Assessing extremal dependence of North Sea storm severity

Monika Kereszturi [a,*], Jonathan Tawn [a], Philip Jonathan [b]

[a] *Mathematics and Statistics, Lancaster University, UK*
[b] *Shell Research Ltd., Manchester, M22 0RR, UK*

A B S T R A C T

Extreme value theory provides asymptotically motivated methods for modelling the occurrences of extreme values of oceanographic variables, such as significant wave height at a single location. Problems are often multivariate or spatial in nature, where interest lies in the risk associated with joint occurrences of rare events, e.g. the risk to multiple offshore structures from a storm event. There are two different classes of joint tail behaviour that have very different implications: asymptotic independence suggesting that extreme events are unlikely to occur together, and asymptotic dependence implying that extreme events can occur simultaneously. It is vital to have good diagnostics to identify the appropriate dependence class. If variables are asymptotically independent, incorrectly assuming an asymptotically dependent model can lead to overestimation of the joint risk of extreme events, and hence to higher than necessary design costs of offshore structures. We develop improved diagnostics for differentiating between these two classes, which leads to increased confidence in model selection. Application to samples of North Sea sea-state and storm-peak significant wave height suggests that tail dependence changes with wave direction and distance between spatial locations, and that in most cases asymptotic independence seems to be the more appropriate assumption.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Metocean and coastal engineers are generally interested in the estimation of design conditions for a specific spatial location. Of primary concern is the estimation of return values for wind-, wave- and current-related variables corresponding to a long return period. However, there are instances when the engineer is concerned with characterising joint occurrences of rare events at different locations or multiple hazard occurrence at a single location. Examples of the former type might include (i) risk to multiple offshore facilities from a single cyclone or hurricane event, (ii) impact of a single storm surge event on multiple coastal locations, and (iii) insurance risk to a moving vessel traversing an ocean basin. Examples of the latter type include the combined risk of wind-, wave- and current-induced forces on a fixed structure, or motions of a floating structure. In such circumstances, a description of the joint structure of extreme events is necessary. This paper develops diagnostic methods for assessing the appropriate form of dependence structure to model metocean data sets and explores the implications of getting this decision right and wrong in application to significant wave height at neighbouring locations.

To help formulate the issues, consider the simplified example of two identically distributed random variables X and Y. For example, the random variables X and Y could represent significant wave height at neighbouring locations. We are typically interested in cases where both variables are large together, i.e., we need to find the probability $\Pr\{X > x, Y > x\}$ for large x, where x could correspond to the 50-year return level of each variable marginally. For large x, this joint probability will be small, so for studying dependence between extremes, it is typical to focus on the conditional probability $\Pr\{Y > x | X > x\}$, recognising that $\Pr\{X > x, Y > x\}$ can be recovered from this conditional probability by multiplying by the marginal probability $\Pr\{X > x\}$. Two special cases of dependence structure between (X,Y) correspond to the variables being independent (which we term perfect independence) or perfectly dependent (i.e., $X=Y$ always). Then,

$$\Pr\{Y > x | X > x\} = \begin{cases} \Pr\{Y > x\} & \text{if perfect independence} \\ 1 & \text{if perfect dependence.} \end{cases} \quad (1)$$

As x tends to the upper endpoint x^* of the common marginal distribution of X and Y, the limiting probability χ then will satisfy

$$\chi = \lim_{x \to x_*} \Pr\{Y > x | X > x\} = \begin{cases} 0 & \text{if perfect independence} \\ 1 & \text{if perfect dependence.} \end{cases} \quad (2)$$

The chance of joint occurrence of extremes is completely different in these cases. We cannot assume observations of variables,

---

* Corresponding author.
  E-mail address: m.kereszturi@lancaster.ac.uk (M. Kereszturi).

such as significant wave height at different locations, to be perfectly independent nor perfectly dependent of one another. An occurrence of an extreme event at one location influences the chance of an extreme event at a location in its neighbourhood. The most well known dependence model is the multivariate normal distribution, or equivalently the Gaussian copula, which is widely used across fields of empirical modelling, including spatial interpolation using Kriging or Gaussian process regression. Consideration of the properties of this bivariate distribution with correlation parameter ρ_P (where we use the subscript P to denote the Pearson correlation coefficient) suggests an alternative form for $\Pr\{Y > x | X > x\}$ for large x, which is intermediate to perfect dependence and perfect independence for all $0 < \rho_P < 1$, but has $\chi = 0$.

Hence, to model tail dependence reliably we need models which will account for the different possibilities for how the conditional probability $\Pr\{Y > x | X > x\}$, and its extensions for our metocean examples, vary for large x. Clearly, models which capture intermediate forms between perfect independence and perfect dependence are required, and we will explain that these correspond to two broad classes, with either $0 < \chi < 1$ or $\chi = 0$, termed asymptotic dependence and asymptotic independence, respectively. For asymptotic independence, it is impossible for the most extreme events for both variables (X, Y) to occur simultaneously. For asymptotic dependence, if X is extreme it is also possible for Y to be simultaneously extreme. In theory, there are also cases where $\Pr\{Y > x | X > x\}$ decays to 0 faster than $\Pr\{Y > x\}$ as $x \to x_*$, corresponding to negative dependence between (X, Y), but for our metocean examples this typically does not occur. Hence, the key focus in the paper will be on positively dependent variables.

As we are interested in joint probabilities for large events, it is natural to look for asymptotic arguments that provide guidance regarding the possible form that these probabilities can take. Extreme value theory provides this framework for one variable. For example, although the distribution of a continuous random variable can take any form, subject to weak convergence conditions, above a sufficiently high threshold, peaks over threshold of this variable follow the generalised Pareto distribution approximately (Pickands, 1975). This motivates the adoption of a specific tail functional form which can be estimated, and from which extrapolations to long return period events can be obtained. For metocean data this has been found to provide a reliable method for return level estimation of a single variable. Furthermore, the parameters of this distribution may be allowed to vary with one or more covariates, such as direction or season (e.g., Jonathan et al., 2008).

Unfortunately, analogous specific asymptotic results are not available from extreme value theory for the joint distribution of extremes of two or more variables. For this reason, a number of different approaches, none of which is completely satisfactory, have emerged for modelling multivariate and spatial extremes. In the applied statistics literature popular methodologies are motivated by consideration of componentwise maxima and regular variation (e.g., Coles and Tawn, 1994; Cooley et al., 2012; Davis et al., 2013; Davison et al., 2012; de Haan and de Ronde, 1998). These lead to tractable models for joint extremes. However, they require the assumption of asymptotic dependence between extremes of pairs of random variables, i.e, $\chi > 0$ (as defined in (2)).

Suppose that X and Y represent significant wave height at two different locations $\boldsymbol{r}_X$ and $\boldsymbol{r}_Y$, respectively. Then if the distance $d_{XY} = |\boldsymbol{r}_Y - \boldsymbol{r}_X|$ between these locations is small relative to the spatial extent of an ocean storm, the assumption of $\chi > 0$ may well be valid. However, as d_{XY} increases, the assumption appears increasingly untenable, since the same physical cause (i.e., a low-pressure field causing wind causing waves) is increasingly unlikely

to produce simultaneous extreme values of significant wave height at both locations despite there being some dependence between observed values of X and Y. Thus, it seems reasonable for these d_{XY} that $\chi = 0$ corresponding to asymptotic independence. If we continue to increase d_{XY}, eventually the locations are so distant that the characteristics of X and Y cannot possibly be influenced by the same physical cause. In this limit, we can say that X and Y exhibit perfect independence of extremes.

The difference between joint extremes estimated under asymptotic dependence and asymptotic independence assumptions can be very large. For example, for storm-peak significant wave height from a typical location in North Sea application we present in Section 5, marginal extreme value analysis suggests that the 100-year return level is approximately 15.5 m and that this does not change much over neighbouring sites. If a 100 year event occurs at one location, the median of the conditional distribution at a neighbouring location will also be approximately 15.5 m, under strong asymptotic dependence. However, under the assumption of an asymptotically independent Gaussian process model with correlation parameter 0.8, the median of the conditional distribution at the neighbouring location reduces to approximately 13.5 m (see Section 5.3 for details). In fact all quantiles of the conditional distribution reduce by approximately 2 m; a reduction of 2 m in the conditional distribution is of considerable practical importance.

The majority of applicable models for multivariate and spatial extremes assume asymptotic dependence, leading potentially to over-estimation of the risk of joint occurrences of extremes of X and Y if asymptotic dependence is incorrectly assumed. On the other hand, if asymptotic independence is assumed when the data are in fact asymptotically dependent, the risk of the extremes will be underestimated. Conservatism in design is desirable, provided that the extent of conservatism is itself reasonable and well bounded. For extreme dependence modelling this conservatism may be unbounded. For instance, if the variables are falsely assumed to be asymptotically dependent (with $\chi > 0$) when they were actually perfectly independent, then $\Pr\{X > x, Y > x\}$ can be over-estimated by the factor $\chi / \Pr\{X > x\} \to \infty$ as $x \to x_*$. This is an issue, as over-conservatism in some aspects of design can lead to increased risks in other aspects, particularly when resources are constrained. If asymptotic dependence is not a reasonable assumption for a particular application, alternative approaches admitting asymptotic independence must be used. The conditional extremes model of Heffernan and Tawn (2004) and the hybrid model of Wadsworth and Tawn (2012) are possibilities, but the field of modelling asymptotically independent spatial extremes is considerably less well developed in general.

These risks of over- and under-design illustrate that we must justify the specific assumptions made concerning extremal dependence in any metocean application. Therefore reliable diagnostic methods for identifying between the asymptotic independence and asymptotic dependence classes of extremal dependence based on a sample of data are required. Tawn and co-workers (e.g., Ledford and Tawn, 1997; Heffernan and Tawn, 2004; Wadsworth and Tawn, 2013) have developed the description of asymptotic independence, and have offered different characterisations. A number of diagnostic statistics for the form of extremal dependence have been proposed (see e.g., Coles et al., 1999) and applied in metocean settings (see e.g., Eastoe et al., 2013). In ideal situations, these provide good guidance regarding the nature of extremal dependence in play in a particular application. However, for a typical metocean application, where sample size might be small, the distributions of X and Y vary with respect to multiple covariates, and are observed only with error (either from a physical hindcast model or from observation), such diagnostics tend to be inconclusive.

This paper attempts to refine the way diagnostics for extremal

dependence are employed in practice, to improve their interpretability. We develop improved diagnostics for differentiating between asymptotic independence and dependence, which leads to increased assurance in model selection. In turn, we hope this facilitates better understanding of the dependence of extremes of ocean environmental variables, so that joint design conditions may be estimated with improved confidence. A key component of our method is the novel idea to supplement measures of extremal dependence with a measure of the dependence for the body of the data. This enables the assessment of extremal dependence in the light of the general dependence and we find that it noticeably increases diagnostic performance.

The layout of the paper is as follows: in Section 2 we summarise measures of asymptotic independence and illustrate their value for a range of the widely used extremal models; in Section 3 we explore a range of inference methods to aid the diagnostic process; in Section 4 we conduct a simulation study to illustrate the performance of the diagnostic methods for a range of distributions and samples sizes; and in Section 5 we apply the methods to North Sea significant wave height data; show that the tail dependence changes with wave direction and distance between spatial locations; and infer that in many cases the assumption of asymptotic independence is more plausible, and in all cases that the presence of asymptotic independence cannot be dismissed. In Section 6 we conclude with a discussion.

2. Extremal dependence

2.1. Different dependence measures for typical and extreme value data

In standard statistical analysis the Pearson correlation coefficient ρ_P is often used to determine the dependence between two variables. A weakness of this measure is that it is not invariant to non-linear monotone increasing transformations of the marginal variables. For oceanographic applications this means we have different correlations between wave heights at two locations and their associated wave powers. To overcome this issue we use Spearman's rank correlation coefficient ρ (Spearman, 1904), but we could have also used the broadly equivalent Kendal τ. In the context of oceanographic variables it is reasonable to interpret a value of ρ close to 0 (and 1) as indicating near independence (and near perfect dependence) respectively.

Spearman's rank correlation measures dependence for the body of the distribution, i.e., typical values, as all observations are given equal weighting in the assessment of dependence. When dealing with extremes, the dependence can no longer be measured using any form of correlation estimate, as the dependence in the body can be quite different from the extremal or tail dependence. This property is illustrated in Sections 4 and 5. A nice illustration of this feature arises in time series: a Gaussian autoregressive process has temporal dependence as measured by the autocorrelation function but no tail dependence as measured by χ (Sibuya, 1960); whereas the reverse holds for the ARCH and GARCH processes (de Haan and Resnick, 1989; Laurini and Tawn, 2012).

Therefore in Section 2.2 we introduce some measures of dependence structure specific to extreme values and illustrate these on different probabilistic models for spatial processes. As there is limited information in the extreme values, there is typically not sufficient evidence from these extremal dependence measures for them to be effective diagnostic tools when identifying the form of the extremal dependence, e.g., whether the processes the data are observed on are asymptotically dependent or asymptotically independent. Therefore, the novel approach we take in this paper is

to supplement these measures of extremal dependence with ρ, measuring dependence of typical values. The combined information from these two dependence measures improves diagnostic performance for extremal dependence, as ρ provides a context in which to assess estimated values of extremal dependence.

2.2. Measures of extremal dependence

There exist several summary statistics for extremal dependence, see Coles et al. (1999), Beirlant et al. (2004) and Wadsworth and Tawn (2013). Here we focus on four measures which identify different characteristics of the joint tail for asymptotically independent and asymptotically dependent distributions.

In Section 1 we introduced χ, with $0 \leq \chi \leq 1$, as a measure of extremal dependence for variables which are identically distributed. This measure holds more generally when X and Y have marginal distribution functions F_X and F_Y, and associated inverses F_X^{-1} and F_Y^{-1}, respectively, with

$$\chi = \lim_{q \to 1} \Pr\{Y > F_Y^{-1}(q) | X > F_X^{-1}(q)\} = \lim_{q \to 1} \Pr\{F_Y(Y) > q | F_X(X) > q\}. \tag{3}$$

The interpretation of χ is as in Section 1. From an oceanographic context this corresponds to the limit of the probability that one variable exceeds its T year return level given that the other variable exceeds its T year return level as $T \to \infty$, thus there is a greater risk for metocean designs the larger the value of χ.

All measures of extremal dependence can be specified on general and different marginal forms, like for χ, but it is often more simple mathematically to present them for identical marginal distributions; compare expressions (2) and (3). The concept of copulas (Nelsen, 2006), shows there is no loss of generality in presenting extremal measures for a specific marginal distribution choice. Specifically when presenting the measure for variables $\tilde{X}$ and $\tilde{Y}$ with common marginal distribution choice F, then we need to transform (X,Y) to $(\tilde{X}, \tilde{Y})$, by

$$\tilde{X} = F^{-1}\{F_X(X)\} \quad \text{and} \quad \tilde{Y} = F^{-1}\{F_Y(Y)\}. \tag{4}$$

To apply this transformation in practice, we estimate F_X and F_Y using the empirical distribution functions. Alternatively, we could replace the empirical distributions in the upper tail with a generalised Pareto distribution (Coles and Tawn, 1991). Various measures are more simple to present mathematically for a specific marginal choice, which then dictate the choice of common distribution function F. Specifically in this paper we use Pareto margins, $F(x) = 1 - x^{-1}$ for $x \geq 1$ with $(\tilde{X}, \tilde{Y})$ then denoted by (X_P, Y_P), Fréchet margins, $F(x) = \exp(-x^{-1})$ for $x > 0$ with $(\tilde{X}, \tilde{Y})$ then denoted by (X_F, Y_F), and Laplace margins, $F(x) = \exp(x)/2$ for $x \leq 0$ and $1 - \exp(-x)/2$ for $x > 0$ with $(\tilde{X}, \tilde{Y})$ then denoted by (X_L, Y_L). Note that $\Pr\{X_F > x\} = 1 - \exp(-x^{-1}) = x^{-1} + x^{-2} + O(x^{-3}) \approx x^{-1} = \Pr\{X_P > x\}$ for large x. Hence, the marginal tails of the Pareto and Fréchet distributions are approximately equivalent and so these two margins can be used inter-changeably when focusing on the upper tail.

