

Modelling spatially-dependent non-stationary extremes with application to hurricane-induced wave heights

Paul Northrop and Philip Jonathan

March 5, 2010

Abstract

In environmental applications it is found frequently that the extremes of a variable of interest are non-stationary, varying systematically in space, time or with the values of covariates. Multi-site datasets are common, and in such cases there is likely to be non-negligible inter-site dependence. We consider applications in which multi-site data are used to infer the marginal behaviour of the extremes at individual sites, while making proper adjustment for inter-site dependence. For reasons of statistical efficiency, modern extreme value analyses often model exceedances of a high threshold. Choosing an appropriate threshold can be problematic, particularly if the extremes are non-stationary. We propose a method for setting covariate-dependent threshold using quantile regression. We consider how the quantile regression model and extreme value models fitted to threshold exceedances should be parameterised, in order that they are compatible. These considerations also suggest a new technique for selecting the level of extreme value thresholds. We adjust estimates of uncertainty for spatial dependence using methodology proposed recently. These methods are illustrated using time series of storm peak significant wave heights from 72 sites in the Gulf of Mexico.

Keywords: Extreme value regression modelling; dependent data; quantile regression; threshold exceedances; threshold selection; wave heights.

1 Introduction

This article deals with regression modelling of extreme values in the presence of spatial dependence. In environmental applications it is found frequently that the extremes of a variable of interest are non-stationary, varying systematically in space, time or with the values of covariates. Regression modelling is a natural way to account for such effects. It is also common to find that the response data are not independent, even once the effects of covariates have been taken into account. Multi-site datasets - containing time series of responses from a number of sites, perhaps with associated covariate values - are common, and in such cases there is likely to be non-negligible inter-site dependence. Since extremal behaviour, and systematic effects on extremal behaviour, will tend to be similar at neighbouring sites, it is advantageous to model simultaneously data from different sites and thus improve precision of estimation. However, inferences must take proper account of spatial dependence.

One way to account for spatial dependence is to model it explicitly. Recent work in this area includes Naveau et al. (2009), Casson and Coles (1999), Cooley et al. (2007). The multivariate

extreme value methodology of Heffernan and Tawn (2004) is also applicable to spatial problems. However, it is often the case that interest is in the marginal distributions of the response variable at the sites, not in the dependence between the time series of responses at the different sites. We may not wish to estimate the strength or form of the spatial dependence, but rather to ensure that our inferences have taken proper account of the fact that the response data are not independent. In such situations it is common to carry out a marginal analysis: ignoring the dependence initially, and then making adjustments to test statistics and estimates of parameter uncertainty to account for the dependence. Fawcett and Walshaw (2007) argue in favour of this approach, in the context of dealing with serial dependence in extreme values. They find that little precision of estimation is gained by modelling extremal dependence explicitly, concluding that the increased complexity, and issues of potential bias resulting from model specification and selection, are not justified unless temporal dependence is of specific interest. Chandler and Bate (2007) gives a good summary of issues pertinent to the choice between explicit modelling of dependence structures and marginal modelling.

For reasons of statistical efficiency, modern extreme value analyses often model exceedances of a high threshold. In this article we argue that if there is clear non-stationarity in extremes a non-constant threshold should be set that reflects the non-stationarity. Even in the stationary case choosing an appropriate threshold can be problematic. We propose a method for setting covariate-dependent threshold using quantile regression. We also consider how the quantile regression model and extreme value models fitted to threshold exceedances should be parameterised, in order that they are compatible. We discuss threshold selection in the presence of covariates in more detail in section 1.1.

When developing environmental design criteria for marine structures it is important to quantify the stochastic behaviour of extreme sea states. A standard measure of sea surface roughness is significant wave height, defined as the mean of the highest one third of waves. In this paper we model time series of significant wave height hindcast data from 72 locations in the Gulf of Mexico, with the aim of quantifying the heights of large waves in this region over future long time periods. Extrapolation of this kind is required to mitigate against anticipated future extremes. In regions such as the Gulf of Mexico extreme sea states are associated with hurricanes, but these occur relatively infrequently. It is desirable, therefore, to combine wave data from different spatial locations. However, in section 2 we find that there is strong positive association between the time series at different sites. Therefore inferences must take account of the strong spatial association between extreme wave heights. We also find that the data appear to exhibit systematic spatial variation. When modelling these data it is important that we capture systematic spatial effects and that we account of the strong spatial dependence in the data (Jonathan and Ewans, 2007; Jonathan et al., 2008; Jonathan and Ewans, 2009).

Jonathan and Ewans (2007) performed a marginal analysis of the wave hindcast data considered in this paper. They use bootstrap techniques (Davison and Hinkley, 1997) to produce confidence intervals for quantities of interest. Based on a simulation study they find that their bootstrapping scheme performed better, in term of producing the desired coverage probabilities, than an implementation of Heffernan and Tawn (2004). However, they do not model spatial variation in the wave data. In this paper we fit regression models in which extreme value parameters depend on location and use the methodology of Chandler and Bate (2007) as a simple and computationally convenient way to adjust our inferences for the presence of spatial dependence in the data. We ignore the potential covariates of season and hurricane direction.

Extreme value theory provides an asymptotic justification for particular families of models for extreme data, analogous to the way that the Central Limit Theorem suggests the normal

distribution as an approximate model for averages of large numbers of random variables. For an accessible introduction to modern statistical methods for extremes see Coles (2001). The theory for the simplest situation, namely the distribution of $M_m = \max(Y_1, \dots, Y_m)$, where Y_1, Y_2, \dots are independent, identically distributed random variables, dates back to Fisher and Tippet (1928). For the purposes of this discussion Y_1, Y_2, \dots occur at a series of time points. After appropriate linear rescaling of M_m , if a non-degenerate limiting distribution for M_m exists as $m \rightarrow \infty$, then it must be a member of the Generalized Extreme Value (GEV) family (Jenkinson, 1955), with distribution function

$$G(y) = \exp \left\{ - \left[1 + \xi \left(\frac{y - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad (1)$$

where $1 + \xi(y - \mu)/\sigma > 0$ and $\sigma > 0$. The case of $\xi = 0$ is interpreted as the limit of (1) as $\xi \rightarrow 0$. The parameter ξ summarises the shape of the GEV distribution, while μ and σ are location and scale parameters respectively. This suggests the $\text{GEV}(\mu, \sigma, \xi)$ distribution as an approximate model for maxima of long sequences of i.i.d. random variables. For example, if m is taken to be the number of observations in one year we model annual maxima. The value of ξ is determined by the tail behaviour of the common distribution of Y_1, Y_2, \dots . In terms of extrapolation the value of ξ is crucial: if $\xi < 0$ the distribution of M_n is bounded above at $\mu - \sigma/\xi$, while increasingly large positive values of ξ correspond to an increasingly heavy upper tail.

If an entire sequence Y_1, Y_2, \dots, Y_n is available it will often be more informative to model the number and magnitude of exceedances of some large threshold u . Suppose that the limiting distribution of M_m is $\text{GEV}(\mu, \sigma, \xi)$. Then, for sufficiently large u , given that there is an exceedance, $Z = Y - u$ has approximately a Generalized Pareto, $\text{GP}(\sigma_u, \xi)$ distribution, where the scale parameter $\sigma_u = \sigma + \xi(u - \mu)$ (Pickands, 1975). Under the assumed independence of Y_1, Y_2, \dots the number of exceedances has a binomial distribution with parameters n and $p = P(Y > u)$, leading to the binomial-GP model for threshold exceedances. Graphical methods can be used to inform the choice of u (Davison and Smith, 1990).

The extreme value models outlined above can be viewed as special cases of a point process model based on a non-homogeneous Poisson process characterisation of threshold exceedances, developed by Pickands (1971) and first used for practical application by Smith (1989). The parameterisation of this model, in terms of the GEV parameters μ, σ and ξ of the distribution of annual maxima, is invariant to u , a property which is advantageous if a non-constant threshold is used. In simple terms, the point process model is a reparameterisation of the binomial-GP model from (p, σ_u, ξ) to (μ, σ, ξ) , in which a binomial process of occurrences of exceedances in time is approximated by a Poisson process.

There are two main ways in which the theory outlined above may be deficient for practical application. Firstly, Y_1, Y_2, \dots may not be independent. However, providing that a weak condition limiting long-range dependence at extreme levels is satisfied, the GEV family can still be used to model M_m . Local dependence means that threshold exceedances occur in clusters in time. However, as we observe in section 2, temporal dependence is not an issue for the wave height data considered in this paper.

