

Revised version submitted to Ocean Engineering, December 2017

## Estimating surge in extreme North Sea storms

Emma Ross<sup>a</sup>, Samta Sam<sup>b</sup>, David Randell<sup>a</sup>, Graham Feld<sup>c</sup>, Philip Jonathan<sup>d,\*</sup>

<sup>a</sup>Shell Projects & Technology, 1031 HW Amsterdam, The Netherlands.

<sup>b</sup>Department of Computer Science, Royal Holloway, University of London TW20 0EX, United Kingdom.

<sup>c</sup>Shell Projects & Technology, Aberdeen AB12 3FY, United Kingdom.

<sup>d</sup>Shell Projects & Technology, London SE1 7NA, United Kingdom.

---

### Abstract

Characterising storm surge in extreme sea states is important in offshore design. Here we estimate key surge design characteristics such as the maximum surge observed during a storm event of given extreme storm severity and covariates such as storm direction. Inferences are made using a simple non-stationary implementation of the conditional extremes model of Heffernan and Tawn (2004), comprising a set of coupled piecewise stationary marginal and dependence models defined on a partition of the covariate domain.

The approach uses samples consisting of pairs of values for peaks over threshold of a conditioning variate, namely storm peak significant wave height, and the corresponding peaks over threshold of an associated value of a key surge characteristic. Each pair is allocated to a particular storm direction covariate interval, and all pairs within the same interval are assumed to exhibit common stationary marginal and conditional extreme value behaviour. Non-stationary marginal extreme value characteristics for each variate are estimated using maximum roughness-penalised generalised Pareto likelihood estimation over covariate intervals. Extremal dependence between variates on a transformed standard Gumbel scale is then estimated using maximum roughness-penalised likelihood estimation for a conditional extremes model, also piecewise stationary with respect to covariates. Sample and threshold uncertainties are quantified using a bootstrapping scheme. Marginal and conditional return value distributions, estimated using numerical integration, incorporate these uncertainties.

From offshore radar-based measurements at locations in the northern, central and southern North Sea, we observe that characteristics of surge and significant wave height vary with wave direction at all locations. Surge is a larger contributor to extreme seas in the southern North Sea than in the northern North Sea in particular. There is evidence that extremal dependence between surge characteristics and significant wave height also varies with storm direction. Hence, the size of surge contribution to return values of total water level also varies with storm direction and location.

*Keywords:* surge, significant wave height, conditional extremes, non-stationary, return value.

---

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

## 1. Introduction

The total extreme water level (TEWL) at a location during a storm event determines whether waves will impact topsides and supporting beams of an offshore structure at that location, causing large structural loads of considerable concern. Since TEWL is the sum of mean sea level, tide, storm surge and wave crest elevation, a joint model for extremes of all of these components is required in order to predict its extremal characteristics well. The physics of mean sea level, tide and surge has been studied for decades (e.g. Pugh 1987), as has that of wind waves (e.g. Kinsman 2012).

Considerable attention has also been devoted to estimation of storm surge during tropical storms (e.g. Needham et al. 2015) and extra-tropical storms (e.g. Bernier and Thompson 2006), and the interaction between surge and tide (e.g. Horsburgh and Wilson 2007, Olbert et al. 2013, Williams et al. 2016). Physically, storm surge is a large-scale increase in sea level due to a storm, lasting from hours to days over areas of hundreds of km<sup>2</sup>. The magnitude of storm surge is influenced by many physical processes, and those same processes control surge in both extra-tropical and tropical cyclones. Atmospheric pressure differences in ocean storms cause the water level in the open ocean to rise in regions of relatively low atmospheric pressure; Harris (1963) estimated an increase in sea level of the order of 1cm for every hPa reduction in atmospheric pressure. High wind stress at the sea surface and the horizontal gradient of atmospheric pressure are other important contributors, as is storm intensity and trajectory. In shallow and coastal waters, bathymetry and topography become important. In regions of high tidal range, the joint occurrence of storm and tidal high water is sometimes referred to as “storm-tide”. The SLOSH model of hurricane surge (Jelesnianski et al. 1992) predicts the maximum envelope of water at a location given hurricane characteristics such as central pressure, storm size, forward motion, track and maximum sustained winds. However, the SLOSH model does not explicitly model the aggregate impact of surge and crest. Bernier and Thompson (2006) seeks to predict the frequency of storm surges and extreme sea levels in the north-west Atlantic using a 2-D non-linear barotropic ocean model to construct a 40-year hindcast for the region. Where tide-surge interactions exist, they lead to lower total water levels than in the case of independence. The degree of decrease varies with extent of dependence, magnitude of surge peak at a particular phase of tide and the distribution of peaks over a tidal cycle. Coles and Tawn (2005a) identify strong seasonal effects in measurements of extreme surges at locations on the eastern coastline of the United Kingdom. Subsequently Coles and Tawn (2005b) propose a Bayesian model for extreme surge, and outline a procedure to estimate the distribution of the sum of surge and tide. The effect of future climate on North Sea storm surge is also of considerable interest (e.g. Gaslikova et al. 2013; Vousdoukas et al. 2017). Butler et al. (2007a) conduct trend estimation in synthetic storm surges. Butler et al. (2007b) reports extreme value analysis of decadal variation in storm surges, including a discussion of the effect of North Atlantic oscillation and tide-surge interaction.

The joint statistics of extreme storm surge and sea state severity at an arbitrary location in the ocean has also received some attention (e.g. Hawkes et al. 2002, Rueda et al. 2016 and Mazas and Hamm 2017). The objective of the current work is to establish a statistical model for surge characteristics conditional on the occurrence of extreme values of significant wave height ( $H_S$ ). Methods of multivariate extreme value analysis are required to describe the joint characteristics of extremes of multiple variables. The conditional extremes model of Heffernan and Tawn (2004) provides a flexible estimation framework. Given that the characteristics of extreme  $H_S$  at a location vary systematically with covariates such as storm direction, it is likely that the joint statistics of

surge and  $H_S$  are non-stationary with respect to covariates. Extensions to the conditional extremes model, such as Jonathan et al. (2014) allow estimation of multivariate extremal dependence in the presence of covariates, and hence estimation of tails of distributions of structure variables and design contours of interest. The conditional extremes model is motivated by an asymptotic form for the limiting conditional distribution of one or more conditioned random variables given a large value of a conditioning variable. An outline of the approach is given by Jonathan et al. (2010) in application to wave spectral characteristics. Conditions for the asymptotic argument to hold have been explored by Heffernan and Resnick (2007). An alternative to the conditional extremes model is a suitable extreme value copula (e.g. Gudendorf and Segers (2010)). Bender et al. (2014) and Serafin and Ruggiero (2014) propose bivariate extreme value models incorporating non-stationary marginal and dependence inference. The main advantage of the conditional extremes model compared with models such as those of Bender et al. (2014) and Serafin and Ruggiero (2014) is that the former incorporates a full class of asymptotic extremal dependence, and also allows relatively straightforward extension to higher dimensions.

The layout of the article is as follows. Section 2 introduces the North Sea application, and illustrates the data used in this work. Section 3 describes the piecewise stationary extreme value model; mathematical detail corresponding to the model are given in the Appendix. A conditional model for surge characteristics given extreme storm peak  $H_S$  is established in Section 4. Section 5 provides discussion and conclusions.

## 2. Motivating application

Data for the current work comes from two sources. The first is measurements of sea surface elevation at four offshore platforms in the North Sea using SAAB Rex wave radars. Measurements were taken at one location in the northern and central North Sea, and two locations in the southern North Sea (henceforth referred to as NNS, CNS, SNS1 and SNS2 for brevity). At each location, harmonic analysis of surface elevation time-series identified mean sea level, tidal variation and a non-tidal residual component interpreted as “surge” in the current work (see, e.g. Pugh 1987). Significant wave height  $H_S$  for sea states of 30 minutes duration was also estimated from the surface elevation time-series. Between 12 and 16 years of measurements were available each of the NNS, CNS, SNS1 and SNS2 locations. The second data source is the WAM North Sea hindcast (Reistad et al. 2009), from which time-series of wave direction corresponding to the surface elevation time-series at the four locations were extracted. Knowledge of wave direction is essential for reasonable characterisation of extreme storms in the North Sea; wave direction is used as a covariate in the analysis below.

Time-series of sea state  $H_S$  was then used to identify storm events, following the approach described in Ewans and Jonathan (2008); a total of between 1000 and 1100 storms were isolated at each of the NNS, CNS, SNS1 and SNS2 locations. For each storm event, (a) storm peak  $H_S$ , (b) wave direction at the time of storm peak, referred to as storm peak direction  $\theta$ , and (c) time-series of surge for the duration of the storm, referred to “the surge trajectory” were isolated. Figure 1 shows typical surge trajectories for each of the four locations.

The relative importance of surge effects is known to vary between the four locations. In the northern North Sea, surge generally makes a small contribution to severe seas, whereas the surge contribution is generally larger in the southern North Sea. Figure 1 suggests that the magnitude of surge increases with decreasing latitude in the North Sea. The relative importance of surge in the estimation of TEWL is therefore clearly greater for locations SNS1 and SNS2. Figure 1 also shows

that the lengths of storms, and hence of surge trajectories, varies considerably within and between locations. Moreover, it is clear that the time of maximum surge does not typically coincide with the time of peak storm severity, corresponding to zero on the x-axis in Figure 1.

[Figure 1 about here.]

For simplicity, each surge trajectory was next summarised in terms of four statistics for subsequent extreme value analysis, namely the maximum value ( $\text{SrgMxm}$ ), the negative of the minimum value, ( $\text{SrgNgtMnm}$ ), the median value ( $\text{SrgMdn}$ ) and the range ( $\text{SrgRng}$ ) of values per surge trajectory. We chose to model the negative of minimum surge since this shows non-negative dependence with increasing storm severity. Storm peak  $H_S$  is known to vary with storm peak direction  $\theta$ , and the importance of accommodating this non-stationarity in extreme value models has been established. For this reason, each bivariate sample of storm peak  $H_S$  and surge characteristic was partitioned by inspection of plots and consideration of likely fetch effects into covariate intervals (or directional sector for the current covariate) within which distributions of storm peak  $H_S$  and surge characteristic were found to be approximately stationary. Figure 2 shows scatter plots of the maximum surge  $\text{SrgMxm}$  observed per storm event against storm peak  $H_S$  per covariate interval for each of the four locations. Four covariate intervals were used for the NNS location, and three elsewhere. There are noteworthy differences between the characteristics of the various scatter plots. For SNS2,  $\text{SrgMxm}$  increases with storm peak  $H_S$  for directional sector  $[240, 0)$  to values above 1m. However, for the other two directional sectors, the mean value of  $\text{SrgMxm}$  corresponding to large  $H_S$  appears to be around zero.

[Figure 2 about here.]

Figure 3 shows scatter plots of  $\text{SrgRng}$  on storm peak  $H_S$  per directional sector per location. Again, there are differences between scatter plots due to location and storm direction.

[Figure 3 about here.]

Extreme value analysis of the samples of  $\text{SrgMxm}$ ,  $\text{SrgNgtMnm}$ ,  $\text{SrgMdn}$ ,  $\text{SrgRng}$ , storm peak  $H_S$  and direction  $\theta$  per location is discussed in Section 4.

### 3. The piecewise stationary extreme value model

#### *Overview*

Consider a sample  $\dot{D} = \{\dot{y}_{i1}, \dot{y}_{i2}\}_{i=1}^n$  of  $n$  pairs of values of peaks over threshold for a conditioning variate  $\dot{Y}_1 > 0$ , and an associated conditioned variate  $\dot{Y}_2 > 0$ . Further, let  $\{\mathbf{x}_i\}_{i=1}^n$  be the corresponding values of a covariate  $X$  on some domain  $\mathcal{X}$ ; in this work, we assume a single directional covariate with  $\mathcal{X} = [0^\circ, 360^\circ)$ . Extension to more complex covariate domains is straightforward compared to other approaches as discussed in Section 5. Our objective, using the sample, is to make inferences about joint structure of  $\dot{Y}_1$  and  $\dot{Y}_2$  for large values of  $\dot{Y}_1$  and likely non-stationarity with respect to  $X \in \mathcal{X}$ . The modelling procedure presented here is a refinement of the conditional extremes model of Heffernan and Tawn (2004), and proceeds in four steps: (a) marginal extreme value modelling of  $\dot{Y}_1$  and  $\dot{Y}_2$  given  $X$ , followed by (b) transformation to  $Y_1$  and  $Y_2$  given  $X$  with standard marginal distributions, (c) dependence modelling of  $Y_2|Y_1, X$  for large  $Y_1$  and (d) estimation of return values of  $\dot{Y}_2|\dot{Y}_1, X$  for large  $\dot{Y}_1$ . The procedure is outlined here, first for the general non-stationary case, then for the particular form of non-stationarity assumed in the piecewise stationary model. We then elaborate on the four stages individually for the piecewise stationary model in Sections 3.1-3.3. Mathematical details are provided in the Appendix.

