Accepted by Ocean Engineering, June 2017

Efficient estimation of return value distributions from non-stationary marginal extreme value models using Bayesian inference

Emma Ross[a], David Randell[b], Kevin Ewans[c], Graham Feld[d], Philip Jonathan[a,*]

[a]*Shell Projects & Technology, London SE1 7NA, United Kingdom.*
[b]*Shell Projects & Technology, 1031 HW Amsterdam, The Netherlands.*
[c]*Metocean Research Limited, New Plymouth 4310, New Zealand.*
[d]*Shell Projects & Technology, Aberdeen AB12 3FY, United Kingdom.*

Abstract

Extreme values of an environmental response can be estimated by fitting the generalised Pareto distribution to a sample of exceedances of a high threshold. In oceanographic applications to responses such as ocean storm severity, threshold and model parameters are typically functions of physical covariates. A fundamental difficulty is selection or estimation of an appropriate threshold or interval of thresholds, of particular concern since inferences for return values vary with threshold choice. Historical studies suggest that evidence for threshold selection is weak in typical samples.

Hence, following Randell et al. (2016), a piecewise gamma-generalised Pareto model for a sample of storm peak significant wave height, non-stationary with respect to storm directional and seasonal covariates, is estimated here using Bayesian inference. Quantile regression (for a fixed quantile threshold probability) is used to partition the sample prior to independent gamma (body) and generalised Pareto (tail) estimation. An ensemble of independent models, each member of which corresponds to a choice of quantile probability from a wide interval of quantile threshold probabilities, is estimated. Diagnostic tools are then used to select an interval of quantile threshold probabilities corresponding to reasonable model performance, for subsequent inference of extreme quantiles incorporating threshold uncertainty.

The estimated posterior predictive return value distribution (for a long return period of the order of 10,000 years) is a particularly useful diagnostic tool for threshold selection, since this return value is a key deliverable in metocean design. Estimating the distribution using Monte Carlo simulation becomes computationally demanding as return period increases. We present an alternative numerical integration scheme, the computation time for which is effectively independent of return period, dramatically improving computational efficiency for longer return periods.

The methodology is illustrated in application to storm peak and sea state significant wave height at a South China Sea location, subject to monsoon conditions, showing directional and seasonal variability.

Keywords: extreme, significant wave height, non-stationary, splines, Bayesian inference, return value, numerical integration.

*Corresponding author. Email: philip.jonathan@shell.com

1. Introduction

1.1. Threshold selection

Threshold selection in practical application of extreme value analysis is almost always problematic. Even in the absence of covariate effects, it is rarely clear where the threshold should be set, or indeed if setting a single threshold is even desirable. A review of threshold selection for extreme value analysis is given by Scarrott and MacDonald (2012). The generalised Pareto (GP) model for peaks-over-threshold is motivated by asymptotic arguments: the threshold needs to be set high enough so that a generalised Pareto model fits threshold exceedances reasonably, to reduce bias. Yet the threshold should be set low enough that there are sufficient exceedances to estimate generalised Pareto model parameters, to reduce variance: a typical bias-variance trade-off. Graphical techniques such as the mean excess plot (Ghosh and Resnick 2010) can be of some use in aiding a sensible threshold choice, as can inspection of the stability of generalised Pareto shape parameter estimate or other key inferences, such as estimates for return values or other structure variables as a function of threshold. Here, a structure variable is any variable defined in terms of one or more responses modelled by extreme value analysis. For example, the load on an offshore structure can be considered a structure variable, defined in terms of extreme value responses including significant wave height, current speed, wind speed etc. Graphical techniques are rarely conclusive however. Some authors, including Sanchez-Archilla et al. (2008), Thompson et al. (2009), Northrop and Coleman (2014) and Wadsworth (2016), have proposed procedures for estimating good thresholds, but these all contain subjective elements. In the presence of covariate effects, threshold selection is even more problematic. Typically, the threshold is set as a local (covariate-dependent) quantile of the response, and the problem of threshold selection transformed into one of specifying the appropriate threshold quantile level for the covariates used.

1.2. Threshold estimation

One approach to overcoming the need to specify an extreme value threshold ψ before extreme value inference is to make ψ a model parameter to be estimated. To achieve this, the extreme value model must be extended so that it describes part or all of the body of the sample, as well as extreme value threshold exceedances, such as in Tancredi et al. (2006), Wadsworth et al. (2010), MacDonald et al. (2011) and Randell et al. (2016). Yet in such a model, incorporating extreme value inference and estimation of ψ, it is not clear whether is is desirable that estimation of ψ be influenced by model fit to threshold non-exceedances: adequate fit of the generalised Pareto tail should take priority, since tail estimation is our primary concern. From this perspective, models for which ψ is pre-specified with no regard to threshold non-exceedances would seem advantageous.

1.3. Incorporating threshold uncertainty

Values $q(\psi)$ for return values (and other structure variables of interest) can be found corresponding to any threshold ψ. A final "preferred" return value q_1 might then be selected corresponding a single "best" threshold $\widehat{\psi}$, found either from inspection of diagnostics (or model estimation), such that

$$q_1 \triangleq q\left(\widehat{\psi}\right).$$

Alternatively, we can provide final return values q_2 by integrating over suitable values of ψ specified by some density $f(\psi)$, itself inferred from inspection of diagnostics (or directly estimated). Then

$$q_2 \triangleq \int_{\psi} q(\psi) f(\psi) d\psi.$$

The advantage of using q_2 over q_1 is that uncertainty in ψ is propagated through to return values. When pre-specification or estimation of $\widehat{\psi}$ is problematic, it seems reasonable to prefer q_2 over q_1. When pre-specification or estimation of $\widehat{\psi}$ is straightforward, we expect q_2 and q_1 to be similar since $f(\psi)$ provides probability mass only around $\widehat{\psi}$. Northrop et al. (2016) discuss cross-validatory threshold selection, including a method for incorporating threshold uncertainty.

In this work, we choose to employ estimates of the form q_2 for return values, but note that there is still considerably subjectivity in the choice of $f(\psi)$ to be used. We seek to inform this choice by consideration of various model diagnostics. We thereby hopefully introduce some rationality, but make no claim to have removed all subjectivity.

1.4. Non-stationary Bayesian extreme value modelling

In Randell et al. (2016), a model for the distribution of independent observations of peaks-over-threshold of a response such as storm peak significant wave height given multidimensional covariates is developed, incorporating the generalised Pareto distribution for exceedances of some estimated non-stationary threshold. A truncated Weibull distribution characterises values below the threshold. The model is used primarily to estimate distributions of return values corresponding to long return periods, for use in the design and reliability assessment of marine and coastal structures. The methodology is intended to be easy to use, and computationally efficient for full-scale oceanographic applications for sample sizes from 10^3 to 10^7 with at least a two-dimensional (e.g. directional-seasonal) covariate domain. Practical applications' experience using the model suggests that (a) there is often little or no evidence in the sample to inform threshold selection, that (b) the model form itself is perhaps too restrictive to facilitate easy threshold estimation using Bayesian inference, and that (c) the gamma distribution provides an equally flexible form for the distribution of non-exceedances, together with more stable inference when likelihood gradients are exploited (Section 3.5).

We have concluded, for typical applications to ocean storm severity, that there is little value in seeking to estimate extreme value threshold directly. Instead, making inferences using an ensemble of piecewise gamma-generalised Pareto models, each of which corresponds to a specific threshold choice, is less problematic inferentially and more useful in practice. Ensemble members (and hence thresholds) are selected to give reasonable performance, as judged by inspection of relevant model diagnostics. Specifically, ensemble threshold choices typically correspond to an interval of threshold probabilities for a non-stationary quantile regression threshold, as in Randell et al. (2015). The model is described in Section 3.

1.5. Return value estimation

Return value estimates are the key deliverable of metocean design. In the presence of covariates and multiple extreme value thresholds, closed form expressions for return values are not available. As a result, Monte Carlo simulation is typically used. This involves simulating under the estimated model, generating thousands of realisations of sets of extreme values corresponding to the return period of interest. Monte Carlo simulation also provides an intuitive framework for estimating

return value distributions for "dissipated" or sea state significant wave height from an extreme value model for storm peak events. The computational burden increases approximately linearly with return period. Parallel computation is useful, but nevertheless return value estimation by Monte Carlo simulation remains computationally intensive, accounting for the vast majority of computing resource required for a typical study (Section 4).

We have developed numerical integration algorithms to replace all Monte Carlo simulations previously performed, resulting in a huge reduction in time required for analysis. The approach is described in Section 4.

1.6. Outline of article

The objective of this article is to improve the applicability of the methodology presented in Randell et al. (2016), in two major respects. Firstly, we simplify the extreme value model by performing prior non-stationary quantile regressions to estimate possible extreme value thresholds, and use each of these thresholds in turn to partition the sample into "body" and "tail". Independent gamma and generalised Pareto models are next estimated for body and tail respectively using Bayesian inference. Finally, an ensemble of models (for different plausible threshold choices) is adopted for return value inference. Secondly, we replace return value inference using computationally-intensive Monte Carlo simulation by a computationally-efficient numerical integration scheme, which amongst other things, improves the accessibility of model diagnostics based on return value distributions.

We motivate the work in Section 2 by considering estimation of return values for ocean storm severity at a location in the South China Sea, dependent on storm direction and season, corresponding to long return periods. In Section 3 we summarise the non-stationary piecewise gamma-generalised Pareto model. Given covariates, the model consists of a truncated gamma distribution to characterise storm peak events at or below an extreme value threshold (i.e. in the body of the sample), and a generalised Pareto distribution for threshold exceedances (in the tail). Section 3 also outlines the spline parameterisation adopted for covariate effects, prior model parameter specification and the inference scheme. Section 4 discusses the numerical integration procedure introduced to estimate return value distributions efficiently. In Section 5, we apply the model to estimation of return values for storm peak and sea state significant wave height in the South China Sea. Findings are discussed and conclusions drawn in Section 6. We note that significant wave height is reported in units of metres throughout.

