



**HAL**  
open science

## Efficient techniques for fast uncertainty propagation in an offshore wind turbine multi-physics simulation tool

E Fekhari, B Iooss, V Chabridon, J Muré

► **To cite this version:**

E Fekhari, B Iooss, V Chabridon, J Muré. Efficient techniques for fast uncertainty propagation in an offshore wind turbine multi-physics simulation tool. 5th International Conference on Renewable Energies Offshore, Nov 2022, Lisbon, Portugal. hal-03757979

**HAL Id: hal-03757979**

**<https://hal.science/hal-03757979>**

Submitted on 22 Aug 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Efficient techniques for fast uncertainty propagation in an offshore wind turbine multi-physics simulation tool

E. Fekhari<sup>1,2</sup>, B. Iooss<sup>1,2</sup>, V. Chabridon<sup>1</sup>, J. Muré<sup>1</sup>

<sup>1</sup>EDF R&D, 6 Quai Watier, Chatou, 78401, France

<sup>2</sup>Université Nice Côte d'Azur, 28 Avenue de Valrose, Nice, 06103, France

August 22, 2022

## Abstract

Offshore wind turbines (OWT) are subject to uncertain environmental conditions (wind and tidal), making long-term investment decisions riskier. To study the impact of the environmental uncertainties on output variables of interest, we propagate them through costly multi-physics numerical models, simulating the OWT. Two main frameworks are possible to retrieve the random variables of interest. First, an efficient sampling method can improve the Monte Carlo reference convergence rate. Second, a regression model can be fitted over a few samples before using it as inexpensive approximation of the numerical model. Some advanced methods offer a sequential improvement of this strategy by iteratively adding samples enhancing the regression model. In this work, our aim is to perform a numerical comparison between various propagation methods to estimate the expected value of the mechanical loads of an OWT over environmental random variables. Additionally, theoretical equivalences between Bayesian quadrature and Kernel herding using maximum mean discrepancy shall be verified on an industrial use-case.

## 1 Introduction

Offshore wind turbine (OWT) new technologies tend to reach for more difficult and uncertain environmental conditions. This industry needs probabilistic tools to manage risks associated with OWT operation and maintenance. For this paper, the OWT behavior is computed by a costly multi-physics numerical simulation code developed by EDF R&D (Milano et al. 2019) and deployed on a high-performance computer facility.

To propagate the various sources of uncertainties through such numerical models, sampling methods such as low-discrepancy sequences (e.g., Sobol' sequences) were proven to improve the Monte Carlo reference convergence rate. An alternative strategy is to emulate the costly function by a regression model. For instance, using a Gaussian process regression model provides an estimation of the regression error represented by the variance of the Gaussian process conditioned to the learning sample. This property is extensively exploited by adaptive methods for optimization, rare event, and quadrature estimation to iteratively pick samples with respect to a specific goal.

Uncertainty propagation techniques have been widely applied to OWT simulation codes to study the mechanical fatigue damage of the structure in a central tendency study (Müller & Cheng 2018, Van den Bos 2020), a sensitivity analysis (Velarde et al. 2019), a static reliability analysis (Zwick & Muskulus 2015, Teixeira, O'Connor, Nogal, Nandakumar, & Nichols 2017, Huchet 2019, Slot, Sørensen, Sudret, Svenningsen, & Thøgersen 2020, Wilkie & Galasso 2021) or a time-dependent reliability analysis (Abdallah, Lataniotis, & Sudret 2019, Lataniotis 2019). Since the environmental conditions are one of the main sources of uncertainty, properly quantifying the site-specific environmental joint distribution is essential to ensure a relevant uncertainty propagation (Dimitrov et al. 2018). Dealing with a sequence of numerical approximations, specific approaches can be used to refine the mechanical model in a smart

Table 1: List of abbreviations.

BQ	Bayesian quadrature
DoE	design of experiments
GP	Gaussian process
KH	Kernel herding
MMD	maximum mean discrepancy
OWT	offshore wind turbine
QoI	quantity of interest
RKHS	reproducing kernel Hilbert space
SP	Support points

way (Mell et al. 2021). Probabilistic designs of OWT also show a large industrial interest applying the so-called reliability-based design optimization method (Cousin 2021, Stieng & Muskulus 2020). As a remark, note that, from the OWT asset management viewpoint, another important output variable of interest might be the annual energy production.

This work presents a numerical comparison of different methods to estimate the expected value of the mechanical loads of an OWT over environmental random variables. Then, theoretical equivalences between Bayesian Quadrature (Briol et al. 2019) and kernel-based sampling methods are studied.

This paper will first give an overview of the industrial problem; then introduce kernel-based sampling methods and adaptive sampling methods for central tendency estimation; before presenting a numerical study comparing their performance on two analytical toy-cases and one industrial use-case.

## 2 Wind turbine fatigue damage estimation

The first variable of interest studied is the damage, a physical measure of the effects of material fatigue accumulated during the OWT operation. Resulting from a large number of stress cycles of magnitude smaller than the elastic limit, it can lead to a material fracture and crack growth. Standards give recommendations for OWT design regarding the fatigue phenomenon (DNV-GL 2016a), also called “fatigue limit states”. More precisely, the damage studied is a scalar quantity, computed by post-processing the output of a 10-minutes simulation using the Rainflow Counting method associated with Miner’s rule (Dowling 1972). The 10-minutes damage is then extrapolated to the OWT’s life span using a simple cross-multiplication rule.

### 2.1 Wind turbine numerical simulation model

EDF R&D has developed a computer simulation chain composed of two simulation tools: the first one, called TurbSim (developed by NREL (Jonkman 2009)) is a stochastic wind generator; the second one, called DIEGO (for “Dynamique Intégrée des Éoliennes et Génératrices Offshore”, developed by EDF R&D (Milano et al. 2019)), is a multi-physics code describing the hydro-aero-servo-elastic behavior of an OWT. As the wind generator is a stochastic code (i.e., two realizations from the same input sample give two different output values), the standards suggest repeating each simulation a finite number of times and averaging the realizations (Slot et al. 2020). For the sake of simplicity, our study will first consider the numerical simulation code to be deterministic by fixing the pseudo-random seed. Since  $g$  is composed of two chained models, the input variables can be divided into two independent groups: environmental random vector  $\mathbf{X} \in \mathcal{D}_{\mathbf{X}} \subset \mathbb{R}^p$  and system random vector  $\mathbf{Z} \in \mathcal{D}_{\mathbf{Z}} \subset \mathbb{R}^q$ . The block diagram in Figure 1 describes the chain of numerical models implicated to assess the damage on the OWT structure (standards recommend to simulate over a period of 10 minutes).

