



**Title:** Spatial Bayesian hierarchical modelling of extreme sea states

**Author(s):** Colm Clancy<sup>1,2,4</sup>, John O'Sullivan<sup>1,2</sup>, Conor Sweeney<sup>1,2</sup>, Frédéric Dias<sup>1,2,3</sup>, Andrew C. Parnell.<sup>1,2</sup>

1. School of Mathematics and Statistics, University College Dublin, Ireland
2. UCD Earth Institute, University College Dublin, Ireland
3. CMLA, ENS Cachan, CNRS, Université Paris-Saclay, 94235 Cachan, France
4. Met Éireann, Glasnevin Hill, Dublin 9, Ireland

This article is provided by the author(s) and Met Éireann in accordance with publisher policies. Please cite the published version.

**NOTICE:** This is the author's version of a work that was accepted for publication in *Ocean Modelling*. Changes resulting from the publishing process such as editing, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *Ocean Modelling*, 107, pp.1-13.

**Citation:** Clancy, Colm, John O'Sullivan, Conor Sweeney, Frédéric Dias, and Andrew C. Parnell. "Spatial Bayesian Hierarchical Modelling of Extreme Sea States." *Ocean Modelling* 107 (2016): 1–13. doi: [10.1016/j.ocemod.2016.09.015](https://doi.org/10.1016/j.ocemod.2016.09.015).

This item is made available to you under the Creative Commons Attribution-Non commercial-No Derivatives 3.0 License.



# Spatial Bayesian hierarchical modelling of extreme sea states

Colm Clancy<sup>a,b</sup>, John O’Sullivan<sup>a,b,\*</sup>, Conor Sweeney<sup>a,b</sup>, Frédéric Dias<sup>a,b,c</sup>,  
Andrew C. Parnell<sup>a,b</sup>

<sup>a</sup>*School of Mathematics and Statistics, University College Dublin, Ireland*

<sup>b</sup>*UCD Earth Institute, University College Dublin, Ireland*

<sup>c</sup>*CMLA, ENS Cachan, CNRS, Université Paris-Saclay, 94235 Cachan, France*

---

## Abstract

A Bayesian hierarchical framework is used to model extreme sea states, incorporating a latent spatial process to more effectively capture the spatial variation of the extremes. The model is applied to a 34-year hindcast of significant wave height off the west coast of Ireland. The generalised Pareto distribution is fitted to declustered peaks over a threshold given by the 99.8th percentile of the data. Return levels of significant wave height are computed and compared against those from a model based on the commonly-used maximum likelihood inference method. The Bayesian spatial model produces smoother maps of return levels. Furthermore, this approach greatly reduces the uncertainty in the estimates, thus providing information on extremes which is more useful for practical applications.

*Keywords:* Bayesian hierarchical modelling, spatial modelling, extreme value analysis, ocean waves, significant wave height

---

## 1. Introduction

A detailed knowledge of the extreme sea states affecting a region is essential for any marine activity. For shipping, offshore and coastal installations, or the

---

\*Corresponding author at: School of Mathematics and Statistics, University College Dublin, Belfield, Dublin 4, Ireland.

*Email address:* john.o-sullivan.2@ucdconnect.ie (John O’Sullivan)

deployment of devices such as wave energy converters, it is crucial to have accurate information on the extremes likely to be encountered during operational lifetimes. These are typically expressed in terms of return levels and periods; for example, the level of significant wave height which is likely to occur on average once every 100 years. Extreme value theory provides statistical tools for such an analysis (Coles, 2001) and the methods have been widely applied in studies of ocean waves; reviews may be found in Vanem (2011) and Jonathan and Ewans (2013). The background theory for this extreme value analysis is outlined in Section 2 below.

Models of extremes are often fitted to data-sets using a maximum likelihood approach. Although straightforward to implement, this can lead to large uncertainties in the parameter estimations and subsequent return levels (Vanem, 2015). Obviously, we wish to reduce the levels of uncertainty and obtain meaningful results which are of practical use. Bayesian inference allows for a more detailed analysis of this uncertainty, by providing complete probability distributions for the parameters given the data (Gelman et al., 2013).

Our aim in this paper is to use Bayesian techniques to model the spatial variability of ocean wave extremes. We follow the approach of Cooley et al. (2007), who include a latent spatial process within a Bayesian hierarchical framework to capture the spatial dependence of precipitation extremes. This is described in detail in Section 3. Such a model has been applied to the study of temperature extremes in the ocean by Oliver et al. (2014) but not to ocean wave data, to the best of the authors' knowledge.

We apply the statistical model to significant wave height data off the west coast of Ireland, obtained from a spectral wave model hindcast. Recently, O'Brien et al. (2013) provided a history of extreme wave events around Ireland, revealing an often severe environment. On the other hand, the seas off the west coast of Ireland have attracted interest due to their potential wave energy resources (Gallagher et al., 2016) and so an accurate description of the likely extremes is of both theoretical and practical relevance.

A description of the domain and data under study, along with model imple-

35 mentation details, is given in Section 4. The results are presented in Section 5  
with a discussion and conclusions in Section 6.

## 2. Extreme Value Analysis

### 2.1. Background theory

There are a number of possible approaches to extreme value analysis. An  
40 introduction to the field may be found in Coles (2001). One fundamental  
method is the block maxima approach. We consider a sequence of independent  
and identically-distributed random variables,  $Z_1, Z_2, \dots$ , and let  $M_n =$   
 $\max(Z_1, \dots, Z_n)$  be the maximum over a block of  $n$  values; for example, we  
may take  $M_n$  to be the annual maxima in a multi-year set of significant wave  
45 height data. The extremal types theorem states that, under certain regularity  
conditions, the distribution function of the  $M_n$  will converge to a specific three-  
parameter form, known as the generalised extreme value (GEV) distribution.

A major disadvantage to this approach is the fact that, by using only the  
maxima from a given block size, we are discarding a lot of data. In this work  
50 we consider a data-set of hourly significant wave height,  $H_s$ . Modelling with,  
for example, annual maxima would be quite wasteful. An alternative is to  
model the excesses over a given threshold (Pickands, 1975). We assume that  
our sequence of independent random variables,  $Z_1, Z_2, \dots$ , satisfies the extremal  
types theorem described above. For large enough threshold  $u$ , the distribution  
55 function of the exceedances  $Y = Z - u$ , conditional on  $Z > u$ , is approximately  
given by the generalised Pareto distribution (GPD)

$$F(y) = 1 - \left(1 + \frac{\xi y}{\sigma}\right)^{-1/\xi} \quad (1)$$

defined on the set  $\{y : y > 0 \text{ and } (1 + \xi y/\sigma) > 0\}$ . Here,  $\xi$  and  $\sigma$  are known as  
the shape and scale parameters, respectively, and have ranges  $-\infty < \xi < \infty$  and  
 $\sigma > 0$ . For the limiting value when  $\xi = 0$ , we get the exponential distribution

$$F(y) = 1 - \exp\left(-\frac{y}{\sigma}\right)$$



These two methods of extreme value analysis have been applied extensively to ocean wave data from different sources. Examples of GEV models include Menéndez et al. (2009), who use monthly maxima of  $H_s$  from observational buoy data, and Izaguirre et al. (2011), in which monthly maxima are obtained from satellite altimeter missions. Threshold exceedance models of  $H_s$  with the GPD may be found in Caires and Sterl (2005), Vinoth and Young (2011) and Nicolae Lerma et al. (2015). In addition, a number of papers have compared the various approaches; see, for example, Caires (2011), Vinoth and Young (2011), Aarnes et al. (2012), Vanem (2015) and Clancy et al. (2015).

Once we have the parameters of a distribution, we may compute the  $N$ -year return levels. For the GPD in (1), we have

$$P(Z > z | Z > u) = \left(1 + \frac{\xi(z - u)}{\sigma}\right)^{-1/\xi}. \quad (2)$$

We write  $\zeta_u = P(Z > u)$  and can then find the return level  $z_m$ , the level which is exceeded on average once every  $m$  observations, by solving

$$P(Z > z_m) = \zeta_u \left(1 + \frac{\xi(z_m - u)}{\sigma}\right)^{-1/\xi} = \frac{1}{m}.$$

Letting  $m = N n_y$ , where  $n_y$  is the number of observations per year, we arrive at the following expression for the  $N$ -year return level:

$$z_N = u + \frac{\sigma}{\xi} \left[ (N n_y \zeta_u)^\xi - 1 \right] \quad (3)$$

For the case of the exponential distribution with  $\xi = 0$ , we have

$$z_N = u + \sigma \log(N n_y \zeta_u)$$

## 70 2.2. Model fitting

Given a set of data, we may fit one of the models described above. The maximum likelihood (ML) estimation method is commonly used. We can consider a set of  $n$  independent values,  $z_1, \dots, z_n$ , to which we wish to fit a probability density function  $f(z; \theta)$ , where  $\theta$  is a parameter of the distribution. The likelihood function is given by

$$L(\theta) = f(z|\theta) = \prod_{i=1}^n f(z_i; \theta)$$

The maximum likelihood estimator  $\hat{\theta}$  is found by maximising the above likelihood function or, more usually, the logarithm of  $L(\theta)$ . Asymptotic properties of the ML estimate, which assume Gaussian behaviour, may then be used to compute confidence intervals. Furthermore, the so-called delta method provides  
75 confidence intervals for quantities derived from the parameter estimates; for example, the return levels in (3). Details of these are given in Coles (2001), along with a discussion of other methods for fitting and analysing uncertainty, such as the profile likelihood method.