2. Motivating application

The application sample is composed of time series (from the SEAFINE 2012 hindcast) for sea state significant wave height (H_S, measured in metres), (dominant) wave direction θ and season ϕ (defined as day of the year, for a notional year consisting of 360 days), for consecutive three hour sea states over the period August 1956 to July 2012, at a location offshore Borneo at a water depth of approximately 30m. With direction *to which* waves travel expressed in degrees clockwise with respect to north, Figure 1 shows plots of storm peak H_S and (sea state) H_S versus direction θ and season ϕ.

[Figure 1 about here.]

The climate is monsoonal, with Southwest Monsoon between June and September and Northeast Monsoon between December and March. At this location, due to atmospheric circulation and

topographical effects, the Southwest Monsoon generates increased H_S with waves propagating in a north-northwesterly direction, and the Northeast Monsoon generates severest sea states with waves propagating towards the south-southwest. The largest value of H_S in the sample is approximately 3.1m. The absence of severe events with directions corresponding approximately to $[225°, 315°]$ coincides with the land shadow of Borneo. A total of 4631 storm events are isolated from these time-series using the procedure described in Ewans and Jonathan (2008). Briefly, contiguous intervals of H_S above a low peak-picking threshold are identified, each interval now assumed to correspond to a storm event. The peak-picking threshold corresponds to a directional-seasonal quantile of H_S with specified non-exceedance probability, estimated using quantile regression. The maximum of H_S during the storm interval is taken as the storm peak significant wave height. The values of directional and seasonal covariates at the time of the storm peak H_S are referred to as storm peak values of those variables. Each storm event is thus characterised in terms of the storm peak H_S, and corresponding storm peak direction and season.

The left-hand side of Figure 2 shows the mean annual count rate for occurrences of storm peak events within each directional-seasonal bin of the covariate domain, as defined in Section 3. The covariate domain is relatively sparsely occupied, with the Northwest Monsoon accounting for the highest rate of occurrence of events. The right-hand side of Figure 2 shows the maximum value of storm peak H_S per directional-seasonal bin. The largest events correspond to the Southwest Monsoon.

[Figure 2 about here.]

For each storm event we also isolate a storm trajectory, namely the time-series $\{H_S(t), \theta(t)\}$ for values of t corresponding to the storm time interval only. The maximum of H_S on the storm trajectory is the storm peak value. Knowledge of storm trajectories is essential for estimation of return values for sea state H_S in Section 4.

[Figure 3 about here.]

3. Models

3.1. Model for size of occurrence

The magnitude y of peaks-over-threshold events is assumed to follow a non-stationary two-part distribution. Asymptotic theory suggests that an extreme value model is suitable to describe the largest values in the sample. Exceedances of some covariate-dependent threshold ψ (> 0) are therefore assumed to follow a generalised Pareto distribution with shape ξ $(\in \mathbb{R})$ and scale σ (> 0), conditional on exceeding the threshold. A truncated gamma distribution with shape α (> 0), (inverse) scale ζ (> 0) and threshold ψ provides a flexible unimodal density for threshold non-exceedances. The supports of the truncated gamma and generalised Pareto distributions are respectively $[0, \psi]$ and (ψ, y^+) (where $y^+ = \psi - \frac{\sigma}{\xi}$ for $\xi < 0$ and $= \infty$ otherwise). All of $\alpha, \zeta, \xi, \sigma$ and ψ in principle are smoothly-varying functions of direction θ and season ϕ. The corresponding probability density function is

$$f(y|\alpha, \zeta, \sigma, \xi, \psi, \tau) \quad = \quad \begin{cases} \tau \times f_{TG}(y|\alpha, \zeta, \psi) & \text{for } y \in [0, \psi] \\ (1-\tau) \times f_{GP}(y|\sigma, \xi, \psi) & \text{for } y \in (\psi, y^+) \end{cases}$$

where f_{TG} and f_{GP} are the densities of the truncated gamma and generalised Pareto distributions. Parameter τ corresponds to the extreme value threshold non-exceedance probability assumed stationary with respect to covariates, for the transition from truncated gamma to generalised Pareto; the factors τ and $(1-\tau)$ weight the relative contributions of the two parts. Below extreme value threshold ψ, y follows a truncated gamma distribution with density

$$f_{TG}\left(y|\alpha,\zeta,\psi\right) \;=\; \frac{f_G\left(y|\alpha,\zeta\right)}{F_G\left(\psi|\alpha,\zeta\right)} \text{ for } y \in [0,\psi],$$

where $f_G\left(y|\alpha,\zeta\right) = \zeta^{\alpha} y^{\alpha-1} \exp\left(-\zeta y\right)/\Gamma(\alpha)$, $F_G\left(\psi|\alpha,\zeta\right) = \gamma\left(\alpha,\zeta\psi\right)/\Gamma(\alpha)$, γ is the incomplete gamma function, and F_G is the cumulative distribution function of the gamma distribution. Above extreme value threshold ψ, y follows a generalised Pareto distribution with density

$$f_{GP}(y|\xi,\sigma) = \frac{1}{\sigma}\left(1+\frac{\xi}{\sigma}\left(y-\psi\right)\right)^{-1/\xi-1} \text{ for } y \in (\psi, y^{+}) \; .$$

In practice, we choose to carry out computations with $\nu = \sigma(1+\xi)$ in place of σ. This choice is motivated by fact that, with $\xi > -0.5$, maximum likelihood estimates for ν and ξ are asymptotically independent, simplifying their joint estimation compared with joint estimation of σ and ξ. Nevertheless, (derived) estimates for σ are reported and visualised below for ease of interpretation. In greatest generality, for a sample $\{y_i\}_{i=1}^n$ of n storm peak events, with corresponding directions $\{\theta_i\}_{i=1}^n$ and seasons $\{\phi_i\}_{i=1}^n$, sample likelihood $\mathcal{L}(\alpha,\zeta,\xi,\sigma,\psi,\tau|\{y_i,\theta_i,\phi_i\}_{i=1}^n)$ is

$$\mathcal{L} = \left[\tau^{n_B} \prod_{i:y_i \leqslant \psi(\theta_i,\phi_i)} f_{TG}(y_i|\alpha,\zeta,\psi)\right] \times \left[(1-\tau)^{(n-n_B)} \prod_{i:y_i > \psi(\theta_i,\phi_i)} f_{GP}(y_i|\xi,\sigma,\psi)\right]$$

for (spline-parameterised) functions $\alpha,\zeta,\xi,\sigma,\psi$ of θ and ϕ, and stationary τ to be estimated, where $n_B = \sum_{i:y_i \leqslant \psi(\theta_i,\phi_i)} 1$ is the number of events at or below threshold ψ. Note that, in the description below, statements along the lines of "estimation of η" for model parameter $\eta \in \{\alpha,\zeta,\nu,\xi\}$ with a spline representation, refer to "estimation of the spline coefficients corresponding to η on the index set of covariates".

As noted in Section 1, estimation of extreme value threshold ψ is typically problematic in practical application. For this reason, we simplify the model somewhat by specifying a fixed prior distribution for τ. We typically assume τ to be uniformly-distributed on an interval of non-exceedance probabilities, the end points of which are informed by inspection of various model diagnostic plots. Given τ, threshold ψ, a function of covariates θ and ϕ, is estimated from the sample using (non-stationary) quantile regression (e.g. Koenker 2005). We seek a smooth function ψ of covariates corresponding to a non-exceedance probability τ of the response for any combination of θ and ϕ. We estimate ψ by minimising the quantile regression lack of fit criterion

$$\ell_{\psi} = \{\tau \sum_{i,r_i \geq 0}^{n} |r_i| + (1-\tau) \sum_{i,r_i<0}^{n} |r_i|\}$$

for residuals $r_i = y_i - \psi(\theta_i,\phi_i;\tau)$, penalised with respect to the spline roughness R of ψ with respect to covariates (Section 3.3) for different plausible values of penalty coefficients. We then specify the optimal penalty coefficient as that which yields best predictive performance in 10-fold cross-validation, and use the corresponding ψ to partition the sample prior to gamma and generalised Pareto fitting.

Threshold ψ is then used to partition the sample into sub-samples for threshold non-exceedances and exceedances respectively, to which truncated gamma and generalised Pareto models are fitted independently using Bayesian inference. Markov chain Monte Carlo with the mMALA algorithm (manifold Metropolis Adjusted Langevin Algorithm, Girolami and Calderhead 2011) is used to estimate α and ζ for the sub-sample below threshold ψ, and independently to estimate ξ and ν (and hence σ) above ψ. Return value distributions are then estimated by numerical integration.

The model is advantageous in that inferences above threshold ψ are not influenced at all by inferences made at or below ψ. Specifically, the sample likelihood for fitting the truncated gamma - generalised Pareto for different non-exceedance probabilities τ is used to guide the choice of a sensible ensemble of models, corresponding to an interval of values for τ yielding good whole-sample fit. This choice is informed by inspection of diagnostic plots like those discussed in the application in Section 5. Once this choice is made, the quality of fit of the whole-sample model, i.e. the sample likelihood, plays no further direct role. Instead, models based on all thresholds in the chosen interval are treated as equally plausible, and carry equal weight for estimation of return values. Only the generalised Pareto tail for each ensemble member is then used to estimate return values. The model is also particularly suitable for parallel computation since Markov chain Monte Carlo estimation for different draws of τ can be performed independently. The disadvantage of the model is that the fixed prior distribution for τ dictates the extreme value thresholds considered appropriate. Prior specification of (e.g. an interval of plausible values for) τ therefore assumes increased importance, and is informed by inspection of model diagnostics evaluated for the widest possible range of values for τ.

3.2. Model for rate of occurrence

The rate of occurrence of all events is estimated using a non-stationary Poisson process with rate ρ varying with θ and ϕ. We approximate the Poisson process likelihood by evaluating it on the set $\{S_j\}_{j=1}^m$ of m directional-seasonal sub-intervals (referred to colloquially as "bins") of area d of covariate domain $\mathcal{D}$ centred on $\{I_j\}_{j=1}^m$ (Section 3.3), assuming that d is small enough that ρ is approximately constant on each bin, following Chavez-Demoulin and Davison (2005) and Randell et al. (2016). Then, for a vector of counts $\boldsymbol{c} = \{c_k\}_{k=1}^m$ of occurrences on the index set, the probability density of counts conditional on the rate of occurrence is:

$$f(\boldsymbol{c}|\rho) = \exp\left(-d\sum_{k=1}^{m}\rho_k\right)\prod_{k=1}^{m}\rho_k^{c_k} \ ,$$

where $\boldsymbol{\rho} = \{\rho_k\}_{k=1}^m$ is the corresponding Poisson count rate to be estimated.

