

# covXtreme : MATLAB software for non-stationary penalised piecewise constant marginal and conditional extreme value models

(Second submission, March 2024)

Ross Towe<sup>a</sup>, Emma Ross<sup>b</sup>, David Randell<sup>b</sup>, Philip Jonathan<sup>a,c</sup>

<sup>a</sup>Shell Research Limited, London SE1 7NA, United Kingdom.

<sup>b</sup>Shell Global Solutions International BV, 1031 HW Amsterdam, The Netherlands.

<sup>c</sup>Department of Mathematics and Statistics, Lancaster University LA1 4YF, United Kingdom.

---

## Abstract

The `covXtreme` software provides functionality for estimation of marginal and conditional extreme value models, non-stationary with respect to covariates, and environmental design contours. Generalised Pareto (GP) marginal models of peaks over threshold are estimated, using a piecewise-constant representation for the variation of GP threshold and scale parameters on the (potentially multidimensional) covariate domain of interest. The conditional variation of one or more associated variates, given a large value of a single conditioning variate, is described using the conditional extremes model of Heffernan and Tawn (2004), the slope term of which is also assumed to vary in a piecewise constant manner with covariates. Optimal smoothness of marginal and conditional extreme value model parameters with respect to covariates is estimated using cross-validated roughness-penalised maximum likelihood estimation. Uncertainties in model parameter estimates due to marginal and conditional extreme value threshold choice, and sample size, are quantified using a bootstrap resampling scheme. Estimates of environmental contours using various schemes, including the direct sampling approach of Huseby et al. 2013, are calculated by simulation or numerical integration under fitted models. The software was developed in MATLAB for metocean applications, but is applicable generally to multivariate samples of peaks over threshold data. The software and case study data can be downloaded from GitHub, with an accompanying user guide.

---

## 1. Introduction

Reasonable estimation of the characteristics of extreme environments is essential to quantify natural hazards, and assure the reliability and safe operability of man-made infrastructure. For example, the extreme ocean environment can be thought of as a multivariate random process in space and time, describing interacting wind, wave, current, tide and surge phenomena. Accurate theoretical modelling of *extremes* from the ocean environment is problematic, because of the complexity of the numerical calculations involved. Practical risk assessment for marine and coastal structures and operations therefore typically requires the development of a statistical model from observational data for the joint behaviour of multiple meteorological-oceanographic (“metocean”) variables.

Physical intuition and previous work has demonstrated that the marginal distribution of a variable such as ocean storm severity (quantified in terms of significant wave height,  $H_S$ ) is dependent on covariates such as the direction in which the storm evolves, and the time of year in which the storm occurs (e.g. Randell et al. 2016). Over the long term, there is evidence that storm severity also evolves slowly in time due to anthropogenic climate change as well as natural climate cycles (e.g. Ewans and Jonathan 2023). It is important therefore that any statistical model for the extreme environment captures covariate non-stationarity. The nature and extent of covariate dependence varies across metocean variables. Experience also suggests that it is reasonable to assume that typical metocean variables are drawn from so-called “max-stable” distributions (e.g. Coles 2001). Therefore, typically conditional on covariates, the conditional distribution of threshold exceedances will be well approximated by the generalised Pareto (GP) distribution, provided that the threshold level is sufficiently high. This suggests that the marginal characteristics of metocean variables can be described reasonably using a non-stationary GP model (e.g. Davison and Smith 1990, Chavez-Demoulin and Davison 2005, Randell et al. 2016).

In addition we require that the statistical model also describes the *joint* tail of all metocean variables in general. However the nature of extremal dependence between different metocean variables is generally unknown. The specification of the statistical model therefore needs to be sufficiently general to admit different extents of extremal dependence, the specifics of which are then estimated by fitting the model to data. The conditional extremes model of Heffernan and Tawn (2004) is an attractive candidate, because it admits different classes of extremal dependence, it has a simple

form and is relatively easily interpretable. There is also evidence that the nature and extent of extremal dependence also varies systematically with covariates (e.g. Jonathan et al. 2013, Ross et al. 2018, Shooter et al. 2021).

The sizes of data samples available (e.g. from direct observation of the environment) for modelling are typically small relative to the time-scales of the inference task. For example, in engineering design we typically need to characterise events which occur on average once in 1000 or 10000 years based on a sample of at most the order of 100 years of observation. In these circumstances, our estimate of the joint tail of the distribution of metocean variables is typically uncertain. It is critical that this uncertainty is captured and incorporated appropriately in our risk assessment.

For individual metocean variables, the size of a rare occurrence is often quantified in terms of a *return value* associated with some return period  $T$  (=1000 years, say). The return value is defined straightforwardly as the quantile of the distribution of the annual maximum of the metocean variable with exceedance probability  $1/T$ , or (for large  $T$ , see Jonathan et al. 2021) the quantile of the distribution of the  $T$ -year maximum with non-exceedance probability  $\exp(-1)$ . In engineering design, the joint tail of the distribution of two (or sometimes three) metocean variables is often quantified in terms of an *environmental design contour*; points on the contour are “equally rare” with respect to some criterion, related to the joint cumulative distribution function or density of the variables (e.g. Ross et al. 2020, Haselsteiner et al. 2021). Therefore estimates for the distribution of the annual maximum or that of the  $T$ -year maximum for individual metocean variables, and environmental contours for pairs or triplets of variables, are key outputs from the statistical analysis.

There are many software tools available for applied extreme value analysis; Stephenson and Gilleland (2006), Gilleland et al. (2013), Gilleland (2016) and Belzile et al. (2023) provide reviews. In the statistical software literature, the R packages `texmex` (Southworth et al. 2024), `evgam` (Youngman 2022) and `mgcv` (Wood 2023) provide useful functionality for extreme value analysis. The discussion of Belzile et al. (2023) notes the lack of good “off the shelf” software for practitioners, particularly incorporating more recent advances in methodology. Some software has been published for environmental contour estimation (e.g. Haselsteiner et al. 2019, Mackay and de Hauteclocque 2023). The purpose of the `covXtreme` software introduced here is to provide the environmental practitioner with a straightforward means to estimate marginal and joint tails of distributions of random variables, and quantify extremes in terms of return values and environmental design contours. `covXtreme` accommodates covariate non-stationarity in both marginal and dependence behaviour, provides flexibility in estimating extremal dependence, and careful uncertainty quantification in inferences.

The sophistication of the methodology underpinning `covXtreme` has been set deliberately, based on the authors’ experience of metocean applications, to accommodate necessary effects in a pragmatic manner, whilst avoiding unnecessary complexity. Extreme value analysis is in some senses a less mature research area (e.g. Davison and Huser 2015) with numerous pressing challenges. For example, there are many approaches to modelling non-stationary marginal extremes; the work of Biller (2000), Wood (2004), Brezger and Lang (2006), Huser and Genton (2016), Wood et al. (2016), Wood et al. (2017), Youngman (2019), Youngman (2020), Shao et al. (2022) provide examples. With sufficiently rich data, complex extreme value models can be well estimated. However, in practice, typical samples tend only to support the adoption of relatively simple covariate models; for example, Zanini et al. (2020) show that a simple piecewise constant covariate representation (used by `covXtreme`) provides marginal inferences competitive with those from more sophisticated P-spline (e.g. Eilers and Marx 2010) and Bayesian adaptive regression spline models (e.g. Biller 2000) for a metocean application.

Elements of the `covXtreme` methodology, such as marginal and dependence modelling for extreme values, uncertainty quantification and environmental risk decision support, are reflected in recent developments in the environmental software and modelling literature. These typically address modelling of extremes of rainfall, river levels, temperatures and heatwaves, but sometimes less common environmental variables such as urban black carbon (Maciejewska et al. 2015). For example, regarding marginal extreme value analysis, De Luca and Napolitano (2023) introduce the `EXTRASTAR` software for modelling of time-series of annual maxima, applied to rainfall. Zhao et al. (2022) use the generalised extreme value distribution to estimate return values of rainfall. Diez-Sierra and del Jesus (2017) use extreme value analysis of block maxima within their `MENSEI-L` software. Souaissi et al. (2023) exploit extreme value analysis of block maxima in their model for river temperatures. Cobos et al. (2022) offer extreme value functionality within their `MarineTools.temporal` package. Minguez et al. (2010) discuss non-stationary extreme value analysis of block maxima, using the Akaike Information Criterion (AIC) for optimal covariate parameterisation. Zhang et al. (2022) consider non-stationary extreme value analysis of block maxima for heat wave tracking. Regarding joint modelling, Cheng et al. (2023) consider both marginal extreme value analysis and dependence modelling using copulas in their development of a drought index, and Liu et al. (2021) consider vine copulas in their model for water level prediction. Hao et al. (2017) develop the R package `drought` incorporating marginal and joint extreme value modelling using copulas, for applications to modelling and assessment. Regarding uncertainty quantification, Villa et al. (2023) present a stochastic model for future extreme temperatures, for decision support in infrastructure analysis, and the `ValueDecisions` software of Haag et al. (2022) considers uncertainty quantification for decision making, including

some elements of extreme value analysis. More generally, Čampulová et al. (2022) discuss extreme value analysis in the context of identifying outliers in environmental time-series, and Vrugt (2016) offers extreme value functionality in their DREAM software for Bayesian inference using Markov chain Monte Carlo.

An early version of the `covXtreme` software was published alongside Ross et al. (2018) for modelling of storm surge, and Ross et al. (2020) for estimation of environmental contours. Bore et al. (2019) suggest that `covXtreme` is useful for analysis of extreme ocean current profiles with depth. Vanem et al. (2020) use `covXtreme` to compare different contour methods with response-based methods for extreme ship response analysis. Guerrero et al. (2023) have used their enhancement of an earlier version of `covXtreme` for analysis of electrical signals in the human brain. The software has motivated the development of analogous prototype software PPL for marginal extreme value modelling with penalised piecewise-linear covariate representations (Barlow et al. 2023). The software has also been used as a pre-processor for transformation of data to standard marginal scale, allowing joint and conditional extreme value analysis (e.g. Shooter et al. 2021; Shooter et al. 2022), and in metocean consultancy work.

### *Target audience for covXtreme*

`covXtreme` was developed for oceanographic applications, but it can and has been used more widely. We believe that the reader will find `covXtreme` potentially useful, if (a) there is interest in quantifying extremes using statistical analysis of a data set, and possibly estimation of extreme quantiles or “return levels”; (b) the data are not stationary, and vary with known covariates; (c) there is interest in quantifying extremes of multiple variables at the same time; and (d) there is a need to quantify uncertainty carefully. Under these conditions, `covXtreme` offers a pragmatic but statistically sound analysis.

### *Objectives and layout of article*

The objective of the current article is to provide motivation and description of the `covXtreme` software, and illustrations of its use in the development of design conditions for ocean engineering. The layout of the article is as follows. Section 2 provides an overview of the software and the statistical methodology on which it is based. Sections 3 and 4 present case studies, involving a bivariate response and single covariate (Section 3), and trivariate response with 2-D covariate (Section 4). An accompanying user guide (available at Towe et al. 2023b) provides a detailed step-by-step description of developing a `covXtreme` model for ocean engineering data sets provided with the software.

## **2. Overview of software and statistical methodology**

`covXtreme` provides simple software for multivariate non-stationary extreme value analysis of peaks over threshold. The modelling framework of `covXtreme` (a) is statistically straightforward to understand and use without sacrificing rigour, (b) is computationally efficient, (c) provides good estimation of key quantities of interest to the extreme value practitioner, and (d) provides realistic quantification of modelling uncertainties.

Extreme value analysis of peaks over threshold of metocean variables requires the satisfactory completion of a number of analysis steps. Inference typically involves using a sample of time-series data for a number of variables (corresponding to some period  $P$  years of observation) as the basis for estimating the characteristics of extreme environments. These might include return values, associated values (e.g. Towe et al. (2023a)) and design contours (e.g. Haselsteiner et al. 2021) corresponding to a return period of  $T$  years,  $T \gg P$ ). For reasonable extreme value inference of a single variable (such as significant wave height), key features of the data must be accommodated carefully in the analysis, including the serial correlation of time-series, the dependence of values of metocean variables on covariates such as direction and season. Appropriate model forms for tails of distributions should also be adopted. For peaks over threshold analysis of a single variable, choice of threshold level can be source of considerable uncertainty. Moreover, different metocean variables (such as wind speed and significant wave height) are likely to be intercorrelated, and the nature of this dependence must be characterised carefully especially for extremes of one or more variables, using appropriate model forms; in these models, choice of threshold is again not always straightforward. Further, since samples of peaks over threshold are generally small (e.g. for high thresholds), it is critical to quantify uncertainty in return values and associated extremal quantities thoroughly using well-understood estimators.

The structure of the `covXtreme` software is written to reflect the sequential nature of practical extreme value analysis, as a set of MATLAB functions (see Section 2.6), illustrated in the workflow of Figure 1. For a given application, the five stages are executed sequentially, leading the user through the inference, addressing important analysis considerations in order. A typical analysis stage involves the specification of tuning parameters, the appropriate choice of which is informed by diagnostic information generated at that stage; generally it will therefore be necessary to repeat the analysis stage until satisfactory diagnostics are achieved, before moving on to the next analysis stage.

Given a sample of multivariate time-series of metocean variables, a typical analysis would proceed as described below. Note, for ease of reference, that subsection numbering 2.1-2.5 here corresponds to the numbering of stages 1-5 of the analysis procedure, and also to the numbering of sections in the `covXtreme` user guide.

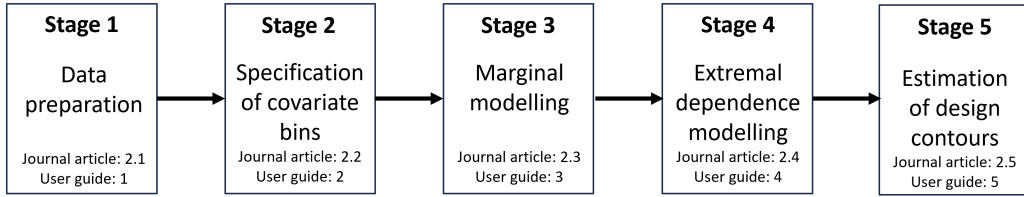


Figure 1: Schematic of `covXtreme` workflow. The five analysis stages are executed in sequence. Generally, at each stage, model tuning based on refinement of modelling hyper-parameters is necessary, guided by diagnostic information, before the next stage is attempted. Section numbers in the main article and user guide, providing relevant details for each stage of analysis, are given.

### 2.1. Data preparation

The main aspect of data preparation is the isolation of storm peak events. Storm peak events correspond to peaks over threshold for a dominant variable (e.g. significant wave height for a wave-dominated metocean application), *associated* values of all other variables of interest (e.g. wind speed) per event, and storm peak covariate values for all relevant covariates (e.g. direction and season). Storm peak events are isolated using the procedure described by Ewans and Jonathan (2008), and discussed in Section 1 of the `covXtreme` user guide. Storm peak events are then assumed to be conditionally independent given (storm peak) covariates.

