Directional-seasonal extreme value analysis of North Sea storm conditions

Hans Fabricius Hansen[a,b,*], David Randell[c], Allan Rod Zeeberg[d], Philip Jonathan[e,f]

[a] *DHI Ports and Offshore Technology, DK-2970 Hørsholm, Denmark*
[b] *HAW MetOcean ApS, DK-1860 Frederiksberg C, Denmark*
[c] *Shell Global Solutions International B.V., 1031 HW Amsterdam, Netherlands.*
[d] *Total E&P Danmark A/S, Britanniavej 10, 6705 Esbjerg, Denmark.*
[e] *Shell Research Limited, London SE1 7NA, United Kingdom.*
[f] *Department of Mathematics and Statistics, Lancaster University LA1 4YW, United Kingdom.*

Abstract

Design and re-analysis of offshore structures requires the joint estimation of extreme values for a set of environmental variables, representing so-called long-term and short-term characteristics of the environment, subject to sources of systematic variation including directionality and seasonality. Estimation is complicated by numerous sources of uncertainty, typically including limited sample size and the specification of a number of analysis parameters (such as thresholds for peaks over threshold analysis). In this work, we present a model to estimate joint extremal characteristics of the ocean environment incorporating non-stationary marginal and conditional extreme value analysis, and thorough uncertainty quantification, within a Bayesian framework. The model is used to quantify the joint directional-seasonal structure of extremes waves, winds and currents at a location in the Danish sector of the North Sea.

Keywords: offshore design, extremes, non-stationary, conditional extremes, Bayesian, uncertainty, extreme value

1. Introduction

Offshore oil and gas installations must be designed to withstand environmental loads with annual probabilities of exceedance less than some small value (e.g. 10^{-4}) when lives are at risk or severe pollution possible. Reliable estimation of rare load levels requires careful analysis of the ocean environmental and structural loading. Environmental data from hindcasts and measurements are typically available for time periods of the order of decades. Extrapolation far beyond the sample is therefore necessary to estimate appropriate design conditions. The quantities of interest from a design perspective, generically called "responses", include maximum wave height, crest elevation or load level. These random quantities are dependent on the environmental conditions for a sea state, typically with length of the order of 30 minutes to 3 hours, summarised in terms of sea-state significant wave height, mean wind speed, etc. Some "short-term" distributions (e.g. of maximum wave height in a sea state) are well-studied, whereas others (e.g. of maximum structural load in a sea state, for a given structure) need to be estimated (e.g. using approximate load models or full time-domain simulation of environmental loads in a given sea state).

Variables summarising sea states themselves also vary in time, but more slowly than short-term variables; their variation also depends on covariates such as wave direction, season, etc. In particular, extreme values of responses tend to be associated with extreme values of sea state variables; it is therefore important to characterise the tails of the distribution of sea state variables well. Statistical analysis of variables summarising consecutive sea states is problematic because of temporal dependence. Hence, it is typical to partition sea states into events, referred to here as "statistical storms", corresponding to contiguous intervals of time, and to then estimate (a) summary variables for whole storms (which can be considered independent given covariates) and (b) the local evolution of sea state variables within a storm given its summary variables. Thus for example, a typical storm might be summarised in terms of the storm peak significant wave height, storm peak direction, season and spectral peak period, and the within-storm evolution of these variables in time given storm peak variables. In the current work, we also explicitly model storm duration.

*Corresponding author `hfh@haw-metocean.com`

Accepted by Ocean Engineering, November 2019

To estimate the "long-term" distribution of response, convolution of the "short-term" distribution of response given sea state with the slowly-varying distribution of sea state variables is required. The method of Tromans and Vanderschuren (1995) convolves the short-term distribution of wave heights with the long-term distribution of most-probable maximum wave height in a storm to provide an estimate of the long-term distribution of the maximum individual wave height. The distribution of individual wave height in a storm conditional on its most probable value is governed by the shape parameter of the Weibull distribution of individual wave heights, and the number N of waves (defined in terms of zero-crossings of surface elevation) over the period of the storm. The value of N and the most probable maximum wave height of historical storms are determined through least-squares regression. The most probable maximum wave heights of historical storms are treated as random variables on which extreme value analysis is performed.

The tail of the distribution of storm peak significant wave height can be described using the generalised Pareto distribution (Pickands 1975). The wind and wave climate typically varies in severity with direction because of variation in available fetch, and directional variation of wind severity. These directional variations are also often associated with seasonal variations, and can be accounted for by fitting extreme value models on a partition of the directional, seasonal or directional-seasonal domain into bins, each of which is assumed to be homogeneous. Specification of appropriate bin sizes is a trade-off between the need for sufficient data per bin for model inference and the requirement for approximate within-bin stationarity. Models estimated using independent bin-by-bin analysis are more uncertain, because of reduced sample size per bin. However, bin-by-bin models are also potentially less biased than a model which ignores covariate effects, because the model can accommodate variation of response with covariates more adequately. Better recent approaches (e.g. Davison and Smith 1990) introduce extreme value models incorporating covariates, but avoiding the need for covariate binning. Jonathan et al. (2014b) use extreme value models with direction, season, longitude and latitude as covariates, imposing smoothness constraints on extreme value parameter variation with covariates.

Storm summary variables (e.g. the storm peak significant wave height and associated spectral peak period) exhibit dependence, as do extreme values of these variables. Since the distribution of response (e.g. maximum individual wave height) depends typically on multiple storm summary variables, it is critical to be able to simulate accurately from the joint tails of distributions of storm summary variables. This requires a suitable model for joint extremes. There are many potential choices, including traditionally those motivated by the work of Longuet-Higgins (1952) and Haver (1987). The conditional extremes model of Heffernan and Tawn (2004) is advantageous because it is suitable to describe a wider class of extremal dependence. The parameters of the conditional extremes model can also be considered as smooth functions of covariates (e.g. Jonathan et al. 2014a).

In estimating the value of response corresponding to an event with annual probability of exceedance of 10^{-4} or smaller, conventionally the effect of uncertainties in the underlying contributing models is neglected. Historically, the mathematical and computational tools to quantify the effect of model uncertainty on predictions of extreme values were not available. The basis for design was a combination of observations, simple statistical and physical modelling approaches, safety factors and good engineering judgement. Specifically, uncertainties were not and could not be accommodated systematically and coherently. Today the situation is different: the most prominent methodology in the statistics literature to quantify uncertainty is Bayesian uncertainty analysis (see e.g. Berger 1985). For example, our uncertainty regarding (e.g.) wind-shear and bottom friction coefficients in a wave hindcast simulator, or in estimating extreme value models from data, can be captured within the statistical analysis, and reflected in estimates for extreme responses, as shown by Jones et al. (2018) in related North Sea work.

Objectives

The objective of the current research is to establish a model for the joint tails of peaks over threshold of storm summary variables, non-stationary with respect to direction and season. The model uses penalised B-spline representations of model parameters for marginal and conditional extremes to accommodate non-stationarity. Bayesian inference is used to estimate parameters, and to propagate sources of uncertainty such as choices of thresholds for marginal and conditional models. Simulation under the estimated models, incorporating between-sea-state variability within a storm, and short-term variability of responses within sea-state, permits estimation of joint long-term distributions of responses.

Layout

The article is laid out as follows. Section 2 introduces the North Sea data used to illustrate the model, and presents the "storm model" used to summarise the characteristics of ocean storms, and outlines statistical approaches to

marginal and conditional extreme value analysis adopted in this work. The procedure used for parameter estimation using Bayesian inference is described in Section 3. Section 4 outlines how simulation under the model is used to estimate distributions of responses corresponding to very long return periods. Results from an application to hindcast data from a location in the Danish sector of the North Sea are presented in Section 5. Section 6 provides discussion and conclusions.

Readers primarily interested in how the model is implemented, might consider reading Section 4 first, before reading from Section 2 in order.

2. The model

In this section, we describe the "storm model" used to isolate so-called "characteristic variables" X, namely important summary statistics of a whole storm used for extreme value modelling. Characteristic variables are isolated from the corresponding time-series of sea state variables $\tilde{X}(s)$. For the application discussed in Section 5, characteristic variables are listed in the third column of Table 1 below, and sea state variables in the fourth column. We then describe a statistical model for the joint distribution of characteristic variables, in particular of their extreme values, varying with directional and seasonal covariates.

2.1. The storm model

The distribution of a response R such as maximum individual wave height in a storm S depends on a set $\tilde{\boldsymbol{X}}(s)$ of sea state variables for sea state s in storm S, such as significant wave height, peak period and directional spreading. The long-term distribution of R also depends on the evolution of sea states within a particular storm, as well as the joint variability of a set of storm characteristic variables. To estimate the long-term distribution of R, we therefore need a hierarchy of models to describe (a) $R|\tilde{\boldsymbol{X}}(s)$, the response given a single sea state s; (b) $\{\tilde{\boldsymbol{X}}(s)\}_{s\in S}|\boldsymbol{X}$, the sea state variables and covariates in time given storm characteristic variables and covariates; and (c) $\boldsymbol{X}$, the storm characteristic variables and covariates.

The proposed storm model is an extension of the response-based approach of Tromans and Vanderschuren (1995). They characterise random independent storm events in terms of the most probable maximum response R_{mpm} in the storm, and an associated equivalent number of waves N in the storm. Unlike Tromans and Vanderschuren (1995), we characterise the storm magnitude, not by the most probable maximum response, but rather by the storm peak significant wave height $H_{m0,p,eq}$ of an "equivalent storm" exhibiting a Gaussian bell-shaped profile in time. Storm duration is then quantified using the standard deviation σ_{eq} of the Gaussian bell, expressed in multiples of the spectral zero-crossing period T_Z. That is, σ_{eq} expresses the "number of waves" in a storm; this parameter plays a similar role in our model to the parameter N of Tromans and Vanderschuren (1995). The most probable response in a storm is therefore a function of both $H_{m0,p,eq}$ and σ_{eq}. The significant wave height at time t relative to the storm peak (at $t = 0$) in the equivalent storm is hence given by

$$H_{m0}(t) = H_{m0,p,eq} \exp\left(-\frac{1}{2}\left(\frac{t}{T_Z \sigma_{eq}}\right)^2\right) . \tag{1}$$

For a given historical storm event, the characteristic variables $H_{m0,p,eq}$ and σ_{eq} are estimated by least-squares minimization of the difference between the most probable maximum response per sea state $R_{mpm}(s)$ for the actual historical and equivalent storms. Figure 1 shows two examples of a true time history of hourly values of H_{m0} (vertical green bars) and equivalent storm representations (black lines). The Forristall (1978) wave height distribution has been chosen here as a representative short-term response distribution, but the estimated values of $H_{m0,p,eq}$ and σ_{eq} are rather insensitive to the choice of short-term distribution. The contribution to maximum short-term response from sea states with H_{m0} less than $75\% - 80\%$ of storm peak H_{m0} ($H_{m0,p}$) is negligible and the "statistical storm´´ may therefore be confined to the sea states with H_{m0} above $75\% - 80\%$ of storm peak. The filled bars in figure 1 mark the sea states of the storms confined to $H_{m0} \geqslant 0.75 H_{m0,p}$.

[Figure 1 about here.]

The distribution of a response R such as maximum individual wave height is typically also dependent on other sea state variables $\tilde{X}$ such as spectral peak period. We therefore include such variables in our storm model, defining their characteristic variables X as weighted averages, with sea-state weights based on the sea state's contribution to

the most probable response $R_{mpm}(S)$ within the whole storm, as follows. The set $\{w_s\}$ of weight factors for each sea states $s \in S$ are computed from the contribution of individual sea states to the most probable maximum response $R_{mpm}(S)$ for the complete storm as

$$w_s = c_w \left(R_{mpm}(S) - R_{mpm}(S \setminus s) \right) \ , \tag{2}$$

where $R_{mpm}(S)$ is the most probable maximum response in the storm considering all sea states, $R_{mpm}(S \setminus s)$ is the most probable maximum response when sea state s is omitted from the storm, and c_w is a normalisation constant set to ensure $\sum_{s \in S} w_s = 1$. Using these weights, characteristic variables X of the form

$$X = \sum_{s \in S} w_s \tilde{X}(s) \tag{3}$$

are estimated and indicated using overbars (e.g. $\overline{T_p}$, $\overline{\text{PWD}}$, etc.) as shown in Table 1 and illustrated as horizontal blue lines in Figure 1. This table gives the full set of characteristic and sea state variables being modelled in the present application.

[Table 1 about here.]

Jones et al. (2018) illustrate the estimation of a discrepancy model for characteristic variables, combining data from a continuous hindcast covering the entire spatial domain of interest with partial measurements at selected locations. The discrepancy model is used to correct for bias in hindcast values, and to estimate uncertainties associated with prediction of measurements at arbitrary locations. We note, as necessary, that we are careful to adjust variables (e.g. for H_{m0} from one-hour and three-hour sea-states) so that they can be fairly compared.

2.2. A statistical model for characteristic variables

Estimation of the statistical model for the joint distribution of characteristic variables X, non-stationary with respect to characteristic covariates, is performed in two stages, as in Ross et al. (2018). First we independently estimate non-stationary marginal models for each characteristic variable in turn. Then we estimate non-stationary conditional extremes models describing the extremal dependence between characteristic variables.

