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 by application to 17,230 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 is found to have a significant effect on extremal storm behaviour especially over N. Europe and the Iberian peninsula.

The model fits suggest that there is a high probability of having minimum pressures lower than the observed minima over 1979-2009, especially when the NAO index is positive. Estimates of lower bounds on minimum sea-level pressure are typically 10-50hPa below the minimum values found for cyclones over the period 1979-2009. For example, lower bounds for cyclone nadirs near London were found to be (943.2, 943.0, 941.8) when the NAO index is -2,0,2 standard deviations above normal, compared to the empirical minimum value of 975.0hPa (occurred during normal NAO) from 1979-2009.

Keywords and phrases: Bayesian Hierarchical Model, spatial random effects, natural hazards, extra-tropical cyclones, extreme value distribution, European windstorms
1. Introduction. Extreme weather events have significant impact on society, e.g. storms, heat waves, droughts. The general problem with analysing weather phenomena is their complex nature: both occurrence (frequency) and intensity (a measure of strength) are irregular, spatially heterogeneous and temporally modulated by climate variability (e.g. through variations in large-scale background atmospheric flow). Irrespective of how intensity is defined, these generic issues exist for all weather systems including extreme ones. Furthermore, extreme phenomena are rare by definition, so data availability is generally poor. Extreme value theory provides an elegant tool for mitigating data limitations, and in this article, we propose a spatio-temporal extreme value framework as a tool 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 type of weather extreme.

We demonstrate our framework by applying it to the societally-relevant problem of N. Atlantic and European extra-tropical cyclones. These natural hazards cause much damage and insurance loss in Europe due to extreme wind speeds but also due to extreme flooding. Recent examples of damaging extra-tropical cyclones 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, the Netherlands, Belgium, Austria, the Czech Republic and Poland. In this study, we use cyclone tracks objectively identified in 6 hourly re-analysis data (see Section 3.2 for details). Figure 1a shows some of these tracks for the period Oct 1989 – Mar 1990. As each cyclone moves eastward across the Atlantic, its central pressure decreases (deepens) until it reaches a
minimum (the nadir) and then it increases again (fills in). We shall use the maximum depth below 960hPa (i.e. 960hPa minus the sea-level pressure at the nadir) as a simple measure of the maximum intensity of each cyclone (see Section 3.2 for details). The nadirs in Figure 1a are denoted with solid circles, and in Figure 2a we plot all cyclone nadirs in the period 1979-2009. The plot illustrates the spatial heterogeneity in the occurrence (i.e. clustering/repulsion). In Figure 2b, we plot only the extreme nadirs where the definition of “extremeness” varies across each grid cell. This illustrates both the decrease in data when analysing extremes but also the spatial heterogeneity in nadir intensity (larger circles indicate lower pressure). In section 3, we address questions such as “How low can sea-level pressure become in extra-tropical cyclones?” (i.e. question 1 above), using the proposed methodology to allow for the fact that the answer to this question will depend both on time and spatial location. For instance, we would expect extreme low pressures associated with cyclones over Scandinavia to be lower than ones in the Mediterranean.

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 dependence, 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, on 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, which is preferable to the GEV and the GPD 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 (both for extremes and non-extremes) 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 fixed location in space. Weather extremes however, do not tend to occur at fixed locations in space. They 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 $\Delta$ 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 $\Delta$ is a crucial part of the methodology (see section 3.3).

For spatial marked point processes, estimation is only possible after making assumptions about spatial ‘smoothness’. Our spatial discretisation ap-
proach assumes that within each grid cell, events are spatially independent, so that spatial dependence only exists between grid cells rather than events. The framework we propose here is one where conditional on the spatial or spatio-temporal dependence structure between cells, independent extreme value models may be assumed for the events in each cell. Such models are often termed hierarchical since, conditional on the spatio-temporal dependence of the cells, events are modelled using an appropriate extreme value model at the top of the hierarchy (data level) whereas at the second level (spatio-temporal process level) the spatial dependence model is defined.

