

Accepted July 2020

Basin-wide spatial conditional extremes for severe ocean storms

Rob Shooter · Jonathan Tawn · Emma Ross · Philip Jonathan

Received: date / Accepted: date

Abstract Physical considerations and previous studies suggest that extremal dependence between ocean storm severity at two locations exhibits near asymptotic dependence at short inter-location distances, leading to asymptotic independence and perfect independence with increasing distance. We present a spatial conditional extremes (SCE) model for storm severity, characterising extremal spatial dependence of severe storms by distance and direction. The model is an extension of Shooter et al. (2019) and Wadsworth and Tawn (2019), incorporating piecewise linear representations for SCE model parameters with distance and direction; model variants including parametric representations of some SCE model parameters are also considered. The SCE residual process is assumed to follow the delta-Laplace form marginally, with distance-dependent parameter. Residual dependence of remote locations given conditioning location is characterised by a conditional Gaussian covariance dependent on the distances between remote locations, and distances of remote locations to the conditioning location. We apply the model using Bayesian inference to estimate extremal spatial dependence of storm peak significant wave height on a neighbourhood of 150 locations covering over 200,000 km² in the North Sea.

Keywords Spatial conditional extremes · Extremal dependence · Covariate effects · Ocean storms

R. Shooter would like to acknowledge financial support from Engineering and Physical Sciences Research Council grant EP/L015692/1 (STOR-i Centre for Doctoral Training) and Shell Research Ltd.

R. Shooter

Met Office Hadley Centre, FitzRoy Road, Exeter, EX1 3PB, United Kingdom.

STOR-i Centre for Doctoral Training, Department of Mathematics and Statistics, Lancaster University, LA1 4YW, United Kingdom.

E-mail: robert.shooter@metoffice.gov.uk

J. Tawn

Department of Mathematics and Statistics, Lancaster University, LA1 4YW, United Kingdom.

E. Ross

Shell Global Solutions International BV, Amsterdam, The Netherlands.

P. Jonathan

Shell Research Ltd., London SE1 7NA, United Kingdom.

Department of Mathematics and Statistics, Lancaster University, LA1 4YW, United Kingdom.

1 Introduction

A key issue in modelling spatial extremes is assessing the nature of dependence between extreme events. If we observe an extreme event at a location, we are interested in the information provided by this event about the probability of observing extreme events simultaneously at other locations. We expect that over short distances, an extreme event being observed at one location may be related to an extreme observation at another. However extremes observed at distant locations are likely to be less related to each other. To describe extremal dependence, Coles et al. (1999) introduce the measures χ and $\bar{\chi}$, most easily calculated through their sub-asymptotic forms $\chi(u)$ and $\bar{\chi}(u)$, where $u \in [0, 1]$. For bivariate uniform random variables (U, V) these are defined as $\chi(u) = 2 - \log \mathbb{P}(U < u, V < v) / \log \mathbb{P}(U < u)$ and $\bar{\chi}(u) = 2 \log \mathbb{P}(U > u) / \log \mathbb{P}(U > u, V > v) - 1$. χ and $\bar{\chi}$ may be obtained by taking the respective limits of these functions, as $u \rightarrow 1$. The nature of extremal dependence between U and V may then be described by considering χ and $\bar{\chi}$ together. If $\chi = 0$ and $-1 \leq \bar{\chi} < 1$, the random variables are asymptotically independent, and the value of $\bar{\chi}$ signifies the level of dependence. On the other hand, if $\bar{\chi} = 1$ and $0 < \chi \leq 1$, then the pair (U, V) exhibit asymptotic dependence, with χ providing a measure of this. For a full description of extremal dependence types, we refer the reader to Ledford and Tawn (1996) and Coles et al. (1999). Spatial extensions of these measures are discussed in Tawn et al. (2018).

Models for spatial extremes typically require that the type of extremal dependence exhibited be decided before model fitting. Here, we describe a spatial conditional extremes (SCE) model able to capture different types of asymptotic behaviour without prior assumption of dependence class. The SCE framework is useful to assess the risk involved in the construction of coastal and marine structures, enabling improved description of extremal dependence between variables characterising the meteorological-oceanographic (“metocean”) environment, and better joint quantification of risk for severe ocean storms.

The current study involves characterisation of extremal spatial dependence of ocean storm severity, quantified for a storm event using storm peak significant wave height (H_S). For two sampling locations a short distance apart (relative to the size of a storm), we expect that an extreme value of H_S arises at each location from the same storm event, characteristic of asymptotic dependence. In contrast, since it is unlikely that extremes occurring at two distant locations would be simultaneously large, we expect observations from distant locations to exhibit asymptotic independence. Previous studies (Kereszturi et al., 2016 and Ross et al., 2017a) have shown that the nature and extent of extremal dependence in an ocean basin changes with distance between locations, and also potentially with their orientation.

The traditional approach to spatial extremes is to consider max-stable processes, as introduced by Brown and Resnick (1977), Smith (1990) and Schlather (2002). Max-stable process models typically make the assumption that the spatial process is asymptotically dependent, and hence such models may be inappropriate for modelling H_S over an ocean basin. Other asymptotically dependent spatial extremes models have been proposed, such as the hierarchical max-stable model of Reich and Shaby (2012) and the Pareto process approach of Ferreira and de Haan (2014) or de Fondeville and Davison (2020), a method motivated by peaks-over-threshold analysis. The link between Pareto processes and the SCE model is described by Tawn et al. (2018); in particular, these processes exhibit asymptotic dependence. More recently, a number of models able to describe either class of extremal dependence have been proposed, including Wadsworth and Tawn (2012), Wadsworth et al. (2017) and Huser and Wadsworth (2019). These models suffer from the restriction that either the fitted model must assume a certain type of extremal dependence across the entire spatial domain in which it is fitted (Huser and Wadsworth 2019), or the model is rather computationally challenging to fit (Wadsworth and Tawn 2012). More detailed overviews of spatial extremes modelling may be found in Davison et al. (2012), Ribatet (2013) and Tawn et al. (2018), for example.

The SCE model presented here overcomes these issues. Careful parameterisation of distance and directional effects enables the spatial extremes problem to be described parsimoniously, even when the number of measurement locations is large. This reduces computational burden when fitting across hundreds of sampling locations compared to broadly equivalent max-stable processes, although Pareto processes are less computationally expensive than max-stable processes. Moreover, the model admits both asymptotic dependence and asymptotic independence, evolving with distance and direction on the spatial domain.

The model builds upon the work of Wadsworth and Tawn (2019) and Shooter et al. (2019), adopting functional forms for SCE parameters. Shooter et al. (2019) found that assuming a simple two-parameter for the decay of SCE slope parameter, α with distance was adequate to capture parameter behaviour adequately whilst reducing computational time for parameter estimation, since α would otherwise need to be estimated separately for pair of conditioning and remote locations. We build upon this idea by imposing either parametric or piecewise-linear forms for SCE parameters with distance and direction, motivated by theoretical considerations and evidence from preliminary analysis. The model is applied using Bayesian inference to a neighbourhood of 150 locations, to describe the joint extremal dependence of 149 remote locations on a central conditioning location. Such an analysis would be computationally demanding in the max-stable process framework. The current SCE model also adopts a delta-Laplace (or generalised Gaussian) form for the marginal distribution of model residuals, coupled to a spatial conditional Gaussian copula for residual dependence.

The article is presented as follows. In Section 2, we discuss the conditional extremes model of Heffernan and Tawn (2004) and the spatial extension of Wadsworth and Tawn (2019). Section 3 summarises the inferential scheme used for parameter estimation. Section 4 outlines the performance of our model in application North Sea storm severity. Discussion and conclusions are given in Section 5. An outline of the Markov chain Monte Carlo (MCMC) scheme used for inference, and the conditional quantile constraints of Keef et al. (2013), is given in the Appendix.

2 Model

2.1 Multivariate conditional extremes

Suppose we have a vector of random variables (X_0, \mathbf{X}) , where X_0 and $\mathbf{X} = (X_1, \dots, X_p)$ have Gumbel marginal distributions, and random variables \mathbf{Z} defined by

$$\mathbf{Z} = \frac{\mathbf{X} - \mathbf{a}(X_0)}{\mathbf{b}(X_0)}$$

for functions $\mathbf{a} : \mathbb{R} \rightarrow \mathbb{R}^p$ and $\mathbf{b} : \mathbb{R} \rightarrow \mathbb{R}^p$, where all operations are taken to be component-wise. Heffernan and Tawn (2004) assume the existence of \mathbf{a} and \mathbf{b} such that, for $x > 0$,

$$\lim_{u \rightarrow \infty} \mathbb{P}(\mathbf{Z} \leq \mathbf{z}, X_0 - u > x | X_0 > u) = G(\mathbf{z}) \exp(-x) \quad (1)$$

where G is a joint distribution with non-degenerate margins. This form for the conditional extremes model is asymptotically justified; see Heffernan and Tawn (2004) and Heffernan and Resnick (2007) for details. Keef et al. (2013) show that if the margins of \mathbf{X} are assumed to be Laplace-distributed (achievable using the probability integral transform), then canonical functional forms for $\mathbf{a}(\cdot)$ and $\mathbf{b}(\cdot)$ are $\mathbf{a}(x) = \boldsymbol{\alpha}x$ and $\mathbf{b}(x) = x^\beta$ (for $x > 0$), where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_p)$ and $\beta = (\beta_1, \dots, \beta_p)$. In this representation, each $\alpha_i \in [-1, 1]$ and $\beta_i \in (-\infty, 1]$, with different parameter values corresponding to different classes of extremal dependence. In the present spatial context, we assume positive dependence and restrict $\alpha_i \in [0, 1], \beta_i \in [0, 1]$ for all $i \in \{1, \dots, p\}$. Motivated by Equation (1), for some high threshold u and all $x_0 > u$, we assume the model form

$$\mathbf{X} | \{X_0 = x_0\} = \boldsymbol{\alpha}x_0 + x_0^\beta \mathbf{Z} \quad (2)$$

where $\mathbf{Z} \sim G$ is independent of X_0 .

2.2 Spatial conditional extremes

We extend the model described in Equation (2) to a spatial context, as described by Tawn et al. (2018) and Wadsworth and Tawn (2019). Consider a stationary spatial process $X(\cdot)$, on a domain \mathcal{S} , which has Laplace marginal distributions and locations $r, r_0 \in \mathcal{S}$. Then for distance $d \in \mathbb{R}_{\geq 0}$ and heading $\theta \in [0, 2\pi)$ between sites, assuming positive dependence between variables, we have in general that, for all $x_0 > u$,

$$X(r) | \{X(r_0) = x_0\} = \alpha(d, \theta)x_0 + x_0^{\beta(d, \theta)} Z(r - r_0)$$

where $\alpha : \mathbb{R}_{>0} \times [0, 2\pi) \rightarrow [0, 1]$, $\beta : \mathbb{R}_{>0} \times [0, 2\pi) \rightarrow [0, 1]$ and $Z(\cdot)$ is a residual process independent of $X(\cdot)$. For inference on a sample of spatial data (X_0, \mathbf{X}) observed at a finite set of locations r_0, r_1, \dots, r_p , $X(\cdot)$ is treated as a finite-dimensional process. We set $d_j = \text{dist}(r_j, r_0)$ for $j = 1, 2, \dots, p$ and $\theta_j = \text{head}(r_j, r_0)$, for distance and heading metrics $\text{dist}(\cdot, \cdot)$ and $\text{head}(\cdot, \cdot)$. Then θ describes the relative heading of a remote location from the conditioning site; note that θ is defined using the mathematical convention, anti-clockwise from due East, for which $\theta = 0$. We then set $\alpha_j = \alpha(d_j, \theta_j)$ and $\beta_j = \beta(d_j, \theta_j)$.