### *General non-stationary case*

We estimate a marginal generalised Pareto model, non-stationary with respect to  $X$ , for threshold exceedances for each of  $\dot{Y}_1$  and  $\dot{Y}_2$ . For general variable  $\dot{Y}$  and threshold  $\psi > 0$ , the generalised Pareto cumulative distribution function is  $F_{GP}(\dot{y}; \xi, \sigma, \psi) = \mathbb{P}(\dot{Y} \leq \dot{y} | \dot{Y} > \psi, X = x) = 1 - (1 + (\xi/\sigma)(\dot{y} - \psi))^{-1/\xi}$  with  $\xi \in \mathbb{R}$ ,  $\sigma > 0$  and  $\dot{y} \in (\psi, \dot{y}^+)$  where  $\dot{y}^+ = \psi - \sigma/\xi$  when  $\xi < 0$  and  $\infty$  otherwise. In practice, threshold  $\psi$  might correspond to a local quantile of the sample given covariate with specified non-exceedance probability  $\tau \in (0, 1)$  estimated using non-stationary quantile regression. In practice, plots of parameter estimates from models corresponding to different thresholds aid the selection of  $\tau$  per variable showing appropriate behaviour, including approximate stability of the estimate of  $\xi$  for thresholds larger than the chosen  $\tau$ . Parameter estimates for  $\xi$ ,  $\sigma$  and  $\psi$  per variable are all in general functions of covariate  $X$ . Fitted marginal models are used to transform the variables  $\dot{Y}_1$  and  $\dot{Y}_2$  to  $Y_1$  and  $Y_2$  on standard Gumbel scale, and hence sample  $\dot{D}$  to  $D$ , using the probability integral transform.

Then we fit a conditional extremes model for  $Y_2|Y_1, X$  for various choices of high threshold  $\phi \in \mathbb{R}$  for the conditioning variate  $Y_1$ , retaining the estimated model parameters and residuals.  $\phi$  is typically a quantile of the standard Gumbel distribution with non-exceedance probability  $\kappa \in (0, 1)$ . Plots of model parameter estimates and residual distributions for different threshold choices aid selection of  $\kappa$  consistent with modelling assumptions. The conditional extremes model form is  $(Y_2|Y_1 = y, X = x) = \alpha y + y^\beta W$  for  $y > \phi$ , where  $\alpha \in [0, 1]$  and  $\beta \in (-\infty, 1]$ .  $W$  is a random variable with unknown distribution, the density of which we estimate using residuals from the fitted model. For fitting purposes only, we assume that  $W \sim N(\mu, \zeta^2)$  with  $\mu \in \mathbb{R}$  and  $\zeta^2 > 0$ . The model fitting procedure corresponds to estimating  $\{\alpha, \beta, \mu, \zeta\}$  given a sample of values for  $Y_1, Y_2, X$  and threshold  $\phi$ . All of  $\phi, \alpha, \beta, \mu$  and  $\zeta$  are in general functions of covariate  $X$ . Using the marginal and conditional extremes models for the distribution of threshold exceedances, and empirical sample-based distributions for non-exceedances, numerical integration is used to estimate the distribution of  $Y_2|Y_1, X$  for large  $Y_1$ , and hence that of  $\dot{Y}_2|\dot{Y}_1, X$  for large values of  $\dot{Y}_1$ . This permits estimation of the conditional return value distribution, namely the distribution of  $\dot{Y}_2$  given occurrences of the  $T$ -year return values of  $\dot{Y}_1$ .

### *Piecewise stationary case*

The piecewise stationary model uses this approach for a particularly simple description of non-stationarity with respect to covariates in marginal and dependence models. For each observation in the sample, the value of covariate  $x_i$  is used to allocate the observation to one and only one of  $m$  covariate intervals  $\{C_k\}_{k=1}^m$  by means of an allocation vector  $A$  such that  $k = A(i)$  and  $\mathcal{X} = \bigcup C_k$ . For each  $k$ , all observations in the set  $\{\dot{y}_{1i'}, \dot{y}_{2i'}\}_{A(i')=k}$  with the same covariate interval  $C_k$  are assumed to have common joint extreme value characteristics.

#### *3.1. Marginal extremes and transformation to Gumbel scale*

Non-stationary marginal extreme value characteristics of each variate are estimated in turn using a generalised Pareto model and cross-validated roughness-penalised maximum likelihood estimation, as outlined below and in the Appendix. For variable  $j \in \{1, 2\}$  and covariate interval  $C_k$ , the extreme value threshold  $\psi_{jk} > 0$  is assumed to be a quantile of the empirical distribution of the variate in that interval, with specified non-exceedance probability  $\tau_j \in (0, 1)$ , with  $\tau_j$  constant across intervals, and estimated by counting. Threshold exceedances are assumed to follow the

generalised Pareto distribution with shape  $\xi_j \in \mathbb{R}$  and scale  $\sigma_{jk} > 0$ , with cumulative distribution function

$$F_{GP}(\dot{y}; \xi_j, \sigma_{jk}, \psi_{jk}) = 1 - (1 + (\xi_j/\sigma_{jk})(\dot{y} - \psi_{jk}))^{-1/\xi_j}$$

where  $\dot{y} \in (\psi_{jk}, \dot{y}_{jk}^+)$  with  $\dot{y}_{jk}^+ = \psi_{jk} - \sigma_{jk}/\xi_j$  when  $\xi_j < 0$  and  $\infty$  otherwise. Since estimation of shape parameter is particularly problematic,  $\xi_j$  is assumed constant (but unknown) across covariate intervals, and the reasonableness of the assumption assessed by inspection of diagnostic plots. Parameters  $\xi_j, \{\sigma_{jk}\}$  are estimated by maximising the predictive performance of a roughness-penalised model, optimally regulating the extent to which  $\{\sigma_{jk}\}$  varies across interval, using a cross-validation procedure. Equations for marginal model estimation are given in the Appendix.

The standard Gumbel distribution has cumulative distribution function  $F_G(y) = \exp(-\exp(-y))$  for  $y \in \mathbb{R}$ . We transform the non-stationary sample  $\dot{D} = \{\dot{y}_{i1}, \dot{y}_{i2}\}$  on generalised Pareto scale to the corresponding stationary standard Gumbel sample  $D = \{y_{i1}, y_{i2}\}$  using the probability integral transform, by defining  $\{y_{i1}, y_{i2}\}$  such that  $F_G(y_{ij}) = F_{GP}(\dot{y}_{ij}; \xi_j, \sigma_{jA(i)}, \psi_{jA(i)})$  per covariate interval; the transformed sample is then assumed to be a stationary sample from a standard Gumbel distribution. Estimation of marginal extreme value models for the North Sea application is discussed in Section 4.1. We note that transformation to Laplace scale provides a useful alternative to Gumbel transformation, especially when considering negative tail dependence (e.g. Keef et al. 2013).

### 3.2. Conditional extremes

The Gumbel-scale sample  $D = \{y_{i1}, y_{i2}\}$  above some threshold of the conditioning variate  $Y_1$  is used to estimate a conditional extremes model with parameters  $\alpha_k \in [0, 1]$  and  $\beta \in (-\infty, 1]$  given by

$$(Y_2|Y_1 = y_{i1}) = \alpha_{A(i)}y_{i1} + y_{i1}^\beta W \text{ for } y > \phi,$$

where  $W \sim N(\mu, \zeta^2)$  with  $\mu \in \mathbb{R}$  and  $\zeta^2 > 0$  is assumed for model estimation only. Threshold  $\phi \in \mathbb{R}$  is defined as the quantile of the standard Gumbel distribution with non-exceedance probability  $\kappa \in (0, 1)$ . Note that, since identifying  $\beta$  is particularly problematic, it is assumed constant (but unknown) across covariate intervals. Similarly, since they are intended to describe a generic residual structure, both  $\mu$  and  $\zeta^2$  are also assumed to be stationary with respect to covariate. Once more, the reasonableness of these modelling assumptions is assessed by inspection of diagnostic plots. Estimation of the conditional extremes model for the North Sea application is discussed in Section 4.2.

### 3.3. Marginal and conditional return values

From an engineering perspective, the main inferences from the current analysis are estimates for marginal and conditional return values corresponding to some long return period  $T$ . Typically, these are obtained by Monte Carlo simulation under the models outlined in Sections 3.1 and 3.2. Specifying a Monte Carlo scheme, even for relatively complex simulations, is generally straightforward. When  $T$  is small (of the order of 100 years, say), simulation is also computationally easy: Monte Carlo estimation might therefore be preferred. When  $T$  is large however, with  $T = 10^7$  years of interest for some analysis, Monte Carlo simulation becomes computationally tedious and demanding. In these circumstances, numerical integration schemes yield dramatic reductions in the computational complexity of return value estimation, as outlined in Ross et al. (2017). Specification

of the numerical integration scheme usually requires some effort (as illustrated in the Appendix for the current work), but computationally this effort pays dividends.

We use numerical integration to estimate cumulative distribution functions for marginal and conditional return values as follows: we estimate (a) the marginal return value distributions  $F_{M_{1k}}$  and  $F_{M_{2k}}$  for a  $T$ -year return period of  $\dot{Y}_1$  and  $\dot{Y}_2$  per covariate interval  $C_k$ , (b) the corresponding ‘‘omni-directional’’ return value distributions  $F_{M_1}$  and  $F_{M_2}$  over all covariate intervals, (c) the conditional return value distribution  $F_{2|M_{1k}}$  of  $\dot{Y}_2$  given the occurrence of a  $T$ -year event of  $\dot{Y}_1$  in covariate interval  $C_k$ , and (d) the conditional return value distribution  $F_{2|M_1}$  of  $\dot{Y}_2$  given the occurrence of an omni-directional  $T$ -year event of  $\dot{Y}_1$  (regardless of the covariate interval in which the  $T$ -year event of  $\dot{Y}_1$  occurs). Details of the numerical integration schemes used to estimate (a)-(d) are given in the Appendix; estimates of the different return value distributions for the North Sea application are illustrated in Section 4.3.

### 3.4. Uncertainty quantification

In practice, the complete modelling procedure is repeated for  $n_{BS}$  bootstrap resamples  $\{\dot{D}^b\}$  of the original sample  $\dot{D}$  to capture sampling uncertainty. For each sample  $\dot{D}^b$ , to capture marginal threshold specification uncertainty, marginal models for  $\dot{Y}_1, \dot{Y}_2$  are evaluated for marginal thresholds  $\psi_1, \psi_2$  with non-exceedance probabilities  $\tau_1, \tau_2$  drawn at random from the intervals  $\mathcal{I}_{\tau_1}, \mathcal{I}_{\tau_2}$  (both  $\subseteq (0, 1)$ ) on which model performance is deemed reasonable from inspection of diagnostics. Following transformation to standard Gumbel scale sample  $D^b$ , to capture conditional threshold uncertainty, a conditional extremes model is then evaluated for dependence threshold  $\phi$  with non-exceedance probabilities  $\kappa$  drawn at random from interval  $\mathcal{I}_{\kappa} \subseteq (0, 1)$  on which the conditional model fit is deemed adequate. Computationally, thus, the estimate for each marginal and dependence parameter corresponds to a  $n_{BS}$  array of values capturing sampling and threshold specification uncertainty, used for uncertainty quantification in parameter and return value inference as described in the Appendix.

## 4. Application to North Sea locations

The characteristics of **SrgMxm**, **SrgNgtMnm**, **SrgMdn** and **SrgRng** given large values of storm peak  $H_S$  and direction  $\theta$  per location are estimated using the piecewise stationary model. For brevity, the analysis procedure from Section 3 is described in Sections 4.1 and 4.2 and illustrated in Figures 4-10 and Figure 12 for the case of **SrgMxm** at the CNS location only. Estimates for marginal and conditional return value distributions are discussed in Section 4.3 and illustrated in Figures 11 for all surge characteristics at all locations. Further, since the analysis involves storm peak significant wave height, as opposed to significant wave height for individual sea states, we use ‘‘ $H_S$ ’’ to indicate storm peak significant wave height, unless clearly indicated to the contrary.

### 4.1. Marginal extremes

We conduct conditional extreme value analysis of **SrgMxm** and (storm peak)  $H_S$ , non-stationary with respect to storm peak direction  $\theta$ . As outlined in Section 3, the first stage of analysis involves marginal extreme value modelling for  $H_S$  and **SrgMxm**, illustrated in Figure 4. To facilitate this, covariate intervals were selected within which the joint characteristics of **SrgMxm** and  $H_S$  were observed, from inspection of diagnostic plots, to be approximately stationary. In this case, the three covariate intervals selected correspond approximately to the land shadow of the United Kingdom

( $\theta \in [160, 270)$ ), open water to the Atlantic and Norwegian Seas ( $\theta \in [270, 40)$ ), and the land shadow of the European mainland from Norway to France ( $\theta \in [40, 160)$ ). The piecewise stationary generalised Pareto model is applied independently to threshold exceedances for each variable, with thresholds per variable specified as quantiles of the marginal distribution per directional interval of  $\theta$ , with non-exceedance probabilities  $\tau_1 \in \mathcal{I}_{\tau_1}$  and  $\tau_2 \in \mathcal{I}_{\tau_2}$ . Intervals  $\mathcal{I}_{\tau_1}$ ,  $\mathcal{I}_{\tau_2}$  were specified by inspection of diagnostic plots for model adequacy over a wide range of values of  $\tau$  (e.g. Figure 7), as discussed below.

