On the Estimation of Ocean Engineering Design Contours

Philip Jonathan
Shell Projects & Technology,
Brabazon House,
Manchester M22 0RR, UK
e-mail: philip.jonathan@shell.com

Kevin Ewans
Sarawak Shell Bhd,
eTiQa Twins (Level 23, Tower 1),
Kuala Lumpur 50450, Malaysia
e-mail: kevin.ewans@shell.com

Jan Flynn
Shell International Exploration and Production,
P.O. Box 60,
AB Rijswijk 2280, The Netherlands
e-mail: jan.flynn@shell.com

Understanding extreme ocean environments and their interaction with fixed and floating structures is critical for offshore and coastal design. Design contours are useful to describe the joint behavior of environmental, structural loading, and response variables. We compare different forms of design contours, using theory and simulation, and present a new method for joint estimation of contours of constant exceedance probability for a general set of variables. The method is based on a conditional extremes model from the statistics literature, motivated by asymptotic considerations. We simulate under the conditional extremes model to estimate contours of constant exceedance probability. We also use the estimated conditional extremes model to estimate other forms of design contours, including those based on the first-order reliability method (FORM), without needing to specify the functional forms of conditional dependence between variables. We demonstrate the application of new method in estimation of contours of constant exceedance probability using measured and hindcast data from the Northern North Sea, the Gulf of Mexico, and the North West Shelf of Australia, and quantify their uncertainties using a bootstrap analysis. [DOI: 10.1115/1.4027645]

1 Introduction

We are interested in joint estimation of design contours for a general set of p variables $\mathbf{X} = \{X_j\}_{j=1}^p$, characterizing the ocean environment, and structural loading and response in offshore design applications. Development of design contours using methods related to the FORM, see, e.g., Refs. [1–4], have proved popular with practitioners since extreme environments are characterized independently of structural loading and response, in a manner which lends itself to subsequent interpretation in terms of structural loading and response. For example, Ref. [5] describes the estimation of inverse-FORM contours for winds, waves, and current off West Africa. In such analyses, we suppose it is possible to transform $\mathbf{X}$ into a set of independent random variables, $\mathbf{V} = \{V_j\}_{j=1}^p$. Using the probability integral transform, we further transform these independently to standard normal random variables $\mathbf{U} = \{U_j\}_{j=1}^p$. This set has a joint probability density function $f(\mathbf{u}) = (2\pi)^{-p/2} \exp(-\frac{1}{2}\|\mathbf{u}\|^2)$, where $\|\mathbf{u}\|^2 = \mathbf{u}^T\mathbf{u}$ is the Euclidean norm of $\mathbf{u}$. The surface $S(\beta)$ defined by $\|\mathbf{u}\| = \beta$ (a circle in two-dimensions, $p = 2$) corresponds to a surface of constant probability density. We select the value of β such that $S(\beta)$ encloses a set with given probability of occurrence, for example, or corresponds to a specified return period marginally (i.e., along any axis).

Contours of constant probability density are not the only possibilities, however. For example, consider the definition of a return value in the univariate case. A return value x corresponding to a T-year return period of random variable X is defined by $\Pr(X > x) = (1/n_T)$, where n_T is the number of events in the period. That is, x is defined in terms of an exceedance probability. Extending the analogy to two-dimensions, the set of points $\mathbf{x}(\theta) = \{x_j(\theta)\}_{j=1}^2, \theta \in [0, 360)$ (say), on a contour of constant exceedance probability α can be defined by the expression

$$\Pr\left(\bigcap_{j=1}^2 (r_j(\theta; \mathbf{r}_0)X_j > r_j(\theta; \mathbf{r}_0)x_j(\theta))\right) = \alpha$$

where $\mathbf{r}(\theta; \mathbf{r}_0) = \{r_j(\theta; \mathbf{r}_0)\}_{j=1}^2$ is defined by $\mathbf{r}(\theta; \mathbf{r}_0) = \mathbf{x}(\theta) - \mathbf{r}_0$. Here, $\mathbf{r}_0$ is a reference location for the distribution under consideration. In some cases, it is appropriate that $\mathbf{r}_0$ refer to some central feature (e.g., mean, median, or mode). In other situations, when we are interested solely in the large values of a variable $X_k, k = 1, 2$, (i.e., in the right hand tail of its distribution), it is appropriate to set $\mathbf{r}_{0k} = 0$. We might construct the contour in terms of the original variables $\mathbf{X}$, or any transformation thereof, in particular, the independent standard normals $\mathbf{U}$.

A key assumption in the estimation of FORM-type contours is problem representation in terms of a set of independent random variables $\mathbf{V}$. Consider the case $p = 2$. The probability density $f(x_1, x_2)$ can be factorized as $f(x_1)f(x_2|x_1)$, in principle. In some applications, a model based on $V_1 = X_1$ and $V_2 = X_2|X_1 = x_1$ is justified on physical grounds, so that $\mathbf{V}$ are independent. For example, it is common (e.g., Ref. [2]) to model the pair of variables significant wave height H_S and peak period T_P in terms of H_S (assumed to be Weibull-distributed) and $T_P|H_S = h$ (assumed to be independently log-normally distributed). If this model can be shown to fit well throughout the sample, we can estimate contours of constant probability density. Whereas the model appears to explain the body of sample data well, it is not clear (e.g., Ref. [6]) whether this model is appropriate for extremes of H_S and T_P. The generalized Pareto (GP) distribution has a stronger asymptotic motivation for modeling (peaks over threshold) of H_S than the Weibull. The log-normal parameter estimates (for $T_P|H_S$) fitted within the body of the sample may not be relevant for extrapolation to extremes. The conditional form $V_2 = (X_2|X_1 = x_1)$ will not be known in general and might not be easily or adequately approximated.

The conditional extremes model of Ref. [7] provides a means to model the marginal and dependence structure of extremes of the set $\mathbf{X}$, by: (1) (marginal) modeling of threshold exceedances of each variable independently (using the generalized Pareto distribution) and (2) (dependence) modeling of pairs $Y_j|Y_k = y_k$ for threshold exceedances of $Y_k, (j = 1, 2, ..., p, j \neq k; k = 1, 2, ..., p)$, following a transformation of variables from the original (X) scale to standard Gumbel (Y) scale. We estimate contours of constant exceedance probability by simulation under the model. The conditional extremes model is motivated by asymptotic considerations, and has been reported by the current authors in the past [6]. It provides a conditional decomposition of variables, given threshold exceedances of the conditioning variate, which is generally applicable for

estimation of design contours. It can complement (or replace) FORM-type analysis when functional forms of conditional relationships are not known and has a strong underpinning in multivariate extreme value theory. It is particularly relevant to estimation of contours of constant exceedance probability on the original (X) scale of variables, as will be illustrated below.

The layout of the article is as follows: In Sec. 2, we introduce four applications motivating the current work. In Sec. 3, we overview the conditional extremes model and its application in estimation of design contours of different forms. In Sec. 4, we evaluate the performance of the conditional extremes model in estimating all design contours for a simulated example. In Sec. 5, we estimate contours of constant exceedance probability for the four applications introduced in Sec. 2. Conclusions are made in Sec. 6. The Appendix gives a fuller description of the conditional extremes model.

We note that, for $p > 2$, the terms surface or manifold would be more appropriate than contour. We persist with the latter for ease of understanding. Further, in two-dimensions, we note that the parameter $\theta \in [0, 360)$ used to reference the contour is defined in the usual mathematical sense, anticlockwise from the abscissa. Significant wave height (H_S) and peak spectral period (T_P) have units of meters and seconds, respectively, throughout.