Spatial independence within cells implies two things: first, the occurrence of events within cells is spatially random, and second, the distribution of event intensity within cells is independent of space. This is a reasonable assumption given that the size of the cells is adequately small. Nevertheless, in the subsequent section we allow for covariates that vary spatially within cells to relax this assumption.

2.3. Model formulation. Data level:
For events within each cell, we assume the point process model for extremes (Coles, 2001). 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 \( \mu, \sigma \) and \( \xi \) (see Appendix A for more details). For clarity, we use the notation \( X|X > u \sim PP(\mu, \sigma, \xi, u) \) to mean that the magnitude \( X \) of an event can be described probabilistically by the point process (PP) model for extremes.

Introducing spatial and temporal variation, let the variable \( X(s, t) \) characterise intensity of some weather event occurring in grid cell \( s \in S \) at time \( t \in T \) where \( S \) and \( T \) are the space and time domains, each a fixed subset of 2-dimensional and 1-dimensional Euclidean space respectively. Suppose there are \( N \) grid cells and define extreme events in each cell as values over \( u(s) \), which is an appropriately chosen threshold, e.g. a high empirical quantile. Note that the threshold may vary between cells but is assumed constant in time. Extending
Cooley and Sain (2008) we model the extremes in the following general way:

\begin{equation}
X(s,t) \mid X(s,t) > u(s) \sim \text{PP}(\mu(s,t), \sigma(s,t), \xi(s), u(s))
\end{equation}

\begin{equation}
\mu(s,t) = \beta^\mu_0 + \sum_{k=1}^{K} \beta^\mu_k(s)z_k(s,t) + \theta^\mu(s) + \gamma^\mu(s,t)
\end{equation}

\begin{equation}
\log(\sigma(s,t)) = \beta^\sigma_0 + \sum_{k=1}^{K} \beta^\sigma_k(s)z_k(s,t) + \theta^\sigma(s) + \gamma^\sigma(s,t)
\end{equation}

\begin{equation}
\xi(s,t) = \beta^\xi_0 + \sum_{k=1}^{K} \beta^\xi_k(s)z_k(s,t) + \theta^\xi(s) + \gamma^\xi(s,t)
\end{equation}

where \(z_1(s,t), \ldots, z_K(s,t)\) 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 \(\beta^\mu_k(s), \beta^\sigma_k(s), \beta^\xi_k(s)\) may be spatially varying or constant, i.e. \(\beta^\mu_k(s) = \beta^\mu_k\), 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 \(\mu, \sigma\) and \(\xi\) across the cells. In addition, \(\gamma^\mu(s,t), \gamma^\sigma(s,t)\) 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 spatio-temporal random effects may capture unobserved effects from missing covariates or as we discuss further on, allow the introduction of particular spatio-temporal behaviour in the model.

The model can be used to calculate the \(r\)-year return level, defined as the \((1 - 1/r)^{th}\) conditional quantile of \(X(s,t)\) in cell \(s\) given a specific value of \(t\):

\begin{equation}
X_{1-1/r}(s,t) = \mu(s,t) + \frac{\sigma(s,t)}{\xi(s,t)}((-\log(1 - 1/r))^{-\xi(s,t)} - 1).
\end{equation}

