

Analysing multivariate extreme conditions using environmental contours and accounting for serial dependence

Erik Vanem*

DNV Group Research and Development, Høvik, Norway
Department of Mathematics, University of Oslo, Oslo, Norway

ARTICLE INFO

Keywords:

Extreme environmental conditions
Environmental contours
Multivariate extremes
Serial dependence
Simulate time series
Ocean environment

ABSTRACT

Accurate statistical description of extreme environmental conditions is needed for risk assessment and management of marine structures and is a crucial input to design of any structure that need to withstand loads from environmental forces. Such descriptions are essentially multivariate extreme value problems, where the environmental loads are due to concurrent extreme combinations of several environmental variables. Typically, in coastal and ocean engineering applications, the simultaneous joint behaviour of significant wave height and wave period is of particular interest and is needed to describe the wave loads on marine structures. Environmental contours are often used to explore the extreme wave loads, and essentially consider extreme combinations of simultaneous significant wave height and wave period, and are based on a joint statistical distribution fitted to relevant metocean data. However, typical applications of environmental contours do not account for temporal dependencies in the environmental variables, and this may lead to an overestimation of extreme conditions. In this paper, an approach for partially accounting for serial dependence in the construction of environmental contours is proposed, based on simulating time-series of a primary variable which preserves both its marginal distribution and auto-correlation structure. It is shown that this gives lower estimates of extreme environmental conditions compared to conventional application of environmental contours that do not account for serial dependence. Hence, more accurate description of the extreme environment can be available for design and construction of structures exposed to environmental loads.

1. Introduction and background

Environmental contours are used to identify extreme environmental conditions governed by the simultaneous behaviour of several correlated environmental variables. Typical examples are extremes related to the joint behaviour of waves, wind and current for ocean and coastal engineering applications. Hence, environmental contours are tools for multivariate extreme value analysis often used in long-term extreme response analysis of marine structures, and represent a simple and approximate alternative to more computationally demanding full long-term analyses. One of the main advantages is that the structural response is decoupled from the environmental description so that only a limited number of short term response calculations are required for a long-term analysis. That is, the environmental contours describe critical conditions for which short term analysis is performed [1].

Typically, environmental contours will be based on a set of environmental data, for example metocean data, and the application of environmental contours often consists of four main steps. The first step relates to collecting and preparing such data. This also include

possible validation and pre-processing of the data. The second step is to fit a joint statistical distribution to the data. Often, a hierarchical conditional model will be assumed, where one variable deemed as most influential on the structural response is selected as the primary variable, and the other variables are modelled conditionally on this. This approach has become an industry standard [2]. For example, a three-parameter Weibull distribution is often used for the marginal distribution of significant wave height (H_S) and a conditional lognormal distribution is often used for concurrent wave period [3,4]. Conditional models for several other metocean variables are proposed in [5], the bivariate lognormal distribution is discussed in [6] and copula-based methods are explored in [7–10]. The third step consist of constructing environmental contours based on the fitted distributions and the final step is to apply the environmental contours for extreme response estimation. All the steps are associated with modelling choices and uncertainties that may influence the results.

Multivariate extreme value analysis is not straightforward, and there exist several definitions of what to understand as a multivariate

* Correspondence to: DNV Group Research and Development, Høvik, Norway.
E-mail address: Erik.Vanem@dnv.com.

extreme event. In the bivariate case, a quantile will typically be vector-valued and can be expressed on the following form [11]

$$Q_{X,Y}(p, \epsilon) = \{(x, y) \in \mathbb{R}^2 : F_{\epsilon}(x, y) = p\}, \quad (1)$$

where $F_{\epsilon}(x, y)$ corresponds to some exceedance or non-exceedance event. Different definitions of such events yields different types of extremes. Some examples are illustrated in the bivariate case in [12–14], i.e., cases where one of the variables is extreme, where both are extreme and where the value of one variable is extreme conditioned on the value of the other variable. Environmental contours represent other definitions of multivariate extremes, often described in terms of exceedance hyperplanes. However, there exist different definitions of environmental contours that do not refer to hyperplanes.

Even with a given definition for multivariate extremes, corresponding to a specific choice of $F_{\epsilon}(x, y)$, there will not be a unique solution to the extreme quantile problem, and there is a continuum of solutions to the extreme value problem in the multivariate case. For example, in the bivariate case, if a solution (x_0, y_0) to the extreme value problem exist, one could always find other solutions by slightly perturbing x_0 and y_0 so that Eq. (1) is still fulfilled [15]. Thus, the solution will be a curve in the variable space associated with a non-exceedance probability defined in some way.

Environmental contours are one way of defining multivariate extreme sets. Environmental contours were initially proposed in [16,17] as equi-density lines, and the definition of environmental contours has continued to be an area of active research. Contours based on the inverse first order reliability method (IFORM) were introduced in [18, 19], and remain the most commonly used approach. Such contours are based on exceedance hyperplanes in standard normal space and transformed to physical variable space. More recently, environmental contours defined in terms of exceedance hyperplanes in the physical variable space based on Monte Carlo sampling were proposed in [20, 21]. These contours, often referred to as direct sampling contours, have well-defined probabilistic properties, and will in some cases be quite different from the IFORM-contours, see the comparison studies in [22,23]. In particular, the direct sampling contours will always be convex [24]. Other approaches for defining environmental contours exist [e.g.,15,25–31], and several studies have investigated uncertainties associated with different aspects of environmental contours [23,32–35]. In this paper, direct sampling contours will be used, which correspond to exceedance regions in the form of exceedance hyperplanes. According to common industry practice, environmental contours are often based on a fitted joint distribution, fitted to all data and ignoring serial correlation [2,36].

Most applications of environmental contours are for two-dimensional problems, but extensions to higher dimensions are in principle straightforward, see [37–39]. However, the study presented in this paper is restricted to two dimensions and considers environmental contours for significant wave height (H_S) and zero up-crossing wave period (T_Z).

Given that environmental contours have been an active area of research in recent years, a benchmark study was recently announced where practitioners and researchers of environmental contours were provided with metocean data and invited to construct contours for comparison [40]. Several contributions were submitted for this exercise; results are presented in [41]. One important observation from this benchmarking exercise is that serial correlation in the metocean variables is often neglected in the construction of environmental contours, and that this may introduce notable biases. This point was further elaborated in [42], where it was shown that neglecting serial dependence will lead to over-estimation of return values. Hence, this paper offers an approach to take serial correlation into account in the construction of environmental contours based on the direct sampling approach. This is done by establishing a statistical model that accounts for serial dependence in the primary variable, which will yield more accurate estimates of the extremes.

In univariate extreme value analysis, the effect of serial correlation is often accounted for by modelling the distribution of peak values, either using a block-maxima or peaks-over-threshold (POT) approach [43]. For the latter method, de-clustering techniques must be applied to ensure that the samples from the tail are reasonably independent and identically distributed (IID) (see e.g. [43,44]). However, with limited amount of data, such approaches are wasteful since it only uses a very few extreme data points in the analysis and disregard most of the data. With the statistical model presented in this paper, all the data can be exploited and the effect of serial dependence is accounted for to reduce the bias. It is believed that this represents an advancement in state of the art, with the benefit of more accurate extreme value estimates.

Identifying peak events is not straightforward in the multivariate case, where combinations of the variables that may not be extreme in any one variable may be regarded as extreme when their joint probability is considered. Indeed, environmental contours typically range over both extreme and non-extreme values of all variables involved (including extremely small values). Moreover, the interest is in environmental loads and responses, which means that the concurrent behaviour of the environmental variables is needed. Hence, some multivariate extreme value methods based on component-wise extremes will not be applicable [45,46]. Some methods, such as the bivariate ACER model [47], explicitly account for the serial correlation in bivariate data.

An approach for defining environmental contours that accounts for serial correlations in the observations was recently proposed in [48]. The method does not require fitting a joint distribution model. Since defining exceedance regions in terms of hyperplanes is equivalent to defining univariate exceedance regions under rotations of the axes [49], the problem of estimating environmental contours can be reduced to a series of univariate problems under various rotations of the axes. A standard univariate POT approach with de-clustering is then applied to the rotated data, to account for serial correlations.

This paper proposes an alternative approach to take serial correlation into account in the construction of environmental contours, based on a direct sampling approach. This approach to construct environmental contours is based on Monte Carlo samples from the joint distribution of environmental variables, where the data is projected in selected directions and hyperplanes are defined in all directions based on the target exceedance probability. The environmental contours are then constructed based on the intersection points of these hyperplanes; see [20,21] for details. Hence, this paper proposes to construct environmental contours on simulated data from a time-series model that preserves the marginal distribution and the serial correlation in the data. A conditional modelling approach, in line with current industry practice, will be employed where the zero up-crossing wave period is modelled conditionally on significant wave height, but significant wave height will be modelled by a time-series model with a fitted marginal distribution and temporal correlation rather than independently from the marginal distribution. In this way, environmental contours that accounts for serial correlation of significant wave height can be constructed. The main challenge with this approach is to establish a statistical model that allows simulation from such a time series with the desired properties.

The remainder of the paper is organized as follows: Section 2 reviews how return periods and return values are defined, and discusses the biases caused by neglecting serial correlation. Section 3 presents the metocean data used in this study and Section 4 presents the statistical models fitted to the data. This includes the time-series model with desired marginal and correlation structure for H_S as well as the conditional model for T_Z . Section 5 presents the resulting contours, and a discussion of the results are given in 6. Finally, Section 7 gives a summary and some concluding remarks.

2. Return periods and return values for serially correlated data

Informally speaking, the return period of a value of a parameter is the average time between exceedances of that value. The value corresponding to a specific return period is called the return value. To understand the effect that serial correlation has on estimates of return periods and return values, consider two variables X and \tilde{X} , with the same marginal distribution, F_X , but where observations of X have some level of serial correlation and observations of \tilde{X} are independent. Serial correlation effectively reduces the number of independent observations in a sequence of a given length, and therefore reduces the probability of large values. Since both X and \tilde{X} have the same marginal distribution, but there is a lower probability of large values in X in a given period, we see that defining return values in terms of the marginal distribution of all observations, F_X , cannot be correct when serial correlation is present.