2 Data

We motivate current work by considering applications for estimation of extreme values of H_S and T_P at a location in the Northern North Sea (NNS) for which both measured and hindcast data are available, to buoy data from a location in the Gulf of Mexico (GoM), and to hindcast data on the North West Shelf (NWS) of Australia. Data correspond to storm peak values for H_S over threshold, observed during periods of storm events, and corresponding values for T_P. For the measured NNS example, illustrated in Fig. 6, the sample corresponds to 620 storm peak pairs for the period (March 1973 and December 2006) measured using a laser device. Hindcast data for the same NNS location (see Fig. 7) corresponds to 827 pairs of values for the period (November 1964 and April 1998). GoM data in Fig. 8 are National Data Buoy Center measurements from buoy 42,002 corresponding to 505 pairs for the period (January 1980 and December 2007). Finally, NWS hindcast data (shown in Fig. 9) correspond to 145 pairs of storm peak values for the period (February 1970 and April 2006). All samples exhibit positive dependence between H_S and T_P. The four application data sets are chosen to illustrate methods of estimation of design contours, not for development of design criteria. In general, sample sizes are only of the order of 30 yr in length, and data include strong covariate effects due, for example, to seasonality. No attempt has been made at this point to accommodate these effects in the analysis.

3 Illustration and Evaluation

3.1 Conditional Extremes Model.
Here, we outline the conditional extremes model in two-dimensions. A fuller description of the model in p dimensions is given in the Appendix. For pairs (Y_1, Y_2) of random variables with marginal Gumbel distributions, Ref. [7] derives a parametric form for the conditional distribution of one variable given a large value of the other. This parametric form, motivated by the assumption of a particular limit representation for the conditional distribution (see Ref. [6]) is appropriate to characterize the conditional behavior of a wide range of theoretical examples of bivariate (and higher-dimensional) distributions for extremes. Model form for positively associated pairs of variables (Y_1, Y_2) takes a particularly simple form

$$(Y_2 | Y_1 = y_1) = a y_1 + y_1^b Z$$

where a and b are location and scale parameters, respectively, to be estimated, with $a \in [0, 1]$ and $b \in (-\infty, 1)$, and y_1 is large. Z is

a random variable, independent of Y_1, converging with increasing y_1 to a nondegenerate limiting distribution G. Joint tail behavior is then characterized by a, b, and G. The form of distribution G is not specified by theory. For a sample $\{y_{i1}, y_{i2}\}_{i=1}^{n}$ of values from (Y_1, Y_2) with values of Y_1 exceeding an appropriate threshold w_{Y1}, the values of a, b, and G are estimated using regression. For simplicity and computational ease during model fitting, G is assumed to be a Gaussian distribution with mean μ_Z and variance σ_Z^2 treated as nuisance parameters. Fitted values

$$\hat{z}_i = \frac{(y_{i2} - \hat{a} y_{i1})}{y_{i1}^{\hat{b}}}, \quad i = 1, 2, 3, ..., n$$

are used to estimate distribution G, and the estimate $\hat{G}$ is then sampled in subsequent simulations. The adequacy of model fit can be assessed by (1) demonstrating that the values $\{\hat{z}_i\}_{i=1}^{n}$ and $\{y_{i1}\}_{i=1}^{n}$ are not obviously dependent, (2) exploring the effect of varying w_{Y1} on $\hat{a}$, $\hat{b}$, and subsequent estimates (e.g., of probabilities associated with extreme sets), and (3) bootstrap resampling to estimate the uncertainty of estimates for a, b, and subsequent estimates for a given threshold choice.

The GP form is appropriate for model marginal distributions of peaks over threshold, rather than the Gumbel distribution. Therefore, in such cases, to use the conditional model above, we transform original variables (X_1, X_2) from GP to Gumbel using the probability integral transform as follows. Suppose we fit the GP distribution (to the sample from X_1 without loss of generality)

$$F_{GP}(x_1; \xi, \beta, u) = 1 - (1 + \frac{\xi}{\beta}(x_1 - u))_+^{-\frac{1}{\xi}}$$

and estimate cumulative probabilities $\{F_{GP}(x_{i1}; \hat{\xi}, \hat{\beta}, u)\}_{i=1}^{n}$. The standard Gumbel distribution has cumulative distribution function $F_G(x) = \exp(-\exp(-x))$; thus, if we define the transformed sample $\{y_{i1}\}_{i=1}^{n}$ such that $F_G(y_{i1}) = F_{GP}(x_{i1}; \hat{\xi}, \hat{\beta}, u)$ or

$$y_{i1} = -\log(-\log(F_{GP}(x_{i1}; \hat{\xi}, \hat{\beta}, u))) \quad \text{for} \quad i = 1, 2, 3, ..., n$$

the transformed sample will be consistent with a sample from a Gumbel distribution. Similarly, given a value y_1 (of Gumbel variate), we can calculate the corresponding value x_1 on the original GP scale. It is essential to demonstrate the adequacy of GP marginal fits, e.g., by using the mean residual life plot and stability of marginal shape parameter as a function of threshold choice. Finally, to simulate a random drawing from the conditional distribution $Y | X > u$, the following procedure can then be followed: (1) Draw a value y_1 of Y_1 at random from its standard Gumbel distribution, given that the value exceeds threshold w_{Y1}, (2) draw a value of z of Z at random from the set $\{\hat{z}_i\}_{i=1}^{n}$, (3) calculate that value of $y_2 | y_1 = \hat{a} y_1 + y_1^{\hat{b}} z$, and (4) transform (y_1, y_2) to (x_1, x_2) using the probability integral transform and the estimated GP marginal model parameters. Using simulation, estimates for various extremal statistics (e.g., values associated with long return periods) can be obtained routinely. Modeling and simulation of $(Y_1 | Y_2 = y_2)$ can be achieved in the same way. Extension to p-dimensional variates is outlined in the Appendix.

3.2 Contours of Constant Exceedance Probability.
Contours of constant exceedance probability can be estimated for a general set of variables $M = \{M_j\}_{j=1}^{p}$ (e.g., on original (X), independent (V), independent standard normal (U) or Gumbel (Y) scales). The contour is defined as the set $\mathbf{m}(\theta) = \{m_j(\theta)\}_{j=1}^{p}$, $\theta \in [0, 360)^{p-1}$ such that

$$\Pr\left(\bigcap_{j=1}^{p}(r_{Mj}(\theta; \mathbf{r}_{M0})M_j > r_{Mj}(\theta; \mathbf{r}_{M0})m_j(\theta))\right) = \alpha$$

where $\mathbf{r}_M(\boldsymbol{\theta}; \mathbf{r}_{M0}) = \{r_{Mj}(\boldsymbol{\theta}; \mathbf{r}_{M0})\}_{j=1}^p$ is defined by $\mathbf{r}_M(\boldsymbol{\theta}; \mathbf{r}_{M0}) = \mathbf{m}(\boldsymbol{\theta}) - \mathbf{r}_{M0}$, where $\mathbf{r}_{M0}$ is a reference location for the distribution under consideration. Thus, $\mathbf{r}_M(\boldsymbol{\theta}; \mathbf{r}_{M0})$ is the position vector of the contour, as a function of angle θ, relative to reference location $\mathbf{r}_{M0}$, beyond which the probability of nonexceedance is a constant in the sense defined.

We now define four forms of design contour for further consideration.

3.3 Design Contour C1: Constant Probability Density on Independent Standard Normal Scale. We assume that FORM-type models are available to transform the original set of variables $\mathbf{X}$ to independent standard normal random variables $\mathbf{U}$. Following the description in Sec. 1, we evaluate the hypersphere $\|\mathbf{u}\| = \beta$ for which the joint probability density function $f(\mathbf{u}) = (2\pi)^{-p/2}\exp(-(1/2)\|\mathbf{u}\|^2)$ is a constant. We select the value of β corresponding to a specific return value of U_1. We then back transform the contour to the original (X) scale for interpretation.

