Accepted for publication in J OMAE in January 2015

Estimation of storm peak and intra-storm directional-seasonal design conditions in the North Sea

Graham Feld
Shell Projects & Technology
Aberdeen AB12 3FY
United Kingdom
graham.feld@shell.com

David Randell
Shell Projects & Technology
Manchester M22 0RR
United Kingdom
david.randell@shell.com

Yanyun Wu
Shell Projects & Technology
Manchester M22 0RR
United Kingdom
yanyun.wu@shell.com

Kevin Ewans
Sarawak Shell Bhd
50450 Kuala Lumpur
Malaysia
kevin.ewans@shell.com

Philip Jonathan
Shell Projects & Technology
Manchester M22 0RR
United Kingdom
philip.jonathan@shell.com

ABSTRACT

Specification of realistic environmental design conditions for marine structures is of fundamental importance to their reliability over time. Design conditions for extreme waves and storm severities are typically estimated by extreme value analysis of time series of measured or hindcast significant wave height, H_S. This analysis is complicated by two effects. Firstly, H_S exhibits temporal dependence. Secondly, the characteristics of H_S^{sp} are non-stationary with respect to multiple covariates, particularly wave direction and season.

We develop directional-seasonal design values for storm peak significant wave height (H_S^{sp}) by estimation of, and simulation under a non-stationary extreme value model for H_S^{sp}. Design values for significant wave height (H_S) are estimated by simulating storm trajectories of H_S consistent with the simulated storm peak events. Design distributions for individual maximum wave height (H_{max}) are estimated by marginalisation using the known conditional distribution for H_{max} given H_S. Particular attention is paid to the assessment of model bias and quantification of model parameter and design value uncertainty using bootstrap resampling. We also outline existing work on extension to estimation of maximum crest elevation and total extreme water level.

1 Introduction

Specification of realistic environmental design conditions for marine structures is of fundamental importance to their reliability over time. Design conditions for extreme waves and storm severities are typically estimated by extreme value analysis of time series of measured or hindcast significant wave height, H_S. This analysis is complicated by two effects.

Firstly, H_S exhibits temporal dependence, invalidating naive application of extreme value analysis. Instead, time series must be de-clustered into observations of (independent) storm peak significant wave

height H_S^{sp}, and (intra-storm) directional dissipation of H_S conditional on H_S^{sp}. Extreme value analysis is then performed on H_S^{sp} providing a mechanism to simulate storm peak events for arbitrary return periods. Design values for H_S (for an arbitrary storm sea-state) are next estimated by incorporation of dissipation effects within the simulation. Design distributions for individual maximum wave height H_{max} can then be estimated by marginalisation using the known conditional distribution for H_{max} given H_S. Design values for other intra-storm variables such as maximum crest elevation and total extreme water level can be estimated similarly.

Secondly, the characteristics of H_S^{sp} are non-stationary with respect to multiple covariates, particularly wave direction and season. Failure to accommodate non-stationarity can lead to incorrect estimation of design values. As shown in OMAE2013-10187, covariate effects in peaks over threshold of H_S^{sp} can be modelled in terms of non-stationary models for extreme value threshold (using quantile regression model), the rate of occurrence of threshold exceedances (using a Poisson model), and the sizes of exceedances (using a generalised Pareto model). Model parameters are described as smooth functions of covariates using appropriate multidimensional penalised B-splines. Optimal parameter smoothness is estimated using cross-validation.

In this work, we develop directional-seasonal design values for H_S^{sp}, H_S and H_{max} for a location in the North Sea. Particular attention is paid to the assessment of model bias and quantification of model parameter and design value uncertainty using bootstrap resampling. We also outline existing work on extension to estimation of maximum crest elevation and total extreme water level.

The use of design criteria varying with direction is well-established, particularly for in-place reassessments and reliability studies of fixed jacket structures. However, there are certain situations where design criteria varying with both season and direction may be more appropriate. One example is site-specific assessments of jack-up or mobile offshore drilling units which will only operate through the summer. The estimation of extreme value models which accommodate directional and seasonal variability is therefore of considerable interest.

There is a large literature on applied extreme value analysis relevant to ocean engineering. Threshold methods in extreme value analysis are reviewed by Scarrott and MacDonald [2012]. Tancredi et al. [2006] considers accounting for threshold uncertainty in extreme value anlaysis. Wadsworth and Tawn [2012] presents likelihood-based procedures for threshold diagnostics and uncertainty. Thompson et al. [2009] proposes automatic threshold selection for extreme value analysis. Thompson et al. [2010] reports Bayesian non-parametric regression using splines. Muraleedharan et al. [2012] and Cai and Reeve [2013] model significant wave height distributions with quantile functions for estimation of extreme wave heights. Scotto and Guedes-Soares [2000] and Scotto and Guedes-Soares [2007] discuss the long-term prediction of significant wave height. Methods for analysis of time-series extremes are reviewed by Chavez-Demoulin and Davison [2012]. Ferro and Segers [2003] and Fawcett and Walshaw [2007] discuss modelling of clustered extremes. Mendez et al. [2006] considers long-term variability of extreme significant wave height using a time–dependent POT model. Ruggiero et al. [2010] reports increasing wave heights and extreme value projections for the US Pacific Northwest. Calderon-Vega et al. [2013] models seasonal variation of extremes in the Gulf of Mexico using a time-dependent GEV model. Mendez et al. [2008] considers the seasonality and duration in extreme value distributions of significant wave height. Mackay et al. [2010] discusses discrete seasonal and directional models for the estimation of extreme wave conditions. Eastoe and Tawn [2012] models non-stationary extremes with application to surface level ozone. Chavez-Demoulin and Davison [2005] provides a nice introduction to modelling non-stationary extremes using splines, and Davison et al. [2012] is a good introduction to spatial extremes. Jonathan and Ewans [2013] overviews extreme value analysis from a met–ocean perspective.

Extreme value models for storm severity are generally estimated using storm peak significant wave height H_S^{sp} (see, for example, Jonathan and Ewans 2013), so that each independent storm event is repre-

sented just once in the sample for statistical modelling. Simulation under this model allows estimation of the distribution of maximum storm peak significant wave height in any return period of interest. To account for within-storm (henceforth intra-storm) evolution of significant wave height H_S (as opposed to H_S^{sp}), simulation of H_S for all storm sea-states is necessary.

Capturing covariate effects of extreme sea states is important when developing design criteria. In previous work (see, for example, Jonathan and Ewans [2007], Ewans and Jonathan [2008]) it has been shown that omni-directional design criteria derived from a non-stationary model which adequately incorporates covariate effects can be materially different from a stationary model which ignores those effects (see, for example, Jonathan et al. [2008]). Similar effects have been demonstrated for seasonal covariates (see, for example, Anderson et al. [2001], Jonathan et al. [2008]). Randell et al. [2013] (and Jonathan et al. 2014) report a spatio-directional model for storm peak significant wave height, H_S^{sp} in the Gulf of Mexico, in which the characteristics of extreme values vary with storm direction and location.

A non-stationarity extreme value model is generally superior to the alternative "partitioning" method sometimes used within the ocean engineering community. In the partitioning method, the sample of storm peak significant wave heights H_S^{sp} is partitioned into subsets corresponding to approximately constant values of covariates; independent extreme value analysis is then performed on each subset. For example, in the current application we might choose to partition the sample into directional octants and seasonal quarters, and then estimate (stationary) extreme value models for each of the 32 ($= 8 \times 4$) subsets. There are two main reasons for favouring a non-stationarity model over the partitioning method. Firstly, the partitioning approach incurs a loss in statistical efficiency of estimation, since parameter estimates for subsets with similar covariate values are estimated independently of one another, even though physical insight would require parameter estimates to be similar. In the non-stationary model, we *require* that parameter estimates corresponding to similar values of covariates be similar, and optimise the degree of similarity using cross-validation. For this reason, parameter uncertainty from the non-stationary model is generally smaller than from the partitioning approach. Secondly, the partitioning approach assumes that, within each subset, the sub-sample for extreme value modelling is homogeneous with respect to covariates. In general it is difficult to estimate what effect this assumption might have on parameter and return value estimates (especially when large intervals of values of covariates are combined into a subset). In the non-stationary model, we avoid the need to make this assumption completely.

Whilst the extreme significant wave height is an important parameter in the process of deriving extreme loads on an offshore structure, the largest load experienced by a structure will usually be due to the effect of a single wave rather than to the whole sea state. In fact, for offshore platforms the most significant characteristics are: (a) the return period maximum wave height and its associated wave period, from which extreme kinematics can be derived (in conjunction with a wave theory such as Stokes Fifth Order or NewWave, Tromans et al. 1991 and Jonathan et al. 1994), and (b) the return period total extreme water level, namely the sum of tide, surge and wave crest, used to determine whether there is wave-in-deck loading, typically at the 10,000-year level. Estimation of return values for maximum individual wave H_{max} and maximum crest C_{max} per sea-state requires the consideration of their intra-storm probability distributions for wave height H and crest elevation C (Forristall 1978 and Forristall 2000 respectively) given sea-state characteristics including H_S.

The cumulative distribution function for the maximum wave height H_{max} in a sea-state of n_s waves with significant wave height $H_S = h_s$ is taken (see, for example, Prevosto et al. 2000) to be given by:

$$P(H_{max} \leq h_{max} | H_S = h_s, M = n_s) = (1 - \exp(-\frac{1}{\beta}(\frac{h_{max}}{h_s/4})^{\alpha}))^{n_s}$$

with $\alpha = 2.13$ and $\beta = 8.42$. The number of waves n_s in a particular sea state is estimated by dividing the length of the sea-state (in seconds) by its zero-crossing period, T_Z.

The objective of the the current work is to estimate 100-year design values for significant wave height H_S and maximum wave height H_{max} based on a sample of oceanographic time-series for a North Sea location (introduced in Section 2). There are two key components of the modelling procedure, the first being the estimation of a directional-seasonal extreme value model for storm peak significant wave height H_S^{sp} (discussed in Section 3). The second component is the simulation of realisations of H_S^{sp} (for the storm peak sea-state) under the model, and thereby simulation of H_S and H_{max} for all storm sea-states (all outlined in Section 4). Simulation of H_S for all sea-states is achieved using so-called *intra-storm trajectories* isolated from the original time-series (see Section 2 and the appendix). Simulation of H_{max} requires the incorporation of the intra-storm probability distributions for H_{max} given H_S. Diagnostic plots for validation of the estimated model are presented in Section 5. Current and future developments are outlined in the discussion (Section 6).

2 Data

The application data consist of hindcast time-series (from Reistad et al. 2011) for significant wave height H_S, (dominant) wave direction θ, season ϕ (defined as day of the year, for a standardised year consisting of 360 days), mean zero up-crossing period T_Z (required for sampling from the distribution of maximum wave height H_{max} for a given sea-state, as outlined above) and period T_{01} (required for the Forristall crest height distribution, see below) for three hour sea-states for the period September 1957 to December 2012 at a northern North Sea location. Aarnes et al. [2012] and Breivik et al. [2013] have studied extreme value characteristics of storm severities from the hindcast.

Storm peak characteristics and intra-storm trajectories 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 corresponding to a storm event. The peak-picking threshold corresponds to a directional quantile of H_S with specified non-exceedance probability, estimated using quantile regression. The maximum of significant wave height during the interval is taken as the storm peak significant wave height for the storm. The value of other variables at the time of the storm peak significant wave height are referred to as storm peak values of those variables. Consecutive storms within 24 hours of one another are combined. The resulting storm peak sample consists of 2761 values of H_S^{sp}. With direction from which a storm travels expressed in degrees clockwise with respect to north, Figure 1 consists of scatter plots of H_S^{sp} versus storm peak direction θ^{sp} and storm peak season ϕ^{sp}. Figure 2 shows empirical quantiles of H_S^{sp} by θ^{sp} and θ^{sp}.