Marginal models

Marginal distributions are estimated for each characteristic variables X (except for covariates, see Table 1). We assume that the marginal probability distribution of X can be expressed as the sum of three parts. Upper and lower tails (defined as exceedances of upper and lower quantile thresholds of the marginal distribution given covariates with specified non-exceedance probabilities) are assumed to follow generalised Pareto (GP) distributions. The remaining central "bulk" of the distribution is described by a truncated gamma distribution. The marginal cumulative distribution function for characteristic variable X can thus be written

$$F_X(x|\alpha,\mu,\xi_1,\zeta_1,\xi_2,\zeta_2) = \begin{cases} F_\Gamma\left(\psi_1|\alpha,\mu\right)\left(1+\frac{\xi_1}{\zeta_1}(\psi_1-x)\right)^{-1/\xi_1} \ , & x \leqslant \psi_1 \\ F_\Gamma(x|\alpha,\mu) \ , & \psi_1 < x \leqslant \psi_2 \\ 1-\left(1-F_\Gamma\left(\psi_2|\alpha,\mu\right)\right)\left(1+\frac{\xi_2}{\zeta_2}(x-\psi_2)\right)^{-1/\xi_2} \ , & x > \psi_2. \end{cases} \tag{4}$$

Subscripts 1 and 2 refer to the lower and upper tails of the distribution respectively and $F_\Gamma(x|\alpha,\mu)$ to the cumulative distribution function of the gamma distribution given by

$$F_\Gamma\left(x|\alpha,\mu\right) = \frac{1}{\Gamma(\alpha)}\gamma\left(\alpha,\frac{\alpha}{\mu}x\right) \ , \tag{5}$$

where $\Gamma(\bullet)$ is the complete gamma function and $\gamma(\bullet,\bullet)$ the lower incomplete gamma function. This particular model parameterisation was motivated by the work of Cox and Reid (1987). Model parameters defining the marginal distributions are therefore the gamma shape α and mean μ, the lower generalised Pareto shape ξ_1 and scale ζ_1, and the corresponding upper tail shape ξ_2 and scale ζ_2. Note specifically that all of these model parameters vary as a function of the relevant covariates; thus for example the model for $H_{m0,p,eq}$ is non-stationary with respect to wave direction and season.

Thresholds ψ_1 and ψ_2 are also non-stationary with respect to covariates, and are set to quantiles of a gamma distribution *fitted to the complete sample*. This fit is performed using Bayesian inference (see Section 3), with sample log-likelihood

$$\ell_\Gamma = \sum_{i=1}^{n_S} \left((\alpha_i - 1) \ln x_i - \tfrac{\alpha_i}{\mu_i} x_i - \ln \Gamma(\alpha_i) - \alpha_i (\ln \mu_i - \ln \alpha_i) \right) , \tag{6}$$

where the additional index i on parameters α and μ indicates that these quantities are evaluated at values of covariates corresponding to storm S_i from a total of n_S storms. Values ψ_1 and ψ_2 are quantiles of the fitted distribution with fixed non-exceedance probabilities κ_1, κ_2 ($> \kappa_1$) such that

$$\begin{aligned} \psi_1 &= F_\Gamma^{-1}(\kappa_1 | \alpha, \mu) , \\ \psi_2 &= F_\Gamma^{-1}(\kappa_2 | \alpha, \mu) . \end{aligned} \tag{7}$$

Threshold uncertainty due to (a) imprecise estimation of α and μ from the sample and (b) lack of knowledge of κ_1, κ_2 is included in subsequent modelling by ensemble forecasting over a range of values for κ_1 and κ_2, sampled at random from uniform distributions over pre-set non-exceedance probability limits, specified following inspection of diagnostic plots for upper and lower tail fits.

Full marginal inference is performed as follows. First α and μ are estimated by fitting the gamma distribution as described above. Then, for κ_1, κ_2 sampled randomly from prior distributions, lower and upper generalised Pareto tails are fitted independently using Bayesian inference to estimate the remaining parameters. Sample log-likelihoods ℓ_1 and ℓ_2 for the lower and upper intervals of the distribution following Equation (4) are therefore

$$\begin{aligned} \ell_1 &= \sum_{i : x_i \leqslant \psi_{1i}} \left(\ln \kappa_1 - \ln \zeta_{1i} - \left(1 + \tfrac{1}{\xi_{1i}} \right) \ln \left(1 + \tfrac{\xi_{1i}}{\zeta_{1i}} (\psi_{1i} - x_i) \right) \right) , \\ \ell_2 &= \sum_{i : x_i > \psi_{2i}} \left(\ln(1 - \kappa_2) - \ln \zeta_{2i} - \left(1 + \tfrac{1}{\xi_{2i}} \right) \ln \left(1 + \tfrac{\xi_{2i}}{\zeta_{2i}} (x_i - \psi_{2i}) \right) \right) . \end{aligned} \tag{8}$$

For computational efficiency, we choose to infer these GP models in terms of parameter set (ξ, ζ^*) in place of (ξ, ζ), where $\zeta^* = \zeta(1 + \xi)$, since maximum likelihood estimates of ξ and ζ^* are asymptotically independent.

Rate of occurrence

The rate ρ of occurrence of a storm event of any magnitude is estimated using an approximation to the Poisson process, following Chavez-Demoulin and Davison (2005) and Jonathan et al. (2014b). Rate ρ is non-stationary with respect to storm covariates. The sample log-likelihood is

$$\ell_\rho = \sum_{k=1}^{n_B} g_k \ln(\rho_k) - \Delta \sum_{k=1}^{n_B} \rho_k , \tag{9}$$

where we assume that the covariate domain has been partitioned into n_B "bins" of (small) constant volume Δ within which rate is stationary. The set $\{g_k\}_{k=1}^{n_B}$ is then the number of occurrences of storm events in each bin, and $\{\rho_k\}_{k=1}^{n_B}$ is the set of corresponding Poisson rates, where ρ_k is evaluated at the values of covariate corresponding to bin k.

Conditional extremes models

The conditional extremes model of Heffernan and Tawn (2004) is defined for variables on a standard marginal Laplace (or alternatively Gumbel) scale. Therefore, having estimated marginal models for each of a set $\boldsymbol{X} = \{X_j\}$ of characteristic variables X, non-stationary with respect to the relevant covariates (see Table 1), we proceed to transform the samples $\{x_{ji}\}_{i=1}^{n_S}$ for each X_j independently to the corresponding standard Laplace samples $\{y_{ji}\}_{i=1}^{n}$ for random variable Y_j following Keef et al. (2013)

$$y_{ji} = \begin{cases} \ln(2F_{X_j}(x_{ji})), & F_{X_j}(x_{ji}) \leqslant 0.5 , \\ -\ln(2((1 - F_{X_j}(x_{ji}))), & F_{X_j}(x_{ji}) > 0.5 . \end{cases} \tag{10}$$

On Laplace Y scale, the conditional extremes model takes the following form for both positive and negative dependence

$$\left(Y_{j'}|Y_j = y\right) = a_{j'|j,i}y + y^{b_{j'|j,i}}\left(m_{j'|j,i} + s_{j'|j,i}Z_{j'|j}\right) \ , \quad y > \nu_j \tag{11}$$

for conditioning characteristic variable Y_j and conditioned characteristic variable $Y_{j'}$, where ν_j is a high quantile of the marginal Laplace distribution with non-exceedance probability λ_j. Parameters $a \in [-1, 1]$, $b \in (-\infty, 1]$, $m \in \mathbb{R}$ and $s > 0$ are non-stationary with respect to covariate, and must be estimated for each combination j, j' of interest. Here for example, $a_{j'|j,i}$ refers to the a parameter for conditional extremes modelling of characteristic variable $Y_{j'}$ conditioned on large values of characteristic variable Y_j, evaluated using covariates for storm S_i. $Z_{j'|j}$ is a residual random variable from an unknown distribution. For parameter estimation purposes only, we assume that $Z_{j'|j} \sim N(0, 1)$ independently for each combination j, j'. Once we have estimated parameters, we estimate the distribution of $Z_{j'|j}$ using the empirical distribution of fit residuals $\{r_{j'ji}\}_{i:y_{ji}>\nu_j}$

$$r_{j'ji} = \frac{1}{s_{j'|j,i}}\left(\left(y_{j'i} - a_{j'|j,i}y_{ji}\right)y_{ji}^{-b_{j'|j,i}} - m_{j'|j,i}\right) \ , \tag{12}$$

where estimates from the posterior distribution of parameters (see Section 3) are substituted for the parameters themselves on the right hand side. Joint extremal dependencies $\{Y_{j'_1}, Y_{j'_2}, ...\}|\{Y_j = y\}$ are estimated by first fitting each of the pairwise dependencies $\{Y_{j'_k}\}|\{Y_j = y\}$, $k = 1, 2, ...$, for some fixed choice of ν_j, and then ensuring that the joint residual set $\{r_{j'_1ji}, r_{j'_2ji}, ...\}_{i:y_{ji}>\nu_j}$ is assembled and sampled appropriately (i.e. using the same storm S_i to assemble and sample across residuals for models of different characteristic variables) to retain the dependence between conditioned variates. The log-likelihood for sample $\{y_{ji}, y_{j'i}\}_{i:y_{ji}>\nu_j}$ is given by

$$\ell_{CE,j'|j} = -\sum_{i:y_{ji}>\nu_j(\lambda_j)}\left\{\frac{1}{2}\ln(2\pi) + \ln s_{j'|j,i}y_{ji}^{b_{j'|j,i}} + \frac{\left(y_{j'i} - \left(a_{j'|j,i}y_{ji} + m_{j'|j,i}y_{ji}^{b_{j'|j,i}}\right)\right)^2}{2\left(s_{j'|j,i}y_{ji}^{b_{j'|j,i}}\right)^2}\right\} \ . \tag{13}$$

Threshold ν_j is set independently of the generalised Pareto threshold ψ_{2j}. As for marginal thresholds, uncertainty in the specification of ν_j is incorporated in inference by sampling non-exceedance probability λ_j from a uniform distribution over a pre-specified range of reasonable threshold non-exceedance probabilities.

P-spline representation of covariate effects

A full description of the covariate representation used in this work is given in Randell et al. (2016). This section provides a motivating summary; readers are referred to Randell et al. (2016) for details. Penalised B-splines (also referred to as P-splines) are used to describe model parameter variation with covariate on some domain. The basic idea of penalised B-splines, originally introduced by Eilers and Marx (1996), is to use a basis set of B-splines with a moderately large number of evenly-spaced knots to characterise an arbitrary function flexibly, but then to control spline smoothness by penalising function roughness. B-spline regression can be explained as follows. First we partition the covariate domain into n' equal intervals by specifying the position of $n' + 1$ knots. B-spline basis functions $\{B_k\}_{k=1}^{n'+q}$ are then constructed as a sequence of polynomial functions of degree q connected at the knots. Each B_k is positive on an interval spanning $q + 2$ knots (for an aperiodic domain), and is zero elsewhere. Parameter estimation using B-splines then consists of finding coefficients $\{\beta_k\}_{k=1}^{n'+q}$ such that the value of any function $\eta(\theta)$ (or any of the model parameters and thresholds from Equation (4) and Equation (11) in the present work) of interest at covariate value θ is expressed as the linear combination

$$\eta(\theta) = \sum_{k=1}^{n'+q}\beta_k B_k(\theta) = \boldsymbol{B}(\theta)\boldsymbol{\beta} \ , \tag{14}$$

where $\boldsymbol{B}(\theta) = \{B_k(\theta)\}$ and $B_k(\theta)$ is the value of the k^{th} B-spline basis at θ, and $\boldsymbol{\beta} = \{\beta_k\}$ is the vector of spline coefficients. Roughness is quantified in terms of the quadratic form

$$\sum_{k=1}^{n'+q}\sum_{k'=1}^{n'+q}\beta_k K_{kk'}\beta_{k'} = \boldsymbol{\beta}'\mathbf{K}\boldsymbol{\beta} \ , \tag{15}$$

where $\mathbf{K} = \{K_{kk'}\}$ is a penalty matrix. In the current work using Bayesian inference, the quadratic form in Equation (15) motivates the prior specification for $\boldsymbol{\beta}$, discussed in Section 3.2. The first order penalty matrix for knots on an arbitrary covariate domain is given by:

$$\mathbf{K} = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix}. \tag{16}$$

This penalty penalises differences between adjacent values of β_k. In the current work we deal with periodic covariates such as direction and season, for which periodic spline bases and penalties are required. In periodic cases only q spline basis functions are required, and the first order penalty matrix becomes

$$\mathbf{K} = \begin{bmatrix} 2 & -1 & & & \cdots & -1 \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ \vdots & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ -1 & & & & -1 & 2 \end{bmatrix}. \tag{17}$$

B-splines can be extended to higher dimensions as tensor-product B-splines (see e.g. Eilers and Marx (2003)). A tensor-product B-spline in two dimensions is illustrated in Figure 2, for estimation of directional-seasonal quantile ψ for $H_{m0,p,eq}$, for different choices of directional and seasonal roughness coefficients τ^2. Coloured shapes underlying the surface are individual tensor-product B-spline basis functions scaled by the respective coefficients. The total number of β coefficients to be estimated in two dimensions is given by the product of the number of spline coefficients in each of the one-dimensional margins. Different numbers of knots, spline orders and roughness penalties may be used on different margins.

[Figure 2 about here.]

Note that inference described in the next section exploits generalised linear array models (GLAMs; Currie et al. 2006, Eilers et al. 2006) permitting computationally-efficient analysis of tensor products of splines; see Randell et al. 2016) for further information. A full list of model parameters estimated in terms of tensor-product B-splines is given in Table 2.

[Table 2 about here.]

Covariate transformation

We also note that individual directional and seasonal covariates are not typically uniformly distributed on their marginal domains. For this reason, is it helpful to transform covariates marginally following Jonathan et al. (2013), so that the original set of covariate values $\{\theta_i\}_{i=1}^{n_S}$ is related to the transformed set $\{\theta_i^*\}$ by

$$\theta_i^* = \frac{360}{n_S}\left(r(\theta_i) - 1\right) \ , \tag{18}$$

where $r(\theta_i)$ is the rank of θ_i in the set of covariates, namely the position of θ_i in the set of covariate values sorted in ascending order. The transformed set is uniformly distributed on $[0, 360)$ by design, stabilising parameter estimation on the transformed scale. Interpreted on the original scale, the transformation imposes greater smoothness for values of covariate which are less frequently observed in the sample, and allows greater parameter flexibility for more frequently observed values, in a natural way according to the rate of occurrence of events for different covariate values.

3. Inference

A posterior estimate for the joint distribution of all marginal (gamma and GP) and conditional extremes model parameters, as described in Section 2.2 and Table 2, is estimated using Bayesian inference. Likelihoods for all models, and a description of the spline parametrisation used, is provided in Section 2. Here, we provide a brief discussion of prior specification and the inference procedure.