Practically, the OWT modeled is an operating bottom fixed 2.3MW from an EDF wind field. This work is part of the HIPERWIND European projet for which this OWT model leads to successful results between three independent simulation codes (Capaldo et al. 2021). Obviously, the uncertainty quantifi-

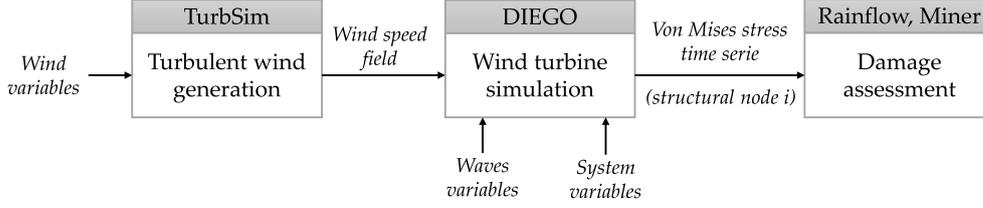


Figure 1: Chained simulation model for OWT damage assessment.

Table 2: Marginals of the environmental joint distribution.

Mean wind speed	$U$	Weibull	10-min. average horizontal wind speed at 10m
Turbulence	$\sigma_s$	Log-normal	10-min. standard deviation of the wind speed
Significant wave height	$H_s$	Weibull	Significant wave height per hour
Peak wave period	$T_p$	Log-normal	Peak 1-hour spectral wave period
Wind-wave misalignment	$\beta$	von Mises	Difference between wind and wave directions

cation of the input variables is also a critical step. As this model is somehow costly to evaluate (about 15 minutes per run), for a given simulation budget, the uncertainty propagation quality relies on the selected design of numerical experiments to represent the input uncertainty.

## 2.2 Input variables uncertainty quantification

The inputs are grouped into two random vectors  $\mathbf{X}$  and  $\mathbf{Z}$  considered independent, respectively with the joint probability density functions  $f_{\mathbf{X}}(\cdot)$  and  $f_{\mathbf{Z}}(\cdot)$ . In order to quantify this uncertainty, onsite environmental data over several years of operation is available. This uncertainty quantification step should be as much as possible site-specific. However, the proximity to the shore of this site makes the definition of the dependency structure within the environmental random vector  $\mathbf{X}$  challenging. Ultimately, using the marginal distributions described in Table 2, one should be able to define a parametric model of the joint environmental distribution as a product of conditional distributions:

$$f_{U, \sigma_s, H_s, T_p, \beta}(U, \sigma_s, H_s, T_p, \beta) = f_U(U) f_{\sigma_s|U}(\sigma_s|U) f_{H_s|U}(H_s|U) f_{T_p|H_s}(T_p|H_s) f_{\beta|U}(\beta|U). \quad (1)$$

To avoid an intricate parametric fit of the environmental random vector, this paper will explore sampling methods based on empirical distributions (e.g., a large i.i.d. sample coming from measured data). Furthermore, various system parameters may also present uncertainties. When studying the damage output variable, most sensitivity analysis studies (Huebler et al. 2017) outline the soil properties over material properties. However, according to Drexler & Muskulus (2021), fatigue computation parameters (such as the Wöhler curve parameters) are also important.

## 2.3 Output quantity of interest

The damage at specific nodes of the structural mesh (over the intended OWT life span) is our variable of interest. On this random variable, many statistics can be estimated such as moments, quantiles, and threshold exceedance probabilities. The standards suggest averaging this damage over the environmental conditions leading to the so-called “global mean damage”. Since the two random vectors  $\mathbf{X}$  and  $\mathbf{Z}$  are considered as independent, the transfer theorem can be written over  $\mathbf{X}$  conditionally to  $\mathbf{Z}$ :

$$\begin{aligned} \mathbb{E}[Y|\mathbf{Z} = \mathbf{z}] &= \mathbb{E}[g(\mathbf{X}, \mathbf{Z})|\mathbf{Z} = \mathbf{z}] \\ &= \int_{\mathcal{D}_{\mathbf{X}}} g(\mathbf{x}, \mathbf{z}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \phi(\mathbf{z}). \end{aligned} \quad (2)$$

The standards recommend estimating the global mean damage by computing realizations of the damage for a set of inputs discretized over a regular grid (see, e.g., the design load cases defined in Section 4.4 in (DNV-GL 2016b)). Yet, regular grids are known to provide poor probabilistic design of experiments for any input dimension higher than one. Once this quantity is estimated over the domain  $\mathcal{D}_{\mathbf{X}}$ , one could define a reliability analysis problem, which consists of estimating the exceedance probability of a given threshold over the domain  $\mathcal{D}_{\mathbf{Z}}$ .

### 3 Sampling methods for design of experiments and uncertainty propagation

Let us denote by  $\mathbf{X}_n = \{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}\} \in \mathcal{D}_{\mathbf{X}} \subset \mathbb{R}^p$  the  $n$ -sample of input realizations, also called input “design of experiments” (DoE) or simply design. Considering a costly function  $g : \mathcal{D}_{\mathbf{X}} \rightarrow \mathbb{R}$ , a first goal of the following DoE methods is to explore the input space in a space-filling way (Franco 2008) (e.g., to build a generic metamodel of  $g$ ). However, this work will exploit these methods with a specific purpose in mind: to numerically integrate  $g$  against the probability density function  $f_{\mathbf{X}}$  which relates to a central tendency estimation of an output random variable  $Y = g(\mathbf{X})$ , resulting from an uncertainty propagation step.

In the recent literature, (Van den Bos 2020) applies a first family of numerical integration methods based on tensor products of quadrature rules to a similar industrial OWT use case. These quadrature rules directly generate a set of pairs of nodes (i.e., points) and associated weights. The weighted sum of the observations at the nodes gives an approximation of the integral. Unfortunately, some of these techniques are limited to inputs without any strong dependency structure and will not be studied in this paper.

Alternatively, sampling methods rely on generating a set of points and approximating the integral by the arithmetic mean of the output realizations (uniform weights). Among them, low-discrepancy sequences (e.g., Sobol’ or Halton sequences) are known to improve the reference Monte Carlo convergence rate and will be used as a deterministic reference method in the upcoming numerical experiments.