For serially correlated data return periods can be defined in terms of the distribution of peak values in the time series [43]. Suppose that local peaks, X_p , have been identified in a time series, using some declustering criteria, such that consecutive observations of X_p are approximately independent. Then the return period of level x is

$$T(x) = \frac{1}{m [1 - F_{X_p}(x)]}, \quad mT > 1, \tag{2}$$

where F_{X_p} is the distribution of the peaks and m is the average number of peaks per year. This definition is also applicable when annual maxima are used instead of local peaks. In this case F_{X_p} is replaced by the distribution of annual maxima and we set $m = 1$. Alternatively, the T -year return value can be defined as the solution of [50]

$$F_T(x) = \left(1 - \frac{1}{T}\right)^T \approx \exp(-1), \tag{3}$$

where F_T is the distribution of the maximum value of X in a T -year period. It can be shown that definitions (2) and (3) are asymptotically equivalent (see e.g. [42,50]). For the sequence \tilde{X} , since all observations are independent, the return value of level x can be defined using (2) as

$$\tilde{T}(x) = \frac{1}{n [1 - F_X(x)]}, \quad nT > 1, \tag{4}$$

where n is the (fixed) number of observations per year.

It was shown in [43] that the ratio of return periods for the independent and dependent sequences is less than or equal to 1. That is

$$\theta_T(x) = \frac{\tilde{T}(x)}{T(x)} \leq 1, \tag{5}$$

The ratio θ_T quantifies the bias introduced when (4) is used to define return values for serially correlated data. It can be interpreted as a sub-asymptotic extremal index, which converges to the usual extremal index, which by definition is an asymptotic quantity, used to quantify the effect of serial correlation in extreme events (see [43,51] for details).

3. Metocean data description

The data used in this study are one of the datasets provided for the recent benchmarking study on environmental contours, i.e., dataset A, from [40]. Several environmental contours established based on these data are presented in [41]. The data contains hourly observations of significant wave height (H_S) (in m) and zero upcrossing wave period (T_Z) (in s) over a 10-year period from 1996 to 2005. It is noted that 10 years of data is rather short for robust extreme value estimation, and uncertainties will be large see e.g. the discussions in [52,53].

A scatterplot, trace plots and empirical densities for these data are shown in Fig. 1, along with summary statistics in Table 1. These plots and summary statistics reveal that H_S and T_Z are positively correlated,

Table 1
Summary statistics for the metocean data.

	H_S (m)	T_Z (s)	(H_S, T_Z)
Minimum	0.0981	2.3104	
Median	0.7702	5.0742	
Mean	0.9444	5.3409	
0.90 quantile	1.6902	7.3226	
0.99 quantile	3.4496	9.3891	
0.99989 quantile	6.4867	12.6341	
Maximum	7.0994	13.1326	
Standard deviation	0.6419	1.4195	
Skewness	2.4696	0.8619	
Kurtosis	12.6331	3.7387	
Pearson (linear) correlation coefficient r			0.3132
Kendall rank correlation coefficient τ			0.1642
Spearman rank correlation coefficient ρ			0.2352

Table 2
10 year return value estimates from marginal extreme value analyses.

	H_S (m)	T_Z (s)
10 year return value (GPD)	7.1624	13.5059
10 year return value (GEV)	6.9630	13.0208

and that both marginal distributions are right skewed and leptokurtic, i.e., that they have fatter tails than a normal distribution. The 0.99989 quantile of the data would correspond to the one-year return value if not considering serial dependence, i.e. $1 - 1/(365.24 * 24) \approx 0.99989$. The traceplots illustrate that there are notable serial dependencies in the data. For further details and description of the data, reference is made to [40,41]. For reference, results from two crude marginal extreme value analyses are shown in Table 2, where the 10-year return value is estimated for both variables. The first is obtained by fitting a generalized Pareto distribution to peaks over threshold data with a threshold of, respectively, 5 m for H_S and 11 s for T_Z and a cluster separation distance of three days (denoted as GPD in the table). The second is based on fitting a generalized extreme value distribution to semi-annual maxima (denoted as GEV in the table).

4. Statistical modelling

For the purpose of structural reliability analysis and design of marine structures, the joint distribution of governing environmental variables such as significant wave height and wave period is often modelled by a conditional model. Thus, one variable is considered the primary variable, typically the one with the largest influence on the environmental load. A joint distribution is constructed based on fitting a marginal distribution to the primary variable, and then conditional models to the other relevant variables [3–5]. Such a model was fitted to the data used in this study, as reported in [54], but without considering serial correlation.

In order to be able to construct environmental contours that account for the serial correlation in the data using the direct sampling approach, one may fit a time series model with appropriate marginal distribution and temporal correlation structure. This is not straightforward, and there are different approaches that could be used [55]. For example, [56] proposes a diffusion-type model with given marginal distribution and autocorrelation function, and [57] describes a diffusion model with Weibull marginals for modelling wind speed. Models based on stochastic differential equations are presented in [58,59]. Weibull and gamma autoregressive processes are presented in [60], and Gaussian and non-Gaussian autoregressive models are discussed in [61]. However, in this paper, an approach based on transforming a Gaussian parent autoregressive time series to have the desired marginal distribution and temporal correlation structure will be taken, as suggested in [62]; see also [63–65]. In the following, a recipe for constructed such a model is outlined. A joint statistical model is

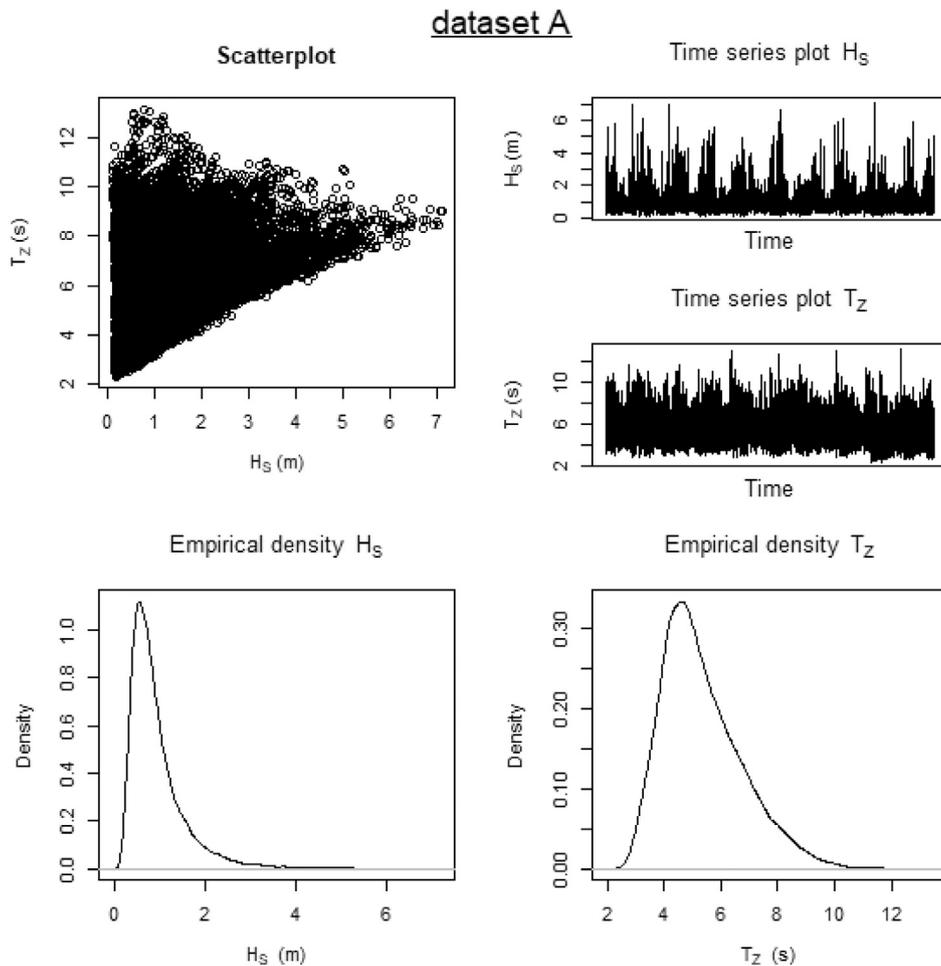


Fig. 1. Plots illustrating the 10-year metocean data used in this study; Scatterplot (top left), trace plots (top right) and empirical densities (bottom).

constructed by fitting a time-series model that preserves the marginal distribution and the correlation structure for the primary variable – significant wave height – and a conditional model for zero up-crossing wave period conditioned on significant wave height. It should be noted that with this approach, the serial correlation of T_Z is not modelled directly but is a result of a conditional distribution conditioned on H_S and would only be partially preserved.

The time-series of interest, $X(t)$, has a marginal distribution, $F_X(x)$ and an autocorrelation function, $\rho_X(\tau)$ and it is assumed that this has a parent Gaussian time-series, $Z(t)$ with a standard Gaussian marginal $\Phi_Z(z)$ and another autocorrelation function $\rho_Z(\tau)$. Hence, modelling and simulation of the time-series $X(t)$ correspond to modelling and simulating the parent Gaussian time-series and finding the transformations that transform the parent time-series to the desired one, i.e., determine functions g and \mathcal{T} such that

$$\begin{aligned} X(t) &= g\{Z(t)\} \\ \rho_Z(\tau) &= \mathcal{T}[\rho_X(\tau)] \end{aligned} \tag{6}$$

Simulating from a Gaussian time-series is straightforward, and the normal distribution have several well-known properties (see e.g. [62,66]). Hence, if one can find a parent Gaussian time-series that can be transformed into the time-series $X(t)$, simulation of $X(t)$ is straightforward.