Figure 2 shows that storm intervals and storm peak values are identified for most directions and seasons using the peak-picking procedure. The effect of fetch variability with direction on storm peak values is clear from the upper panel of Figure 1. For storms emanating from the north-east (i.e. from approximately 45^o), there is only one occurrence of an event appreciably above $4m$ regardless of season. Further inspection of Figure 2 shows that, even during winter months, storm severities from $[0, 90)$ are low compared with events from other directions. These storms are very unlikely to influence estimates for omni-directional or omni-seasonal return values, but they will influence estimation of directional and seasonal return values for the directions and seasons concerned.

Corresponding to each storm and storm peak quadruplet H_S^{sp}, θ^{sp}, ϕ^{sp} and T_Z^{sp}, the within-storm time-series of H_S, θ, ϕ and T_Z are together referred to as the *intra-storm trajectory* for the storm. Intra-storm trajectories are essential for estimation of design values for intra-storm characteristics H_S and H_{max} in Section 4. Figure 3 shows intra-storm trajectories of significant wave height, H_S, on wave direction θ for 30 randomly-chosen storm events (in different colours). The variability in storm length, and storm directions covered is clear.

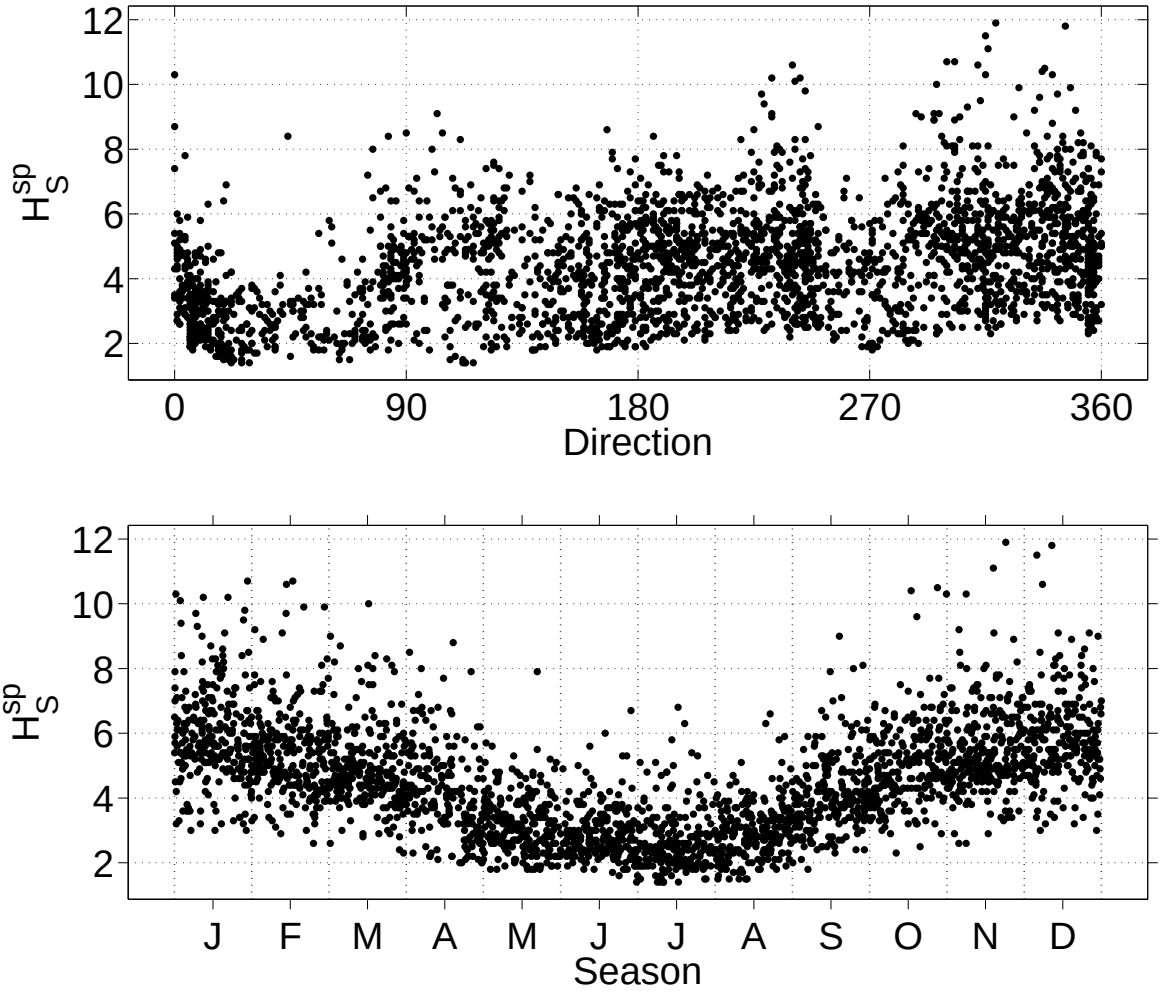

Fig. 1. Storm peak significant wave height H_S^{sp} on storm direction θ^{sp} (upper panel) and storm season ϕ^{sp} (lower panel).

3 Extreme value model

We seek to estimate a non-stationary extreme value model for storm peak significant wave height H_S^{sp}, the parameters of which vary smoothly with respect to storm peak direction θ_{sp} and season ϕ_{sp}.

3.1 Model components

Following Randell et al. [2013], for a sample $\{\dot{z}_i\}_{i=1}^{\dot{n}}$ of $\dot{n}$ storm peak significant wave heights observed with storm peak directions $\{\dot{\theta}_i\}_{i=1}^{\dot{n}}$ and storm peak seasons $\{\dot{\phi}_i\}_{i=1}^{\dot{n}}$ (henceforth together referred to as covariates), we proceed using the peaks over threshold approach as follows.

Threshold: We first estimate a threshold function ψ above which observations $\dot{z}$ are assumed to be extreme. The threshold varies smoothly as a function of covariates ($\psi \triangleq \psi(\theta,\phi)$) and is estimated using quantile regression. We retain the set of n threshold exceedances $\{z_i\}_{i=1}^{n}$ observed with storm peak directions $\{\theta_i\}_{i=1}^{n}$ and storm peak seasons $\{\phi_i\}_{i=1}^{n}$ for further modelling.

Rate of occurrence of threshold exceedance: We next estimate the rate of occurrence ρ of threshold exceedance using a Poisson process model with Poisson rate $\rho(\triangleq \rho(\theta,\phi))$.

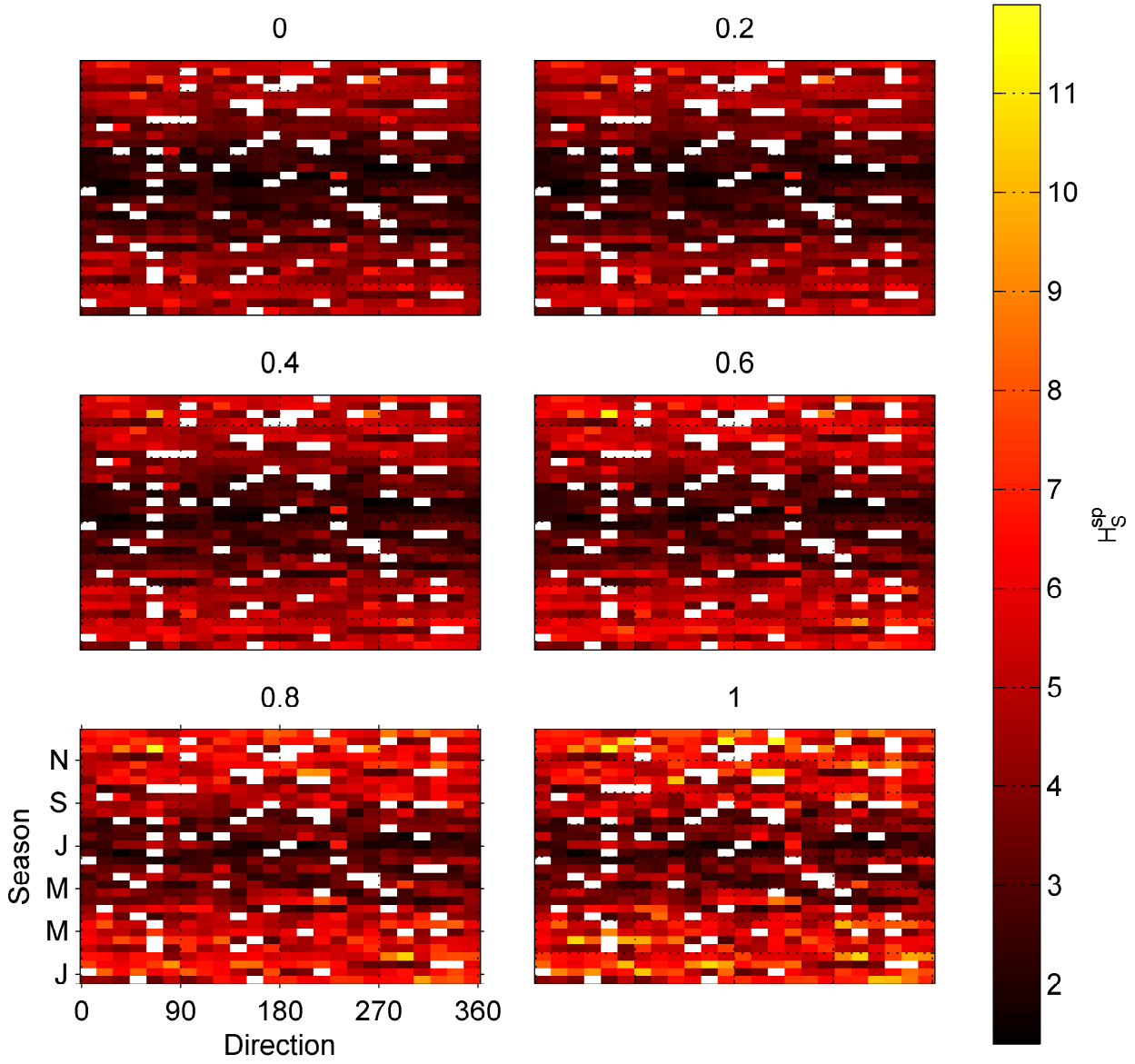

Fig. 2. Empirical quantiles of storm peak significant wave height, H_S^{sp} by storm direction, θ^{sp}, and storm season, ϕ^{sp}. Panel titles indicate quantile non-exceedance probability. Empty bins are coloured white.

Size of occurrence of threshold exceedance: We estimate the size of occurrence of threshold exceedance using a generalised Pareto (henceforth GP for brevity) model. The GP shape and scale parameters ξ and σ are also assumed to vary smoothly as functions of covariates.

This approach to extreme value modelling follows that of Chavez-Demoulin and Davison [2005] and is equivalent to direct estimation of a non-homogeneous Poisson point process model (see, for example, Dixon et al. 1998, Jonathan and Ewans [2013]).

