Efficient sampling for extreme event statistics of the wave loads on an offshore platform

THE 30th AMERICAN TOWING TANK CONFERENCE
WEST BETHESDA, MARYLAND, OCTOBER 2017

Efficient sampling for extreme event statistics of the wave loads on an offshore platform

Mustafa A. Mohamad, Themistoklis P. Sapsis*

Department of Mechanical Engineering,
Massachusetts Institute of Technology,
77 Massachusetts Ave., Cambridge, MA 02139

September 14, 2017

Abstract

We develop a method for the evaluation of extreme event statistics associated with nonlinear dynamical systems, using a very small number of samples. From an initial dataset of design points, we formulate a sequential strategy that provides the 'next-best' data point (set of parameters) that when evaluated results in improved estimates of the probability density function (pdf) for a scalar quantity of interest. The approach utilizes Gaussian process regression to perform Bayesian inference on the parameter-to-observation map describing the quantity of interest. We then approximate the desired pdf along with uncertainty bounds utilizing the posterior distribution of the inferred map. The 'next-best' design point is sequentially determined through an optimization procedure that selects the point in parameter space that maximally reduces uncertainty between the estimated bounds of the pdf prediction. Since the optimization process utilizes only information from the inferred map it has minimal computational cost. Moreover, the special form of the criterion emphasizes the tails of the pdf. The method is applied to estimate the extreme event statistics for a very high-dimensional system with millions degrees of freedom: an offshore platform subjected to three-dimensional irregular waves. It is demonstrated that the developed approach can accurately determine the extreme event statistics using orders of magnitude smaller number of samples compared with traditional approaches.

Keywords Extreme events; Gaussian processes regression; Sequential experimental design.

For many natural and engineering systems, extreme events, corresponding to large excursions, have significant consequences and are important to predict. Examples include extreme economic events, such as credit shocks [1], rogue waves in the ocean [2], and extreme climate events [3]. Extreme events 'live' in the tails of the probability distribution function (pdf). For most real-world problems, the underlying processes are far too complex to enable estimation of the tails through direct simulations or repeated experiments. This is a result of the low probabilities of extreme events, which necessitates a large number of experiments or ensembles to resolve their statistics. For random dynamical systems with inherently nonlinear dynamics (expressed through intermittent events, nonlinear energy transfers, broad energy spectrum, and large intrinsic dimensionality) we are usually limited to a few ensemble realizations.

The setup in this article involves a stochastic dynamical system that depends on a set of random parameters with known probability distribution. Because of the inherent stochastic and transient character of extreme responses, it is not sufficient to consider the dynamical properties of the system independently from the statistical characteristics of solutions. A statistical approach to this problem has important limitations, such as requiring various extrapolation schemes due to insufficient sample numbers (see extreme value theorems [4]). Another strategy is large deviations theory [5, 6], a method for the probabilistic quantification of large fluctuations in systems, which involves identifying a large deviations principle that explains the least unlikely rare event. While applied to many problems, for complex systems estimating the

rate function can be very costly and the principle does not characterize the full probability distribution. The resulting distributions via such approaches cannot always capture the non-trivial shape of the tail, dictated by physical laws in addition to statistical characteristics. On the other hand, in a dynamical systems approach there are no sufficiently generic methods to infer statistical information from dynamics. For example, the Fokker-Planck equation [7] is formulated for white-noise driven systems, which is nonetheless challenging to solve in moderate-dimensions [8]. To this end, it is essential to consider a blended strategy. The utilization of combined dynamic-stochastic models for the prediction of extreme events have also been advocated and employed in climate science and meteorology by others [9, 10, 11]. In [12, 13] a probabilistic decomposition of extreme events was utilized to efficiently characterize the probability distribution of complex systems, which considered both the statistical characteristics of trajectories and the mechanism triggering the instabilities (extreme events). While effective, the proposed decomposition of intermittent regimes requires explicit knowledge of the dynamics triggering the extremes, which may not be easily identifiable or obtainable for general systems with high complexity.

Here we formulate a sequential quantification method for capturing the statistics of an observable that has the form of a functional of the dynamical system state. The dynamical properties of the system are modeled through a machine learning scheme that infers the quantity of interest (i.e. the observed variable) by utilizing only a few strategically sampled numerical simulations or experiments. Combining these posterior predictions from the machine learning model (via Gaussian process regression) with available statistical information of the random parameters, we formulate an optimization problem that provides the 'next-best' or most informative experiment that should be evaluated to maximally constrain the pdf prediction. The optimization process relies exclusively on the inferred properties of the parameter-to-observation map and no additional simulations are required in the search for the 'next-best' parameter set. The proposed method allows us to sequentially select where to sample in parameter space in order to rapidly capture the pdf and, in particular, the tails of the distribution of the observable.

