this work it is chosen to tailor the input domain specifically to the 99 available sites by using all data including wake added turbulence. This is illustrated in Figure 1 for turbulence intensity[d] (i.e. 10min wind speed coefficient of variation) and wind shear as function of wind speed. The bounds (blue lines) are based on approximations to the extreme quantiles of all data (red lines) with a slight conservative offset.

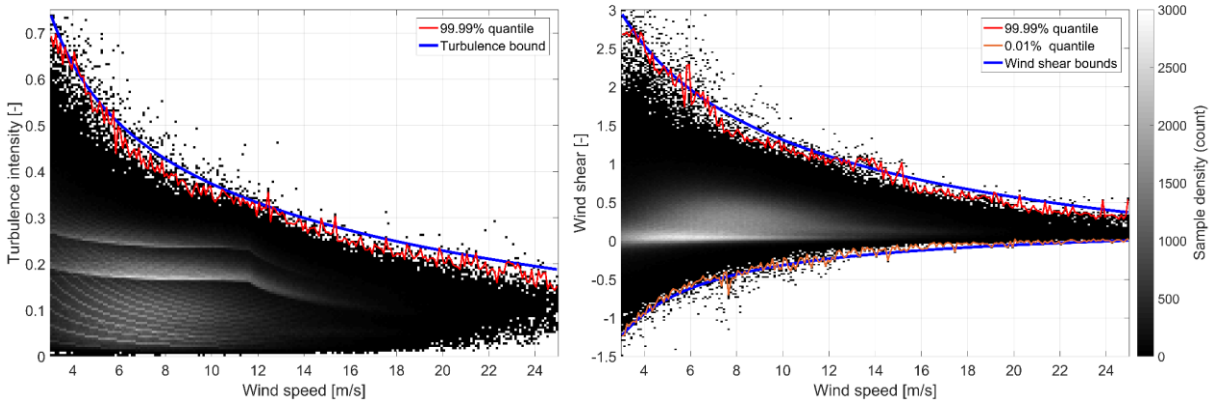

Figure 1: Turbulence bound (left) and wind shear bounds (right) as function of wind speed based on all available measurements. The pronounced clear lines inside the turbulence samples at 0.2 and 0.3 correspond to the wake added turbulence at 3 RD and 5 RD. The smaller clear lines are a product of the decimal truncation when the wind measurements are logged.

The bounds on air density are based on the meso-scale modelled data as shown in Figure 2, and the bounds on flow inclination are based on engineering judgement in the interval from -16° to 16°.

[d] Turbulence intensity is used such that the slight offset of the bound account for the increasing scale of turbulence with increasing wind speed.

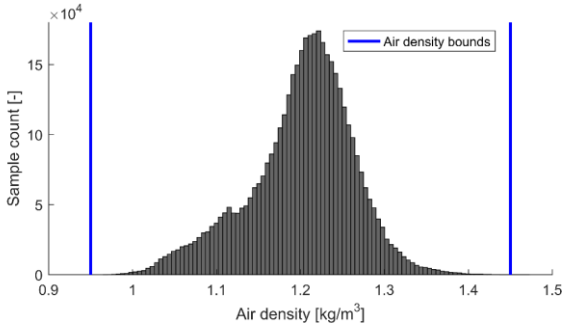

Figure 2: Air density bounds based on all available measurements.

Together, the defined bounds enclose a hyper-volume in the 5-dimensional space of the considered wind climate parameters. An experimental design consisting of 625 is sampled uniformly inside this hyper-volume using a quasi-random Halton sequence. This ensure a good space-filling both for the entire experimental design but also when only a subset of the experimental design is considered. To avoid any clear patterns in the Halton sequence, which may compromise the accuracy of the surrogates, a reverse-radix scrambling is performed as described by Kocis and Whiten[27]. The experimental design is shown in Figure 3. Each of the corresponding DELs are estimated using 100 seeds, resulting in a total of 62,500 aero-elastic simulations.

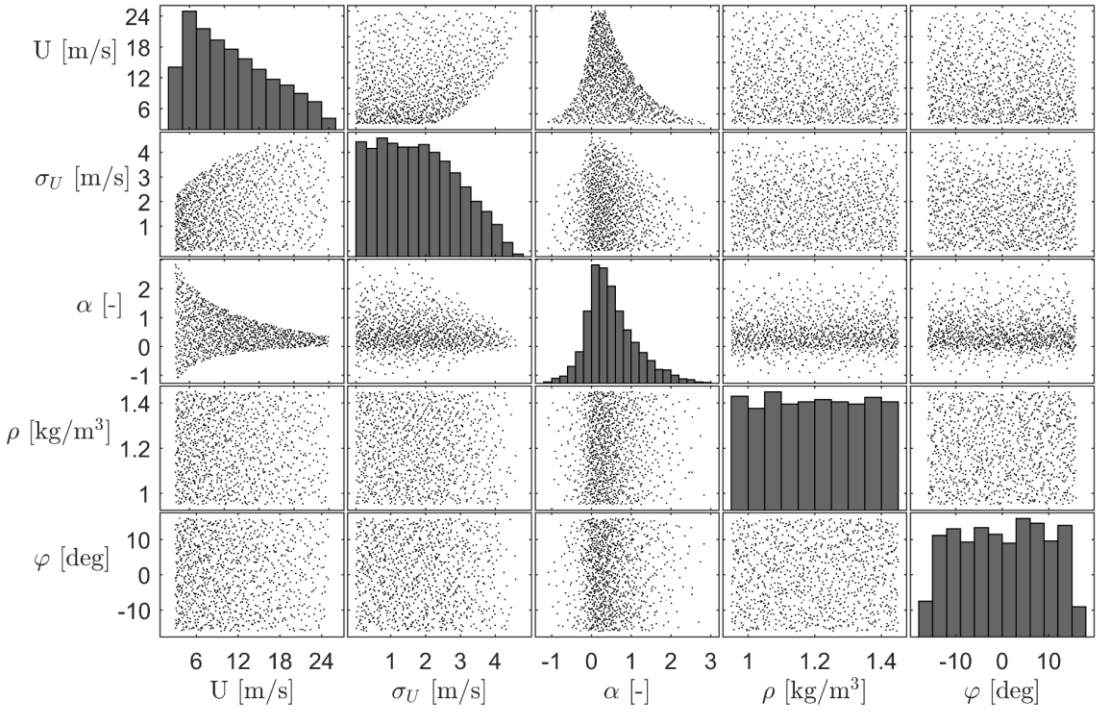

Figure 3: Experimental design with 625 samples drawn uniformly inside the bounded hyper-volume of the wind climate parameters.