3.4 Design Contour C2: Constant Exceedance Probability on Independent Standard Normal Scale. As for contour C1, we assume we can transform the original set of variables $\mathbf{X}$ to independent standard normal random variables $\mathbf{U}$. We evaluate contours of constant exceedance probability using the equation in Sec. 3.2 with $\mathbf{M} = \mathbf{U}$, with a reference location $\mathbf{r}_{U0} = \mathbf{0}$. We then back-transform the contour to the original (X) scale for interpretation.

In $p = 2$ dimensions, in the first quadrant ($\theta \in (0, 90)$), the contour corresponds to $\Pr(U_1 > u_1(\theta), U_2 > u_2(\theta)) = \alpha$. In the second, third, and fourth quadrants ($\theta \in (90, 180)$, $\theta \in (180, 270)$, and $\theta \in (270, 360)$, respectively), corresponding contour definitions are $\Pr(U_1 \leq u_1(\theta), U_2 > u_2(\theta)) = \alpha$, $\Pr(U_1 \leq u_1(\theta), U_2 \leq u_2(\theta)) = \alpha$, and $\Pr(U_1 > u_1(\theta), U_2 \leq u_2(\theta)) = \alpha$.

3.5 Design Contour C3: Constant One-Sided Exceedance Probability on Independent Standard Normal Scale. This contour differs from C2 only in the choice of reference location $\mathbf{r}_{U0}$. To achieve a constant one-sided exceedance contour with respect to any subset $\mathbf{X}_K$, $K \subseteq \{j\}_{j=1}^p$, we set the corresponding components of $\mathbf{r}_{U0}$ to the minimum values of $\mathbf{U}_K$ observed in the data sample. Again, we back transform the contour found to the original (X) scale for interpretation.

In $p = 2$ dimensions, in the first and second quadrants ($\theta \in (0, 180)$), contour definition is $\Pr(U_1 > u_1(\theta), U_2 > u_2(\theta)) = \alpha$, whereas in the third and fourth quadrants ($\theta \in (180, 360)$), contour definition is $\Pr(U_1 > u_1(\theta), U_2 \leq u_2(\theta)) = \alpha$.

3.6 Design Contour C4: Constant One-Sided Exceedance Probability on Original Scale. We estimate contours using the equation in Sec. 3.2 with $\mathbf{M} = \mathbf{X}$, i.e., using the original variables, by direct simulation under the conditional extremes model in p dimensions, necessitating the estimation of conditional extremes models with respect to each of the p variables, as outlined in the Appendix. The reference location definition is analogous to that for C3.

In $p = 2$ dimensions, in the first and second quadrants ($\theta \in (0, 180)$), the contour definition is $\Pr(X_1 > x_1(\theta), X_2 > x_2(\theta)) = \alpha$, whereas in the third and fourth quadrants ($\theta \in (180, 360)$), the contour definition is $\Pr(X_1 > x_1(\theta), X_2 \leq x_2(\theta)) = \alpha$.

4 Design Contours Using the Conditional Extremes Model

We illustrate the characteristics of the four design contours C1–C4 introduced in Sec. 3 above, using a two-dimensional model for the pair $(X_1, X_2) = (H_S, T_P)$ discussed in Ref. [8]. We also quantify the performance of the conditional extremes model
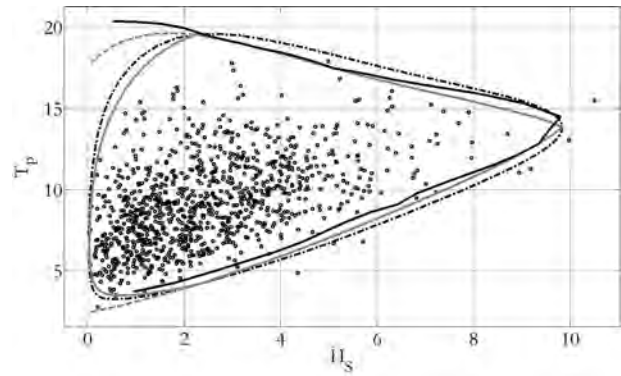

Fig. 1 Contours C1-C4 (see Sec. 3) corresponding to the H_S–T_P model (see Sec. 4) (equivalent to a 1 in 1000 event of H_S marginally): C1 (dashed black), C2 (solid gray), C3 (dashed gray), and C4 (solid black). Also shown is a random sample of 1000 values from the model.

in approximating each of these contours. We assume that H_S is Weibull-distributed as follows:

$$F_{Hs}(h) = \exp\left(-\left(\frac{h}{2.822}\right)^{1.547}\right)$$

Conditionally on H_S, T_P is assumed independently log-normally distributed

$$((\log_e T_P)|H_S = h) = 1.59 + 0.42\log_e(H_S + 2) + \sqrt{0.005 + 0.085\exp(-0.13 H_S^{1.34})}\varepsilon$$

where ε is a standard normal random variate.

4.1 True Model Design Contours. Given these functional forms, we can evaluate true contours C1–C4 for any return value. Figure 1 shows the contours corresponding to a 1 in 1000 event of H_S marginally, together with a random sample of 1000 values from the model (for illustration only).

A number of interesting features can be observed from the figure. The maximum value of H_S for each of the contours is the same by construction (corresponding to the 1 in 1000 event). The constant density contour C1 (dashed black) is continuous and smooth (i.e., continuous first derivative) whereas the constant exceedance probability contours C2–C4 are continuous but not smooth. Contours C2 (solid gray) and C3 (dashed gray) are identical for values of H_S above its median by definition. For this example, the locations of the four contours are relatively similar.

4.2 Design Contours Estimated Using the Conditional Extremes Model. Now we use the conditional extremes model to estimate each of the contours C1–C4 by fitting to a sample. A random sample of 1000 values was drawn from the H_S–T_P model and the conditional extremes model estimated using a threshold corresponding to the 80‰ used for marginal modeling, and the 95‰ for conditional model. The estimated conditional extremes model is then used to estimate the four contours C1–C4 in turn. For estimation of C1–C3, only the estimated conditional extremes model parameters are required for contour estimation. For contour C4, we simulate a sample of 10,000 values under the conditional extremes model, and estimate the contour empirically (by counting). Results are shown in Figs. 2–5, respectively, in terms of estimates for bootstrap median (thick black) and 95% band (thin black) for $H_S(\theta)$ (solid black) and $T_P(\theta)$ (dashed black) on the contour for $\theta \in [0, 360)$. Also shown, for each figure, is the corresponding true contour (in gray). The bootstrap analysis is based
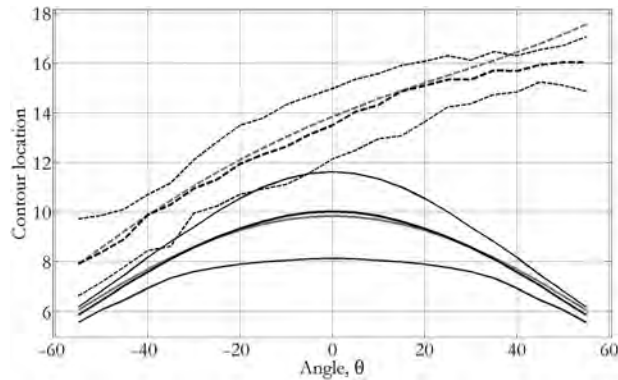
Fig. 2 True (gray) and estimated (black) contour C1 corresponding to the H_S–T_P model (equivalent to a 1 in 1000 event of H_S marginally), as a function of angle θ, for $\theta \in [0, 360)$. $H_S(\theta)$ is the solid line, and $T_P(\theta)$ is the dashed line. Thick lines correspond to the bootstrap median, thin lines to a bootstrap 95% uncertainty band.
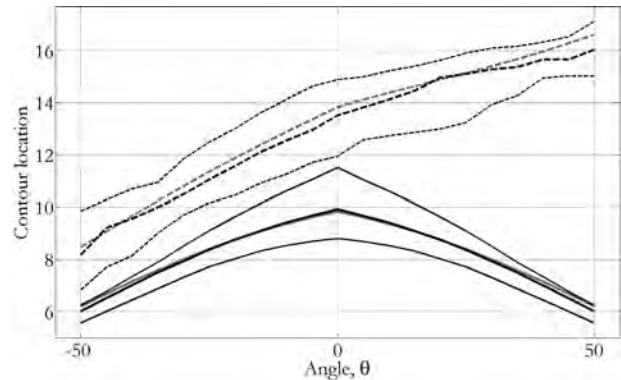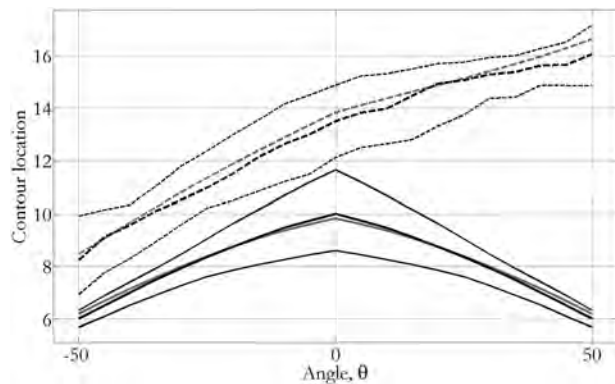