The use of sequential strategies have been considered in previous works for experimental design selection without considering the possible existence of an underlying dynamical model with random parameters nor the subsequent estimation of the target observables distribution, e.g. [14]; such an approach, while useful in a purely black-box setting does not consider the properties of the random system parameters and can thus lead to infective and uninformative sampling, e.g. experimental parameters with zero probability. The blended approach we propose can therefore more effectively and efficiently determine the desired distribution of the observable by minimizing the number of simulations required.

1 Problem setup

We consider a random, nonlinear dynamical system with state variable $u \in \mathbb{R}^n$ with dynamics

$$\frac{du}{dt} = f(t, u, \theta(\omega)), \quad \omega \in \Omega,$$

where Ω is the sample space in an appropriate probability space (we denote the density of the random variable X by f_X). The random variable $\theta : \Omega \to U \subset \mathbb{R}^m$ parameterizes sources of uncertainty, such as stochastic forcing terms, initial conditions, or system parameters with *a priori* known distribution f_θ. For fixed $\omega \in \Omega$, the response u_θ is a deterministic function in time. We are interested in estimating the pdf $\hat{f}$ of a scalar observable $q \in \mathbb{R}$ given by

$$q = \hat{T}(\theta) \triangleq \mathcal{G}(u_\theta) + \varepsilon$$

where $\hat{T} : \mathbb{R}^m \supset U \to \mathbb{R}$ is the *parameter-to-observation* map of the quantity of interest, $\mathcal{G}$ is an arbitrary functional of u_θ, and ε is some (e.g Gaussian) observational or numerical noise term, which we take as zero, without loss of generality. In our setup the parameter-to-observation-map $\hat{T}$ is *expensive* to compute (e.g. a large scale numerical simulation or a costly physical experiment), so we seek to minimize evaluations of this map. Our objective is to estimate the statistics (especially non-Gaussian features) of the observable q described by the map $\hat{T}$:

Given an observable $q = \hat{T}(\theta)$ with unknown distribution $\hat{f}$, where the distribution of θ is a priori known f_θ and a dataset $\mathcal{D} = \{\hat{\theta}_i, \hat{T}(\hat{\theta}_i)\}_i^n$ of small size n (so that the estimated distribution of q is f_n), find the next best experimental parameter θ_{n+1} (without evaluating $\hat{T}$) such that when this new simulation is evaluated the error between the resulting estimate of the distribution f_{n+1} and $\hat{f}$ is minimized, with particular emphasis on minimizing the error in the tail features of the distribution.

The search for the next best experiment θ_{n+1} should not involve direct evaluation of the true map $\hat{T}$ (expensive) and also cannot involve $\hat{f}$ since this is unknown. The exact statistics of q are given by,

$$\hat{f}(s) = \frac{dF_\theta}{ds} = \frac{d}{ds}\int_{A(s)} f_\theta(\theta)\, d\theta,$$

where $A(s) = \{\theta \in U : \hat{T}(\theta) \leq s\}$ and F_θ is the cumulative distribution function of $\theta \in \mathbb{R}^m$. Our aim is to determine the (non-Gaussian) statistics of $\hat{f}$ through a minimum number of observations of q. The key idea behind the method is the observation that we do not need to densely sample all regions in θ space, since not all regions have significant probability (i.e. f_θ may be negligible) or importance. Specifically, we formulate a sampling strategy that aims to accurately predict the tail region of the pdf by taking into account both the magnitude of the map $|\hat{T}|$ (through a surrogate approximation) *and* the value of f_θ.

In Fig. 1 we illustrate this point graphically: the proposed sampling will not in general be uniform in the whole set U; especially if we emphasize 'extreme magnitude events', where $|\hat{T}|$ is large. The strategy that we formulate is based on the construction of a surrogate for the map $\hat{T}$ using Gaussian process regression. Based on this surrogate we then estimate the pdf for the quantity of interest, as well as the pdf for its bounds. The selection of the 'next-best' point is based on the minimization of this estimation error, in the form of an optimization problem.