A further alternative is to use Bayesian inference for parameter estimation  
80 (Gelman et al., 2013). Continuing the above example, we use Bayes' Theorem to write

$$f(\theta|z) \propto f(z|\theta) f(\theta) \quad (4)$$

Thus, we arrive at a *posterior distribution*,  $f(\theta|z)$ , from a combination of the likelihood of the data and a given prior distribution  $f(\theta)$ . Whereas the ML method gives a point estimate of a parameter, with the Bayesian approach the  
85 parameter is described by a complete distribution. This allows us to characterise the uncertainty in a natural way. Rather than appealing to asymptotic theory for confidence intervals, we may use, for example, the percentiles of the posterior distribution.

A detailed treatment of Bayesian methods may be found in Gelman et al.  
90 al. (2013). Coles (2001) provides a brief introduction to their use in extreme value analysis while Coles et al. (2003) further discuss their benefits over likelihood-based inference in modelling extremes. In the context of ocean wave modelling, Egozcue et al. (2005) and Scotto and Guedes Soares (2007) were among the first to apply a Bayesian approach; see Vanem (2011) for a review of various  
95 models of ocean extremes. The review of Jonathan and Ewans (2013) points to the growing use of Bayesian methods and their potential for ocean engineering applications.

Practical implementation of Bayesian inference can be computationally intensive, in particular the calculation of the proportionality constants in (4).

100 The development of the Markov chain Monte Carlo (MCMC) technique has  
been hugely successful in making these methods viable. This algorithm may  
be used to draw simulated samples from the desired posterior distributions  
(Geyer, 2011).

### *2.3. Spatial modelling of extremes*

105 A number of authors have examined the spatial variation of extreme sea  
states, rather than focussing on one particular location. Fedele (2012), for ex-  
ample, considered space-time extremes of individual crest heights over a spatial  
region.

For studies involving extreme value modelling of significant wave height, the  
110 global or regional data-sets used have come from satellites (Vinoth and Young,  
2011; Izaguirre et al., 2011) or model hindcasts and reanalyses (Caires and  
Sterl, 2005; Cañellas et al., 2007; Aarnes et al., 2012; Guo and Sheng, 2015).  
This extreme value analysis has been carried out at each individual point on  
some given spatial grid. However, this approach does not explicitly model the  
115 spatial variations since each point is treated independently. This current work  
improves upon this by incorporating a spatial model within a Bayesian hierar-  
chical framework.

Previously, Vanem et al. (2012) used a Bayesian hierarchical structure in a  
spatio-temporal model of significant wave height data. Here, our specific inter-  
120 est is extremes. Cooley et al. (2007) presented a method for producing maps  
of extreme precipitation return levels in Colorado, using separate hierarchical  
models to model the intensity and frequency of events. Within both models,  
it was assumed that regional extreme precipitation is driven by a latent spatial  
process, defined by geographical and climatological covariates, and that effects  
125 not fully captured by the covariates are captured by the spatial structure in the  
hierarchies, using Gaussian processes. Inference was then conducted using an  
MCMC algorithm. This approach has since been used in oceanographic appli-  
cations by Oliver et al. (2014), to analyse extremes of sea surface temperatures.

In this current work, we apply a similar model to extreme significant wave

130 heights off the Irish west coast. This continues on from the work of Clancy  
 et al. (2015), in which this region was examined using extreme value analysis  
 applied independently at each point. In the next section we describe the model  
 in detail.

### 3. Model Details

135 The aim is to produce  $N$ -year return levels of significant wave height,  $H_s$ .  
 We described earlier how these could be computed using (3). Given a data-  
 set, we require a suitable threshold  $u$ , the parameters from the generalised  
 Pareto distribution (GPD) for modelling the exceedances, and their probability  
 of occurrence  $\zeta_u$ . The choice of threshold will be discussed later in Section 4.

140 For the exceedances and the probability  $\zeta_u$ , we follow the approach of Cooley  
 et al. (2007) and employ a Bayesian hierarchical model with three layers. The  
 first layer consists of linking the data to some parameters through the GPD.  
 The second describes the latent spatial process underlying the extremes in the  
 region while the third layer consists of the prior distributions on the parameters  
 145 controlling the second.

Using Bayes theorem, under a three-layer hierarchical model the inference  
 for the vector of parameters  $\theta_1$  (for the GPD of exceedances or the probabilities  
 $\zeta_u$ ) is given by

$$p(\theta_1|Z(x)) \propto p_1(Z(x)|\theta_1)p_2(\theta_1|\theta_2)p_3(\theta_2) \quad (5)$$

where the  $p_j$  are the probability densities with indices associated with the levels  
 150 of the hierarchy and  $Z(x)$  specifies the data at a given location  $x$ . We now  
 describe the two hierarchical models. A directed acyclic graph (DAG) depicting  
 the hierarchal structure of the models in detail is given in Fig. 1.

#### 3.1. Modelling the threshold exceedances

##### 3.1.1. Data layer

A GPD given by (2) is used to model the data at the first layer of the  
 hierarchy. To ensure a positive scale parameter throughout the computations,

we reparameterise with  $\phi = \log \sigma$ . At this level we thus have two spatially-varying parameters for the distribution, which we collectively write as  $\theta_1 = [\phi(x), \xi(x)]^T$ . The first term in the hierarchy (5) is then derived from the density function for the GPD and given by the likelihood function

$$p_1(Z(x)|\theta_1) = \prod_{i=1}^{n_x} \prod_{k=1}^{n_i} \frac{1}{\exp \phi(x_i)} \left( 1 + \frac{\xi(x_i) z_k(x_i)}{\exp \phi(x_i)} \right)^{-1-1/\xi(x_i)}$$

155 where the indices  $i$  and  $k$  are such that  $z_k(x_i)$  refers to the  $k$ -th exceedance at grid-point  $x_i$ . We have denoted the number of grid-points by  $n_x$  and the number of exceedances at each point  $x_i$  is then  $n_i$ .

### 3.1.2. Process layer

Both  $\phi(x)$  and  $\xi(x)$  are modelled as Gaussian processes (Banerjee et al., 2014) and so the second term in (5) will take the form

$$p_2(\theta_1|\theta_2) = p_\phi(\phi(x)|\mu_\phi, \Sigma_\phi) p_\xi(\xi(x)|\mu_\xi, \Sigma_\xi)$$

where

$$p_\phi(\phi(x)|\mu_\phi, \Sigma_\phi) = \frac{1}{\sqrt{(2\pi)^{n_x} |\Sigma_\phi|}} \exp \left[ -\frac{1}{2} (\phi - \mu_\phi)^T \Sigma_\phi^{-1} (\phi - \mu_\phi) \right].$$

A similar expression is used for  $p_\xi(\xi(x)|\mu_\xi, \Sigma_\xi)$ . Here  $|\cdot|$  denotes the determinant and  $\theta_2$  above represents all of the hyperparameters for  $\mu_\phi$ ,  $\mu_\xi$ ,  $\Sigma_\phi$  and  $\Sigma_\xi$ ,  
160 to be discussed below.

A Gaussian process characterises an infinite-dimensional smooth surface such that any finite collection of  $n_x$  points on the surface follows a multivariate normal distribution (above) of dimension  $n_x$ . Such a smooth surface is an appropriate  
165 choice for the model parameters as we expect similar wave climates at nearby locations.

In addition to distance, the effect of any other covariates may be readily incorporated into the model. For  $m$  covariates  $c^{(1)}, \dots, c^{(m)}$ , we write the mean vector in the general form

$$\mu_\phi = \mathcal{C} \alpha_\phi \tag{6}$$

where  $\mathcal{C}$  is the  $n_x \times (m + 1)$  matrix

$$\mathcal{C} = \begin{pmatrix} 1 & c_1^{(1)} & c_1^{(2)} & \dots & c_1^{(m)} \\ 1 & c_2^{(1)} & c_2^{(2)} & \dots & c_2^{(m)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & c_{n_x}^{(1)} & c_{n_x}^{(2)} & \dots & c_{n_x}^{(m)} \end{pmatrix}$$

170 and the vector of coefficients is  $\alpha_\phi = (\alpha_{\phi,0}, \alpha_{\phi,1}, \dots, \alpha_{\phi,m})^T$ . In this work we have used  $m = 3$ : the longitude, latitude and depth of a grid-point, and therefore we will have four coefficients  $\alpha_{\phi,j}$ .

The use of the Gaussian process also offers great flexibility through the choice of the covariance matrix  $\Sigma_\phi$ . Here, we use the matrix given by

$$\Sigma_\phi = \zeta_\phi^2 \mathcal{E} + \tau_\phi^2 I \tag{7}$$

175 where  $I$  is the identity matrix. The matrix  $\mathcal{E}$  is given by an exponential correlation function and has components

$$\mathcal{E}_{i,j} = \exp(-d(i,j)^T \beta d(i,j)). \tag{8}$$

For two grid-points  $x_i$  and  $x_j$ , the vector  $d(i,j)$  has two components given by the differences in longitude and latitude between  $x_i$  and  $x_j$ .

The  $2 \times 2$  matrix  $\beta$  is symmetric positive definite. Its entries measure how quickly spatial dependence drops off in the two different directions. The other 180 parameters of the covariance matrix appearing in (7) are the partial sill  $\zeta_\phi^2$  and the nugget parameter  $\tau_\phi^2$ . Further details on variogram analysis may be found in Zimmerman and Li (2012).

As mentioned, we assume the same Gaussian process model for the shape 185 parameter  $\xi$ . Similar expressions as those above are used for  $\mu_\xi$  and  $\Sigma_\xi$ .

### 3.1.3. Priors layer

Finally for the third layer in the hierarchical model, priors must be assigned to the hyperparameters, which are assumed to be independent. For those in (6), a normal distribution with large variance was selected: the covariates were

re-scaled to be centred on zero and priors  $\alpha_{\phi,i}$ ,  $\alpha_{\xi,i} \sim N(0, 50)$  were used. A lognormal prior was employed for the positive  $\zeta_{\phi}^2$  and  $\tau_{\phi}^2$  parameters in (7); that is, their logarithm was assumed to have the normal distribution  $N(0,10)$ .