We follow Wadsworth and Tawn (2019) and suppose that process $Z(\cdot)$ has delta-Laplace margins with parameters μ, σ and δ dependent on d and θ . That is, the marginal density f_{Z_j} of Z at distance d_j is

$$f_{Z_j}(z_j) = \frac{\delta_j}{2\kappa_j\sigma_j\Gamma\left(\frac{1}{\delta_j}\right)} \exp\left\{-\left|\frac{z - \mu_j}{\kappa_j\sigma_j}\right|^{\delta_j}\right\}$$

for $j = 1, 2, \dots, p$, $\delta_j, \sigma_j, \kappa_j \in \mathbb{R}_{>0}, \mu_j \in \mathbb{R}, z_j \in \mathbb{R}$, where $\kappa_j^2 = \Gamma(1/\delta_j)/\Gamma(3/\delta_j)$ and $\Gamma(\cdot)$ represents the gamma function. The mean and variance of this distribution are respectively μ_j and σ_j^2 , regardless of the choice of δ_j , and we denote the distribution as $\text{DL}(\mu_j, \sigma_j^2, \delta_j)$. With this notation, $\delta_j = 2$ corresponds to a Gaussian distribution, whereas $\delta_j = 1$ yields a Laplace distribution; the standard Laplace distribution with variance 2 corresponds to $\sigma_j^2 = 2$ in our notation. Of particular importance is the requirement that as $d \rightarrow \infty$, we approach perfect independence between $X(r)$ and $X(r_0)$. Inspection of the equations above suggests that we should simply be left with standard Laplace random variables; i.e., $\lim_{d \rightarrow \infty} \delta(d, \theta) = 1$, $\lim_{d \rightarrow \infty} \alpha(d, \theta) = \lim_{d \rightarrow \infty} \beta(d, \theta) = \lim_{d \rightarrow \infty} \mu(d, \theta) = 0$, $\lim_{d \rightarrow \infty} \sigma(d, \theta) = \sqrt{2}$. The model is not informative for δ and ρ and $d = 0$.

Now consider the vector of random variables (X_0, \mathbf{X}) corresponding to $p + 1$ spatial locations, with standard Laplace marginal distributions, i.e., $X_j \sim \text{DL}(0, 2, 1)$ for $j = 0, 1, \dots, p$. We assume, conditional on $X_0 = x_0$, for $x_0 > u$, that \mathbf{X} follows a multivariate extension of the delta-Laplace distribution

$$(X_1, \dots, X_p) | \{X_0 = x_0\} = \boldsymbol{\alpha}x_0 + x_0^\beta \mathbf{Z}$$

where $\mathbf{Z} \sim \text{DL}_p(\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\delta}; \boldsymbol{\Sigma})$. $\boldsymbol{\Sigma}$ is the correlation matrix for a Gaussian dependence structure between residual components, discussed further in Section 2.4, such that

$$F_{\mathbf{Z}}(\mathbf{z}) = \Phi_p(\Phi^{-1}(F_{Z_1}(z_1)), \Phi^{-1}(F_{Z_2}(z_2)), \dots, \Phi^{-1}(F_{Z_p}(z_p)); \mathbf{0}, \boldsymbol{\Sigma}) \quad (3)$$

where F represents a cumulative distribution function and Φ is the cumulative distribution function of a standard Gaussian distribution. $\Phi_p(\mathbf{0}, \boldsymbol{\Sigma})$ is the cumulative distribution function of a p -dimensional Gaussian distribution with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$. Hence marginally, $Z_j \sim \text{DL}(\mu_j, \sigma_j^2, \delta_j)$, so that $Z_j = Z_j^* \sigma_j + \mu_j$, where $Z_j^* \sim \text{DL}(0, 1, \delta_j)$. Writing X_j^c to represent $X_j | \{X_0 = x_0\}$, we have $X_j^c = \alpha_j x_0 + x_0^{\beta_j} Z_j \sim \text{DL}(m_j, s_j^2, \delta_j)$, for conditional means m_j and standard deviations s_j given by

$$m_j = \alpha_j x_0 + x_0^{\beta_j} \mu_j \quad \text{and} \quad s_j = x_0^{\beta_j} \sigma_j \quad (4)$$

for $j = 1, 2, \dots, p$. Hence

$$\mathbf{X}^c = (\mathbf{X} | \{X_0 = x_0\}) \sim \text{DL}_p(\mathbf{m}, \mathbf{s}^2, \boldsymbol{\delta}; \boldsymbol{\Sigma}), \quad (5)$$

a p -dimensional delta-Laplace distribution with mean $\mathbf{m} = (m_1, m_2, \dots, m_p)$, variance $\mathbf{s}^2 = (s_1^2, s_2^2, \dots, s_p^2)$, $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_p)$ and (standard Gaussian-scale) covariance $\boldsymbol{\Sigma}$.

The model in Equation (5) describes different types of extremal dependence, inferred from the values of parameters (α_j, β_j) for $j \in \{1, \dots, p\}$. If $(\alpha_j, \beta_j) = (1, 0)$, then random variables X_0 and X_j are asymptotically dependent, whereas if $\alpha_j < 1$, the random variables exhibit asymptotic independence. Further discussion of this can be found in Tawn et al. (2018).

2.3 Model parameter variation with distance and direction

The $p + 1$ measurement locations are assumed to have coordinates \mathbf{r}_j , $j = 0, 1, \dots, p$. Parameters $\alpha, \beta, \mu, \sigma, \delta$ are assumed to be continuous functions of distance with $d_j = \text{dist}(\mathbf{r}_j, \mathbf{r}_0)$ and heading $\theta_j = \text{head}(\mathbf{r}_j, \mathbf{r}_0)$ for remote location \mathbf{r}_j , $j = 1, 2, \dots, p$ given conditioning location \mathbf{r}_0 , defined using a local Cartesian frame discussed further in Section 2.5. Two specific forms for the distance and directional dependence of SCE model parameters are considered. The first representation is piecewise linear in d and θ . The second representation, adopted to reduce model size and improve computational performance, assumes parametric variation of α and β with d .

For the piecewise linear representation, we specify a set of n_d equally-spaced nodes d_k , $k = 1, 2, \dots, n_d$, covering the domain radially from \mathbf{r}_0 , and a corresponding set of n_θ directions $\theta_\ell \in [0, 2\pi)$, $\ell = 1, 2, \dots, n_\theta$ to codify angular variation. The value of any SCE model parameter $\eta(d, \theta)$ on the spatial domain can then be expressed in terms of the set of values $\{\eta_{k,\ell}\}$ as

$$\begin{aligned} \eta(d, \theta) &= \frac{\Delta_L^d \eta_{k^*,1} + \Delta_U^d \eta_{k^*+1,1}}{\Delta_F^d} \quad \text{for } n_\theta = 1 \\ &= \frac{\Delta_L^d \Delta_L^\theta \eta_{k^*,\ell^*} + \Delta_U^d \Delta_L^\theta \eta_{k^*+1,\ell^*} + \Delta_L^d \Delta_U^\theta \eta_{k^*,\ell^*+1} + \Delta_U^d \Delta_U^\theta \eta_{k^*+1,\ell^*+1}}{\Delta_F^d \Delta_F^\theta} \quad \text{for } n_\theta = 2, 3, \dots \end{aligned}$$

with $\eta \in \{\alpha, \beta, \mu, \sigma, \delta\}$, where $k^* = \underset{k}{\text{argmax}} \{d_k : d_k < d\}$ and $\ell^* = \underset{\ell}{\text{argmax}} \{\theta_\ell : \theta_\ell < \theta\}$. Further $\Delta_L^d = d - d_{k^*}$, $\Delta_U^d = d_{k^*+1} - d$ and $\Delta_F^d = d_{k^*+1} - d_{k^*}$, with analogous definitions for Δ^θ terms. The values of $\eta_{k,\ell}$, $k = 1, 2, \dots, n_d$, $\ell = 1, 2, \dots, n_\theta$ are estimated during inference. We also consider parametric representations of α and β with d for each θ_ℓ , $\ell = 1, 2, \dots, n_\theta$, taking the forms

$$\begin{aligned} \alpha(d, \theta_\ell) &= \exp \left\{ - \left(\frac{d}{A_{1\ell}} \right)^{A_{2\ell}} \right\} \quad \text{for } A_{1\ell}, A_{2\ell} \in \mathbb{R}_{>0}, \text{ and} \\ \beta(d, \theta_\ell) &= \frac{B_{1\ell} d^{B_{2\ell}} \exp \left(- \frac{d}{B_{3\ell}} \right)}{\max_{d>0} \left\{ d^{B_{2\ell}} \exp \left(- \frac{d}{B_{3\ell}} \right) \right\}} \quad \text{for } B_{1\ell} \in (0, 1), B_{2\ell}, B_{3\ell} \in \mathbb{R}_{>0}. \end{aligned}$$

These choices were motivated by previous application studies including Shooter et al. (2019), together with the known theoretical behaviours of α and β for $d = 0$ and $d \rightarrow \infty$. This parametric form does not admit asymptotic dependence, however, since $\alpha(d, \theta_t) \neq 1$ for $d > 0$, although for suitably large $A_{1\ell}$ and $A_{2\ell} \geq 2$ then $\alpha(d, \theta_t) \approx 1$ for $d \approx 0$; a possible parametric form for $\alpha(d, \theta_t)$ which permits asymptotic dependence is given by Wadsworth and Tawn (2019). Angular effects continue to be parameterised using piecewise linear forms analogous to that in the equation above. Using these parametric forms, the full behaviour of α and β with distance and direction can be inferred by estimating just $5n_\theta$ parameters rather than $2n_d n_\theta$ for the piecewise linear representation. With $n_d = 7$ and $n_\theta = 6$, this corresponds to a reduction of $(2n_d - 5)n_\theta = 54$ in the number of model parameters to be estimated, improving the computational efficiency of inference.

2.4 Gaussian residual correlation matrix with distance

The j, j' element $\Sigma_{jj'}$ of residual correlation matrix Σ , $j, j' = 1, 2, \dots, p$ quantifies the dependence between SCE residuals (on standard Gaussian-scale) at locations \mathbf{r}_j and $\mathbf{r}_{j'}$ given conditioning on location \mathbf{r}_0 . It is reasonable to expect that the value of $\Sigma_{jj'}$ depends on both the distance $\text{dist}(\mathbf{r}_j, \mathbf{r}_{j'})$ between remote locations, and distances $\text{dist}(\mathbf{r}_j, \mathbf{r}_0)$ and $\text{dist}(\mathbf{r}_{j'}, \mathbf{r}_0)$ from remote locations to conditioning site. When $\text{dist}(\mathbf{r}_j, \mathbf{r}_0)$ and $\text{dist}(\mathbf{r}_{j'}, \mathbf{r}_0)$ are large relative to $\text{dist}(\mathbf{r}_j, \mathbf{r}_{j'})$, the conditional correlation between remote locations will be similar to the unconditional correlation. However, when $\text{dist}(\mathbf{r}_j, \mathbf{r}_0)$ and $\text{dist}(\mathbf{r}_{j'}, \mathbf{r}_0)$ are of comparable size to $\text{dist}(\mathbf{r}_j, \mathbf{r}_{j'})$, we expect conditioning to influence correlation considerably.

