Submitted to the Annals of Applied Statistics

SPATIO-TEMPORAL MODELLING OF EXTREME WEATHER EVENTS

By Theodoros Economou*, David B. Stephenson* and Christopher A. T. Ferro*

*University of Exeter**

This study presents a flexible spatio-temporal framework for modelling extreme weather events that occur at irregular spatial locations and that can depend on time-varying background conditions. The framework introduces a spatial regularisation and covariates to extend a recently published Bayesian hierarchical model for spatial extremes.

The framework is illustrated **on** extra-tropical cyclones objectively identified over the Atlantic and Europe in 6-hourly re-analyses from 1979-2009. Spatial variation in the extremal properties of the cyclones is well-captured using a 150 cell spatial regularisation, latitude as a covariate, and spatial random effects. Local spatial pooling across neighbouring cells gives much narrower credible intervals than a model having no spatial random effects, and ensures that the intervals converge as the cell size is reduced.

The North Atlantic Oscillation (NAO) is found to have a significant effect on intensifying extremal storm behaviour especially over N. Europe and the Iberian peninsula. Estimates of lower bounds on minimum sea-level pressure are typically 10-50hPa below the minimum values observed for historical storms with largest differences occuring when the NAO index is positive.

Keywords and phrases: Bayesian Hierarchical Model, spatial random effects, natural hazards, extra-tropical cyclones, extreme value distribution, European windstorms

1. Introduction. Extreme weather events are a major source of risk for society that need to better understood and quantified. They pose an interesting challenge for statistical modelling for several reasons: the events occur irregularly in space and time with rates and magnitudes that are spatially heterogeneous and non-stationary in time (due to modulation by large-scale climate conditions). Furthermore, at any one location, very few extreme events are generally observed in short historical weather data sets.

To address these issues, we propose and test a spatio-temporal extreme value framework suitable for analysing weather extremes. Our framework is designed to address important generic questions such as:

1. How extreme can weather systems become? Or more precisely, how much more extreme can events become, compared to the most extreme values recorded in short series of historical observations/analyses?
2. How does the extreme behaviour vary spatially?
3. How does the extreme behaviour vary in time due to modulation by large-scale climate patterns?

The framework is an extension of the methodology in Cooley and Sain (2008), which was based on a spatial extreme value point process model for extremes on a regular grid. The novelty here is the inclusion of temporal covariates in the model formulation, the adaptation to irregularly occurring (i.e. random occurrence rather than fixed locations) extremes in space, and the application to an important natural hazard.

We demonstrate the framework by applying it to Northern Atlantic and European extra-tropical cyclones. These natural hazards cause much damage and insurance loss in Europe due to extreme wind speeds/flooding. Recent examples include the December 1999 windstorms Anatol and Lothar (Ulbrich et al., 2001), which resulted in large loses in France, Germany and Switzerland; and windstorm Kyrill in 2007 which affected the UK, Germany, Holland, Belgium, Austria, Czech Republic and Poland. An important scientific question is: how low can sea-level pressure become in extra-tropical cyclones? Unfortunately, there are no simple physical arguments for how deep an extra-tropical cyclone can become. The most extreme events often deepen *explosively* with rapid decreases in central pressure e.g. storms known as *bombs* having pressure drops of more than 24hPa in 24 hours at 60N. Explosive cyclogenesis depends on many factors, for example, the deepest recorded 20th century low of 913hPa (the Braer cyclone of January 1993) deepened 78hPa in 24 hours due to a combina-

tion of several factors such as available moisture and stratospheric conditions (Odell et al., 2013). In the unlikely possibility that such conditions could be maintained for 2 days gives a minimum value of SLP of around 990-156=834hPa starting from a typical background state of 990hPa. It should also be noted that SLP less than 650hPa would correspond to mid-latitude geostrophic wind speeds faster than the speed of sound, which due to shock wave dissipation would be impossible to maintain energetically. In the absence of any more rigorous physical bounds, it is of interest to estimate bounds empirically using statistical approaches such as extreme value theory.

There has been surprisingly little use of extreme value theory to investigate extreme cyclones (see Katz (2010) for a discussion about the lack of extreme value theory in climate applications). Lionello et al. (2008) investigated changes in future cyclone climatology over Europe using the Generalised Extreme Value (GEV) distribution to model pressure minima. Return levels were calculated over the whole North Atlantic domain without explicit characterisation of spatial or temporal heterogeneity. Della-Marta and Pinto (2009) and Della-Marta et al. (2009) used a Generalised Pareto Distribution (GPD) model to analyse extreme wind intensity and how this might change in the future. Three different non-overlapping areas of large size were considered, however there was no formal consideration of spatial or temporal variation in the extremes. Sienz et al. (2010) used GPD models extending the work by Della-Marta and Pinto (2009) to include temporal covariates such as the North Atlantic Oscillation (NAO) and a linear trend but did not account for spatial variability. Bonazzi et al. (2012) used bivariate extreme value copulas to model the spatial dependence in footprints of peak gust wind speeds from a set of 135 damaging European cyclones. However, this study did not explicitly model the magnitude of many cyclones and so does not answer the question about upper bounds on cyclone magnitudes.

2. Statistical methodology.

2.1. *Brief review of spatial extreme models.* Davison et al. (2012) identified three main classes of statistical models for spatial extremes: Bayesian hierarchical models (BHM), copula based models and max-stable process models. Although max-stable processes explicitly characterise spatial depen-

dence, BHM can be more flexible and pragmatic by allowing for inclusion of physical mechanisms in terms of covariates and random effects. The major issue with BHM is the conditional independence assumption of the extremes whereas for max-stable processes it is model implementation and flexibility. Copula models lie somewhere in-between since the dependence of the extremes is modelled by the copula assuming that the marginal distributions are separable from this dependency structure (Sang and Gelfand, 2010).

In this paper, we adapt BHM as the modelling framework mainly because of their flexibility in allowing for (temporal) covariate effects along with a versatile spatial dependency structure through spatial random effects. BHM generally assume independence of the extremes for given values of the covariates (conditional independence), although they can be extended to model spatial extremal dependence by including max-stable processes (Reich and Shaby, 2012). For our application to extra-tropical cyclones, we believe conditional independence to be a reasonable working assumption. Much of the dependency between successive cyclones has been shown to be induced by modulation of rates by time-varying climate modes and so can be accounted for by the inclusion of appropriate covariates (Mailier et al., 2006; Vitolo et al., 2009).

There has been recent interest, in spatial and spatio-temporal BHM for extremes since their introduction by Casson and Coles (1999). In Cooley et al. (2007) and Cooley and Sain (2008), a Generalised Pareto Distribution (GPD) and a point process model for extremes are used to model extreme precipitation where the spatial dependence is characterised by including Gaussian random effects in the formulation of both the GPD and the point process parameters. These models however did not incorporate any temporal non-stationarity. Gaetan and Grigoletto (2007), Heaton et al. (2009) and Sang and Gelfand (2009) incorporated temporal structure in BHM through time-varying covariates where the conditional model is a Generalised Extreme Value (GEV) distribution. Turkman et al. (2010) used a similar model where the conditional model is a GPD. In this paper, we use the computationally efficient MCMC algorithm from Cooley and Sain (2008) based on recent work on Markov random fields (Rue and Held, 2005) and add temporal covariates and random effects, to account for temporal trends and variations. We use the point process model for extremes at the data level of the hierarchy, as it utilises more of the data than the GEV approach, and unlike the GPD approach, inference is invariant to the choice of threshold.

2.2. *Spatial discretisation.* Conventional Bayesian spatial models generally rely on the assumption that data are either gridded or they come from

fixed locations in space (see Banerjee et al. (2004)), where one or more observations are available at each location. Weather extremes however, behave like a spatial marked point process where both location of occurrence and magnitude are random. To utilise such Bayesian models, we propose for simplicity to discretise space by imposing a finite grid and to consider the minimum possible size Δ for each grid cell, to ensure that enough data are available for estimation in each cell. Inference should not be sensitive to the choice of grid spacing, provided it is fine enough (in the limit $\Delta \to 0$ one should obtain the original marked point process). Sensitivity analysis for Δ is a crucial part of the methodology (see section 4.1).

For spatial marked point processes, estimation is only possible after making assumptions about spatial (and temporal) structure. The assumption made by discretising is that conditional on a cell-specific random effect and possible covariates, the extreme events within each cell come from the same distribution. The cell-specific random effects are spatially dependent to allow for correlation between events in neighbouring cells. We also assume that the events can occur anywhere within the cells, with equal probability. Note that by re-defining space into discrete grid cells also provides a way of defining extremes in space: as values over a cell-varying threshold or as the $r = 1, 2, \ldots$ largest values, in fixed time periods.

More generally, conditional on a given spatial or spatio-temporal dependence structure between cells, events are modelled using an appropriate extreme value model. This is a hierarchical model where at the top of the hierarchy, random effects define a spatio-temporal process which modulates the process giving rise to the extremes.

2.3. *Model formulation.* **Data level:** For events within each cell, we assume the point process model for extremes (Coles, 2001) **(conditional on spatio-temporal random effects)**. For some high threshold u of the intensity, this model is parametrised in terms of the location, scale and shape parameters of the GEV distribution, namely μ, σ and ξ (see Appendix A). **For clarity, we use the notation $X \sim \mathbf{PP}(\mu, \sigma, \xi, u)$ when the magnitude X of an event can be described probabilistically by the point process (PP) model for extremes.**

Introducing spatial and temporal variation, let $X(s, t)$ characterise intensity of some weather event in grid cell $s \in \mathbb{S}$ at time $t \in \mathbb{T}$ where $\mathbb{S}$ and $\mathbb{T}$ are the space and time domains, each a fixed subset of 2-dimensional and 1-dimensional Euclidean space respectively. Define extreme events in each cell

as values over $u(s)$, an appropriately chosen threshold, e.g. a high empirical quantile. Note that the threshold is assumed constant over time. Extending Cooley and Sain (2008) we model the extremes in the following general way:

$$X(s,t)|\theta^{\psi},\gamma^{\psi},\beta_k^{\psi} \sim \mathrm{PP}(\mu(s,t),\sigma(s,t),\xi(s),u(s)) \tag{1}$$

$$\mu(s,t) = \beta_0^{\mu}+\sum_{k=1}^{K}\beta_k^{\mu}(s)z_k(s,t)+\theta^{\mu}(s)+\gamma^{\mu}(s,t) \tag{2}$$

$$\log(\sigma(s,t)) = \beta_0^{\sigma}+\sum_{k=1}^{K}\beta_k^{\sigma}(s)z_k(s,t)+\theta^{\sigma}(s)+\gamma^{\sigma}(s,t) \tag{3}$$

$$\xi(s,t) = \beta_0^{\xi}+\sum_{k=1}^{K}\beta_k^{\xi}(s)z_k(s,t)+\theta^{\xi}(s)+\gamma^{\xi}(s,t) \tag{4}$$

where $\psi=\mu,\sigma,\xi$ and $z_1(s,t),\ldots,z_K(s,t)$ are possible covariates that may depend on space (e.g. elevation, orography) or time (e.g. time trend, climate patterns) or both (e.g. temperature). The associated parameters $\beta_k^{\mu}(s)$, $\beta_k^{\sigma}(s)$, $\beta_k^{\xi}(s)$ may be spatially varying or constant, e.g. $\beta_k^{\mu}(s)=\beta_k^{\mu}$, depending on the application. $\theta^{\mu}(s)$, $\theta^{\sigma}(s)$ and $\theta^{\xi}(s)$ are spatial random effects that define the spatial variability of each parameter μ, σ and ξ across the cells, **after accounting for covariates**. In addition, $\gamma^{\mu}(s,t)$, $\gamma^{\xi}(s,t)$ and $\gamma^{\xi}(s,t)$ are random effects that introduce spatio-temporal variability which is additive to the purely spatial variability of $\theta^{\mu}(s)$, $\theta^{\sigma}(s)$ and $\theta^{\xi}(s)$. These random effects also allow for unobserved effects from missing covariates.

