Ocean Engineering 62 (2013) 68–77

Contents lists available at SciVerse ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Statistical measures of extremal dependence illustrated using measured sea surface elevations from a neighbourhood of coastal locations

Emma Eastoe [a,*], Sotirios Koukoulas [b], Philip Jonathan [c,**]

[a] Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YF, UK
[b] Department of Geography, University of the Aegean, Mytilene 81100, Greece
[c] Shell Technology Centre, Thornton, Chester CH1 3SH, UK

ARTICLE INFO

Article history:
Received 7 November 2011
Accepted 5 January 2013

Keywords:
Extreme waves
Extremal dependence
Statistical models
Non-stationarity
Peaks over threshold
Generalised Pareto
Conditional extremes

ABSTRACT

The extremal dependence of stationary time-series at pairs of locations can be summarised using one or more of a number of statistics. We illustrate the application of the coefficient of tail dependence, the χ and $\overline{\chi}$ statistics, and the conditional extremes model of Heffernan–Tawn to estimate the extremal dependence in time-series of 3-h maxima of sea surface elevation across a spatial array of measurement gauges at the US Army Corps of Engineers' Field Research Facility on the Atlantic coast of North Carolina. Although the original data are non-stationary, we induce stationarity on a site-by-site basis using a non-parametric model to remove the mean trend. Subsequently, we find that pairs of locations are generally asymptotically dependent. Parameter estimates for the Heffernan–Tawn model, although uncertain, suggest that characteristics of conditional extremes vary systematically with distance from the conditioning site.

1. Introduction

In this paper, we discuss a variety of statistical approaches to quantify the extremal dependence between measured sea surface elevations at pairs of locations, illustrated using measurements from an Atlantic coastal location near Duck, North Carolina. Statistical models which accurately incorporate extremal dependence can provide more accurate estimates for site-specific return levels than those provided from models fitted on a site-by-site basis. In addition, an understanding of the extremal dependence structure between sea surface elevations at different locations is important for the siting and maintenance of off-shore structures. In particular, it is useful to understand whether or not particularly large waves are likely to occur concurrently at two (or more) sites. If only two locations are to be jointly modelled, we refer to a *bivariate* model, and, more generally, to a multivariate model.

When modelling multivariate data, it is important to distinguish between two features of the data: the marginal and dependence structures. By marginal behaviour, we refer to the characteristics of the data at a given site. This might include non-stationarity in the mean and/or variance, and within-series correlation. There may also be non-homogeneity in the extremal behaviour. By dependence structure, we refer to the joint characteristics of the data at two or more sites. In particular, we are interested in the joint characteristics of large values. Our approach follows a standard two-stage procedure, first modelling the margins, and then, conditional on the marginal models, modelling the dependence structure between pairs of sites.

Our main focus is on the extremal dependence between surface elevation at pairs of adjacent sites, viewed as a pair of random variables. Extremal dependence is a specific mathematical feature of all multivariate probability distributions. For pairs of random variables, it describes the joint behaviour of two variables as one or other, or both, becomes large. Extremal dependence takes one of two forms: asymptotic dependence or asymptotic independence. If variables are asymptotically dependent, large values of the variables tend to occur simultaneously. If they are asymptotically independent, large values never occur together. Even within these two classes, there are varying strengths of association.

Identification of which class of extremal dependence structure a data set falls into is critical to the correct choice of multivariate extreme value model. Further, if an incorrect assumption is made on the form of the asymptotic dependence, then any inferences drawn from the model may be invalid. For example, suppose that surface elevations at two sites were assumed to be asymptotically dependent, when in fact they are asymptotically independent. Then the model would overestimate the probability of a jointly extreme events at the two sites. Summary statistics, such as the

* Corresponding author. Tel. +44 1524 593954.
** Corresponding author.
E-mail addresses: e.eastoe@lancaster.ac.uk (E. Eastoe), skouk@geo.aegean.gr (S. Koukoulas), philip.jonathan@shell.com (P. Jonathan).

0029-8018/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.oceaneng.2013.01.002

χ and $\overline{\chi}$ measures of Coles et al. (1999) and the coefficient of tail dependence (Ledford and Tawn, 1996, 1997), can tell us to which class of extremal dependence a data set belongs; and, given the class, the strength of association. These statistics are discussed in Section 3.2.

A range of models for multivariate extreme values exists; selection of an appropriate model depends on whether the data is asymptotically dependent or asymptotically independent. For asymptotically dependent data, models based on the multivariate extreme value distribution and associated point process representation are appropriate (Coles and Tawn, 1991; Ledford and Tawn, 1997). However, such models are of not appropriate for asymptotically independent data, since they reduce the data to the case of independence. Recent work by Heffernan and Tawn (2004) and Ramos and Ledford (2009) attempts to redress this. We follow the approach taken by Heffernan and Tawn (2004); they suggest a conditional representation, in which, given that one variable is large, the distribution of the remaining variable(s) is modelled. Unlike the traditional models, this approach extends easily to higher dimensional variables and incorporates both classes of extremal dependence.

An outline of the paper is as follows. The data are introduced in Section 2. Section 3 discusses the approach that we take to measuring and modelling the extremal dependence structure of sea surface elevations at pairs of sites. The data analysis is presented in Section 4, with a discussion in Section 5.

2. Data

2.1. Description

Occurrences of large waves are typically caused by one or more of wind forcing (e.g. due to differences in atmospheric pressure during a storm, hurricane or cyclone) and periodic tidal forcing (generated by the gravitational interaction of the Sun, Moon and Earth). For an ocean engineer, reliable structural design relies on adequate understanding of (the extremes of) waves, winds and currents and their interactions. In this work we consider time-series of measured 3-h maximum sea surface elevations from the *8 m array* at the Field Research Facility (FRF), the US Army Corps of Engineers' Atlantic coastal observatory near Duck, North Carolina, for the period from January to December 2005. The 8 m array provides 2 Hz measurements of surface elevation (in mm) from 15 pressure gauges mounted approximately 0.5 m above the sea bed, approximately 900 m offshore, arranged approximately along two intersecting lines (one longshore, one cross-shore), as illustrated in Fig. 1 along with bathymetry. Spacings between gauges are irregular by design, varying from a minimum of 5 m in both long- and cross-shore directions, to maxima of 255 m long-shore and 120 m cross-shore. Further details are available from the USACE web page. Excellent images of the location are available from the USACE Field Research Facility web page. Fig. 2 shows time-series plots of 3-h measured surface elevation at sites 4 and 7, and a scatter plot of the data at these sites. The scatter plot suggests a strong dependence between surface elevation at the two sites; the (Pearson) correlation is 0.97.

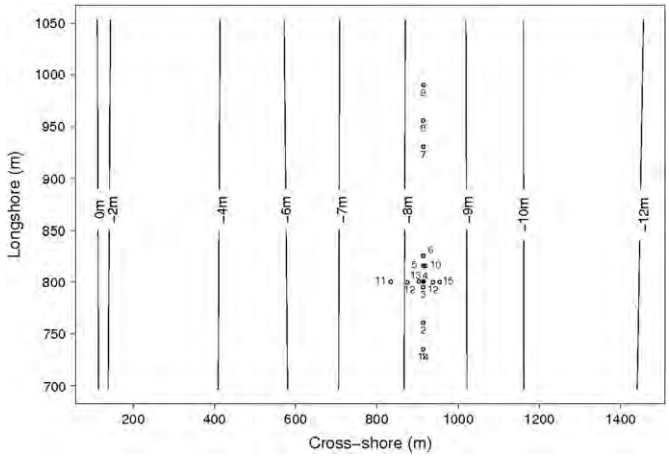

Fig. 1. Locations of data collection sites and contours of water depth.

