Active-Subspace Analysis of Exceedance Probability for Shallow-Water Waves

Šehić, Kenan; Bredmose, Henrik; Sørensen, John D.; Karamehmedović, Mirza

Published in:
Journal of Engineering Mathematics

Link to article, DOI:
10.1007/s10665-020-10080-5

Publication date:
2021

Document Version
Peer reviewed version

Noname manuscript No.
(will be inserted by the editor)

Active-Subspace Analysis of Exceedance Probability for Shallow-Water Waves

Kenan Šehić · Henrik Bredmose · John D. Sørensen · Mirza Karamehmedović

Received: DD Month YEAR / Accepted: DD Month YEAR

Abstract We model shallow-water waves using a one-dimensional Korteweg-de Vries equation with the wave generation parameterized by random wave amplitudes for a predefined sea state. These wave amplitudes define the high-dimensional stochastic input vector for which we estimate the short-term wave crest exceedance probability at a reference point. For this high-dimensional and complex problem, most reliability methods fail, while Monte Carlo methods become impractical due to the slow convergence rate. Therefore, first within offshore applications, we employ the dimensionality reduction method called *Active-Subspace Analysis*. This method identifies a low-dimensional subspace of the input space that is most significant to the input-output variability. We exploit this to efficiently train a Gaussian process (i.e., a kriging model) that models the maximum 10-minute crest elevation at the reference point, and to thereby efficiently estimate the short-term wave crest exceedance probability function. The active low-dimensional subspace for the Korteweg-de Vries model also exposes the expected incident wave groups associated with extreme

Kenan Šehić
Department of Applied Mathematics and Computer Science, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark
Tel.: +4545 25 30 74
Fax: +4545 25 30 74
E-mail: kense@dtu.com

Henrik Bredmose
Department of Wind Energy, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

John D. Sørensen
Department of Civil Engineering, Aalborg University, DK-9100 Aalborg, Denmark
Department of Wind Energy, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

Mirza Karamehmedović
Department of Applied Mathematics and Computer Science, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

waves and loads. Our results show the advantages and the effectiveness of the active-subspace analysis against the Monte Carlo implementation for offshore applications.

Keywords Active subspaces · offshore applications · Monte Carlo methods · Probability of exceedance · Reliability analysis

1 Introduction

Nonlinear hydrodynamic effects are a major concern in bottom-fixed offshore structures at shallow and intermediate depth. Structures such as wind turbines must be designed to withstand extreme waves with strongly nonlinear behavior. The simplest model of the waves would stem from linear wave theory and use a Gaussian stochastic model for the wave surface, resulting in a Gaussian response. However, this approach ignores the nonlinearity that leads to a marked asymmetry in the waves, which means the wave crest elevation systematically exceeds the trough depths at the same probability level [1]. The asymmetry increases with decreasing water depth, which eventually produces substantial instabilities resulting in breaking waves and extreme loads. A number of uncertainty sources need to be accounted for when applying numerical wave simulations as an attempt to represent the real offshore conditions [2–7]. These uncertainties are related to the long-term representation of sea-state parameters, wave surface elevation, kinematics, and estimation of wave loads.

For structural reliability analysis, the probability of failure, in general, is written as a d-fold integral

$$P_F = \int_{g(\theta)\leq 0} \pi_d(\theta) d\theta, \tag{1.1}$$

where $\theta \in \mathbb{R}^d$ is the uncertain input of a numerical model for the limit-state function $g(\theta)$, π_d is the joint probability density function (PDF) for θ, and $g(\theta) \leq 0$ is the failure criterion.

For failure modes within the offshore engineering framework, $g(\theta)$ can model failure events related to wave load effects exceeding arbitrary specified resistances. We here assume the failure event to be related to the maximum crest elevation exceeding a critical level within a certain sea state. We choose the sea-state duration of 10 minutes, which can be relevant for offshore wind turbines. Eq. (1.1) is related to the short-term exceedance probability as standard normal random variables θ construct random wave amplitudes for the wave generation with a predefined ocean-wave spectrum and the wave propagation time. If we would additionally include uncertainties/variability related to the sea state, we would evaluate the long-term exceedance probability P_L related to, e.g., one year as

$$P_L = \int_{\text{state}} P_F(\text{state})\pi(\text{state})\, d\,\text{state}, \tag{1.2}$$

where $P_F(\text{state})$ is the probability of failure for a given sea state obtained by Eq. (1.1) and $\pi(\text{state})$ accounts for the long-term stochastic modeling of the sea state parameters. In Eq. (1.1), we assume θ are defined within the standard normal space. If this is not the case, the Rosenblatt transformation [8] or the Nataf distribution [9] can be used to transfer a non-standard input distribution to the standard normal space. In our study, we focus on the short-term exceedance probability P_F, Eq. (1.1), for a predefined sea state with independent and identically distributed (iid) random variables θ drawn from the standard normal distribution $\mathcal{N}(0,1)$. We model the wave surface elevation but do not include the effects of model uncertainties in the estimations. Further, we formulate the limit-state function such that failure corresponds to the 10-minute maximum crest elevation exceeding a threshold value γ, and the failure condition is rewritten as $g(\theta) \geq \gamma$, where $g(\theta)$ is recognized as the quantity of interest (e.g., the maximum crest elevation) of a numerical wave model, $g : \mathbb{R}^d \to \mathbb{R}$.

The standard reliability approach based on Taylor expansion (the first-order reliability method - FORM and the second-order reliability method - SORM) fails for multiple design points and high-dimensional cases [10]. A more robust approach would be to use the simple Monte Carlo (MC) method that can handle any numerical model. The simple MC approximates Eq. (1.1) by the sample mean of the indicator function $\mathbb{I}(\theta)$, where $\mathbb{I}(\theta) = 1$ if $g(\theta) \geq \gamma$ and $\mathbb{I}(\theta) = 0$ otherwise. The major disadvantage of MC is its inefficiency. Following the mean squared error indicator for a finite sampling of Eq. (1.1) [11], we would need to evaluate a numerical model $5 \cdot 10^4$ times to estimate the exceedance probability of $2 \cdot 10^{-3}$ with the relative error less than 0.1. It would take approximately 35 days to estimate the sample mean of Eq. (1.1) for a numerical model that runs for 1 minute. Specific variance reduction and surrogate approximation methods such as Polynomial Chaos expansion [7, 12] and Gaussian (Kriging) process [13] were proposed to improve the performance. However, their requirements would exponentially grow with the dimension. For Gaussian process regression, a large covariance matrix would need to be inverted several times to produce a prediction.

Therefore, a solution is to search for and exploit a low-dimensional subspace of the input space of initial uncertainties that captures the variability of the quantity of interest and that constitutes a suitable low-dimensional foundation for surrogate models. This method is called active-subspace analysis (ASA) [14]. Previously, similar work had been done in the Ph.D. thesis by Trent M. Russi [15]. It is based on the gradients of the system output, in our case the gradients of the maximum crest elevation, and it can be seen as a principal component analysis in the input space. The gradients can reveal hidden correlation between linear combinations of the input parameters θ of a numerical wave model g and the variability of the quantity of interest $g(\theta)$. We hence determine a low-dimensional subspace by rotating the input space, separating the directions of substantial variability from directions where the quantity of interest changes insignificantly on average [16]. Gradients can be estimated numerically by adjoint methods [17, 18], finite difference approxima-

tions or automatic differentiation [17, 19]. For this paper, we employ forward automatic differentiation (F-AD). In high-dimensional numerical experiments, F-AD is inefficient as it requires one realization per input parameter. However, combining F-AD with an adjoint equation, gradients of the quantity of interest with respect to the input parameters can be estimated within one numerical realization. We here do not include the adjoint approach. We apply the active-subspaces method on a simplistic, yet nonlinear, shallow-water wave model that is a reasonable intermediate step towards a fully nonlinear model. This model is thus used here to test the advantages and disadvantages of the active-subspace analysis within offshore applications against the standard methods.

We examine the implementation of the active-subspace analysis within Gaussian process regression to efficiently and accurately evaluate the short-term exceedance probability for the maximum 10-minute crest elevation at the reference point. Section 2 briefly introduces Gaussian process regression, which we use to define a surrogate model for the numerical wave model, while Section 3 outlines the theoretical background of active-subspace analysis utilized to reduce the computational load of the Gaussian process regression by obtaining a low-dimensional subspace within the input parameters of the wave generation. In Section 4, we describe the shallow-water wave model, and Section 5 contains the numerical results. There, we demonstrate that Gaussian process regression based on the active-subspace analysis can estimate the exceedance probability based on only 1% of the required Monte Carlo evaluations. The paper closes with the conclusions in Section 6.

2 Gaussian process regression - Kriging

Quantification of extreme events for expensive numerical models such as OceanWave3D [20] or other high-order numerical models [7] is generally impractical due to the intensive computation load. A standard approach would be to evaluate an expensive numerical model at only a few carefully designed points, that are then used within simple polynomial regression, to formulate (i.e., to train) a cheap replacement (i.e., a surrogate model). However, the limitation of simple polynomial regression is the assumption that the errors of observations are independent, which may not be true in most cases. In addition, it requires an optimal polynomial degree that is difficult to know prior to training. Therefore, Gaussian process regression (GP, also known as kriging) utilizes a probabilistic framework via a kernel function to describe the correlation (i.e., interaction) of observations within simple regression, which results eventually in the most probable behaviour of a function (i.e., an expensive numerical model) based on the training data (i.e., the sample set). The approach typically does not require a large sample set, even for high-dimensional problems, as a GP model is primarily built on simple regression. The probabilistic framework within a GP model provides an uncertainty measure for predictions via a confidence interval, which is a very powerful property that can be employed for the sequential design and generally as an error indicator.