Mathematically, consider a sample  $\dot{Y}_1(t_i), \dot{Y}_2(t_i), \dots, \dot{Y}_D(t_i)$  of multivariate time-series for  $D$  metocean variables observed at regularly-spaced time points  $t_i, i = 1, 2, 3, \dots$  over some period. Data preparation requires the isolation of a sample of values  $\{\dot{y}_{i1}, \dot{y}_{i2}, \dots, \dot{y}_{iD}\}_{i=1}^N$  of  $N$  storm peak events (by convention for the first variable) and associated events (for the remaining variables) summarising the peak characteristics of each of  $N$  storms observed, and corresponding storm peak covariate values  $\{x_{i1}, x_{i2}, \dots, x_{iC}\}_{i=1}^N$  for  $C$  storm peak covariates  $X_1, X_2, \dots, X_C$  defined on some domain  $\mathcal{X}$ . (Note that the “double dot” notation (e.g.  $\ddot{Y}$ ) indicates *time-series* variables from which *storm peak* events, indicated by “single dot” notation (e.g.  $\dot{Y}$ ) must be isolated.) Storm peak events are identified as local maxima (between successive up- and down-crossings of a given level) of  $\dot{Y}_1$ . Associated values for a storm are the values of the other random variables at the time of occurrence of the storm peak event. Storm peak and associated values are taken to be conditionally-independent given covariates, in the sense that  $\dot{y}_{id}$  can be viewed as an independent draw from  $\dot{Y}_d | (X_1 = x_{i1}, X_2 = x_{i2}, \dots, X_C = x_{iC})$ , for  $i = 1, 2, \dots, N, d = 1, 2, \dots, D$ .

Note that `covXtreme` also provides functionality to simulate data with known characteristics for checking of the performance of the statistical methodology.

### 2.2. Specification of covariate bins

`covXtreme` adopts a piecewise-constant parameterisation for the distribution of peaks over threshold on a partition of the covariate domain comprised of elements referred to as “covariate bins”. Covariate bins are not estimated as part of the analysis, but must be specified carefully by the user prior to the analysis, for reasonable inference. Briefly, the user is required to partition the domain of each covariate (independently) into a set of intervals. The resulting covariate bins over all covariates are then simply Cartesian products of these intervals, and hence a rectangular grid of covariate bins. `covXtreme` assumes, *within* a given covariate bin, that the statistical properties of a storm peak or associated variable are no longer dependent on covariates. Diagnostic plots illustrating features such as the variation of storm peak and associated variables with individual covariates, are provided to inform reasonable choice of marginal partitions. Illustrations of the specification of covariate bins are given in Section 2 of the `covXtreme` user guide.

Mathematically, each observation  $\dot{y}_{i1}, \dot{y}_{i2}, \dots, \dot{y}_{iD}$  is allocated to one of  $B$  covariate bins by means of an allocation vector  $A$ , with  $A(i) = b$  mapping observation (with index)  $i$  to bin (with index  $\mathcal{B} =$ )  $b, i = 1, 2, \dots, N, b = 1, 2, \dots, B$ . All observations within a specified covariate bin are assumed to have common extreme value characteristics, specified in terms of the parameters of the marginal model for peaks over threshold from that bin for each of the  $D$  components of the observation. Hence in particular, all threshold exceedances of storm peak variable  $\dot{Y}_d, d = 1, 2, \dots, D$  from covariate bin  $b$  can be viewed as independent draws from a generalised Pareto distribution with common shape, scale and threshold parameter values.

### 2.3. Marginal modelling

A two-stage modelling procedure is used to describe the marginal distribution of storm peak and associated variables. First a three-parameter gamma distribution is fitted independently to all the data for each covariate bin in turn, providing a good description of the bulk of the distribution of storm peak and associated variables on the covariate domain. The extreme value threshold for each covariate bin is then set to the quantile of the corresponding

gamma distribution with pre-specified non-exceedance probability. Finally the distribution of threshold exceedances per covariate bin is estimated using a non-stationary generalised Pareto distribution. Since generalised Pareto shape parameter is typically more difficult to estimate than scale, its value is assumed fixed but unknown across all covariate bins. Moreover, variation of the generalised Pareto scale parameter on the covariate domain is regulated using roughness penalisation, set to maximise the predictive performance of the generalised Pareto model on a hold-out sample within a cross-validation scheme. We judge from prior experience that these assumptions regarding the variation of generalised Pareto parameters are reasonable, and confirm using diagnostics during analysis that the assumptions give reasonably well-fitting models. By reducing the number of degrees of freedom for model fitting, the assumptions are a possible source of estimation bias, but also of a reduction in variance in estimated model parameters and return values. Further details and illustrations of marginal modelling are given in Section 3 of the `covXtreme` user guide.

Mathematically, for each variable  $\dot{Y}_d$ ,  $d = 1, 2, \dots, D$ , and covariate bin  $\mathcal{B} = b$ ,  $b = 1, 2, \dots, B$  independently, we estimate a three-parameter gamma distribution using maximum likelihood estimation with sample likelihood

$$\mathcal{L}_{\text{Gmm}}(\omega_{bd}, \kappa_{bd}, l_{bd}; \{\dot{y}_{id}\}_{i=1}^N) = \prod_{i:A(i)=b} f_{\text{Gmm}}(\dot{y}_{id}; \omega_{bd}, \kappa_{bd}, l_{bd}) \quad (1)$$

for sample  $\{\dot{y}_{id}\}_{i=1}^N$ , where  $f_{\text{Gmm}}$  is the density of the gamma distribution with shape  $\omega_{bd} \in \mathbb{R}$ , scale  $\kappa_{bd} > 0$  and location  $l_{bd} \in \mathbb{R}$  given by

$$f_{\text{Gmm}}(\dot{y}_{id}; \omega_{bd}, \kappa_{bd}, l_{bd}) = (\kappa_{bd}^{-\omega_{bd}} / \Gamma(\omega_{bd})) (\dot{y}_{id} - l_{bd})^{\omega_{bd}-1} \exp(-(\dot{y}_{id} - l_{bd}) / \omega_{bd}) \quad (2)$$

where  $\Gamma(\bullet)$  is the gamma function. In practice for fitting the gamma model, the location parameters  $\{l_{bd}\}$  are first estimated using a low empirical quantile of the sample per variate per covariate bin, and the remaining shape and scale parameters estimated using maximum likelihood. Next we calculate the extreme value threshold as  $\psi_{bd} = F_{\text{Gmm}}^{-1}(\tau_d; \hat{\omega}_{bd}, \hat{\kappa}_{bd}, \hat{l}_{bd})$  using the estimated gamma parameters, for the pre-specified non-exceedance probability  $\tau_d$ , where  $F_{\text{Gmm}}$  is the cumulative distribution function of the gamma distribution. The marginal sample likelihood for threshold exceedances of  $\psi_{bd}$  for variable  $\dot{Y}_d$  over all covariate bins can therefore be written

$$\mathcal{L}_{\text{GP}}(\xi_d, \{\nu_{bd}\}_{b=1}^B; \{\dot{y}_{id}\}_{i=1}^N, \{\psi_{bd}\}_{b=1}^B) = \prod_{b=1}^B \prod_{i:A(i)=b; \dot{y}_{id} > \psi_{bd}(\tau_d)} f_{\text{GP}}(\dot{y}_{id}; \xi_d, \nu_{bd}, \psi_{bd}) \quad (3)$$

where  $f_{\text{GP}}$  is the density of the generalised Pareto distribution with shape  $\xi_d \in \mathbb{R}$  and scale  $\nu_{bd} > 0$  given by

$$f_{\text{GP}}(\dot{y}_{id}; \xi_d, \nu_{bd}, \psi_{bd}) = (1/\nu_{bd}) [1 + (\xi_d/\nu_{bd})(\dot{y}_{id} - \psi_{bd})]_+^{-1/\xi_d - 1} \quad (4)$$

where  $[A]_+ = A$  when  $A > 0$ , and  $= 0$  otherwise. The corresponding cumulative distribution functions of the gamma and generalised Pareto distributions are

$$F_{\text{Gmm}}(\dot{y}_{id}; \omega_{bd}, \kappa_{bd}, l_{bd}) = (1/\Gamma(\alpha)) \gamma(\omega_{bd}, \kappa_{bd}(\dot{y}_{id} - l_{bd})) \quad (5)$$

where  $\gamma(\bullet, \bullet)$  is the lower incomplete gamma function, and

$$F_{\text{GP}}(\dot{y}_{id}; \xi_d, \nu_{bd}, \psi_{bd}) = 1 - [1 + (\xi_d/\nu_{bd})(\dot{y}_{id} - \psi_{bd})]_+^{-1/\xi_d} \quad (6)$$

**Optimal predictive performance** for the marginal generalised Pareto model is achieved using roughness-penalisation for the scale parameters  $\{\nu_{bd}\}_{b=1}^B$  across covariate bins. The corresponding penalised negative log likelihood takes the form

$$\begin{aligned} & -\log \mathcal{L}_{\text{GP}}^*(\xi_d, \{\nu_{bd}\}_{b=1}^B; \{\dot{y}_{id}\}_{i=1}^N, \{\psi_{bd}\}_{b=1}^B) \\ &= -\log \mathcal{L}_{\text{GP}}(\xi_d, \{\nu_{bd}\}_{b=1}^B; \{\dot{y}_{id}\}_{i=1}^N, \{\psi_{bd}\}_{b=1}^B) + \lambda_d \left( \frac{1}{B} \sum_{b=1}^B \nu_{bd}^2 - \left[ \frac{1}{B} \sum_{b=1}^B \nu_{bd} \right]^2 \right). \end{aligned} \quad (7)$$

The smoothness penalty  $\lambda_d$  controls the extent to which the generalised Pareto scale varies across covariate bins, and the penalty is proportional to the variance of  $\nu$  over covariate bins. Parameters can be estimated to minimise the penalised negative log likelihood for each variable  $\dot{Y}_d$ ,  $d = 1, 2, \dots, D$  in turn for each of a set of pre-specified values for  $\lambda_d$ . The optimal value  $\lambda_d^\circ$  of  $\lambda_d$  is chosen to maximise predictive likelihood for a hold-out sample within a  $k$ -fold cross-validation procedure. Gamma parameter estimates per covariate bin, and generalised Pareto parameter estimates evaluated using the full sample for  $\lambda_d = \lambda_d^\circ$  are carried forward to subsequent inference.

Since the numbers of parameters in the various marginal models above are relatively small, a simplex search procedure provides a straightforward approach to parameter estimation by minimisation of (penalised) negative log likelihoods.

**Threshold selection** for extreme value analysis of peaks over threshold is an important consideration (e.g. Northrop and Jonathan 2011, Scarrott and MacDonald 2012, Wadsworth 2016). Within a `covXtreme` analysis, multiple marginal models based on different random choices of threshold non-exceedance probabilities  $\tau_d \in \mathcal{I}_{\tau_d} \subseteq [0, 1)$  are constructed,  $d = 1, 2, \dots, D$ . The user’s task is to specify the interval  $\mathcal{I}_{\tau_d}$  for each variable  $\dot{Y}_d$ . As explained in Section 3 of the user guide, this choice is aided by numerous diagnostic plots, including examination of the stability of the estimated value of  $\xi_d$  as a function of  $\tau_d$ .

**Uncertainty quantification** for marginal inference is performed using a non-parametric bootstrap procedure. The original sample of storm peak and associated values is resampled with replacement. For each variable  $\dot{Y}_d$ ,  $d = 1, 2, \dots, D$ , the full marginal extreme value analysis is then repeated using the bootstrap resample together with a new random selection of  $\tau_d$ . The outcome of the complete marginal inference is then quantified in terms of sets of parameter estimates (  $\{\xi_d^r, \{\nu_{bd}^r, \psi_{bd}^r, \omega_{bd}^r, \kappa_{bd}^r, l_{bd}^r\}_{b=1}^B; \tau_d^r, \lambda_d^o\}_{d=1}^D$ ) for each of  $R$  (typically = 100 or 250) bootstrap resamples (where superscript  $r$  indicates an estimate from a resample). Typically, the value of  $\lambda_d^o$  estimated using the original sample is adopted for all bootstrap resamples also, although the option to estimate a new optimal roughness coefficient per bootstrap resample is provided.

The **distribution of the  $T$ -year maximum event** per covariate bin, and over all covariate bins can then be estimated (or sampled) under the fitted marginal model. From above, the full marginal model for storm peak or associated variate  $\dot{Y}_d$  in covariate bin  $\mathcal{B} = b$  is

$$F_{\text{GmmGP}}(y; \omega_{bd}, \kappa_{bd}, l_{bd}, \xi_d, \nu_{bd}, \psi_{bd}) = \begin{cases} F_{\text{Gmm}}(y; \omega_{bd}, \kappa_{bd}, l_{bd}) & \text{for } A(i) = b, \dot{y}_{id} \leq \psi_{bd}, \\ \tau_d + (1 - \tau_d)F_{\text{GP}}(y; \xi_d, \nu_{bd}, \psi_{bd}) & \text{for } A(i) = b, y > \psi_{bd}. \end{cases} \quad (8)$$

Then under the model, the distribution of a random occurrence of  $\dot{Y}_d$  from any covariate bin is

$$F_{\dot{Y}_d}(y; \{\omega_{bd}, \kappa_{bd}, l_{bd}, \nu_{bd}, \psi_{bd}\}_{b=1}^B, \xi_d) = \sum_{b=1}^B p_b F_{\text{GmmGP}}(y; \omega_{bd}, \kappa_{bd}, l_{bd}, \xi_d, \nu_{bd}, \psi_{bd}) \quad (9)$$

where  $p_b$  is an empirical estimate of the probability of observing a storm event in covariate bin  $b$ . If we further assume that the number  $N$  of storms in  $T$ -years is Poisson-distributed with mean  $T\rho$ , where  $\rho$  is an empirical estimate for the number of storms per annum, and suppressing parameter dependence for brevity, the distribution of the  $T$ -year maximum is simply

$$\begin{aligned} F_{\dot{Y}_d T\text{-year}}(y) &= \sum_{k=0}^{\infty} \mathbb{P}(N = k) F_{\dot{Y}_d}^k(y) = \sum_{k=0}^{\infty} (\exp(-T\rho)(T\rho)^k / k!) F_{\dot{Y}_d}^k(y) \\ &= \exp(-T\rho (1 - F_{\dot{Y}_d}(y))). \end{aligned} \quad (10)$$

We can use a similar approach to estimate and sample from the distribution of the  $T$ -year maximum of  $\dot{Y}_d$  for any combination of covariate bins, by restricting the set of covariate bins over which the summation is performed in Equation 9 (and linearly scaling the values of  $\{p_b\}$  so that they sum to unity).