**Spatio-temporal process level:**

Following Cooley and Sain (2008), we define the purely spatial effects in the model as follows:

\begin{equation}
(\theta^\mu(s), \theta^\sigma(s), \theta^\xi(s))' \sim \text{N}((U^\mu(s), U^\sigma(s), U^\xi(s))', \text{diag}(\tau)^{-1})
\end{equation}

\begin{equation}
(U^\mu, U^\sigma, U^\xi)' \sim \text{N}(0, \Omega^{-1})
\end{equation}

\begin{equation}
\beta^\psi_k(s) \sim \text{N}(\nu^\psi_k, \phi^\psi_k), \quad k = 1, \ldots, K; \quad \psi = \mu, \sigma, \xi
\end{equation}

where \(U^\psi = (U^\psi(1), \ldots, U^\psi(N))'\) for all \(\psi = \mu, \sigma, \xi\). The vectors \(U^\mu, U^\sigma, U^\xi\) each follow an Intrinsic AutoRegressive (IAR) spatial model (Besag et al., 1995) which is effectively the multivariate Gaussian in (7), whose covariance


matrix $\Omega^{-1}$ is based on an $N \times N$ spatial proximity matrix $W$ between cells. In the application of this model to cyclone extremes in section 3, spatial proximity is based on east-west and north-south neighbouring cells so that off-diagonal elements of $W$ 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 structures). As a model, the IAR uses a single unknown parameter and the proximity matrix to control the spatial autocorrelation or ‘smoothness’ between cells.

When the IAR model is used as a prior, like here, then it is improper: the distribution does not integrate to 1. So, to identify the intercept terms $\beta^\psi_0$, $W$ is constructed as above implying that row sums are zero. This in turn imposes the restriction that $\sum_s U^\psi(s) = 0$ for all $\psi = \mu, \sigma, \xi$, making the intercept terms identifiable. The parameters $\beta^\psi_k(s)$, are chosen to be spatially variable but in an unstructured way: as random slopes with common global mean $\nu^\psi_k$ and variance $\phi^\psi_k$. This ensures that $\beta^\psi_k(s)$ share information to aid estimation in cells with few events but less so compared to using a structured (IAR) spatial prior. Parameter $\nu^\psi_k$ reflects the overall effect of the associated covariate on the extreme events. There is no reason however, why parameters $\beta^\psi_k(s)$ cannot be spatially varying according to an IAR model. In that case, $\beta^\psi_k(s) \sim N(\nu^\psi_k + U^\psi_k(s), \phi^\psi_k)$ where $U^\psi_k = (U^\psi_k(1), \ldots, U^\psi_k(N))'$ is modelled as an IAR process and can be incorporated as an element to the left hand side of (7).

As in Cooley and Sain (2008), vectors $U^\psi$ are modelled jointly using a separable formulation (Banerjee et al., 2004) so that the precision matrix $\Omega = T \otimes W$ where $T$ is a $3 \times 3$ positive definite symmetric matrix (that needs to be estimated), meaning that $\Omega$ is $3N \times 3N$. This formulation explicitly allows for dependence between $\mu, \sigma$ and $\xi$ using just 3 extra parameters (the off-diagonals of $T$). Allowing for this dependence can also aid the MCMC estimation discussed in section 2.4, in terms of convergence to the posterior but also mixing of the MCMC samples.

The diagonal elements of $T$ relate to the variance parameter in each IAR controlling both the strength of spatial dependence and the variability in each $U^\psi$. The off-diagonal elements relate to the dependence between each $U^\psi$. Parameter $\tau$ in (6), is a vector of precisions reflecting the variability of each $\theta^\psi$, where $\theta^\psi = (\theta^\psi(1), \ldots, \theta^\psi(N))'$, other than that induced from the IAR model in (7) (sometimes referred to as the ‘nugget’ effect). In this model

\[ \sum_{s} U^\psi(s) = 0 \text{ for all } \psi = \mu, \sigma, \xi, \text{ making the intercept terms identifiable.} \]

\[ \beta^\psi_k(s) \sim N(\nu^\psi_k + U^\psi_k(s), \phi^\psi_k) \text{ where } U^\psi_k = (U^\psi_k(1), \ldots, U^\psi_k(N))' \text{ is modelled as an IAR process and can be incorporated as an element to the left hand side of (7).} \]

\[ U^\psi_k = (U^\psi_k(1), \ldots, U^\psi_k(N))' \text{ is modelled as an IAR process and can be incorporated as an element to the left hand side of (7).} \]