3.1. Distribution for hyper-parameters

Distributions for marginal threshold non-exceedance probabilities κ_1, κ_2 are set by inspection of GP tail diagnostic plots. Conditional extremes non-exceedance probabilities $\{\lambda_j\}$ are similarly set by inspection of conditional extremes model diagnostics; we do not learn about these hyper-parameters during inference.

Specification of a reasonable distribution for B-spline roughness coefficients τ^2 is challenging. For example, inference for each of the gamma and GP marginal directional-seasonal models involves estimation of two parameters, each of which has a roughness coefficient in direction and season. Thus, specifically, the GP high tail marginal model requires specification of a joint prior for four roughness coefficients $\tau^2_{\xi\theta}$, $\tau^2_{\xi\phi}$, $\tau^2_{\zeta^*\theta}$ and $\tau^2_{\zeta^*\phi}$, where θ and ϕ refer to direction and season respectively. We estimate reasonable combinations for these hyper-parameters using a cross-validation scheme. The intention of the scheme is to identify combinations of τ^2 which yield good predictive performance of any model of interest. We partition the sample into n_{CV} blocks, withhold one block and estimate the model for the remaining blocks using Bayesian inference, for specified choices of τ^2 parameters. Then we evaluate the predictive log-likelihood using the withheld block: large log-likelihood values indicate good predictive performance. We repeat this procedure until each block has been withheld exactly once, and use the sum of predictive log-likelihoods over all blocks as an estimate of the relative predictive performance for different choices of τ^2. We exponentiate the predictive log-likelihood, and normalise its integral to unity over the four-dimensional domain of τ^2, using the resulting empirical density as a prior for the joint distribution of $\tau^2_{\xi\theta}$, $\tau^2_{\xi\phi}$, $\tau^2_{\zeta^*\theta}$ and $\tau^2_{\zeta^*\phi}$. Since this scheme is computationally slow for four or more τ^2s, we use an approximate calculation which proceeds as follows. First, (a) we use the cross-validation scheme to estimate a common τ^2 across all four margins. Then (b) we use the cross-validation scheme to find an optimal ratio of $\tau^2_\xi/\tau^2_{\zeta^*}$ for both direction and season starting from the solution to (a). Then finally (c) use the cross-validation scheme to evaluate the predictive performance on a two-dimensional $\tau^2_{\xi\theta} \times \tau^2_{\xi\phi}$ domain with fixed (optimal) $\tau^2_\xi/\tau^2_{\zeta^*}$ ratio. Figure 3 shows the variation of predictive log-likelihood with $\tau^2_{\xi\theta}$ and $\tau^2_{\xi\phi}$, with optimal ratio $\log_{10}(\tau^2_\xi) - \log_{10}(\tau^2_{\zeta^*}) = -1.7$. We do not learn about τ^2 in the Bayesian inference.

[Figure 3 about here.]

3.2. Prior for B-spline coefficients

All other model parameters, described using tensor-product B-splines, require specification of a prior distribution for the corresponding vector of spline coefficients $\boldsymbol{\beta}$. Motivated by the outline in Section 2, following Green and Silverman (1994), the prior density is given by

$$f(\boldsymbol{\beta}|\tau^2) \propto \frac{1}{(\tau^2)^{\frac{\text{rk}(\mathbf{K})}{2}}} \exp\left(-\frac{1}{2\tau^2}\boldsymbol{\beta}'\mathbf{K}\boldsymbol{\beta}\right) , \tag{19}$$

where $\text{rk}(\mathbf{K})$ is the rank of the penalty matrix. Thus as the value of τ^2 increases, the prior distribution for $\boldsymbol{\beta}$ becomes narrower and the resulting spline estimate less variable.

3.3. MCMC proposal generation

Posterior distributions are approximated using Markov Chain Monte Carlo with a Metropolis-Hastings (MH) sampling scheme. For example, for the GP high tail, the basic MH scheme seeks to estimate the posterior distribution

$$f(\boldsymbol{\beta}_\xi, \boldsymbol{\beta}_{\zeta^*}|\text{Data}) \propto f(\text{Data}|\boldsymbol{\beta}_\xi, \boldsymbol{\beta}_{\zeta^*}) \times f(\boldsymbol{\beta}_\xi, \boldsymbol{\beta}_{\zeta^*})$$

by iteratively sampling from the set of full conditionals

$$\begin{aligned} f(\boldsymbol{\beta}_\xi|\text{Data}, \boldsymbol{\beta}_{\zeta^*}) &\propto f(\text{Data}|\boldsymbol{\beta}_\xi, \boldsymbol{\beta}_{\zeta^*}) \times f(\boldsymbol{\beta}_\xi) \\ f(\boldsymbol{\beta}_{\zeta^*}|\text{Data}, \boldsymbol{\beta}_\xi) &\propto f(\text{Data}|\boldsymbol{\beta}_\xi, \boldsymbol{\beta}_{\zeta^*}) \times f(\boldsymbol{\beta}_{\zeta^*}) \end{aligned}$$

for specified choices of hyper-parameters κ_1, κ_2 and $\{\tau^2\}$. For example, a proposal $\boldsymbol{\beta}_\xi^\dagger$ for $\boldsymbol{\beta}_\xi$ is generated from the current state using a proposal density g. In this work, we assume g represents a random walk from a multivariate Gaussian density with covariance matrix $\boldsymbol{C}$. Well-chosen correlated proposals improve the rate of convergence, and mixing of MCMC chains. Therefore, following Rue (2001) and Lang and Brezger (2004) for starting iterations, we set $\boldsymbol{C} = (\boldsymbol{B}_2^T\boldsymbol{B}_2 + \boldsymbol{P})^{-1}$ where matrix $\boldsymbol{P}$ incorporates four τ^2s and four penalty matrices $\boldsymbol{K}$, and $\boldsymbol{B}_2$ represents a directional-seasonal tensor-product B-spline basis matrix as explained in Randell et al. (2016). After sufficient iterations, we then follow Roberts and Rosenthal (2009) in setting

$$\boldsymbol{C} = 2.38^2(1-\epsilon)^2\frac{\boldsymbol{\Sigma}}{p} + 0.01\epsilon^2\frac{\boldsymbol{I}_p}{p} \ , \tag{20}$$

where p is the number of parameters being estimated, $\boldsymbol{\Sigma}$ is the empirical covariance matrix estimated from previous iterations of the Markov chain, and $\boldsymbol{I}_p$ is a $p \times p$ identity matrix. We use $\epsilon = 0.05$ as recommended by Roberts and Rosenthal (2009). The proposed state is then accepted with probability

$$\min\left(1, \frac{f(\boldsymbol{\beta}_\xi^\dagger|\text{Data}, \boldsymbol{\beta}_{\zeta^*})g(\boldsymbol{\beta}_\xi^\dagger \to \boldsymbol{\beta}_\xi)}{f(\boldsymbol{\beta}_\xi|\text{Data}, \boldsymbol{\beta}_{\zeta^*})g(\boldsymbol{\beta}_\xi \to \boldsymbol{\beta}_\xi^\dagger)}\right) \ .$$

3.4. Full model inference

The procedure detailed above is appropriate for estimating the posterior distribution of parameters for any one of the marginal gamma, lower or upper GP tail for any characteristic variable X_j, or any conditional extremes model $X_{j'}|X_j$. Full model inference requires estimation of parameters of all model components in a hierarchical order. As outlined in Section 2, marginal distributions are first estimated for each X_j in turn; conditional extremes models are then estimated for $\{X_{j'}\}_{j'\neq j}|X_j$ for each X_j in turn. This is described in more detail below.

Initially (a) the rate of occurrence model for storm events is fitted across the covariate domain. Then marginal analysis is made for each characteristic variable (e.g. $H_{m0,p,eq}$, $\overline{T_p}$), involving the following steps: (b) Fit the gamma distribution to all events and save a number of independent posterior samples of parameters from the MCMC chain. (c) For each sample of the gamma parameters from (b), sample a low threshold probability, compute the extreme value threshold, perform GP inference, and save a number of independent posterior samples of parameters. (d) Repeat (c) for the high GP tail. Steps (b)-(d) results in the generation of n_Γ samples of gamma parameters, and $n_\Gamma \times n_{GP}$ GP samples.

For a given conditioning variate, all required conditional extremes models can then be estimated contemporaneously to accumulate vectors of residuals, preserving dependencies between residuals corresponding to different choices of marginal and conditional extremes parameters; these dependencies can then be carried over into storm simulations. Uncertainty in the conditional extremes model threshold probabilities $\{\lambda_j\}$ is accounted for by sampling a new value for each new selection of marginal GP model parameters used. For conditioning variate X_j, the inference procedure is as follows: (e) Sample a threshold non-exceedance probability λ_j and identify events for which the value of X_j exceeds this. (f) Estimate conditional extremes models for $\{X_{j'}\}_{j'\neq j}|X_j$ for each conditioning variate X_j in turn. Retain posterior parameter estimates and sets of residuals for the last iteration of the MCMC only. (g) Accumulate an array of conditional extremes residuals. Steps (e)-(g) generate $n_\Gamma \times n_{GP}$ sets of conditional extremes model parameters and residuals.

The above procedure (a)-(g) yields an equal number of posterior samples of marginal and conditional model parameters, and associated residuals, providing a characterisation of the marginal and joint structure of the set of characteristics variables X. A number of threshold choices for both marginal tails and conditional extremes are incorporated in this sample, accounting for threshold uncertainty, on a specified prior interval of threshold non-exceedance probabilities. It is our experience that this straightforward approach to incorporation of threshold variability is useful, and superior to ignoring the effects of threshold choice.

4. Estimation of extreme values

The procedure described in the previous section is used to simulate realisations of characteristic variables X corresponding to long return periods; this allows estimation of extreme values for those characteristic variables. Historical storm trajectories are then allocated to realisations of characteristic variables from simulation, by matching values of simulated and historical characteristics variables; this allows simulation of full intra-storm time-series

$\{\tilde{\boldsymbol{X}}(s)\}|\boldsymbol{X}$ for arbitrary return periods for each sea state s in a storm. As storm trajectories are simulated for each storm event, this procedure further allows convolution of long-term distributions of sea state variables with short-term (within sea state) distributions of one or more responses, and hence the estimation of the long-term distribution of response: a typical example in this respect is convolution of the long-term distribution of sea states H_{m0} with the short-term distribution of maximum crest height (within sea state) to obtain the long-term distribution of the maximum crest elevation. Extreme values are thence obtained by straightforward simulation. In layman's terms, such a simulation simply consists in sampling a very large number of storms, storm trajectories and short-term responses, reading off the T-year extreme value as the $\lceil P/T \rceil^{\text{th}}$ largest value in a simulation of P years. We note alternative methods for deriving extreme values from numerical integration, presented in Ross et al. (2017), and from importance sampling.

4.1. Simulation of storm characteristic variables

The simulation procedure followed to simulate characteristic variables in one year of (statistical storm) events is as follows: (a) Sample a joint set of marginal and conditional extremes model parameters to use; (b) Sample the number of events to be simulated from a Poisson distribution with expected annual rate of occurrence; (c) Assign covariate values to each storm event using the fitted non-stationary rate function for each conditioning variable; (d) Sample the magnitude of the conditioning variable from its marginal non-stationary distribution; (e) Estimate magnitudes of conditioned variables above the conditional extreme model quantile threshold using the conditional extremes model. For non-exceedances of the conditional extremes threshold, estimate the magnitude of conditioned variables by sampling with replacement from the original sample of threshold non-exceedances; and (f) store the annual maximum value observed, and values of other variables given occurrences of annual maxima of conditioning variates. The T-year extreme values are typically estimated as a quantile of the distribution of the annual maximum with non-exceedance probability $\exp(-1/T)$. The period P of simulation (in years) needs to be considerably longer than the return period T of interest. In the current work, we impose $P \geqslant 100T$; in other words, a 100 year extreme value requires simulation of around 10.000 years.

Extreme values found in this way incorporate sources of epistemic uncertainty from model estimation as described above. Larger model-fitting uncertainty will inflate extreme values; inflation can be large for return periods considerably longer than the period P_0 of the historical data.

4.2. Simulation of storm trajectories

The hourly evolution of each storm event is required to estimate the long-term distribution of short-term responses. For each storm event simulated using the procedure described in Section 4.1, we achieve this by adopting a historical storm trajectory, the values of characteristic variables for which are similar to those of the simulated storm event, and associate it after some adjustment with the simulated storm, following Feld et al. (2015). In short, we (a) identify a sample of historical storms similar to the simulated storm event; (b) select one of these at random, the "matched" storm, and (c) scale, stretch and rotate the matched historical storm trajectory such that characteristic storm variables calculated from the modified trajectory matches those of the simulated event.

Historical storm events, similar to the simulated storm, are found by computing a "storm dissimilarity" for all historical storms given the simulated storm event. For historical observations $\{x_{ji}^{\text{H}}\}_{i=1}^{n}$ of each of p characteristic variables X_j, and the corresponding simulated value x_{jk}^{S}, dissimilarity is calculated using

$$d_{jik} = \frac{|x_{ji}^{\text{H}} - x_{jk}^{\text{S}}|}{\sigma_j^{\text{H}}} \tag{21}$$

for historical storm S_i and simulated storm S_k. σ_j^{H} is the standard deviation of X_j over all historical storms, used to standardise the dissimilarity for each X_j. Then the overall dissimilarity of S_k from S_i is calculated as

$$d_{ik}^2 = \sum_{j=1}^{p} d_{jik}^2 \ . \tag{22}$$

For each S_k, a matched historical storm S_{i^*} is then selected from the set of historical storms yielding one of the lowest values of d_{ik}. This is then adjusted (as described in the next paragraph) and adopted as the storm trajectory

for simulated storm S_k. Typically, the matched storm is selected from the 20 least dissimilar storms; experience suggests that inferences are not overly sensitive to this number.