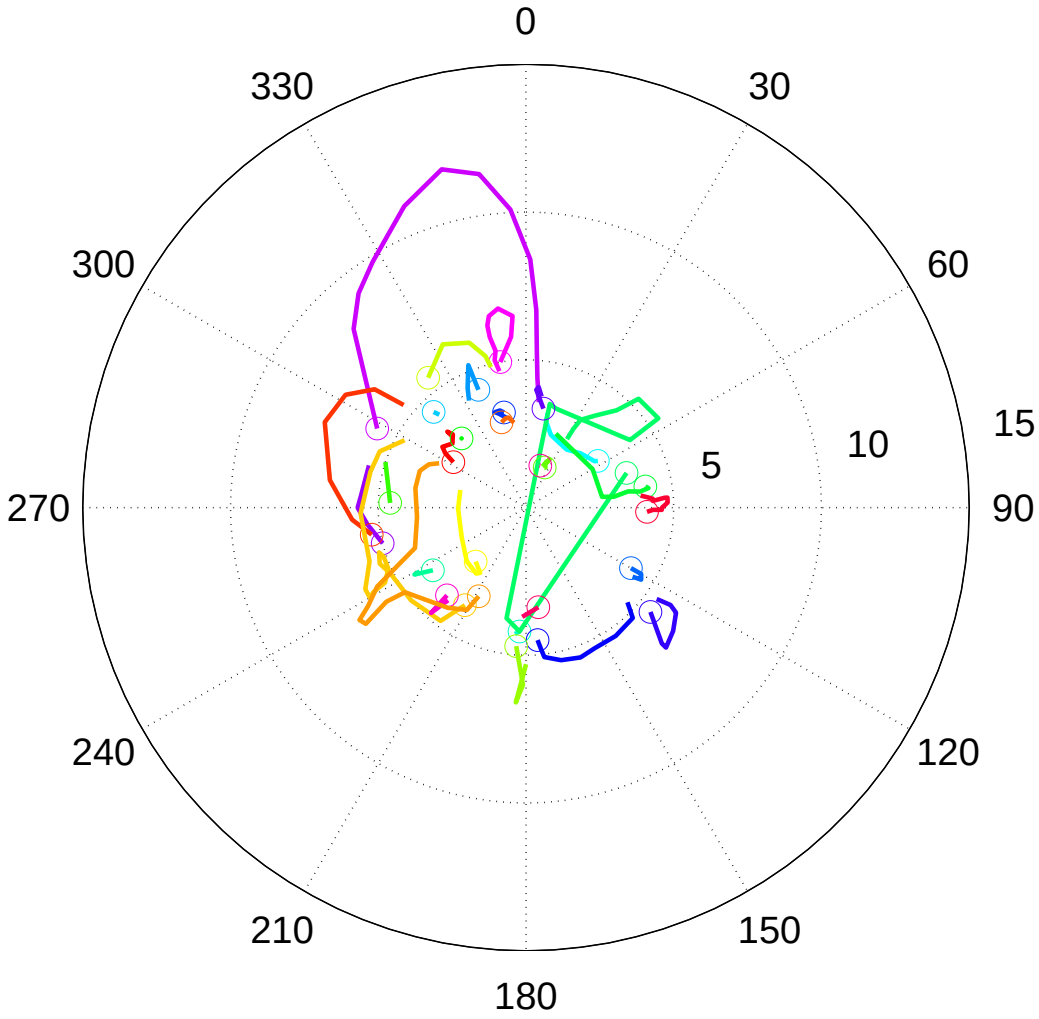

Fig. 3. Storm trajectories of significant wave height, H_S, on wave direction θ for 30 randomly-chosen storm events (in different colours). A circle marks the start of each intra-storm trajectory.

3.2 Parameter estimation

For quantile regression, 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 cross-validation to provide good predictive performance.

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

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

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

$$\ell_{\xi,\sigma} = \sum_{i=1}^{n} \log \sigma_i + (\frac{1}{\xi_i} + 1) \log(1 + \frac{\xi_i}{\sigma_i}(z_i - \psi_i))$$

where $\psi_i = \psi(\theta_i, \phi_i)$, $\xi_i = \xi(\theta_i, \phi_i)$ and $\sigma_i = \sigma(\theta_i, \phi_i)$, and a similar expression is used when $\xi_i = 0$ (see Jonathan and Ewans 2013). In practice, we set $\lambda_\xi = \kappa\lambda_\sigma$ for prespecified 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.

3.3 Parameter smoothness

Physical considerations suggest that we should expect the model parameters ψ, ρ, ξ and σ to vary smoothly with respect to covariates θ, ϕ. For estimation, this can be 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 under consideration. 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 the marginal basis matrices. Thus

$$B = B_\phi \otimes B_\theta$$

provides a $(m \times p)$ 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 therefore reduces to estimating appropriate sets of basis coefficients for each of ψ, ρ, ξ and σ.

The roughness R of any function can be easily 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 can be easily evaluated for the marginal and three dimensional domains. The form of P is motivated by taking differences of neighbouring values of β, thereby penalising lack of local smoothness. The values of p_θ and p_ϕ are functions of the number of spline knots for each marginal domain, and also depend on whether spline bases are specified as periodic (which is the case for both marginal bases in this application).

3.4 Uncertainty quantification

Bootstrap resampling is used for uncertainty quantification. 95% bootstrap uncertainty bands are estimated by repeating the full extreme value analysis for 1000 resamples of the original storm peak sample. In particular, estimation of optimal roughness penalties is performed independently for each bootstrap resample, so that uncertainty bands also reflect variability in these choices. It was also confirmed that 1000 resamples was sufficient to ensure stability of bootstrap confidence intervals.

3.5 Estimated parameters

Figure 4 shows plots for extreme value threshold ψ, corresponding to non-exceedance probability 0.5 of H_S^{sp}. The upper panel shows the bootstrap median threshold on storm peak direction θ^{sp}, and storm peak season ϕ^{sp}. The lower panels show 12 monthly directional thresholds in terms of bootstrap median (solid) and 95% bootstrap uncertainty band (dashed). From inspection of the upper image, it is clear that summer periods are relatively calm, as are storm events from directions in $[0, 90)$. Figure 5 shows plots for rate of threshold exceedance ρ of H_S^{sp}. The upper panel shows the bootstrap median rate on θ^{sp} and ϕ^{sp}. The lower panels show 12 monthly directional rates in terms of bootstrap median (solid) and 95% bootstrap uncertainty band (dashed). The rate of occurrence of threshold exceedances is largest for winter storms from either around 180^o or 360^o.

Figure 6 shows plots for generalised Pareto shape ξ. The upper panel shows the bootstrap median shape on θ^{sp} and ϕ^{sp}. The lower panels show 12 monthly directional shapes in terms of bootstrap median (solid) and 95% bootstrap uncertainty band (dashed). The corresponding plots for generalised Pareto scale σ are given in Figure 7. ξ shows greatest directional variability in the months of October - December, but the uncertainty in the estimates of ξ are relatively large (so that a constant model for would suffice for this application). The estimates of σ show greater variation; largest values are observed for winter storms emanating from directions in $[270, 360)$.

4 Estimation of return values

Return values corresponding to some return period P of interest are estimated by simulation under the model developed in Section 3. The procedure is as follows, for each of a large number N of realisations of storms:

1. Select a bootstrap resample and the corresponding estimated directional-seasonal extreme value model for storm peak significant wave height.
2. For each directional-seasonal covariate bin, estimate the number of storm peak realisations to be drawn at random using the estimated directional-seasonal rate of threshold exceedance, ρ, for that

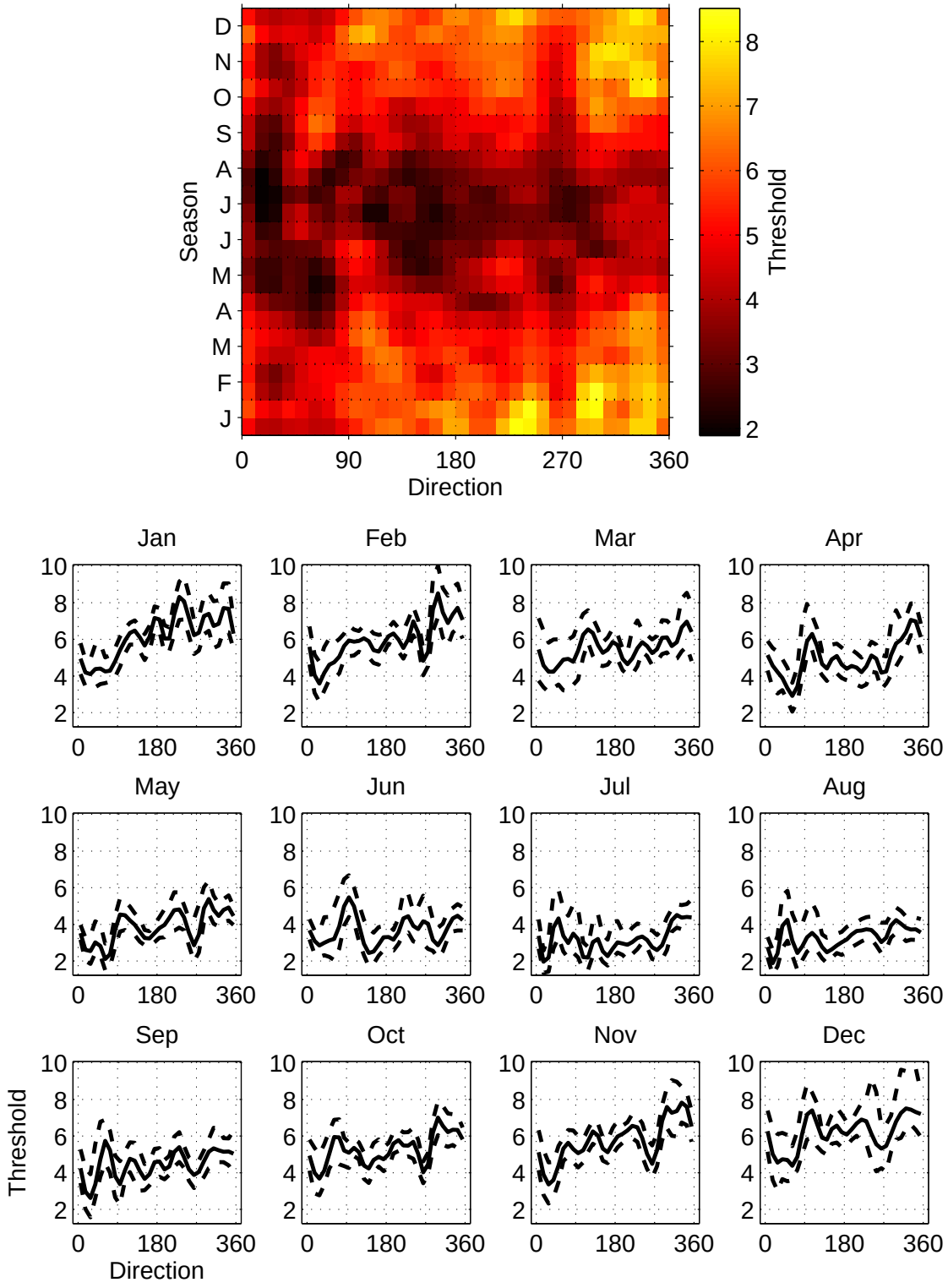

Fig. 4. Directional-seasonal parameter plot for extreme value threshold, ψ, corresponding to non-exceedance probability 0.5 of H_S^{sp}. The upper panel shows the bootstrap median threshold on storm peak direction, θ^{sp}, and storm peak season, ϕ^{sp}. The lower panels show 12 monthly directional thresholds in terms of bootstrap median (solid) and 95% bootstrap uncertainty band (dashed).