Secondly, Y_1, Y_2, \dots may not be identically distributed but rather they form a non-stationary process, indexed by covariates. The standard approach in such instances is to use regression modelling (Davison and Smith, 1990), in which the parameters of the extreme value model, and perhaps the threshold, depend on the values of the covariates. The idea is that we appeal to standard extreme value arguments conditionally on the values of covariates. It is common to assume that ξ is constant. In some applications it is found that there are insufficient data

to estimate reliably covariate effects on shape.

The main aim of an extreme value analysis is the estimation of the levels exceeded with specified small probabilities over a given period of time. Suppose that, conditional on the values of covariates, annual maxima Y_m have a $\text{GEV}(\mu, \sigma, \xi)$ distribution. The conditional N year return level q_N is defined as the $100(1 - 1/N)\%$ quantile of the distribution of annual maxima, that is, q_N such that $P(Y_m > q_N) = 1/N$. It is approximately the level exceeded once, on average, every N years. Under the assumed $\text{GEV}(\mu, \sigma, \xi)$ distribution for Y_m , q_N is given by

$$q_N = \mu + \frac{\sigma}{\xi} \left[\left\{ -\log \left(1 - \frac{1}{N} \right) \right\}^{-\xi} - 1 \right].$$

1.1 Threshold selection in the presence of covariates

If covariate effects are anticipated or observed it seems reasonable to consider setting a covariate-dependent threshold, since what is ‘extreme’ depends on the value of the covariates. For example, for seasonal data it is common to split the data into months or seasons and set a different threshold for each subset of the data (see, for example, Smith (1989)). A potentially more satisfactory approach is to specify a threshold that is a smooth function of time. Coles (2001) took this approach for temperature data from the USA, selecting a cyclical threshold by trial-and-error for which the rate of exceedance seemed constant through the year.

More generally, it seems sensible to set a threshold for which the probability p of threshold exceedance is constant across different values of the covariates. In this paper we use quantile regression (Koenker and Bassett, 1978; Koenker, 2005) to achieve approximate constancy of exceedance probability. We give further details, and consider consequences for model parameterisation, in sections 3.2.1 and 3.2.2. We are not aware of other work in which quantile regression is used to set thresholds for extreme values in this way. Jagger and Elsner (2009) carry out separate analyses of wind speeds near the USA using quantile regression and generalised Pareto modelling. They make a comparison of the respective parameter estimates (similar to the comparison we make in section 3.4.1) but otherwise the two analyses are unconnected. Eastoe and Tawn (2009) apply a constant threshold to residuals from a Box-Cox normal regression model fitted to all the data, effectively setting a covariate-dependent threshold to achieve approximate constancy of p . However, since the threshold is set using a model fitted to all the data, the threshold set is driven largely by trends in typical values of the response which is undesirable. Quantile regression is an obvious alternative to this. It models directly covariate effects at the desired level of extremity, does not require a distributional assumption for the response data, is more robust to the presence of outliers than the a Box-Cox normal regression model and retain their statistical properties under any monotonic transformation of the response. The latter property may be useful in applications where a parameter-dependent non-linear transformation is applied to the response data (Wadsworth et al., 2010).

Eastoe and Tawn (2009) also argue against standard approaches to extreme value regression modelling based on a constant threshold. One objection is that the standard binomial-GP regression models, for example Davison and Smith (1990) in which $\log \sigma_u$ is linear in the covariates and ξ is constant, do not respect a threshold stability property of the GP distribution. The problem is that this model for σ_u does not retain the same functional form over different thresholds. We note that this particular issue could be skirted by taking $\log(\sigma_u + a)$ to be linear in the covariates, where a is an unknown parameter to be estimated. Problems related to threshold stability do not exist in the point process model as all the parameters are invariant to threshold.

In section 2 we describe the GOMOS wave hindcast dataset and produce some preliminary plots to guide later analyses. In section 3 we consider how to use the point process model to analyse the hindcast data. Section 3.1 gives a summary of the inferential adjustments made for spatial dependence in the data. Sections 3.2.1 , 3.2.2 and 3.2.3 consider how to use quantile regression to set an appropriate threshold, how to parameterise appropriately the quantile regression model and the point process model and how to choose a suitable overall level for the threshold. The extreme value analysis of the hindcast data is presented in section 3.3. Inferences about return levels are presented in section 4. All the calculations and plots are produced using R (R Development Core Team, 2009).

2 Wave hindcast data

The data are derived from significant wave height (H_s) data produced by the proprietary GOMOS Gulf of Mexico hindcast study by Oceanweather Inc. (2005), for the period September 1900 to September 2005 inclusive, at 30 minute intervals. The hindcasts are produced by a physical model, calibrated to observed hurricane data. For a typical Gulf of Mexico location we selected 72 sites arranged on a 6×12 rectangular lattice with spacing 0.125 degrees (approximately 14km). For reasons of confidentiality we code longitude as $1, \dots, 12$ and latitude as $1, \dots, 6$, rather than giving their true values. Over this region there are 315 hurricane events during the period in question: an average of three events per year. For each event for each site, we isolate the storm peak significant wave height H_s^{sp} , that is, the largest H_s value observed during the event, and treat the data at each site as a discrete time series indexed by event number.

Prior to modelling we produce some preliminary plots to assess informally whether the storm peak data exhibit systematic variation or strong dependence in space or time. Figure 1 shows how the site maximum, that is, the largest storm peak observed at each site, varies with longitude and latitude. The plot suggests that the storm peak data may vary non-linearly over space. Similar effects are observed in plots of at-site estimates of high quantiles.

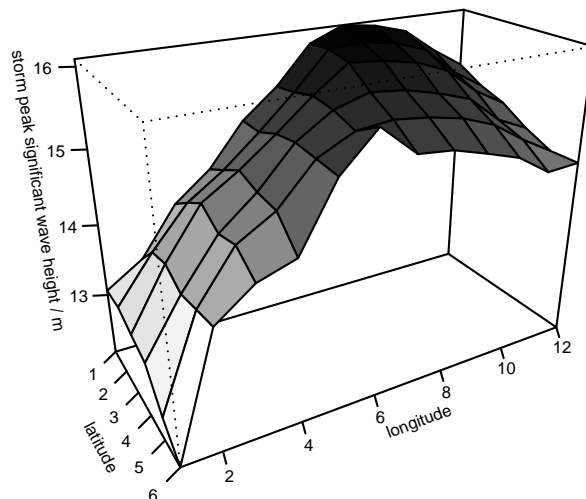


Figure 1: Site maximum of H_s^{sp} against longitude and latitude. The grey scale indicates the value: the larger the storm peak the darker the shading.

Figure 2 is plot of the network maximum, that is, the largest of the 72 observed peaks for each hurricane event, against event number. The LOESS smooth of the series (produced using the

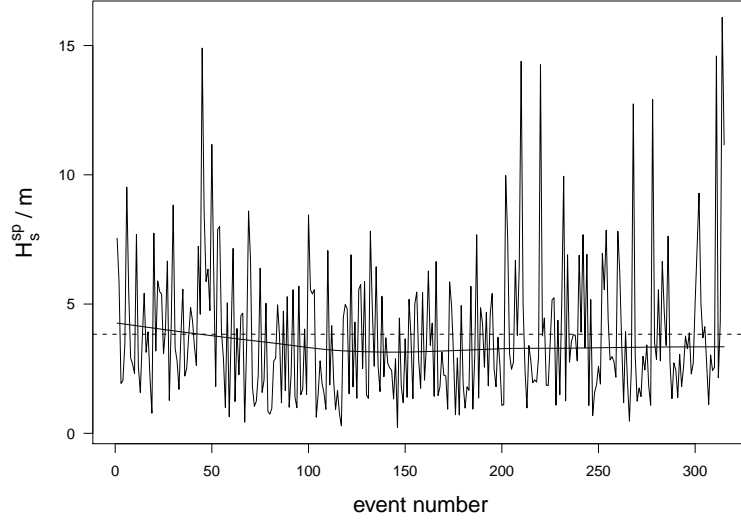


Figure 2: Network maximum of H_s^{sp} against hurricane event number. The horizontal dashed line indicates the mean of the series. The solid line is a LOESS smooth of the series.

R's loess() function with its default settings) suggests a slight decrease in average storminess over the earliest third of the data, but any trend in time appears less pronounced than over space. Therefore, for the most part, we will assume in this paper that the data are temporally homogeneous. However, there is interest in examining temporal variability of wave hindcasts, in part because methods for measuring hurricane strength have changed over the 20th century (see, for example, Killick et al. (2009)).