4.2. Polynomial chaos expansion

In this work the Wiener-Askey generalized PCE is considered.[28] Generally, it may be used to approximate a function (Y) of a random vector ($\boldsymbol{X}$) in dimension D ($\boldsymbol{X} \in \mathbb{R}^D$) by an infinite expansion of a multivariate orthonormal polynomial basis (Ψ) with respect to the joint probability density function of the input. This is outlined in Eq. (11) where j is a multi-index of the components of the multivariate polynomials.

$$\mathrm{Y}(\boldsymbol{X}) = \sum_{j \in \mathbb{N}^D} c_j \Psi_j(\boldsymbol{X}) \tag{11}$$

The polynomial basis is built from a set (family) of univariate orthonormal polynomials with respect to each input variable, and classical families have been developed which cover common input distributions.[28] For simplicity, the experimental design is therefore transformed into the standard uniform space by Rosenblatt transformation[29]. Consequently, only the Legendre family of orthonormal polynomials is required to build the PCE.

For practical application the infinite sum of polynomials in Eq. (11) is truncated using a hybrid least angle regression algorithm to penalize higher order terms together with a hyperbolic truncation scheme to disregard insignificant interactive terms, see details in Blatman and Sudret[30]. By considering polynomial degrees up to 20 the PCE which minimize the leave one out cross-validation error (ϵ_{LOO}) is selected following the implementation in UQLab[31]. The ϵ_{LOO} is chosen as optimization metric to increase robustness towards over-fitting of the PCE when high order polynomials are considered.

4.3. Kriging

Kriging is a stochastic interpolation technique which assumes the model output (Y) to be a realization of a deterministic mean defined by a regression model ($\boldsymbol{\beta}_{KRG}\boldsymbol{f}_{KRG}^{\mathrm{T}}$) and a correlated stochastic process (Z), see Eq. (12).[25]

$$\mathrm{Y}(\boldsymbol{X}) = \boldsymbol{\beta}_{KRG}\boldsymbol{f}_{KRG}^{\mathrm{T}}(\boldsymbol{X}) + Z(\boldsymbol{X}) \tag{12}$$

The first term models the trend (mean) of the output by a set of basis functions $\boldsymbol{f}_{KRG}(\boldsymbol{X}) = [f_1(\boldsymbol{X}), \dots, f_n(\boldsymbol{X})]$ and associated regression coefficients $\boldsymbol{\beta}_{KRG} = [\beta_1, \dots, \beta_n]$. The second term is interpolating the known residuals at the experimental design by a stationary zero mean Gaussian process (Z) fully described by its covariance (cov):

$$cov(\boldsymbol{X}, \boldsymbol{X}') = \sigma_{KRG}^2 R(\boldsymbol{X}, \boldsymbol{X}', \boldsymbol{\theta}_{\boldsymbol{R}}) \tag{13}$$

Here σ_{KRG}^2 is the overall process variance (assumed constant) and R models the correlation between $Z(\boldsymbol{X})$ and $Z(\boldsymbol{X}')$ by their inter-distance and a correlation function defined by the hyper parameters $\boldsymbol{\theta}_{\boldsymbol{R}}$. Once a suitable basis of functions and a correlation model is chosen $\boldsymbol{\beta}_{KRG}$, σ_{KRG}^2, and $\boldsymbol{\theta}_{\boldsymbol{R}}$ may be estimated by maximizing the likelihood of observing the output at the experimental design.[32]

A priori it is known that the sensitivity between fatigue loads and the different wind climate parameters vary significantly. Therefore, an anisotropic separable correlation formulation is considered as shown in Eq. (14).

$$R(\boldsymbol{X}, \boldsymbol{X}', \boldsymbol{\theta_R}) = \prod_{i=1}^{D} R(X_i, X_i', \theta_{R,i}) \tag{14}$$

A main challenge when calibrating an accurate Kriging model is to select an appropriate trend and correlation function. By a combinatorial approach similar to Morató et al.[9] we found universal Kriging with a quadratic trend and the Matérn 3/2 correlation function to yield the best results overall.

5. Method for Assessment of Surrogate Model Uncertainty

The surrogate model uncertainties of PCE and Kriging are estimated according to EN 1990[33]. In this section the method is briefly outlined followed by a description of the numerical integration scheme that is used to assess the site-specific fatigue loads.

5.1. EN 1990 method

The model uncertainty X_{proxy} is estimated by rewriting it in terms of a unit mean lognormal error term ($X_{proxy,EN}$) and a mean value correction factor to account for the model bias (b_{proxy}), see Eq. (15).

$$F_{eq,direct} = b_{proxy} X_{proxy,EN} F_{eq,proxy} \tag{15}$$

Given the available set of 99 statistically independent[e] joint wind climates the bias may be estimated using a least squares approach as shown in Eq. (16).

$$b_{proxy} = \frac{\sum_{i=1}^{99} F_{eq,direct,i} F_{eq,proxy,i}}{\sum_{i=1}^{99} F_{eq,proxy,i}^2} \tag{16}$$

Next, the logarithm of the residuals at each site is estimated by Eq. (17).

$$\delta_{EN,i} = \ln\left(\frac{F_{eq,direct,i}}{b_{proxy,EN} F_{eq,direct,i}}\right) \tag{17}$$

An estimate of the standard deviation of the residuals ($\sigma_{\delta,EN}$) is then assessed by Eq. (18) where $\delta_{EN,\mu}$ is the mean value of all error realizations.

[e] The diversity of the included sites in terms of complexity and geographical spread validate the assumption of independence.

$$\sigma_{\delta,EN} = \sqrt{\frac{1}{99-1}\sum_{i=1}^{99}\left(\delta_{EN,i} - \delta_{EN,\mu}\right)^2} \tag{18}$$

The coefficient of variation of the lognormal surrogate model uncertainty (V_{proxy}) is then obtained by Eq. (19).

$$V_{proxy} = \sqrt{e^{\sigma_{\delta,EN}^2} - 1} \tag{19}$$

5.2. Numerical fatigue load integration

To estimate the surrogate model uncertainty, it is required to assess F_{eq} by direct simulation. This is not trivial and involves hundreds of thousands of aero-elastic simulations. It is therefore relevant to discuss the applied method for integrating the fatigue load in detail in the following.

