View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by MURAL - Maynooth University Research Archive Library

arXiv:1906.06744v1 [stat.AP] 16 Jun 2019

Bayesian spatial extreme value analysis of maximum temperatures in County Dublin, Ireland

John O'Sullivan, Conor Sweeney, Andrew C. Parnell

Abstract

In this study, we begin a comprehensive characterisation of temperature extremes in Ireland for the period 1981-2010. We produce return levels of anomalies of daily maximum temperature extremes for an area over Ireland, for the 30-year period 1981-2010. We employ extreme value theory (EVT) to model the data using the generalised Pareto distribution (GPD) as part of a three-level Bayesian hierarchical model. We use predictive processes in order to solve the computationally difficult problem of modelling data over a very dense spatial field. To our knowledge, this is the first study to combine predictive processes and EVT in this manner. The model is fit using Markov chain Monte Carlo (MCMC) algorithms. Posterior parameter estimates and return level surfaces are produced, in addition to specific site analysis at synoptic stations, including Casement Aerodrome and Dublin Airport. Observational data from the period 2011-2018 is included in this site analysis to determine if there is evidence of a change in the observed extremes. An increase in the frequency of extreme anomalies, but not the severity, is observed for this period. We found that the frequency of observed extreme anomalies from 2011-2018 at the Casement Aerodrome and Phoenix Park synoptic stations exceed the upper bounds of the credible intervals from the model by 20% and 7% respectively.

Keywords: Bayesian, Gaussian Processes, Predictive Processes, Ireland, spatial, extreme value analysis, climate extremes, GPD

1 Introduction

In this study, we produce return-levels of anomalies of daily maximum temperature extremes for an area over Ireland, for the 30-year period of 1981-2010. We apply Bayesian hierarchical modelling (Gelman et al., 2013) combined with a reduced-rank method called predictive processes to solve the computationally difficult problem of modelling data over a very dense spatial field (Banerjee et al., 2008). The role of extreme value theory (EVT) (Coles et al., 2001) is increasingly important in furthering our understanding of climate change and climate

extremes. With an increase in maximum temperature extremes through the current century projected by the Intergovernmental Panel on Climate Change (IPCC), it is important to have a better understanding of recent observations of climate extremes (IPCC, 2013). Ireland is "completely off course in terms of achieving its 2020 and 2030 emission reduction targets", according to the Climate Change Advisory Council's Annual Review 2018. Without urgent action leading to substantial reductions in greenhouse gas emissions, they warn that Ireland is unlikely to deliver on its national, EU and international obligations (Climate Change Advisory Council, 2018). A better understanding of our current climate extremes is essential to enable both the general public and policymakers to plan better in order to mitigate or avoid the many impacts of climate change; whether social, infrastructural, environmental, or economic (IPCC, 2014).

In order to model extremes of daily maximum temperature, the chosen statistical models must use EVT because extremes of temperature are rare (by definition), and occur in the tails of the distribution. Since extremes of temperature vary by location on any given day, the statistical models used should account for this spatial dependence. And since these distributions are governed by parameters that depend not only on the data, but on other (e.g., physical or mathematical) principles and constraints, it is natural to use a Bayesian framework. With this in mind, we apply a Bayesian hierarchical spatial extreme value model to a dataset of daily maximum temperatures in Ireland.

Estimating hierarchial Bayesian models using Markov chain Monte Carlo (MCMC) methods involves matrix factorisations of the order of n^3, where n is the number of locations (Guhaniyogi et al., 2011). In our case, with a dataset of daily maximum temperature from 1981 to 2010 at more than $\sim$72,000 locations, this becomes computationally infeasible. In order to overcome this problem, we focus our attention on the capital city of Ireland and its surroundings, and employ reduced-rank spatial models. These are very popular models for analysing large spatial datasets (Banerjee et al., 2008). In particular, we work with a flexible class of low-rank models called predictive process models.

We aim to begin a comprehensive characterisation of temperature extremes in Ireland for the period 1981-2010. Our contribution here expands upon existing research on historical temperature extremes in Ireland, which to our knowledge, consists only of site-specific analysis with no spatial component to the models used (*e.g.*, Walsh (2012) and Osman et al. (2015) - described in the next section). All code needed to reproduce our analysis is available in a public repository on GitHub at https://github.com/jackos13/extremes. We note that this repository does not include the data; they are available upon request from Met Éireann.