The model can be used to calculate the r-year return level, defined as the $(1-1/r)$th conditional (on t) quantile of $X(s,t)$ in cell s:

$$X_{1-1/r}(s,t)=\mu(s,t)+\frac{\sigma(s,t)}{\xi(s,t)}((-\log(1-1/r))^{-\xi(s,t)}-1). \tag{5}$$

Spatio-temporal process level: Following Cooley and Sain (2008), we define the purely spatial effects in the model as follows:

$$\big(\theta^{\mu}(s),\theta^{\sigma}(s),\theta^{\xi}(s)\big)'|\boldsymbol{U} \sim \mathbf{N}\big((U^{\mu}(s),U^{\sigma}(s),U^{\xi}(s))',\mathbf{diag}(\boldsymbol{\tau})^{-1}\big) \tag{6}$$

$$\boldsymbol{U}=(\boldsymbol{U}^{\mu},\boldsymbol{U}^{\sigma},\boldsymbol{U}^{\xi})' \sim \mathbf{N}(\mathbf{0},\boldsymbol{\Omega}^{-1}) \tag{7}$$

$$\beta_k^{\psi}(s) \sim \mathbf{N}(\nu_k^{\psi},\phi_k^{\psi}), \quad k=1,\ldots,K; \quad \psi=\mu,\sigma,\xi \tag{8}$$

where $\boldsymbol{U}^{\psi}=\big(U^{\psi}(1),\ldots,U^{\psi}(N)\big)$ for all $\psi=\mu,\sigma,\xi$. The mean of the random effects $\big(\theta^{\mu}(s),\theta^{\sigma}(s),\theta^{\xi}(s)\big)'$ depends on the spatially structured random effects $\boldsymbol{U}$. Allowing $\big(\theta^{\mu}(s),\theta^{\sigma}(s),\theta^{\xi}(s)\big)'$ to inherit their

spatial dependence through $\boldsymbol{U}$ in this way, greatly aids computation (see section 2.4) but also allows for unusual behaviour in certain cells, which would otherwise be smoothed by $\boldsymbol{U}$. The parameter $\boldsymbol{\tau}$ is a vector of precisions reflecting the spatial variability of each $\theta^{\psi}(s)$ other than that induced from $\boldsymbol{U}$.

As in Cooley and Sain (2008), vectors $\boldsymbol{U}^{\psi}$ are modelled jointly using a separable formulation (Banerjee et al., 2004, ch. 7), so that the precision matrix $\boldsymbol{\Omega} = \boldsymbol{T} \otimes \boldsymbol{W}$. $\boldsymbol{T}$ is a 3×3 positive definite symmetric matrix (that needs to be estimated) and $\boldsymbol{W}$ is an $N \times N$ proximity matrix defining the spatial proximity between the N cells. Therefore, $\boldsymbol{\Omega}$ is $3N \times 3N$.

Each vector $\boldsymbol{U}^{\mu}, \boldsymbol{U}^{\sigma}, \boldsymbol{U}^{\xi}$ is modelled by an Intrinsic AutoRegressive (IAR) spatial model (Besag et al., 1995). The IAR model uses the proximity matrix and a single unknown parameter to control the spatial dependency structure (see Appendix A.3 for the specification of an IAR prior). Here, there are three such parameters for each of $\boldsymbol{U}^{\mu}, \boldsymbol{U}^{\sigma}, \boldsymbol{U}^{\xi}$ and are found in the diagonal of $\boldsymbol{T}$. (Note that the value of $\boldsymbol{\tau}$ is conventionally fixed beforehand to avoid non-identifiability between $\boldsymbol{\tau}$ and the diagonal of $\boldsymbol{T}$ (Banerjee et al., 2004).) Dependence between $\boldsymbol{U}^{\mu}, \boldsymbol{U}^{\sigma}, \boldsymbol{U}^{\xi}$ is modelled using 3 parameters, the off-diagonals of $\boldsymbol{T}$, each controlling the strength of dependence. Allowing explicitly for this dependence can aid the MCMC estimation discussed in section 2.4, in terms of convergence to the posterior but also mixing of the MCMC samples.

The parameters $\beta_k^{\psi}(s)$, are chosen to be spatially variable but in an unstructured way: as random slopes with common global mean ν_k^{ψ} and variance ϕ_k^{ψ}. This ensures that $\beta_k^{\psi}(s)$ share information to aid estimation in cells with few events but less so compared to using a structured (IAR) spatial prior. The parameter ν_k^{ψ} reflects the overall effect of the associated covariate on the extreme events. The parameters $\beta_k^{\psi}(s)$ can of course be also modelled using an IAR model. In that case, $\beta_k^{\psi}(s) \sim \mathbf{N}(\nu_k^{\psi} + U^{\beta_k^{\psi}}(s), \phi_k^{\psi})$ where $\boldsymbol{U}^{\beta_k^{\psi}} = (U^{\beta_k^{\psi}}(1), \ldots, U^{\beta_k^{\psi}}(N))'$ has an IAR model and can be incorporated as an element to the left hand side of (7).

The spatio-temporal effects are modelled as follows: for $s \in \mathbb{S}$ and $t \in \mathbb{T}$,

$$\gamma^{\psi}(s,t)|\alpha^{\psi}(s), \gamma^{\psi}(s,t-1) \sim \mathrm{N}(\alpha^{\psi}(s)\gamma^{\psi}(s,t-1), \omega) \tag{9}$$

$$\gamma^{\psi}(s,0) \sim \mathrm{N}(0,\omega) \tag{10}$$

$$\alpha^{\psi}(s)|\boldsymbol{U}^{\alpha} \sim \mathrm{N}(\alpha_0^{\psi} + U^{\alpha^{\psi}}(s), \chi^{\psi}) \tag{11}$$

$$\boldsymbol{U}^{\alpha} = (\boldsymbol{U}^{\alpha^{\mu}}, \boldsymbol{U}^{\alpha^{\sigma}}, \boldsymbol{U}^{\alpha^{\xi}})' \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Phi}^{-1}). \tag{12}$$

for all $\psi = \mu, \sigma, \xi$. This is an autoregressive model in time where the autocorrelation parameter $\alpha^{\psi}(s)$ is spatially varying. This is a flexible way of capturing spatio-temporal behaviour in weather extremes that is present due to unmeasured factors. The autoregressive behaviour can emulate temporal dependence which may be spatially varying. The model for the spatial random effects in (12) is the same as in (7) with a different precision matrix $\boldsymbol{\Phi} = \boldsymbol{P} \otimes \boldsymbol{W}$, where $\boldsymbol{P}$ is a 3×3 precision matrix. Of course, the model in (9) may be simpler, i.e. the autocorrelation parameter may be spatially unstructured as in (8) or even constant in space. The simplest model is $\gamma^{\psi}(s,t) \sim \mathrm{N}(0,\omega)$ where both temporal and spatial effects are unstructured but still constrained by the single variance parameter ω. Note that for the same argument as with $\boldsymbol{\tau}$ in (6), the variance parameters χ^{ψ} in (12) would also need to be fixed before estimation.

2.4. *Estimation by Markov Chain Monte Carlo.* To complete the model specification, one needs to define the prior distribution for each parameter in the model. At the data level, we have parameters β_0^{ψ}, $\theta^{\psi}(s)$ and $\gamma^{\psi}(s,t)$ for all $\psi = \mu, \sigma, \xi$ as well as $\beta_k^{\psi}(s)$ for $k = 1, \ldots, K$. The prior distributions for $\theta(s)^{\psi}$, $\beta_k^{\psi}(s)$ and $\gamma^{\psi}(s,t)$ are given in (6), (8) and (9) respectively. For spatially constant parameters β_k^{ψ} we can assume flat Gaussian prior distributions with zero mean and large variance. Random effects $\theta^{\psi}(s)$, $\beta_k^{\psi}(s)$ and $\gamma^{\psi}(s,t)$ are sampled by Metropolis-Hastings, specifically using a random walk sampler. The intercepts β_0^{ψ} are given Gaussian prior distributions with large variance, and means $(\bar{\mu}, \log\bar{\sigma}, \bar{\xi})$, where each of these is the mean of the maximum likelihood estimates per cell of each parameter, obtained by fitting independent point process models. The intercepts can be sampled from their full conditionals using Gibbs sampling, by considering them as intercepts in the mean for each $\theta^{\psi}(s)$.

The prior distributions for $\left(\boldsymbol{U}^{\mu}, \boldsymbol{U}^{\sigma}, \boldsymbol{U}^{\xi}\right)'$ and $\left(\boldsymbol{U}^{\alpha^{\mu}}, \boldsymbol{U}^{\alpha^{\sigma}}, \boldsymbol{U}^{\alpha^{\xi}}\right)'$ are given in (7) and (12) respectively. For $\boldsymbol{T}$ and $\boldsymbol{P}$ we use a Wishart prior with 3 degrees of freedom (uninformative) and a mean that relates to the variability of μ, σ and ξ across cells (see section 4 for particular details). Using these priors one may draw samples of $(\boldsymbol{U}^{\mu}, \boldsymbol{U}^{\sigma}, \boldsymbol{U}^{\xi})'$, $(\boldsymbol{U}^{\alpha^{\mu}}, \boldsymbol{U}^{\alpha^{\sigma}}, \boldsymbol{U}^{\alpha^{\xi}})'$, $\boldsymbol{T}$ and

$\boldsymbol{P}$ from their full conditionals using a Gibbs sampler, utilising the specific techniques in Cooley and Sain (2008). Parameter ν_k^ψ is given a flat Gaussian prior with zero mean and large variance and the prior distribution $\pi()$ for ϕ_k^ψ is chosen so that $\pi(\phi_k^\psi) \propto 1/\phi_k^\psi$ which enables one to draw both parameters from their full conditionals: Gaussian for ν_k^ψ and scaled inverse-χ^2 for ϕ_k^ψ (Gelman et al., 2004).

Note that when the IAR model is used as a prior like here, then it is improper: the distribution does not integrate to 1. So, to make the intercept terms β_0^ψ identifiable, the rows of $\boldsymbol{W}$ must sum to zero. This in turn imposes the restriction that $\sum_s U^\psi(s) = 0$ for all $\psi = \mu, \sigma, \xi$. In the application of this model to cyclone extremes in section 4, spatial proximity is based on nearest neighbours so that off-diagonal elements of $\boldsymbol{W}$ are $w_{i,j} = -1$ if cells i and j are adjacent and $w_{i,j} = 0$ otherwise, whereas diagonal elements $w_{i,i} = -\sum_{i \neq j} w_{i,j}$ (see Bailey and Gatrell (1995) p.261–262 for other examples of proximity measures).