3.3. Spline parameterisation

We assume that model parameters ρ, α, ζ, ξ, ν and ψ vary smoothly with θ and ϕ, expressing each in terms of a basis for B-splines for the domain $\mathcal{D}$ of covariates, where $\mathcal{D} = \mathcal{D}_2 \times \mathcal{D}_1$, and $\mathcal{D}_1 = \mathcal{D}_2 = [0, 360)$ is the (marginal) domain of both θ and ϕ. We define a $m \times p$ basis matrix $\boldsymbol{B}$ for $\mathcal{D}$ using the tensor product $\boldsymbol{B} = \boldsymbol{B}_2 \otimes \boldsymbol{B}_1$ of marginal bases, for an index set $\{I_j\}_{j=1}^m$ of $m = m_1 \times m_2 = 32 \times 24 = 768$ directional-seasonal locations. The numbers of directional and seasonal subintervals m_1, m_2 are specified from physical and engineering considerations. Experience suggests that the wave environment is approximately homogeneous within 11.25° directional intervals and 2-weekly seasonal intervals; this implies that choices of $m_1 = 32, m_2 = 24$ are reasonable. 10 basis parameters are estimated for each marginal basis, so that $p = p_1 \times p_2 = 10 \times 10 = 100$. The values

of any of ρ, α, ζ, ξ, ν and ψ on the 2D index set then take the form $\boldsymbol{B}\boldsymbol{\beta}$ for some $p \times 1$ vector $\boldsymbol{\beta}$ of spline coefficients to be estimated. For any model parameter, spline roughness on $\mathcal{D}$ is $R = \beta' \boldsymbol{P} \beta$ for directional-seasonal penalty matrix $\boldsymbol{P} = \lambda_\theta \boldsymbol{P}_\theta + \lambda_\phi \boldsymbol{P}_\phi$ and penalty coefficients $\boldsymbol{\lambda} = \{\lambda_\theta, \lambda_\phi\}$. $\boldsymbol{P}_\theta$ and $\boldsymbol{P}_\phi$ are marginal directional and seasonal roughness penalty matrices incorporating random roughness perturbations $\boldsymbol{\delta}_\theta$ and $\boldsymbol{\delta}_\phi$ (following Brezger and Lang 2006). For Bayesian inference (Section 3.5), $\boldsymbol{P}$ is used as the prior precision matrix for spline coefficients $\boldsymbol{\beta}$. Generalised linear array methods (GLAM, Currie et al. 2006) greatly reduce the computational burden of spline manipulations. A fuller description of the spline parameterisation is given in Randell et al. (2016).

3.4. Prior specification for Bayesian inference

For each of ρ, α, ζ, ξ and ν, spline roughness penalty coefficients $\boldsymbol{\lambda}$ are assumed to be gamma-distributed $\Gamma(a_\lambda, b_\lambda)$ independently and identically, so that full conditional distributions are available in closed form, enabling Gibbs' sampling, with $a_\lambda = b_\lambda = 10^{-3}$ throughout. Spline coefficients $\boldsymbol{\beta}$ are distributed with density proportional to $|\boldsymbol{P}|^{1/2} \exp(-\frac{1}{2}\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta})$ (Section 3.3); this density is Gaussian-like, but not Gaussian since $\boldsymbol{P}$ is not a valid covariance matrix in general. For each model parameter, the distributional characteristics of roughness perturbations $\boldsymbol{\delta}_\theta$ and $\boldsymbol{\delta}_\phi$ are not inferred, as explained in Randell et al. (2016). Instead, all elements of roughness perturbations are sampled identically and independently from the prior distribution $\Gamma(a_\delta, b_\delta)$ with $a_\delta = b_\delta = \frac{1}{2}$.

3.5. Bayesian inference

Bayesian inference is used to estimate the posterior distributions of spline coefficients $\boldsymbol{\beta}_\eta$ and directional and seasonal roughnesses $\boldsymbol{\lambda}_\eta$ corresponding to model parameters η of the three models: gamma body ($\eta \in \{\alpha, \zeta\}$), generalised Pareto tail ($\eta \in \{\nu, \xi\}$) and rate of occurrence ($\eta = \rho$). Unfortunately, none of these posteriors is available in closed form. We therefore use Markov chain Monte Carlo (MCMC) to sample iteratively from full conditionals independently for each η in each model in turn

$$\begin{aligned} f(\boldsymbol{\beta}_\eta | \boldsymbol{y}, \Omega \setminus \boldsymbol{\beta}_\eta) &\propto f(\boldsymbol{y} | \boldsymbol{\beta}_\eta, \Omega \setminus \boldsymbol{\beta}_\eta) \times f(\boldsymbol{\beta}_\eta | \boldsymbol{\lambda}_\eta) \ , \\ f(\boldsymbol{\lambda}_\eta | \boldsymbol{y}, \Omega \setminus \boldsymbol{\lambda}_\eta) &\propto f(\boldsymbol{\beta}_\eta | \boldsymbol{\lambda}_\eta) \times f(\boldsymbol{\lambda}_\eta) \ , \end{aligned}$$

where Ω is the full set of parameters for the relevant model. Since full conditionals are not available in closed form, we use Metropolis-Hastings within Gibbs, exploiting function gradient and curvature information (using mMALA, following Girolami and Calderhead 2011) to generate Metropolis-Hastings proposals efficiently. For further details, see Randell et al. (2016).

4. Return value estimation by numerical integration

Typically, return value distributions are estimated by Monte Carlo simulation under the estimated model (e.g. Randell et al. 2016). For long return periods $T \gg 1000$ years however, this simulation can be computationally the most time-consuming stage of analysis: numerical integration is then a more attractive strategy.

4.1. Storm peak return values

Consider directional-seasonal bin S_j $(j = 1, 2, ..., m)$ centred on location I_j in the index set of covariate locations (Section 3.3). S_j is sufficiently small that all model parameters $\rho, \alpha, \zeta, \xi, \nu$ and ψ are assumed constant within it. For bin S_j, with $\omega_j = \{\alpha_j, \zeta_j, \xi_j, \nu_j, \psi_j\}$ for brevity, we define $F(y|\omega_j)$ to be the cumulative distribution function of any storm peak event given ω_j. If ρ_j is interpreted as the number of storm peak events in S_j per annum, we estimate the cumulative distribution function $F_{M_T}(y|\omega_j)$ of the maximum M_T observed in a period of T years in S_j as

$$\begin{aligned} F_{M_T}(y|\omega_j) &= \mathbb{P}\left(M_T < y\right) \\ &= \sum_{k=0}^{\infty} \mathbb{P}\left(k \text{ events in } S_j \text{ in } T \text{ years }\right) \times \mathbb{P}^k\left(\text{size of an event in } S_j < y\right) \\ &= \sum_{k=0}^{\infty} \frac{(T\rho_j)^k}{k!} \exp(-T\rho_j) \times F^k(y|\omega_j) \\ &= \exp\left(-T\rho_j\left(1 - F(y|\omega_j)\right)\right). \end{aligned}$$

Since storm peak events are independent given covariates, we estimate the cumulative distribution function $F_{M_T}(y|\boldsymbol{\omega})$ (where $\boldsymbol{\omega} = \{\omega_j\}_{j=1}^m$) of the "omni-directional omni-seasonal" storm peak maximum M_T over all directional-seasonal bins, by taking the product

$$F_{M_T}(y|\boldsymbol{\omega}) = \prod_{j=1}^{m} F_{M_T}(y|\omega_j).$$

The final estimate for $F_{M_T}(y)$, unconditional on $\boldsymbol{\omega}$, is estimated by marginalising over $\boldsymbol{\omega}$

$$F_{M_T}(y) = \int_{\boldsymbol{\omega}} F_{M_T}(y|\boldsymbol{\omega}) f(\boldsymbol{\omega}) d\boldsymbol{\omega},$$

where $f(\boldsymbol{\omega})$ is the estimated posterior density for $\boldsymbol{\omega}$. In practice, we sample values of $F_{M_T}(y|\omega_j)$ and $F_{M_T}(y|\boldsymbol{\omega})$ for some suitable set of storm peak values y from the Markov chain Monte Carlo simulation following convergence for all members of the ensemble over τ, providing immediate posterior predictive estimates for distributions of storm peak maxima corresponding to return period T years.

4.2. Sea state return values

For applications, it is also necessary to estimate the distribution of return value $M_{T\mathbb{S}}(y)$ for maxima of sea state (as opposed to storm peak) H_S events for arbitrary directional sector $\mathbb{S}$ (Ewans and Jonathan 2008). To achieve this, we need to accommodate contributions of so-called "dissipated" storms, the storm peak directions for which do not correspond to $\mathbb{S}$, to $M_{T\mathbb{S}}(y)$, again using numerical integration.

First we estimate the storm dissipation function $\delta(\mathbb{S}; j, y)$ for sea state H_S in directional sector $\mathbb{S}$ from a single storm with storm peak direction in directional-seasonal bin $S_j, j = 1, 2, ..., m$ and storm peak value y. $\delta(\mathbb{S}; j, y)$ is estimated empirically from the sample of storm trajectories described in Section 2 and illustrated in Figure 3. Next we estimate the cumulative distribution

function $F_{D_{\mathbb{S}}}(d|\omega_j)$ of $D_{\mathbb{S}}$, the dissipated sea state H_S in sector $\mathbb{S}$ from a random storm dissipating from directional-seasonal bin S_j

$$F_{D_{\mathbb{S}}}(d|\omega_j) = \mathbb{P}(D_{\mathbb{S}} \leqslant d|\omega_j) \quad = \quad \int_y \mathbb{P}(\delta(\mathbb{S}; j, Y) \leqslant d|Y = y) f(y|\omega_j) dy,$$

where $f(y|\omega_j)$ is the marginal directional density of storm peak H_S in directional-seasonal bin S_j corresponding to cumulative distribution function $F(y|\omega_j)$. The cumulative distribution $F_{D_{\mathbb{S}}^T}(d|\omega_j)$ of $D_{\mathbb{S}}^T$, the T-year maximum of $D_{\mathbb{S}}$ from directional-seasonal bin S_j, is evaluated using the same argument as for storm peaks in Section 4.1, as $\exp\left(-T\rho_j\left(1 - F_{D_{\mathbb{S}}}(d|\omega_j)\right)\right)$. The cumulative distribution $F_{M_{T\mathbb{S}}}(d|\boldsymbol{\omega})$ of $M_{T\mathbb{S}}$, the T-year maximum of sea state H_S in directional sector $\mathbb{S}$ from all seasonal-directional bins, for each value of d, is therefore given by $\prod_{j=1}^{m} \mathbb{P}(M_{T\mathbb{S}} \leqslant d|\omega_j)$. Note that, by restricting this product to directional-seasonal bins corresponding to a particular season, seasonal estimates can be obtained. Finally, as for the storm peak case above, the estimate for $F_{M_{T\mathbb{S}}}(y)$ is found by marginalising over $\boldsymbol{\omega}$ as $\int_{\boldsymbol{\omega}} F_{M_{T\mathbb{S}}}(y|\boldsymbol{\omega}) f(\boldsymbol{\omega}) d\boldsymbol{\omega}$.

