Environmetrics 26 442-450 (2015)

Distributions of return values for ocean wave characteristics in the South China Sea using directional-seasonal extreme value analysis

D. Randell, G. Feld, K. Ewans and P. Jonathan*

Abstract

Estimation of ocean environmental return values is critical to the safety and reliability of marine and coastal structures. For ocean waves and storm severity, return values are typically estimated by extreme value analysis of time series of measured or hindcast *sea state* significant wave height H_S. For a single location, this analysis is complicated by the serial dependence of H_S in time and its non-stationarity with respect to multiple covariates, particularly direction and season.

Here, we report a non-stationary extreme value analysis of *storm peak* significant wave height H_S^{sp}, assumed temporally independent given covariates, incorporating directional and seasonal effects using a spline-based methodology incorporating an ensemble of models for different extreme value thresholds. Quantile regression is used to estimate suitable thresholds. For each threshold, a Poisson process is used to estimate the rate of occurrence of threshold exceedances, and a generalised Pareto model characterises the magnitude of threshold exceedances. Covariate effects are incorporated at each stage using penalised tensor products of B-splines to give smooth model parameter variation as a function of covariates. Optimal smoothing penalties are selected using cross-validation, and uncertainty quantified using BCa bootstrap re-sampling.

We use the model to estimate environmental return values for a location in the Makassar Strait, in the South China Sea. Return values distributions for H_S^{sp} are estimated by simulation under the *threshold ensemble* model. Return values for H_S are then estimated by simulating intra-storm trajectories of H_S consistent with the characteristics of the simulated storm peak events using a matching procedure. Return values for maximum individual crest elevation C are estimated by marginalisation using pre-specified conditional distribution for C given H_S and other sea state parameters. Model validation is performed by comparing confidence intervals for cumulative distribution functions of H_S^{sp} and H_S for the period of the data with empirical sample-based estimates.

1 Introduction

Safe design and operation of marine and coastal structures requires reliable estimation of extreme storm environments corresponding to return periods of hundreds or thousands of years, yet data for extreme environments is only available for historical periods typically of tens of years. Extreme value analysis is therefore essential to characterise the tails of distributions of environmental variables, and estimate their extreme quantiles. However, the characteristics of extreme ocean storm events are also known to vary systematically with covariates such as direction (e.g. Mackay et al. 2010) and season (e.g. Mendez et al. 2008). It is therefore critical to employ non-stationary extreme value models to estimate extreme quantiles accurately (e.g. Jonathan et al. 2008). These models enable consistent estimation of return value distributions for storm variables corresponding to arbitrary combinations of storm directions and seasons, also required in design. Moreover, design specifications (e.g. ISO19901-1 (2005)) require that return values be estimated not only for overall storm severity, but also for specific storm characteristics such as the maximum wave height (from trough to following crest) and the maximum crest elevation (from mean sea level to crest). This inference is severely complicated by the serial dependence of oceanographic variables in time, and their dependence on covariates.

The variability of the ocean's surface in a particular interval of time (typically 0.5, 1 or 3 hours, referred to as a sea state) is quantified using significant wave height H_S, which can be defined as 4 times the standard deviation of the ocean surface elevation (in metres) for the sea state. Within a sea state, wave characteristics are assumed to be stationary. However, H_S for consecutive sea states (referred to as sea state H_S, or simply H_S) exhibits serial dependence. Ocean storm events, composed of multiple consecutive sea states, are assumed to be statistically independent, since they are produced by independent occurrences of atmospheric pressure systems. Hence extreme value analysis of storm peak significant wave height H_S^{sp} is typically performed, where H_S^{sp} is the maximum value of sea state H_S over constituent sea states. Storm peak events at a location are assumed to be independent given covariates such as direction θ and

*Accepted by Environmetrics, June 2015; philip.jonathan@shell.com

season ϕ. Chavez-Demoulin and Davison (2005) suggest how extreme value analysis of H_S^{sp} might be achieved using extreme value models whose parameters are functions of covariates. To impose smooth variation of extreme value characteristics of H_S^{sp} with direction and season, we use penalised B-spline representations for all model parameters following the work of Eilers and Marx (2010). Efficient inference is achieved using ideas from generalised linear array models (Currie et al. 2006).

The distributional characteristics of sea state parameter set $\mathcal{S}$ (which includes H_S) given storm peak characteristics can be estimated from a historical sample of sea states corresponding typically to around 50 years. Given these, and given realisations of storm peak characteristics corresponding a longer return period of interest, return value distributions for parameter set $\mathcal{S}$ corresponding to the longer return period can also be estimated. Return value distribution for other variables (such as maximum crest elevation, C) whose distributional characteristics given $\mathcal{S}$ have been extensively studied can hence be estimated, as described in the online Supplementary Material (SM) accompanying this article.

Choice of threshold for extreme value modelling is problematic (e.g. Scarrott and MacDonald 2012), and in the authors' experience a major source of uncertainty in practical application of extreme value analysis. The novelty of the current work rests in avoiding a specific threshold selection. Instead we estimate an ensemble of non-stationary extreme value models, corresponding to different threshold choices, specified in terms of a plausible interval of non-exceedance probabilities for directional-seasonal quantile regression thresholds. We simulate under the ensemble to produce realisations of storm peak characteristics corresponding to any return period of interest. This article defines the *threshold ensemble* model (Section 3), evaluates it in application (Section 4) to estimation of return values for H_S^{sp}, H_S and C for a location in the South China Sea (Section 2), and presents careful uncertainty analysis of model parameters and return value distributions using the bias-corrected and accelerated (BCa) bootstrap scheme of Efron (1987). Section 4 also presents return value distributions estimated using an extended ensemble over bootstrap resamples of the original data. Section 5 provides a brief discussion, particularly of opportunities for further work.

2 Motivating application

The application sample includes hindcast time-series for sea state significant wave height H_S, (dominant) wave direction θ, season ϕ (defined as day of the year, for a notional year consisting of 360 days), for three hour sea states over the period August 1956 to July 2012.

The sample corresponds to a location in the Makassar Strait between the islands of Borneo and Sulawesi in Indonesia at a water depth of approximately 500m. The main climatic features are monsoonal: the southwest monsoon occurs between July and September and the northeast monsoon between December and March. At the location of interest, due to atmospheric circulation and topographical effects, the southwest monsoon actually generates increased H_S with waves propagating in a northwesterly direction during the northern summer. The northeast monsoon generates more severe sea states with waves propagating towards the east in the northern winter. Compared with other locations (for example in the northern North Sea or Gulf of Mexico), the location of interest is relatively benign. The largest value of H_S in the sample is approximately 3.5m.

Storm characteristics are isolated from these time-series using the procedure described in Ewans and Jonathan (2008). The resulting storm peak sample consists of 3095 values of H_S^{sp}. With direction *to which* a storm travels expressed in degrees clockwise with respect to north, Figure 1 shows plots of H_S^{sp} and H_S versus direction θ and season ϕ. Figure S1 (in online Supplementary Material, SM) shows empirical quantiles of H_S^{sp} by θ and ϕ.

For each storm, the *within-storm* time-series of H_S, θ, ϕ are referred to as the *intra-storm trajectory* for the storm. Intra-storm trajectories are essential for estimation of design values for intra-storm characteristics such as H_S and C. Figure 2 shows intra-storm trajectories of significant wave height, H_S, on wave direction θ for 30 randomly-chosen storm events (in different colours). Variability in the joint evolution of H_S and θ from storm to storm is clear.

3 Model

We estimate a non-stationary extreme value model for storm peak significant wave height H_S^{sp}, the parameters of which vary smoothly with direction θ and season ϕ. For a sample $\{\dot{z}_i\}_{i=1}^{\dot{n}}$ of $\dot{n}$ storm peak significant wave heights observed with directions $\{\dot{\theta}_i\}_{i=1}^{\dot{n}}$ and seasons $\{\dot{\phi}_i\}_{i=1}^{\dot{n}}$, we proceed as follows. We first estimate a set of threshold functions $\psi = \psi(\theta, \phi)$ above which observations $\dot{z}$ are assumed to be extreme using quantile regression. For each threshold in turn, we then estimate (a) the rate of occurrence ρ of threshold exceedance using a Poisson process model, and (b) estimate the distribution of threshold exceedances using a generalised Pareto (GP) model with shape parameter ξ and