With the assumption that the quantity of interest is a smooth function, Gaussian process regression describes $g(\theta)$ as a realization of an underlying Gaussian process [13]

$$g(\theta) \approx \hat{g}(\theta) = \beta^T \cdot f_T(\theta) + \sigma^2 Z_{\mathrm{GP}}(\theta, \omega_z), \tag{2.1}$$

where $\beta^T \cdot f_T(\theta)$ is the trend of the GP which is a simple regression form, e.g., linear or quadratic, σ^2 is the Gaussian process variance and $Z_{\mathrm{GP}}(\theta, \omega_z)$ is a zero-mean, unit-variance stationary Gaussian process with ω_z an elementary event (i.e., a sample point) in the probability space. The trend describes the global behavior of the quantity of interest $g(\theta)$. The probabilistic foundation of a Gaussian process is a kernel matrix $\mathbf{K}_{ij} = K(|\theta_i - \theta_j|; \Theta)$ with hyperparameters Θ (such as the overall correlation of samples or smoothness). The overall performance is sensitive to the selection of the optimal kernel function and of the design points. Generally, finding an optimal number of design points N for Gaussian process regression is a standard challenge. Gramacy and Apley [21] suggested selecting the number of design points which minimizes the mean squared predictive error.

Define the input matrix $\mathbf{X} = (\theta_{ij}) \in \mathbb{R}^{N \times d}$ and write the corresponding evaluations of a numerical wave model g as $Y = (Y_i = g(\theta_i)) \in \mathbb{R}^{N \times 1}$, where N is the number of observations and corresponding numerical evaluations with d as the dimension of the input parameters $\theta \in \mathbb{R}^d$. Firstly, the parameters β, σ^2 are generated by a generalized least-squares regression [13]. For a kernel matrix $\mathbf{K}_{ij}$, the hyperparameters Θ are estimated by the maximum likelihood estimation. Finally, for predictions, we define the prediction mean $\mu_g(\theta)$ and the corresponding variance $\sigma_g^2(\theta)$ for the quantity of interest $g(\theta)$ as [13]

$$\mu_g(\theta) = f_T(\theta) \cdot \beta + k(\theta)^T \mathbf{K}^{-1}(Y - \mathbf{F}_T \beta), \tag{2.2}$$

$$\sigma_g^2(\theta) = \sigma^2 \left(1 - \langle f_T(\theta)^T k(\theta)^T \rangle \begin{bmatrix} 0 & \mathbf{F}_T^T \\ \mathbf{F}_T & \mathbf{K} \end{bmatrix}^{-1} \begin{bmatrix} f_T(\theta) \\ k(\theta) \end{bmatrix} \right). \tag{2.3}$$

Here $k(\theta)$ is the correlation between the prediction and the rest of the samples within the set and $\mathbf{F}_T$ is the information matrix regarding the GP trend. Now, instead of using an expensive numerical model g to evaluate, e.g., the maximum crest elevation at an offshore application, we can use a cheap surrogate model, Eq. (2.2), and estimate the short-term exceedance probability, Eq. (1.1), by simple MC. The second moment, Eq. (2.3), quantifies uncertainties in the predictions. The MATLAB function `fitrgp` from the Statistics and Machine Learning Toolbox trains a Gaussian process regression model based on design points.

However, for higher dimensions, e.g., $d = 100$, the process of estimating the GP parameters becomes time-consuming as it requires repeated inversion of the $N \times N$ kernel matrix, incurring a $\mathcal{O}(N^3)$ cost. Also, to estimate the hyperparameters with the maximum likelihood approach, the kernel matrix K_{ij} needs to be inverted. The process can be improved if we find a low-dimensional

optimal representation of θ for $g(\theta)$. We assume that it is inexpensive to estimate gradients numerically for the quantity of interest $g(\theta)$.

3 Active-subspace analysis

Active-subspace analysis (ASA) is a dimensionality reduction approach that has been studied in detail in the book [14] by Paul G. Constantine. It is based on identifying and exploiting the most important linear combinations of the input parameters concerning the quantity of interest, e.g., in our case, the maximum crest elevation at the reference point. A split between important and less important linear directions in the input space is usually defined by a spectral gap in the eigenvalues of the gradient data.

We assume that $g : \mathbb{R}^d \to \mathbb{R}$ is a differentiable function that is square integrable with respect to a probability density π_d for the initial uncertainties θ. An active subspace, i.e., a subspace of the input space with significant variation of the output, is ideally spanned by a relatively small number ($\ll d$) of eigenvectors of the symmetric positive semi-definite $d \times d$ matrix $\mathbf{C}$, which is an uncentered covariance matrix of the output gradients. Thus, we write the expected value of the outer product of the gradients as [14, 16, 22]

$$\mathbf{C} = \int \nabla_\theta g(\theta) \nabla_\theta g(\theta)^T \pi_d(\theta)\, d\theta = \mathbf{W} \boldsymbol{\Lambda} \mathbf{W}^T, \tag{3.1}$$

where $g(\theta)$ is the quantity of interest, $\nabla_\theta g$ is the gradient of $g(\theta)$ with respect to θ, the non-negative eigenvalues of $\mathbf{C}$ are sorted in descending order along the diagonal of the diagonal matrix $\boldsymbol{\Lambda}$, and $\mathbf{W}$ is the orthogonal matrix of eigenvectors $d \times d$.

As shown in **Lemma 3.1.** [14, p. 23], each eigenvalue λ_i is the expected squared directional derivative of $g(\theta)$ along the corresponding eigenvector $\mathbf{w}_i$,

$$\lambda_i = \int (\nabla_\theta g(\theta)^T \mathbf{w}_i)^2 \pi_d(\theta)\, d\theta. \tag{3.2}$$

Hence, if there is a significant spectral gap after the first r largest eigenvalues of $\mathbf{C}$, with $\mathbf{W}_r$ being the first r columns of the orthogonal eigenvector matrix $\mathbf{W}$, then it should be possible to construct a reasonable approximation of $g(\theta)$ in terms of [14, 16]

$$g(\theta) \approx \hat{g}(\mathbf{W}_r^T \theta), \tag{3.3}$$

where $\hat{g}$ is a surrogate model obtained using, e.g., a regression. The reduction of the input space dimension helps to build efficiently a surrogate model $\hat{g}$ and eventually to quantify uncertainties in an otherwise infeasible setting.

3.1 Active subspace estimation

The covariance matrix $\mathbf{C}$, Eq. (3.1), cannot be computed exactly. Therefore, we employ the simple Monte Carlo method to approximate it as [14, 16]

$$\mathbf{C} \approx \hat{\mathbf{C}} = \frac{1}{M} \sum_{i=1}^{M} (\nabla_{\theta_i} g(\theta_i))(\nabla_{\theta_i} g(\theta_i))^T. \tag{3.4}$$

The estimation of how many samples are required to approximate the covariance matrix accurately is heuristic. At least, to have full rank, we need to have $M = d$. Constantine [14, p. 35] recommends $M = \alpha_{\mathrm{A}} k_{\mathrm{A}} \log(d)$, where α_{A} is an oversampling factor between 2 and 10, and k_{A} is the number of eigenvalues to approximate. If we can evaluate the gradients analytically, it is straightforward to use Eq. (3.4). However, this is not the case with numerical models in general. At least, we can approximate the gradients. First-order finite differences (FD) require $M \cdot (d + 1)$ model evaluations per gradient evaluation, which is infeasible for high-dimensional computations. Instead, to employ the FD approach, we use forward automatic differentiation, as described in Section 3.1.1. The active-subspace analysis based on singular value decomposition is outlined in **Algorithm 1**.

Algorithm 1 Monte Carlo Estimation of Active Subspace [14, 16]

1: **procedure** ASA($g(\theta)$,π_d)
2: Draw M iid θ_i from π_d. // Use $M = \alpha_{\mathrm{A}} \cdot k_{\mathrm{A}} \cdot \log(d)$. To have at least a full matrix rank, we should have $M \geq d$.
3: For each θ_i, define $\nabla_\theta g_i = \nabla_\theta g(\theta_i)$. // Use an adjoint solver or a finite difference approach.
4: Define the matrix $\mathbf{G}$ following the relation $\hat{\mathbf{C}} = \mathbf{G}\mathbf{G}^T$ as

$$\mathbf{G} = \frac{1}{\sqrt{M}} \left[\nabla_{\theta_1} g(\theta_1), \nabla_{\theta_2} g(\theta_2), ..., \nabla_{\theta_N} g(\theta_N) \right]. \tag{3.5}$$

5: Compute the singular value decomposition $\mathbf{G} = \widehat{\mathbf{W}} \sqrt{\widehat{\boldsymbol{\Lambda}}} \hat{\mathbf{V}}^T$.
6: **end procedure**

Following **Line 5** in **Algorithm 1**, we search for a spectral gap in the singular values of the matrix $\mathbf{G}$ as a means of identifying the important (active) and the unimportant (inactive) directions in the input space θ. If the singular values do not present a significant spectral gap, an alternative is to estimate the distance between the true r-dimensional active subspace and the estimated (i.e., approximated) r-dimensional active subspace as follows [14, p. 32]

$$\mathrm{dist}(\mathrm{ran}(\mathbf{W}_r), \mathrm{ran}(\widehat{\mathbf{W}}_r) \leq \frac{4\lambda_1 \epsilon}{\lambda_r - \lambda_{r+1}}, \tag{3.6}$$

where $\mathbf{W}_r$ is the true subspace, $\widehat{\mathbf{W}}_r$ is the estimated subspace, the denominator is the spectral gap, 'ran' with a matrix argument is a shorthand notation for

the range of the columns of the matrix, and ϵ is the relative accuracy. The distance, Eq. (3.6), also depends on the spectral gap $\lambda_r - \lambda_{r+1}$ meaning the larger this gap is, the better the estimate. The relative accuracy ϵ can be estimated, following **Corollary 3.10** [14, p. 32], by

$$\epsilon \leq \frac{\lambda_r - \lambda_{r+1}}{5\lambda_1}. \tag{3.7}$$

As pointed out by Constantine [14, p. 32], this bound could perhaps be improved. Evidently, Eq. (3.6) requires $\mathrm{ran}(\mathbf{W}_r)$ and $\mathrm{ran}(\hat{\mathbf{W}}_r)$ to be uniquely defined, i.e., $\lambda_r \geq \lambda_{r+1}$. The spectral gap equal to zero would result in infinitely many solution, because the subspace approximation error is inversely proportional to the corresponding gaps in the singular values. Furthermore, this means also that, for example, the estimate of a three-dimensional active subspace is more accurate than the estimate of a two-dimensional active subspace, if the spectral gap is larger between λ_3 and λ_4 than between λ_2 and λ_3.