Next we sample the matched historical storm trajectory $\tilde{x}^{\mathrm{H}}_{ji^*}(s)$, adjusting it such that the characteristic variable from the adjusted storm trajectory is equal to that of the simulated storm. Adjustment is performed using a constant linear additive term (for directional covariates only) and a constant linear scaling factor (otherwise) applied to the entire storm trajectory. Thus for a (non-covariate) characteristic variable X_j such as $H_{m0,eq}$ and simulated storm S_k, we define scale factor υ_{jk}

$$\upsilon_{jk} = \frac{x^{\mathrm{S}}_{jk}}{x^{\mathrm{H}}_{ji^*}} \tag{23}$$

and use it to adjust the matched historical storm trajectory $\tilde{x}^{\mathrm{H}}_{ji^*}(s)$ such that

$$\tilde{x}^{\mathrm{S}}_{jk}(s) = \upsilon_{jk} \times \tilde{x}^{\mathrm{H}}_{ji^*}(s) \ . \tag{24}$$

Wave, wind and current directional covariates are corrected in the analogous manner. Typically, peak or mean wave direction is used as characteristic covariate for marginal and conditional extremes models; wind and current directions are not modelled statistically. The same additive directional correction is used to rotate all of wave, wind and current directions, such that wind-wave and current-wave misalignment from the historical storm is maintained in the simulated storm.

The duration of the simulated storm event is a function of the simulated characteristic storm variables σ_{eq} (number of waves in the storm) and $\overline{T_{02}}$ (average duration in seconds of each wave). So that the simulated storm trajectory has the right duration, the time axis for the simulated storm is adjusted according to

$$t^{\mathrm{S}}_{jk} = t^{\mathrm{H}}_{ji^*} \times \upsilon_{jk\overline{T_{02}}} \times \upsilon_{jk\sigma_{eq}} \ , \tag{25}$$

with $\upsilon_{jk\overline{T_{02}}}$ and $\upsilon_{jk\sigma_{eq}}$ being the scaling factors applicable for $\overline{T_{02}}$ and storm duration σ_{eq}, respectively.

5. Application to North Sea

The marginal and conditional extremes models introduced in Section 2 are estimated for hindcast time-series of wave, current and wind. A large number of marginal and conditional extremes models are estimated; only a small subset of these are presented here, corresponding to non-stationary directional-seasonal models, with $H_{m0,p,eq}$ used as conditioning variate for conditional extremes models. Space also prevents presentation of all the diagnostic information generated during model estimation. Nevertheless we hope that the illustrations in this section give at least a flavour of the analysis.

5.1. Hindcast data and exploratory data analysis

The sample used for model estimation is taken from a hindcast at a location in the Danish sector of the North Sea, for a period of approximately 37 years. The hindcast model uses CFSR wind fields (Saha et al. 2010, Saha et al. 2014) to force local models for waves, water levels and currents using MIKE21 Spectral Waves (Sørensen et al. 2005) and MIKE21 Hydrodynamics (DHI 2017) respectively. Independent statistical storm events are then isolated and characteristic variables and covariates estimated.

Figure 4 illustrates the directional and seasonal variation of $H_{m0,p,eq}$. Directions of increased rate of occurrence of storms are apparent, reflecting fetch effects; the most severe events emerge from approximately 315°N. $H_{m0,p,eq}$ shows smooth seasonal variation in storm severity; the most severe events occur in the winter months from November to February. Figure 5 shows the characteristic peak period $\overline{T_p}$ and residual water level (or "surge") $\overline{\mathrm{WL}_{resi}}$ on $H_{m0,p,eq}$, and Figure 6 shows the same parameters on the characteristic wave direction $\overline{\mathrm{PWD}}$. The relationship between $\overline{\mathrm{WL}_{resi}}$ and $H_{m0,p,eq}$ appears rather unclear from inspection of Figure 5. However, comparing Figure 6 with Figure 4 suggests that, given $\overline{\mathrm{PWD}}$, this relationship is somewhat more straightforward. Large negative surge amplitudes are related to events from the south-east, whereas the largest positive surge amplitudes occur for waves from the north-west. This direction also coincides with the largest values of both $H_{m0,p,eq}$ and $\overline{T_p}$ and corresponds to the direction of largest free fetch.

[Figure 4 about here.]

[Figure 5 about here.]

[Figure 6 about here.]

5.2. Marginal modelling

Posterior median estimates for marginal model parameters of $H_{m0,p,eq}$ are shown in Figure 7 as function of the direction and season. Estimates for gamma shape α and scale μ both exhibit minima for summer storms from approximately 180°N. The estimate for upper tail generalised Pareto shape ξ_2 does not show much variability with covariates. The corresponding scale parameter ζ_2 exhibits a clear maximum for winter storms from approximately 315°N, reflecting evidence from Figure 4. We note that, to generate this estimate, we have adjusted samples of generalised Pareto scale estimates from the MCMC so that they all correspond to a marginal non-exceedance probability ψ_2 of 0.4 .

[Figure 7 about here.]

Posterior distributions of parameter estimates from Figure 7 are further illustrated in Figure 8, for "season" corresponding to the 15th January. From the figure, it is clear that directional variation is of importance for all parameters except ξ_2 in this case.

[Figure 8 about here.]

Figure 9 compares empirical tails of the distribution of $H_{m0,p,eq}$, generated from Monte-Carlo simulations under the model, with the actual hindcast sample for different directional intervals. A large number of realisations of the same length (37 years) as the hindcast were simulated. Marginal parameter uncertainty was included by randomly drawing new parameter estimates from the sample of posterior estimates for every realisation of 37 years. Posterior median and 95% credible intervals for the tail are illustrated in the figure. The hindcast tail is generally found to lie within the 95% uncertainty interval as would be expected from a reasonable model.

[Figure 9 about here.]

5.3. Conditional extremes modelling

Posterior median parameter estimates for a, b, m and s from the conditional extremes model of characteristic residual water level $\overline{\mathrm{WL}_{resi}}$ conditioned on $H_{m0,p,eq}$ are shown in Figure 10 as functions of direction and season. For all parameters, it is clear that directional variability is more prominent than seasonal variability; further, directions around 315° show the greatest directional variability.

[Figure 10 about here.]

Posterior distributions of parameters from Figure 10 as a function of direction are illustrated in Figure 11, again for 15$^{\mathrm{th}}$ January.

[Figure 11 about here.]

There is directional variation in conditional extremes model parameter estimates, but practically no seasonal variation. This tallies with physical understanding of processes at play in a semi-enclosed basin like the North Sea. Strong winds result in large significant wave heights for effectively all directions; the same winds drive negative surge for some directions and positive surge for others. It appears that these processes do not result in seasonal variation in extremal dependence, over and above that already captured by marginal extremal models for $H_{m0,p,eq}$ and $\overline{\mathrm{WL}_{resi}}$. The directional dependence between storm surge and extreme significant wave height is further illustrated in Figure 12. This figure illustrates the modelled distribution of storm surge conditional on $H_{m0,p,eq}$ exceeding the 1-year and 100-year quantile respectively, as function of peak wave direction. The model appears to capture both the negative dependence for wave directions in the interval of 225°N-270°N and the positive dependence for wave directions around 315°N. It also provides reasonable estimates of the variance of the storm surge associated with extremes of $H_{m0,p,eq}$, as illustrated by the quantiles of the conditioned distributions.

[Figure 12 about here.]

Figures 13, 14 and 15 illustrate the output of simulations under the fitted model. The left hand panel of Figure 13 shows characteristic $H_{m0,p,eq}$ against $\overline{\mathrm{PWD}}$; the right hand panel shows the corresponding relationship for hourly sea-state variable H_{m0} on PWD. As described in the figure caption, the colour of a simulated event varies with the rate of occurrence of that event (thereby making it possible to perceive variations in rate of occurrence even

when individual events are no longer identifiable in the figure). The original hindcast sample has been added to the plots for reference. The simulation corresponds to 50,000 years with only significant wave heights above 4 meters included. Contours of constant probability density are also shown as solid lines; the density for all points on this contour is equal to the maximum density (i.e. rate of occurrence, over all directions) corresponding to the marginal 10- and 100-year extreme value estimates.

[Figure 13 about here.]

Figure 14 shows realisations of characteristic storm length σ_{eq} (on natural logarithmic scale) against $H_{m0,p,eq}$ (left) and $\overline{\text{PWD}}$ (right). σ_{eq} decreases with increasing storm severity. Directional variation is rather weak, with some evidence that storms from approximately 135°N are more persistent; no meteorological explanation for this is offered here.

[Figure 14 about here.]

The left hand panel of Figure 15 shows characteristic spectral peak period $\overline{T_p}$ on $H_{m0,p,eq}$; the right hand panel gives the corresponding plot of T_p on H_{m0} for hourly sea states. The variance of T_p for given H_{m0} is generally larger than that of $\overline{T_p}$ for the same value of $H_{m0,p,eq}$, both for hindcast and simulated data. This effect is caused by the fact that sea states during storm rise are steeper (i.e. larger H_{m0}/T_p^2) than at storm peak, and less steep during storm decay; this feature may be important to capture well in storm modelling since wave breaking is more frequent in steep sea states. For certain short-term responses, this effect might be as important as variation of significant wave height around the storm peak.

[Figure 15 about here.]

6. Discussion and conclusions

We present an approach to joint modelling of multiple features of wind-driven storm events in a statistically-rigorous and physically-reasonable manner. The method is comprised of (a) estimation of summary variables for storms, referred to in this work as characteristic variables (b) marginal extreme value modelling of characteristic variables, non-stationary with respect to covariates, (c) non-stationary conditional extremes modelling of characteristic variables given extreme values of a different conditioning characteristic variable, (d) subsequent modelling of within-storm evolution of related sea state variables, corresponding to storm rise and decay, and (e) simulation under the fitted model to estimate joint samples of storm characteristic and sea state variables corresponding to arbitrary return periods, and thereby estimation of extreme values for structural design and reassessment.

Inference is performed within a Bayesian framework, allowing estimation of full joint posterior distributions, specification of appropriate distributions for hyper-parameters and prior distributions for model parameters encoding prior engineering knowledge, and propagation and quantification of uncertainty in a consistent fashion. Tensor products of penalised B-splines are use to provide general representations of the domain of covariates, providing flexible modelling of non-stationarity. The storm model used provides a reasonable approach to isolation of statistically independent storms, represented by characteristic variables and covariates, facilitating modelling of independent events. No prior assumptions regarding storm shape, the directional and seasonal dependence of storms, or the (extremal) dependence between characteristic variables is necessary. Given characteristic variables, historical storm trajectories representing storm rise and decay are exploited to incorporate the effects of storm duration and within-storm variation. Seasonal-directional design criteria may be obtained from the model by simulation, and simulated sea state data may also be used directly for subsequent structural reliability assessment, such as that carried out for the Tyra field (Tychsen et al. 2016).

The model has been applied to hindcast data for a location in the central North Sea. Diagnostics demonstrate that the model makes reasonable predictions of extreme events for return periods of the order of the length of the data record for the hindcast sample.

Acknowledgements

We thank Graham Feld and Emma Ross of Shell for assistance at various stages of this work. We further acknowledge the support of colleagues at DHI, Total E&P Danmark A/S and Shell.

References

Berger, J.O., 1985. Statistical decision theory and Bayesian analysis. Springer-Verlag, Berlin.

Chavez-Demoulin, V., Davison, A., 2005. Generalized additive modelling of sample extremes. J. Roy. Statist. Soc. Series C: Applied Statistics 54, 207–222.

Cox, D.R., Reid, N., 1987. Parameter orthogonality and approximate conditional inference. J. Roy. Statist. Soc. B 49, 1–39.

Currie, I.D., Durban, M., Eilers, P.H.C., 2006. Generalized linear array models with applications to multidimensional smoothing. J. Roy. Statist. Soc. B 68, 259–280.

Davison, A., Smith, R.L., 1990. Models for exceedances over high thresholds. J. R. Statist. Soc. B 52, 393.

DHI, 2017. MIKE 21 and MIKE 3 Flow Model FM, Hydrodynamic and Transport Module, Scientific Documentation. Technical Report. DHI.

Eilers, P.H., Marx, B.D., 2003. Multivariate calibration with temperature interaction using two-dimensional penalized signal regression. Chemometrics and Intelligent Laboratory Systems 66, 159 – 174.

Eilers, P.H.C., Currie, I.D., Durban, M., 2006. Fast and compact smoothing on multi-dimensional grids. Comput. Statist. Data Anal. 50, 61–76.

Eilers, P.H.C., Marx, B.D., 1996. Flexible smoothing with b-splines and penalties. Statistical Science 11, 89–102.

Feld, G., Randell, D., Wu, Y., Ewans, K., Jonathan, P., 2015. Estimation of storm peak and intra-storm directional-seasonal design conditions in the North Sea. J. Offshore. Arct. Eng. 137, 021102:1–021102:15.

Forristall, G.Z., 1978. On the statistical distribution of wave heights in a storm. J. Geophys. Res. 83, 2353–2358.

Green, P.J., Silverman, B., 1994. Nonparametric regression and generalised linear models: A roughness penalty approach. Chapman and Hall, London, UK.

Haver, S., 1987. On the joint distribution of heights and periods of sea waves. Ocean Eng. 14, 359–376.

Heffernan, J.E., Tawn, J.A., 2004. A conditional approach for multivariate extreme values. J. R. Statist. Soc. B 66, 497–546.

Jonathan, P., Ewans, K., Randell, D., 2013. Joint modelling of extreme ocean environments incorporating covariate effects. Coastal Engineering 79, 22 – 31.

Jonathan, P., Ewans, K.C., Randell, D., 2014a. Non-stationary conditional extremes of northern North Sea storm characteristics. Environmetrics 25, 172–188.

Jonathan, P., Randell, D., Wu, Y., Ewans, K., 2014b. Return level estimation from non-stationary spatial data exhibiting multidimensional covariate effects. Ocean Eng. 88, 520–532.

Jones, M., Hansen, H.F., Zeeberg, A.R., Randell, D., Jonathan, P., 2018. Uncertainty quantification in estimation of ocean environmental return values. Coastal Engineering 141, 36 – 51.

Keef, C., Papastathopoulos, I., 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. Mult. Anal. 115, 396–404.