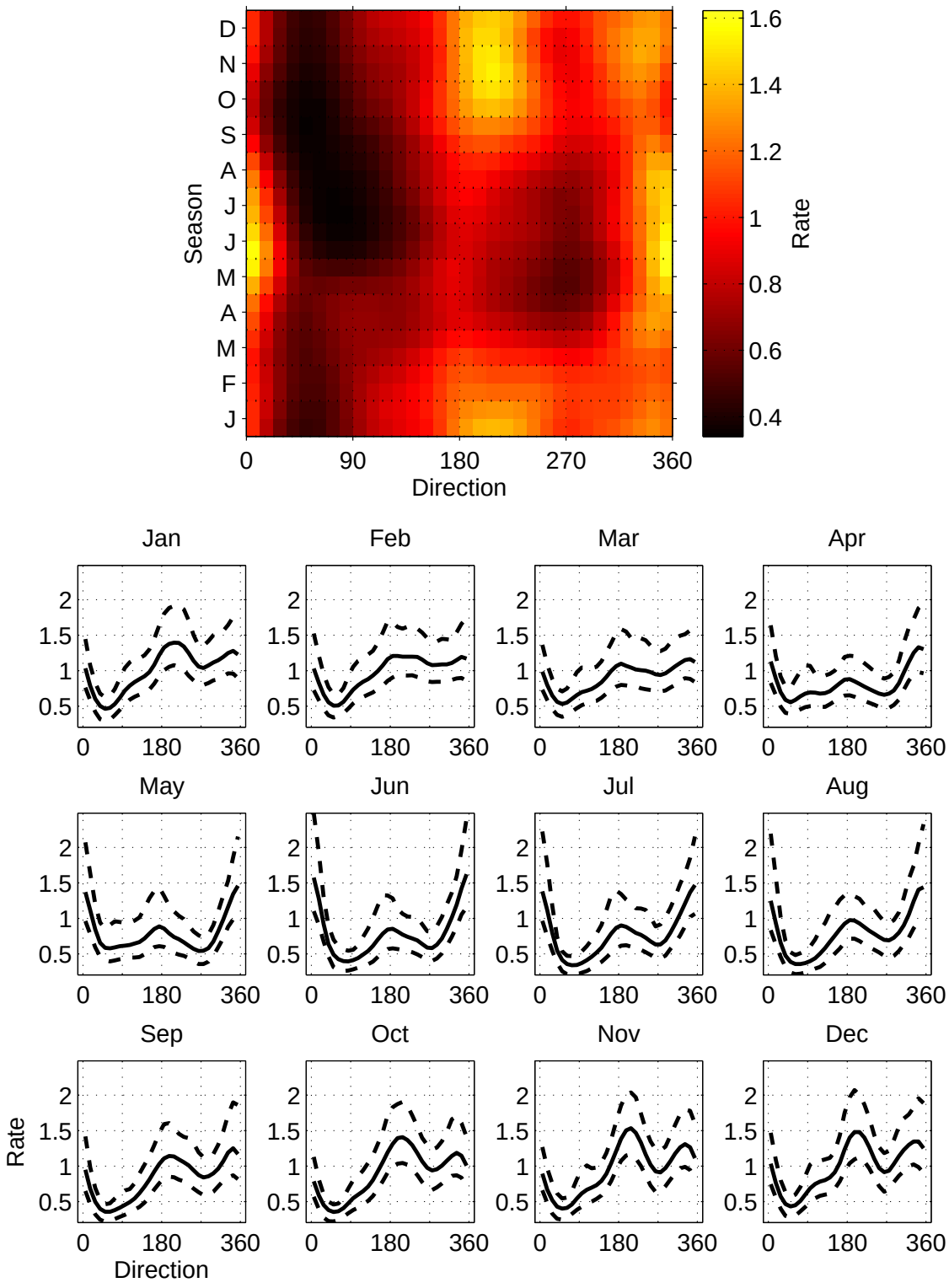

Fig. 5. Directional-seasonal parameter plot for rate of threshold exceedance, $\rho \times 10^4$, of H_S^{sp}. The upper panel shows the bootstrap median rate on θ^{sp} and ϕ^{sp}. The lower panels show 12 monthly directional rates in terms of bootstrap median (solid) and 95% bootstrap uncertainty band (dashed). Unit of rate ρ is number of occurrences per annum per directional-seasonal covariate bin.

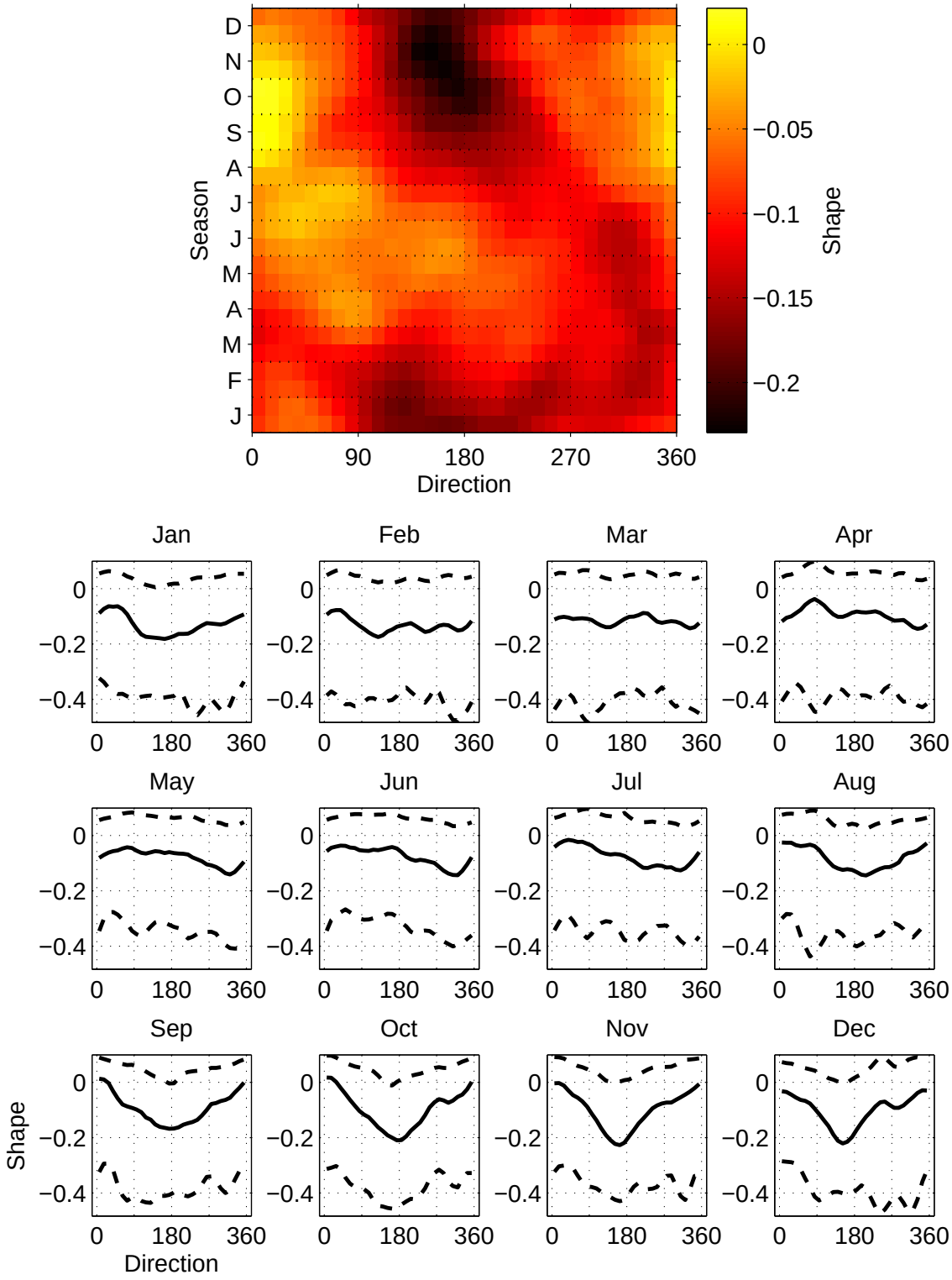

Fig. 6. Directional-seasonal parameter plot for generalised Pareto shape, ξ. The upper panel shows the bootstrap median shape on θ^{sp} and ϕ^{sp}. The lower panels show 12 monthly directional shapes in terms of bootstrap median (solid) and 95% bootstrap uncertainty band (dashed).

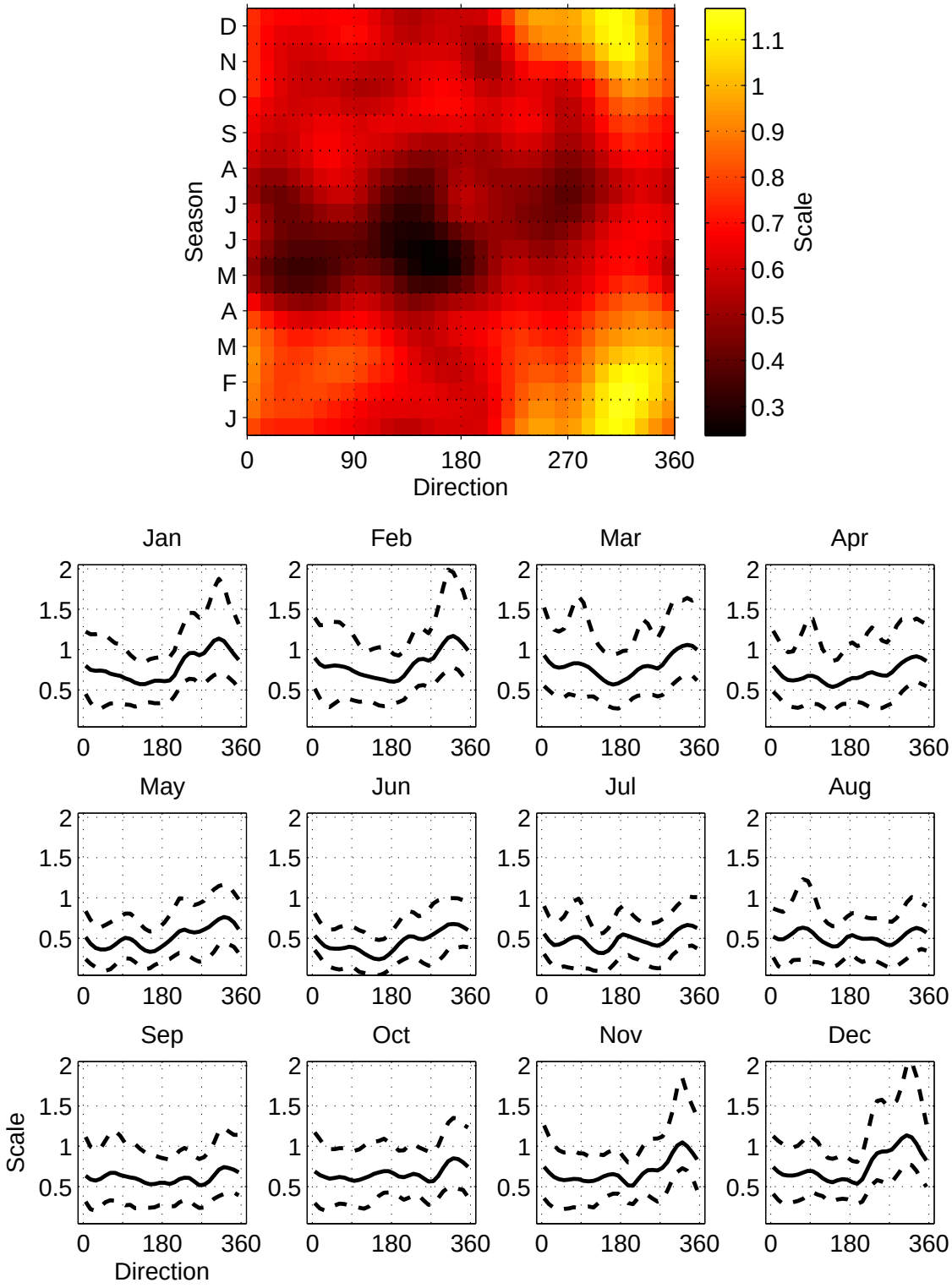

Fig. 7. Directional-seasonal parameter plot for generalised Pareto scale, σ. The upper panel shows the bootstrap median scale on θ^{sp} and ϕ^{sp}. The lower panels show 12 monthly directional scales in terms of bootstrap median (solid) and 95% bootstrap uncertainty band (dashed).

bin, scaled to return period, P. If T is the period of the original sample, the scaled rate is $\rho * P/T$). Then, for each storm peak realisation:

(a) Draw a pair of values for storm peak direction θ^{sp*} and storm peak season ϕ^{sp*} at random from the volume corresponding to the covariate bin.
(b) Draw a value of storm peak significant wave height H_S^{sp*} corresponding to θ^{sp*} and ϕ^{sp*}, at random from corresponding the generalised Pareto model.
(c) Draw an intra-storm trajectory corresponding to the triplet (H_S^{sp*}, θ^{sp*}, H_S^{sp*}, ϕ^{sp*}), using the procedure described in the appendix.
(d) For each sea-state in the intra-storm trajectory, use closed-form distributions for maximum wave height H_{max} to sample values H^*_{max}.