We adopt a parameterisation for Σ equivalent to the correlation function for a standard Gaussian field evaluated at $p + 1$ locations conditioned on its value at one location, with powered exponential dependence. With Σ^* representing the correlation matrix of the unconditioned field (and matrix indexing starting from zero for convenience), the correlation matrix Σ for the conditional field has elements given by

$$\Sigma_{jj'} = \frac{\Sigma_{jj'}^* - \Sigma_{j0}^* \Sigma_{j'0}^*}{(1 - \Sigma_{j0}^{*2})^{1/2} (1 - \Sigma_{j'0}^{*2})^{1/2}} \quad (6)$$

for $j, j' = 1, 2, \dots, p$ with conditioning location indexed by zero (see e.g. Mardia et al. 1979, page 63). Further, it is reasonable to assume that the correlation between observations at different locations in the unconstrained field reduces as a function of the distance between the locations. For this reason, we adopt an isotropic powered exponential form for the correlation of the unconditional field, given by

$$\Sigma_{jj'}^* = \rho(\text{dist}(\mathbf{r}_j, \mathbf{r}_{j'})) = \exp\left(-\left(\frac{\text{dist}(\mathbf{r}_j, \mathbf{r}_{j'})}{R_1}\right)^{R_2}\right) \text{ for } R_1, R_2 \in \mathbb{R}_{>0}$$

and $j, j' = 0, 1, 2, \dots, p$. Parameters R_1 and R_2 are to be estimated.

2.5 Calculating distances and directions

Locations of points on the surface of the Earth are typically specified in terms of longitude-latitude coordinates. Temporarily adopting oceanographic notation, the shortest distance (e.g. in metres) on the surface of a spherical Earth between locations with longitude-latitude coordinates (λ, φ) and (λ', φ') can be calculated using the spherical law of cosines. For calculations of local distances and directions between locations, it is convenient to adopt a local Cartesian description of displacement on the surface of a sphere, (see e.g. Vallis 2017, page 69). For locations with longitude-latitude coordinates $\mathbf{r} = (\lambda, \varphi)$ and $\mathbf{r}' = (\lambda', \varphi')$, we locate the local origin of coordinates at $(\bar{\lambda}, \bar{\varphi}) = ((\lambda + \lambda')/2, (\varphi + \varphi')/2)$, with x axis running West-East (in the Northern Hemisphere) and y axis poleward. Then, to a good approximation when $|\lambda' - \lambda|$ and $|\varphi' - \varphi|$ are small, the local Cartesian displacement \mathbf{d} between the points can be written

$$\mathbf{d} = (a \cos(\bar{\varphi})(\lambda' - \lambda), a(\varphi' - \varphi))$$

where a is the radius of the spherical Earth. We adopt this model to estimate distances and directions for all pairs of locations considered, with $\text{dist}(\mathbf{r}, \mathbf{r}') = \|\mathbf{d}\|$ and $\text{head}(\mathbf{r}, \mathbf{r}') = \arg(\mathbf{d})$.

3 Inference

3.1 Sample likelihood

Differentiating Equation (3), the joint density $f_{\mathbf{X}^c}(\mathbf{x}) = \mathbb{P}(\mathbf{X} = \mathbf{x} | X_0 = x_0)$, with marginal distributions $f_{X_j^c}$ can be written

$$f_{\mathbf{X}^c}(\mathbf{x}) = \phi_p(w_1, w_2, \dots, w_p; \mathbf{0}, \Sigma) \prod_{j=1}^p \frac{f_{X_j^c}(x_j)}{\phi(w_j)} \quad (7)$$

where ϕ denotes the standard Gaussian density, ϕ_p represents the p -dimensional Gaussian density with given mean vector and correlation matrix, and $w_j = \Phi^{-1}\{F_{X_j^c}(x_j)\}$ with $x_j \in \mathbb{R}$ for $j = 1, 2, \dots, p$. From (7) above, the negative log-density is

$$-\log\{f_{\mathbf{X}^c}(\mathbf{x})\} = -\log \phi_p(\mathbf{w}; \mathbf{0}, \Sigma) - \sum_{j=1}^p \log f_{X_j^c}(x_j) + \sum_{j=1}^p \log \phi(w_j)$$

where $\mathbf{w}' = (w_1, w_2, \dots, w_p)$. Hence, the sample negative log-likelihood given a sample $\{x_{ij}\}_{i=1, j=0}^{n,p}$ is

$$\begin{aligned} & \frac{np}{2} \log(2\pi) + \frac{n}{2} \log |\boldsymbol{\Sigma}| + \sum_{j=1}^p \log \left(2s_j \kappa_j \Gamma \left[\frac{1}{\delta_j} \right] \right) - n \sum_{j=1}^p \log \delta_j \\ & + \sum_{i=1}^n \left\{ \mathbf{w}'_i \boldsymbol{\Sigma}^{-1} \mathbf{w}_i + \sum_{j=1}^p \left| \frac{x_{ij} - m_j}{\kappa_j s_j} \right|^{\delta_j} + \sum_{j=1}^p \log \phi(w_{ij}) \right\} \end{aligned} \quad (8)$$

where $\kappa_j^2 = \Gamma(1/\delta_j)/\Gamma(3/\delta_j)$ for each j and $\mathbf{w}'_i = (w_{i1}, w_{i2}, \dots, w_{ip})$.

3.2 Parameter estimation

We use Bayesian inference to estimate the joint posterior distribution of SCE model parameters, with sample negative log-likelihood as defined above. In practice we find that allowing directional variation of SCE residual parameters μ , σ and δ cannot be justified given data examined to date. Similarly, we assume that the residual process depends on distances between locations only. Any directional effects in the Laplace-scale data can be accommodated by parameterisation of SCE parameters α and β as described below. Prior marginal transformation of data per location to Laplace scale using directional extreme value models estimates marginal directional variation of the original sample on physical scale.

For SCE models with non-parametric representations for α and β terms (henceforth ‘ABN’ models), the parameter set Ω is $\{\{\{\alpha_{k,\ell}, \beta_{k,\ell}\}_{\ell=1}^{n_\theta}, \mu_k, \sigma_k, \delta_k\}_k^{n_d}, R_1, R_2\}$, consisting of $2n_d n_\theta + 3n_d + 2$ parameters. In this case, a grouped adaptive MCMC algorithm based on Roberts and Rosenthal (2009) is used for parameter inference, described in Shooter et al. (2019). Briefly, random search is used to find a reasonable starting solution. Then a Metropolis-within-Gibbs algorithm is used iteratively to sample each of the elements of Ω in turn for a total of n_{MiG} iterations. Subsequently we use the grouped adaptive Gaussian random walk Metropolis within-Gibbs algorithm iteratively for a further n_{GA} iterations. Within the grouped adaptive algorithm, we jointly update groups $\{\{\alpha_{k\ell}, \beta_{k\ell}, \mu_{k\ell}, \sigma_{k\ell}\}_{\ell=1}^{n_\theta}\}$ of parameters (for each $d = 1, 2, \dots, n_d$) following the adaptive Metropolis scheme of Roberts and Rosenthal (2009). Parameters $\{\delta_{k\ell}\}_{k,\ell=1}^{n_d, n_\theta}$, R_1 and R_2 are updated using Gaussian random walk Metropolis within-Gibbs, with automatic proposal step-size adjustment to achieve proposal acceptance rates near 0.25. Uniform prior distributions on plausible domains are used for model parameters. Chain convergence is judged to have occurred when trace plots for parameters and their dependence stabilise. Typically, prior MCMC analyses on multiple chains is performed for burn-in and to identify ‘hot’ starting solutions for subsequent analysis; it was ensured that the number of MCMC iterations after burn-in was at least 30,000. Mean effective sample size (across all parameters) was approximately 400, with a minimum effective sample size in excess of 200. A summary of the MCMC scheme and prior specification are given in the Appendix.

For SCE models with parametric representations for α and β terms (henceforth ‘ABP’ models), the parameter set Ω is $\{\{A_{1\ell}, A_{2\ell}, B_{1\ell}, B_{2\ell}, B_{3\ell}\}_{\ell=1}^{n_\theta}, \{\mu_k, \sigma_k, \delta_k\}_k^{n_d}, R_1, R_2\}$, consisting of $5n_\theta + 3n_d + 2$ parameters. In this case, we apply the adaptive Metropolis scheme of Roberts and Rosenthal (2009) directly on the full set of parameters. Uniform prior distributions on plausible domains are used for model parameters; details are given in the Appendix.

We further optionally restrict the space of feasible sets of SCE parameters using the conditional extremes model constraints of Keef et al. (2013). The constraints are motivated by the fact that conditional quantiles (of one variable given a large value of another) under any form of asymptotic independence ($\alpha < 1$) cannot be larger than under asymptotic positive dependence ($\alpha = 1$). Details of the constraints applied in a spatial setting are given in the supplementary material of Shooter et al. (2019), and summarised here in the Appendix.

4 Application

4.1 North Sea storm peak significant wave height

We apply the SCE model to data for ocean storm severity from a spatial lattice (see Figure 1) consisting of 150 locations in the North Sea. The sample corresponds to hindcast values of winter storm peak significant wave height from the NEXTRA hindcast (Oceanweather 2002). Storm intervals for a total of 1680 storms during the period 1 Oct 1964 to 31 Mar 1995 were isolated from up- and down-crossings of a sea state significant wave height threshold for a central conditioning location, using the procedure outlined in Ewans and Jonathan (2008). Storm peak significant wave height for each storm interval at each location provided a sample of 1680×150 observations for further analysis. For each storm-location combination, the direction (from which waves emanate, measured clockwise from North) at the time of the storm peak, referred to as the storm direction, was also retained. The spatial extremal characteristics of this sample have been examined previously in Ross et al. (2017a) and Shooter et al. (2019). At each location, marginal directional extreme value analysis of storm peak values was performed, and storm peak data transformed to standard Laplace marginal scale (see Ross et al. 2017b). The central location was adopted as the conditioning site, an SCE model is estimated to describe the joint characteristics of the remaining 149 locations given a large value at the conditioning site.

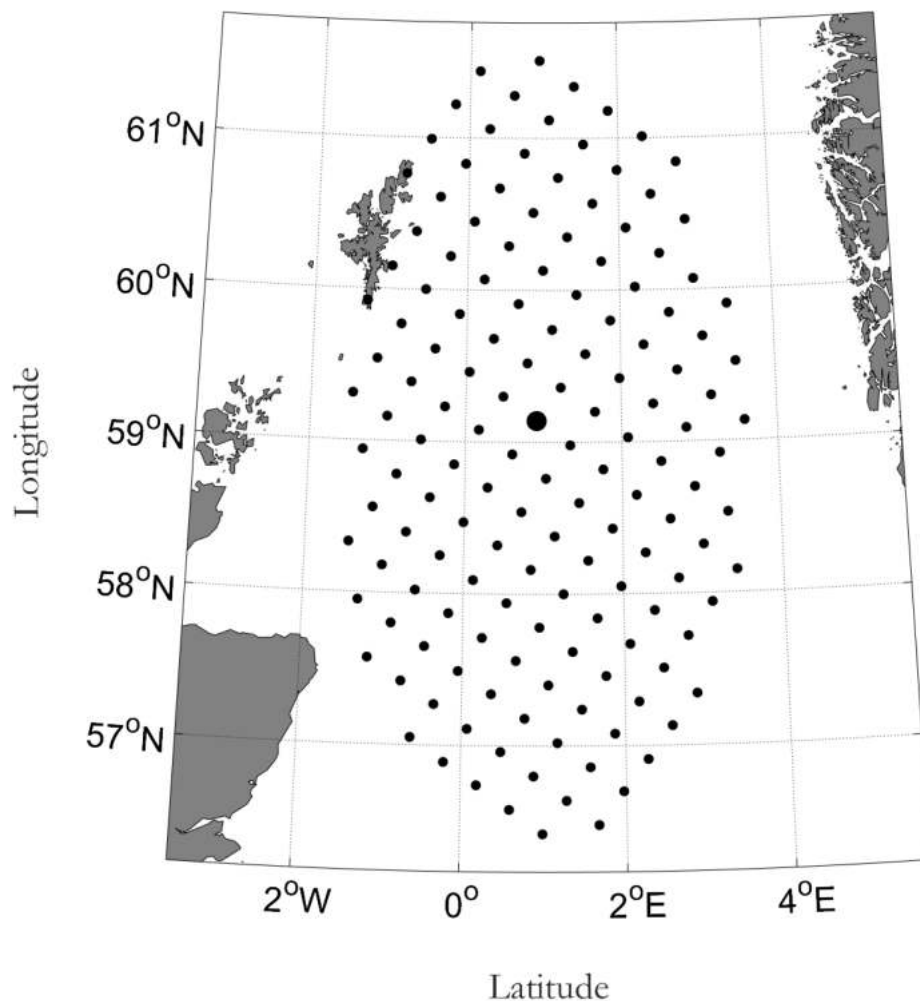


Fig. 1 Map of the 150 locations in the North Sea ocean. The larger disc represents the conditioning location. The neighbourhood extends approximately ± 150 km West-East, and ± 300 km North-South relative to the conditioning location. UK and Norway land masses also shown.