Lang, S., Brezger, A., 2004. Bayesian p-splines. J. Computnl Graph. Statist. 13, 183–212.

Longuet-Higgins, M.S., 1952. On the statistical distribution of the height of sea waves. J. Mar. Res. 11, 245–266.

Pickands, J., 1975. Statistical inference using extreme order statistics. Ann. Statist. 3, 119–131.

Randell, D., Turnbull, K., Ewans, K., Jonathan, P., 2016. Bayesian inference for non-stationary marginal extremes. Environmetrics 27, 439–450.

Roberts, G.O., Rosenthal, J.S., 2009. Examples of adaptive MCMC. J. Comp. Graph. Stat. 18, 349–367.

Ross, E., Randell, D., Ewans, K., Feld, G., Jonathan, P., 2017. Efficient estimation of return value distributions from non-stationary marginal extreme value models using Bayesian inference. Ocean Eng. 142, 315–328.

Ross, E., Sam, S., Randell, D., Feld, G., Jonathan, P., 2018. Estimating surge in extreme North Sea storms. Ocean Eng. 154, 430–444.

Rue, H., 2001. Fast sampling of gaussian markov random fields. Journal of the Royal Statistical Society. Series B (Statistical Methodology) 63, 325–338.

Saha, S., Moorthi, S., Pan, H.L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.T., Chuang, H.y., Juang, H.M.H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R.W., Rutledge, G., Goldberg, M., 2010. The NCEP climate forecast system reanalysis. Bulletin of the American Meteorological Society 91, 1015–1058.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.T., Chuang, H.y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M.P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., Becker, E., 2014. The NCEP climate forecast system version 2. Journal of Climate 27, 2185–2208.

Sørensen, O.R., Kofoed-Hansen, H., Rugbjerg, M., Sørensen, L.S., 2005. A third-generation spectral wave model using an unstructured finite volume technique, in: Proceedings of the 29th International Conference on Coastal Engineering, pp. 894–906.

Tromans, P.S., Vanderschuren, L., 1995. Variable based design conditions in the North Sea: application of a new method. Offshore Technology Confernence, Houston (OTC–7683) .

Tychsen, J., Risvig, S., Hansen, H.F., Ottesen Hansen, N.E., Stevanato, F., 2016. Summary of the impact on structural reliability of the findings of the Tyra field extreme wave study 2013-15, in: The 3rd Offshore Structural Reliability Conference OSRC2016, pp. 120–131.

List of Figures

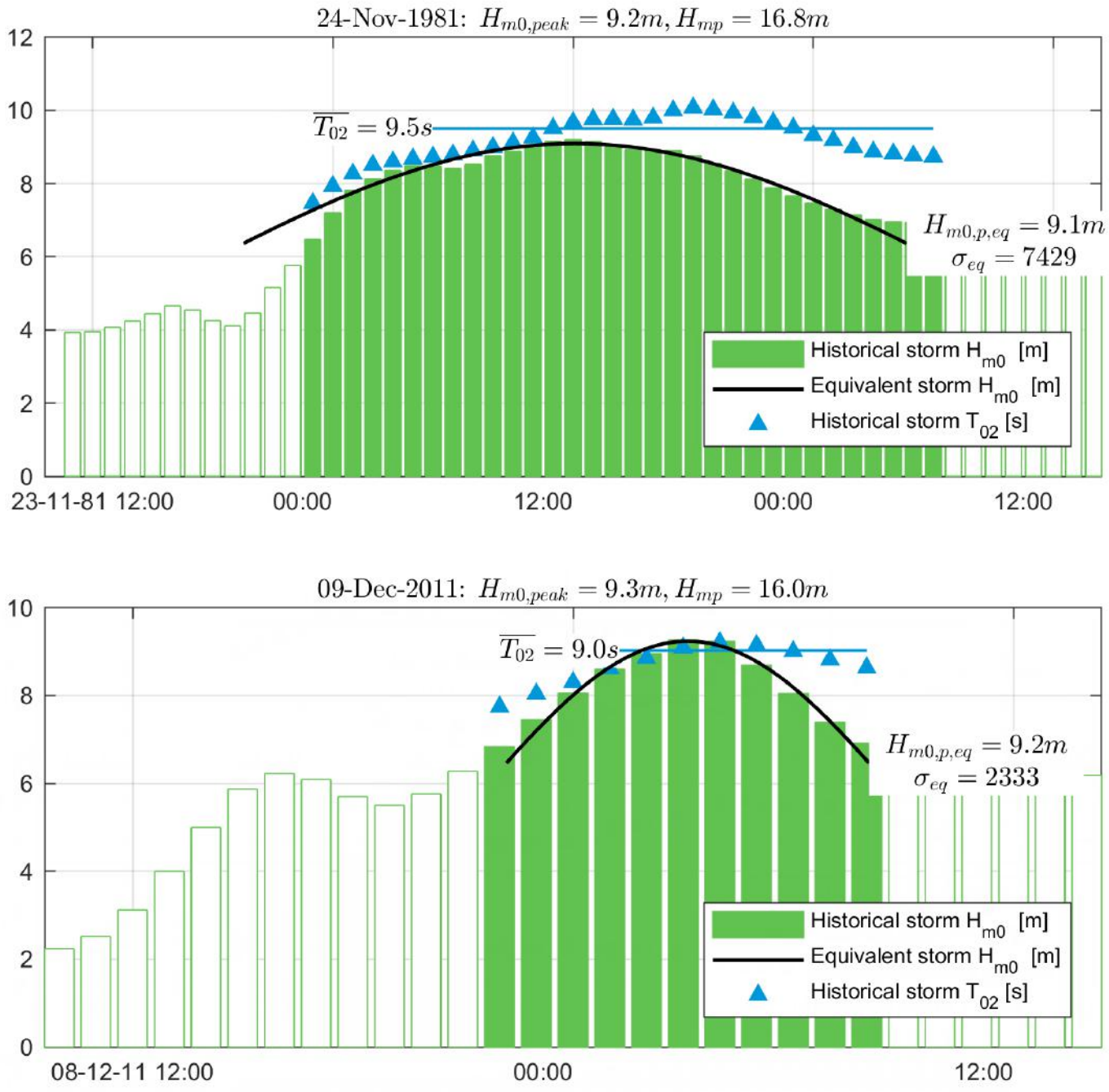

Figure 1: Two examples of time series from storms of hindcast H_{m0} (vertical bars) and T_{02} (blue triangles), both with 1-hour time step resolution. Solid black lines show the Gaussian bell shaped equivalent storm time series and horizontal blue lines show characteristic storm $\overline{T_{02}}$. The confined storms containing only sea states with $H_{m0} \geqslant 0.75 H_{m0,p}$ are marked by the filled green bars.

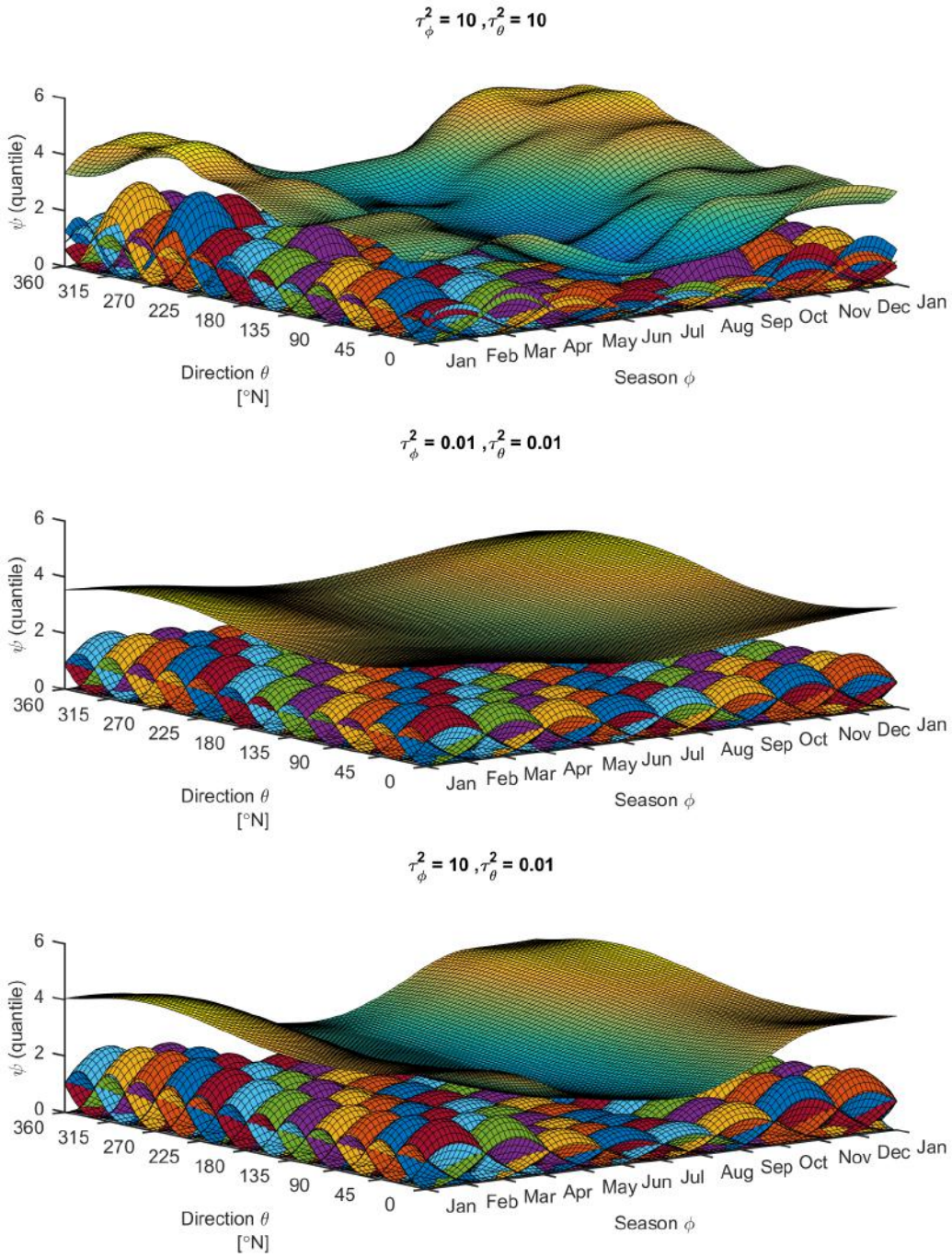

Figure 2: Quantile regression example illustrating the components of tensor-product P-splines in 2 dimensions and the effect of roughness penalty. The coloured surfaces show the individual Tensor-product B-splines each multiplied by its respective β-coefficient. Quadratic B-splines ($q = 2$) and first order penalty have been used.

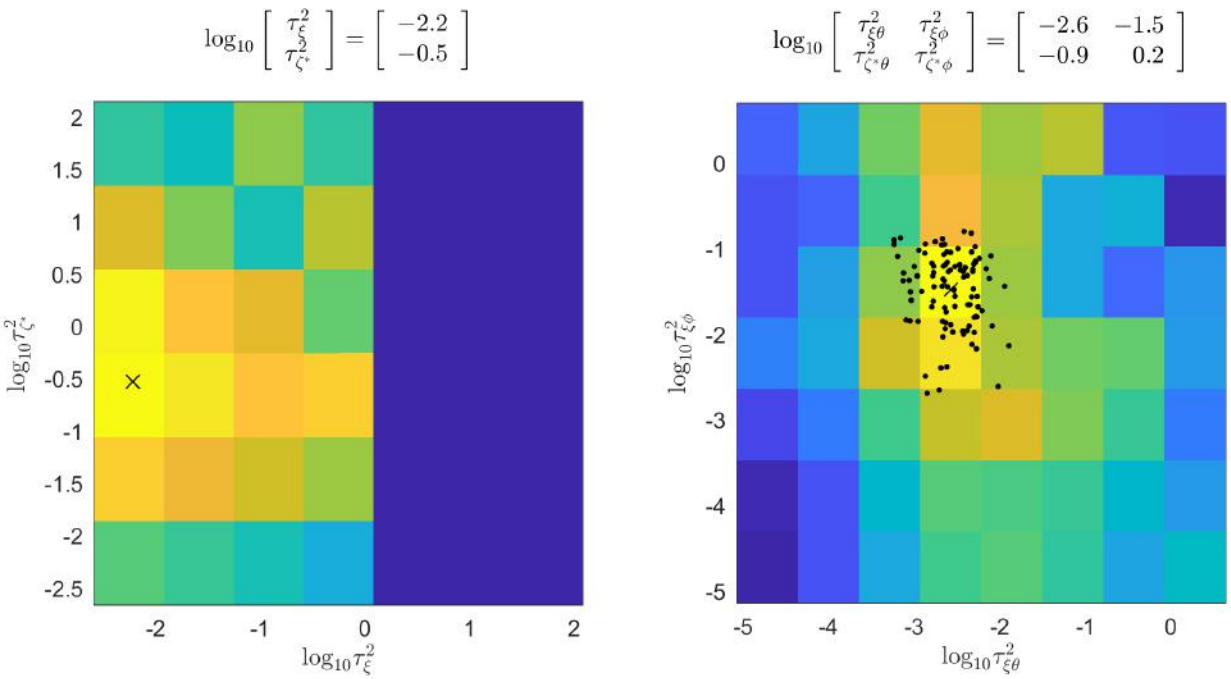

Figure 3: Predictive log-likelihoods for GP high tail estimation for $H_{m0,p,eq}$. Yellow indicates better predictive performance. Left panel shows cross-validation for optimal ratio $\tau^2_\xi/\tau^2_{\zeta^*}$. Right panel shows cross-validation across the two-dimensional $\tau^2_{\xi\theta} \times \tau^2_{\xi\phi}$ domain using the optimal ratio $\log_{10}(\tau^2_\xi) - \log_{10}(\tau^2_{\zeta^*}) = -1.7$. A random sample of combinations of τ^2 from the resulting empirical density is shown as black dots.