The measure χ has limitations: all asymptotically independent variables have $\chi = 0$; it does not provide information about the dependence at levels of interest, e.g., at the 50 year return level; and it is hard to estimate as it involves an extrapolation. In particular, there could be stronger dependence at the 50 year level for asymptotically independent variables than for asymptotically dependent variables. Therefore we need a more refined measure than χ. The second extremal dependence measure $\chi(x)$ is presented for Pareto marginal variables, with

$$\chi(x) = \Pr\{Y_P > x | X_P > x\} \quad \text{for } x \geq 1 \tag{5}$$

and $\chi = \lim_{x \to \infty} \chi(x)$. On this scale x corresponds to the return period on the marginal scale, so $x = 50$ is the one in every 50

observations return level. It follows from Section 1 that for positively dependent variables $x^{-1} \leq \chi(x) \leq 1$, the bigger the value of $\chi(x)$, for large x, the stronger the level of extremal dependence.

A complication with $\chi(x)$ is that it is a function over $x > 1$ and so not parsimonious. Ideally what is required is a formulation for how $\chi(x)$ behaves for large x. Ledford and Tawn (1996) prove that under weak conditions

$$\Pr\{X_P > x, Y_P > x\} = \mathcal{L}(x)x^{-1/\eta} \quad \text{for } x \geq 1, \tag{6}$$

where $\eta \in (0, 1]$ is termed the coefficient of the tail dependence, and $\mathcal{L}$ is any positive slowly varying function at infinity, i.e.,

$$\lim_{t \to \infty} \mathcal{L}(tx)/\mathcal{L}(t) = 1, \tag{7}$$

provided the derivative with respect to x of the right hand side of expression (6) is non-negative for all $x \geq 1$. This formulation gives that

$$\chi(x) = \mathcal{L}(x)x^{-(1-\eta)/\eta} \quad \text{for } x \geq 1. \tag{8}$$

From expression (8) the relevance of η is apparent. When the variables (X_P, Y_P) are perfectly independent (perfectly dependent) expression (1) shows that $\eta = \frac{1}{2}$ and $\mathcal{L}(x) = 1$ ($\eta = 1$ and $\mathcal{L}(x) = 1$), respectively. More generally, when $\eta = 1$ and $\mathcal{L}(x) \to c$ with $0 < c \leq 1$, it follows that $\chi = c$ and the variables are asymptotically dependent. Furthermore, $\chi(x)$ converges to c slower than any power of x for large x, and hence, for practical cases, can be viewed as constant for large x, i.e., $\chi(x) = c$ for $x > u$ for some large u. When $0 < \eta < 1$ then $\chi(x) \to 0$ as $x \to \infty$, so $\chi = 0$ and the variables are asymptotically independent. However, the value of η provides additional information about the level of asymptotic independence as η describes the rate of convergence to 0, and therefore is key in determining the difference between $\chi(x)$ and χ. Specifically $\frac{1}{2} < \eta < 1$ corresponds to positive extremal dependence and $0 < \eta < \frac{1}{2}$ to negative extremal dependence. An example when $0 < \eta < 1$ occurs is for the Gaussian dependence structure, with $\eta = (1 + \rho_P)/2$, where ρ_P is the Pearson correlation coefficient, so $0 < \rho_P < 1$ implies $\frac{1}{2} < \eta < 1$, i.e., all non-perfectly dependent Gaussian dependence structure variables are asymptotically independent.

Although the characterisation of Ledford and Tawn (1996) provides a more concrete formulation for the structure of asymptotic independence, it is restricted to situations where both variables are large. A more general formulation, based on slightly different assumptions, has been proposed by Heffernan and Tawn (2004) and found to have substantial uses in metocean applications (see e.g., Jonathan et al., 2013 and Eastoe et al., 2013). The formulation is most apparent in Laplace margins, i.e., we have (X_L, Y_L), with the connection for large values of these variables with the Pareto variable case being that $X_L \approx \log(X_P/2)$ and $Y_L \approx \log(Y_P/2)$. Under weak assumptions, a combination of Heffernan and Tawn (2004) and Heffernan and Resnick (2007) show that there exist values α and β, with $-1 \leq \alpha \leq 1$, $0 \leq \beta < 1$, for $x > 0$ and $-\infty < z < \infty$, such that

$$\lim_{u \to \infty} \Pr\left\{ \frac{Y_L - \alpha X_L}{X_L^\beta} < z, X_L - u > x | X_L > u \right\} = \exp(-x)G(z), \tag{9}$$

where G is a non-degenerate distribution function. It follows that α, β and G jointly determine the form of extremal dependence. Here, if $\alpha = 1$ and $\beta = 0$ the variables are asymptotically dependent with $\chi = \int_0^\infty [1 - G(-t)] \exp(-t)\, dt > 0$, but otherwise $\chi = 0$. Specifically, the form of asymptotic independence is given by

$$\chi(x) = \Pr\{Y_L > x_L | X_L > x_L\}$$
$$= \int_0^\infty \left[1 - G\left(\frac{x_L - \alpha(x_L + t)}{(x_L + t)^\beta} \right) \right] \exp(-t)\, dt, \tag{10}$$

where $x_L = \log(x/2)$. For statistical purposes, Heffernan and Tawn (2004) assume that limit (9) holds exactly for a large finite u. As a consequence, conditionally on $X_L > u$, $X_L - u$ and $(Y_L - \alpha X_L)/X_L^\beta$ are independent random variables, and conditionally on X_L, with $X_L > u$, we have

$$Y_L = \alpha X_L + X_L^\beta Z, \tag{11}$$

where Z has mean μ and variance σ^2, with distribution function G. It follows that $E[Y_L|X_L] = \alpha X_L + X_L^\beta \mu$ and $\text{Var}(Y_L|X_L) = X_L^{2\beta}\sigma^2$ for $X_L > u$. Then maximum likelihood methods can be used, under a working assumption of normality for G, to estimate $\hat{\alpha}$, $\hat{\beta}$, $\hat{\mu}$ and $\hat{\sigma}$. Rearranging (11) gives

$$\hat{Z} = \frac{Y_L - \hat{\alpha}X_L}{X_L^{\hat{\beta}}}, \quad \text{for } X_L > u,$$

and using the empirical distribution of such $\hat{Z}$'s gives an estimate of the distribution G, i.e., we estimate the distribution of Z using all pairs (X_L, Y_L) with $X_L > u$.

2.3. Extremal dependence models

2.3.1. Introduction

We will consider three classes of spatial processes: max-stable process, Gaussian processes and inverted max-stable processes. Max-stable processes can be asymptotically dependent or perfectly independent only. In contrast, Gaussian processes and inverted max-stable processes are asymptotically independent always. The max-stable family is the most studied asymptotically dependent model, but it has no finite parametrisation. To illustrate some of the max-stable process features we will present a range of parametric sub-models that are widely used in applications. For each of the three classes of processes we will present the dependence measures χ, $\chi(x)$ for large x, η and α and β to reveal different features of their dependence structures. These processes are also used in Section 4 in a simulation study for assessing the performance of the estimation methods of Section 3.

2.3.2. Max-stable processes

Max-stable processes arise as the class of marginally non-degenerate limit distributions for componentwise maxima of independent and identically distributed replicates of a spatial process. For such processes all finite dimensional distributions are multivariate extreme value distributions (Beirlant et al., 2004, Ch. 8). Max-stable fields are typically presented via the following constructive approach (see e.g., de Haan, 1984; Schlather, 2002; de Haan and Ferreira, 2006, Ch. 9).

For a max-stable process on the space S, let $\{r_i\}$, $i = 1, ...,$ be the points of a Poisson process on $(0, \infty)$, with intensity dr, and define the spectral function to be $\{W(s) \geq 0; s \in S\}$, a continuous random function on S, satisfying $E\{W(s)\} = 1$ for all $s \in S$. Then if the processes $W_i(\cdot)$ for $i = 1, ...$ are independent and identically distributed copies of $W(\cdot)$, then

$$X_F(s) = \max_i W_i(s)/r_i \quad \text{for } s \in S$$

is a max-stable random field with unit Fréchet margins and d-dimensional distribution function

$$\Pr\{X_F(s_1) \leq x_1, ..., X_F(s_d) \leq x_d\} = \exp\left(E\left[\max\left\{ \frac{W(s_1)}{x_1}, ..., \frac{W(s_d)}{x_d} \right\} \right] \right)$$

for $x_i > 0$, $i = 1, ..., d$.

To motivate this representation from a practical point of view, we can view the ith event having magnitude $1/r_i$, and event profile/shape over space S of $W_i(\cdot)$, so the ith event for the process is $W_i(s)/r_i$ for $s \in S$. Thus, $X_F(s)$ is the componentwise maxima of these events at location $s \in S$. The dependence structure of the max-stable process is determined by the form of the exponent function, V,

$$V(\mathbf{x}) = E\left[\max\left\{\frac{W(s_1)}{x_1}, ..., \frac{W(s_d)}{x_d}\right\}\right],$$

with $\mathbf{x} = (x_1, ..., x_d)$ for all $\mathbf{x}$ and all d. Here $1 \le V(\mathbf{x}) \le d$, with the lower and upper bounds achieved when the variables are perfectly dependent and perfectly independent respectively. Key to the properties of max-stable processes is homogeneity of order -1 of V, i.e., $V(n\mathbf{x}) = n^{-1}V(\mathbf{x})$ for all n and $\mathbf{x}$. Thus

$$\Pr\{X_F(s)/n \le x(s), s \in S\}^n = \Pr\{X_F(s) \le x(s), s \in S\},$$

i.e., $X_F(\cdot)$ and the maximum of n independent copies of $X_F(\cdot)/n$ have the same distribution.

All max-stable processes are pairwise asymptotically dependent or perfectly independent. Dropping the index notation for sites, we have the joint survivor function of $(X_F, Y_F) = (X_F(s_1), X_F(s_2))$ as

$$\Pr\{X_F > x, Y_F > x\} = 1 - 2\exp(-1/x) + \exp\{-V(x, x)\}$$
$$= 1 - 2\exp(-1/x) + \exp\{-V(1, 1)/x\}$$
$$= \frac{2 - V(1, 1)}{x} + \left[\frac{V(1, 1)^2}{2} - 1\right]\Big/x^2 + O(x^{-3}),$$
$$\text{as } x \to \infty. \tag{12}$$

It follows that

$$\chi(x) = 2 - V(1, 1) + \left[\frac{V(1, 1)^2}{2} - 1\right]\Big/x + O(x^{-2}), \quad \text{as } x \to \infty.$$

Hence, $\chi = 2 - V(1, 1) > 0$ if $V(1, 1) < 2$, i.e., the variables are asymptotically dependent when they are not perfectly independent, and $\chi = 0$ when $V(1, 1) = 2$, i.e., the variables are only asymptotically independent when they are perfectly independent. For the bivariate extreme value distribution with $V(1, 1) < 2$, $\eta = 1$, $\alpha = 1$ and $\beta = 0$, with $G(z) = -V_1(1, e^z)$, where $V_1(\cdot, \cdot)$ is the first derivative of $V(\cdot, \cdot)$ (see Heffernan and Tawn, 2004, Section 8.4).

Typically, it is difficult to derive V explicitly, and this is the limiting factor in finding tractable max-stable models. There exists a slowly growing set of models for which the bivariate distributions are available; see, e.g., Smith (1990), Schlather (2002), Kabluchko et al. (2009), Davison et al. (2012) and Wadsworth and Tawn (2012). We list the details of the most widely used below:

Smith process: Smith (1990) takes $W_i(s) = \phi_d(s - t_i, \Sigma)$, where $d = \dim(S)$, $\phi_d(\cdot, \Sigma)$ is the d-dimensional multivariate normal density function where the vector mean is $\mathbf{0}$ and variance is Σ, and where t_i have measure dt over $\mathbb{R}^d$. For this model the pairwise exponent measure for locations o, the origin, and h is

$$V_h(x, y) = x^{-1}\Phi\left\{\frac{\psi(h)}{2} + \frac{1}{\psi(h)}\log\left(\frac{y}{x}\right)\right\}$$
$$+ y^{-1}\Phi\left\{\frac{\psi(h)}{2} + \frac{1}{\psi(h)}\log\left(\frac{x}{y}\right)\right\}, \tag{13}$$

where $x > 0$ and $y > 0$, Φ is the standard normal distribution function, and $\psi(h) = \sqrt{h^T\Sigma^{-1}h}$ is the Mahalanobis distance between h and o. It follows that the level of asymptotic dependence at separation h is $\chi_h = 2[1 - \Phi\{\psi(h)/2\}]$ for which $\chi_h \to 0$ as $h \to \infty$, and $\chi_h \to 1$ as $h \to 0$, spanning the range of possible asymptotic

dependencies. At all finite separations of the locations we have asymptotic dependence.

Schlather process: Schlather (2002) took $W_i(s) = \max\{0, Z_i(s)\}$ where the $Z_i(s)$ are independent replicates of a stationary process $Z(s)$ on $\mathbb{R}^d$, satisfying $E[\max\{0, Z(o)\}] = 1$, where o denotes the origin. When $\sqrt{2\pi}Z(s)$ is a stationary Gaussian process with mean zero, variance 2π and correlation function $\rho(h)$, then the pairwise exponent measure for locations o, the origin, and h is

$$V_h(x, y) = \frac{1}{2}\left(\frac{1}{x} + \frac{1}{y}\right) \times \left(1 + \left[1 - 2\frac{(\rho(h) + 1)xy}{(x + y)^2}\right]^{1/2}\right),$$

where $x > 0$ and $y > 0$. Here $\chi_h = 1 - \{[1 - \rho(h)]/2\}^{1/2}$. When $d = 2$, because $\rho(h)$ is a positive definite function, $\chi_h \ge 0.162$. This means that the model cannot account for processes with low asymptotic dependence at any separation.

Brown–Resnick process: Brown and Resnick (1977) let $W(s) = \exp\{\epsilon(s) - \gamma(s)\}$, where $\epsilon(s)$ is a fractional Brownian process with semivariogram $\gamma(h) = (h/c)^\alpha$, $0 < \alpha \le 2$, $c > 0$, $h > 0$, and $\epsilon(0) = 0$ almost surely. Here $\chi_h = 2[1 - \Phi\{\{2\gamma(h)\}^{1/2}/2\}]$. A special case is the Smith model when $\alpha = 2$.