In the metocean community, extreme value analysis is typically conducted on “storm peak” events, taken to be temporally independent. Storm peak events are isolated from (temporally dependent) time-series of storm severity at a location using physical considerations. The fundamental motivation is that storm peak events are temporally sufficiently separated that the causative physical processes responsible for them are reasonably taken to be independent. Therefore the storm peak significant wave heights are assumed independent; storm peak events are typically separated by at least a day (see Ewans and Jonathan 2008). The characteristics of the ocean environment within storm (including the evolution of sea-state significant wave height, and the dependence between it and other oceanographic and structural loading variables of interest) are then inferred conditional on storm peak characteristics. In other applications, when a physical basis for declustering of time-series of extremes is not available, Ferro and Segers (2003) provides one possible approach to characterising that dependence. Covariates also potentially play a role in the selection of storm periods and storm peaks, and are critical to reasonable transformation of storm peak data to standard marginal scale. Scatter plots of Laplace-scale data for selected locations on West-East and South-North transects are given in Figure 2. For locations within 50 kilometres (km) of the conditioning site, scatter plots suggest that values of α near unity are expected, decaying with increasing distance.

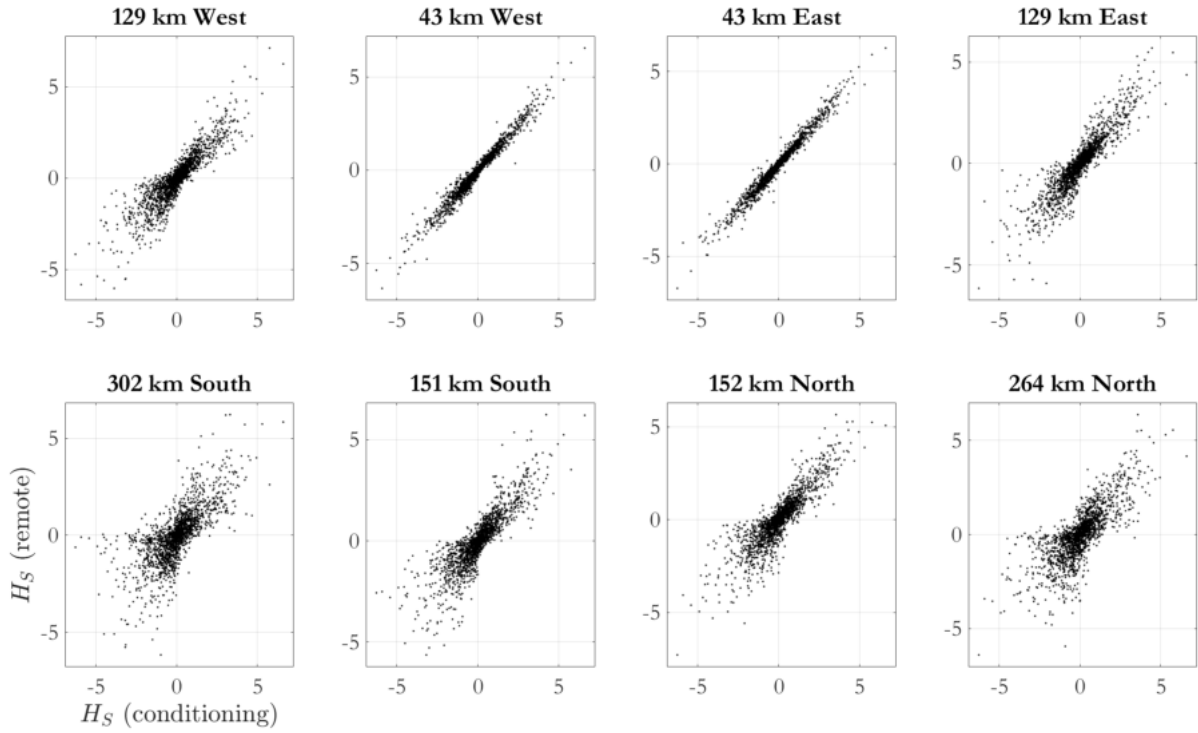


Fig. 2 Scatter plots for Laplace-scale storm peak significant wave height H_S ; H_S at the conditioning location is plotted on the x-axis in each case. Panels represent different locations on West-East transect (top) and South-North transect (bottom) at different distances from the conditioning site.

4.2 Results

For the sample of storm peak significant wave height data transformed to standard Laplace margins, the SCE model was estimated as follows.

Using a threshold u at the conditioning location with marginal non-exceedance probability of 0.95, joint posterior distributions of SCE model parameters were estimated based on both the ABN and ABP parameter representations. For the ABN model, a distance-direction lattice of $n_d \times n_\theta = 7 \times 6$ was adopted to estimate the distance-direction evolution of α and β parameters, with n_d nodes used to characterise the evolution of residual parameters μ , σ and δ . The Gaussian-scale residual correlation was estimated in terms of two parameters R_1 , R_2 . The ABP model is similarly parameterised, except that evolution of α and β with distance is described parametrically for each of n_θ directions. Results illustrated here were obtained with the conditional quantile constraints of Keef et al. (2013) active. A discussion of the corresponding results obtained without conditional quantile constraints is given in Section 4.3, together with an outline of results for different conditioning locations, and different choices of conditioning threshold u .

Marginal summaries of posterior distributions of SCE model parameters with distance and direction are shown in Figures 3-5 for both the ABN and ABP models. In each figure, solid lines indicate the variation of posterior medians with distance in km, with 95% credible intervals shown by dashed lines. Black lines indicate estimates from the ABN representation, and orange lines from the ABP model.

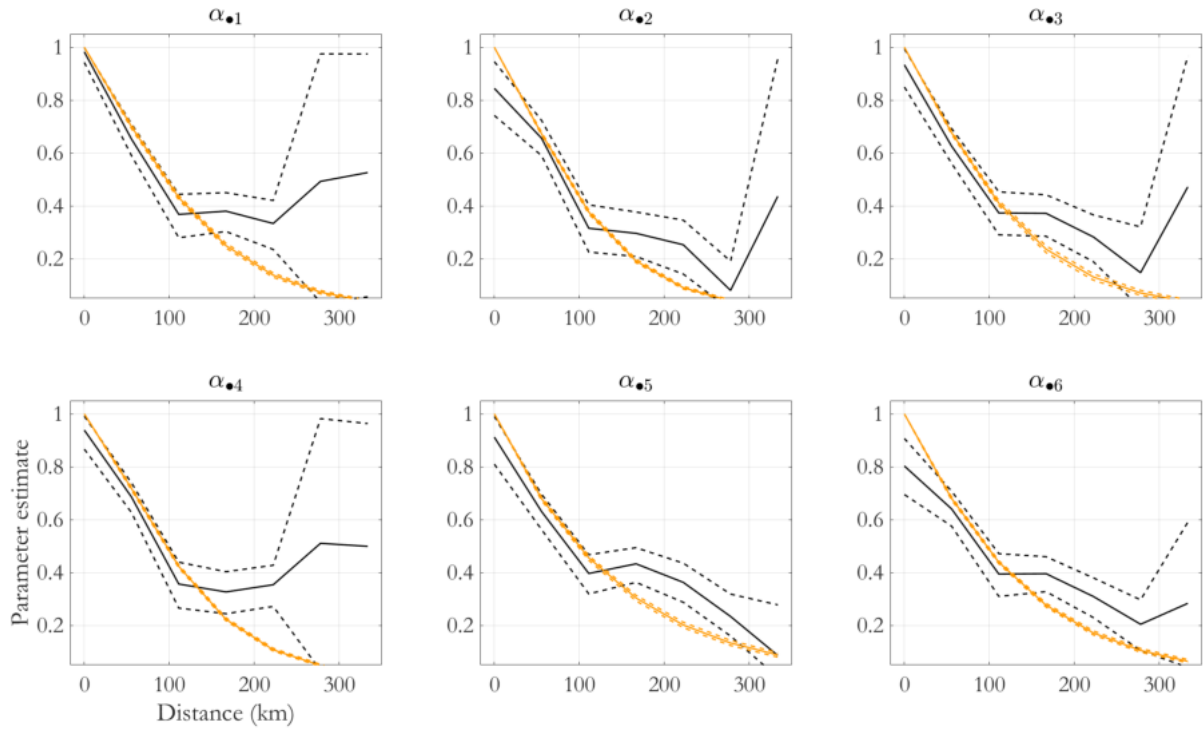


Fig. 3 Estimates for α with distance d (km) for each of $n_\theta = 6$ directions $0, \pi/3, 2\pi/3, \dots$ defined in the normal mathematical sense (anticlockwise, with 0 equivalent to East). Solid lines represent posterior medians, with dashed lines representing the upper and lower limits of empirical 95% posterior credible intervals. ABN model in black, ABP model in orange (grey in greyscale).

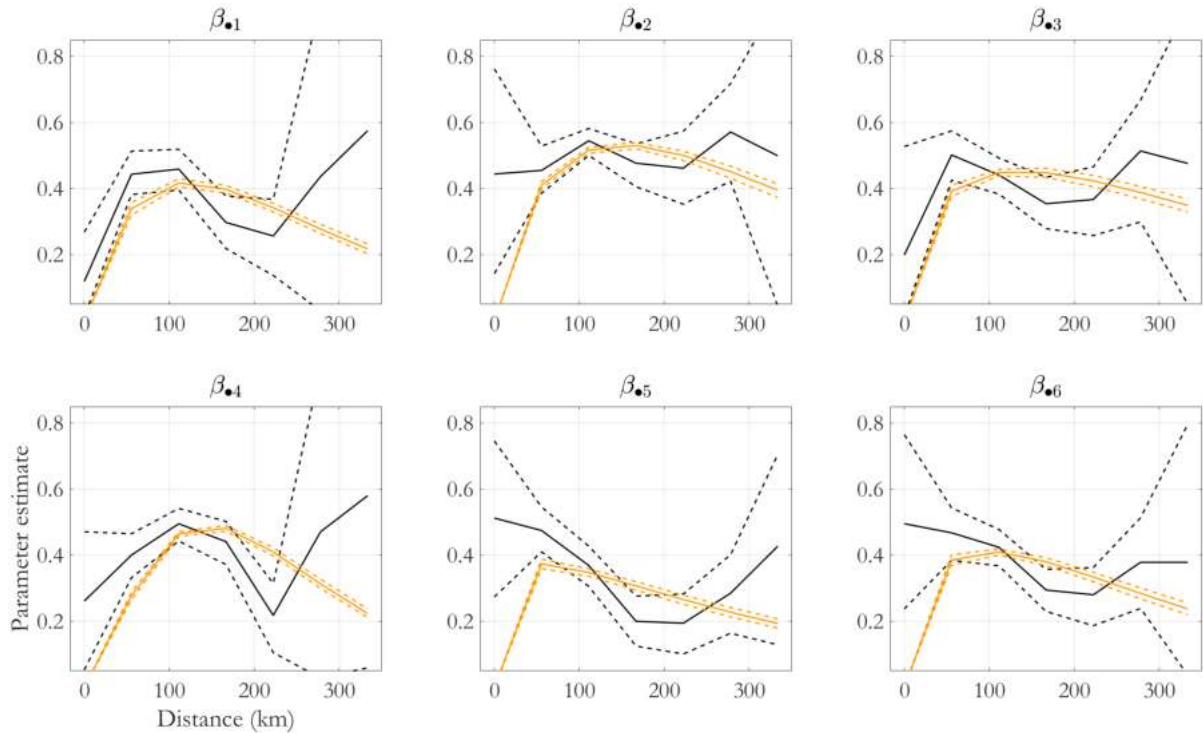


Fig. 4 Estimates for β with distance d (km) for each of $n_\theta = 6$ directions $0, \pi/3, 2\pi/3, \dots$. Solid lines represent posterior medians, with dashed lines representing the upper and lower limits of empirical 95% posterior credible intervals. ABN model in black, ABP model in orange (grey in greyscale).