Recently, other sampling methods based on the notions of discrepancy between distributions in a kernel-based functional space were used to approximate integrals Pronzato & Zhigljavsky (2020). More precisely, one can mention the use of the distance called the *maximum mean discrepancy* (MMD) as a core ingredient of advanced sampling methods such as the *Support points* by Mak & Joseph (2018) and the *Kernel herding* by Chen et al. (2010). The MMD is convenient to manipulate since it can simply be expressed using the underlying kernel arbitrarily chosen (note that it can be also used for an advanced sensitivity analysis such as shown in Da Veiga (2015)). Let us setup the introduction of the Kernel herding and Support points methods by briefly defining a few mathematical concepts.

**Reproducing kernel Hilbert space.** Assuming that  $k$  is a symmetric and positive definite function  $k : \mathcal{D}_{\mathbf{X}} \times \mathcal{D}_{\mathbf{X}} \rightarrow \mathbb{R}$ , latter called a “reproducing kernel” or simply a “kernel”. A *reproducing kernel Hilbert space* (RKHS) is an inner product space  $\mathcal{H}(k)$  of functions  $g : \mathcal{D}_{\mathbf{X}} \rightarrow \mathbb{R}$  with the following properties:

- $\langle g, k(\cdot, \mathbf{x}) \rangle_{\mathcal{H}(k)} = g(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{D}_{\mathbf{X}}, \forall g \in \mathcal{H}(k)$
- $k(\cdot, \mathbf{x}) \in \mathcal{H}(k), \quad \forall \mathbf{x} \in \mathcal{D}_{\mathbf{X}}$ .

Notice that for a defined reproducing kernel, a unique RKHS exists and vice versa.

**Potential.** For any target distribution  $\mu$ , its *potential* (also called “kernel mean embedding”) associated with the kernel  $k$  is defined as:

$$P_{\mu}(\mathbf{x}) := \int_{\mathcal{D}_{\mathbf{X}}} k(\mathbf{x}, \mathbf{x}') d\mu(\mathbf{x}'). \quad (3)$$

Let us also define the potential of a discrete distribution  $\zeta_n = \sum_{i=1}^n w_i \delta(\mathbf{x}^{(i)})$ ,  $w_i \in \mathbb{R}$  (weighted sum of Dirac distributions at the design points  $\mathbf{X}_n$ ) associated with the kernel  $k$  as:

$$P_{\zeta_n}(\mathbf{x}) = \sum_{i=1}^n w_i k(\mathbf{x}, \mathbf{x}^{(i)}). \quad (4)$$

A possible interpretation is that the closer these potentials gets, the more the design  $\mathbf{X}_n$  brings in average a relevant information to represent the target distribution  $\mu$  (and therefore to compute a quantity such as Eq. (2)). The following scalar product can be used to bound the quadrature error committed by using the design  $\mathbf{X}_n$  by writing the Cauchy-Schwarz inequality (see in Briol et al. (2019)):

$$\sum_{i=1}^n w_i g(\mathbf{x}^{(i)}) - \int_{\mathcal{D}_{\mathbf{x}}} g(\mathbf{x}) d\mu(\mathbf{x}) = \langle g, (P_{\zeta_n}(\mathbf{x}) - P_{\mu}(\mathbf{x})) \rangle_{\mathcal{H}(k)}. \quad (5)$$

**Maximum mean discrepancy.** A recent metric of discrepancy and quadrature error is offered by the MMD. This distance between two distributions  $\mu$  and  $\zeta$  is given by the maximal quadrature error committed for any function within a defined RKHS:

$$\text{MMD}_k(\mu, \zeta) := \sup_{\|g\|_{\mathcal{H}(k)} \leq 1} \left| \int_{\mathcal{D}_{\mathbf{x}}} g(\mathbf{x}) d\mu(\mathbf{x}) - \int_{\mathcal{D}_{\mathbf{x}}} g(\mathbf{x}) d\zeta(\mathbf{x}) \right|. \quad (6)$$

Using the Cauchy-Schwartz inequality, one can demonstrate that  $\text{MMD}_k(\mu, \zeta) = \|P_{\mu}(\mathbf{x}) - P_{\zeta}(\mathbf{x})\|_{\mathcal{H}(k)}$ . Moreover, kernels are called characteristic when the following implication is true,  $\text{MMD}_k(\mu, \zeta) = 0 \Rightarrow \mu = \zeta$ .

### 3.1 Kernel herding

In this section we introduce the Kernel herding (KH) (Chen et al. 2010), a sampling method which intends to minimize a squared MMD by adding points iteratively. Considering a design  $\mathbf{X}_n$  and its corresponding discrete distribution  $\zeta_n = \frac{1}{n} \sum_{i=1}^n \delta(\mathbf{x}^{(i)})$ , a KH iteration can be written as an optimization over the point  $\mathbf{x}^{(n+1)} \in \mathcal{D}_{\mathbf{x}}$  of the following criterion:

$$\mathbf{x}^{(n+1)} \in \arg \min_{\mathbf{x} \in \mathcal{S}} (P_{\zeta_n}(\mathbf{x}) - P_{\mu}(\mathbf{x})) \quad (7)$$

considering a kernel  $k$  and a given set  $\mathcal{S} \subseteq \mathcal{D}_{\mathbf{x}}$  of candidate points (e.g., a fairly dense finite subset with size  $N \gg n$  that emulates the target distribution). This compact criterion derives from the expression of a descent algorithm with respect to  $\mathbf{x}_{n+1}$  (see Pronzato & Rendas (2021) for the full proof).

In practice,  $P_{\mu}(\mathbf{x})$  can be expressed analytically in the specific cases of input distribution and kernel (e.g., for independent uniform or normal inputs and a Matérn 5/2 kernel (Fekhari et al. 2022)), making the computation very fast. Alternatively, the potential can be evaluated on an empirical measure  $\mu_N$ , substituting  $\mu$ , formed by a dense and large-size sample of  $\mu$  (e.g., the candidate set  $\mathcal{S}$ ).  $P_{\mu}(\mathbf{x})$  is then approached by  $P_{\mu_N}(\mathbf{x}) = (1/N) \sum_{j=1}^N k(\mathbf{x}, \mathbf{x}'^{(j)})$ , which can be injected in Eq. (7) to solve the following optimization:

$$\mathbf{x}^{(n+1)} \in \arg \min_{\mathbf{x} \in \mathcal{S}} \left( \frac{1}{n} \sum_{i=1}^n k(\mathbf{x}, \mathbf{x}^{(i)}) - \frac{1}{N} \sum_{j=1}^N k(\mathbf{x}, \mathbf{x}'^{(j)}) \right). \quad (8)$$

When no observation is available, which is the common situation at the design stage, the kernel hyperparameters (e.g., correlation lengths) have to be set to heuristic values. MMD minimization is quite versatile and was explored in more details by Teymur et al. (2021) or Pronzato (2022), however the method is very sensitive to the kernel chosen and its tuning. Support points is a closely related method with a more rigid mathematical structure but interesting performances.