Extremal-t process: Demarta and McNeil (2005) assume that

$$W(s) = \sqrt{\pi}2^{-\nu/2+1}[\Gamma\{(\nu + 1)/2\}]^{-1}\max\{\epsilon(s), 0\}^\nu,$$

where $\nu \ge 1$, Γ is the Gamma function, and $\epsilon(x)$ is a stationary Gaussian random field with mean zero and correlation function $\rho(h)$. The case when $\nu = 1$ corresponds to the Schlather process. The level of asymptotic dependence between locations o and h is $\chi_h = 2[1 - T_{\nu+1}(\sqrt{(\nu + 1)[1 - \rho(h)]/[1 + \rho(h)]})]$, where T_ν denotes the cumulative distribution function of a student-t random variable with ν degrees of freedom.

2.3.3. Gaussian processes

Gaussian processes are the simplest and most well-known asymptotically independent processes. Let $\{Z(s), s \in S\}$ be a stationary Gaussian process with $E\{Z(s)\} = 0$, $\text{Var}\{Z(s)\} = 1$ and $\text{Corr}\{Z(o), Z(h)\} = \rho(h)$. Define

$$X_F(s) = -1/\log\Phi(Z(s)),$$

then $X_F(s)$ has Fréchet margins and a Gaussian process copula. For this process, $\chi_h = 0$ for all $h \ne o$, and $\eta_h = \{1 + \rho(h)\}/2$, where η_h is the measure in (6) for the process at the pair of sites o and h. For large x we have

$$\chi_h(x) = \mathcal{L}_h(x)/x^{1/\eta_h},$$

with $\mathcal{L}_h$ a slowly varying function. Furthermore, in the Heffernan and Tawn formulation, the dependence parameters (α, β) for the process at the pair of sites o and h are

$$\alpha_h = (\rho(h))^2\text{sign}(\rho(h)) \quad \text{and} \quad \beta_h = 1/2.$$

The limiting conditional distribution G_h (corresponding to G in Eq. (9)) is Gaussian with variance $2\rho(h)^2[1 - \rho(h)^2]$. This shows that simply using a Gaussian copula for asymptotically independent cases is not a sufficiently general approach as both β_h and G_h have restrictive forms relative to the more general Heffernan and Tawn class.

2.3.4. Inverted max-stable processes

Wadsworth and Tawn (2012) introduce an alternative class of asymptotically independent processes – the inverted max-stable processes. Essentially, these processes have the same copula as max-stable processes, but with the copula inverted, so that lower tail dependence translates to upper tail dependence, and vice versa. Formally, if $X_{MS}(\cdot)$ and $X_{IMS}(\cdot)$ are a max-stable process and an inverted max-stable process, respectively, both with Fréchet

margins, then

$$X_{IMS}(s) = -1/\log[1 - \exp\{-1/X_{MS}(s)\}], \quad \text{for } s \in S. \tag{14}$$

This process has $\chi_h = 0$ for all $h \neq o$, and

$$\chi_h(x) = x^{-1/V_h(1,1)} + O(x^{-1-1/V_h(1,1)}) \quad \text{for large } x,$$

so $\eta_h = 1/V_h(1, 1)$, where V_h is the exponent measure for $X_{MS}(o)$ and $X_{MS}(h)$. The form of the conditional extremes parameters (α_h, β_h) is complicated in general, see Papastathopoulos and Tawn (2015). We focus in particular on the example where $V(x, y)$ corresponds to the logistic model (Tawn, 1988)

$$V(x, y) = (x^{-1/\gamma} + y^{-1/\gamma})^\gamma$$

where $x > 0$, $y > 0$ and $\gamma \in (0, 1]$, with $\gamma \to 1$ implying independence and $\gamma \to 0$ implying perfect dependence. Then the joint distribution function for the inverted bivariate logistic extreme value distribution is

$$F(x, y) = \exp(-1/x) + \exp(-1/y) - 1$$
$$+ \exp\left\{ -\left[\left(\frac{-1}{\log(1 - e^{-1/x})} \right)^{-1/\gamma} + \left(\frac{-1}{\log(1 - e^{-1/y})} \right)^{-1/\gamma} \right]^\gamma \right\}. \tag{15}$$

For this distribution $\eta = 2^{-\gamma}$, $\alpha = 0$ and $\beta = 1 - \gamma$.

3. Estimation

In this section we introduce several methods for estimating the dependence measures described in Section 2. The performance of these methods will be explored in Section 4. Our particular focus is on the estimation of η and $\chi(x)$; the former as it gives a single number summary indicating how close the joint tail is to asymptotic dependence, and the latter as it provides a clear picture of the implications of the nature of the extremal dependence on actual joint occurrences. We base our estimation of η on methods of Ledford and Tawn (1996), and propose new models and constraints. Eqs. (8) and (10) present different formulations for $\chi(x)$; we present inference based on both of these.

3.1. Estimating η

To estimate the tail dependence given in representation (6) it is useful to define $T = \min(X_P, Y_P)$ and note that

$$\Pr\{T > x\} = \Pr\{\min(X_P, Y_P) > x\} = \Pr\{X_P > x, Y_P > x\}$$
$$= \mathcal{L}(x)x^{-1/\eta} \quad \text{for } x \geq 1. \tag{16}$$

Thus, $0 < \eta \leq 1$ is the tail index of the variable T. This means that, even though dependence is a bivariate feature, univariate techniques can be used to estimate η simply by using data on T constructed from (X_P, Y_P). To estimate η using the tail form (16), we need to make some modelling assumptions about the positive slowly varying function $\mathcal{L}(x)$ above a high threshold u, in addition to the property (7) and expression (16) representing a valid survivor function. In previous studies the only constraint imposed when parametric models have been specified for $\mathcal{L}(x)$ for $x \geq u$ is that $0 < \Pr\{T > u\} \leq 1$, so $0 < \mathcal{L}(u) \leq u^{1/\eta}$. However, from conditional probability (8) we have the stronger condition $0 < \mathcal{L}(u) \leq u^{(1-\eta)/\eta}$. For the tail of T to have a density $f_T(x) > 0$ for $x > u$, we also require that

$$\mathcal{L}(x) > \eta x \mathcal{L}'(x) \quad \text{for all } x \geq u.$$

Possible choices for $\mathcal{L}(x)$, for $x > u$, that have been suggested in the past are:

$$\mathcal{L}_1(x) = c \tag{17}$$

$$\mathcal{L}_2(x) = c + dx^{1/\eta - 2} \tag{18}$$

$$\mathcal{L}_3(x) = \lambda \left[x \left(1 + \frac{\eta(x - u)}{\sigma} \right) \right]^{-1/\eta}. \tag{19}$$

Model (17) leads to the Hill estimator for η and model (19) corresponds to a generalised Pareto tail with the shape $\eta > 0$, scale $\sigma > 0$ and threshold exceedance rate $0 < \lambda < 1$. In contrast, the two-term model (18) attempts to capture a natural second order decay rate with the tail $cx^{-1/\eta} + dx^{-2}$, so for this to be the second term, we additionally require $\eta > 1/2$, corresponding to positive extremal dependence. Our constraints impose that $0 \leq c \leq u^{(1-\eta)/\eta}$ for model (17), and $cu^{-(1-\eta)/\eta} + du^{-1} \leq 1$ and if $d < 0$ that $c > 2\eta|d|u^{-2+1/\eta}$ for model (18). Additionally, we imposed that η must be in the range $(0, 1]$.

The models given above can be fitted using a censored likelihood approach. We take the likelihood to be

$$L(\theta) = \left\{ \prod_{i=1}^{n_u} f_T(t_i) \right\} \Pr\{ T \leq u \}^{n - n_u}, \tag{20}$$

where θ is the vector parameter for the tail of T, $T = \min(X_P, Y_P)$, n_u is the number of points above the threshold u, and $t_1, \ldots, t_{n_u}$ is an enumeration of the values of $T > u$. We use constrained maximum likelihood to impose our restrictions on the parameter space when estimating $\theta = (c, \eta)$, $\theta = (c, d, \eta)$ or $\theta = (\lambda, \sigma, \eta)$ for models (17), (18) and (19), respectively. For model (17) analytical estimates can be obtained

$$\hat{\theta} = (\hat{c}, \hat{\eta}) = \left(\frac{n_u}{n} u^{1/\hat{\eta}}, \ \min\left\{ \frac{1}{n_u} \sum_{i=1}^{n_u} \log\left(\frac{t_i}{u} \right), 1 \right\} \right).$$

The other two models require constrained numerical maximisation of the likelihood to obtain parameter estimates as no tractable solutions exist.

Simply using the likelihood function does not exploit knowledge about the processes involved, e.g., they are positively dependent in the extremes, nor our caution in the inference about being conservative from the outset, i.e., expecting the extremal dependence structure to be more like asymptotic dependent than asymptotic independent prior to seeing data. To incorporate this information it is natural to adopt a Bayesian framework to the inference. Although much work (see e.g., Lee, 2012) has been undertaken on selecting priors to represent our lack of knowledge (i.e., non-informative priors), less work exists for guiding the incorporation of informative prior knowledge. A key development in the latter has been through the approach of Simpson et al. (2014). They introduce the penalised complexity (PC) prior as a way of constructing prior distributions for situations where there is a natural nested structure in the model components; i.e., there is a more flexible model that is an extension of a base model. PC priors penalise the complexity induced by deviating from the simpler base model. For our context, a natural base model is asymptotic dependence and the more flexible model covers both asymptotic independence and dependence parametrised by η. Focusing on exceedances of the threshold u by T and adopting model (17) for $\mathcal{L}(x)$, the base model is

$$g(x) = \frac{u}{x^2}, \quad x > u,$$

and the more flexible model is

$$f(x) = \frac{u^{1/\eta}}{\eta x^{1/\eta+1}}, \quad x > u, \; 0 < \eta \le 1, \; u > 0.$$

So, $f(x) = g(x)$ when $\eta = 1$. The penalised complexity prior for η (derived in Appendix A) is then

$$\pi(\eta; \lambda) = \frac{\lambda(1 - \eta)\exp(-\lambda\sqrt{-2\log\eta + 2(\eta - 1)})}{\eta\sqrt{-2\log\eta + 2(\eta - 1)}}, \quad \text{for } 0 < \eta \le 1,$$

where the parameter $\lambda > 0$ controls the shape of the prior. The value of λ determines the strength of the prior belief that a model close to an asymptotically dependent model is appropriate to the data. In the studies in this paper we chose λ such that $\pi(\eta > 0.5; \lambda) = 0.99$, which essentially ensures positive extremal dependence. From this it follows that $\lambda = 7.409$. We use this prior with the likelihood given in (20) to obtain a maximum a posteriori (MAP) estimate of η. Although the prior is only presented for model (17), a similar prior form is given in Kereszturi (2016) for model (18).

3.2. Estimating $\chi(x)$

Though it is possible to estimate $\chi(x)$, defined by probability (5), empirically, this restricts estimates to x in the range of the data. To estimate beyond the data we need to use a model; here we use asymptotic forms (8) and (10).

Using the Ledford and Tawn model (8), with estimates $\hat{\eta}$ and $\hat{\mathcal{L}}(x)$ for each of the models (17)–(19), substituting the maximum likelihood estimates of the parameters, gives

$$\hat{\chi}_i(x) \approx \frac{\hat{\mathcal{L}}_i(x)}{x^{1/\hat{\eta}-1}}, \quad \text{for } x \ge u, \; i = 1, 2, 3. \tag{21}$$

Now consider inference for $\chi(x)$ using the conditional extremes model (9), with $\chi(x)$ given by the integral expression (10). The simplest way to evaluate this integral is through Monte Carlo methods using estimated values $\hat{\alpha}$, $\hat{\beta}$ and $\hat{G}$ of α, β and G, as shown in the algorithm below.

Algorithm 1. Algorithm for Monte Carlo evaluation of $\chi(x)$ using the conditional extremes model.

1 Simulate X_L above $x_L = \log(x/2)$ as $X_L = x_L + v$, where v is a realisation of an Exponential(1) random variable;

2 Simulate Z from $\hat{G}$ independently of X_L;

3 Set $Y_L = \hat{\alpha}X_L + (X_L)^{\hat{\beta}}Z$;

4 Repeat steps 1–3 m-times;

5 Set $\hat{\chi}_{HT}(x)$ as the proportion of the m simulated pairs (X_L, Y_L) with $Y_L > x_L$.

4. Simulation study

Here we compare different methods for the estimation of η and $\chi(x)$ presented in Section 3 for samples from asymptotically dependent and asymptotically independent models given in Section 2.3. We further investigate how sample size and threshold choice affect estimation.

4.1. Design

4.1.1. Simulation processes

We simulate samples from the max-stable Smith, Schlather, Brown–Resnick and extremal-t processes, the Gaussian process, and from the inverted logistic bivariate extreme value distribution.

Simulating from a max-stable process is relatively straightforward; for details, see Schlather (2002). In this paper we used the *SpatialExtremes* package in *R* (Ribatet, 2008), which gives samples on Fréchet margins that can then be transformed to the desired margins using the probability integral transform. We simulated from the processes over 10 equally spaced locations on a straight line. For the Smith process, we set $\Sigma = \{\sigma_{ij}\}$ with $\sigma_{11} = \sigma_{22} = 2$ and $\sigma_{12} = \sigma_{21} = 0$, in order to give a good range of Spearman's ρ on (0,1) between pairs of locations. For similar reasons, we simulated the Schlather process with correlation function $\rho_p(h) = \exp(-h/4)$ and the Brown–Resnick process with semivariogram $\gamma(h) = \exp(-h/0.8)$. For the extremal-t simulation, the correlation function used was $\rho_p(h) = \exp(-h/3)$ with $\nu = 5$ degrees of freedom.

The Gaussian process samples were simulated over 15 equally spaced locations on a line, with a distance of one unit between the closest sites. The process has correlation function $\rho_p(h) = e^{-\lambda h}$, where h is the distance between two locations, and λ was chosen such that $\rho_p(1) = 0.6$. Hence, $\rho_p(mh) = \rho_p(h)^m$, for $m = 1, \ldots, 15$, leading to a process at the observed sites being a first order Markov process. We generated the samples using the *mvtnorm* package in *R* (Genz et al., 2014). The chosen combination of number of sites and first order dependence parameter λ ensures a good coverage of the range (0,1) for Spearman's ρ.
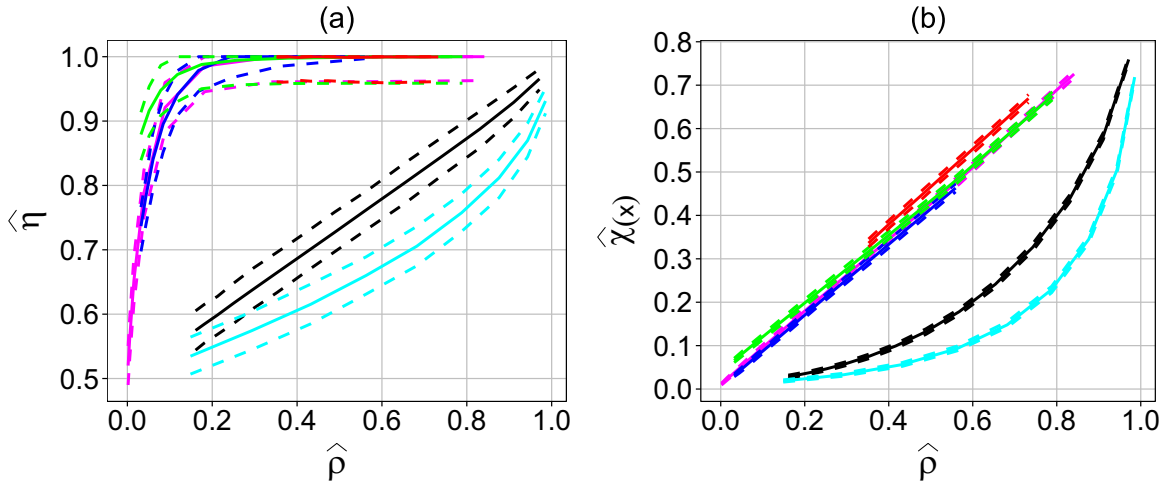
Samples from the inverted logistic distribution were obtained by simulating from the bivariate extreme value distribution using the *evd* package in *R* (Stephenson, 2002), and then inverting the lower tail of this using (14). We simulated over 10 equally spaced locations on a line, and the dependence parameter γ was chosen such that a good range of Spearman's ρ is obtained over (0,1).

