

# A novel method for analysis of offshore sustainable energy systems

Energy Exploration &amp; Exploitation

2023, Vol. 41(3) 1033–1050

© The Author(s) 2023

DOI: 10.1177/01445987231157792

journals.sagepub.com/home/eea

Yingguang Wang<sup>1,2</sup> 

## Abstract

In the present research, we have proposed a new adaptive kernel density estimation method formulated on the theory of linear diffusion processes. By examining the calculation results we have found that in the tail region, our proposed new adaptive kernel density estimation distribution curve becomes very smooth and fits quite well with the histogram of the measured ocean wave dataset at the US National Data Buoy Center station 46026. Carefully studying the calculation results also reveals that the 50-year extreme Power-Take-Off heaving force value forecasted based on the environmental contour derived using the new method is 3021700N, which is much larger than the value 2458700N forecasted via the Rosenblatt-inverse second-order reliability method contour method. Consequently, our proposed new adaptive kernel density estimation method formulated on the theory of linear diffusion processes can forecast well and efficiently forecast the 50-year extreme design force values for offshore sustainable energy systems.

## Keywords

Offshore sustainable energy systems, adaptive kernel density estimation distribution, linear diffusion processes, extreme design force

## Introduction

Designing, installing, and operating offshore sustainable energy systems need the vital information provided by statistically describing the sea severity. Typically, a measured data set of significant

<sup>1</sup>State Key Laboratory of Ocean Engineering, Shanghai Jiao Tong University, Shanghai, P. R. China

<sup>2</sup>School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai, P. R. China

### Corresponding author:

Yingguang Wang, State Key Laboratory of Ocean Engineering, Shanghai Jiao Tong University, Shanghai, 200240, P. R. China; School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai, 200240, P. R. China.

Email: [wyl10@sjtu.edu.cn](mailto:wyl10@sjtu.edu.cn)



wave heights ( $H_S$ ) at a specific sea location spanning a relatively short time period (e. g., less than 25 years) are the available information regarding the sea severity. In practice, researchers and development engineers need to know the roughest sea state (possessing a highest  $H_S$  value) an offshore sustainable energy system will experience during its whole life cycle time (e.g., 50 years) in order to successfully design such a system. In order to accurately forecast this roughest sea state (with the highest  $H_S$  value in a 50-year time period), people need first to find a probability distribution model that can fit the measured ocean wave data well.

Currently, the most frequently used probability distribution model for  $H_S$  is the 3-parameter Weibull distribution (also called the translated Weibull distribution in some literature). Having obtained the fitted marginal distributions of the significant wave height and a specific wave period together with their correlation structures, a 50-year environmental contour line can subsequently be derived by conducting a proper transformation between the Gaussian space and the physical parameter space. Having obtained an environmental contour line by using the inverse first-order reliability method (IFORM) and utilizing the information on the obtained contour line as a basis, Clarindo et al. (2021) further studied to reduce the variance on the probability of failure estimates by applying the Monte Carlo Importance Sampling technique. In this study, the marginal probability distribution of  $H_S$  was described by using a 3-parameter Weibull distribution model fitted to a large set of simulated ocean wave data. In the study of Haselsteiner et al. (2021) for deriving highest density environmental contour lines, the  $H_S$  was also assumed to follow a 3-parameter Weibull distribution. Mackay and Haselsteiner (2021) used a sea state model in which the  $H_S$  is modeled using a 3-parameter Weibull distribution and subsequently obtained environmental contour lines by using a Rosenblatt transformation based on the IFORM method. Wrang et al. (2021) applied the IFORM method to generate 50-year environmental contours based on observed and hindcast data at several sea locations in the Baltic Sea, North Sea, and Skagerrak. For implementing the IFORM analysis, Wrang et al. (2021) assumed that the marginal distribution of  $H_S$  follows a 3-parameter Weibull distribution. In order to study the system reliability of an offshore platform, Zhao and Dong (2022a) calculated environmental contour lines using the IFORM method in which the marginal distribution of  $H_S$  was fitted using the 3-parameter Weibull model. Zhao and Dong (2022b) extended the alternative contour line approach based on the IFORM method in a three-dimensional model that accounting for uncertainties for the short-term extreme responses. In Zhao and Dong (2022b), the marginal distribution of  $H_S$  was fitted to a 3-parameter Weibull model in order to construct the environmental contour lines. Chai and Leira (2018) estimated inverse second-order reliability method (ISORM) environmental contour lines by extending the traditional IFORM method. Nevertheless, when performing the Rosenblatt transformation in the ISORM method the marginal distribution of  $H_S$  is still assumed to follow the 3-parameter Weibull model.

All the aforementioned studies applied the 3-parameter Weibull model for fitting the  $H_S$  during the process of calculating the environmental contour lines. Unfortunately, the suitability of the 3-parameter Weibull model for fitting the  $H_S$  distribution tails has not been investigated in these studies. To investigate this suitability, Haselsteiner and Thoben (2020) analyzed three wave datasets measured at the US Atlantic Coast and another three hindcast-simulated wave datasets covering the North Sea. They derived the concluding remarks that for all six datasets the 3-parameter Weibull model seriously underestimated the probability density values at the tail region.

We can notice that the parametric method for fitting the  $H_S$  probability distribution had been used by all the above-mentioned publications. In the literature, a nonparametric ordinary kernel density estimation (KDE) method was used in Haselsteiner et al. (2017) for calculating the  $H_S$

probability distributions in order to predict environmental contour lines of extreme sea states. However, it was stated in Ross et al. (2020) that the ordinary KDE method behaves poorly in predicting the probability density values at the tail region. To overcome the deficiencies of the ordinary KDE method and improve the performance of capturing the probability distribution tails, Eckert-Gallup and Martin (2016) adopted an adaptive bandwidth selection procedure in their bivariate KDE method for estimating the environmental contour lines of extreme sea states. Nevertheless, it is usually very difficult and time-consuming to implement Abramson's adaptive bandwidth selection procedure proposed in Eckert-Gallup and Martin (2016). After reviewing all the aforementioned literature in its current form, it is obvious that there is a knowledge gap that all the existing methods (parametric or nonparametric) can't predict accurately and efficiently the sea state parameter probability distribution tails. Therefore, finding a highly efficient adaptive bandwidth selection procedure is imperative so that people can predict the probability distribution tails and environmental contour lines more efficiently and accurately.