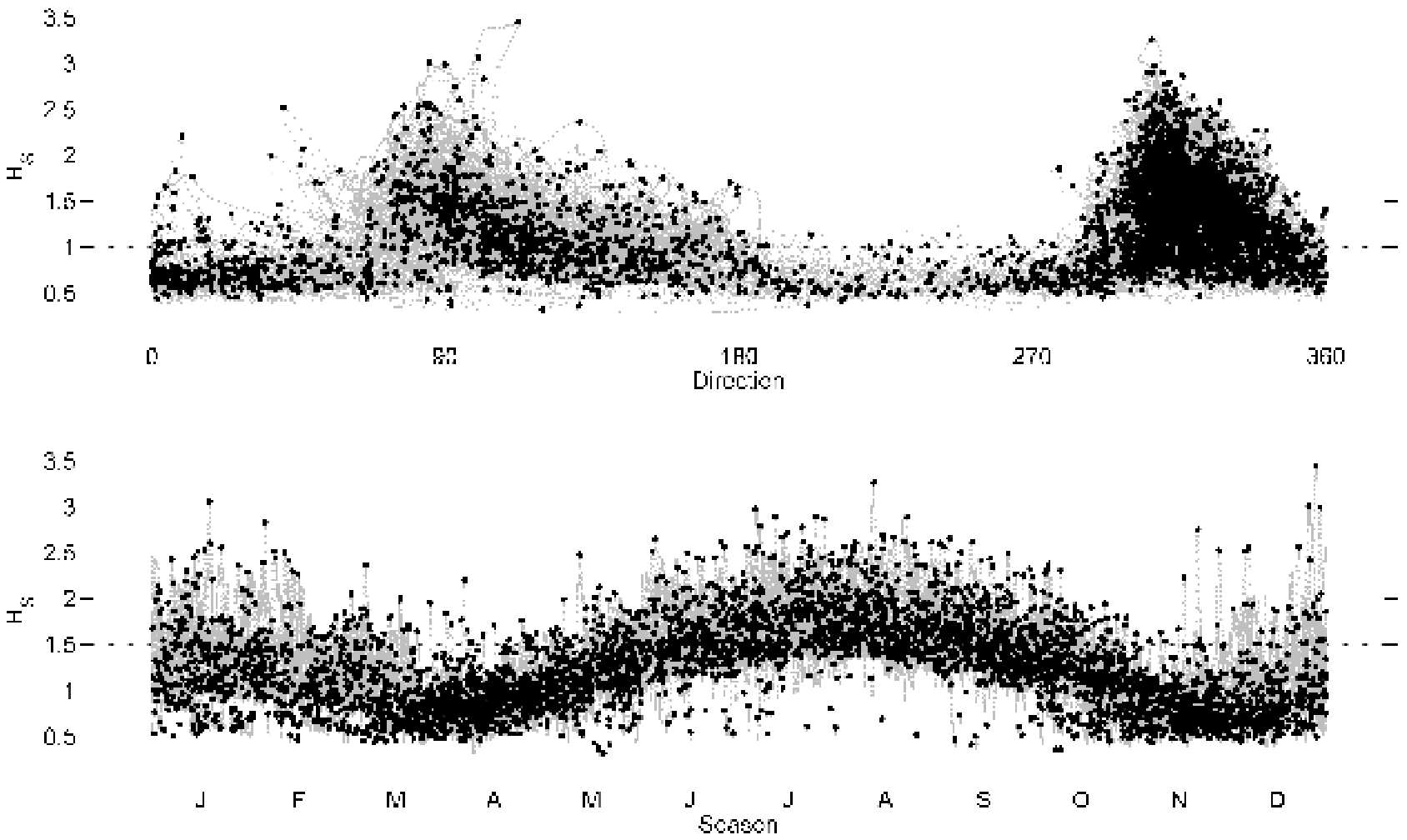

Figure 1: Storm peak significant wave height H_S^{sp} (black) on direction θ (upper panel) and season ϕ (lower panel). Also shown is sea state significant wave height H_S (grey) on direction θ (upper panel) and season ϕ (lower panel).

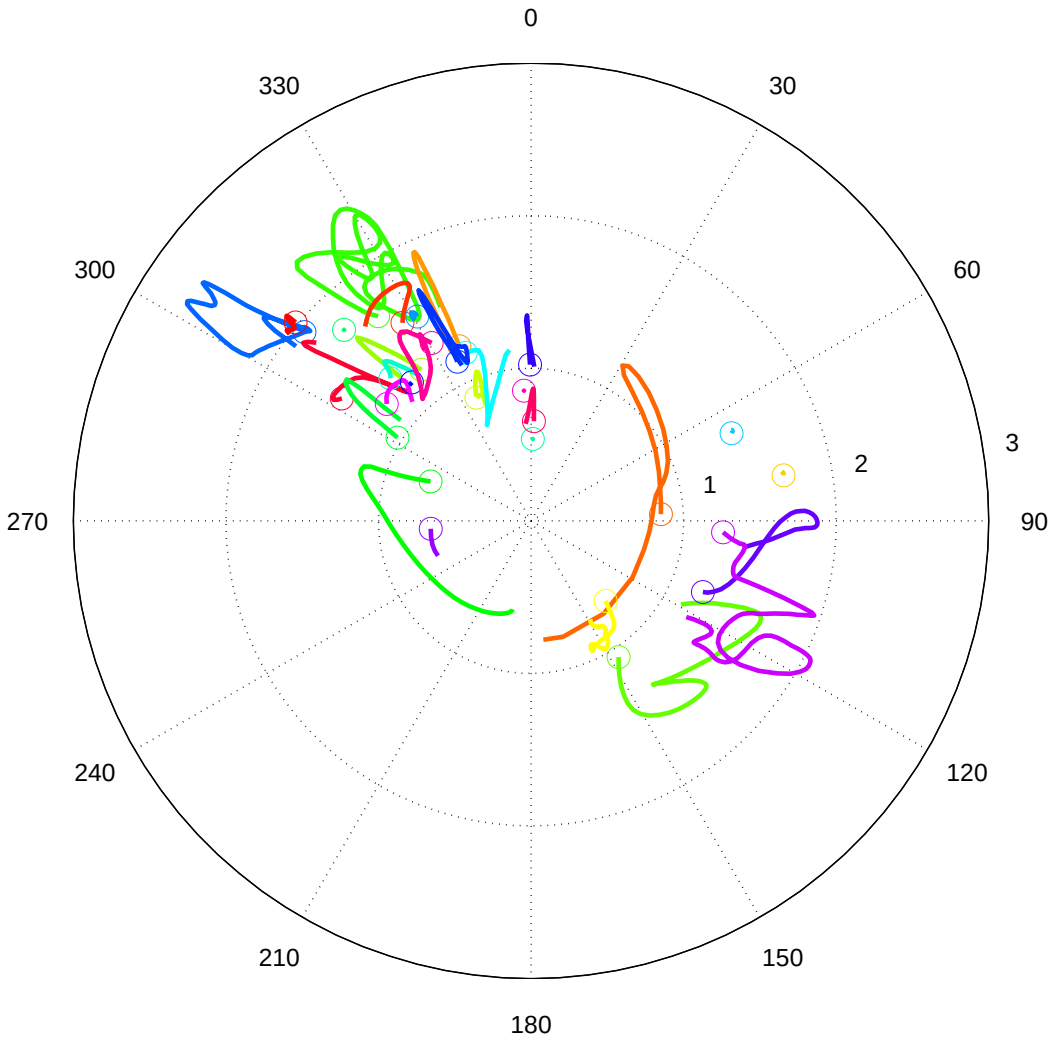

Figure 2: Intra-storm trajectories of significant wave height H_S on wave direction θ for 30 randomly-chosen storm events (in different colours). A circle marks the start of each intra-storm trajectory.

scale σ. All of ψ, ρ, ξ and σ are assumed smooth functions of covariates. This approach to extreme value modelling follows that of Chavez-Demoulin and Davison (2005).

3.1 Incorporating uncertainty in choice of extreme value threshold

Choice of threshold for extreme value analysis is generally problematic (e.g. Scarrott and MacDonald 2012). Here, we admit for return value inference a number of models, each corresponding to a different threshold choice. Inspection of diagnostic plots (see Section 4) might suggest that models corresponding to a certain interval J_τ of threshold non-exceedance probabilities τ yield plausible models. We then proceed to estimate an ensemble E_τ of n_τ models, each model corresponding to a systematically- (or alternatively, randomly-) chosen threshold non-exceedance probability $\tau_u, u = 1, 2, ..., n_\tau$ on J_τ. A *threshold ensemble* return value distribution of interest can then be estimated by repeated simulation under systematically- (or randomly-) chosen models from ensemble E_τ (see Section 3), as opposed to repeated simulation under a single model corresponding to a single extreme value threshold choice. In this work, we specify $n_\tau = 9$ threshold non-exceedance probabilities equally spaced on $J_\tau = [0.5, 0.9]$, and estimate extreme value models for each, obtaining sets of parameter estimates ψ_u, ρ_u, σ_u and ξ_u for $u = 1, 2, ..., n_\tau$. We confirmed that this choice of threshold non-exceedance probabilities provided stable inference. In the following outline of parameter estimation, threshold "u" subscripts are suppressed for conciseness unless necessary for clarity of explanation.

3.2 Parameter estimation

For quantile regression, for each threshold non-exceedance probability, we seek a smooth function ψ of covariates corresponding to non-exceedance probability τ of storm peak H_S for any combination of θ, 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 $r_i = z_i - \psi(\theta_i, \phi_i; \tau)$. We regulate the smoothness of the quantile function by penalising lack of fit for parameter roughness R_ψ (with respect to all covariates), by minimising the penalised criterion $\ell^*_\psi = \ell_\psi + \lambda_\psi R_\psi$ where the value of roughness coefficient λ_ψ is selected using leave-one-out cross-validation to provide good predictive performance. Roughness R (for ψ, ρ, ξ and σ) is discussed further below.

For Poisson modelling, we use penalised likelihood estimation. The rate ρ of threshold exceedance is estimated by minimising the roughness-penalised (negative log) likelihood $\ell^*_\rho = \ell_\rho + \lambda_\rho R_\rho$ where R_ρ is parameter roughness with respect to all covariates, λ_ρ is again evaluated using cross-validation, and Poisson (negative log) likelihood is given by

$$\ell_\rho = -\sum_{i=1}^{n} \log \rho(\theta_i, \phi_i) + \int \rho(\theta, \phi) d\theta d\phi \ .$$

The generalised Pareto model of size of threshold exceedance is estimated in a similar manner by minimising the roughness penalised (negative log) GP likelihood $\ell^*_{\xi,\sigma} = \ell_{\xi,\sigma} + \lambda_\xi R_\xi + \lambda_\sigma R_\sigma$ where R_ξ and R_σ are parameter roughnesses with respect to all covariates, λ_ξ and λ_σ are evaluated using cross-validation, and GP (negative log) likelihood is given by

$$\ell_{\xi,\sigma} = \sum_{i=1}^{n} \log \sigma(\theta_i, \phi_i) + (\frac{1}{\xi(\theta_i, \phi_i)} + 1) \log(1 + \frac{\xi(\theta_i, \phi_i)}{\sigma(\theta_i, \phi_i)}(z_i - \psi(\theta_i, \phi_i)))$$

for $\xi(\theta_i, \phi_i) \neq 0$, with a similar expression when $\xi(\theta_i, \phi_i) = 0$ (see Jonathan and Ewans 2013). In practice, we set $\lambda_\xi = \kappa \lambda_\sigma$ for pre-specified constant κ, so that only one cross-validation loop is necessary. The value of κ is estimated by inspection of the relative smoothness of ξ and σ with respect to covariates.

Return value simulation entails generation of samples corresponding to (a typically long) time period P. This is achieved by sampling under the estimated model for threshold exceedance, and by resampling the original data below the extreme value threshold. To perform the latter reliably, we find it useful to estimate an additional non-stationary Poisson model for rate of occurrence of threshold non-exceedances.