Figure 3 shows that, even for pairs of sites situated at opposite points of the network, there is strong spatial association in the storm peak values. This is because contemporary values at different sites are produced by the same hurricane event. Storm peak values at neighbouring sites are very similar.

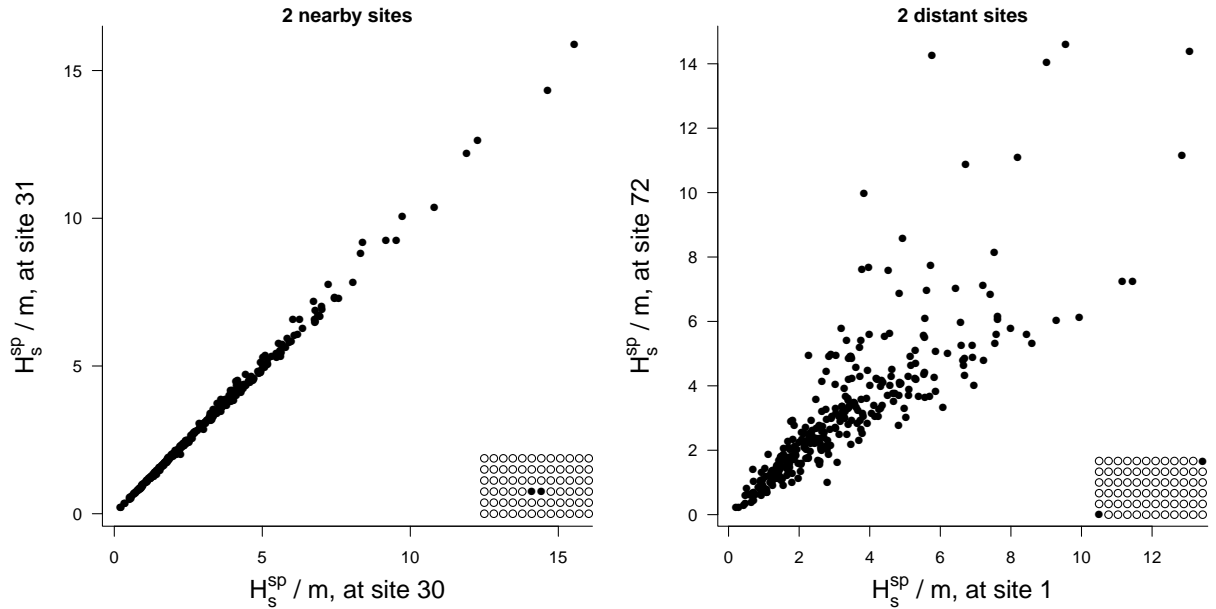


Figure 3: Scatter plots of contemporaneous H_s^{sp} values at pairs of sites. Left: contiguous sites (site 30: latitude=3, longitude=6 and site 31: latitude=3 and longitude=7). Right: distant sites (site 1: latitude=1, longitude=1 and site 72: latitude=6, longitude=12).

Storm peaks have been isolated with the aim of eliminating temporal dependence in the data. We check this informally at each site by plotting successive H_s^{sp} values against each other.

Figure 4 shows this plot for one of the centrally located grid points. Plots for other sites

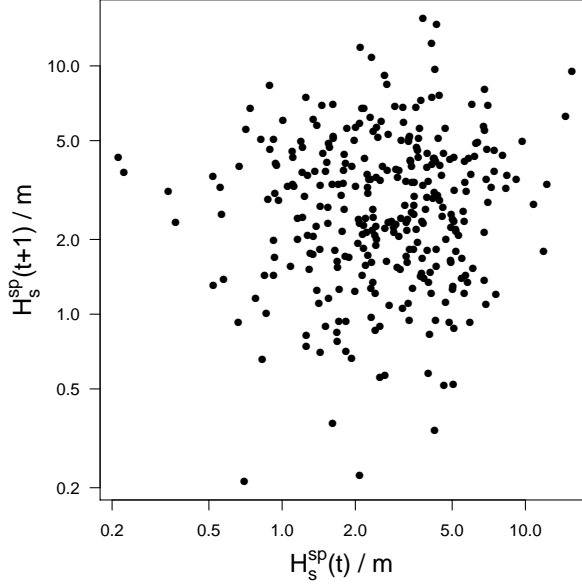


Figure 4: Scatter plot of pairs of successive H_s^{sp} values, that is, $H_s^{sp}(t+1)$ against $H_s^{sp}(t)$, at site 30 (longitude=6, latitude=3), on log scales.

are similar. Based on these plots it seems reasonable to assume that H_s values from different hurricane events are independent.

3 Extreme value modelling

Our main aim is to fit to the GOMOS wave data spatial regression models based on the point process characterisation of exceedances of a high threshold developed by Pickands (1971). We consider models in which the parameters μ, σ and perhaps ξ are related to longitude and latitude of the site. The results are given in section 3.3. We use quantile regression to set a threshold for which the probability p of threshold exceedance is constant for all sites. The details are given in section 3.2.1. We see in section 3.2.2 that the covariates included in the quantile regression model fitted are determined by the covariates included in the process model. We choose a suitable value of p in section 3.2.3 by looking for stability in parameter estimates from the point process model. The value of p , and the form of the quantile regression used to set the threshold, depend on the form of the point process model. Therefore, we use a preliminary GEV analysis of approximate annual maxima to suggest a tentative model from which to start the point process modelling.

Let the response Y_{ij} be the storm peak significant wave height at site i in storm j , where $i = 1, \dots, 72$ and $j = 1, \dots, 315$, and let \mathbf{x}_{ij} be a vector of covariates relevant to Y_{ij} . In particular, \mathbf{x}_{ij} contains functions of the longitude and latitude of site i . Let $u(\mathbf{x}_{ij})$ be the threshold applied for a given \mathbf{x}_{ij} . Let θ be a vector containing the parameters that relate the GEV parameters $\mu(\mathbf{x}_{ij})$, $\sigma(\mathbf{x}_{ij})$ and $\xi(\mathbf{x}_{ij})$ to the covariates and let $f_{ij}(y_{ij} | \mathbf{x}_{ij}; \theta)$ be the conditional density of y_{ij} given \mathbf{x}_{ij} . If, conditional on the covariate values, the responses are independent, the likelihood function under the point process model is given by

$$L_I(\theta) = \prod_{j=1}^{315} \prod_{i=1}^{72} f_{ij}(y_{ij} | \mathbf{x}_{ij}; \theta)$$

$$\begin{aligned}
&= \prod_{j=1}^{315} \prod_{i=1}^{72} \exp \left\{ -\frac{1}{\lambda} \left[1 + \xi(\mathbf{x}_{ij}) \left(\frac{u(\mathbf{x}_{ij}) - \mu(\mathbf{x}_{ij})}{\sigma(\mathbf{x}_{ij})} \right) \right]^{-1/\xi(\mathbf{x}_{ij})} \right\} \\
&\quad \times \prod_{j=1}^{315} \prod_{i: y_{ij} > u(\mathbf{x}_{ij})} \frac{1}{\sigma(\mathbf{x}_{ij})} \left[1 + \xi(\mathbf{x}_{ij}) \left(\frac{y_{ij} - u(\mathbf{x}_{ij})}{\sigma(\mathbf{x}_{ij})} \right) \right]^{-1/\xi(\mathbf{x}_{ij})-1}, \tag{2}
\end{aligned}$$

where $\sigma(\mathbf{x}_{ij}) > 0$, $1 + \xi(\mathbf{x}_{ij}) [u(\mathbf{x}_{ij}) - \mu(\mathbf{x}_{ij})] / \sigma(\mathbf{x}_{ij}) > 0$ and $1 + \xi(\mathbf{x}_{ij}) [y_{ij} - \mu(\mathbf{x}_{ij})] / \sigma(\mathbf{x}_{ij}) > 0$ for all i, j and λ is the average number of observations per year. For the hindcast data $\lambda = 3$.

We fit the point process model underlying (2) by maximising $\log L(\theta)$ with respect to θ . However, the plots in section 2 show clearly that it is inappropriate to assume that the response data are spatially independent. In section 3.1 we consider how to adjust the ‘independence’ likelihood $L_I(\theta)$ to account for spatial dependence.

Following Chandler (2005) we represent systematic spatial variation using a basis of Legendre polynomials: $\phi_i(\cdot), i = 0, 1, \dots$ (Abramowitz and Stegun, 1965). For example, if l_x represents longitude and l_y represents latitude, a quadratic representation of the effects of longitude and latitude on the GEV parameter μ is given by

$$\mu = \mu_0 + \mu_1 \phi_1(l_x) + \mu_2 \phi_1(l_y) + \mu_3 \phi_2(l_x) + \mu_4 \phi_1(l_x) \phi_1(l_y) + \mu_5 \phi_2(l_y),$$

where $\phi_1(l_x) = (l_x - 6.5)/5.5$, $\phi_1(l_y) = (l_y - 3.5)/2.5$ and $\phi_2(\cdot) = 3(\phi_1^2(\cdot) - 1)/2$. A linear representation is obtained if μ_3, μ_4 and μ_5 are set to zero.