#### 2.4. Extremal dependence modelling

Given non-stationary marginal models for storm peak and associated variables, we next seek to describe the nature of the dependence between them for large values of the storm peak variable. This is achieved using the conditional extremes model of Heffernan and Tawn (2004). Under the fitted conditional extremes model, we can then estimate the characteristics of the joint distribution of all associated variables given a large storm peak, and thereby estimate joint environmental design conditions and design contours. The conditional extremes model is specified for sets of variables on a common standard Laplace marginal scale, rather than their original physical scales. For this reason, a necessary first step is to transform the sample of storm peak and associated values to this scale, using the fitted marginal models. Incorporation of covariate effects is generally important for extremal dependence modelling, and these can be accommodated using conditional extremes models of different complexities with respect to covariates. The `covXtreme` software allows any number of the parameters of the conditional extremes model to vary with covariates. Nevertheless, experience suggests that the data indicate the need for covariate non-stationarity of just one model parameter (the “slope” parameter, see below) for typical metocean applications. We therefore describe this specific model form as a recommended “default” approach here. Further details and illustrations of extremal dependence modelling are given in Section 4 of the `covXtreme` user guide.

Mathematically, `covXtreme` provides a model for the joint conditional tail  $(\dot{Y}_2, \dot{Y}_3 \dots \dot{Y}_D | \dot{Y}_1 = \dot{y})$  for large  $\dot{y}$ , using an extension of the conditional extremes model of Heffernan and Tawn (2004) on standard Laplace (and optionally, as originally, on standard Gumbel) scale. Inference therefore requires that we transform the storm peak and associated value sample  $\{\dot{y}_{i1}, \dot{y}_{i2}, \dots, \dot{y}_{iD}\}_{i=1}^N$  of variables  $\{\dot{Y}_i, \dot{Y}_2, \dots, \dot{Y}_D\}$  to the corresponding Laplace scale sample  $\{y_{i1}, y_{i2}, \dots, y_{iD}\}_{i=1}^N$  for variables  $\{Y_1, Y_2, \dots, Y_D\}$ .

#### 2.4.1. Marginal transformation to standard Laplace scale

The marginal transformation to standard Laplace scale is achieved using the probability integral transform such that for  $i = 1, 2, \dots, N$ ,  $d = 1, 2, \dots, D$ ,  $b = 1, 2, \dots, B$

$$F_{\text{GmmGP}}(\dot{y}_{id}; \omega_{bd}, \kappa_{bd}, l_{bd}, \xi_d, \nu_{bd}, \psi_{bd}) = F_{\text{Lpl}}(y_{id}) \text{ for } A(i) = b \quad (11)$$

where  $F_{\text{GmmGP}}$  is the marginal cumulative distribution function of storm peak variable  $\dot{Y}_d$ , for sets of parameters of marginal gamma and generalised Pareto models from Equation 8, and  $F_{\text{Lpl}}$  is the cumulative distribution function of the standard Laplace distribution given by  $F_{\text{Lpl}}(y) = 0.5 \exp(-|y|)$  for  $y \leq 0$  and  $= 1 - 0.5 \exp(-|y|)$  otherwise.

#### 2.4.2. Conditional extremes modelling

The Laplace-scale sample  $\{y_{i1}, y_{i2}, \dots, y_{iD}\}_{i=1}^N$  from random variables  $\{Y_1, Y_2, \dots, Y_D\}$  is next characterised using the conditional extremes model, for values  $y$  of the conditioning variate  $Y_1$  above a dependence threshold  $\phi(\tilde{\tau}) = F_{\text{Lpl}}^{-1}(\tilde{\tau})$ , for which the conditional extremes model is assumed to hold, for carefully specified non-exceedance probability  $\tilde{\tau}$ . For  $y > \phi(\tilde{\tau})$ , in covariate bin indexed  $\mathcal{B} = b$ , the recommended non-stationary model takes the form

$$(Y_2, Y_3 \dots Y_D) | (Y_1 = y, \mathcal{B} = b) = (\alpha_{b2}, \alpha_{b3}, \dots, \alpha_{bD})y + y^{(\beta_2, \beta_3, \dots, \beta_D)} \mathbf{Z} \quad (12)$$

with linear slope parameters  $\alpha_{bd'} \in [-1, 1]$ ,  $d' = 2, 3, \dots, D$  varying across covariate bins, scalar exponent parameters  $\beta_{d'} \in [-\infty, 1]$  common to all covariate bins, and residual random variable  $\mathbf{Z} = (Z_2, Z_3, \dots, Z_D) \in \mathbb{R}^{D-1}$  whose distribution is unknown.

For estimation of slope and exponent parameters only, we assume that each component  $Z_{d'}$  of  $\mathbf{Z}$  is independently distributed according to

$$Z_{d'} = \mu_{d'} + \sigma_{d'} W_{d'} \quad (13)$$

for mean  $\mu_{d'} \in \mathbb{R}$ , scale  $\sigma_{d'} > 0$  and random variable  $W_{d'} \in \mathbb{R}$  following a generalised Gaussian (or delta-Laplace) distribution with zero mean, unit variance and shape  $\delta_{d'}$ .  $\mu_{d'}$  and scale  $\sigma_{d'}$  are treated as nuisance parameters in the estimation, and are not needed for subsequent inferences under the fitted model. For general mean  $m$  and variance  $s^2$ , the corresponding density  $f_{\text{GG}}$  of the generalised Gaussian distribution is

$$f_{\text{GG}}(w; m, s^2, \delta_{d'}) = \frac{\delta_{d'}}{2\kappa(\delta_{d'})s\Gamma(1/\delta_{d'})} \exp \left\{ - \left| \frac{w - m}{\kappa(\delta_{d'})s} \right|^{\delta_{d'}} \right\}, \quad (14)$$

where  $\kappa(\delta_{d'})^2 = \Gamma(1/\delta_{d'})/\Gamma(3/\delta_{d'})$ . The value of exponent  $\delta_{d'} \in \{1, 2\}$  is user-specified in `covXtreme`.  $\delta_{d'} = 1$  imposes a Laplace distribution (with zero mean and unit variance) on  $W_{d'}$ , appropriate when the dependence between  $Y_{d'}$  and  $Y_1$  is low.  $\delta_{d'} = 2$  imposes a standard Gaussian distribution on  $W_{d'}$ , the original assumption of Heffernan and Tawn (2004), appropriate otherwise. For estimation purposes therefore, from the properties of the generalised Gaussian distribution,  $Y_{d'} | (Y_1 = y)$  is assumed to follow a generalised Gaussian distribution with mean  $m_{bd'} = \alpha_{bd'}y + \mu_{d'}y^{\beta_{d'}}$  and standard deviation  $\zeta_{d'} = y^{\beta_{d'}}\sigma_{d'}$ .

The conditional dependence likelihood  $\tilde{\mathcal{L}}_{d'}$  for the model  $Y_{d'} | (Y_1 = y)$  and  $y > \phi(\tilde{\tau})$  can then be written

$$\tilde{\mathcal{L}}_{d'} \left( \{\alpha_{bd'}\}_{b=1}^B, \beta_{d'}, \mu_{d'}, \sigma_{d'}; \{y_{id}\}_{i=1, d=1}^{N, D}, \delta_{d'}, \tilde{\tau} \right) = \prod_{b=1}^B \prod_{\substack{i: b=A(i) \\ y_{i1} > \phi(\tilde{\tau})}} f_{\text{GG}} \left( y_{id}; \alpha_{bd'}y_{i1} + \mu_{d'}y_{i1}^{\beta_{d'}}, (y_{i1}^{\beta_{d'}}\sigma_{d'})^2, \delta_{d'} \right). \quad (15)$$

As for marginal models, we regulate the smoothness of  $\{\alpha_{bd'}\}_{b=1}^B$  optimally on the covariate domain using a cross-validation procedure, selecting a value  $\tilde{\lambda}_{d'}^*$  for roughness coefficient  $\tilde{\lambda}_{d'}$  in roughness-penalised negative log likelihood

$$\begin{aligned} & -\log \tilde{\mathcal{L}}_{d'} \left( \{\alpha_{bd'}\}_{b=1}^B, \beta_{d'}, \mu_{d'}, \sigma_{d'}; \{y_{id}\}_{i=1, d=1}^{N, D}, \delta_{d'}, \tilde{\tau} \right) \\ &= -\log \tilde{\mathcal{L}}_{d'} \left( \{\alpha_{bd'}\}_{b=1}^B, \beta_{d'}, \mu_{d'}, \sigma_{d'}; \{y_{id}\}_{i=1, d=1}^{N, D}, \delta_{d'}, \tilde{\tau} \right) + \tilde{\lambda}_{d'} \left( \frac{1}{B} \sum_{b=1}^B \alpha_{bd'}^2 - \left[ \frac{1}{B} \sum_{b=1}^B \alpha_{bd'} \right]^2 \right). \end{aligned} \quad (16)$$

to maximise predictive performance on cross-validation hold-out samples. Further, it is important to confirm reasonable choice of the dependence threshold  $\tilde{\tau}$  by inspection of diagnostic plots, for example of stability of estimated parameters as a function of threshold.

Two algorithms are provided to minimise the penalised negative log likelihood in Equation 16. The default is a Newton-Raphson approach exploiting gradients; the alternative is a simplex search procedure.

Once parameter estimates are available, we represent the distribution of the  $(D - 1)$ -dimensional residual random variable  $\mathbf{Z}$  using the set  $\mathcal{E}_b = \{e_{i'bd'}\}$ , for all  $d' = 2, 3, \dots, D$ , and all  $i' = 1, 2, \dots, N$  such that  $y_{i'1} > \phi(\tilde{\tau})$ ,  $b = A(i')$ , of residuals from the fit per covariate bin, with elements

$$e_{i'bd'} = \left( y_{i'd'} - \hat{\alpha}_{bd'} y_{i'1} - \hat{\mu}_{d'} y_{i'1}^{\hat{\beta}_{d'}} \right) / \left( \hat{\sigma}_{d'} y_{i'1}^{\hat{\beta}_{d'}} \right). \quad (17)$$

During subsequent simulation under the fitted conditional extremes model, these residuals are resampled jointly as  $\{e_{i'b2}, e_{i'b3}, \dots, e_{i'bD}\}$ , thereby preserving the dependence between them, and hence the dependence between variables  $Y_2, Y_3, \dots, Y_D | (Y_1 = y, y > \phi(\tilde{\tau}), \mathbf{B} = b)$  per covariate bin. Similarly, simulation below threshold  $\phi(\tilde{\tau})$  is achieved simply by resampling from the original Laplace-scale storm peak sample. We note that `covXtreme` also facilitates estimation of variants of this dependence model, for which any number of conditional extremes model parameters  $\alpha$ ,  $\beta$ ,  $\mu$  and  $\sigma$  are allowed to vary between covariate bins, and their overall roughness penalised for good performance by extending the penalty term in Equation 16 to include all “non-stationary” parameters; see Section 4.2 of the `covXtreme` user guide. It is also possible to pool estimates of residuals across covariate bins, useful when covariate bins with low occupancy are present.

**Uncertainty quantification** for extremal dependence inference is performed by extending the bootstrap procedure described in Section 2.3. Conditional extremes models are estimated for each of the bootstrap resamples based on which marginal models were estimated. The combined marginal and conditional extremes analysis therefore produces the set  $\{ \{ \xi_d^r, \{ \nu_{bd}^r, \psi_{bd}^r, \omega_{bd}^r, \kappa_{bd}^r, l_{bd}^r \}_{b=1}^B \}_{d=1}^D, \{ \{ \alpha_{bd'}^r \}_{b=1}^B, \beta_{d'}^r, \mu_{d'}^r, \sigma_{d'}^r \}_{d'=2}^D, \{ \mathcal{E}_b^r \}_{b=1}^B; \{ \tau_d^r, \lambda_d^o \}_{d=1}^D, \{ \tilde{\tau}_{d'}^r, \tilde{\lambda}_{d'}^o \}_{d'=2}^D \}$  of parameter estimates and residuals for each of  $R$  bootstrap resamples indexed by superscript  $r$ . As for marginal modelling, the value of roughness coefficient  $\tilde{\lambda}_{d'}^o$  is typically estimated using the original sample only, and adopted for all bootstrap resamples also, although the option to estimate a new optimal roughness coefficient per bootstrap resample is provided.

A sampling procedure is used to estimate the **conditional return value** distribution of the associated value for  $\dot{Y}_{d'}$  given a  $T$ -year maximum event for  $\dot{Y}_1$  per covariate bin, and over unions of covariate bins. The approach combines importance sampling from the marginal  $T$ -year maximum distribution of  $\dot{Y}_1$  (exploiting Equation 10) with Monte Carlo sampling under the fitted conditional extremes model, managing transformations between marginal and standard Laplace scales. Further discussion of conditional return values can be found in Towe et al. (2023a).

## 2.5. Estimation of design contours

Using the estimated marginal and conditional extremes models, `covXtreme` allows estimation of quantities such as return values for the dominant variable, conditional return values for associated variables, and environmental design contours. Haselsteiner et al. (2017), Ross et al. (2020) and Haselsteiner et al. (2021) provide recent reviews of contour estimation, encompassing a range of approaches. Algorithms to estimate three types of design contour are implemented in `covXtreme`. These are (a) constant exceedance contours; (b) direct sampling contours; and (c) conditional extremes (or “HT”, in acknowledgement of the authors of the conditional extremes model, Heffernan and Tawn) constant density contours. Each of the three contour methods estimates a line (in 2-D) or surface (in 3-D) on which certain characteristics of the distributions of  $\dot{Y}_2 | (\dot{Y}_1 = y)$  (in 2-D, or  $(\dot{Y}_2, \dot{Y}_3) | (\dot{Y}_1 = y)$  in 3-D) for large  $y$  are preserved. For example, the constant exceedance contour in 2-D is a line on which the probability of exceedance in an “outward” sense (e.g.  $\mathbb{P}(\dot{Y}_1 > \zeta_1, \dot{Y}_2 > \zeta_2)$  in the first quadrant) is constant, for points  $(\zeta_1, \zeta_2)$  on the contour; see Jonathan et al. (2014) for details. The direct sampling contour of Huseby et al. (2015) is similar, except that now the probability in the half plane defined by the tangent to the contour at any point of interest is constant, and the contour itself must be convex. In 2-D, the conditional extremes constant density contour defines a line on which the joint density of the pair  $(\dot{Y}_1, \dot{Y}_2)$  is constant. By construction, all contours pass through a so-called “lock point”, defined as an extreme quantile (e.g. the  $T$ -year return value) of dominant variable  $Y_1$  and the corresponding conditional median of associated variate  $Y_2 | (Y_1 = T\text{-year value})$ . The lock point defines the value of the distributional characteristic preserved on the contour. Further discussion and illustrations are provided in Section 5 of the `covXtreme` user guide. The contour methods discussed in this section are intended to illustrate the kinds of inferences which can be made routinely using simulation under a fitted `covXtreme` model. All the approaches discussed provide means to quantify the extent of some or all of a cloud of realisations. No claims are made regarding their usefulness in any particular field of application (e.g. Mackay and Haselsteiner 2021, Hafver et al. 2022). The user interested in making specific



inferences for particular structure variables, or choices of design contour, is free to implement these as alternatives or extensions to the existing Stage 5 of the analysis procedure

Mathematically, in 2-D, the constant exceedance contour  $\zeta(\theta) = (\zeta_1(\theta), \zeta_2(\theta))$  for  $\theta \in \Theta \subseteq [0, 360)$  can be defined by the equation

$$\mathbb{P} \left( \bigcap_{d=1}^2 \left( r_d(\theta) \dot{Y}_d > r_d(\theta) \zeta_d(\theta) \right) \right) = p \quad (18)$$

where  $\mathbf{r}^\circ = (r_1^\circ, r_2^\circ)$  is a reference location (see e.g. Jonathan et al. 2014), and  $r_d(\theta) = \zeta_d(\theta) - r_d^\circ$ ,  $d = 1, 2$ , for some small probability  $p > 0$ . We set the value of  $r_2^\circ$  equal to the conditional mean  $\dot{Y}_2 | (\dot{Y}_1 = T\text{-year value})$ , and the value of  $r_1^\circ$  to the minimum value of  $\dot{Y}_1$  in the sample, to ensure continuity of the  $T$ -year contour when  $\dot{Y}_1$  exceeds the sample minimum.