Fig. 4 True (gray) and estimated (black) contour C3 corresponding to the H_S–T_P model (equivalent to a 1 in 1000 event of H_S marginally), as a function of angle θ, for $\theta \in [0, 360)$. $H_S(\theta)$ is the solid line, and $T_P(\theta)$ is the dashed line. Thick lines correspond to the bootstrap median and thin lines to a bootstrap 95% uncertainty band.
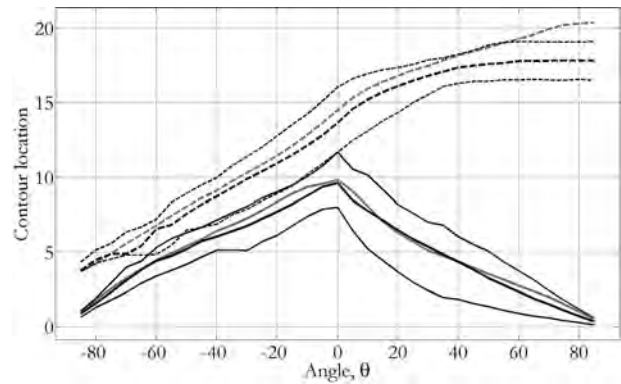

Fig. 3 True (gray) and estimated (black) contour C2 corresponding to the H_S–T_P model (equivalent to a 1 in 1000 event of H_S marginally), as a function of angle θ, for $\theta \in [0, 360)$. $H_S(\theta)$ is the solid line, and $T_P(\theta)$ is the dashed line. Thick lines correspond to the bootstrap median and thin lines to a bootstrap 95% uncertainty band.

Fig. 5 True (gray) and estimated (black) contour C4 corresponding to the H_S–T_P model (equivalent to a 1 in 1000 event of H_S marginally), as a function of angle θ, for $\theta \in [0, 360)$. $H_S(\theta)$ is the solid line, and $T_P(\theta)$ is the dashed line. Thick lines correspond to the bootstrap median and thin lines to a bootstrap 95% uncertainty band.

on full modeling of 100 resamples of the original data, and quantifies the sampling uncertainty of the whole modeling procedure. We note that, for contours C1–C3, the conditional exceedance model can only be used to estimate the section of the contour corresponding to exceedances of the H_S threshold, since the conditional extremes model applies only to extremes of H_S and corresponding conditional values of T_P. For contour C4, however, simulation under the conditional extremes model allows the full contour to be estimated. We see from the figures that agreement between true contours and their estimates is good in each case. Moreover, contour estimates from the conditional extremes model were achieved without explicit knowledge of the H_S–T_P model. This suggests that the conditional extremes model is applicable for evaluation of design contours for arbitrary sets of variables, with no requirement for knowledge of conditional relationships between variables (other that that provided by the sample data).

5 Application

The four samples of measured and hindcast (storm peak) H_S and T_P data selected to illustrate development of design contours using the conditional extremes model are introduced in Sec. 2. Marginal and conditional extremes modeling of the samples has been reported in Ref. [6]. Threshold selection in marginal

modeling was noted to be problematic in some cases. Based on assessment of various diagnostics plots, particularly the variability of marginal and conditional modeling parameters, and marginal return value estimates, with threshold, a 60‰ threshold was adopted for marginal modeling and a 70‰ threshold for conditional modeling in all four cases. Uncertainties in model parameter estimates, quantified using a bootstrap analysis, have also been reported in Ref. [6]. Here, we focus on the estimation of contours of constant exceedance probability (of type C4), one-sided with respect to H_S, corresponding to return periods of 10, 100, and 1000 yr. Figures 6–9 show estimates for measured NNS, hindcast NNS, measured GoM, and hindcast NWS, respectively, together with a scatter plot of the corresponding samples. All contour estimates are based on a 10,000 yr full simulation using the conditional extremes model. The measured and hindcast NNS contours (for the same nominal location) would appear to be qualitatively similar. The spacing of contours corresponding to different return periods is larger for the measured GoM indicates long marginal tails, particularly in T_P. The hindcast NWS contours show a complex structure.

Figures 10–13 show 100 yr contour estimates in terms of bootstrap median (thick) and 95% uncertainty band (thin) for $H_S(\theta)$ (solid) and $T_P(\theta)$ (dashed), for $\theta \in [0, 360)$. There is considerable uncertainty in the location of the contour, reflecting uncertainties in estimating marginal and dependence model parameters.
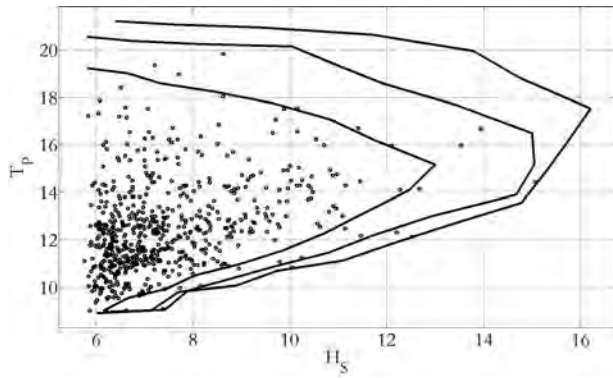
Fig. 6 Estimated contours of constant exceedance probability, one-sided in H_S, (equivalent to 10-, 100-, and 1000-yr H_S marginally), for measured Northern North Sea sample. Also shown is the sample.
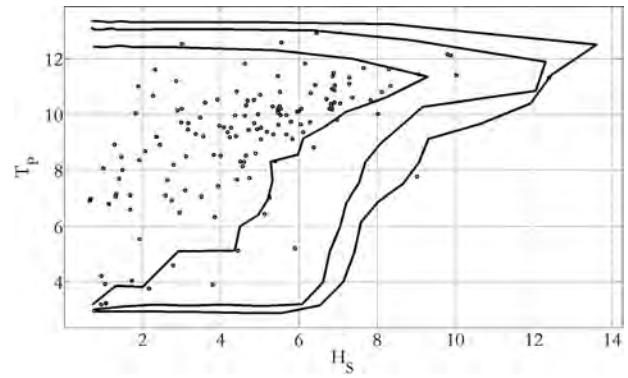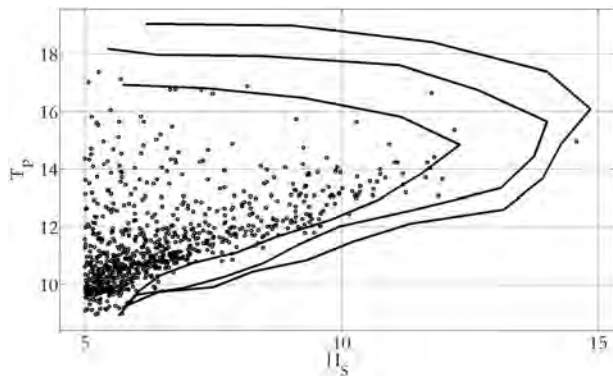