For the entries of the matrix  $\beta$  in (8), a discrete uniform prior is assumed. We begin with a set of proposal values  $v_{\beta} = \{0.001, 0.005, 0.01, 0.05, \dots, 100, 500, 1000\}$ , with all values being considered equally likely *a priori*. Within the MCMC algorithm, the entries of  $\beta$  are randomly proposed from  $v_{\beta}$  and accepted or rejected accordingly by the algorithm.

### 3.2. Modelling the frequency of exceedances

We now turn to  $\zeta_u$ , which is defined as the probability that the threshold  $u$  is exceeded and is needed in (3) to compute return levels. For a given choice of threshold (discussed in Section 4), we let  $\zeta(x_i)$  be the exceedance probability at the location  $x_i$ . It is again assumed that there is a latent spatial process driving this and a hierarchical model is used, with data, process and prior layers.

At the data layer it is assumed that, at each grid-point  $i$ , the number of declustered threshold exceedances  $N_i$  is a binomial random variable with  $m_i$  trials (the total number of observations), each with a probability  $\zeta(x_i)$  of being a cluster maximum. That is,  $N_i \sim \text{Bin}(m_i, \zeta(x_i))$ , where

$$P(N = N_i) = \binom{m_i}{N_i} \zeta(x_i)^{N_i} (1 - \zeta(x_i))^{m_i - N_i}.$$

The process layer is similar to that of the GPD parameter  $\phi(x)$ . Following Diggle et al. (1998),  $\zeta(x_i)$  is first transformed using a logit transformation, where

$$\zeta'(x_i) \equiv \text{logit}(\zeta(x_i)) = \log \left( \frac{\zeta(x_i)}{1 - \zeta(x_i)} \right).$$

This is then modelled as a Gaussian process as before, with mean vector  $\mu_{\zeta'}$  and covariance matrix  $\Sigma_{\zeta'}$  taking the same form as in (6) and (7), respectively. The hyperparameters are then given the same prior distributions as described above in Section 3.1.3.

## 4. Implementation Details

### 4.1. Data

210 The data used in this study comes from the 34-year hindcast described in Gallagher et al. (2014). The third-generation spectral WAVEWATCH III version 4.11 model (Tolman, 2014) was used with an unstructured grid (Roland, 2008) to simulate the wave climate around Ireland for the period from 1979 to 2012. The model was forced with directional spectra and 10m wind data from  
215 the ERA-interim reanalysis of the European Centre for Medium-Range Weather Forecasts (Dee et al., 2011). The unstructured triangular grid consisted of approximately 15,000 nodes (Fig. 2(a)) with horizontal resolution varying from 250m in the nearshore to 10km further offshore. The hourly fields produced were validated with observations from wave buoys and satellite altimeter data.  
220 The simulation was found to give excellent agreement for the significant wave height, defined here as  $H_s = 4\sqrt{m_0}$ , where  $m_0$  is the zeroth moment of the directional wave spectrum.

Gallagher et al. (2014) reported a strong spatial variability in  $H_s$ . A preliminary extreme value analysis of this data-set using the annual maxima and  
225 GEV approach was carried out in Gallagher (2014). The highest extremes of  $H_s$  were found to occur off the west coast of Ireland along with a high level of uncertainty in the estimates. Similar results were found in Clancy et al. (2015), using a data-set from a coarser-resolution hindcast.

Following on from this, we will focus on the region off the west coast indicated  
230 in Fig. 2. This domain contains  $n_x = 334$  nodes and has a depth ranging from 39m to 1902m (see Fig. 2(b)). In Fig. 3 we show the mean, 99.8th percentile and maximum of the hindcasted  $H_s$  fields. We can see that, even not far from the coast, we have maxima in excess of 14m, a level of sea state categorised as ‘phenomenal’ by the World Meteorological Organization (WMO, 2009).

### 235 4.2. Threshold choice and declustering

We wish to apply the GPD model of threshold exceedance to this data-set. The choice of an appropriate threshold is a non-trivial issue in extreme value



modelling and the subject of much ongoing research (Scarrott and MacDonald, 2012). With a low threshold, the asymptotic validity of the GPD may be violated, leading to bias. On the other hand, if the threshold is set too high we will be left with few data points for fitting the model, resulting in large variances. Coles (2001) describes commonly-used graphical methods for choosing the threshold for a single time series, based on the asymptotic properties of the GPD given in (1). However these are rather subjective and furthermore are not suited to modelling over a region with multiple locations.

Numerous other methods have been described in the literature. Thompson et al. (2009) automate the choice using a goodness-of-fit test for the distribution of successive parameter estimate differences as the threshold is increased. A quantile regression model was employed by Northrop and Jonathan (2011), while Dupuis (1998) used optimal bias robust estimators to fit the model and test for the validity of the threshold. In Tancredi et al. (2006), the authors, using a Bayesian framework, incorporate the uncertainty from the threshold choice in their model.

Here we adopt a more straightforward approach. At each grid-point, the 99.8th percentile of the  $H_s$  data series is selected as a threshold for modelling at that point. Taking a percentile-based threshold is convenient when dealing with a spatial array of data. In Clancy et al. (2015), the 97th percentile was used. Caires and Sterl (2005) examined both the 93rd and 97th percentile and found the higher to be more appropriate in general. Vanem (2015) tested thresholds based on even higher percentiles and found, in some cases, a value around the 99.5th to be too low. The validity of the threshold may be assessed *a posteriori* by examining the fit of the model, and we discuss this choice further below.

Once we have chosen our threshold, we need to decluster the data to be used. This is necessary because the theoretical basis for the use of the GPD assumes that the exceedances are independent. Caires (2011) retains only the maxima of clusters of successive exceedances and additionally removes any peaks which occur less than 48 hours from another, these regarded as having been caused by the same storm system. Nicolae Lerma et al. (2015) varied this time between

48 and 72 hours and found no significant differences in their final results.

270 In this study we apply a similar, though slightly stricter, method of declustering to that of Caires (2011). Two successive sequences of exceedances are considered to be part of the same cluster and system if the time series drops below the threshold for 48 hours or less. We then use the peaks of each cluster for modelling with the GPD.

### 275 4.3. Model fitting

The GPD was fitted to the data-set discussed above using the spatial Bayesian hierarchical model (BHM). Approximate draws from the posterior distribution of each parameter in the hierarchical model were obtained via the MCMC algorithm. Metropolis-Hastings steps were employed to update each parameter  
280 in turn, for each iteration of the MCMC algorithm. This involves drawing a potential value from an appropriate distribution and accepting or rejecting it according to the Hastings ratio. It is a standard method particularly suited to high-dimensional distributions, such as we have here (Geyer, 2011).

Three parallel chains were run for each model. Each simulation consisted  
285 of 20,000 iterations, of which the first 2,000 were considered as burn-in and consequently discarded. In order to reduce dependence amongst the remaining values, only every 10th was kept. Convergence of the resulting chains was established using the  $\hat{R}$  criterion recommended by Gelman (1996), with values below the suggested criterion of 1.2 taken to imply convergence.

290 The model was implemented in R using a package called Rcpp (Eddelbuettel, 2013). This interface allows integration of R with C++ code, leading to appreciable reduction in the computational burden of the Metropolis-Hastings MCMC algorithm used.

We compare the results with those obtained by fitting the distribution independently at each grid-point with inference from a maximum likelihood (ML)  
295 method. For this we have used the Wave Analysis for Fatigue and Oceanography (WAFO) toolbox in Matlab (WAFO, 2011). In addition to spatial maps of the output, we have chosen four locations to focus on in more detail. These

are marked in Fig. 2(b). Their locations and some hindcast details are listed in  
300 Table 1.

As discussed in Section 2.2 above, the ML approach produces a single estimate of a given parameter and confidence intervals may be derived from its asymptotic properties. We will consider 95% confidence intervals for our estimates and present the lower and upper bounds of these intervals. The Bayesian  
305 model, on the other hand, yields a distribution for the parameter. From the values simulated by the MCMC algorithm, we will present the median value of this distribution and, again, confidence intervals bounded by the 2.5th and 97.5th percentiles.

## 5. Results

### 310 5.1. Parameters of the GPD

In Figs. 4 and 5 we compare the estimates of the GPD shape,  $\xi$ , and scale,  $\sigma$ , parameters, respectively. In both cases, the BHM shows less variation over the domain when compared with the ML. Additionally, in particular for  $\sigma$ , the spatial variation is noticeably smoother in the BHM. This is to be expected given  
315 the latent spatial processes in the BHM, with the covariances of parameters at different grid-points, given by (7) and (8), based on the distances between them. With the ML model we have simply fitted the distribution independently at each grid-point, with no relationship between neighbouring points.

The ML estimates of the shape and scale parameters (in Figs. 4 and 5 respectively) display a strong negative correlation between the two surfaces. This suggests that the ML method struggles to decide whether the observed extremes are better modelled either by a large value of the shape or a large value of the scale. This uncertainty is absent from the BHM method - both median surfaces are smoother than their ML counterparts, and lack the negative correlation of  
325 the latter.

The lower and upper bounds of the 95% confidence intervals for the estimates of  $\xi$  and  $\sigma$  are plotted in Figs. 6 and 7, respectively. The benefits of the Bayesian

approach are immediately clear: there is much less uncertainty in the parameter value estimates as evident from the much narrower confidence intervals. This can be seen in more detail in Fig. 8, in which we show parameter estimates with confidence intervals for both models for the four points detailed in Table 1. Looking at Figs. 8(a) and 8(b) for shape and scale, respectively, we see that at each of these four points the confidence intervals for the BHM are contained within the larger intervals for the ML; in fact, this is true for  $\sigma$  at every grid-point and for  $\xi$  at all but five grid-points.