4.3. Computational efficiency

In typical applications, numerical integration yields large improvements in computational efficiency even for moderate values of T (e.g. 10 years) compared to Monte Carlo simulation for estimation of return value distributions. For example, Monte Carlo simulation of 800 realisations of length of 1, 10, 100, 1,000 and 10,000 years and subsequent estimation of the corresponding return value distributions, for standard directional-seasonal partitions as illustrated in Section 5, for both storm peak and sea state H_S requires 6 hours 15 minutes computation on an Intel Xeon 2.7GHz workstation with 40 cores and 192GB RAM. The same inferences are made using numerical integration in 2 minutes to approximately the same accuracy, as judged from the results of a number of studies using "toy" models. In this case, time to estimate the extreme value model prior to return value estimation, is 18 minutes. Thus, complete analysis time is reduced from around 6 hours 30 minutes to 20 minutes.

5. Application to South China Sea significant wave height

The model described in Section 3 is estimated for storm peak H_S on direction θ and season ϕ for a regularly spaced sequence of 20 values of quantile regression non-exceedance probability $\tau \in [0, 1]$. For each value of τ, posterior distributions for rate parameter ρ, gamma parameters α and ζ jointly, and generalised Pareto parameters ξ and ν (and hence σ) jointly are estimated using Bayesian inference. To confirm convergence we perform inference for each value of τ twice, and ensure that the resulting Markov chain Monte Carlo traces are similar (in the sense of having similar first and second-order structure) following burn-in. The posterior predictive distribution of full sample negative log likelihood is summarised in Figure 4(a) in terms of a median and 95% uncertainty band; this is a predictive estimate of how well the whole gamma-generalised Pareto model fits the sample. Good model fit is desirable, hence we prefer values of τ corresponding to low values of this likelihood. Figure 4(a) exhibits a minimum at $\tau = 0.2$, and an approximately constant value for $\tau \in [0.1, 0.8]$; the negative log likelihood increases outside this interval (as noted in the figure caption). The plot suggests that values of $\tau \in [0.1, 0.8]$ would be reasonable choices for subsequent analysis, since the estimate of negative log likelihood is relatively low here. Figure 4(b) illustrates percentiles of the distribution of the 10,000-year return value, estimated using the

fitted generalised Pareto model; it is desirable that the estimate of this return value be relatively constant on the interval of τ chosen for subsequent inference. The median return value reduces with increasing τ in general, as does the width of the distribution. However, for $\tau \in [0.2, 0.8]$ the return value is relatively stable. Based on the evidence from Figure 4 (a) and (b), the interval $\tau \in [0.2, 0.8]$ appears reasonable for subsequent analysis.

[Figure 4 about here.]

Figure 5 compares empirical sample-based and model-based density estimates for the two most populous regions of the covariate domain, for $\tau = 0.2, 0.5$ and 0.8. Agreement is reasonable for all threshold choices in both Figure 5(a,b) (corresponding to $\theta \times \phi = [310, 350] \times [\text{Jun}, \text{Sep}]$) and Figure 5(c,d) (corresponding to $\theta \times \phi = [150, 190] \times [\text{Dec}, \text{Feb}]$). These covariate intervals correspond to combinations of the underlying directional-seasonal bins $\{S_j\}_{j=1}^{m}$. Plots for other values of $\tau \in [0.2, 0.8]$ are similar. There is evidence in Figure 5 that for $\tau \approx 0.2$ the overall quality of model fit is dominated by the gamma distribution fit to non-exceedances. The peak in the empirical sample density at around 1m is described better using $\tau = 0.2$ than when using $\tau > 0.2$; there is some evidence for this effect also in Figure 12 of the appendix. In this sense, the good overall model fit at $\tau = 0.2$ could be regarded as an "artefact" of the specific model form chosen for threshold non-exceedances. We conclude that, whereas all threshold choices in $[0.2, 0.8]$ provide reasonable models, proceeding using an ensemble of models corresponding to the restricted interval $\tau \in [0.3, 0.7]$ avoids the extremities of the interval $[0.2, 0.8]$ at which diagnostic plots indicate that model performance begins to degrade.

[Figure 5 about here.]

The ensemble model is then re-estimated for storm peak H_S on direction θ and season ϕ for a regularly spaced sequence of 20 values of quantile regression non-exceedance probabilities τ in the restricted interval $[0.3, 0.7]$. Figure 6 shows directional-seasonal variation of the set of posterior median parameter estimates over MCMC chains and the ensemble of τ values. Extreme value threshold ψ, gamma scale α and generalised Pareto scale σ show similar directional-seasonal variability; this reflects the strong monsoonal influence on the response. The Poisson rate ρ is relatively similar. Gamma shape ζ varies less with covariates, and the variability in median generalised Pareto shape ξ is small given that $\xi \in (-0.5, 0)$ is typically observed for studies of storm peak H_S.

[Figure 6 about here.]

The quality of model fit, particularly for the generalised Pareto component, is judged by comparing empirical sample-based and model-based tail estimates for a time period of the original sample, as shown in Figure 7. Agreement between the empirical and model-based tails is excellent in general. Poorer performance for the eastern octant (E) is attributed to a small number of large events in this region which are not adequately captured in smooth model parameter variation with covariates. The Poisson model characterises the rate of occurrence of events well directionally. The corresponding plot exploring seasonal fit shows equally good agreement.

[Figure 7 about here.]

To this point, all inferences made have concerned storm peak events. From an engineering perspective however, it is critical to capture the effects of storm dissipation in return value distributions.

Specifically, directional return values estimated from storm peaks alone ignore the influence of dissipation, and will in general be biased low, particularly in directional sectors for which the rate of occurrence of storm peaks is relatively low. Figure 8 illustrates the typical characteristics of dissipation at the Southern China Sea location under consideration here. It summarises the distribution of fractional dissipation, namely the value of the ratio of sea state H_S to storm peak value, for sea state direction relative to storm peak direction, for all values of storm peak and all seasons. The figure suggests that dissipation effects extend to $\pm 45^\circ$ for all directional sectors, but that their influence beyond $\pm 45^\circ$ is small, although the angular extent of dissipation does vary between octants. The northeastern (NE) and eastern (E) octants show asymmetry, suggesting that storms have larger dissipative influence when sea state direction lies anticlockwise of storm peak direction.

[Figure 8 about here.]

Using dissipation and the approach outlined in Section 4, return value distributions for sea state H_S corresponding to a return period of 10,000-year are estimated for different standard directional-seasonal partitions, as illustrated in Figure 9 and Figure 10; the former explores the return value distribution's seasonal variability, and the latter directional variability. Monsoonal characteristics are obvious in both figures: December, January and February plots in Figure 11 are dominated by the Northeast Monsoon propagating approximately southwards (with $\theta \approx 180^\circ$). July to August is similarly dominated by the Southwest Monsoon.

[Figure 9 about here.]

[Figure 10 about here.]

6. Discussion and conclusions

Extreme value models are motivated by asymptotic arguments. The generalised Pareto form is justified only for peaks-over-threshold modelling for high threshold ψ. Pre-specifying ψ is in general problematic, yet key inferences such as return values for long return periods can vary considerably with choice of ψ. A whole sample model is desirable since it allows estimation of ψ as a model parameter. In a whole sample model incorporating extreme value inference and estimation of ψ, it is not clear whether estimation of ψ should be influenced by model fit to threshold non-exceedances: adequate fit of the generalised Pareto tail should take priority, since tail estimation is our primary concern. From this perspective, whole sample models for which ψ is pre-specified with no regard to the values of threshold non-exceedances would seem advantageous. When there is compelling motivation for the form of the model for threshold non-exceedances, then perhaps allowing fit to body to influence estimation of ψ is more acceptable. For example, a Rayleigh distribution for individual wave heights can be derived theoretically given appropriate assumptions; it would therefore seem sensible to value the fit of a Rayleigh model to the body of a distribution of individual wave heights as much as the generalised Pareto distribution for extreme wave heights. Wadsworth et al. (2010) use similar logic to motivate their two part model. An alternative approach is to specify a model for the body of the distribution sufficiently flexibly that the body fit will not materially influence the estimation of ψ or the tail; this is the approach taken by Tancredi et al. (2006) and MacDonald et al. (2011). In the current work, we estimate plausible thresholds ψ by inspecting the quality of fit of gamma-generalised Pareto models, each incorporating a prior quantile regression

specified in terms of quantile probability τ. We then retain an ensemble of plausible models for return value inference. We find this approach to be a useful compromise between the statistical ideal of estimating ψ as a model parameter, the concerns above regarding the influence of model body fit on ψ and hence the tail fit, and the general weakness of evidence for ψ in typical met-ocean samples.

We also find that numerical integration provides a much improved approach, particularly in terms of computational efficiency, to estimation of return value distributions, even when non-trivial effects such storm non-stationarity with respect of covariates, and storm dissipation are involved. Numerical integration also leads to efficient estimation of the effects of within-sea state variability, e.g. of return value distributions for wave height or crest elevation. The availability of fast estimation for return value distributions corresponding to long return periods makes their use in model diagnosis far more tractable and appealing. Since the 10,000-year return value in particular is a key design quantity, it is natural to assess model performance in terms of the characteristics of this statistic. In estimating sea state return values, we are currently working to improve the model for the storm dissipation function underlying the estimation in Section 4.2 using Markov extremal models following Winter and Tawn (2016) and Winter and Tawn (2017).