3.1.1 Gradient approximations by Automatic Differentiation (AD)

To construct the gradients for the active-subspace analysis without using the FD approach, forward automatic differentiation (AD) is applied on subroutine by subroutine basis to the code required to compute the quantity of interest. The main strategy behind AD is to define the input parameter θ with an additional second component, $\theta + \dot{\theta}\Gamma$. Here, Γ is a symbol distinguishing the second component analogous to the imaginary unit $i = \sqrt{-1}$, but in the AD case $\Gamma^2 = 0$ as opposed to $i^2 = -1$. The input parameters have been converted from type "real" to type "complex". The "real" part will remain unchanged, and the "imaginary" part can be used to approximate the derivative of variables for a single design variable. We add an imaginary perturbation to the desired complex input parameters to determine the corresponding imaginary part of the quantity of interest. When the process is generated and validated, forward differentiation can be performed. In this paper, the forward differentiation is done using the ADiMat software from the Institute for Scientific Computing of TU Darmstadt [23].

3.1.2 Constructing a regression surface

Once the spectral gap is identified, the function $g(\theta)$ is replaced by its low-dimensional surrogate by expressing the initial uncertainties $\theta \in \mathbb{R}^d$ in terms of the active part $\{y_\mathrm{A}\}$ and inactive part $\{z_\mathrm{A}\}$, [14, p. 24]

$$\theta = \mathbf{W}\mathbf{W}^T\theta = \mathbf{W}_r\mathbf{W}_r^T\theta + \mathbf{W}_{d-r}\mathbf{W}_{d-r}^T\theta = \mathbf{W}_r y_\mathrm{A} + \mathbf{W}_{d-r} z_\mathrm{A}. \tag{3.8}$$

In particular, this means that $g(\theta)$ is expressed as $g(\mathbf{W}_n y_\mathrm{A} + \mathbf{W}_{m-n} z_\mathrm{A})$. Small perturbations of z_A change the quantity of interest $g(\theta)$ insignificantly on

average. Thus, the optimal approximation of $g(\theta)$ is to calculate the conditional expectation for each fixed active point y_{A}, and we define $\hat{g}(y_{\mathrm{A}})$ as

$$\hat{g}(y_{\mathrm{A}}) = \int g(\mathbf{W}_r y_{\mathrm{A}} + \mathbf{W}_{d-r} z_{\mathrm{A}}) \pi_{z_{\mathrm{A}}|y_{\mathrm{A}}}(z_{\mathrm{A}}) dz_{\mathrm{A}}, \tag{3.9}$$

where $\pi_{z_{\mathrm{A}}|y_{\mathrm{A}}}(z_{\mathrm{A}})$ is a conditional probability density [14, p. 49]. One can argue that we are going back to multidimensional integration again, however using MC has its advantages in this specific case as the variation of $g(\theta)$ in the inactive subspace is significantly small and requires only a small number of samples. Therefore, we write $g(y_{\mathrm{A}})$ based on MC as

$$\hat{g}(y_{A,j}) \approx \frac{1}{Z} \sum_{i=1}^{Z} g(\mathbf{W}_r y_{A,j} + \mathbf{W}_{d-r} z_{A,i}), \tag{3.10}$$

where Z is the number of samples in the inactive directions and $\{z_{A,i}\}$ are random sample points from the conditional probability density $\pi_{z_{\mathrm{A}}|y_{\mathrm{A}}}(z_{\mathrm{A}})$ [14, p. 51]. If the function $g(\theta)$ is constant in an inactive direction, meaning that the eigenvalue for this direction is zero, then we need to sample only once to account properly for the variation of $g(\theta)$ along this direction.

Hence, to construct a low-dimensional approximation of $g(\theta)$, we generate a number N_y of fixed active points $y_{A,j}$ in the active subspace and collect their corresponding conditional expectations $\{\hat{g}(y_{A,j})\}$ (i.e., Eq. (3.10)). Based on the pairs $\{y_{A,j}, \hat{g}(y_{A,j})\}$ along the active directions $\mathbf{W}_r$, we generate a regression surface for $\hat{g}(y_{\mathrm{A}})$ that is a low-dimensional approximation of the function $g(\theta)$,

$$g(\theta) \approx \hat{g}(\mathbf{W}_r^T \theta). \tag{3.11}$$

Thus, instead of training a Gaussian process (i.e., a kriging model) in the original, high-dimensional space $\mathbb{R}^d$, we first project the training set onto the active, low-dimensional subspace $\mathbb{R}^r$ ($r \ll d$) using the active directions $\mathbf{W}_r^T$ and then train a Gaussian process efficiently and accurately between $\mathbf{W}_r^T \theta \in \mathbb{R}^r$ as the input parameters and the corresponding evaluations $Y \in \mathbb{R}$ as demonstrated in Section 2.

4 A simple 1D Korteweg-de Vries model

While our long-term goal is the accelerated load statistics for fully nonlinear models, we here use a much simpler wave model to investigate the feasibility of the active-subspace analysis for rare events. This section describes the background of the simple wave model g used to propagate waves until a reference location. Initially, we describe the numerical implementation of the model, for which the goal is to define a cheap replacement (i.e., a surrogate model) with Gaussian process regression. The second part explains the wave generation (i.e., the boundary condition for the model) that defines a random surface elevation used for the propagation.

We consider unsteady water waves defined by the Korteweg-de Vries equation (KdV) for uni-directional nonlinear surface waves under the influence of gravity. The original KdV equation, derived by Korteweg and de Vries (1895), describes weakly nonlinear and weakly dispersive shallow-water waves by adding one dispersive term to the nonlinear shallow water equation. There are different modifications of the KdV equation, and we here use KdV22 [24],

$$\eta_t(x,t) + \sqrt{g \cdot h} \cdot \eta_x(x,t) + \frac{3}{2}\sqrt{\frac{g}{h}}\eta(x,t)\eta_x(x,t) + \\ + (\beta + \frac{1}{6})\sqrt{\frac{g}{h}}h^3\eta_{xxx}(x,t) + \beta h^2 \eta_{xxt}(x,t) = 0, \quad (4.1)$$

with $\beta = 19/60$. The KdV22 equation is derived by multiplication of the original KdV equation by $(1 + \beta\partial_{xx})$ and retainment of only the leading order nonlinear and dispersive terms. The β factor allows for adjustment of the dispersive behaviour of the equation. For $\beta = 19/60$, the linear phase speed is the Padé [2,2] approximation of the fully dispersive result. Further, h is the seabed depth, x is the spatial-domain variable, g is the gravitational acceleration, η_t represents $(\partial\eta/\partial t)(x,t)$ and η_x represents $(\partial\eta/\partial x)(x,t)$. The nonlinear term $\eta\eta_x$ accounts for the steepening of the waves, while η_{xxx} and η_{xxt} are the leading-order dispersive terms, that describe the effect of frequency dispersion, that makes long waves travel faster than short waves. For the KdV22 model, which does not describe breaking waves, we assume inviscid and irrotational flow. The seabed is assumed to be flat at the depth of $h = 20$m as illustrated in Fig. 1.

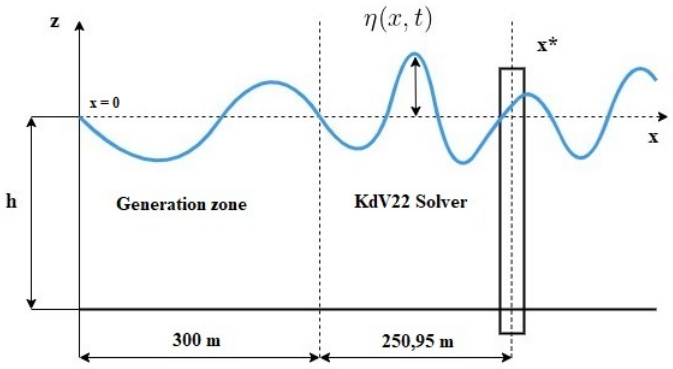

Fig. 1 Illustration of the propagation of the surface elevation $\eta(x,t)$ via the generation zone, Eq. (4.2), and the KdV22 wave model, Eq. (4.1), up to the fixed spatial location $x = x^*$.

4.1 Numerical solution and wave generation

The KdV22 equation is solved numerically by the method of lines. The spatial domain is extended periodically along the x-axis, to allow spectral computation of spatial derivatives by forward and inverse fast Fourier transform (FFT). To this end, we neglect the spectral content above 60% of the Nyquist frequency to avoid aliasing from the quadratic nonlinearity. The time integration is carried out by the classical fourth-order Runge-Kutta method.

We are interested in the nonlinear wave propagation of a known wave field η_{BC}, Eq. (4.4), from $x = 0$ to a down-wave location $x = x^*$, that represents a potential wind turbine location as illustrated on Fig. 1. In a classical non-periodic wave model, this is achieved by either imposing η_{BC} as a boundary condition or by enforcing it in a generation zone (see e.g. [20]). We use the latter technique here, and enforce a known wave solution η_{zone} in an embedded generation zone from $x = 0$m to $x = 300$m.