The structure of the rest of the article is as follows. Section 2 reviews the body of literature of previous work in the area. Section 3 describes the dataset and covariates used in this study. Section 4 comprises four subsections describing

the methodology: the first subsection provides an overview of EVT; the second provides an overview of standard spatial statistics theory and predictive processes; the third discusses the issue of threshold selection for our data; and in the fourth subsection, we detail our selected model. Section 5 illustrates the main results from the model-fitting stage. Section 6 discusses these results, putting them in context with other studies, and concludes the article.

2 Previous relevant work in climate extremes and spatial modelling

2.1 Extreme climate work in Ireland

In a study of Ireland's climate from 1981 to 2010, Walsh (2012) produces a table of monthly values of mean daily maximum temperatures for two synoptic stations, one at Casement Aerodrome and one at Valentia Observatory. At Casement Aerodrome (in Dublin), the annual mean daily maximum temperature is found to be 13.4°C. The monthly mean daily maximum temperature at this station ranges from 8°C in January to 19.8°C in July. A further breakdown by individual days is not provided. Osman et al. (2015) investigate temperature extremes at six different locations in Ireland. They fit a generalised Pareto distribution (GPD) to the observed extremes at each location (for the period 1961-2000), allowing the scale parameter to vary as a function of large-scale climate variables derived from reanalysis data. From these models, they then produce return-level plots for all six locations, leading them to conclude that significant changes in extreme temperature events are projected to occur in Ireland over the course of the 21st century. These include hotter summers and milder winters.

2.2 Spatial modelling for large data sets

In the context of using Bayesian hierarchical spatial models to analyse the large spatial datasets that occur frequently in environmental sciences, Guhaniyogi et al. (2011) discuss predictive process models. They point out the flexibility of these models, explaining how they can overcome the 'big n' problem of large datasets, where matrix factorisations of the order of n^3 (where n is the number of locations) can make the direct estimation of hierarchical spatial models infeasible. Finley et al. (2012), using a Bayesian hierarchical framework, apply a predictive process model to mean temperature data from 2000-2005 from weather stations situated across the northeastern U.S.A. They found that the low-dimensional predictive process model was highly effective at borrowing information over space to make accurate and precise predictions for new locations. In addition to this, the choice of a Bayesian framework meant that the authors could determine the full posterior predictive distribution at any location.

2.3 Spatial extreme value analysis

In order to produce maps of precipitation return levels and uncertainty measures over a region in Colorado in the U.S.A., Cooley et al. (2007) construct two separate Bayesian hierarchical models for extreme precipitation events: one for the intensity (using the GPD), and another for the frequency (using the Binomial distribution) of such events. The assumption underlying both models is that regional extreme precipitation is driven by a latent spatial process, characterised by geographical and climatological covariates. Their approach involves pooling all of the information from different stations in order to produce a 25-year daily precipitation return level map, which directly takes into account the parameter and interpolation uncertainty in the method itself. Shaby and Reich (2012) apply a Bayesian hierarchical spatial extreme value model to temperature extremes across Europe in order to assess the changes in the risk of widespread extremely high temperatures across agricultural land. For their data, they use annual maximum temperatures on a selected subgrid of 985 locations. They find that the risk of large percentages of cropland exceeding a high temperature threshold has probably increased in the last century, but only slightly so. Lehmann et al. (2016) use a Bayesian hierarchical framework and a block-maxima approach to model extremes of precipitation of different durations from over 1,300 weather stations in two different regions in Australia. The parameters of the distribution are modelled as spatial Gaussian processes. They found that pooling the data across space, and thus borrowing strength from neighbouring stations, leads to more precise parameter estimates and therefore superior posterior inference. This borrowing of strength is particularly important when dealing with extremes, which are, by definition, rare. In a similar study, Dyrrdal et al. (2015) also used spatial Gaussian processes in a Bayesian hierarchical model in order to produce spatial maps of extreme hourly precipitation over Norway. They specify a generalised extreme value distribution at each point in space, and allow the parameters to depend on location-specific geographic and meteorological variables, a structure similar to generalised linear modelling. Variable uncertainty is accounted for using Bayesian Model Averaging (BMA). They find that their approach performs well at estimating extreme hourly precipitation return levels, both in terms of magnitude and spatial distribution. Ghosh and Mallick (2011), using a Bayesian hierarchical approach, compare two separate data models (a GEV distribution and a variation of the GPD) to model precipitation extremes over continental U.S.A. from 1900 to 1998. They find that the peaks-over-threshold approach (i.e., using the GPD) shows better fit and improved forecast ability to the block-maxima approach (i.e., the GEV).