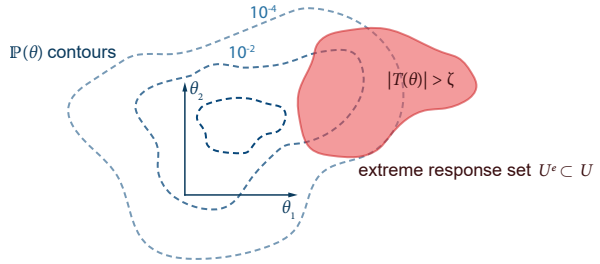

Figure 1 Areas with large probability in θ are not necessarily associated to regions with large $\hat{T}$. The proposed algorithm focuses on sampling regions where the probability of occurrence is important and also takes into account the expected magnitude of $|\hat{T}|$.

2 Method description

An important ingredient of our algorithm is the construction of a surrogate for the map $\hat{T}$. To this end, we utilize Gaussian process regression (GPR) method. An overview of GPR is given in appendix A. An important property of GPR is that it provides a posterior distribution and the variance of the posterior can be used as an error estimate, which will guide the selection of the next point to sample.

We estimate the parameter-to-observation map, $\hat{T}(\theta)\colon U \to \mathbb{R}$, with $U \in \mathbb{R}^m$, via a GPR scheme from an observed dataset $\mathcal{D}_{n-1} = \{\hat{\theta}_i, \hat{T}(\theta_i)\}_{i=1}^{n-1}$. These are the points that we have already processed. To estimate $\hat{T}$ we place a Gaussian process prior over $\hat{T}$ and consider the function values as a realization of a GP. This gives us the posterior mean $T_{n-1}(\theta)$ and variance $\sigma_{n-1}(\theta)$ (see, e.g. [15]).

We then write the sequence of derived distributions based on the GPR surrogate T_{n-1} as,

$$f_{n-1}(s) = \frac{dF_{n-1}}{ds} = \frac{d}{ds}\int_{A_{n-1}(s)} f_\theta(\theta)\, d\theta,$$

where $A_{n-1}(s) = \{\theta \in U : T_{n-1}(\theta) \leq s\}$. F_{n-1} is the corresponding cumulative distribution function and the pdf f_{n-1} is our estimation (the indexed $n-1$, denotes the current number of data points).

Next we formulate the optimization problem that will provide the 'next-best' experiment. This will be based on minimizing the distance between the pdfs of the upper and lower bounds of the surrogate map. Specifically, defining as the upper and lower bounds of the surrogate map through the α-scaled standard deviation of the posterior of the Gaussian process regressor, we have the following sequence of pdfs that represent the map bounds:

$$f^{\pm}_{n-1}(s) = \frac{d}{ds}\int_{A^{\pm}_{n-1}(s)} f_\theta(\theta)\, d\theta, \qquad (1)$$

where $A^{\pm}_{n-1}(s) = \{\theta \in U : T_{n-1}(\theta) \pm \alpha\sigma_{n-1}(\theta) \leq s\}$. Throughout this work we employ the 95% interval bounds, so that the standard deviation is scaled by a factor $\alpha = 1.96$, i.e. $T_n \pm 1.96\sigma_n$. Note again, the surrogate map T_{n-1} and the corresponding standard deviation are based on the points $\mathcal{D}_{n-1} \triangleq \{\hat{\theta}_i, \hat{T}(\hat{\theta}_i)\}_i^{n-1}$.

Denote the point that is sampled next by θ^*. We seek a criterion for the selection of θ^* that does not involve the computation of the exact map $\hat{T}$ at this stage. The exact map value will be computed *after* we have selected the optimal experimental parameter value θ^*. To this end, we assume that the exact value of the map is given by the best linear unbiased estimator, using the $n-1$ points we have already analyzed through the GPR surrogate, i.e. $\hat{T}(\theta^*) \simeq$

$T_{n-1}(\theta^*)$. In addition, the variance of the updated map, now consisting of n points ($n-1$ real points and one hypothetical) will vanish at θ^*. Note that the mean value of the map based on the n points ($n-1$ real points and one hypothetical) is identical to the map based on the $n-1$ real points. However, what differs, is the variance which now vanishes at the additional point θ^*. Based on this construction we have the pdfs for the upper and lower bounds that are based on the $n-1$ real points, $\hat{\theta}_i, i = 1, \ldots, n-1$ and the one hypothetical point θ^*. These are different from the bounds based on the $n-1$ points. We denote these pdfs as $\tilde{f}_n^{\pm}(s;\theta^*)$, which are computed using Eq. 1 with the additional point $(\theta^*, T_{n-1}(\theta^*))$ appended to the dataset $\mathcal{D}_{n-1}$.