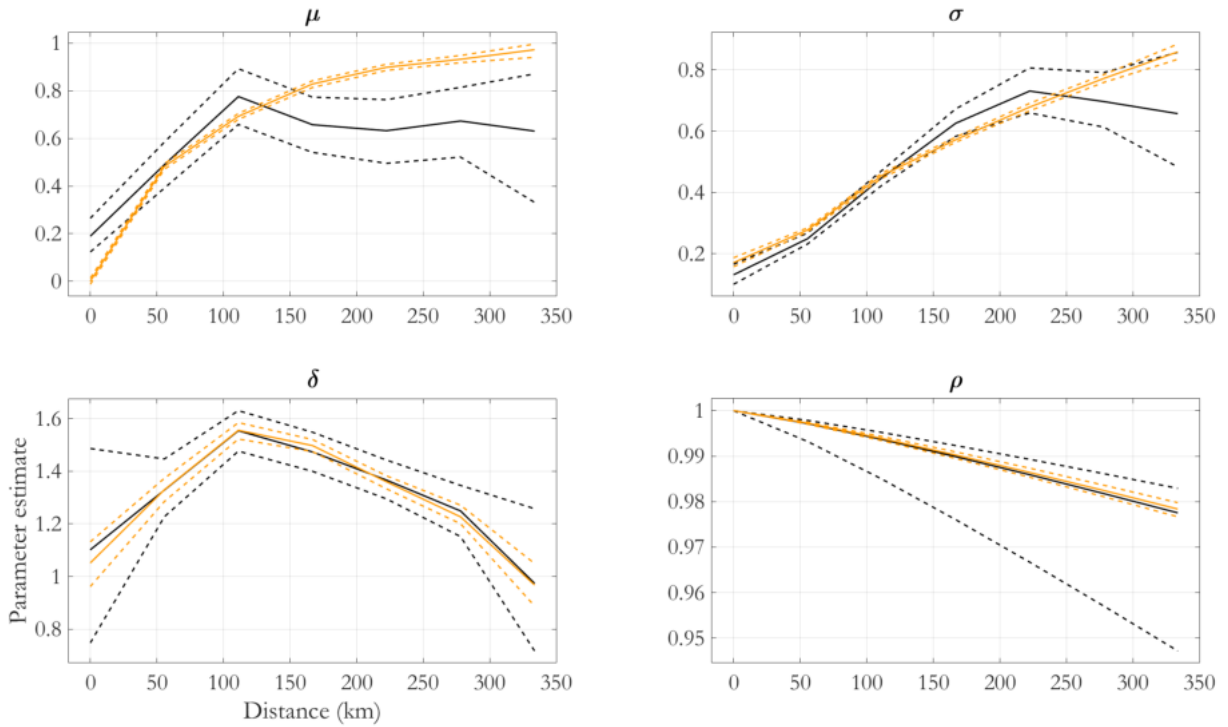


Fig. 5 Estimates for SCE residual parameters μ , σ , δ and ρ with distance d (km). Solid lines represent posterior medians, with dashed lines representing upper and lower limits of empirical 95% posterior credible intervals. ABN model in black, ABP model in orange (grey in greyscale).

There is generally good agreement between estimates for the distance-direction behaviour of SCE model parameters using the ABN and ABP parameterisations. Both show decay of α with distance (Figure 3), and that the rate of decay reduced with increasing distance. At distances between 150km and 250km, the parametric form for α from the ABP parameterisation suggests smaller values; this is accompanied by correspondingly larger values of μ (Figure 5). There is little evidence for directional variation in the decay of α with distance. For $\alpha_{\bullet\ell}$, $\ell = 1, 2, 3, 4$, the estimate reverts to the prior $\text{Unif}(0, 1)$ at distances larger than 250km, since there are no remote locations present at these distance-direction combinations. Figure 4 suggests that the parametric form for β with distance is also generally appropriate. Estimates of β around 0.4 are prevalent for intermediate distances. There is some support for β approaching zero with decreasing distance. Estimates for $\beta_{\bullet 2}(d)$ and $\beta_{\bullet 5}(d)$ in particular appear somewhat different to the others. Estimates for σ and δ , and powered-exponential ρ with distance in Figure 5 show good agreement between ABN and ABP parameterisations. That $\sigma(d=0)$ is greater than zero suggests it may be playing the role of a “nugget effect” at zero distance; the same might be true for $\beta(d=0)$. It is noteworthy that δ is near unity (corresponding to a standard Laplace distribution) for both small and large distances, increasing to around 1.5 at intermediate distances; that is, marginal residual distributions appear more Laplace-like ($\delta = 1$) than Gaussian ($\delta = 2$) for the current application. Since the pair of values $(\alpha_{\bullet\ell}(d), \beta_{\bullet\ell}(d)) = (1, 0)$, for $\ell = 1, \dots, 6$, is not observed for any $d \neq 0$ then the data do not exhibit asymptotic dependence, and Pareto process models would be inappropriate here (Tawn et al., 2018). Achieved sample negative log-likelihoods for ABN and ABP parameterisations were similar relative to the variation in negative log-likelihood MCMC traces, with ABN providing a somewhat lower value. In terms of the deviance information criterion (DIC, Spiegelhalter et al. 2002), the ABP model provided somewhat lower values. A table of DIC values for different comparable models is given in the Appendix.

For practical purposes of simulation, it is useful to examine distance-direction variation of conditional means and standard deviations (Equation (4)) for the ABN and ABP parameterisations. For a conditioning value x_0 at the conditioning site with non-exceedance probability 0.95, these are shown by distance and direction in Figures 6 and 7. There is again good agreement between the two parameterisations. The narrow width of credible intervals for conditional means and standard deviations, for a given distance, using the ABP model reflects the constraining influence of the parametric forms assumed; this is also clear in Figures 3-5 for the underlying model parameters. As previously mentioned, the ABP model has 53 degrees of freedom compared with 107 for the ABN model.

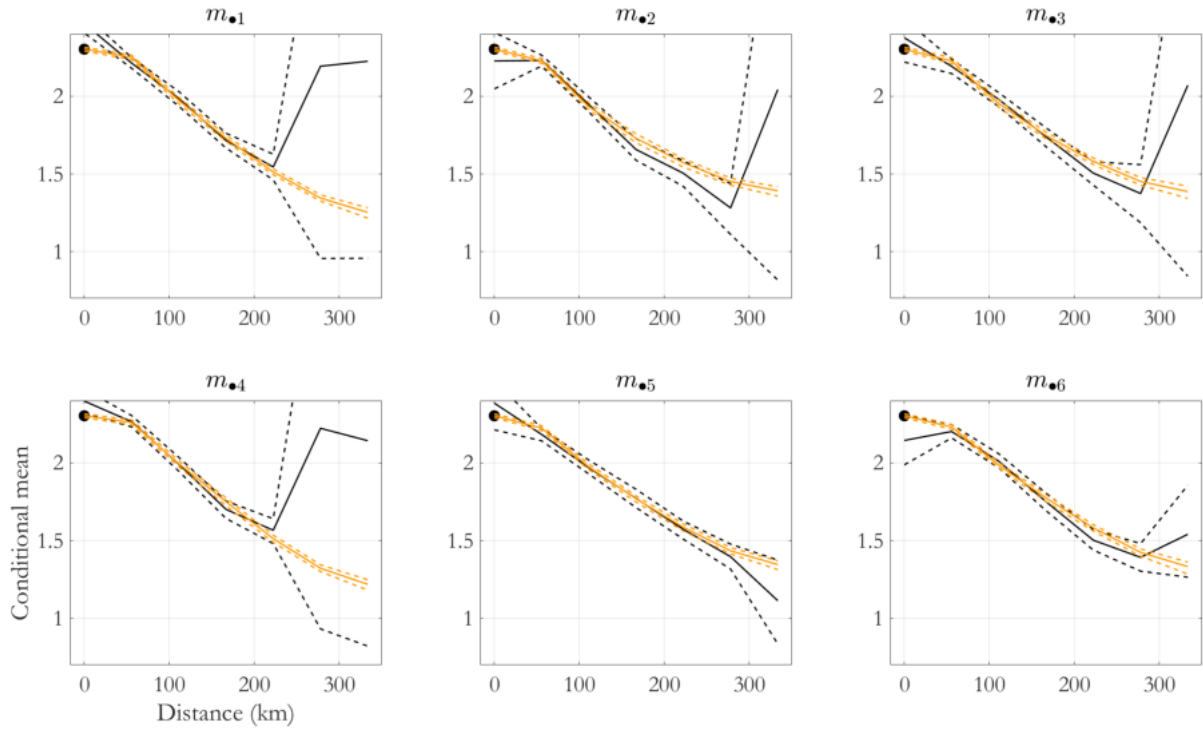


Fig. 6 Estimates for conditional means m with distance d (km) for each of $n_\theta = 6$ directions $0, \pi/3, 2\pi/3, \dots$ for conditioning x_0 with non-exceedance probability 0.95. Solid lines represent posterior medians, with dashed lines representing the upper and lower limits of empirical 95% posterior credible intervals. ABN model in black, ABP model in orange (grey in greyscale).