3. Accumulate maximum values for storm peak (H_S^{sp*}) and intra-storm H^*_{max} variables per directional-seasonal covariate bin.

Empirical cumulative distribution functions for storm peak and intra-storm maxima are then trivially estimated by sorting the values for each variable for arbitrary combinations of covariate bins. In this way, for example, cumulative distribution functions for directional return values each month of the year, or seasonal return values for directional octants can be estimated. By retaining only maxima over all covariate bins, omni-directional omni-seasonal are obtained. Importantly, since realisations based on models from different bootstrap resamples of the original sample are used, the resulting cumulative distribution functions incorporate both the (aleatory) inherent randomness of return values and the extra (epistemic) uncertainty introduced by model parameter estimation from a sample of data. Figure 8 shows cumulative distribution functions (cdfs) for 100-year storm peak significant wave height H_{S100} from simulation under the directional-seasonal model, incorporating uncertainty in parameter estimation using bootstrap resampling as explained above. Upper panel shows cdfs for directional octants and lower panel for months of year. The common omni-directional omni-seasonal cdf is shown in both panels (in black). It is clear that the severest storms come from the north - west in winter months. The median omni - directional omni - seasonal 100 - year storm peak value is approximately 12.2m. Figure 9 shows return value plots for 100-year significant wave height H_{S100}. The upper panel shows omni-seasonal return values on wave direction θ, in terms of directional octant median (solid black), most-probable (dot-dashed black), 2.5%ile and 97.5%ile (both dashed black) and the corresponding omni-directional omni-seasonal estimates (in red, common to Figure 10). The lower panels show 12 monthly directional octant return values (in black) in terms of median (solid), most-probable (dot-dashed), 2.5%ile and 97.5%ile (both dashed). Also shown are the corresponding omni-directional estimates (in red). Figure 10 also shows return value plots for H_{S100}. But now the upper panel shows omni-directional return values on wave season ϕ, in terms of monthly median (solid black), most-probable (dot-dashed black), 2.5%ile and 97.5%ile (both dashed black) and the corresponding omni-directional omni-seasonal estimates (in red, common to Figure 9). The lower panels show seasonal return values for directional octants in terms of median (solid), most-probable (dot-dashed), 2.5%ile and 97.5%ile (both dashed). Also shown are the corresponding omni-seasonal estimates (in red). There are obvious, and statistically significant differences between return values for different directions and seasons. It is important to note that all omni = directional and omni - seasonal estimates here are calculated from the directional - seasonal model; estimating these from models which ignore directional and seasonal variation in extremes would be inappropriate.

Figure 11 shows directional-seasonal return value plots for 100-year maximum wave height H_{max100}. The upper panel shows omni-seasonal return values on wave direction θ, in terms of directional octant median (solid black), most-probable (dot-dashed black), 2.5%ile and 97.5%ile (both dashed black) and the corresponding omni-directional omni-seasonal estimates (in red, common to Figure 12). The lower panels show 12 monthly directional octant return values (in black) in terms of median (solid),

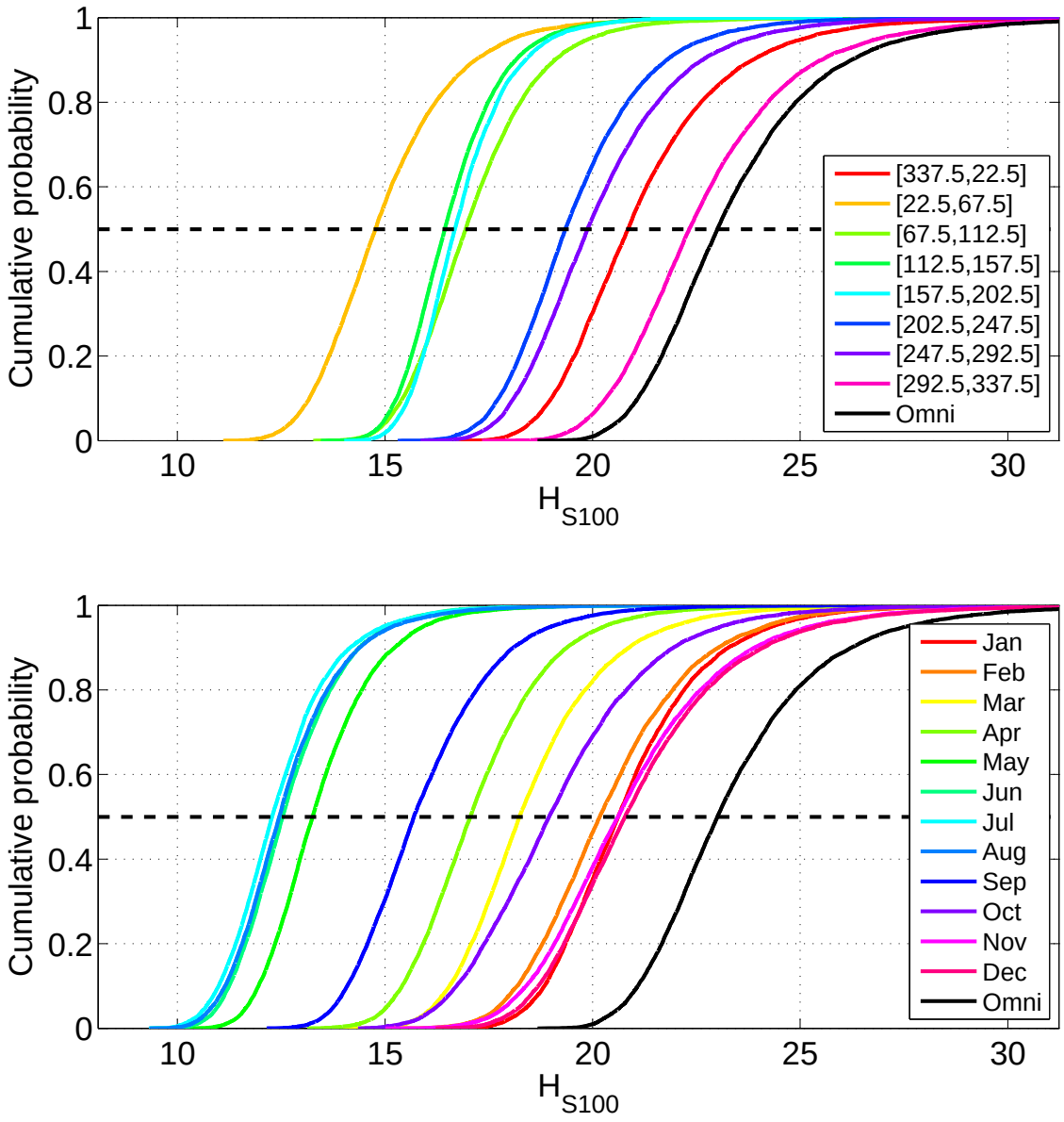

Fig. 8. Cumulative distribution functions (cdfs) for 100-year storm peak significant wave height, H_{S100} from simulation under the directional-seasonal model, incorporating uncertainty in parameter estimation using bootstrap resampling. Upper panel shows cdfs for directional octants and lower panel for months of year. The common omni-directional omni-seasonal cdf is shown in both panels (in black).

most-probable (dot-dashed), 2.5%ile and 97.5%ile (both dashed). Also shown are the corresponding omni-directional estimates (in red). Figure 12 shows the corresponding plots for omni-directional and directional octant extremes as a function of season. Again, there is statistically significant variation in the estimates values for H_{max100}. It is unsurprising that the directional and seasonal profiles of H_{max100} closely mimic those of H_{S100}, since the intra-storm conditional distribution for H_{max} given H_S is stationary with respect to both direction and season.

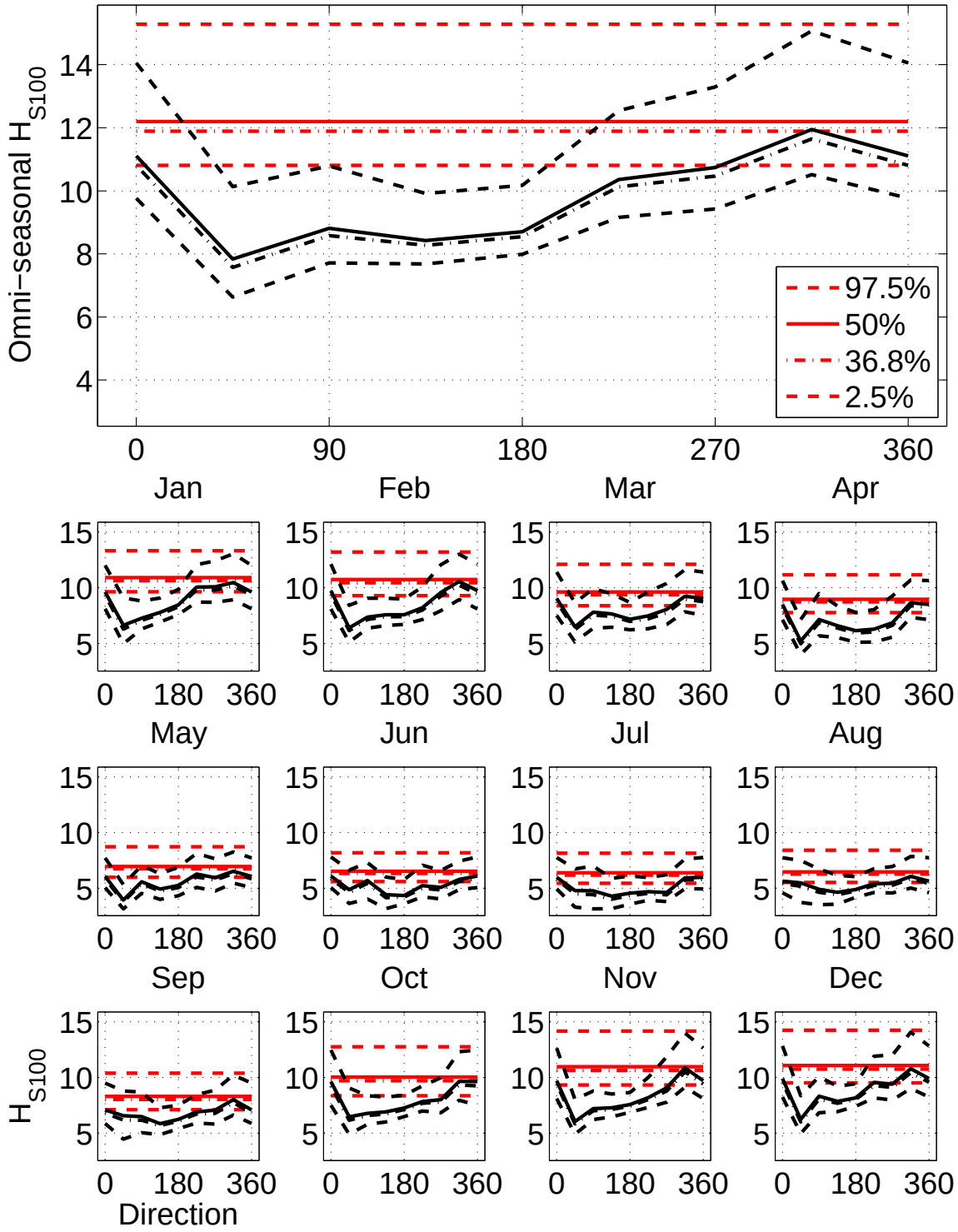

Fig. 9. Directional-seasonal return value plot for 100-year significant wave height, H_{S100}. The upper panel shows omni-seasonal return values on wave direction, θ, in terms of directional octant median (solid black), most-probable (dot-dashed black), 2.5%ile and 97.5%ile (both dashed black) and the corresponding omni-directional omni-seasonal estimates (in red, common to Figure 10). The lower panels show 12 monthly directional octant return values (in black) in terms of median (solid), most-probable (dot-dashed), 2.5%ile and 97.5%ile (both dashed). Also shown are the corresponding omni-directional estimates (in red).