Fig. 7 Estimated contours of constant exceedance probability, one-sided in H_S, (equivalent to 10-, 100-, and 1000-yr H_S marginally), for hindcast Northern North Sea sample. Also shown is the sample.
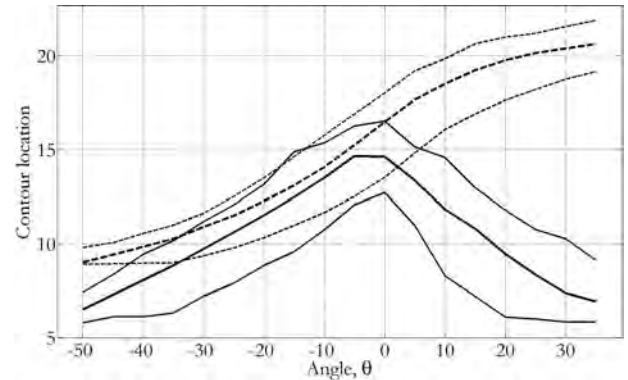

Fig. 8 Estimated contours of constant exceedance probability, one-sided in H_S, (equivalent to 10-, 100-, and 1000-yr H_S marginally), for measured Gulf of Mexico sample. Also shown is the sample.

Fig. 9 Estimated contours of constant exceedance probability, one-sided in H_S, (equivalent to 10-, 100-, and 1000-yr H_S marginally), for hindcast North West Shelf of Australia sample. Also shown is the sample.

Fig. 10 Estimated contour of constant exceedance probability, one-sided in H_S (equivalent to the 100-yr H_S marginally), for measured Northern North Sea sample, as a function of angle θ, for $\theta \in [0, 360)$. $H_S(\theta)$ (solid), and $T_P(\theta)$ (dashed). Thick lines correspond to the bootstrap median and thin lines to a bootstrap 95% uncertainty band.
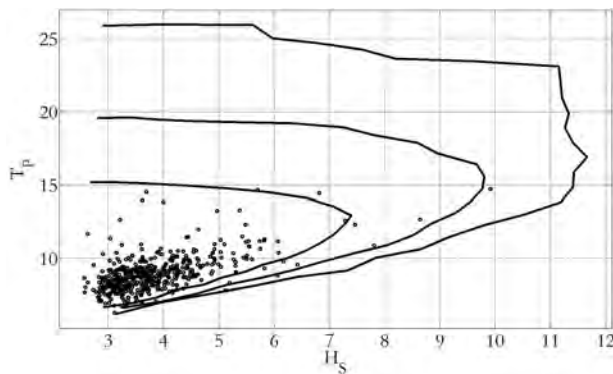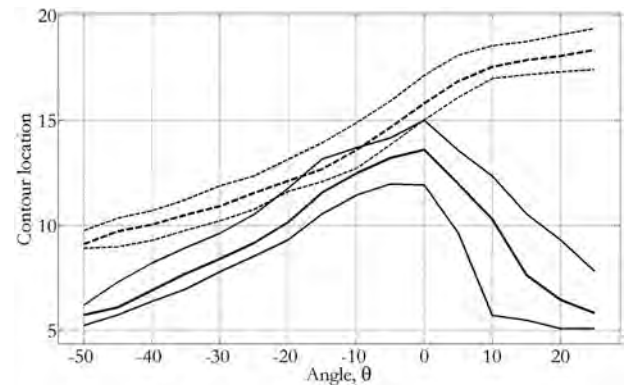

Fig. 11 Estimated contour of constant exceedance probability, one-sided in H_S (equivalent to the 100-yr H_S marginally), for hindcast Northern North Sea sample, as a function of angle θ, for $\theta \in [0, 360)$. $H_S(\theta)$ (solid), and $T_P(\theta)$ (dashed). Thick lines correspond to the bootstrap median and thin lines to a bootstrap 95% uncertainty band.

6 Discussion and Conclusions

For contours of types C1–C3 (see Sec. 3), based on an independent standard normal representation, the conditional extremes model can only provide an estimate for the part of the contour corresponding to large values of H_S (for which the conditional extremes model is valid). Nevertheless, in practice, interest in design contours usually focuses on sections of the contour corresponding to large values of at least one variable. Figures 14–17 show portions of the conventional FORM contours C1, estimated using the conditional extremes model, directly comparable to Figs. 10–13, in terms of median and bootstrap 95% uncertainty bands as a function of angle. For contours based on
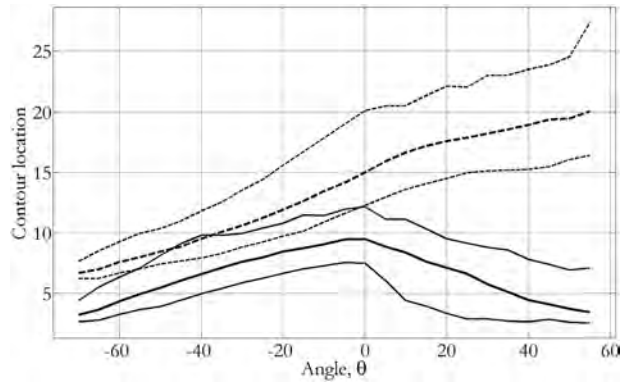
Journal of Offshore Mechanics and Arctic Engineering

Fig. 12 Estimated contour of constant exceedance probability, one-sided in H_S (equivalent to the 100-yr H_S marginally), for measured Gulf of Mexico sample, as a function of angle θ, for $\theta \in [0, 360)$. $H_S(\theta)$ (solid), and $T_P(\theta)$ (dashed). Thick lines correspond to the bootstrap median and thin lines to a bootstrap 95% uncertainty band.

Fig. 14 Estimated contour of constant probability density on standard normal scale (C1), equivalent to the 100-yr H_S marginally, for measured Northern North Sea sample, as a function of angle θ, for $\theta \in [0, 360)$. $H_S(\theta)$ (solid), and $T_P(\theta)$ (dashed). Thick lines correspond to the bootstrap median and thin lines to a bootstrap 95% uncertainty band.
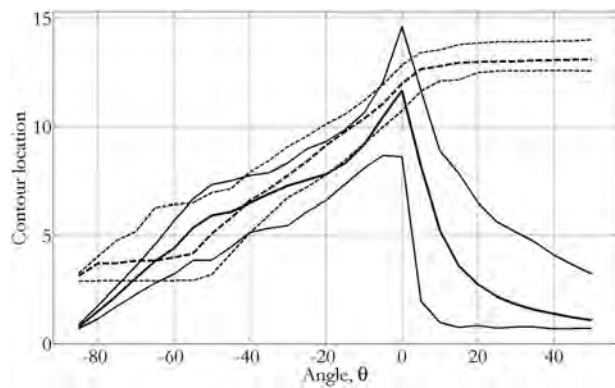

Fig. 13 Estimated contour of constant exceedance probability, one-sided in H_S (equivalent to the 100-yr H_S marginally), for hindcast North West Shelf of Australia sample, as a function of angle θ, for $\theta \in [0, 360)$. $H_S(\theta)$ (solid), and $T_P(\theta)$ (dashed). Thick lines correspond to the bootstrap median and thin lines to a bootstrap 95% uncertainty band.