### 3.2 Greedy support points

Support points (SP) Mak & Joseph (2018) are such that their associated empirical distribution  $\zeta_n$  has minimum energy distance with respect to a target distribution  $\mu$ . This criterion can be seen as a particular case of the MMD for the characteristic “energy-distance” kernel Székely & Rizzo (2013) given by:

$$k_E(\mathbf{x}, \mathbf{x}') = \frac{1}{2} (\|\mathbf{x}\| + \|\mathbf{x}'\| - \|\mathbf{x} - \mathbf{x}'\|). \quad (9)$$

Compared to more heuristic methods for solving quantization problems, Support points benefit from the theoretical guarantees of MMD minimization in terms of convergence of  $\zeta_n$  to  $\mu$  as  $n \rightarrow \infty$ . At first sight, this optimization problem seems intractable, although Mak & Joseph (2018) propose to rewrite the function as a difference of convex functions in  $\mathbf{X}_n$ , which yields a difference-of-convex program. To simplify the algorithm and keep an iterative design, a different approach will be used here. At iteration  $n + 1$ , the algorithm solves greedily the MMD minimization between  $\zeta_n$  and  $\mu$  for the candidate set  $\mathcal{S}$ :

$$\mathbf{x}^{(n+1)} \in \arg \min_{\mathbf{x} \in \mathcal{S}} \left( \frac{1}{N} \sum_{j=1}^N \|\mathbf{x} - \mathbf{x}'^{(j)}\| - \frac{1}{n+1} \sum_{i=1}^n \|\mathbf{x} - \mathbf{x}^{(i)}\| \right). \quad (10)$$

For this criterion, one can notice that it is almost identical to the KH one in Eq. (7) when taking as kernel the energy-distance kernel given in Eq. (9). These two iterative methods were exploited in Fekhri et al. (2022) to study new ways to construct a validation set for machine learning models by conveniently selecting a test set for a better model performance estimation.

So far, the methods previously mentioned generate an input DoE which is used to approximate the integral given in Eq. (2) by an arithmetic mean of the function observation on this DoE (i.e., uniformly weighted observations). One can notice that the DoE construction can be done without using any output observation (i.e., without any call to the function  $g$ ). Other approaches can take advantage of the progressive knowledge acquired sequentially on  $g$  to select the following points in the DoE. These methods are sometimes called “active learning” or “adaptive strategies” and, for a large panel of them, rely on a Gaussian process (or Kriging) metamodel sequentially updated.

## 4 Bayesian quadrature

Before introducing active strategies, let us define the framework around Bayesian Quadrature (BQ) Huszár & Duvenaud (2012). For the sake of clarity, a few supplementary notations are introduced hereafter. The objective is to estimate the following quantity of interest, denoted by  $a_\mu(g)$ , integrating a costly function with scalar output  $g : \mathcal{D}_X \rightarrow \mathbb{R}$  over a probability measure  $\mu$ :

$$a_\mu(g) = \int_{\mathcal{D}_X} g(\mathbf{x}) d\mu(\mathbf{x}). \quad (11)$$

Note that, from a general point of view, this quantity can be different than an expected value by composing  $g$  with another function (e.g., other moments, quantiles, exceedance probabilities).

A common approach is to first approximate the costly function  $g$  by a cheap-to-run metamodel (or surrogate model)  $\xi$  before computing a fine estimation of the quantity. Thus, approximating the true function introduces a metamodeling error, carried out when using the metamodel to estimate the quantity  $a_\mu(g)$ . Let us assume, adopting a Bayesian point of view, that  $\xi$  is a stochastic process describing our uncertain knowledge about the true function  $g$ . Let  $\xi$  be a Gaussian process (GP) prior with mean  $\mathbf{0}$  (to ease the calculation) and covariance kernel  $k$ . The posterior  $\xi_n = (\xi | \mathbf{y}_n) \sim \mathcal{GP}(\eta_n, s_n^2)$  has been conditioned on the observations  $\mathbf{y}_n = [g(\mathbf{x}^{(1)}), \dots, g(\mathbf{x}^{(n)})]^\top$  at the input design  $\mathbf{X}_n$  and is fully defined by the so-called Kriging equations (see, e.g., Dubourg (2011)):

$$\xi_n : \begin{cases} \eta_n(\mathbf{x}) &= \mathbf{k}_n^\top(\mathbf{x}) \mathbf{K}_n^{-1} \mathbf{y}_n \\ s_n^2(\mathbf{x}) &= k_n(\mathbf{x}, \mathbf{x}) - \mathbf{k}_n^\top(\mathbf{x}) \mathbf{K}_n^{-1} \mathbf{k}_n(\mathbf{x}) \end{cases} \quad (12)$$

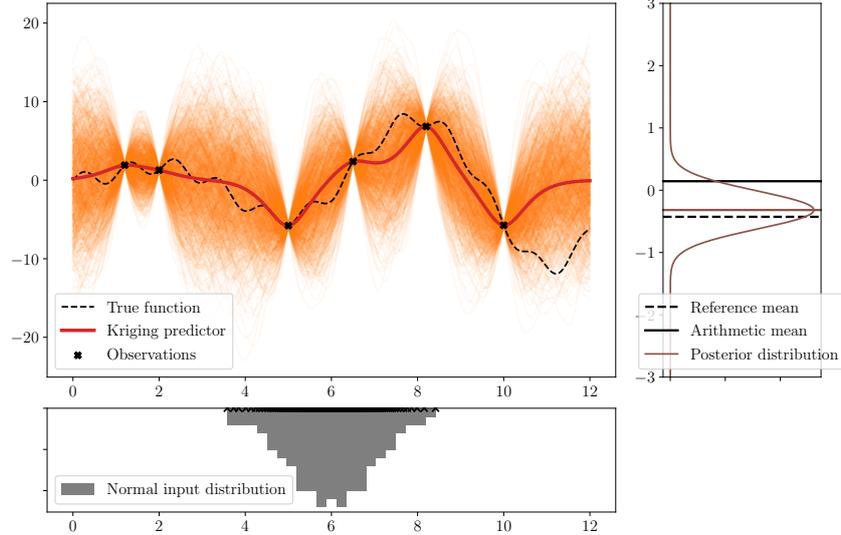


Figure 2: Bayesian quadrature on a one-dimensional case.