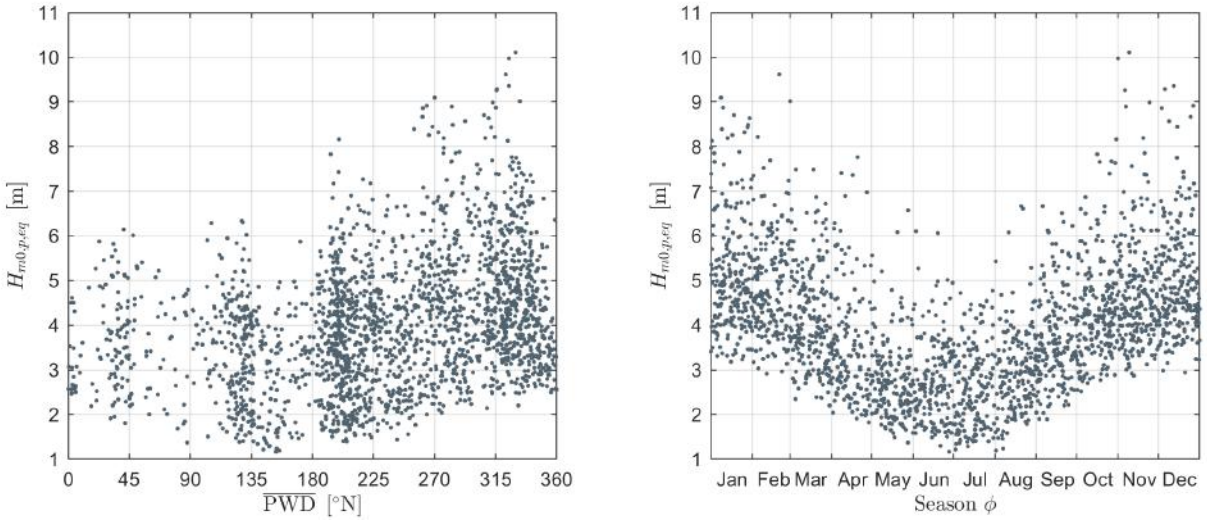

Figure 4: Characteristic variable $H_{m0,p,eq}$ against covariates $\overline{\text{PWD}}$ (left) and season ϕ (right).

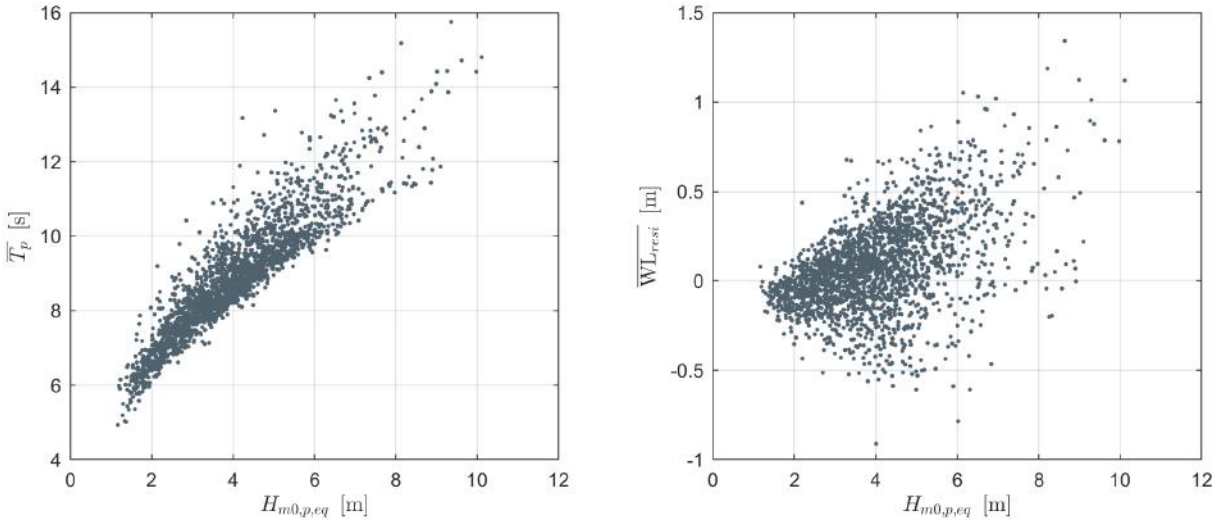

Figure 5: Characteristic variables $\overline{T_p}$ and $\overline{\mathrm{WL}_{resi}}$ on $H_{m0,p,eq}$.

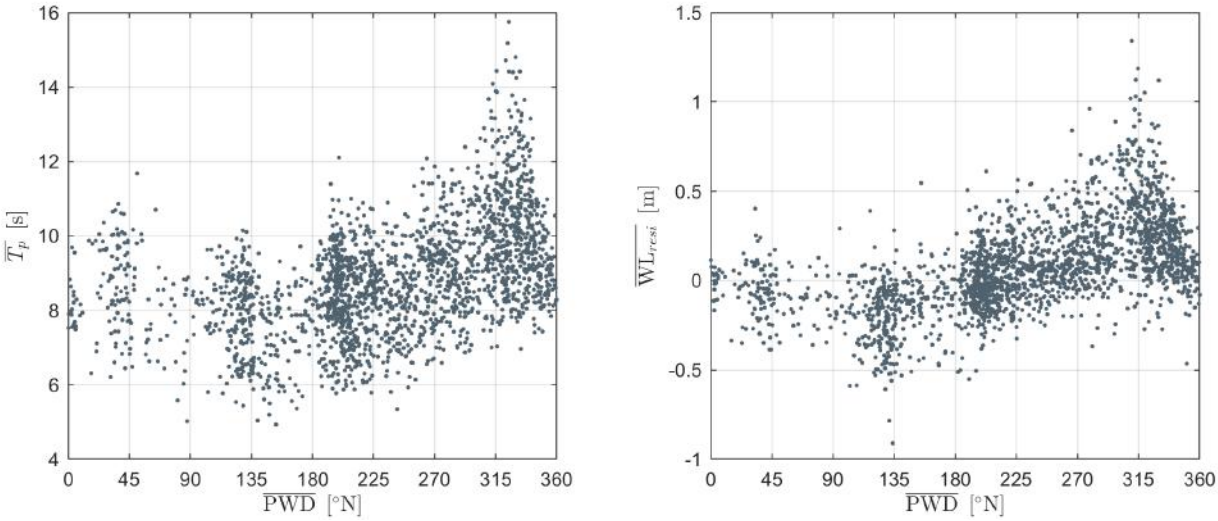

Figure 6: Characteristic variables $\overline{T_p}$ and $\overline{\text{WL}_{resi}}$ on covariate $\overline{\text{PWD}}$.

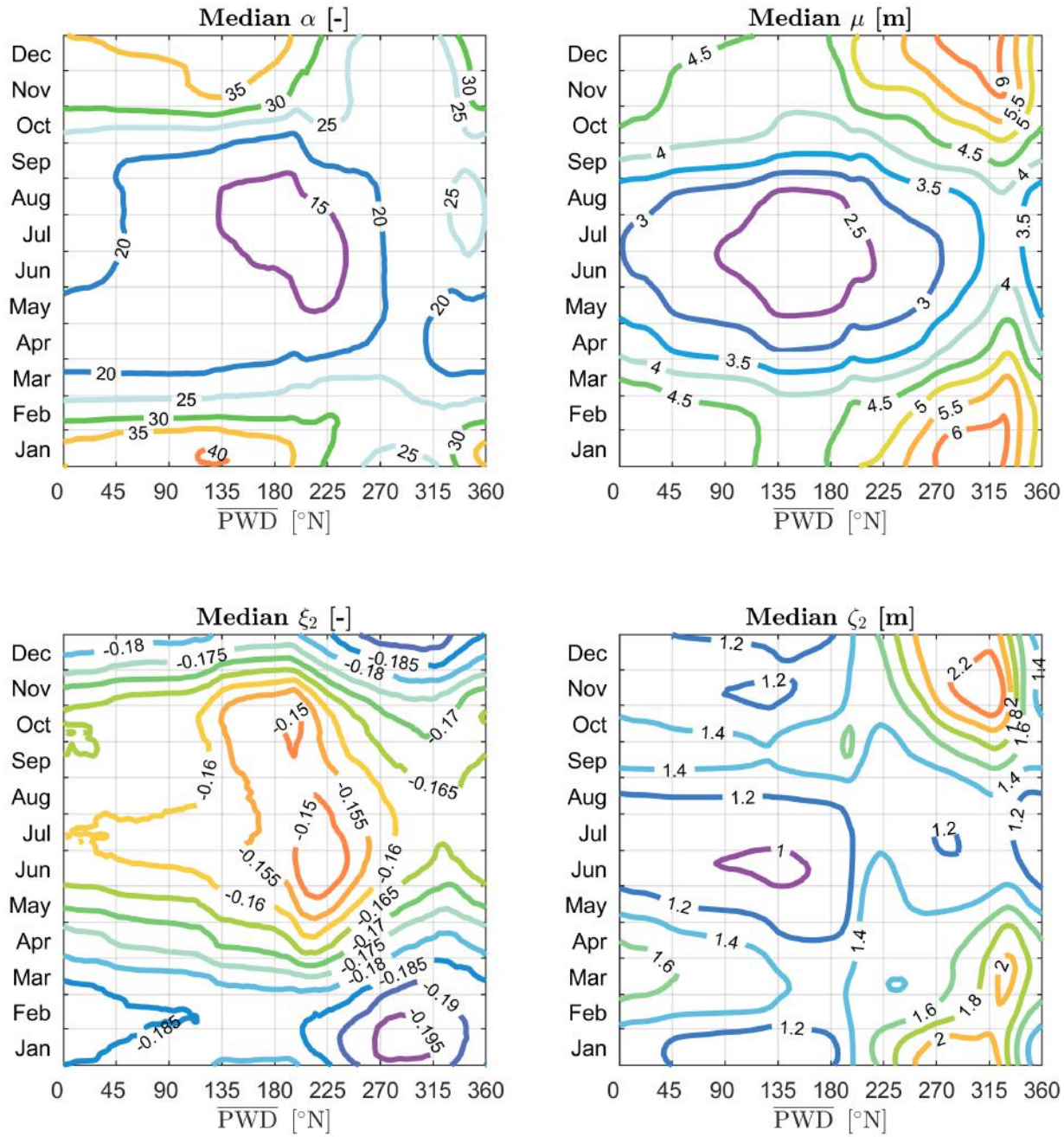

Figure 7: Posterior median parameter estimates for the marginal model of $H_{m0,p,eq}$. Estimates for the upper tail generalised Pareto scale parameter ζ_2 are adjusted to a nominal non-exceedance probability of 0.4 .

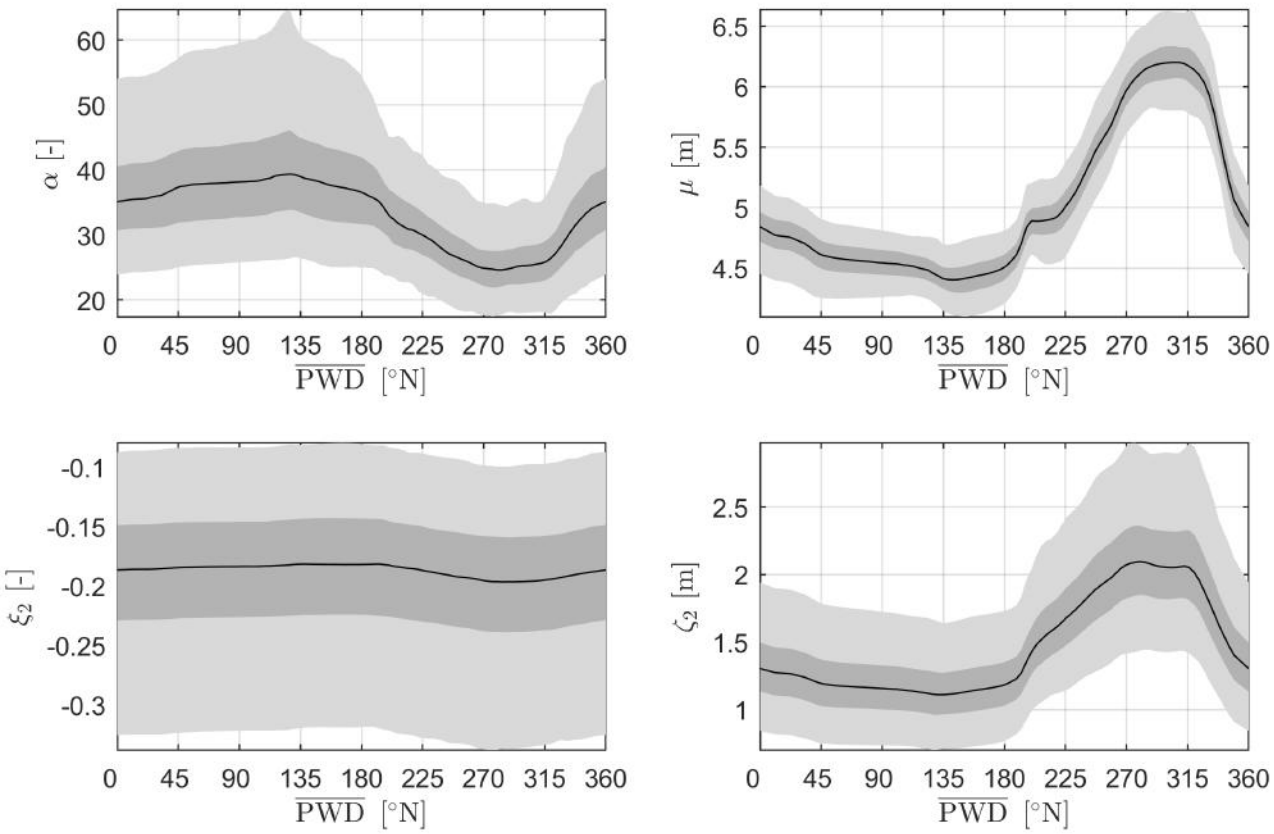

Figure 8: Directional variation of posterior marginal model parameter estimates for $H_{m0,p,eq}$, on 15th January. Shown are median (black line), 50% credible interval (dark grey band) and 95% credible interval (light grey band).