With the expectation to overcome the disadvantages of all the above-mentioned non-parametric and parametric methods, a new adaptive KDE methodology based on linear diffusion processes will be proposed to help researchers and engineers to predict the probability density values in the tail region more efficiently and accurately. This new adaptive KDE methodology will subsequently be applied to calculate the  $H_s$  probability density values in the tail region based on a measured dataset at the US National Data Buoy Center (NDBC) Station 46026. The efficiency and accuracy of the proposed new adaptive KDE methodology will be validated by comparing its prediction results with those forecasted by using a parametric method, an ordinary KDE method and Abramson's adaptive KDE method. Subsequently, 50-year environmental contours will be respectively forecasted by the new adaptive KDE methodology and the ISORM method with Rosenblatt transformation based on the aforementioned measured dataset at the US NDBC Station 46026. The obtained 50-year environmental contours will then be used for the offshore sustainable analysis (i.e., predicting the 50-year dynamic response values for a typical offshore sustainable energy system (in this case, a floating point absorber wave energy converter (WEC))). By carrying out the aforementioned procedures we have succeeded to bridge the offshore sustainable analysis and our objective of utilizing the proposed new adaptive KDE method based on significant wave height and sea state prediction. By carefully investigating the prediction results, the benefits of utilizing our proposed new adaptive KDE methodology for safety analysis of offshore sustainable energy systems will be demonstrated.

## **The theories corresponding to various analysis methods**

### *The theoretical backgrounds of the IFORM and ISORM environmental contour line approaches*

For the performance analysis of a specific offshore sustainable energy system, the long-term (e.g., 50 years) dynamic response extremes of the sustainable energy system can be predicted by carrying out a short-term (e.g., 3 h) analysis based on the environmental contour line method as specified in, e.g., Haver and Winterstein (2009). The environmental contour lines at a specific ocean site can be calculated when the  $H_S$  marginal probability distribution and the wave period conditional probability distribution are both available.

In order to derive an environmental contour line, one should first transform the marginal  $H_S$  and conditional wave period probability distributions to a standard normal plane (i.e., the  $(u_1, u_2)$  plane) utilizing the Rosenblatt transformations as follows:

$$F_{H_S}(h) = \Phi(u_1) \quad (1)$$

$$F_{T_p|H_S}(t|h) = \Phi(u_2) \quad (2)$$

In equation (1)  $F_{H_S}(h)$  is the  $H_S$  marginal probability distribution,  $F_{T_p|H_S}(t|h)$  in equation (2) is the wave peak spectral period ( $T_p$ ) conditional probability distribution.  $\Phi()$  in the above two equations is the standard univariate Gaussian cumulative distribution function. In the standard univariate Gaussian plane, an environmental contour line corresponding to the  $q$  annual exceedance probability will be a circle possessing a radius  $r = \Phi^{-1}(1 - q/2920)$  in which 2920 equals the number of the 3-h stationary sea states in 1 year.

The radius  $r$  (also call the reliability index and denoted by  $\beta_F$  in the classical literatrary works) is measured from the design point to the origin of the circle. A relation  $\Phi(\beta_F) = 1 - \alpha$  can be obtained if we set  $\alpha = q/2920$ . Utilizing the inverse Rosenblatt transformations (see e.g., Manuel et al. (2018)) as specified in equations (3–4), one can then derive the  $q$ -probability environmental contours by backward transforming the circles in the standard Gaussian plane to the original physical parameter plane:

$$H_S = F_{H_S}^{-1}(\Phi(u_1)) = F_{H_S}^{-1}(\Phi(r \cos \theta)) \quad (3)$$

$$T_e = F_{T_e|H_S}^{-1}(\Phi(u_2)|H_S) = F_{T_e|H_S}^{-1}(\Phi(r \sin \theta)|H_S) \quad (4)$$

One can derive the entire environmental contour line by varying the angle  $\theta$  in equations (3–4) in the interval  $[0, 2\pi]$ . The process of deriving an environmental contour line by using equations (1–4) is called the IFORM (inverse first-order reliability method) in the existing literature. However, the IFORM method has the following disadvantages when being used for deriving environmental contour lines. (a) for an IFORM environmental contour line corresponding to an exceedance probability of  $\alpha$ , the standard Gaussian variables in standard Gaussian space possess maximum values with an exceedance probability of  $\alpha$ . However, the original physical random variables along the IFORM environmental contour line in the original physical parameter space do not have maximum values with an exceedance probability of  $\alpha$  (Mackay and Haselsteiner (2021)). (b) For the case with a convex failure region, the predicted results by using the IFORM contour method are conservative. In contrast, for the case with a concave failure region, the predicted results by using the IFORM contour method are not conservative. In Chai and Leira (2018) an always conservative ISORM was proposed by conducting a second-order approximation to the failure surface. This ISORM method assumes that in the standard Gaussian space the failure surface encloses a circle with a radius  $\beta_{S2}$  centered at the origin. The value of the radius  $\beta_{S2}$  is chosen such that the probability that an observation falls outside the circular space equals  $\alpha$ . It is noticed in Chai and Leira (2018) that because the sum of the squares of two independent standard normal variables follows a Chi-square distribution with two degrees of freedom,  $\chi_2^2$ , the radius  $\beta_{S2}$  satisfies the following mathematical relation:

$$\chi_2^2(\beta_{S2}^2) = 1 - \alpha \quad (5)$$

From the above relation we can find that in the ISORM method  $\beta_{S2}$  is a function of  $\alpha$  (the exceedance probability). As with the IFORM environmental contour lines, the ISORM environmental

contour lines in the physical parameter space are obtained by conducting the inverse Rosenblatt transformation to the contour lines in standard normal space.

### *The theories for the ordinary and Abramson's adaptive KDE methods*

The IFORM and ISORM approaches for deriving environmental contour lines are both parametric methods. For implementing these two methods, some specific probability distribution models regarding the sea state parameters should be predefined. However, the parametric modeling approach has a serious drawback, i.e., the predefined probability model might be too rigid and restrictive to accurately estimate the true underlying function. In order to overcome the rigidity of the parametric models of the sea state parameter probability distributions, one can utilize the non-parametric KDE method. In probability and statistics, KDE is the application of kernel smoothing for probability density estimation, i.e., a non-parametric method to estimate the probability density function of a random variable based on kernels as weights. The ordinary univariate KDE is carried out as follows (Silverman (1986)):

$$\hat{f}(x; h) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right), \quad (6)$$

In equation (6) ( $X_1 \dots X_i \dots X_n$ ) are  $n$  data samples drawn from an unknown density  $f$  at any given point  $x$ .  $K$  is a non-negative real-valued integrable kernel function that satisfies  $\int K(x)dx = 1$ .  $h > 0$  is a smoothing parameter called the window width parameter or bandwidth parameter. Several types of kernel functions are commonly used: uniform, triangle, Epanechnikov, quartic (biweight), tricube, tri-weight, Gaussian, quadratic, and cosine. In this study, we choose to use the Epanechnikov kernel because it is optimal in a mean square error sense and has convenient mathematical properties.