The generation zone damps the numerical solution that propagates into the zone as its 'outer edge' and transforms it continuously to the η_{zone} field, Eq. (4.5), out of the zone by enforcing the correction

$$\eta_t = [\eta_t]_{\mathrm{KdV22}} - \gamma_{\mathrm{force}}\chi(\xi)(\eta - \eta_{\mathrm{zone}}), \tag{4.2}$$

where $\gamma_{\mathrm{force}} = 3.5$, χ is the spatial weighting factor [25]

$$\chi(\xi) = 1 - \frac{\exp(\xi^{\beta_{\mathrm{shape}}}) - 1}{\exp(1) - 1}, \tag{4.3}$$

and $\beta_{\mathrm{shape}} = 3.5$ is a factor that governs the spatial variation of the weighting factor. Finally, $\xi \in [0, 1]$ is a local coordinate, equal to zero at the outer edge and to one at the inner edge of the generation zone.

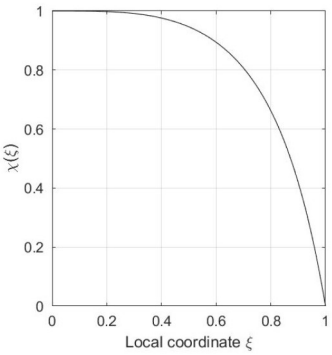

Fig. 2 Graph of the spatial weighting factor, $\chi(\xi)$, used for matching the periodic boundary condition [25].

Ocean waves are stochastic and can be reasonably well described as Gaussian and ergodic processes. This description provides a good starting point for

numerical wave simulations. Therefore, the unidimensional free surface elevation used in this paper as the boundary condition is

$$\eta_{\mathrm{BC}} := \eta(0,t) = \sum_{n=1}^{d/2} \sqrt{S(f_n)\cdot \Delta f}\Big[A_n \cos(2\pi \cdot f_n t) + \\ + B_n \sin(2\pi \cdot f_n t)\Big]. \tag{4.4}$$

Here, $S(f_n)$ is the JONSWAP spectrum value (Section 4.1.1) for the corresponding frequency value $f_n = n\Delta f$, $\Delta f = 1/T$ is the inverse of the wave simulation duration T, and A_n and B_n are independent and uncorrelated random variables drawn from the standard normal distribution $\mathcal{N}(0,1)$. In order to simplify the notation, we write $\theta = (A_1, \ldots, A_{d/2}, B_1, \ldots, B_{d/2}) \in \mathbb{R}^d$ for the active-subspace analysis with π_d as the standard normal density function, i.e., defined for $\mathcal{N}(0,1)$.

Inside the wave generation, see Fig. 1, η_{BC} is extended with its solution of the linearized version of the KdV22 equation such that

$$\eta_{\mathrm{zone}} := \eta(x,t) = \sum_{n=1}^{d/2} \sqrt{S(f_n)\cdot \Delta f}\Big[A_n \cos(2\pi \cdot f_n t - k_n x) + \\ + B_n \sin(2\pi \cdot f_n t - k_n x)\Big], \tag{4.5}$$

which satisfies $\eta_{\mathrm{BC}} = \eta_{\mathrm{zone}}(0,t)$ and is further a solution of the linearized field equation Eq. (4.1). Outside of the generation zone, the wave field is propagated nonlinearly towards $x = x^*$, where it is collected. The wave numbers k_n are the solutions of the linear dispersion relation for Eq. (4.1).

A power spectrum such as the JONSWAP is defined within a certain frequency range. Thus, the frequency step Δf determines the number of elements n within Eq. (4.4), i.e., the parameter d. Therefore, 1-hour of wave propagation, Eq. (4.4), with the frequency interval from 0.05Hz to 0.3Hz generates $d \approx 1802$ number of elements, which results in a highly complex uncertainty quantification problem.

4.1.1 Wave spectrum

The wave spectrum density $S(f)$ describes the power spectrum of the free surface elevation. There are many wave spectra used for offshore applications in deep water. A fundamental spectrum is the PiersonMoskowitz spectrum (PM), which describes a fully developed sea. PM is used for fatigue analysis and extreme analysis. We write [26]

$$S_{\mathrm{PM}}(f) = 0.3125 \cdot H_S^2 \cdot f_P^4 \cdot f^{-5} \cdot \exp\left(-1.25 \cdot \left(\frac{f_P}{f}\right)^4\right), \tag{4.6}$$

where H_S is the significant wave height [m], f_P is the peak frequency [Hz] related to the peak period T_P by $f_P = 1/T_P$ and f is the frequency [Hz].

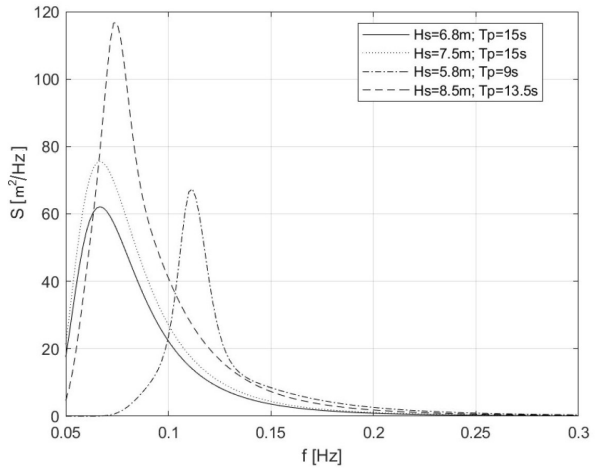

Fig. 3 JONSWAP spectrum for different significant wave heights H_S and wave periods T_P.

The JONSWAP (JS) spectrum is a modification of the PM spectrum for a developing sea state in a fetch limited region. JS accounts for a higher peak and a narrower spectrum in a storm situation. Hence, it is often used for extreme events analyses [7]. JS has additional two parameters: a peak enhancement factor γ^α and a normalizing factor $C_{\mathrm{JS}}(\gamma)$. Here γ^α increases the peak and narrows the spectrum, and $C_{\mathrm{JS}}(\gamma)$ modifies the spectral amplitude to maintain the area under the spectrum. Thus, we write [26]

$$S(f) = C_{\mathrm{JS}}(\gamma) \cdot S_{\mathrm{PM}}(f) \cdot \gamma^\alpha. \tag{4.7}$$

Figure 3 shows examples of the JS spectrum energy distribution curve with different significant wave heights H_S and time periods T_P. We can see that JS is a narrow-banded spectrum. Its energy is mainly focused in a certain frequency band. The JS spectrum was developed to describe storm waves in the North Sea. Since the active subspace method is intended for eventual use in offshore design, we use it in the present study, although the simplistic KdV22 model has its origin in shallow water wave propagation.

5 Results

The KdV22 wave model, Eq. (4.1), is not fully nonlinear but still represents a first natural step towards a fully nonlinear and dispersive model. Expensive numerical wave models such as OceanWave3D [20] require a substantial computational effort to produce reference results, due to the slow convergence rate

of MC methods. Thus, we have chosen a simple but representative replacement such as KdV22 to test and investigate the active-subspace analysis.

In our study, unidirectional water waves propagate in a predefined sea state for $T = 600$ seconds. The usual length of a predefined sea state is 1 hour or 3 hours. We use the length of 10 minutes due to computation limitations. The idea is to have a fast solver to test different approaches before implementing an expensive, fully nonlinear model with longer realisations. Usually, 10 minutes are used for wind load modeling as a time interval with stationary conditions for the wind field turbulence. The significant wave height and the peak period have been specified as $H_s = 6.8$ meters and $T_p = 15$ seconds, as these conditions describe a typical storm conditions with 100-year return period at a typical site of interest.

At the start of the simulation, the d random variables θ are drawn from the standard normal distribution $\mathcal{N}(0,1)$ to determine A and B of η_{BC} for Eq. (4.4). The time integration is then started with the initial condition $\eta(x,0) = 0$. While solving Eq. (4.1), η_{BC} is enforced through the generation zone (4.2). The KdV22 model thereby propagates the input from η_{BC} in space until the reference point $x = x^*$, illustrated in Fig. 1. As we are interested in estimating the short-term exceedance probability, i.e. Eq. (1.1), the quantity of interest $g(\theta)$ is defined here as the maximum crest elevation $\eta_{\max}$ of the surface elevation η within $T = 600$s at the reference point $x = x^*$,

$$\eta_{\max} = \max\{\eta(x^*, t), 0 \leq t \leq T\}. \tag{5.1}$$

As explained previously in the section 4.1, the dimensionality d of the input parameter θ (i.e., the total number of A_n and B_n in Eq. (4.4)) is derived based on a predefined frequency range and the duration of wave propagation. In our case, we have a 10-minute wave simulation with the frequency interval between 0.05Hz and 0.3Hz, which results in $d = 302$. It represents a complex and high-dimensional problem for which standard reliability and surrogate methods become impractical. Initially, we generate $N_{\mathrm{MC}} = 5 \cdot 10^4$ evaluations of KdV22 for iid θ drawn from $\mathcal{N}(0,1)$ to produce the reference probability density function for $\eta_{\max}$ as well the reference short-term exceedance probability for $\eta_{\max}$ as demonstrated in Fig. 4. In Fig. 4a we recognize a heavy-tailed distribution with $\mu_{\eta_{\max}} \approx 5.7$m. The exceedance probability levels lower than 10^{-4} with respect to the maximum crest elevations above 10m as shown in Fig. 4b require more numerical evaluations to have the confidence interval narrowed. The confidence interval of 95% is defined as approximately two standard deviations from the mean estimation. More numerical evaluations for the simple Monte Carlo method would reduce the variance and eventually improve the confidence interval. However, for offshore wind turbines, the short-term exceedance probability typically ranges between 10^{-2} and 10^{-4}. The wind turbine standards [27, 28] indicate a target reliability level corresponding to an annual probability of failure $5 \cdot 10^{-4}$ and 10^{-4}, respectively. These reliability levels are related to the convolution of long-term statistics of mean wind speed, significant wave height, short-term statistics of maximum, e.g., 1-hour

response and etc. As we consider the short-term statistics, which range of exceedance probabilities depends on the relative magnitude of all uncertainties (short-term, long-term, and resistances), the exceedance levels between 10^{-2} to 10^{-4} are sufficient for analysis. In addition, similar considerations for the reliability of offshore oil-gas platforms with a higher required reliability level indicate that this range is acceptable.