\[ (U^\psi_k(1), \ldots, U^\psi_k(N))' \text{ is modelled as an IAR process and can be incorporated as an element to the left hand side of (7).} \]
formulation, the spatial random effects \( (\theta^\mu(s), \theta^\sigma(s), \theta^\xi(s))^T \) have a Gaussian distribution whose mean is governed by \((U^\mu, U^\sigma, U^\xi)^T\). Allowing \((\theta^\mu(s), \theta^\sigma(s), \theta^\xi(s))^T\) to inherit their spatial dependence through \((U^\mu, U^\sigma, U^\xi)^T\) 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 the IAR model in (7). The values for \(\tau\) are conventionally fixed beforehand to avoid non-identifiability between \(\tau\) and the diagonal of \(T\) (Banerjee et al., 2004).

The spatio-temporal effects are modelled as follows: for all \(s \in S\) and \(t \in T\),

\[
\begin{align*}
\gamma^\psi(s, t) & \sim N(\alpha^\psi(s)\gamma^\psi(s, t-1), \omega) \\
\gamma^\psi(s, 0) & \sim N(0, \omega) \\
\alpha^\psi(s) & \sim N(\alpha^\psi_0 + U^\alpha^\psi(s), \chi^\psi) \\
(U^\alpha^\mu, U^\alpha^\sigma, U^\alpha^\xi)^T & \sim N(0, \Phi^{-1}).
\end{align*}
\]

This is an autoregressive model in time where the autocorrelation parameter \(\alpha^\psi(s)\) is spatially varying. It is potentially a useful formulation as it can capture spatio-temporal behaviour in the weather extremes that are 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 \(\Phi = P \otimes W\), where \(P\) is another \(3 \times 3\) 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 \(\gamma^\psi(s, t) \sim N(0, \omega)\) where both temporal and spatial effects are unstructured but still constrained by the single variance parameter \(\omega\). Note that for the same argument as with \(\tau\) in (6), the variance parameters \(\chi^\psi\) 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 \(\beta^\psi_0, \theta^\psi(s)\) and \(\gamma^\psi(s, t)\) for all \(\psi = \mu, \sigma, \xi\) as well as \(\beta^\psi_k(s)\) for \(k = 1, \ldots, K\). The prior distributions for \(\theta(s)^\psi, \beta^\psi_k(s)\) and \(\gamma^\psi(s, t)\) are given in (6), (8) and (9) respectively. For spatially constant parameters \(\beta^\psi_k\) we can assume flat Gaussian prior distributions with zero mean and large variance. Random effects \(\theta^\psi(s), \beta^\psi_k(s)\) and \(\gamma^\psi(s, t)\) are sampled by Metropolis-Hastings, specifically using a random walk sampler. The intercepts \(\beta^\psi_0\) are given Gaussian prior distributions.
with large variance, and means \((\bar{\mu}, \bar{\sigma}, \bar{\xi})\), 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 \(\theta^\psi(s)\).

At the spatio-temporal process level, the prior distributions for \((U^\mu, U^\sigma, U^\xi)^t\) and \((U^{\alpha^\mu}, U^{\alpha^\sigma}, U^{\alpha^\xi})^t\) are given in (7) and (12) respectively. For \(T\) and \(P\) we use a Wishart prior with 3 degrees of freedom (uninformative) and a mean that relates to the variability of \(\mu\), \(\sigma\) and \(\xi\) across cells (see section 3 for particular details). Using these priors one may draw samples of \((U^\mu, U^\sigma, U^\xi)^t\), \((U^{\alpha^\mu}, U^{\alpha^\sigma}, U^{\alpha^\xi})^t\), \(T\) and \(P\) from their full conditionals using a Gibbs sampler, utilising the specific techniques in Cooley and Sain (2008). Parameter \(\nu_k^\psi\) is given a flat Gaussian prior with zero mean and large variance and the prior distribution \(\pi()\) for \(\phi_k^\psi\) 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 \(\nu_k^\psi\) and scaled inverse-\(\chi^2\) for \(\phi_k^\psi\) (Gelman et al., 2004).