As pointed out in [62], g is easily identified as $g(Z) = F_X^{-1}(\Phi_Z(Z))$, which will give the desired marginal distribution. However, this transformation of the marginal distribution does not preserve the autocorrelation structure, which depends on the marginal distribution $F_X(x)$. Hence, one needs to establish the autocorrelation structure of the parent Gaussian time-series, $\rho_Z(\tau)$ that yields the required autocorrelation structure for the target time-series $X(t)$ after transformation.

In the bivariate case, it can be shown that for an arbitrary transformation, the correlation between the transformed variables will be smaller than the correlation between the initial standard normal variables (see [62]), i.e.,

$$\rho_X = \text{cor}[g\{Z(t_1)\}, g\{Z(t_2)\}] \leq \text{cor}[Z(t_1), Z(t_2)] = \rho_Z, \tag{7}$$

meaning that the ρ_Z values needs to be inflated to obtain the target ρ_X . Moreover, the correlation coefficient of two lagged variables $X(t_1)$ and $X(t_2)$ can be expressed as

$$\rho_X(\tau) = \frac{E\{X(t_1)X(t_2)\} - \mu_X^2}{\sigma_X^2}, \tag{8}$$

where E is the expectation operator, $\tau = t_2 - t_1$ and μ_X and σ_X are the mean and the standard deviation of X . Now, assuming that $X(t) = g\{Z(t)\} = F_X^{-1}[\Phi_Z\{Z(t)\}]$,

$$\begin{aligned} E\{X(t_1)X(t_2)\} &= E\{F_X^{-1}(\Phi_Z[Z(t_1)])F_X^{-1}(\Phi_Z[Z(t_2)])\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_X^{-1}\{\Phi_Z(u)\} F_X^{-1}\{\Phi_Z(v)\} \\ &\quad \times \phi_{u,v}[u, v; \rho_Z(t_2 - t_1)] dudv, \end{aligned} \tag{9}$$

where $\phi(u, v; \rho)$ denotes the density function of the bivariate standard normal distribution with correlation ρ . This gives a relationship between $\rho_Z(\tau)$ and $\rho_X(\tau)$ and can be used to calculate $\rho_X(\tau)$ from known $\rho_Z(\tau)$, i.e.

$$\rho_X(\tau) = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_X^{-1}(\Phi_Z(u))F_X^{-1}(\Phi_Z(v))\phi_{u,v}(u, v; \rho_Z(\tau))dudv - \mu_X^2}{\sigma_X^2}. \tag{10}$$

Note however, that in order to model the time series with known $\rho_X(\tau)$, one needs the inverse transformation to find $\rho_Z(\tau)$ from $\rho_X(\tau)$. Even though the double integral above does not have an analytical solution in general, its numerical estimation is straightforward.

Hence, the modelling problem is solved by defining a correlation transformation function to estimate the autocorrelation of the parent Gaussian time-series from a given target autocorrelation structure that may be estimated from the data. The following parametric form of such an autocorrelation transformation function is proposed in [62]:

$$\rho_Z = \mathcal{T}(\rho_X) = \frac{(1 + \zeta \rho_X)^{1-\eta} - 1}{(1 + \zeta)^{1-\eta} - 1}. \tag{11}$$

This can be used to estimate a parametric autocorrelation function of the parent Gaussian time-series, $\rho_Z(\tau)$ based on a parametric autocorrelation function for the target time-series, $\rho_X(\tau)$. The parameters ζ and η can be estimated by calculating ρ_X for $\rho_Z = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95)$ according to Eq. (10) and then conducting a least-squares minimization of the assumed model (11) to the set of (ρ_X, ρ_Z) -points (see [62] for details).

With this approach, modelling a time series with desired marginal and correlation structure involves fitting a marginal parametric distribution and estimating the empirical autocorrelation structure from the time-series, fitting the parametric autocorrelation transformation function to obtain the autocorrelation structure of the parent Gaussian time-series, fit an autoregressive model, for example by solving the Yule Walker equations [e.g., 67], to generate the Gaussian parent time-series and transforming the Gaussian time-series using the transformation $X(t) = g[Z(t)] = F_X^{-1}[\Phi\{Z(t)\}]$.

Hence, in the following, environmental contours for significant wave height (H_S) and zero up-crossing wave period (T_Z) will be based on simulated time-series for H_S with prescribed marginals and autocorrelation structure and concurrent values for T_Z simulated as independent values from a conditional model for T_Z given H_S . Therefore, any serial correlation in T_Z is assumed to be, at least partly, accounted for by the serial correlation in H_S and the dependence structure between H_S and T_Z . The validity of this assumption is discussed in Section 6.

4.1. Joint distribution

First, a joint distribution will be fitted to the data disregarding the serial correlation. In this study, a conditional model with a three-parameter Weibull distribution for H_S and a conditional log-normal model for concurrent T_Z will be assumed. That is,

$$f_{H_S, T}(h_s, t) = f_{H_S}(h_s) f_{T_Z|H_S}(t_z|h_s) \tag{12}$$

where

$$f_{H_S}(h_s) = \frac{\beta}{\alpha} \left(\frac{h_s - \gamma}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{h_s - \gamma}{\alpha}\right)^\beta\right], \tag{13}$$

$$f_{T_Z|H_S}(t_z|h_s) = \frac{1}{\sigma(h_s)t_z\sqrt{2\pi}} \exp\left[-\frac{\{\ln t_z - \mu(h_s)\}^2}{2\sigma(h_s)^2}\right] \tag{14}$$

and

$$\begin{aligned} \mu(h_s) &= E(\ln T_Z|H_S = h_s) = a_1 + a_2 h_s^{a_3}, \\ \sigma(h_s) &= \text{std}(\ln T_Z|H_S = h_s) = b_1 + b_2 e^{b_3 h_s}. \end{aligned} \tag{15}$$

This is the same model that was used in [54], where environmental contours were established for this dataset without considering serial correlation. Fitted model parameters are presented in Table 3. It is noted that this particular model was fitted to the data by way of minimizing the second order Anderson–Darling statistic. This is emphasizing on the upper tail, and may not fit the body of the distribution as well as more well-known methods as the maximum likelihood or the method of moments (see e.g. [44]).

4.2. Temporal correlation structure

The temporal correlation structure is estimated from the empirical time series. Then, a number of alternative parametric autocorrelation functions are fitted to this, and the one providing the best fit is selected for further analysis. There are many parametric autocorrelation functions to choose from, see e.g. [62] and in this study, six different parametric functions are tried out, i.e. the Weibull, the Burr, the logarithmic, the fGn, the generalized fGn (gfGn) and the Pareto autocorrelation functions. The empirical autocorrelation function is shown together with the fitted parametric candidate models in Fig. 2. The functions are fitted by least squares, and the residual sum of squares (RSS) are indicated for each candidate model in the plots. It is observed that the 3-parameter Burr-type autocorrelation function fits best to the data, and this would typically be preferred. However, it is noted that this autocorrelation function is known to sometimes yield not positive definite covariance matrices (see [68]), and therefore this model may not always be appropriate. The preferred autocorrelation function must be determined on a case-by-case manner, according to how well it fits the data, and it is also possible to use the non-parametric empirical autocorrelation function. The parametric forms of the autocorrelation functions explored in this study are given in Eq. (16).

$$\begin{aligned} \rho_{Burr}(\tau; \zeta, \eta, \theta) &= \left(1 + \theta \left(\frac{\tau}{\zeta}\right)^\eta\right)^{-\frac{1}{\theta\eta}} \\ \rho_{Weibull}(\tau; \zeta, \eta) &= e^{-\left(\frac{\tau}{\zeta}\right)^\eta} \\ \rho_{Log}(\tau; \zeta, \eta) &= \left(1 + \ln\left(1 + \eta \frac{\tau}{\zeta}\right)\right)^{-\frac{1}{\eta}} \\ \rho_{fGn}(\tau; \zeta) &= \frac{1}{2} (|\tau - 1|^{2\zeta} - 2|\tau|^{2\zeta} + |\tau + 1|^{2\zeta}) \sim \tau^{-2(1-\zeta)} \\ \rho_{gfGn}(\tau; \rho_1, \zeta) &= \zeta(2\zeta - 1) \left\{ \tau - 1 + \left[\frac{\rho_1}{\zeta(2\zeta - 1)}\right]^{-\frac{1}{2(1-\zeta)}} \right\}^{-2(1-\zeta)} \sim \tau^{-2(1-\zeta)} \\ \rho_{Pareto}(\tau; \zeta, \eta) &= \left(1 + \eta \frac{\tau}{\zeta}\right)^{-\frac{1}{\eta}} \end{aligned} \tag{16}$$

It can be observed that most of the parametric autocorrelation functions fit reasonably well to the data for lags up to 100, except for two of the parametric alternatives. It is also noted that the choice of autocorrelation function may influence return value estimates, as seen in [65], so some efforts should be made to find the best possible fit to the data. However, if a good fit cannot be found, it will be possible to simply assume the non-parametric autocorrelation observed in the data.

4.3. Simulated time-series

Having estimated all model parameters, one may simulate time-series from the statistical model corresponding to any desired number of years. The simulated data would then have the desired marginal distribution and the desired autocorrelation structure. One may also simulate from the marginal distribution assuming IID. Fig. 3 shows about 1 year of empirical data for H_S and compares to simulated data based on the time-series model and simulated data assuming IID. The bottom plot in Fig. 3 contains first IID simulated data corresponding to half the empirical data duration and then simulated time-series data for the remainder of the empirical data length. It is clearly seen that the data simulated when accounting for the serial correlation better resemble the observed data.

One may also compare the densities of simulated data with the density of the empirical data, as shown in Fig. 4, where 500 time series of 10 years' length are simulated, assuming IID and with the

Table 3
Fitted model parameters for the joint distribution of H_S and T_Z .