We note that the estimates in Fig. 4 show a positive shape parameter in general. Looking at the confidence intervals in Fig. 6 and also, in particular, Fig. 8(a), we see that the ML model allows for a wide range of values, both positive and negative. The narrower intervals of the Bayesian model favour a positive shape.

Some previous studies have fixed the range of the shape parameter in advance: to be, for example, strictly negative (Ortego et al., 2012) or equal to zero (Caires and Sterl, 2005). We have set no such restrictions *a priori*. This allows the model the flexibility to determine the best values of the parameters, given the data. The appropriateness of this can be addressed when validating the model. This will be discussed in detail in Section 6 below.

## 5.2. Return levels

We now turn to the  $N$ -year return levels of significant wave height, which we denote by  $H_{sN}$ . In Figs. 9 and 10, we present spatial plots of  $H_{s20}$  and  $H_{s100}$ , respectively. The estimates are given along with the lower and upper bounds of the 95% confidence interval. The overall patterns of the return levels are broadly similar for the estimates with both models (top panels of Figs. 9 and 10). The main differences can be seen in the size of the confidence intervals, which again are much narrower in the BHM. This is even more evident for  $H_{s100}$  in Fig. 10.

As we extrapolate in time to longer return periods, we expect the uncertainty to grow. However, with the BHM we still have a much tighter confidence interval

even at 100 years. Examining the values in Figs. 8(c) and 8(d) makes this more explicit. At Point 4, for example, the ML method produces a 100-year return  
 360 level in the range of 12.15 to 21.19 metres. The BHM gives a sharper result of  
 14.90 to 19.42 metres.

We find that the highest extremes of significant wave height are to be expected roughly between 53°N and 54.5°N, with 100-year levels close to 17m. This spatial pattern is consistent with the annual maxima GEV analysis of the  
 365 data-set in Gallagher (2014), although this approach resulted in considerably higher levels of uncertainty in the estimates, due to the much smaller number of data-points used. In Clancy et al. (2015), a larger area was studied using a coarser-resolution hindcast but, nevertheless, this region (and further west) showed the most extreme sea states.

### 370 5.3. Validation

Return level plots, as described in Coles (2001), are a useful diagnostic tool for assessing the fit of a model, in addition to illustrating the model estimates. The return level estimates, along with the bounds of the confidence intervals, are plotted against the return period. Empirical return level estimates from the data are added as follows. Given any value  $z_{(i)}$  in the data set, the return period is given by  $N_{(i)}$  if

$$P(Z > z_{(i)}) = 1 - \tilde{F}(z_{(i)}) = 1/N_{(i)}$$

where  $\tilde{F}(z_{(i)})$  is the empirical distribution function; c.f. Sections 2.6.7 and 3.3.5 of Coles (2001) for more details.

These are shown in Fig. 11 for the four locations in Table 1 which we have been examining. The black curves are for the ML model, while the BHM is  
 375 shown in red. The return levels from both models are quite similar for all return periods shown. However, once again we see clearly how the Bayesian model yields estimates with much less uncertainty. This is more evident as we extrapolate to longer return periods, such as 1000 years. Comparing with

the empirical estimates shows a satisfactory fit, particularly for shorter return  
380 periods.

The Bayesian spatial model is further validated in Fig. 12 using a posterior predictive distribution (Gelman et al., 1996). This has been constructed as follows: at a given grid-point  $i$ , we have  $n_i$  exceedances over the threshold. We then randomly draw  $n_i$  values for the shape and scale parameters from the  
385 respective estimated posterior distributions; these describe  $n_i$  separate GPDs. For each of these, one value is generated. We now have  $n_i$  observed and  $n_i$  predicted exceedances. After ordering, the two sets should match since a good model will have predicted the observed values.

This is repeated for each of the  $n_x = 334$  grid-points and a scatter-plot of  
390 the results is shown in Fig. 12. The combination of positive values of the shape parameter and randomly selecting larger values of the scale leads to some quite high predicted values. But in general we see quite a good match between the observed and predicted exceedances. The correlation coefficient is 0.95 and the bias is 0.43m, giving confidence in the validity of the fitted Bayesian model.

## 395 6. Discussion and Conclusions

We have applied a Bayesian hierarchical model (BHM) to a hindcast dataset in order to study the extremes of significant wave height off the west coast of Ireland. Exceedances of  $H_s$  over a high threshold are modelled with the generalised Pareto distribution. The hierarchical model includes a latent spatial  
400 process which allows us to more effectively capture the spatial variation of the extremes. This approach was compared with a model which uses maximum likelihood (ML) inference and simply carries out an independent extreme value analysis on the time-series at each grid-point in a given domain.

The parameters of the fitted generalised Pareto distribution (GPD) were  
405 used to produce spatial maps of extrapolated return levels of  $H_s$ . Consistent with previous studies, we found that the highest extremes are to be expected in the latitude band roughly between  $53^\circ\text{N}$  and  $54.5^\circ\text{N}$  to the west of Ireland, with

phenomenal sea states of around 17m estimated for the 100-year return level. A comparison of the two methods showed that the BHM produces smoother  
410 estimates, as a result of the latent spatial process whereby the model parameters at a given grid-point are influenced by the neighbouring points.

As we have used data from a wave model hindcast, we should consider how well the model is capturing extremes when interpreting the results. A slight underestimation of the extremes of significant wave height has been reported  
415 (Gallagher et al., 2014), consistent with an underestimation in ERA-interim winds driving the model (Stopa and Cheung, 2014). As a check on the validity of our model-derived results, we may compare with those derived from satellite data in Izaguirre et al. (2011). Although at much lower resolution, for the regions immediately to the west of Ireland they reported wintertime 20-year  
420 return levels of around 14-16m (c.f. their Fig. 2). These are consistent with our return levels in Fig. 9. In future work, nevertheless, it may be worthwhile to consider any deficiencies in numerical model output; see, for example, Caires and Sterl (2005) and Hanafin et al. (2012).

A major advantage of the Bayesian approach is the fact that it formally handles  
425 parameter uncertainty, rather than relying on the approximate normality of the ML estimate. Parameter and return level estimates were analysed with confidence intervals bounded by the 2.5th and 97.5th percentiles. The BHM showed much narrower intervals throughout, yielding much higher levels of certainty with the results. This is of crucial importance, as a single point estimate  
430 alone is of little practical value without a meaningful measure of uncertainty. Indeed, looking at the 1000-year return levels in Fig. 11, we see upper bounds of nearly 35m for the ML, which seem physically unrealistic.

The threshold chosen for this work was the 99.8th percentile of the  $H_s$  data at each grid-point. As noted, this approach has been used by a number of authors,  
435 with various ranges of percentiles tested. Initially, we fitted the model using lower percentiles: the 97th, following previous work in Clancy et al. (2015), and then the 99th. However, when analysing the validity of the fit, as discussed in Section 5.3, we found that both models (ML and BHM) greatly underestimated

the higher return levels when using the 97th. Subsequently in an early draft  
440 of this work, modelling using the 99th percentile still showed a bias with the  
return level intervals for the Bayesian model still failing to capture the higher  
extremes. Moving to the higher 99.8th as a threshold improved this greatly, as  
seen from Fig. 11.

As discussed earlier, a very high threshold is necessary for the asymptotic  
445 validity of the GPD model, but can result in too few points with which to fit.  
Mazas and Hamm (2011), for example, recommended a threshold yielding an  
average of 2-5 excesses per year. Here we have an average of 51 excesses at each  
location, after declustering, for a 34-year data-set. This still leaves us with many  
more data points than if modelling with the annual maxima approach, and the  
450 validating results outlined above in Section 5.3 attest to the appropriateness of  
the high threshold.

Increasing the threshold was also found to have a noticeable effect on the  
shape parameter. With the lower thresholds, the shape was negative, similar  
to other reported results (for example Caires, 2011). But with the 99.8th per-  
455 centile threshold eventually chosen for this work, we found a generally positive  
shape parameter. As noted above, the lower thresholds were failing to capture  
the highest extremes, which could be seen a result of the negative shape which  
describes a distribution with an upper bound. With the higher threshold and  
resulting positive shape parameter, the higher values are represented more ac-  
460 curately. We note, in passing, that positive values of this parameter have been  
reported elsewhere: in other geographical regions (Izaguirre et al., 2011), for  
wave data-sets from future climate projections (Vanem, 2015) and with the an-  
nual maxima GEV model in a similar domain off the west of Ireland (Clancy et  
al., 2015).

465 In addition to a correlation between points based on the distance between  
them, we have included depth as a covariate in the mean of all second-layer  
parameters, as outlined in 6. Other covariates may be added to the model. For  
example, time-dependent covariates can be used to study the trends in extremes  
with a changing climate (Caires et al., 2006; Méndez et al., 2006).



470 Different spatio-temporal models may also be explored, in which extremes  
at a point are affected not just by neighbouring locations, but by recent extreme  
events. These more complex models may potentially further reduce the uncer-  
tainty levels in the Bayesian model, to yield estimates of increasing practical  
value. Such an investigation is a possible extension to this current work.

#### 475 **Acknowledgements**

The authors are grateful to Sarah Gallagher for her helpful advice on the  
hindcast data. The comments from the reviewers were of great benefit and led to  
tests with the higher threshold and an overall improvement in the manuscript.  
This work was supported by the European Research Council (ERC) under the  
480 research project ERC-2011-AdG 290562-MULTIWAVE and Science Foundation  
Ireland under grant number SFI/12/ERC/E2227, as well as by the Environmen-  
tal Protection Agency grant number 2012-CCRP-PhD.3.