With flow inclination being fixed (φ_{fix}) the dimension of the integration problem in Eq. (7) is reduced as shown in Eq. (20). Here the 12 discrete directions are directly introduced and $\boldsymbol{C}_R = [U, \alpha, \rho]$ contain the reduced set of wind climate parameters of wind speed, wind shear and air density.

$$\begin{aligned} &F_{eq}(\varphi_{fix}) \\ &= \left(c_{sim} \sum_{k=1}^{12} p_\theta(\theta_k) \underbrace{\int_{\sigma_U} \int_{\boldsymbol{C}_R \in R^3} f_{Site}(\boldsymbol{C}_R, \sigma_U | \, \theta_k) DEL\left(\boldsymbol{C}_R, \sigma_{U,T}, \varphi_{fix}, m\right)^m d\boldsymbol{C}_R \, d\sigma_U}_{F_{eq,sect,k}^m} \right)^{\frac{1}{m}} \end{aligned} \tag{20}$$

To find an optimal numerical integration scheme it is sufficient to consider the sector wise fatigue loads ($F_{eq,sect,k}$), which mathematically is the m'th order weighted Hölder mean of DEL with respect to f_{Site}. In turn, $F_{eq,sect,k}^m$ is the expected value of DEL^m and may be estimated approximatively by Monte-Carlo (MC) sampling, which avoids the "curse of dimensionality" associated to traditional grid-based integration.[6] The MC-integration is outlined in Eq. (21) where N is the number of samples and p models the probability of generating sample i in terms of ambient turbulence.

$$F_{eq,sect,k}^m(\varphi_{fix}) \approx \frac{1}{N}\sum_{i=1}^{N} \frac{f_{Site}\left(\boldsymbol{C}_{R,i}, \sigma_{U,i} | \, \theta_k\right) DEL\left(\boldsymbol{C}_{R,i}, \sigma_{U,T,i}, \varphi_{fix}, m\right)^m}{p(\boldsymbol{C}_{R,i}, \sigma_{U,i})} \tag{21}$$

The convergence behaviour depends on the choice of the MC-sampling distribution (h_{MC}). In principle convergence is obtained faster if it resembles the product of DEL^m and f_{Site} to concentrate samples in the region which contributes most to the integral (i.e. importance sampling). This leads to the following main considerations for h_{MC}:

1. Since DELs are raised to the power of m it is important to sample high fatigue load events with a low probability of occurrence.
2. Wind climate combinations with a high probability of occurrence have to be sampled to capture the majority of the wind turbines lifetime.

The importance of point 1 was clearly shown by Graf et al.[6] where MC-integration was benchmarked by sampling from the joint wind climate distribution (i.e. $h_{MC} = f_{Site}$) which lead to very slow convergence rates for components with high Wöhler exponents.

Optimizing h_{MC} with respect to point 1 is sensor-specific but in general it requires unlikely wind climates to be sampled (e.g. very high turbulence or wind shear). On the other hand, optimization of h_{MC} with respect to point 2 is site-specific and requires likely wind climates to be sampled. To cover all 99 sites and all considered sensors a straight-forward compromise is then to sample equally dense across the entire input domain defined in Section 4 (i.e. h_{MC} uniformly distributed).

Several techniques can be used to sample from h_{MC} but in low dimensions (lower than 6) it was shown by Morokoff and Catflisch[34] that quasi-random numbers from low-discrepancy sequences provide fast convergence. Instead of a convergence ratio $\propto N^{-0.5}$ for crude MC a Halton sequence obtain a convergence ratio $\propto N^{-\lambda}$ where $0.5 \leq \lambda \leq 1$ with $\lambda \to 1$ for low dimensions.

A main question to answer is how many samples that are needed to assess F_{eq}. Based on the results of Graf et al.[6] and a preliminary convergence study using a surrogate it was found reasonable to evaluate the integral by 25,000 DELs. This results in an accuracy within approximately 1% of the converged value obtained at one million samples for all considered sensors. Given that the surrogate model uncertainty is obtained by a relative comparison of $F_{eq,direct}$ and $F_{eq,proxy}$, both estimated by the same MC-samples, the small error is assumed to be insignificant.

Finally, it is noted that each sample in the MC-integration correspond to one specific flow inclination so three databases of fatigue loads were simulated to cover all 99 sites. Using 100 seeds to estimate each DEL this resulted in a total of 7.5 million simulations to accurately assess $F_{eq,direct}$.

6. Quantification of the Surrogate Model Uncertainty

Using the presented methods, the surrogate model uncertainty across all considered sensors is quantified. First an example is shown for the blade root flap-wise bending moment to clearly outline how the results are obtained step-by-step. Then the surrogate model uncertainty is

quantified for all considered sensors and presented as the worst case across the sensors as it is impractical to differentiate the surrogate model uncertainty for each sensor[f].

6.1. Blade root bending moment example

Using the full experimental design, a Kriging and PCE model is trained to capture fatigue loads of the blades. Both surrogates are then used to estimate the site-specific fatigue loads across all 99 sites by Eq. (20) which is compared to direct simulation as shown in Figure 4.

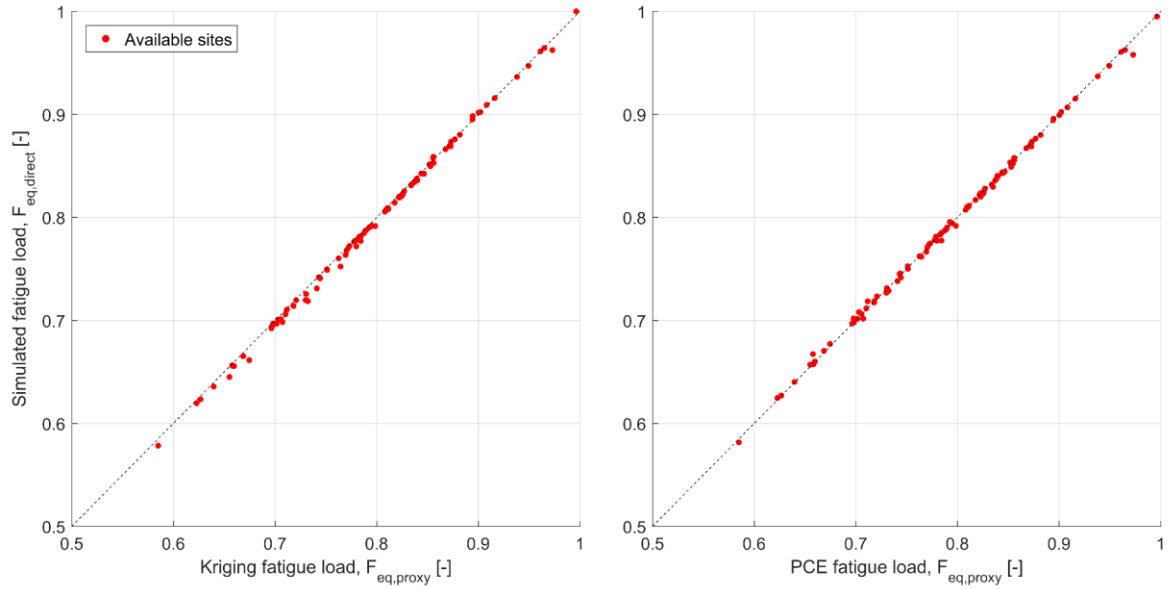

Figure 4: Normalized site-specific fatigue loads on the blades predicted by Kriging and PCE and compared to direct simulation. The dashed lines indicate the perfect model where $F_{eq,proxy} = F_{eq,direct}$

Based on this comparison the surrogate model uncertainty is estimated by Eqs. (16) and (19). The results are presented in Table 5.