3 Data

For this study, we make use of a gridded dataset of daily maximum temperatures over Ireland. The data is on a 1×1 km^2 grid ($\sim$72,000 gridpoints), consisting

of daily values of maximum temperature for the 30-year period 1981-2010. This was produced by Met Éireann's Climatology and Observation division, using observational data from 138 stations and independent variables available at each grid point, such as elevation, latitude, and longitude. This observational data was interpolated onto the grid, using an inverse distance weighted algorithm (Walsh, 2017) (available following a suitable request to the Climate Enquiries Office at Met Éireann: https://www.met.ie/). Figure 1 shows the domain of the dataset. The image on the left illustrates the extent of the full dataset, which covers the state of Ireland. Geographic covariates of latitude, longitude and altitude are available on the same 1×1 km^2 grid as the temperature data. The plot on the right of Figure 1 shows the region we chose to focus on in our research.

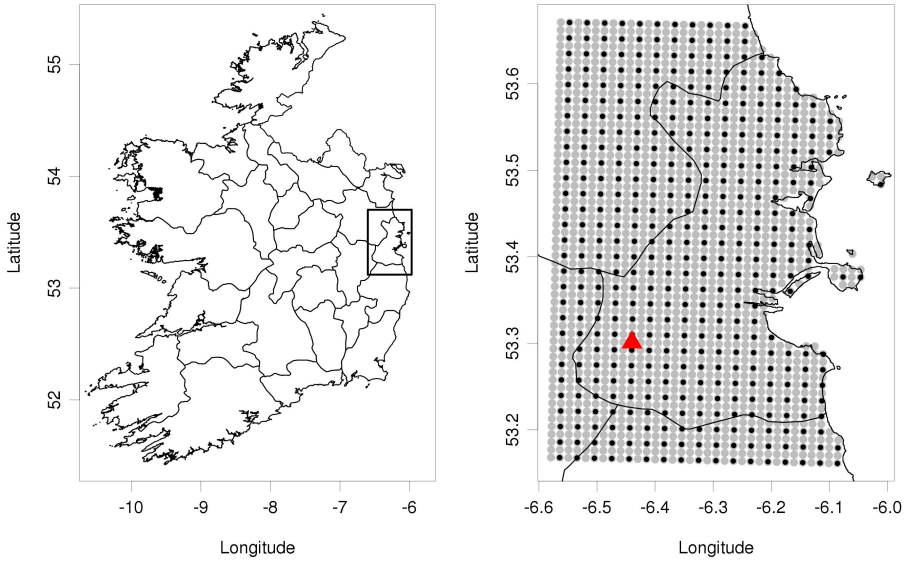

Figure 1: The data domain showing the outline of the state of Ireland (with county outlines included) is on the left of the image. The area within the black box is our study region, and is blown-up and shown on the right. Here, the full grid (grey circles) and the selected subgrid (black circles) are illustrated. A red triangle marks the location of Casement Aerodrome, for which we have synoptic station data that is used later in the paper for model evaluation purposes.

3.1 Study domain

Due to the large number of gridpoints in the dataset, we decided to focus on a smaller section to ensure that running the model was computationally feasible. We focused on an area covering County Dublin, which includes the capital city and its surroundings, located in the east of the island. This study domain is

shown in Figure 1(b). It contains approximately 1,700 gridpoints. The full grid is displayed with grey circles and the selected subgrid is shown with black circles. The subgrid is constructed by selecting every second gridpoint in both the horizontal and vertical directions. Also marked here with a red triangle is the location of Casement Aerodrome, where a synoptic meteorological station is situated. This is used in the specific-site analysis later in this section and in section 4.

3.2 Data verification

We performed various checks of the gridded data in order to compare it to the underlying synoptic station observations which were used in generating it. Two of these site-specific checks are illustrated in Figures 2 and 3. The observations here are taken from the synoptic station at Casement Aerodrome (location shown in Figure 1(b)), while the gridded data is taken from the nearest gridpoint (approximately 540 m away).