#### **References**

- [1] S. Coles, An introduction to statistical modeling of extreme values,  
485 Springer-Verlag London, 2001.
- [2] E. Vanem, Long-term time-dependent stochastic modelling of extreme  
waves, *Stochastic Environmental Research and Risk Assessment* 25 (2011)  
185–209.
- [3] P. Jonathan, K. Ewans, Statistical modelling of extreme ocean environ-  
490 ments for marine design: a review, *Ocean Engineering* 62 (2013) 91–109.
- [4] E. Vanem, Uncertainties in extreme value modelling of wave data in a cli-  
mate change perspective, *Journal of Ocean Engineering and Marine Energy*  
1 (2015) 339–359.
- [5] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, D. B.  
495 Rubin, *Bayesian data analysis*, 3rd Edition, Chapman & Hall, 2013.

- [6] D. Cooley, D. Nychka, P. Naveau, Bayesian Spatial Modeling of Extreme Precipitation Return Levels, *Journal of the American Statistical Association* 102 (2007) 824–840.
- [7] E. C. Oliver, S. J. Wotherspoon, N. J. Holbrook, Estimating extremes from global ocean and climate models: A Bayesian hierarchical model approach, *Progress in Oceanography* 122 (2014) 77–91.
- [8] L. O’Brien, J. M. Dudley, F. Dias, Extreme wave events in Ireland: 14 680 BP – 2012, *Natural Hazards and Earth System Science* 13 (2013) 625–648.
- [9] S. Gallagher, R. Tiron, E. Whelan, E. Gleeson, F. Dias, R. McGrath, The nearshore wind and wave energy potential of Ireland: a high resolution assessment of availability and accessibility, *Renewable Energy* 88 (2016) 494–516.
- [10] J. Pickands, Statistical inference using extreme order statistics, *Annals of Statistics* 3 (1975) 119–131.
- [11] M. Menéndez, F. J. Méndez, C. Izaguirre, A. Luceño, I. J. Losada, The influence of seasonality on estimating return values of significant wave height, *Coastal Engineering* 56 (2009) 211–219.
- [12] C. Izaguirre, F. J. Méndez, M. Menéndez, I. J. Losada, Global extreme wave height variability based on satellite data, *Geophysical Research Letters* 38 (2011) 1–6.
- [13] S. Caires, A. Sterl, 100-year return value estimates for ocean wind speed and significant wave height from the ERA-40 data, *Journal of Climate* 18 (2005) 1032–1048.
- [14] J. Vinoth, I. R. Young, Global estimates of extreme wind speed and wave height, *Journal of Climate* 24 (2011) 1647–1665.
- [15] A. Nicolae Lerma, T. Bulteau, S. Lecacheux, D. Idier, Spatial variability of extreme wave height along the Atlantic and channel French coast, *Ocean Engineering* 97 (2015) 175–185.

- [16] S. Caires, Extreme value analysis: wave data. JCOMM Technical Report  
525 No. 57, Tech. rep., World Meteorological Organization (2011).
- [17] O. J. Aarnes, Ø. Breivik, M. Reistad, Wave Extremes in the Northeast  
Atlantic, *Journal of Climate* 25 (2012) 1529–1543.
- [18] C. Clancy, Belissen V, R. Tiron, S. Gallagher, F. Dias, Spatial variability  
of extreme sea states on the Irish west coast, in: Proceedings of the  
530 ASME 2015 34th International Conference on Ocean, Offshore and Arctic  
Engineering, St John’s, NL, Canada, 2015.
- [19] S. Coles, L. R. Pericchi, S. Sisson, A fully probabilistic approach to extreme  
rainfall modeling, *Journal of Hydrology* 273 (2003) 35–50.
- [20] J. J. Egozcue, V. Pawlowsky-Glahn, M. I. Ortego, Wave-height hazard  
535 analysis in Eastern Coast of Spain - Bayesian approach using generalized  
Pareto distribution, *Advances in Geosciences* 2 (2005) 25–30.
- [21] M. G. Scotto, C. Guedes Soares, Bayesian inference for long-term prediction  
of significant wave height, *Coastal Engineering* 54 (2007) 393–400.
- [22] C. J. Geyer, Introduction to MCMC, in: S. Brooks, A. Gelman, G. L. Jones,  
540 X.-L. Meng (Eds.), *Handbook of Markov Chain Monte Carlo*, Chapman &  
Hall/CRC, 2011.
- [23] F. Fedele, Space-Time Extremes in Short-Crested Storm Seas, *Journal of  
Physical Oceanography* 42 (2012) 1601–1615.
- [24] B. Cañellas, A. Orfila, F. Méndez, M. Menéndez, J. Tintoré, Application of  
545 a POT model to estimate the extreme significant wave height levels around  
the Balearic Sea (Western Mediterranean), *Journal of Coastal Research  
Special Issue* 50 (2007) 329–333.
- [25] L. Guo, J. Sheng, Statistical estimation of extreme ocean waves over the  
550 eastern Canadian shelf from 30-year numerical wave simulation, *Ocean Dy-  
namics* 65 (2015) 1489–1507.

- [26] E. Vanem, A. B. Huseby, B. Natvig, A Bayesian hierarchical spatio-temporal model for significant wave height in the North Atlantic, *Stochastic Environmental Research and Risk Assessment* 26 (2012) 609–632.
- [27] S. Banerjee, B. P. Carlin, A. E. Gelfand, *Hierarchical Modeling and Analysis for Spatial Data*, 2nd Edition, Chapman & Hall/CRC, 2014.
- [28] D. L. Zimmerman, J. Li, Model-Based Frequentist Design for Univariate and Multivariate Geostatistics, in: J. Mateu, W. G. Muller (Eds.), *Spatio-temporal Design: Advances in Efficient Data Acquisition*, Wiley, 2012, p. 370.
- [29] P. J. Diggle, J. A. Tawn, R. A. Moyeed, Model-based geostatistics, *Applied Statistics* 47 (1998) 299–350.
- [30] S. Gallagher, R. Tiron, F. Dias, A long-term nearshore wave hindcast for Ireland: Atlantic and Irish Sea coasts (1979-2012), *Ocean Dynamics* 64 (2014) 1163–1180.
- [31] H. Tolman, User manual and system documentation of Wavewatch III version 4.18. Technical report 316, Tech. rep., NOAA/NWS/NCEP/MMAB (2014).
- [32] A. Roland, Development of WWM (Wind Wave Model) II: Spectral wave modelling on unstructured meshes., Ph.D. thesis, Technical University Darmstadt, Germany (2008).
- [33] D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J. J. Morcrette, B. K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J. N. Thépaut, F. Vitart, The ERA-Interim re-

- analysis: Configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society 137 (2011) 553–597.
- 580 [34] S. Gallagher, The Nearshore Wave Climate and Wave Energy Resource of Ireland : Past , Present and Future, Ph.D. thesis, University College Dublin (2014).
- [35] WMO, Guide to Meteorological Instruments and Methods of Observation, Tech. rep. (2009).
- 585 [36] C. Scarrott, A. MacDonald, A review of extreme value threshold estimation and uncertainty quantification, Revstat Statistical Journal 10 (2012) 33–60.
- [37] P. Thompson, Y. Cai, D. Reeve, J. Stander, Automated threshold selection methods for extreme wave analysis, Coastal Engineering 56 (2009) 1013–1021.
- 590 [38] P. J. Northrop, P. Jonathan, Threshold modelling of spatially dependent non-stationary extremes with application to hurricane-induced wave heights, Environmetrics 22 (2011) 799–809.
- [39] D. J. Dupuis, Exceedances over High Thresholds: A Guide to Threshold Selection, Extremes 1 (1998) 251–261.
- 595 [40] A. Tancredi, C. Anderson, A. O’Hagan, Accounting for threshold uncertainty in extreme value estimation, Extremes 9 (2006) 87–106.
- [41] A. Gelman, Inference and monitoring convergence, in: W. R. Gilks, S. Richardson, D. J. Spiegelhalter (Eds.), Markov Chain Monte Carlo in Practice, Chapman & Hall, 1996.
- 600 [42] D. Eddelbuettel, Seamless R and C++ Integration with Rcpp, Springer-Verlag New York, 2013.
- [43] WAFO-group, WAFO - a Matlab Toolbox for Analysis of Random Waves and Loads, Tech. rep., Math. Stat., Center for Math. Sci., Lund Univ.,

Lund, Sweden (2011).

605 URL <http://www.maths.lth.se/matstat/wafo>

- [44] M. I. Ortego, R. Tolosana-Delgado, J. Gibergans-Báguena, J. J. Egozcue, A. Sánchez-Arcilla, Assessing wavestorm hazard evolution in the NW Mediterranean with hindcast and buoy data, *Climatic Change* 113 (2012) 713–731.
- 610 [45] A. Gelman, X.-L. Meng, H. Stern, Posterior predictive assessment of model fitness via realized discrepancies, *Statistica Sinica* 6 (1996) 733–807.
- [46] J. E. Stopa, K. F. Cheung, Intercomparison of wind and wave data from the ECMWF Reanalysis Interim and the NCEP Climate Forecast System Reanalysis, *Ocean Modelling* 75 (2014) 65–83.
- 615 [47] J. A. Hanafin, Y. Quilfen, F. Ardhuin, J. Sienkiewicz, P. Queffelec, M. Obrebski, B. Chapron, N. Reul, F. Collard, D. Corman, E. B. De Azevedo, D. Vandemark, E. Stutzmann, Phenomenal Sea States and Swell from a North Atlantic Storm in February 2011: A Comprehensive Analysis , *Bulletin Of The American Meteorological Society* 93 (2012) 1825–1832.
- 620 [48] F. Mazas, L. Hamm, A multi-distribution approach to POT methods for determining extreme wave heights, *Coastal Engineering* 58 (2011) 385–394.
- [49] S. Caires, V. R. Swail, X. L. Wang, Projection and analysis of extreme wave climate, *Journal of Climate* 19 (2006) 5581–5605.
- 625 [50] F. J. Méndez, M. Menéndez, A. Luceño, I. J. Losada, Estimation of the long-term variability of extreme significant wave height using a time-dependent peak over threshold (pot) model, *Journal of Geophysical Research: Oceans* 111 (C7).