3. Data. The methodology is illustrated by application to extreme extra-tropical cyclones. Objective feature-identification software (Hodges (1994, 1995, 1999)) was used to extract cyclone tracks from 6-hourly National Center for Environmental Prediction Climate Forecast System (NCEP-CFS) re-analysis data (Saha and Co-authors, 2010) available over the period 1979-2009. Individual cyclone tracks are identified by tracking local maxima in relative vorticity just above the boundary layer (at around 1.5km altitude above sea-level). The minimum sea-level pressure (MSLP) and its location are recorded every 6 hours throughout the lifecycle of each cyclone. We use sea-level pressure as a measure of cyclone intensity mainly because this variable is well-observed and has smooth variation during the lifetime of a cyclone, unlike other possible variables such as wind speed or vorticity. Figure 1a shows a map of cyclone tracks defined by 6-hourly MSLP recordings for a period with high cyclone activity. Only a subset of tracks is plotted: ones where any 6-hourly MSLP value reached below 960hPa, considered to be an extreme (low) value. Typical damaging cyclones over Europe reach values in the range 940-970hPa, for instance Anatol: 953hPa (Ulbrich et al., 2001) and Kyrill: 962hPa (Mitchell-Wallace and Mitchell, 2007), whereas the lowest ever recorded Braer cyclone reached 913hPa off the North-West of Scotland in January 1993 (Odell et al., 2010).

Note that wind speed could also have been used, but exploratory analysis suggests that extreme sea-level pressure and wind speed are strongly dependent, as to be expected from simple balance

arguments. Above the surface boundary layer outside equatorial regions, centrifugal and Coriolis forces are almost completely balanced by the pressure gradient force. Hence, wind speeds above the boundary layer in extra-tropical cyclones are proportional to pressure gradients (gradient wind balance). Pressure gradients in turn are strongly related to the minimum sea-level pressure of the storm since extra-tropical cyclones have similar synoptic spatial dimensions (the so-called Rossby scale). Hence, from such simple dynamical meteorology arguments, minimum pressure and maximum wind speeds are expected to be extremally dependent and so will convey similar information. Figure 3a shows a scatter plot of maximum (6-hourly) wind speeds at about 1.5km above the surface (on the 925hPa pressure surface) against negated MSLP (obtained by multiplying MSLP by -1) - there is strong positive association with the loess smoother indicating a nearly linear relationship. A line at -960hPa has been added along with a line at 45m/s, where this value is the equivalent quantile of maximum wind speed that 960hPa is for MSLP. To better visualise extremal dependence, Figure 3b shows the empirical copula obtained by producing a scatter plot of the empirical probabilities $q_t = (\mathbf{rank}(z_t)-1)/(n-1)$ **(Stephenson et al., 2008), where** z_t **is the 6-hourly measure of intensity and** n **is the total number of 6-hourly measurements from all cyclone tracks. This transforms out the margins to uniform distributions since the empirical probabilities are estimates of the cumulative distribution functions,** $F_Z(z)$ **and** $F_W(w)$**. Strong dependence of the extremes is evident from the convergence of the points in the upper right hand corner of the graph.**

Figure 3c shows estimates of the extremal dependence measure χ (Coles et al., 1999) which is defined as $\lim_{p\to 1} \Pr(F_Z(z) > p | F_W(w) > p)$. As $p \to 1$, $\chi \to 0$ implying asymptotic independence, so we also show $\bar{\chi}$, a measure of strength of extremal dependence in Figure 3d. Since $\bar{\chi}$ does not tend zero, we conclude that there is strong positive association at the extremes of negated MSLP and maximum wind speed, so either variable could be used to investigate extremes (see Appendix A.2 for details on χ and $\bar{\chi}$).

Figures 1b and 1c show plots of MSLP against latitude and longitude respectively, for two particular cyclone tracks in the 1989-1990 winter (Figure 1a). The plots illustrate the tendency of intense cyclones to move in a west-to-north direction but also the fact that MSLP decreases (cyclone deepening) as the cyclone propagates in space and time, to reach a mini-

mum (which we assume approximates the unobserved value of the cyclone nadir) before it starts increasing again until the end of the life cycle. Understanding the limiting strength of the nadirs is an important aspect in the study of extra-tropical cyclones. However, the rate of growth of cyclones depends on the large scale atmospheric environment that they pass from, so the pressure limit of cyclone nadirs will vary with the spatial location of the cyclone. By only considering the nadir from each cyclone track, we are focusing on a fundamental limiting property of cyclones, namely how deep they can get in general rather than how deep they can get in specific spatial locations. Of course, how deep the cyclones get is spatially variable for the aforementioned reasons, which is why we need to adopt a model with spatial components. In other words, we are focusing on spatial variation in cyclone intensity rather than maximum local cyclone impact.

The analysis of nadirs only also helps to eliminate dependency between successive 6-hourly MSLP measures and reduces the amount of data from 313,557 6-hourly measurements to 17,230 nadirs. Figure 2a shows the (re-analysis) nadir from each track, in the Atlantic region where dots in black are nadirs with sea-level pressure lower than 960hPa. However, a single value for the threshold defining the extremes is not appropriate and the definition of extremeness should vary spatially. For example, a damaging cyclone in the Mediterranean is likely to be considered a weak one over Scandinavia.

Note that we use cyclone tracks from a re-analysis data set mainly because generally cyclone track observations for the extra-tropics are not readily available. However, re-analysis data are output from climate models with assimilated historical observational data. There is much smoothing/interpolation of the observational data when creating a re-analysis dataset so the interpretation of any results obtained here, is conditional on the effects of such smoothing.

4. Application to extreme extra-tropical cyclones.

4.1. *Spatial Grid.* Conventionally, extreme value modelling is applied to the upper tails so the nadirs are negated to obtain variable $X(s,t)$, where s refers to the grid cell and t refers to time. We may think of $X(s,t)$ as the depth of a cyclone so that high values of $X(s,t)$ correspond to low values of MSLP. We divided the domain in Figure 2a into $N = 150$ $5^\circ \times 10^\circ$ grid cells. The threshold $u(s)$ in each cell was defined as the empirical 90th quantile of $X(s,t)$. We performed exploratory threshold analysis using mean residual life plots (Coles, 2001), ensuring that the 90th empirical quantile was an appropriate threshold choice. Figure 2b shows the map of the extremes (1,736 nadirs) and Figure 2c shows the map of $u(s)$. Note that in Figure

2c, three cells are highlighted: cells containing coordinates (5.2°E, 60.2°N), (0°E, 5°N) and (3.5°W, 40.2°N) marked in white crosses. These coordinates relate to the cities of Bergen, London and Madrid respectively, and will be used throughout the paper for illustration of results as they adequately span Europe in terms of latitude.

We explored the sensitivity of the results to cell size. Both a purely spatial model (i.e. model (1) with just the spatial random effects $\theta(s)^\psi$), and a stationary model (i.e. one without any random effects) were implemented for different grid configurations. For each model, the 100-year return level **(i.e. the level exceeded by the annual maximum in any particular year with probability 0.01)** of $X(s,t)$ was calculated **using (5)**, for each of the three coordinates marked in Figure 2c. For the stationary model, Figures 4a, 4b and 4c show the posterior mean of the 100-year return level against the number of cells in each grid configuration, along with 95% credible intervals. Figures 4d, 4e and 4f show the equivalent plots for the spatial model. The spatial model shows convergence of the return value as the number of cells increases, although this varies slightly due to different information available for estimation in the particular grid cell. The spatial random effects pool information spatially, ensuring the evident convergence whereas the stationary model ignores neighbouring cells resulting in failure to converge, especially over London. Finally, credible intervals from the spatial model are notably smaller which again reflects the fact that information from neighbouring cells was used; note that the intervals are skewed. We chose $N = 150$ cells for the analysis which ensured that all cells have an adequate number of nadirs (ranging from 14 to 376).

4.2. *Model specification.* The following model formulation was used:

$$X(s,t)|\theta^\psi(s),\beta_2(s) \sim \mathrm{PP}(\mu(s,t),\sigma(s,t),\xi(s),u(s)) \tag{13}$$

$$\mu(s,t) \sim \beta_0^\mu + \beta_1^\mu z_1(t) + \beta_2(s) z_2(t) + \theta^\mu(s) \tag{14}$$

$$\log(\sigma(s,t)) \sim \beta_0^\sigma + \beta_1^\sigma z_1(t) + \theta^\sigma(s) \tag{15}$$

$$\xi(s) \sim \beta_0^\xi + \theta^\xi(s) \tag{16}$$

for $\psi = \mu, \sigma, \xi$. At the spatial process level:

$$\big(\theta^\mu(s),\theta^\sigma(s),\theta^\xi(s)\big)'|\boldsymbol{U} \sim \mathrm{N}((U^\mu(s),U^\sigma(s),U^\xi(s))', \mathrm{diag}(\boldsymbol{\tau})^{-1}) \tag{17}$$

$$\boldsymbol{U} = (\boldsymbol{U}^\mu,\boldsymbol{U}^\sigma,\boldsymbol{U}^\xi)' \sim \mathrm{N}(\mathbf{0},\boldsymbol{\Omega}^{-1}) \tag{18}$$

$$\beta_2(s) \sim \mathrm{N}(\nu,\phi^2). \tag{19}$$

We chose not to include any spatio-temporal random effects. Cyclone nadirs occur irregularly in time so the autoregressive model seemed redundant and

in addition, long-term temporal variability was allowed through appropriate covariates. Covariate selection was obtained by adding explanatory variables to a "null" model (the model in (13) without any temporal covariates). Model comparison was based on the Deviance Information Criterion (DIC), a model selection criterion for Bayesian models implemented using MCMC (Spiegelhalter et al., 2002), and by investigating whether posterior distributions of associated parameters are centred at zero with relatively large variance. The model in (13) was first implemented with the addition of latitude, longitude, latitude squared, longitude squared and an interaction term between longitude and latitude as covariates in both $\mu(s,t)$ and $\log(\sigma(s,t))$. This allows for large scale spatial trends leaving the local spatial dependence to the random effects. It also relaxes the assumption of complete spatial randomness of extreme events within a cell, both in terms of occurrence and intensity.

The only significant reduction in the DIC occurred when latitude was included in both $\mu(s,t)$ and $\log(\sigma(s,t))$, denoted by $z_1(t)$ in (14) and (15). The posterior distributions of parameters relating to longitude, the squared terms and the interaction, had means and medians very close to zero. In principle, non-parametric surfaces can also be considered for smoothing large scale spatial trends (see Davison et al. (2012) p.173 for references) but this was not deemed necessary here. No covariates were considered for the shape parameter $\xi(s)$ since this is a particularly difficult parameter to estimate, however it was allowed to vary between cells.

To quantify the effect of large scale climate patterns, two climate indices were considered as covariates: the North Atlantic Oscillation (NAO) and the East Atlantic Pattern (EAP), both of which have been shown to be influential for extra-tropical cyclones (Mailier et al., 2006; Seierstad et al., 2006; Pinto et al., 2009; Nissen et al., 2010). Variable $z_2(t)$ in (14) is the NAO which was given a cell-varying parameter $\beta_2(s) \sim \mathrm{N}(\nu, \phi^2)$ so that ν is the overall NAO effect on extreme cyclones. The DIC reduction for the model with just $z_1(t)$ and the model with both $z_1(t)$ and $z_2(t)$ was about 100 points implying model improvement when NAO is included (the standard error of the deviance form both models was around 30 indicating that the difference is substantial). The EAP was also included as $\beta_3(s)z_3(t)$, but was deemed not significant. Lastly, both the NAO and EAP were included in the formulation for $\sigma(s,t)$ in (15), however the DIC and posteriors of associated parameters suggested this did not improve model fit.