The bivariate KDE can be carried out through extending the aforementioned univariate KDE process by using two bandwidths ( $h_1$  in one coordinate's direction and  $h_2$  in another coordinate's direction). Suppose  $X_{11}, \dots, X_{n1}$  are a sample of dataset in one coordinate's direction and  $X_{12}, \dots, X_{n2}$  are a sample of dataset in another coordinate's direction, the ordinary bivariate KDE calculation can be conducted as follows:

$$\hat{f}(x_1, x_2; h_1, h_2) = \frac{1}{n h_1 h_2} \sum_{i=1}^n K\left(\frac{x_1 - X_{i1}}{h_1}, \frac{x_2 - X_{i2}}{h_2}\right) \quad (7)$$

The ordinary KDE method (either by implementing equation (6) or equation (7)) usually can only predict well the mode of a probability density function. Unfortunately, the ordinary KDE method is not able to satisfactorily forecast the tails of a probability density function. To overcome this deficiency, Eckert-Gallup and Martin (2016) suggested to use Abramson's adaptive KDE method by varying the bandwidth of the kernels over the domain of estimation. The first step to use Abramson's adaptive KDE method is to find a pilot estimate  $\tilde{f}(x_1, x_2)$  that satisfies  $\tilde{f}(X_{1i}, X_{2i}) > 0$  with respect to all  $i$ . Then the local bandwidth parameters  $\lambda_i$  can be calculated using the following equation (Wand and Jones (1995)):

$$\lambda_i = \{\tilde{f}(X_{1i}, X_{2i}) / g\}^\alpha \quad (8)$$

In equation (8)  $\alpha$  is the sensitivity parameter satisfying  $0 \leq \alpha \leq 1$ .  $g$  is calculated as follows:

$$\log(g) = n^{-1} \sum \log(\tilde{f}(X_{1i}, X_{2i})) \quad (9)$$

The adaptive kernel estimate  $\hat{f}(x_1, x_2)$  can subsequently be calculated using the following equation (Silverman (1986)):

$$\hat{f}(x_1, x_2) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_1 h_2 \lambda_i} K\left(\frac{x_1 - X_{i1}}{\lambda_i h_1}, \frac{x_2 - X_{i2}}{\lambda_i h_2}\right) \tag{10}$$

However, it is usually very difficult and time-consuming to implement Abramson’s adaptive KDE method. In the following, we will propose to use a highly efficient new adaptive bandwidth selection procedure so that people can predict the probability distribution tails and environmental contour lines more efficiently and accurately.

**Adaptive KDE method formulated on the theory of linear diffusion processes**

Suppose once again that  $(X_1, \dots, X_n)$  are a dataset sampled based on an unknown univariate probability density distribution  $f$  at any given support value  $x$ . Botev et al. (2010) showed that the Gaussian kernel density estimator can be derived using the following formulation:

$$\hat{f}(x; t) = \frac{1}{n} \sum_{i=1}^n \phi(x, X_i; t) \tag{11}$$

The  $\phi(\bullet)$  function in the above equation is defined as follows:

$$\phi(x, X_i, t) = \frac{1}{\sqrt{2\pi t}} e^{-(x-X_i)^2/2t} \tag{12}$$

By carefully studying equation (12) we can find that the bandwidth parameter is  $\sqrt{t}$  in this case. Botev et al. (2010) showed that by evolving the following diffusion equation and obtaining its solution up to time  $t$  the Gaussian kernel density estimator as expressed in equation (12) can be uniquely derived:

$$\frac{\partial}{\partial t} \hat{f}(x; t) = \frac{1}{2} \frac{\partial^2}{\partial x^2} \hat{f}(x; t), \quad t > 0, \tag{13}$$

The initial condition of equation (13) is expressed as the average of the sum of the series of the Dirac delta function  $\delta(x - X_i)$  at  $X_i$ :

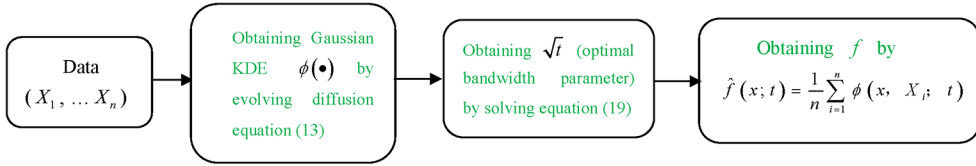
$$\hat{f}(x; 0) = \frac{1}{n} \sum_{i=1}^n \delta(x - X_i) \tag{14}$$

Botev et al. (2010) showed that the diffusion model (13) can be extended on the basis of the smoothing properties of the following linear diffusion equation:

$$\frac{\partial}{\partial t} g(x; t) = Lg(x; t) \tag{15}$$

$L$  in equation (15) is a linear differential operator possessing the form  $\frac{1}{2} \frac{d}{dx} \left( a(x) \frac{d}{dx} \left( \frac{\bullet}{p(x)} \right) \right)$  in which  $a(x)$  and  $p(x)$  are arbitrarily chosen positive functions having bounded second derivatives. The initial condition of equation (15) is expressed as the average of the sum of the series of the Dirac Delta function  $\delta(x - X_i)$  at  $X_i$ :

$$g(x; 0) = \frac{1}{n} \sum_{i=1}^n \delta(x - X_i) \tag{16}$$



**Figure 1.** A flowchart describing procedures for obtaining  $f$  by the proposed new adaptive kernel density estimation method.

Botev et al. (2010) showed that the following relation can be derived by solving equation (15):

$$g(x; t) = \frac{1}{n} \sum_{i=1}^n \kappa(x, X_i; t) \tag{17}$$

The term  $\kappa(x, X_i; t)$  in equation (17) can be asymptotically approximated as follows:

$$\kappa(x, X_i; t) = \frac{p(x)}{\sqrt{2\pi t} [p(x)a(x)a(X_i)p(X_i)]^{1/4}} \exp \left\{ \frac{-1}{2t} \left[ \int_{X_i}^x \frac{p(s)}{a(s)} ds \right]^2 \right\}, \quad t \downarrow 0 \tag{18}$$

Botev et al. (2010) showed that the square value of the asymptotically optimal bandwidth parameter  $\sqrt{t}$  can be calculated as follows:

$$t = \left( \frac{E_f \left[ (\sqrt{a(x)/p(x)})^{-1} \right]}{2n\sqrt{\pi} \|L_f\|^2} \right)^{2/5} \tag{19}$$

In equation (17)  $E$  is the symbol for the expected value in probability theory.  $\|\bullet\|$  is the symbol for the Euclidean distance from the origin. It is easy to see that the value of  $t$  (i.e., the square value of the bandwidth parameter  $\sqrt{t}$ ) varies with the value of  $x$  (the support value). This fact leads us to call this proposed new method an adaptive KDE method formulated on the theory of linear diffusion processes.