An alternative is to use non-parametric regression modelling, for example by extending the spline-based generalized additive modelling approach of Chavez-Demoulin and Davison (2005) to the spatial situation. However, in Chavez-Demoulin and Davison (2005) the increased complexity and computational demands mean that thresholds were chosen in a rather arbitrary way. The extension to the spatial context is non-trivial as smoothing over two dimensions is more demanding of data and an additive representation. Another possibility is the local-likelihood approach used by Butler et al. (2007) to smooth temporal trends over space. The simpler polynomial basis approach used in the current paper is designed to produce a reasonable and computationally tractable approximation to the spatial structure seen in figure 1. The model diagnostics described in section 3.4 suggest that this is adequate for the data considered in this paper.

3.1 Adjusting for spatial dependence

Figure 3 suggests that the values of H_s are strongly spatially dependent within hurricane events. It is important that to take account of this dependence if inferences are to be meaningful. Following Chandler and Bate (2007), we first define an ‘independence’ log-likelihood

$$l_I(\theta) = \sum_{j=1}^{315} \sum_{i=1}^{72} \log f_{ij}(y_{ij} \mid \mathbf{x}_{ij}; \theta) = \sum_{j=1}^{315} l_j(\theta; \mathbf{y}_j),$$

where $\mathbf{y}_j = (y_{1j}, \dots, y_{72j})'$, under the assumption that all observations are independent, and then adjust $l_I(\theta)$ to account for spatial dependence. In the context of Chandler and Bate (2007) each of the 315 hurricane events constitutes a cluster of 72 observations. We assume (supported by plots such as figure 4) that, given \mathbf{x}_{ij} , data from different clusters (hurricane events) are independent.

The maximum likelihood estimator (MLE) $\hat{\theta}$ maximizes $l_I(\theta)$, that is, it is the root of

$$U(\theta) = \frac{\partial l_I(\theta)}{\partial \theta} = \sum_{j=1}^{315} \frac{\partial}{\partial \theta} l_j(\theta; \mathbf{y}_j) = \sum_{j=1}^{315} U_j(\theta) = 0.$$

Suppose that θ_0 is the true value of θ . We define H_I to be the expected Hessian of $l_I(\theta)$ at θ_0 and V to be the covariance matrix of $U(\theta_0)$. H_I is estimated by the observed Hessian, \hat{H}_I , at $\hat{\theta}$. V is estimated by $\hat{V} = \sum_j U_j(\hat{\theta})$, justified by the independence of the score contributions from different clusters. In regular problems as the number of clusters tends to infinity, in distribution,

$$\hat{\theta} \rightarrow N(\theta_0, H_I^{-1} V H_I^{-1}). \quad (3)$$

Therefore, to account for dependence within clusters Chandler and Bate (2007) advocate scaling the independence log-likelihood so that it has Hessian $\hat{H}_A = (\hat{H}_I^{-1} \hat{V} \hat{H}_I^{-1})^{-1}$ at $\hat{\theta}$. The idea is to adjust $l_I(\theta)$ so that it has the appropriate curvature at $\hat{\theta}$, namely while retaining the general shape of $l_I(\theta)$. For the extreme value models considered in this paper, (3) holds provided that $\xi > -1/2$ (Smith, 1985).

In this article we use the vertically-adjusted log-likelihood

$$l_A(\theta) = l_I(\hat{\theta}) + \frac{(\theta - \hat{\theta})' \hat{H}_A (\theta - \hat{\theta})}{(\theta - \hat{\theta})' \hat{H}_I (\theta - \hat{\theta})} \left(l_I(\theta) - l_I(\hat{\theta}) \right),$$

proposed in the discussion section of Chandler and Bate (2007). This has the advantage, over a horizontal adjustment, that the adjusted log-likelihood is always defined, even though there are constraints on the parameter space of extreme value models.

The covariance matrix in (3) is estimated by $\hat{H}_I^{-1} \hat{V} \hat{H}_I^{-1}$, from which adjusted standard errors for the parameter estimates can be obtained. Approximate confidence intervals for model parameters can be obtained by profiling $l_A(\theta)$, if a symmetric interval based on (3) is not appropriate. Nested models, with a difference in dimensionality q , can be compared using a χ_q^2 distribution for the adjusted likelihood ratio statistic (ALRS)

$$\Lambda_A = 2 \left\{ l_A(\hat{\theta}) - l_A(\tilde{\theta}) \right\},$$

where $\tilde{\theta}$ maximizes $l_A(\theta)$ subject to the constraint imposed on the unrestricted model to obtain the restricted model.

The idea of adjusting the inferences from extreme value models to account for dependence is not new. Smith (1990), who adjusts for spatial dependence, and Fawcett and Walshaw (2007), who adjust for temporal dependence, work with $l_I(\theta)$ and adjust its null distribution for the presence of dependence. In contrast, $l_A(\theta)$ is defined in order to preserve the usual asymptotic distribution of the likelihood ratio statistic. It leads to more powerful tests when more than one parameter is considered simultaneously, because the profiles have the correct orientation in the parameter space, and is easier to compute.

In section 3.2 we fit the point process models to the hindcast data, using the adjusted log-likelihood $l_A(\theta)$ to guide the model-building process. Unless stated otherwise, all standard errors and likelihood ratio statistics are based on $l_A(\theta)$.

3.2 Modelling of threshold exceedances

We use a threshold-based modelling approach, with a threshold set using quantile regression. In the presence of covariates, standard methods to choose an appropriate level for the threshold depend on which covariates are to be included in the point process model for threshold

exceedances. Therefore, it is desirable to have some idea of which model(s) to fit before selecting a threshold. We do this by defining, at each site, approximate annual maxima as the largest value of successive triplets of storm peaks (recall that there are on average 3 hurricane events per year) and fitting GEV regression models with covariates of Legendre polynomials of longitude and latitude. We find that a model in which μ is quadratic in longitude and latitude, but σ and ξ are constant, is suggested.

3.2.1 Quantile regression

Recall that we wish to set a threshold for which the probability of exceedance is constant across different values of covariates. Quantile regression (Koenker and Bassett, 1978) is used to quantify how a given conditional quantile (or quantiles) of a response variable Y depends on the observed values of covariates. Let y^τ denote the conditional τ quantile of Y , satisfying $P(Y \leq y^\tau) = \tau$. In the current example, we assume that

$$y^\tau = \beta_0 + \beta_1 \phi_1(l_x) + \beta_2 \phi_1(l_y) + \beta_3 \phi_2(l_x) + \beta_4 \phi_1(l_x) \phi_1(l_y) + \beta_5 \phi_2(l_y), \quad (4)$$

that is, we assume that y^τ is quadratic in longitude and latitude. Suppose that we have responses y_1, \dots, y_n with associated predictors $y_1^\tau, \dots, y_n^\tau$. The regression parameters are estimated by minimising

$$\min_{\boldsymbol{\beta}} \left\{ (1 - \tau) \sum_{y_i < y_i^\tau} (y_i^\tau - y_i) + \tau \sum_{y_i > y_i^\tau} (y_i - y_i^\tau) \right\},$$

with respect to $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5)'$. We use the R package `quantreg` (Koenker, 2009) to estimate $\boldsymbol{\beta}$. The presence of spatial dependence complicates model selection and estimation of parameter uncertainty. This is the subject of current research (see, for example Hallin et al. (2010)). As we will see in section 3.2.2, if the probability of threshold exceedance is constant, the form of the point process model fitted implies a particular form for the quantile regression model used to set the threshold. Therefore, for a given point process model, we know which quantile regression model to fit, but subsequent inferences neglect the uncertainty due to the estimation of $\boldsymbol{\beta}$. In section 3.2.3 we present the results of fitting model (4) to estimate a threshold for extreme value modelling.