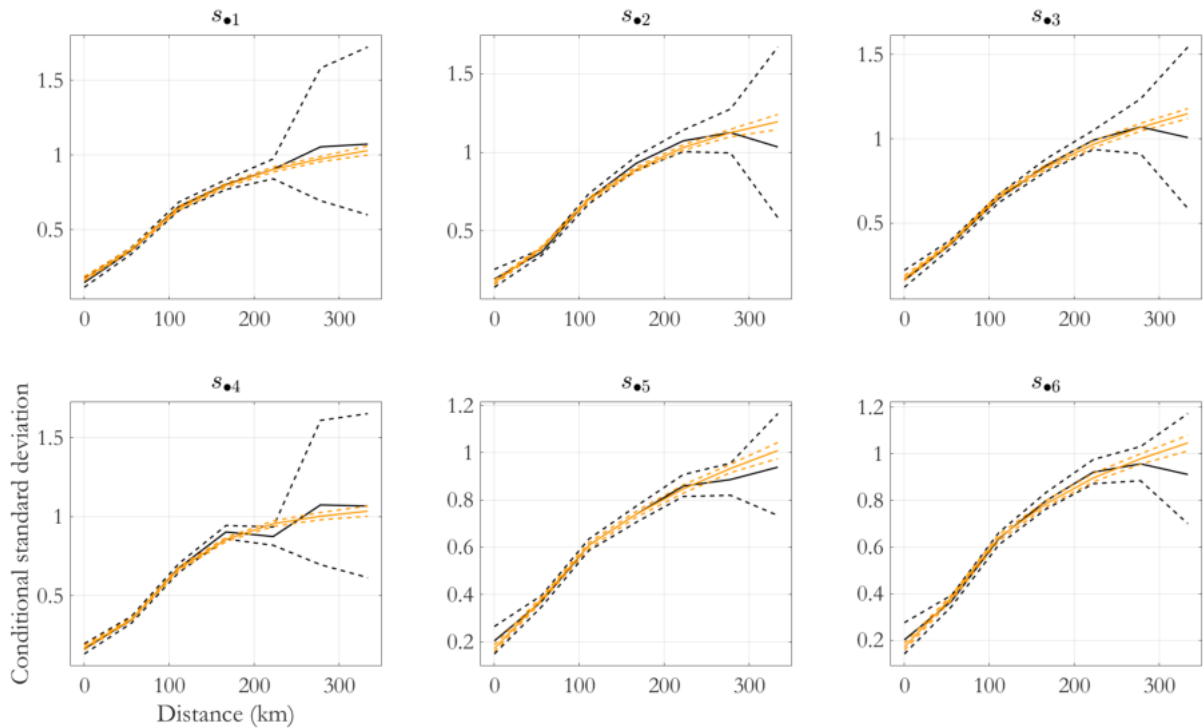


Fig. 7 Estimates for conditional standard deviations s with distance d (km) for each of $n_\theta = 6$ directions $0, \pi/3, 2\pi/3, \dots$ for conditioning x_0 with non-exceedance probability 0.95. Solid black lines represent posterior medians, with dashed lines representing the upper and lower limits of empirical 95% posterior credible intervals. ABN model in black, ABP model in orange (grey in greyscale).

Comparison of observed spatial trajectories with those generated by simulation under the fitted model is illustrated in Figure 8. The figure shows conditional quantiles with non-exceedance probabilities as described in the figure caption, for conditioning values x_0 with non-exceedance probability $\in (0.95, 0.96]$ at the conditioning site, for two specific transect directions, using the ABN model. The left hand transect emanates from the conditioning site to the North-East; the right hand transect emanates from the conditioning site to the North-West. The number of observed trajectories available for estimation of conditional quantiles is small (22 and 20 respectively); 1000 realisations were used to estimate simulated conditional quantiles. Agreement between observation and simulation is good. Figure A1 plots the observed trajectories together with 100 simulated trajectories for each transect. Visual inspection of similar plots generated using the ABP model indicated a similar level of agreement.

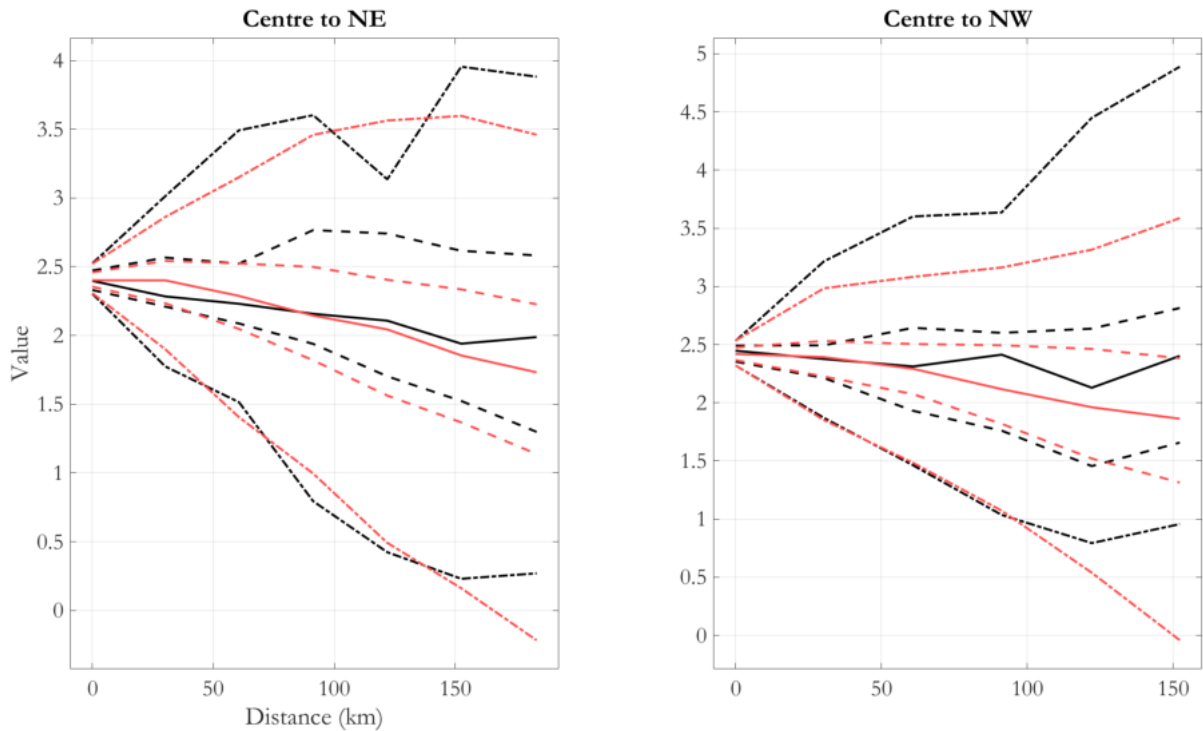


Fig. 8 Quantiles of observed trajectories (black) and corresponding simulated trajectories (red; grey in greyscale) from the fitted ABN model, for conditioning values x_0 with non-exceedance probabilities $\in (0.95, 0.96]$. Quantile probabilities shown are 0.025, 0.25, 0.5, 0.75, 0.975. Left: transect emanating from conditioning site to North-East; right: transect emanating from conditioning site to North-West. Numbers of observed trajectories are 22 (NE) and 20 (NW); 1000 simulated trajectories used in each case to estimate trajectory quantiles.

Figure 9 provides scatter plots of the correlation of residuals at pairs of remote locations falling within different distance intervals with from the conditioning site. The x-coordinate in each panel gives the value of residual correlation, on standard Gaussian scale, calculated using the estimated parameters R_1 and R_2 , and Equation (6). The y-coordinate shows the residual correlation estimated empirically by (a) calculating fit residuals on delta-Laplace scale using $r_{ij}^{DL} = (x_{ij} - m_j)/s_j$ (for i^{th} observation x_{ij} at location j , see Equation (4)), (b) transforming to standard Gaussian scale residuals r_{ij}^N using the probability integral transform and the appropriate estimated marginal delta-Laplace model (see Section 2.2) and (c) estimating the correlation of samples r_{ij}^N and $r_{i'j'}^N$ for appropriate choices of j and j' . Again, the correspondence between residual correlations is good for all distances from the conditioning site. The characteristics of estimated residual correlation structure are illustrated further in Figure A2.

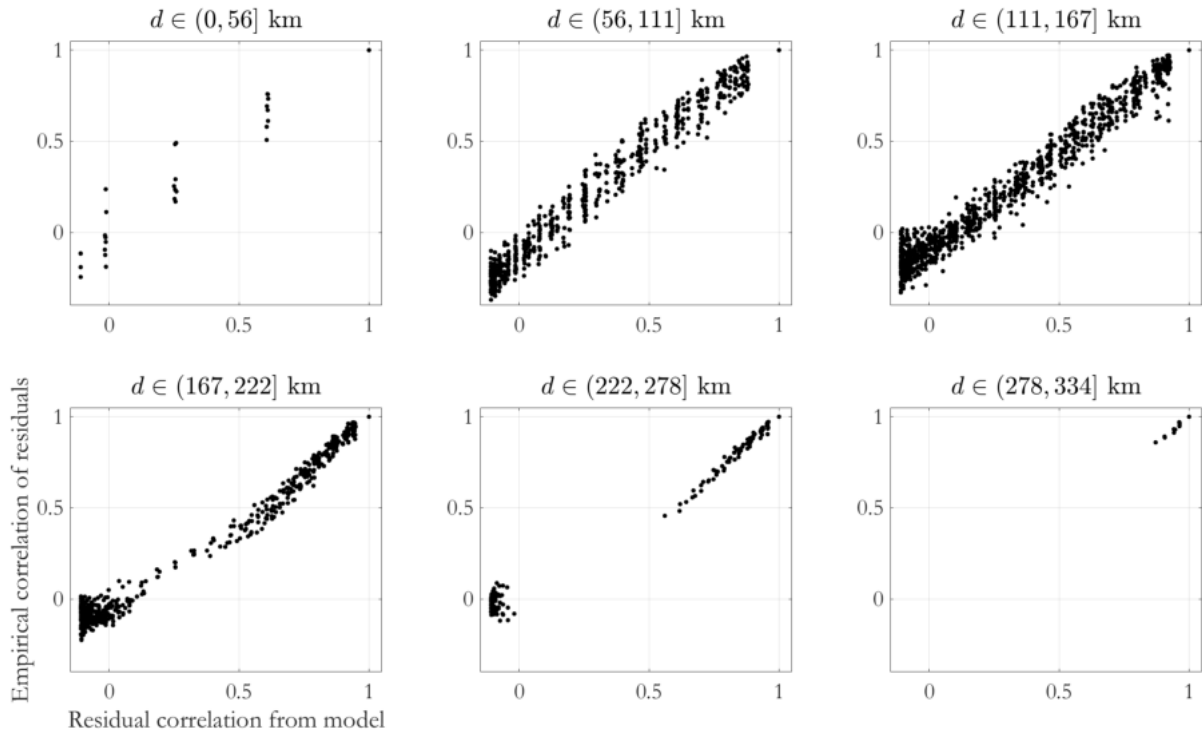


Fig. 9 Observed and estimated residuals, on standard Gaussian scale, for pairs of remote location with distances d from the conditioning site falling within the interval indicated in each panel title. See 4.2 for further discussion.

4.3 Sensitivity studies

The ABN analysis above was repeated for different conditioning threshold choices (corresponding to non-exceedance probabilities of 0.90, 0.95 and 0.975). Estimated characteristics of SCE parameters with distance and direction were very similar to those reported above. The main distinguishing features for the three thresholds examined is the anticipated increase in estimated parameter uncertainty with increasing threshold.

The ABN and ABP models were also estimated without imposition of the conditional quantile constraints. In this case, the decay of α with distance was considerably slower, decreasing to approximately 0.7 at 200 km, than for the constrained analysis. However, the peak value of β at around 100 km to 200 km was larger, at around 0.7. However, these changes were compensated for by smaller estimates for μ very close for zero, and more gradual growth of σ with distance. Therefore overall, conditional means and standard deviations for both ABN and ABP models with and without constraints were found to be very similar. Removal of conditional quantile constraints reduces the negative log-likelihood for posterior median parameters, as might be expected. However, there was little difference in the performance of ABN and ABP models, with or without constraints, in terms of agreement with observed trajectories.

The ABN model was also used with each of a number of different conditioning sites. The general characteristics of parameter estimates with distance are similar to those reported above. Again, there was no evidence for systematic directional variation of α . Other models for the residual correlation were considered. We found that the conditional Gaussian correlation adopted above performed better, in terms of negative log-likelihood for models of similar complexity, and better trajectory simulation.