α	β	γ	a_1	a_2	a_3	b_1	b_2	b_3
0.4983	0.8573	0.4187	1.4306	0.2561	0.5556	0.0150	0.3004	-0.2884

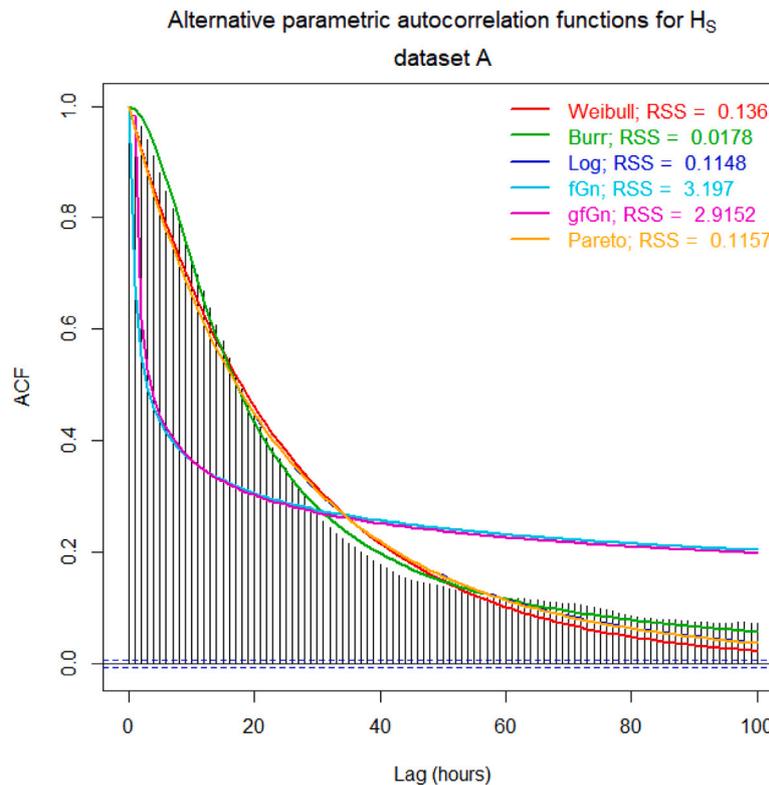


Fig. 2. Fit of alternative autocorrelation functions.

time-series model that preserves the autocorrelation, respectively, and the simulated densities are compared to the empirical density of the observed data. The densities have been fitted based on a goodness-of-fit measure that focuses on the upper tails of the distribution, i.e., the second order Anderson–Darling statistic [44,69], and therefore there is a notable mismatch for small values of H_S . However, since the upper tail of the significant wave height distribution is more important for structural reliability of marine structures, the poor fit on the lower tail is acceptable. It is noted, however, that better fit of this part of the distribution could presumably be obtained if other fitting methods were employed, and this is recommended in applications where the lower tail of the distribution is of importance. Exceedance probabilities (on logarithmic scale) are also compared to highlight tail behaviour. In the upper plots, the densities and exceedance probabilities are estimated from the complete simulated datasets of 5000 years, whereas the lower plots show the densities and exceedance probabilities for individual 10-year periods of simulated data. The top plots clearly demonstrate that the marginal distributions are the same, regardless of whether data are simulated IID from the fitted distribution or from the time-series model with the same marginal distribution. However, there are some differences in the simulated individual 10-year datasets, due to sampling effects.

The sampling uncertainty associated with the largest observation in a sample is very large. As the observations in a sample are random quantities, the return periods associated with each observation are also random quantities. It can be shown that a $1 - 2\alpha$ confidence interval (CI) for the return period associated with the largest observation in an N -year time series tends to $(-N/\log(\alpha), -N/\log(1 - \alpha))$ as $N \rightarrow \infty$ [70], and the approximation is still reasonable for $N = 10$ years. So, for individual 10-year time series, a 95% confidence interval for the

return period associated with the largest observation would be (2.7, 395) years. The width of the confidence interval decreases as the ratio N/T increases, where N is the length of the time series and T is the target return period (see [70] for details).

Fig. 5 shows the distribution of the 10-year maxima from the 500 simulated 10-year time-series. The 10-year return values estimated from these distributions are also shown, where the return values are estimated as the $(1 - 1/10)^{10} \approx 0.3487$ quantiles of the maximum-value distributions according to Eq. (3). It is observed that the return values for the 10-year maxima are higher for the IID data compared to the serially-correlated time series. The vertical lines for the combined data correspond to ignoring serial correlation and estimating the return value according to Eq. (4) as the quantile from the combined simulated datasets from the IID and TS simulations, respectively. It is observed that these generally agrees with the return value estimate from the distribution of 10-year IID datasets, apart from small effects due to sampling variability, and the difference can hardly be recognized in the figure. It is also interesting to compare the return value estimates from the simulated data with the summary statistics and the estimated return values from classical extreme value methods presented in Tables 1–2: 10-year return value estimated from simulations assuming independent and identically distributed data is 8.967 m whereas 10-year return value estimate from simulated time series is 7.229 m. It is observed that the return value estimated based on the simulated data from the time series model is much closer to the return value estimates from the classical extreme value analysis, and also closer to the maximum value observed in the data. Hence, the positive bias from ignoring serial correlation is avoided.

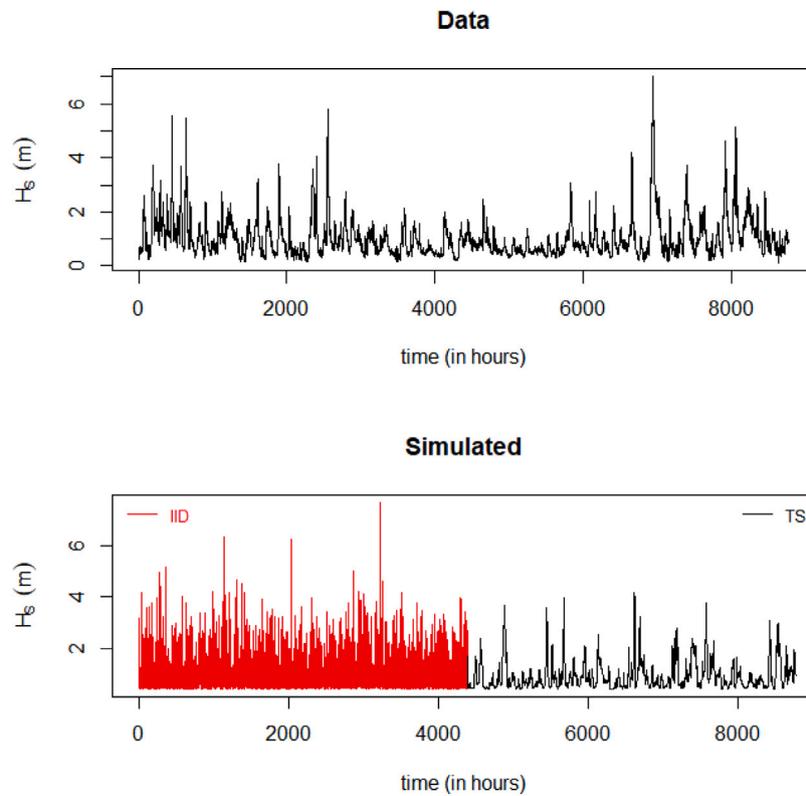


Fig. 3. Empirical (top) and simulated (bottom) time-series of significant wave height. Index represents time in hours.

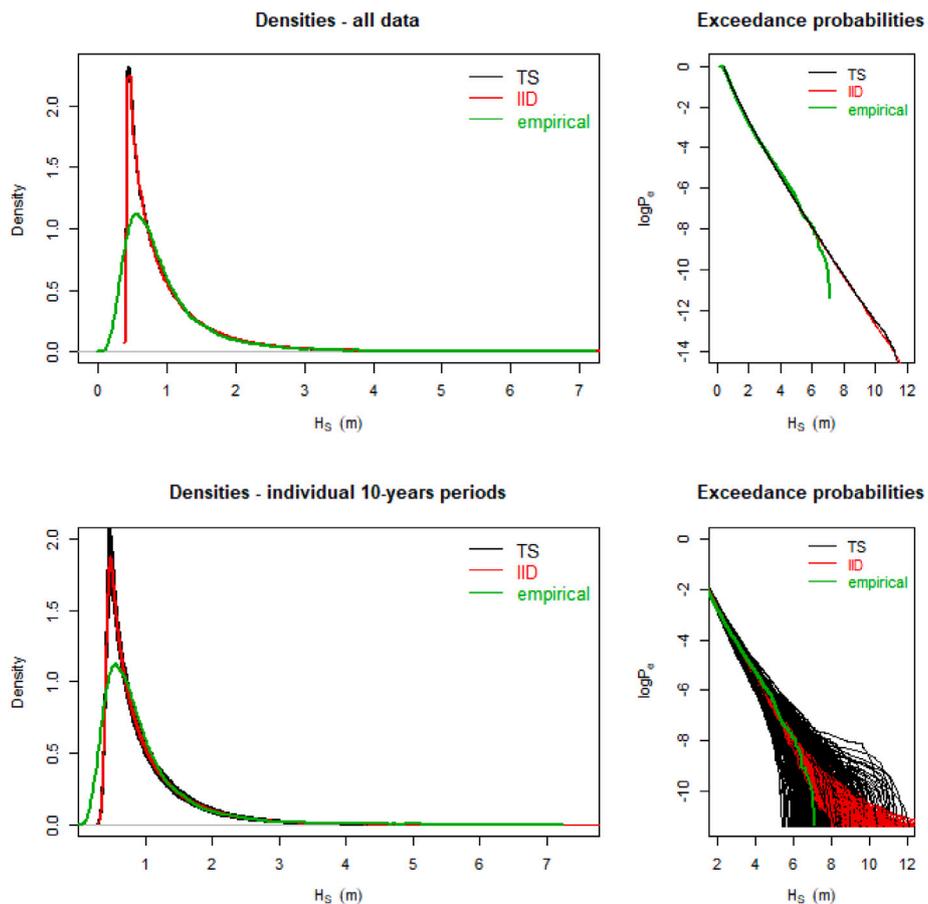


Fig. 4. Densities and exceedance probabilities of empirical and simulated data; based on combining all simulated data (top) and for individual simulated 10-year periods (bottom).