where  $\mathbf{k}_n(\mathbf{x})$  designates the column vector gathering the covariance kernel evaluations  $[k_n(\mathbf{x}, \mathbf{x}^{(1)}), \dots, k_n(\mathbf{x}, \mathbf{x}^{(n)})]$  and  $\mathbf{K}_n$  is the  $(n \times n)$  variance-covariance matrix such as the  $(i, j)$ -element is given by  $\{\mathbf{K}_n\}_{i,j} = k_n(\mathbf{x}^{(i)}, \mathbf{x}^{(j)})$ .

Bayesian quadrature, sometimes referred to as “probabilistic integration” (Briol et al. 2019) is a family of numerical integration methods well adapted to costly integrand. It does not only provide an estimator, but a full posterior distribution on the integral value which can be either used to optimize weights on the current observations or to iteratively decide where one needs to add points to a given design. Then, the quantity of interest in Eq. (11) expressed on  $\xi_n$  becomes a random variable (a.k.a., the posterior distribution):

$$a_\mu(\xi_n) = \int_{\mathcal{D}_X} (\xi(\mathbf{x}) | \mathbf{y}_n) d\mu(\mathbf{x}). \quad (13)$$

Figure 2 provides a one-dimensional illustration of the Bayesian quadrature of an unknown function (dashed black curve) against a given input distribution (grey histogram of a normal distribution). For an arbitrary DoE, one can fit a GP model, interpolating the function observations (black crosses). Then, multiple trajectories of this conditioned GP are drawn (orange curves) whilst its mean function, also called “predictor”, is represented by the red curve. Therefore, the input distribution is propagated through the conditioned GP to obtain the posterior distribution represented on the right plot (brown curve). Still on the right plot, remark how the mean of the posterior distribution (brown line) is closer to the reference output expected value (dashed black line) than the arithmetic mean of the observations (black line).

#### 4.1 Optimal weights for quadrature

As exposed in Huszár & Duvenaud (2012), the expected value of the posterior distribution minimizes a Bayes risk with a squared loss (see Bect et al. (2012)).

Working with GP conveniently allows us to express in a closed form the first moments of the posterior distribution (in brown in Figure 2). For instance, its expected value can be written using the expression of the Kriging predictor given in Eq. (12) and the Fubini-Lebesgue theorem:

$$\begin{aligned} \mathbb{E}[a_\mu(\xi_n)] &= \int_{\mathcal{D}_X} \eta_n(\mathbf{x}) d\mu(\mathbf{x}) = \left[ \int_{\mathcal{D}_X} \mathbf{k}_n^\top(\mathbf{x}) d\mu(\mathbf{x}) \right] \mathbf{K}_n^{-1} \mathbf{y}_n \\ &= P_\mu(\mathbf{X}_n) \mathbf{K}_n^{-1} \mathbf{y}_n, \end{aligned} \quad (14)$$

where  $P_\mu(\mathbf{X}_n)$  is given by the row vector of potentials  $[\int k_n(\mathbf{x}, \mathbf{x}^{(1)})d\mu(\mathbf{x}), \dots, \int k_n(\mathbf{x}, \mathbf{x}^{(n)})d\mu(\mathbf{x})]$ . Immediately, the so-called Bayesian quadrature estimator appears to be as a simple linear combination of the observations by taking the row vector  $\mathbf{w}_{\text{BQ}} = P_\mu(\mathbf{X}_n)\mathbf{K}_n^{-1}$ , designated as “optimal weights” for the quadrature.

Another way to recover this result is to minimize the MMD between a target distribution  $\mu$  and a discrete distribution  $\zeta_n$  w.r.t. the weights associated with  $\zeta_n$  (Briol et al. 2019).

As a remark, note that these weights can be computed for any arbitrary kernel  $k$ , without needing any output observation. To illustrate this result, Figures 3 and 4 represent a bivariate random mixture with a nonlinear dependency structure (iso-probability contours of its probability density function in dark grey). On Figure 3 are plotted two samples: first, a regular grid in  $[0, 1]^2$  (red dots); second, a Kernel herding sample of the target distribution (blue crosses). Unsurprisingly, the regular grid is not suited for sampling any non-uniform distribution. On Figure 4, the only difference is that the markers sizes are proportionate to the optimal weights  $\mathbf{w}_{\text{BQ}}$  computed for both samples (for an isotropic Matérn kernel with regularity parameter  $5/2$ ). After all, for the regular grid, the weights of some points are so small that they disappeared. Meanwhile, most of the total weight spreads between six points among twenty-five. This extreme example highlights the limits of regular grids when trying to sample a non-uniform distribution and how the optimal weights aim at enhancing the representativity of any DoE.

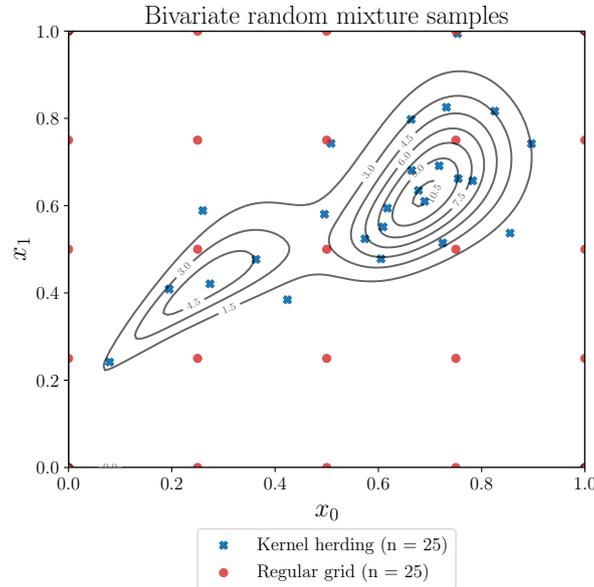


Figure 3: Two samples of a random mixture (grey contours) before applying weights, built by regular grid (red dots) and Kernel herding (blue crosses).