Fig. 15 Estimated contour of constant probability density on standard normal scale (C1), equivalent to the 100-yr H_S marginally, for hindcast Northern North Sea sample, as a function of angle θ, for $\theta \in [0, 360)$. $H_S(\theta)$ (solid), and $T_P(\theta)$ (dashed). Thick lines correspond to the bootstrap median and thin lines to a bootstrap 95% uncertainty band.

exceedances on the original variable scale (see Sec. 3.6), some choices of reference location $\mathbf{r}_{X0}$ (e.g., the sample median) will result in contour estimates, which are discontinuous.

A comparison of the H_S–T_P model and conditional extremes models for joint estimation of extremes of (H_S, T_P) is given in Ref. [6].

Contours of constant exceedance probability estimated using the conditional extremes models (e.g., Figs. 6–9) are not smooth in general, since the conditional extremes model simulation is based on a resampling of the set of model residuals $\{\hat{z}_i\}_{i=1}^n$ (see Sec. 3.1 for $p = 2$), where n is the number of threshold exceedances available for modeling. Increasing simulation size will not produce smoother contours in general. These could be achieved by fitting a parametric model to residuals $\{\hat{z}_i\}_{i=1}^n$ and sampling from it, at the price of increased modeling complexity and additional modeling uncertainty which we have chosen to avoid.

As explained in the Appendix, the conditional extremes model can be used to model and simulate from p-dimensional distributions, $p > 2$; Ref. [9] uses such analysis for joint modeling of extreme current profiles (with $p = 32$). Contours of constant exceedance probability can be estimated in higher dimensions; computationally, the size of simulations increases with p and in this sense the analysis becomes more complex. To estimate a FORM-type contour in higher dimensions would require the specification and/or estimation of an increasing number of (usually
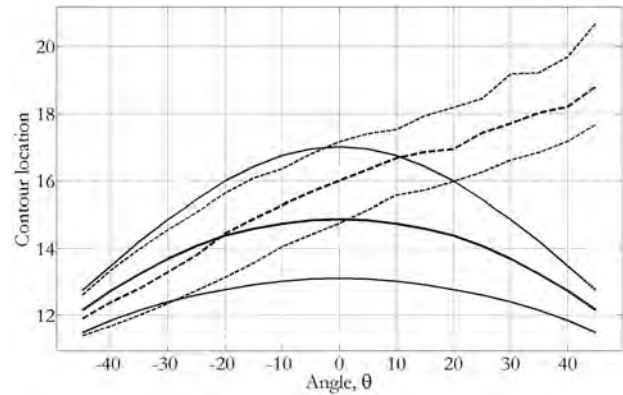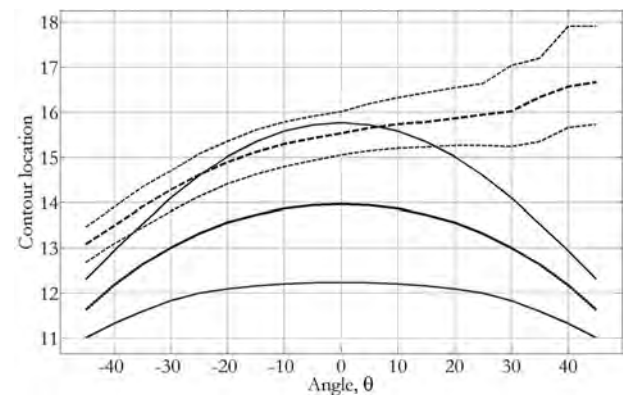
unknown) conditional distributions. Reference [10] discusses issues in application of first- and second-order reliability models in higher dimensions.

For the two independent standard normal variates U_1 and U_2, it is useful to note that a contour of constant probability density is defined by pairs of values (u_1, u_2) satisfying $u_1^2 + u_2^2 = c_D$, whereas a contour of constant exceedance probability is defined by expressions $(1 - \Phi(u_1))(1 - \Phi(u_2)) = c_E$, $(1 - \Phi(u_1))\Phi(u_2) = c_E$, $\Phi(u_1)(1 - \Phi(u_2)) = c_E$, and $\Phi(u_1)\Phi(u_2) = c_E$, where Φ is the cumulative distribution function of the standard normal distribution. c_D and c_E are a related pair of constants.

A key design objective is the estimation of structural integrity. Given the joint probability density $f_\mathbf{X}(\mathbf{x})$ of environmental and structural loading variables, and a failure surface (or limit-state function) $g(\mathbf{X})$, such that the structure failures when $g(\mathbf{X}) \leq 0$ and is safe otherwise, the failure probability of the structure can be expressed as $\int_{g(\mathbf{X}) \leq 0} f_\mathbf{X}(\mathbf{x})d\mathbf{x}$. In FORM analysis (see, e.g., Ref. [11]), this integral is approximated in terms of a single design point. Assuming that transformation of variables to independent standard normals $\mathbf{U}$ is possible (see, e.g., contour C1 in 3.3 above), the FORM design point is given by $\mathbf{u}^* = \arg \min(\mathbf{u}^T\mathbf{u}) s.t. g(\mathbf{u}) = 0$, and the failure probability by $\Phi(-\beta)$, where $\beta = \mathbf{u}^T\mathbf{u}$, where Φ is the cumulative distribution function
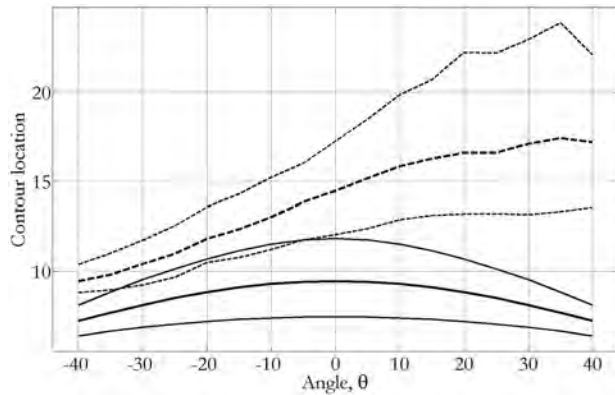
Fig. 16 Estimated contour of constant probability density on standard normal scale (C1), equivalent to the 100-yr H_S marginally, for measured Gulf of Mexico sample, as a function of angle θ, for $\theta \in [0, 360)$. $H_S(\theta)$ (solid), and $T_P(\theta)$ (dashed). Thick lines correspond to the bootstrap median and thin lines to a bootstrap 95% uncertainty band.

Fig. 17 Estimated contour of constant probability density on standard normal scale (C1), equivalent to the 100-yr H_S marginally, for hindcast North West Shelf of Australia sample, as a function of angle θ, for $\theta \in [0, 360)$. $H_S(\theta)$ (solid), and $T_P(\theta)$ (dashed). Thick lines correspond to the bootstrap median and thin lines to a bootstrap 95% uncertainty band.
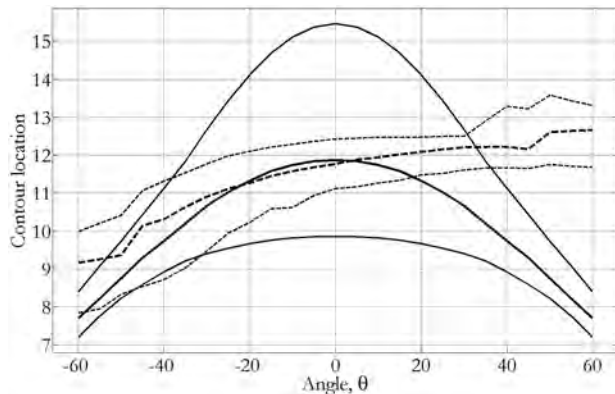
of the standard normal distribution. Using full simulation under the conditional extremes model, direct estimation of failure probability (and its uncertainty) is possible avoiding the need for design points. Were an estimate of probability density $f_\mathbf{X}(\mathbf{x})$ required, this could be achieved from the simulation (e.g., by estimating, smoothing and differentiating an empirical copula [12]).