3.2.2 Parameterisation of covariate effects

The assumption that the conditional probability of threshold exceedance is constant across all observations has consequences for the parameterisation of the effects of the covariates on the extreme value parameters. Let $Y(\mathbf{x}_{ij})$ denote the response, and $u(\mathbf{x}_{ij})$ the threshold applied, for covariate \mathbf{x}_{ij} . The probability of exceedance $p(\mathbf{x}_{ij}) = P(Y(\mathbf{x}_{ij}) > u(\mathbf{x}_{ij}))$ for covariate \mathbf{x}_{ij} is related approximately to the GEV parameters via

$$p(\mathbf{x}_{ij}) \approx \frac{1}{\lambda} \left[1 + \xi(\mathbf{x}_{ij}) \left(\frac{u(\mathbf{x}_{ij}) - \mu(\mathbf{x}_{ij})}{\sigma(\mathbf{x}_{ij})} \right) \right]^{-1/\xi(\mathbf{x}_{ij})}. \quad (5)$$

Suppose that $\xi(\mathbf{x}_{ij}) = \xi$ is constant. If we use quantile regression to set $u(\mathbf{x}_{ij})$ so that $p(\mathbf{x}_{ij}) = p$ is constant then, inverting (5),

$$u(\mathbf{x}_{ij}) = \mu(\mathbf{x}_{ij}) + c \sigma(\mathbf{x}_{ij}), \quad (6)$$

where $c = [(\lambda p)^{-\xi} - 1]/\xi$.

Suppose that we have set $u(\mathbf{x}_{ij})$ using a linear quantile regression on \mathbf{x}_{ij} . To have the correct functional form on the right hand side of (6) we should specify $\mu(\mathbf{x}_{ij})$ and $\sigma(\mathbf{x}_{ij})$ to be linear in \mathbf{x}_{ij} . For example, if $\mu(\mathbf{x}_{ij})$ and/or $\sigma(\mathbf{x}_{ij})$ are quadratic in longitude and latitude then quantile regression model (4) is indicated. A potential problem is that $\sigma(\mathbf{x}_{ij})$ is not constrained to be positive. However, this has not been a problem for the dataset analysed in this paper. Taking $\log \sigma(\mathbf{x}_{ij})$ to be linear in \mathbf{x}_{ij} , to ensure positivity of $\sigma(\mathbf{x}_{ij})$, results in a model that is inconsistent with the quantile regression used to set the threshold. If we wish to use a log link to relate $\sigma(\mathbf{x}_{ij})$ to the covariates, we should use non-linear quantile regression (Koenker and Park, 1994) to set $u(\mathbf{x}_{ij})$.

3.2.3 Threshold selection

Since we set the threshold $u(\mathbf{x}_{ij})$ using quantile regression to achieve constant exceedance probability p , choosing the level of the threshold amounts to choosing a value for p . In the presence of covariates the standard approach to threshold selection is to fit covariates model(s) over a range of values of p and look for stability in the parameter estimates. We choose the largest value of p above which, taking into account the uncertainty in the estimates summarised by 95% confidence intervals, the estimates appear approximately constant. For this purpose, we work with the point process model suggested by the preliminary GEV analysis, that is, a model in which $\mu(\mathbf{x}_{ij})$ is quadratic in longitude and latitude and σ and ξ are constant. Having identified a suitable threshold p , we then revisit this choice of exceedance model.

Figure 5, for different thresholds p , shows MLEs for this initial point process model. Inevitably there is some subjectivity in the choice of p . Here we choose $p = 0.4$. This is a relatively large value but we should bear in mind that the raw data contain only the largest values from each hurricane event. Indeed Jonathan and Ewans (2007) use a constant threshold of 2.5m for these data, which equates to an exceedance probability of 0.54.

If $\mu(\mathbf{x}_{ij})$ is quadratic in \mathbf{x}_{ij} then we should set the threshold using quantile regression model 4, using $\tau = 0.6$. The resulting threshold

$$\hat{u}_{QR} = 3.310 + 0.106 \phi_1(l_x) - 0.153 \phi_1(l_y) - 0.081 \phi_2(l_x) + 0.098 \phi_1(l_x) \phi_1(l_y) - 0.025 \phi_2(l_y), \quad (7)$$

varies between 2.85m and 3.48m and produces 9075 threshold exceedances out of 22680 observations. The number of exceedances per site varies between 123 and 130, and shows no obvious systematic variation with longitude and latitude.

3.3 Modelling spatial variation

Having set a suitable threshold, we now aim to select an appropriate point process model for threshold exceedances. Firstly, we consider exceedance models in which the location parameter $\mu(\mathbf{x}_{ij})$ is modelled as a Legendre polynomial function of longitude and latitude and the scale σ and shape parameters are constant. The quadratic representation of $\mu(\mathbf{x}_{ij})$ is consistent with (6). The results in table 1 suggest that we use a model in which $\mu(\mathbf{x}_{ij})$ is quadratic in longitude and latitude. The effect of adjusting the independence log-likelihood for spatial dependence is to make the regression effects in μ more statistically significant. For example, when comparing the quadratic model to the linear model, the unadjusted LR statistic is 11.00, compared to the ALRS of 20.50. This is analogous to the greater power possessed by a paired test compared to an unpaired two-sample t -test. Extension to include a quadratic form for $\sigma(\mathbf{x}_{ij})$

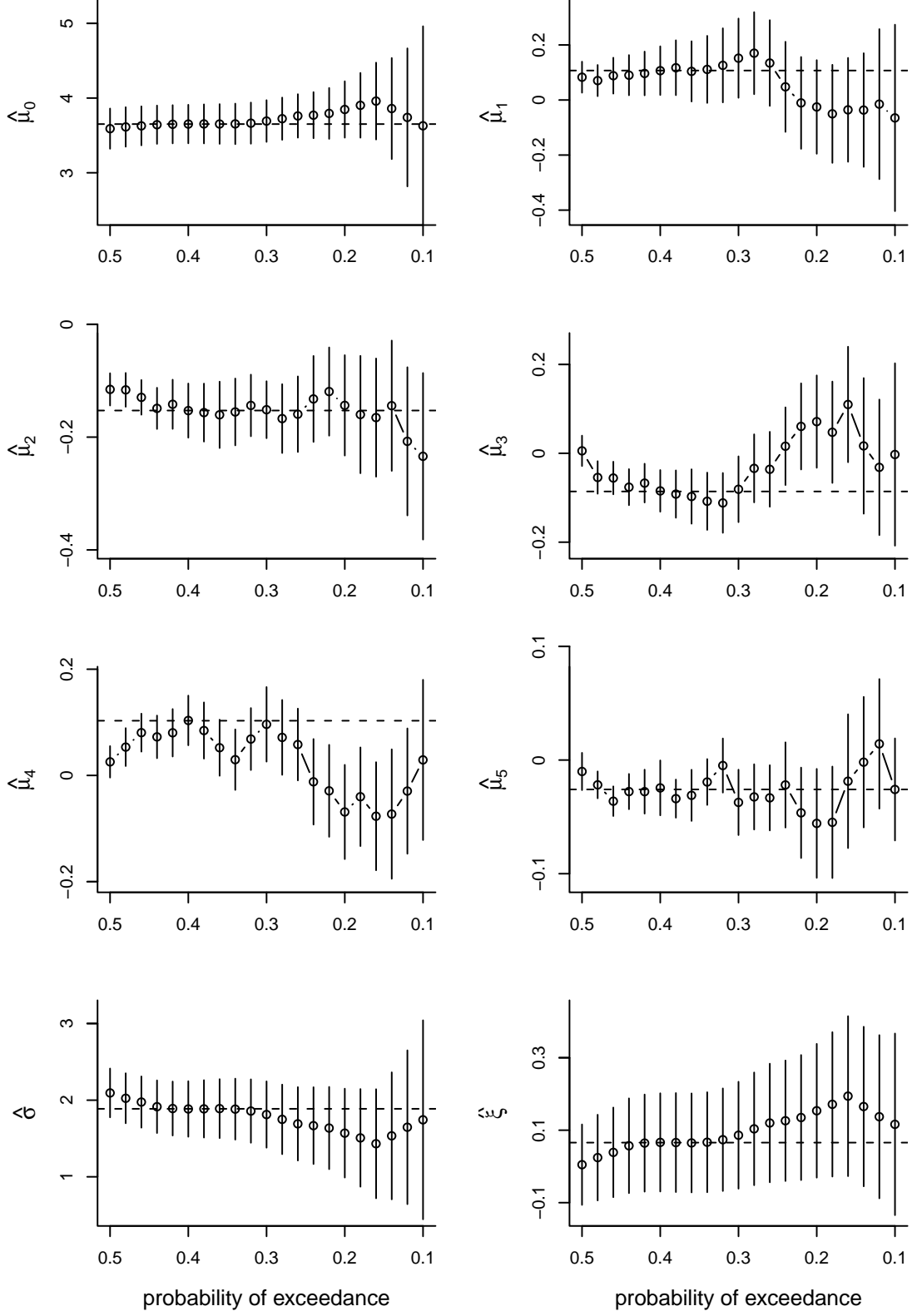


Figure 5: MLEs against probability of exceedance p for a point process model in which μ is quadratic in longitude and latitude. The vertical lines indicate approximate (symmetric) 95% confidence intervals, based on adjusted standard errors. The threshold, also quadratic in longitude and latitude, is estimated using quantile regression with exceedance probability p . The horizontal dashed lines indicate the MLEs for $p = 0.4$.