The following Figure 1 shows a flowchart describing procedures for obtaining  $f$  by the proposed new adaptive KDE method.

### Theories for the dynamic analysis of an offshore sustainable energy system

WEC Simulator (WEC-Sim), an open-source software for simulating WECs has been utilized in this study. WEC-Sim is developed in MATLAB/SIMULINK using the multi-body dynamics solver Simscape Multibody. WEC-Sim has the ability to model devices that are comprised of hydrodynamic bodies, joints and constraints, power take-of systems, and mooring systems. Simulations are performed in the time-domain by solving the governing WEC equations of motion in the six rigid Cartesian degrees-of-freedom. For the dynamic analysis of a specific offshore sustainable energy system (a point-absorber WEC), the following equations of motion can be utilized (Wang (2018, 2019, 2020)):

$$\mathbf{M}_{RB}\ddot{\mathbf{x}}(t) = -\mathbf{A}(\infty)\dot{\mathbf{x}}(t) - \int_0^t \bar{\mathbf{K}}(t-\tau)\dot{\mathbf{x}}(\tau)d\tau + \mathbf{P}_{wave}(t) + \mathbf{P}_{PTO}(t) + \mathbf{P}_{visc}(t) - \mathbf{P}_{hs} \tag{20}$$

In equation (20)  $\mathbf{M}_{RB}$  is a matrix of the WEC inertia.  $\mathbf{x}(t)$  is the linear and angular WEC position.  $\ddot{\mathbf{x}}(t)$  is the (translational and rotational) acceleration vector of the device.  $\mathbf{A}(\infty)$  is the matrix of the device added mass at the infinite frequencies.  $\bar{\mathbf{K}}(t)$  is a matrix of the kernel functions.  $\mathbf{P}_{wave}(t)$  is the wave excitation force and torque (6-element) vector.  $\mathbf{P}_{visc}(t)$  denotes the hydrodynamic viscous damping force and torque vector.  $\mathbf{P}_{PTO}(t)$  is the power-take-off (PTO) force and torque vector, and  $-\mathbf{P}_{hs}$  denotes the net buoyancy restoring force and torque vector. The WEC PTO mechanism is represented as a linear spring-damper system where the reaction force  $\mathbf{P}_{PTO}(t)$  is given by:

$$\mathbf{P}_{PTO}(t) = -K_{PTO}\mathbf{x}_{rel} - C_{PTO}\dot{\mathbf{x}}_{rel} \quad (21)$$

In equation (21)  $K_{PTO}$  denotes the PTO stiffness and  $C_{PTO}$  denotes the PTO damping.  $\mathbf{x}_{rel}$  and  $\dot{\mathbf{x}}_{rel}$  denote the relative motion and relative velocity between the WEC spar and float.

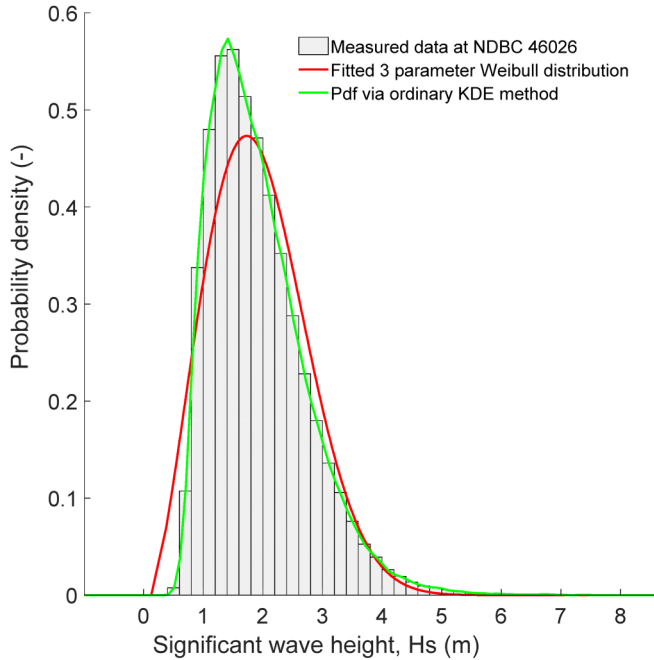
### Calculation examples regarding the $H_S$ probability distribution tails

In the following, we provide our calculation examples regarding the  $H_S$  probability distribution tails based on a wave dataset measured at the US NDBC Station 46026. The NDBC station 46026 is a 3-meter discus buoy with a seal cage whose picture is shown in the following Figure 2. The location of this data buoy is at a sea site 18 nautical miles west of San Francisco, in the US state of California.



**Figure 2.** The picture of the US National Data Buoy Center (NDBC) 46026 station data buoy ([https://www.ndbc.noaa.gov/station\\_page.php?station=46026](https://www.ndbc.noaa.gov/station_page.php?station=46026)).





**Figure 3.** Model fit between the measured data at US National Data Buoy Center (NDBC) 46026 and the two considered models respectively.

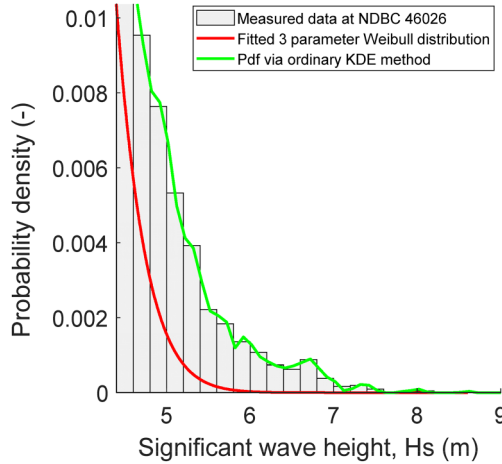
This wave dataset was measured hourly during the time span of January 1, 1996 to December 31, 2021 and contains 210240  $H_S$  values and 210240 corresponding energy periods ( $T_e$ ) values.