[Figure 4 about here.]

Figures 4(a, d) show estimated thresholds for  $H_S$  and **SrgMxm** in terms of bootstrap median and 95% uncertainty bands (over  $n_{BS} = 500$  bootstrap resamples of the original sample, with  $\mathcal{I}_{\tau_1} = [0.7, 0.8]$  and  $\mathcal{I}_{\tau_2} = [0.65, 0.75]$ ). The original sample is shown in terms of threshold exceedances (black) and non-exceedances (grey). For  $H_S$ , there appears to be little variability in threshold level across covariate intervals, The largest two  $H_S$  events correspond to storms emanating from the east. The threshold for **SrgMxm** shows clearer directional variability. Surge effects are larger when storms emanate from approximately the south, and are smaller for storms from the east. For each bootstrap resample of each variable, generalised Pareto models are estimated. Figures 4(b, e) give bootstrap distributions for the estimated generalised Pareto shape parameter  $\xi$ , set constant with respect to  $\theta$ . For storm peak  $H_S$ ,  $\xi$  is centred at approximately -0.15; but for **SrgMxm**, the bootstrap distribution is centred at approximately zero, indicating that the distribution of **SrgMxm** has a longer tail than that of  $H_S$ . Figures 4(c, f) show the estimate for generalised Pareto scale parameter  $\sigma$  in terms of its bootstrap median and 95% uncertainty band per covariate interval. For  $H_S$ , the estimate for  $\sigma$  is larger in the covariate interval corresponding to largest values of  $H_S$ . For **SrgMxm**, the estimate of  $\sigma$  is lower in the sector corresponding to smaller values of **SrgMxm**.

To confirm model fit, tail diagnostic plots such as those illustrated in Figures 5 and 6 for  $H_S$  and **SrgMxm** are inspected. Figure 5 shows that the tail of empirical distribution of  $H_S$  from the sample is in agreement with the estimated generalised Pareto tail, represented in terms of a bootstrap median and 95% uncertainty band, for each covariate interval and ‘‘omni-directionally’’ over all covariate intervals (in the bottom right hand plot of Figure 5).

[Figure 5 about here.]

[Figure 6 about here.]

It is important also to establish that the estimate for generalised Pareto shape  $\xi$  is approximately constant with threshold for all thresholds corresponding to intervals  $\mathcal{I}_{\tau_1}$ ,  $\mathcal{I}_{\tau_2}$ . Figure 7 shows this variation for the **SrgMxm** model as a function of threshold non-exceedance probability  $\tau_2$ . For values of  $\tau_2 \in [0.65, 0.75]$ , the estimate of  $\xi$  appears to be constant for larger values of  $\tau_2$  to within bootstrap uncertainty. For this reason, we set  $\mathcal{I}_{\tau_2} = [0.65, 75]$  for marginal modelling of **SrgMxm** at the CNS location. Diagnostic plots such as those illustrated in Figures 5-7 were inspected to confirm adequacy of marginal fits for  $H_S$  and all surge characteristics at all locations.

[Figure 7 about here.]

#### 4.2. Conditional extremes

For each bootstrap resample  $\hat{D}^b$  of the original sample  $\hat{D}$  and random choices of  $\tau_1 \in \mathcal{I}_{\tau_1}$  and  $\tau_2 \in \mathcal{I}_{\tau_2}$ , marginal models for **SrgMxm** and  $H_S$  are estimated and used to transform the resample



to standard Gumbel scale as explained in Section 3. The Gumbel-scale sample  $D^b$  is then used to estimate the conditional extremes model for thresholds of the conditioning variate corresponding with non-exceedance probabilities  $\kappa \in \mathcal{I}_\kappa$ . Inspection of fit residuals (Figure 8) and parameter stability with  $\kappa$  (Figure 10) suggested setting  $\mathcal{I}_\kappa = [0.55, 0.65]$  for **SrgMxm** on  $H_S$  at the CNS location. With this choice, the characteristics of residuals  $\{r_h\}$  (see Appendix) are illustrated in Figure 8. The left hand panel suggests that residuals are centred around zero, with a unimodal distribution showing a longer right hand tail. The right hand panel suggests that the distribution of residuals is approximately stationary with respect to  $\theta$ . These diagnostic plots suggest that the piecewise stationary model framework, with  $\beta$ ,  $\mu$  and  $\zeta$  constant with respect to  $\theta$ , is reasonable.

[Figure 8 about here.]

Parameter estimates for the conditional extremes model are shown in Figure 9; only slope parameter  $\alpha$  ( $= \{\alpha_k\}$ ) varies between intervals in the piecewise stationary model. The top left panel summarises  $\alpha$  in terms of bootstrap median and 95% uncertainty bands per as a function of  $\theta$ . There is evidence that  $\alpha$  is lower for the covariate interval  $[45, 150]$  (corresponding to the land-shadow of mainland Europe) than elsewhere. The estimate of  $\alpha$  is considerably less than unity everywhere, suggesting that **SrgMxm** and  $H_S$  show asymptotic independence (e.g. Eastoe et al. 2013). The estimate of scale exponent parameter  $\beta$  is centred around 0.3, suggesting that the width of the distribution of **SrgMxm** given  $H_S$  increases with increasing  $H_S$ . Residual mean  $\mu$  and scale  $\zeta$  are centred around 0.2 and 1.1.

[Figure 9 about here.]

Figure 10 explores the stability of estimates for  $\{\alpha_k\}$  and  $\beta$  as a function of conditional extremes threshold non-exceedance probability  $\kappa \in \mathcal{I}_\kappa$ . Estimates for  $\alpha_k$  in each covariate interval, and  $\beta$  are all stationary to within bootstrap uncertainty, suggesting that a reasonable choice of  $\mathcal{I}_\kappa$  has been made.

[Figure 10 about here.]

#### 4.3. Marginal and conditional return values

Using the numerical integration approach outlined in Section 3.3 and the Appendix, marginal and conditional return values were estimated for a  $T = 100$  year return period. The left hand panel of Figure 11 gives estimates for the marginal omni-directional 100-year return value distribution  $F_{M_1}$  of storm peak  $H_S$  at each of the NNS (red), CNS (green), SNS1 (cyan) and SNS2 (purple) locations. The median 100-year return value is approximately 15m for NNS, 12m for CNS and 4m for SNS1 and SNS2 as might be expected. The return value distributions are of course rather broad, reflecting both inherent aleatory uncertainty as well as epistemic uncertainty (estimated by bootstrapping); for example, we would not be too surprised to observe 100-year return values of storm peak  $H_S$  at the NNS location in excess of 17m. The centre and right hand panels of the figure give the corresponding omni-directional marginal ( $F_{M_2}$ , solid) and conditional ( $F_{2|M_1}$ , dashed) return value estimates, one panel for each of the four surge characteristics of interest.

For **SrgMxm** (top centre of Figure 11), marginal surge is greatest in the southern North Sea, with median 100-year values approaching 2m, compared with approximately 1m at CNS and 0.5m at NNS; these observations are consistent with the trends in Figure 2 and physical intuition. It is noteworthy however that the conditional distribution of **SrgMxm** given the occurrence of omni-directional

100-year  $H_S$  yields considerably lower estimates; for NNS and CNS, the median conditional surge is approximately zero. However, for SNS the median conditional surge remains at around 1.5m, approximately 80% of the marginal 100-year maximum surge of 1.8m. This demonstrates that consideration of surge (and conditional surge in particular) is much more important for design in extreme seas of this region of the North Sea. Characteristics of the corresponding return value distributions for  $\text{SrgMdn}$ ,  $\text{SrgNgtMnm}$  and  $\text{SrgRng}$  are generally rather similar. Of particular interest is the fact that the conditional distributions of  $\text{SrgNgtMnm}$  at SNS1 and SNS2 are considerably to the left of corresponding marginal distributions; this is not the case for other surge characteristics.

[Figure 11 about here.]

Figure 11 shows that differences exist between omnidirectional marginal and corresponding conditional return value distributions. Figure 12 illustrates, at the CNS location, that directional marginal and conditional return value distributions have interesting differences also. The top panel for the figure gives directional marginal return value distributions  $\{F_{M_{1k}}\}$  for (storm peak)  $H_S$  for  $T = 100$  years, together with the omnidirectional distribution ( $F_{M_1}$ , black). The omnidirectional return value almost coincides with the directional distribution for the  $\theta \in [45, 150)$  covariate interval. Return values for the two remaining covariate intervals are smaller. The second panel shows that the omnidirectional return value distribution for  $\text{SrgMxm}$  has large contributions from all covariate intervals except for  $\theta \in [45, 150)$ . This has an interesting effect on the location of the omnidirectional conditional return value distribution (dashed black). Because the most severe storms correspond to  $\theta \in [45, 150)$ , for which  $\text{SrgMxm}$  is relatively small, the omnidirectional conditional return value distribution is itself similar to that for the  $[45, 150)$  covariate interval, with median value around 0.1m. However, the conditional return value distributions of  $\text{SrgMxm}$  for covariate intervals  $[270, 45)$  and  $[150, 270)$  are considerably to the right of the omnidirectional conditional distribution. Informally, we might interpret this information for design in the following manner. For estimation of TEWL (ignoring mean sea level and tide) omnidirectionally,  $H_S$  has a 100-year median of around 11.3m, to which a median associated  $\text{SrgMxm}$  contribution of 0.1m is added, yielding TEWL of 11.4m. However, for the  $[270, 45)$  covariate interval,  $H_S$  has a median return value of approximately 10.0m, to which a median associated  $\text{SrgMxm}$  contribution of 0.5m is added, yielding TEWL of 10.5m. That is,  $\text{SrgMxm}$  makes a greater associated contribution to TEWL in covariate interval  $[270, 45)$  than it does omnidirectionally. We also see, informally, that naive addition of median omnidirectional marginal  $H_S$  of around 11.3m to median omnidirectional marginal  $\text{SrgMxm}$  of around 1.1m yields TEWL of 12.4m; this estimate is 1.0m larger than the more realistic estimate based on median marginal  $H_S$  and median associated  $\text{SrgMxm}$ .

[Figure 12 about here.]

## 5. Discussion and conclusions

In this work, we examine the statistical properties of storm surge in extreme North Sea storms using a non-stationary conditional extreme value model consisting of a set of coupled piecewise stationary models on a partition of the domain of a storm direction covariate. We find evidence for variation of marginal and conditional extreme surge characteristics with storm direction and location. In particular, we provide an efficient numerical integration algorithm to estimate marginal and conditional return value distributions of surge characteristics. Sampling and threshold selection uncertainties are incorporated within inferences using a bootstrapping scheme. The piecewise

stationary model appears to be adequate in terms of its complexity for the samples examined here; the whole modelling procedure can be completed for one surge characteristic at one location in approximately 5 minutes.

The approach requires specification of covariate partition  $\{C_k\}$  prior to analysis. In the current work, choice of covariate partition boundaries was informed by inspection of data, and elementary physical understanding of likely directional effects on significant wave height and surge at each location. Sensitivity of inferences to small changes in locations of partition boundaries was investigated empirically, and found generally to be small. An improved procedure for specification of the covariate partition could be considered, including estimation of the partition boundaries (and in principle the number of boundaries) as part of the extreme value inference. This was not attempted here, to avoid introducing undue complexity in model estimation.

Extension of the piecewise stationary model to higher  $p$ -dimensional covariates ( $p > 1$ ) is relatively straightforward: simply define  $\{C_k\}$  as a partition of the covariate domain in  $p$ -dimensions. For the typical size  $n$  and quality of samples available for analysis, identification of models with  $p > 2$  is likely to be problematic. We anticipate that inclusion of season at covariate may be of benefit. Currently, the marginal directional threshold for  $H_S$  effectively eliminates all but winter storms. Adopting a directional-seasonal threshold (and piecewise stationary model) might provide the opportunity to examine seasonality in more detail.

Inferences concerning key surge characteristics such as **SrgMxm**, **SrgMdn** and **SrgNgtMnm** can be used to forecast a complete surge trajectory  $s^*(t)$  for some interval of time  $t$  corresponding to an extreme storm with specified storm peak  $H_S$  and direction  $\theta$ . We might choose to do this by adjusting a historical surge trajectory  $s(t)$  for a storm with similar  $H_S$  and  $\theta$  linearly using

$$s^*(t) = \begin{cases} s_o^* + \frac{s(t) - s_o}{s_+ - s_o}(s_+^* - s_o^*) & s(t) > s_o \\ s_o^* - \frac{s(t) - s_o}{s_o - s_-}(s_-^* - s_o^*) & s(t) \leq s_o, \end{cases}$$

where  $s_-$ ,  $s_o$  and  $s_+$  are (negative) **SrgNgtMnm**, **SrgMdn** and **SrgMxm** for the selected historical surge trajectory, and  $s_-^*$ ,  $s_o^*$  and  $s_+^*$  are the corresponding quantities for the forecast surge trajectory, estimated jointly using the piecewise stationary model. An alternative strategy would be to estimate a model similar to that of Winter and Tawn (2016) and Winter and Tawn (2017) for the growth and decay of surge in time with respect to its maximum value in an extreme storm. In this case, the piecewise stationary model would be used to estimate the surge maximum relative to which surge decays (forwards and backwards in time), and the surge minimum necessary for estimating a threshold below which the surge trajectory should not extend.