Choice of number n_d of distance nodes and n_θ of directional nodes was based on preliminary analyses, and physical considerations. For example, the minimum distance between locations on the lattice is approximately 30 km, but physically we do not expect large changes in extremal dependence over such short distances. The maximum distance between a remote location and conditioning site is just over 300 km. With $n_d = 7$, an inter-distance node interval of 56 km provides a reasonably smooth description. Similarly, $n_\theta = 6$ corresponds to directional nodes at $\pi/3$ or 60° intervals, sufficient from a physical perspective to characterise any directional variability. A non-directional ($n_\theta = 1$) ABN model was also estimated (with and without conditional quantile constraints) yielding values of negative log-likelihood and DIC considerably higher than for $n_\theta = 6$. Truncated zero-mean Gaussian priors for μ (on domain $[-1, 1]$) and σ (on domain $(0, \sqrt{2})$) were also evaluated, for different choices of Gaussian standard deviation, for the ABN model; no appreciable influence on parameter estimates was observed for reasonable choices of standard deviation.

5 Discussion

In this paper we develop a two-dimensional spatial conditional extremes model incorporating distance-directional variation of model parameters, with parametric and non-parametric model parameter representations. The SCE model allow asymptotic dependence at short inter-location distances, leading to asymptotic independence, and eventually perfect independence, as distance increases. The SCE residual process is given delta-Laplace (generalised Gaussian) marginal distributions, and a (Gaussian-scale) conditional Gaussian correlation structure with powered exponential correlation function. The model is applied to estimate the conditional extremal dependence of a sample of hindcast storm peak significant wave height data for a lattice of 150 locations in the North Sea. The model fit is found to be reasonable with or without applied conditional quantile constraints, and is relatively insensitive to conditioning threshold choice, within the interval considered. Prior transformation of data to standard Laplace marginal scale is achieved using non-stationarity directional marginal extreme value analysis.

For the current analysis, an occurrence of a large extra-tropical cyclone would cover the entire spatial domain considered. We therefore expect to find some dependence of extreme values of significant wave height over the whole domain. The nature of that dependence is less clear a priori. The importance of allowing for different forms of extremal dependence with distance is illustrated in Figure 2, for example, which shows that SCE parameter α decays from approximately unity to zero over approximately 300 km for North Sea storm peak significant wave height. The current analysis suggests that there is little directional variation of extremal dependence of storm peak significant wave height for the lattice considered. There is some evidence supporting relatively large values of β for direction $\pi/3$ and small values at $4\pi/3$. There are a number of physical features of the North Sea environment which might explain such differences, including prevailing wind and storm directions, and bathymetric effects.

In the absence of conditional quantile constraints, the values of α parameter estimates are generally higher than when then constraints are applied. The values of parameter estimates for μ are near zero in the absence of conditional quantile constraints, but positive when constraints are applied. The net result of these competing effects is that conditional means and standard deviations are very similar for lower conditioning values x_0 ($x_0 > u$) at the conditioning site. Nevertheless, for conditioning values x_0 corresponding to large non-exceedance probabilities, we might expect that parameter estimates from unconstrained models would provide larger values for conditional means, since αx_0 for $\alpha \in [0, 1]$ increases more quickly with $x_0 \in \mathbb{R}_{>0}$ than μx_0^β , $\mu \in [0, 1]$ and $\beta \in [0, 1]$. However, no obvious evidence for this difference was observed for the current sample.

One source of uncertainty not explored in the current work is the effect of uncertainty in marginal fitting of directional generalised Pareto models for the 150 locations. Transformation to standard Laplace marginals was performed using posterior median estimates. Exploratory analysis (and plots such as Figure 2) suggest that marginal fitting is reasonable. It would be interesting however to propagate uncertainty from marginal fitting through to SCE estimation, or better to jointly estimate marginal and SCE models.

Our results suggest a number of potential avenues for further method development and application. From a physical perspective, we are keen to ensure that the SCE model provides an adequate representation of the conditional spatial structure of large ocean storms. For the sizes of spatial domains examined here (and in Shooter et al. 2019, Shooter 2020), this would appear to be the case. Current work is considering yet larger spatial domains. The approach would appear to be ideally suited for characterisation of spatial ocean surface roughness, e.g. as measured by satellite altimetry (Young and Ribal, 2019).

The SCE model allows multiple different classes of extremal dependence. Our results indicate that $\alpha \approx 1$ occurs for small distances d , but that $\alpha < 1$ otherwise. This suggests that a modelling strategy admitting both asymptotic dependence and asymptotic independence is essential to characterise large ocean storms. In particular, competitor approaches involving max-stable and Pareto process models, which admit only asymptotic dependence, are not appropriate. Further, the SCE model is computationally relatively efficient to estimate, and does not require the adoption of composite likelihoods for tractable inference. We believe that the model proposed in this work is a flexible and practical approach, with some obvious advantages to estimation of max-stable and Pareto processes for relatively large neighbourhoods of spatial locations.

Acknowledgements The authors would like to thank Jenny Wadsworth (Lancaster University) and Matthew Jones, David Randell and Ross Towe (Shell) for helpful discussions. We also thank two anonymous reviewers and an associate editor for comments on an earlier draft of the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Brown BM, Resnick SI (1977) Extreme values of independent stochastic processes. *J Appl Probab* 14:732–739
- Coles S, Heffernan J, Tawn J (1999) Dependence measures for extreme value analyses. *Extremes* 2:339–365
- Davison AC, Padoan SA, Ribatet M (2012) Statistical modelling of spatial extremes. *Statist Sci* 27:161–186
- Ewans KC, Jonathan P (2008) The effect of directionality on northern North Sea extreme wave design criteria. *J Off-shore Arct Eng* 130:041604:1–041604:8

- Ferreira A, de Haan L (2014) The generalized Pareto process; with a view towards application and simulation. *Bernoulli* 20(4):1717–1737
- Ferro CAT, Segers J (2003) Inference for clusters of extreme values. *J Roy Statist Soc B* 65:545–556
- de Fondeville R, Davison AC (2020) Functional peaks-over-threshold analysis. 2002.02711
- Heffernan JE, Resnick SI (2007) Limit laws for random vectors with an extreme component. *Ann Appl Probab* 17:537–571
- Heffernan JE, Tawn JA (2004) A conditional approach for multivariate extreme values. *J R Statist Soc B* 66:497–546
- Huser RG, Wadsworth JL (2019) Modeling spatial processes with unknown extremal dependence class. *J Am Statist Soc* 114:434–444
- Keef C, Papastathopoulos I, Tawn JA (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
- Kereszturi M, Tawn J, Jonathan P (2016) Assessing extremal dependence of North Sea storm severity. *Ocean Eng* 118:242–259
- Ledford AW, Tawn JA (1996) Statistics for near independence in multivariate extreme values. *Biometrika* 83:169–187
- Mardia KV, Kent JT, Bibby JM (1979) *Multivariate Analysis*. Academic Press
- Oceanweather (2002) North European Storm Study User Group Extension and Reanalysis Archive. Oceanweather Inc.
- Reich BJ, Shaby BA (2012) A hierarchical max-stable spatial model for extreme precipitation. *Ann Appl Stat* 6:1430–1451
- Ribatet M (2013) Spatial extremes: max-stable processes at work. *J Soc Fr Statistique* 154:156–177
- Roberts GO, Rosenthal JS (2009) Examples of adaptive MCMC. *J Comp Graph Stat* 18:349–367
- Ross E, Kereszturi M, van Nee M, Randell D, Jonathan P (2017a) On the spatial dependence of extreme ocean storm seas. *Ocean Eng* 145:1–14
- Ross E, Randell D, Ewans K, Feld G, Jonathan P (2017b) Efficient estimation of return value distributions from non-stationary marginal extreme value models using Bayesian inference. *Ocean Eng* 142:315–328
- Schlather M (2002) Models for stationary max-stable random fields. *Extremes* 5:33–44
- Shooter R (2020) *Conditional Models for Spatial Extremes*. PhD Thesis, Lancaster University, UK
- Shooter R, Ross E, Tawn JA, Jonathan P (2019) On spatial conditional extremes for ocean storm severity. *Environmetrics* 30:e2562
- Smith RL (1990) Max-stable processes and spatial extremes. Unpublished article, available electronically from [www-statuncedu/postscript/rs/spatexpdf](http://www.statuncedu/postscript/rs/spatexpdf)
- Spiegelhalter DJ, Best N, Carlin BP, van der Linde A (2002) Bayesian measures of model complexity and fit (with discussion). *J Roy Statist Soc B* 64:1–34
- Tawn JA, Shooter R, Towe R, Lamb R (2018) Modelling spatial extreme events with environmental applications. *Spat Stat* 28:39–58
- Vallis G (2017) *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation*. Cambridge University Press
- Wadsworth J, Tawn J (2012) Dependence modelling for spatial extremes. *Biometrika* 99:253–272
- Wadsworth JL, Tawn JA (2019) Higher-dimensional spatial extremes via single-site conditioning. In submission
- Wadsworth JL, Tawn JA, Davison AC, Elton DM (2017) Modelling across extremal dependence classes. *J Roy Statist Soc C* 79:149–175
- Young IR, Ribal A (2019) Multiplatform evaluation of global trends in wind speed and wave height. *Science* 364:548–552

Appendices

A Grouped adaptive MCMC

Under the ABN model (with piecewise linear distance-direction representations for α and β) the parameter set Ω to be estimated is $\{\{\alpha_{k\ell}, \beta_{k\ell}\}_{\ell=1}^{n_\theta}, \mu_k, \sigma_k, \delta_k\}_k^{n_d}, R_1, R_2\}$ with $n_P = 2n_d n_\theta + 3n_d + 2$ parameters.

Prior specification: Impose uniform prior distributions for each parameter. Explicitly, $\alpha_{k\ell} \sim \text{Unif}(0, 1)$, $\beta_{k\ell} \sim \text{Unif}(0, 1)$, $\mu_k \sim \text{Unif}(-1, 1)$, $\sigma_k \sim \text{Unif}(0, \sqrt{2} + 0.1)$ and $\delta_k \sim \text{Unif}(0.5, 2.5)$ for all $k = 1, 2, \dots, n_d$, $\ell = 1, 2, \dots, n_\theta$, and $R_1 \sim \text{Unif}(0, 100)$, $R_2 \sim \text{Unif}(0, 2)$,

Starting solution: We obtain a random starting solution $\Omega^{(0)}$ by sampling the elements of Ω from their prior distributions, verifying that the starting solution has a valid likelihood (satisfying the conditional quantile constraints if applied).

Metropolis-within-Gibbs for n_{MiG} iterations: Writing $\Omega_k^{(i)}$ as the value of the k th parameter of Ω at the i th iteration, use adaptive random walk Metropolis-within-Gibbs scheme for n_{MiG} ($> 2n_P$) iterations. That is, for $i = 2, \dots, n_{MiG}$, update each $\Omega_k^{(i)}$ in turn, proposing candidate value $\Omega_k^{(i)c}$ from distribution

$$Q_{MiG} = N(\Omega_k^{(i-1)}, 0.1^2).$$

Grouped adaptive Metropolis-within-Gibbs for n_{GA} iterations: For $i > n_{MiG}$, use a grouped adaptive random walk Metropolis-within-Gibbs scheme, updating groups $\Omega_{G_k}^{(i)} = \{\{\alpha_{k\ell}, \beta_{k\ell}\}_{\ell=1}^{n_\theta}, \mu_k, \sigma_k\}$ jointly for $k = 1, 2, \dots, n_d$ in turn, before

updating $\{\delta_k\}_{k=1}^{n_d}$, R_1 and R_2 separately as before. Propose candidate $\boldsymbol{\Omega}_{G_k}^{(i)c}$ from distribution

$$Q_{GA} = (1 - \beta)N(\boldsymbol{\Omega}_{G_k}^{(i-1)}, 2.38^2 \mathbf{C}_i) + \beta N(\boldsymbol{\Omega}_{G_k}^{(i-1)}, 0.1^2 / (4n_\theta))$$

where $\beta = 0.05$, as suggested by Roberts and Rosenthal (2009), and \mathbf{C}_i is the empirical variance-covariance matrix of the parameters $\boldsymbol{\Omega}_{G_k}^{(i)}$ from the previous i iterations.