Point	Lon.	Lat.	Depth	Mean	99.8th	Max.
1	-9.66	52.88	68.84	2.42	8.70	13.51
2	-11.22	53.10	134.74	3.12	10.03	15.21
3	-11.60	53.97	329.90	3.21	10.24	15.46
4	-10.70	53.91	130.17	3.12	10.08	15.24

Table 1: The longitude, latitude (both in degrees) and depth (metres) of each of the four points indicated in Fig. 2(b), along with the mean, 99.8th percentile and maximum significant wave height (metres) from the 34-year hindcast.

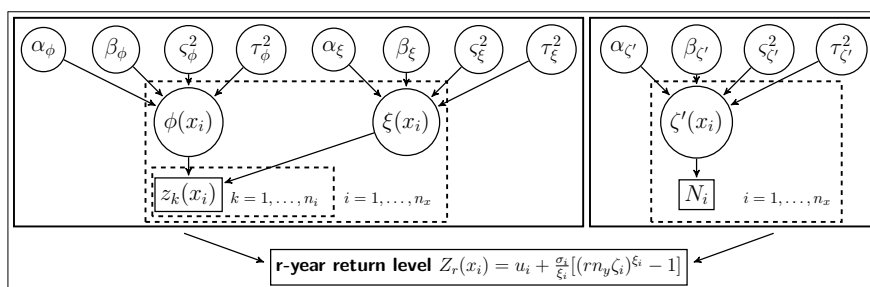


Figure 1: A directed acyclic graph (DAG) of the Bayesian hierarchical models fitted to the wave data. On the left is the model for the excesses (using the GPD to model the data) and on the right is the model for the probability of an observation exceeding the threshold (using the Binomial distribution). The parameters of the distributions are represented as circles in the middle layer, with the hyperparameters controlling these represented in the top layer. The data is represented in the bottom layer (in rectangles). Arrows run into nodes from their direct predecessors (often called *parents*). Given its parents, each node is independent of all other nodes in the graph except its descendants (often called *children*). Posterior estimates of the parameters' distributions can be used to form quantities of interest - typically return levels, as illustrated. Further details of each layer and the parameters involved may be found in the text.

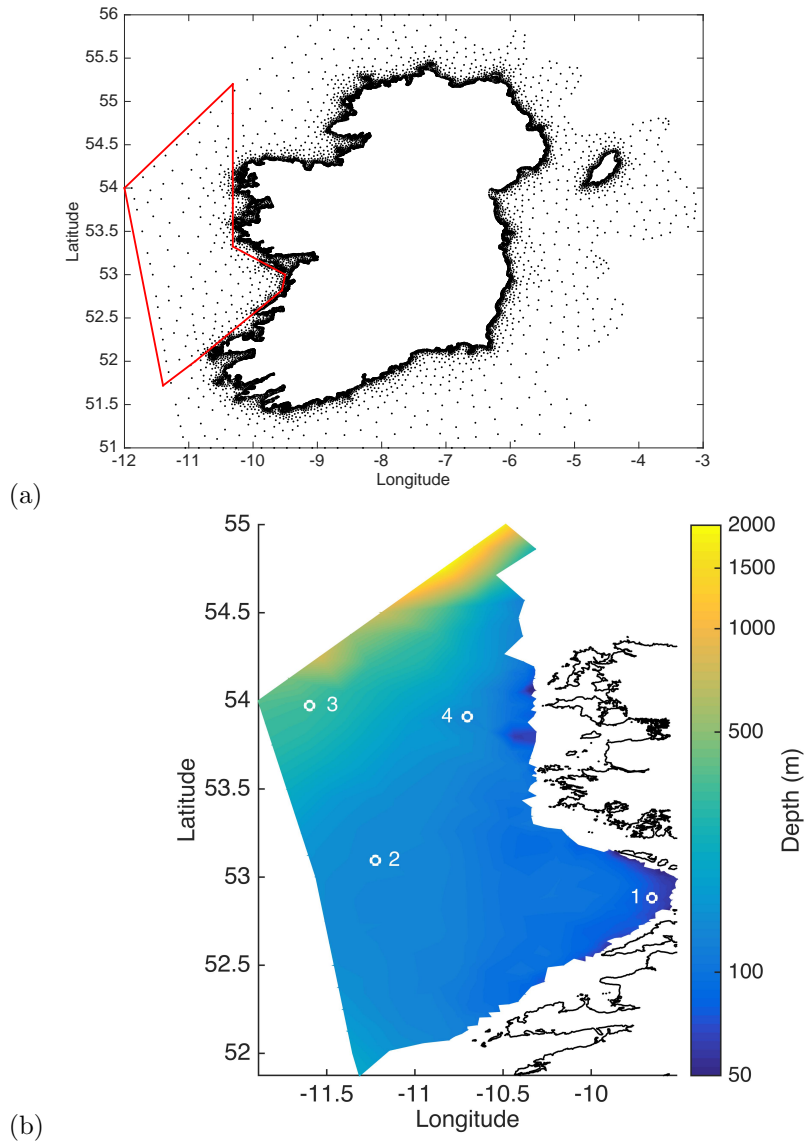


Figure 2: (a) the computational grid used for the hindcast, with our region of study outlined. (b) the depth of this region in metres, plotted on a logarithmic scale, with the four locations from Table 1 marked.



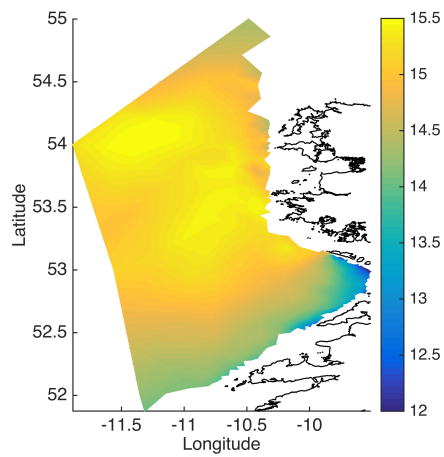
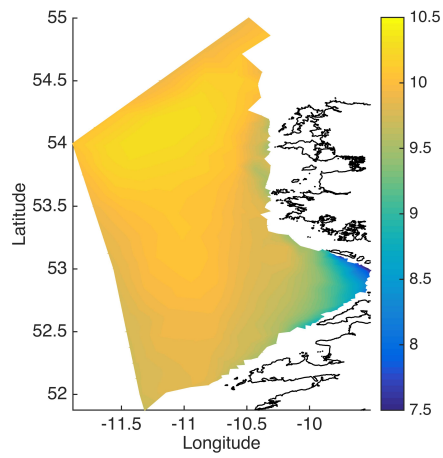
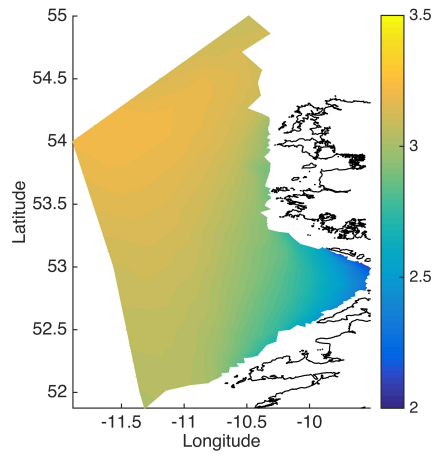


Figure 3: Top to bottom: mean, 99.8th percentile and maximum significant wave heights in metres, for the 1979–2012 hindcast. Note the differing scales.

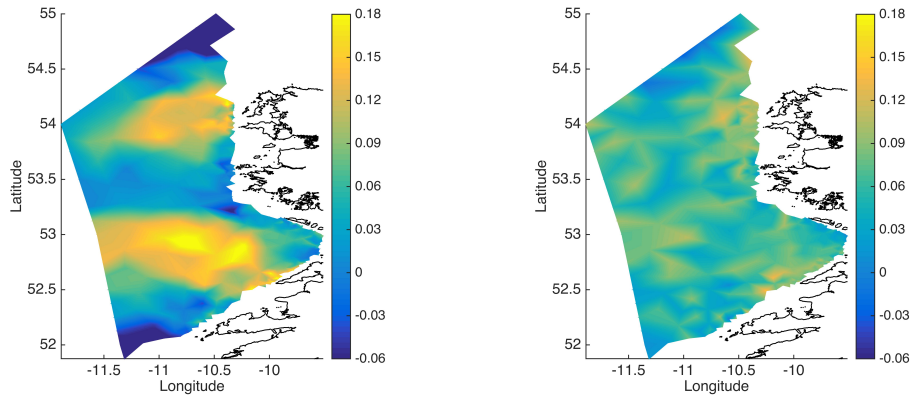


Figure 4: Estimates of the shape parameter  $\xi$ , using (left) the maximum likelihood (ML) method and (right) the Bayesian (BHM) model.