The probability density distributions calculated based on the aforementioned  $H_S$  values are presented in Figure 3 in which the red curve shows the calculation results by fitting the aforementioned wave dataset to a 3-parameter Weibull model expressed as follows:

$$f_{H_S}(x_1) = \frac{\beta}{\alpha} \left( \frac{x_1 - \gamma}{\alpha} \right)^{\beta-1} e^{-\left(\frac{x_1 - \gamma}{\alpha}\right)^\beta} \quad (20)$$

The values of the parameters in the above model had been obtained by utilizing a maximum likely method as follows:  $\alpha = 2.0290$ ,  $\beta = 2.3435$  and  $\gamma = 0.1266$ . By examining the calculation results in Figure 3 we can find that the mode of the fitted 3-parameter Weibull probability density distribution is much lower than that of the histogram of the measured ocean wave dataset. In order to investigate the performances of the 3-parameter Weibull model for fitting the probability density distribution tails, a zoom-in plot had been made and presented in Figure 4. By examining the calculation results in Figure 4 we can find that in the tail region the fitted 3-parameter Weibull probability density distribution is also much lower than the histogram of the measured ocean wave dataset. These calculation results clearly demonstrate that the parametric 3-parameter Weibull model is a poor choice for fitting the  $H_S$  probability density distribution.

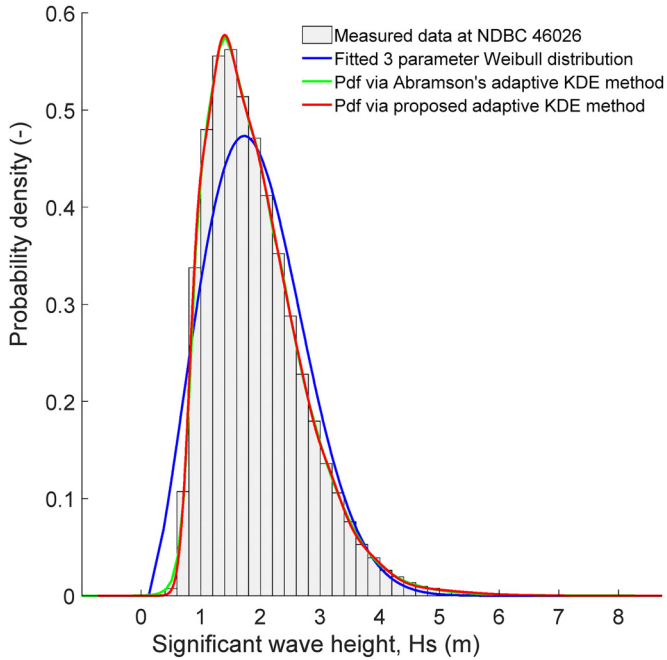
The green distribution curve in Figure 3 is the calculation results obtained by fitting the aforementioned wave dataset using the ordinary KDE method.



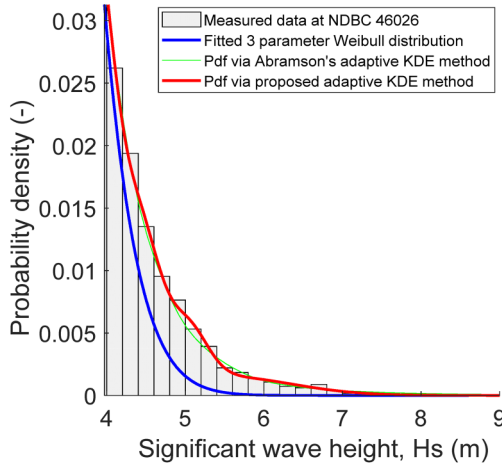
**Figure 4.** Model fit at the tail region between the measured data at US National Data Buoy Center (NDBC) 46026 and the two considered models respectively.

By examining the calculation results in Figure 3 we can find that the mode of the ordinary KDE probability density distribution almost coincides with that of the histogram of the measured ocean wave dataset. In order to investigate the performances of the ordinary KDE method for fitting the probability density distribution tails, the aforesaid zoom-in plot in Figure 4 was once again utilized. This time, by examining the calculation results in Figure 4 we can find that in the tail region the ordinary KDE distribution curve is zigzagged. These calculation results clearly demonstrate that the ordinary KDE method is also a poor choice for fitting the  $H_S$  probability density distribution.

Realizing that accurate results of the distribution tails cannot be predicted by using the ordinary KDE method, Abramson's adaptive KDE method had subsequently been used in this study in order to obtain better results for the probability density values at the tail region. The corresponding calculation results have been obtained by running a MATLAB program and summarized in Figures 5 and 6. By examining the calculation results in Figure 5 we can find that the mode of Abramson's adaptive KDE probability density distribution (the green curve) almost coincides with that of the histogram of the measured ocean wave dataset. In order to investigate the performances of Abramson's adaptive KDE method for fitting the probability density distribution tails, the zoom-in plot in Figure 6 was utilized. This time, by examining the calculation results in Figure 6 we can find that in the tail region Abramson's adaptive KDE distribution curve becomes very smooth and fits well with the histogram of the measured ocean wave dataset. However, it had been very computationally expensive (costing more than 2128 s) to run the MATLAB program to obtain the green curve in Figure 5 on a Dell Precision 5820 high-performance tower desktop workstation. This clearly indicates that Abramson's adaptive KDE method is not a good choice when being used for some time-constrained real-world practical engineering projects. With the expectation of further improving the computational accuracy and efficiency, we have resorted to our proposed new adaptive KDE method formulated on the theory of linear diffusion processes to predict the  $H_S$  probability density distribution tails. Our calculation had been based on the same set of measured wave data mentioned previously, and our computational results are presented in Figure 5 by the red curve of the probability density distribution.



**Figure 5.** Model fit between the measured data at US National Data Buoy Cente (NDBC) 46026 and the three considered models respectively.



**Figure 6.** Model fit at the tail region between the measured data at US National Data Buoy Cente (NDBC) 46026 and the three considered models respectively.

By examining the calculation results in Figure 5 we can find that the mode of our proposed new adaptive KDE probability density distribution (the red curve) almost coincides with that of the histogram of the measured ocean wave dataset. In order to investigate the performances of our

proposed new adaptive KDE method for fitting the probability density distribution tails, the zoom-in plot in Figure 6 was utilized. This time, by examining the calculation results in Figure 6 we can find that in the tail region, our proposed new adaptive KDE distribution curve becomes very smooth and fits quite well with the histogram of the measured ocean wave dataset. If we take a closer look at the zoom-in plot in Figure 6, we can also find that the red curve also fits the histogram of the measured ocean wave dataset slightly better than the green distribution curve. Furthermore, it had been very computationally efficient (costing only about 40 s) to run the MATLAB program to obtain the red curve in Figure 5 on the same Dell Precision 5820 high-performance tower desktop workstation mentioned before. Comparing to Abramson's adaptive KDE method, our proposed new adaptive KDE method is about 53 times faster, indicating it is a better choice to capture the sea state parameter distribution tails when being utilized in some time-constrained real-world practical engineering projects.