## 4.2 Sequential Bayesian quadrature criterion

An alternative way to building the most relevant input sample for quadrature is to exploit the output observations so as to iteratively select the next design points. Usually, this smart selection procedure relies on a sequentially updated GP conditioned to the observations coupled with a dedicated “acquisition function” (or “learning criterion”) outlining the following point to be added to the design. This section mentions adaptive strategies to estimate the central tendency of an output random variable.

Assuming that  $(\mathbf{X}_n, \mathbf{y}_n)$  is an initial design representative of the input distribution (e.g., generated by a low discrepancy sequence), various criteria can be imagined to select the next point  $\mathbf{x}_{n+1}$  to be added to the design using the conditioned GP  $\xi_n$ . Considering a learning function  $\mathcal{A}$ , each point selection can be done by a greedy optimization on a fairly dense finite sample  $\mathcal{S} \subset \mathcal{D}_{\mathbf{X}}$  with size  $N \gg n$  that represents

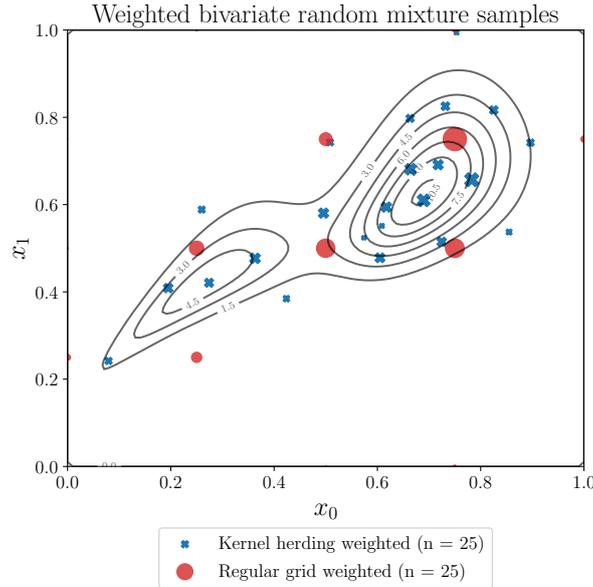


Figure 4: The same samples as Figure 3 with markers' sizes proportional to optimal weights  $\mathbf{w}_{\text{BQ}}$ .

the distribution  $\mu$  as follows:

$$\mathbf{x}_{n+1} \in \arg \max_{\mathbf{x} \in \mathcal{S}} \mathcal{A}(\mathbf{x} | (\mathbf{X}_n, \mathbf{y}_n)). \quad (15)$$

To do so, various learning functions can be considered. Most of them rely on minimizing the posterior variance (Osborne, Garnett, Ghahramani, Duvenaud, Roberts, & Rasmussen 2012, Huchet 2019), which can be addressed as a *Sequential Bayesian quadrature* (SBQ). The posterior variance writes:

$$\text{Var}[a_\mu(\xi_n)] = \int_{\mathcal{D}_X} \int_{\mathcal{D}_X} k_n(\mathbf{x}, \mathbf{x}') d\mu(\mathbf{x}) d\mu(\mathbf{x}'). \quad (16)$$

Ultimately, note that this posterior variance does not depend on the output observations, but only on the location of the design points. Then, this adaptive method offers an iterative design without taking into account any output observation, which can be a desirable property in practice. Therefore, one could then argue that these methods are not properly adaptive, however, updating the kernel's hyper-parameters tailors the GP to the unknown function by using observations. Finally, Huszár & Duvenaud (2012) also show that SBQ is equivalent to applying the optimal weights obtained previously for a Kernel herding design.

## 5 Numerical results

This section presents numerical results computed on two different analytical toy-cases, respectively in dimension 2 (toy-case 1) and dimension 5 (toy-case 2), with easy to evaluate functions  $g(\mathbf{x})$  and associated input distributions. Therefore, we could precisely compute reference values for each toy-case with a large Monte Carlo sample ( $M = 10^7$ ). To estimate the expected values of both toy-cases, designs built by Sobol' sequences, Support points and Kernel herding were used. We compared the performances of each methods for both uniform and optimally weighted estimators. All numerical experiments were computed with the Python package OpenTURNS for uncertainty quantification (Baudin et al. 2017), while the two kernel-based methods were implemented in an open-source Python package named `otkerneldesign`<sup>1</sup>. Finally, the Kernel herding and Support points were applied to the industrial use-case in the last paragraph.

<sup>1</sup><https://github.com/efekhari27/otkerneldesign>

## 5.1 Toy-cases

**Toy-case 1.**  $g_1(\mathbf{x}) = 10 \exp\left(-25 \sum_{i=1}^2 (x_i - 0.5)^2\right)$  also referred to as the ‘‘Gaussian peak function<sup>2</sup>’’, with an input distribution which density is represented by the iso-probability contours in Figure 3.

**Toy-case 2.**  $g_2(\mathbf{x}) = \prod_{i=1}^5 \frac{|4x_i-2|+a_i}{1+a_i}$ ;  $\mathbf{a} = \{1, 2, 3, 4, 5\}$  also referred to as ‘‘G-Sobol<sup>2</sup>’’ function, with a normal input distribution  $\mathcal{N}(\mathbf{0.5}, \mathbf{I}_5)$ .

## 5.2 Results and analysis

The toy-cases results obtained in this section are presented in Figure 5 and Figure 6. Each figure illustrates the expected value estimation for increasing sample sizes (in log scale), knowing that the reference values are represented by the black horizontal lines. The two kernel-based methods are compared to a quasi-Monte Carlo method, widely used for numerical integration (solid lines). Additionally, optimal weights introduced in Section 4.1 are computed for each numerical experiment (dashed lines). Hereafter, the Support points mostly shows better performances whilst Kernel herding surprisingly suffers from this high-dimensional toy-case. Remember that the KH requires to choose a kernel, which offers versatility but reduces its robustness. With a different kernel, KH might perform better but this tuning is hard to set up without prior information. Undoubtedly, the association the optimal weights with any sampling method consequently improves the estimation performances. The optimal weights are particularly relevant for small samples since they have a bigger added value and are easier to compute (the variance-covariance squared matrix can be ill-conditioned and harder to inverse when its size increases).