It is well known that the NAO has influence on the development of extra-tropical cyclones (Pinto et al., 2009). By definition, the NAO index is standardised to have zero mean and unit variance, and here we define NAO in

terms of 5-day non-overlapping averages from 1979-2009. Figure 5 shows the time series of NAO and Figure 5b shows the histogram of NAO where the values of 2 and -2 are marked, as we consider these as high and low NAO threshold values throughout the rest of this paper. Figure 5c and 5d show extreme values of $X(s,t)$ for which NAO ≥ 2 and NAO ≤ -2 respectively. There is a clear North-South pattern in the Central Atlantic implying NAO has a notable effect on extreme cyclones.

4.3. *Model implementation.* As mentioned in section 2.3, the values for $\boldsymbol{\tau}$ are fixed and here we set $\boldsymbol{\tau} = (0.1, 10, 100)'$. These values were chosen by fitting independent point process models in each cell and investigating the level of variability between cells for $\mu(s)$, $\log(\sigma(s))$ and $\xi(s)$, to reflect the difference in scale for the three parameters but also to make sure that most of the variability is modelled by the random effects $\boldsymbol{U}^{\mu}, \boldsymbol{U}^{\sigma}, \boldsymbol{U}^{\xi}$ and not $\boldsymbol{\tau}$. If values in $\boldsymbol{\tau}$ are too small, then the variability in each $\theta^{\psi}(s)$ is forcibly large and may cause problems in estimating the diagonal of T which relates to the variability of each $\boldsymbol{U}^{\psi}$. Sensitivity analysis was performed to ensure these values have little effect on inference (not shown for conciseness).

For the 3×3 precision matrix $\boldsymbol{T}$, the Wishart prior was given the following mean: $\text{diag}(0.02, 4, 40)'$. As with $\boldsymbol{\tau}$, these values were calibrated by fitting independent point process models and were chosen to reflect the associated levels of variability for each of $\mu(s)$, $\log(\sigma(s))$ and $\xi(s)$.

The model in (13) was implemented in R (R Core Team, 2012) using three parallel MCMC chains. These were ran on a workstation with a 3.07GHz i7 processor and the processing speed for each chain was 30 seconds for 1000 samples. A total of 50,000 samples were collected per chain and thinned by 5 to reduce auto-correlation. After thinning, the first 3000 samples from each chain were discarded based on a trace plot of deviance (minus twice the log-likelihood) shown in Figure 6a. Convergence in the deviance is a good indication of convergence to the joint posterior of all parameters (Gelman et al., 2004). Summarising, 21,000 posterior samples were used to calculate posterior distribution statistics for the parameters. Figure 6b shows an example trace plot of $\xi(s)$ for the grid cell containing the London coordinate.

5. Results. Posterior distributions for global parameters are summarised in Table 1. Latitude has a positive linear effect on both the location and log-scale parameters of extreme cyclone depth $X(s,t)$. The overall NAO effect ν is positive in agreement with findings from previous studies (Pinto et al., 2009). To formally assess MCMC convergence, the Gelman and Rubin $\hat{R}$ multi-chain diagnostic was used for each of our model parameters (Gelman et al.,

2004). The $\hat{R}$ values for each parameter in Table 1 are all close to unity suggesting convergence.

Figure 7 shows posterior means and standard deviations of $\mu(s,t)$, $\sigma(s,t)$ and $\xi(s)$. Much of the spatial structure in the extreme nadirs comes from the location and scale parameters. The posterior means for the shape parameter $\xi(s,t)$ are more uniform and generally negative apart from one cell over Iceland. Exploring this further, the two deepest nadirs in the re-analysis occurred in this cell, and they are considerably lower than the rest of the nadirs in the vicinity. A return level plot from the particular cell indicated that the two nadirs (one of them being from the record breaking Braer cyclone) unduly influenced the sign of the shape parameter significantly. This has been quantified by removing those two points and re-fitting the model, however this being an analysis of extremes, it makes little sense to remove such values.

A negative shape parameter implies that the distribution of extreme cyclone depth $X(s,t)$, at time t and cell s, has an upper bound given by $\sigma(s,t)/\xi(s) - \mu(s,t)$. Here this corresponds to a lower limit on nadir sea-level pressure. Many of the posteriors for $\xi(s)$ do have some mass over the positive real line (see for instance Figure 6b). However, except for the Iceland cell, the negative masses for $\xi(s)$ are all greater than 0.5 therefore we can use the negative posterior $\xi(s)$ samples, to obtain a conditional posterior distribution for the estimated lower limit. The posterior means of the estimated lower limits are shown in Figure 8c for NAO $= 0$. The estimated lower limit for the cell containing Bergen is 890.0hPa [193.0, 932.6], for the London cell it is 943.0hPa [714.8, 959.4] whereas in the Madrid cell it is 953.5hPa [537.9, 978.7]. The 95% credible intervals are skewed and noticeably wide which is to be expected given the quantity we are trying to estimate is effectively the 100th percentile. **Note that the lower bounds on some of these intervals are too low to be physically plausible and this reflects the fact that the statistical model is not constrained by physical mechanisms**. Note also that there is considerable literature focusing on the problem of estimating upper/lower bounds of distributions. See de Haan and Ferreira (2006, ch. 4) for a detailed discussion and a description of both maximum likelihood and moment estimators for bounds, arising from extreme value distributions. In addition, Einmahl and Magnus (2008) provide refined estimators for bounds of world records in athletics and their respective sampling distributions.

The posterior means of the NAO effects $\beta_2(s)$, are shown in Figure 7g along with associated standard errors in Figure 7h. A positive effect is prominent in the area where cyclones deepen the most: in the vicinity of Iceland,

northern Europe and Scandinavia. A negative effect is also apparent, effectively over Spain and the Azores. This North-South NAO effect in the central Atlantic is consistent with the exploratory diagnostics in Figures 5c and 5d. Maps of the estimated lower limit for NAO $= -2$ and NAO $= 2$ are given in Figures 8a and 8b. To better see the effect of NAO on the estimated lower limit, Figure 8d shows the difference in hPa between the estimated lower limits for NAO $= 2$ and NAO $= -2$. The difference can get up to 25hPa in the area where NAO has a the biggest effect, i.e. northern Europe and Scandinavia.

Figure 9 shows return level plots of $X(s,t)$ for some selected grid cells. The cells are those shown in Figure 2c, namely the cells containing Bergen, London and Madrid. There are two plots for each cell, one for NAO $= 2$ (top panel) and one for NAO $= -2$ (bottom panel). Note that this is not a goodness of fit test, as each point in these plots (the recorded value) is associated with a different value of NAO, namely the one that occurred at that time of recording, whereas the return level curves are calculated at NAO $= \pm 2$. A positive NAO effect is noticeable in the Bergen cell whereas a negative effect is prominent in the Madrid cell confirming the NAO North-South effect. No distinguishable NAO effect is evident in the London cell. The horizontal line in each plot is the estimated cyclone depth limit suggesting that for all three cells, nadirs can get much deeper than the ones that have been recorded in the re-analysis tracks.

The fitted model, may be used to quantify the possibility of observing deeper nadirs than the ones in the re-analysis tracks. Specifically, we consider the quantity $\Pr\big(X(s,t) > x_m(s)\big)$ where $x_m(s)$ is the negated minimum recorded nadir in grid cell s for the 30-year period. Note that this is equivalent to describing how unusual the recorded depth was, rather than the probability of ever getting deeper that the recorded **30-year minimum nadir**. We transform the GEV parameters (Katz et al., 2005) to reflect the distribution of 30-year, rather than yearly depth values: $\tilde{\sigma} = \sigma\delta^{\xi}$ and $\tilde{\mu} = \mu + \tilde{\sigma}(1-\delta^{-\xi})/\xi$ where $\delta = 30$. Figure 10a shows $\Pr(X(s,t) > x_m(s))$ for values of NAO associated with $x_m(s)$. There is a high probability of cyclone nadirs being deeper than what has been seen in the re-analysis data especially over western Europe. Figure 10b shows $\Pr(X(s,t) > x_m(s))$ for NAO $= 2$ indicating that for a positive NAO phase there is high probability of deeper nadirs over Europe, Iceland and Scandinavia. For NAO $= -2$, Figure 10c attributes high probability of deeper nadirs in Southern central Atlantic, specifically over Spain, Portugal, west of France but also over the Azores region. Furthermore, Figure 10d shows the difference in hPa between the estimated depth limit for sea-level pressure (Figure 10a) and $x_m(s)$ for

each cell. For most cells, the difference is in the range of 10–50hPa while for cells over Iceland the range is 80-110hPa reinforcing the fact that the 30-year re-analysis is not long enough to include nadir depths that are close to the estimated limit.

For model checking, predictions of $X(s,t)$ were compared to recorded values. The posterior predictive distribution of $X(s,t)$ is defined as $\Pr(\tilde{X}(s,t)|X(s,t))$, where $\tilde{X}(s,t)$ refers to future observable values. For our purposes, predictions $\tilde{X}(s,t)$ are obtained from the model by using the exact same covariate and random effect values associated with the observed $X(s,t)$. We obtained samples from the posterior predictive distribution of a) the deepest 30-year nadir and b) the deepest yearly nadir, in each cell. This is equivalent to sampling from a GEV distribution, at each posterior MCMC sample of the model parameters $\mu(s,t), \sigma(s,t)$ (or $\tilde{\mu}(s,t), \tilde{\sigma}(s,t)$ for the yearly values) and $\xi(s)$, where s and t correspond to the associated recorded values. In Figure 11a, recorded values are plotted against predictions (calculated as means from the posterior predictive distributions) of the 30-year deepest nadirs in each cell, along with 95% prediction intervals. Figure 11b is the equivalent plot for the yearly deepest nadirs in each cell. Both the 30-year and yearly deepest nadirs are adequately predicted by the model. We also calculate the probabilities $\Pr(X(s,t) < x_{obs}(s,t))$ where $x_{obs}(s,t)$ are the observed values in cell s. If the model is a good fit, these probabilities should be uniformly distributed on (0,1) and to assess that, posterior means along with 95% credible intervals of these probabilities are used to produce empirical quantile-quantile plots in Figure 12 for the Bergen, London and Madrid cells. The plots suggest adequate model fit, the London cell having the most deviations from the 45° line. Note that what we are plotting here are commonly known as probability integral transform (PIT) values, often used in forecast verification, see for instance Gneiting et al. (2007) and references therein. Although histograms are the more conventional way of displaying PIT values, here we only have a few data points for each cell so we use quantile-quantile plots.

6. Conclusions. We have identified a class of problems relevant to environmental extreme events. Firstly, extreme events are inherently rare so data coverage in many applications is poor. In addition, both occurrence and intensity of these extreme phenomena are spatially heterogeneous but also temporally varying, possibly even modulated by large-scale climate patterns. As a solution, we have extended the extreme Bayesian hierarchical methodology proposed in Cooley and Sain (2008). Our model allows for both spatial and temporal components in terms of covariates and random effects, and it

is based on discretising space to deal with the issue of irregular occurrence in space.

The model was successfully applied to re-analysis cyclone data in what we believe to be the first study that simultaneously models both the spatial and temporal structure of extreme extra-tropical cyclones. The data consisted of 17,230 extra-tropical cyclones objectively identified over the Atlantic in 6-hourly re-analyses from 1979-2009. Using 1) spatial random effects, 2) latitude as a covariate, and 3) a 150 cell spatial regularisation, spatial variation was adequately modelled in the extremal behaviour of the cyclones. Local spatial pooling across neighbouring cells gave much narrower credible intervals than a model having no spatial random effects, and ensured that the intervals converged as the cell size was reduced. The North Atlantic Oscillation was used as a covariate in the model and was was found to have a significant effect on extremal storm behaviour especially over N. Europe and the Iberian peninsula. The model was used to infer that there is a high probability of having nadirs with sea-level pressure lower than that corresponding to the observed deepest nadirs over 1979-2009, especially when the NAO index is positive. Estimates of lower bounds on nadir sea-level pressure are typically 10-50hPa below the minimum values found for cyclones over the period 1979-2009. The model fits suggest that in northern Europe, there is high probability of nadir sea-level pressure going below the deepest recorded values for a positive NAO phase, whereas for a negative NAO phase it is in southern Europe that this probability is high. Note that this is not the probability of ever experiencing deeper nadirs, but rather the probability of observing deeper nadirs in any 30-year period where the atmospheric conditions are the same as in the 30-year period of re-analysis data.