Table 5: Surrogate model uncertainty for blade flap-wise bending.

Surrogate model	Bias [-]	Coefficient of variation [-]
PCE	1.001	0.004
Kriging	1.004	0.004

Given the typical scale of the other uncertainties in the LSE listed in Table 4 the resulting surrogate model uncertainties of both PCE and Kriging are insignificant and may be neglected.

[f] This representation of the surrogate model uncertainty is in line with the model uncertainty for wind load effects being identical across all considered sensors.

Next, it is investigated how the results change with the amount of simulations that are invested in training the surrogate models. This is illustrated in Figure 5 where the coefficient of variation is plotted as function of the number of samples in the experimental design and how many seeds that are used to evaluate the DELs. Similarly, the model bias is shown as function of invested simulations in Figure 6. The number of samples starts at 100 as both models become significantly inaccurate for smaller experimental designs. The noise in the PCE results is explained by the adaptive scheme to select the optimal polynomial degree which may change for each combination of samples/seeds in order to minimize ϵ_{LOO}.

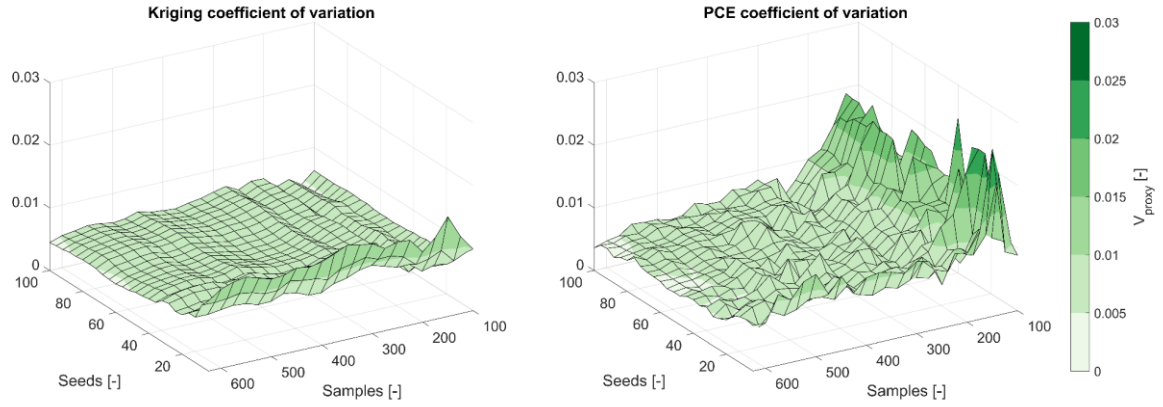

Figure 5: Coefficient of variation of the surrogate model uncertainty for blades using Kriging and PCE shown as function of the number of samples used for training and number of seeds used to evaluate DELs at each sample.

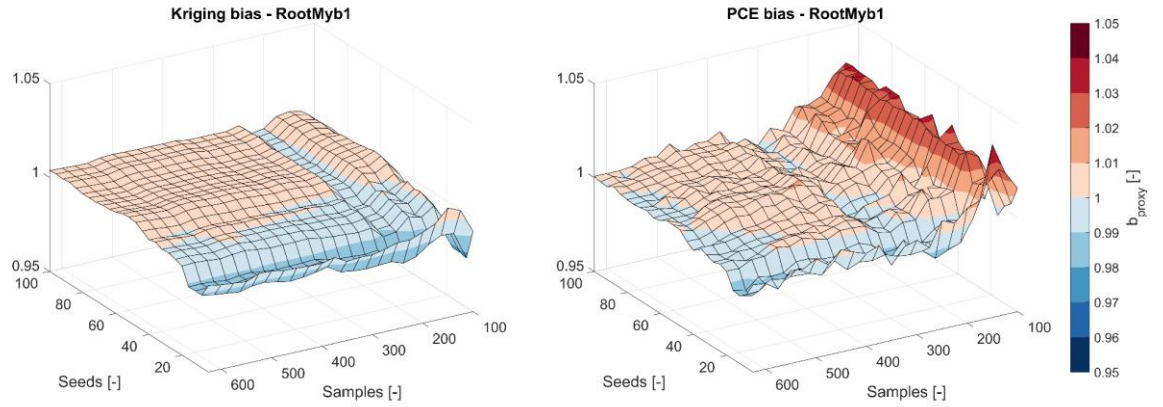

Figure 6: Surrogate model bias for blades using Kriging and PCE shown as function of the number of samples used for training and number of seeds used to evaluate DELs at each sample. The apparent "valley" in the Kriging results at approximately 200 samples is due to the oscillating convergence behaviour of the bias across samples. This is clearly seen when less than 100 samples are considered, but to keep the figures consistent and clean this is not included here.

6.2. Surrogate model uncertainty across all sensors

Following the described procedure for the blade flap-wise bending the surrogate model uncertainty has been computed for all sensors in Table 3. The maximum model uncertainty across the sensors is plotted in Figure 7 for both surrogates.