For the industrial application, the QoI studied is the expected value of the damage (e.g., at the OWT mud-line) against the environmental conditions (e.g., empirically modeled by a large i.i.d. sample coming from measured data). Each realization is computed on a 10 minutes TurbSim-DIEGO simulation then extrapolated to the OWT life span. Figure 7 illustrates the mean damage estimation for increasing samples sizes (up to  $n = 10^3$ ). We take as reference a Monte Carlo sample ( $M = 2000$ ) on which we compute a reference mean (horizontal black line). The convergence of the Monte Carlo mean estimator (in grey) with its corresponding 95% confidence interval (in light grey) is compared with kernel-based methods introduces earlier. Both Support points (in orange) and Kernel herding (in blue) converge faster than Monte Carlo in this case. One can notice that the Monte Carlo confidence interval for small sample sizes is quite large while KH and SP are fully deterministic which guaranties a repeatability of the estimation. If the optimal weights were very efficient on the toy-cases, their effect on this case is negligible. This might be due to the properties of the function or the complex input distribution. In fact, the empirical damage distribution is heavily-tailed, making its mean estimation harder than the toy-cases.

## 6 Conclusion

Wind energy infrastructures are designed to be long-term assets facing various risks. In this work, the proposed approach relies on physical simulation models of wind turbines and their environment to perform the propagation of uncertain inputs. Directly using environmental data as an empirical probabilistic distribution, the first goal is to quickly estimate the expected value of the output fatigue damage in the structure.

This work contributes to solving this issue in several ways: the industrial problem definition and the deployment of a numerical simulation model on a high-performance facility; the study of kernel-based sampling methods and their implementation in a dedicated Python package, called `otkerneldesign`; the illustration of theoretical equivalences between methods of Sections 3 and 4 in analytical toy-cases and an application on the OWT industrial use-case.

Subsequently, with a method for fast central tendency estimation, this work should focus on the next steps of the industrial use-case. Among other ideas, our upcoming work could first continue with

---

<sup>2</sup><http://www.sfu.ca/ssurjano/gaussian.html>

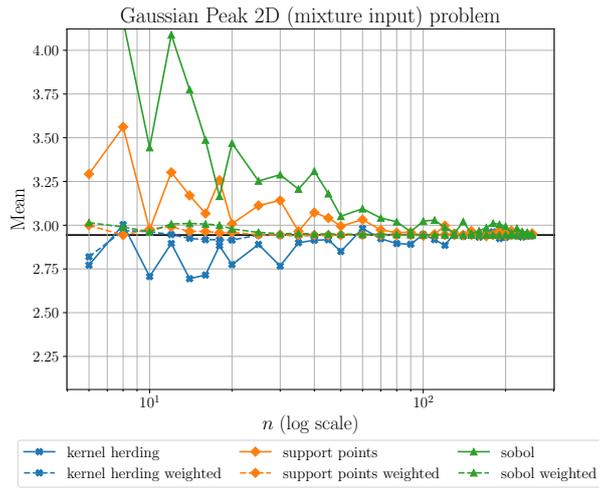


Figure 5: Expected value assessment for the toy-case 1 with Sobol’ sequences, Support points and Kernel herding design.

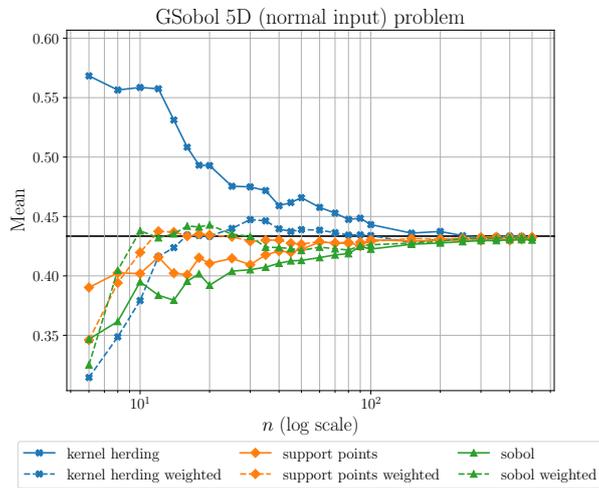


Figure 6: Expected value assessment for the toy-case 2, with Sobol’ sequences, Support points and Kernel herding design.

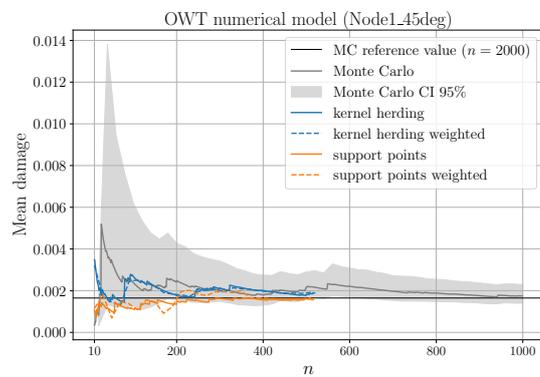


Figure 7: Mean damage estimation at the OWT mud-line, with Monte Carlo, Support points and Kernel herding design.

a reliability analysis of the system together with intending a reliability-oriented sensitivity analysis by adapting recent kernel-based sensitivity indices Marrel & Chabridon (2021) to the sensitivity of a failure probability.

## Acknowledgments.

This study was part of HIPERWIND project which has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement No. 101006689. The authors are grateful to Matteo Capaldo, Anais Lovera for their cooperation on the industrial use-case and Emmanuel Ardillon for fruitful discussions.