From Fig. 2 it is also clear that the marginal behaviour is not straightforward. The data are non-stationary; non-parametric smoothing of the mean using locally weighted regression (loess) (Cleveland and Devlin, 1988), see Fig. 2, shows that there are a number of cyclic trends present in the mean surface elevation, with periods extending from approximately 12 h to 1 month, attributed to tidal effects.

The loess estimate requires selection of a smoothing parameter, which specifies the number of neighbouring observations used to obtain an estimate of the mean at a given time point; we selected this parameter to use 2% of the observations, since this value provided a reasonably smooth moving mean together with approximately stationary residuals time-series. Non-stationarity in the extremes is discussed in Section 4.1.

2.2. Within-series dependence

Due to tidal influences, the data show strong within-series correlation. For example, at site 4, the correlations at lags 1, 2, 4 and 8 are 0.70, 0.43, 0.83 and 0.76. Note, lag 4(8) corresponds to a lag of 12(24) h, so the strong correlations at these lags reflect lunar semi-diurnal tidal constituents. Similar behaviour is observed at the other sites.

This marginal dependence extends to the extremes. If we consider extremes to be all exceedances of a high threshold, the extremal index $\theta \in [0,1]$ (Smith and Weissman, 1994) is a measure of within-series dependence in these extremes. Independent data have $\theta = 1$, although $\theta = 1$ does not necessarily imply independence in the full data set, merely independence between extremes. The closer θ is to zero, the stronger the dependence between extremes. For series in which the extremes are not independent, the extreme points will tend to occur in independent clusters; the mean cluster size is given by the reciprocal θ^{-1}.

A number of methods are available to estimate the extremal index, including runs and blocks estimators (Smith and Weissman, 1994), an automatic scheme proposed by Ferro and Segers (2003) and the two-threshold approach suggested by Laurini and Tawn (2003). We use the simple runs estimator: for some choice of run length m, threshold exceedances separated by a run of at least m consecutive non-exceedances are considered to be (approximately) independent. Since the value of m is user-defined, the sensitivity of the estimate of the extremal index to the choice of m should be checked.

For our data, we consider extreme values to be exceedances above a site-specific threshold. For each site, we set the threshold to be the empirical 90% quantile of the data at that site, giving 292 exceedances per site. Assuming that any exceedances separated by a run of 9 consecutive non-exceedances are independent[1], the extremal index is 0.09 at site 4 and 0.10 at site 7. Equivalently threshold exceedances occur, on average, in clusters of sizes 11 and 10, respectively. This corresponds to clusters of 33 and 30 h, respectively.

A common approach to dealing with extremal dependence is to extract clusters and analyse the independent cluster maxima

[1] Estimates of the extremal index stabilised for run lengths greater than 9.

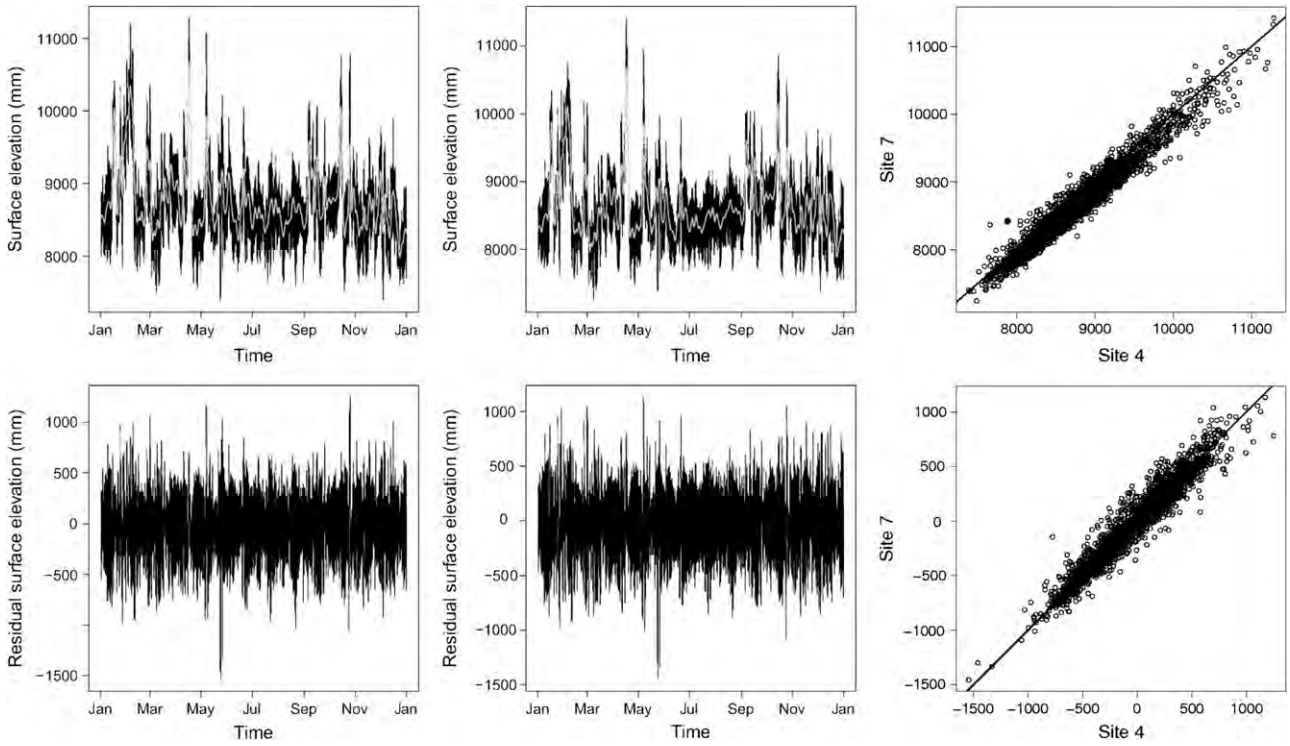

Fig. 2. Top: 3-h measured surface elevation at sites 4 (left) and 7 (centre); scatter plot of surface elevation at the two sites (right). Gray lines show loess estimate of the mean. Bottom: residual surface elevation at sites 4 (left) and 7 (centre); scatter plot of residual surface elevation at the two sites (right). Residual surface elevation is surface elevation after removal of the loess mean.

only. Although asymptotic theory states that the distribution of the arbitrary and cluster maxima exceedances are the same (Leadbetter, 1991), recent work by Fawcett and Walshaw (2007) has suggested that a bias is obtained when estimating the distribution of all threshold exceedances using the cluster maxima only. Eastoe and Tawn (2012) show why this bias occurs and suggest an alternative model for cluster maxima. Assuming independent extremes, when in fact they are clustered, has implications on inference. In particular, standard errors of parameter estimates will be under-estimated.

As we show below, pre-whitening our data to remove non-stationarity has the beneficial side-effect of markedly reducing the within-series dependence. Thus we can avoid the issue of clustering.

2.3. Pre-whitening

Various methods for dealing with non-stationarity in models for extreme values have been proposed. A common approach to handling non-stationarity is to include covariates in the extreme value model parameters (Davison and Smith, 1990; Chavez-Demoulin and Davison, 2005; Mendez et al., 2006, 2008; Jonathan and Ewans, 2011). Instead, we follow a pre-whitening approach (Eastoe and Tawn, 2009), and remove the mean structure at each site, before modelling the dependence structure between residuals at pairs of sites. Covariates can additionally be included in the extreme value model, if this is thought to be necessary, i.e. if the extremes are thought to have some additional (e.g. directional) structure.

The intention of pre-whitening is to create a "residual" time-series which shows no variability with respect to one or more covariates (such as time or wave direction in an ocean engineering context) by subtraction of a "mean" time-series which is itself a relatively smooth function of the covariates of interest. Extreme value analysis then proceeds on the residual time-series which can then be assumed to be stationary. The pre-whitening approach has a number of advantages over the direct covariate method; a full list is provided by Eastoe and Tawn (2009). In particular, since the physical reasons for non-stationarity in the body of the data and the extremes are linked, the pre-processing approach, which utilises all available information on non-stationarity, is more likely to identify the appropriate form of non-stationarity. Additionally, it can be computationally simpler and more efficient to fit a pre-whitening model; in particular, if a large number of covariates are required to explain the non-stationarity.