Up until the present, we have graphically proven the accuracy and effectiveness of our proposed new adaptive KDE method based on linear diffusion processes. Despite the clarity about the accuracy and efficiency of this new method graphically, we continue to prove the accuracy of this new method by using goodness of fit tests such as the Crámer-von Mises statistic. The Crámer-von Mises statistic could be used to assess and compare the fits. In statistics, the Crámer-von Mises is a criterion used for judging the goodness of fit of a cumulative distribution function  $F^*$  compared to a given empirical distribution function  $F_n$ . It is defined as

$$\omega^2 = \int_{-\infty}^{\infty} [F_n(x) - F^*(x)]^2 dF^*(x) \quad (23)$$

Let  $x_1, x_2 \dots x_n$  be the observed values, in increasing order. Then, the Crámer-von Mises test statistic is

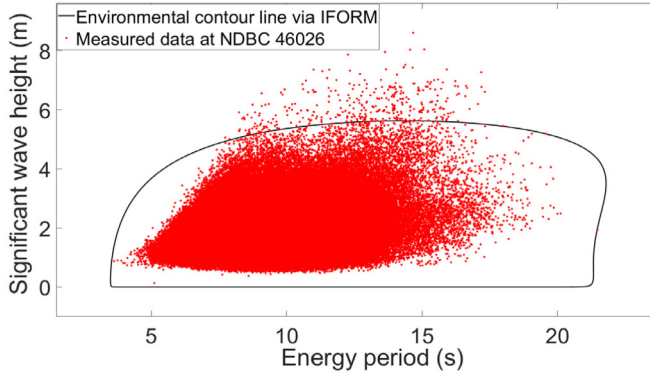
$$T = n\omega^2 = \frac{1}{12n} + \sum_{i=1}^n \left[ \frac{2i-1}{2n} - F^*(x_i) \right]^2 \quad (24)$$

If the Crámer-von Mises test statistic value is larger, then the fit of the cumulative distribution function  $F^*$  to the given empirical distribution function  $F_n$  is poorer. We have written a MATLAB code to calculate the Crámer-von Mises test statistic value. Based on the measured data (the  $H_S$  data in Figures 5 and 6 in this paper), we have calculated and obtained that the Crámer-von Mises test statistic value for the proposed adaptive KDE estimated distribution (the red curve in Figure 5) is 0.0947. Similarly, we have also calculated and obtained that the Crámer-von Mises test statistic value for the fitted 3-parameter Weibull distribution in Figure 5 in this paper is 250.9687. Therefore, we have quantitatively verified that in Figures 5 and 6 in this paper the proposed adaptive KDE estimation fits the data better than the 3-parameter Weibull distribution.

## Calculation examples regarding the 50-year environmental contour lines

In this section, we present the computational results for the 50-year environmental contour lines based on the measured ocean wave data at NDBC 46026.

There are 210240 red "dots" in Figure 7, and each red dot represents a measured  $H_S$  value and the corresponding  $T_e$  value at the NDBC station 46026 mentioned previously. These 210240 hourly measurements of  $H_S$  and  $T_e$  had been taken from January 1, 1996 to December 31, 2021. The black Rosenblatt-IFORM environmental contour line had been derived based on the mathematical



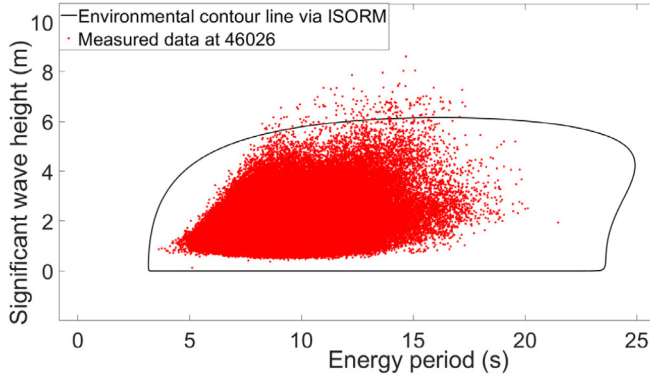
**Figure 7.** The 50-year environmental contour line created by the traditional IFORM method for NDBC 46026. IFORM: inverse first-order reliability method; NDBC: US National Data Buoy Center.

theories as elucidated in “The theoretical backgrounds of the IFORM and ISORM environmental contour line approaches” section. Specifically, the first step in deriving the black contour was to fit the measured 210240  $H_S$  values with a 3-parameter Weibull probability density distribution and also to fit the measured 210240  $T_e$  values with a conditional lognormal distribution. Next, equations (1–4) in “The theoretical backgrounds of the IFORM and ISORM environmental contour line approaches” section were applied for obtaining the black Rosenblatt-IFORM environmental contour in Figure 7.

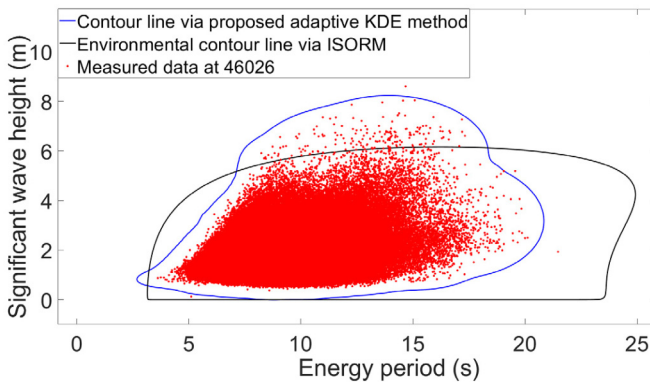
Common sense would be that a 50-year environmental contour should contain most of the measured 26-year wave data at the NDBC station 46026 mentioned previously. However, the black Rosenblatt-IFORM environmental contour line is unexpectedly exceeded by a large quantity of measured data points having high  $H_S$  values. This indicates that the black Rosenblatt-IFORM contour will provide unconservative results and lead to the design of unsafe offshore sustainable energy systems.