3.3 Parameter smoothness

Physical considerations suggest that we should expect model parameters ψ, ρ, ξ and σ to vary smoothly with respect to covariates θ, ϕ. This is achieved by expressing each parameter in terms of an appropriate basis for the domain D of covariates, where $D = D_\theta \times D_\phi$. $D_\theta = D_\phi = [0, 360)$ are the (marginal) domains of storm peak direction and season respectively. We calculate a periodic marginal B-spline basis matrix B_θ for an index set of 32 directional knots, and a periodic marginal B-spline basis matrix B_ϕ for an index set of 24 seasonal bins, yielding a total of $m(= 32 \times 24)$ combinations of covariate values. Then we define a basis matrix for the two dimensional domain D using Kronecker products of marginal basis matrices. Thus $B = B_\phi \otimes B_\theta$ provides an $(m \times p)$ basis matrix (where $m = 32 \times 24$ and $p = p_\theta p_\phi$) for modelling each of ψ, ρ, ξ and σ, any of which can be expressed in the form $B\beta$ for some $(p \times 1)$ vector of basis coefficients. Model estimation reduces to estimating appropriate sets of basis coefficients for each of ψ, ρ, ξ and σ. In this work, marginally for periodic directional and seasonal bases, we allocate one B-spline function to each covariate bin and co-locate their centres, so that $p_\theta = 32$ and $p_\phi = 24$, although this is not necessary in general. The roughness R of any function η can be evaluated on the index set (at which $\eta = B\beta$). Following the approach of Eilers and Marx (see, for example, Eilers and Marx 2010), we define roughness using $R = \beta' P \beta$ where P is a $p \times p$ penalty matrix. P is constructed such that R penalises the difference between marginally neighbouring values of spline coefficients, thereby penalising the overall lack of local spline smoothness (see Currie et al. 2006 for details). It is possible further to penalise directional roughness differently to seasonal, and to optimise the relative marginal penalty coefficient using cross-validation. In this work, we select and fix the relative marginal penalty coefficients for each of ψ, ρ, ξ and σ for a model using the non-exceedance probability $\tau = 0.7$.

3.4 Illustration of parameter estimates

We use median threshold ensemble estimates to illustrate model parameter variation with θ, ϕ. These estimates are medians over the n_τ estimates for different extreme value thresholds non-exceedance probabilities τ_u, adjusted as necessary to correspond to a specific non-exceedance probability $\tau_{\tilde{u}}$, since estimates of ρ_u and σ_u are threshold-dependent. Specifically

$$\tilde{\rho} \;=\; \underset{u}{\text{med}} \left\{ \rho_u \frac{\tau_{\tilde{u}}}{1 - \tau_u} \right\}, \qquad \tilde{\sigma} = \underset{u}{\text{med}} \{ \sigma_u + \xi_u (\psi_{\tilde{u}} - \psi_u) \}, \qquad \tilde{\xi} = \underset{u}{\text{med}} \{ \xi_u \} \ ,$$

where the reference non-exceedance probability is taken arbitrarily to be $\tau_{\tilde{u}} = 0.5$.

3.5 Uncertainty quantification

Bootstrap resampling is used for uncertainty quantification of the median threshold ensemble parameter estimates $\tilde{\rho}, \tilde{\sigma}$ and $\tilde{\xi}$. 95% bootstrap confidence intervals are estimated using the BCa approach of Efron (1987), by repeating the full extreme value analysis for 2000 resamples of the original storm peak sample, and estimating leave-one-out jack-knife parameters. In particular, estimation of optimal roughness penalties is performed independently for each bootstrap and jack-knife sample, so that uncertainty bands also reflect variability in these choices. It was also confirmed that 2000 resamples was sufficient to ensure stability of bootstrap confidence intervals. BCa confidence intervals are also provided in Section 4 for estimates of quantiles $\tilde{q}$ of return value distributions from the threshold ensemble model.

3.6 Return values

For a specified threshold non-exceedance probability τ, return values for variables of interest are estimated by simulation. Estimation of return value distributions in closed form is not possible since model parameter estimates are non-stationary with respect to covariates. First, we simulate sets of storm peak events (H_S^{sp}, and associated θ and ϕ per event) corresponding to any return period of P years of interest. By accumulating maximum values from multiple P-year realisations, potentially restricted to some directional-seasonal interval A of covariates, we estimate a return value distribution $Q|\tau$ for H_S^{sp}. For each storm peak event simulated, we can also simulate an intra-storm trajectory consisting of a time-series of sea state variables $\mathcal{S}$ per storm peak event using a matching procedure described in Feld et al. (2014). A return value distribution for sea state H_S, potentially restricted to covariate interval A, can therefore be accumulated. Hence we can further simulate a value of maximum crest elevation C as described in SM.

The distribution of *threshold ensemble* return value $\tilde{Q}$ for any variable of interest (H_S^{sp}, H_S and C here) can now be defined in terms of the distribution of the return value $Q|\tau$ corresponding to threshold with non-exceedance probability

τ as follows

$$\Pr(\tilde{Q} \leq x) \quad = \quad \int_{\tau \in J_\tau} \Pr(Q \leq x|\tau) dF(\tau) = \frac{1}{n_\tau} \sum_{u=1}^{n_\tau} \Pr(Q \leq x|\tau_u)$$

where we assume that the distribution $F(\tau)$ consists of point masses of $\frac{1}{n_\tau}$ at each of the n_τ non-exceedance probabilities in the threshold ensemble. The quantiles $\tilde{q}(p)$, $p \in [0, 1]$ of this distribution are solutions to the equation $\Pr(\tilde{Q} \leq x) = p$. In particular, $\tilde{q}(0.5)$ refers to the median threshold ensemble return value. BCa confidence intervals can be estimated for $\tilde{q}(p)$ in the usual way.

4 Application

4.1 Selection of suitable interval J_τ of threshold non-exceedance probability

A preliminary modelling study is undertaken to specify an appropriate interval J_τ of threshold non-exceedance probabilities τ for use in the threshold ensemble. For each of a suitable number of values for τ, we estimate a non-stationary extreme value model (as outlined in Section 3) corresponding to each of a large number of bootstrap resamples B of the original sample D. We simulate under the estimated model for each τ and B to estimate the 100-year storm peak event H_S^{sp}. We then examine the stability of the estimated 100-year event as a function of τ and B. We seek an interval of values for τ within which estimates of return value are relatively insensitive to change in τ. As can be seen in Figure S6, the return value estimates are relatively stable in $[0.5, 0.9]$, motivating this choice for J_τ.

4.2 Model parameters

To illustrate model parameter estimates, Figure 3(a) shows the extreme value threshold ψ corresponding to the reference non-exceedance probability $\tau_{\tilde{u}} = 0.5$ of H_S^{sp} as a function of direction θ and season ϕ. The southwest monsoon (July to September for directions around 300^o) and northeast monsoon (in December and January for directions around 90^o) are clear. Figure S2 shows plots of extreme value thresholds ψ corresponding to the set of n_τ non-exceedance probabilities of H_S^{sp}. Figure 3(b) shows directional-seasonal plots for the *median threshold ensemble* rate $\tilde{\rho}$ ($\times 1000$) of threshold exceedance of H_S^{sp}. Rate is largest by far for the southwest monsoon. Figure S3 also shows 12 monthly directional $\tilde{\rho}$ ($\times 1000$) estimates (solid) and 95% BCa confidence intervals (dashed). Figures 3(c) and 3(d) show median threshold ensemble generalised Pareto scale $\tilde{\sigma}$ and shape $\tilde{\xi}$. $\tilde{\sigma}$ shows local maxima corresponding to the northeast monsoon (in particular) and the southwest monsoon. $\tilde{\xi}$ exhibits predominantly seasonal variation, with lower values during the northern summer. Values of $\tilde{\xi} > 0$ suggest that the distribution of H_S^{sp} is unbounded above (e.g. Jonathan and Ewans 2013). Figures S4 and S5 also show the corresponding 12 monthly directional estimates (solid) and 95% BCa confidence intervals (dashed).

4.3 Model diagnostics

We use the return value simulation described Section 3 to assess model adequacy, by comparing the characteristics of realisations under the model with those of the original sample, possibly restricted to some covariate interval A. Here, for illustration, we compare the distribution of sea state H_S from simulation (incorporating intra-storm evolution) and original data. We simulate a set of sea state H_S events corresponding exactly to the period of the original sample under the threshold ensemble model. We estimate the quantiles $\tilde{q}(p)$, $p \in [0, 1]$ of the empirical cumulative distribution function of H_S from the set. Then we estimate the 2.5%, median and 97.5% percentiles of $\tilde{q}(p)$ as a function of p; 95% BCa confidence intervals for these percentiles are illustrated in black in Figure 4, as a logarithmic tail to emphasise tail behaviour. We then estimate the quantiles $q_0(p)$ of sea state H_S directly from the original sample. Using 1000 bootstrap resamples of the original sample, we also estimate an empirical 95% bootstrap uncertainty band for $q_0(p)$ as a function of p. This is illustrated in red in the figure. Although there is some tendency for the model to overestimate sea state H_S for the very largest events, agreement is good for omni-directional omni-seasonal and monthly omni-directional estimates shown. Comparison of omni-seasonal estimates per directional octant (not shown) also indicate good agreement between the original sample and simulation. Analogous diagnostics plots for storm peak H_S show good agreement also.