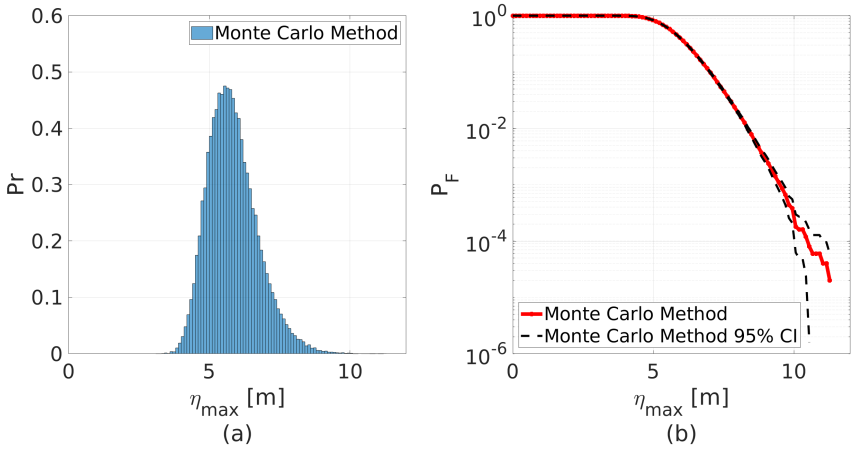

Fig. 4 (a) The probability density function of $\eta_{\max}$ for 10-minutes, based on $N_{\mathrm{MC}} = 5 \cdot 10^4$ numerical evaluations of the KdV22 model. (b) The corresponding short-term exceedance probability of $\eta_{\max}$.

5.1 Dimensionality Reduction

We employ the dimensionality reduction approach outlined in **Algorithm 1** for the maximum crest elevation $\eta_{\max}$ and estimate the corresponding matrix $\mathbf{G}_{\max}$, by using the relation

$$M = \alpha_{\mathrm{A}} k_{\mathrm{A}} \log(d)$$

proposed by Constantine [14, p. 35] for the number of samples M sufficient to estimate the covariance matrix adequately. We define heuristically the oversampling factor $\alpha_{\mathrm{A}} = 2.45$ and are interested in the first 100 eigenvalues, therefore $M = 544$ for $d = 302$. We experimented with different numbers of samples and found $M = 544$ to provide a good balance between the number of gradient evaluations and the approximation quality of the matrix $\mathbf{G}_{\max}$.

Fig. 5a shows the singular values of the matrix $\mathbf{G}_{\max}$, and its corresponding bootstrap replicates. We notice relatively insignificant values, less than 0.2, for all singular values. Thus, the quantity of interest ($\eta_{\max}$) has low variability in each subspace direction. (Recall that a singular value expresses the expected variation of the square of the quantity of interest in the direction of its singular

vector in input space.) Hence we do not need to sample significantly in these directions to have a good overall estimate of $\eta_{\max}$.

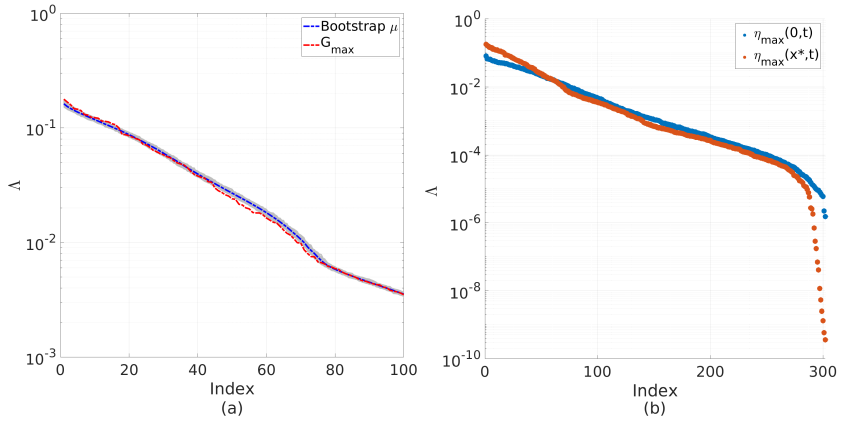

Fig. 5 (a) The singular values $\boldsymbol{\Lambda}$ for the matrix $\mathbf{G}_{\max}$ from the active-subspace analysis with the 500 bootstrap replicates. (b) The singular values $\boldsymbol{\Lambda}$ of the active-subspace analysis for the gradient matrices at the wave generation, Eq. (4.4), $x = 0$ and the reference point $x = x^*$.

Fig. 5b shows the effect of the weakly nonlinear propagation on the singular matrix spectrum of $\mathbf{G}_{\max}$. Compared with the matrix spectrum of the boundary condition, the tail of the propagated matrix spectrum shows an earlier decay at high frequencies and a seemingly more pronounced variation at low frequencies. This may indicate that a matrix spectrum propagated by a fully nonlinear model, such as OceanWave3D [20], will feature a prominent spectral gap. As the maximum frequency for the KdV22 wave model is $f_{\max} = 0.2889$, the input parameters θ above $j = 149$ in Eq. (4.4) are disregarded. This property is recognized as well in Fig. 5b as the singular values above the index value of 289 are insubstantial. The bootstrap replicates in Fig. 5a show the insignificant sample variation of the estimation, which indicates that the singular values are well estimated.

As we cannot find a clear spectral gap in Fig. 5a, we need to estimate the subspace errors using Eq. (3.6). The upper bounds on the subspace errors, Fig. 6, suggest that the 17-dimensional subspace might be the optimal choice for $\eta_{\max}$ as it generates the lowest distance between the true active subspace and its approximation. The bootstrap procedure in Figure 6 reveals a linear increase in the approximation error with increasing dimension. In view of Eq. (3.6), this may be due to the overall flattening of the singular value spectrum with increasing index, that, due to $\lambda_r - \lambda_{r+1}$ approaching zero (on average) with increasing r.

To additionally support our choice of the 17-dimensional subspace, we employ the coefficients of the singular vectors as the design parameters for the boundary condition in Eq. (4.4). Figure 7 reveals the first singular vector $\mathbf{w}_1$

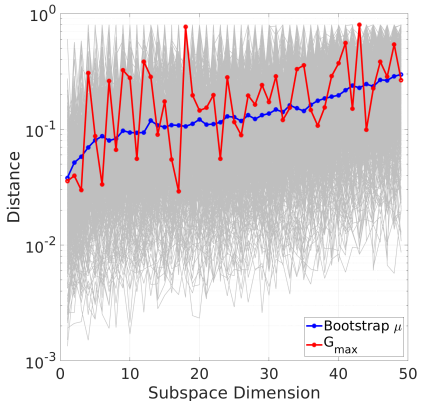

Fig. 6 The estimated error Eq. (3.6) in subspaces of dimension 1 to 49 with the 500 bootstrap replicates.

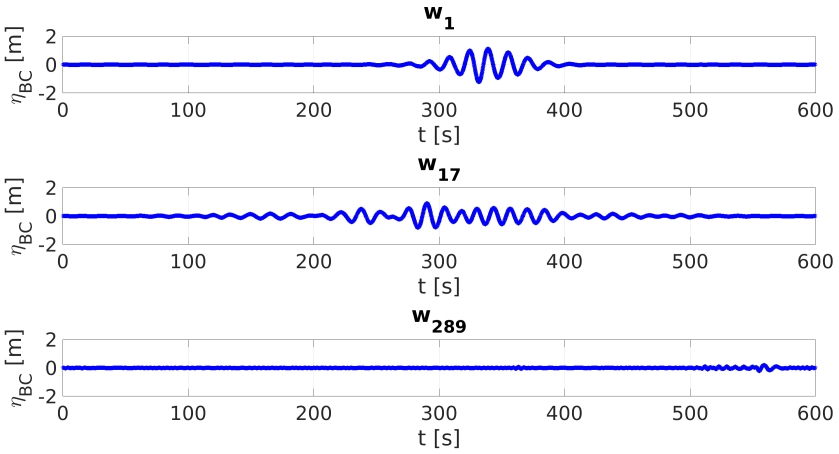

Fig. 7 The boundary condition signals for free surface elevation $\eta_{\mathrm{BC}}(t)$, Eq. (4.4), for the singular vectors $\mathbf{w}_1$, $\mathbf{w}_{17}$ and $\mathbf{w}_{289}$.

(the most active direction in the input space) to be a focused wave group, while the effect diminishes in singular vectors such as $\mathbf{w}_{17}$ that correspond to smaller singular values (the less important directions). The singular vector $\mathbf{w}_{289}$ corresponds to the insignificant singular value λ_{289}, and it therefore represents insignificant free surface variations. It is well-known that extreme waves are associated with wave groups, cf. New Wave theory [29–33]. The ability of the active-subspace analysis to pick out initial conditions that produce a high degree of wave grouping at the structure confirms the relevance of the method. This way, we can construct active focused wave groups for future numerical or even laboratory measurements. The singular vector $\mathbf{w}_{17}$ retains some of the localization, and it makes sense to keep it as an active direction.

The matrix spectrum above the index 17 is treated as measurement noise, for which Gaussian process regression is suitable [16].