Similarly, there are 210240 red “dots” in Figure 8, and each red dot represents a measured  $H_S$  value and the corresponding  $T_e$  value at the NDBC station 46026 mentioned previously. These 210240 hly measurements of  $H_S$  and  $T_e$  had been taken from January 1, 1996 to December 31, 2021. The black Rosenblatt-ISORM contour in Figure 8 had been derived based on the equations (1–5) in “The theoretical backgrounds of the IFORM and ISORM environmental contour line approaches” section. Theoretically, the Rosenblatt-ISORM method is more conservative than the Rosenblatt-IFORM method, and this is confirmed by the fact that in Figure 8 the ISORM contour has missed less of the measured wave data (comparing to the Rosenblatt-IFORM contour in Figure 7). However, taking a closer look at Figure 8 reveals that the Rosenblatt-ISORM environmental contour line is still exceeded by some measured wave data points having high  $H_S$  values. This indicates that the Rosenblatt-ISORM environmental contour line will also provide unconservative results and lead to the design of unsafe offshore sustainable energy systems.

Realizing that accurate results of the environmental contour lines cannot be predicted by using the Rosenblatt-IFORM method and the Rosenblatt-ISORM method based on fitted parametric distributions, our proposed new adaptive KDE method had subsequently been resorted to in this study in order to obtain better results for the 50-year environmental contour lines. The corresponding



**Figure 8.** The 50-year environmental contour line created by the traditional ISORM method for NDBC 46026. ISORM: inverse second-order reliability method; NDBC: US National Data Buoy Center.



**Figure 9.** 50-year extreme sea state contours created by the traditional method and the proposed new method presented in this paper for US National Data Buoy Center (NDBC) 46026.

calculation results have been obtained by running a MATLAB program and presented in Figure 9 as the blue environmental contour line.

We recall that in Figures 5 and 6 we have already shown that our proposed new adaptive KDE method formulated on the theory of linear diffusion processes can forecast well the  $H_S$  (or  $T_e$ ) probability distribution tail, and because in Figure 9 the blue environmental contour line was directly derived based on the nonparametric KDE  $H_S$  and  $T_e$  probability distributions, we can infer that the blue environmental contour is definitely more accurate than the black Rosenblatt-ISORM contour derived based on the parametric  $H_S$  and  $T_e$  models with inaccurate probability distribution tails.

This indicates that the blue environmental contour directly derived based on our proposed new adaptive KDE method formulated on the theory of linear diffusion processes would be superior to the black Rosenblatt-ISORM contour Rosenblatt-ISORM and lead to the design of safe and reliable offshore sustainable energy systems.

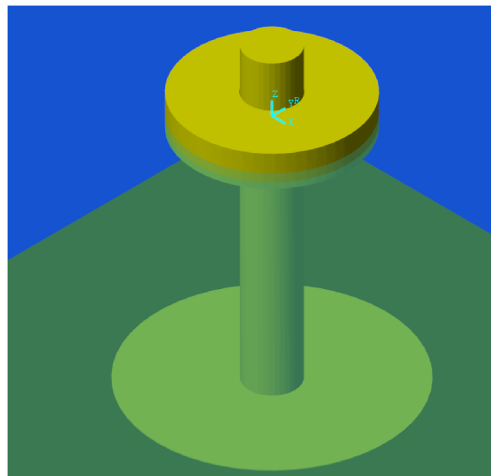
## The extreme dynamic responses of a specific offshore sustainable energy system

### *The selected offshore sustainable energy system*

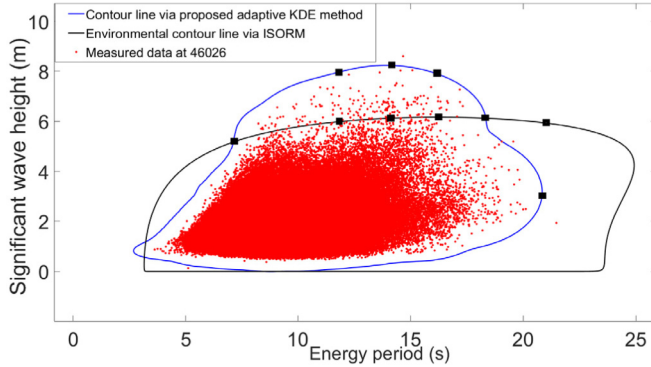
A specific offshore sustainable energy system (a two-body floating point absorber WEC) was chosen in this study. We have created the WEC simulation model using the open-source software WEC-Sim (<https://wec-sim.github.io/WEC-Sim/master/introduction/overview.html>) and have shown the picture of this model in Figure 10. WEC-Sim (WEC-Sim) is an open-source software for simulating WECs. It is developed in MATLAB/SIMULINK using the multi-body dynamics solver Simscape Multibody. WEC-Sim has the ability to model devices that are comprised of hydrodynamic bodies, joints, and constraints, power take-of systems, and mooring systems. Simulations are performed in the time domain by solving the governing WEC equations of motion in the six rigid Cartesian degrees-of-freedom.

This specific offshore sustainable energy system contains a spar possessing a diameter of 6 m and a float having a diameter of 20 m. The spar has been designed with a height of 38 m and a mass of 878.3 t. The float has been designed with a thickness of 5 m and a mass of 727.01 t. This specific offshore sustainable energy system is installed at an ocean site with a draft of 82 m. WEC-Sim dynamic simulations have been executed so that we could forecast the 50-year extreme PTO heaving forces of this floating WEC in six chosen fully developed sea states characterized with a Pierson-Moskowitz ocean wave spectrum. Six  $T_e$  values characterizing these fully developed sea states had been selected along each of the two contours shown in Figure 11: 7.191, 11.792, 14.131, 16.194, 18.305, and 20.942 s. These 6  $T_e$  values have been selected so that the peak  $H_S$  values are included. These selected fully developed sea states are represented by the black solid “squares” in Figure 11 and the corresponding values of the parameters of these sea states are summarized in Table 1.

These selected fully developed sea states characterized by different Pierson-Moskowitz ocean wave spectra were inputted to the WEC-Sim time-domain simulations for obtaining the WEC



**Figure 10.** The WEC-Sim simulation model of the selected WEC.  
WEC: wave energy converter; WEC-Sim: wave energy converter SIMulator.



**Figure 11.** 50-year extreme sea state contours created by the traditional method and the proposed new method presented in this paper for US National Data Buoy Center (NDBC) 46026.