Although this is a first step towards studying the spatio-temporal behaviour of extreme cyclones, the analysis relies on assumptions which may over-simplify the problem: 1) the creation of an artificial grid, 2) the choice of threshold in each cell and 3) the subjective choice of spatial proximity. The choice of the grid is a potential weakness which can introduce bias, as both the number of cells and their shape are subjectively chosen. Techniques such as Dirichlet tessellation or Delaunay triangulation (Illian et al., 2008) may be useful for defining a more optimal grid in the sense of discretising space according to the data rather than subjectively. The shape of the cells is particularly important if one is interested in modelling data along cyclone tracks rather than individual points as in our application. If interest was in the relative cyclone impact at locations, one might want to use cell-specific rather than cyclone-specific nadirs rendering the rectangular cells inappropriate. Hexagonal cells would be more appropriate since they

avoid the problem of a track passing very closely to the corners of the rectangular cells, a technique successfully applied to tropical cyclone analysis in Elsner et al. (2012). Threshold choice in each cell may also prove to be an issue. Ideally, model fit should be one of the criteria for choosing the threshold. For the application in this paper, three different thresholds were considered: the 85%, 90% and 95% quantile in each cell. The diagnostic plots of model fit in Figure 11 were effectively the same between the 90% and 95% quantile thresholds indicating that higher thresholds did not necessarily improve model fit. Comparing the 85% quantile with the 90% quantile indicated worse model fit for the lower threshold, as it became apparent from checking individual return level plots for each cell, like the ones in Figure 9. We selected the 90% quantile for model application. A possible way to avoid choosing the threshold altogether is to use a technique where the threshold is estimated from the data. For example, we have explored the possibility of using a mixture model as in Frigessi et al. (2003) where the threshold is estimated but which also allows all available data (not just the extremes) to be used for each cell, which in turn allows the implementation of a finer grid. The proximity structure used to define the covariance matrix of the spatial random effects is also an assumption which can affect the amount of spatial smoothing. The spatially random occurrence of cyclone nadirs was 'marginalised' here by dividing the region into grid cells, whereas one should ideally try to model both the spatial occurrence and intensity at the same time (spatial marked point process).

Despite these caveats, the methodology developed in this study is generic and so will hopefully be useful for studying extreme behaviour of other extreme weather events such as tropical cyclone wind speed maxima, storm-related peak precipitation, etc.

APPENDIX A

A.1. Point process model for extremes. The point process model for extremes, involves a bivariate variable $Y = (X, T)$ with $T \in [0, 1]$ being a scaled random variable associated with time and $X \in \mathbb{R}$ a random variable associated with intensity. The model is a marked point process which for $X > u$ under some linear normalisation and mixing criteria (Smith, 1989), behaves like a non-homogeneous Poisson process with intensity function

$$\lambda(x,t) = \frac{1}{\sigma}\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi-1}$$

provided that $1 + (\xi/\sigma)(x - \mu) > 0$. The intensity function $\lambda(x, t)$ is zero for $1 + (\xi/\sigma)(x - \mu) < 0$. Note that u is a high threshold for x that defines the

extremes. Like the GPD, the point process approach uses more data than the GEV which uses only block maxima, however unlike the GPD inference is invariant to the threshold u (Coles, 2001). This approach combines the desired effects of the GEV and GPD with an additional benefit that the exceedance rate is explicitly modelled in terms of the mean number of exceedances in the time interval $[t_1, t_2]$:

$$\Lambda([t_1, t_2] \times (u, \infty)) = (t_2 - t_1) \left[1 + \xi \left(\frac{u - \mu}{\sigma}\right)\right]^{-1/\xi}.$$

The likelihood of the point process, given observations $y_i = (x_i, t_i)$ in the region $[0, 1] \times (u, \infty)$, is given by

$$L(\mu, \sigma, \xi; \boldsymbol{x}, \boldsymbol{t}) = \exp\left\{-n_y \int_0^1 \int_u^\infty \lambda(x, t) dx dt\right\} \prod_i \lambda(x_i, t_i) \tag{20}$$

$$= \exp\left\{-n_y \left[1 + \xi \left(\frac{u - \mu}{\sigma}\right)\right]^{-1/\xi}\right\} \prod_i \lambda(x_i, t_i) \tag{21}$$

where n_y is the number of years of observed data so that parameters μ, σ and ξ correspond to the GEV distribution of yearly maxima. The likelihood contribution from a single event (x_i, t_i) is

$$\mathrm{PP}(\mu, \sigma, \xi, u) = \exp\left\{-n_y [t_i - t_{i-1}] \left[1 + \xi \left(\frac{u - \mu}{\sigma}\right)\right]^{-1/\xi}\right\} \lambda(x_i, t_i),$$

for $i = 0, \ldots, n$ where n is the number of events. Note that $t_0 = 0$ and that the likelihood contribution, for the time interval between the last event occurrence and $t = 1$, is the probability of no events in the interval, i.e.

$$\exp\left\{-n_y [1 - t_n] \left[1 + \xi \left(\frac{u - \mu}{\sigma}\right)\right]^{-1/\xi}\right\}$$

For clarity of exposition, we use the symbolic notation $Y \sim \mathrm{PP}(\mu, \sigma, \xi, u)$ to mean that random variable Y behaves probabilistically according to the point process (PP) model for extremes above threshold u.

The conditional model in (1) is implemented using the likelihood (21) for each cell. However, because of the temporal covariates, the outermost integral over time in (20) is impossible to calculate analytically unless one knows explicitly how the covariates evolve in time. A remedy is to approximate the

integral: divide the time range in small intervals $0 = k_1, k_2, \ldots, k_J = 1$ and assume the function is constant in each interval. The integral

$$\int_0^1 \left[1 + \xi(s)\left(\frac{u(s) - \mu(s,t)}{\sigma(s,t)}\right)\right]^{\frac{-1}{\xi(s)}} dt$$

is thus approximated by the Riemann sum

$$\frac{1}{J}\sum_{i=1}^{J}\left[1 + \xi(s)\left(\frac{u(s) - \mu(s,k_i)}{\sigma(s,k_i)}\right)\right]^{\frac{-1}{\xi(s)}}$$

where J is the number of intervals. In practice, J is determined by observations of the covariates for all data not just the extremes.

A.2. Measures of extremal dependence. Consider random variables Z and W. The measure of spatial dependence $0 < \chi < 1$ is defined as

$$\chi = \lim_{p\to 1} \Pr(F_Z(z) > p | F_W(w) > p)$$

where and F_Z and F_W are the respective distribution functions of Z and W. If $\chi = 0$, the two variables are asymptotically independent. However, even if $\chi = 0$, the two variables may still exhibit extremal dependence which can be summarised by $-1 \leq \bar{\chi} \leq 1$, where

$$\bar{\chi} = \lim_{p\to 1} \frac{2\log \Pr\big(F_Z(z) > p\big)}{\log \Pr\big(F_Z(z) > p, F_W(w) > p\big)} - 1.$$

Roughly, $\bar{\chi}$ measures the 'speed' at which χ approaches zero and if $\chi = 0$ and $\bar{\chi} \neq 0$, the variables are asymptotically independent. **Coles et al. (1999) advocate the use of both χ and $\bar{\chi}$ as indicators of extremal dependence, providing complementary information on different aspects of that dependence.**

A.3. Intrinsic AutoRegressive priors. Consider a grid with N cells. If the random effect $\boldsymbol{\phi} = (\phi(1), \ldots, \phi(N))'$ is assumed to have an IAR prior then $\boldsymbol{\phi} \sim N(0, (\tau \boldsymbol{W})^{-1})$ where $\boldsymbol{W}$ is the adjacency matrix and the conditional distribution for each $\phi(s)$ given the rest is given by:

$$\phi(s)|\boldsymbol{\phi}(-s) \sim N\left(\bar{\phi}(s), \frac{1}{\tau m(s)}\right)$$

where $\boldsymbol{\phi}(-s)$ is $\boldsymbol{\phi}$ excluding $\phi(s)$; $\bar{\phi}(s)$ is the average of $\boldsymbol{\phi}(-s)$ that are adjacent to $\phi(s)$ and $m(s)$ is the number of those adjacencies.

ACKNOWLEDGEMENTS

This research was kindly supported by the project RACEWIN, which was funded by AXA research fund. We kindly thank Dan Cooley for providing the R code which was adapted for this study and for insightful comments. We thank Kevin Hodges for providing the cyclone track data for NCEP-CFS and also Simon Brown and Phil Sansom for their constructive and stimulating comments.

TABLE 1
Parameter summary.

Parameter	Prior	Posterior Mean (s.e.)	95% Cr.I.	$\hat{R}$
β_1^{μ}: Latitude	$\mathrm{N}(0, 100)$	4.71 (0.62)	[3.61,5.94]	1.13
β_1^{σ}: Latitude	$\mathrm{N}(0, 100)$	0.12 (0.07)	[0.00,0.25]	1.03
overall NAO effect ν	$\mathrm{N}(0, 100)$	1.21 (0.24)	[0.77,1.66]	1.01
variance NAO effect ϕ^2	$\propto 1/\phi^2$	5.6 (1.85)	[3.09,10.25]	1.00
β_0^{μ}	$\mathrm{N}(-944.1, 100)$	-987.4 (0.51)	[-988.5,-986.5]	1.04
β_0^{σ}	$\mathrm{N}(5.7, 100)$	2.03 (0.06)	[1.91,2.15]	1.05
β_0^{ξ}	$\mathrm{N}(-0.19, 100)$	-0.13 (0.03)	[-0.18,-0.07]	1.10