$\mu(\mathbf{x}_{ij})$	neg. log-lik	d.f.	ALRS	p -value
constant	22763.20			
linear	22742.59	2	34.23	3.7×10^{-8}
quadratic	22737.09	3	20.50	1.3×10^{-4}
cubic	22737.02	4	2.09	0.72

Table 1: Summary of point process modelling in which the location parameter μ is modelled as a Legendre polynomial function of longitude and latitude and σ and ξ are constant. The likelihood ratio tests compare the model with the model in the row above. d.f. = degrees of freedom; neg. log-lik = negated maximised log-likelihood.

was supported by an ALRS (p -value 0.10), but the resulting fits made little difference to the diagnostic plots presented in section 3.4. Therefore, we proceed with the model in which $\sigma(\mathbf{x}_{ij})$ is constant. The estimates and adjusted standard errors are given in table 2. The adjusted standard errors for $\hat{\mu}_0, \hat{\sigma}$ and $\hat{\xi}$ are much larger than their unadjusted versions of 0.022, 0.031 and 0.011 respectively. We found that the additional complexity of allowing the shape parameter ξ to vary spatially seems not to be warranted (for example, extension to include a quadratic form for $\xi(\mathbf{x}_{ij})$ produced an ALRS with an associated p -value of 0.30). We have checked that these conclusions are not sensitive to the choice of exceedance probability. An approximate 95%

	$\hat{\mu}_0$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$	$\hat{\mu}_5$	$\hat{\sigma}$	$\hat{\xi}$
MLE	3.652	0.107	-0.153	-0.085	0.104	-0.024	1.885	0.066
adjusted s.e.	(0.132)	(0.045)	(0.025)	(0.024)	(0.024)	(0.012)	(0.185)	(0.069)

Table 2: Maximum likelihood estimates and adjusted standard errors for a point process model in which μ is quadratic in longitude and latitude and σ and ξ are constant. The threshold is quadratic in longitude and latitude.

confidence interval for ξ , calculated by profiling the adjusted log-likelihood, is $(-0.052, 0.223)$. Jonathan and Ewans (2007) obtained a point estimate of ξ of -0.098 (compared to 0.066 in the current analysis) and a bootstrap 95% confidence interval of $(-0.164, 0.015)$ using the same raw data. The latter confidence interval is narrower at least partly because the threshold is set lower than in the current analysis. In addition Jonathan and Ewans (2007) use a constant threshold and do not model spatial variation. The differences in the inferences about ξ may not seem large in magnitude but we see in section 4 that, because positive values of ξ produce return levels q_N that increase quickly as return period N increases, they translate to much larger differences in inferences about return levels.

3.4 Model checking

We check three aspects of the model fit: the compatability of the threshold with the fitted threshold exceedance model, where threshold exceedances occur and the values of threshold exceedances.

3.4.1 Compatability of the threshold with the exceedance model

In section 3.2.2 we observed that if the threshold $u(\mathbf{x}_{ij})$ is set so that the probability p of exceedance is constant for all observations then

$$u(\mathbf{x}_{ij}) = \mu(\mathbf{x}_{ij}) + c\sigma, \quad (8)$$

where $c = [(\lambda p)^{-\xi} - 1]/\xi$. Based on the threshold diagnostic plots in section 3.2.3 we chose $p = 0.4$. We substitute the MLEs given in table 2 into equation (8) to obtain a threshold \hat{u}_{PP} implied by the fitted point process model. We would hope that \hat{u}_{PP} is similar to the threshold \hat{u}_{QR} estimated using quantile regression given in equation (7). Since we have used the same covariates in the quantile regression and for $\mu(\mathbf{x}_{ij})$ in the point process model, $\hat{u}_{PP} = \hat{u}_{QR}$ if $\hat{\mu}_0 + \hat{c}\hat{\sigma} = \hat{\beta}_0$ and $\hat{\mu}_i = \hat{\beta}_i, i = 1, \dots, 5$. Comparing equation (7) and table 2, and noting that $\hat{\mu}_0 + \hat{c}\hat{\sigma} = 3.310$, shows that this is approximately the case.

We treat the threshold \hat{u}_{QR} as fixed and examine whether u_{PP} is significantly different from \hat{u}_{QR} . An adjusted likelihood ratio test of $\mu_0 + c\sigma = \hat{\beta}_0, \mu_i = \hat{\beta}_i, i = 1, \dots, 5$ produces an ALRS of 0.07 on 6 d.f. and a p -value of 0.99999. This suggests that the threshold \hat{u}_{QR} is compatible with the inferences made about the parameters of the point process model, based on this threshold. If there are covariate effects in both $\mu(\mathbf{x}_{ij})$ and $\sigma(\mathbf{x}_{ij})$ a similar comparison and test can be constructed based on equation (6). The effect on this test of not treating \hat{u}_{QR} as fixed is not clear cut. Accounting for the uncertainty in β will tend to increase the p -value, but accounting for the positive association between \hat{u}_{QR} and \hat{u}_{PP} - resulting from the fact that they are estimated from (different aspects of) the same data - will tend to reduce the p -value.

The comparison of u_{PP} to \hat{u}_{QR} , could supplement the methods used in section 3.2.3 to select the level of the threshold, that is, the value of p . Over a range of values of p , we could inspect the p -value from the test of $u_{PP} = \hat{u}_{QR}$, and perhaps compare graphically the closeness of \hat{u}_{QR} to \hat{u}_{PP} , to suggest values of p for which the threshold set is consistent with the extreme value model fitted to the data.

3.4.2 Occurrence of threshold exceedances

In section 3.2.3 we observed that the number of threshold exceedances does not seem to vary systematically across the sites. We have assumed that data from different hurricane events are independent. Therefore, at each site, the numbers of events between successive threshold exceedances should be geometrically distributed with mean $1/0.4$. Performing Pearson chi-squared goodness-of-fit tests at each of the sites results in p -values that vary from 0.18 to 0.87. Due to the strong dependence between the time series at the sites, we expect these p -values to vary less than a random sample from a standard uniform distribution. The p -values show no obvious systematic variation with longitude and latitude.

We also check whether the rate of exceedance seems to vary systematically over time. Figure 6 shows how the number of exceedances increases with event number n_s for site 1 (longitude=1, latitude=1), with 95% envelopes based on a binomial($n_s, 0.4$) distribution. Due to a series of large events around event number 50, evident in figure 2, a few points lie above the envelope. Otherwise, this plot does not reveal any clear time trend. The plots for the other sites are very similar.

3.4.3 Values of threshold exceedances

We inspect individual QQ-plots of threshold exceedances at each of the 72 sites. Figure 7 shows QQ-plots from six of the sites. We plot $u + z_{(r)}$ against $u + \hat{F}^{-1}[(r - 1/3)/(n + 1/3)]$, where $z_{(1)}, \dots, z_{(n)}$ are the ordered threshold exceedances at the site in question, $\hat{F}(\cdot)$ is the c.d.f. of the fitted GP($\sigma_u, \xi(\mathbf{x}_{ij})$) distribution at this site and u is the threshold applied at the site. For this choice of plotting positions the line of equality is virtually coincident with the medians of the fitted distributions of the order statistics (Hyndman and Fan, 1996). We superimpose 95%

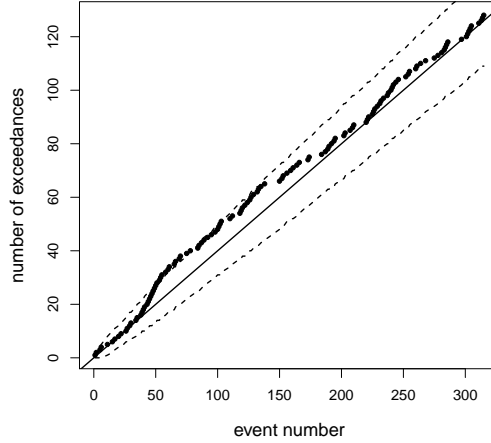


Figure 6: Number of exceedances against event number n_s for site 1. The solid line gives the expected number ($0.4n_s$) of exceedances under the fitted model. The dashed lines give the 97.5% and 2.5% quantiles of the binomial($n_s, 0.4$) distribution.

envelopes on the plots. The envelopes have been adjusted for the uncertainty in estimating the GEV parameters, using simulation from the estimated normal distribution in (3). The strong spatial dependence means that neighbouring sites have very similar QQ plots. None of the QQ-plots reveal any clear lack-of-fit.