The method introduced here is applicable to estimation of conditional characteristics of (in principle arbitrary numbers of) environmental variables. It can be used for specification of associated design conditions for (e.g.) wind speed or peak wave period given extreme  $H_S$ , and in the construction of design contours (e.g. Huseby et al. 2015, Vanem 2017). The method is currently being used to estimate surge characteristics for extreme storms in other ocean basins.

## Acknowledgements

We acknowledge early contributions to this work from Mihaela Paun (University of Glasgow, UK) and Shiraz Basheer (Royal Holloway, University of London, UK and latterly Shell). We

further thank Vadim Anokhin (Shell), Jonathan Tawn (Lancaster University, UK), Stan Tendijck (Technical University of Delft, The Netherlands) and Yanyun Wu (Shell) for their support. The development of the piecewise stationary extreme value model was partly motivated and funded by the European Union ERA-NET project entitled “Environmental Contours for SAfe DEsign of Ships and other marine structures (ECSADES)”. MATLAB (2017) code is available from the authors.

## Appendix

This section contains details of the underlying statistical calculations made. Using the notation of Section 3, we first describe the estimation of marginal and conditional extreme value models, then the estimation of marginal and conditional return value distributions.

### *Marginal extreme value model*

Continuing from Section 3.1, the marginal sample generalised Pareto likelihood  $\mathcal{L}_j$  for variable  $j$  under the piecewise stationary model is

$$\mathcal{L}_j = \prod_{k=1}^m \prod_{\substack{i:A(i)=k; \\ \hat{y}_{ij} > \psi_{jk}}} \frac{1}{\sigma_{jk}} \left[ 1 + \frac{\xi_j}{\sigma_{jk}} [\hat{y}_{ij} - \psi_{jk}] \right]^{-1/\xi_j - 1},$$

where  $\mathcal{L}_j$ ,  $\{\psi_{jk}\}$ ,  $\xi_j$  and  $\{\sigma_{jk}\}$  are functions of marginal extreme value threshold non-exceedance probability  $\tau_j$ , and  $\xi_j$  is constant across the  $m$  intervals  $\{C_k\}$ . The negative log likelihood, penalised for the roughness of  $\{\sigma_{jk}\}$  across intervals, is then

$$\ell_j^* = -\log_e \mathcal{L}_j + \lambda_{\sigma_j} \left( \frac{1}{m} \sum_{k=1}^m \left[ \sigma_{jk} - \frac{1}{m} \sum_{k'=1}^m \sigma_{jk'} \right]^2 \right),$$

where  $\ell_j^*$  is a function of both  $\tau$  and roughness coefficient  $\lambda_{\sigma_j}$ . For given  $\tau$  and  $\lambda_{\sigma_j}$ , estimates for marginal model parameters  $\xi_j$  and  $\{\sigma_{jk}\}$  are found by minimising  $\ell_j^*$ . A random 10-fold cross-validation is then used to select the value  $\hat{\lambda}_{\sigma_j}$  of  $\lambda_{\sigma_j}$  and corresponding  $\hat{\xi}_j$ ,  $\{\hat{\sigma}_{jk}\}$  which, for each  $\tau_j$ , maximise predictive performance. These parameter estimates are used for subsequent inference. Note that the “hat” notation is suppressed below in reference to the use of estimates of marginal parameters for subsequent inference.

### *Conditional extreme value model*

Continuing from Section 3.2, since  $W$  is assumed to be Gaussian-distributed for fitting, the negative log-likelihood of the piecewise stationary conditional extremes model is

$$\tilde{\ell} = \frac{1}{2} \sum_{\substack{i: \\ y_{i1} > \phi}} \left[ \log(2\pi\zeta^2) + \frac{1}{\zeta^2} \left( \frac{y_{i2} - \alpha_{A(i)}y_{i1}}{y_{i1}^\beta} - \mu \right)^2 \right]$$

for each value of threshold  $\phi$  with non-exceedance probability  $\kappa$ . As for the marginal case, we penalise the roughness of  $\{\alpha_k\}$  using roughness penalisation and 10-fold cross-validation. The penalised negative log-likelihood, for given  $\kappa$ , is

$$\tilde{\ell}^* = \tilde{\ell} + \lambda_\alpha \left( \frac{1}{m} \sum_{k=1}^m \left[ \alpha_k - \frac{1}{m} \sum_{k'=1}^m \alpha_{k'} \right]^2 \right).$$

We estimate optimal roughness coefficient  $\hat{\lambda}_\alpha$  and corresponding parameter estimates  $\{\hat{\alpha}_k\}$ ,  $\hat{\beta}$ ,  $\hat{\mu}$  and  $\hat{\zeta}^2$ . For all subsequent inference, for given  $\kappa$ , the distribution of  $W$  is then represented empirically by the sample  $\{r_h\}$  of  $n_r$  residuals from the fit

$$r_h = y_{h1}^{-\hat{\beta}} (y_{h2} - \hat{\alpha}_{A(h)}y_{h1}),$$

for all  $h$  such that  $y_{h1} > \phi$ . In particular, the cumulative distribution function of  $W$  is assumed to be  $F_W(w) = 1/n_r \sum_h I(r_h \leq w)$ , where  $I$  is an indicator function returning unity when its argument is true, and zero otherwise.

*Marginal return value estimation*

Continuing from Section 3.3, let  $F_{jk}(\dot{y})$  be the cumulative distribution function of a storm peak event  $\dot{Y}_j$  in covariate interval  $C_k$  for response indexed by  $j \in \{1, 2\}$ . For threshold non-exceedance probability  $\tau_j$ , this is estimated as

$$F_{jk}(\dot{y}) = \begin{cases} \tau_j F_{Ej}(\dot{y}|C_k) & \dot{y} \leq \psi_{jk} \\ \tau_j + (1 - \tau_j) F_{GjP}(\dot{y}|\xi_j, \sigma_{jk}, \psi_{jk}) & \text{otherwise,} \end{cases}$$

where  $F_{Ej}(\dot{y}|C_k) = 1/\tilde{n}_k \sum_{i:k=A(i)} I(\dot{y}_{ij} \leq \psi_{jk})$  is an empirical estimate for threshold non-exceedances, and  $\tilde{n}_k$  is the number of non-exceedances in interval  $C_k$ . If  $\rho_{jk} > 0$  is the number of storm peak events in the interval per annum, estimated empirically from the sample, we estimate the cumulative distribution function  $F_{M_{jk}}(y)$  of the maximum  $M_{jk}$  observed in a period of  $T$  years in  $C_k$  as

$$\begin{aligned} F_{M_{jk}}(y) &= \mathbb{P}(M_{jk} < y) \\ &= \sum_{p=0}^{\infty} \mathbb{P}\left(p \text{ peaks of } \dot{Y}_j \text{ in } C_k \text{ in } T \text{ years}\right) \times \mathbb{P}^p(\text{size of an event in } C_k < y) \\ &= \sum_{p=0}^{\infty} \frac{(T\rho_{jk})^p}{p!} \exp(-T\rho_{jk}) \times F_{jk}^p(y) \\ &= \exp(-T\rho_{jk}(1 - F_{jk}(y))). \end{aligned}$$

Since storm peak events are assumed to be independent given covariate, we estimate the cumulative distribution function  $F_{M_j}(\dot{y})$  of the ‘‘omni-directional’’ storm peak maximum  $M_j$  over all intervals, by taking the product

$$F_{M_j}(\dot{y}) = \prod_{k=1}^m F_{M_{jk}}(\dot{y}).$$

As noted in Section 3.4, penalised maximum likelihood estimates  $\widehat{F}_{M_j}(\dot{y}|D^b)$  for  $F_{M_j}(\dot{y})$  are available for  $n_{BS}$  bootstrap resamples  $\{D^b\}$  of the original sample  $D$ , each  $D^b$  using a random selection of thresholds  $\tau_1 \in \mathcal{I}_{\tau_1}$  and  $\tau_2 \in \mathcal{I}_{\tau_2}$  for marginal estimation. The final estimate  $\widehat{F}_{M_j}(\dot{y})$  is then a model average with respect to these, given by

$$\widehat{F}_{M_j}(\dot{y}) = \frac{1}{n_{BS}} \sum_b \widehat{F}_{M_j}(\dot{y}|D^b).$$

*Conditional return value estimation*

We estimate  $F_{2|M_1}(\dot{y}_2) = \mathbb{P}(\dot{Y}_2 \leq \dot{y}_2|M_1)$ , namely the conditional distribution of  $\dot{Y}_2$  given an occurrence of an omni-directional  $T$ -year maximum of  $\dot{Y}_1$ . Note that the covariate interval within which  $M_1$  occurs is not known, and neither is the value of  $M_1$ . We can write  $F_{2|M_1}(\dot{y}_2)$  as

$$F_{2|M_1}(\dot{y}_2) = \int_{\dot{y}_1} \mathbb{P}(\dot{Y}_2 \leq \dot{y}_2|M_1 = \dot{y}_1) f_{M_1}(\dot{y}_1) d\dot{y}_1,$$

where  $f_{M_1}(\dot{y}_1)$  is the probability density function of  $M_1$ , which can be seen by differentiation of the cumulative distribution function  $F_{M_j}(\dot{y})$  given in the previous section to be

$$f_{M_1}(\dot{y}_1) = \sum_k \left[ f_{M_{1k}}(\dot{y}_1) \prod_{k' \neq k} F_{M_{1k'}}(\dot{y}_1) \right].$$

The  $k^{\text{th}}$  term in the square brackets above corresponds to the occurrence of  $M_1 = \dot{y}_1$  in the  $k^{\text{th}}$  covariate interval  $C_k$ . To see this informally, note that this event requires  $M_{1k} = \dot{y}_1$  and  $M_{1k'} \leq \dot{y}_1$  for  $k' \neq k$ ; the corresponding probability is  $f_{M_{1k}}(\dot{y}_1) d\dot{y}_1 \times \prod_{k' \neq k} F_{M_{1k'}}(\dot{y}_1)$ . The expression for  $F_{2|M_1}(\dot{y}_2)$  now becomes

$$\begin{aligned} F_{2|M_1}(\dot{y}_2) &= \int_{\dot{y}_1} \left( \sum_k \mathbb{P}(\dot{Y}_2 \leq \dot{y}_2 | M_1 = \dot{y}_1) \left[ f_{M_{1k}}(\dot{y}_1) \prod_{k' \neq k} F_{M_{1k'}}(\dot{y}_1) \right] \right) d\dot{y}_1 \\ &= \int_{\dot{y}_1} \left( \sum_k \mathbb{P}(\dot{Y}_2 \leq \dot{y}_2 | M_1 = \dot{y}_1 \text{ in } C_k) \left[ f_{M_{1k}}(\dot{y}_1) \prod_{k' \neq k} F_{M_{1k'}}(\dot{y}_1) \right] \right) d\dot{y}_1. \end{aligned}$$

Using  $\{g_{jk}\}$  to represent the set of functions which map  $\dot{Y}_j$  to  $Y_j$  for intervals  $\{C_k\}$  in Section 3.1, we have

$$\begin{aligned} \mathbb{P}(\dot{Y}_2 \leq \dot{y}_2 | M_1 = \dot{y}_1 \text{ in } C_k) &= \mathbb{P}(g_{2k}(\dot{Y}_2) \leq g_{2k}(\dot{y}_2) | g_{1k}(M_1) = g_{1k}(\dot{y}_1) \text{ in } C_k) \\ &= \mathbb{P}(\alpha_k g_{1k}(\dot{y}_1) + g_{1k}(\dot{y}_1)^\beta W \leq g_{2k}(\dot{y}_2)) \\ &= F_W \left( \frac{g_{2k}(\dot{y}_2) - \alpha_k g_{1k}(\dot{y}_1)}{g_{1k}(\dot{y}_1)^\beta} \right), \end{aligned}$$

where residuals from fitting the conditional extremes model provide an estimate for  $F_W$ .