3. Application to extreme extra-tropical cyclones.

3.1. Background. 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 i.e. pressure drops of more than 24hPa in 24 hours. Explosive cyclogenesis depends on many factors, for example, the deepest recorded recent low of 913hPa (the Braer cyclone of January 1993) deepened 78hPa in 24 hours due to a combination of several factors such as available moisture and stratospheric conditions (Odell et al., 2013). Assuming the unlikely possibility of a maximum deepening rate of 5hPa/hour sustained for 2 days would give a minimum bound of SLP of around 750hPa for the typical background state of 990hPa. 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 general 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 do-
main without explicit characterisation of spatial or temporal heterogeneity. Della-Marta and Pinto (2009) used the GPD model to examine future changes in European and North Atlantic wind-cyclone intensity and a Poisson model for changes in cyclone frequency, using data from a coupled general circulation model (GCM) of climate. Although their study considered three different non-overlapping areas of large size, there was no formal consideration of spatial (or temporal) dependence. Furthermore, Della-Marta et al. (2009) used the GPD and a Poisson model for the exceedances, to investigate extreme wind-gusts associated with North Atlantic cyclones in the re-analysis data set ERA40, however spatial heterogeneity was not explicitly modelled. 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.

3.2. Data. 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 1.5 km 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 October 1989 to March 1990 (a notable period with many damaging cyclones). For clarity, only a subset of tracks is plotted: ones where any 6-hourly MSLP value reached below 960hPa, where 960hPa is considered an extreme (low) value for extra-tropical cyclone sea-level pressure. 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 cyclone (the 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 geostrophic wind balance. Outside equatorial regions, the Coriolis force almost completely balances the pressure gradient force (geostrophic balance), and so wind velocities in extra-tropical cyclones are linearly determined by pressure gradients. 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 these simple dynamical meteorology arguments, minimum pressure and maximum wind speeds are expected to be extremally dependent and so convey similar information. Figure 3a plots maximum (6-hourly) wind speed against negated MSLP (obtained by multiplying MSLP by -1) showing a strong 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 making a scatter plot of the empirical probabilities $q_t = (\text{rank}(z_t) - 1)/(n - 1)$ (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 (Stephenson et al., 2008) of the extremal dependence measure $\chi$ (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 extreme levels of negated MSLP and maximum wind speed, so either variable could be used to effectively investigate extremes (see Appendix A.2 for a formal definition of $\chi$ and $\bar{\chi}$).

During their life cycle, extra-tropical cyclones start (cyclogenesis) by rapidly deepening to the nadir which is when the cyclone starts to weaken and deteriorate until the end of the life cycle (cyclolysis). 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 Atlantic extratropical 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 the deepest point (which we assume approximates the value of the nadir) before it starts increasing again until cycolysis. 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 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.

By only analysing nadirs, 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 strong cyclone in the Mediterranean can cause damage even if it is classified as weak in relation to cyclones over Scandinavia.

A potential limitation in using cyclone tracks derived from a re-analysis dataset, is that these tracks are not the same as using observations (generally cyclone track observations for the extratropics are not readily available). Re-analysis data is output from climate models with assimilated historical observational data (i.e. constrained by observations). There is much smoothing/interpolation of the observational data when creating a re-analysis dataset so the interpretation of any results obtained from implementing the statistical model proposed here, is conditional on the effects of such smoothing.