4 Return level estimation

The solid lines in figure 8 show how the estimates, and 95% confidence intervals, of the 500 year return level q_{500} vary with site longitude and latitude under the fitted regression model. The dashed lines give the corresponding values obtained by fitted the point process model individually at each of the sites, using the same threshold values set for the regression model. The uncertainty in estimating the threshold, and setting its general level, is neglected. Figure 8 shows the main advantage of pooling information over sites: typically the confidence intervals are narrower under the regression model than the at-site intervals. Exceptions to this occur where the at-site estimate of ξ is negative, because smaller estimates of ξ tend to produce narrower confidence intervals for large return levels. For these data Jonathan and Ewans (2007) give an estimate of 13.0m and bootstrap 95% confidence interval of (11.5, 16.3)m for the 100 year return q_{100} . These values apply to all sites because spatial variation is not modelled. The larger point estimate of, and slightly greater uncertainty about, ξ in the current analysis produces point estimates and 95% confidence intervals for q_{100} , that vary across the sites, but are of the order of 13.8m and (11.6, 18.4)m.

5 Discussion

We have argued that when modelling threshold exceedances from a non-stationary process it seems sensible to set a threshold that depends on the values of important covariates, by aiming to achieve a constant probability p of exceedance. In section 1.1 we gave several reasons to support quantile regression as the natural way to achieve this. If the covariates take a small number of distinct levels (for example, covariates of season, year, spatial location) then it is

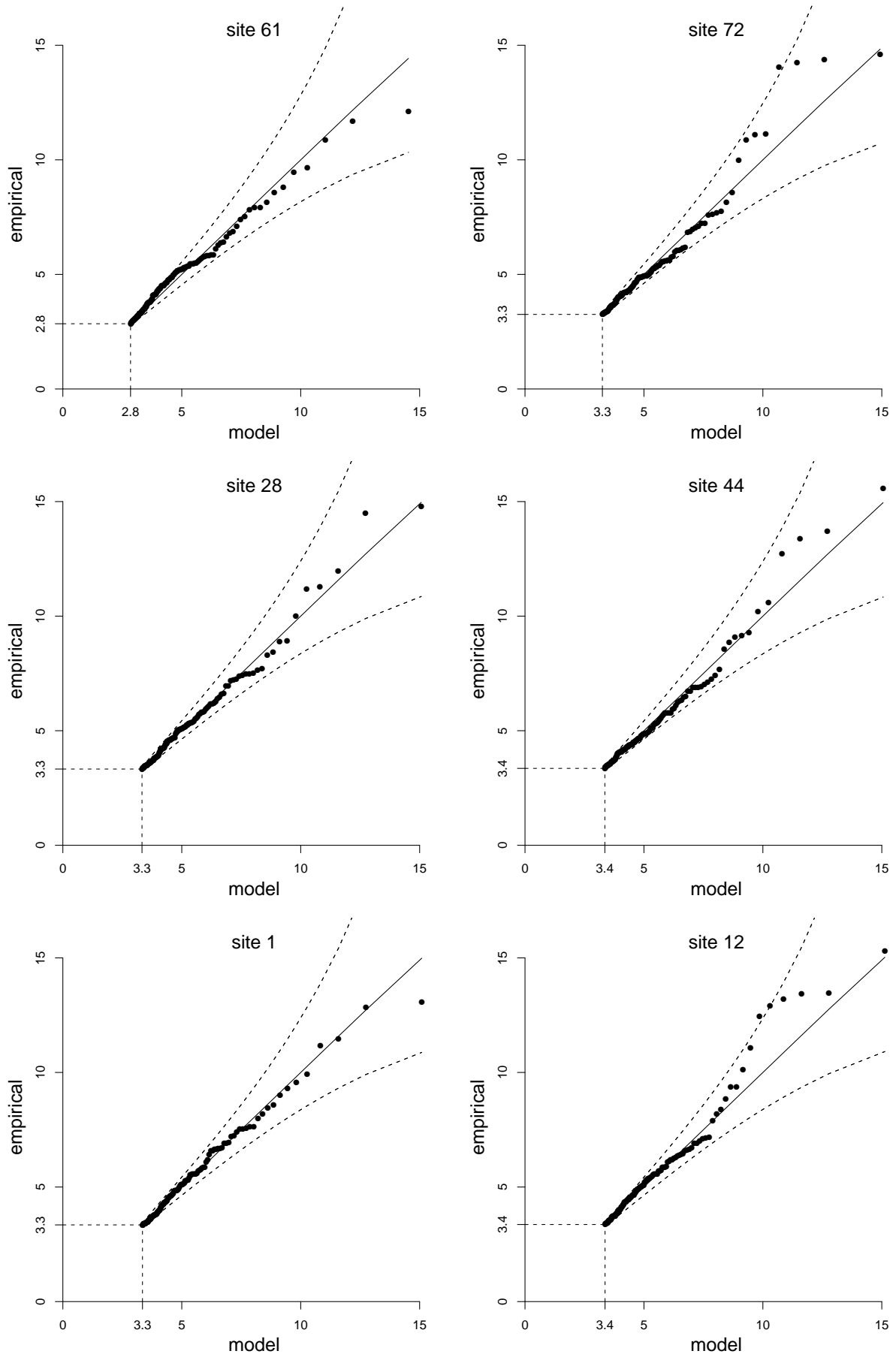


Figure 7: GP QQ-plots at ‘corner’ sites 1, 12, 61 and 72, and two sites, 28 and 44, nearer the centre of the region. The dashed curves give the 2.5% and 97.5% quantiles of the fitted distributions of the order statistics. The threshold is shown using dashed lines.

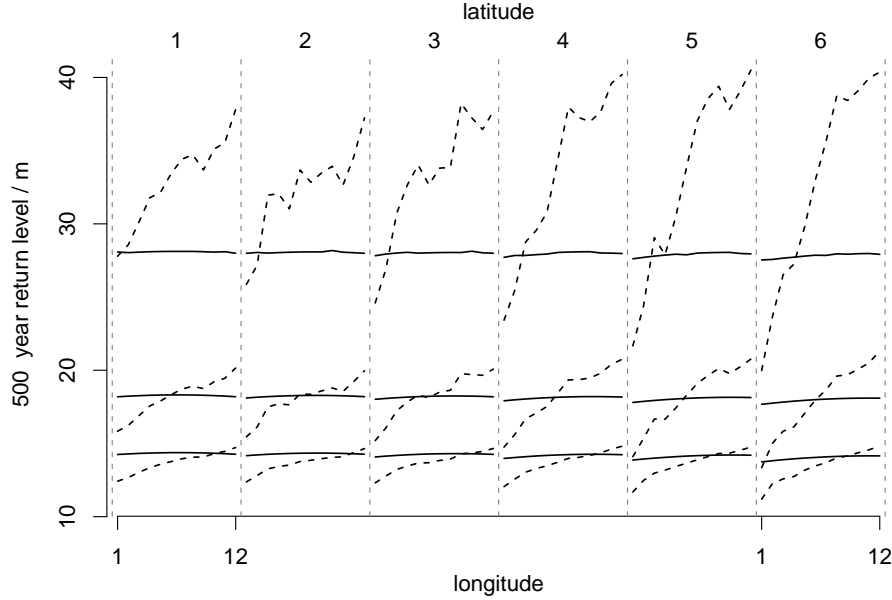


Figure 8: Point estimates and 95% confidence intervals for 500 year return levels, by longitude and latitude, based on the regression model (solid lines) and at-site estimates (dashed lines). The 95% confidence intervals are produced by profiling the (adjusted) log-likelihood.

possible to set a separate threshold for each possible combination of the covariates. However, this may result in over-parameterisation of the threshold. Quantile regression provides a means of reducing the number of parameters to be estimated and producing a more realistic threshold. In particular, if extreme value parameters are modelled as a smooth function of space then it makes sense to approach the setting of a threshold in the same way.

If a constant threshold is set then it is common to model $p(\mathbf{x})$ as a function of covariates \mathbf{x} , using, for example, logistic regression. We argue that it is more logical to model the threshold for constant p than to set a constant the threshold and model how $p(\mathbf{x})$ varies with \mathbf{x} . If the covariate effects are strong then $p(\mathbf{x})$ may vary greatly across different combinations of the covariates. In this event, a logistic regression will be detecting covariate effects on both extreme and non-extreme responses, and will tend to include more covariates than a quantile regression with p set to a single value.