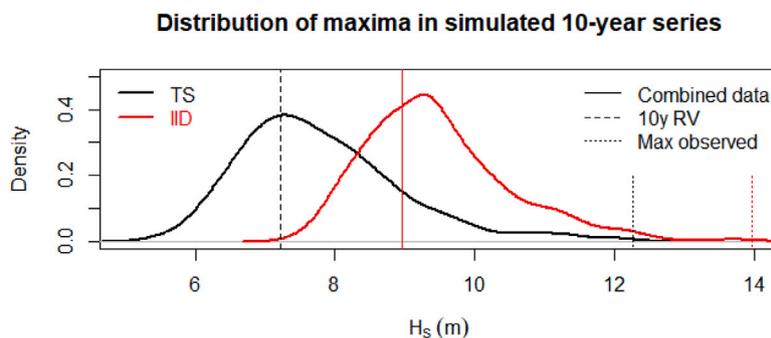


Fig. 5. Distribution of extremes from individual 10-year sets of simulated data; IID and time-series simulations. The vertical lines correspond to different estimates of the 10-year return value: Combined data correspond to return value estimates based on the combined datasets of 500×10 years of data and ignoring serial correlation according to Eq. (4); the other return value estimates are based on the distribution of 10-year maxima, for simulated data assuming IID and the time-series model, respectively, according to Eq. (3). The maximum observed values are also indicated by vertical line segments, illustrating that accounting for serial correlation effectively corresponds to simulating shorter time periods compared to IID.

5. Environmental contours accounting for serial correlation

Environmental contours are often used for describing extreme environmental conditions. With current industry practice, such contours are typically constructed from exceedance probabilities based on Eq. (4), ignoring the effect of serial correlation. An alternative approach to construct environmental contours that takes the effect of serial correlation into account is proposed. With this approach, a total of M T -year contour estimates can be obtained by simulating M T -year time-series from the statistical model and estimate the contours for the most extreme conditions in each individual simulated T -year period of data. This will give a Monte Carlo distribution of T -year contours which can be used to construct environmental contours corresponding to T -year return values according to Eq. (3). The main steps involved in estimating T -year environmental contours using this approach are as follows assuming an appropriate time-series model has been established (see [20,21] for further details on how direct sampling contours are constructed):

1. Simulate M T -year time series from the statistical model
2. For selected angles of rotation θ , get the M angular maxima from the simulated time-series to estimate the distribution of T -year angular maxima, $F_T(x_\theta)$, where x_θ denotes the data projected in direction θ : $X_\theta = X \cos \theta + Y \sin \theta$ for bivariate data (X, Y) , see, e.g., [20,21].
3. Calculate return values, in each direction, according to Eq. (3)
4. Calculate T -year environmental contours from the angular return values by the direct sampling approach [20,21]

The results of simulating $M = 500$ 10-year time-series from the statistical models described in Section 4 and calculating environmental contours for the maxima in each individual time-series are shown in Fig. 6. Results are shown for both data simulated from the joint model assuming IID and for data simulated from the time-series model accounting for autocorrelation. The results from the standard way of constructing environmental contours (ignoring serial correlation) based on the combined 5000-year datasets are also indicated in these plots (referred to as “combined”). Contours corresponding to the 10-year return values, calculated from Eq. (3) based on the simulated 10-year maxima, are also included in the figures, for both the IID and the time-series simulated data, together with 95% confidence bounds of the 10-year maxima. Vertical lines indicate the corresponding values for the estimated 10-year return values for H_S by the different approaches. It is observed that the results from the standard approach on the combined data are the same as the results for the 10-year return value based on individual 10-year datasets assuming IID. However, a notable difference is observed for contours based on time-series simulations taking autocorrelation into account. This corresponds to the bias in neglecting serial correlation, and these results indicate that this bias is

notable and may have practical consequences in design and assessment of marine structures.

The environmental contours corresponding to the 10-year return value are compared in the same plot as shown in Fig. 7. The contours are based on the individual simulated 10-year periods and calculated according to Eq. (3), assuming IID and autocorrelation. Again, a notable difference can be seen, and in particular, contours accounting for serial dependence correct for the positive bias under the IID assumption. This is most pronounced in the primary variable, significant wave height, and not in the secondary variable T_Z . This is presumably because the serial correlation in the secondary variable is not well captured by the statistical model (see also the discussion in Section 6 below). Notwithstanding, it is suggested that the approach outlined in this paper can be adopted when calculating environmental contours as one way of accounting for serial dependence in observed data.

An alternative way of estimating environmental contours based on time-series simulations would be to, rather than simulating M T -year time series and calculating T -year contours from a distribution of simulated T -year maxima according to Eq. (3), to simulate N one-year time series from the statistical model and then calculate T -year contours from the distribution of simulated annual maxima according to Eq. (2). Typically, $N > M$ but since the individual time-series would be much shorter this could be computationally more efficient. Moreover, the simulated data could be used to estimate environmental contours for any return period $T > 1$. However, the overall idea is the same and the results would presumably be similar. The uncertainty of the return periods associated with the contours could then be calculated explicitly as a function of N and T as described in [70]. Fig. 8 show the distribution of maxima from $N = 1000$ simulated 1-year time series, assuming IID data and accounting for the autocorrelation, respectively. Vertical bars in the plot indicate the estimated 1-year return values from Eq. (3) as well as estimated 10- and 100-year return values from Eq. (2). It is clearly seen that the IID assumption leads to a positive bias in the return value estimates. The corresponding 1-, 10- and 100-year environmental contours based on the $N = 1000$ simulated yearly time series are shown in Fig. 9

One advantage of the proposed compared to the approach proposed in [48] is that no de-clustering is required. However, a disadvantage is the need for fitting a time-series model, and the results will be depend on how well such a time-series model capture the temporal correlations. An alternative approach would be to use a block-resampling approach, rather than a time-series model, as described in [71]. The same method for constructing contours as described above could be used for time series generated from any model or resampling approach.

6. Discussion

This paper has presented a way of defining environmental contours that take serial correlation properly into account. Essentially, time

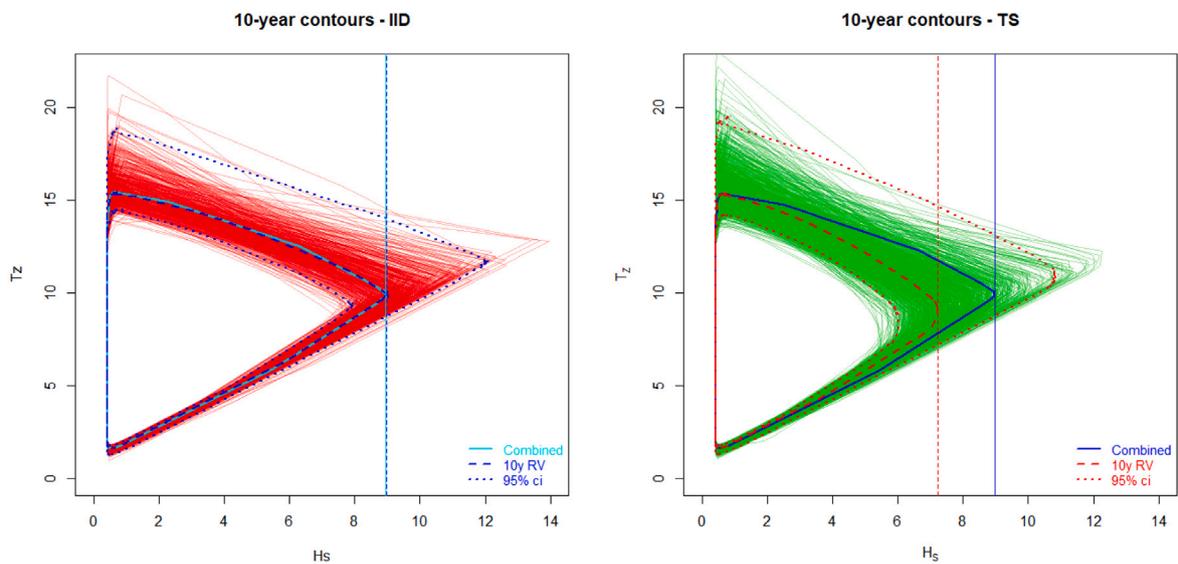


Fig. 6. Contours for the 10-year return value based on simulated time series; assuming IID (left) and accounting for autocorrelation (right). The plots show $M = 500$ individual environmental contours calculated from independent 10-year simulations as well as contours corresponding to the 10-year return value from all the simulations (dashed lines). Direction-wise 95% confidence intervals are included (dotted lines). Contours calculated from the combined datasets ignoring serial correlation are also shown in the Figures. Vertical lines correspond to return value estimates of the primary variable, H_s and show results from the combined dataset (solid lines) as well as results based on the distribution of 10-year maxima (dashed lines) from simulations assuming IID and accounting for serial dependence, respectively.