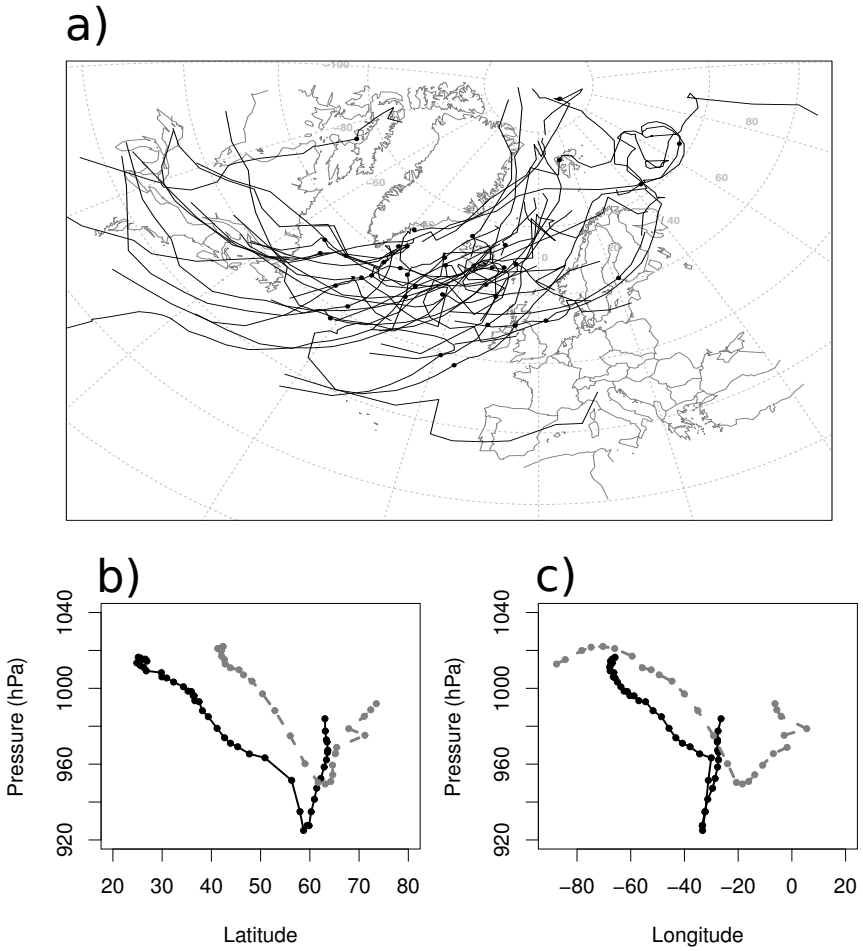

FIG 1. *a) Cyclone tracks for the October 1989 to March 1990 extended winter. Only a subset of tracks is plotted: ones where any 6-hourly MSLP value reached below 960hPa. Nadir positions are denoted with solid circles. b) Sea-level pressure versus latitude and c) latitude for two of the cyclone tracks in a).*

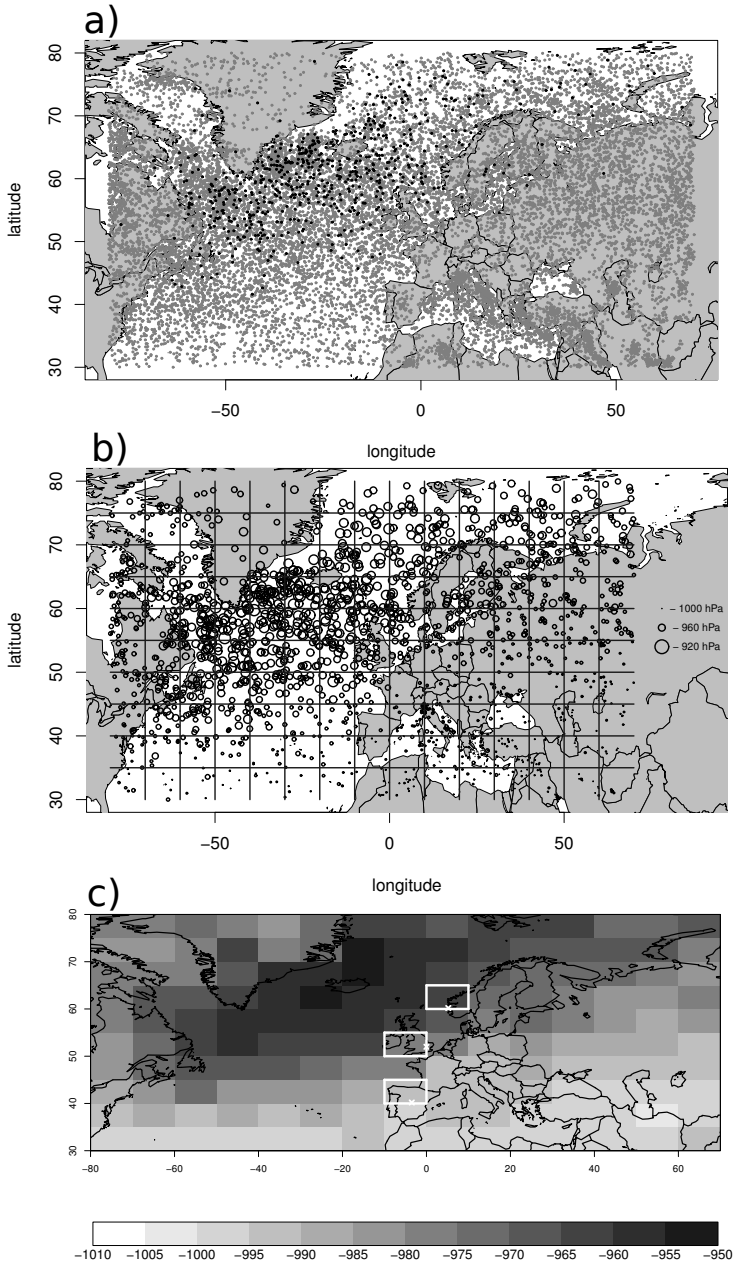

FIG 2. *a) Map of all cyclone nadirs: dots in black represent nadirs deeper than 960hPa, b) map of recorded $X(s,t)$ that are greater than the threshold (90th empirical quantile) in each grid cell and c) map of thresholds in each cell.*

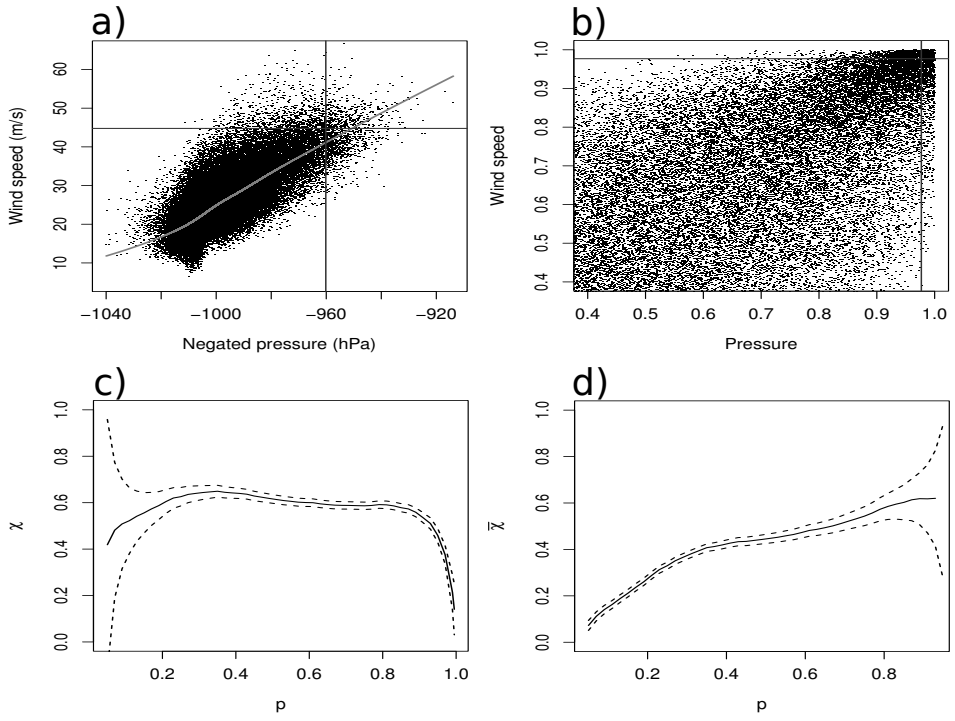

FIG 3. *a) Wind speed against negated sea-level pressure with an associated loess fit (grey line). The intersecting lines are the values -960hPa and 45m/s for pressure and wind speed respectively, representing the same high empirical quantile for each variable. b) empirical copula of wind speed and pressure along with the associated quantile lines from a). c) the associated extremal dependency measure* χ *and d)* $\bar{\chi}$*. The* 95% *confidence intervals in c) and d) are based on the Normal approximation to proportions and are calculated as introduced in Coles et al. (1999).*

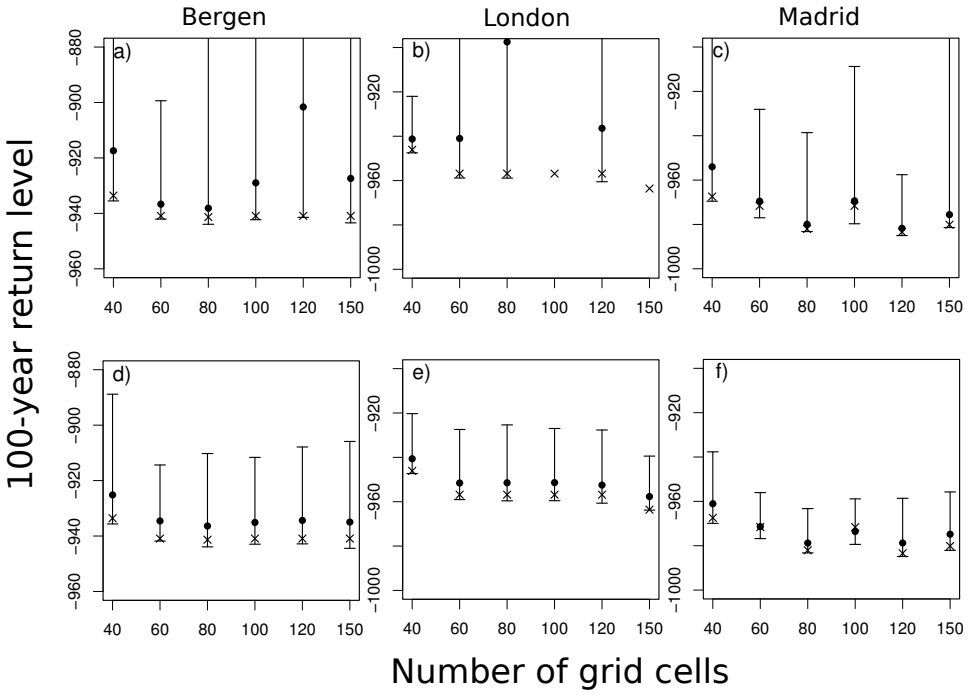

FIG 4. *Dots are posterior means of the 100-year return level of $X(s,t)$ versus number of cells in different grid specifications, along with 95% credible intervals. Left (a and d), middle (b and e) and right (c and f) panels refer to the Bergen, London and Madrid cells respectively. Top (a, b, c) and bottom (d, e, f) panels refer to the stationary and the spatial model respectively. For reference, the deepest recorded value of $X(s,t)$ in each cell is shown with a cross symbol.*

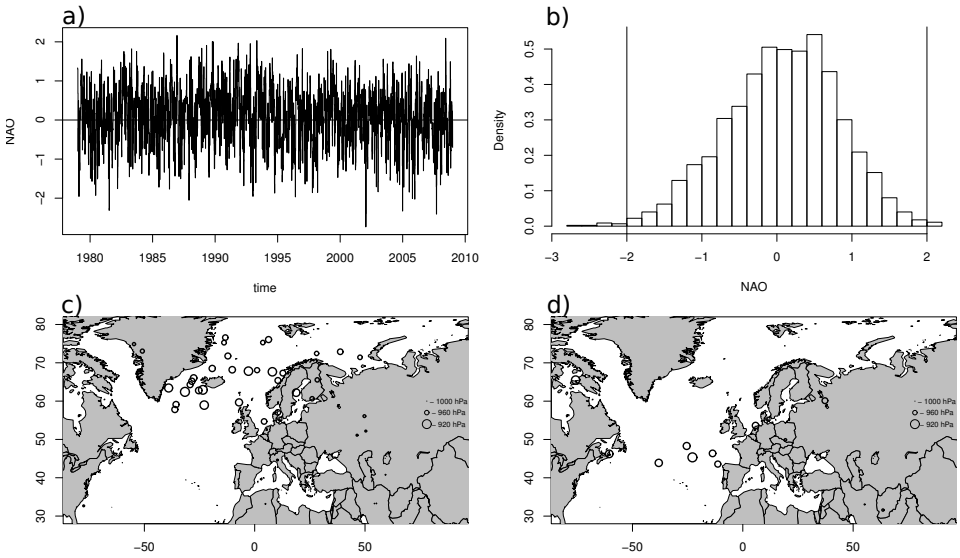

FIG 5. *a) Time series of NAO defined as a 5-day average of daily NAO, b) Histogram of NAO along with vertical lines marking the values -2 and 2, c) occurrences of recorded nadirs where the associated NAO value was greater than 2 and d) less than -2.*