Using the ensemble model, a complete non-stationary extreme value analysis for a sample of the order of 5000 storms on 32×24 directional-seasonal bins and 10×10 tensor product of second-order spline bases, including all model estimation using Bayesian inference, ensemble assessment and refinement, and calculations of return value distributions for storm peak and sea state significant wave height corresponding to return periods of arbitrary length, can be completed in a matter of minutes in a statistically rigorous manner.

Acknowledgements

We acknowledge the generous support of Vadim Anokhin, Shejun Fan, Laks Raghupathi and Yanyun Wu at Shell, and Matthew Jones at Durham University, UK.

Appendix: Supplementary model diagnostics

Using the procedure described in Section 4, we estimate the cumulative distribution of sea state significant wave height, and use it to compare tails from the ensemble model with empirical estimates directly from the sample. Figure 11 illustrates the tail fit, in a similar vein to Figure 7, and shows good agreement. Figures 11 and 7 are generally similar; the left hand plots are identical (as should be the case by definition). The right hand directional plots for sea state H_S (Figure 11) extend to more negative values of $\log_{10}(1-F)$ as would be expected, since the number of sea state H_S values per directional octant exceeds the corresponding number of storm peak events. The location of the tails for corresponding directional octants is very similar, indicating that dissipated values do not contribute to the largest values of sea state H_S in Figure 11.

[Figure 11 about here.]

Figure 12 shows a whole sample quantile-quantile plot on standard Gaussian scale for the estimated ensemble model. There is reasonable agreement except perhaps for the smallest values of response. This lack of fit is also visible in Figure 5(a,c), and suggests that the gamma model has some difficulty in accommodating the rapid increase in probability mass with increasing response for the smallest values of response. We judge this not to be of major concern since the objective of this work is characterising the right hand tail of the distribution.

[Figure 12 about here.]

To illustrate the quality of model fit further, Figure 13 compares estimates of seasonal return values for storm peak H_S, using a procedure analogous to that used to generate Figure 7.

[Figure 13 about here.]

References

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: Applied Statistics 54, 207–222.

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.

Ewans, K. C., Jonathan, P., 2008. The effect of directionality on northern North Sea extreme wave design criteria. J. Offshore. Arct. Eng. 130, 041604:1–041604:8.

Ghosh, S., Resnick, S., 2010. A discussion on mean excess plots. Stoch. Proc. Appl. 120, 1492–1517.

Girolami, M., Calderhead, B., 2011. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. J. Roy. Statist. Soc. B 73, 123–214.

Koenker, R., 2005. Quantile regression. Cambridge University Press, New York.

MacDonald, A., Scarrott, C. J., Lee, D., Darlow, B., Reale, M., Russell, G., 2011. A flexible extreme value mixture model. Comput. Statist. Data Anal. 55, 2137–2157.

Northrop, P., Attalides, N., Jonathan, P., 2016. Cross-validatory extreme value threshold selection and uncertainty with application to ocean storm severity. J. R. Stat. Soc. C. doi:10.1111/rssc.12159.

Northrop, P. J., Coleman, C. L., 2014. Improved threshold diagnostic plots for extreme value analyses. Extremes 17, 289–303.

Randell, D., Feld, G., Ewans, K., Jonathan, P., 2015. Distributions of return values for ocean wave characteristics in the South China Sea using directional-seasonal extreme value analysis. Environmetrics 26, 442–450.

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

Sanchez-Archilla, A., Aguar, J. G., Egozcue, J. J., Prinos, P., 2008. Extremes from scarce data: the role of Bayesian and scaling techniques in reducing uncertainty. J. Hydraul. Res. 46, 224–234.

Scarrott, C., MacDonald, A., 2012. A review of extreme value threshold estimation and uncertainty quantification. Revstat 10, 33–60.

SEAFINE, 2012. South East Asia Meteorological and Oceanographic Hindcast (SEAMOS) Fine Grid Hindcast Study Joint Industry Project. Oceanweather Inc.

Tancredi, A., Anderson, C., O'Hagan, A., 2006. Accounting for threshold uncertainty in extreme value estimation. Extremes 9, 87–106.

Thompson, P., Cai, Y., Reeve, D., Stander, J., 2009. Automated threshold selection methods for extreme wave analysis. Coastal Eng. 56, 1013–1021.

Wadsworth, J. L., 2016. Exploiting structure of maximum likelihood estimators for extreme value threshold selection. Technometrics 58 (1), 116–126.

Wadsworth, J. L., Tawn, J. A., Jonathan, P., 2010. Accounting for choice of measurement scale in extreme value modelling. Annals of Applied Statistics 4, 1558–1578.

Winter, H. C., Tawn, J. A., 2016. Modelling heatwaves in central France: a case-study in extremal dependence. J. Roy. Statist. Soc. C 65, 345–365.

Winter, H. C., Tawn, J. A., 2017. kth-order Markov extremal models for assessing heatwave risks. Extremes 20, 393–415.

List of Figures

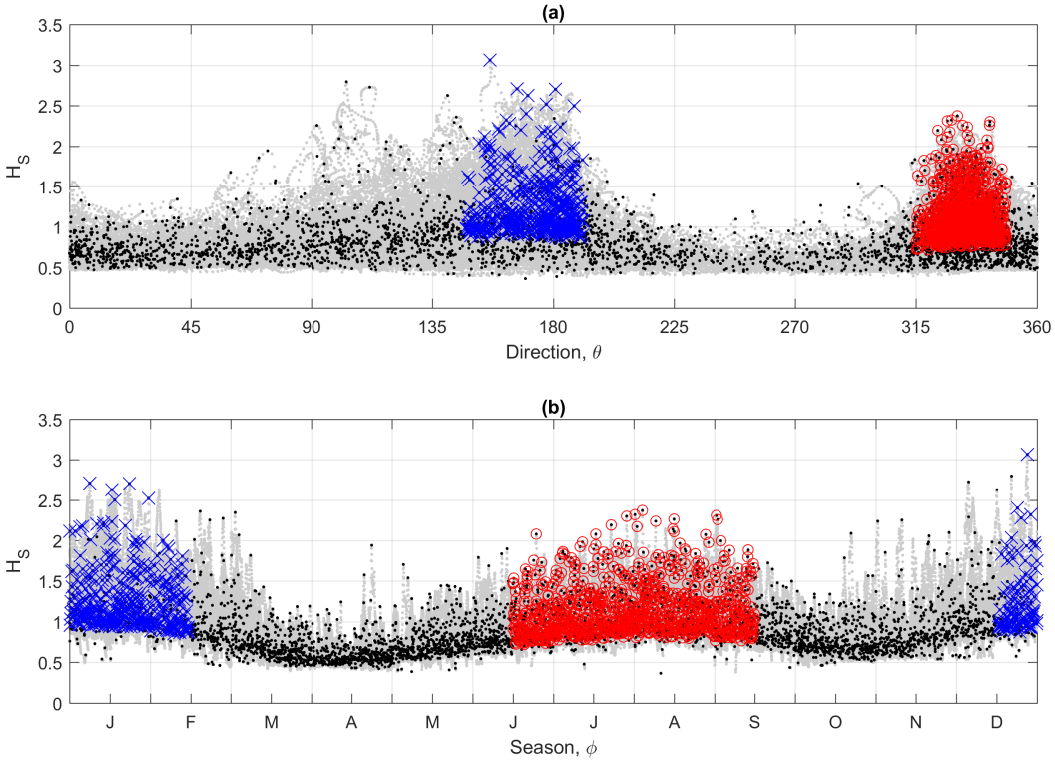

Figure 1: South China Sea sample. Significant wave height (H_S, grey) and storm peak H_S (black) on (a) direction θ and (b) season ϕ. Coloured storm peak events correspond to the Southwest Monsoon (red circles) and the Northeast Monsoon (blue crosses).

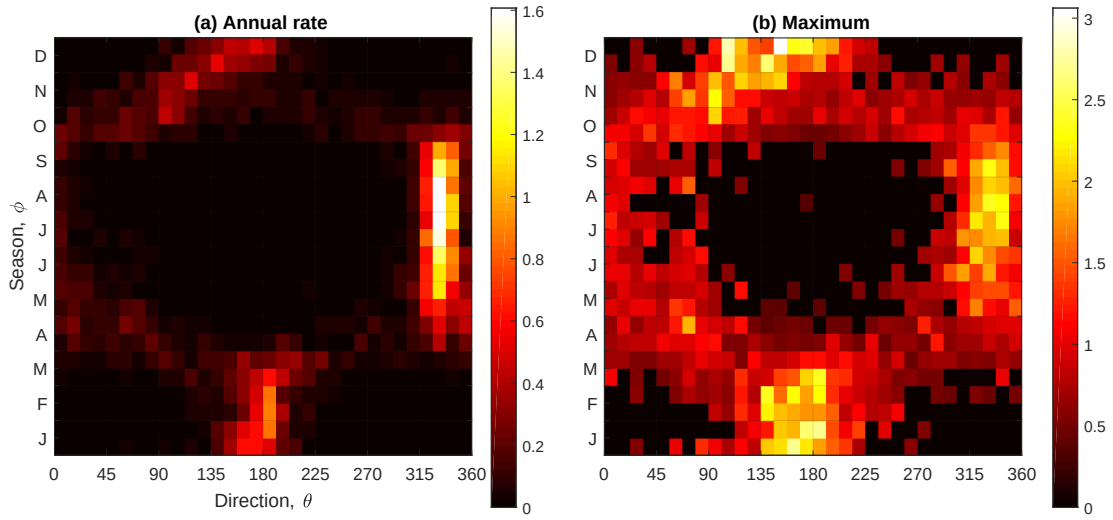

Figure 2: South China Sea sample. Panel (a) shows mean annual count rate of occurrences of storm peak H_S per directional-seasonal bin, and panel (b) the maximum observed value of storm peak H_S per directional-seasonal bin.