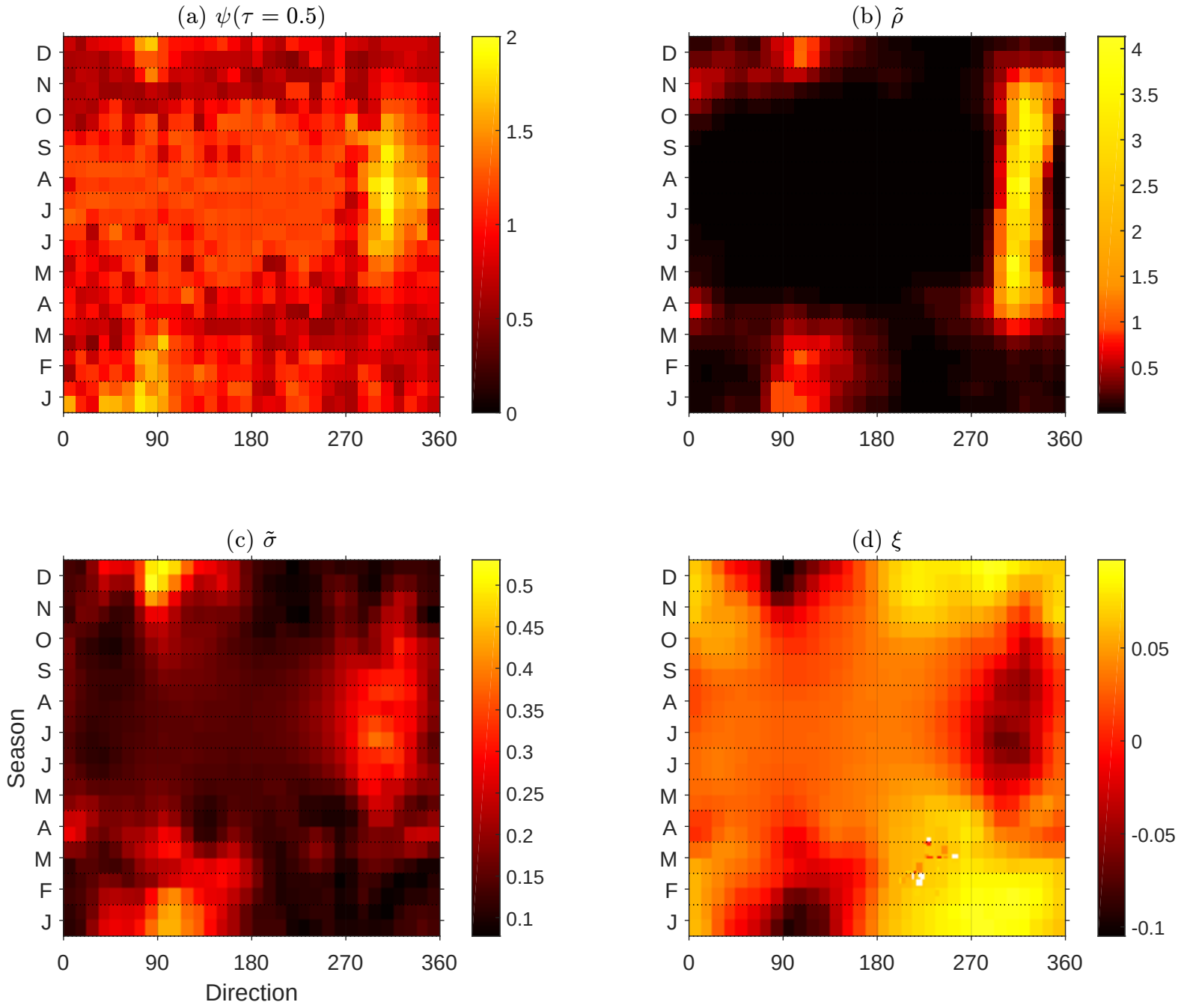

Figure 3: Directional-seasonal plots for extreme value model parameters on on θ and ϕ. (a) Extreme value threshold, ψ, for $\tau = 0.5$, (b) median threshold ensemble rate $\tilde{\rho}(\times 1000)$ of threshold exceedance, (c) median threshold ensemble generalised Pareto scale, $\tilde{\sigma}$, and (d) median over threshold generalised Pareto shape, $\tilde{\xi}$.

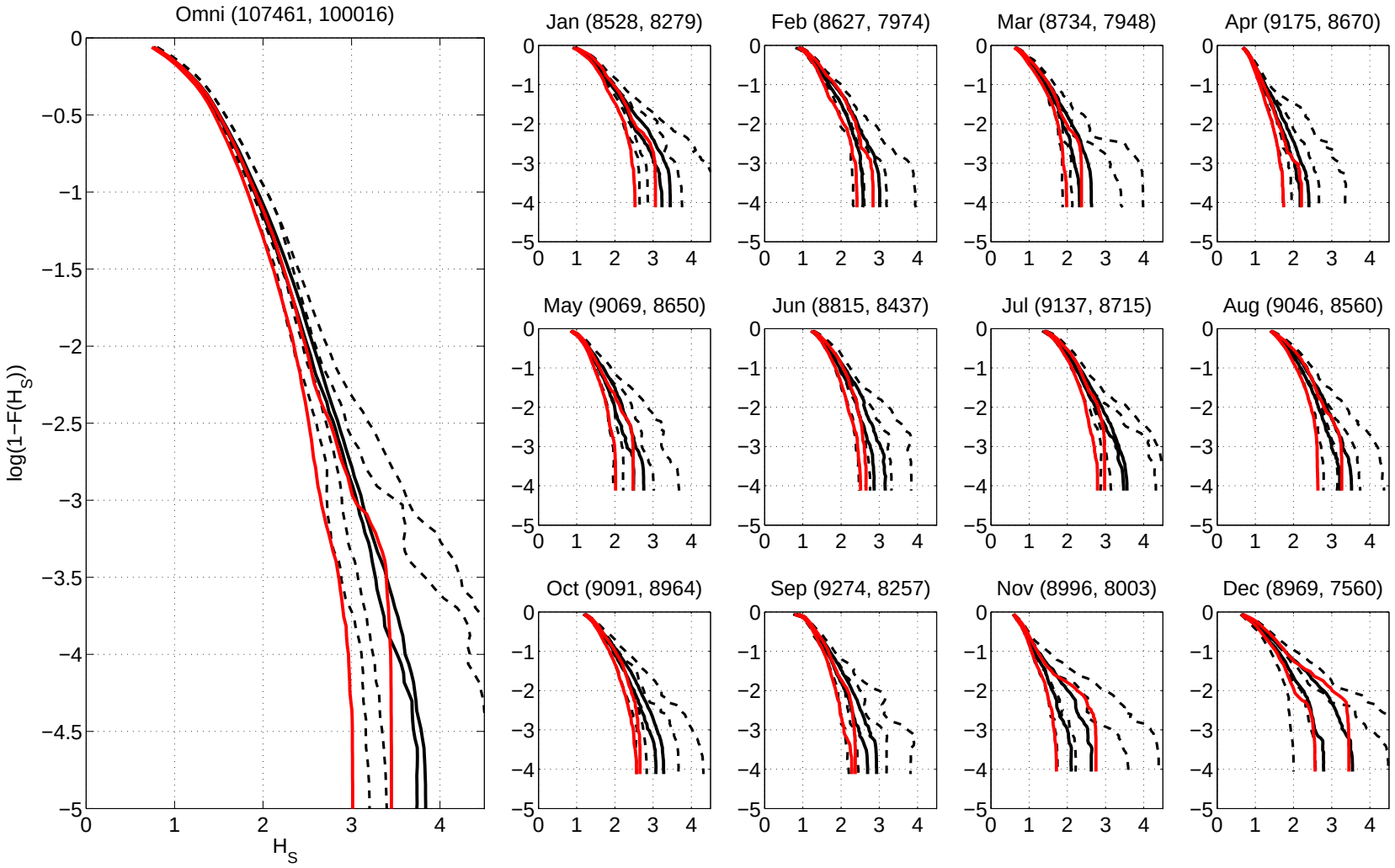

Figure 4: Illustration of validation of return value estimation for sea state H_S. The red curves represent an empirical 95% bootstrap uncertainty band for the quantile function $q_0(p)$ of the original sample. The solid black curves represent a BCa 95% confidence band for the median (for each given probability $p = F(H_S)$) of the threshold ensemble quantile function $\tilde{q}(p)$ estimated under the threshold ensemble model. The dashed black lines show the corresponding BCa confidence band for the 2.5% and 97.5% percentiles of $\tilde{q}(p)$. The left hand plot corresponds to the omni-directional, omni-seasonal case, and the 12 right hand panels to monthly omni-directional estimates. Titles for plots, in brackets, are the numbers of actual and simulated events.

4.4 Return values

Figure 5 shows directional and seasonal variability of the median threshold ensemble estimate $\tilde{q}(0.5)$ for H_S and a return period of 100 years, estimated by return value simulation. Monsoonal effects are prominent. Figure S7 shows the corresponding estimates for 100-year crest elevation C, which is qualitatively very similar to Figure 4.

4.5 Bootstrap threshold ensemble return value distributions

In the spirit of the work of Davison (1986), Hall et al. (1999) and Fushiki et al. (2005) we define the *bootstrap threshold ensemble* return value $\breve{Q}$ in a similar fashion to $\tilde{Q}$. First we estimate the return value distribution $Q|B,\tau$ by simulation as outlined in Section 4 for each pair of threshold non-exceedance probability $\tau \in J_\tau$ and bootstrap resample $B \in \{B_b\}_{b=1}^{n_B} = \mathcal{B}$ of the original sample D. Then we define the cumulative distribution of $\breve{Q}$ as

$$\Pr(\breve{Q} \leq x) \quad = \quad \int_{\tau \in J_\tau} \int_{B \in \mathcal{B}} \Pr(Q \leq x|B,\tau) dF(B) dF(\tau) = \frac{1}{n_\tau} \frac{1}{n_B} \sum_{u=1}^{n_\tau} \sum_{b=1}^{n_B} \Pr(Q \leq x|B_b, \tau_u)$$

where the distributions $F(\tau)$ and $F(B)$ are discrete with point masses of size $\frac{1}{n_\tau}$ and $\frac{1}{n_B}$ for each $\tau \in J_\tau$ and $B \in \mathcal{B}$ respectively. The quantiles $\breve{q}(p)$ of $\breve{Q}$ for $p \in [0,1]$ are defined in the usual way.

The bootstrap threshold ensemble propagates both sample uncertainty and uncertainty in threshold specification into return values $\breve{Q}$. In contrast, the threshold ensemble encapsulates threshold uncertainty only; BCa confidence intervals on $\tilde{Q}$ represent sample uncertainty. Figure 6 compares threshold ensemble estimates $\tilde{Q}$ (with 95% BCa confidence intervals for its quantiles) and bootstrap threshold ensemble estimates $\breve{Q}$ for the 100-year return value at Makassar. Since $\breve{Q}$ incorporates an extra source of uncertainty compared to $\tilde{Q}$, its estimated distribution is wider than that of the sample estimate of $\tilde{Q}$ (shown with solid lines in the upper plots). However, the confidence band for $\tilde{Q}$ (which accounts for sampling uncertainty, shown with dashed lines in the upper plots) covers a wider interval of H_S than the corresponding $\breve{Q}$ in general.