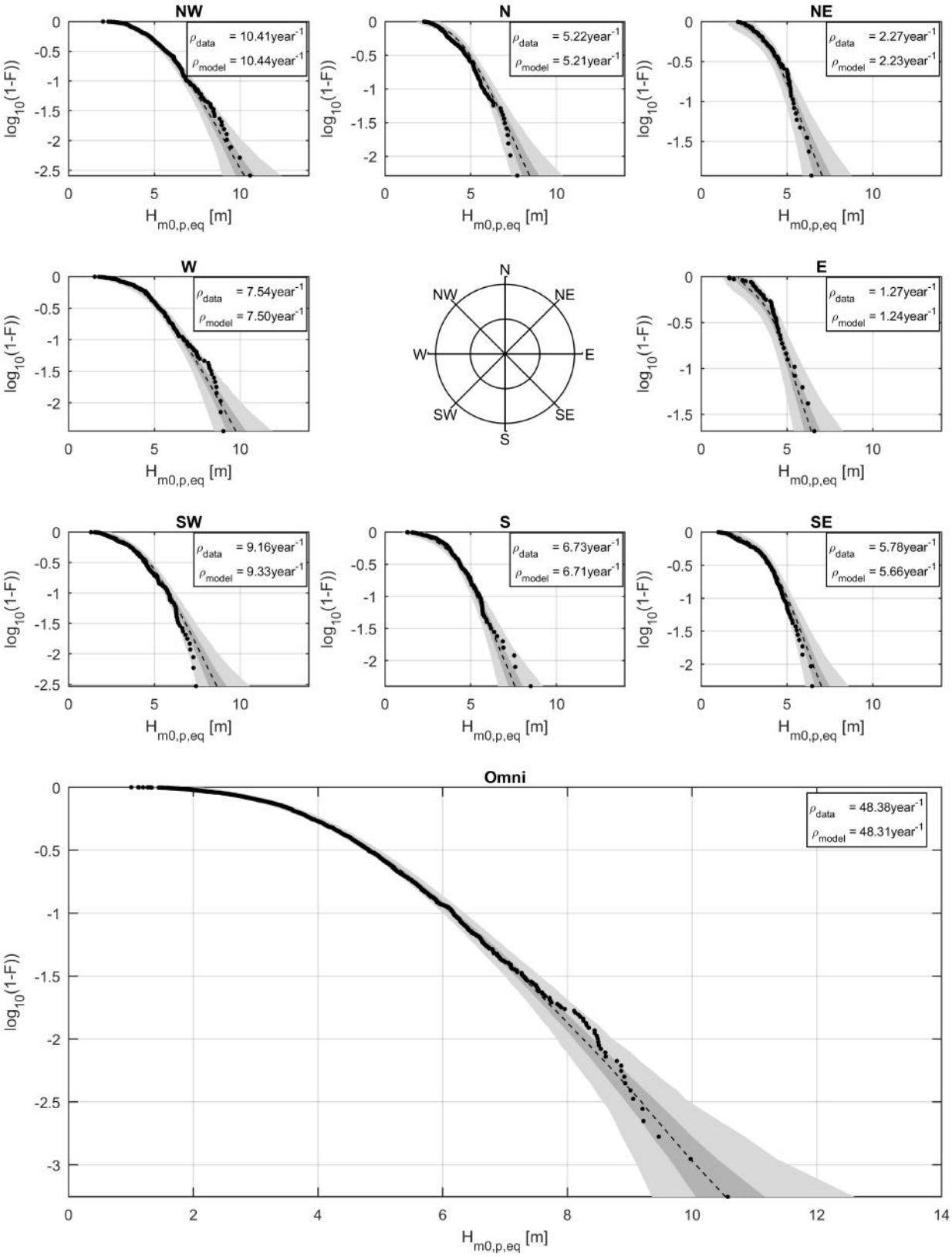

Figure 9: Tails of the distribution of $H_{m0,p,eq}$ for different directional intervals, empirically estimated from the original hindcast sample (black dots), and from simulation under the fitted marginal mode. For each direction, black dashed lines show the simulated median, dark grey band the central 50% credible interval and the light grey band the corresponding 95% credible interval. Empirical and model estimates for the rate of occurrence are given in the legend.

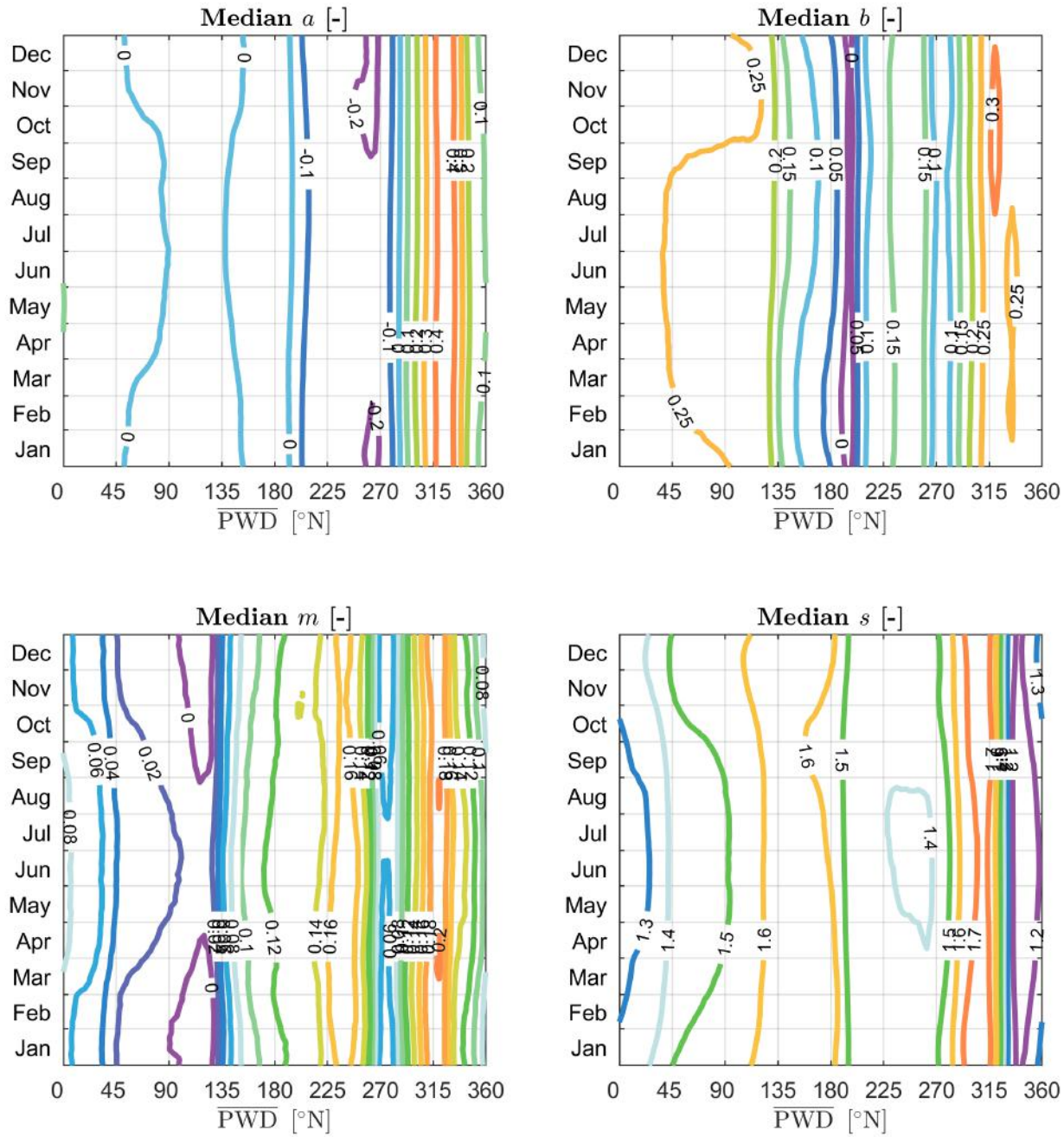

Figure 10: Posterior median parameter estimates from conditional extremes model parameters of $\overline{\text{WL}_{resi}}|H_{m0,p,eq}$.

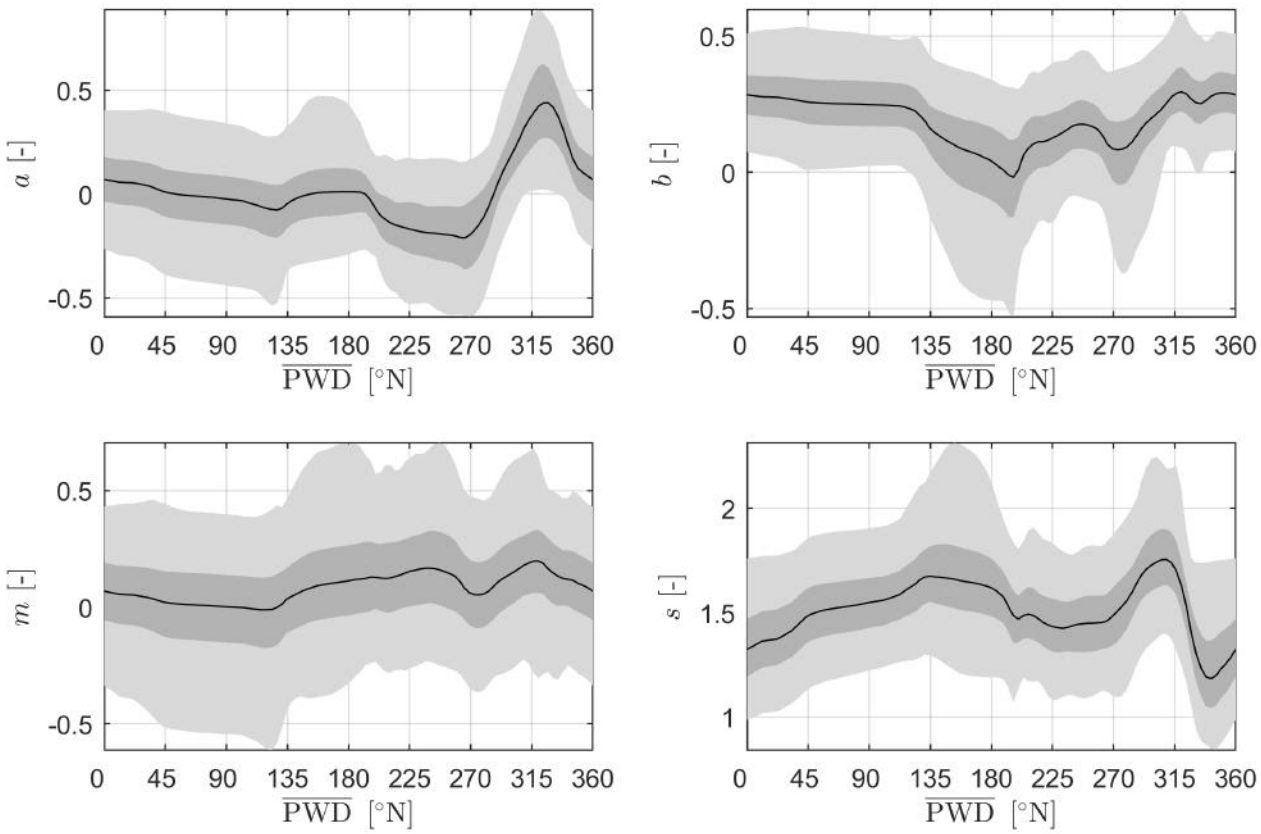

Figure 11: Posterior distributions of conditional extremes model parameters for $\overline{\mathrm{WL}_{resi}}|H_{m0,p,eq}$ on 15$^{\mathrm{th}}$ January as a function of $\overline{\mathrm{PWD}}$. Black line show median, dark grey patch the 50% band and the light grey patch the 95% confidence band.