The criterion we utilize to select the 'next-best' point is based on the following L_1 distance between the pdfs $\tilde{f}_n^{\pm}(s;\theta^*)$,

$$\hat{Q}_n(\theta^*) \triangleq \frac{1}{2}\int (\log \tilde{f}_n^{+}(s;\theta^*) - \log \tilde{f}_n^{-}(s;\theta^*))\, ds,$$

which we write as $d_{L_1}(\log(\tilde{f}_n^{+}), \log(\tilde{f}_n^{-}))$. The integral is computed over the intersection of the two domains that the pdfs $\tilde{f}_n^{\pm}$ are defined over. The next sample point $\hat{\theta}_n$ is chosen so that Q_n is minimized. In summary, as it is not feasible to compute the criterion $\hat{Q}_n$ based on $f_n^{\pm}$, we instead use $\tilde{f}_n^{\pm}$, whose evaluation only requires the GPR emulator T_{n-1} at the current iteration. This provides a practical and efficient strategy to perform the optimization for the next, near-optimal[1] design point $\hat{\theta}_n$.

The starting design plan size n_s, if not a priori given, should be small and a simple strategy such a Latin hypercube sampling plan may be utilized if it can be freely specified. We also recommend to process a few initial iterations using a d_{L_2} metric in Q to quickly capture the main mass of the probability density function, i.e. low order moments, before utilizing the proposed metric that emphasizes extreme event statistics. In addition, it is not necessary to retrain the GPR hyperparameters after every iteration, which can remain fixed after being calibrated from a few iterations. Updating the Gaussian process emulator after the addition of new data points can be done in $\mathcal{O}(n^2)$ if the hyperparameters are fixed, otherwise the GPR emulator must be performed anew in $\mathcal{O}(n^3)$ [2]. The strategy described above is summarized in pseudocode in appendix B. The dominant cost is the optimization of the objective Q_n, the details of which depend upon the chosen optimizer. In the applications section we utilize a particle swarm global optimizer.

[1] Near optimal in the sense that we are using the best linear predictor of the map.

[2] For low-dimensional θ, since we presume the dataset size is small, the cost difference may be negligible.

3 Application: Hydrodynamic forces and moments on an offshore platform

Here we apply the sampling algorithm to compute the probability distributions describing the loads on an offshore platform in irregular seas. The response of the platform is quantified through direct, three-dimensional numerical simulations of Navier-Stokes utilizing the smoothed particle hydrodynamics (SPH) method [16]. We demonstrate that the developed framework allows for the accurate quantification of the extreme event statistics through a small number of numerical experiments, showing that the proposed scheme can realistically take into account extreme events in the design and optimization processes for systems with this degree of complexity.

Our numerical setup parallels that of a physical wave tank experiment and consists of a wave maker on one end and a sloping 'beach' on the other end of the tank to quickly dissipate the energy of incident waves and avoid wave reflections.

Wind generated ocean waves are empirically described by their energy spectrum. Here, we consider irregular seas with JONSWAP spectral density:

$$S(f) = \frac{\alpha g^2}{(2\pi)^4 f^5} \exp\Big[-\tfrac{5}{4}\big(\tfrac{f_p}{f}\big)^2\Big] \cdot \gamma^{\exp\left[\frac{-(f-f_p)^2}{2\delta^2 f_p^2}\right]}, \quad (2)$$

where $\delta = 0.07$ for $k \leq k_0$ and $\delta = 0.09$ for $k > k_0$ and f_p is the peak frequency. In the original formulation α is related to the fetch and the mean wind speed, however for offshore applications, especially in the North Sea, the following modified version is often adopted [17] $\alpha = 5.058 H_s^2 f_p^4 (1 - 0.287 \log\gamma)$. The random wave field can be described by a superposition of primary wave groups, each characterized by a group length scale L and height A. Following [18] we describe these primary wavegroups by the representation $u(x) = A\,\mathrm{sech}(x/L)$, which is an explicit parameterization in terms of L and A. Thus, L and A correspond to θ_1 and θ_2 in the notation of Eq. 1. The statistical characteristics of the wave groups associated with a random wave field (such as the one given by the JONSWAP spectrum in Eq. 2) can be obtained by applying the scale-selection algorithm described in [19]. Specifically, by generating many realizations

of the spectrum in Eq. 2 we use a group detection algorithm to identify coherent group structures in the field along with their lengthscale and amplitude (L and A). This procedure provides us with the empirical probability distribution f_θ of the wave field and thus a nonlinear parametrization of the randomness in this system.