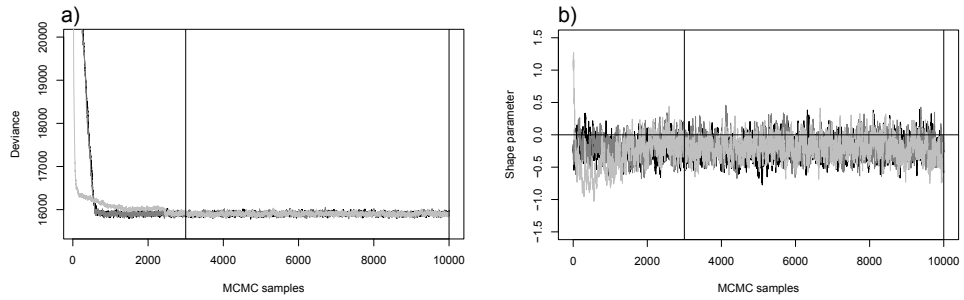

FIG 6. *a) Deviance samples from each of the three MCMC chains. Vertical lines denote the burn-in and the total number of simulations. Samples between the two lines are used for inference. b) Samples of the shape parameter $\xi(s)$ for the grid cell containing London.*

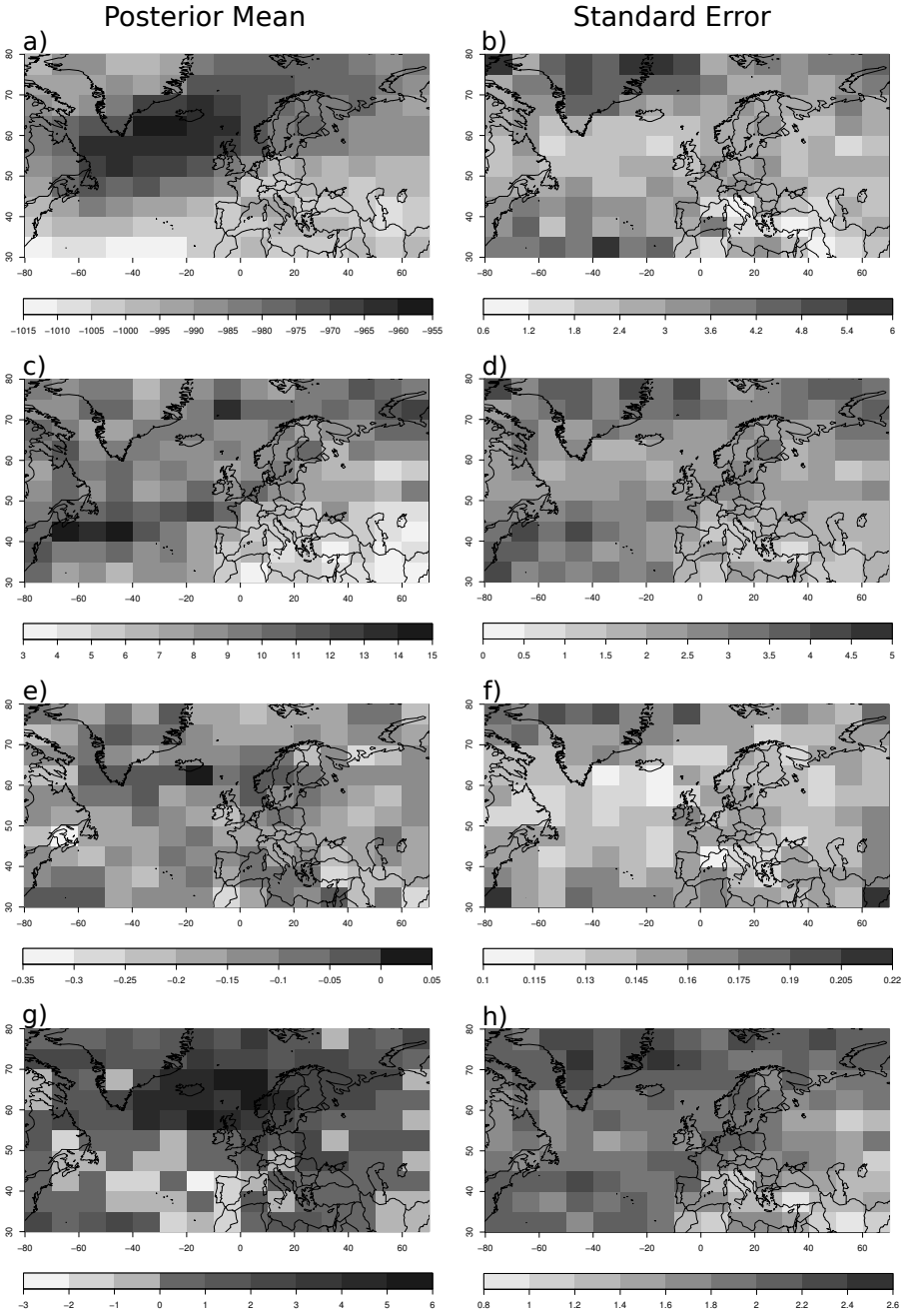

FIG 7. *Posterior means for a) $\mu(s,t)$, c) $\sigma(s,t)$, e) $\xi(s)$ and g) $\beta_2(s)$ and standard errors in b), d), f) and h) respectively, where $z_1(t)$ is latitude at centre of grid cell and $z_2(t) = 0$.*

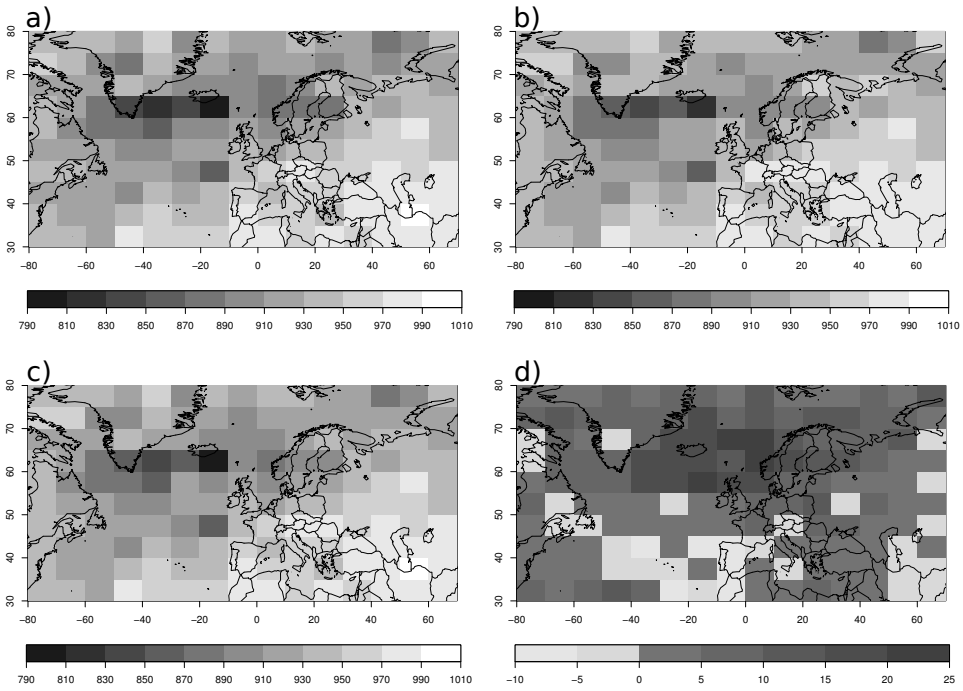

FIG 8. *Estimated lower limits of nadir sea-level pressure for a) $NAO = 2$, b) $NAO = -2$ and c) $NAO = 0$. d) shows the difference between a) and b).*

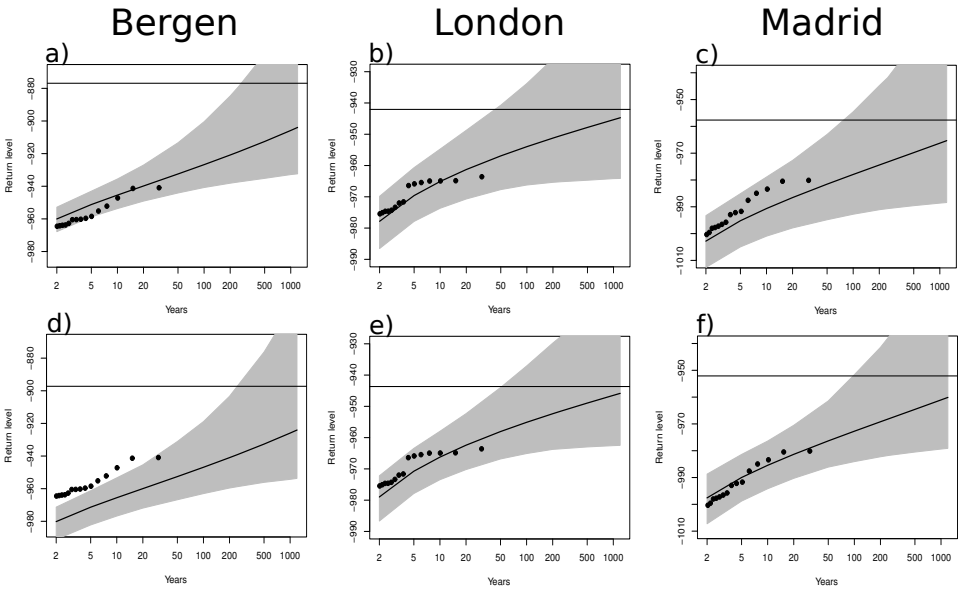

FIG 9. *Individual grid cell return level plots (posterior means) with* 95% *credible intervals. Observed values shown in solid circles. Top panel:* $NAO = 2$*, bottom panel:* $NAO = -2$*. Left panel: Bergen cell, middle panel: London cell, right panel: Madrid cell. Horizontal lines are estimated upper bounds of* $X(s,t)$ *for* $NAO = 2$ *(top) and* $NAO = -2$ *(bottom).*