As for marginal return values, penalised maximum likelihood estimates  $\widehat{F}_{2|M_1}(\dot{y}_2 | D^b)$  (and also  $\{\widehat{F}_{2|M_{1k}}(\dot{y}_2 | D^b)\}$ ) of conditional return values are available for  $n_{BS}$  bootstrap resamples  $\{D^b\}$  of the original sample  $D$ , each  $D^b$  using random selections of thresholds  $\tau_1 \in \mathcal{I}_{\tau_1}$  and  $\tau_2 \in \mathcal{I}_{\tau_2}$  for marginal estimation, and  $\kappa \in \mathcal{I}_\kappa$  for conditional extremes estimation. The final estimate  $\widehat{F}_{2|M_1}(\dot{y}_2)$  is again a model average given by

$$\widehat{F}_{2|M_1}(\dot{y}_2) = \frac{1}{n_{BS}} \sum_b \widehat{F}_{2|M_1}(\dot{y}_2 | D^b),$$

with analogous expressions for  $\{\widehat{F}_{2|M_{1k}}(\dot{y}_2)\}$ .

## References

- Bender, J., Wahl, T., Jensen, J., 2014. Multivariate design in the presence of non-stationarity. *Journal of Hydrology* 514, 123–130.
- Bernier, N. B., Thompson, K. R., 2006. Predicting the frequency of storm surges and extreme sea levels in the northwest atlantic. *J. Geophys. Res.* 111, C10009.
- Butler, A., Heffernan, J. E., Tawn, J. A., Flather, R. A., 2007a. Trend estimation in extremes of synthetic North Sea surges. *J. Roy. Statist. Soc. C* 56, 395–414.
- Butler, A., Heffernan, J. E., Tawn, J. A., Flather, R. A., Horsburgh, K. J., 2007b. Extreme value analysis of decadal variations in storm surge elevations. *Journal of Marine Systems* 67, 189–200.
- Coles, S. G., Tawn, J. A., 2005a. Bayesian modelling of extreme sea surges on the UK east coast. *Phil. Trans. R. Soc. A* 363, 1387–1406.
- Coles, S. G., Tawn, J. A., 2005b. Seasonal effects of extreme surges. *Stoch. Environ. Res. Risk Assess.* 19, 417–427.
- Eastoe, E., Koukoulas, S., Jonathan, P., 2013. Statistical measures of extremal dependence illustrated using measured sea surface elevations from a neighbourhood of coastal locations. *Ocean Eng.* 62, 68–77.
- Ewans, K. C., Jonathan, P., 2008. The effect of directionality on northern North Sea extreme wave design criteria. *J. Offshore. Arct. Eng.* 130, 041604:1–041604:8.
- Gaslikova, L., Grabemann, I., Groll, N., 2013. Changes in North Sea storm surge conditions for four transient future climate realizations. *Natural Hazards* 66, 1501–1518.
- Gudendorf, G., Segers, J., 2010. Extreme value copulas. In: Jaworski, P., Durante, F., Hardle, W. K., Rychlik, T. (Eds.), *Copula theory and its applications. Lecture notes in statistics.* Springer Berlin Heidelberg, pp. 127–145.
- Harris, D. L., 1963. *Characteristics of the hurricane storm surge.* Washington, D.C.: Dept. of Commerce, Weather Bureau.
- Hawkes, P. J., Gouldby, B. P., Tawn, J. A., Owen, M. W., 2002. The joint probability of waves and water levels in coastal defence design. *Journal of Hydraulic Research* 40, 241–251.
- Heffernan, J. E., Resnick, S. I., 2007. Limit laws for random vectors with an extreme component. *Ann. Appl. Probab.* 17, 537–571.
- Heffernan, J. E., Tawn, J. A., 2004. A conditional approach for multivariate extreme values. *J. R. Statist. Soc. B* 66, 497–546.
- Horsburgh, K. J., Wilson, C., 2007. Tide-surge interaction and its role in the distribution of surge residuals in the north sea. *J. Geophys. Res. C* 112, CO8003.
- Huseby, A., Vanem, E., Natvig, B., 2015. Alternative environmental contours for structural reliability analysis. *Struct. Saf.* 54, 32–45.



- Jelesnianski, C. P., Chen, J., Shaffer, W. A., 1992. SLOSH: Sea, lake, and overland surges from hurricanes. NOAA Technical Report NWS 48, National Oceanic and Atmospheric Administration, U. S. Department of Commerce.  
URL <http://www.nhc.noaa.gov/surge/slosh>
- Jonathan, P., Ewans, K. C., Randell, D., 2014. Non-stationary conditional extremes of northern North Sea storm characteristics. *Environmetrics* 25, 172–188.
- Jonathan, P., Flynn, J., Ewans, K. C., 2010. Joint modelling of wave spectral parameters for extreme sea states. *Ocean Eng.* 37, 1070–1080.
- Keef, C., Papastathopoulos, I., Tawn, J. A., 2013. Estimation of the conditional distribution of a vector variable given that one of its components is large: additional constraints for the Heffernan and Tawn model. *J. Mult. Anal.* 115, 396–404.
- Kinsman, B., 2012. *Wind waves: Their generation and propagation on the ocean surface*. Dover, New York.
- MATLAB, 2017. Version 9.2. The MathWorks Inc., Natick, MA.
- Mazas, F., Hamm, L., 2017. An event-based approach for extreme joint probabilities of waves and sea levels. *Coastal Engineering* 122, 44–59.
- Needham, H. F., Keim, B. D., Sathiaraj, D., 2015. A review of tropical cyclone-generated storm surges: global data sources, observations, and impacts. *Rev. Geophys.* 53, 545–591.
- Olbert, A. I., Nash, S., Cunnane, C., Hartnett, M., 2013. Tide–surge interactions and their effects on total sea levels in Irish coastal waters. *Ocean Dynamics* 63, 599–614.
- Pugh, D. T., 1987. *Tides, surges and mean sea level. A handbook for engineers and scientists*. Wiley.
- Reistad, M., Breivik, O., Haakenstad, H., Aarnes, O. J., Furevik, B. R., 2009. A high-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea and the Barents Sea. Norwegian Meteorological Institute.
- Ross, E., Randell, D., Ewans, K., Feld, G., Jonathan, P., 2017. Efficient estimation of return value distributions from non-stationary marginal extreme value models using Bayesian inference. *Ocean Eng.* 142, 315–328.
- Rueda, A., Camus, P., Toms, A., Vitousek, S., Mndez, F., 2016. A multivariate extreme wave and storm surge climate emulator based on weather patterns. *Ocean Modelling* 104, 242–251.
- Serafin, K. A., Ruggiero, P., 2014. Simulating extreme total water levels using a time-dependent, extreme value approach. *Journal Geophys. Res. - Oceans* 119 (9), 6305–6329.
- Vanem, E., 2017. A comparison study on the estimation of extreme structural response from different environmental contour methods. *Mar. Struct.* 56, 137–162.
- Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., Feyen, L., 2017. Extreme sea levels on the rise along europe’s coasts. *Earth’s Future* 5, 304–323.

- Williams, J., Horsburgh, K. J., Williams, J. A., Proctor, R. N. F., 2016. Tide and skew surge independence: new insights for flood risk. *Geophys. Res. Lett.* 43, 6410–6417.
- Winter, H. C., Tawn, J. A., 2016. Modelling heatwaves in central France: a case-study in extremal dependence. *J. Roy. Statist. Soc. C* 65, 345–365.
- Winter, H. C., Tawn, J. A., 2017.  $k$ th-order Markov extremal models for assessing heatwave risks. *Extremes* 20, 393–415.

## List of Figures

1	Surge trajectories for a northern North Sea location (NNS), a central North Sea location (CNS) and two southern North Sea locations (SNS1, SNS2). Time on the x-axis is given in days relative to the time of occurrences of storm peak $H_S$ . . . . .	21
2	Scatter plots of surge maximum (SrgMxm) on storm peak $H_S$ for each of NNS, CNS, SNS1 and SNS2 locations, for each directional sector identified. Four covariate intervals used for the NNS location, and three elsewhere. . . . .	22
3	Scatter plots of surge range (SrgRng) on storm peak $H_S$ for each of NNS, CNS, SNS1 and SNS2 locations, for each directional sector identified. Two covariate intervals used for the NNS location, and three elsewhere. . . . .	23
4	Marginal piecewise stationary modelling for storm peak $H_S$ ((a)-(c)) and maximum surge (SrgMxm, (d)-(f)) at the CNS location. Panel (a) shows the estimated non-stationary extreme value threshold $\psi$ for storm peak $H_S$ with direction $\theta$ in terms of its bootstrap median and 95% uncertainty bands in blue; storm peak $H_S$ events exceeding $\psi$ with $\theta$ are shown as black discs; non-exceedances are shown in grey. Boundaries of covariate intervals are shown in red. Panel (b) gives a histogram of bootstrap estimates for the generalised Pareto shape parameter $\xi$ for storm peak $H_S$ , assumed stationary with $\theta$ . Panel (c) shows the estimated non-stationary generalised Pareto scale parameter $\sigma$ for storm peak $H_S$ with $\theta$ in terms of its bootstrap median and 95% uncertainty bands in black. Panels (d)-(f) show the corresponding plots for SrgMxm. . . . .	24
5	Marginal model validation for storm peak $H_S$ by comparison of estimates for the distribution of storm peak $H_S$ corresponding to the period of the original sample, plotted as $\log_{10}(1 - F)$ to accentuate tail behaviour, for cumulative distribution function $F$ at the CNS location. Panels give comparisons for three directional sectors $[270, 45)$ , $[45, 150)$ and $[150, 270)$ and omni-directionally $[0, 360)$ . In each panel, the dotted red curve is an empirical estimate. The black curves summarise the predictive distribution of the quantile estimate (for given tail probability $1 - F$ ) under the piecewise stationary model, as the median (solid) and 2.5% and 97.5% values (dashed), estimated using numerical integration. . . . .	25
6	Marginal model validation for maximum surge (SrgMxm) by comparison of estimates for the distribution of storm peak $H_S$ corresponding to the period of the original sample, plotted as $\log_{10}(1 - F)$ to accentuate tail behaviour, for cumulative distribution function $F$ at the CNS location. Panels give comparisons for three directional sectors $[270, 45)$ , $[45, 150)$ and $[150, 270)$ and omni-directionally $[0, 360)$ . In each panel, the dotted red curve is an empirical estimate. The black curves summarise the predictive distribution of the quantile estimate (for given tail probability $1 - F$ ) under the piecewise stationary model, as the median (solid) and 2.5% and 97.5% values (dashed), estimated using numerical integration. . . . .	26
7	Variation of estimated generalised Pareto shape parameter $\xi$ for SrgMxm with non-exceedance threshold $\tau$ at the CNS location, estimated using 500 bootstrap resamples uniformly sampled on the interval $\tau \in [0.5, 0.95]$ , with bootstrap median and 95% uncertainty band in red. . . . .	27

8	Conditional extremes model validation for maximum surge ( <b>SrgMxm</b> ) given storm peak $H_S$ at the CNS location. The left hand panel shows a histogram of residuals $r = \{r_h\}$ from the conditional extremes model fit. The right hand panel shows residuals with direction $\theta$ . . . . .	28
9	Piecewise stationary conditional extremes modelling for maximum surge ( <b>SrgMxm</b> ) on storm peak $H_S$ at the CNS location. Top left panel shows estimated slope parameter $\alpha = \{\alpha_k\}$ with direction $\theta$ in terms of its bootstrap median and 95% uncertainty band. Subsequent panels show histograms of bootstrap estimates for scale exponent parameter $\beta$ , and for residual mean $\mu$ and scale $\zeta$ , all assumed stationary with $\theta$ . . .	29
10	Piecewise stationary conditional extremes modelling for maximum surge ( <b>SrgMxm</b> ) on storm peak $H_S$ at the CNS location. Estimates of the conditional extremes slope parameter $\alpha = \{\alpha_k\}$ as a function of conditional extremes threshold non-exceedance probability $\kappa$ in each of the directional sectors [270, 45), [45, 150) and [150, 270), and estimate of conditional extremes scale exponent $\beta$ assumed stationary with direction. . . . .	30
11	Omni-directional marginal and conditional cumulative distribution functions for the $T = 100$ -year return period estimated by numerical integration for all locations. Left: marginal storm peak $H_S$ . Centre and right: marginal ( $Y$ , solid) and conditional ( $Y$ given 100-year $H_S$ , dashed) return value distributions for maximum ( <b>SrgMxm</b> ), median ( <b>SrgMdn</b> ) and negative minimum ( <b>SrgNgtMnm</b> ) surge and surge range ( <b>SrgRng</b> ). Colours indicate locations as explained in the left hand panel. . . . .	31
12	Directional marginal and conditional cumulative distribution functions for the 100-year return period estimated by numerical integration at the CNS location. First row: marginal storm peak $H_S$ . Subsequent rows: marginal ( $Y$ , solid) and conditional ( $Y$ given 100-year $H_S$ , dashed) return value distributions for maximum ( <b>SrgMxm</b> ), median ( <b>SrgMdn</b> ) and negative minimum ( <b>SrgNgtMnm</b> ) surge and surge range ( <b>SrgRng</b> ). Colours indicate covariate intervals. . . . .	32