Since environmental covariates are unavailable, we model the mean structure using the loess fits shown in Fig. 2, where time is a proxy for covariates. Fig. 2 shows the residuals at sites 4 and 7 following removal of the mean trend. The auto-correlation functions (not shown) have been considerably reduced; at site 4, the correlations at lags 1, 2, 4 and 8 are 0.23, −0.46, 0.61 and 0.61, respectively. The extremes of the residuals also seem to cluster less, with average cluster sizes of 3.5 and 3.3 at the two sites. Further, a scatter plot of the residuals suggests that the dependence between the two sites is not much changed by considering the residuals only; the correlation remains high at 0.97.

We proceed with the rest of our analysis by considering the residuals only. We do this under the assumptions that, at any given site, the residuals are independent in time and are stationary in mean and variance. Although the independence assumption could be questioned, we wish to avoid issues that arise from using marginally de-clustered data for multivariate analyses if possible. In particular, since de-clustering occurs marginally, i.e. on a site by site basis, the resulting multivariate data are unlikely to have been observed at concurrent time points. Therefore, since dependence is quite weak, we feel justified in using the full data sets. Note that estimation of return values using the pre-whitening approach described here would require simulation from and aggregation of estimated "mean" and residuals, as explained for example in Jonathan et al. (2012).

3. Methods

Almost all methods for modelling the joint extremal behaviour of two or more random variables, for example surface elevations at different locations, involve separate models for marginal and multivariate features. In this section we discuss a possible approach for each of these steps.

3.1. Methods for univariate extremes

Given observations from a process $\{x_t\}$ indexed by time t, a common statistical assumption is that these observations are taken from an independent and identically distributed sequence of random variables $\{X_t\}$ with common distribution function F. The distribution function completely characterises the behaviour of the variables, in particular the upper tail of F tells us about the extremal behaviour. In practice, F is unknown and so must be estimated from the sample $\{x_t\}$. The simplest estimate is the empirical distribution function $\tilde{F}$; if $x^{(i)}$ is the ith largest observation, then $\tilde{F}(x^{(i)}) = n-i+1/(n+1)$, where n is the length of the series.

Since we are specifically interested in the tail behaviour, and in particular we wish to extrapolate beyond the highest observations, we would instead like a parametric model for F—or at least for the upper tail. Extreme value theory provides us with exactly this. Asymptotic theory (Pickands, 1975) motivates the use of the generalised Pareto (GP) distribution as a model for the upper tail of F. This model is suitable regardless of the true form of F and corresponds to modelling all exceedances of some high threshold u by the GP distribution,

$$\overline{F}_u(x) = \Pr[X > x | X > u] = \left[1+\xi\left(\frac{x-u}{\psi}\right)\right]_+^{-1/\xi}, \quad x > u, \tag{3.1}$$

where $z_+ = \max\{z,0\}$. Note that this is a conditional distribution, with two parameters, referred to as scale, $\psi > 0$, and shape, ξ. The shape parameter is determined by how fast the underlying distribution F decays. Estimates of the parameters can be obtained using any method of statistical inference, including maximum likelihood (Pawitan, 2001), L-moments (Hosking, 1990) or Bayesian inference (Coles and Powell, 1996; Gelman et al., 2004).

Combining the GP model for the upper tail, with the empirical distribution for observations below the threshold, gives an improved model $\hat{F}$ for the distribution F,

$$\hat{F}(x) = \begin{cases} 1-\phi\overline{F}_u(x), & x > u; \\ \tilde{F}(x), & x \le u. \end{cases} \tag{3.2}$$

Here $\phi = \Pr[X > u]$ is required to undo the conditioning in Eq. (3.1) and is estimated by the proportion of observations exceeding u. Once the full model for F has been obtained the data can be transformed to have any marginal distribution by application of the probability integral transform. In particular, in Section 3.2 we assume that data have a Gumbel distribution. Let y denote the data on the Gumbel scale, then

$$y = -\log[-\log \hat{F}(x)]. \tag{3.3}$$

3.2. Summary statistics for extremal dependence

Various summary statistics exist to assess whether two processes $\{Y_{1,t}\}$ and $\{Y_{2,t}\}$ on marginal Gumbel scales are asymptotically dependent or asymptotically independent; for example χ and $\overline{\chi}$ (Coles et al., 1999) and the coefficient of tail dependence η (Ledford and Tawn, 1996, 1997). Since they have a natural interpretation, we focus on the χ and $\overline{\chi}$ statistics. In fact, the coefficient of tail dependence is related to $\overline{\chi}$ by

$$\overline{\chi} = 2\eta - 1. \tag{3.4}$$

We can exploit this relationship to estimate $\overline{\chi}$ (see below). The statistic χ is defined as

$$\chi = \lim_{y\to\infty} \Pr[Y_2 > y | Y_1 > y] = \begin{cases} 0 & \text{if asymptotically independent;} \\ c & \text{if asymptotically dependent,} \end{cases}$$

where $0 < c < 1$. In practice, we estimate $\chi(y) = \Pr[Y_2 > y | Y_1 > y]$ for a range of y and consider the behaviour of $\chi(y)$ with increasing y. Increasing values of χ correspond to stronger asymptotic dependence. Estimation is non-parametric, using sample proportions.

In the same way that χ tells us about the strength of asymptotic dependence, $\overline{\chi}$ tells us about the strength of asymptotic independence between Y_1 and Y_2. The definition of $\overline{\chi}$ is not so straightforward, see Coles et al. (1999) for details, but its interpretation is as follows,

$$\overline{\chi} = \begin{cases} 1 & \text{if asymptotically dependent;} \\ c_1 & \text{if asymptotically independent with positive association;} \\ 0 & \text{if independent;} \\ c_2 & \text{if asymptotically independent with negative association,} \end{cases}$$

where $0 < c_1 < 1$ and $-1 < c_2 < 0$. Again, $\overline{\chi}$ is calculated at a range of levels y, and the behaviour of $\overline{\chi}(y)$ considered as y increases. We estimate $\overline{\chi}$ by first estimating the coefficient of tail dependence η and then exploiting relationship (3.4). We use a model-based approach for the estimation of η, as proposed by Ledford and Tawn (1996) (see Appendix A for details).

Uncertainty in both χ and $\overline{\chi}$ can easily be assessed using a non-parametric bootstrap (Davison and Hinkley, 1997). Under the assumption that the multivariate data are independent and identically distributed, we sample with replacement from the original multivariate data set, to create a bootstrapped sample of the same size as the original data set. Estimates of χ and $\overline{\chi}$ are obtained for each bootstrapped data set, and these estimates are treated as samples from the sampling distributions of the respective parameters.

3.3. Modelling extremal dependence using the Heffernan–Tawn model

In an ocean engineering context, the Heffernan–Tawn conditional extremes model has been used by Jonathan et al. (2010, 2012) to model wave spectral parameters and current profiles with depth, respectively. In a hydrological context, the model has been used by Keef et al. (2009) to model spatial dependence in extreme river flows and Mendes and Perrichi (2009) to model flood risk across multiple rivers. We provide a summary of the model in the bivariate case. For further details, and the extension to the multivariate setting, see Heffernan and Tawn (2004).

The Heffernan–Tawn model assumes Gumbel margins; since data are very rarely on Gumbel margins, the first step is to fit the marginal model described in Eq. (3.2). The data can then be transformed to the required margins using transformation (3.3).