In practice, the best choice of contour (including C1–C4 and many other possibilities) to employ depends on the requirements of the application. Analogy with the univariate definition of return value suggests that contours should be based on constant one-sided nonexceedance probability (such as C4). Within FORM-type procedures, involving the (often unjustified) assumption of a Gaussian dependence structure, contours of constant nonexceedance probability coincide with those of constant probability density. The analysis in this paper concentrates on characterizing the extremal dependence between storm peak events of H_S and T_P. From a structural risk perspective, particularly for floating structures exhibiting resonant responses to the ocean environment, we note that important combinations of environmental variables might occur away from the peaks of storms. For this reason, we might choose to include short-term environmental variability and examine extremal dependence structure on a sea-state basis (see, e.g., Ref. [13]). Alternatively, we might include structural

response variables (estimated as in Ref. [14]) within the conditional extremes model.

Bootstrap uncertainty analysis quantifies (sampling) uncertainties associated with model fitting but cannot estimate the inadequacy of model form (regardless of the model being fitted). Simply stated, we have no reason in general to believe that a model (such as the H_S–T_P model above) estimated using the body of a multivariate sample will be appropriate to characterize its extremes. However, as we move into the tail of the sample away from its body, it is difficult to estimate when exactly that model becomes unreliable. Conversely, the conditional extremes model has a solid foundation in the asymptotic theory of multivariate extreme values. However, it is nevertheless difficult to estimate how far into the tail of the data we must venture (with ever reducing sample size for modeling) so that the model becomes sufficiently reliable. These issues become more pressing as p increases—the curse of dimensionality. For serious application of the conditional extremes model, careful assessment of the sensitivity of estimates to fundamentals such as the choice of extreme value threshold (for all marginal and pairwise dependence estimation) and variable scale (for marginal modeling) are crucial. Incorporation of the effects of covariates such as location, direction, season, and time is vital for reliable application of both conditional extremes and FORM-type models.

In this paper, we consider various approaches to the specification of design contours for ocean engineering applications. We illustrate and compare these approaches under known model conditions. We introduce design contours based on the conditional extremes model of Ref. [7]. We show that the conditional extremes model can be used to estimate different forms of design contours at least in part, without the need for model form assumptions required for FORM-type analysis. Because of its asymptotic motivation, the conditional extremes model is likely to give better representations of design contours for joint extremes, compared to contours based on modeling the nonextreme body of the sample. In particular, we show that a full simulation under the conditional extremes model can be used to estimate contours of constant exceedance probability for the original variables (of type C4), and quantify the uncertainty with which that contour is estimated. We consider this to be an interesting supplement to ocean engineering methods.

Acknowledgment

Appendix

Conditional Extremes Model in p Dimensions. This section describes the conditional extremes model in p dimensions, extending the outline (for $p = 2$) in Sec. 3.1 of the main text. For a vector $\mathbf{Y}\ (= \{Y_j\}_{j=1}^p)$ of random variables with marginal Gumbel distributions, Ref. [7] derives a parametric form for the conditional distribution of the remaining variables $\mathbf{Y}_{-k}(= \{Y_j\}_{j=1, j\neq k}^p)$, given a large value y_k of the variable Y_k, at least as large as threshold w_{Y_k}, $k = 1, 2, ..., p$. The conditional model form for positively dependent variables $\mathbf{Y}$ is

$$(\mathbf{Y}_{-k}|Y_k = y_k) = \mathbf{a}_k y_k + y_k^{\mathbf{b}_k} \mathbf{Z}_k, \text{ where } y_k \geq w_{Y_k}, \text{ for } k = 1, 2, ..., p$$

where $\mathbf{a}_k\ (= \{a_{jk}\}_{j=1, j\neq k}^p)$ and $\mathbf{b}_k\ (= \{b_{jk}\}_{j=1, j\neq k}^p)$ are vectors of location and scale parameters, respectively, to be estimated, with element $a_{jk} \in [0, 1]$ and $b_{jk} \in (-\infty, 1)$, and y_k large, for $j = 1, 2, ..., p, j \neq k$, and $k = 1, 2, ..., p$, and componentwise multiplication is understood (i.e., for p-dimensional vectors $\mathbf{r} = \{r_i\}_{i=1}^p$ and $\mathbf{s} = \{s_i\}_{i=1}^p$, then $\mathbf{rs} = \{r_i s_i\}_{i=1}^p$). $\mathbf{Z}_k$ is a vector random

variable, independent of Y_k, converging with increasing y_k to a nondegenerate limiting distribution G_k. For each k, $k = 1,2, ..., p$, joint tail behavior conditional on Y_k is characterized by $\mathbf{a}_k$, $\mathbf{b}_k$, and G_k. The form of distribution G_k is not specified by theory. When two variables Y_j, Y_k are negatively dependent (i.e., $a_{jk} = 0, b_{jk} < 0$), the form of the conditional relationship is slightly different, but still amenable to empirical modeling

$$(\mathbf{Y}_{-k}|Y_k = y_k) = \mathbf{c}_k - \mathbf{d}_k \log_e(y_k) + y_k^{\mathbf{b}_k}\mathbf{Z}_k, \text{ where } y_k \geq w_{Y_k}$$

for $k = 1,2, ..., p$, where $\mathbf{c}_k \,(= \{c_{jk}\}_{j=1,j\neq k}^p)$ and $\mathbf{d}_k \,(= \{d_{jk}\}_{j=1,j\neq k}^p)$ have elements such that $c_{jk} \in (-\infty, \infty)$ and $d_{jk} \in [0,1]$. For a sample $\{y_{ij}\}_{i=1,j=1}^{n,p}$ of values from $\mathbf{Y}$ with values of conditioning variate Y_k exceeding an appropriate threshold w_{Y_k}, the values of $\mathbf{a}_k$, $\mathbf{b}_k$, and G_k are estimated using regression. For simplicity and computational ease during model fitting, G_k is assumed to be a multivariate normal distribution with mean $\mu_{\mathbf{Z}k}$ and diagonal variance–covariance matrix with elements $\sigma_{\mathbf{Z}k}^2$ treated as nuisance parameters, for $k = 1,2, ..., p$. These assumptions resolve the modeling procedure into fitting of $p-1$ pairwise models for pairs (Y_j, Y_k), $j = 1,2, ..., p, j \neq k$, for each of p conditioning variates Y_k, $k = 1,2, ..., p$. Fitted residuals

$$\hat{z}_{ijk} = \frac{(y_{ij} - \hat{a}_{jk}y_{ik})}{y_{ik}^{\hat{b}_{jk}}}, \quad i = 1,2,3,...,n, \quad j,k = 1,2,...,p, j \neq k$$

provide a dependent sample from the multivariate distribution G_k, which is then resampled (with replacement) in subsequent simulations. The adequacy of model fit can be assessed by (1) demonstrating that the values $\{\hat{z}_{ijk}\}_{i=1}^n$ and $\{y_{ik}\}_{i=1}^n$ are not obviously dependent, (2) exploring the effect of varying w_{Y_k} on $\hat{a}_{jk}, \hat{b}_{jk}$ and subsequent estimates (e.g., of probabilities associated with extreme sets), and (3) bootstrap resampling to estimate the uncertainty of estimates for $\mathbf{a}_k$, $\mathbf{b}_k$, and subsequent estimates for a given threshold choice w_{Y_k} (see e.g., Ref. [6]).