4.1.2. General behaviour of η and $\chi(x)$ with ρ

First we explore how values of η and $\chi(x)$ link to Spearman's ρ for the different processes we have introduced above. Here we evaluate each of these properties using estimates obtained from replicates of very large samples.

Fig. 1 illustrates the relationship between median estimates for η and $\chi(x)$ and values for ρ from 1000 realisations of samples of size 10^6 from each of the processes, with η estimated using model (19) with threshold probability $q = \Pr\{T \le u\} = 0.99$, and $\chi(x)$ estimated empirically using Eq. (5) with $x = 100$. The point-wise (2.5%, 97.5%) uncertainty band is also shown. Here $\hat{\rho}$ denotes the median estimate of ρ over the 1000 sample realisations. The width of the 95% confidence interval for $\hat{\rho}$ is at most 0.01 for this large sample size, so the ρ estimates are very precise. The relationship between $\hat{\eta}$ and $\hat{\rho}$ is very similar for all asymptotically dependent max-stable processes, as is the relationship between $\hat{\chi}(x)$ and $\hat{\rho}$. The behaviour of the asymptotically independent Gaussian process and inverted logistic distribution is clearly different for both $\hat{\eta}$ with $\hat{\rho}$ and $\hat{\chi}(x)$ with $\hat{\rho}$. However, for asymptotic dependence, the known limiting value of unity for η is only achieved (approximately) for max-stable processes with $\hat{\rho}$ exceeding 0.2. Similarly, for asymptotic independence, the known limiting value of zero for $\chi(x)$ is never achieved for either the Gaussian process or inverted logistic distribution for our choices of x. This suggests that relying purely on sample estimates for η and $\chi(x)$ approximating limiting values, even for large samples, is not a useful diagnostic for extremal dependence. However, the obvious differences in behaviour between estimates of η with ρ, and $\chi(x)$ with ρ suggest that distinction between asymptotic dependence and asymptotic independence is possible.

We confirmed by simulation that the functional relationships between η and ρ, and $\chi(x)$ and ρ, are not strongly dependent on the values of model parameters for a specific model. The results we present in the remainder of this section are therefore effectively independent of the model parameter specification.

Fig. 1. Estimates for (a) η and (b) $\chi(x)$ against estimates for Spearman's ρ for sample size $n = 10^6$, from the Smith (magenta), Schlather (red), Brown–Resnick (blue), extremal-t (green) and Gaussian (black) processes, and the inverted logistic distribution (cyan). Estimation methods use model (19) for η with $q=0.99$, and the empirical estimate for $\chi(x)$ with $x=100$. Solid lines are median estimates from 1000 sample replications, dashed lines give 2.5% and 97.5% quantiles. (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this paper.)

4.1.3. Estimators, sample sizes and thresholds

We investigate the performance of the set of estimators for η and $\chi(x)$ for samples from different processes, with different sample sizes and threshold choices. Since the characteristics of all max-stable models (for given sample size and threshold) are found to be very similar, we choose here to illustrate their behaviour using results for the Smith process only; corresponding results for other max-stable processes are given in Kereszturi (2016).

In Sections 4.2 and 4.3, we examine the performance of different estimators for η and $\chi(x)$, respectively, for samples from the Smith max-stable process and the Gaussian process. In Section 4.4, we assess the effect of sample size on estimates for η and $\chi(x)$, for samples from the Smith and Gaussian processes, and the inverted logistic distribution. We also consider in Section 4.4 the effect of the threshold probability q and level x on estimates for η and $\chi(x)$, respectively, for these samples.

We consider three sample sizes, motivated by the North Sea application discussed in Section 5: a large sample with $n = 10^6$ observations (probably unrealistically large for typical metocean applications, but useful to assess large sample performance), and two more realistic sample sizes of $n=58,585$ (medium size, equivalent to the size of the sea-state significant wave height sample), and $n=916$ (small size, equivalent to the size of the storm-peak significant wave height sample). For each sample size, we adjust the threshold probability q and the level x to achieve acceptable numbers of observations for estimation of η and $\chi(x)$, respectively. For the large sample, high threshold probabilities and levels are chosen, whereas for the more realistically-sized medium and small samples we are forced to choose lower threshold probabilities.

The key feature of figures shown in this section is the degree of agreement between sample estimates for η and $\chi(x)$ and their known limiting behaviour given the underlying process (or distribution) used to simulate the sample. For asymptotically dependent samples, we expect estimates for η close to unity, and estimates for $\chi(x)$ bounded away from zero even for large x. For asymptotically independent samples, we expect estimates for η less than unity, and estimates for $\chi(x)$ approaching zero for large x. We expect that agreement between sample estimate and limiting behaviour further depends on the overall extent of dependence in the full sample, hence we plot these estimated extremal dependence features against the Spearman's rank correlation ρ.

4.2. Comparison of estimation methods for η

We denote the estimators of η given by the censored likelihood approach (20) for the models (17)–(19) by $\hat{\eta}_1$, $\hat{\eta}_2$ and $\hat{\eta}_3$ respectively. We also estimated models (17) and (18) with the Bayesian penalised complexity (PC) prior described in Section 3.1. However, we found that the estimates were effectively unchanged unless very strong priors were used. Hence, we only show these results for the smallest sample size, where the prior is not dominated by the sample likelihood. See Kereszturi (2016) for the full results.

We compare estimators for the following combinations (n,q) of sample size n and threshold probability q: $(10^6, 0.99)$, $(58, 585, 0.90)$ and $(916, 0.90)$. For each sample size, 1000 replicate samples are generated from each of the Smith and Gaussian processes to allow estimation of 2.5% and 97.5% quantiles of the sampling distribution of the estimator of η. Since ρ is relatively precise to estimate, with the width of the 95% confidence interval around $\hat{\rho}$ at most 0.01 for the large sample and at most 0.1 for the small sample, we use the median of $\hat{\rho}$ over the 1000 sample replications. Fig. 2 shows estimates for η using the different methods, represented by lines of different colours, for the large sample from the Smith (see Fig. 2(a)) and Gaussian processes (see Fig. 2(b)), plotted against Spearman's ρ.

The Gaussian process is asymptotically independent, with $\eta = (1 + \rho_P)/2 < 1$. The estimates in Fig. 2(b) are consistent with this. There is no material difference between the means of the sampling distributions for the different methods for estimating η here. However, note that the variance of the sampling distributions is slightly larger for model (18). This is likely due to the additional term in the tail model.

The Smith model is asymptotically dependent over all distances, so the true value of η is unity for all pairs. Whilst there is some difference between the different estimation methods, all estimators underestimate η for asymptotically dependent samples when dependence is low and the sample size is relatively small (see Fig. 2(a)). All η estimators are close to unity for large $\hat{\rho}$, but $\hat{\eta}_2$ and $\hat{\eta}_3$ seem to perform better than $\hat{\eta}_1$ for pairs with lower $\hat{\rho}$, as these better capture the higher order features of the tail decay.

The following argument helps understand why η estimates start to perform poorly for the Smith process with weak dependence. For the Smith process, using expression (12), we know that
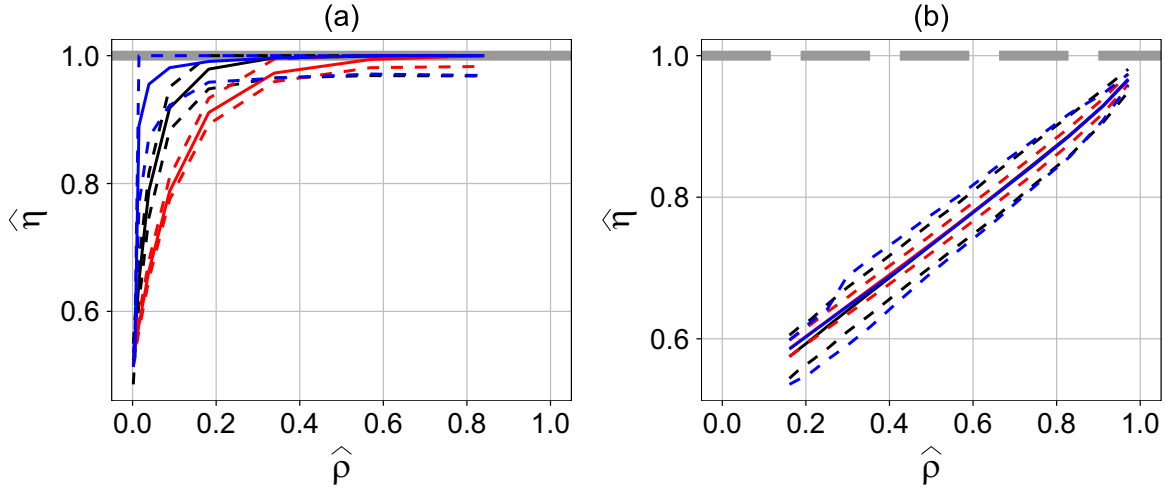
Fig. 2. Estimates for η against Spearman's ρ for sample size $n = 10^6$ with $q=0.99$, from (a) Smith process and (b) Gaussian process. Estimation methods are model (17) (red), model (18) (blue), and model (19) (black). Solid lines give the median from 1000 sample replications, dashed lines give 2.5% and 97.5% quantiles. The solid grey line shows a correct limiting value for η. The dashed grey line shows a limiting value that η should not take. (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this paper.)

$$\bar{F}_{T,h}(x) = \Pr(\min(X(s), X(s+h)) > x) = \Pr(X(s) > x, X(s+h) > x)$$

$$= \frac{(2 - 2\Phi\{a(h)/2\})}{x} + \frac{2\Phi^2\{a(h)/2\} - 1}{x^2} + O(x^{-3}),$$

for large x, \hfill (22)

where Φ is the standard normal distribution function and $a(h)$ is the Mahalanobis distance between h and the origin o. For two locations with Spearman's ρ approximately equal to 0.1, the upper tail in Eq. (22) becomes $\bar{F}_{T,h}(x) \approx cx^{-1} + dx^{-2} + O(x^{-3})$, with $c=0.004$ and $d=0.991$. In order for the first term to dominate we would need $cx^{-1} \gg dx^{-2}$, but this can only be achieved if x is very large. Even to just have $cx^{-1} > dx^{-2}$, for $x > u$, we need $u \approx 250$ on the Fréchet scale, which means a threshold quantile of approximately 99.995% for T. Hence, for practical sample sizes and threshold choices, the estimator $\hat{\eta}_1$ will be biased for pairs with low Spearman's ρ.

Figs. 3 and 4 present analogous estimates of η for the medium and small sample sizes, respectively. Note that the smaller the sample size, the η estimates for the Smith model have more bias for pairs with low dependence, and show more uncertainty (see Figs. 3(a) and 4(a)). However, when the sample size is small, the prior is more important and its inclusion gives better estimates, drawing $\hat{\eta}$ towards unity (see Fig. 4(a)). For the Gaussian process the estimates are consistent with the true η value, but here too

there is more uncertainty in the estimates due to the small sample sizes. The increased uncertainty means that unity lies within the confidence bounds for pairs with $\rho > 0.6$ (see Fig. 4(b)), so in these cases asymptotic dependence might be selected incorrectly.

4.3. Comparison of estimation methods for $\chi(x)$

For samples of size $n=58,585$ from the Smith and Gaussian processes, we compare different methods for estimating $\chi(x)$ outlined in Section 3.2. Specifically, model-based estimates $\hat{\chi}_1(x)$, $\hat{\chi}_3(x)$ and $\hat{\chi}_{HT}(x)$ are compared with empirical estimate $\hat{\chi}_{EMP}(x)$. We estimate $\hat{\chi}_{EMP}(x)$ using the conditional probability given in (5), by simply counting the points in the region $\{Y_P > x, X_P > x\}$ and dividing by the number of points in the region $\{Y_P > x\}$. Estimates for $\hat{\chi}_2(x)$ are omitted in figures for clarity due to their close similarity to those for $\hat{\chi}_1(x)$. For the model-based estimates $\hat{\chi}_1(x)$ and $\hat{\chi}_3(x)$ a modelling threshold u was chosen such that $\Pr\{T \le u\} = 0.90$. For the third model-based estimate $\hat{\chi}_{HT}(x)$ we estimated the model parameters α, β, μ, and σ above a 90th percentile threshold (see Section 2.2). We estimate $\chi(x)$ at the level $x=10,000$. This level was chosen, since it is a very high quantile of the sample, with only 5 data points above this level, so the empirical estimate is expected to be poor. Hence, we can better
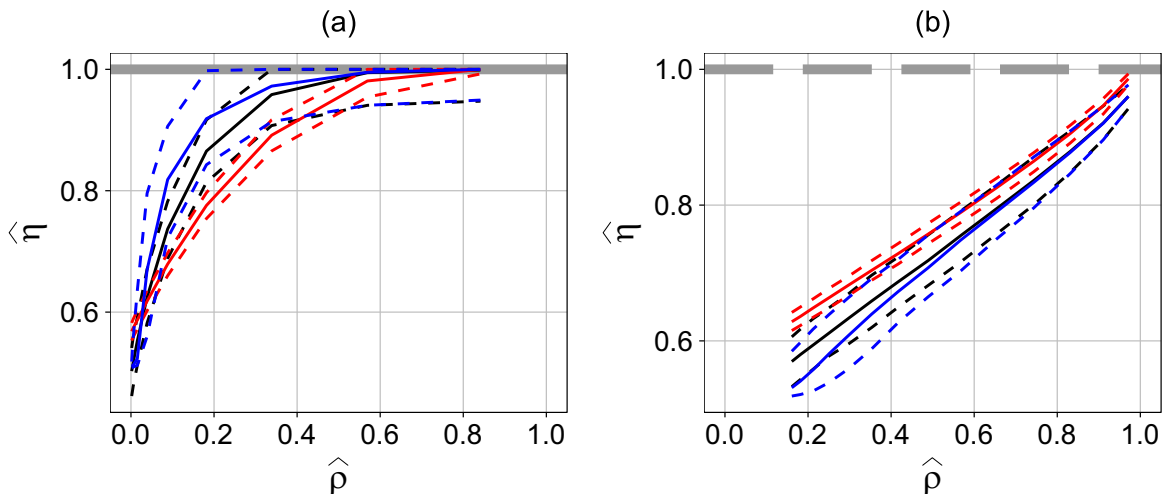
Fig. 3. Estimates for η against Spearman's ρ for sample size $n=58,585$ with $q=0.90$, from (a) Smith process and (b) Gaussian process. Lines are as described in Fig. 2.