Let the random variable Y_j represent surface elevation at site j, $1 \le j \le p$, on Gumbel margins. Let v_j be a high threshold for Y_j. Then, conditional on $Y_j > v_j$ being large, we assume that, for $i \ne j$,

$$\begin{aligned} &\mathbb{E}[Y_i | Y_j = y_j] = \alpha y_j + \mu y_j^{\beta}, \\ &\mathrm{sd}(Y_i | Y_j = y_j) = \sigma y_j^{\beta}, \end{aligned} \tag{3.5}$$

where $0 \le \alpha \le 1$, $\beta < 1$ and $\sigma > 0$. The model is essentially a regression model, in which α tells us how fast Y_i grows with Y_j, and β tells us how the variability of Y_i changes as Y_j increases. To

express this in terms of extremal dependence, α tells us the overall strength of the extremal dependence between Y_i and Y_j, whereas β tells us how the strength changes with Y_j. The combination $\alpha = 1$ and $\beta = 0$ corresponds to asymptotic dependence. The parameters μ and σ are considered to be nuisance parameters.

This model makes no assumption about the distribution of the conditional variable $Y_i | Y_j > y_j$, and neither does it allow for a negative extremal association between the variables. If a negative association exists, then $\alpha = 0$ and $\beta < 0$ and an additional additive term, taking the form $\gamma - \delta \log(y_j)$, is required in the expectation. The $\log(y_j)$ term accounts for the fact that the lower tail of the Gumbel distribution decays far more rapidly than the upper tail.

In order to estimate the parameters, first define the residuals

$$Z_i = \frac{Y_i - \alpha y_j}{y_j^{\beta}}. \tag{3.6}$$

Heffernan and Tawn (2004) use the working assumption that, conditional on $Y_j > v_j$, the distribution of Z_i is Gaussian, with mean and variance given by μ and σ^2 from Eq. (3.5). Implicit in this is the assumption that the residuals are independent of the conditioning variable Y_j. For inferential purposes, independence is also assumed, both in time and between sites. This allows the use of a straightforward maximum likelihood function to obtain parameter estimates.

Uncertainty in the estimates is obtained through a semi-parametric bootstrap, in order to account for both dependence and marginal uncertainty. The basic scheme, for two variables, is as follows. For each bootstrap sample,

Step 1: Sample with replacement from the bivariate data on the Gumbel scale, to obtain a sample of the same size as the original.

Step 2: Take two independent samples from the Gumbel distribution; use these to replace the data in the sample from step 1, but ensuring that the bivariate ordering remains the same. By this, if in the original sample the largest value of variable 1 occurs with the fourth largest value of variable 2, then the same should be true in the new sample, and so forth.

Step 3: Transform to the original margins, using the inverse of transformation (3.3).

Step 4: Refit the entire model to the bootstrapped data set.

This gives a set of parameter estimates which can be assumed to be taken from the sampling distribution of the parameters.

To assess model fit, Heffernan and Tawn (2004) suggest checking that, for $Y_j > v_j$, the residuals Z_i defined in Eq. (3.6) are independent of the conditioning variable Y_j. This can be assessed through a simple plot of the residuals against the conditioning variable. In addition, as with all threshold-based extreme value models, it might be desirable to assess how sensitive the model parameters are to changes in the threshold v_j used for the conditioning variable.

4. Data analysis

To illustrate the methods discussed in Section 3, we first fit marginal models at each site. We then model the bivariate extremal dependence between each of the sites and the 'central site', site 4.

4.1. Testing for homogeneity in the extremes

The pre-whitening method discussed in Section 2.3 does not guarantee that non-stationarity in the extremes will be removed. As noted in Eastoe and Tawn (2009), the trends in the extremes of a process may be quite different in form to those in the main body. Before fitting the marginal model discussed in Section 3.1 we first use various *ad hoc* measures to check that the extremes are homogeneous in time. A simple approach is to split the process into blocks according to the value of the covariate of interest (here time), to assess whether or not the rate and size of exceedances differs between blocks.

Consider site 4, with a 90% threshold, corresponding to 292 exceedances. Splitting the series into 12 blocks of consecutive observations (approximately months), we can calculate a point estimate and confidence interval for the numbers of exceedances per month (see Appendix B). If exceedances occur homogeneously in time, we would expect approximately 24 (15,34) exceedances per month; figures in brackets give an asymptotic 95% confidence interval. In practice, the observed number of exceedances per month all lie within the 95% confidence interval, varying from 15 to 29. This suggests that the rate of exceedances is homogeneous, at least on a monthly scale. Further, there are no obvious trends in the monthly exceedance counts.

To investigate trends in the size of extremes we consider running quantiles. Similar to a running mean, for each time point one calculates the local quantile using only observations in the moving time window of pre-specified size centred at the time

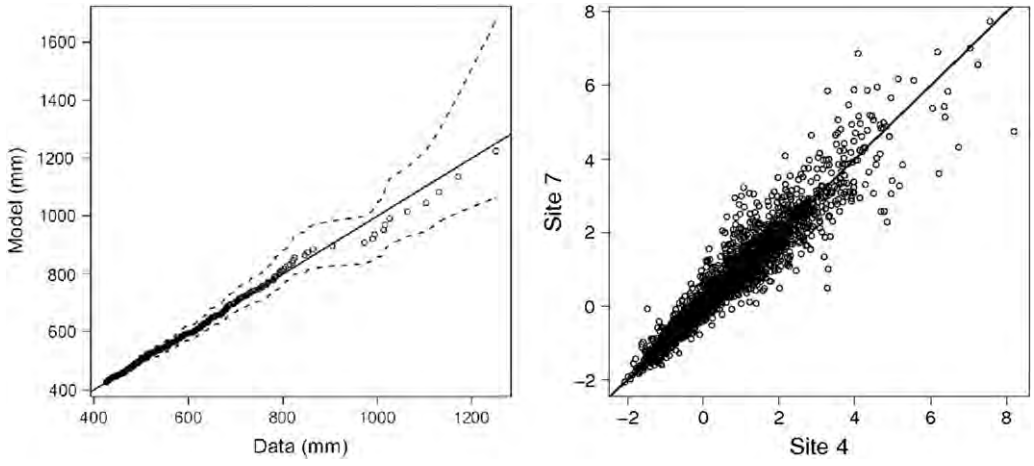

Fig. 3. Left: quantile plot to illustrate fit of GP distribution to the exceedances of a 90% threshold at site 4. Dashed lines show 95% bootstrapped confidence intervals. Right: scatter plot of residual surface elevation at sites 4 and 7, on the Gumbel scale.

point under consideration. Although there is evidence of more variability in the quantiles as the level is increased, for example the running 95% quantile is more variable than the running median, no interpretable trends are discernible. We therefore continue our analysis under the assumption that the extremes are time homogeneous, both in rate and size.

Table 1
Maximum likelihood estimates of generalised Pareto parameters, with 95% confidence intervals in brackets. Confidence intervals were calculated using asymptotic properties of the likelihood function.

Site	ψ	ξ
1	154 (126,182)	−0.010 (−0.15,0.13)
2	161 (134,188)	−0.037 (−0.16,0.086)
3	160 (135,186)	−0.039 (−0.15,0.073)
4	156 (131,182)	−0.038 (−0.15,0.078)
5	130 (106,154)	0.076 (−0.073,0.22)
6	150 (123,177)	0.054 (−0.085,0.19)
7	158 (131,185)	−0.074 (−0.20,0.052)
8	147 (122,172)	−0.034 (−0.16,0.092)
9	146 (123,169)	0.081 (−0.029,0.19)
10	127 (105,148)	0.15 (0.016,0.28)
11	131 (109,154)	0.15 (0.025,0.28)
12	135 (113,157)	0.16 (0.040,0.27)
13	139 (115,162)	0.123 (−0.00037,0.25)
14	141 (116,166)	0.043 (−0.092,0.18)
15	130 (108,153)	0.10 (−0.029,0.23)

4.2. Marginal models

We first fit model (3.2) at each site, using maximum likelihood inference for parameter estimation (Pawitan, 2001). A site-specific threshold is used, corresponding to the 90% quantile calculated individually for each site. Graphical diagnostics, such as probability and quantile plots (Coles, 2001), show that the generalised Pareto distribution fits well to the threshold exceedances, at all sites. An example of the quantile plot for site 4 is shown in Fig. 3. The points follow the 45° slope well, showing close agreement between model and data; further all the points lie within the 95% parametric bootstrapped confidence intervals.