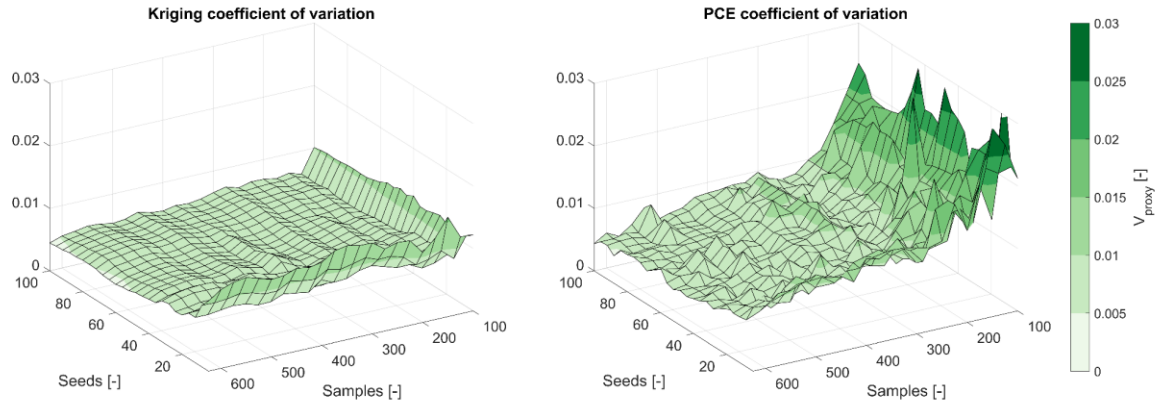

Figure 7: Surrogate model uncertainty coefficient of variations across all considered sensors shown as function of the number of samples used for training and number of seeds used to evaluate DELs at each sample.

Relatively few samples and seeds are required for the surrogate models to obtain a converged model uncertainty of ~0.5%.

To analyse the surrogate model bias it is necessary to consider both the minimum and maximum across all sensors as illustrated in Figure 8 and Figure 9, respectively. This show that for sparse designs (low number of samples) PCE tends to be non-conservative while Kriging seems to be almost exclusively conservative across all sensors.

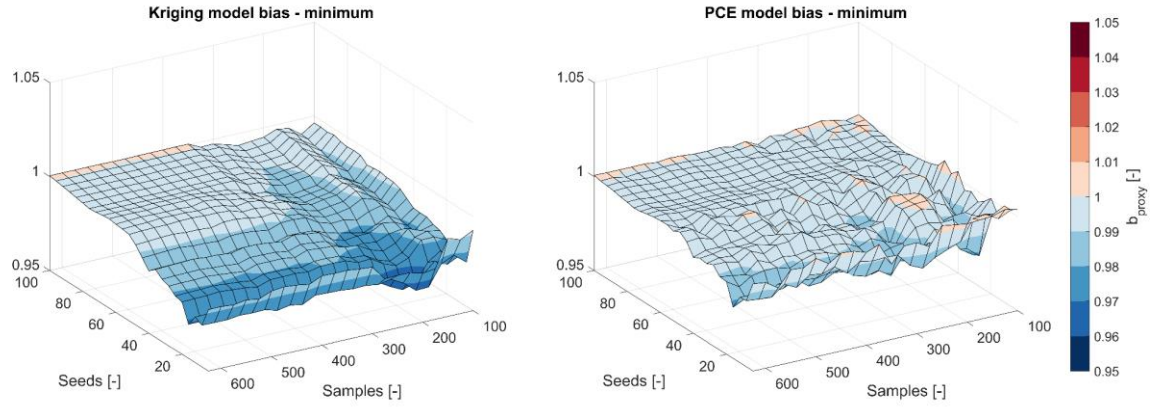

Figure 8: Minimum surrogate model bias across all considered sensors shown as function of the number of samples used for training and number of seeds used to evaluate DELs at each sample.

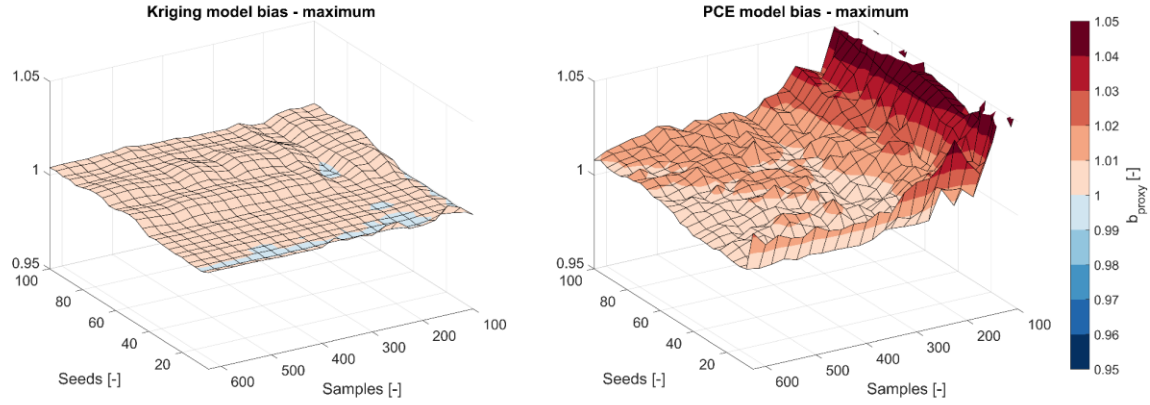

Figure 9: Maximum surrogate model bias across all considered sensors shown as function of the number of samples used for training and number of seeds used to evaluate DELs at each sample. The PCE model bias at 100-150 samples is missing due to a sharp increase to 1.10 which would otherwise distort the figure.

7. Sensitivity Analysis and General Recommendations

Both Kriging and PCE can be trained to accurately predict fatigue loads with a relatively small computational investment. In this section a sensitivity analysis is conducted to assess how the surrogate model uncertainty affects the reliability level of the considered components. This is used to define three levels of accuracy followed by a set of recommendations by the authors on training PCE and Kriging for fatigue reliability analysis of wind turbines.

7.1. Reliability sensitivity analysis

To estimate the sensitivity of the surrogate model uncertainty a reference design is established by assuming that each component is designed to the limit using direct aero-elastic simulation. This implies that the reliability index at the last year of service is exactly the target of $\Delta\beta_t = 3.3$ [1], where a lifetime of 20 years is assumed. Here the yearly probability of failure is approximated as $\Delta P_{f,20} \cong P_{f,20} - P_{f,19}$ where $P_{f,20}$ and $P_{f,19}$ represent the failure probability at years 19 and 20, respectively. The failure probabilities are estimated by the LSE presented in Eq. (8) where the uncertainties are modelled according to Table 4.

Using the reference design of each component the reliability index at the last year of service ($\Delta\beta$) is calculated by the LSE in Eq. (10) where b_{proxy} and V_{proxy} are varied individually. The sensitivity of the surrogate model uncertainty is then quantified as the ratio of the reliability index with respect to the target reliability ($\Delta\beta/\Delta\beta_t$). In Figure 10 this is shown for a representative range of V_{proxy} for both PCE and Kriging where the relative change of the reliability is less than 0.5%. Figure 11 shows the results of varying b_{proxy} within a representative range, demonstrating that $\Delta\beta/\Delta\beta_t \cong b_{proxy}$ for all components considered.