(a)   (b)
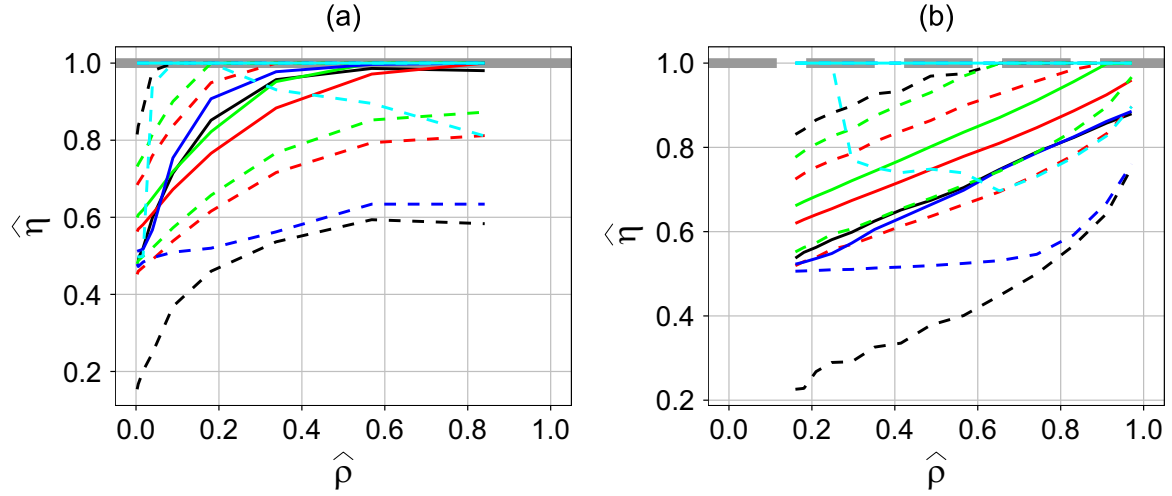
Fig. 4. Estimates for η against Spearman's ρ for sample size $n=916$ with $q=0.90$ from (a) Smith process and (b) Gaussian process. Lines are as described in Fig. 2. Additionally, estimation methods using model (17) with PC prior (green) and model (18) with PC prior (cyan) are shown. (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this paper.)

compare the model-based estimates. We have also compared the estimators at the level $x=100$, which is within the data, but at this level all the estimators were performing very similarly, so these results are not shown here.

The different estimates are shown against estimates of Spearman's ρ in Fig. 5. The empirical estimate is not reliable, due to the small number of points above this level. The model-based estimates perform similarly for the Gaussian sample (see Fig. 5(b)) However, for the Smith model, $\hat{\chi}_3(x)$ gives higher mean values and wider uncertainty bands than the other model-based estimators. This might be due to the tail model (19) that this estimator is based on. Note the different patterns of relationship between $\hat{\chi}(x)$ and $\hat{\rho}$ for the two processes, as already noted in Fig. 1; the asymptotically dependent Smith process shows higher levels of tail dependence than the asymptotically independent Gaussian process across all levels of Spearman's $\hat{\rho}$. At the high level $x=10,000$, $\hat{\chi}(x)$ approaches 0 for the Gaussian process sample, which is the limiting value as $x \to \infty$.

In summary, it appears that the different estimation methods lead to fairly similar estimates for both processes, when x is set to

a level that is within the data. For this reason, in Section 4.4, we adopt the empirical estimate only due to ease of computation. We note however that model-based estimates offer the additional benefit of extrapolation to rare levels (beyond the sample) when necessary.

4.4. Effect of sample size and threshold effect

Here we assess the effect of sample size, threshold probability and level x on selected estimates for η and $\chi(x)$ for samples from the Smith and Gaussian processes, and the inverted logistic distribution. The estimator $\hat{\eta}_3$ as given by model (19) is used for η estimation, and the empirical estimate $\hat{\chi}_{EMP}(x)$ for $\chi(x)$. For estimating η, the following pairs (n,q) of sample size n and threshold probability q are considered: $(10^6, 0.99)$, $(10^6, 0.999)$, $(10^6, 0.9999)$, $(58, 585, 0.80)$, $(58, 585, 0.90)$ and $(58, 585, 0.99)$. For estimating $\chi(x)$, the level x was set so that the same number of observations would be used for estimation as for the corresponding η estimation. This gives the following pairs (n,x) of sample size n and level x: $(10^6, 100)$, $(10^6, 1000)$, $(10^6, 10000)$, $(58, 585, 5)$, $(58, 585, 10)$ and $(58, 585, 100)$.

(a)   (b)
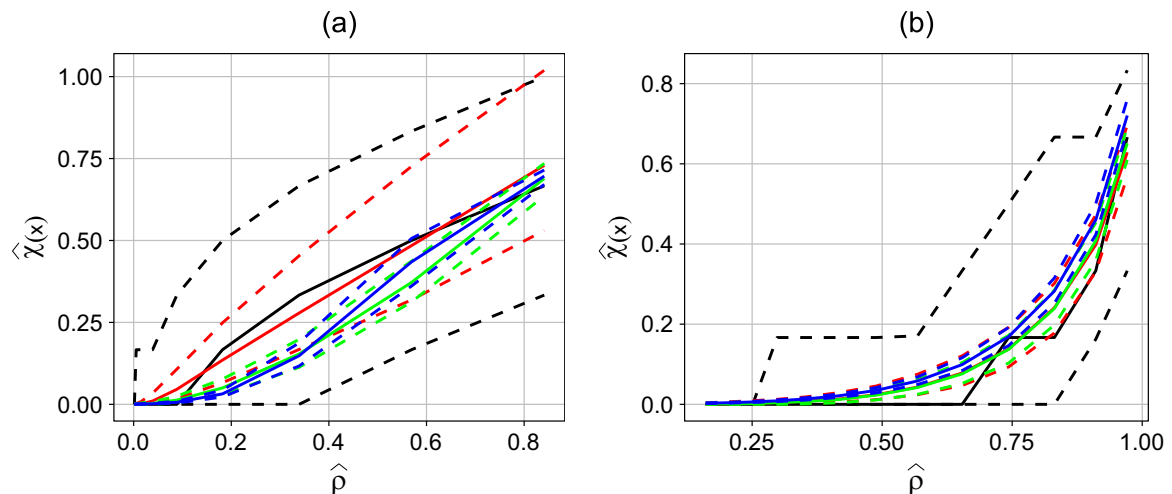
Fig. 5. Estimates for $\chi(x)$ against Spearman's ρ for sample size $n=58,585$ with $x=10,000$ from (a) Smith process and (b) Gaussian process. Estimation methods are empirical $\chi_{EMP}(x)$ (black), $\chi_1(x)$ (blue), $\chi_3(x)$ (red) and $\chi_{HT}(x)$ (green). Solid lines give the median from 1000 sample replications, dashed lines give 2.5% and 97.5% quantiles. (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this paper.)
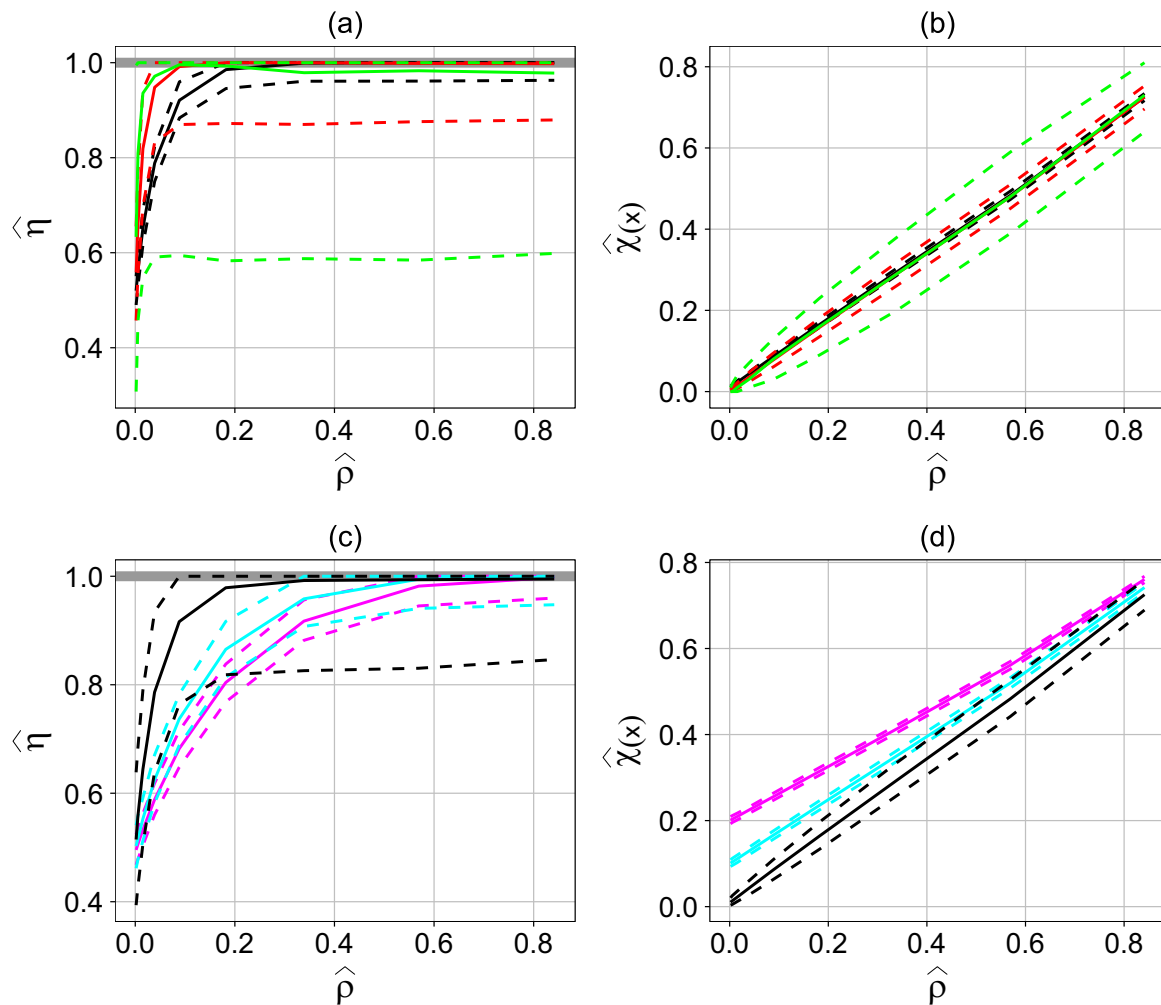
Fig. 6. Estimates of η and $\chi(x)$ plotted against Spearman's ρ for simulated data from the Smith model. Sample size is $n = 10^6$ (top row) and $n=58,585$ (bottom row). Threshold probabilities for the η estimation are $q=0.9999$ (green line), $q=0.999$ (red line), $q=0.99$ (black line), $q=0.90$ (cyan line) and $q=0.80$ (magenta line). Corresponding levels for the $\chi(x)$ estimation are $x=10,000$ (green line), $x=1000$ (red line), $x=100$ (black line), $x=10$ (cyan line) and $x=5$ (magenta line). Solid lines give the median from a 1000 sample replications, dashed lines give the 2.5% and 97.5% quantiles. The solid grey line shows a correct limiting value for η. (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this paper.)

For asymptotically dependent samples, we expect η estimates close to unity, and $\chi(x)$ estimates bounded away from zero for all x. Results for the Smith process clearly show that $\hat{\chi}(x) > 0$ even for large x and small ρ (see Fig. 6(b) and (d)). Further, $\eta \approx 1$ for pairs with high dependence in the body (as measured by Spearman's ρ), however, η is underestimated for pairs with weak dependence in the body, particularly for smaller sample sizes and lower thresholds. Fig. 6(a) shows that the η estimate is good for our large sample, but this sample size and threshold probability are not typically achievable in practice. For realistic sample sizes and threshold probabilities, η is underestimated for low levels of dependence (see Fig. 6(c)). This is the case for other asymptotically dependent models as well. We conducted the same analysis for the other max-stable processes described in Section 2.3 with similar results (see Kereszturi, 2016).

For asymptotically independent samples, the limiting value of η is less than 1 for all pairs, and so $\chi = 0$. For both the Gaussian and the inverted logistic samples, $\hat{\eta} < 1$ for all pairs (see Figs. 7(a,c) and 8(a,c)). However, $\hat{\chi}(x)$ is approximately 0 only for pairs with low dependence in the body of the data for all levels of x considered here. At high levels x, estimates approach the limiting value (see Figs. 7(b,d) and 8(b,d)). Face-value interpretation of $\chi(x)$ estimates suggests asymptotic dependence; and we know this is not the case. This shows the critical importance η plays in distinguishing

asymptotic independence from asymptotic dependence.

We note also that $\hat{\eta}$ and $\hat{\chi}(x)$ appear to decrease faster with decreasing $\hat{\rho}$ for the inverted logistic sample than for the Gaussian sample (compare Figs. 7 and 8). This behaviour is not affected by the distributional parameters chosen for sample simulation from these processes; these only affect the range of ρ, not the relationship between ρ and η and $\chi(x)$. These features therefore provide a valuable tool for model diagnostics for asymptotically independent processes, and will be used in Section 5.