5 Discussion

We present a methodology for estimation of return value distributions for ocean wave characteristics in environments which exhibit directional and seasonal non-stationarity. This allows estimation of return values for storm peak, sea state and individual wave characteristics given arbitrary sets of directional and seasonal covariates in a straightforward but rigorous manner without the need for specific extreme value threshold selection. Uncertainties are quantified by bootstrap resampling over the whole inference procedure.

From an ocean engineering perspective, there are a number of refinements possible to the methodology presented here, to improve its applicability: a) For some locations, rapid directional or seasonal changes in extreme ocean wave characteristics might be expected on physical grounds; in such cases, it is desirable that the model allows greater non-stationary for part of the covariate domain. This might be achieved by using a non-uniform spline roughness penalisation. b) For some covariate intervals, the rate of occurrence of exceedances of extreme value threshold is low or zero; in such cases, the extreme value model is not informative for return values, and other approaches to return value inference must be adopted. In the current work, we simply resample non-exceedances of the extreme value threshold. c) The effect of intra-storm variability in return value simulation is currently incorporated by an empirically matching procedure. A statistical model for intra-storm variability would be a valuable contribution to the literature, but a challenging task. d) The bootstrap scheme used here is computationally demanding, since it encompasses all analysis steps. Bootstrapping lends itself naturally to parallel computation. In the current work we use a simple parallel scheme across 50 processors, but we could do better (e.g. see Raghupathi et al. 2015).

Adequate incorporation on non-stationarity in extreme value models requires realistic methods for modelling non-stationarity of extreme value model parameters. We find that penalised B-splines provide good bases for covariate modelling, since they are relatively low-dimensional in terms of the number of parameters to be estimated (compared, for example, to Gaussian process models). However, other approaches including Gaussian processes and Markov random fields have their merits.

We find that cross-validation in general provides the most reliable method for estimation of optimal roughness penalisation. AIC is suitable in some situations, but difficult to apply well in others; for example, in spatial applications where observations of storm severities at neighbouring locations are not independent - but the extent of dependence is not known. Spatial block cross-validation provides a straightforward approach to roughness estimation. Methods

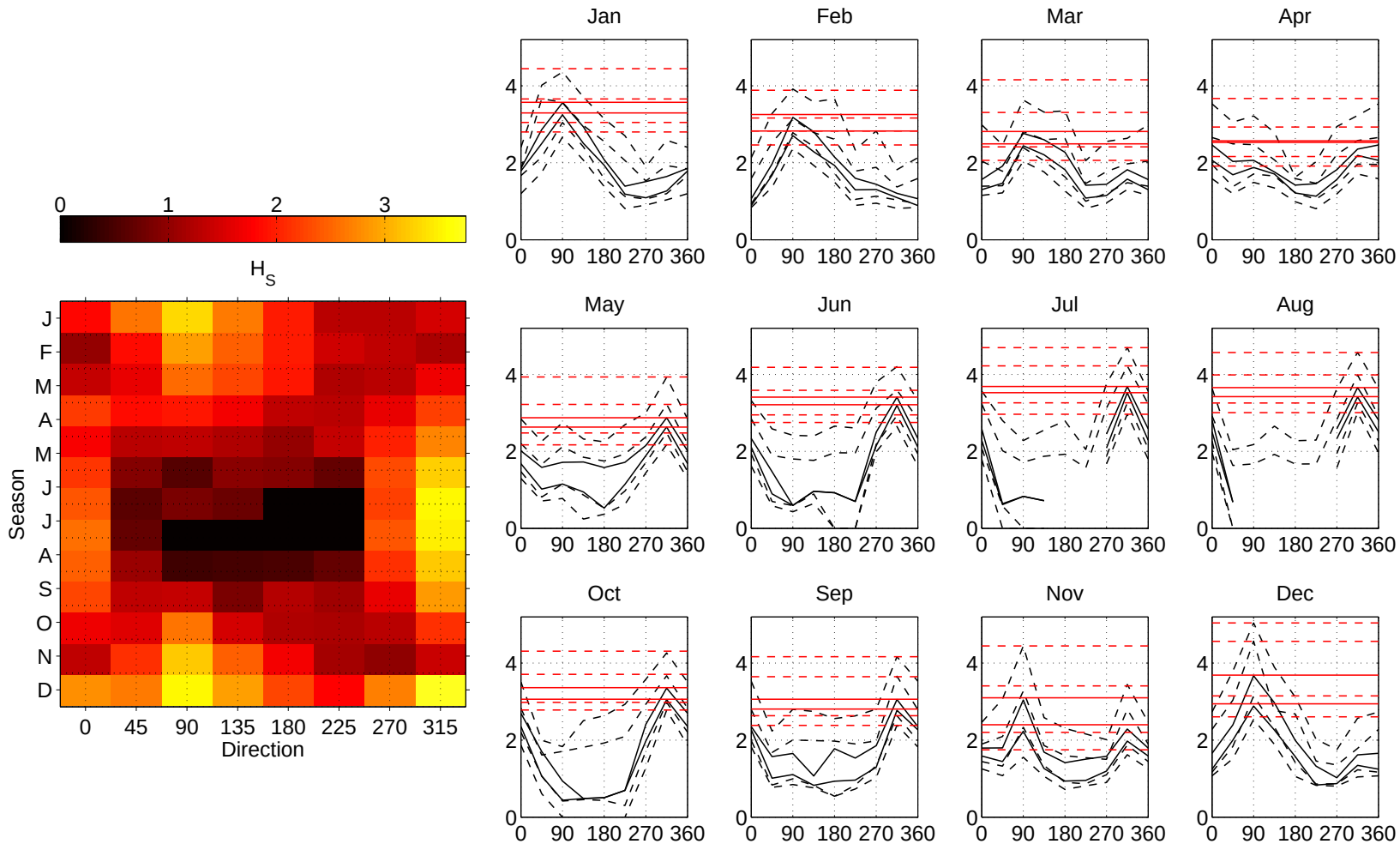

Figure 5: Directional-seasonal return value plot for 100-year significant wave height (in metres). The left-hand panel shows directional and seasonal variability of the median threshold ensemble estimate $\tilde{q}(0.5)$ for H_S. The right hand panel shows 12 monthly directional octant return values (in black) in terms of BCa 95% confidence limits for $\tilde{q}(0.5)$ (solid), $\tilde{q}(0.025)$ (dashed) and $\tilde{q}(0.975)$ (dashed). Also shown are the corresponding omni-directional estimates (in red).

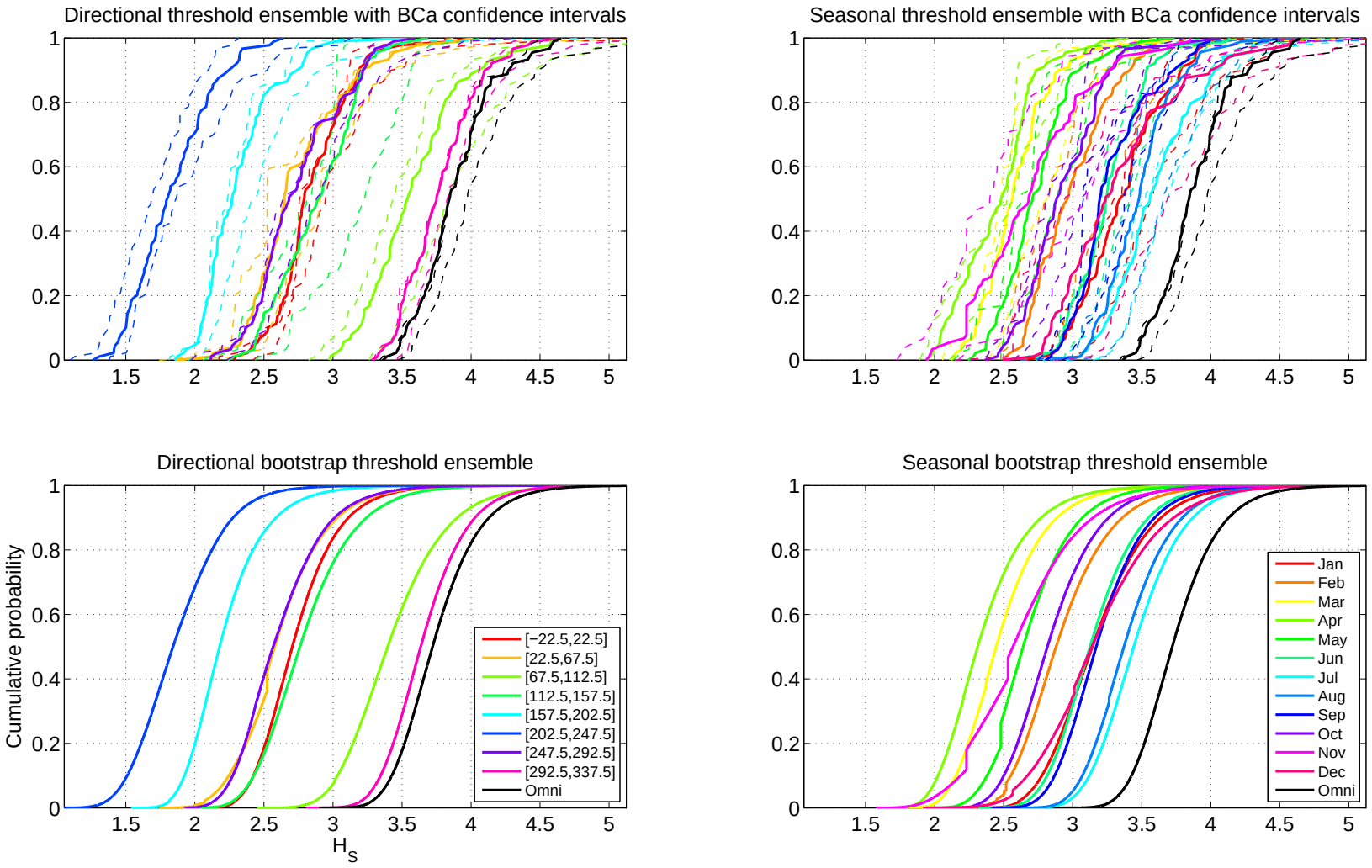

Figure 6: Empirical cumulative distribution functions (cdfs) for 100-year significant wave height from simulation under the directional-seasonal model. Left hand and right hand panels show directional and seasonal cdfs respectively. Upper panels shows threshold ensemble estimates $\tilde{Q}$ with 95% BCa confidence intervals, and lower panels bootstrap threshold ensemble estimates $\breve{Q}$.

employing adjusted likelihoods to accommodate spatial dependence provide an alternative (e.g. Chandler and Bate 2007). Applying cross-validation well is more problematic for likelihoods with bounded domains (such as the generalised Pareto with negative shape parameter); Northrop et al. (2015) discuss possible modifications of the standard cross-validation scheme.