To estimate the direct sampling contour in two-dimensions for probability level  $\alpha$ , following Huseby et al. (2015), we first find the function  $C(\theta)$ , the  $(1 - \alpha)$ -quantile of the distribution of the projection  $\dot{Y}_1 \cos(\theta) + \dot{Y}_2 \sin(\theta)$  for each value of  $\theta \in [0, 2\pi)$

$$C(\theta) = \inf \left\{ c : \mathbb{P} \left( \left[ \dot{Y}_1 \cos(\theta) + \dot{Y}_2 \sin(\theta) \right] > c \right) = \alpha \right\}. \quad (19)$$

Then we estimate the contour  $\mathcal{C} = \{(\zeta_1(\theta), \zeta_2(\theta)) : \theta \in [0, 2\pi)\}$  using

$$\zeta_1(\theta) = C(\theta) \cos(\theta) - \frac{dC}{d\theta} \sin(\theta), \quad (20)$$

$$\zeta_2(\theta) = C(\theta) \sin(\theta) + \frac{dC}{d\theta} \cos(\theta), \quad (21)$$

and potentially further smooth  $\mathcal{C}$  as a function of  $\theta$ . Following Winterstein et al. (1993), for at  $T$ -year return period, it is recommended that the value of  $\alpha$  be set to  $1/T$ .

The conditional extremes constant density contour consists of the set of points  $\zeta \in \mathbb{R}^2$  for which the joint density  $f_{\mathbf{Y}}(\zeta) = c$  for some small value  $c > 0$ . Algorithms for estimation of the three contour types exploit importance sampling (e.g. Section 3.4 of Towe et al. 2021) for computational efficiency where possible. Contours can be estimated per covariate bin (e.g. as “directional” or “seasonal” contours), or integrated over covariate bins (e.g. to provide “omni” estimates over all bins).

The direct sampling contour is convex by definition, in the sense that it encloses a convex region (of  $\mathbb{R}^2$  for 2-D contours), possibly with the help of the coordinate axes. Depending on the nature of the joint density, the conditional extremes constant density “contour” could exist in the form of a set of disjoint curves, some of which may be closed. The constant exceedance and direct sampling contours are invariant to transformations of variables, whereas the conditional extremes constant density contour is not (e.g. Ross et al. 2020).

## 2.6. Accessing the software, and performing a typical analysis

The `covXtreme` is available for download from GitHub at Towe et al. (2023b). Software is written using MATLAB object-oriented programming. Full specifications of classes, properties and methods are therefore provided, and available for user inspection.

A typical analysis involves executing each of the MATLAB scripts `Stage1_PeakPicking`, `Stage2_SetBinEdges`, `Stage3_FitMargin` (for each variable marginally), `Stage4_FitHeffernanTawn` and `Stage5_Contour` in sequence, as illustrated by the workflow in Figure 1. As explained in detail in the `covXtreme` user guide, each stage of the analysis involves the specification of control parameters for that stage. Default values for control parameters are provided, but it is necessary for the user at each stage to assess whether values are appropriate by confirming that diagnostic plots generated have reasonable characteristics. It may be necessary to repeat a given stage multiple times until acceptable diagnostic characteristics are found, before proceeding to the next stage.

Sections 3 and 4 below provide brief descriptions of the application of `covXtreme` to a pair of random variables (specifically significant wave height and spectral peak period) varying with directional covariate (Section 3) and to four random variables (significant wave height, spectral peak period, wind speed and overturning moment) varying with 2-D directional-seasonal covariate.

## 3. Case study : Single covariate, bivariate response

This section outlines to application of `covXtreme` to estimation of design contours for extreme storm peak significant wave height and associated spectral peak period, both varying with storm direction. The analysis follows the five stages

discussed in Section 2. The data correspond to approximately 35 years of time-series output from a hindcast simulator for a location in the northern North Sea.

Following isolation of storm peak significant wave height  $H_S$ , associated spectral peak period  $T_P$  and storm peak direction using Stage 1, Figure 2 illustrates the output of Stage 2 of the analysis, showing storm peak  $H_S$  and associated  $T_P$  as a function of the direction from which storms emanate, in degrees clockwise from north. The figure also shows the user-input bin edges at  $0^\circ$ ,  $20^\circ$ ,  $60^\circ$ ,  $225^\circ$ ,  $270^\circ$  and  $315^\circ$ , specified so that the variation of  $H_S$  and  $T_P$  is approximately independent of direction within each bin, but different between bins. Thus, for example, the bin including  $45^\circ$  corresponds to the land shadow of Norway, with resulting low values of storm peak  $H_S$ , whereas the bins including  $250^\circ$  and  $340^\circ$  contain large storms from the Atlantic Ocean and Norwegian Sea. Figure 3 provides scatter

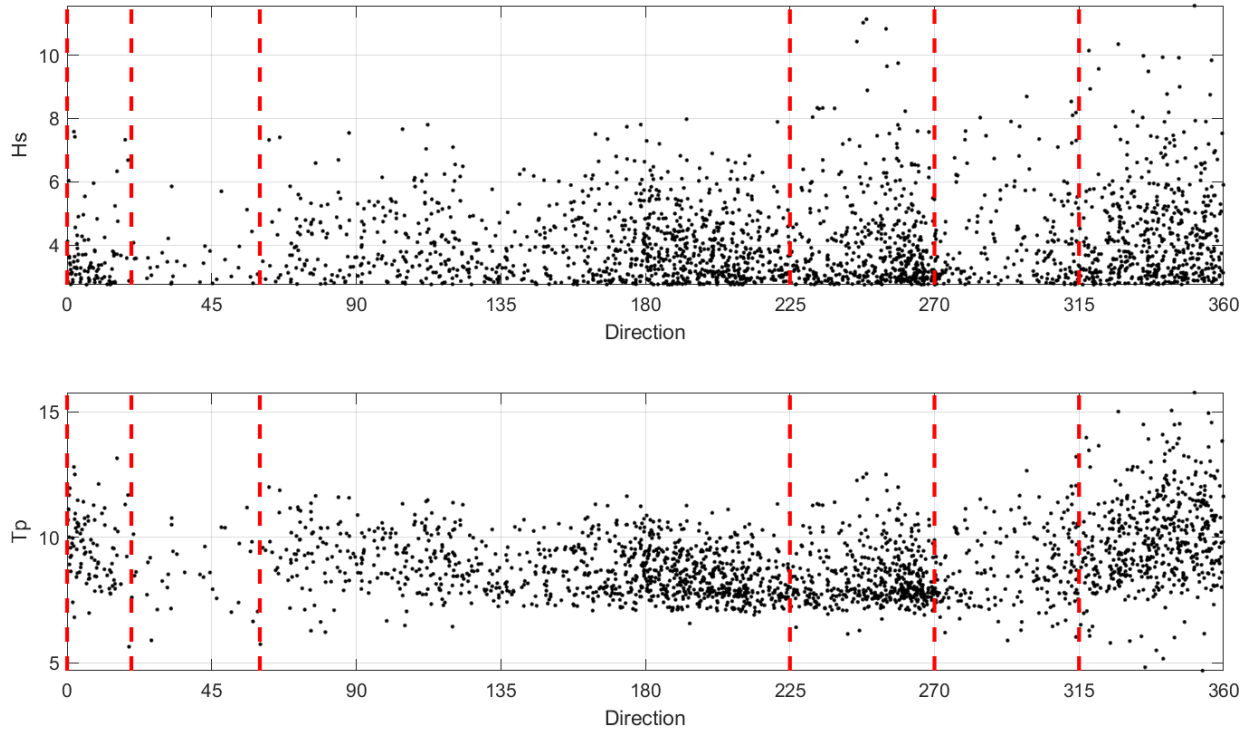


Figure 2: Directional variation of storm peak significant wave height ( $H_S$  in metres, top) and associated spectral peak period ( $T_P$  in seconds, bottom). Directions (in degrees) indicate the angle from which a storm emanates, defined clockwise from North. Also shown are directional bin edges (red) for 6 bins. The variation of  $H_S$  and  $T_P$  is approximately independent of direction within each bin, but different between bins.

plots of associated  $T_P$  on storm peak  $H_S$  for each of the six directional bins. The rates of occurrence of storms, and the marginal characteristics of  $H_S$  and  $T_P$  are clearly different between bins. In directional bin  $[225, 270)$ , the dependence between  $T_P$  and  $H_S$  appears to be particularly strong.

Figure 4 illustrates marginal extreme value models for storm peak  $H_S$ , the panels describing the variation of the estimates of generalised Pareto scale  $\nu$ , gamma shape  $\omega$  and scale  $\kappa$  with direction, in terms of means (solid) and 95% uncertainty bands (dashed) over bootstrap resamples and random choices of marginal threshold non-exceedance probabilities drawn from the interval  $[0.7, 0.9]$ . The empirical density of corresponding (stationary) estimates of generalised Pareto shape  $\xi$  is shown in Figure 9 of the user guide to be Gaussian-like with mean at approximately  $-0.2$ . The variation of  $\nu$  and  $\kappa$  with direction appears consistent with expectations given Figures 2 and 3. In particular, given estimated  $\nu$ , the largest return values for storm peak  $H_S$  would be expected to emanate from the Atlantic and Norwegian Sea. Indeed, inspection of directional maxima for storm peak  $H_S$  for return periods of 10 and 100 years in Figure 5 confirms this: the largest contributors to the omni-directional maximum (in black) are the covariate bins corresponding to the Norwegian Sea ( $[315, 0)$ , magenta) and Atlantic ( $[225, 270)$ , cyan).

It is critical to assess the diagnostic plots generated by `covXtreme` to confirm that model fit is adequate. A number of illustrative diagnostic plots corresponding to this case study are shown in the user guide. For the current application, a plot of the predictive negative log likelihood from the cross-validation procedure (see Equation 7) suggests that the optimal choice  $\lambda^\circ$  of roughness penalty lies at around three. The goodness of fit of the marginal model is assessed by examining the stability of the estimate for generalised Pareto shape parameter  $\xi$  as a function of extreme value threshold  $\tau$ . Comparison of empirical tails directly from the sample, with corresponding tails (and their uncertainty) estimates under the extreme value model, per covariate bin and omni-directionally, also suggests the marginal model

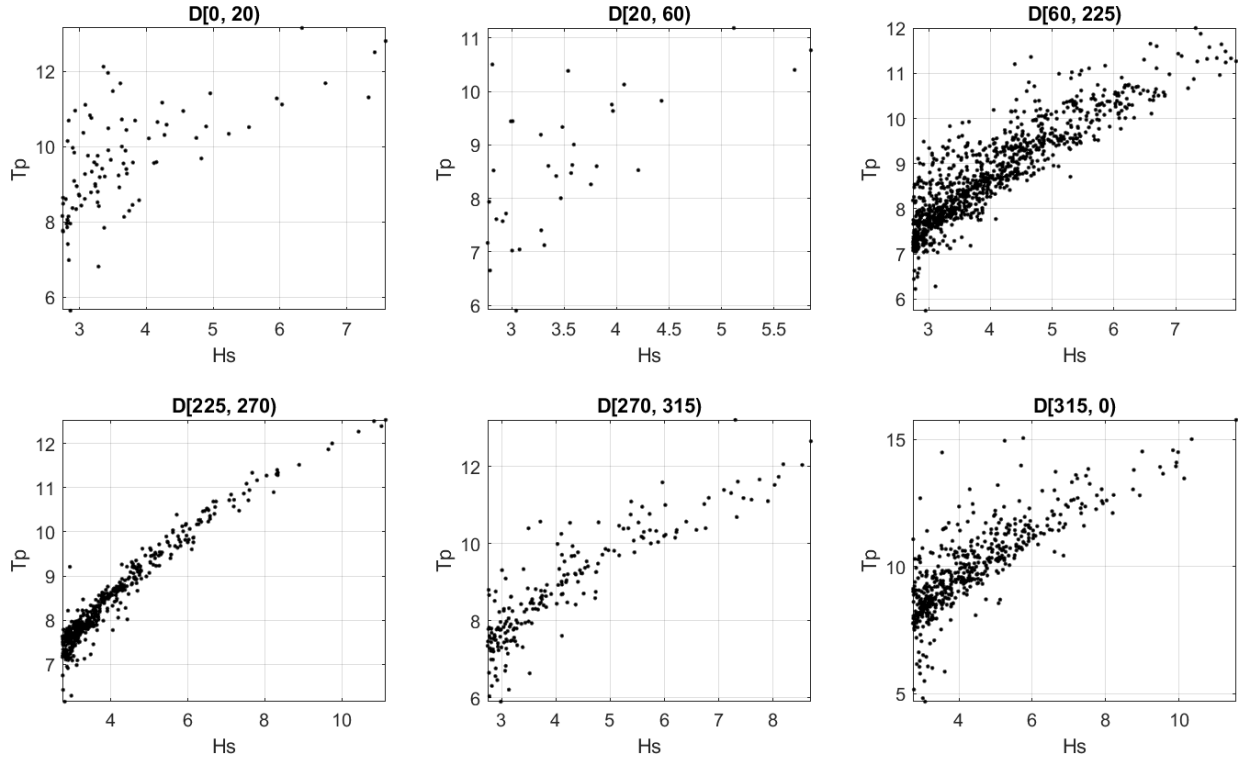


Figure 3: Associated spectral peak period (s) on storm peak significant wave height (m) per directional bin. Panel titles indicate that the covariate is direction “D”, and give the angular interval corresponding to the bin. It is apparent that the dependence between  $H_S$  and  $T_P$  varies between bins.