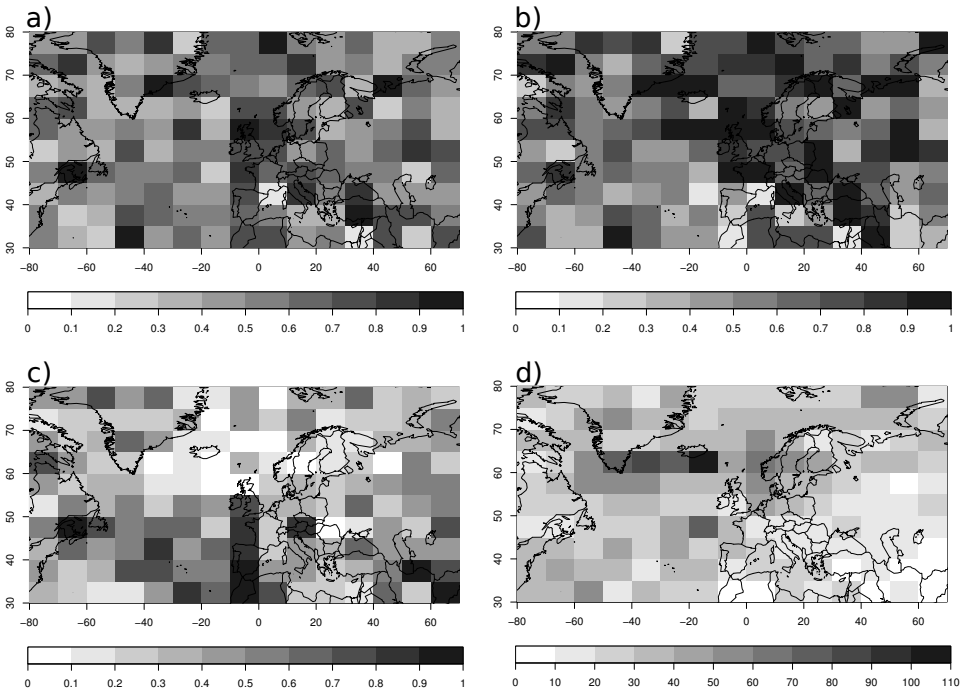

FIG 10. *Probability of observing a deeper nadir than the recorded 30-year deepest nadir in each cell: a) Calculated for NAO values associated with the recorded values nadirs, b) $NAO = 2$ and c) $NAO = -2$. d) the difference in hPa between the estimated depth limit and the deepest recorded 30-year nadirs in each cell.*

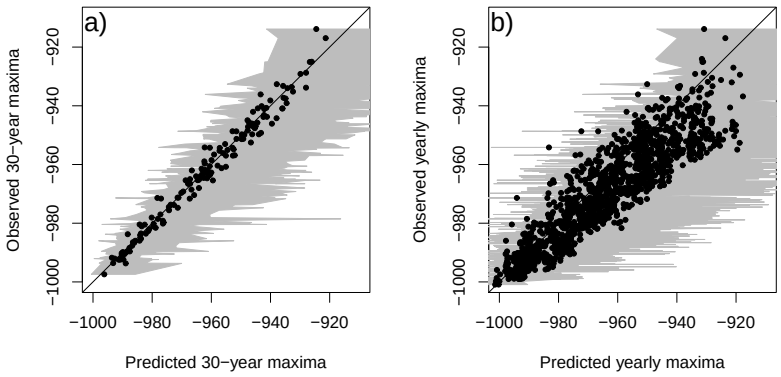

FIG 11. *Recorded versus predicted values of $X(s,t)$: a) 30-year deepest nadirs in each cell and b) yearly deepest nadirs in each cell. Predicted values are the means of the posterior predictive distributions of $X(s,t)$ while the grey shaded area represent the associated* 95% *prediction intervals.*

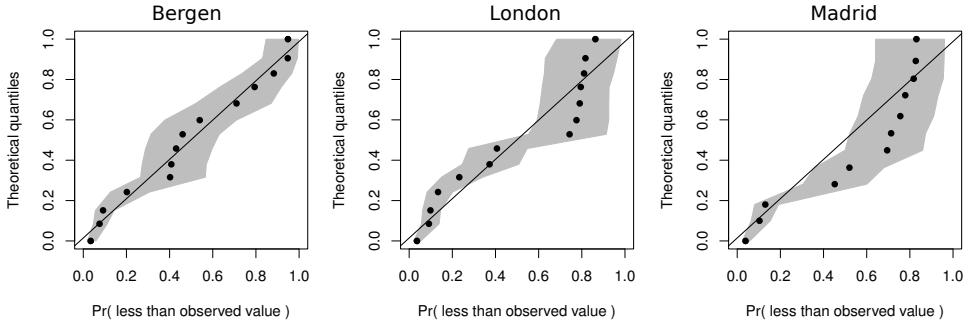

FIG 12. *Quantile-quantile plots of the posterior probability of nadirs being deeper than the observed values in each of: the Bergen cell, the London cell and the Madrid cell. The y-axis are theoretical quantiles from a uniform distribution on (0,1). The* 95% *credible intervals reflect estimation uncertainty.*

REFERENCES

BAILEY, T. C. AND GATRELL, A. C. (1995). *Interactive Spatial Data Analysis*. Longman.

BANERJEE, S., CARLIN, B. P., AND GELFAND, A. E. (2004). *Hierarchical modelling and analysis for spatial data*. Chapman and Hall.

BESAG, J., GREEN, P. J., HIGDON, D., AND MENGERSEN, K. (1995). Bayesian computation and stochastic systems. *Statistical Science*, 10:3–66.

BONAZZI, A., CUSACK, S., MITAS, C., AND JEWSON, S. (2012). The spatial structure of European wind cyclones as characterized by bivariate extreme-value copulas. *Nat. Hazards Earth Syst. Sci*, 12:1769–1782.

CASSON, E. AND COLES, S. (1999). Spatial regression models for extremes. *Extremes*, 1(4):449–468.

Coles, S. (2001). *An introduction to Statistical Modeling of Extreme Values*. Springer-Verlag.

Coles, S., Heffernan, J., and Tawn, J. (1999). Dependence measures for extreme value analyses. *Extremes*, 2(4):339–365.

Cooley, D., Nychka, D., and Naveau, P. (2007). Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistcal Association*, 102(479):449–468.

Cooley, D. and Sain, S. R. (2008). Spatial hierarchical modeling of precipitation extremes from a regional climate model. *Journal of Agricultural, Biological and Environmental Statistics*, 15(3):381–402.

Davison, A. C., Padoan, S. A., and Ribatet, M. (2012). Statistical modelling of spatial extremes. *Statistcal Science*, 27(2):161–186.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. New York:Springer.

Della-Marta, P. M., Mathis, H., Frei, C., Liniger, M. A., Kleinn, J., and Appenzeller, C. (2009). The return period of wind cyclones over Europe. *International Journal of Climatology*, 29(3):437–459.

Della-Marta, P. M. and Pinto, J. G. (2009). Statistical uncertainty of changes in winter cyclones over the North Atlantic and Europe in an ensemble of transient climate simulations. *Geophysical Research Letters*, 36:L14703:1–5.

Einmahl, J. H. J. and Magnus, J. R. (2008). Records in athletics through extreme-value theory. *Journal of the American Statistical Association*, 103(484):1382–1391.

Elsner, J. B., Hodges, R. E., and Jagger, T. H. (2012). Spatial grids for hurricane climate research. *Climate Dynamics*, 39:21–36.

Frigessi, A., Haug, O., and Rue, H. (2003). A dynamic mixture model for unsupervised tail estimation without threshold selection. *Extremes*, 5:219–235.

Gaetan, C. and Grigoletto, M. (2007). A hierarchical model for the analysis of spatial rainfall extremes. *Journal of Agricultural, Biological, and Environmental Statistics*, 12:434–449.

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

Gneiting, T., Balabdaoui, F. and Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B*, 69(2):243–268.

Heaton, M. J., Katzfuss, M., Ramachandar, S., Pedings, K., Gilleland, E., Shamseldin, E. M., and Smith, R. L. (2009). Spatio-temporal models for large scale indicators of extreme weather. *Environmetrics*, 22:294–303.

Hodges, K. (1994). A general method for tracking analysis and its application to meteorological data. *Monthly Weather Review*, 122(11):2573–2585.

Hodges, K. (1995). Feature tracking on the unit sphere. *Monthly Weather Review*, 123(12):3458–3465.

Hodges, K. (1999). Adaptive constraints for feature tracking. *Monthly Weather Review*, 127(6):13621373.

Illian, J., Penttinen, A., Stoyan, H., and Stoyan, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*. Wiley.

Katz, R. W. (2010). Statistics of extremes in climate change. *Climate Change*, 100:71–76.

Katz, R. W., Brush, G. S., and Parlange, M. B. (2005). Statistics of extremes: Modeling ecological disturbances. *Ecology*, 86(5):1124–1134.

Lionello, P., Boldrin, U., and Giorgi, F. (2008). Future changes in cyclone climatology over Europe as inferred from a regional climate simulation. *Climate Dynamics*,

30:657–671.

Mailier, P. J., Stephenson, D. B., Ferro, C. A. T., and Hodges, K. I. (2006). The seriality of extratropical cyclones. *Bulletin of the American Meteorological Society*, 87(10):1317–1318.

Mitchell-Wallace, K., Mitchell, A. (2007). Willis' detailed report for wintercyclone Kyrill. *http://www.willisresearchnetwork.com*.

Nissen, K. M., Leckebusch, G. C., Pinto, J. G., Renggli, D., Ulbrich, S., and Ulbrich, U. (2010). Cyclones causing wind cyclones in the Mediterranean: characteristics, trends and links to large-scale patterns. *Natural Hazards and Earth System Science*, 10(7):1379–1391.

Odell, L., Knippertz, P., Pickering, S., Parkes, B., and Roberts, A. (2013). The Braer cyclone revisited. *Weather*, 68(4):105–111.

Pinto, J. G., Zacharias, S., Fink, A. H., Leckenbusch, G. C., and Ulbrich, U. (2009). Factors contributing to the development of extreme North Atlantic cyclones and their relationship with the NAO. *Climate Dynamics*, 32:711–737.

R Core Team (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Reich, B. J. and Shaby, B. A. (2012). A hierarchical max-stable spatial model for extreme precipitation. *Annals of Applied Statistics*, 6(4):1430–1451.

Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields: Theory and Applications*. Chapman and Hall.

Saha, S. and Co-authors (2010). The NCEP climate forecast system reanalyses. *Bulletin of the American Meteorological Society*, 91:1015–1057.

Sang, H. and Gelfand, A. E. (2009). Hierarchical modeling for extreme values observed over space and time. *Environ. Ecol. Stat.*, 16:407–426.

Sang, H. and Gelfand, A. E. (2010). Continuous spatial process models for spatial extreme values. *Journal of Agricultural, Biological, and Environmental Statistics*, 15(1):49–65.

Seierstad, I. A., Stephenson, D. B., and Kvamsto, N. G. (2006). How useful are teleconnection patterns for explaining variability in extratropical cycloneiness. *Tellus A*, 59(2).

Sienz, F., Schneidereit, A., Blender, R., Fraedrich, K., and Lunkeit, F. (2010). Extreme value statistics for North Atlantic cyclones. *Tellus A*, 62(4):347–360.

Smith, R. L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone (with discussion). *Statistical Science*, 4:367–393.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639.

Stephenson, D. B., Casati, B., Ferro, C. A. T., and Wilson, C. A. (2008). The extreme dependency score: a non-vanishing measure for forecasts of rare events. *Meteorological Applications*, 15:41–50.

Turkman, K. F., Turkman, M. A. A., and Pereira, J. M. (2010). Asymptotic models and inference for extremes of spatio-temporal data. *Extremes*, 13:375–397.

Ulbrich, U., Fink, A. H., Klawa, M., and Pinto, J. G. (2001). Three extreme cyclones over Europe in December 1999. *Weather*, 56(3):70–80.

Vitolo, R., Stephenson, D. B., Cook, I. M., and Mitchell-Wallace, K. (2009). Serial clustering of intense European storms. *Meteorologische Zeitschrift*, 18(4):411–424.

T. Economou, D. B. Stephenson and C. A. T Ferro
College of Engineering, Mathematics and Physical Sciences
Exeter Climate Systems
North Park Road
EX4 4QF, Exeter, UK
E-mail: t.economou@ex.ac.uk
d.b.s.stephenson@ex.ac.uk
c.a.t.ferro@ex.ac.uk