The methodology outlined here is applicable to spatial extremes (e.g. Davison et al. (2012)) and multivariate extremes (e.g. using the conditional extremes model of Heffernan and Tawn (2004) and developments thereof). It seems inevitable and desirable that model estimation be eventually performed within a Bayesian framework (Lang and Brezger 2004).

Additional information and supplementary material for this article is available online at the journal's website.

Acknowledgement

We acknowledge useful discussions on computational aspects with Laks Raghupathi of Shell, Bangalore, and the thoughtful comments of two reviewers.

References

Chandler, R. E., Bate, S., 2007. Inference for clustered data using the independence loglikelihood. Biometrika 94, 167–183.

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.

Davison, A. C., 1986. Approximate predictive likelihood. Biometrika 73, 323–32.

Davison, A. C., Padoan, S. A., Ribatet, M., 2012. Statistical modelling of spatial extremes. Statist. Sci. 27, 161–186.

Efron, B., 1987. Better bootstrap confidence intervals. J. Am. Statist. Soc. 82, 171–185.

Eilers, P. H. C., Marx, B. D., 2010. Splines, knots and penalties. Wiley Interscience Reviews: Computational Statistics 2, 637–653.

Ewans, K. C., Jonathan, P., 2008. The effect of directionality on northern North Sea extreme wave design criteria. J. Offshore. Arct. Eng. 130, 10.

Feld, G., Randell, D., Wu, Y., Ewans, K., Jonathan, P., 2014. Estimation of storm peak and intra-storm directional-seasonal design conditions in the North Sea. Proceedings of 33nd International Conference on Ocean, Offshore and Arctic Engineering OMAE2014-23157, and accepted for publication in J. Offshore. Arct. Eng.

Fushiki, T., Komaki, F., Aihara, K., 2005. Nonparametric bootstrap prediction. Bernoulli 11, 293–307.

Hall, P., Peng, L., Tajvidi, N., 1999. On prediction intervals based on predictive likelihood or bootstrap methods. Biometrika 86, 871–80.

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

ISO19901-1, 2005. Petroleum and natural gas industries. Specific requirements for offshore structures. Part 1: Metocean design and operating considerations. International Standards Organisation.

Jonathan, P., Ewans, K. C., 2013. Statistical modelling of extreme ocean environments with implications for marine design : a review. Ocean Eng. 62, 91–109.

Jonathan, P., Ewans, K. C., Forristall, G. Z., 2008. Statistical estimation of extreme ocean environments: The requirement for modelling directionality and other covariate effects. Ocean Eng. 35, 1211–1225.

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

Mackay, E. B. L., Challenor, P. G., Bahaj, A. S., 2010. On the use of discrete seasonal and directional models for the estimation of extreme wave conditions. Ocean Eng. 37, 425–442.

Mendez, F. J., Menendez, M., Luceno, A., Medina, R., Graham, N. E., 2008. Seasonality and duration in extreme value distributions of significant wave height. Ocean Eng. 35, 131–138.

Northrop, P., Attalides, N., Jonathan, P., 2015. Cross-validatory extreme value threshold selection and uncertainty with application to wave heights. (In preparation, draft at www.lancs.ac.uk/∼jonathan).

Raghupathi, L., Randell, D., Ewans, K., Jonathan, P., 2015. Fast computation of large scale marginal extremes with multi-dimensional covariates. (Submitted to Comp. Stat. Data Anal., draft at www.lancs.ac.uk/∼jonathan).

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

Supplementary Material

Distributions of return values for ocean wave characteristics in the South China Sea using directional-seasonal extreme value analysis

by

D. Randell, G. Feld, K. Ewans and P. Jonathan

Please address all correspondence to *philip.jonathan@shell.com* or *p.jonathan@lancaster.ac.uk*

Distribution of maximum crest elevation in a sea state

The distributions of maximum wave height and maximum crest elevation for a stationary sea state with parameters $\mathcal{S}$ have been extensively studied (e.g. Forristall 1978 and Forristall 2000). Given a simulation of realisations of sea states $\mathcal{S}$ corresponding to a return period of interest, we can therefore also estimate directional-seasonal return value distributions for these and other wave properties, required for structural design. In this work, we focus on estimation of maximum crest elevation C, the cumulative distribution function of which, given sea state $\mathcal{S}$, has been estimated to have the Weibull form

$$\Pr(C \leq \eta|\mathcal{S}) = (1 - \exp(-\frac{\eta}{\alpha_{\mathcal{S}} H_S})^{\beta_{\mathcal{S}}})^{n_{\mathcal{S}}}$$

where all of the Weibull parameters $\alpha_{\mathcal{S}}$, $\beta_{\mathcal{S}}$ and $n_{\mathcal{S}}$ are functions of the sea state parameters $\mathcal{S}$, the forms of which are motivated by second-order wave theory, the parameters of which were estimated from historical measurements (see Forristall 2000 for details).

Figures

Figure S1 illustrates empirical quantiles of storm peak significant wave height H_S^{sp} by direction θ and season ϕ, for threshold non-exceedance probabilities τ as listed. Figure S2 shows plots of extreme value thresholds ψ corresponding to the set of n_τ non-exceedance probabilities of H_S^{sp}. Each panel shows threshold ψ on direction θ and season ϕ. The southwest monsoon (in June and July for directions around 300^o) and northeast monsoon (in December and January for directions around 90^o) are clear, especially for larger thresholds.

Figure S3 shows directional-seasonal plots for the *median threshold ensemble* rate $\tilde{\rho}$ of threshold exceedance of H_S^{sp}. The left-hand panel shows $\tilde{\rho}$ ($\times 1000$) on θ and ϕ. The right hand panel shows 12 monthly directional $\tilde{\rho}$ ($\times 1000$) estimates (solid) and 95% BCa confidence intervals (dashed). The rate of occurrence is largest by far for the southwest monsoon.

Figure S4 shows plots of median threshold ensemble generalised Pareto scale $\tilde{\sigma}$. The left-hand panel shows $\tilde{\sigma}$ on θ and ϕ. The right hand panel shows 12 monthly directional shapes in terms of bootstrap median (solid) and 95% bootstrap uncertainty bands (dashed). The corresponding plots for generalised Pareto shape $\tilde{\xi}$ are given in Figure S5. $\tilde{\sigma}$ shows local maxima corresponding to the northeast monsoon (in particular) and the southwest monsoon. $\tilde{\xi}$ exhibits predominantly seasonal variation, with lower values during the northern summer.

Figure S6 illustrates a preliminary modelling study undertaken to specify an appropriate interval J_τ of threshold non-exceedance probabilities τ for use in the threshold ensemble. For each of a suitable number of values for τ, we estimate a non-stationary extreme value model (as outlined in Section 3) corresponding to each of a large number of bootstrap resamples B of the original sample D. We simulate under the estimated model for each τ and B to estimate the 100-year storm peak event H_S^{sp}. We then examine the stability of the estimate of the 100-year event as a function of τ and B. We seek an interval of values for τ within which estimates of return value are relatively insensitive to change in τ. As can be seen in the figure, the return value estimates are relatively stable in $[0.5, 0.9]$, motivating this choice for J_τ.

Figure S7 shows directional and seasonal variability of the median threshold ensemble estimate $\tilde{q}(0.5)$ for maximum crest elevation C and a return period of 100 years. Monsoonal effects are prominent. Figure S7 is qualitatively very similar to Figure 4 of the main text.

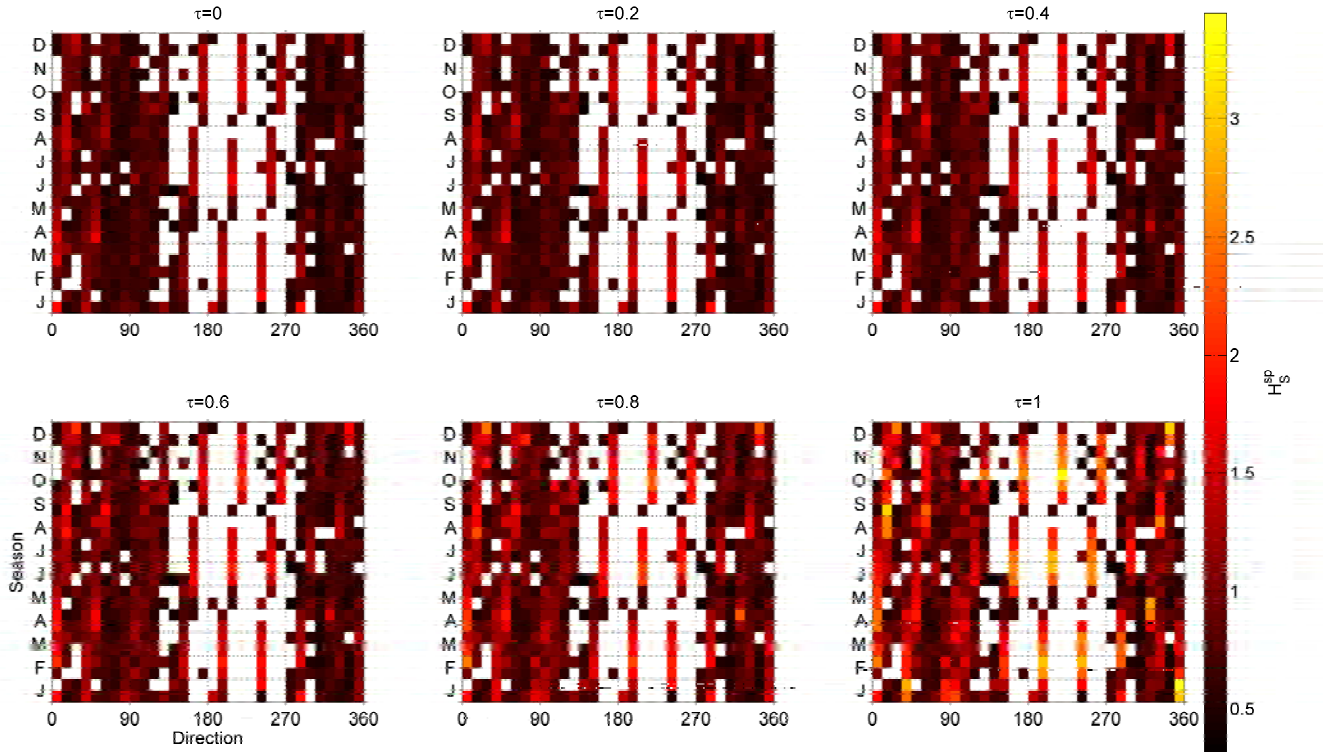

Figure S1. Empirical quantiles of storm peak significant wave height H_S^{sp} by direction θ and season ϕ, for threshold non-exceedance probabilities τ as listed. Empty bins are coloured white.