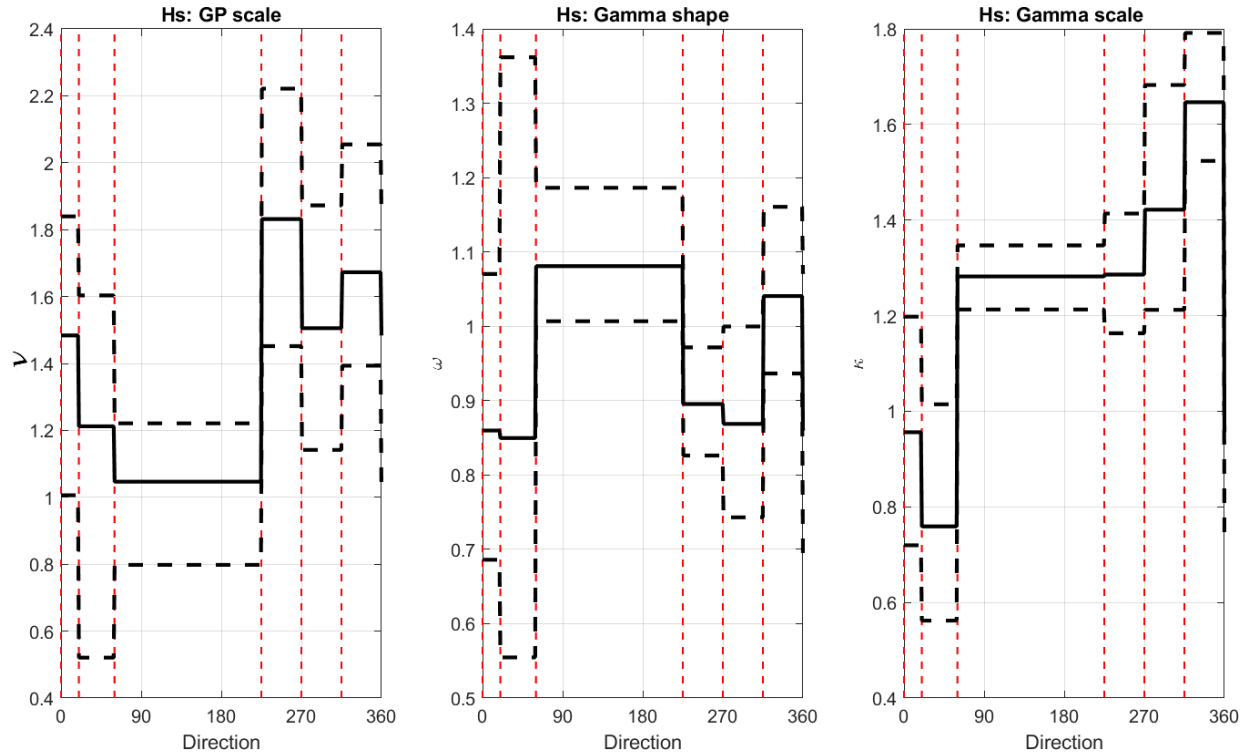


Figure 4: Marginal directional extreme value model for storm peak significant wave height (m). Variation of parameter estimates for GP scale  $\nu$ , gamma shape  $\omega$  and scale  $\kappa$  with direction. Mean estimates in solid black, and bootstrap 95% uncertainty bands in dashed black. Also shown are directional bin edges (red). Directional dependence is particularly clear for  $\nu$  and  $\kappa$ . See Equations 2 and 4 for model forms and parameter definition.

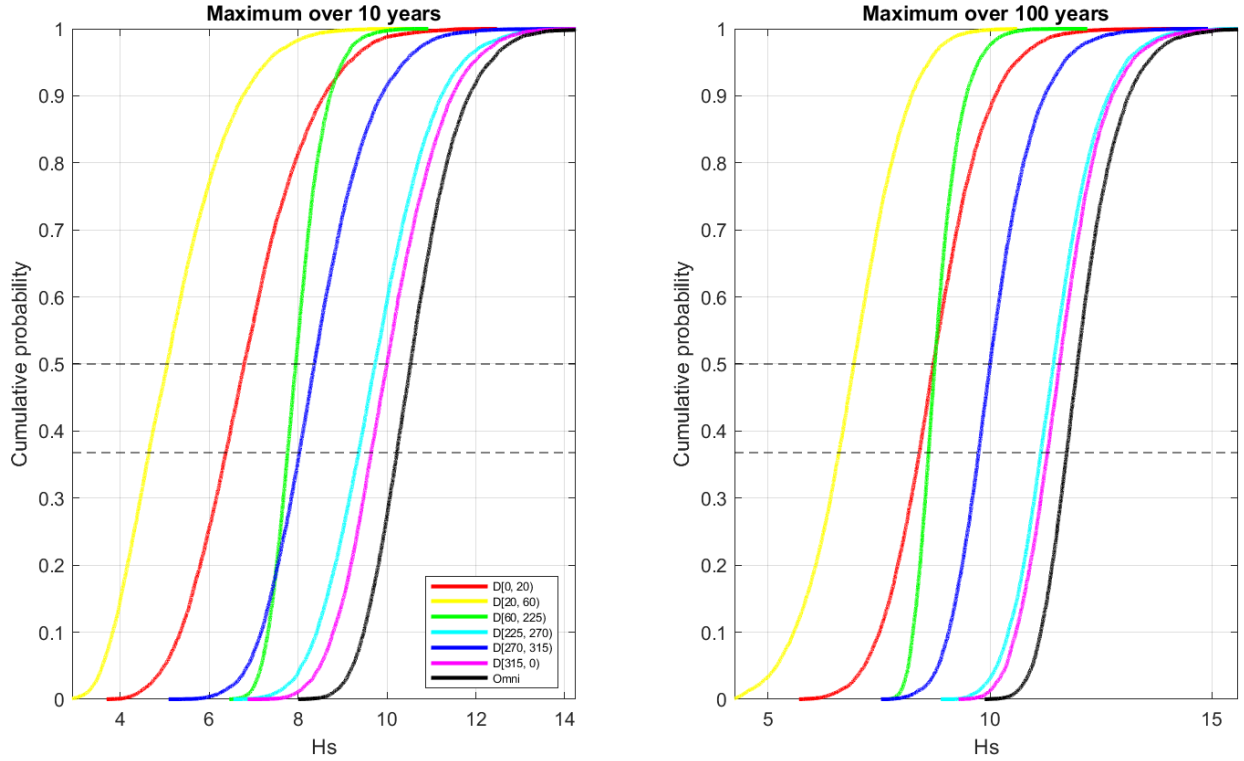


Figure 5: Cumulative distribution functions for the 10-year (left) and 100-year (right) maximum of storm peak significant wave height (m) per directional bin and over all bins (“omni”, black). Horizontal dashed lines drawn at the  $\exp(-1)$  quantile and median.

is reasonable. The corresponding full marginal analysis must also be performed for associated  $T_P$  with direction.

Using the marginal models, Figure 6 illustrates parameter estimates for the conditional extremes model for associated  $T_P$  given storm peak  $H_S$  on standard Laplace margins. The non-stationary estimate for slope parameter  $\alpha$  suggests that dependence between the variables is high (near the maximum possible of unity), especially in the covariate bin corresponding to Atlantic storms. The uncertainty band for the slope parameter included unity, admissible at sub-asymptotic levels (Tendijck et al. 2023). The value of exponent parameter  $\beta$  is slightly larger than zero, indicating that the sizes of residuals “ $y^\beta Z$ ” from the conditional extremes model fit (Equation 12) grow very slowly with the conditioning value  $y$  on Laplace scale. The empirical densities for residual parameters  $\mu$  and  $\sigma$  are typical in our experience; unimodal densities, with approximately Gaussian shape.

Diagnostic plots are again assessed to confirm reasonable goodness of fit for the conditional extremes model. For the current application, parameter estimates from  $\alpha$  per covariate bin are reasonably stable as a function of dependence threshold  $\tilde{\tau}$  on the interval  $[0.7, 0.85]$  (see Figure 20 of the user guide). Further, the distribution of residuals  $\mathcal{E}$  from the model fit do not appear to be obviously dependent on the directional covariate (see Figure 18 of the user guide). The overall distribution of residuals from the model fit (see Figure 17 of the user guide) is typical; a Gaussian-like density with positive skew.

We simulate under the fitted models to generate the environmental design contours (see Section 2.5) shown in Figure 7 omni-directionally and Figure 8 per directional bin. The constant exceedance, direct sampling and HT density contours are labelled as “Exc”, “Hus” and “HTDns” in the figures. The three different contour methods produce estimates which have similar characteristics in terms of describing the “extent” of the data cloud, reflecting the positive dependence between  $H_S$  and  $T_P$ , and passing through the appropriate lock points (shown in green in the figures). However the methods also use different definitions of the environmental contour; it is not surprising therefore that the contours estimates do not agree fully. The HTDns contour in particular produces a more variable estimate, especially when sample size is small (e.g. in the [20,60] directional bin in Figure 8). For engineering design, points on the contour with large values of  $H_S$  would typically be adopted to test the integrity of models for the offshore or coastal structure of interest.

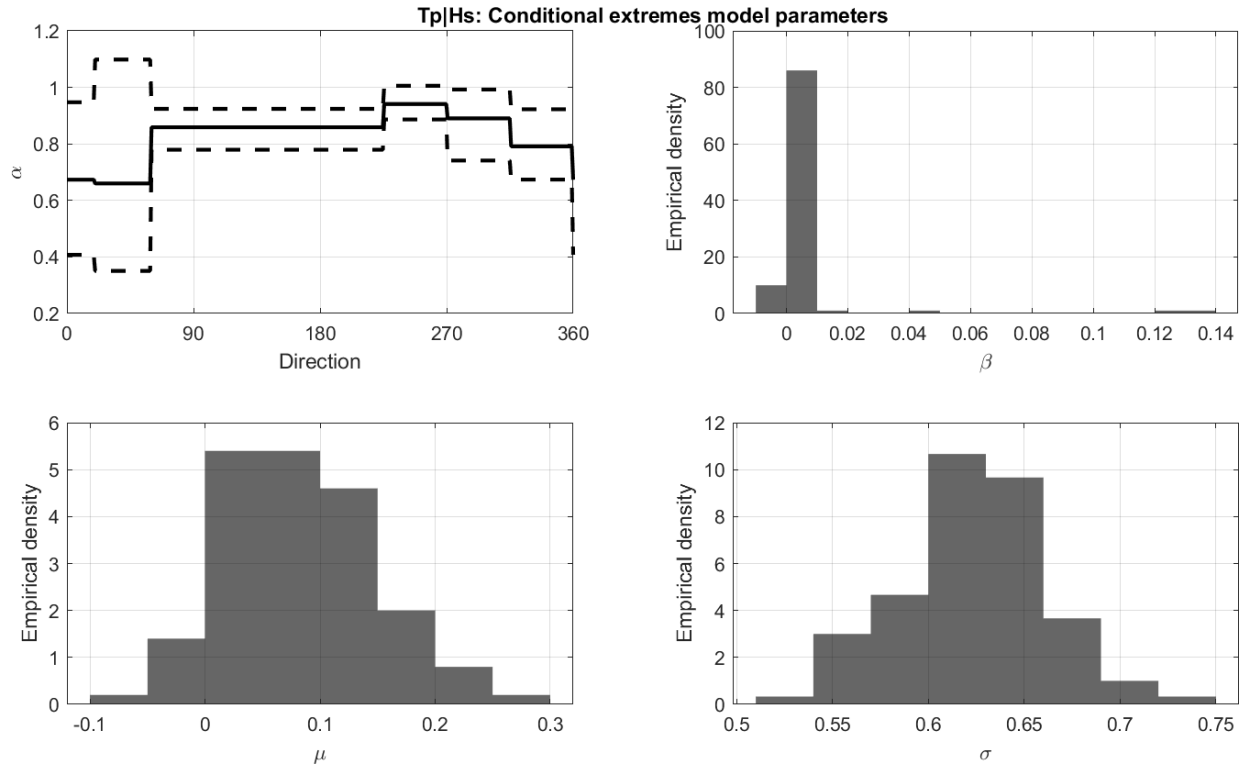


Figure 6: Parameter estimates from the conditional extreme value model for associated peak period given storm peak significant wave height (m). Top left: directional variation of estimated linear slope parameter  $\alpha$  summarised as mean and 95% bootstrap uncertainty band; top right: histogram of estimated exponent parameter  $\beta$ ; bottom left: histogram of estimated residual mean  $\mu$ ; bottom right: histogram of estimated residual standard deviation  $\sigma$ . See Equations 12 and 13 for model form and parameter definition.

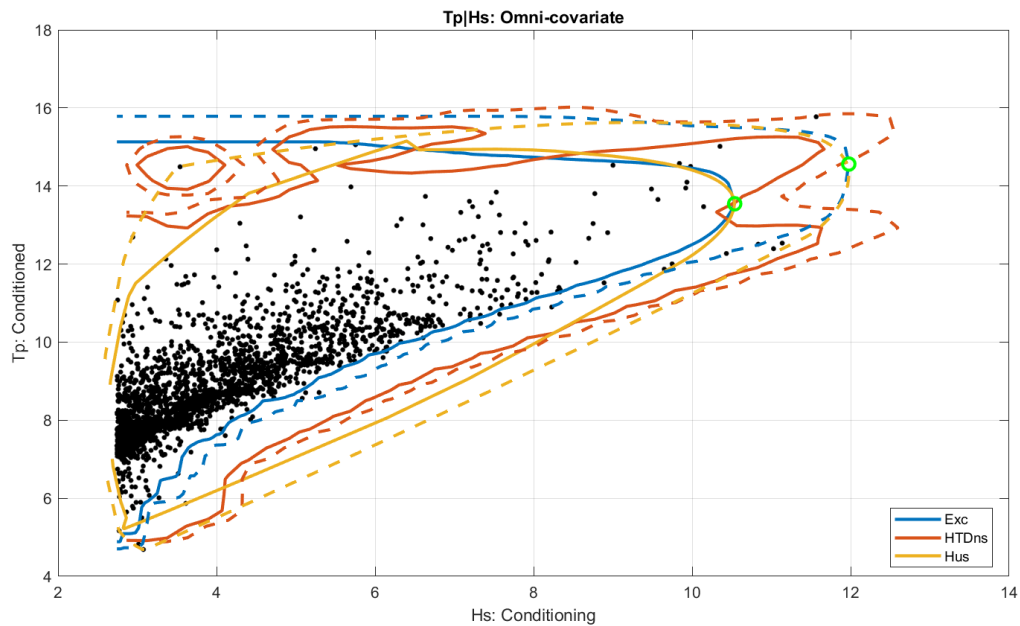


Figure 7: Omni-directional environmental contours of associated peak period (s) and storm peak significant wave height (m), for 10- and 100-year maximum values of storm peak significant wave height shown as solid and dashed lines respectively, corresponding to the Exceedance (Exc, blue), Heffernan-Tawn density (HTDns, orange) and Huseby (Hus, yellow) contours. Lock points for the respective return periods are shown in green.