For site 4, the GP parameter estimates, with 95% confidence intervals in brackets, are, $\hat{\psi} = 160(130,182)$ and $\hat{\xi} = -0.038$ $(-0.15,0.078)$. Confidence intervals were obtained using an asymptotic property of the likelihood function (Pawitan, 2001). At all sites, the confidence interval for the shape parameter contains zero, suggesting that it is not significantly different to zero, except at sites 10, 11 and 12, where ξ is significantly positive (see Table 1). At the sites with shape not significantly different to zero, we could instead fit an exponential or Weibull model. This would reduce the number

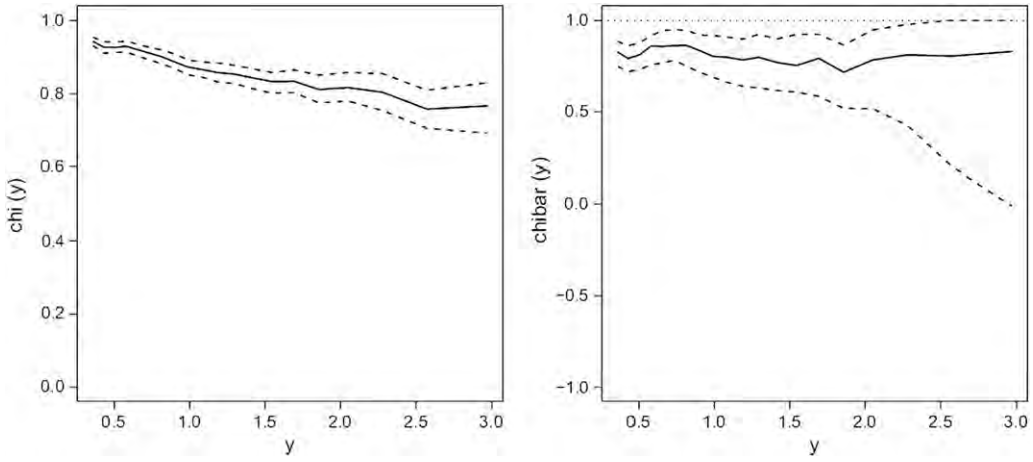

Fig. 4. Plots of $\chi(y)$ (left) and $\overline{\chi}(y)$ (right) for sites 4 and 7, as a function of y on the Gumbel scale. Full lines show point estimates, dashed lines 95% bootstrapped confidence intervals. The dotted line on the plot of $\overline{\chi}(y)$ indicates asymptotic dependence.

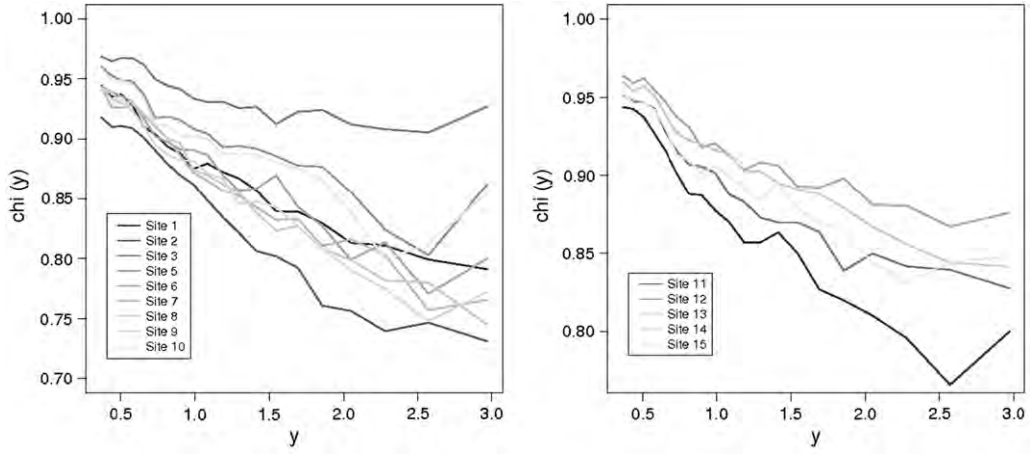

Fig. 5. Variation of $\chi(y)$ with y (on the Gumbel scale) for each site conditional on site 4, in long-shore (left) and cross-shore (right) directions. The value of $\chi(y)$ for long-shore location 3 is particularly large.

of parameters and thus remove some model complexity, however, we prefer to retain the greater flexibility of the generalised Pareto model fit. Further, as Table 1 shows, there is not one pair of sites for which the 95% confidence intervals for ψ (or ξ) do not overlap. Consequently, we prefer to fit the same model at all sites.

Combining the fitted GP models with the empirical distribution function for non-exceedances, we transform the data to Gumbel margins on a site-by-site basis, using Eq. (3.3). Fig. 3 shows a scatter plot of the data at sites 4 and 7 on these margins. This plot emphasises the usefulness of the Gumbel scale in magnifying the joint tail behaviour.

4.3. Extremal dependence

The extremal dependence of sites 4 and 7 can be seen in Fig. 4, which shows plots of $\chi(y)$ and $\overline{\chi}(y)$ for $y \in [0.37,3.2]$, i.e. from the 50% to the 95% quantile of the Gumbel distribution. Smaller sample sizes corresponding to higher thresholds did not allow

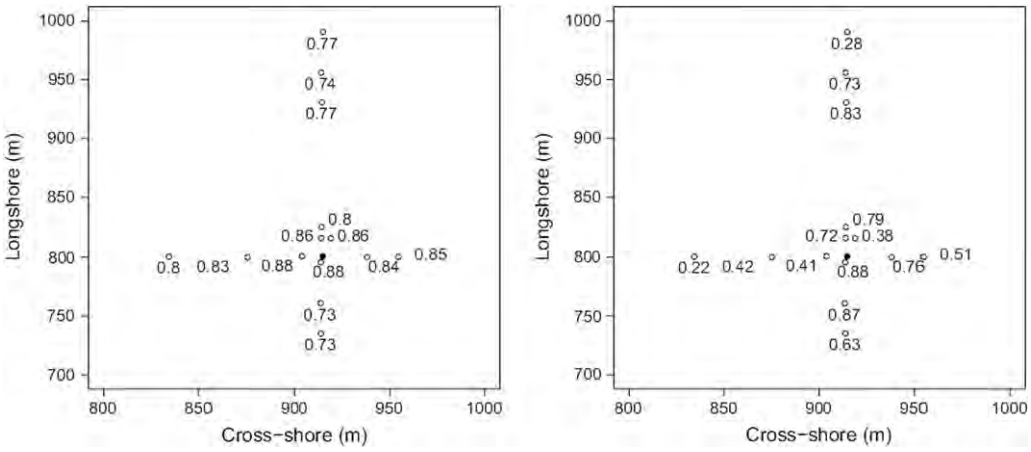

Fig. 6. Plots of χ (left) and $\overline{\chi}$ (right) for site 4 with all remaining sites. All estimates were calculated using the 95% Gumbel quantile.