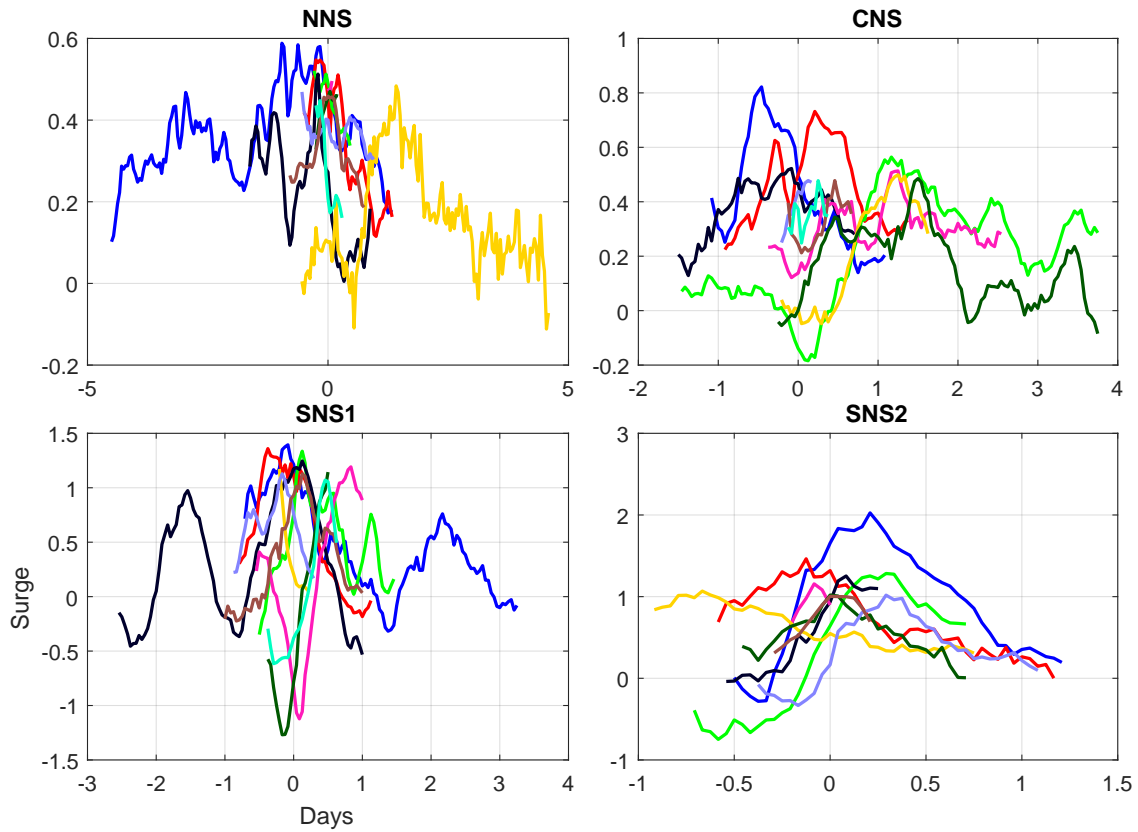


Figure 1: Surge trajectories for a northern North Sea location (NNS), a central North Sea location (CNS) and two southern North Sea locations (SNS1, SNS2). Time on the x-axis is given in days relative to the time of occurrences of storm peak  $H_S$ .

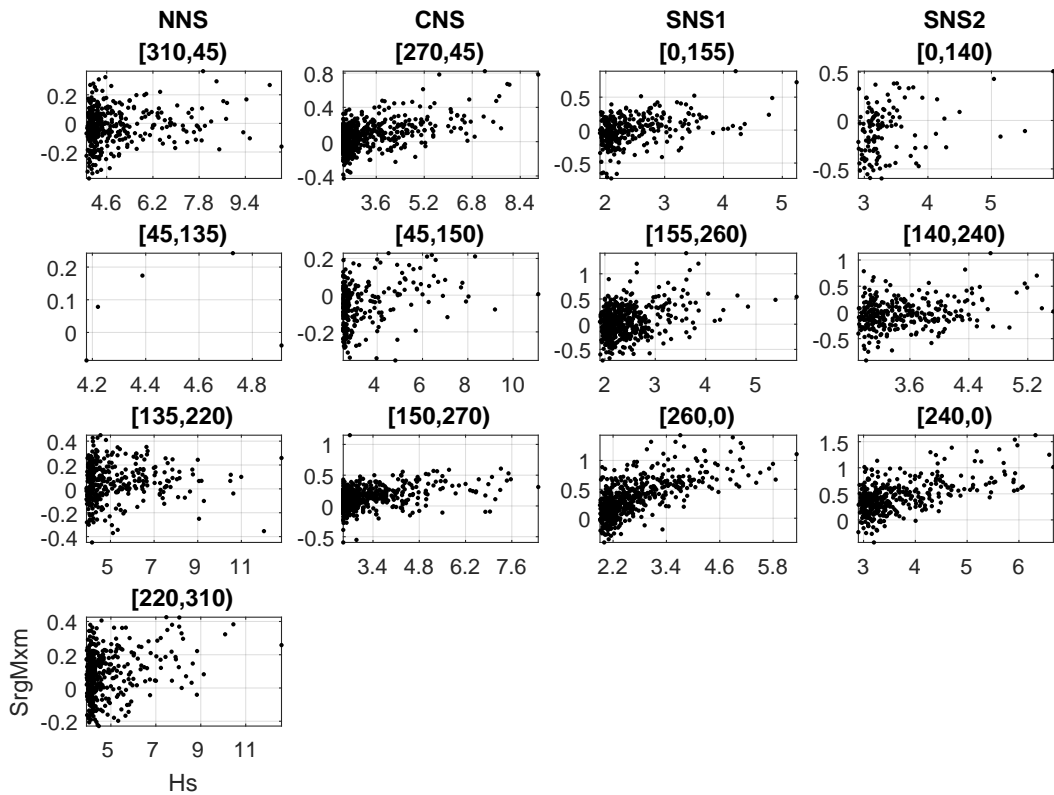


Figure 2: Scatter plots of surge maximum ( $SrgMxm$ ) on storm peak  $H_S$  for each of NNS, CNS, SNS1 and SNS2 locations, for each directional sector identified. Four covariate intervals used for the NNS location, and three elsewhere.

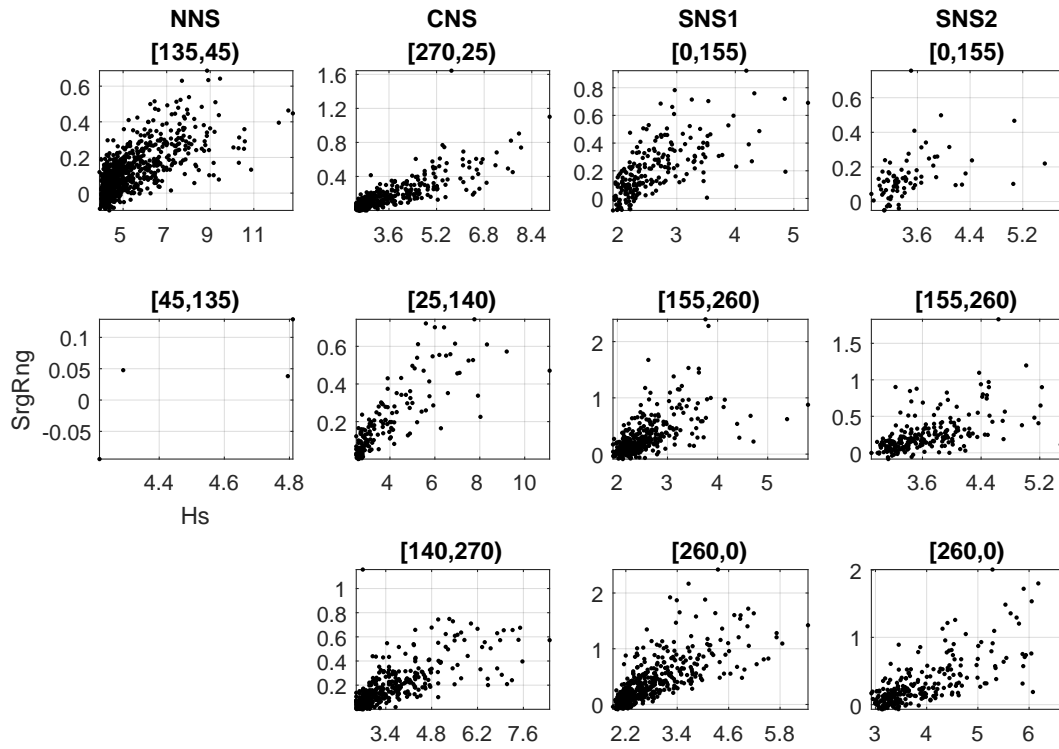


Figure 3: Scatter plots of surge range (SrgRng) on storm peak  $H_S$  for each of NNS, CNS, SNS1 and SNS2 locations, for each directional sector identified. Two covariate intervals used for the NNS location, and three elsewhere.

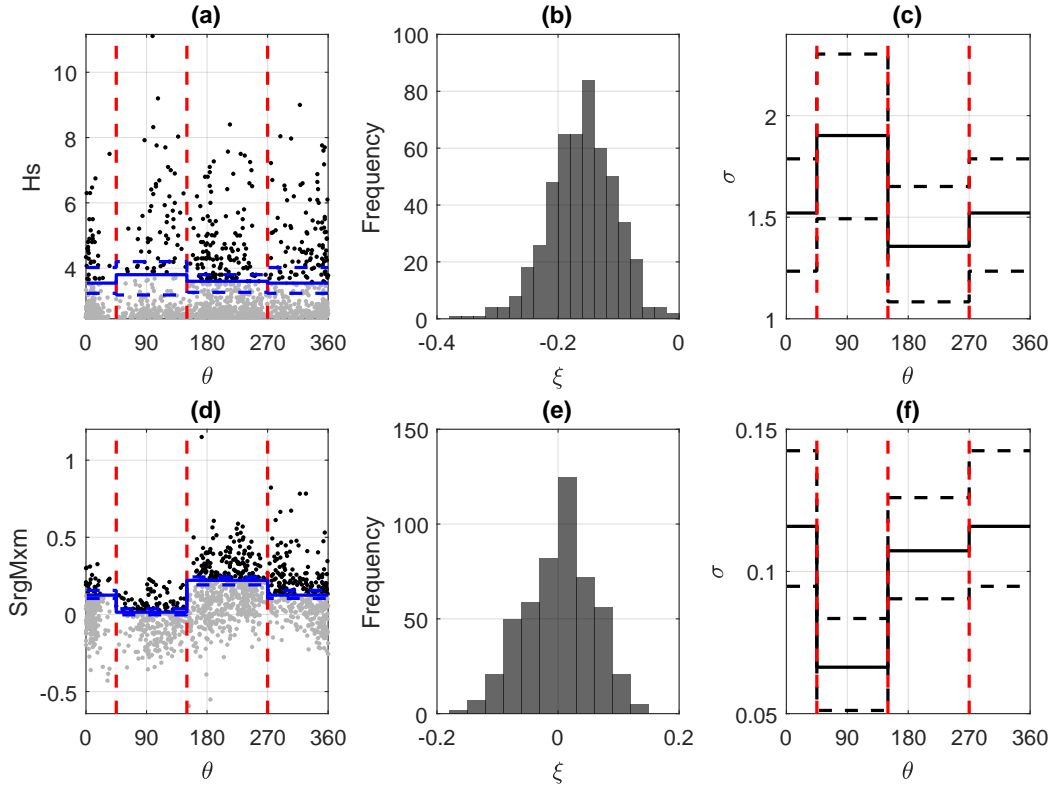


Figure 4: Marginal piecewise stationary modelling for storm peak  $H_S$  ((a)-(c)) and maximum surge (SrgMxm, (d)-(f)) at the CNS location. Panel (a) shows the estimated non-stationary extreme value threshold  $\psi$  for storm peak  $H_S$  with direction  $\theta$  in terms of its bootstrap median and 95% uncertainty bands in blue; storm peak  $H_S$  events exceeding  $\psi$  with  $\theta$  are shown as black discs; non-exceedances are shown in grey. Boundaries of covariate intervals are shown in red. Panel (b) gives a histogram of bootstrap estimates for the generalised Pareto shape parameter  $\xi$  for storm peak  $H_S$ , assumed stationary with  $\theta$ . Panel (c) shows the estimated non-stationary generalised Pareto scale parameter  $\sigma$  for storm peak  $H_S$  with  $\theta$  in terms of its bootstrap median and 95% uncertainty bands in black. Panels (d)-(f) show the corresponding plots for SrgMxm.