10-year environmental contours

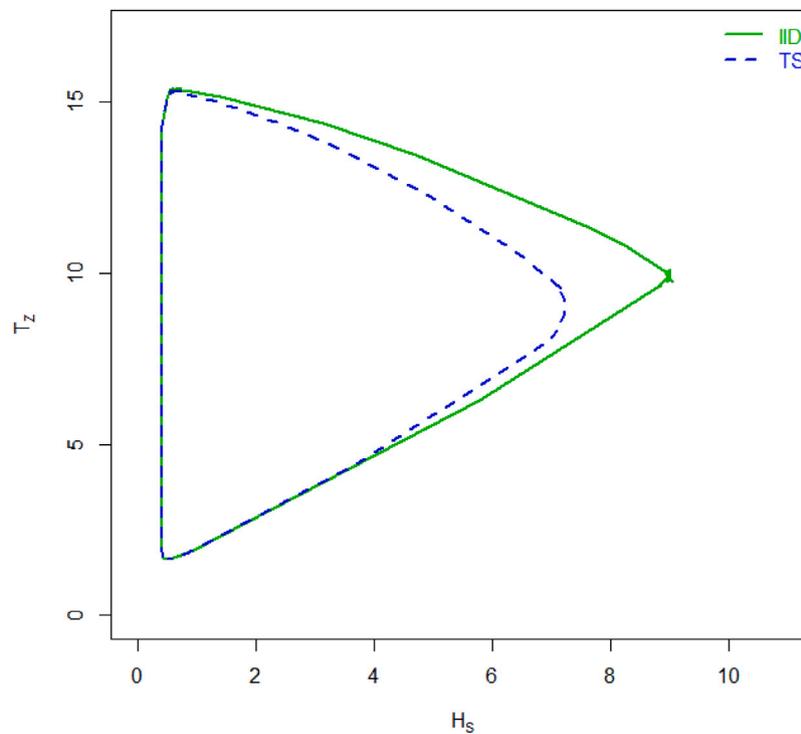


Fig. 7. Comparing contours based on time-series modelling, assuming IID or accounting for autocorrelation in the data; based on simulated 10-year time series.

series are simulated by transformation of a parent Gaussian autoregressive model to yield the target marginal distribution and serial correlation. It has been demonstrated that this yields different contour estimates compared to the commonly used IID assumption, and that the positive bias in IID contours can be avoided. Even though this is believed to be an improvement, some aspects have not been accounted for.

In the statistical models used to construct the environmental contours, seasonal effects have been ignored. Possibly, such effects could also be accounted for in more advanced time-series models, e.g. assuming seasonal ARIMA models rather than simple autoregressive models as the parent time-series. Alternatively, seasonal effects could be taken into account by preprocessing of the data before the statistical models are estimated, as suggested in e.g. [13]. Notwithstanding, seasonality

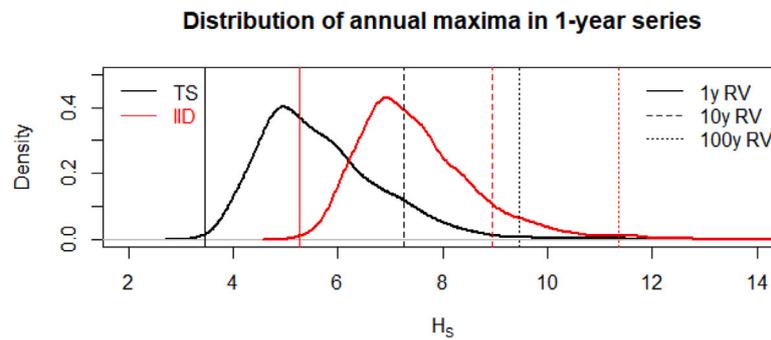


Fig. 8. Distribution of maxima from simulated one-year time series.

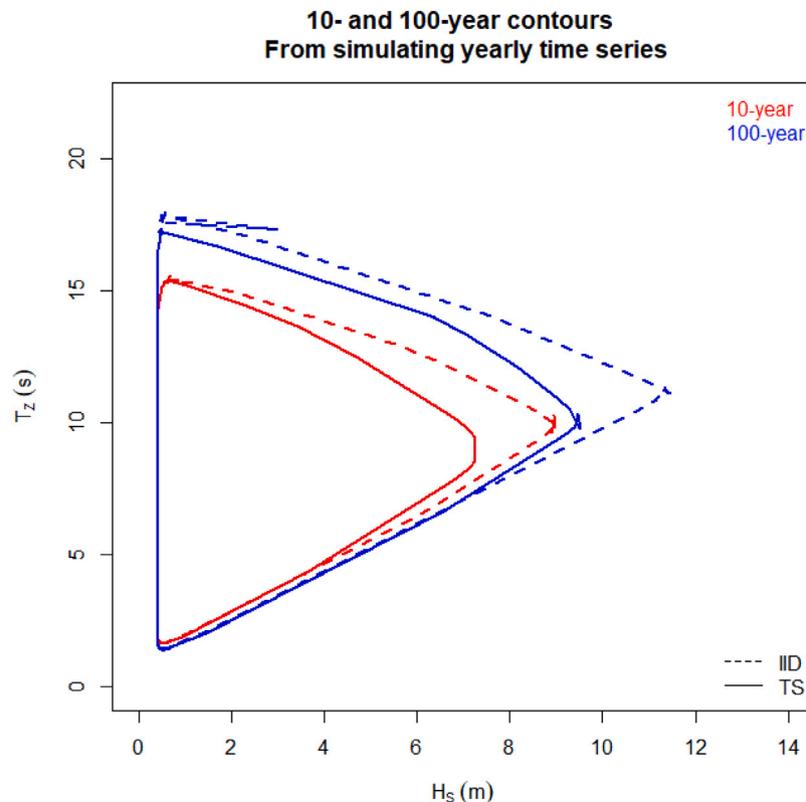


Fig. 9. Environmental contours based on simulating $N = 1000$ one-year series of data.

has not been considered in the current work, and remains an area for further research and potential for improvement of the modelling. Other deterministic components of the observed time-series, for example possible long-term trends due to climate change [72,73], should also ideally be filtered out before the autocorrelation function is estimated [74–76] but this is left for further work. If climate change can be expected to change the future environment, one may model this explicitly in the statistical models, or one may want to add additional safety factors accounting for this to ensure the necessary conservatism and safety level of the design [77]. At any rate, the effect of climate change can potentially be included in the statistical models, for example as an additive deterministic component.

It is noted that only 10 years of data have been available for this study, and that this is generally too short to give robust estimates of extreme conditions; there will be a need to extrapolate the probability distributions beyond the support of the data. Hence, the actual 10-year return values is not known, making it difficult to evaluate directly the impact of accounting for serial correlation on the extreme estimates. Notwithstanding, it is known that neglecting serial correlation yields

a positive bias, and it is observed that return value estimates without accounting for serial dependence gives higher return value estimates than the ones obtained from the time series model accounting for it. Therefore, it is assumed that accounting for serial dependence gives an improved extreme value estimate that is closer to the true value, and that corrects for this positive bias. However, exact quantification of this improvement is difficult without longer datasets and knowledge of the true return values.

The statistical model assumed in this paper consists of a time-series model for the primary variable, H_S , which preserves the desired marginal distribution and autocorrelation structure, and a conditional distribution for the concurrent value of T_Z . It is assumed that this is sufficient and that preserving the autocorrelation structure of the primary variable as well as the dependence between the primary and secondary variable by way of the conditional distribution yields improved description of the extreme environmental conditions. However, it is acknowledged that the autocorrelation structure of the secondary variable will not necessarily be preserved. Hence, further research can be directed towards time series models that preserves all marginals,

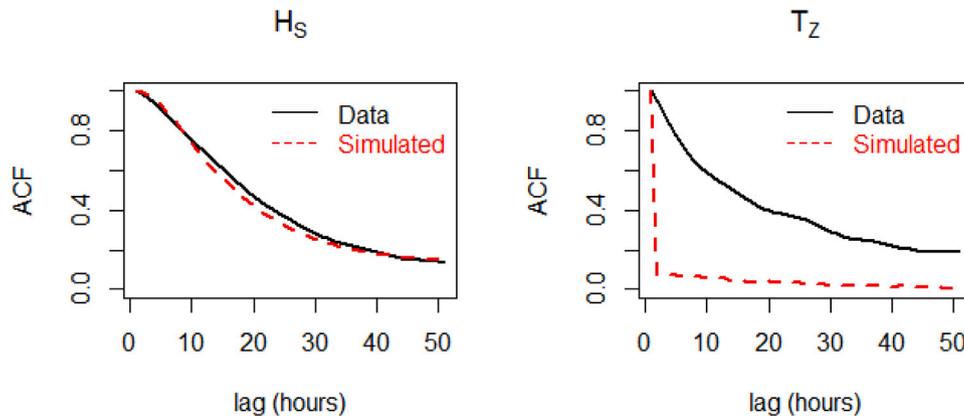


Fig. 10. Autocorrelation structure of observed and simulated time series for both primary and secondary variable.

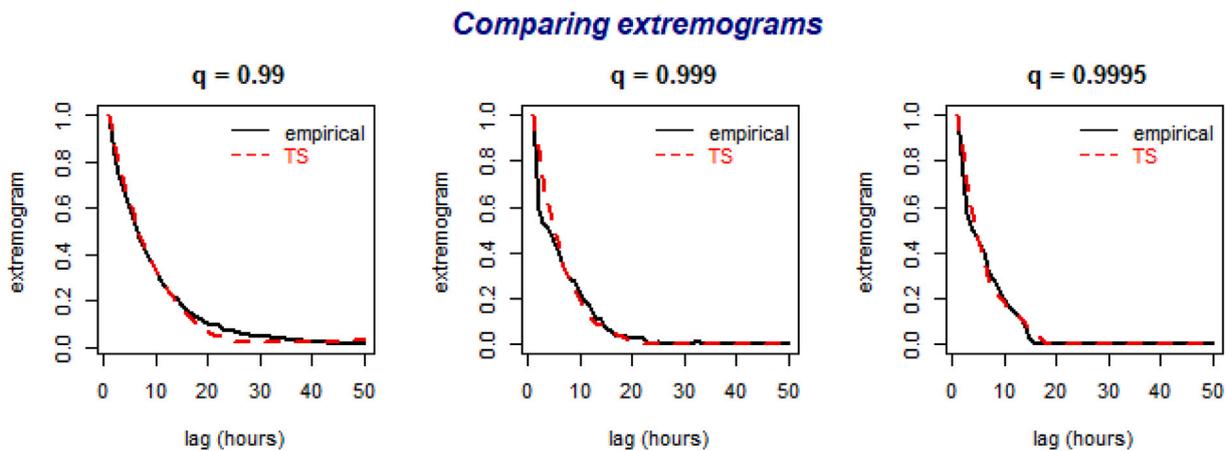


Fig. 11. Comparing the extremogram of the data with that of simulated time-series data for selected quantile levels of significant wave height.

the joint distribution and the autocorrelation structures of multivariate time-series (see, e.g., [78–80]). The empirical autocorrelations are shown for the empirical and simulated data in Fig. 10 for both the primary and the secondary variable. It is observed that the autocorrelation is preserved reasonably well for the primary variable, but not for the secondary. It is observed that some serial correlation is obtained from the conditional modelling approach, but that this is far less than in the observed time-series. Notwithstanding, in practical engineering applications, the extremes of the primary variable, significant wave height, will be most important, so it is believed that the model proposed in this paper represent an improvement that is practically significant in ocean engineering applications.