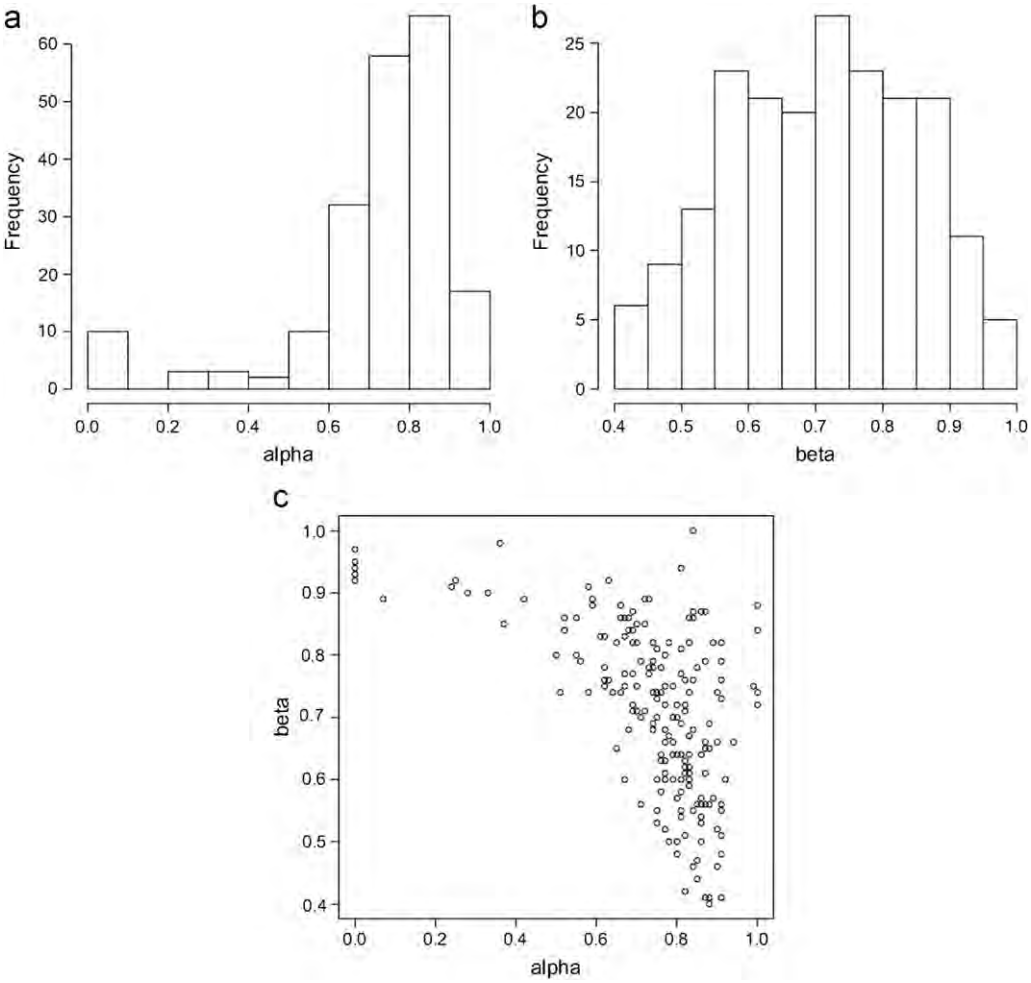

Fig. 7. For the Heffernan–Tawn model fit to site 1, conditional on site 4 being greater than the 85% quantile, plots of the marginal sampling distributions for α (top left) and β (top right) as well as for the joint sampling distribution of (α,β) (bottom). The sampling distributions were obtained using 200 bootstrap replicate data sets.

reliable estimation of summary statistics. Confidence intervals were obtained using 200 bootstrapped samples. Since the 95% confidence intervals for $\chi(y)$ do not contain 0, it would appear that the sites are asymptotically dependent. However, $\chi(y)$ decreases as y increases, with no evidence of attaining a limiting value. More conclusive evidence on the limiting behaviour of $\chi(y)$ would require a larger sample, so that we could consider higher y. The point estimate of $\overline{\chi}(y)$ is close to 1 for all y. At higher values of y the 95% confidence interval contains 1, suggesting clear evidence of asymptotic dependence.

It is also interesting to consider how χ and $\overline{\chi}$ vary with distance. Since site 4 is the central site, we estimate the extremal dependence between this site and each of the remaining 14 sites. Plots of $\chi(y)$ (Fig. 5) show that dependence decays with distance and is weaker in the long-shore than the cross-shore direction.

We now obtain point estimates for both χ and $\overline{\chi}$, using the 95% quantile of the Gumbel distribution and consider how these estimates change with distance and direction. Plots of these estimates are given in Fig. 6. The estimates of χ fall in the range 0.73–0.88, suggesting reasonably strong asymptotic dependence between site 4 and the remaining sites. Bootstrapped confidence intervals (not shown) suggest that χ is significantly greater than zero for all the pairs. A less clear pattern emerges from the estimates of $\overline{\chi}$, which also have greater uncertainty. Given that the estimates of χ suggest asymptotic dependence, we would expect the $\overline{\chi}$ estimates to fall around 1. In fact, this is not always the case, with a range of $\overline{\chi}$ between 0.22 (the furthest site to the west of site 4) and 0.88 (intuitively, the closest site to site 4). The upper bound for the confidence interval is 1 for all sites except sites 9–12 (where it varies from 0.66 to 0.93).

4.4. Heffernan–Tawn model

Conditioning on site 4 being large, we fit the Heffernan–Tawn model to the remaining 14 sites. We consider only the bivariate extremal dependence structure, that is we fit 14 separate models.

Let us first consider the model fit for site 1, conditional on site 4 being large. For the conditional model, we use the 85% quantile as a threshold for site 4, resulting in 438 observations being used for inference. The choice of threshold requires the usual trade-off between bias (if the threshold is taken too low, the asymptotic motivation fails to hold) and variance (if the threshold is too high, there will be too few observations to make sensible inference). Additionally, a lower threshold (equivalently more data) is generally required to estimate dependence parameters than would be required to estimate marginal parameters, hence our move from the 90 to the 85% quantile. At this threshold, point estimates are $\hat{\alpha}=0.82$ and $\hat{\beta}=0.65$. We would expect α to be close to one, given that the estimate of χ for this pair (0.73) suggested asymptotic dependence.

Fig. 7 shows the sampling distributions of the two parameters obtained using the bootstrap scheme described in Section 3.3 with 200 bootstrapped replicates. These plots show considerable uncertainty in both parameters. Further, the joint sampling distribution shows that the two parameters are not independent.

From the marginal sampling distributions we can derive 95% confidence intervals of (0.00,0.94) for α and (0.44,0.95) for β. These confidence intervals, in particular those for α, are not highly informative. This is in part due to the fact that, when $\beta=1$, α becomes unidentifiable. The reason for this can be seen by substituting $\beta=1$ into Eq. (3.5). Then the conditional expectation for $Y_i(i\neq 4)$ becomes $(\alpha+\mu)y_4$, so that α and μ are redundant, explaining the wide range of values taken by α when β is close to 1 in Fig. 7.

Retaining the 85% threshold for the conditional modelling, Fig. 8 shows the estimates of α and β for each site, conditional on site 4 being large. As expected, since it measures the overall strength of extremal dependence, α is highest for those sites closest to site 4. In fact, $\alpha=1$ for the closest site, site 3, and $\beta=0.13$, strongly suggesting asymptotic dependence. However, for the remaining sites, the estimates of β are not particularly close to zero. Table 2 gives interquartile ranges and 95% confidence intervals for each of the estimates of α and β. Fig. 9 shows convex hulls for the sampling distributions of α and β, conditional

Table 2
Point estimates for α and β for each site, conditional on site 4 exceeding the 85% quantile. Figures in brackets are interquartile ranges and 95% confidence intervals (both bootstrapped).