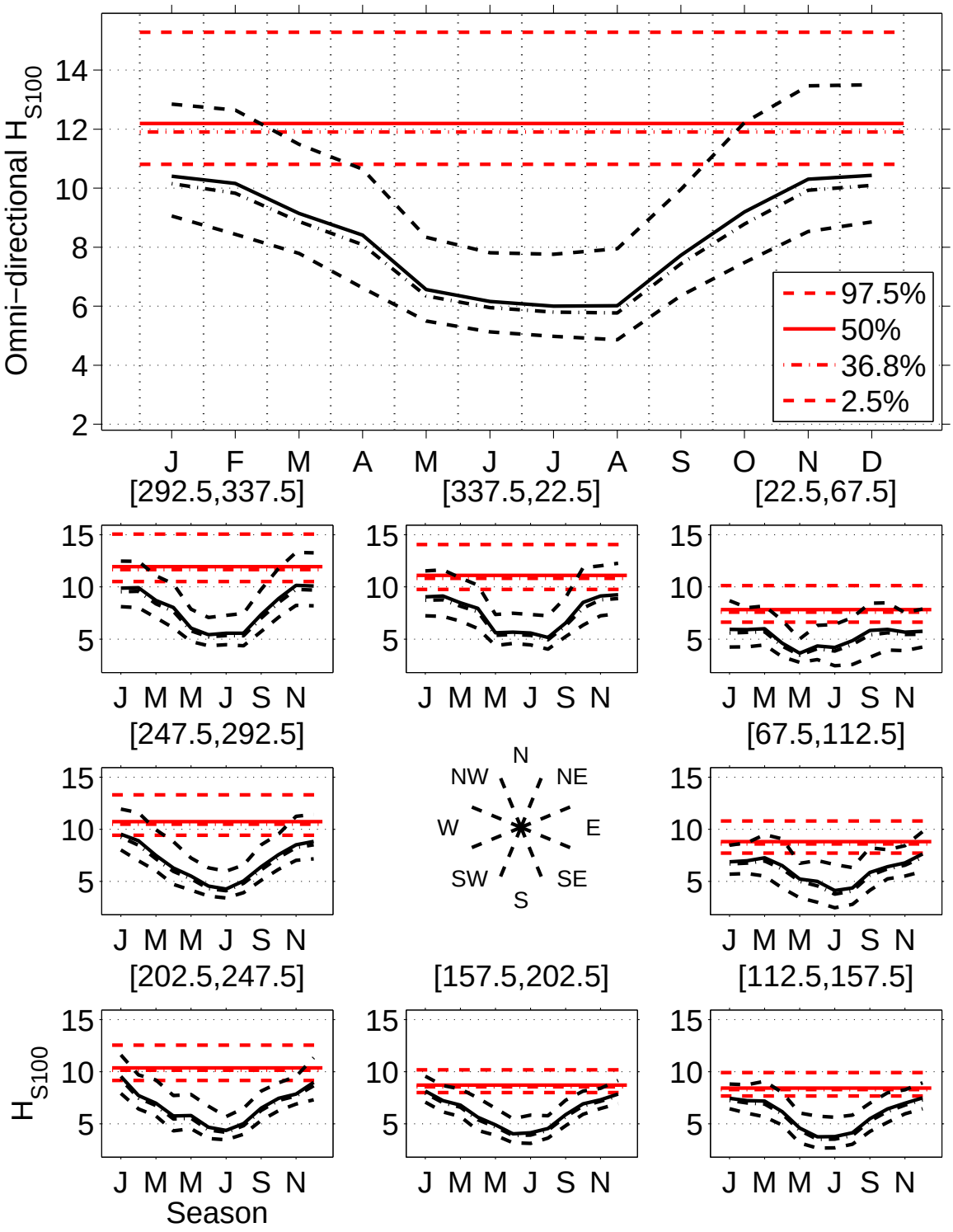

Fig. 10. Directional-seasonal return value plot for 100-year significant wave height, H_{S100}. The upper panel shows omni-directional return values on wave season, ϕ, in terms of monthly median (solid black), most-probable (dot-dashed black), 2.5%ile and 97.5%ile (both dashed black) and the corresponding omni-directional omni-seasonal estimates (in red, common to Figure 9). The lower panels show seasonal return values for directional octants in terms of median (solid), most-probable (dot-dashed), 2.5%ile and 97.5%ile (both dashed). Also shown are the corresponding omni-seasonal estimates (in red).

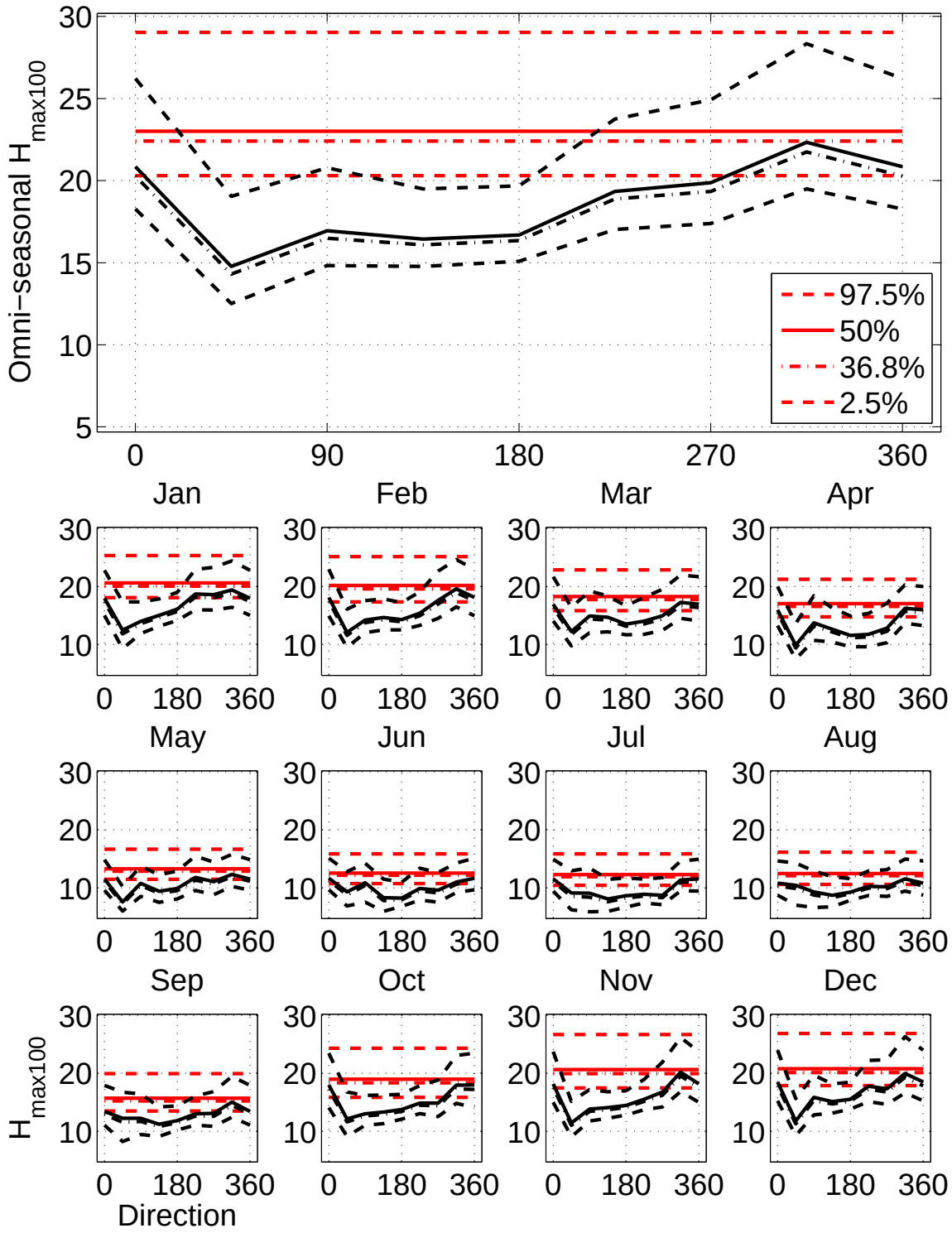

Fig. 11. Directional-seasonal return value plot for 100-year maximum wave height, H_{max100}. The upper panel shows omni-seasonal return values on wave direction θ, in terms of directional octant median (solid black), most-probable (dot-dashed black), 2.5%ile and 97.5%ile (both dashed black) and the corresponding omni-directional omni-seasonal estimates (in red, common to Figure 12). The lower panels show 12 monthly directional octant return values (in black) in terms of median (solid), most-probable (dot-dashed), 2.5%ile and 97.5%ile (both dashed). Also shown are the corresponding omni-directional estimates (in red).