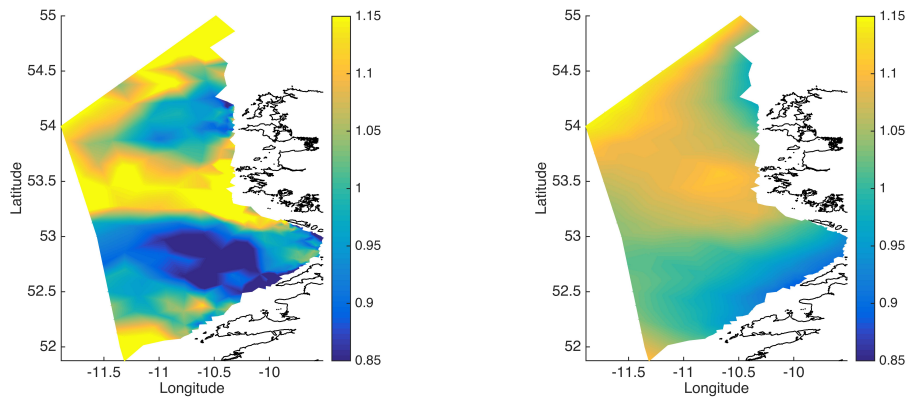


Figure 5: Estimates of the scale parameter  $\sigma$ , using (left) the maximum likelihood (ML) method and (right) the Bayesian (BHM) model.

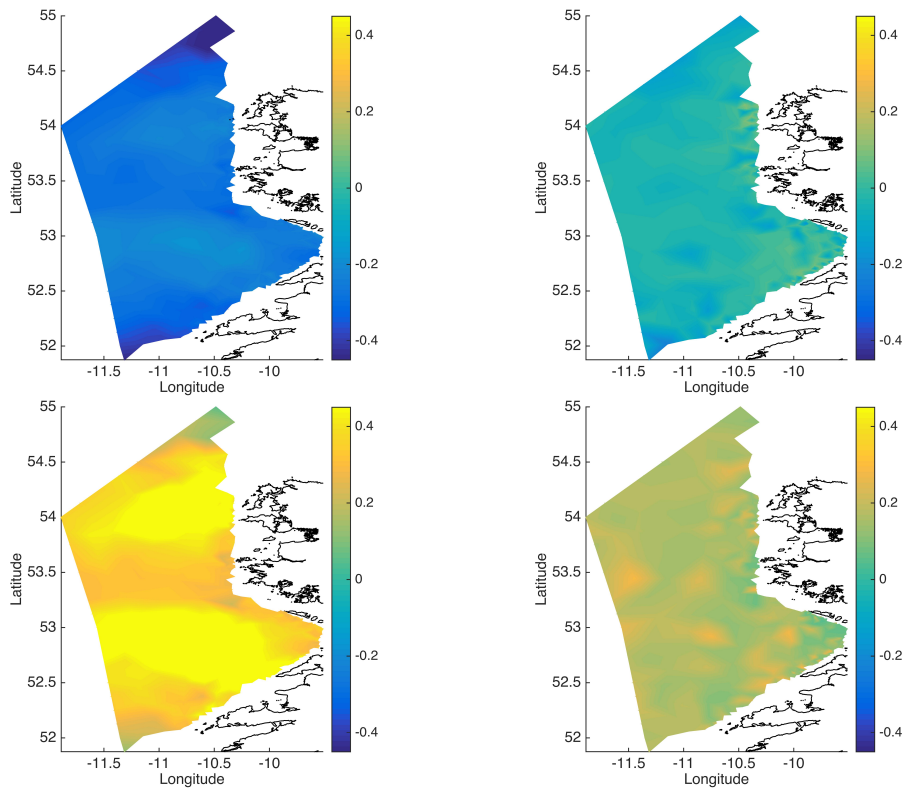


Figure 6: Confidence interval bounds for the shape parameter  $\xi$  using the ML method (left) and the BHM (right), given by the 2.5th (above) and 97.5th (below) percentiles of the estimates.

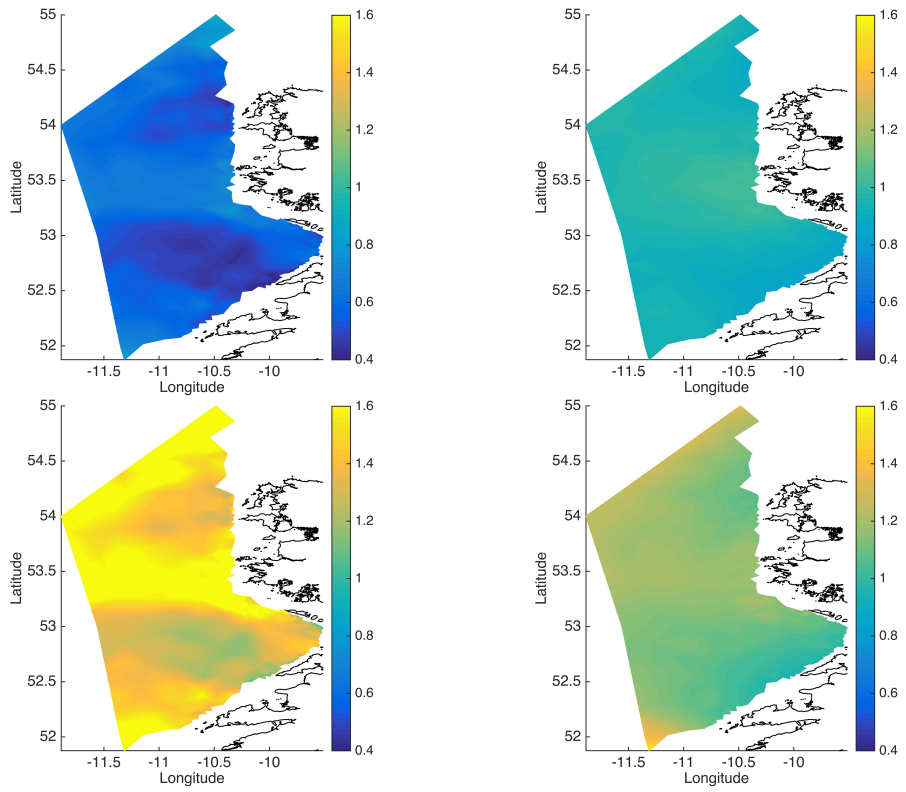


Figure 7: Confidence interval bounds for the scale parameter  $\sigma$  using the ML method (left) and the BHM (right), given by the 2.5th (above) and 97.5th (below) percentiles of the estimates.

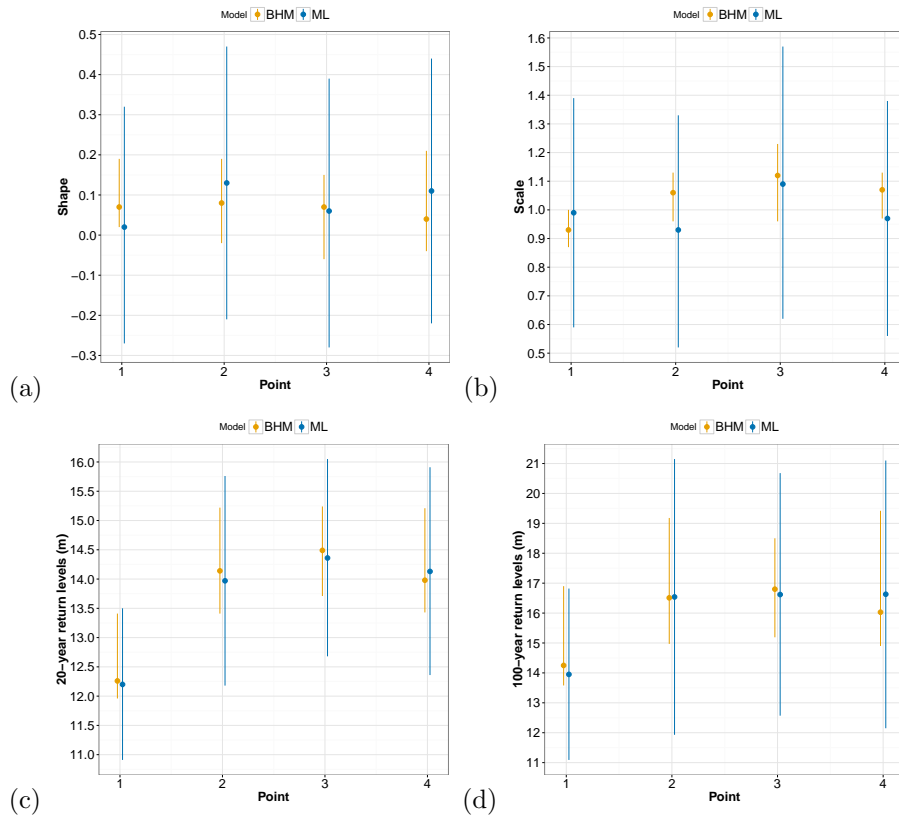


Figure 8: Comparison of estimates from both models for the points marked in Fig. 2(b) and detailed in Table 1: (a) shape and (b) scale parameters, along with the (c) 20-year and (d) 100-year return levels. The dot marks the estimate, with the vertical lines indicating the confidence interval bounded by the 2.5th and 97.5th percentile.