Site	α	β
1	0.82 (0.69,0.84) (0.00,0.94)	0.65 (0.60,0.82) (0.44,0.95)
2	0.95 (0.91,0.97) (0.82,1.0)	0.29 (0.29,0.48) (0.10,0.66)
3	1.0 (0.99,1.0) (0.91,1.0)	0.13 (0.21,0.64) (0.09,0.95)
5	0.95 (0.90,0.96) (0.81,1.0)	0.49 (0.48,0.69) (0.24,0.84)
6	0.90 (0.82,0.91) (0.49,0.98)	0.57 (0.56,0.75) (0.33,0.91)
7	0.38 (0.01,0.80) (0.00,0.95)	0.93 (0.82,0.96) (0.61,1.0)
8	0.66 (0.27,0.78) (0.00,0.92)	0.82 (0.75,0.93) (0.61,1.0)
9	0.69 (0.45,0.75) (0.00,0.91)	0.63 (0.63,0.86) (0.41,0.98)
10	0.87 (0.85,0.91) (0.76,0.99)	0.34 (0.40,0.59) (0.17,0.81)
11	0.84 (0.81,0.92) (0.60,1.0)	0.57 (0.55,0.73) (0.33,0.93)
12	0.85 (0.84,0.91) (0.74,0.99)	0.31 (0.33,0.57) (0.08,0.88)
13	0.91 (0.89,0.97) (0.76,1.0)	0.50 (0.47,0.74) (0.21,0.97)
14	0.99 (0.96,1.0) (0.88,1.0)	0.47 (0.44,0.67) (0.11,0.86)
15	0.89 (0.80,0.93) (0.56,1.0)	0.59 (0.57,0.79) (0.40,1.0)

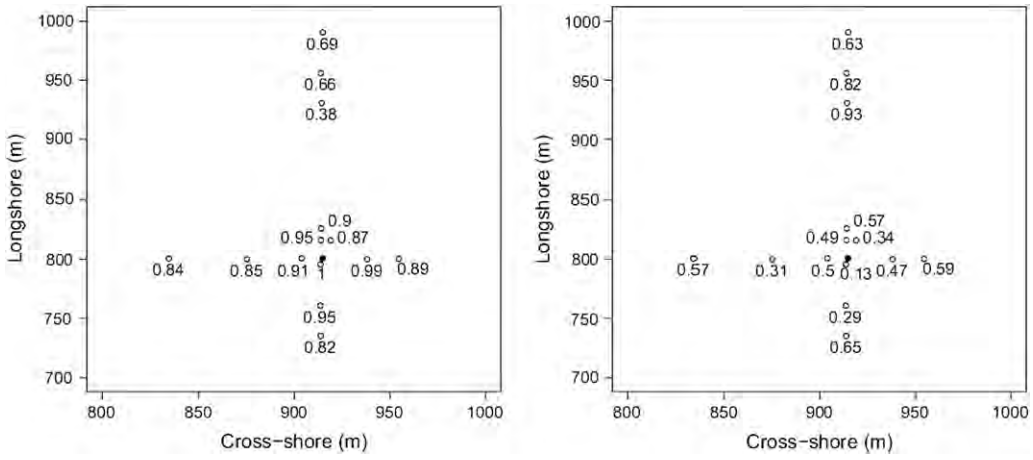

Fig. 8. Plots of α (left) and β (right) for each site conditional on site 4 exceeding the 85% quantile.

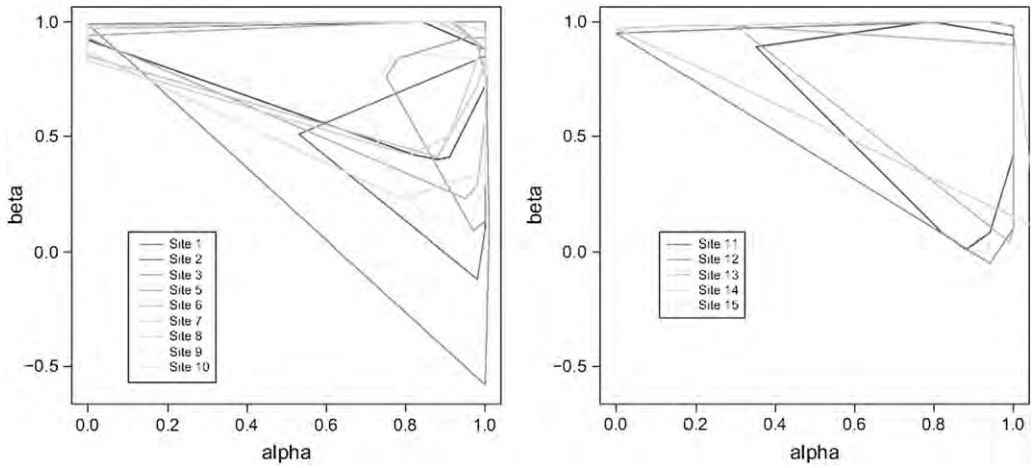

Fig. 9. Convex hulls of α and β for each site conditional on site 4 exceeding the 85% quantile, in long-shore (left) and cross-shore (right) directions.

on site 4 being large, in long- and cross-shore directions. There is some suggestion that estimates for β are larger at sites 7, 8 and 9 than elsewhere.

For all 14 models, plots (not shown) of the residuals against the conditioning variable showed no evidence of any trends in the residuals with the conditioning variable. This shows that the modelling assumption that the residuals are conditionally independent of the conditioning site is satisfied. Although, for all sites, estimates of α and β were found to vary across thresholds and there were no large inconsistencies with respect to threshold.

5. Discussion

For all 15 sites, marginal generalised Pareto models fit well. The estimated shape parameters ξ are generally slightly negative, but bootstrapped 95% confidence intervals include zero. The exceptions are sites 10, 11 and 12, at which ξ is significantly positive (see Table 1). This suggests that the marginal tails are relatively heavy, consistent with a Weibull distribution (or ever heavier). Physical considerations suggest that there are upper limits to the size of sea waves due to steepness constraints, and for shallow waters due to water depth (see, e.g. Battjes and Groenendijk, 2000, however, it is unlikely that either of these constraints is active for the current measurements. For example, the largest wave heights in the data are approximately 3 m in 8 m of water, yielding a ratio of 0.375, considerably smaller that the Miche limit (Battjes and Groenendijk, 2000), with limiting ratio of approximately 0.8). Similarly, there is no evidence of (steepness-induced) wave breaking in the measurements. It is interesting to note that the 8 m array was established to estimate directional wind-wave spectra, under the assumption that the wave field is spatially homogeneous. The current analysis suggests that are no clear spatial trends in marginal extreme characteristics of residual 3-h maxima surface elevation.

Estimates of the extremal dependence χ statistic conditional on site 4 being large are significantly different from zero over the range of thresholds examined, suggesting that all pairs of locations involving site 4 exhibit asymptotic dependence. Uncertainty bands for estimates of $\overline{\chi}$ are generally sufficiently close to unity that asymptotic dependence is also plausible. These findings are of practical significance in that they support the application of multivariate extreme value models assuming asymptotic dependence, e.g. composite likelihood methods motivated by consideration of component-wise maxima for these data. There is no strong indication that the values of χ or $\overline{\chi}$ show distance- or direction-related trends, except perhaps that χ values are larger for cross-shore dependence with site 4 than long-shore.

For the conditional extremes model of Heffernan–Tawn, notwithstanding the fact that estimates of α and β are generally subject to considerable uncertainty, in part due to issues with identifiability, estimates of β shown in Fig. 8 appear to increase as a function of distance from conditioning site 4, with the possible exception of site 9. This concords with physical intuition with Eq. (3.5) in mind: variability of conditional extremes increases with distance from the conditioning site. Similarly intuitively pleasing trends are observed in estimates of α as shown in Fig. 8 and Table 2. These are lower for the more long-shore distant sites 7–9 than for nearer neighbours. The bootstrapped 95% confidence bands for α in Table 2 include zero for sites 7, 8 and 9 and the most distant of the remaining sites long-shore (site 1) and cross-shore (site 11). There is no indication that the characteristics of conditional extremes vary systematically with direction.