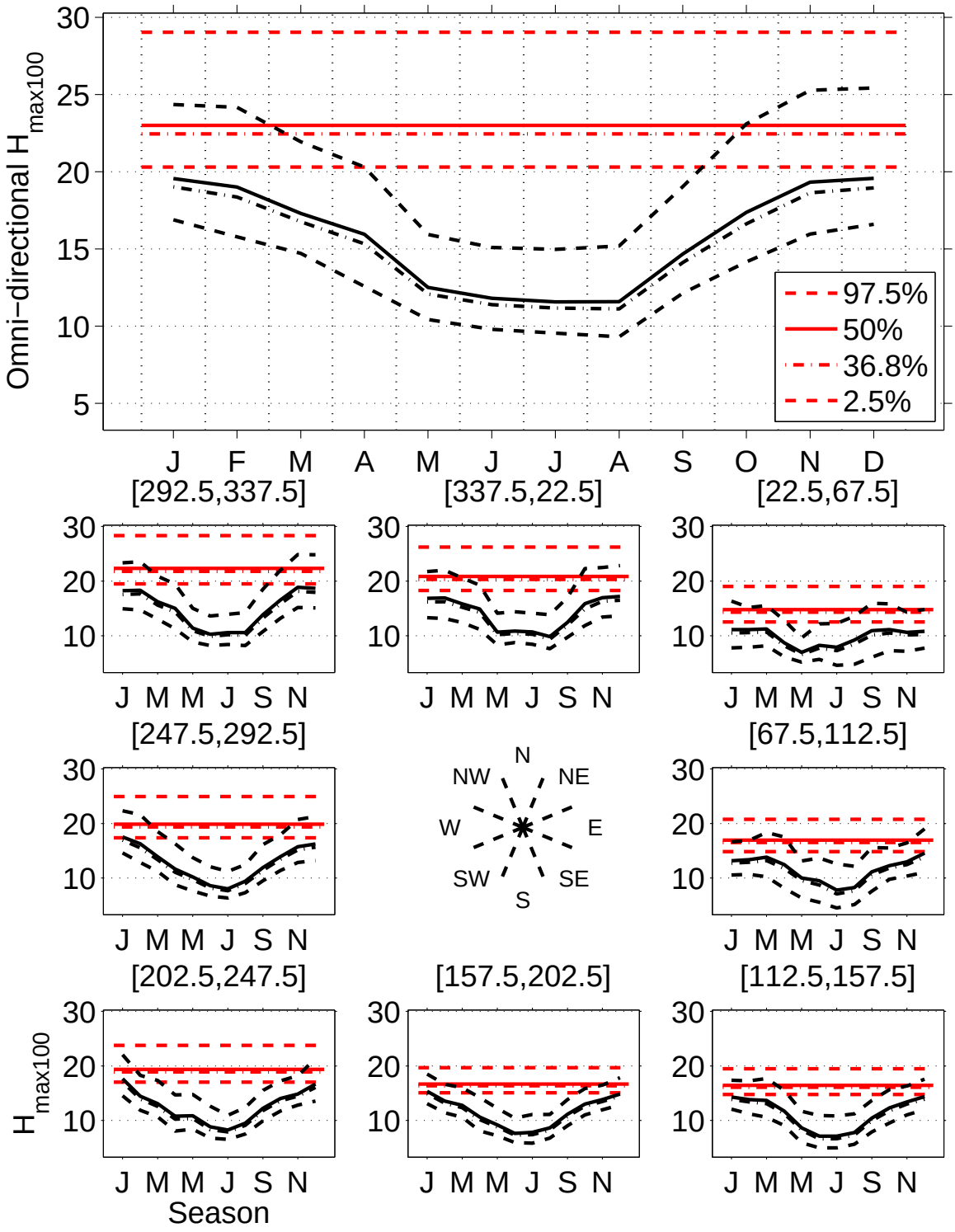

Fig. 12. Directional-seasonal return value plot for 100-year maximum wave height, H_{max100}. The upper panel shows omni-directional return values on wave season, ϕ, in terms of monthly median (solid black), most-probable (dot-dashed black), 2.5%ile and 97.5%ile (both dashed black) and the corresponding omni-directional omni-seasonal estimates (in red, common to Figure 11). The lower panels show seasonal return values for directional octants in terms of median (solid), most-probable (dot-dashed), 2.5%ile and 97.5%ile (both dashed). Also shown are the corresponding omni-seasonal estimates (in red).

5 Validation

Model diagnostics are essential to demonstrate adequate model fit. Of primary concern is that (a) the estimated storm peak extreme value model generates directional - seasonal distributions of H_S^{sp} consistent with observed storm peak data, and that (b) the simulation procedure for estimation of return values (in Section 4) generates directional - seasonal distributions of H_S (for all storm sea-states) consistent with observed data. To quantify this, we use the simulation procedure to generate 1000 realisations of storms, each realisation for the same period (of 55.3 years) as the original data. We then construct 95% uncertainty bands for cumulative distribution functions (cdfs) of H_S^{sp} and H_S, partitioned by direction and season as appropriate. Then we confirm that empirical cdfs for the actual data, for the same directional - seasonal partitions, are consistent with the simulated cdfs. Figure 13 illustrates this for H_S^{sp}. The upper panel shows the omni-directional omni-seasonal cdf for the original sample (red), the corresponding median from simulation (solid black), together with 2.5%ile and 97.5%ile from simulation (both dashed). The lower panels compare 12 monthly cdfs in the same way. There is reasonable agreement.

Figure 14 illustrates the validation of directional-seasonal model for significant wave height, H_S, by comparison of cumulative distribution functions (cdfs) for original sample with those from 1000 sample realisations under the model (incorporating intra-storm evolution of H_S) corresponding to the same time period as the original sample. The upper panel shows the omni-directional omni-seasonal cdf for the original sample (red), the corresponding median from simulation (solid black), together with 2.5%ile and 97.5%ile from simulation (both dashed). The lower panels compare 12 monthly cdfs in the same way. Again, agreement is good.

Plots similar to Figures 13 and 14 showing cdfs per directional octant suggest model fit of similar quality. Since we do not have access to data for maximum wave height, we cannot apply the diagnostic procedure directly.

6 Discussion

In this work, we develop a procedure based on non-stationary extreme value analysis to estimate the distributions of storm peak significant wave height H_S^{sp} and maximum wave height H_{max} corresponding to arbitrary long return periods. The approach exploits recent advances in extreme value analysis with multidimensional covariates to characterise return value characteristics for H_S^{sp} with direction and season (for the storm peak sea state only), and simulation under the extreme value model (a) incorporating intra-storm trajectories to estimate return value characteristics for H_S for all storm sea-states, and (b) known conditional distributions for H_{max} given H_S to estimate return value characteristics for H_{max}. Diagnostic tests demonstrate that the approach performs will in application to North Sea hindcast data.

According to ISO19901-1 [2005], a convolution approach should be used to correctly account for the possibility of a large wave resulting from a sea-state with relatively low severity. The simulation approach used here is similar to the numerical approach described in Tromans and Vanderschuren [1995] but has a number of advantages. The current approach readily accommodates different storm characteristics from different directions, as well as seasonal variability. Furthermore, there is no need to define a single, representative storm shape; instead, actual storm histories are used reflecting natural variability within real storms.

Figures 8-12 report design values resolved into (directional) octants and (seasonal) monthly octants from simulation under the model. Design values for arbitrary directional-seasonal partitions can be estimated in the same way by simulation, in an entirely consistent fashion. For example, omni-directional design values corresponding to the May-September period might be estimated and exploited by the designer for short-term offshore activities. In stark contrast, the lack of consistency in engineering specification of directional design criteria in particular has been the subject of some debate (see, e.g., Forristall 2004). Guidelines such as API [2005] and ISO19901-1 [2005] provide recommendations on

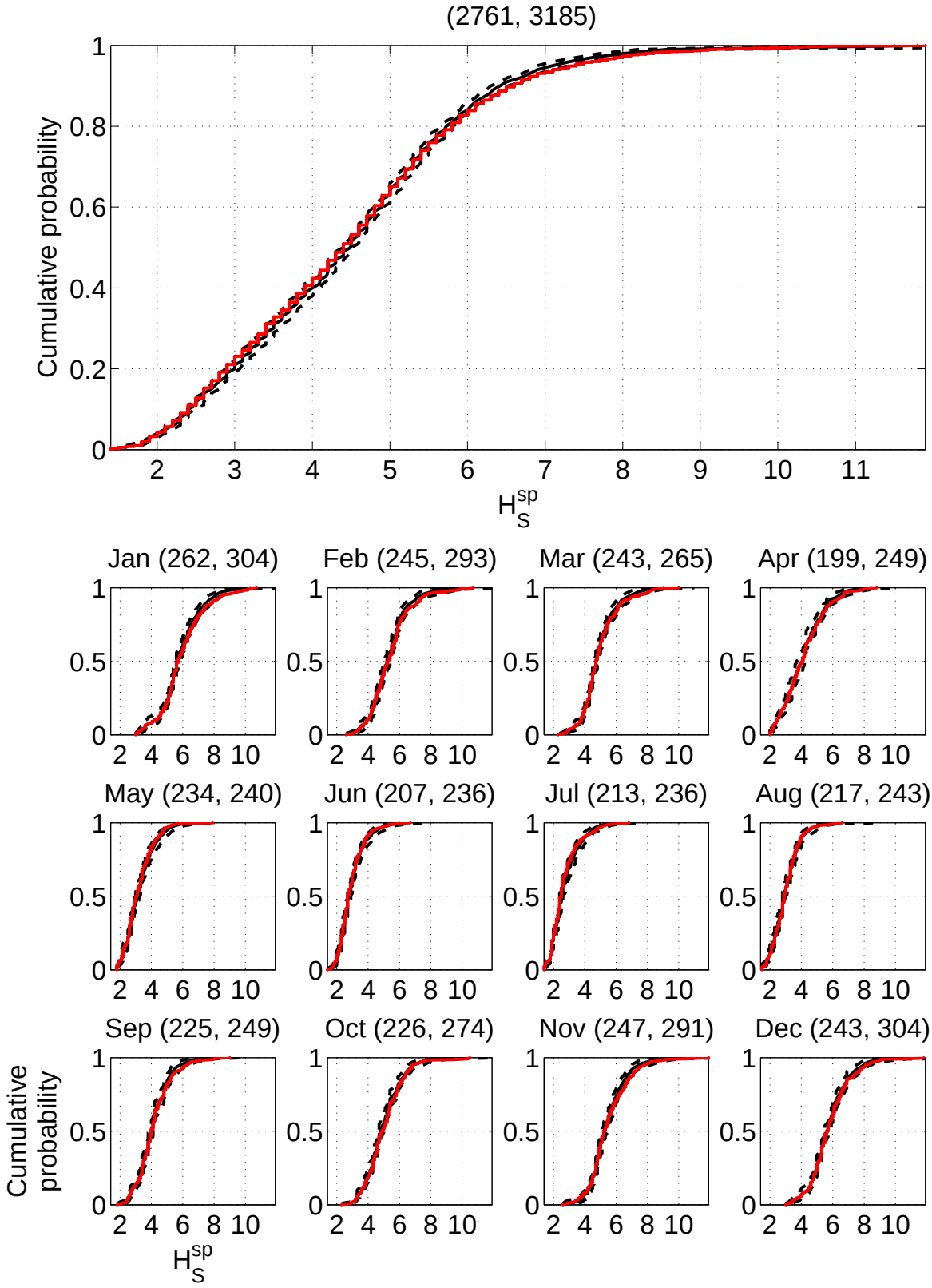

Fig. 13. Validation of directional-seasonal model for storm peak significant wave height, H_S^{sp}, by comparison of cumulative distribution functions (cdfs) for original storm peak sample with those from 1000 sample realisations under the model corresponding to the same time period as the original sample. The upper panel shows the omni-directional omni-seasonal cdf for the original sample (red), the corresponding median from simulation (solid black), together with 2.5%ile and 97.5%ile from simulation (both dashed). The lower panels compare 12 monthly cdfs in the same way. Titles for plots, in brackets following the month name, are the numbers of actual and simulated events in each month.