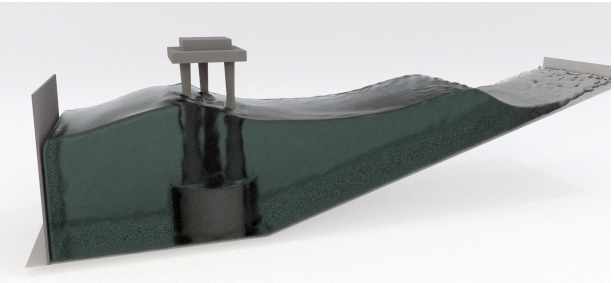

Figure 2 Rendered snapshot of the SPH simulation at $t = 103.5\,\mathrm{s}$ with $\theta_1 = 4.63$ and $\theta_2 = 0.662$ in non-dimensional units.

The quantities of interest in this problem are the forces and moments acting on the platform. The incident waves primary impact is in the x direction and as such we consider the pdf of the force in the x direction F_x and the moment M_y about the bottom-center of the platform:

$$q_f = \max_{t\in[0,T]} |F_x(t)| \quad \text{and} \quad q_m = \max_{t\in[0,T]} |M_y(t)|.$$

We consider a spectrum with parameters $\alpha = 0.060, \gamma = 3.0, H_s = 13.2\,\mathrm{m}$ (in the appendix C we provide an additional case). These parameters are representative of North Sea conditions. The chosen peak wave period is $T_p = 10\,\mathrm{s}$. The forces and moments are normalized by $k_0^3/\rho g$ and $k_0^2/\rho g$, respectively. A snapshot in time from a value at a sampled grid point of the SPH simulation is shown in Fig. 2.

In Fig. 3 we show the results of the algorithms progression. We begin by arbitrary selecting 4 initial sample points from a Latin Hypercube sampling strategy. Next, we perform 4 iterations using the L_1 distance metric to quickly capture the main mass of the distribution before focusing on the distribution away from the mean that utilizes the L_1 metric of the logarithmic of the pdf. The lightly shaded red region in the pdf plots is a visualization of the uncertainty in the pdf, obtained by sampling the GPR prediction and computing the pdf for 200 realizations and then computing the upper (lower) locus of the maximum (minimum) value of the pdf at each value. The figures demonstrate that with 15 (i.e 14 total sample points) iterations (together with the 4 samples in the initial configuration) we are able to approximate the pdf to good agreement with the 'exact' pdf, which was computed from a densely sampled grid. We note that the sampling selects points in the upper right half of the map, where the values of the force becomes large but which also have non-negligible probability of occurrence. We emphasize that for this problem the GPR operates on the logarithm of the observable because the underlying function is always positive.

4 Conclusions

We developed and analyzed a computational algorithm for the evaluation of extreme event statistics associated with nonlinear dynamical systems that depend on a set of random parameters. The algorithm, provides a sequence of points that lead to improved estimates of the probability distribution for a scalar quantity of interest. The criterion for the selection of the next design point emphasizes the tail statistics. We have demonstrated its applicability through a problem involving a demanding system with millions degrees of freedom.

Future work will explore extensions of the proposed method by including multifidelity models. Here, lower resolution but cheaper to compute models that are correlated with the high-fidelity model can be utilized to enhance GPR predictions of the parameter-to-observation-map (see [20]). Incorporating multiple levels of fidelity and utilizing additional samples from cheaper models would lead to improved convergence rates. Such extensions could greatly benefit complex and expensive computer codes such as the three-dimensional hydrodynamic wave loads application we considered.

Acknowledgments

This work has been supported through the ONR grant N00014-15-1-2381, the AFOSR grant FA9550-16- 1-0231, and the ARO grant 66710-EG-YIP. We also thank Alex Crespo for helpful comments regarding the smooth particle hydrodynamic simulations.