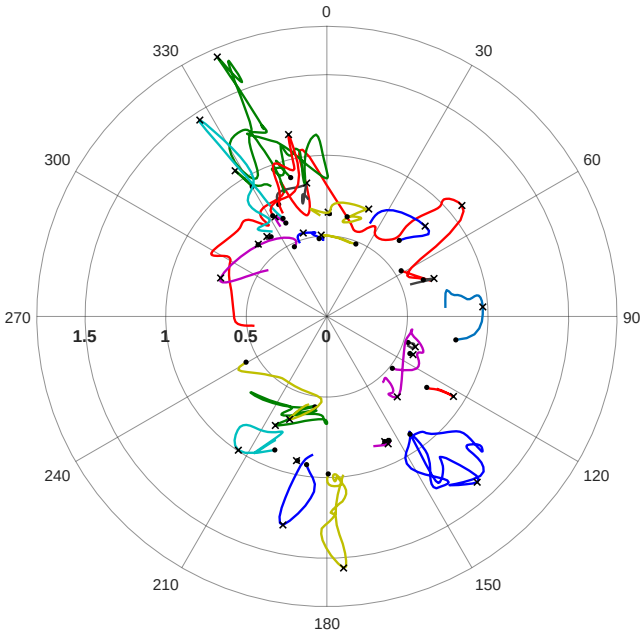

Figure 3: South China Sea sample. Polar plot of 30 typical storm trajectories. The start of each trajectory is marked with a dot, and the storm peak with a cross.

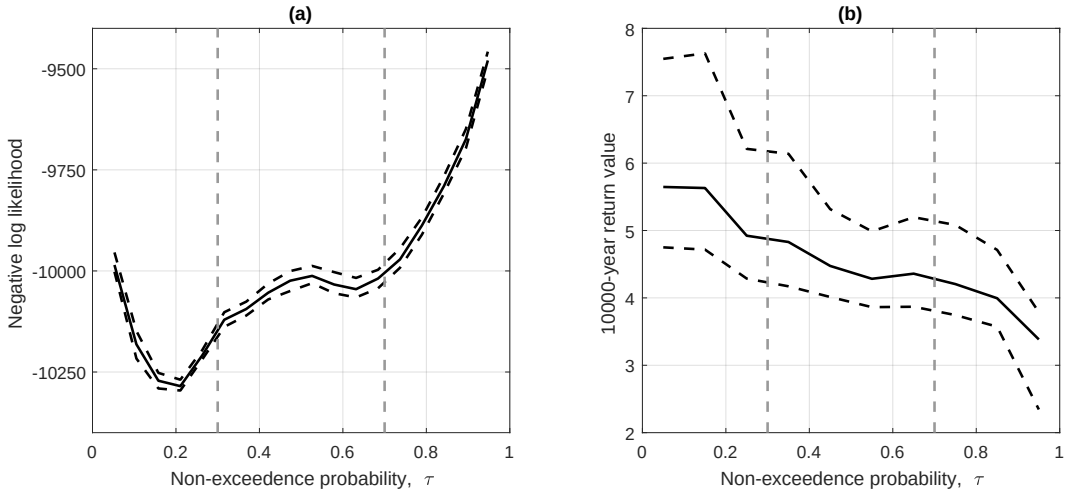

Figure 4: Model diagnostics for different quantile thresholds. Panel (a) summarises the posterior predictive distribution of negative log-likelihood as a function of quantile regression non-exceedance probability τ, in terms of the 2.5%, 50% and 97.5% percentiles. Panel (b) summarises the distribution of posterior predictive 10,000-year return value as a function of τ in terms of the same percentiles. In panel (a), omitted median values of negative log likelihood at $\tau = 0$ and $\tau = 1$ respectively were estimated as -8,185 and -8,232; the interval $[0.2, 0.8]$ therefore represent a deep well. The boundaries of the restricted interval $[0.3, 0.7]$ used for subsequent inference are shown as vertical dashed grey lines in both panels.

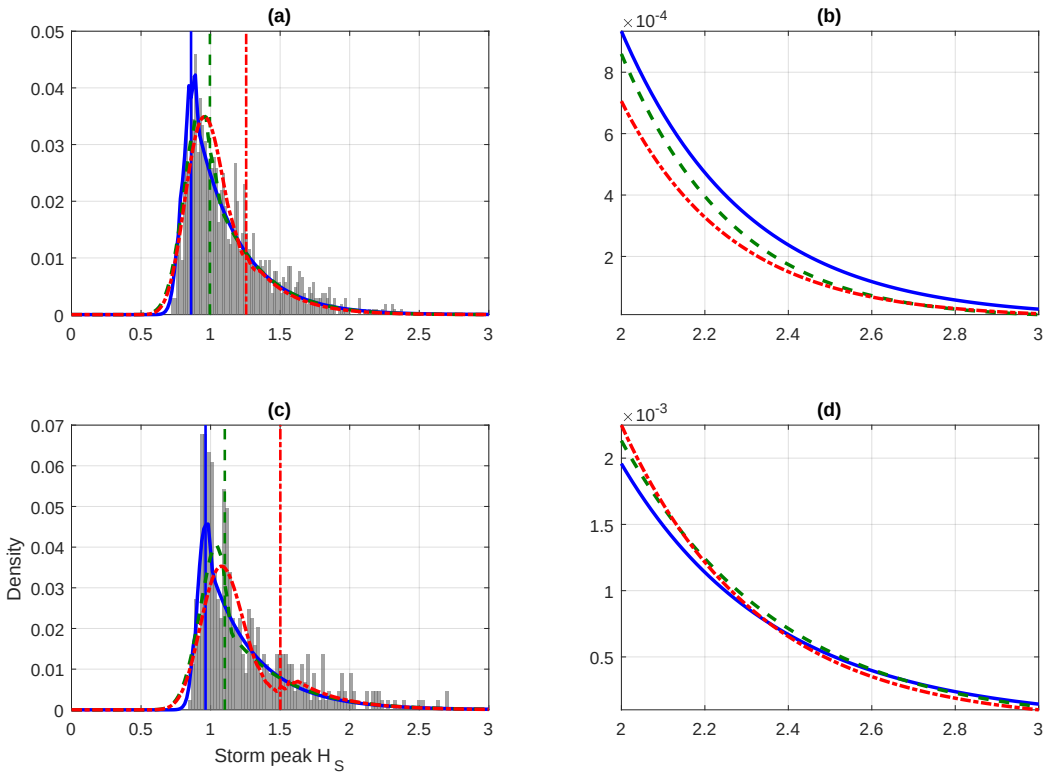

Figure 5: Densities for most populous regions of covariate domain. Empirical density (from sample, grey) and model-estimated densities with $\tau = 0.2$ (blue), $\tau = 0.5$ (green) and $\tau = 0.8$ (red). Panels (a,c) show the full density, and panels (b,d) focus on the tail. Threshold levels on physical scale are shown as vertical lines with corresponding colours in panels (a,c). Panels (a,b) show covariate interval $\theta \times \phi = [310, 350] \times [\text{Jun}, \text{Sep}]$, and panels (c,d) show $\theta \times \phi = [150, 190] \times [\text{Dec}, \text{Feb}]$.

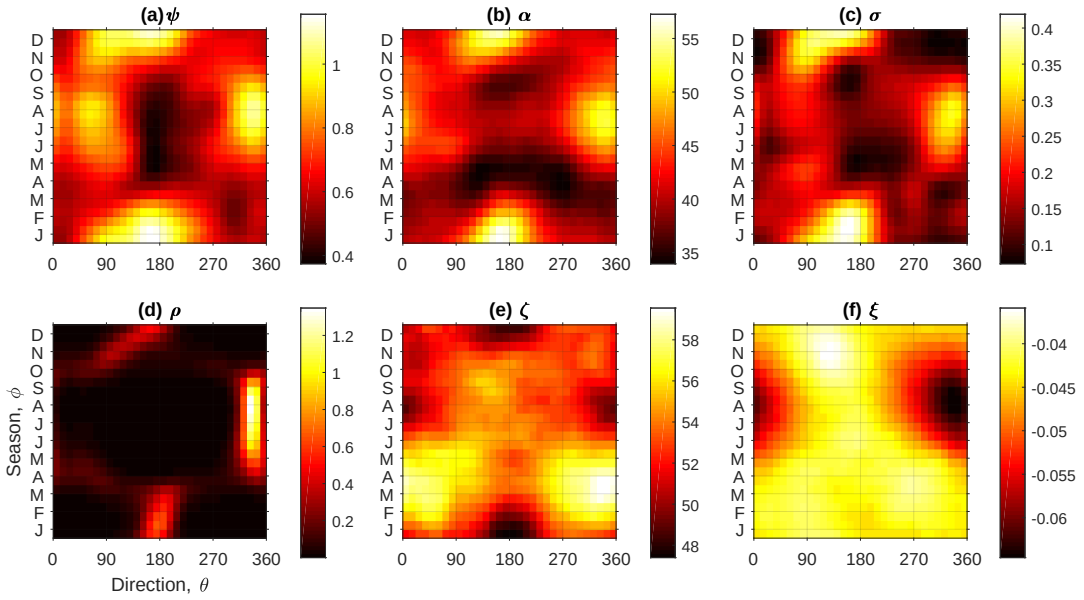

Figure 6: Median posterior parameter estimates with direction and season for (a) extreme value threshold ψ, (b) truncated gamma shape α, (c) generalised Pareto scale σ, (d) rate of occurrence ρ, (e) truncated gamma scale ζ and (f) generalised Pareto shape ξ.