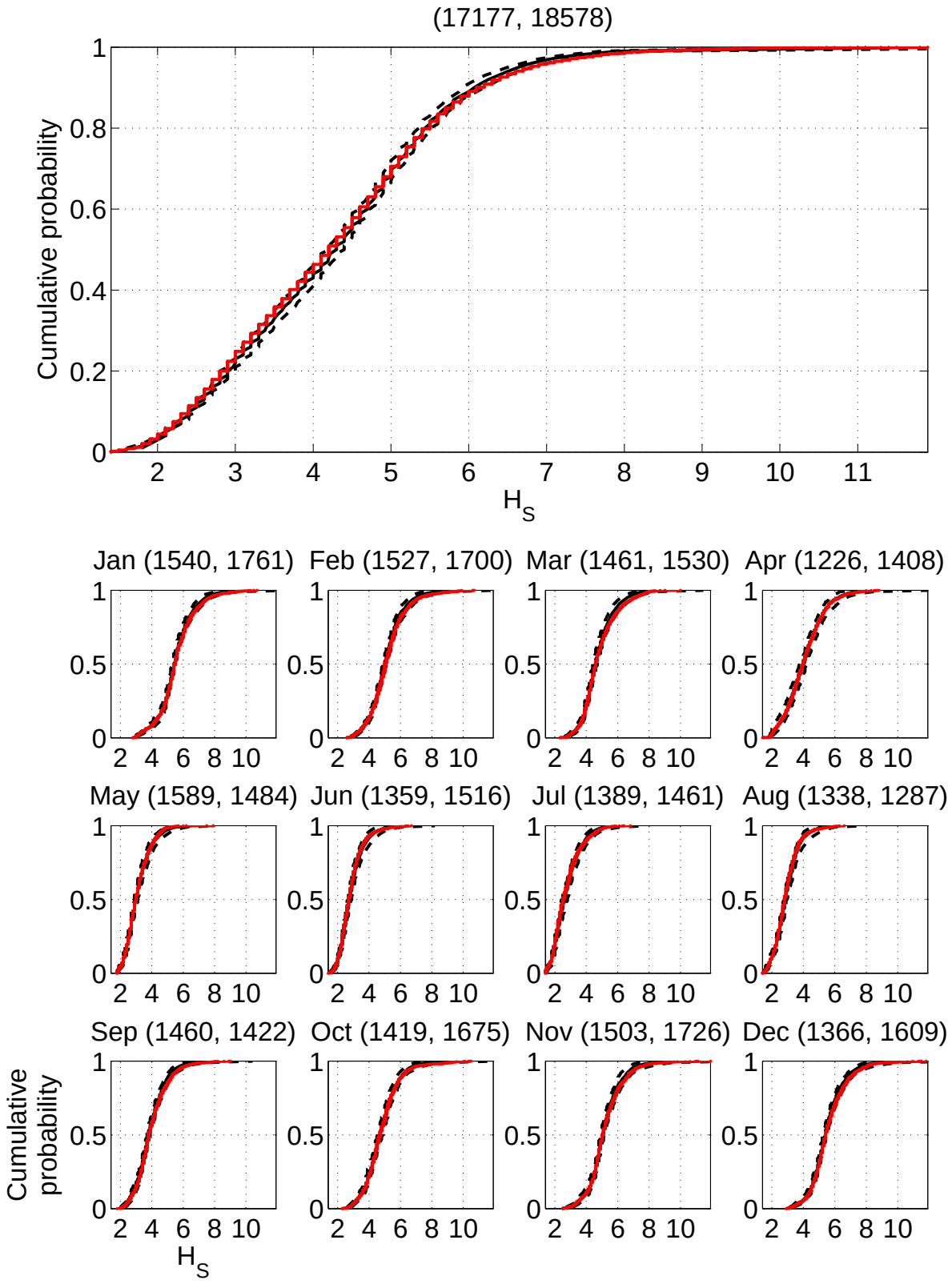

Fig. 14. Validation of directional-seasonal model for significant wave height, H_S, by comparison of cumulative distribution functions (cdfs) for original sample with those from 1000 sample realisations under the model (incorporating intra-storm evolution of H_S) corresponding to the same time period as the original sample. The upper panel shows the omni-directional omni-seasonal cdf for the original sample (red), the corresponding median from simulation (solid black), together with 2.5%ile and 97.5%ile from simulation (both dashed). The lower panels compare 12 monthly cdfs in the same way. Titles for plots, in brackets following the month name, are the numbers of actual and simulated events (in each month).

treating directional criteria, but even when these are followed, either inconsistency remains (in the case of API), or insufficient detail is given on how to make the criteria consistent (in the case of ISO).

The method as described here has focussed on the estimation of estimation of extreme storm peak significant wave height and maximum wave height. Extension to estimation of maximum crest elevation and total extreme water level is the subject of current work and a companion publication in preparation. Since the method of incorporation of non-stationary within the extreme value modelling framework is quite general, extensions to spatial and temporal covariates (for example) are straightforward.

Acknowledgement

We acknowledge useful discussions on computational aspects with Laks Raghupathi of Shell, Bangalore.

Appendix: Selecting intra-storm trajectories for simulated storm events

The directional-seasonal extreme value model is estimated for storm peak significant wave height H_S^{sp}, since storm peak events provide independent events for statistical modelling. However, we require return values for significant wave height H_S from any sea-state (not just the storm peak). We also require return values for maximum wave height H_{max} and maximum crest elevation, which may or may not correspond to the storm peak sea-state. Therefore, in the simulation procedure for return value estimation described in Section 4 we need to generate realisations of whole intra-storm trajectories as defined in Section 3 not just storm peak events. We achieve this by selecting an appropriate intra-storm trajectory from the original sample, with storm peak characteristics in good agreement with those of the current storm peak realisation.

Let $\{\eta_j^*\}_{j=1}^3$ represent the values of H_S^{sp*}, θ^{sp*} and ϕ^{sp*} respectively for the current realisation, and let $\{\eta_{ij}\}_{i=1,j=1}^{n,3}$ represent the corresponding values n values for the original storm peak sample. We define the dissimilarity d_i between the ith (original) storm and the current storm peak realisation as:

$$\begin{aligned} d_i &= \sum_{j=1}^{3} d_{ij}, \\ d_{ij} &= \frac{\eta_{ij} - \eta_j^*}{\tau_j}, \text{ for } c_j(\eta_{ij} - \eta_j^*) > \tau_j, \\ &= 0 \text{ otherwise.} \end{aligned}$$

where $c_1(\bullet)$ is Euclidean distance and $c_2(\bullet)$, $c_3(\bullet)$ are circular distance functions defined on $[0, 360)$. The cut-off values $\{\tau_j\}_{j=1}^3$ indicate when the difference $c_j(\eta_{ij} - \eta_j^*)$ is sufficiently small that it can be ignored in the specification of dissimilarity. After some experimentation, values of $\tau_1 = 0.5$ (metres, for H_S^{sp}), $\tau_2 = 20$ (degrees, for θ^{sp}) and $\tau_3 = 45$ (degrees, for ϕ^{sp}) were chosen.

The subset of original intra-storm trajectories yielding the smallest values of dissimilarity are deemed good matches to the simulated storm peak event. One of these good matching intra-storm trajectories is selected at random. The intra-storm trajectory is then adjusted (a) so that its storm peak value is equal to H_S^{sp*} (by multiplying the H_S component of the intra-storm trajectory by an appropriate scale factor), and (b) so that its storm peak direction corresponds to θ^{sp*} (by cyclic rotation of the directional component of the intra-storm trajectory), and c) by scaling the wave period T_Z such that the sea-state steepness from the original sample is retained. The adjusted intra-storm trajectory is then allocated to the current simulated storm peak.

Using this procedure, intra-storm trajectories are allocated to simulated storm peaks, ensuring that only (adjusted) intra-storm trajectories from the original sample with similar storm peak characteristics are used, but also incorporating the inherent variability in intra-storm trajectories with respect to given storm peak characteristics.

References

O J Aarnes, O Breivik, and M Reistad. Wave extremes in the northeast atlantic. *J. Climate*, 25:1529–1543, 2012.

C.W. Anderson, D.J.T. Carter, and P.D. Cotton. *Wave climate variability and impact on offshore design extremes*. Report commissioned from the University of Sheffield and Satellite Observing Systems for Shell International, 2001.

API. *API Recommended Practice 2A-WSD (RP 2A-WSD), Recommended Practice for Planning, Designing and Constructing Fixed Offshore Platforms - Working Stress Design*. API, 2005.

0 Breivik, O J Aarnes, J-R Bidlot, A Carrasco, and Oyvind Saetra. Wave extremes in the north east atlantic from ensemble forecasts. *J. Climate*, 26:7525–7540, 2013.

Y. Cai and D. E. Reeve. Extreme value prediction via a quantile function model. *Coastal Eng.*, 77: 91–98, 2013.

F. Calderon-Vega, A. O. Vazquez-Hernandez, and A. D. Garcia-Soto. Analysis of extreme waves with seasonal variation in the Gulf of Mexico using a time-dependent GEV model. *Ocean Eng.*, 73:68–82, 2013.

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

V. Chavez-Demoulin and A.C. Davison. Modelling time series extremes. *REVSTAT - Statistical Journal*, 10:109–133, 2012.

A. C. Davison, S. A. Padoan, and M. Ribatet. Statistical modelling of spatial extremes. *Statistical Science*, 27:161–186, 2012.

J. M. Dixon, J. A. Tawn, and J. M. Vassie. Spatial modelling of extreme sea-levels. *Environmetrics*, 9: 283–301, 1998.

E.F. Eastoe and J.A. Tawn. Modelling non-stationary extremes with application to surface level ozone. *Biometrika, doi: 10.1093/biomet/asr078*, 2012.

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

K. C. Ewans and P. Jonathan. The effect of directionality on northern North Sea extreme wave design criteria. *J. Offshore Mechanics Arctic Engineering*, 130:10, 2008.

L. Fawcett and D. Walshaw. Improved estimation for temporally clustered extremes. *Environmetrics*, 18:173–188, 2007.

C. A. T. Ferro and J. Segers. Inference for clusters of extreme values. *J. Roy. Statist. Soc. B*, 65:545–556, 2003.

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.

G. Z. Forristall. On the use of directional wave criteria. *J. Wtrwy., Port, Coast., Oc. Eng.*, 130:272–275, 2004.

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

P. Jonathan and K. C. Ewans. The effect of directionality on extreme wave design criteria. *Ocean Eng.*,

34:1977–1994, 2007.

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

P. Jonathan, P. H. Taylor, and P. S. Tromans. Storm waves in the northern North Sea. *Proc. 7th Intl. Conf. on the Behaviour of Offshore Structures, Massachusetts, USA*, 2:481, 1994.

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

P. Jonathan, D. Randell, Y. Wu, and K. Ewans. Return level estimation from non-stationary spatial data exhibiting multidimensional covariate effects. *(Accepted by Ocean Engineering July 2014, draft at www.lancs.ac.uk/∼jonathan)*, 2014.

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

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

F.J. Mendez, M. Menendez, A. Luceno, and I.J. Losada. Estimation of the long-term variability of extreme significant wave height using a time–dependent pot model. *Journal of Geophysical Research*, 11:C07024, 2006.

G. Muraleedharan, Claudia Lucas, C. Guedes Soares, N. Unnikrishnan Nair, and P.G. Kurup. Modelling significant wave height distributions with quantile functions for estimation of extreme wave heights. *Ocean Eng.*, 54:119–131, 2012.

M. Prevosto, H. E. Krogstad, and A. Robin. Probability distributions for maximum wave and crest heights. *Coastal Engineering*, 40:329–360, 2000.

D. Randell, Y. Wu, P. Jonathan, and K. C. Ewans. Omae2013-10187: Modelling covariate effects in extremes of storm severity on the Australian North West Shelf. *Proc. 32nd Conf. Offshore Mech. Arct. Eng.*, 2013.

M Reistad, O Breivik, H Haakenstad, O J Aarnes, B R Furevik, and J-R Bidlot. A high-resolution hindcast of wind and waves for the north sea, the norwegian sea, and the barents sea. *J. Geophys. Res.*, 116:1–18, 2011.

P Ruggiero, P D Komar, and J C Allan. Increasing wave heights and extreme value projections: The wave climate of the US pacific northwest. *Coastal Eng.*, 57:539–522, 2010.

C. Scarrott and A. MacDonald. A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT - Statistical Journal*, 10:33–60, 2012.

M.G. Scotto and C. Guedes-Soares. Modelling the long-term time series of significant wave height with non-linear threshold models. *Coastal Eng.*, 40:313, 2000.

M.G. Scotto and C. Guedes-Soares. Bayesian inference for long-term prediction of significant wave height. *Coastal Eng.*, 54:393, 2007.

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

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

P. Thompson, Y. Cai, R. Moyeed, D. Reeve, and J. Stander. Bayesian nonparametric quantile regression using splines. *Computational Statistics and Data Analysis*, 54:1138–1150, 2010.

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

P. S. Tromans, A. Anaturk, and P. Hagemeijer. A new model for the kinematics of large ocean waves - application as a design wave. *Proc. 1st Int. Offshore and Polar Engng. Conf. ISOPE.*, 1991.

J. L. Wadsworth and J. A. Tawn. Likelihood-based procedures for threshold diagnostics and uncertainty

in extreme value modelling. *J. Roy. Statist. Soc. B*, 2012.