5. Application

5.1. Data

Characterisation of marginal and spatial dependence for extremes of significant wave height H_S is essential to reliable design and operation of marine and coastal installations. Sea-state H_S is defined as four times the standard deviation of the ocean surface elevation in a given time period. The data examined here consists of time-series for H_S for three hour sea-states (from the hindcast of Reistad et al., 2011) for the period from September 1957 to December 2012 on a grid of 150 northern North Sea locations, covering an area of approximately 5° longitude by 5° latitude; see
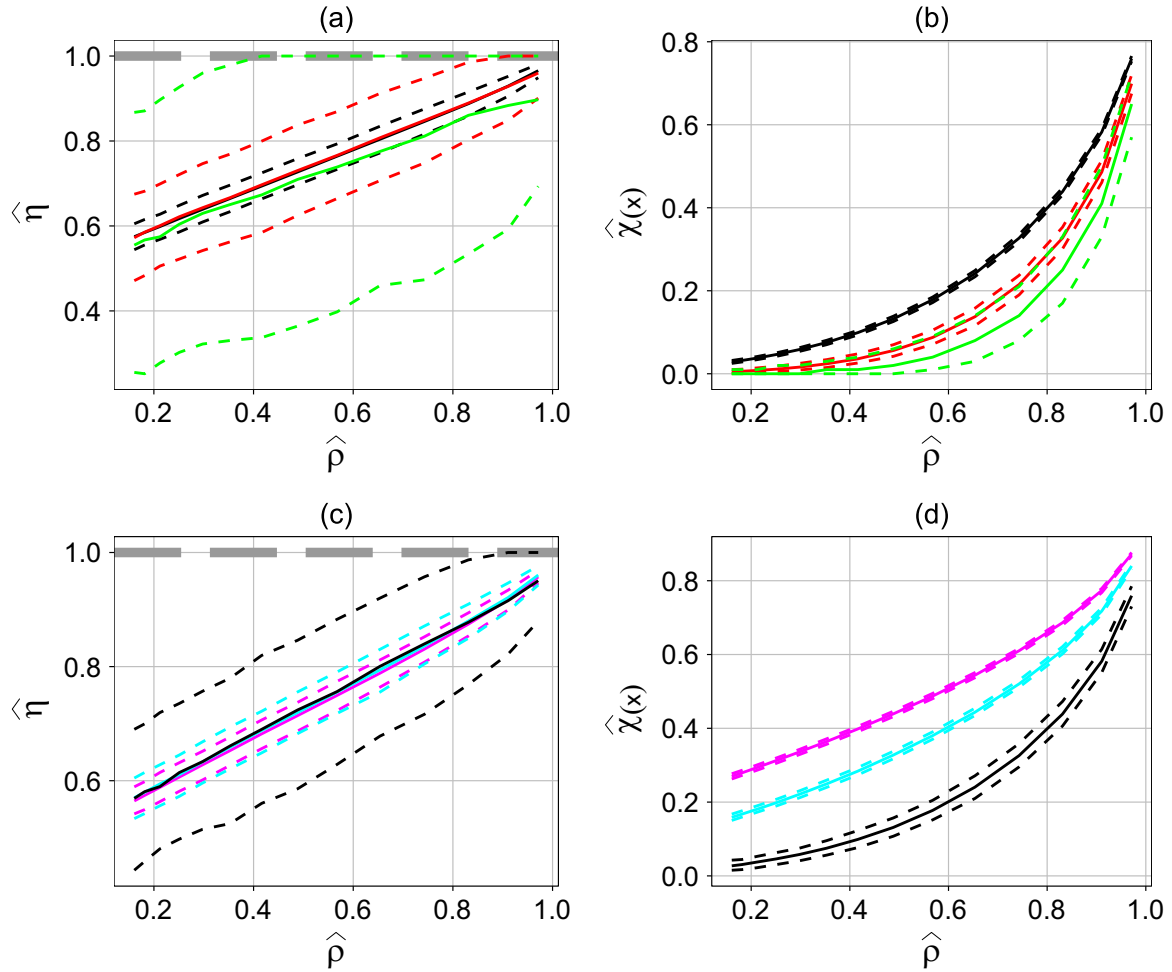
Fig. 7. Estimates of η and $\chi(x)$ plotted against Spearman's ρ for simulated data from the Gaussian process. Sample size is $n = 10^6$ (top row) and $n = 58,585$ (bottom row). Lines are as described in Fig. 6. The dashed grey line shows a limiting value for η that the estimates should not take.

Fig. 9. The univariate extreme value characteristics of H_S from the hindcast have been examined by Aarnes et al. (2012) and Breivik et al. (2013). We analyse this sample of sea-state H_S and a further sample of 916 independent storm-peak H_S values extracted from the sea-state data using the procedure of Ewans and Jonathan (2014): briefly, contiguous intervals of sea-state H_S above a low peak-picking threshold are identified, each interval corresponding to a storm event. The maximum of H_S during the interval is taken as the storm-peak H_S for the storm. Consecutive values of sea-state H_S at a location typically exhibit temporal dependence, whereas consecutive storm-peak values do not.

The distribution of sea-state H_S varies with wave direction θ due to the combination of prevailing wind field and land shadow effects. Fig. 10 shows (θ, H_S) for one location. Note that θ represents the direction from which waves emanate, measured clockwise from North. The density estimates for $\theta | H_S > v$ for $v = 0, 4, 8$ m and 10 m are also shown. It can be seen that as the threshold v increases, the density is increasingly concentrated in the [320, 10] and [120, 200] directional sectors. This suggests that for this particular location the biggest waves come from these directions, corresponding to storms emanating from the North North-West and South South-East approximately, respectively, roughly the orientation of the blue strip in Fig. 9. We therefore might expect directional effects to influence spatial dependence, and that ignoring θ in assessing spatial dependence might not be justified. Focusing on strips of locations with different orientations helps us assess and illustrate directional influences on spatial dependence in a more straightforward way. Hence, we proceed to examine the

extremal dependence between pairs of locations on the four straight-line strips, shown in Fig. 9. The blue strip has approximately North-South orientation with long along-strip fetches in both directions. The red and black strips are both aligned North-West–South-East. The westerly end of the red strip is in the land shadow of the Shetland islands, but the westerly end of the black strip is located between the Shetland and Orkney islands, so has a long fetch from the North-West Atlantic Ocean. The green strip has a South-West–North-East orientation with along-strip fetch limited by the land masses of the Britain Isles and Norway.

5.2. Identification of extremal dependence

Fig. 11 illustrates that the dependence between locations on each of the strips varies with inter-location separation. In Fig. 11 (a), plots of convex hulls for sea-state H_S from pairs of locations on the black strip corresponding to inter-location separations of 1, 5 and 10 are shown (or separations of 0.5°, 2.3° and 4.5° respectively). Fig. 11(a), and corresponding figures for other strips (not shown), suggests that dependence reduces with increasing separation as expected. Because of the orientation of the location grid used for the hindcast, inter-location separation along the red and black strips is approximately the same; however, inter-location separations along the green and blue strips are different. Fig. 11(b) shows convex hulls of sea-state H_S for pairs from all strips with inter-location separation of approximately 2.4°, achieved by taking an inter-location separation of 5 for the red and black strips, 6 for the blue and both 6 (approximately 2.2°) and 7
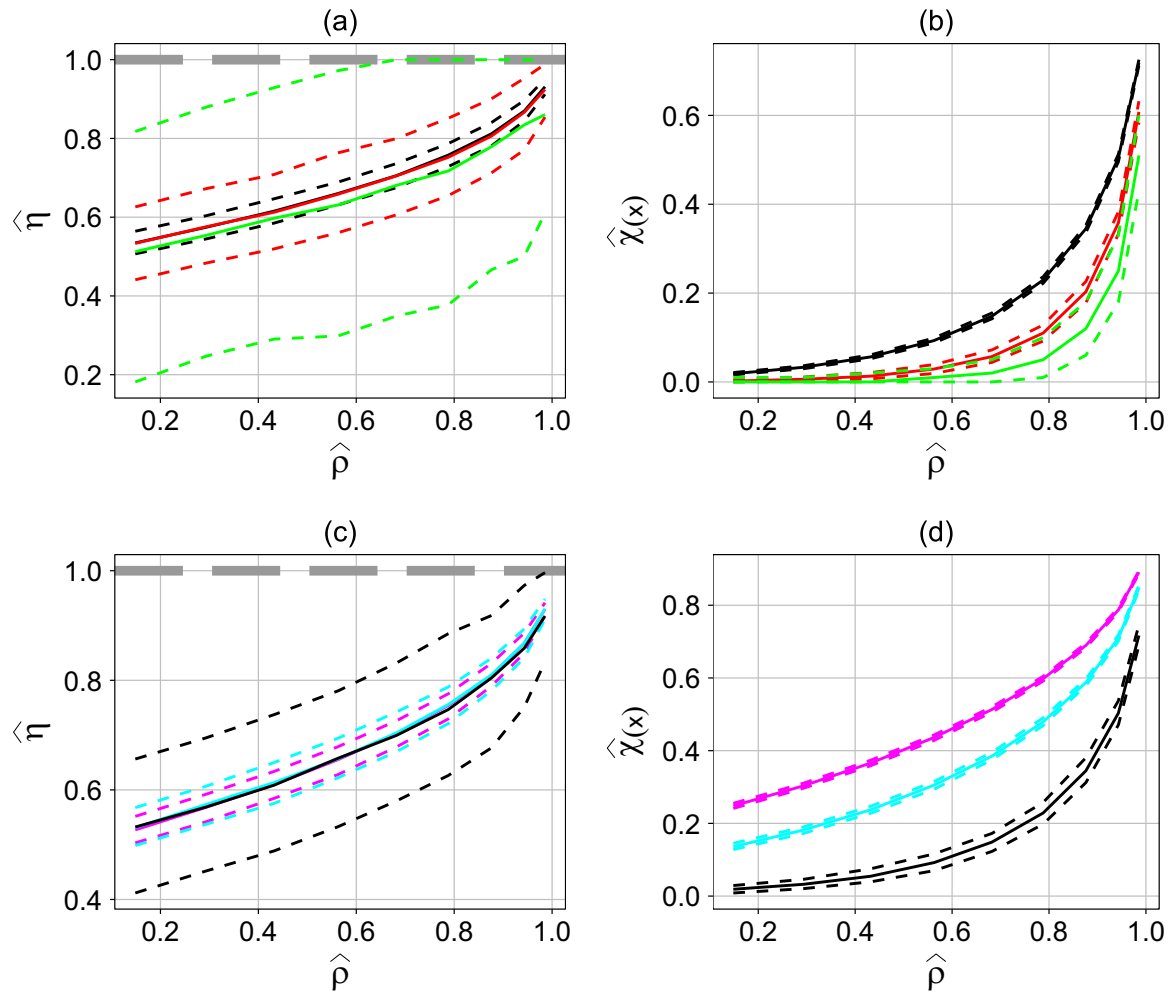
Fig. 8. Estimates of η and $\chi(x)$ plotted against Spearman's ρ for simulated data from the inverted logistic model. Sample size is $n = 10^6$ (top row) and $n=58{,}585$ (bottom row). Lines are as described in Fig. 6. The dashed grey line shows a value for η that the estimates should not take.

(approximately 2.6°) for the green strip. From Fig. 11(b) it is difficult to identify differences in dependence between strips for pairs at a separation of approximately 2.4°; however there is some suggestion that the dependence between sea-state H_S for pairs of locations on the red and black strips is similar. Blue and green strips also appear to be relatively similar. The corresponding plots for the storm-peak sample are given in Fig. 11(c, d), with very similar characteristics.

We next estimate ρ, η and $\chi(x)$ for sea-state and storm-peak H_S for pairs of locations on each of the four strips. Each of η, $\chi(x)$ and ρ were estimated for sea-state H_S at pairs of locations on the four strips shown in Fig. 9. We used the estimator $\hat{\eta}_3$ for η estimation, and $\hat{\chi}_{EMP}(x)$ for $\chi(x)$ estimation, as in Section 4.4. The estimates are shown in Fig. 12. Equivalent estimates for simulated samples of the same size from the Smith and Gaussian processes and the inverted logistic distribution are also shown. This enables us to compare the patterns of relationship between η, $\chi(x)$ and ρ for the different asymptotically dependent and asymptotically independent processes that we observed in Fig. 1 with the pattern the data exhibits. It appears that behaviour on the blue (North–South) and green (South-West–North-East) strips is reasonable well represented by the Gaussian model, whereas the dependence characteristics on black and red strips (North-West–South-East) are better captured by the inverted logistic model. Both the functional relationships between η and ρ and $\chi(x)$ and ρ support this fit. This is particularly clear for higher threshold probabilities q and levels x. Both these best-matching distributions are asymptotically

independent. This clearly shows how our novel approach of supplementing measures of extremal dependence with ρ improves diagnostic performance.

We further estimate η, $\chi(x)$ and ρ for pairs of values of storm-peak H_S from locations on each of the four strips. Fig. 13 illustrates these estimates together with the corresponding estimates from simulated samples of the same size from the Smith and Gaussian processes and the inverted logistic distribution. It appears that storm-peak H_S dependence is consistent with the asymptotically independent Gaussian process for all strips, but due to small sample size there is a large overlap between the estimated uncertainty bands for different models. Inferring the nature of extremal dependence from samples of this size is extremely difficult.

5.3. Design implications

The choice of extremal dependence structure of storm-peak significant wave height values at neighbouring locations has material implications for structural design and reliability assessment. To illustrate the type of difference, consider the storm-peak significant wave heights (X,Y) at two of our North Sea locations separated along a strip by approximately 2.6°. If the risk of joint extremes at the two locations is of interest, then we would want to find the distribution of Y given that X was equal to its 100 year return level, denoted by x_{100}, i.e. we want to find y_p, the pth conditional quantile, such that

$P(Y < y | X = x_{100}) = p$.

For simplicity of calculation we make the assumption that the marginal distribution of X and Y is identical to a typical site in our region shown in Fig. 9. This is a reasonable assumption, as over this spatial scale, away from land shadows, the marginal distributions do not change rapidly. As mentioned in Section 1, a Naïve approach is to assume perfect dependence, and hence $y_p = x_{100}$ for all p. Under strong asymptotic dependence similar values for y_p arise. The question is what reduction can we get in y_p from using our best fitting model for extremal dependence, i.e., an asymptotically independent distribution, which we take here as the Gaussian dependence structure. For the storm-peak significant wave heights we take a threshold u and assume that the marginal

Fig. 9. Map of location of data, showing four colour-coded sets of locations lying on straight lines with particular orientations, referred to in the text as "strips". (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this paper.)

conditional distribution of threshold exceedances is described by a generalised Pareto distribution

$$F(x) = P(X < x | X > u) = P(Y < x | Y > u) = 1 - \left(1 + \xi \frac{x - u}{\sigma}\right)_+^{-1/\xi},$$

for $x > u$,

where ξ and $\sigma > 0$ are the shape and scale parameters respectively and $t_+ = \max(t, 0)$. Threshold u is selected using parameter stability plots (Coles, 2001). Here we choose $P\{X > u\} = \phi = 0.5$, and estimate the parameters ξ and σ using maximum likelihood. For our data, $u = 6.66$ m and the estimates are $(\hat{\xi}, \hat{\sigma}) = (-0.23, 2.48)$. Then the expected number of storm peaks above some value x in 100 years is
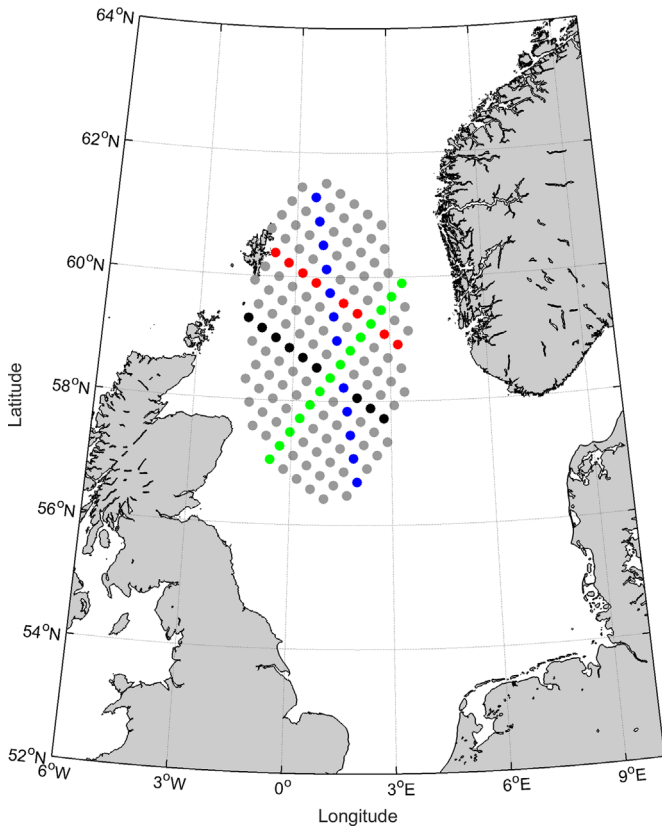
$$100 \times n_{py} \times \phi \times \left[1 + \hat{\xi} \frac{x - u}{\hat{\sigma}}\right]^{-1/\hat{\xi}}, \tag{23}$$

where n_{py} is the average number of storm peaks per year ($n_{py} = 30.11$ for our data). Setting expression (23) to 1 when $x = x_{100}$, and rearranging, we find the marginal 100-year return value x_{100} for location X to be the marginal quantile of storm peaks with non-exceedance probability $F(x_{100}) = 1 - (100 n_{py} \phi)^{-1}$, corresponding to a value $x_{100} \approx 15.5$ m. Thus under perfect dependence $y_p = 15.5$ m.