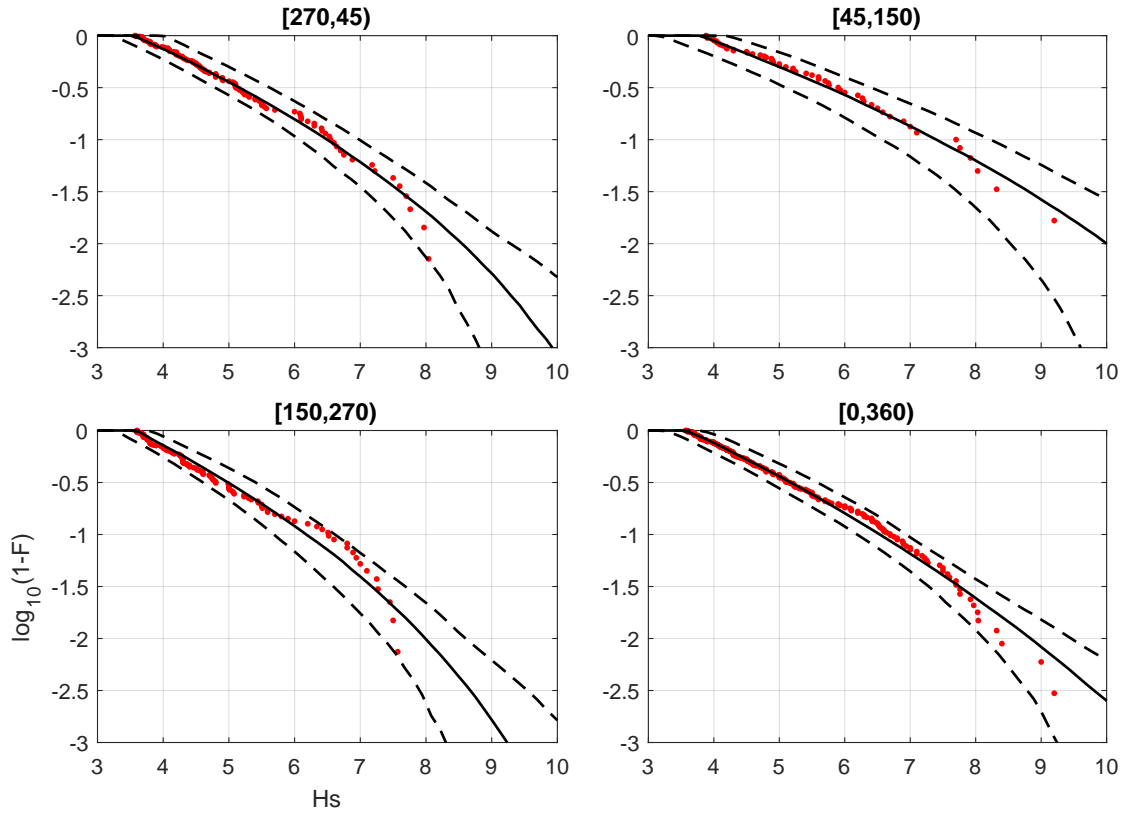


Figure 5: Marginal model validation for storm peak  $H_S$  by comparison of estimates for the distribution of storm peak  $H_S$  corresponding to the period of the original sample, plotted as  $\log_{10}(1 - F)$  to accentuate tail behaviour, for cumulative distribution function  $F$  at the CNS location. Panels give comparisons for three directional sectors [270, 45), [45, 150) and [150, 270) and omni-directionally [0, 360). In each panel, the dotted red curve is an empirical estimate. The black curves summarise the predictive distribution of the quantile estimate (for given tail probability  $1 - F$ ) under the piecewise stationary model, as the median (solid) and 2.5% and 97.5% values (dashed), estimated using numerical integration.

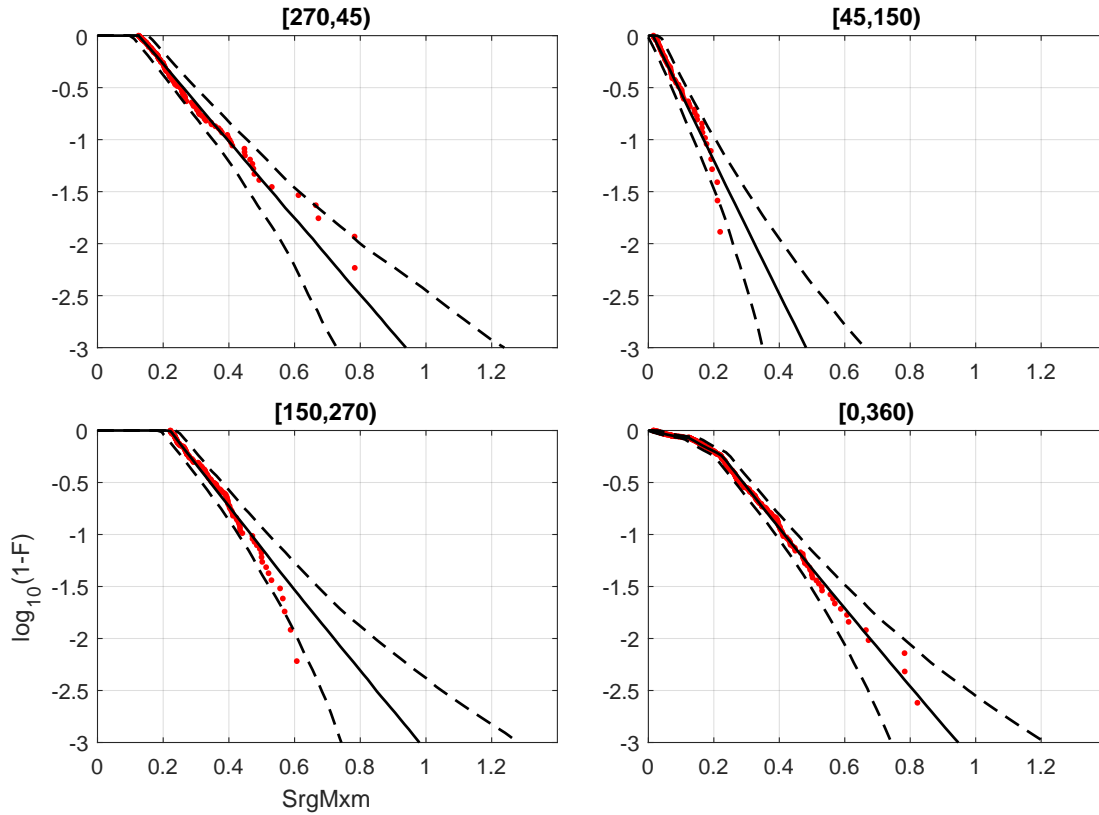


Figure 6: Marginal model validation for maximum surge (SrgMxm) by comparison of estimates for the distribution of storm peak  $H_S$  corresponding to the period of the original sample, plotted as  $\log_{10}(1 - F)$  to accentuate tail behaviour, for cumulative distribution function  $F$  at the CNS location. Panels give comparisons for three directional sectors [270, 45), [45, 150) and [150, 270) and omni-directionally [0, 360). In each panel, the dotted red curve is an empirical estimate. The black curves summarise the predictive distribution of the quantile estimate (for given tail probability  $1 - F$ ) under the piecewise stationary model, as the median (solid) and 2.5% and 97.5% values (dashed), estimated using numerical integration.

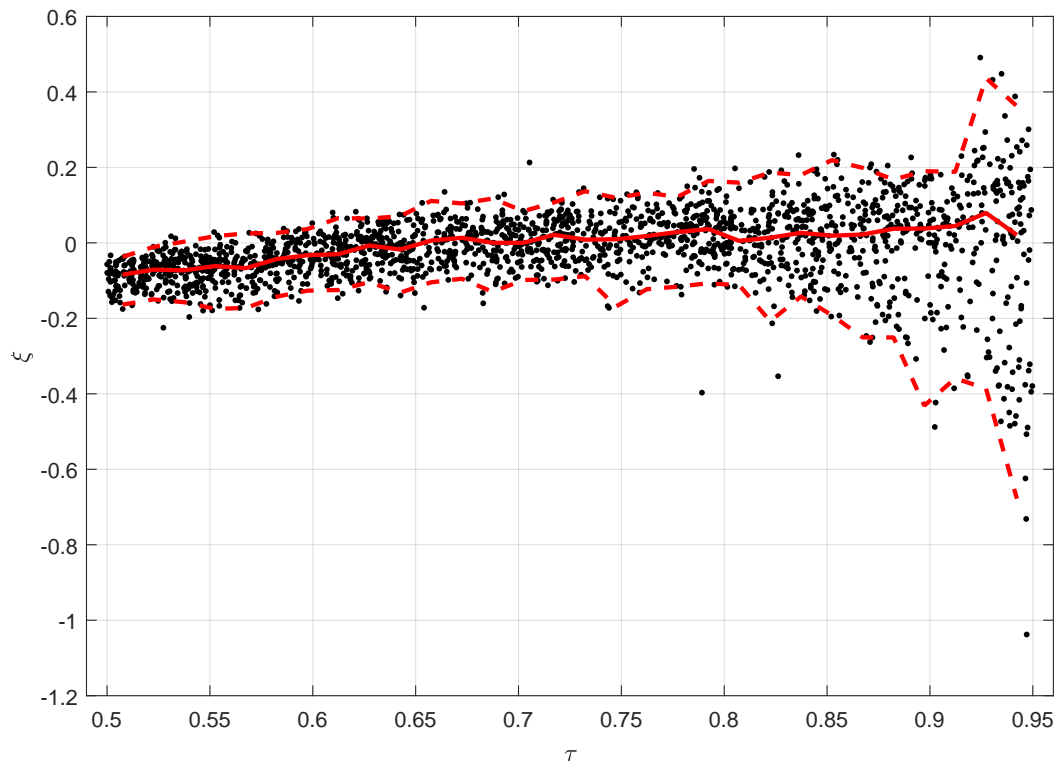


Figure 7: Variation of estimated generalised Pareto shape parameter  $\xi$  for `SrgMxm` with non-exceedance threshold  $\tau$  at the CNS location, estimated using 500 bootstrap resamples uniformly sampled on the interval  $\tau \in [0.5, 0.95]$ , with bootstrap median and 95% uncertainty band in red.

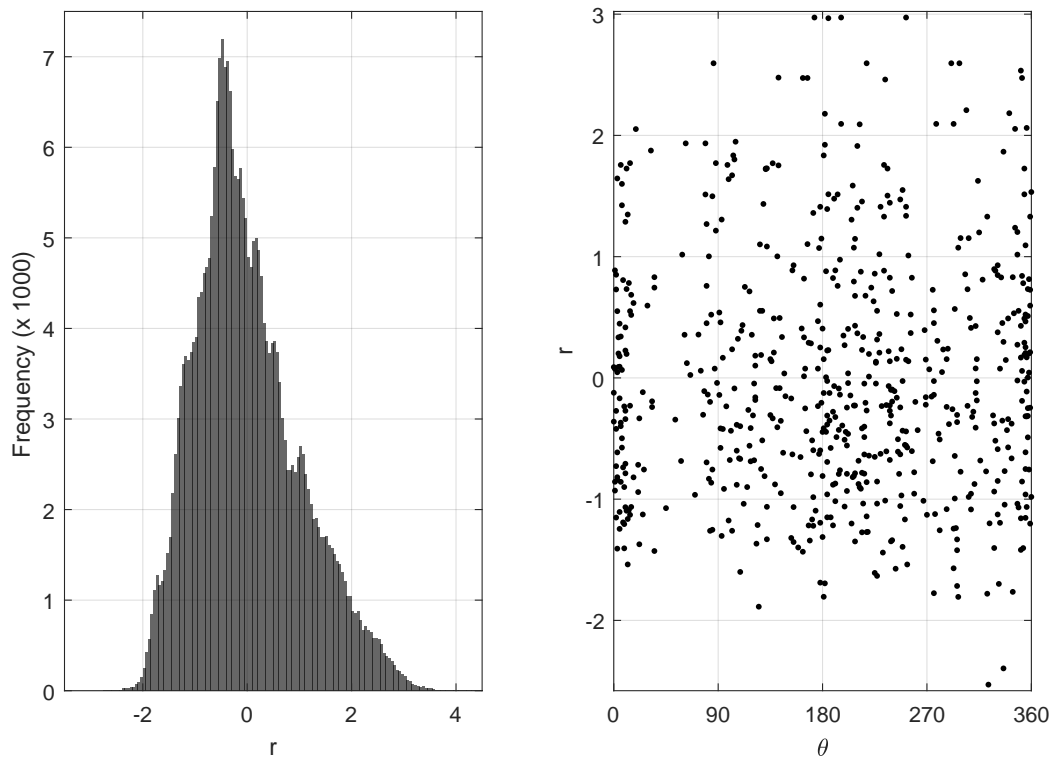


Figure 8: Conditional extremes model validation for maximum surge ( $SrgMxm$ ) given storm peak  $H_S$  at the CNS location. The left hand panel shows a histogram of residuals  $r = \{r_h\}$  from the conditional extremes model fit. The right hand panel shows residuals with direction  $\theta$ .