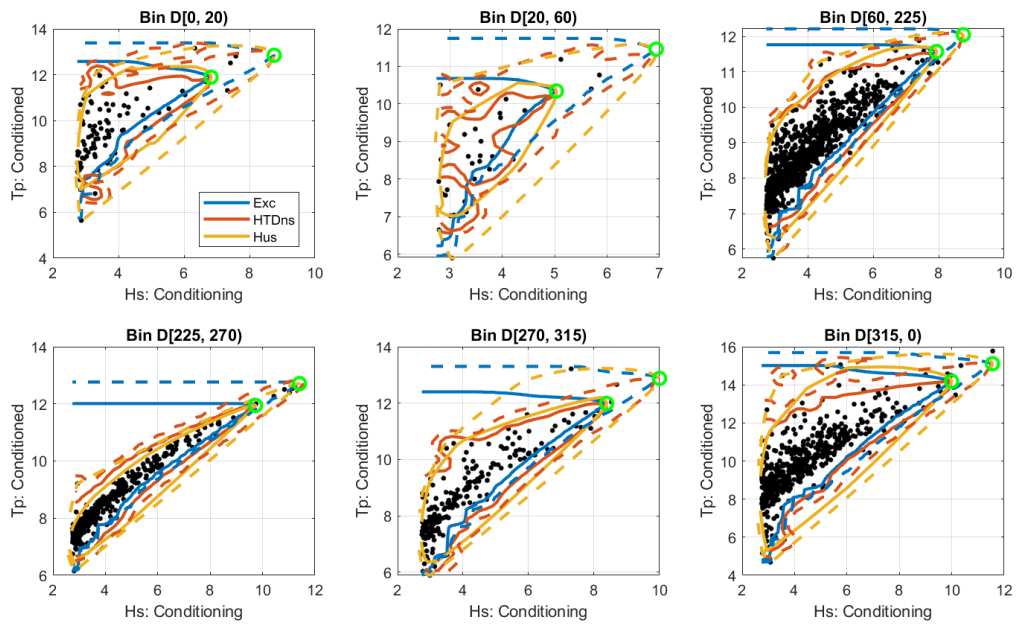


Figure 8: Environmental contours of associated peak period (s) and storm peak significant wave height (m), for individual directional bins. Contours shows for 10- and 100-year maximum values of storm peak significant wave height, shown as solid and dashed lines respectively, corresponding to the Exceedance (Exc, blue), Heffernan-Tawn density (HTDns, orange) and Huseby (Hus, yellow) contours. Lock points for the respective return periods are shown in green.

#### 4. Case study : 2-D covariate, multivariate response

The second case study is an extension of that reported in Section 3, which incorporates the over-turning moment (OTM) experienced by an offshore structure subjected to wind and wave loading. Specifically, we seek to characterise the joint distribution of (storm peak)  $H_S$  and wind speed (WS) conditional on extreme values of OTM, subject to directional and seasonal covariate effects. The analysis again follows the 5 stages described in Section 2.

Stage 1 of the analysis is isolation of storm peak events from the underlying storm  $H_S$  time-series data, following the same procedure as in Section 3. For Stage 2, Figure 9 shows the directional and seasonal variation of the conditioning variate OTM, and associated variates  $H_S$  and WS together with the directional and seasonal bin edges specified. Note that a smaller number of directional bins is used for this case study, to limit the total number of directional-seasonal bins and hence the number of parameters to be inferred in the analysis. Nevertheless, we are careful to allow for possible directional effects due to storms from the Atlantic and Norwegian Sea. There are clear directional and seasonal effects present.

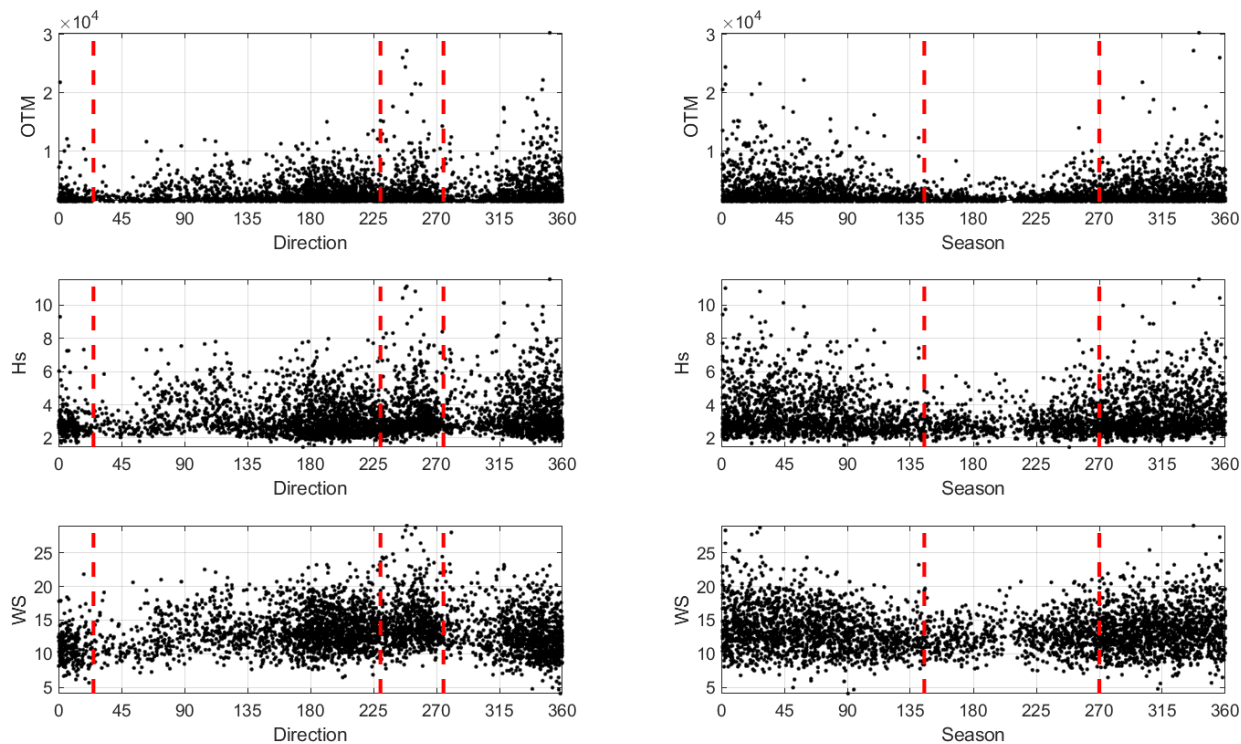


Figure 9: Directional (left) and seasonal (right) variation of overturning moment (OTM, in mega Newton metres, MNm, top), storm peak significant wave height ( $H_S$  in m, second row) and associated wind speed (WS in metres per second,  $\text{ms}^{-1}$ , bottom). Also shown are bin edges for three directional bins coupled with two seasonal bins (and hence a total of  $6 = 4 \times 2$  directional-seasonal bins). The variation of OTM,  $H_S$  and WS is approximately independent of covariates within bins, but different between bins. For the directional convention, see caption of Figure 2. Season (or seasonal degree) is defined by mapping the calendar year linearly into  $(0,360]$ , with 0 indicating midnight on the 1st January.

The resulting scatter plots of WS on OTM per covariate bin are shown in Figure 10 (and the corresponding plot for  $H_S$  on OTM in Figure 28 of the user guide). Illustrative diagnostic plots for the estimation of marginal models for each of  $H_S$ , OTM, and WS with directional and seasonal covariates are given in Figures 30-32 of the user guide. Marginal model parameters show clear directional and seasonal variation, and comparison of empirical and model-based tails suggest reasonable model fit. The fitted marginal models are then used to estimate the cumulative distribution functions for  $T$ -year maxima of interest; estimates for the 10- and 100-year maximum of WS are given in Figure 11. Unsurprisingly, the “omni” distribution (estimated over all directional and seasonal bins) is dominated by winter storms from the Atlantic sector, and the most probable 100-year maximum WS is approximately  $29 \text{ ms}^{-1}$ .

Following transformation to standard Laplace scale, a non-stationary conditional extremes model is estimated for  $H_S$  and WS jointly, conditional on large OTM, given directional and seasonal covariates. Figure 12 gives parameter estimates for WS, showing directional and seasonal dependence of the conditional extremes slope parameter  $\alpha$ ; as would be expected perhaps, the dependence between WS and OTM is strongest for winter storms emanating from the North-West. The estimates for index  $\beta$ , common over covariate bins, is somewhat larger than found in Case Study 1. The empirical densities of nuisance parameters  $\mu$  and  $\sigma$  are approximately Gaussian-shaped. Supporting diagnostic



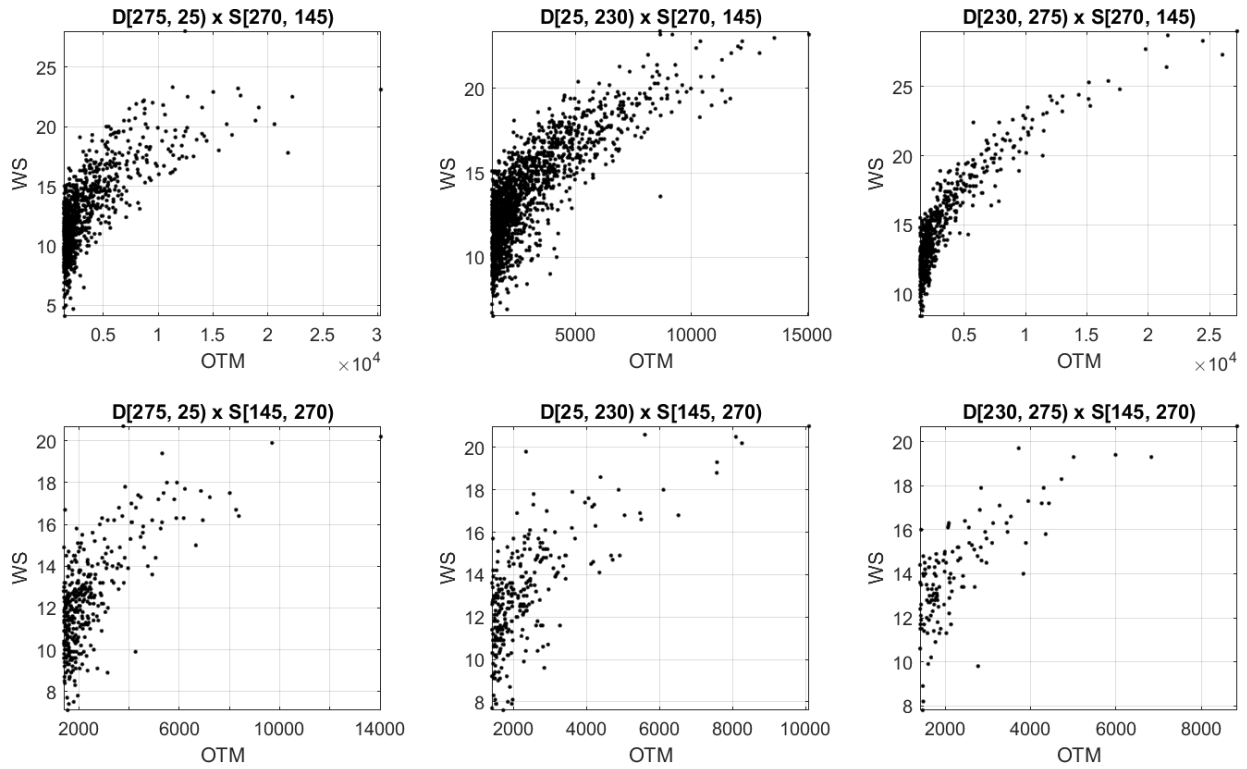


Figure 10: Associated wind speed ( $\text{ms}^{-1}$ ) on storm peak overturning moment (MNm) per directional-seasonal bin. Panel titles indicate that the angular intervals of direction “D” and season “S” for the bin “D”. It is apparent that the dependence between WS and OTM varies between bins, with obvious non-linearity in some bins. A similar plot is generated for associated significant wave height (m) on storm peak overturning moment.

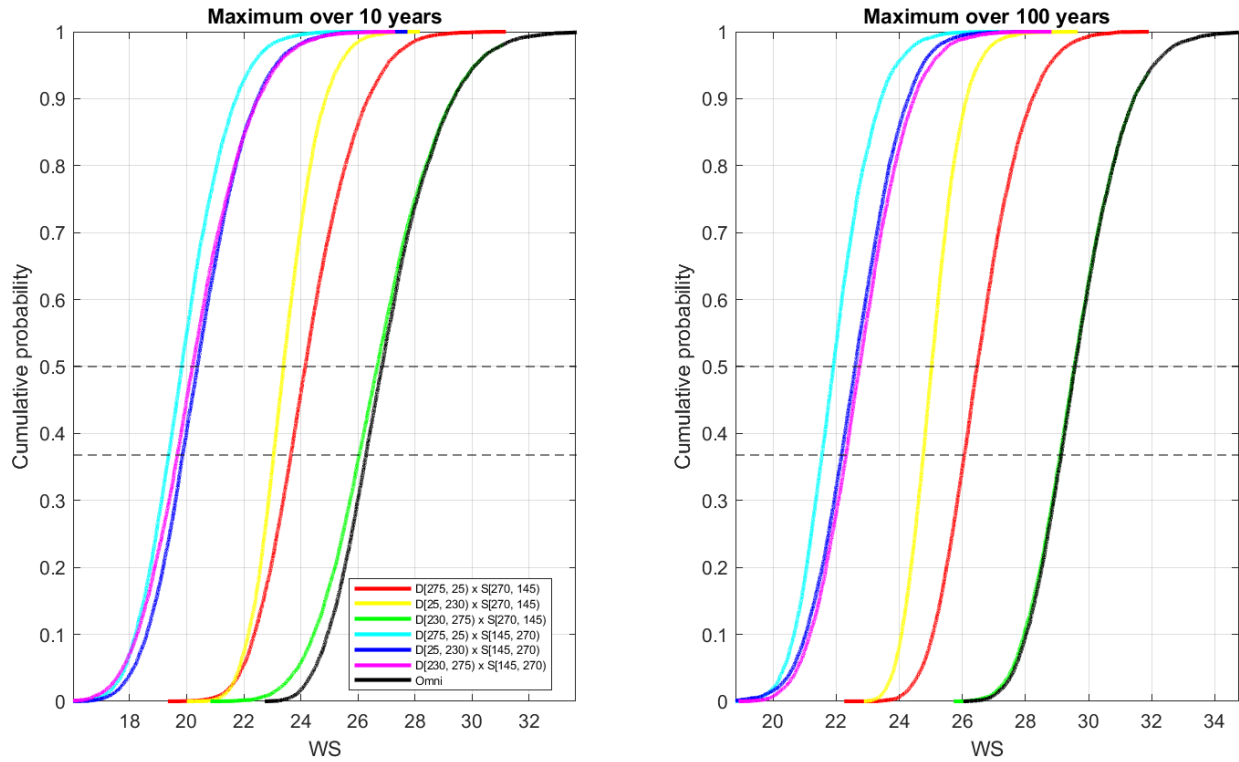


Figure 11: Cumulative distribution functions for the 10-year (left) and 100-year (right) maximum of associated wind speed ( $\text{ms}^{-1}$ ) per covariate bin and over all bins (“omni”, black). Horizontal dashed lines drawn at the  $\exp(-1)$  quantile and median.



plots for the conditional extremes model fit are given in Figures 33-38 of the user guide. In particular, Figure 37 indicates that the value of the  $\alpha$  parameter is relatively stable for conditional extremes threshold levels  $\tilde{\tau} \in [0.8, 0.9]$ , and Figure 38 indicates that a minimum roughness penalty for predictive lack of fit is achieved at around 10.