**Table 1.** Values of significant wave heights from each contour line method as shown in Figure 11.

Energy period $T_e$ , s	Spectral peak period $T_p$ , s	Proposed adaptive KDE $H_s$ , m	Rosenblatt ISORM $H_s$ , m
7.191	8.413	5.194	5.194
11.792	13.797	7.931	5.983
14.131	16.533	8.252	6.123
16.194	18.947	7.931	6.162
18.305	21.417	6.123	6.123
20.942	24.502	3.044	5.932

KDE: kernel density estimation; ISORM: inverse second-order reliability method.

dynamic responses based on the theories specified in the “Theories for the dynamic analysis of an offshore sustainable energy system” section. A specific type of dynamic responses, i. e., the 3-h time series of the Power-Take-Off heaving forces have been acquired. These PTO heaving forces are critical for designing the floating WEC. A most probable maximum value corresponding to each 3-h time series of the Power-Take-Off heaving forces was then calculated utilizing the theories elucidated in Edwards and Coe (2019). Repeating this procedure for all the above-mentioned 6  $T_e$  values at each of the two environmental contours shown in Figure 11, all the most probable maximum Power-Take-Off heaving force values had been obtained and summarized in Table 2. Carefully studying the data in column 2 in Table 2 reveals that the largest value among the six most probable maximum power-take-off heaving force values based on the proposed adaptive KDE contour method is 3021700N, and this value is designated as the 50-year extreme Power-Take-Off heaving force value. Similarly, studying the data in column 3 in Table 2 reveals that the 50-year extreme Power-Take-Off heaving force value predicted via the Rosenblatt-ISORM contour method is 2458700N, which is much smaller than the 3021700N value predicted via the new method, and therefore can lead to the design of an unsafe offshore sustainable energy system (in the present study, a floating WEC). These outcomes demonstrate that our proposed new adaptive KDE method formulated on the theory of linear diffusion processes can forecast well and efficiently forecast the 50-year extreme design force values for offshore sustainable energy systems.



**Table 2.** Most probable maximum PTO force values based on the two contour methods.

Spectral peak period $T_p$ , s	Most probable maximum PTO force values based on the proposed adaptive KDE contour method (N)	Most probable maximum PTO force values based on the Rosenblatt ISORM contour method (N)
8.413	$2.4587 \times 10^6$	$2.4587 \times 10^6$
13.797	$3.0217 \times 10^6$	$2.2886 \times 10^6$
16.533	$2.7302 \times 10^6$	$1.7291 \times 10^6$
18.947	$2.1288 \times 10^6$	$1.6529 \times 10^6$
21.417	$1.4162 \times 10^6$	$1.4162 \times 10^6$
24.502	$5.2457 \times 10^6$	$1.1623 \times 10^6$

KDE: kernel density estimation; ISORM: inverse second-order reliability method; PTO: power-take-off.

## Conclusions

In the present research, we have proposed a new adaptive KDE method formulated on the theory of linear diffusion processes. By examining the calculation results we have found that in the tail region, our proposed new adaptive KDE distribution curve becomes very smooth and fits quite well with the histogram of the measured ocean wave dataset at the NDBC station 46026. Carefully studying the calculation results also reveals that the 50-year extreme power-take-off heaving force value forecasted based on the environmental contour derived using the new method is 3021700N, which is much larger than the value 2458700N forecasted via the Rosenblatt-ISORM contour method. Consequently, our proposed new adaptive KDE method formulated on the theory of linear diffusion processes can forecast well and efficiently forecast the 50-year extreme design force values for offshore sustainable energy systems in renewable energy devices development.

## Declaration of conflicting interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China (grant number 51979165).

## ORCID iD

Yingguang Wang  <https://orcid.org/0000-0003-4629-3243>

## References

- Botev ZI, Grotowski JF and Kroese DP (2010) Kernel density estimation via diffusion. *Annals of Statistics* 38(5): 2916–2957.
- Chai W and Leira BJ (2018) Environmental contours based on inverse SORM. *Marine Structures* 60: 34–51.
- Clarindo G, Teixeira AP and Guedes Soares C (2021) Environmental wave contours by inverse FORM and Monte Carlo Simulation with variance reduction techniques. *Ocean Engineering* 228: 108916.

- Eckert-Gallup A and Martin N (2016) Kernel density estimation (KDE) with adaptive bandwidth selection for environmental contours of extreme sea states. Proceedings of the MTS/IEEE OCEANS 2016, Monterey, CA.
- Edwards SJ and Coe RG (2019) The effect of environmental contour selection on expected wave energy converter response. *Journal of Offshore Mechanics and Arctic Engineering* 141(1): 011901.
- Haselsteiner AF, Mackay E and Thoben KD (2021) Reducing conservatism in highest density environmental contours. *Applied Ocean Research* 117: 102936.
- Haselsteiner AF, Ohlendorf JH and Thoben KD (2017) Environmental contours based on kernel density estimation. In: Proceedings of the 13th German wind energy conference (DEWEK 2017), Bremen, 17–18.
- Haselsteiner AF and Thoben KD (2020) Predicting wave heights for marine design by prioritizing extreme events in a global model. *Renewable Energy* 156: 1146–1157.
- Haver S and Winterstein S (2009) Environmental contour lines: A method for estimating long-term extremes by a short-term analysis. *Transactions - Society of Naval Architects and Marine Engineers* 116: 116–127.
- Mackay E and Haselsteiner AF (2021) Marginal and total exceedance probabilities of environmental contours. *Marine Structures* 75: 102863.
- Manuel L, Nguyen PTT, Canning J, et al. (2018) Alternative approaches to develop environmental contours from metocean data. *Journal of Ocean Engineering and Marine Energy* 4: 293–310.
- Ross E, Astrup OC, Bitner-Gregersen E, et al. (2020) On environmental contours for marine and coastal design. *Ocean Engineering* 195: 106194.
- Silverman BW (1986) *Density Estimation for Statistics and Data Analysis*. New York, USA: Chapman and Hall.
- Wand MP and Jones MC (1995) *Kernel Smoothing*. New York, USA: Chapman and Hall.
- Wang YG (2018) A novel simulation method for predicting power outputs of wave energy converters. *Applied Ocean Research* 80: 37–48.
- Wang YG (2019) Efficient prediction of wave energy converters power output considering bottom effects. *Ocean Engineering* 181: 89–97.
- Wang YG (2020) Predicting absorbed power of a wave energy converter in a nonlinear mixed sea. *Renewable Energy* 153: 362–374.
- Wrang L, Katsidoniotaki E, Nilsson E, et al. (2021) Comparative analysis of environmental contour approaches to estimating extreme waves for offshore installations for the Baltic Sea and the North Sea. *Journal of Marine Science and Engineering* 9(1): 1–24.
- Zhao YL and Dong S (2022a) Comparison of environmental contour and response-based approaches for system reliability analysis of floating structures. *Structural Safety* 94: 102150.
- Zhao YL and Dong S (2022b) Design load estimation with IFORM-based models considering long-term extreme response for mooring systems. *Ships and Offshore Structures* 17(3): 541–554.