The methods demonstrated in the current work are generally applicable to characterise extremal dependence in the ocean environment (see, e.g. discussion of Jonathan et al., 2010). In particular, the Heffernan–Tawn model provides a means to simulate arbitrarily long samples of characteristics of extreme ocean environments. Interesting further applications would be to quantify the extremal dependence of (a) more distant locations (which might be expected to be asymptotically independent), (b) coastal locations with varying water depth, and (c) significant wave heights across an ocean basin (particularly when storms events are localised in space and/or strongly directional).

Acknowledgements

The authors would like to acknowledge discussions with Kevin Ewans of Sarawak Shell Bhd.

Appendix A. Estimating the coefficient of tail dependence

Ledford and Tawn (1996) base their estimation of the coefficient of tail dependence on the following model for the joint survivor function of two unit Fréchet random variables. Suppose that X_1^F and X_2^F have unit Fréchet margins, such that $\Pr(X^F \le x) = e^{-1/x}$ for X_1^F and X_2^F, then for large x,

$$\Pr[X_1^F > x, X_2^F > x] \sim \mathcal{L}(x) x^{-1/\eta}, \tag{A.1}$$

where $\mathcal{L}(\cdot)$ is a slowly varying function at infinity and $0 \le \eta \le 1$ is the coefficient of tail dependence. If $\eta = 1$, the variables are asymptotically dependent.

To estimate η, Ledford and Tawn use the 'structure variable' $T = \min(X_1^F, X_2^F)$. From Eq. (A.1), $\Pr[T > x] \sim \mathcal{L}(x)x^{-1/\eta}$. Consider then the conditional probability $\Pr[T > u+t \mid T > u]$; for large u, this reduces to

$$\Pr[T > u+t \mid T > u] \sim \frac{\mathcal{L}(u+t)}{\mathcal{L}(u)} \frac{(u+t)^{-1/\eta}}{u^{-1/\eta}} \sim \left(1 + \frac{t}{u}\right)^{-1/\eta},$$

since $\mathcal{L}(\cdot)$ is slowly varying. In other words, exceedances of u, by T, follow a generalised Pareto distribution with scale and shape parameters given by ηu and η, respectively. Hence, to estimate η, we fit a generalised Pareto distribution to the exceedances by T of some high threshold and use the estimated shape parameter to give an estimate for η. To estimate $\eta(x)$, we use the shape parameter obtained from fitting the generalised Pareto distribution to the exceedances by T of x.

Transforming between Gumbel and Fréchet marginals is straight forward. If Y is Gumbel-distributed, then $X^F = \log(Y)$ is Fréchet-distributed.

Appendix B. Assessing homogeneity in the rate of threshold exceedances

Suppose that a series of length n is divided into B blocks each of length l. Let n_u represent the total number of exceedances of a threshold u. Assuming that exceedances occur homogeneously in time, the proportion of exceedances in a given block is the same as the proportion p in the complete series. This can be estimated by the observed proportion $\hat{p} = n_u/n$.

Now consider $N_{B,u}$, the random variable representing the number of exceedances of u in a given block. Then

$$N_{B,u} \sim \text{Binomial}(l,p),$$

with expectation $\mathrm{E}[N_{B,u}] = lp$ and variance $\mathrm{Var}(N_{B,u}) = lp(1-p)$. Provided that l is reasonably large and p is small, we can use the asymptotic approximation of the Gaussian to the Binomial distribution, with mean and variance given by the above expression, replacing p with $\hat{p}$, to obtain confidence intervals for $N_{B,u}$.

References

Battjes, J.A., Groenendijk, H.W., 2000. Wave height distributions on shallow foreshores. Coastal Eng. 40, 161–182.

Chavez-Demoulin, V., Davison, A.C., 2005. Generalized additive modelling of sample extremes. J. R. Stat. Soc. C 54, 207–222.

Cleveland, W.S., Devlin, S.J., 1988. Locally weighted regression: an approach to regression analysis by local fitting. J. Am. Stat. Soc. 83, 596–610.

Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. Springer-Verlag, London.

Coles, S., Tawn, J.A., 1991. Modelling extreme multivariate events. J. R. Stat. Soc. B 53, 377–392.

Coles, S.G., Heffernan, J.E., Tawn, J.A., 1999. Dependence measures for extreme value analyses. Extremes 2, 339–365.

Coles, S.G., Powell, E.A., 1996. Bayesian methods in extreme value modelling: a review and new developments. Int. Stat. Rev. 64, 119–136.

Davison, A.C., Hinkley, D.V., 1997. Bootstrap Methods and Their Application. Cambridge University, Press.

Davison, A.C., Smith, R.L., 1990. Models for exceedances over high thresholds. J. R. Stat. Soc. B 52, 393–442.

Eastoe, E.F., Tawn, J.A., 2009. Modelling non-stationary extremes with application to surface level ozone. J. R. Stat. Soc. C 58, 25–45.

Eastoe, E.F., Tawn, J.A., 2012. Modelling the distribution for the cluster maxima of exceedances of sub-asymptotic thresholds. Biometrika 99, 43–55.

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

Ferro, C.A.T., Segers, J., 2003. Inference for clusters of extreme values. J. R. Stat. Soc. B 65, 545–556.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. Bayesian Data Analysis. Chapman and Hall/CRC.

Heffernan, J.E., Tawn, J.A., 2004. A conditional approach for multivariate extremes. J. R. Stat. Soc. B 66, 497–546.

Hosking, J.R.M., 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. J. R. Stat. Soc. B 52, 105–124.

Jonathan, P., Ewans, K., 2011. Modelling the seasonality of extreme waves in the Gulf of Mexico. ASME J. Offshore Mech. Arct. Eng. 133, 021104.

Jonathan, P., Flynn, J., Ewans, K., 2010. Joint modelling of wave spectral parameters for extreme sea states. Ocean Eng. 37, 1070–1080.

Jonathan, P., Ewans, K.C., Flynn, J., 2012. Joint modelling of vertical profiles of large ocean currents. Ocean Eng. 42, 195–204.

Keef, C., Svensson, C., Tawn, J.A., 2009. Spatial dependence in extreme river flows and precipitation for Great Britain. J. Hydrol. 378, 240–252.

Laurini, F., Tawn, J.A., 2003. New estimators for the extremal index and other cluster characteristics. Extremes 6, 189–211.

Leadbetter, M.R., 1991. On a basis for 'peaks over threshold' modeling. Stat. Probab. Lett. 12, 357–362.

Ledford, A.W., Tawn, J.A., 1996. Statistics for near independence in multivariate extreme values. Biometrika 83, 169–187.

Ledford, A.W., Tawn, J.A., 1997. Modelling dependence within joint tail regions. J. R. Stat. Soc. B 59, 475–499.

Mendes, B.V.M., Perrichi, L.R., 2009. Assessing conditional extremal risk of flooding in Puerto Rico. Stoch. Environ. Res. Risk. Assess. 23, 399–410.

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

Mendez, F.J., Menendez, M., Luceno, A., Losada, I.J., 2006. Estimation of the long-term variability of extreme significant wave height using a time-dependent POT model. J. Geophys. Res. 11, C07024.

Pawitan, Y., 2001. In all Likelihood: Statistical Modelling and Inference Using Likelihood. Oxford University Press, Oxford.

Pickands, J., 1975. Statistical inference using extreme order statistics. Ann. Stat. 3, 119–131.

Ramos, A., Ledford, A., 2009. A new class of models for bivariate joint tails. J. R. Stat. Soc. B 71, 219–241.

Smith, R.L., Weissman, I., 1994. Estimating the extremal index. J. R. Stat. Soc. B 56, 515–528.

USACE. Field Research Facility 8m Array, US Army Corps of Engineers, Duck, North Carolina, USA. 〈http://www.frf.usace.army.mil/8mArray/about8mArray.html〉.

USACE. Field Research Facility. US Army Corps of Engineers, Duck, North Carolina, USA. 〈http://www.frf.usace.army.mil/〉.