A Overview of Gaussian process regression

An important ingredient of our algorithm is the construction of a surrogate for the map $\hat{T}$. To this end here we utilize the Gaussian Process Regression (GPR)

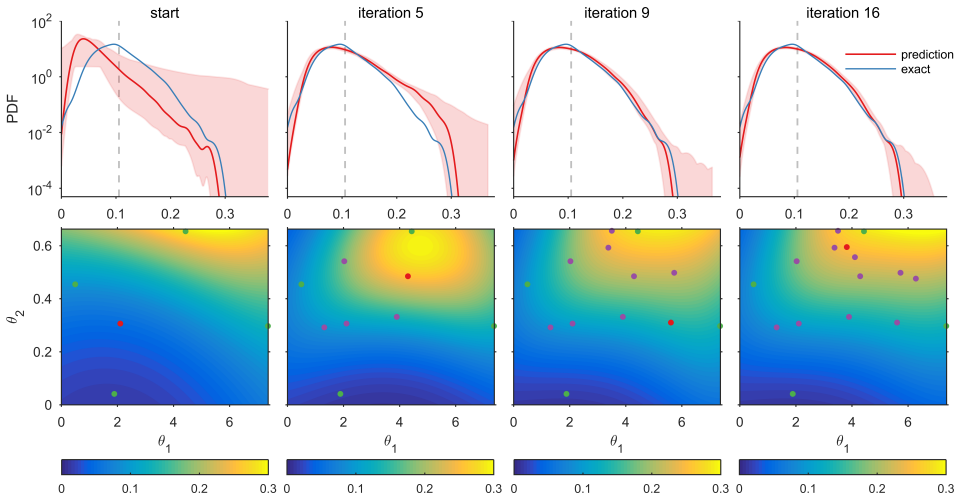

Figure 3 Progression for the force variable. Bottom row show's the map $T(\theta)$. Green points denote the initial configuration (LH sampling), purple points are from the iterative algorithm, and the red point represents the next predicted sample point. Dashed lines denote one standard deviation.

method. A feature of critical importance for GPR is that it provides a posterior distribution and the variance of the posterior can be used as an error estimate, which can in turn be used to guide optimization and explore parameter space. Here, we briefly provide an overview of GPR.

We estimate the parameter-to-observation map, $\hat{T}(\theta)\colon U \to \mathbb{R}$, with $U \in \mathbb{R}^m$, via a GPR scheme from an observed dataset $\mathcal{D}_n = \{\hat{\theta}_i, \hat{T}(\theta_i)\}_{i=1}^n$, using n design points. These are the points that we have already sampled. Specifically, to estimate $\hat{T}$ we place a Gaussian process prior over $\hat{T}$ and consider the function values as a realization of the GP. In particular, with $\Theta = \{\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n\}$, we have the following posterior mean $T_n(\theta)$ and covariance $k_n(\theta, \theta')$ (see, e.g. [15]):

$$T_n(\theta) = \overline{T}(\theta) + k_0(\theta, \Theta)^{\mathsf{T}} k_0(\Theta, \Theta)^{-1} (\hat{T}(\Theta) - \overline{T}(\Theta))$$
$$k_n(\theta, \theta') = k_0(\theta, \theta') - k_0(\theta, \Theta)^{\mathsf{T}} k_0(\Theta, \Theta)^{-1} k_0(\Theta, \theta')$$

where,

- $\overline{T}(\theta)$ is an arbitrary regression mean function, often chosen to be a constant or zero,
- $k_0(\theta, \theta') = \sigma^2 \exp\left(-\frac{|\theta - \theta'|}{2\lambda}\right)$ is the regression covariance function with σ and λ being positive parameters,
- $k_0(\Theta, \Theta) \in \mathbb{R}^{n \times n}$ is the covariance matrix, with the ij^{th} entry given by $k_0(\hat{\theta}_i, \hat{\theta}_j)$, and $k_0(\theta, \Theta)$, $\overline{T}(\Theta)$, $\hat{T}(\Theta)$ are n-dimensional vectors with the i^{th} entries given by $k_0(\theta, \hat{\theta}_i)$, $\overline{T}(\hat{\theta}_i)$, $\hat{T}(\hat{\theta}_i)$, respectively, and
- $\sigma_n^2(\theta) = k_n(\theta, \theta)$ denotes the local variance.

There are several important properties to emphasize in the GPR scheme [21]. Firstly, for any choice of the regression function the GPR mean estimate is an interpolant of the exact map at the design points, $T_n(\Theta) = \hat{T}(\Theta)$. Secondly, for $\overline{T}(\theta) = 0$ the GPR scheme is the best linear estimator in the mean square error sense [22].