3.3. 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 sea-level pressure.** We divide the domain in Figure 2a into $N = 150$ grid cells of $5^\circ$ in latitude and $10^\circ$ in longitude. In each grid cell, the threshold $u(s)$ is defined as the empirical 90th quantile of $X(s,t)$ which was chosen for two main reasons. Firstly, the number of nadirs for some cells was small so the 90th quantile allows more data for estimation compared to higher quantiles and second, we performed exploratory threshold analysis in selected cells, i.e. using mean residual life plots (Coles, 2001), ensuring that the 90th empirical quantile is an appropriate threshold choice. Figure 2b shows the map of the extremes (1,736 nadirs) where the radius of each circle relates to the intensity of $X(s,t)$ and Figure 2c shows the map of thresholds. Note that in Figure 2c, three cells are highlighted in white, specifically from top to bottom: cells containing coordinates $(5.2^\circ E, 60.2^\circ N)$, $(0^\circ E, 5^\circ N)$ and $(3.5^\circ W, 40.2^\circ N)$ 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 (stationary and spatial), the 100-year return level of $X(s,t)$ was calculated for each of the three coordinates marked in Figure 2c. For the stationary model, Figure 4a, 4b and 4c plot the posterior mean of the 100-year return level against the number of cells in each grid configuration, along with 95% credible intervals. Figure 4d, 4e and 4f show the equivalent plots for the spatial model. The left, middle and right panels of Figure 4 reflect the Bergen, London and Madrid coordinates respectively. 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 several nadirs in them (ranging from 14 to 376 nadirs).
3.4. Model specification and selection of covariates. The particular model formulation we propose is the following:

\begin{align}
X(s,t)|X(s,t) > u(s) & \sim \text{PP}(\mu(s,t), \sigma(s,t), \xi(s), u(s)) \\
\mu(s,t) & \sim \beta_0^\mu + \beta_1^\mu z_1(t) + \beta_2(s)z_2(t) + \theta^\mu(s) \\
\log(\sigma(s,t)) & \sim \beta_0^\sigma + \beta_1^\sigma z_1(t) + \theta^\sigma(s) \\
\xi(s) & \sim \beta_0^\xi + \theta^\xi(s)
\end{align}

where at the spatial process level we have:

\begin{align}
(\theta^\mu(s), \theta^\sigma(s), \theta^\xi(s))' & \sim \text{N}((U^\mu(s), U^\sigma(s), U^\xi(s))', \text{diag}(\tau)^{-1}) \\
(U^\mu, U^\sigma, U^\xi)' & \sim \text{N}(0, \Omega^{-1}) \\
\beta_2(s) & \sim \text{N}(\nu, \phi^2)
\end{align}

For this application 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, as we discuss further on. 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 \(\log(\sigma(s,t))\). Hence, \(z_1(t)\) in (14) and (15) is the latitude of each negated nadir \(X(s,t)\) within the grid cell. 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 is given a cell-varying parameter \( \beta_2(s) \sim N(\nu, \phi^2) \) so that \( \nu \) 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 found not to be significant both by using the DIC and by checking that the posteriors of all \( \beta_3(s) \) were centred at zero. Lastly, both the NAO and EAP were included in 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 \( \geq 2 \) and NAO \( \leq -2 \) respectively. There is a clear North-South pattern in the Central Atlantic implying NAO has a notable effect on extreme cyclones.

3.5. Model implementation. As mentioned in section 2.3, the values for \( \tau \) are fixed and here we set \( \tau = (0.1, 10, 100)'^t \). 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 \( U^\mu, U^\sigma, U^\xi \) and not \( \tau \). If values in \( \tau \) 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 \( U^\psi \). Sensitivity analysis was performed to ensure these values have little effect on inference (not shown for conciseness).

For the \( 3 \times 3 \) precision matrix \( T \), the Wishart prior is given the following mean: \( \text{diag}(0.02, 4, 40)' \). As with \( \tau \) these values were calibrated by perform-
ing exploratory analysis using 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) with $z_1(t)$ the latitude at the centre of each cell and $z_2(t)$ the NAO was implemented in R (R Core Team, 2012) using three parallel MCMC chains. These were run on a workstation with an Intel i7 core processor 3.07GHz 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 (burn-in) based on traces 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.