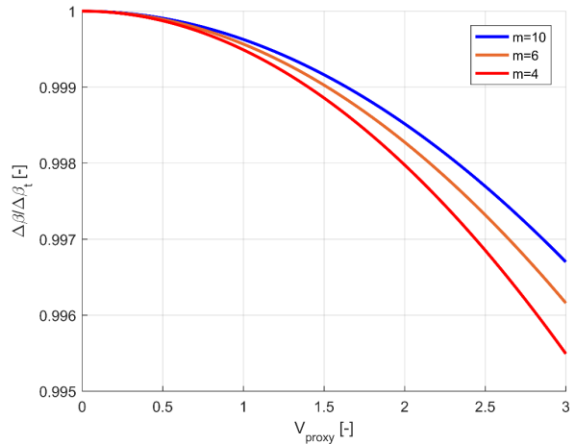

Figure 10: Reliability index sensitivity to the surrogate model uncertainty coefficient of variation.

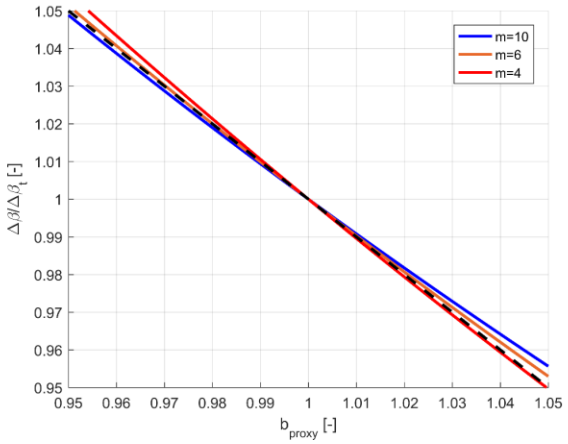

Figure 11: Reliability index sensitivity to surrogate model bias. The dashed line indicates $\Delta\beta/\Delta\beta_t = b_{proxy}$.

7.2. Surrogate model accuracy

Typically, the target reliability index is specified with one decimal suggesting that a change of less than 1% of the reliability index at the limit is negligible. V_{proxy} may therefore be ignored in reliability analysis of wind turbines for both Kriging and PCE when at least 100 samples are considered in the experimental design.

The change in reliability is roughly proportional to b_{proxy} which makes the bias significant when it is outside the range of 0.99 to 1.01. This is made clear by Table 6, which leads to the definition of three accuracy classes.

Table 6: Absolute reliability for varying surrogate model bias.

b_{proxy}	0.95	0.96	0.97	0.98	0.99	1.00	1.01	1.02	1.03	1.04	1.05
$\Delta\beta$	3.1	3.2	3.2	3.2	3.3	3.3	3.3	3.4	3.4	3.4	3.5

High accuracy is obtained when $0.99 \leq b_{proxy} \leq 1.01$ where the estimated reliability index is unchanged on the first decimal. Medium accuracy is obtained when $0.96 \leq b_{proxy,EN} \leq 1.04$ leading to a slight change of ±0.1 of the estimated reliability index. If $0.96 > b_{proxy,EN} > 1.04$ low accuracy is obtained which potentially changes the estimated reliability index significantly with more than ±0.2. The accuracy classes are illustrated for both PCE and Kriging in Figure 12.

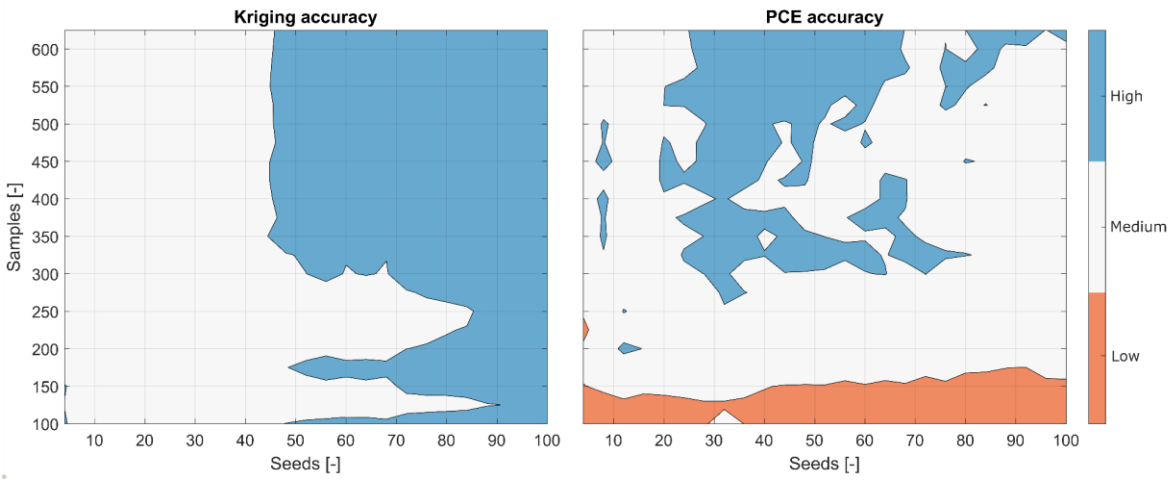

Figure 12: Surrogate model bias divided by three accuracy classes. Notice the clear pattern of increasing accuracy with the number of seeds for Kriging and the increased accuracy with number of samples for PCE.

For Kriging it is seen that the accuracy depends mostly on the number of seeds that are used. The upper right corner confirms a converged plateau of high accuracy which require an experimental design with more than 350 samples using at least 50 seeds to assess each DEL. If a medium accuracy is acceptable it is only required to use more than 100 samples and at least 4 seeds.

For PCE the accuracy depends on the density of the experimental design. A high accuracy may be obtained, but even when the full design is used the accuracy is not fully consistent with an increase in seeds. This indicates that the PCE coefficients may not have converged in the investigated range of samples and seeds. If more than 175 samples are considered the PCE model reach a medium accuracy, and a low accuracy is obtained for experimental designs with less than ~175 samples.

7.3. Recommendations

The results of the optimization analysis are summarized in Tables 7 and 8 as a set of recommendations towards training Kriging and PCE for fatigue reliability analysis. Since only a subset of load-bearing wind turbine components are considered the recommendations are slightly conservative compared to the results shown in Figure 12.