Since $k_0(\Theta, \Theta)$ is positive definite

$$\sigma_n(\theta) \leq \sigma_{n-1}(\theta) \leq \ldots \leq \sigma_0(\theta), \quad \theta \in U,$$

or in other words, additional data points lead to non-increasing local variance. In addition, at each of the design points the estimated variance vanishes, $\sigma_n(\hat{\theta}_i) = 0$, for $i = 1, \ldots, n$.

B Algorithm pseudocode

Below we summarize the main loop of the sequential algorithm in pseudocode:

input initial design plan $\mathcal{D}_n = \{\hat{\theta}_i, \hat{T}(\hat{\theta}_i)\}_{i=1}^n$ of size $n = n_s$
repeat
  $T_i, \sigma_i \leftarrow$ predict Gaussian process mean and variance
  $f_i, f_i^+, f_i^- \leftarrow$ integration using T_i, σ_i, and f_θ
  $\hat{\theta}^* \leftarrow \arg\min_\theta Q_n(\theta; T_i, \mathcal{D}_i)$
  Append $(\theta^*, \hat{T}(\theta^*))$ to dataset $\mathcal{D}_i$
until desired error level $d(f_i^+, f_i^-) < \epsilon$
return f_i

This procedure provides us an iterative scheme that leads to a series of pdfs of the observable f_i which

converges to the true pdf under appropriate conditions. The computation of the distance measure Q_n is summarized below:

function $Q(\theta^*; T_n, \mathcal{D}_n)$
  Append $(\theta^*, T_n(\theta^*))$ to dataset $\mathcal{D}_n$
  $\tilde{\sigma}_{n+1} \leftarrow$ predict Gaussian process variance
  $\tilde{f}^+_{n+1}, f^-_{n+1} \leftarrow$ integration using T_n, $\tilde{\sigma}_{n+1}$, and f_θ
  return $d_{L_1}(\log(\tilde{f}^+_n), \log(\tilde{f}^-_n))$
end function

C Hydrodynamic forces and moments on an offshore platform

C.1 Numerical experiments

The numerical simulations are performed using the open-source code DualSPHysics [16], which utilizes the smoothed particle hydrodynamics (SPH) framework, a meshless Lagrangian method. DualSPHysics has been validated on numerous test cases for offshore engineering applications, including forces on structures and also wave propagation, see e.g. [24].

A sketch of the numerical domain is provided in Fig. 4. We are interested in analyzing the forces and moments on offshore structures in the deep ocean, where the dispersion relation is given by $\omega_0^2 = gk_0$. We consider waves with peak period 10 s so that the characteristic wavelength is $\lambda_0 = \frac{2\pi}{k_0} = 156\,\text{m}$. In addition, the depth of the wave tank is selected so that $\tanh(k_0 h) = 0.99$, thus the water depth is $h = 62\,\text{m}$. The beach is setup at a 17.5° angle and the length of the horizontal tank dimension, that is excluding the sloping beach, is 100 m. The structure we consider is an offshore gravity platform (Fig. 4, right) and the dimensions of the model are based on prototypical values. In particular, the base width of the platform is 45 m with height 25 m, three columns with base diameter 10 m extend from the bottom platform and narrow to a 4 m width at height 42.5 m. To generate waves, we implemented a hinged-type wave maker utilizing the corresponding Biésel transfer function [25] to relate the wave height to the stroke of the paddle. For the flap-type wave maker the utilized Biésel transfer function is given by

$$\frac{H}{S_0} = \frac{2\sinh(kh)(1-\cosh(kh)+kh\sinh(kh)}{kh(\sinh(kh)\cosh(kh)+kh)},$$

where S_0 is the stroke at the free surface, H is the wave height in the far-field, k is the wavenumber, and h is the water depth.

References

[1] Jean-Pierre Fouque, George Papanicolaou, Ronnie Sircar, and Knut Sølna. *Multiscale Stochastic Volatility for Equity, Interest-Rate and Credit Derivatives*. Cambridge University Press, 2011.

[2] Paul C. Liu. A chronology of freaque wave encounters. 24(1):57–70.

[3] Andreas Hense and Petra Friederichs. *Wind and precipitation extremes in the Earth's atmosphere*, pages 169–187. Springer Berlin Heidelberg.

[4] Mario Nicodemi. *Extreme value statistics*, pages 1066–1072. Springer New York.

[5] S. R. Srinivasa Varadhan. *Large Deviations and Applications*. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, 1984.

[6] Giovanni Dematteis, Tobias Grafke, and Eric Vanden-Eijnden. Rogue waves and large deviations in deep sea.