3.6. Results. Summaries of posterior distributions for global parameters are given in Table 1. Latitude has a positive linear effect on both the location and scale parameters of extreme cyclone depth $X(s,t)$, through $\beta_1^\mu$ and $\beta_1^\sigma$. The overall NAO effect $\nu$ is positive, which is in line with much of the physical understanding about cyclones (Pinto et al., 2009). To formally assess MCMC convergence, the Gelman and Rubin $\hat{R}$ diagnostic was used, which takes into account that more than one chain was ran (Gelman et al., 2004). The diagnostic calculates $\hat{R}$ for each parameter and this must be close to 1 if convergence has been reached. In Table 1, $\hat{R}$ values for each parameter are shown and all are sufficiently close to 1 implying convergence. None of the $\hat{R}$ values was greater than 1.20 for any model parameter (random effects etc.) indicating that MCMC samples have converged to the posteriors.

Posterior means and standard deviations of $\mu(s,t)$, $\sigma(s,t)$ and $\xi(s)$ are shown in Figure 7 for $z_2(t) = 0$. Much of the spatial structure in the extreme nadirs comes from the location and scale parameters and it is important to note that the posterior means for the shape parameter $\xi(s,t)$ are all negative but one cell over Iceland. Exploring this further, the two deepest nadirs in the re-analysis occurred in this cell, and these two are considerably lower than the rest of the nadirs in that cell but also the neighbouring cells. 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 grid cell \( s \), has an upper bound defined as \( \frac{\sigma(s,t)}{\xi(s)} - \mu(s,t) \). In this application, 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 all 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 there is considerable literature focusing on the problem of estimating upper/lower bounds of distributions. See (de Haan and Ferreira, 2006, chapter 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, that is 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 = ±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 $\Pr(X(s, t) > x_m(s))$ 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: $	ilde{\sigma} = \sigma \delta \xi$ and $	ilde{\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 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, predictions (calculated as means from the posterior predictive distributions) of the 30-year deepest nadirs in each cell, are plotted against the corresponding recorded values, along with 95% prediction intervals (2.5% and 97.5% quantiles of the posterior predictive distribution). 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 \( \text{Pr}(X(s,t) < x_{\text{obs}}(s,t)) \) where \( x_{\text{obs}}(s,t) \) 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.

3.7. Summary. In this paper we presented a spatio-temporal framework for modelling extreme weather events that occur at irregular spatial locations, modulated by time-varying background conditions. The framework introduces a spatial regularisation and covariates to extend the Bayesian hierarchical modelling framework of Cooley and Sain (2008).

The framework was illustrated by application to 17,230 extratropical 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, thus addressing question 2 in section 1. 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.

To answer question 3, the North Atlantic Oscillation was used as a covariate in the model and was found to have a significant
effect on extremal storm behaviour especially over N. Europe and the Iberian peninsula. In addressing question 1, 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.

Specifically, 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.

Furthermore, the model formulation presented here for extreme nadir sea-level pressure, is applicable to other cyclone intensity measures that can be tracked from re-analysis datasets (or datasets from climate model output), such as wind-speed and vorticity, or even other cyclone-related variables such as precipitation.

4. 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 investigates both the spatial and temporal structure of extreme extra-tropical cyclones. The methodology can be potentially used for weather events at different temporal and spatial scales (e.g. micro-, meso-, synoptic meteorological scale), since it was developed on a generic spatial and temporal domain.

Although this is a first step towards studying the spatio-temporal behaviour of extreme cyclones, the analysis relies on a few 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. More generally, 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. Such a model would be a spatial marked point process for extremes which is a class of models not given much attention in the literature.

With regards to the specific application, taking the nadir of each cyclone eliminates much of the data. This was efficient as data reduction, however
one might be ignoring potentially useful information. The remedy to this
would be to include all 6-hourly values and then model the dependence of
data within each track. This is a possible future extension to the method
presented here.