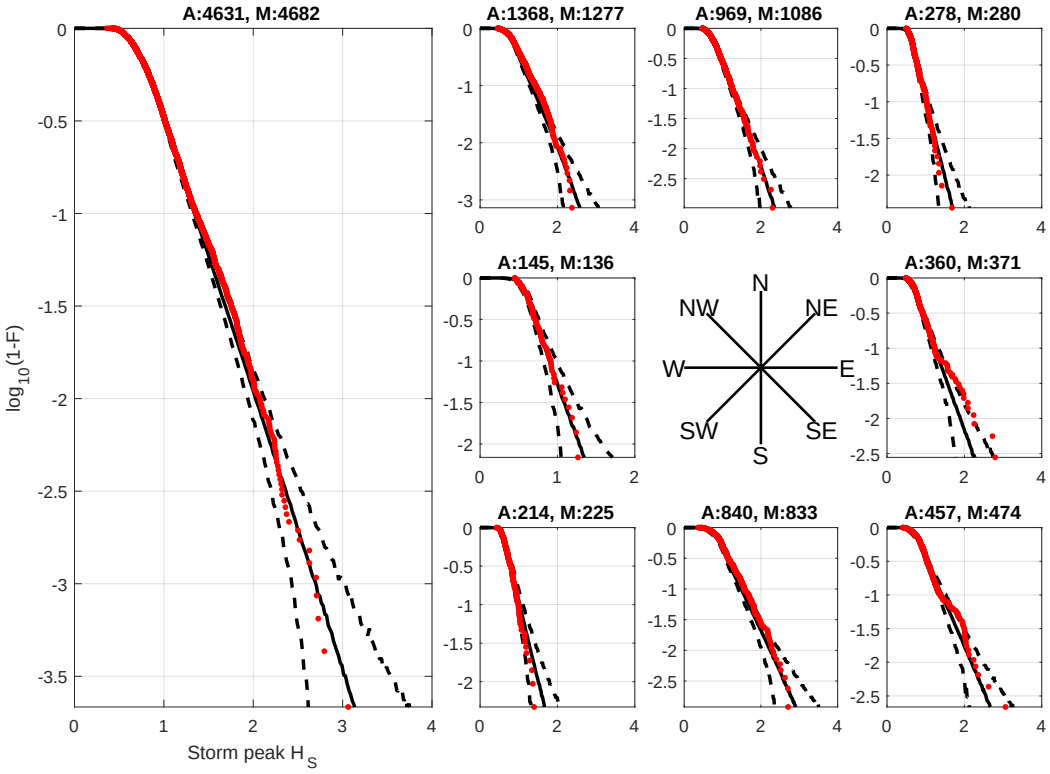

Figure 7: Illustration of model validation by comparison of estimates for the distribution of storm peak H_S corresponding to the period of the original sample, plotted as $\log_{10}(1-F)$ to accentuate tail behaviour, for cumulative distribution function F. The dotted red curve is an empirical estimate. The black curves summarise the predictive distribution of the quantile estimate (for given tail probability $1-F$) under the directional-seasonal model, as the median (solid) and 2.5% and 97.5% values (dashed), estimated using numerical integration. The left hand panel corresponds to the omni-directional omni-seasonal case, and the right hand panels to omni-seasonal estimates for each of 8 directional octants. Titles for plots are the numbers of actual (A:) events and the median number (M:) of events simulated.

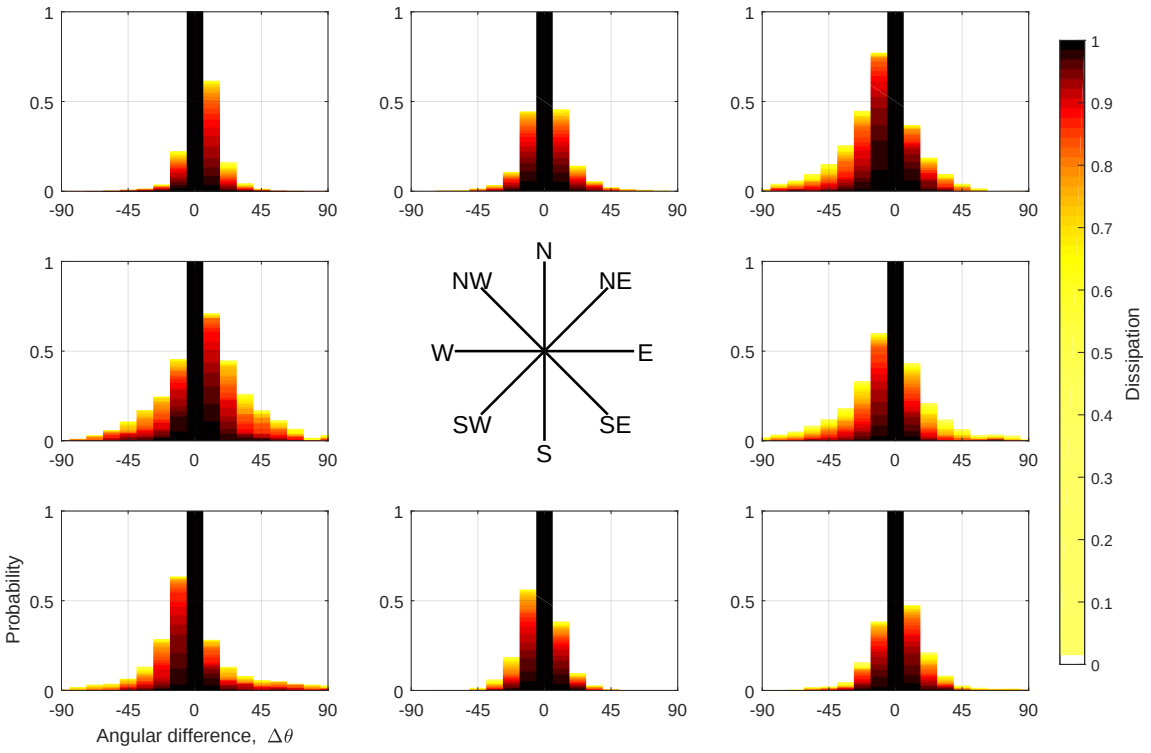

Figure 8: Distribution of directional storm dissipation as a *fraction* of storm peak H_S for directional octants centred on the cardinal directions. Each panel illustrates the directional distribution of the *fractional* dissipation of storms with storm peaks in the corresponding octant, as a function of angular difference from storm peak direction. For each angular difference, the vertical combination of coloured bars provides a visual representation of the distribution of dissipation: colour indicates the value of dissipation decreasing from unity (black) to zero (white), and y-axis value the probability of at least that value of dissipation. In each octant for angular difference 0, dissipation is unity by definition since the storm peak occurs for that angular difference. Further, e.g., in the northeast octant (NE), for angular difference $-45°$, non-zero dissipation is observed with probability approximately 0.2, and total dissipation (to zero) with probability approximately 0.8. For the same octant and angular difference $45°$, total dissipation occurs with probability approximately 0.9.

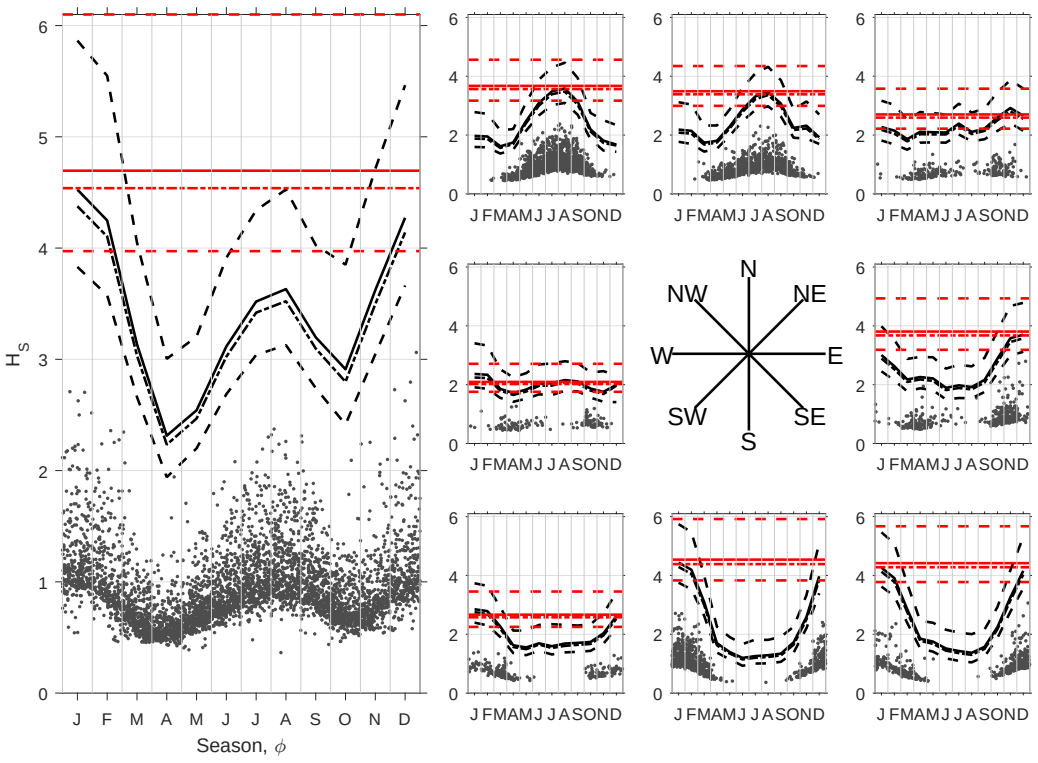

Figure 9: Predictive distribution of sea state H_S corresponding to the 10,000-year return period. The left hand panel summarises seasonal variation of the omni-directional return value in terms of the 2.5%, 37%, median and 97.5% values of the predictive distribution (black). Also shown are the corresponding omni-seasonal estimates (in red). The right hand plot shows the corresponding curves for each of 8 directional octants.