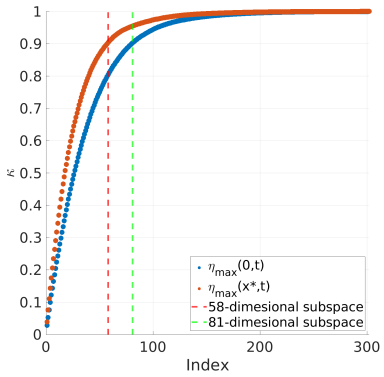

Fig. 8 The ratio κ between $\lambda_1 + ... + \lambda_r$ and $\lambda_1 + ... + \lambda_d$ with the green and red line as the 90% threshold for $\eta_{\max}(0,t)$ and $\eta_{\max}(x^*,t)$.

The identification of the important directions can alternatively be based on a conservative approach [34] that uses the *total variation* of the singular values,

$$\kappa = \frac{\sum_{i=1}^{r} \lambda_i}{\sum_{i=1}^{d} \lambda_i}. \tag{5.2}$$

The active-subspece dimension r is then selected to preserve a certain percentage, say 90%, of the total variation as shown in Fig. 8. It is clear that the singular values with index above 150 are negligible, and the variation κ is preserved 100%. For practical reasons, we might select 90% as our threshold, which would result in a 58-dimensional active subspace for the KdV22 model. In the following, we shall work both with a 17-dimensional and a 58-dimensional active subspace. It is interesting to note in Fig. 8 that the weakly nonlinear wave propagation decreases the dimension of the active subspace for the same level of total variation κ. We expect this effect to also be present when using fully nonlinear models.

Figure 9 includes the boundary condition signals for free surface elevation for the directions of the singular vectors close to index 58. The results expose insignificant permutations of the surface elevation without clear wave groups. Hence, the influence of these singular vectors on the overall result is insignificant as demonstrated in Fig. 8. For their singular values, we can expect that the Gaussian process architecture can easily control the error produced by neglecting the less important directions.

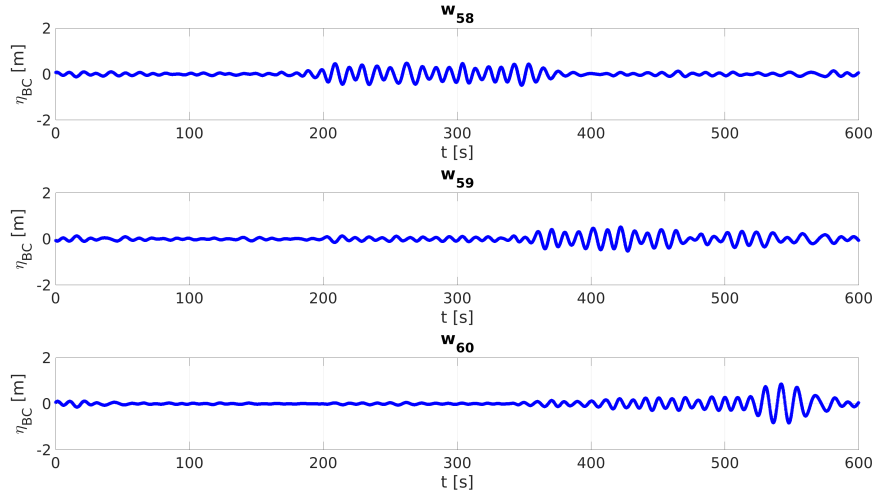

Fig. 9 The boundary condition signals for free surface elevation $\eta_{\mathrm{BC}}(t)$, Eq. (4.4), for the singular vector $\mathbf{w}_{58}$, $\mathbf{w}_{59}$ and $\mathbf{w}_{60}$.

5.2 Active-GP model

The active-subspace analysis based on (3.6) and (5.2) uses 17-dimensional and 58-dimensional active subspaces. We now construct the Gaussian process architecture on these low-dimensional subspaces, selecting the anisotropic squared exponential kernel which for the original high-dimensional space $\mathbb{R}^d$ is defined by

$$K(|\theta_i - \theta_j|;\Theta) = \Theta_0 \exp\left[-\frac{1}{2}\sum_{m=1}^{d} \frac{|\theta_{i,m} - \theta_{j,m}|^2}{\Theta_m} \right], \tag{5.3}$$

where $\Theta = (\Theta_0, \ldots, \Theta_d)$ are the hyperparameters. The hyperparameter Θ_0 is related to the output variance that is how much samples diverge from the mean. While the hyperparametes Θ_m are the length-scale weights. With the active low-dimensional projections $W_r^T\theta$, the kernel is now defined for $\mathbb{R}^r$ by

$$K(|W_r^T\theta_i - W_r^T\theta_j|;\Theta) = \Theta_0 \exp\left[-\frac{1}{2}\sum_{m=1}^{r} \frac{|(W_r^T\theta_i)_m - (W_r^T\theta_j)_m|^2}{\Theta_m} \right], \tag{5.4}$$

which reduces the computation load since $r \ll d$.

The hyperparameters Θ are found using maximum likelihood estimation that is maximizing a likelihood function. This means that under a surrogate model based on Gaussian process regression the observed data within the training set is most probable. As previously explained, typically a specific amount of variation of the quantity of interest is associated with each singular vector, with most variation occurring along the first singular vector. Thus, an anisotropic kernel is a natural choice. A squared exponential part is also an adequate choice as θ is recognized as a set of random Fourier coefficients drawn

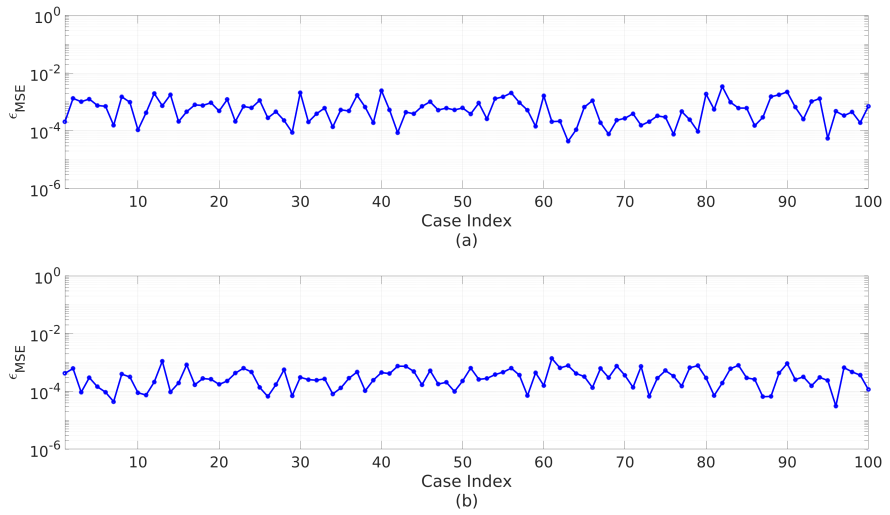

Fig. 10 The mean-squared error estimations ϵ_{MSE} for the cross-validation tests of (a) 17-dimensional active-GP model and (b) 58-dimensional active-GP model.

from the standard normal distribution and the wave generation, Eq. (4.4), as a Fourier series, i.e. a modified white-noise. The trend (i.e, the global behaviour of numerical evaluations) is based on the pure quadratic regression.

Our active-subspace analysis is based on $M = 544$ evaluations of $\eta_{\max,i} = g(\theta_i)$, and their input parameters θ_i are split randomly into the mutually disjoint training set and test set. The size of the training set depends on the active-subspace dimension. The rest of the samples are used as test cases. As we mentioned previously, finding the number of samples to be used for active-GP regression is a well-known problem. We used $N_{\mathrm{GP}}^{17} = 100$ for the 17-dimensional subspace and $N_{\mathrm{GP}}^{58} = 200$ for the conservative approach. We do not claim that this choice is the most efficient and accurate one.

Based on the cross-validation procedure, we draw randomly 100 distinct N_{GP}-combinations of design points θ_i from the M initial observations, and we also record the corresponding evaluations $\eta_{\max,i}$. For each drawn combination, we train an active-GP model and estimate the mean-squared error (MSE) for the short-term exceedance probability based on the test data. We select the optimal design set that achieves a minimal MSE. Here, in both cases, the minimal MSE is achieved with the order of approximately 10^{-4} as demonstrated in Fig. 10. The corresponding active-GP model is kept and used to evaluate all M samples used in the active-subspace analysis. Based on these predictions, Fig. 11 shows the relative error in the predictions against the true evaluations for these M samples. As we need to increase the sample set to $N_{\mathrm{GP}}^{58} = 200$ for the 58-dimensional subspace, the performance of the active-GP models is not directly comparable. However, we can discuss the overall performance. The 17-dimensional active-GP model based on the optimal cross-validation design set attains a relative error of $\approx 13\%$ on average, which is $\approx 15\%$ less than the

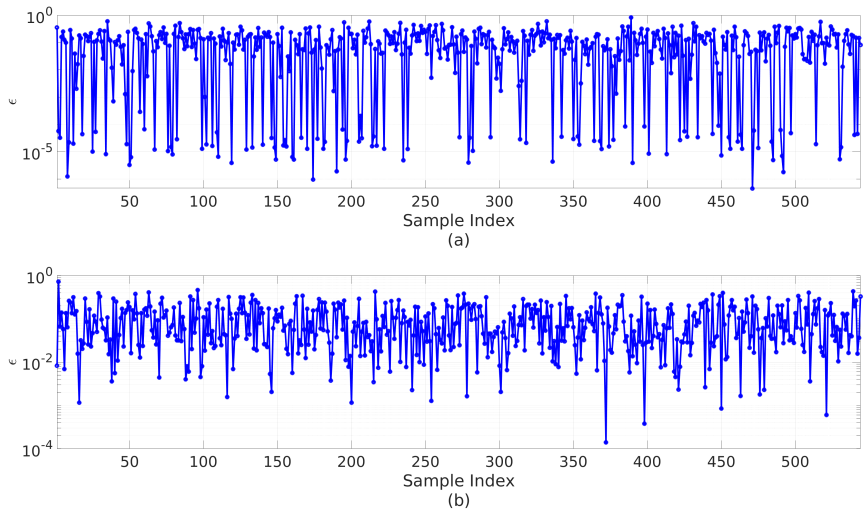

Fig. 11 The relative error estimations ϵ for (a) 17-dimensional active-GP model and (b) 58-dimensional active-GP model.

relative error on average for the 58-dimensional active-GP model. The maximum peaks of the relative error for these GP models are at ≈ 0.85 and ≈ 0.77, respectively. By adding singular vectors up to $\mathbf{w}_{58}$, we build up a Gaussian process architecture that would require a bigger kernel matrix and more design points to describe $g(\theta)$ properly. This can easily give poor performance for relatively small numbers of samples. In light of the singular values in Fig. 5 and of the boundary condition signals for the less important directions in Fig. 7, we know that our quantity of interest changes on average insignificantly in the directions spanned by $\mathbf{w}_j$ with $j > 17$. We can expect that the active-GP architecture will compensate for the errors in the less important directions by treating them as measurement noise, and that an active-subspace analysis based on Eq. (3.6) is sufficient for this work.