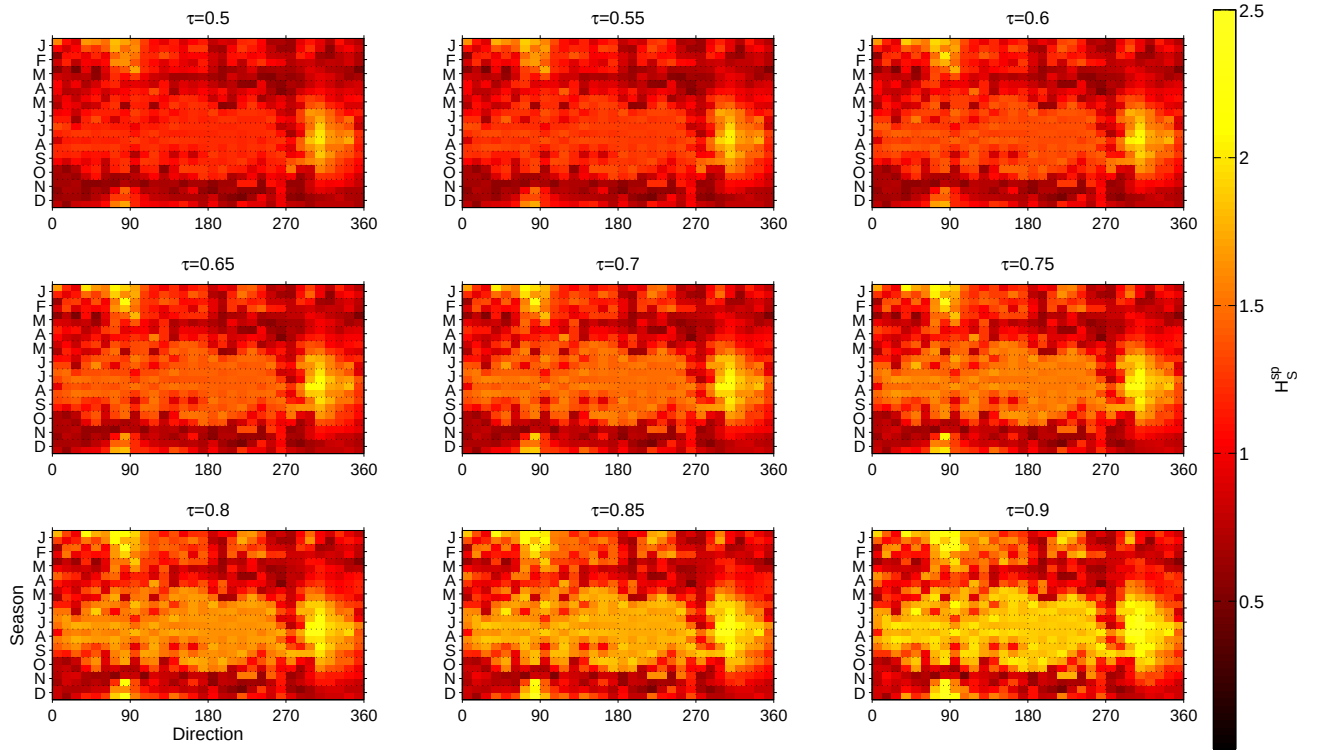

Figure S2. Directional-seasonal plots for extreme value thresholds, ψ, corresponding to equally spaced non-exceedance probabilities of H_S^{sp} on $[0.5, 0.9]$ (left to right, then top to bottom).

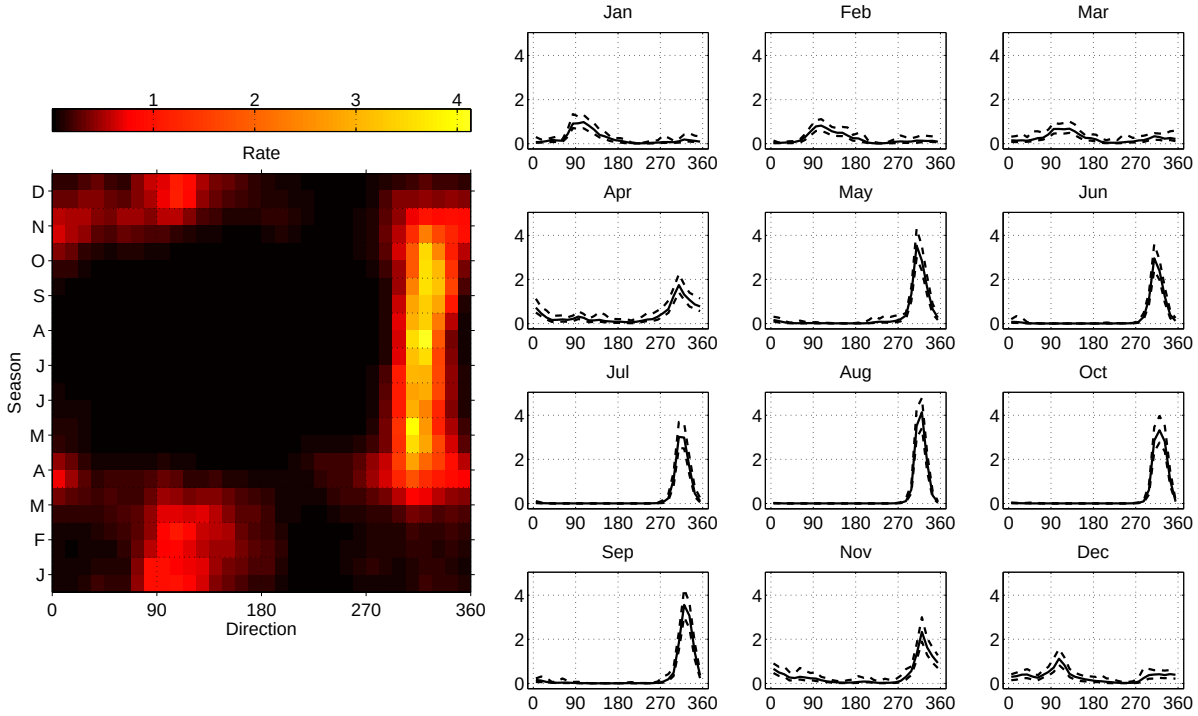

Figure S3. Directional-seasonal plot for median threshold ensemble rate $\tilde{\rho}(\times 1000)$ of threshold exceedance of H_S^{sp}. The left-hand panel shows $\tilde{\rho}(\times 1000)$ on θ and ϕ. The right hand panel shows 12 monthly directional estimates with 95% BCa bootstrap confidence intervals (dashed).

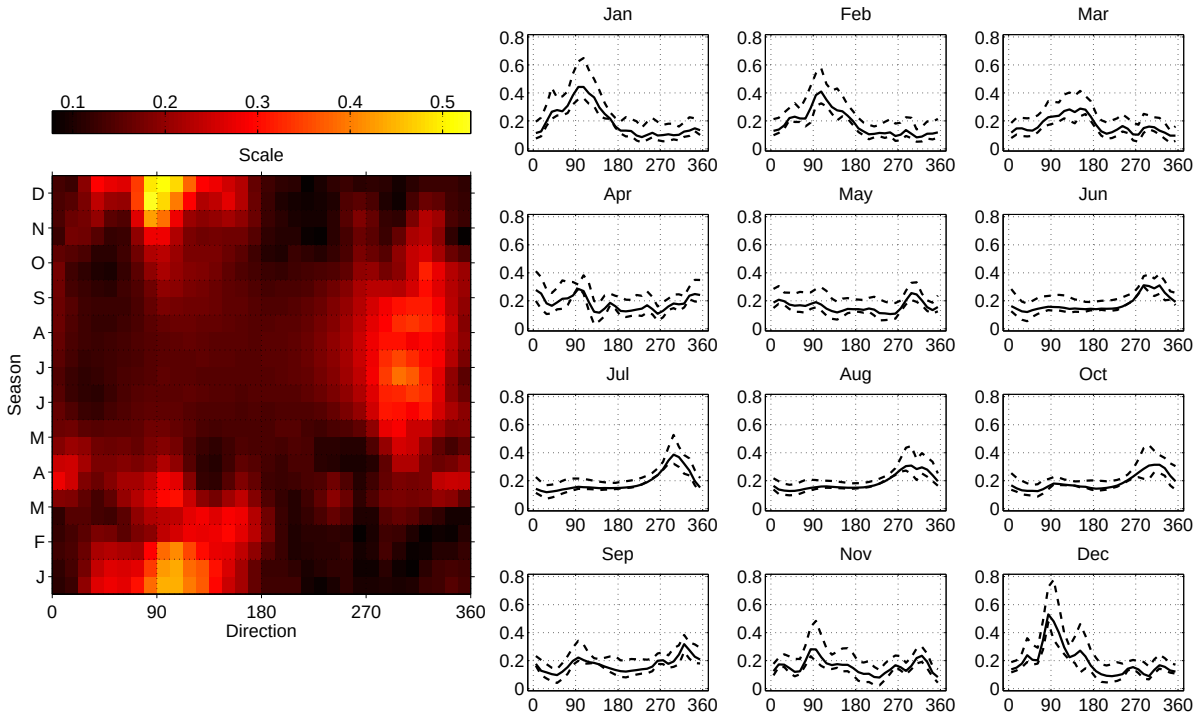

Figure S4. Directional-seasonal plot for median threshold ensemble generalised Pareto scale, $\tilde{\sigma}$. The left-hand panel shows $\tilde{\sigma}$ on θ and ϕ. The right hand panel shows 12 monthly directional estimates with 95% BCa bootstrap confidence intervals (dashed).