## References

- Abdallah, I., C. Lataniotis, & B. Sudret (2019). Parametric hierarchical kriging for multi-fidelity aero-servo-elastic simulators – Application to extreme loads on wind turbines. *Probabilistic Engineering Mechanics* 55, 67 – 77.
- Baudin, M., A. Dutfoy, B. Iooss, & A. Popelin (2017). Open TURNS: An industrial software for uncertainty quantification in simulation. In R. Ghanem, D. Higdon, and H. Owhadi (Eds.), *Springer Handbook on Uncertainty Quantification*, pp. 2001 – 2038. Springer.
- Bect, J., D. Ginsbourger, L. Li, V. Picheny, & E. Vazquez (2012). Sequential design of computer experiments for the estimation of a probability of failure. *Statistics and Computing* 22, 773 – 793.
- Briol, F.-X., C. Oates, M. Girolami, M. Osborne, & D. Sejdinovic (2019). Probabilistic Integration: A Role in Statistical Computation? *Statistical Science* 34, 1 – 22.
- Capaldo, M., M. Guiton, G. Huwart, E. Julian, N. K. Dimitrov, T. Kim, A. Lovera, & C. Peyrard (2021). Design brief of hiperwind offshore wind turbine cases: bottom fixed 10mw and floating 15mw. Technical report, H2020 European project HIPERWIND.
- Chen, Y., M. Welling, & A. Smola (2010). Super-samples from kernel herding. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, pp. 109 – 116. AUAI Press.
- Cousin, A. (2021). *Optimisation sous contraintes probabilistes d’un système complexe : Application au dimensionnement d’une éolienne offshore flottante*. Ph. D. thesis, Institut Polytechnique de Paris.
- Da Veiga, S. (2015). Global sensitivity analysis with dependence measures. *Journal of Statistical Computation and Simulation* 85, 1283 – 1305.
- Dimitrov, N., M. Kelly, A. Vignaroli, & J. Berg (2018). From wind to loads: wind turbine site-specific load estimation with surrogate models trained on high-fidelity load databases. *Wind Energy Science* 3, 767 – 790.
- DNV-GL (2016a). Dnvgl-rp-c203: Fatigue design of offshore steel structures. Technical report, DNVGL.
- DNV-GL (2016b). Dnvgl-st-0437: Loads and site conditions for wind turbines. Technical report, DNVGL.
- Dowling, N. E. (1972). Fatigue Failure Predictions for Complicated Stress-Strain Histories. *Journal of Materials, JMLSA* 7, 71 – 87.
- Drexler, S. & M. Muskulus (2021). Reliability of an offshore wind turbine with an uncertain S-N curve. *Journal of Physics: Conference Series* 2018, 012014.
- Dubourg, V. (2011). *Adaptive surrogate models for reliability analysis and reliability-based design optimization*. Ph. D. thesis, Université Blaise Pascal.
- Fekhari, E., B. Iooss, J. Mure, L. Pronzato, & M. Rendas (2022). Model predictivity assessment: incremental test-set selection and accuracy evaluation. preprint.
- Franco, J. (2008). *Planification d’expériences numériques en phase exploratoire pour la simulation des phénomènes complexes*. Ph. D. thesis, Ecole Nationale Supérieure des Mines de Saint-Etienne.
- Huchet, Q. (2019). *Kriging based methods for the structural damage assessment of offshore wind turbines*. Ph. D. thesis, Université Blaise Pascal.
- Huebler, C., C. Gebhardt, & R. Rolfes (2017). Hierarchical four-step global sensitivity analysis of offshore wind turbines based on aeroelastic time domain simulations. *Renewable Energy* 111, 878–891.
- Huszár, F. & D. Duvenaud (2012). Optimally-Weighted Herding is Bayesian Quadrature. In *Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence*, pp. 377 – 386.
- Jonkman, B. (2009). Turbsim user’s guide: Version 1.50. Technical report.
- Lataniotis, C. (2019). *Data-driven uncertainty quantification for high-dimensional engineering problems*. Ph. D. thesis, ETH Zürich.
- Mak, S. & V. R. Joseph (2018). Support points. *The Annals of Statistics* 46, 2562 – 2592.

- Marrel, A. & V. Chabridon (2021). Statistical developments for target and conditional sensitivity analysis: Application on safety studies for nuclear reactor. *Reliability Engineering & System Safety* 214, 107711.
- Mell, L., V. Rey, & F. Schoefs (2021). A multifidelity approach using discretization error bounds to estimate the probability of failure of structures. In *Proceedings of the 14th World Congress on Computational Mechanics (WCCM) ECCOMAS Congress*.
- Milano, D., C. Peyrard, M. Capaldo, D. Ingram, Q. Xiao, & L. Johanning (2019). Impact of high order wave loads on a 10 mw tension-leg platform floating wind turbine at different tendon inclination angles. In *Proceedings of the ASME 2019 38th International Conference on Ocean, Offshore and Arctic Engineering*.
- Müller, K. & P. W. Cheng (2018). Application of a monte carlo procedure for probabilistic fatigue design of floating offshore wind turbines. *Wind Energy Science* 3, 149 – 162.
- Osborne, M., R. Garnett, Z. Ghahramani, D. K. Duvenaud, S. J. Roberts, & C. Rasmussen (2012). Active learning of model evidence using bayesian quadrature. In *Advances in Neural Information Processing Systems*, Volume 25.
- Pronzato, L. (2022). Performance analysis of greedy algorithms for minimising a Maximum Mean Discrepancy. preprint.
- Pronzato, L. & M. Rendas (2021). Validation design I: construction of validation designs via kernel herding. preprint.
- Pronzato, L. & A. Zhigljavsky (2020). Bayesian quadrature and energy minimization for space-filling design. *SIAM/ASA Journal on Uncertainty Quantification* 8, 959 – 1011.
- Slot, R. M., J. D. Sørensen, B. Sudret, L. Svenningsen, & M. L. Thøgersen (2020). Surrogate model uncertainty in wind turbine reliability assessment. *Renewable Energy* 151, 1150 – 1162.
- Stieng, L. & M. Muskulus (2020). Reliability-based design optimization of offshore wind turbine support structures using analytical sensitivities and factorized uncertainty modeling. *Wind Energy Science* 5, 171–198.
- Székely, G. J. & M. L. Rizzo (2013). Energy statistics: A class of statistics based on distances. *Journal of Statistical Planning and Inference* 143, 1249 – 1272.
- Teixeira, R., A. O'Connor, M. Nogal, K. Nandakumar, & J. Nichols (2017). Analysis of the design of experiments of offshore wind turbine fatigue reliability design with kriging surfaces. *Procedia Structural Integrity* 5, 951 – 958.
- Teymur, O., J. Gorham, M. Riabiz, & C. Oates (2021). Optimal quantisation of probability measures using maximum mean discrepancy. In *International Conference on Artificial Intelligence and Statistics*, pp. 1027 – 1035. preprint.
- Van den Bos, L. (2020). *Quadrature Methods for Wind Turbine Load Calculations*. Ph. D. thesis, Delft University of Technology.
- Velarde, J., C. Kramhøft, & J. Sørensen (2019, 03). Global sensitivity analysis of offshore wind turbine foundation fatigue loads. *Renewable Energy* 140, 177 – 189.
- Wilkie, D. & C. Galasso (2021). Gaussian process regression for fatigue reliability analysis of offshore wind turbines. *Structural Safety* 88, 102020.
- Zwick, D. & M. Muskulus (2015). The simulation error caused by input loading variability in offshore wind turbine structural analysis. *Wind Energy* 18, 1421 – 1432.