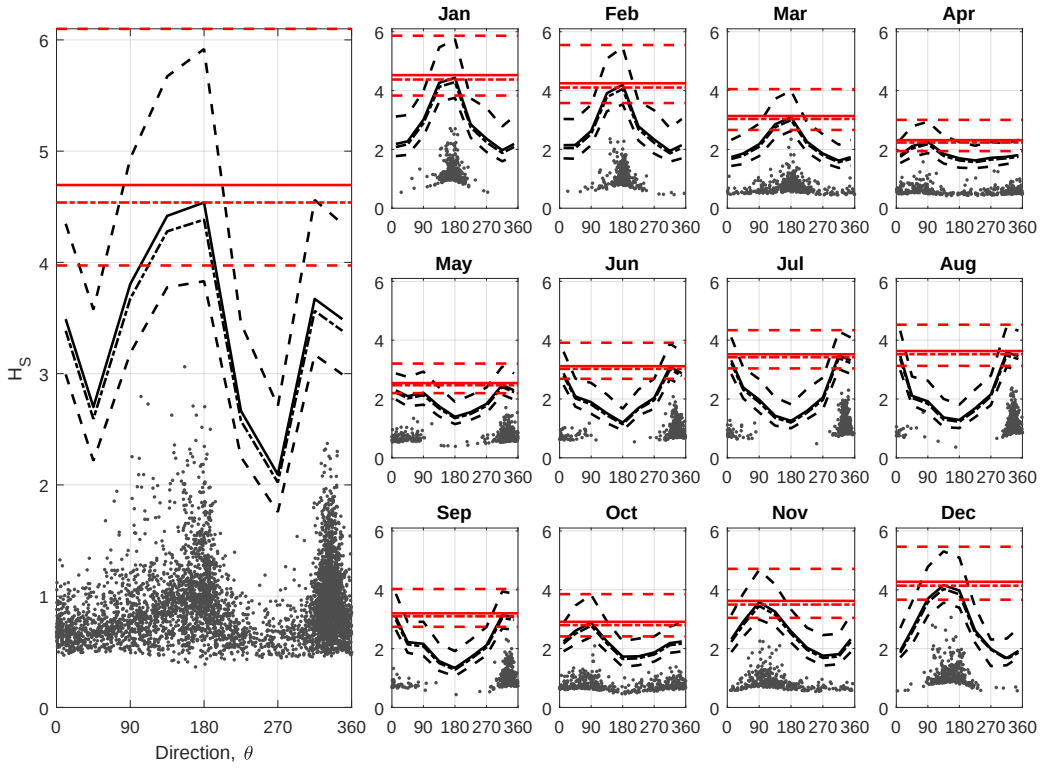

Figure 10: Predictive distribution of sea state H_S corresponding to the 10,000-year return period. The left hand panel summarises directional variation of the omni-seasonal return value in terms of the 2.5%, 37%, median and 97.5% values of the predictive distribution (black). Also shown are the corresponding omni-directional estimates (in red). The right hand plot shows the corresponding curves for each of 12 months.

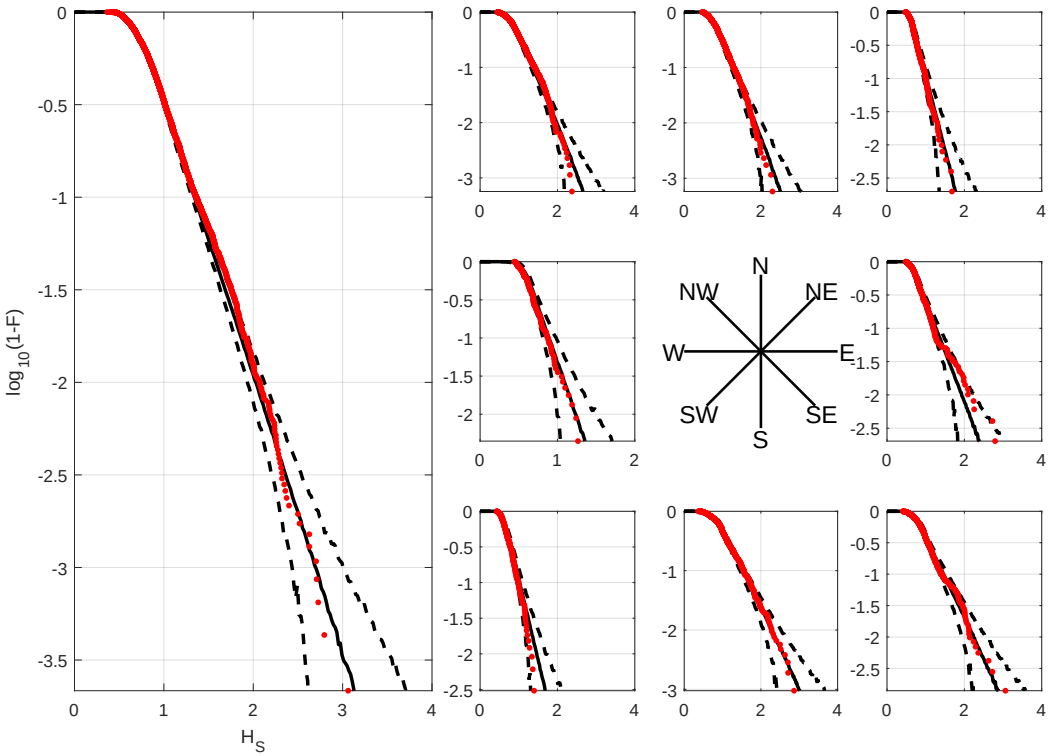

Figure 11: Illustration of model validation by comparison of estimates for the distribution of sea state H_S corresponding to the period of the original sample, plotted as $\log_{10}(1-F)$ to accentuate tail behaviour. For other details, see Figure 7.

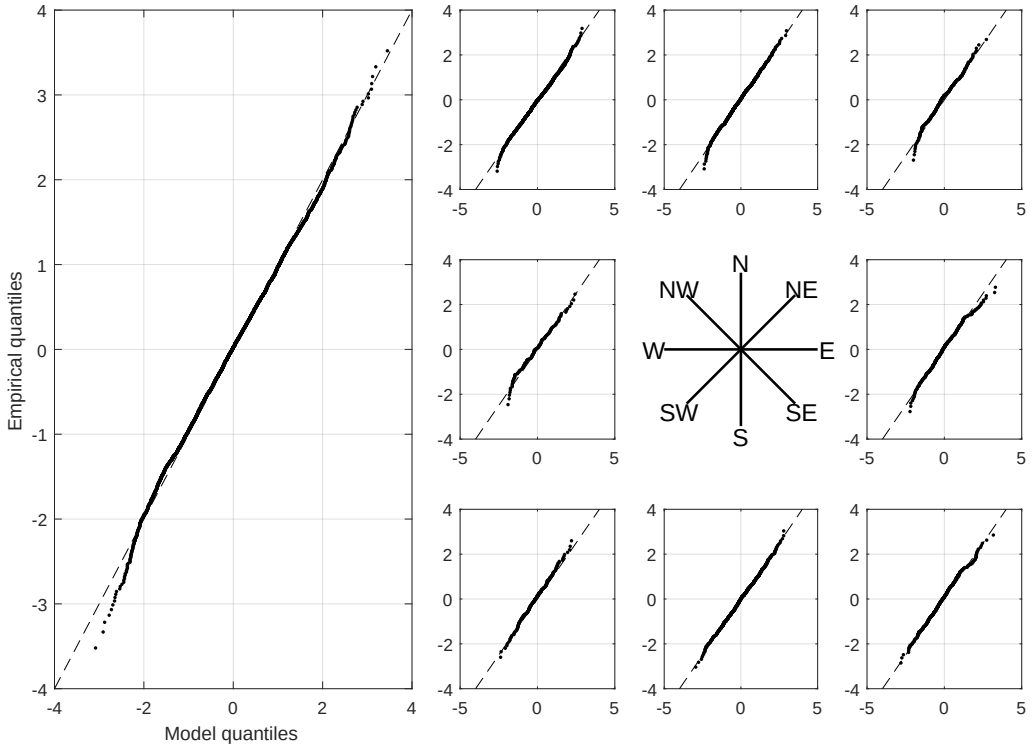

Figure 12: Quantile-quantile plot for estimated model on standard Gaussian scale. The left hand panel shows a scatter plot of empirical quantiles against fitted quantiles for all directions and seasons. The right hand plot shows the corresponding plot for each of 8 directional octants.

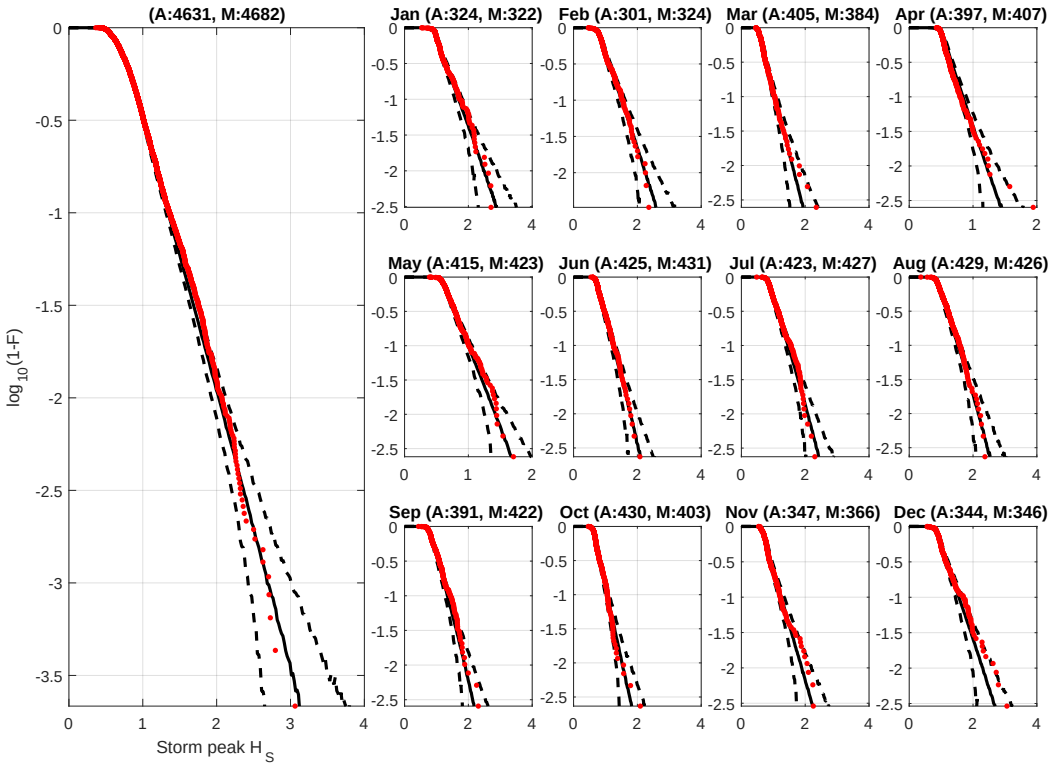

Figure 13: Illustration of model validation by comparison of estimates for the seasonal distribution of storm peak H_S corresponding to the period of the original sample, plotted as $\log_{10}(1-F)$ to accentuate tail behaviour. The titles for panels give the numbers of actual (A:) and estimated (M:) events in the corresponding seasonal period. For other details, see Figure 7.