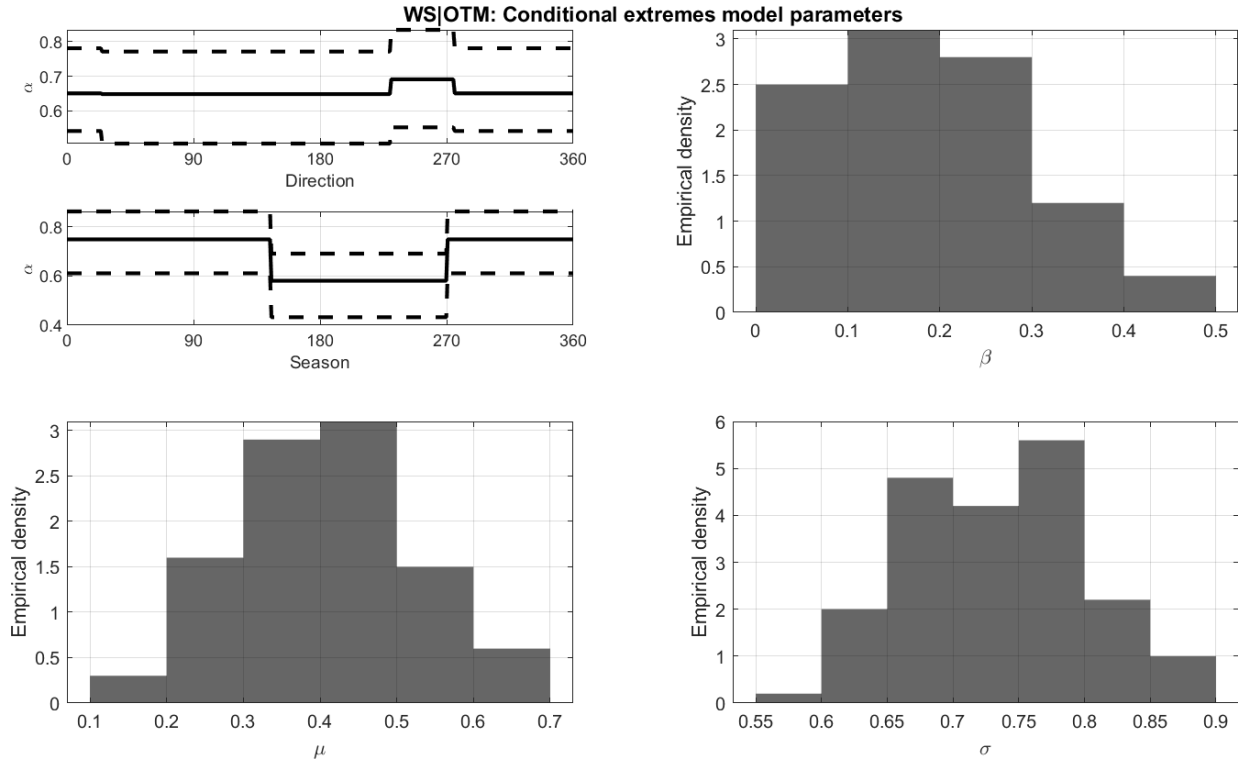


Figure 12: Parameter estimates from the conditional extreme value mode for associated wind speed given storm peak overturning moment. Top left: directional and seasonal variation of  $\alpha$  summarised as mean and 95% bootstrap uncertainty band; top right: histogram of  $\beta$ ; bottom left: histogram of  $\mu$ ; bottom right: histogram of  $\sigma$ . Note the very high directional dependence for  $\alpha$  at around  $280^\circ$ . See Equations 12 and 13 for model form and parameter definition, and Figure 6 for comparison.

Figure 13 gives estimated conditional return value distributions for  $H_S$  and WS given 10-year and 100-year maximum OTM for individual directional-seasonal covariate bins, and “omni” over all covariate bins. For  $H_S$ , conditional return values are again largest for Atlantic winter storms. For WS however, we observe an interesting transition involving winter storms from directional sectors  $[275,315)$  to  $[230,275)$ . The most probable *conditional* value of WS given a 100-year maximum OTM is approximately  $26.5 \text{ ms}^{-1}$ , lower than the marginal most probable maximum 100-year wind speed.

The resulting omni-covariate environmental design contours for  $H_S$  and WS, given occurrences of the 10-year and 100-year maximum OTM, are shown in Figure 14. The general features of the contours are similar to those of Figure 7. Corresponding illustrative plots of contours per covariate bin for WS are given in Figure 40 of the user guide.

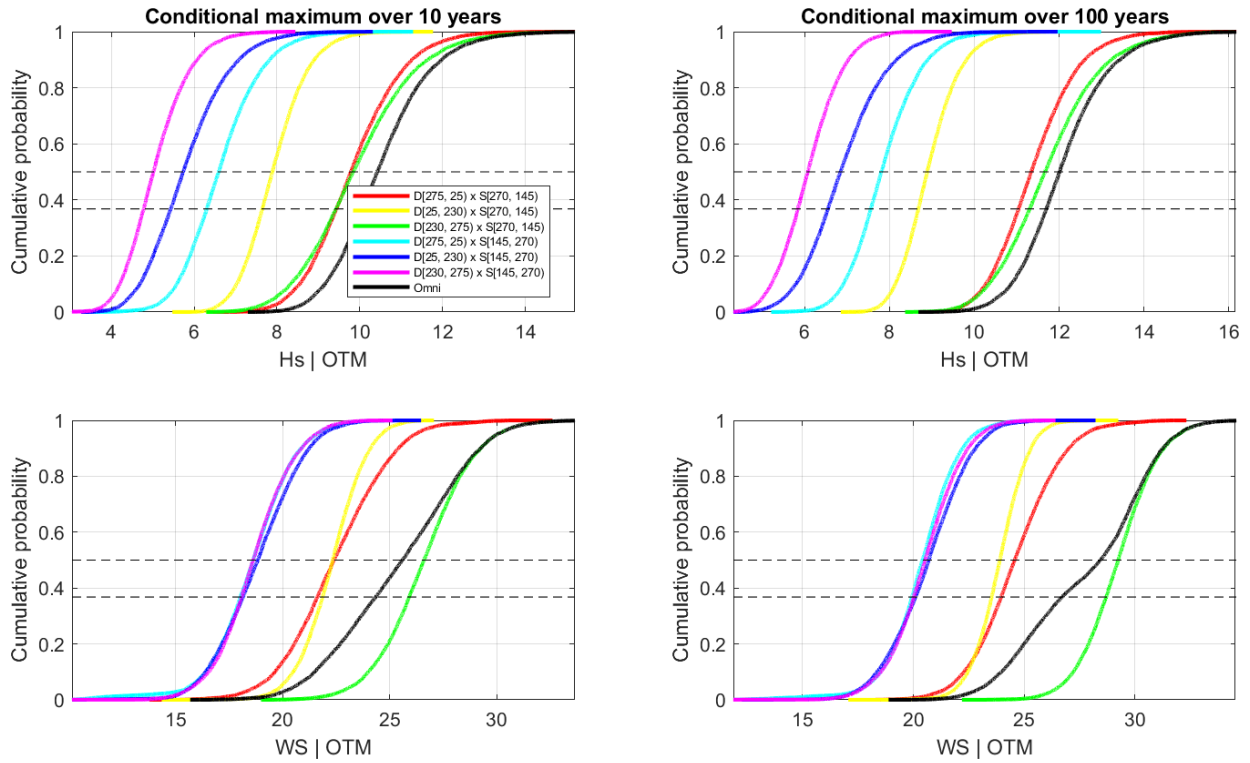


Figure 13: Conditional cumulative distribution functions of associated significant wave height (m, top) and associated wind speed ( $\text{ms}^{-1}$  bottom) per directional bin and over all bins (“omni”, black), conditional on the 10-year (left) and 100-year (right) maximum of storm peak overturning moment (MNm). Horizontal dashed lines drawn at the  $\exp(-1)$  quantile and median.

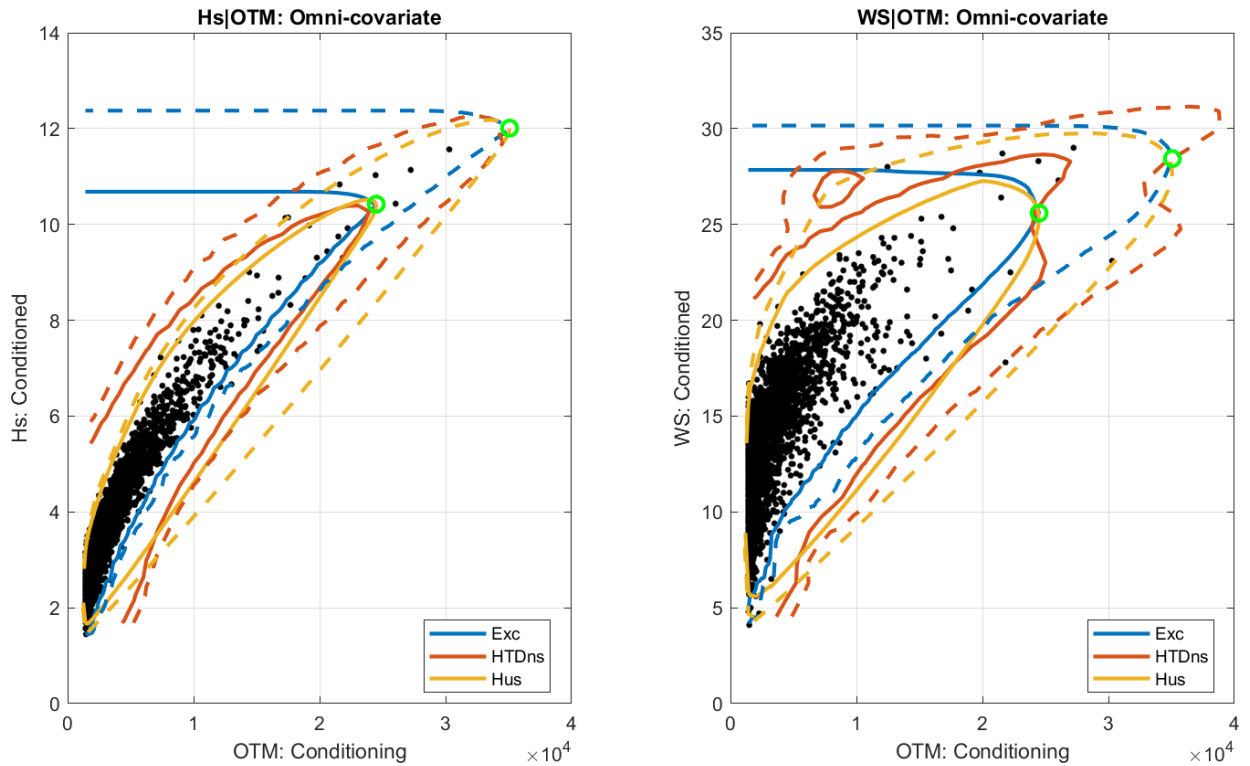


Figure 14: Omni-directional-seasonal environmental contours of associated significant wave height (m, left) and associated wind speed ( $\text{ms}^{-1}$ , right) and storm peak overturning moment (MNm). Contours shows for 10- and 100-year maximum values of storm peak significant wave height, shown as solid and dashed lines respectively, corresponding to the Exceedance (blue), Heffernan-Tawn density (HTDns, orange) and Huseby (yellow) contours. Corresponding contours per covariate bin are also generated.

## 5. Discussion

This article introduces the `covXtreme` software for pragmatic multivariate extreme value analysis with covariate non-stationarity. The MATLAB software provides functionality to isolate temporal peaks from time-series, and to define an appropriate partition of the covariate domain. Using this partition, marginal generalised Pareto models are estimated for each variate independently assuming piecewise constant threshold and scale parameterisations within a penalised likelihood framework; optimal scale parameter roughness is estimated using cross-validation. Marginal models are then used to transform the sample to standard Laplace margins. A non-stationary conditional extremes model is then estimated with piecewise constant parameterisation for the slope (“ $\alpha$ ”) parameter, again using penalised likelihood estimation with cross-validation to estimate optimal roughness. Simulations and importance sampling are then used to estimate the distribution of  $T$ -year maxima for each variate, and environmental design contours conditional on a single conditioning variate. Uncertainty due to marginal and dependence threshold selection is quantified by fitting multiple models with randomly chosen thresholds within user-specified plausible intervals of threshold non-exceedance probability. Uncertainties in parameter estimates and subsequent inferences from model fitting are estimated using bootstrap resampling. As noted in Section 1, the software has already been used in a number of studies, mostly but not exclusively meteocean-related.

The `covXtreme` methodology makes a number of simplifying assumptions, motivated by the authors’ experience of extreme value analysis applied to the ocean environment using a range of methodologies of difference complexities. For example, `covXtreme` relies on sensible user-specified partitioning of the covariate domain into bins within which it is reasonable to assume common marginal tails and a common dependence structure; this simplifies inference considerably compared with competitor approaches. Moreover, we believe that inferences using `covXtreme` with good partitioning are competitive with alternatives using more sophisticated tools. Specifically, `covXtreme` marginally is equivalent to a Voronoi set representation with pre-specified covariate partition, which was demonstrated by Zanini et al. (2020) to be competitive with P-spline and Bayesian adaptive regression spline covariate representations. `covXtreme` further assumes that the generalised Pareto shape parameter  $\xi$  in each marginal model is constant with covariate; because of the relative difficulty of estimating the shape parameter compared with the scale, this would appear reasonable in the absence of strong evidence to the contrary, especially for small samples of data. Likewise, the  $\beta$ ,  $\mu$  and  $\sigma$  parameters of the conditional extremes model are assumed stationary. This appears reasonable since  $\beta$  is an exponent, again difficult to estimate. Moreover,  $\mu$  and  $\sigma$  are essentially nuisance parameters; any model misspecification caused by the assumption of stationarity will be accommodated to some extent by the adoption of residuals from model fitting for inferences under the model. It might be appropriate to relax some of the assumptions for specific applications, for example (a) when there is strong evidence that generalised Pareto  $\xi$  is unlikely to be constant (e.g. due to land shadow and fetch limitation effects on  $H_S$ ), or (b) since parameter estimates for conditional extremes  $\alpha$  and  $\mu$  are highly correlated when  $\beta$  is close to unity. We also note that it is not clear whether non-stationary marginal extreme analysis is necessarily the best approach to estimate marginal return values from non-stationary data. Provided that sufficient sample is available, so that a sufficiently high threshold for peaks over threshold analysis can be set, often a stationary marginal analysis is at least competitive if not preferable; see Mackay and Jonathan (2020) for further discussion.

We anticipate that `covXtreme` might provide a pragmatic starting point to studies of spatial and temporal dependence of extremes, since the underlying methodology of both conditional spatial extremes (e.g. Shooter et al. 2019, Shooter et al. 2022, Wadsworth and Tawn 2022) and Markov extremal and similar time-series models (e.g. Winter and Tawn 2017, Tendijck et al. 2019, Tendijck et al. 2024) involves each of Stages 1-4 of the `covXtreme` approach. We also hope that `covXtreme` might be useful generally in estimating joint characteristics of extremes from multivariate time-series.