When the interest is in the extremes, it could be argued that it is not necessarily the standard autocorrelation function that is of most relevance, but rather the level-dependent serial correlation expressed by the conditional probability $\chi(u, \tau) = Pr(X(t + \tau) > u | X(t) > u)$, which may be illustrated by extremograms [81,82]. The correlation structure for high levels, u , should also be preserved in the time series to ensure the effect of serial correlation on extremes are properly accounted for. In this study, the emphasis has been on simulating time-series that preserves the marginal distribution and the autocorrelation structure, rather than preserving the extremogram, and further studies are recommended to construct models that preserve also extremal serial correlation. However, it is possible to compare the sample extremogram of the data and the simulated time-series, as shown in Fig. 11 for simulated significant wave height for levels corresponding to quantiles $q = 0.99, 0.999$ and 0.9995 . As can be seen in these plots, the simulated time series preserves the extremal correlation quite well, and this is clearly an improvement compared to assuming IID data.

7. Summary and conclusions

This paper considers the construction of environmental contours for the description of multivariate extreme environmental condition, a practice that is often used in design and assessment of marine structures. However, the effect of serial dependence is often neglected in engineering practice, and this paper proposes a way to construct contours while accounting for the serial dependence. A statistical model constructed by a joint conditional model for the variables and a time-series model preserving the autocorrelation structure for the primary variable is established and applied to data of significant wave height and zero up-crossing period. This latter time-series model is constructed by transforming a parent Gaussian autoregressive model using appropriate transformations for the marginals and the autocorrelation structure. It has been demonstrated that environmental contours may be constructed based on such statistical models, using the direct sampling approach. Moreover, results indicate that the resulting environmental contours avoid the positive bias one normally gets when ignoring serial dependence, as is current industry practice. This again, means that a more accurate description of the extreme environment is achieved that can be used for more informed decisions in design and operation of marine structures.

Further research could investigate how to account for seasonal effects and the serial correlation of secondary variables as well as models that better capture the serial correlation of extremes. Notwithstanding, the proposed approach represent a significant improvement compared to current engineering practice of ignoring serial correlation and gives more accurate environmental contours.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work has partly been carried out within the HIPERWIND project, supported by EU H2020. The Author also wants to acknowledge valuable comments and discussions with Dr. Ed Mackay from University of Exeter, UK and Guillaume de Hauteclocque from Bureau Veritas, France.

References

- [1] S. Haver, G. Kleiven, Environmental contour lines for design purposes - why and when? in: Proc. 23rd International Conference on Offshore Mechanics and Arctic Engineering (OMAE 2004), American Society of Mechanical Engineers (ASME), 2004.
- [2] DNV GL, Environmental Conditions and Environmental Loads, DNVGL-RP-C205, 2019 ed, DNV GL, 2019.
- [3] J. Mathisen, E. Bitner-Gregersen, Joint distributions for significant wave height and wave zero-up-crossing period, *Appl. Ocean Res.* 12 (1990) 93–103.
- [4] E.M. Bitner-Gregersen, Joint met-ocean description for design and operation of marine structures, *Appl. Ocean Res.* 51 (2015) 279–292.
- [5] J.-T. Horn, E. Bitner-Gregersen, J. Krokstad, B.J. Leira, J. Amdahl, A new combination of conditional environmental distributions, *Appl. Ocean Res.* 73 (2018) 17–26.
- [6] S. Yue, The bivariate lognormal distribution for describing joint statistical properties of a multivariate storm event, *Environmetrics* 13 (2002) 811–819.
- [7] C. Lucas, C. Guedes Soares, Bivariate distributions of significant wave height and mean wave period of combined seas, *Ocean Eng.* 106 (2015) 341–353.
- [8] E. Vanem, Joint statistical models for significant wave height and wave period in a changing climate, *Mar. Struct.* 49 (2016) 180–205.
- [9] Y. Zhang, C.-W. Kim, M. Beer, H. Dai, C. Guedes Soares, Modeling multivariate ocean data using asymmetric copulas, *Coast. Eng.* 135 (2018) 91–111.
- [10] Y. Lin, S. Dong, S. Tao, Modelling long-term joint distribution of significant wave height and mean zero-crossing wave period using a copula mixture, *Ocean Eng.* 197 (2020) 106856.
- [11] F. Chebana, T. Ouarda, Multivariate quantiles in hydrological frequency analysis, *Environmetrics* 22 (2011) 63–78.
- [12] F. Serinald, Dismissing return periods!, *Stoch. Environ. Res. Risk Assess.* 29 (2015) 1179–1189.
- [13] E. Vanem, A simple approach to account for seasonality in the description of extreme ocean environments, *Mar. Syst. Ocean Technol.* 13 (2018) 63–73.
- [14] E. Vanem, Bivariate regional extreme value analysis for significant wave height and wave period, *Appl. Ocean Res.* 101 (2020) 102266, 1–15.
- [15] E. Ross, O.C. Astrup, E. Bitner-Gregersen, N. Bunn, G. Feld, B. Gouldby, A. Huseby, Y. Liu, D. Randell, E. Vanem, P. Jonathan, On environmental contours for marine and coastal design, *Ocean Eng.* 195 (2019) 106194.
- [16] S. Haver, Wave climate off northern Norway, *Appl. Ocean Res.* 7 (1985) 85–92.
- [17] S. Haver, On the joint distribution of heights and periods of sea waves, *Ocean Eng.* 14 (1987) 359–376.
- [18] S. Winterstein, T. Ude, C. Cornell, P. Bjerager, S. Haver, Environmental parameters for extreme response: Inverse FORM with omission factors, in: Proc. 6th International Conference on Structural Safety and Reliability, 1993.
- [19] S. Haver, S. Winterstein, Environmental contour lines: A method for estimating long term extremes by a short term analysis, *Trans. Soc. Nav. Arch. Mar. Eng.* 116 (2009) 116–127.
- [20] A.B. Huseby, E. Vanem, B. Natvig, A new approach to environmental contours for ocean engineering applications based on direct Monte Carlo simulations, *Ocean Eng.* 60 (2013) 124–135.
- [21] A.B. Huseby, E. Vanem, B. Natvig, Alternative environmental contours for structural reliability analysis, *Struct. Saf.* 54 (2015) 32–45.
- [22] E. Vanem, E.M. Bitner-Gregersen, Alternative environmental contours for marine structural design – a comparison study, *J. Offshore Mech. Arct. Eng.* 137 (2015) 051601, 1–8.
- [23] E. Vanem, A comparison study on the estimation of extreme structural response from different environmental contour methods, *Mar. Struct.* 56 (2017) 137–162.
- [24] A.B. Huseby, E. Vanem, C. Agrell, A. Hafver, Convex environmental contours, *Ocean Eng.* 235 (2021) 109366.
- [25] B.J. Leira, A comparison of stochastic process models for definition of design contours, *Struct. Saf.* 30 (2008) 493–505.
- [26] A.F. Haselsteiner, J.-H. Ohlendorf, W. Wosniok, K.-D. Thoben, Deriving environmental contours from highest density regions, *Coast. Eng.* 123 (2017) 42–51.
- [27] F. Silva-González, E. Heredia-Zavoni, R. Montes-Iturrizaga, Development of environmental contours using Nataf distribution model, *Ocean Eng.* 58 (2013) 27–34.
- [28] Q. Li, Z. Gao, T. Moan, Modified environmental contour method for predicting long-term extreme responses of bottom-fixed offshore wind turbines, *Mar. Struct.* 48 (2016) 15–32.
- [29] R. Montes-Iturrizaga, E. Heredia-Zavoni, Environmental contours using copulas, *Appl. Ocean Res.* 52 (2015) 125–139.
- [30] W. Chai, B.J. Leira, Environmental contours based on inverse SORM, *Mar. Struct.* 60 (2018) 34–51.
- [31] L. Manuel, P.T. Nguyen, J. Canning, R.G. Coe, A.C. Eckert-Gallup, N. Martin, Alternative approaches to develop environmental contours from metocean data, *J. Ocean Eng. Mar. Energy* 4 (2018) 293–310.
- [32] F. Silva-González, Vázquez-Hernández, L. Sagrilo, R. Cuamatzi, The effect of some uncertainties associated to the environmental contour lines definition on the extreme response of an FPSO under hurricane conditions, *Appl. Ocean Res.* 53 (2015) 190–199.
- [33] R. Montes-Iturrizaga, E. Heredia-Zavoni, Assessment of uncertainty in environmental contours due to parametric uncertainty in models of the dependence structure between metocean variables, *Appl. Ocean Res.* 64 (2017) 86–104.
- [34] E. Vanem, O. Gramstad, E.M. Bitner-Gregersen, A simulation study on the uncertainty of environmental contours due to sampling variability for different estimation methods, *Appl. Ocean Res.* 91 (2019) 101870.
- [35] G. Clarindo, A. Teixeira, C. Guedes Soares, Environmental wave contours by inverse FORM and Monte Carlo simulation with variance reduction techniques, *Ocean Eng.* 228 (2021) 108916.
- [36] NORSOK, Action and Action Effects, NORSOK Standard N-003:2017, third ed, 2017.
- [37] E. Vanem, 3-dimensional environmental contours based on a direct sampling method for structural reliability analysis of ships and offshore structures, *Ships Offshore Struct.* 14 (2018) 74–85.
- [38] N. Raillard, M. Prevesto, H. Pineau, 3-d environmental extreme value models for the tension in a mooring line of a semi-submersible, *Ocean Eng.* 184 (2019) 23–31.
- [39] Y. Zhao, S. Dong, Design load estimation with IFORM-based models considering long-term extreme response for mooring systems, *Ships Offshore Struct. Latest Articles* (2021).
- [40] A.F. Haselsteiner, R.G. Coe, L. Manuel, P.T. Nguyen, N. Martin, A. Eckert-Gallup, A benchmark exercise on estimating extreme environmental conditions: Methodology & baseline results, in: Proc. 38th International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2019), American Society of Mechanical Engineers (ASME), 2019.
- [41] A.F. Haselsteiner, R.G. Coe, L. Manuel, W. Chai, B. Leira, G. Clarindo, C. Guedes Soares, Á. Hannesdóttir, N. Dimitrov, A. Sander, J.-H. Ohlendorf, K.-D. Thoben, G. de Hauteclocque, E. Mackay, P. Jonathan, C. Qiao, A. Myers, A. Rode, A. Hildebrandt, B. Schmidt, E. Vanem, A.B. Huseby, A benchmarking exercise for environmental contours, *Ocean Eng.* 236 (2021) 109504.
- [42] E. Mackay, G. de Hauteclocque, E. Vanem, P. Jonathan, The effect of serial correlation in environmental conditions on estimates of extreme events, *Ocean Eng.* 242 (2021) 110092.
- [43] S. Coles, An Introduction To Statistical Modeling of Extreme Values, Springer-Verlag, 2001.
- [44] E. Vanem, Uncertainties in extreme value modeling of wave data in a climate change perspective, *J. Ocean Eng. Mar. Energy* 1 (2015) 339–359.
- [45] B.C. Arnold, E. Castillo, J.M. Sarabina, Some alternative bivariate Gumbel models, *Environmetrics* 9 (1998) 599–616.
- [46] P. Rakonczai, A. Zempléni, Bivariate generalized Pareto distribution in practice: models and estimation, *Environmetrics* 23 (2012) 219–227.
- [47] O. Gaidai, G. Storhaug, A. Naess, Extreme large cargo ship panel stresses by bivariate ACER method, *Ocean Eng.* 123 (2016) 432–439.
- [48] Q. Derbanne, G. da Hauteclocque, A new approach for environmental contour and multivariate de-clustering, in: Proc. 38th International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2019), American Society of Mechanical Engineers (ASME), 2019.
- [49] E. Mackay, A.f. Haselsteiner, Marginal and total exceedance probabilities of environmental contours, *Mar. Struct.* 75 (2021) 102863.
- [50] P. Jonathan, D. Randell, J. Wadsworth, J. Tawn, Uncertainties in return values from extreme value analysis of peaks over threshold using the generalised Pareto distribution, *Ocean Eng.* 220 (2021) 107725.
- [51] J. Beirlant, Y. Goegebeur, J. Segers, J. Teugels, *Statistics of Extremes: Theory and Applications*, Wiley, Chichester, UK, 2004.
- [52] Y. Cai, D. Hames, Minimum sample size determination for generalized extreme value distribution, *Commun. Stat. – Simul. Comput.* 40 (1) (2010) 87–98.
- [53] T.H. Soukissian, C. Tsalis, Effects of parameter estimation method and sample size in metocean design conditions, *Ocean Eng.* 169 (2018) 19–37.
- [54] E. Vanem, A.B. Huseby, Environmental contours based on a direct sampling approach and the IFORM approach: contribution to a benchmark study, in: Proc. 39th International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2020), American Society of Mechanical Engineers (ASME), 2020.
- [55] V. Monbet, P. Ailliot, M. Prevesto, Survey of stochastic models for wind and sea state time series, *Probab. Eng. Mech.* 22 (2007) 113–126.