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 $X > 0$ 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 infer-
ence 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

\begin{align}
(20) \quad L(\mu, \sigma, \xi; x, t) &= \exp \left\{ -ny \int_0^1 \int_u^\infty \lambda(x,t) dx dt \right\} \prod_i \lambda(x_i, t_i) \\
(21) &= \exp \left\{ -ny \left[ 1 + \xi \left( \frac{u - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \prod_i \lambda(x_i, t_i)
\end{align}

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

$$
PP(\mu, \sigma, \xi, u) = \exp \left\{ -ny [t_i - t_{i-1}] \left[ 1 + \xi \left( \frac{u - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \lambda(x_i, t_i),
$$
for \( i = 0, \ldots, n \) where \( n \) is the number of events. Note that \( t_0 = 0 \) and that for the time interval between the last event occurrence and \( t = 1 \), the likelihood contribution 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 \( Y \sim \text{PP}(\mu, \sigma, \xi, u) \) 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 by summation: 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 is thus approximated by

\[
\int_0^1 \left[ 1 + \xi(s) \left( \frac{u(s) - \mu(s, t)}{\sigma(s, t)} \right) \right] \frac{1}{\xi(s)} \, dt = \frac{1}{J} \sum_{i=1}^J \left[ 1 + \xi(s) \left( \frac{u(s) - \mu(s, k_i)}{\sigma(s, k_i)} \right) \right] \frac{1}{\xi(s)},
\]

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

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

Roughly, \( \bar{\chi} \) measures the ‘speed’ at which \( \chi \) approaches zero and if \( \chi = 0 \) and \( \bar{\chi} \neq 0 \), the variables are asymptotically independent and \( \bar{\chi} \) is the appropriate measure of extremal dependence (see Coles et al. (1999) for details). If \( \bar{\chi} = 0 \), the variables are independent.
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.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Prior</th>
<th>Posterior Mean (s.e.)</th>
<th>95% Cr.I.</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1^\mu$: Latitude</td>
<td>N(0, 100)</td>
<td>4.71 (0.62)</td>
<td>[3.61, 5.94]</td>
<td>1.13</td>
</tr>
<tr>
<td>$\beta_1^\sigma$: Latitude</td>
<td>N(0, 100)</td>
<td>0.12 (0.07)</td>
<td>[0.00, 0.25]</td>
<td>1.03</td>
</tr>
<tr>
<td>overall NAO effect $\nu$</td>
<td>N(0, 100)</td>
<td>1.21 (0.24)</td>
<td>[0.77, 1.66]</td>
<td>1.01</td>
</tr>
<tr>
<td>variance NAO effect $\phi^2$</td>
<td>$\propto 1/\phi^2$</td>
<td>5.6 (1.85)</td>
<td>[3.09, 10.25]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\beta_0^\mu$</td>
<td>N(−944.1, 100)</td>
<td>−987.4 (0.51)</td>
<td>[−988.5, −986.5]</td>
<td>1.04</td>
</tr>
<tr>
<td>$\beta_0^\sigma$</td>
<td>N(5.7, 100)</td>
<td>2.03 (0.06)</td>
<td>[1.91, 2.15]</td>
<td>1.05</td>
</tr>
<tr>
<td>$\beta_0^\xi$</td>
<td>N(−0.19, 100)</td>
<td>−0.13 (0.03)</td>
<td>[−0.18, −0.07]</td>
<td>1.10</td>
</tr>
</tbody>
</table>
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 960 hPa, 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 $\chi$ and d) $\bar{\chi}$. 
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) $\text{NAO} = 2$, b) $\text{NAO} = -2$ and c) $\text{NAO} = 0$. 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. Predicted versus recorded 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% credible intervals.

FIG 12. Quantile-quantile plots of the posterior mean of the 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


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