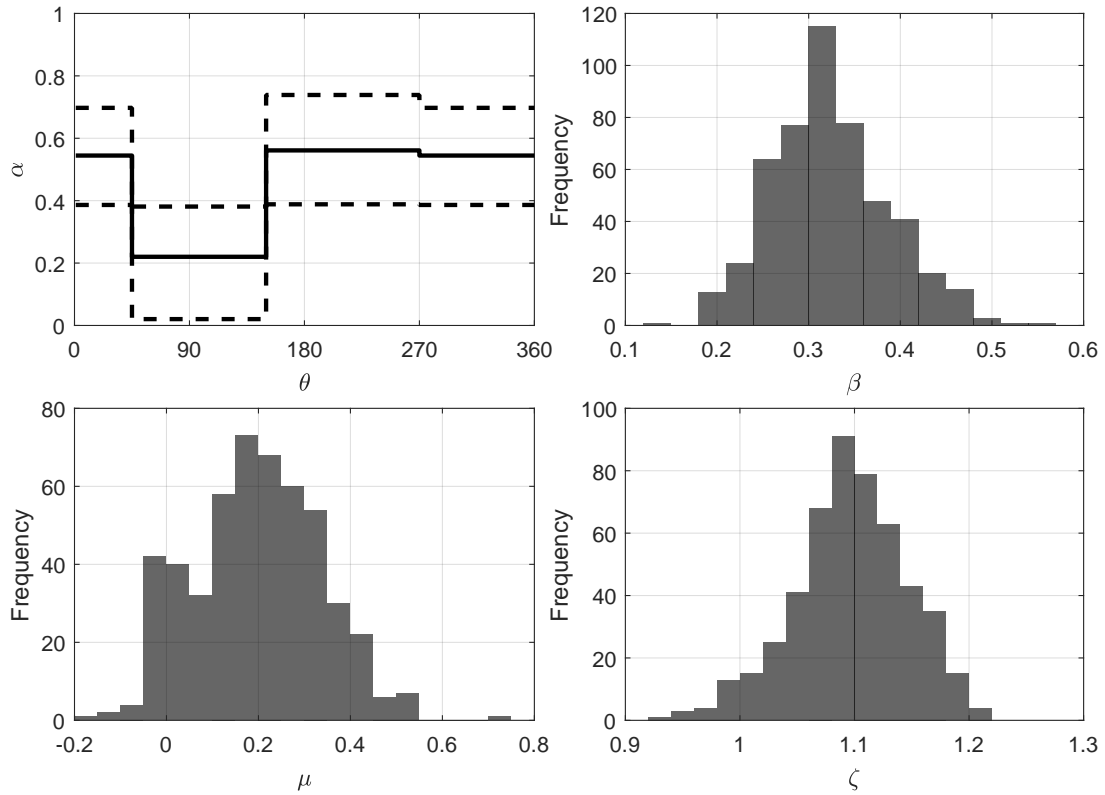


Figure 9: Piecewise stationary conditional extremes modelling for maximum surge (**SrgMxm**) on storm peak  $H_S$  at the CNS location. Top left panel shows estimated slope parameter  $\alpha = \{\alpha_k\}$  with direction  $\theta$  in terms of its bootstrap median and 95% uncertainty band. Subsequent panels show histograms of bootstrap estimates for scale exponent parameter  $\beta$ , and for residual mean  $\mu$  and scale  $\zeta$ , all assumed stationary with  $\theta$ .

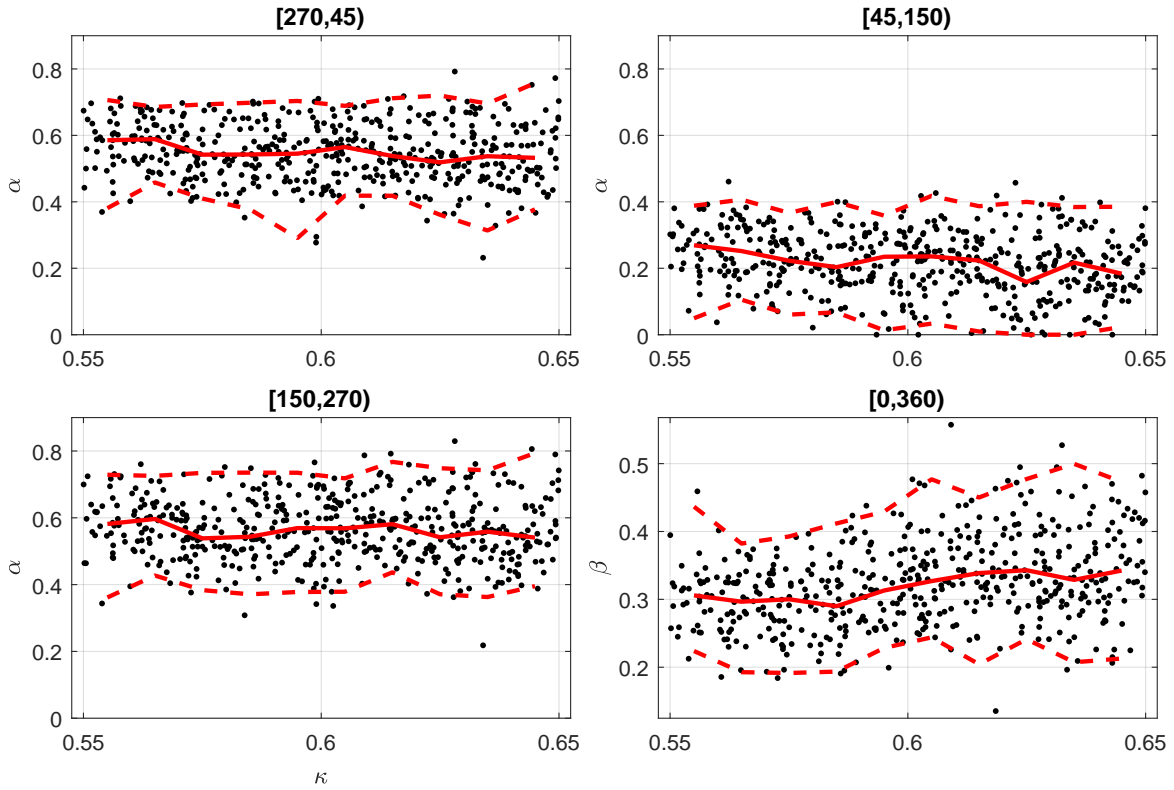


Figure 10: Piecewise stationary conditional extremes modelling for maximum surge ( $\text{SrgMxm}$ ) on storm peak  $H_S$  at the CNS location. Estimates of the conditional extremes slope parameter  $\alpha = \{\alpha_k\}$  as a function of conditional extremes threshold non-exceedance probability  $\kappa$  in each of the directional sectors [270, 45), [45, 150) and [150, 270), and estimate of conditional extremes scale exponent  $\beta$  assumed stationary with direction.

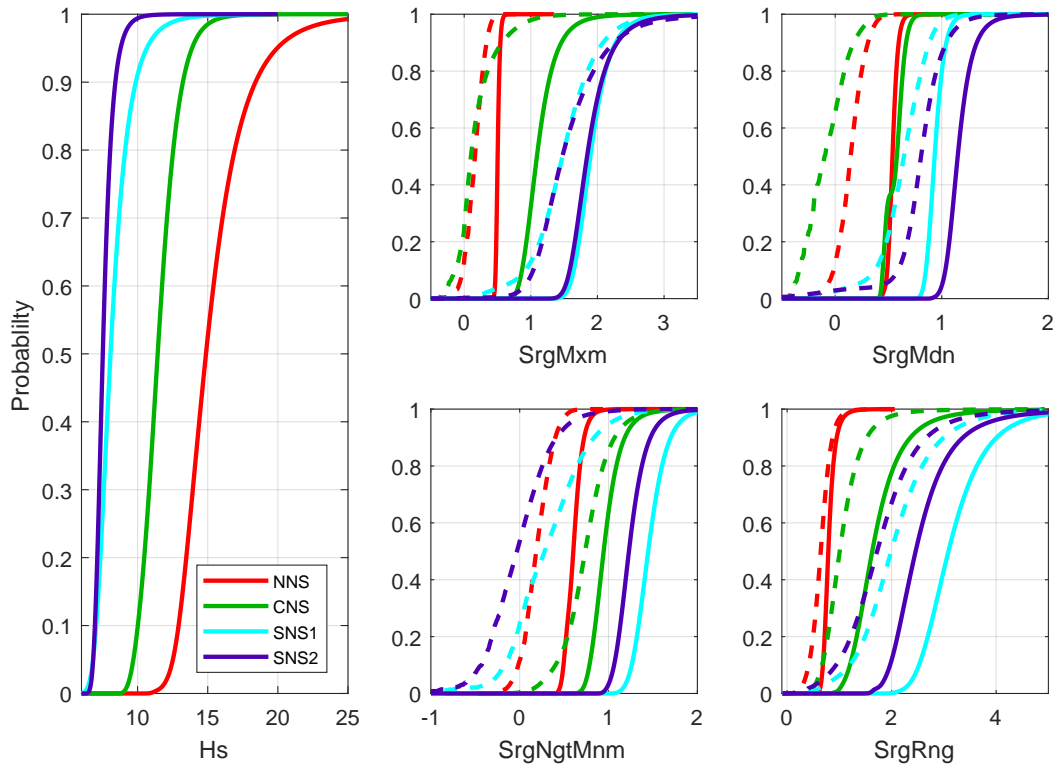


Figure 11: Omni-directional marginal and conditional cumulative distribution functions for the  $T = 100$ -year return period estimated by numerical integration for all locations. Left: marginal storm peak  $H_S$ . Centre and right: marginal ( $Y$ , solid) and conditional ( $Y$  given 100-year  $H_S$ , dashed) return value distributions for maximum ( $SrgMxm$ ), median ( $SrgMdn$ ) and negative minimum ( $SrgNgtMnm$ ) surge and surge range ( $SrgRng$ ). Colours indicate locations as explained in the left hand panel.

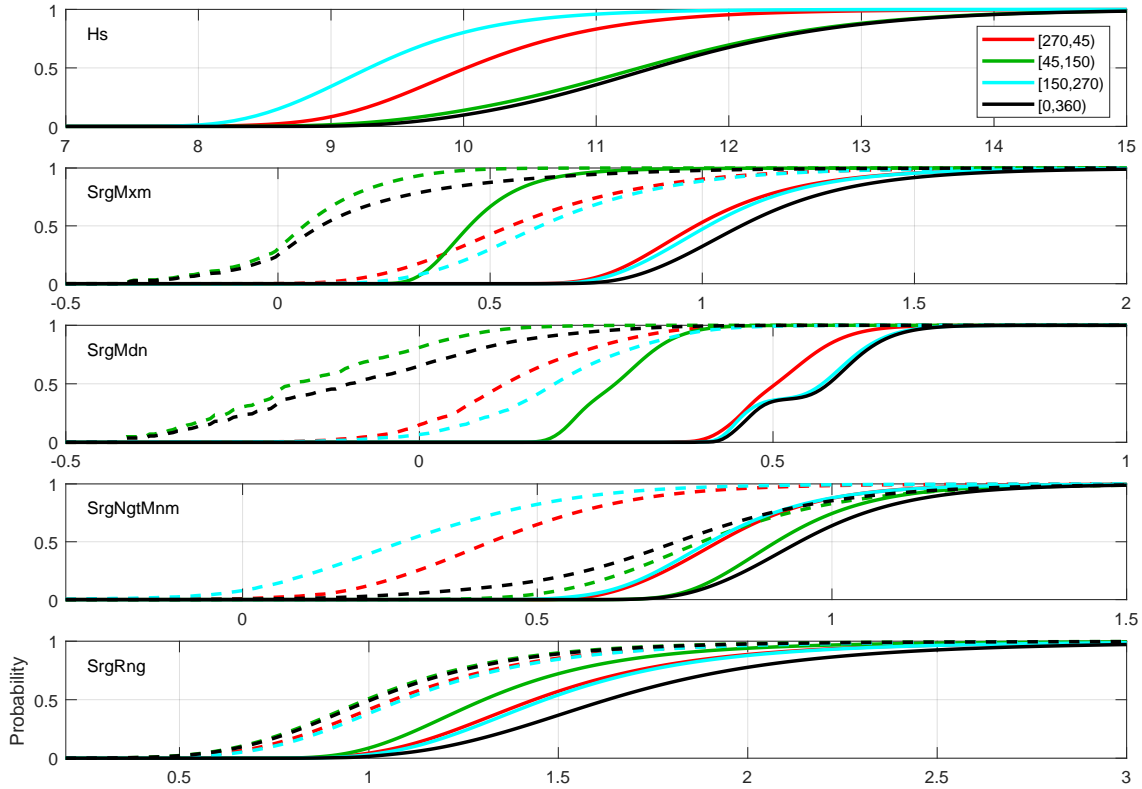


Figure 12: Directional marginal and conditional cumulative distribution functions for the 100-year return period estimated by numerical integration at the CNS location. First row: marginal storm peak  $H_S$ . Subsequent rows: marginal ( $Y$ , solid) and conditional ( $Y$  given 100-year  $H_S$ , dashed) return value distributions for maximum (**SrgMxm**), median (**SrgMdn**) and negative minimum (**SrgNgtMnm**) surge and surge range (**SrgRng**). Colours indicate covariate intervals.