Iteration to convergence: Throughout, a candidate state is accepted using the standard Metropolis-Hastings acceptance criterion. Since prior distributions for parameters are uniform, and proposals symmetric, this is just a likelihood ratio. Candidates lying outside their prior domains, or violating the conditional quantile constraints if applied, are rejected.

Under the ABP model (with parametric forms for α and β with distance only), the estimation scheme is simplified so that the full set of parameters is updated at the grouped adaptive stage. The resulting procedure is the same as the original scheme of Roberts and Rosenthal (2009). Uniform priors $A_{1\ell} \sim \text{Unif}(1, 20)$, $A_{2\ell} \sim \text{Unif}(0.1, 5)$, $B_{1\ell} \sim \text{Unif}(0.1, 1)$, $B_{2\ell} \sim \text{Unif}(1, 5)$ and $B_{3\ell} \sim \text{Unif}(0, 20)$, $\ell = 1, 2, \dots, n_\theta$ were applied. The prior distributions for all other parameters under the ABN and ABP models are the same.

B Conditional quantile constraints

We optionally restrict the space of feasible combinations of α and β to ensure that conditional quantiles from asymptotic independent models do not exceed those from asymptotic dependent models, as proposed by Keef et al. (2013). For any pair α and β corresponding to distance d_k and direction θ_ℓ , $k = 1, 2, \dots, n_d$, $\ell = 1, 2, \dots, n_\theta$ on the distance-direction lattice, we require either

$$\alpha \leq \min\{1, 1 - \beta z(q)\nu^{\beta-1}, 1 - \nu^{\beta-1}z(q) + \nu^{-1}z^+(q)\}$$

or

$$1 - \beta z(q)\nu^{\beta-1} < \alpha \leq 1, \text{ and} \\ (1 - \beta^{-1})\{\beta z(q)\}^{1/(1-\beta)}(1 - \alpha)^{-\beta/(1-\beta)} + z^+(q) > 0.$$

In the above, ν is a value of the conditioning variate (on standard Laplace scale) above the threshold level at which the SCE model is applied. Further $z^+(q)$ is the quantile of the distribution of standardised residuals from the conditional extremes model with non-exceedance probability q , and $z^-(q)$ is quantile of the distribution of standardised residuals from the conditional extremes model assuming asymptotic positive dependence (i.e. by imposing $\alpha = 1$, $\beta = 0$) with non-exceedance probability q . In practice, as suggested by Keef et al. (2013), it is sufficient to satisfy the constraints above for $q = 1$ and ν equal to the maximum observed value of the conditioning variate.

C Supporting diagnostic plots

Figure A1 supports the discussion in Section 4 around Figure 8 regarding agreement between simulated spatial trajectories under the fitted SCE model, and observed trajectories. The figure shows 100 simulated trajectories for transects emanating to the North-East and to the North-West, from the central conditioning location, in black. The trajectories all correspond to a conditioning value x_0 with non-exceedance probability in the interval (0.95, 0.96). Also shown are observed trajectories available (22 and 20 in number respectively for the left and right hand cases), for the same conditioning, in red.

Figure A2 illustrates the fitted conditional Gaussian correlation function of Equation (6). Consider two remote locations, equidistant from a third conditioning location. For each of the curves in the figure, the x-coordinate of the black disc indicates the distance between the remote locations. The curve passing through each disc gives the value of conditional residual correlation as a function of the (common) distance of the remote locations from the conditioning site. Thus, the leftmost curve pertains to two remote locations approximately 30 km apart. When the distance of these locations to the conditioning site is large, the remote locations have large residual correlation since conditioning has effectively no effect: the residual process is essentially unconstrained, and the value of correlation determined by the powered exponential ρ . However, as distance to the conditioning site is reduced, residual correlation decreases, since more sample variation is explained by the SCE model (as opposed to the SCE conditional residual process). The minimum value of distance to conditioning site occurs when the three locations are collinear, with the conditioning site mid-way between remote locations. For this arrangement, conditioning induces a small negative residual correlation between remote locations. As the distance between remote locations increases, for different curves left to right, the maximum conditional residual correlation is reduced.

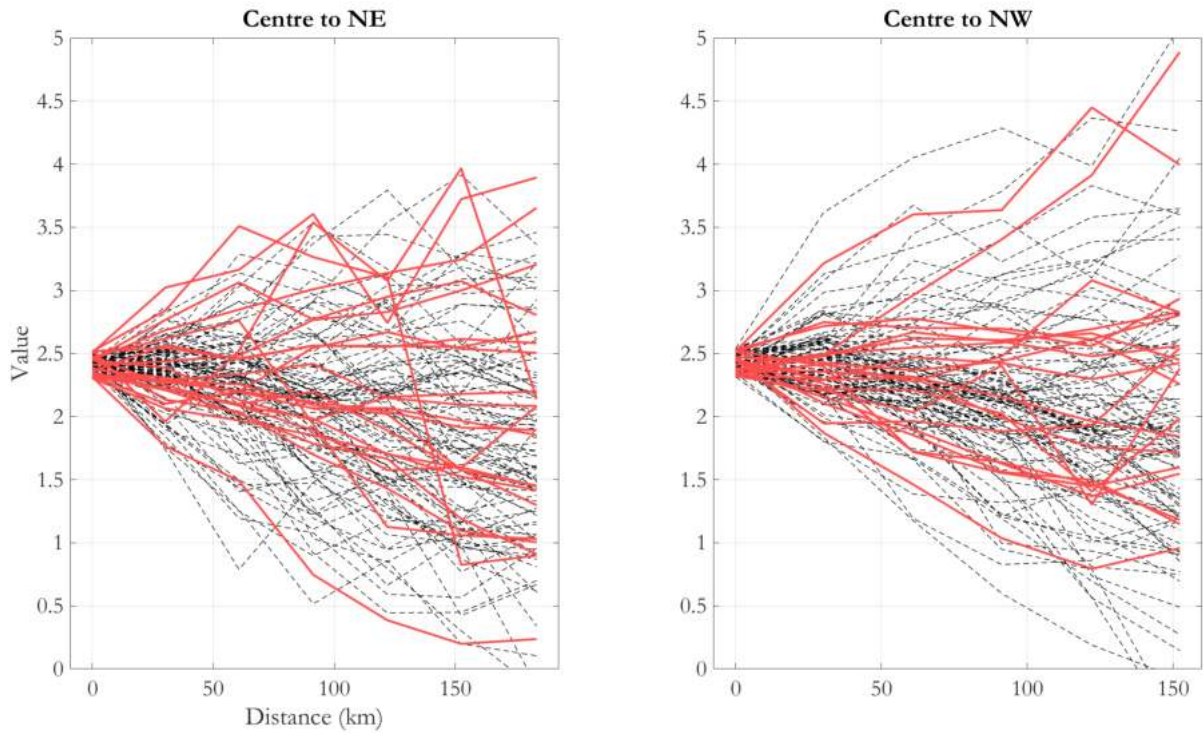


Fig. A1 Observed (black dashed) and simulated (red; grey in greyscale) trajectories from the fitted ABN model, for conditioning value x_0 with non-exceedance probabilities in the interval $(0.95, 0.96]$. Left: transect emanating from conditioning site to North-East; right: transect emanating from conditioning site to North-West. Numbers of observed trajectories are 22 (NE) and 20 (NW); 100 simulated trajectories shown in each case.

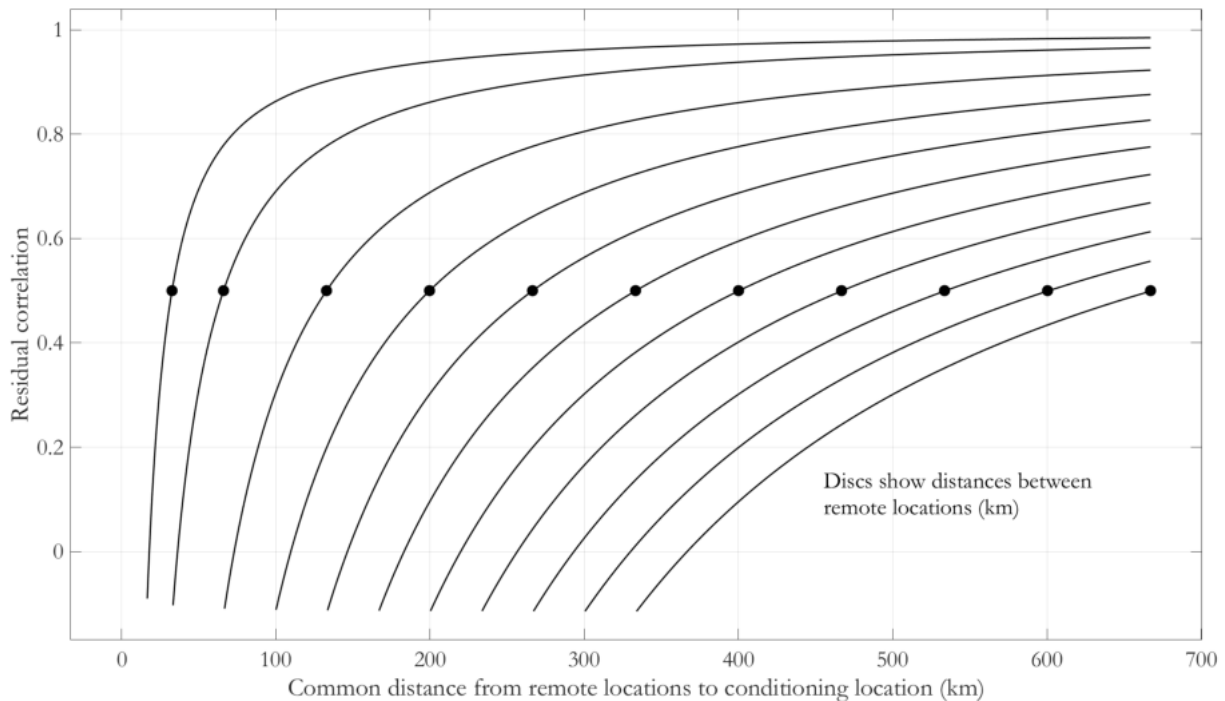


Fig. A2 Illustration of the estimated Gaussian-scale residual correlation function. Consider a pair of remote locations with common distance to conditioning site. Distance between remote locations is indicated by x-coordinate of black disc for each curve. Curve through disc indicates the residual correlation between remote locations as a function of distance to the conditioning site.

D Deviance and DIC

The table below provides MCMC deviance statistics and DIC values to support the discussion in Sections 4.

Model	n_θ	Constraints	$\overline{D(\Omega)}$	$\text{sd}(D)$	$D(\bar{\Omega})$	DIC
ABN	1	Y	1770	8	1740	1800
ABN	6	Y	1550	11	1490	1620
ABP	6	Y	1580	3	1570	1580
ABN	1	N	1730	7	1760	1760
ABN	6	N	1470	10	1422	1520
ABP	6	N	1510	4	1500	1510

Columns in the table are as follows: n_θ is the number of directional parameters in the model, ‘Constraints’ indicates whether the conditional quantile constraints of Keef et al. (2013) were imposed, $\overline{D(\Omega)}$ is the posterior mean deviance, $\text{sd}(D)$ is its standard deviation and $D(\bar{\Omega})$ the deviance calculated using posterior mean parameters. In summary, there is evidence in favour of a directional model, since lower values of deviance and DIC are obtained for $n_\theta = 6$ compared with $n_\theta = 1$ both with and without application of conditional quantile constraints. DIC values for ABP are somewhat lower than for the corresponding ABN, but differences are small.