As we want to recreate the reference short-term exceedance probability, we evaluate the 17-dimensional active-GP model for $N = 5 \cdot 10^4$ and compare the performances with the simple Monte Carlo as shown in Fig. 12. Figure 12b demonstrates how well the active-GP model reproduces the performance of the simple Monte Carlo. The green lines are the 95% confidence interval as a quality prediction measure because the Gaussian process method employs a distribution over the design points. The active-GP model shows slight underprediction around the exceedance order of 10^{-4} with a relative error of 6.3% on average. The histograms, Fig.12a, are also almost identical with the ℓ^2-distance (i.e., the Euclidean distance) of 0.2. As already mentioned, for offshore wind turbines, the short-term exceedance probability typically ranges between 10^{-2} and 10^{-4}. Therefore, the maximum crest elevation $\eta_{\max}$ at 10^{-3} for the simple Monte Carlo is $\eta_{\max}^{MC} \approx 9.45$m. The active-GP model based on $N_{\mathrm{GP}}^{17} = 100$ points estimates the maximum crest elevation as $\eta_{\max}^{\mathrm{GP}} \approx 9.45$m for the same

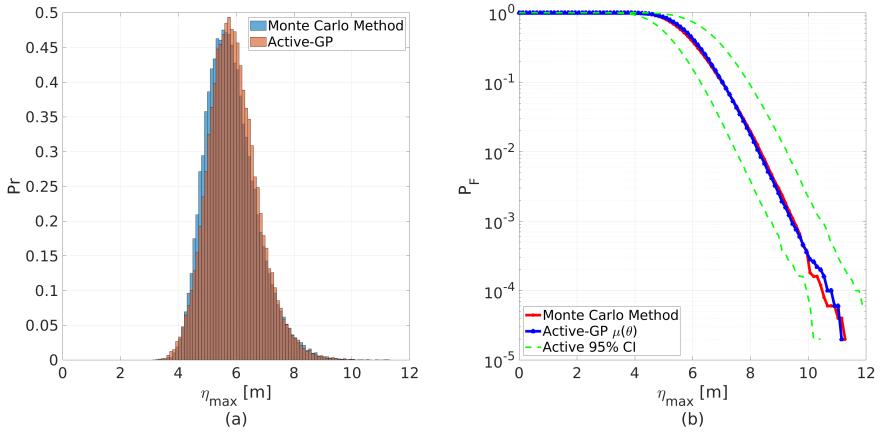

Fig. 12 (a) The short-term probability density function for $\eta_{\max}$ for 10-minutes based on the active-GP model. (b) The associated exceedance probability of $\eta_{\max}$.

exceedance level, which gives a relative error of 0.1%. For the exceedance level of 10^{-4}, the simple MC estimates $\eta_{\max} \approx 10.5$m, while the active-GP model generates $\eta_{\max}^{\mathrm{GP}} \approx 10.65$m with a relative error of 1.4%. These results are collected based on only 544 evaluations, used to estimate the matrix $\mathbf{G}_{\max}$ and to design the active-GP model. This is a reduction in the number of evaluations of 99% compared to simple Monte Carlo.

5.3 A global sensitivity measure

Active-subspace analysis can also provide a sensitivity measure of the quantity of interest, in our case $\eta_{\max}$, regarding the original input parameters θ. In Fig. 13 we plot the components of the singular vectors $\mathbf{w}_j$ and their corresponding frequencies for $j = 1$, 17, 58 and 289. We discover that the frequencies above 0.1Hz are negligible for the singular vectors $\mathbf{w}_{1-17}$ that span the active subspace. This indicates that 66% of the defined JONSWAP spectrum does not significantly affect the quantity of interest. The lower frequencies contribute most to the expectation value of $\eta_{\max}$. While moving in the directions of the less important vectors, e.g., $\mathbf{w}_{58}$ and $\mathbf{w}_{289}$, contributions from the higher frequencies (smaller waves) become more prominent as seen in Fig. 13.

We construct a global sensitivity metric, shown in Fig. 14, by multiplying the singular values λ_j, as the main indicator of the directional importance, with the squared components of the singular vectors. The so-called *activity score* for the j'th component of the input θ, or the j'th initial uncertainty parameter, is then defined by

$$s_j = \sum_{i}^{r} \lambda_i \mathbf{w}_{i,j}^2,$$

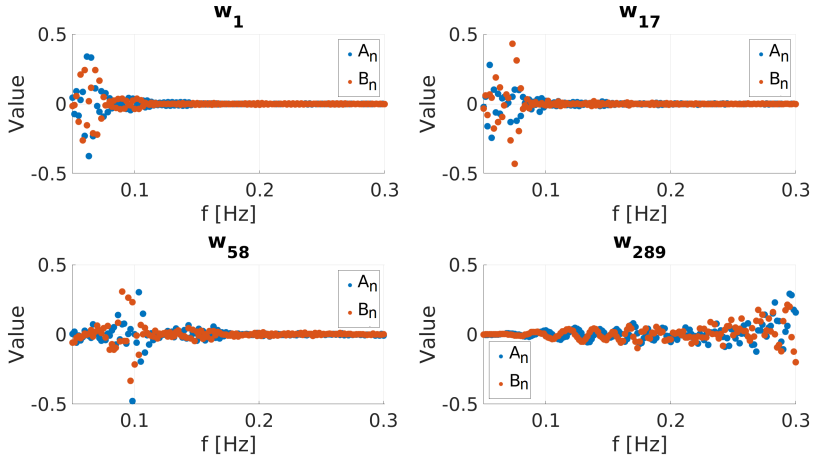

Fig. 13 The components of singular vectors $\mathbf{w}_1$, $\mathbf{w}_{17}$, $\mathbf{w}_{58}$ and $\mathbf{w}_{289}$ for the corresponding A_n (blue) and B_n (orange) as a function of the frequency f.

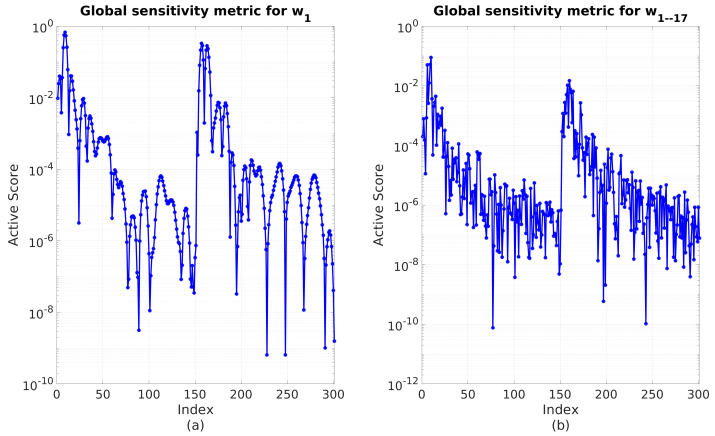

Fig. 14 The activity scores for the input parameters $\theta = (A_1, \ldots, A_{d/2}, B_1, \ldots, B_{d/2}) \in \mathbb{R}^{d=302}$ in the direction of (a) $\mathbf{w}_1$ and (b) $\mathbf{w}_{1-17}$.

where $j \in \mathbb{R}^d$ [35]. It is interesting to notice in Fig. 14 the second shorter peak for the input parameters with the indexes $j = 120$ and $j = 240$ that correspond to the frequency value of 0.11Hz. A wave spectrum, such as a JONSWAP spectrum, is typically a global sensitivity measure with respect to the initial uncertainties $\theta \in \mathbb{R}^d$. We expect those input parameters that correspond to the peak of a wave spectrum to be the most important input parameters as we can see in Fig. 14, which resembles to some extent a wave spectrum Fig. 3. While, for a wave spectrum, higher frequencies over the peak reveal the exponential decay in the importance of the corresponding input parameters, the global sensitivity measure based on the activity scores indicates the second shorter peak of the importance around 0.11Hz (i.e., the index $j = 120$ for

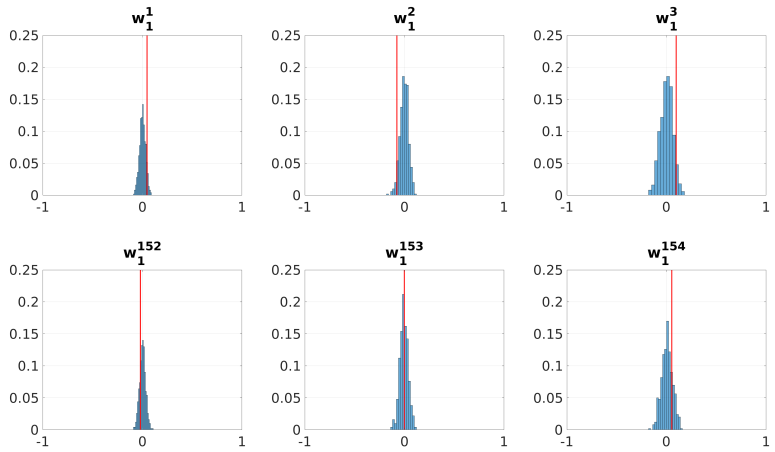

Fig. 15 Bootstrap histograms of the components of the active subspace singular vector $\mathbf{w_1}$ for the maximum crest elevation $\eta_{\max}$.