It is important to note that the recommendations are tied to the current setup of using the 5MW reference turbine by NREL together with the experimental design described in section 4. The minimum samples and seeds are also dependent on the size of the input domain which directly relates to the sample density. If a similar experimental design is used together with a smaller or larger input domain emphasis should therefore be put on the recommended maximum leave one out error. Because the proposed experimental design is uniformly distributed without any emphasis on the specific turbine response it is also expected that the recommended maximum leave one out error can be used as tentative guidance for other (similar) turbines.

Table 7: Kriging recommendations

Accuracy	Minimum samples in DoE [-]	Minimum seeds to assess DEL [-]	Total amount of 10-min simulations [-]	Maximum ϵ_{LOO} [%]
High	400	75	30.000	0.60
Medium	100	10	1.000	2.00
Low	100	4	400	3.25

Table 8: PCE recommendations

Accuracy	Minimum samples in DoE [-]	Minimum seeds to assess DEL [-]	Total amount of 10-min simulations [-]	Maximum ϵ_{LOO} [%]
High	>625	50	>31.250	0.35
Medium	200	10	2.000	1.35
Low	100	4	400	2.20

8. Comparison and Discussion of PCE and Kriging

Overall Kriging obtains a better accuracy than PCE per simulation, but it also requires more computational power to predict new samples compared to PCE.[7] However, it is noted that h_{MC} in the presented numerical integration is completely independent of the sector-wise joint wind climate distributions. A direct advantage of this is that the DEL of each MC-sample has to be evaluated only once per fixed flow inclination to cover all 99 sites. This neglects the difference in computational time between PCE and Kriging when predicting DELs compared to the simulations invested in training the surrogates.

To increase the accuracy of PCE it requires more samples while Kriging requires more seeds. This observation is in line with the fundamentals of the surrogates of regression and interpolation, respectively. If few seeds are used to evaluate the DELs it can be interpreted as noise which Kriging is forced to capture when interpolating the residuals. On the other hand, PCE levels out the noise in a mean sense, but a relatively large number of samples are needed to reliably fit the coefficients of the expansion.

While not shown in this paper, this encouraged an analysis of using Polynomial Chaos Kriging[35] (PCK) as implemented in UQLab[36]. This method combines the two surrogate models by using the PCE as basis functions (trend) for the universal Kriging model. In a best-case scenario this could lead to a surrogate model which is accurate for a small experimental design (as Kriging) using just a few seeds to estimate DELs (as PCE). However, the results of using PCK mostly resembles that of using PCE, and it is therefore not preferred over universal Kriging. The reason is probably that the residuals used to estimate σ^2_{KRG} and $\boldsymbol{\theta_R}$ appear without any significant

correlation as they mostly just represent the noise from using few seeds. In turn, the interpolation only influences predictions at the very vicinity of the experimental design, which was also observed by Dimitrov et al.[7]

Another possibility to increase the accuracy of Kriging is to introduce a so-called "nugget". The nugget models a set of values that are added to the diagonal of $R(\boldsymbol{X}, \boldsymbol{X}', \theta_R)$, thereby allowing a non-zero uncertainty bound around the experimental design. A well optimized nugget could therefore potentially make the Kriging model more robust to the noise from using a small number of seeds.[37] To the best of the authors knowledge a general framework to include and optimize a nugget in universal Kriging is not yet developed, and therefore this approach was not pursued further.

9. Summary and Conclusions

Using wind data from 99 international sites the model uncertainty related to approximating lifetime fatigue loads by Kriging and PCE has been quantified based on aero-elastic simulations of the 5MW reference turbine by NREL. The main components of the turbine were considered namely; blades, low speed shaft, yaw bearing, and tower.

All available data was used to define an experimental design to train the surrogates in terms of wind speed, turbulence, wind shear exponent, air density and flow inclination. Using up to 625 samples, with up to 100 seeds to evaluate fatigue loads at each sample, both surrogate models were calibrated using UQLab. For Kriging, a combinatorial approach was used to conclude that universal Kriging with a second order trend and the Matérn 3/2 correlation is optimal to predict fatigue loads.

The model uncertainty of both surrogates across all sensors was estimated by the recommended approach in Eurocode 1990, Annex D. This showed that the model uncertainty coefficient of variation is less than 2.5%. In contrast, the model bias varied significantly for the two surrogate models between 0.95 to 1.05.

Based on a sensitivity study it was shown that the model uncertainty coefficient of variation changes the structural reliability index less than 0.5% which is negligible compared to the target reliability of 3.3. In addition, the sensitivity study showed that the relative change in reliability index is approximately equal to the model bias across all considered components.

Three accuracy classes were introduced based on the bias of the surrogate models namely high, medium, and low accuracy. Compared to direct simulation a surrogate model with a high accuracy estimate reliability indices within ±0.05, for medium accuracy the error is within ±0.15 and for low accuracy the error on the first decimal is ≥0.15.

It was documented that Kriging obtains a high accuracy for the 5MW turbine when more than 30.000 aero-elastic simulations are performed. In comparison, PCE did not consistently obtain a high accuracy in the investigated range of 625 samples and 100 seeds. However, using more than 200 samples and 10 seeds PCE obtain a medium accuracy.

Overall, Kriging obtains a higher accuracy per invested simulation compared to PCE. Kriging is therefore the preferred method over PCE for reliability analysis of onshore wind turbines, when using the methods described in this paper.

Acknowledgements

The authors wish to thank the data providers: KNMI, ICDC, CliSAP/KlimaCampus, University of Hamburg, DTU, Vattenfall and VENTUS INGENIERÍA. The work presented in this paper is part of the PhD project "From wind climate to wind turbine loads – efficient and accurate decision support and risk analysis" co-funded by EMD International A/S, Aalborg University and the Innovationfund Denmark case number 5189-00022B. Their financial support is highly appreciated.