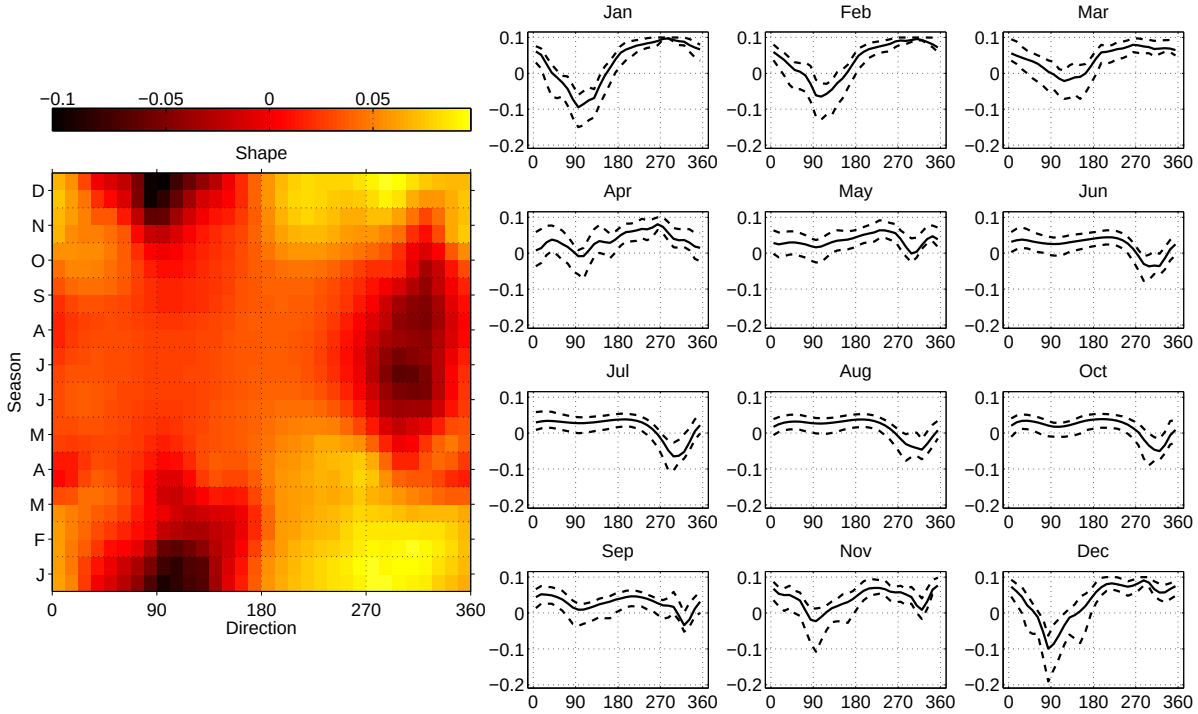

Figure S5. Directional-seasonal plot for median over threshold generalised Pareto shape, $\tilde{\xi}$. The left-hand panel shows $\tilde{\xi}$ on θ and ϕ. The right hand panel shows 12 monthly directional estimates with 95% BCa bootstrap confidence intervals (dashed).

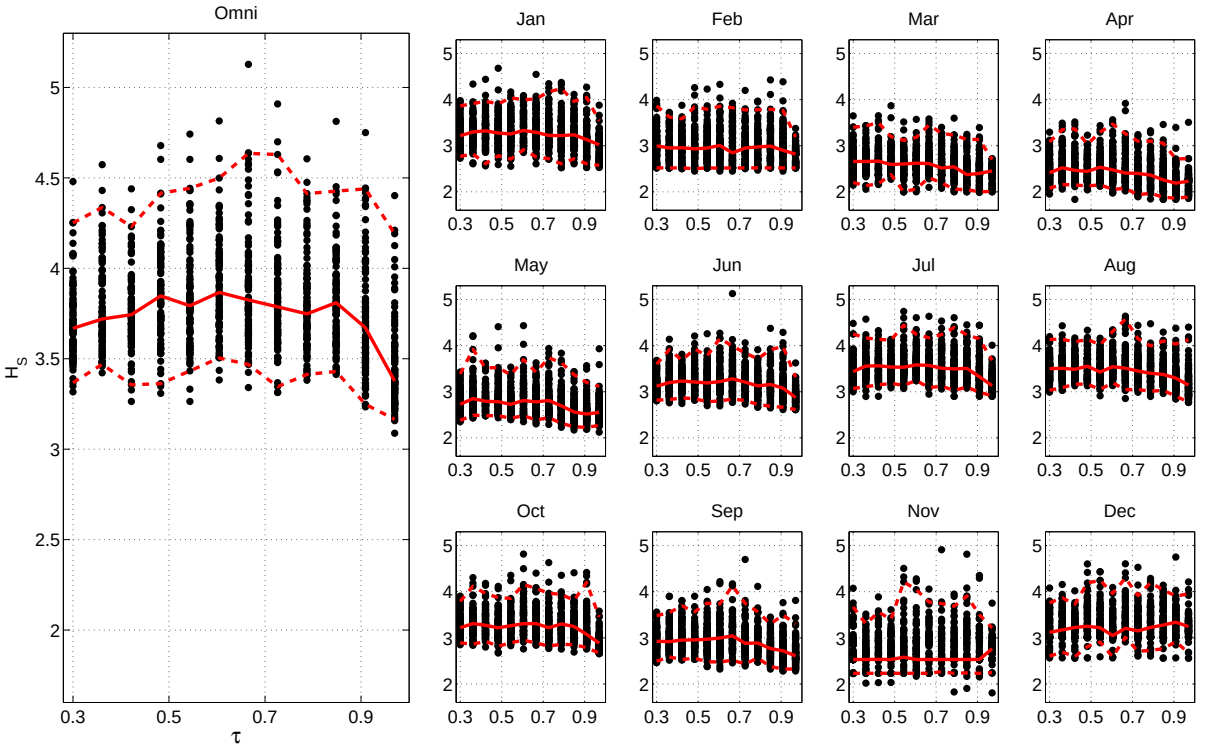

Figure S6. Estimates for 100-year maximum for H_S^{sp} from simulation under models corresponding to 100 bootstrap resamples for each of 12 choices of threshold non-exceedance probability, τ. Median estimates are connected by a solid red line. 2.5% and 97.5% iles are connected by dashed red lines. The left hand panel shows the omni-directional, omni-seasonal estimate. The right hand panels show 12 monthly omni-directional estimates.

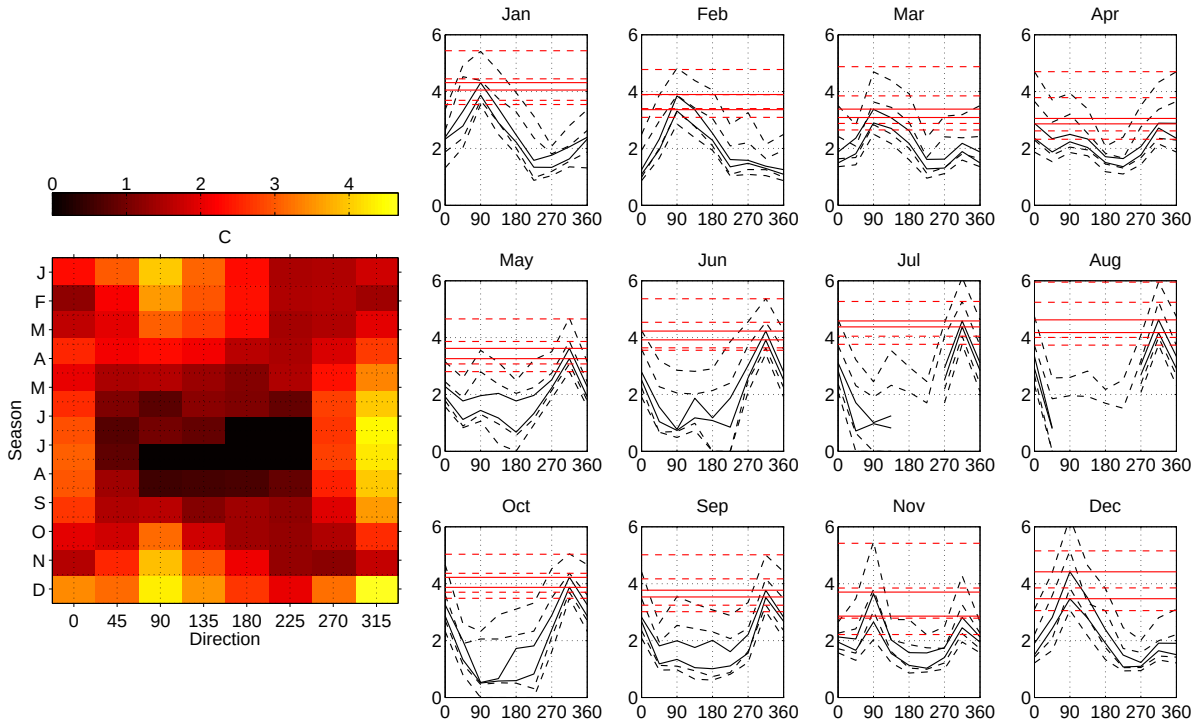

Figure S7. Directional-seasonal return value plot for 100-year crest elevation (in metres). The left-hand panel shows directional and seasonal variability of the median quantile over threshold $\tilde{q}(0.5)$ for C. The right hand panel shows 12 monthly directional octant return values (in black) in terms of BCa 95% confidence limits for $\tilde{q}(0.5)$ (solid), $\tilde{q}(0.025)$ (dashed) and $\tilde{q}(0.975)$ (dashed). Also shown are the corresponding omni-directional estimates (in red).

References

G. Z. Forristall. On the statistical distribution of wave heights in a storm. *J. Geophysical Research*, 83:2353–2358, 1978.

G. Z. Forristall. Wave crest distributions: Observations and second-order theory. *Journal of Physical Oceanography*, 30: 1931–1943, 2000.