A_n and $j = 240$ for B_n), which can be related to the modification of the wave spectrum due to wave propagation. This change within a wave spectrum for higher frequencies can be found also in the offshore literature [7], which additionally proves the value of the active-subspace analysis.

To estimate the variability within, e.g., the components of $\mathbf{w}_1$, we employ the bootstrap approach with 500 replicates for the covariance matrix $\mathbf{C}$ and **Line 5** of **Algorithm 1**. This cost is negligible because the bootstrap approach uses only the available model evaluations. The sharp peaks in the histogram around the expected value suggest confidence in the computed directions [14]. However, the relatively wider histograms as seen in Fig. 15 are due to the insufficient number of gradient evaluations M for the active-subspace analysis.

6 Conclusion

We applied a dimensionality reduction method called the active-subspace analysis (ASA) to a high-dimensional offshore problem. We modeled shallow-water waves using a simple but credible weakly nonlinear numerical model based on the Kortweg-de Vries equation (KdV22) with a high-dimensional initial Gaussian input. Our approach can be seen as an intermediate step towards a fully nonlinear model. For this high-dimensional complex problem, the standard offshore methods have an infeasible convergence rate in providing accurate results. The active-subspace analysis uses gradient evaluations to identify a low-dimensional subspace within the input space that is most significant in terms of the sensitivity of the output.

In contrast to Principal Component Analysis (PCA), the ASA reduces dimensionality while retaining information about the numerical model. However,

estimating gradients is typically challenging and requires an adjoint solver for optimal efficiency. We perform our analysis using forward automatic differentiation despite the large required number of realizations.

We apply the ASA to the maximum crest elevation at the reference point to reduce the uncertainty dimension at the wave generation within 10-minute wave propagation for a predefined sea state. The singular value decomposition (SVD) of the gradient evaluations reveals the slow spectral decay for the singular values without a clear spectral gap, which is crucial for accurate active subspace estimation. However, we can construct the low-dimensional active subspace based on the error bound, which exploits the relation between the true and estimated active subspace. Also, the active subspace exposes a focused wave group associated with extreme waves and loads. The global sensitivity of the ASA demonstrates the wave spectrum modification due to wave propagation. Based on the numerical evaluations used for SVD, we train efficiently Gaussian processes on the active subspace for different batches (i.e., different training and test sets) and select the Gaussian process with the lowest mean-squared error. Finally, by using the simple Monte Carlo method, the trained Gaussian process accurately estimates the short-term exceedance probability with the relative error of around 6% on average. The reference short-term exceedance probability is obtained by $5 \cdot 10^4$ numerical evaluations, while the active-subspace analysis and Gaussian process regression use only 1% of the required Monte Carlo evaluations to provide the comparable result efficiently.

Acknowledgements This research was funded by the DeRisk project of Innovation Fund Denmark, grant number 4106-00038B. This support is gratefully acknowledged.

References

1. Jha AK, Winterstein SR (2000) Nonlinear random ocean waves: prediction and comparison with data. In: Proc. OMAE2000, pp OMAE2000–6125
2. Bigoni D, Engsig-Karup AP, Eskilsson C (2016) Efficient uncertainty quantification of a fully nonlinear and dispersive water wave model with random inputs. J Eng Math 101(1):87–113
3. Bredmose H, Dixen M, Ghadirian A, Larsen TJ, Schler S, Andersen SJ, Wang S, Bingham HB, Lindberg O, Christensen E, Vested MH, Carstensen S, Engsig-Karup AP, Petersen OS, Hansen HF, Mariegaard JS, Taylor PH, Adcock TAA, Obhrai C, Gudmestad OT, Tarp-Johansen NJ, Meyer CP, Krokstad JR, Suja-Thauvin L, Hanson TD (2016) DeRisk accurate prediction of ULS wave loads. Outlook and First Results. Energy Procedia 90:379–387
4. Ge L, Cheung KF, Kobayashi MH (2008) Stochastic solution for uncertainty propagation in nonlinear shallow-water equations. J Hydraul Eng 134(12)
5. Ghadirian A, Bredmose H (2019) Pressure impulse theory for a slamming wave on a vertical circular cylinder. J Fluid Mech 867:R1

6. Schløer S, Bredmose H, Ghadirian A (2017) Analysis of experimental data: The average shape of extreme wave forces on monopile foundations and the NewForce model. Energy Procedia 137:223–237
7. Yildirim B, Karniadakis GE (2015) Stochastic simulations of ocean waves: An uncertainty quantification study. Ocean Model 86:15–35
8. Rosenblatt M (1952) Remarks on a multivariate transformation. Ann Math Stat 23(3):470–472
9. Liu P, Der Kiureghian A (1986) Multivariate distribution models with prescribed marginals and covariances. Probabilistic Eng Mech 1(2):105–122
10. Rackwitz R (2001) Reliability analysis - a review and some perspectives. Struct Saf 23(4):365–395
11. Owen AB (2013) Monte Carlo theory, methods and examples. Open Access
12. Li J, Xiu D (2010) Evaluation of failure probability via surrogate models. J Comput Phys 229(23):8966–8980
13. Schöbi R, Sudret B, Marelli S (2016) Rare event estimation using polynomial-chaos kriging. ASCE ASME J Risk Uncertain Eng Syst A Civ Eng 3(2)
14. Constantine PG (2015) Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies. SIAM, Philadelphia
15. Russi TM (2010) Uncertainty quantification with experimental data and complex system models. PhD thesis, UC Berkeley
16. Constantine PG, Dow E, Wang Q (2014) Active subspace methods in theory and practice: Applications to kriging surfaces. SIAM J Sci Comput 36(4):A1500–A1524
17. Lockwood B, Mavriplis D (2013) Gradient-based methods for uncertainty quantification in hypersonic flows. Comput Fluids 85(2013):27–38
18. Bryson AE, Ho YC (1975) Applied Optimal Control: Optimization, Estimation, and Control. Hemisphere Pub. Corp., New York
19. Griewank A (2000) Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation, 1st edn. SIAM, Philadelphia
20. Engsig-Karup AP, Bingham H, Lindberg O (2009) An efficient flexible-order model for 3D nonlinear water waves. J Comput Phys 228(6):2100–2118
21. Gramacy RB, Apley DW (2015) Local Gaussian process approximation for large computer experiments. J Comput Graph Stat 24(2):561–578
22. Constantine PG, Eftekhari A, Wakin MB (2015) Computing active subspaces efficiently with gradient sketching. arXiv:150604190
23. Bischof CH, Bucker HM, Lang B, Rasch A, Vehreschild A (2002) Combining source transformation and operator overloading techniques to compute derivatives for MATLAB programs. In: Proc. Second IEEE International Workshop on Source Code Analysis and Manipulation, pp 65–72
24. Bredmose H (1999) Evolution equations for wave-wave interaction. Master's thesis, Technical University of Denmark, Lyngby, Denmark
25. Paulsen BT, Bredmose H, Bingham HB (2014) An efficient domain decomposition strategy for wave loads on surface piercing circular cylinders.

736 Coast Eng 86(2014):57–76
737 26. Hasselmann K, Barnett T, Bouws E, Carlson H, Cartwright D, Enke K,
738 Ewing J, Gienapp H, Hasselmann D, Kruseman P, Meerburg A, Mller P,
739 Olbers D, Richter K, Sell W, Walden H (1973) Measurements of wind-
740 wave growth and swell decay during the Joint North Sea Wave Project
741 (JONSWAP). Dtsch Hydrogr Zeitschrift Ergänzung Reihe A(8) 12:1–95
742 27. DNV GL (2018) Support structures for wind turbines (DNVGL-ST-
743 0126:2018). Tech. rep., Oslo, Norway
744 28. International Electrotechnical Commission (2019) Wind energy generation
745 systems - Part 1: Design requirements (IEC 61400-1:2019). Tech. rep.,
746 Geneva, Switzerland
747 29. Boccotti P (1983) Some new results on statistical properties of wind waves.
748 Appl Ocean Res 5(3):134–140
749 30. Cousins W, Sapsis TP (2016) Reduced-order precursors of rare events in
750 unidirectional nonlinear water waves. J Fluid Mech 790:368–388
751 31. Ghadirian A, Bredmose H (2019) Investigation of the effect of the bed
752 slope on extreme waves using First Order Reliability Method. Mar Struct
753 67
754 32. Lindgren G (1970) Some properties of a normal process near a local max-
755 imum. Ann Math Stat 41(6):1870–1883
756 33. Tromans PS, Anaturk AR, Hagemeijer A (1991) A new model for the
757 kinematics of large ocean waves-application as a design wave. In: The
758 First International Offshore and Polar Engineering Conference, pp 64–71
759 34. Karamehmedović M, Šehić K, Damman B, Suljagić M, Karamehmedović
760 E (2019) Autoencoder-aided measurement of concentration from a single
761 line of speckle. Opt Express 27(20):29098–29123
762 35. Constantinea PG, Diaz P (2017) Global sensitivity metrics from active
763 subspaces. Reliab Eng Syst Saf 162:1–13