- [56] B.M. Bibby, I.M. Skovgaard, M. Sørensen, Diffusion-type models with given marginal distribution and autocorrelation function, *Bernoulli* 11 (2) (2005) 191–220.
- [57] A. Bensoussan, A. Brouste, Marginal Weibull diffusion model for wind speed modeling and short-term forecasting, in: P. Drobinski, M. Mougeot, D. Picard, R. Plougonven, P. Tankov (Eds.), *Renewable Energy: Forecasting and Risk Management*, Vol. 254, in: Springer Proceedings in Mathematics & Statistics, Springer Nature, Switzerland, 2018, pp. 3–22.
- [58] R. Zárate-Miñano, F.M. Mele, F. Milano, SDE-based wind speed models with Weibull distribution and exponential autocorrelation, in: 2016 IEEE Power and Energy Society General Meeting, PESGM, 2016, pp. 1–5.
- [59] J.P. Arenas-López, M. Badaoui, The Ornstein–Uhlenbeck process for estimating wind power under a memoryless transformation, *Energy* 213 (2020) 118842.
- [60] C.-H. Sim, Simulation of weibull and gamma autoregressive stationary process, *Comm. Statist. Simulation Comput.* 15 (4) (1986) 1141–1146.
- [61] S. Kaur, M. Rakshit, Gaussian and non-Gaussian autoregressive time series models with rainfall data, *Int. J. Eng. Adv. Technol.* 9 (1) (2019) 6699–6704.
- [62] S.M. Papalexiou, Unified theory for stochastic modelling of hydroclimatic processes: Preserving marginal distributions, correlation structures, and intermittency, *Adv. Water Resour.* 115 (2018) 234–252.
- [63] S.M. Papalexiou, Y. Markonis, F. Lombardo, A. AghaKouchak, E. Foufoula-Georgiou, Precise temporal disaggregation preserving marginals and correlations (DiPMaC) for stationary and nonstationary processes, *Water Resour. Res.* 54 (2018) 7435–7458.
- [64] P. Carpena, P.A. Bernaola-Galván, M. Gómez-Extremera, A.V. Coronado, Transforming Gaussian correlations: Applications to generating long-rang power-law correlated time series with arbitrary distribution, *Chaos* 30 (2020) 083140.
- [65] E. Vanem, Analuzing extreme sea state conditions by time.series simulation, in: *Proc. 41st International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2022)*, American Society of Mechanical Engineers (ASME), 2022.
- [66] W. Feller, *An Introduction To Probability Theory and Its Applications*, John Wiley and Sons, 1971.
- [67] P.J. Brockwell, R.A. Davis, *Time Series: Theory and Methods*, second ed, Springer-Verlag, 1991.
- [68] S.M. Papalexiou, F. Serinaldi, E. Porcu, Advancing space–time simulation of random fields: From storms to cyclones and beyond, *Water Resour. Res.* 57 (2021) e2020WR029466.
- [69] A. Luceño, Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators, *Comput. Statist. Data Anal.* 51 (2) (2006) 904–917.
- [70] E. Mackay, P. Jonathan, Sampling properties and empirical estimates of extreme events, *Ocean Eng.* 239 (2021) 109791.
- [71] E. Mackay, P. Jonathan, Estimation of environmental contours using a block resampling method, in: 39th International Conference on Ocean, Offshore and Arctic Engineering, 2020, pp. OMAE2020–18308.
- [72] M.A. Hemer, Y. Fan, N. Mori, A. Semedo, X.L. Wang, Projected changes in wave climate from a multi-model ensemble, *Nature Clim. Change* 3 (2013) 471–476.
- [73] H. Lobeto, M. Menendez, I.J. Losada, Future behavior of wind wave extremes due to climate change, *Sci. Rep.* 11 (2021) 7869.
- [74] G. Athanassoulis, C.N. Stefanakos, A nonstationary stochastic model for long-term time series of significant wave height, *J. Geophys. Res.* 100 (8) (1995) 16149–16162.
- [75] C.N. Stefanakos, G.A. Athanassoulis, Extreme value predictions based on nonstationary time series of wave data, *Environmetrics* 17 (2006) 25–46.
- [76] C.N. Stefanakos, E. Vanem, Nonstationary fuzzy forecasting of wind and wave climate in very long-term scales, *J. Ocean Eng. Sci.* 3 (2018) 144–155.
- [77] E.M. Bitner-Gregersen, E. Vanem, O. Gramstad, T. Hørte, O. j. Aarnes, M. Reistad, Ø. Breivik, A.K. Magnussen, B. Natvig, Climate change and safe design of ship structures, *Ocean Eng.* 149 (2018) 226–237.
- [78] S.M. Papalexiou, F. Serinaldi, Random fields simplified: Preserving marginal distributions, correlations, and intermittency, with applications from rainfall to humidity, *Water Resour. Res.* 56 (2020) e2019WR026331.
- [79] I. Tsoukalas, C. Makropoulos, D. Koutsoyiannis, Simulation of stochastic processes exhibiting any-range dependence and arbitrary marginal distributions, *Water Resour. Res.* 54 (2018) 9484–9513.
- [80] I. Tsoukalas, A. Efstratiadis, C. Makropoulos, Stochastic periodic autoregressive to anything (SPARTA): Modeling and simulation of cyclostationary processes with arbitrary marginal distributions, *Water Resour. Res.* 54 (2018) 161–185.
- [81] R.A. Davis, T. Mikosch, The extremogram: A correlogram for extreme events, *Bernoulli* 15 (4) (2009) 977–1009.
- [82] R.A. Davis, T. Mikosch, I. Cribben, Towards estimating extremal serial dependence via the bootstrapped extremogram, *J. Econometrics* 170 (2012) 142–152.