Marginal modeling of the original data (in terms of variables $\mathbf{X}$ $(= \{X_j\}_{j=1}^p)$) proceeds independently per variable as described in the main text, as does transformation between original and Gumbel variates.

Conditional Simulation. To simulate a random drawing $\{x_{sj}\}_{j=1}^p$ from the conditional distribution $\mathbf{X}_{-k}|X_k > w_{Xk}$, the following procedure can then be followed:

(1) draw a value y_{sk} of Y_k at random from the standard Gumbel distribution
(2) draw a set of values $\{z_{sjk}\}_{j=1,j\neq k}^p$ of $\mathbf{Z}_{-k}$ at random from the set $\{\hat{z}_{ijk}\}_{i=1,j=1,j\neq k}^{n,p}$ generated during model fitting
(3) calculate that value of $y_{sj}|y_{sk} = \hat{a}_{jk}y_{sk} + y_{sk}^{b_{jk}}z_{sjk}$ for $j = 1,2, ..., p, j \neq k$
(4) transform $\{y_{sj}\}_{j=1}^p$ to $\{x_{sj}\}_{j=1}^p$ using the probability integral transform and the estimated GP marginal model parameters.

To use this scheme to generate a multivariate dependent sample, we must first estimate the relative frequency with which to condition on each variable Y_k (or X_k), $k = 1,2, ..., p$, in turn. This is achieved as follows. Note that the sample $\{y_{ij}\}_{i=1,j=1}^{n,p}$ corresponds to a subset E of the original sample, for which the value of at least one variable (for each observation) exceeds its corresponding threshold (w_{Y_k} for variable k). We can therefore partition E into the union of sets $\{E_k\}_{k=1}^p$

$$E_k = \{\{y_{ij}\}_{i=1,j=1}^{n,p} s.t. y_{ik} > y_{ij}; j = 1,2,...,p, j \neq k; i = 1,2,...,n\}$$

so that for observations in set E_k, the value of variable k is more extreme in its marginal distribution than the value of any other variable. Since the conditional extremes model is motivated asymptotically, it is most appropriately applied to the conditioning variate whose value is most extreme in its distribution. We therefore use the relative frequency of observations in the sets $\{E_k\}_{k=1}^p$ to determine the rate at which we condition on each variable Y_k, $k = 1,2, ..., p$, during conditional simulation. See Ref. [7] for further discussion.

A number of applications and extensions of the conditional extremes approach have been reported. Reference [15] discusses the evaluation of environmental joint extremes for the offshore industry, including the incorporation of covariates (such as storm direction) within the conditional extremes model (see Refs. [16] and [17] for further detail). The conditional extremes model has been used for estimation of extreme flood risk (e.g., Ref. [18]) and extreme weather events (e.g., Ref. [19]). Recently, Ref. [20] proposed an empirical Bayes extension.

References

[1] Madsen, H. O., Krenk, S., and Lind, N. C., 1986, *Methods of Structural Safety*, Prentice-Hall, Englewood Cliffs, NJ.

[2] Winterstein, S. R., Ude, T. C., Cornell, C. A., Bjerager, P., and Haver, S., 1993, "Environmental Parameters for Extreme Response: Inverse FORM With Omission Factors," 6th International Conference on Structural Safety and Reliability, Innsbruck, Austria, August 9–13, pp. 551–557.

[3] Winterstein, S., and Engebretsen, K., 1998, "Reliability-Based Prediction of Design Loads and Responses for Floating Ocean Structures," 27th International Conference on Offshore Mechanics and Arctic Engineering, Lisbon, Portugal.

[4] Winterstein, S. R., Jha, A. K., and Kumar, S., 1999, "Reliability of Floating Structures: Extreme Response and Load Factor Design," J. Waterw., Port, Coastal, Ocean Eng., **125**, pp. 163–169.

[5] Nerzic, R., Frelin, C., Prevosto, M., and Quiniou-Ramus, V., 2007, "Joint Distributions of Wind/Waves/Current in West Africa and Derivation of Multivariate Extreme I-FORM Contours," 16th International Offshore and Polar Engineering Conference, Lisbon, Portugal.

[6] Jonathan, P., Flynn, J., and Ewans, K. C., 2010, "Joint Modeling of Wave Spectral Parameters for Extreme Sea States," Ocean Eng., **37**, pp. 1070–1080.

[7] Heffernan, J. E., and Tawn, J. A., 2004, "A Conditional Approach for Multivariate Extreme Values," J. R. Stat. Soc. B, **66**, pp. 497–546.

[8] Haver, S., and Nyhus, K., 1986, "A Wave Climate Description for Long Term Response Calculations," 5th International OMAE Symposium, Tokyo, Vol. 4, pp. 27–34.

[9] Jonathan, P., Ewans, K. C., and Flynn, J., 2012, "Joint Modeling of Vertical Profiles of Large Ocean Currents," Ocean Eng., **42**, pp. 195–204.

[10] Adhikari, S., 2005, "Asymptotic Distribution Method for Structural Reliability Analysis in High Dimensions," Proc. R. Soc. A, **461**, pp. 3141–3158.

[11] Adhikari, S., and Langley, R. S., 2002, "Reduction of Random Variables in Structural Reliability Analysis," 3rd International Conference on Mathematical Methods in Reliability Methodology and Practice, Trondheim, Norway, June 17–20.

[12] Nelsen, R. B., 1998, "*An Introduction to Copulas (Lecture Notes in Statistics)*," Springer, New York, p. 236.

[13] Feld, G., Randell, D., Wu, Y., Ewans, K., and Jonathan, P., 2014, "Estimation of Storm Peak and Intra-Storm Directional-Seasonal Design Conditions in the North Sea," 33rd International Conference on Ocean, Offshore, and Arctic Engineering, San Francisco, CA, June 8–13.

[14] Tromans, P. S., and Vanderschuren, L., 1995, "Risk Based Design Conditions in the North Sea: Application of a New Method," Offshore Technology Conference, (OTC 7683), Houston, TX.

[15] Ewans, K. C., and Jonathan, P., 2014, "Evaluating Environmental Joint Extremes for the Offshore Industry," J. Mar. Syst., **130**, pp. 124–130.

[16] Jonathan, P., Ewans, K. C., and Randell, D., 2013, "Joint Modeling of Environmental Parameters for Extreme Sea States Incorporating Covariate Effects," Coastal Eng., **79**, pp. 22–31.

[17] Jonathan, P., Ewans, K. C., and Randell, D., 2014, "Non-Stationary Conditional Extremes of Northern North Sea Storm Characteristics," Environmetrics, **25**, pp. 172–188.

[18] Keef, C., Papastathopoulos, I., and Tawn, J. A., 2013, "Estimation of the Conditional Distribution of a Vector Variable Given That One of its Components is Large: Additional Constraints for the Heffernan and Tawn Model," J. Multivar. Anal., **115**, pp. 396–404.

[19] Gilleland, E., Brown, B. G., and Ammann, C. M., 2013, "Spatial Extreme Value Analysis to Project Extremes of Large-Scale Indicators for Severe Weather," Environmetrics, **24**, pp. 418–432.

[20] Cheng, L., Gilleland, E., Heaton, M. J., and AghaKouchak, A., 2013, "Empirical Bayes Estimation for the Conditional Extreme Value Model," http://www.rap.ucar.edu/staff/ericg/ChengEtAl2014.pdf