References

1. IEC. International Standard IEC 61400-1 ed. 4, "Wind Turbines - Part 1 Design Requirements". 2019.
2. ISO. International Standard ISO 2394:2015, "General principles on reliability for structures." 2015.
3. Slot RMM, Svenningsen L, Sørensen JD, Thøgersen ML. Importance of Shear in Site Assessment of Wind Turbine Fatigue Loads. *J Sol Energ.* 2018;140(4):041012. doi:10.1115/1.4039748
4. Stensgaard Toft H, Svenningsen L, Moser W, Dalsgaard Sørensen J, Lybech Thøgersen M. Wind Climate Parameters for Wind Turbine Fatigue Load Assessment. *J Sol Energ.* 2016;138(3). doi:10.1115/1.4033111
5. IEC. International Standard IEC 61400-1 ed. 3, "Wind Turbines - Part 1 Design Requirements". 2010.
6. Graf PA, Stewart G, Lackner M, Dykes K, Veers P. High-throughput computation and the applicability of Monte Carlo integration in fatigue load estimation of floating offshore wind turbines. *Wind Energy*. 2016;19:861-872. doi:110.1002/we.1870
7. Dimitrov N, Kelly M, Vignaroli A, Berg J. From wind to loads: wind turbine site-specific load estimation using databases with high-fidelity load simulations. *Wind Energy Sci*. 2018;3:767-790. doi:10.5194/wes-2018-18
8. Toft HS, Svenningsen L, Sørensen JD, Moser W, Thøgersen ML. Uncertainty in wind climate parameters and their influence on wind turbine fatigue loads. *Renew Energy*. 2016;90:352-361. doi:10.1016/j.renene.2016.01.010
9. Morató A, Sriramula S, Krishnan N. Kriging models for aero-elastic simulations and reliability analysis of offshore wind turbine support structures. *Ships Offshore Struct*. 2018;0(0):1-14. doi:10.1080/17445302.2018.1522738
10. Teixeira R, O'Connor A, Nogal M, Krishnan N, Nichols J. Analysis of the design of experiments of offshore wind turbine fatigue reliability design with Kriging surfaces. *Procedia Struct Integr*. 2017;5:951-958. doi:10.1016/j.prostr.2017.07.132
11. Murcia JP, Réthoré P, Dimitrov N, et al. Uncertainty propagation through an aeroelastic wind turbine model using polynomial surrogates. *Renew Energy*. 2017;119:910-922. doi:10.1016/j.renene.2017.07.070

12. Hu W, Choi KK, Cho H. Reliability-based design optimization of wind turbine blades for fatigue life under dynamic wind load uncertainty. *Struct Multidiscip Optim*. 2016;54(4):953-970. doi:10.1007/s00158-016-1462-x

13. Dimitrov N, Natarajan A, Kelly M. Model of wind shear conditional on turbulence and its impact on wind turbine loads. *Wind Energy*. 2015;18(11):1917-1931. doi:10.1002/we.1797

14. Svenningsen L, Slot RMM, Thøgersen ML. A novel method to quantify atmospheric stability. *J Phys Conf Ser*. 2018;1102. doi:10.1088/1742-6596/1102/1/012009

15. Kelly M, Larsen G, Dimitrov NK, Natarajan A. Probabilistic Meteorological Characterization for Turbine Loads. *J Phys Conf Ser*. 2014;524:012076. doi:10.1088/1742-6596/524/1/012076

16. Miner MA. Cumulative damage in fatigue. *J Appl Mech*. 1945;12:159-164.

17. Jonkman JM, Butterfield S, Musial W, Scott G. *Definition of a 5-MW Reference Wind Turbine for Offshore System Development*. National Renewable Energy Laboratory; 2009.

18. Jonkman J. FAST An aeroelastic computer-aided engineering (CAE) tool for horizontal axis wind turbines. https://nwtc.nrel.gov/FAST [Accessed 2019-06-18]. Published 2015. Accessed April 3, 2017.

19. B. Jonkman NK. TurbSim A stochastic, full-field, turbulence simulator primarialy for use with InflowWind/AeroDyn-based simulation tools. https://nwtc.nrel.gov/TurbSim. Published 2016. Accessed April 3, 2017.

20. Standard. IEC 61400-1 ed. 3, 2005, International Electrotechnical Commission, Wind turbines, Part 1: Design requirements, Edition 3 (2005) incl. Amendment 1 (2010). 2010.

21. ASTM. ASTM No. E1049-85, “Standard Practice for Cycle Counting in Fatigue Analysis”. 2011.

22. Sørensen JD, Frandsen S, Tarp-Johansen NJ. Effective turbulence models and fatigue reliability in wind farms. *Probabilistic Eng Mech*. 2008;23(4):531-538. doi:10.1016/j.probengmech.2008.01.009

23. IEC 61400-1. Safety Factors - IEC 61400-1 ed. 4 - background document. 2014.

24. Marelli S, Sudret B. UQLab: A Framework for Uncertainty Quantification in Matlab. In: *Vulnerability, Uncertainty, and Risk*. Reston, VA: American Society of Civil Engineers; 2014:2554-2563. doi:10.1061/9780784413609.257

25. Santner TJ, Williams BJ, Notz WI. *The Design and Analysis of Computer Experiments*. New York, NY: Springer New York; 2018. doi:10.1007/978-1-4939-8847-1

26. Sudret B. Polynomial chaos expansions and stochastic finite element methods. In: Phoon KK, Ching J, eds. *Risk and Reliability in Geotechnical Engineering*. Taylor and Francis; 2015:265-300.

27. Kocis L, Whiten WJ. Computational investigations of low-discrepancy sequences. *ACM Trans Math Softw*. 1997;23(2):266-294. doi:10.1145/264029.264064

28. Xiu D, Karniadakis GE. The Wiener--Askey Polynomial Chaos for Stochastic Differential Equations. *SIAM J Sci Comput*. 2002;24(2):619-644. doi:10.1137/S1064827501387826

29. Rosenblatt M. Remarks on a Multivariate Transformation. *Ann Math Stat*. 1952;23(3):470-472.

30. Blatman G, Sudret B. Adaptive sparse polynomial chaos expansion based on least angle regression. *J Comput Phys*. 2011;230:2345-2367. doi:10.1016/j.jcp.2010.12.021

31. Marelli S, Sudret B. *UQLab User Manual - Polynomial Chaos Expansions*. ETH Zurich: Report UQLab-V1.0-105; 2017.

32. Lataniotis C, Marelli S, Sudret B. The Gaussian process modelling module in UQLab. September 2017.

33. CEN. EN 1990, Eurocode - Basis of structural design. 2002.

34. Morokoff WJ, Caflisch RE. Quasi-Monte Carlo Integration. *J Comput Phys*. 1995:218-230.

35. Schöbi R, Sudret B, Wiart J. Polynomial-chaos-based Kriging. *Int J Uncertain Quantif*. 2015;5(2):171-193.

36. Marelli S, Sudret B. *UQLab User Manual - PC-Kriging*. ETH Zurich: Report UQLab-V1.0-105; 2017.

37. Andrianakis I, Challenor PG. The effect of the nugget on Gaussian process emulators of computer models. *Comput Stat Data Anal*. 2012;56(12):4215-4228. doi:10.1016/j.csda.2012.04.020