In this paper we have considered the point process model but the general approach of setting a covariate-dependent threshold using quantile regression could be used for the binomial-GP model. If there is a need to interpolate (perhaps between sites in space) or to extrapolate (perhaps in time), then, under the binomial-GP model an estimate of the threshold is required at the values of the covariates not used in the estimation. Quantile regression would provide this, but additional estimation would be required to do this if a separate threshold was set, for example, separately at different spatial locations.

In section 3.2.2 we considered how the quantile regression used to set a threshold should be related to the subsequent point process model for threshold exceedances. This means that we do not need to select empirically the covariates to be included in the quantile regression. As detailed in section 3.4.1 this relationship also suggests a new diagnostic for threshold selection, based on the closeness of the agreement between the threshold set using quantile regression and the threshold implied by the point process model.

We have also investigated the possibility of setting a threshold that agrees exactly with the resulting parameter estimates of the point process model. Suppose that, for a given value of p ,

we set a threshold by substituting point process model parameters $\theta^{(in)}$ into equation 8. Let $\theta^{(out)}$ be the MLEs that result when the point process model is fitted. We wish to find $\theta^{(in)}$ for which $\theta^{(out)} = \theta^{(in)}$. For the model summarised in table 2, we find that this is not possible, perhaps due to the presence of ties in the data. However, starting from the estimates in table 2, very close agreement is possible with $\theta^{(out)}$ differing from $\theta^{(in)}$ by at most 0.75%. Thus, for a given value of p , we could try to set a suitable covariate-dependent threshold without using quantile regression.

For the hindcast data additional model diagnostics (not shown) are not strongly suggestive of a time trend in the location of the data. Adding polynomial terms up to order 3 in event number for μ in the point process model does not improve significantly the fit of the model. Recent research by Killick et al. (2009) considers whether changes in the variability of storm peak significant wave height occur over time in these hindcast data.

Underlying the likelihood (2) is a Poisson(np) approximation to the binomial(n, p) distribution of the total number of exceedances, under the assumption that n is sufficiently large and p is sufficiently small. Since the exceedance probability of 0.4 chosen in section 3.2.3 is relatively large, we checked that it is reasonable to make this approximation for the current data, by refitting the final model (quadratic threshold, μ quadratic in longitude and latitude, σ and ξ constant) without making this approximation. The resulting point estimates are extremely close to the original estimates, certainly close enough to make no material difference to the general conclusions.

The approach described in this paper can be used to model richer sources of significant wave height data, containing information on wave direction (Jonathan et al., 2008) and season, or for a larger number of sites covering a wider spatial domain. The estimated spatial effects are statistically significant but relatively small. These effects are anticipated to be of greater magnitude over wider spatial domains, so it is important that spatial effects can be incorporated in the modelling. The precision of estimates could further be improved by working with the raw, 30 minute interval data, rather than the storm peaks. The approach described in this paper can deal with these extensions. In particular, appropriate estimation of parameter uncertainty in the presence of spatial and temporal dependence, is computable using the methodology of section 3.1, whereas the use of bootstrapping for this purpose may become much more onerous.

Acknowledgements

The authors acknowledge discussions with Kevin Ewans (Shell) and Vince Cardone (Oceanweather) and thank Richard Chandler (UCL) for his extremely helpful comments on a first draft of this paper.

References

- Abramowitz, M. and I. Stegun (1965). *Handbook of Mathematical Functions: With Formulas, Graphs and Mathematical Tables*. New York: Dover.
- Butler, A., J. E. Heffernan, J. A. Tawn, and R. A. Flather (2007). Trend estimation in extremes of synthetic North Sea surges. *Appl. Statist.* 56(4), 395–414.
- Casson, E. and S. Coles (1999). Spatial regression models for extremes. *Extremes* 1, 449–468.

- Chandler, R. E. (2005). On the use of generalized linear models for interpreting climate variability. *Environmetrics* 16, 699–715.
- Chandler, R. E. and S. B. Bate (2007). Inference for clustered data using the independence loglikelihood. *Biometrika* 94(1), 167–183.
- Chavez-Demoulin, V. and A. C. Davison (2005). Generalized additive modelling of sample extremes. *Appl. Statist.* 54(1), 207–222.
- Coles, S. G. (2001). *An introduction to statistical modeling of extreme values*. London: Springer.
- Cooley, D., D. Nychka, and P. Naveau (2007). Bayesian spatial modelling of extreme precipitation return levels. *J. Am. Stat. Assoc.* 102(479), 824–840.
- Davison, A. C. and D. A. Hinkley (1997). *Bootstrap methods and their application*. Cambridge: Cambridge University Press.
- Davison, A. C. and R. L. Smith (1990). Models for exceedances over high thresholds (with discussion). *J. R. Statist. Soc. B* 52, 393–442.
- Eastoe, E. F. and J. A. Tawn (2009). Modelling non-stationary extremes with application to surface level ozone. *Appl. Statist.* 58, 25–45.
- Fawcett, L. and D. Walshaw (2007). Improved estimation for temporally clustered extremes. *Environmetrics* 18, 173–188.
- Fisher, R. A. and L. H. C. Tippett (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Camb. Phil. Soc.* 24, 180–190.
- Hallin, M., Z. Lu, and K. Yu (2010). Local linear spatial quantile regression. *Bernoulli*.
- Heffernan, J. E. and J. A. Tawn (2004). A conditional approach for multivariate extremes. *J. Roy. Statist. Soc. B*, 66, 497–546.
- Hyndman, R. J. and Y. Fan (1996). Sample quantiles in statistical packages. *American Statistician* 50, 361–365.
- Jagger, T. H. and J. B. Elsner (2009). Modeling tropical cyclone intensity using quantile regression. *International Journal of Climatology* 29, 1351–1361.
- Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) of meteorological events. *Q. J. R. Meteorol. Soc.* 81, 158–171.
- Jonathan, P. and K. C. Ewans (2007). Uncertainties in wave height estimates for hurricane-dominated regions. *J. Offshore Mechanics Arctic Engineering* 129(4), 300–305.
- Jonathan, P. and K. C. Ewans (2009). A spatio-directional model for extreme waves in the Gulf of Mexico. In *Proceedings of OMAE 2009: the 28th International Conference on Offshore Mechanics and Arctic Engineering*, 31 May – 4 June 2009, Honolulu, U.S.A.
- Jonathan, P., K. C. Ewans, and G. Z. Forristall (2008). Statistical estimation of extreme ocean environments: the requirement for modelling directionality and other covariate effects. *Ocean Engineering* 35, 1211–1225.
- Killick, R., I. A. Eckley, K. Ewans, and P. Jonathan (2009). Detection of changes in the characteristics of oceanographic time-series using changepoint analysis. (Preprint. <http://www.lancaster.ac.uk/~jonathan/Kll09.pdf>).

- Koenker, R. (2005). *Quantile regression*. Cambridge: Cambridge University Press.
- Koenker, R. (2009). *quantreg: Quantile Regression*. R package version 4.44.
- Koenker, R. and G. Bassett (1978). Regression quantiles. *Econometrica* 46(1), 33–50.
- Koenker, R. and B. Park (1994). An interior point algorithm for nonlinear quantile regression. *Journal of Econometrics* 71(1–2), 265–283.
- Naveau, P., A. Guillo, D. Cooley, and J. Diebolt (2009). Modelling pairwise dependence of maxima in space. *Biometrika* 96(1), 1–17.
- Oceanweather Inc. (2005). GOMOS – USA Gulf of Mexico Oceanographic Study, Northern Gulf of Mexico Archive, Oct.
- Pickands, J. (1971). The two-dimensional Poisson process and extremal processes. *J. App. Prob.* 8, 745–756.
- Pickands, J. (1975). Statistical inference using extreme order statistics. *Ann. Stat.* 3, 119–131.
- R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.
- Smith, R. L. (1985). Maximum likelihood estimation in a class of non-regular cases. *Biometrika* 72, 67–92.
- Smith, R. L. (1989). Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone. *Statistical Science* 4, 367–377.
- Smith, R. L. (1990). Regional estimation from spatially dependent data. (Preprint. <http://www.stat.unc.edu/postscript/rs/regest.pdf>).
- Wadsworth, J. L., J. A. Tawn, and P. Jonathan (2010). Accounting for choice of measurements scale in extreme value modeling. (Preprint. <http://www.lancaster.ac.uk/~jonathan/WdsTwnJnt09.pdf>).