[7] Kazimierz Sobczyk. *Stochastic Differential Equations*. Mathematics and Its Applications (East European Series). Springer Netherlands.

[8] A. Masud and L. A. Bergman. Solution of the four dimensional Fokker-Planck equation: still a challenge. In *ICOSSAR*, volume 2005, pages 1911–1916.

[9] Christian L. E. Franzke. Extremes in dynamic-stochastic systems. 27(1):012101.

[10] Andrew J. Majda. *Introduction to Turbulent Dynamical Systems in Complex Systems*. Springer International Publishing.

[11] Nan Chen and Andrew J. Majda. Simple stochastic dynamical models capturing the statistical diversity of el niño southern oscillation. 114(7):1468–1473.

[12] Mustafa A. Mohamad, Will Cousins, and Themistoklis P. Sapsis. A probabilistic decomposition-synthesis method for the quantification of rare events due to internal instabilities. 322:288–308.

[13] Mustafa A. Mohamad and Themistoklis P. Sapsis. Probabilistic description of extreme events in intermittently unstable dynamical systems excited by correlated stochastic processes. 3(1):709–736.

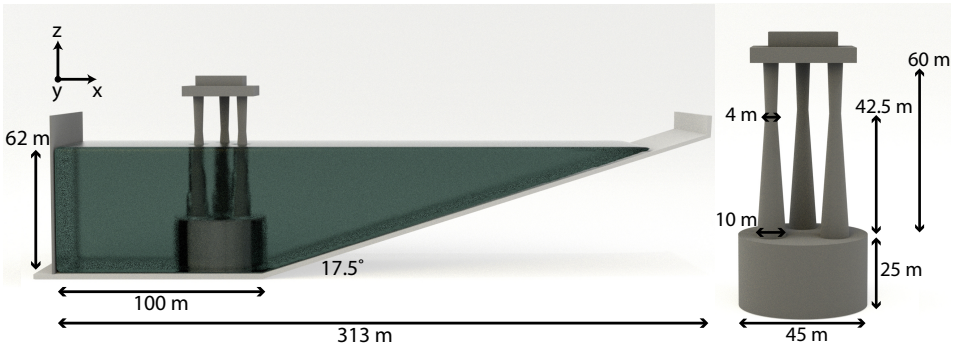

Figure 4 The numerical domain (left) and a close up view of the offshore platform (right).

[14] Loic Le Gratiet and Claire Cannamela. Kriging-based sequential design strategies using fast cross-validation techniques with extensions to multi-fidelity computer codes.

[15] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. MIT Press.

[16] A. J. C. Crespo, J. M. Domnguez, B. D. Rogers, M. Gómez-Gesteira, S. Longshaw, R. Canelas, R. Vacondio, A. Barreiro, and O. Garca-Feal. DualSPHysics: Open-source parallel CFD solver based on smoothed particle hydrodynamics (sph). 187:204–216.

[17] Arvid Naess and Torgeir Moan. *Stochastic Dynamics of Marine Structures*. Cambridge University Press.

[18] Will Cousins and Themistoklis P. Sapsis. Unsteady evolution of localized unidirectional deep-water wave groups. 91:063204.

[19] Will Cousins and Themistoklis P. Sapsis. Reduced-order precursors of rare events in unidirectional nonlinear water waves. 790:368–388.

[20] Marc C. Kennedy and Anthony O'Hagan. Predicting the output from a complex computer code when fast approximations are available. 87(1):1–13.

[21] Andrew M. Stuart and Aretha L. Teckentrup. Posterior consistency for gaussian process approximations of bayesian posterior distributions.

[22] Michael L. Stein. *Interpolation of Spatial Data*. Springer Series in Statistics. Springer-Verlag New York.

[23] Grigorios A. Pavliotis. *Stochastic Processes and Applications*. Number 60 in Texts in Applied Mathematics. Springer-Verlag New York.

[24] Corrado Altomare, Alejandro J. C. Crespo, Jose M. Domnguez, Moncho Gómez-Gesteira, Tomohiro Suzuki, and Toon Verwaest. Applicability of smoothed particle hydrodynamics for estimation of sea wave impact on coastal structures. 96:1–12.

[25] Peter Frigaard, Michael Høgedal, and Morten Christensen. Wave generation theory. Technical report, Hydraulics & Coastal Engineering Laboratory, Department of Civil Engineering, Aalborg University.