The conditional extremes model of Heffernan and Tawn (2004) is most straightforward to apply on Laplace margins. To transform to Laplace margins we use expression (4) which transforms quantiles of F to the same quantile of the Laplace distribution F_L. Hence the transformed 100-year level on Laplace scale is x_{100}^L where $F(x_{100}) = F_L(x_{100}^L)$, giving $x_{100}^L = -\log[2(100 n_{py} \phi)^{-1}]$. Now suppose further that the extremal dependence between the locations corresponds approximately to the asymptotically independent Gaussian process with correlation ρ_P, which is consistent with our findings of Section 5.2. We had found in our analysis that two locations at separation of approximately 2.6° correspond to a Spearman's ρ value of approximately 0.8 (see Fig. 13), which, in turn, corresponds to a Gaussian process with Pearson correlation coefficient of $\rho_P = 0.8$. The conditional extremes model of Heffernan and Tawn then suggests for Gaussian processes, with Laplace margins (X_L, Y_L), that

$$P\left\{\frac{Y_L - \rho_P^2 X_L}{X_L^{1/2}} < z \middle| X_L = x_{100}^L\right\} \approx \Phi\left(\frac{z}{\rho_P(1 - \rho_P^2)^{1/2}}\right); \tag{24}$$
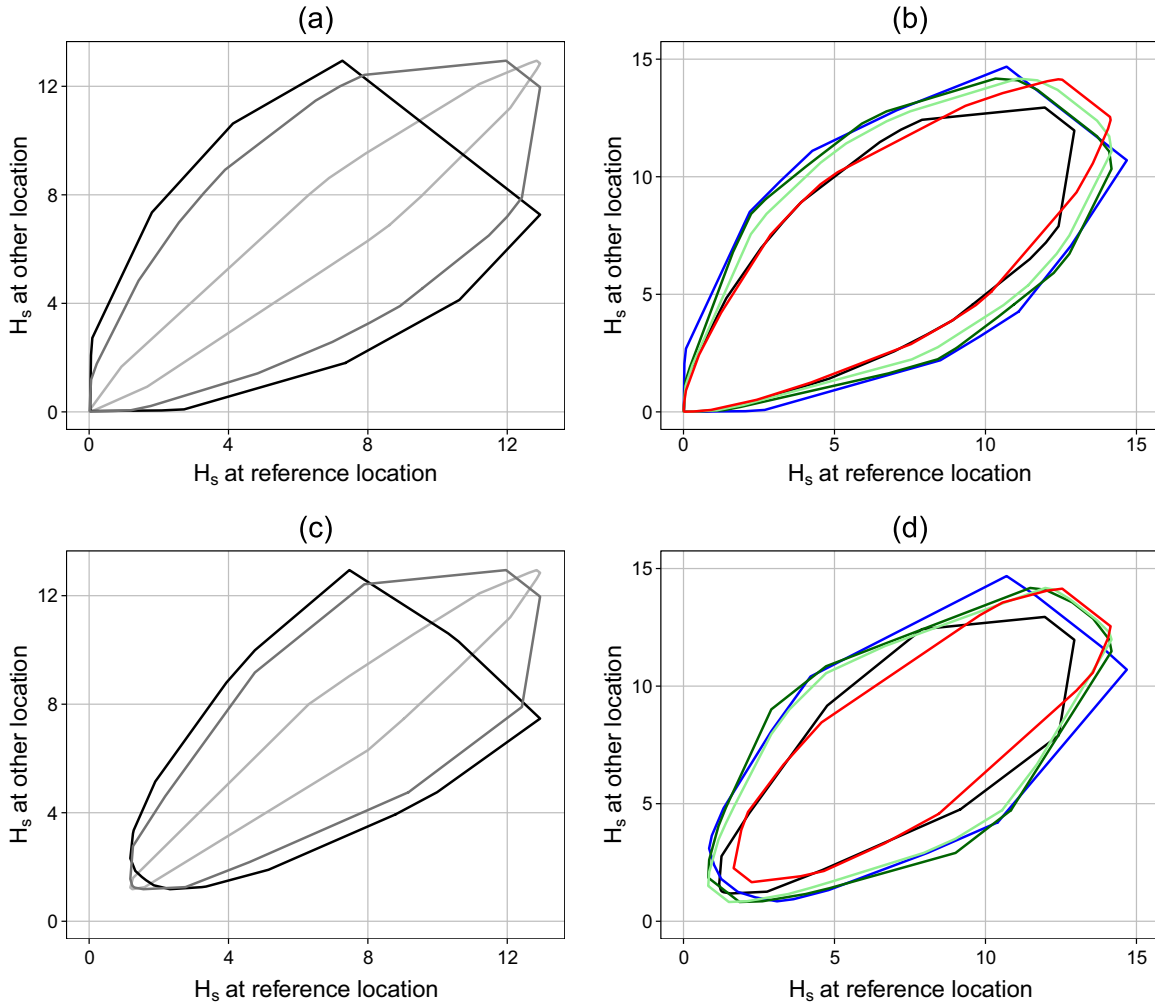
see Section 8 of Heffernan and Tawn (2004) for more details. If we assume that approximation (24) holds for x_{100}^L then the pth quantile of the distribution of $Y_L | (X_L = x_{100}^L)$ is
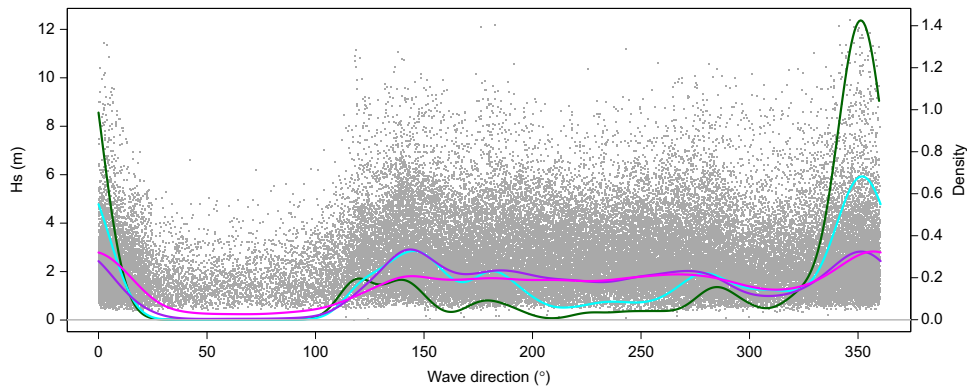
Fig. 10. Sea-state H_S against wave direction θ for a central location, and corresponding density estimates for $\theta | H_S > v$ for $v = 0$ m (magenta), $v = 4$ m (purple), $v = 8$ m (cyan) and $v = 10$ m (dark green). (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this paper.)

Fig. 11. Convex hulls for pairs of sea-state significant wave heights H_S illustrating dependence. (a) Convex hulls for locations from the black strip (see Fig. 9) are plotted, corresponding to inter-location separation of 0.5° (light grey), 2.3° (grey) and 4.5° (black); (b) convex hulls for locations with inter-location separation of approximately 2.4° are plotted, for all strips, coloured accordingly. Since there are no pairs of locations on the green strip corresponding to separation of 2.4°, those with separation 2.2° (light green) and 2.6° (dark green) are shown. The corresponding convex hulls for the storm-peak sample are shown in plots (c) and (d). (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this paper.)

$$y_p^L = \rho_P^2 x_{100}^L + \rho_P (1 - \rho_P^2)^{1/2} (x_{100}^L)^{1/2} \Phi^{-1}(p), \qquad (25)$$

where Φ is the distribution function of a standard normal variable. Here y_p^L is on Laplace scale, and converting back to the original scale, quantile to quantile transformation suggests that y_p satisfies $F(y_p) = F_L(y_p^L)$. This means that the associated quantile of storm-peak significant wave height at the neighbouring location will be

$$y_p = u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left\{ \frac{1}{2\phi} \exp(-y_p^L) \right\}^{-\hat{\xi}} - 1 \right),$$

where y_p^L is given by expression (25). For our data the y_p values are plotted against p (see Fig. 14). For almost all values of p, $y_p \lessapprox 15.5$ m, so the estimated risk is much less than under perfect dependence; for example, the median is 2 m lower. Only for ρ_P very close to 1 is there essentially no practical difference between the two approaches.

6. Discussion and conclusions

The extremal dependence between contemporaneous observations of significant wave height at neighbouring locations determines the rate and size of occurrences of joint extremes at those locations. In this work we examined the extremal dependence of significant wave height for pairs of locations from a North Sea hindcast. We showed that the characteristics of extremal dependence vary systematically with distance between locations and with orientation of locations, given prevailing atmospheric conditions and fetch variation. Using the variation of coefficient of tail dependence η and the $\chi(x)$ statistic with (full-sample) Spearman rank correlation ρ to quantify extremal dependence, we find that a sample of sea-state significant wave height exhibits asymptotic independence consistent with either the Gaussian process or inverted logistic models. The nature of extremal dependence for a sample of storm-peak significant wave height is very difficult to estimate, due to its small size of approximately 1000 in comparison to 58,000 for the full sample. However, again the extremal dependence of the sample examined is consistent with either the Gaussian process or the inverted max-stable model. Had we found instead that our sample exhibits asymptotic dependence, a choice between various max-stable model needs to be made. Our diagnostics are not ideal for this; for more details about model selection for asymptotic dependence see Davison et al. (2012).

We examined the behaviour of different estimators for η and $\chi(x)$ for simulated samples from processes with different known extremal dependencies. We found that, in general, different
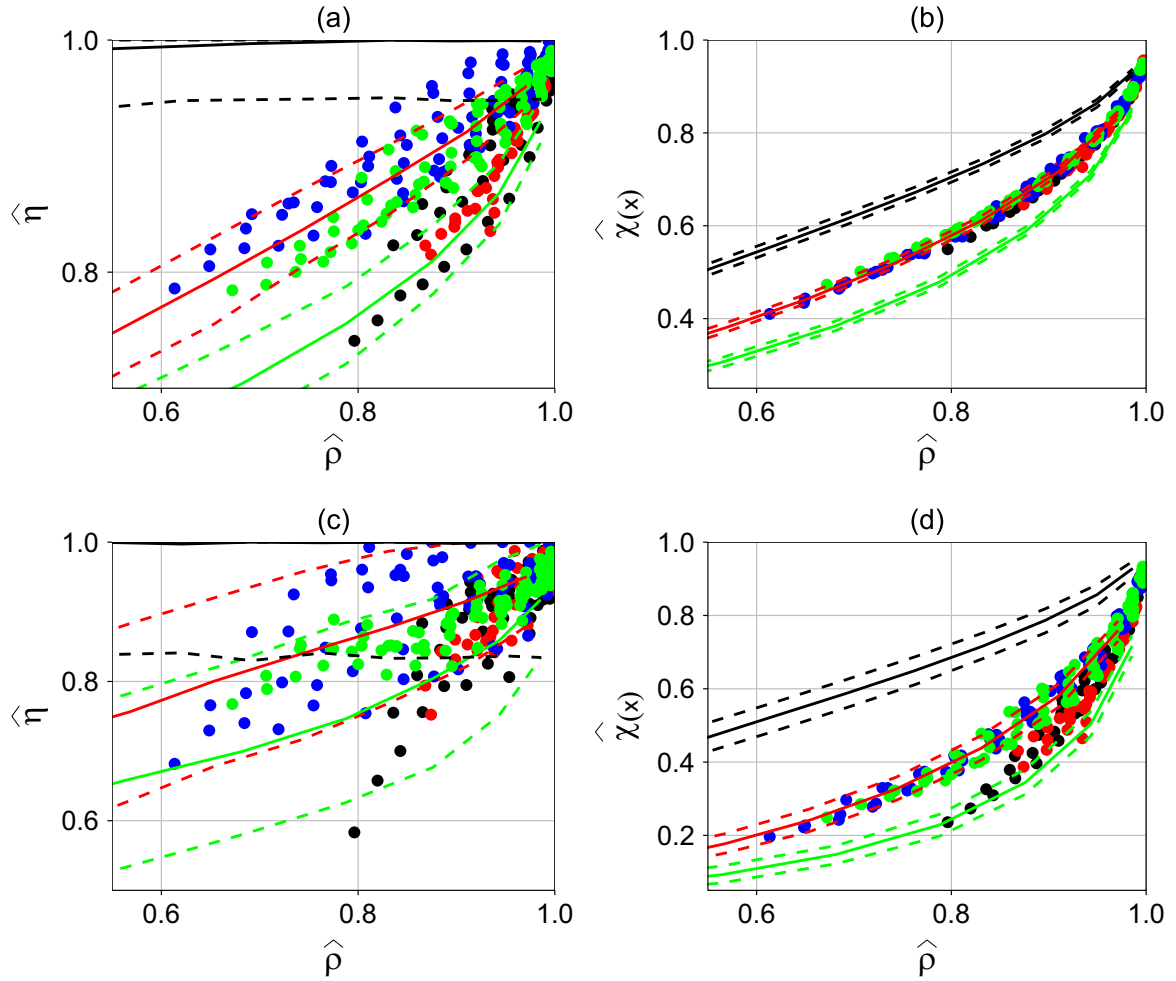
Fig. 12. Estimates of η with (a) $q=0.90$ and (c) $q=0.99$, and $\chi(x)$ with (b) $x=10$ and (d) $x=100$, plotted against Spearman's ρ for sea-state H_S sample of size $n=58,585$. Coloured points identify estimates from corresponding strip. Lines identify estimates using simulated samples of same size from Smith (black) and Gaussian (red) processes, and from the inverted logistic distribution (green). (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this paper.)