### Software and data availability

- Name: `covXtreme`
- Developer: Ross Towe, Emma Ross, David Randell, Philip Jonathan
- Contact: ross.towe@shell.com
- First published: 2023
- Language: MATLAB (only base MATLAB needed)
- GitHub: <https://github.com/sede-open/covXtreme>
- Program size: 300kB
- Licence: apache 2.0
- Free-to-use software, sample data and user guide available on GitHub.

## Credit

Towe: Methodology, Software, Validation, Writing; Ross: Conceptualisation, Methodology, Software, Validation, Writing; Randell: Conceptualisation, Methodology, Software. Jonathan: Conceptualisation, Methodology, Software, Validation, Writing.

## Acknowledgement

The original version of this software was developed as a component of the “Environmental Contours for SAfe DEsign of Ships and other marine structures” (ECSADES) project, part-funded by the European Union ERANET programme, and was summarised in a review paper on the definition and application of environmental contours (Ross et al. 2020). `covXtreme` software, user guide and test data sets are available at Towe et al. (2023b). We would like to thank Jennifer Kensler at Shell for an internal technical review of `covXtreme` methodology and software. At time of publication of this article, the software is hosted by Linux Foundation Energy at <https://lfenergy.org/projects/covxtreme/>, where the user community is encouraged to develop and share enhancements.

## References

- Barlow, A.M., Mackay, E., Eastoe, E., Jonathan, P., 2023. A penalised piecewise-linear model for non-stationary extreme value analysis of peaks over threshold. *Ocean Eng.* 267, 113265.
- Belzile, L.R., Dutang, C., Northrop, P.J., Opitz, T., 2023. A modeler’s guide to extreme value software. *Extremes* 26, 595–638.
- Billier, C., 2000. Adaptive Bayesian regression splines in semiparametric generalized linear models. *J. Comput. Graph. Statist.* 9, 122–140.
- Bore, P.T., Amdahl, J., Kristiansen, D., 2019. Statistical modelling of extreme ocean current velocity profiles. *Ocean Eng.* 186, 106055.
- Brezger, A., Lang, S., 2006. Generalized structured additive regression based on Bayesian P-splines. *Comput. Statist. Data Anal.* 50, 967–991.
- Chavez-Demoulin, V., Davison, A., 2005. Generalized additive modelling of sample extremes. *J. Roy. Statist. Soc. Series C: Appl. Stat.* 54, 207–222.
- Cheng, Y., Zhang, K., Chao, L., Shi, W., Feng, J., Li, Y., 2023. A comprehensive drought index based on remote sensing data and nested copulas for monitoring meteorological and agroecological droughts: A case study on the Qinghai-Tibet Plateau. *Environ. Model. Softw.* 161, 105629.
- Cobos, M., Otinar, P., Magana, P., Lira-Loarca, A., Baquerizo, A., 2022. `MarineTools.temporal`: A Python package to simulate Earth and environmental time series. *Environ. Model. Softw.* 150, 105359.
- Coles, S., 2001. *An introduction to statistical modelling of extreme values*. Springer.
- Davison, A., Smith, R.L., 1990. Models for exceedances over high thresholds. *J. R. Statist. Soc. B* 52, 393.
- Davison, A.C., Huser, R., 2015. Statistics of extremes. *Annual Review of Statistics and Its Application* 2, 203–235.
- De Luca, D.L., Napolitano, F., 2023. A user-friendly software for modelling extreme values: EXTRASTAR (EXTRemes Abacus for STAtistical Regionalization). *Environ. Model. Softw.* 161, 105622.
- Diez-Sierra, J., del Jesus, M., 2017. A rainfall analysis and forecasting tool. *Environ. Model. Softw.* 97, 243–258.
- Eilers, P.H.C., Marx, B.D., 2010. Splines, knots and penalties. *Wiley Interscience Reviews: Computational Statistics* 2, 637–653.
- Ewans, K., Jonathan, P., 2008. The effect of directionality on northern North Sea extreme wave design criteria. *J. Offshore. Arct. Eng.* 130, 041604:1–8.
- Ewans, K., Jonathan, P., 2023. Uncertainties in estimating the effect of climate change on 100-year return period significant wave heights. *Ocean Eng.* 272, 113840:1–17.

- Gilleland, E., 2016. Computing software, in: Dey, D., Yan, J. (Eds.), *Extreme value modeling and risk analysis: methods and applications*. CRC Press, Boca Raton, USA, pp. 505–516.
- Gilleland, E., Ribatet, M., Stephenson, A.G., 2013. A comparative software review for extreme value analysis. *Extremes* 16, 103–119.
- Guerrero, M.B., Huser, R., Ombao, H., 2023. Conex-connect: Learning patterns in extremal brain connectivity from multi-channel eeg data. *Ann. Appl. Stat.* 17, 178–198.
- Haag, F., Aubert, A.H., Lienert, J., 2022. ValueDecisions, a web app to support decisions with conflicting objectives, multiple stakeholders, and uncertainty. *Environ. Model. Softw.* 150, 105361.
- Hafver, A., Agrell, C., Vanem, E., 2022. Environmental contours as voronoi cells. *Extremes* 25, 451–486.
- Hao, Z., Hao, F., Singh, V.P., Ouyang, W., Cheng, H., 2017. An integrated package for drought monitoring, prediction and analysis to aid drought modeling and assessment. *Environ. Model. Softw.* 91, 199–209.
- Haselsteiner, A.F., Coe, R.G., Manuel, L., Chai, W., Leira, B., Clarindo, G., Guedes Soares, C., Hannesdóttir, Á., Dimitrov, N., Sander, A., Ohlendorf, J.H., Thoben, K.D., de Hauteclocque, G., Mackay, E., Jonathan, P., Qiao, C., Myers, A., Rode, A., Hildebrandt, A., Schmidt, B., Vanem, E., Huseby, A.B., 2021. A benchmarking exercise for environmental contours. *Ocean Engineering* 236, 109504.
- Haselsteiner, A.F., Lehmkuhl, J., Pape, T., Windmeier, K.L., Thoben, K.D., 2019. A software to compute multivariate extremes using the environmental contour method. *SoftwareX* 9, 95–101.
- Haselsteiner, A.F., Ohlendorf, J.H., Wosniok, W., Thoben, K.D., 2017. Deriving environmental contours from highest density regions. *Coastal Eng.* 123, 42–51.
- Heffernan, J.E., Tawn, J.A., 2004. A conditional approach for multivariate extreme values. *J. R. Statist. Soc. B* 66, 497–546.
- Huseby, A., Vanem, E., Natvig, B., 2015. Alternative environmental contours for structural reliability analysis. *Struct. Saf.* 54, 32–45.
- Huser, R., Genton, M.G., 2016. Non-stationary dependence structures for spatial extremes. *J. Agric. Biol. Environ.* 21, 470–491.
- Jonathan, P., Ewans, K., Flynn, J., 2014. On the estimation of ocean engineering design contours. *ASME J. Offshore Mech. Arct. Eng.* 136:041101.
- Jonathan, P., Ewans, K.C., Randell, D., 2013. Joint modelling of environmental parameters for extreme sea states incorporating covariate effects. *Coastal Eng.* 79, 22–31.
- Jonathan, P., Randell, D., Wadsworth, J., Tawn, J., 2021. Uncertainties in return values from extreme value analysis of peaks over threshold using the generalised Pareto distribution. *Ocean Eng.* 220, 107725.
- Liu, Z., Cheng, L., Lin, K., Cai, H., 2021. A hybrid Bayesian vine model for water level prediction. *Environ. Model. Softw.* 142, 105075.
- Maciejewska, K., Juda-Rezler, K., Reizer, M., Klejnowski, K., 2015. Modelling of black carbon statistical distribution and return periods of extreme concentrations. *Environ. Model. Softw.* 74, 212–226.
- Mackay, E., de Hauteclocque, G., 2023. Model-free environmental contours in higher dimensions. *Ocean Eng.* 273, 113959.
- Mackay, E., Haselsteiner, A.F., 2021. Marginal and total exceedance probabilities of environmental contours. *Marine Structures* 75, 102863.
- Mackay, E., Jonathan, P., 2020. Assessment of return value estimates from stationary and non-stationary extreme value models. *Ocean Eng.* 207, 107406.
- Minguez, R., Mendez, F.J., Izaguirre, C., Menendez, M., Losada, I.J., 2010. Pseudo-optimal parameter selection of non-stationary generalized extreme value models for environmental variables. *Environ. Model. Softw.* 25, 1592–1607.

- Northrop, P., Jonathan, P., 2011. Threshold modelling of spatially-dependent non-stationary extremes with application to hurricane-induced wave heights. *Environmetrics* 22, 799–809.
- Randell, D., Turnbull, K., Ewans, K., Jonathan, P., 2016. Bayesian inference for non-stationary marginal extremes. *Environmetrics* 27, 439–450.
- Ross, E., Astrup, O.C., Bitner-Gregersen, E., Bunn, N., Feld, G., Gouldby, B., Huseby, A., Liu, Y., Randell, D., Vanem, E., Jonathan, P., 2020. On environmental contours for marine and coastal design. *Ocean Eng.* 195, 106194.
- Ross, E., Sam, S., Randell, D., Feld, G., Jonathan, P., 2018. Estimating surge in extreme North Sea storms. *Ocean Eng.* 154, 430–444.
- Scarrott, C., MacDonald, A., 2012. A review of extreme value threshold estimation and uncertainty quantification. *Revstat* 10, 33–60.
- Shao, X., Hazra, A., Richards, J., Huser, R., 2022. Flexible modeling of nonstationary extremal dependence using spatially-fused LASSO and ridge penalties. URL: <https://arxiv.org/abs/2210.057923>.
- Shooter, R., Ross, E., Ribal, A., Young, I.R., Jonathan, P., 2021. Spatial conditional extremes for significant wave height from satellite altimetry. *Environmetrics* 32, e2674.
- Shooter, R., Ross, E., Ribal, A., Young, I.R., Jonathan, P., 2022. Multivariate spatial conditional extremes for extreme ocean environments. *Ocean Eng.* 247, 110647.
- Shooter, R., Ross, E., Tawn, J.A., Jonathan, P., 2019. On spatial conditional extremes for ocean storm severity. *Environmetrics* 30, e2562.
- Souaïssi, Z., Ouarda, T.B.M.J., St-Hilaire, A., 2023. Regional thermal index model for river temperature frequency analysis in ungauged basins. *Environ. Model. Softw.* 164, 105709.
- Southworth, H., Heffernan, J.E., Metcalfe, P.D., 2024. `texmex`: statistical modelling of extreme values. <https://cran.r-project.org/package=texmex>.
- Stephenson, A., Gilleland, E., 2006. Software for the analysis of extreme events: The current state and future directions. *Extremes* 8, 87–109.
- Tendijck, S., Eastoe, E., Tawn, J., Randell, D., Jonathan, P., 2023. Modeling the extremes of bivariate mixture distributions with application to oceanographic data. *J. Am. Statist. Soc.* 118, 1373–1384.
- Tendijck, S., Jonathan, P., Randell, D., Tawn, J.A., 2024. Temporal evolution of the extreme excursions of multivariate  $k$ th order Markov processes with application to oceanographic data. *Environmetrics* 35, e2834.
- Tendijck, S., Ross, E., Randell, D., Jonathan, P., 2019. A non-stationary statistical model for the evolution of extreme storm events. *Environmetrics* 30, e2541.
- Towe, R., Randell, D., Kensler, J., Feld, G., Jonathan, P., 2023a. Estimation of associated values from conditional extreme value models. *Ocean Eng.* 272, 113808.
- Towe, R., Ross, E., Randell, D., Jonathan, P., 2023b. `covXtreme`: MATLAB software for non-stationary penalised piecewise constant marginal and conditional extreme value models. <https://github.com/sede-open/covXtreme>.
- Towe, R., Zanini, E., Randell, D., Feld, G., Jonathan, P., 2021. Efficient estimation of distributional properties of extreme seas from a hierarchical description applied to calculation of un-manning and other weather-related operational windows. *Ocean Eng.* 238, 109642.
- Vanem, E., Guo, B., Ross, E., Jonathan, P., 2020. Comparing different contour methods with response-based methods for extreme ship response analysis. *Mar. Struct.* 69, 102680.
- Čampulová, M., Čampula, R., Holešovský, J., 2022. An R package for identification of outliers in environmental time series data. *Environ. Model. Softw.* 155.
- Villa, D.L., Schostek, T., Govertsen, K., Macmillan, M., 2023. Position Paper: A stochastic model of future extreme temperature events for infrastructure analysis. *Environ. Model. Softw.* 163, 105663.

- Vrugt, J.A., 2016. Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. *Environ. Model. Softw.* 75, 273–316.
- Wadsworth, J., Tawn, J., 2022. Higher-dimensional spatial extremes via single-site conditioning. *Spatial Statistics* 51, 100677.
- Wadsworth, J.L., 2016. Exploiting structure of maximum likelihood estimators for extreme value threshold selection. *Technometrics* 58, 116–126.
- Winter, H.C., Tawn, J.A., 2017. kth-order Markov extremal models for assessing heatwave risks. *Extremes* 20, 393–415.
- Winterstein, S.R., Ude, T.C., Cornell, C.A., Bjerager, P., Haver, S., 1993. Environmental parameters for extreme variable: Inverse FORM with omission factors, in: *Proc. 6th Int. Conf. on Structural Safety and Reliability*, Innsbruck, Austria.
- Wood, S., 2023. mgcv: mixed GAM computation vehicle with automatic smoothness estimation. <https://cran.r-project.org/package=mgcv>.
- Wood, S.N., 2004. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *J. Am. Statist. Soc.* 99, 673–686.
- Wood, S.N., Li, Z., Shaddick, G., Augustin, N.H., 2017. Generalized additive models for gigadata: Modeling the u.k. black smoke network daily data. *J. Am. Statist. Soc.* 112, 1199–1210.
- Wood, S.N., Pya, N., Sáfken, B., 2016. Smoothing parameter and model selection for general smooth models. *J. Am. Statist. Soc.* 111, 1548–1563.
- Youngman, B., 2022. evgam: generalised additive extreme value models. <https://cran.r-project.org/package=evgam>.
- Youngman, B.D., 2019. Generalized additive models for exceedances of high thresholds with an application to return level estimation for U.S. wind gusts. *J. Am. Statist. Soc.* 114, 1865–1879.
- Youngman, B.D., 2020. Flexible models for nonstationary dependence: methodology and examples. URL: <https://arxiv.org/abs/2001.06642>.
- Zanini, E., Eastoe, E., Jones, M., Randell, D., Jonathan, P., 2020. Covariate representations for non-stationary extremes. *Environmetrics* 31, e2624.
- Zhang, M., Yang, X., Cleverly, J., Huete, A., Zhang, H., Yu, Q., 2022. Heat wave tracker: A multi-method, multi-source heat wave measurement toolkit based on Google Earth Engine. *Environ. Model. Softw.* 147, 105255.
- Zhao, L., Liu, M., Song, Z., Wang, S., Zhao, Z., Zuo, S., 2022. Regional-scale modeling of rainfall-induced landslides under random rainfall patterns. *Environ. Model. Softw.* 155, 105454.