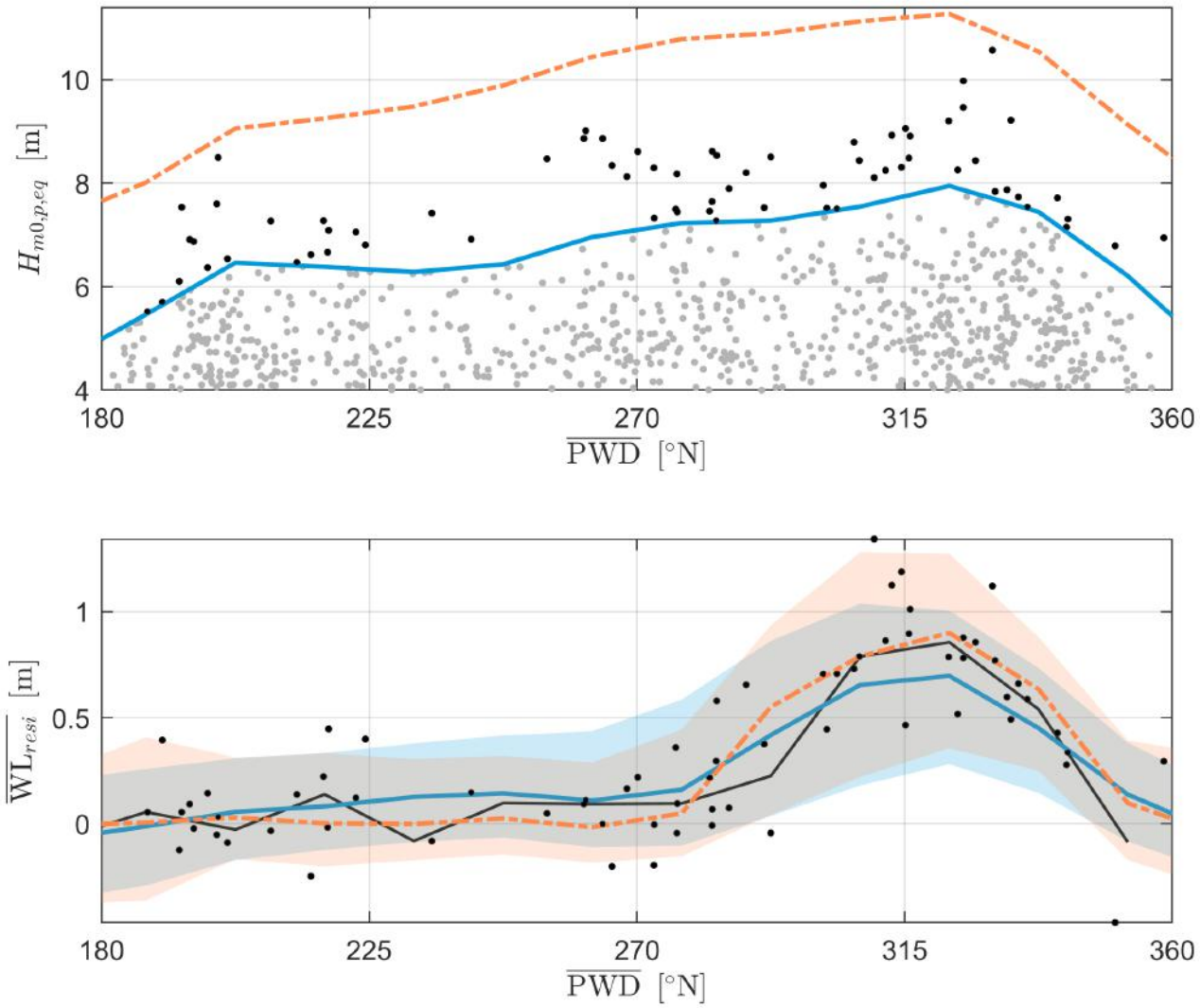

Figure 12: Top panel shows $H_{m0,p,eq}$ on $\overline{\text{PWD}}$. Coloured lines, estimated under the fitted model, have constant annual exceedance probabilities corresponding to return periods of 1 year (blue) and 100 years (orange). Black and grey dots show the original hindcast above and below the blue level respectively. Lower panel shows $\overline{\text{WL}_{resi}}$ on $\overline{\text{PWD}}$ for events with $H_{m0,p,eq}$ above the level marked by the blue line in the top panel. Black dots show the corresponding hindcast data; the solid black line is their directional median. Blue and orange lines mark the median of $\overline{\text{WL}_{resi}}$ associated with exceedances of $H_{m0,p,eq}$ above the blue and orange lines in the top panel; the shaded areas mark intervals bounded by the 10% and 90% quantiles of the corresponding conditional distributions.

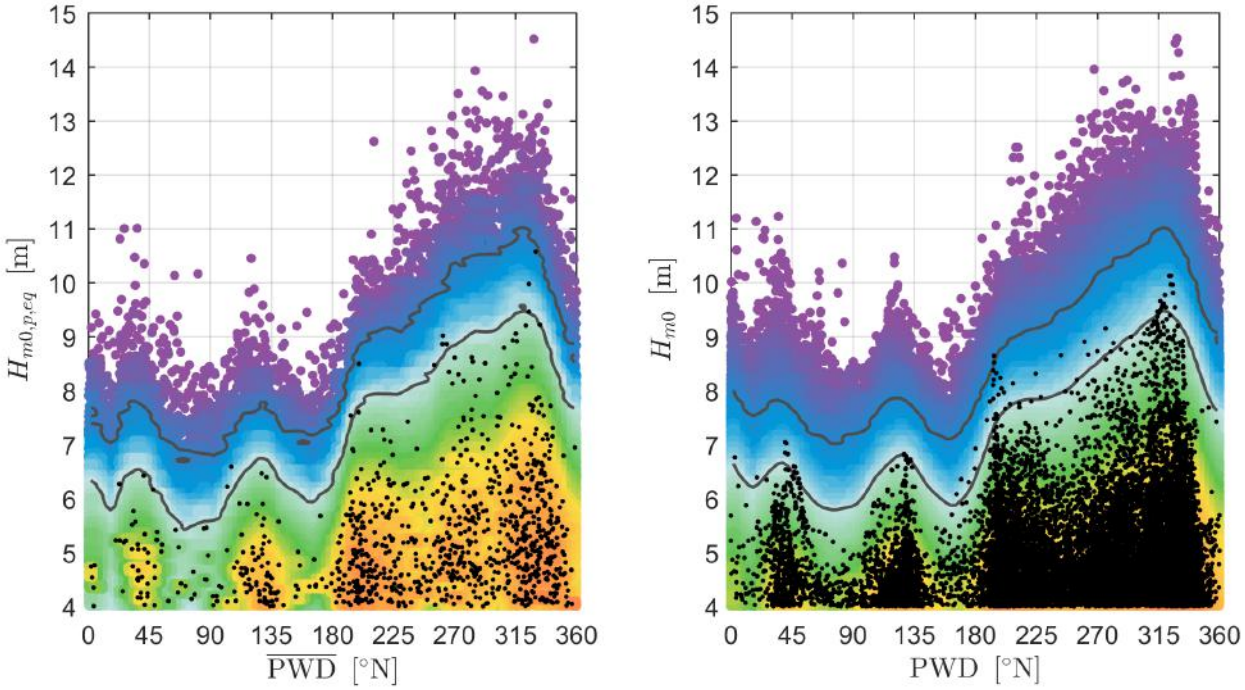

Figure 13: Directional distribution of significant wave height. Scatter plots of 50,000 years of simulated data (coloured round markers) compared to hindcast data (black dots); "warmer" colours indicate higher rate of occurrence of simulated events. Left: Characteristic storm $H_{m0,p,eq}$ vs. $\overline{\text{PWD}}$. Right: Hourly values of H_{m0} vs. PWD. Solid lines represent directional density contours for 10- and 100-year marginal extreme values. Black disks indicate the original hindcast sample.

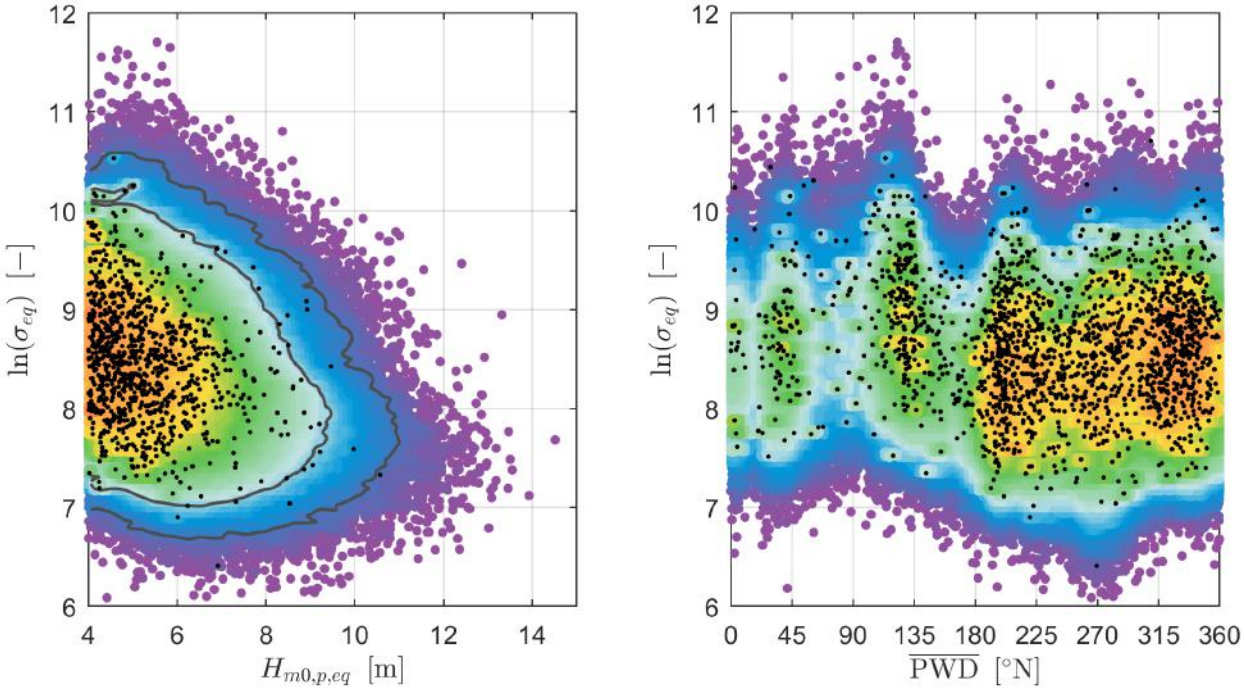

Figure 14: Directional distribution of storm length. Left: $\ln(\sigma_{eq})$ on $H_{m0,p,eq}$ per storm. Right: $\ln(\sigma_{eq})$ on $\overline{\text{PWD}}$ per sea state. For other details, see Figure 13.

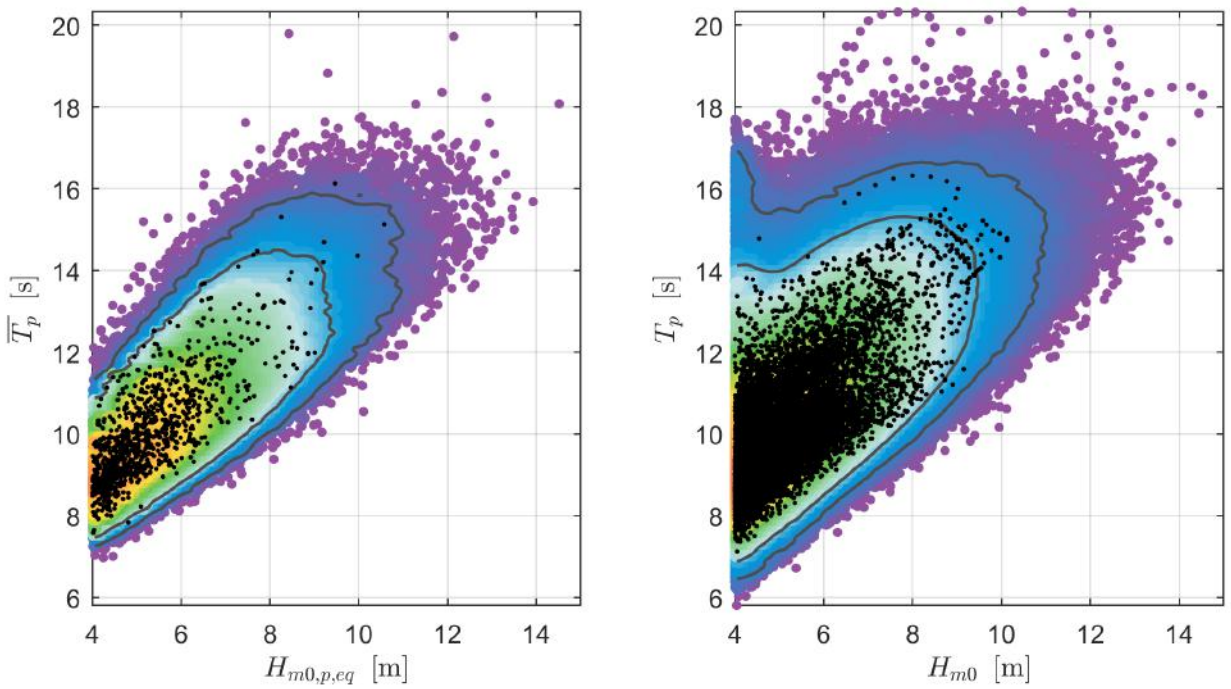

Figure 15: Joint distribution of spectral peak period and significant wave height. Left: Characteristic $\overline{T_p}$ on $H_{m0,p,eq}$ per storm. Right: T_p on H_{m0} per sea state. For other details, see Figure 13.

List of Tables

Description	**Unit**	**Characteristic variable, X**	**Sea state variable, $\tilde{X}$**
Significant wave height	[m]	$H_{m0,p,eq}$	H_{m0}
Storm duration parameter	[−]	σ_{eq}	-
Peak wave period	[s]	$\overline{T_p}$	T_P
Second moment wave period	[s]	$\overline{T_{02}}$	T_{02}
Directional spread at spectral peak	[°]	$\overline{\sigma_{\theta,p}}$	$\sigma_{\theta,p}$
Residual water level	[m]	$\overline{\mathrm{WL}_{resi}}$	WL_{resi}
Residual current speed	[m/s]	$\overline{\mathrm{CS}_{resi}}$	CS_{resi}
Mean wind speed	[m/s]	$\overline{\mathrm{WS}}$	WS
Air density	[kg/m^3]	$\overline{\rho_{air}}$	ρ_{air}
Peak wave direction†	[°N]	$\overline{\mathrm{PWD}}$	PWD
Residual current direction†	[°N]	-	CD_{resi}
Mean wind direction†	[°N]	-	WD

Table 1: Overview of environmental variables. The last three listed variables marked with † are covariates in extreme value models.

Description	**Symbol**	**Type**
Rate of occurrence	ρ	Tensor-product B-spline
Γ shape	α	Tensor-product B-spline
Γ mean	μ	Tensor-product B-spline
GP low tail threshold	ψ_1	Scalar
GP low tail threshold probability	κ_1	Scalar hyper-parameter
GP low tail shape	ξ_1	Tensor-product B-spline
GP low tail scale	ζ_1	Tensor-product B-spline
GP high tail threshold	ψ_2	Scalar
GP high tail threshold probability	κ_2	Scalar hyper-parameter
GP high tail shape	ξ_2	Tensor-product B-spline
GP high tail scale	ζ_2	Tensor-product B-spline
CE threshold	ν	Scalar
CE threshold probability	λ	Scalar hyper-parameter
CE a parameter	a	Tensor-product B-spline
CE b parameter	b	Tensor-product B-spline
CE mean	m	Tensor-product B-spline
CE standard deviation	s	Tensor-product B-spline
Roughness coefficient	τ^2	Scalar hyper-parameter

Table 2: Overview of model parameters. The joint distribution of all parameters (other than hyperparameters) is estimated by MCMC. Hyperparameters for κ are set by hand. Hyperparameters for τ^2 are set to provide reasonable ranges of penalised B-spline flexibility using a pre-analysis described in Section 3