estimators yield consistent estimates for η and $\chi(x)$. As might be expected, the bias of estimates reduces with increasing tail threshold, and variance of estimates reduces with increasing tail sample size. For samples with sizes of the order of 10^4 typical for metocean applications, interpretation of plots of estimated η and $\chi(x)$ as a function of estimated ρ is difficult. In particular, when estimated ρ is small, η estimated from asymptotically dependent

processes is biased, and $\chi(x)$ from asymptotically independent processes only approaches the limiting value of zero for levels x that are unachievable in practice. It appears that comparison of sample-based plots of η and $\chi(x)$ with ρ for significant wave height with those for samples of the same size from bivariate processes with known extremal dependence provides more intuitive interpretation.
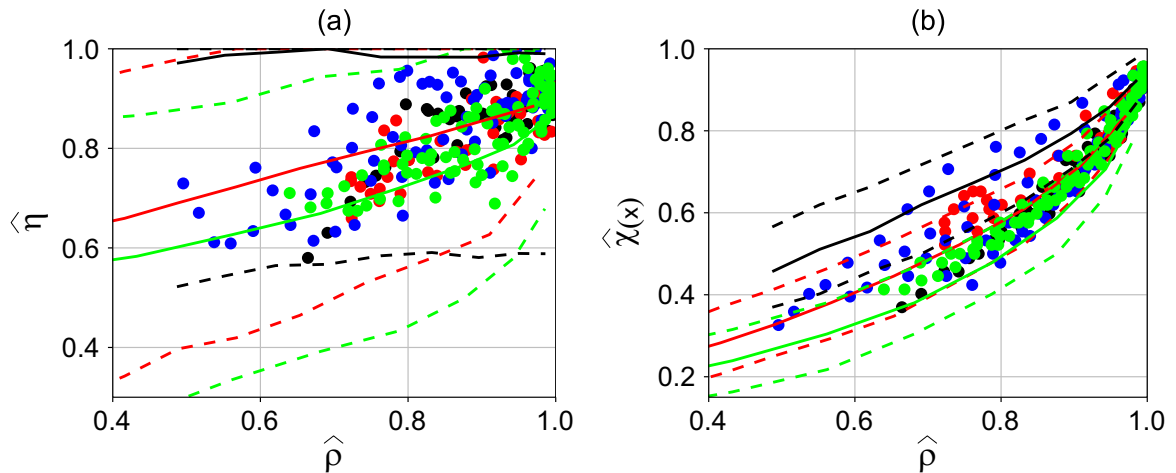
Fig. 13. Estimates of (a) η with $q=0.90$ and (b) $\chi(x)$ with $x=10$, plotted against Spearman's ρ for storm-peak H_S sample of size $n=916$. Points and lines as described in Fig. 12.
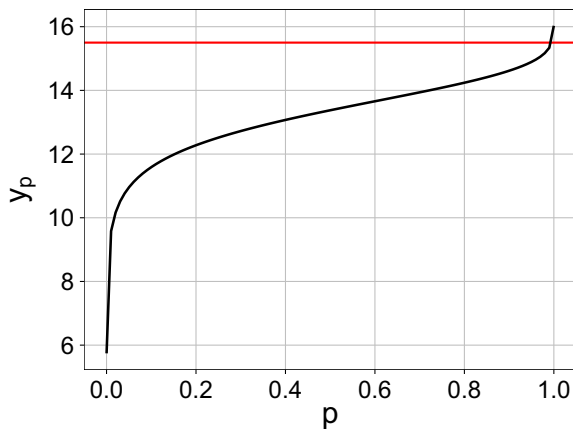
Fig. 14. Values of y_p plotted against values of p.

The extremal spatial dependence of sea-state and storm-peak significant wave height in the North Sea appears, at the very least, not to be inconsistent with asymptotic independence, implying that extreme value models admitting asymptotic independence should be sought for such applications. The great majority of models for spatial extremes which have been applied to environmental problems, motivated by max-stable assumptions, admit asymptotic dependence only; these are not ideal therefore for the current application. The conditional extremes model of Heffernan and Tawn (2004) provides a potential solution, although the approach does not lend itself naturally to spatial modelling over continuous space. The hybrid model of Wadsworth and Tawn (2012) is an alternative more suitable to spatial application; but considerable effort would be needed to develop a reliable engineering implementation. We found that the asymptotically independent Gaussian process and the inverted max-stable processes seem to fit well to the data. As the inverted max-stable models benefit from the recent investment of model development in max-stable processes, we foresee future spatial modelling of extremes of H_S using these models.

Inspection of Fig. 13 shows that identifying the appropriate form of extremal dependence for storm-peak significant wave height is problematic for samples of size approximately 1000. Since this is critical to reliable design, and since samples of size less than 1000 are typical for metocean applications, it is appealing to seek means of improved identification of extremal dependence for storm-peak H_S. The obvious approach is to gather larger samples of storm-peak values, but this is often impossible in practice. Another approach might be to assume that the dependence between contemporaneous values of sea-state H_S (for which sample size is large, and identification of extremal dependence somewhat easier) is similar if not equal to that of contemporaneous peak-over-threshold values. However, in reality, times of storm-peak occurrences at neighbouring locations are not the same; for example, they are obviously associated with the storm trajectory in space and time across the neighbourhood. Perhaps estimates of extremal dependence for time-series of sea-state H_S, appropriately time-lagged such that storm-peak events are contemporaneous, offer an approximate solution.

In some ocean basins, pooling of samples from locations within a spatial neighbourhood is performed routinely to increase sample size for extreme value modelling (see e.g., Heideman and Mitchell, 2009). Data from the different locations that are pooled are assumed to follow a common marginal extreme value distribution, so that any marginal spatial non-stationarity is neglected. Moreover, the extremal dependence of values from different locations is also typically ignored; at best, uncertainties in estimated return values are inflated to reflect dependence using a suitable

bootstrapping scheme. A well-informed spatial extremes model, accommodating the appropriate form of extremal dependence, offers reduced bias and improved uncertainty quantification.

The diagnostic tools explored in this work are of general relevance to any application in which understanding the characteristics of the joint occurrences of extreme events is of interest. For example, the nature of extremal dependence between H_S at two locations may differ from location to location within an ocean basin, and between ocean basins. Application of the diagnostic analysis suggested here for other locations in the North Sea, and for other ocean basins, would quantify this. We might explore the spatial behaviour of joint extremes of variables other than H_S, such as wind speed between locations (or altitudes), or current speed between locations (or water depths). There is no reason to expect that wind speed and current speed exhibit the same form of extremal dependence as H_S, and the form of dependence may differ temporally depending on prevailing conditions at the locations of interest. For example, if we are interested in understanding the loading on risers of a moored vessel, appropriate models for extreme current profiles with depth are required, incorporating appropriate forms for extremal dependence; otherwise biased estimates of riser loads and tensions may result, as explained in Section 5 above.

The current work is also relevant in improving understanding of engineering design practices which implicitly assume, often with little justification, a particular form of extremal dependence between variables: an estimate of return value for a particular oceanographic variable, with corresponding estimate of associated return value for another related variable, falls into this category. For example, any parametric model for the joint structure of two oceanographic variables (e.g., H_S and peak wave period T_P, or H_S and wind speed) is usually estimated by fitting the model to observations, usually not exclusively of jointly extreme events. We therefore expect that the assumed parametric form explains the bulk of observations well. But the parametric form also imposes a specific structure on the extremal dependence between variables which may or may not be justified. Estimating the type of extremal dependence present explicitly using the current diagnostics for jointly extreme events is at least advisable if not essential in demonstrating that the parametric form used is also appropriate for joint extremes.

When there is doubt concerning the nature of extremal dependence present, conventional wisdom is that the metocean engineer should err on the side of caution and assume asymptotic dependence, since this will yield more and larger joint occurrences of large events. However, this approach too may have undesirable knock-on consequences. For example, consider the design of a multi-component system. The incorrect assumption of asymptotic dependence for variables in one component (which actually exhibit asymptotic independence) might result in greater build cost for that component. When overall build cost is constrained, this might lead to less build resources for other system components. That is, over-conservatism in one aspect of design causes lack of conservatism in others, and hence overall increased structural risk. A sensible approach in such circumstances is to balance risk in a statistically valid manner with respect to competing design requirements, requiring consistent assessment of risk throughout design. In a metocean context, specification of consistent return values for a set of oceanographic variables is challenging, because of the difficulty of quantifying the dependence between extreme values of two or more variables. The current work provides diagnostics to assist in quantifying extremal dependence rationally.

Acknowledgement

Appendix A. Derivation of the PC prior

Simpson et al. (2014) use the Kullback–Leibler divergence (KLD) to measure the increased complexity from g to f. In other words, it is a measure of the information lost when the base model g is used to approximate the more flexible model f. This is defined by

$$\text{KLD}(f \parallel g) = \int f(x) \log\left(\frac{f(x)}{g(x)}\right) dx.$$

For $g(x)$ and $f(x)$ as defined in Section 3.1, we can calculate the KLD between g and f as,

$$\text{KLD}(f \parallel g) = -\log \eta + \eta - 1.$$

This can be transformed onto a physically interpretable distance scale:

$$d_{fg} = \sqrt{2\text{KLD}(f \parallel g)} = \sqrt{-2 \log \eta + 2(\eta - 1)}$$

Simpson et al. (2014) assume that the penalisation rate is constant, which implies an exponential prior on the distance scale. Hence, $d_{fg} \sim \text{Exponential}(\lambda)$ and $\pi_\lambda(d_{fg}) = \lambda \exp(-\lambda d_{fg})$, $\lambda > 0$. Since d_{fg} is a function of η $(d = \sqrt{-2 \log \eta + 2(\eta - 1)})$, we can use Jacobian transformation to obtain a prior for η; i.e. $\pi_\lambda(\eta) = \pi_\lambda(d_{fg}) \cdot |J|$, where $J = dd_{fg}/d\eta$. This gives the following prior for η,

$$\pi(\eta; \lambda) = \frac{\lambda(1 - \eta) \exp(-\lambda\sqrt{-2 \log \eta + 2(\eta - 1)})}{\eta\sqrt{-2 \log \eta + 2(\eta - 1)}}.$$

References

Aarnes, A.J., Breivik, O., Reistad, M., 2012. Wave extremes in the North East Atlantic. J. Clim. 25, 1529–1543.
Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. Statistics of Extremes: Theory and Applications. Wiley, Chichester.
Breivik, O., Aarnes, O.J., Bidlot, J.-R., Carrasco, A., Saetra, O., 2013. Wave extremes in the North East Atlantic from ensemble forecasts. J. Clim. 26, 7525–7540.
Brown, B.M., Resnick, S.I., 1977. Extreme values of independent stochastic processes. J. Appl. Probab. 14 (4), 732–739.
Coles, S.G., 2001. An introduction to statistical modelling of extreme values. Springer, Berlin.
Coles, S.G., Heffernan, J.E., Tawn, J.A., 1999. Dependence measures for extreme value analyses. Extremes 2 (4), 339–365.
Coles, S.G., Tawn, J.A., 1991. Modelling extreme multivariate events. J. R. Stat. Soc. B 53 (2), 377–392.
Coles, S.G., Tawn, J.A., 1994. Statistical methods for multivariate extremes: an application to structural design. Appl. Stat. 43 (1), 1–48.
Cooley, D., Cisewski, J., Erhardt, R.J., Mannshardt, E., Omolo, B.O., Sun, Y., 2012. A survey of spatial extremes: measuring spatial dependence and modeling spatial effects. Revstat 10 (1), 135–165.
Davis, R.A., Klüppelberg, C., Steinkohl, C., 2013. Statistical inference for max-stable processes in space and time. J. R. Stat. Soc. B 75 (5), 791–819.
Davison, A.C., Padoan, S.A., Ribatet, M., 2012. Statistical modeling of spatial extremes. Stat. Sci 27 (2), 161–186.
de Haan, L., 1984. A spectral representation for max-stable processes. Ann. Probab. 12 (4), 1194–1204.
de Haan, L., de Ronde, J., 1998. Sea and wind: multivariate extremes at work. Extremes 1 (1), 7–45.
de Haan, L., Ferreira, A., 2006. Extreme Value Theory: An Introduction. Springer, New York.
de Haan, L., Resnick, S.I., 1989. Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. Stoch. Process. Appl. 32 (2), 213–224.
Demarta, S., McNeil, A.J., 2005. The t copula and related copulas. Int. Stat. Rev. 73 (1), 111–129.
Eastoe, E.F., 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., 2014. Evaluating environmental joint extremes for the offshore industry. J. Mar. Syst. 130, 124–130.
Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., Hothorn, T., 2014. mvtnorm: Multivariate Normal and t Distributions. R Package Version 1.0-2.
Heffernan, J.E., Resnick, S.I., 2007. Limit laws for random vectors with an extreme component. Ann. Appl. Probab. 17 (2), 537–571.
Heffernan, J.E., Tawn, J.A., 2004. A conditional approach for multivariate extreme values. J. R. Stat. Soc. B 66 (3), 497–546.
Heideman, J.C., Mitchell, D.A., 2009. Grid point pooling in extreme value analysis of hurricane hindcast data. J. Waterw. Port Coast. Ocean Eng. 135 (2), 31–38.
Jonathan, P., Ewans, K.C., Forristall, G.Z., 2008. Statistical estimation of extreme ocean environments: the requirement for modelling directionality and other covariate effects. Ocean Eng. 35, 1211–1225.
Jonathan, P., Ewans, K.C., Randell, D., 2013. Joint modelling of extreme ocean environments incorporating covariate effects. Coastal Eng. 79, 22–31.
Kabluchko, Z., Schlather, M., de Haan, L., 2009. Stationary max-stable fields associated to negative definite functions. Ann. Probab. 37 (5), 2042–2065.
Kereszturi, M., 2016. Assessing and Modelling Extremal Dependence in Spatial Extremes (Ph.D. thesis). Lancaster University, Lancaster. In preparation.
Laurini, F., Tawn, J.A., 2012. The extremal index for GARCH(1, 1) processes. Extremes 15 (4), 511–529.
Ledford, A.W., Tawn, J.A., 1996. Statistics for near independence in multivariate extreme values. Biometrika 83 (1), 169–187.
Ledford, A.W., Tawn, J.A., 1997. Modelling dependence within joint tail regions. J. R. Stat. Soc. B 59 (2), 475–499.
Lee, P.M., 2012. Bayesian Statistics: An Introduction, fourth ed. Wiley, Chichester.
Nelsen, R.B., 2006. An Introduction to Copulas, second ed. Springer, New York.
Papastathopoulos, I., Tawn, J.A., 2015. Conditioned limit laws for inverted max-stable processes. arXiv:1402.1908v3.
Pickands, J., 1975. Statistical inference using extreme order statistics. Ann. Stat. 3 (1), 119–131.
Reistad, M., Breivik, O., Haakenstad, H., Aarnes, O.J., Furevik, B.R., Bidlot, J.-R., 2011. A high-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea, and the Barents Sea. J. Geophys. Res. 116, 1–18.
Ribatet, M., 2008. SpatialExtremes: A R Package for Modelling Spatial Extremes. R Package Version 2.0-2.
Schlather, M., 2002. Models for stationary max-stable random fields. Extremes 5 (1), 33–44.
Sibuya, M., 1960. Bivariate extreme statistics. Ann. Inst. Stat. Math. 11, 195–210.
Simpson, D.P., Martins, T.G., Riebler, A., Fuglstad, G.-A., 2014. Penalising model component complexity: a principled, practical approach to constructing priors. arXiv:1403.4630v3.
Smith, R.L., 1990. Max-stable processes and spatial extremes. Unpublished article, available electronically from: ⟨http://www.stat.unc.edu/postscript/rs/spatex.pdf⟩.
Spearman, C., 1904. The proof and measurement of association between two things. Am. J. Psychol. 15 (1), 72–101.
Stephenson, A.G., 2002. evd: Extreme Value Distributions. R Package Version 2.3-0.
Tawn, J.A., 1988. Bivariate extreme value theory: models and estimation. Biometrika 75 (3), 397–415.
Wadsworth, J.L., Tawn, J.A., 2012. Dependence modelling for spatial extremes. Biometrika 99 (2), 253–272.
Wadsworth, J.L., Tawn, J.A., 2013. A new representation for multivariate tail probabilities. Bernoulli 19 (5B), 2689–2714.