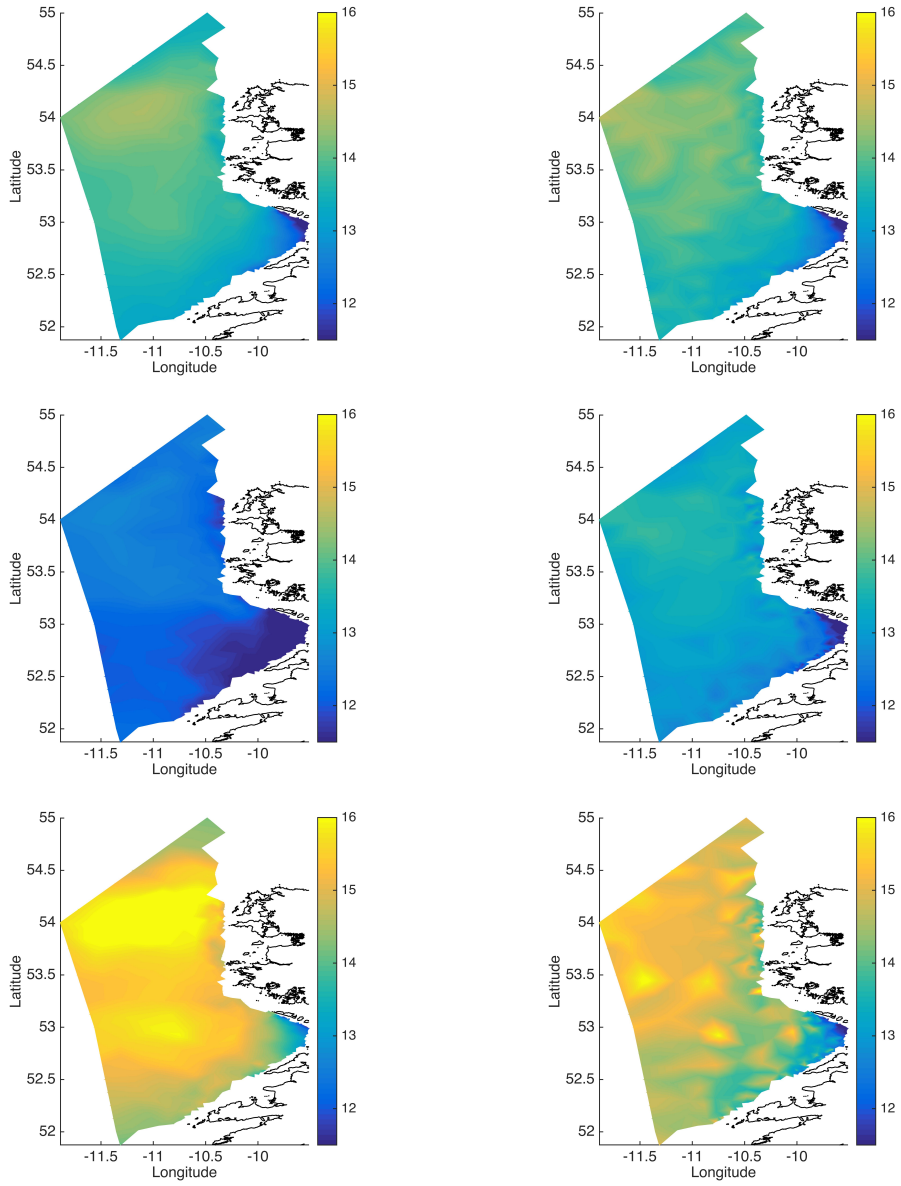


Figure 9: Maximum likelihood (left) and Bayesian (right) estimates of the 20-year return level of  $H_s$ . The estimates are shown (top), along with confidence interval bounds given by the 2.5th (middle) and 97.5th (bottom) percentiles.

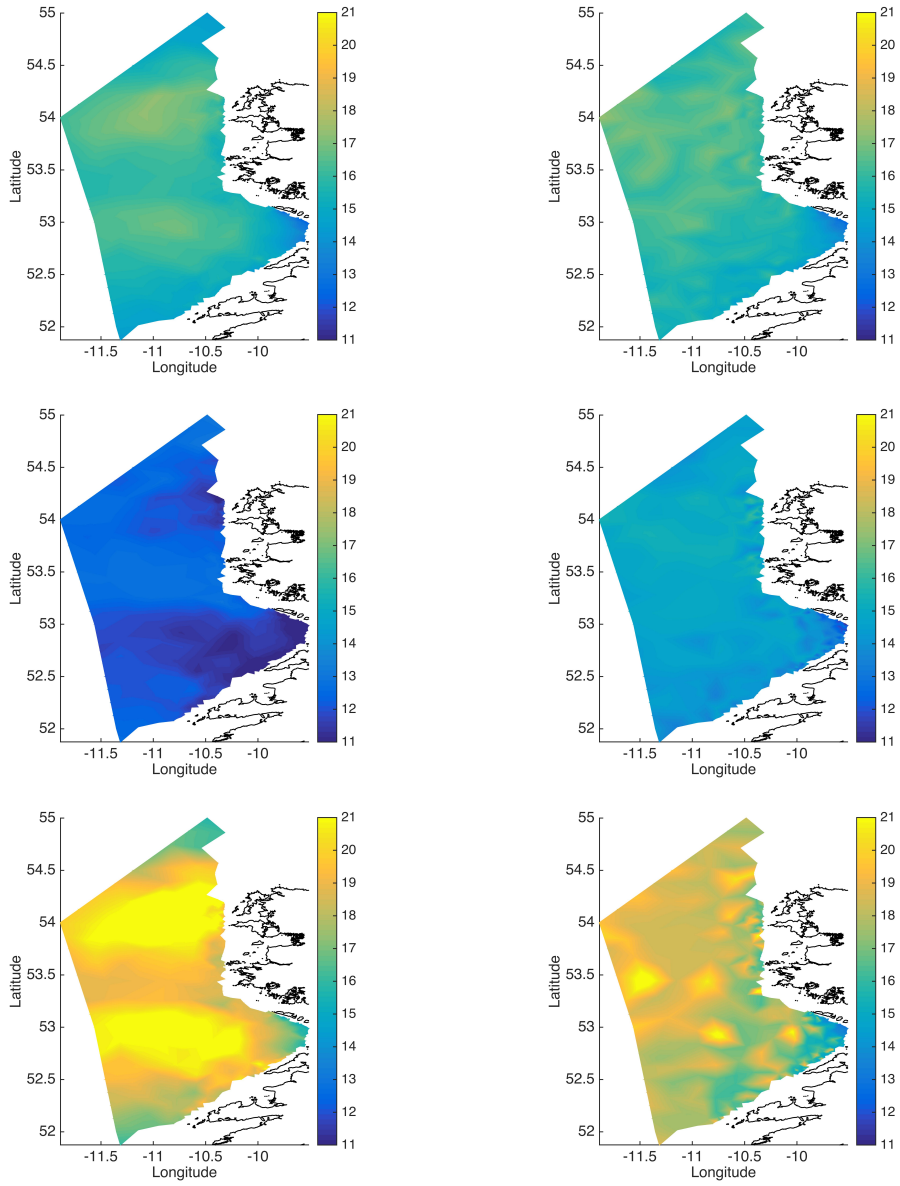


Figure 10: Maximum likelihood (left) and Bayesian (right) estimates of the 100-year return level of  $H_s$ . The estimates are shown (top), along with confidence interval bounds given by the 2.5th (middle) and 97.5th (bottom) percentiles.

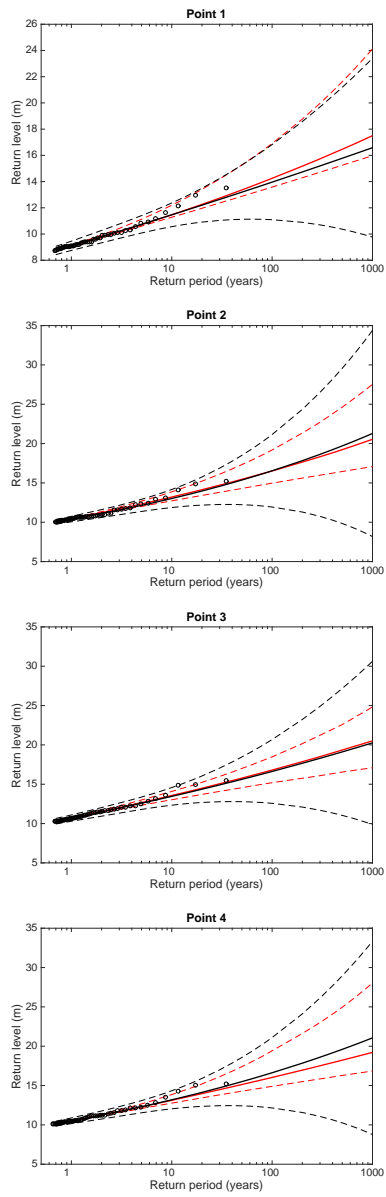


Figure 11: Model return level plots with empirical estimates (circles) for the four points (top to bottom) from Table 1. Continuous curves are the return level estimates for the maximum likelihood (black) and Bayesian (red) models using the 99.8th percentile threshold. The dashed curves are the corresponding lower and upper bounds of the 95% confidence interval. Note the logarithmic scale used for the return period.



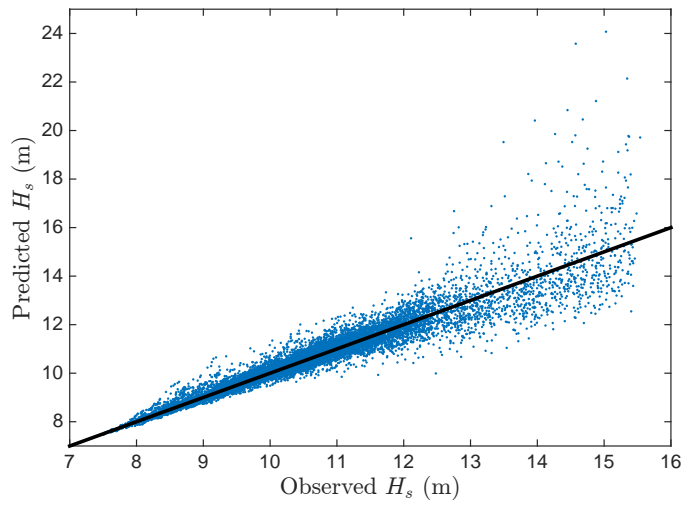


Figure 12: Validating the Bayesian spatial model by comparing observed exceedances of  $H_s$  with those predicted by the model. Further details may be found in the text.