Firstly, we created a scatterplot of observed *vs.* gridded data. This shows (Figure 2) a very strong positive correlation between the observed and the gridded data, as would be expected. At all 9 sites checked across the full domain (i.e., nationwide), there was a correlation of > 0.985 between the observations at the synoptic station there and the gridded observations at the nearest gridpoint. For the data illustrated here, the correlation is 0.997.

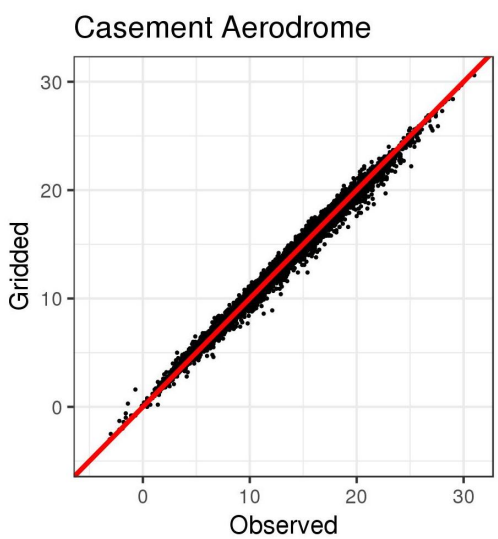

Figure 2: Shown is a scatterplot of observed *vs.* gridded temperatures at Casement Aerodrome. A very strong positive correlation between the observed and the gridded data is clearly visible (correlation = 0.997).

Secondly, in order to focus on the tail of the data (which is where our interest lies), we then considered only the data above the 99th percentile in both cases: observations and gridded data. The second panel (Figure 3) shows these two smoothed densities overlaid (where the respective thresholds were subtracted from the two datasets in order to have them both begin at 0 to aid comparison). A very close relationship can be seen between the two densities: the observations and the gridded data have a very similar shape (area of overlap calculated from the two kernel density estimations for the empirical data = 0.95). For all 9 sites nationwide checked in this manner (namely, Ballyhaise, Belmullet, Carlow Oakpark, Casement Aerodrome, Dublin Airport, Fermoy Moorepark, Malin Head, Roches Point, and Valentia Observatory), there was an overlap of > 0.89 between the overlaid densities of exceedances of the threshold for the observations and the gridded data from the nearest gridpoint.

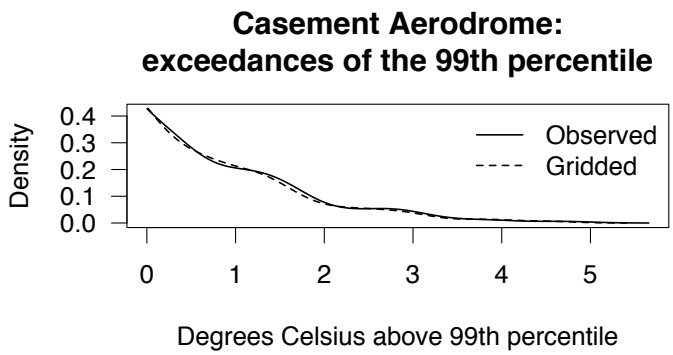

Figure 3: Shown are the two overlaid densities (of the daily maximum temperature gridded data and observations) of the exceedances above the 99th percentile at Casement Aerodrome (area of overlap = 0.95).

3.3 Temperature Anomalies

In order to focus our research on maximum temperatures which would be considered extreme relative to the time of year in which they occurred, we chose to create and then analyse temperature anomalies. Adapting a procedure used by Brown et al. (2008), we created the anomalies by removing the mean annual cycle at each location. To do this, we first calculated a 31-day moving average for each location over the 30 years of data, using only values at that location. We then averaged by calendar date in order to get a climate value for the first of January, a climate value for the second of January *etc.* at that location. Using these 366 values provided a distinct climate curve for each location - a mean annual cycle at that gridpoint. Each of these values were then subtracted from the 30 years of raw data for the corresponding day, resulting in a dataset of anomalies at each location. It is the extreme values of these anomalies that we

wish to analyse.

We first selected those anomalies which exceeded the 99th percentile at their gridpoint, in order to further explore the values which would be considered extreme. Figure 4(a) illustrates this threshold surface: it shows the value at each point above which only 1% of anomalies at that point lie, and ranges here from 5.2 - 6.2°. The threshold is generally lower near the coast - this is unsurprising, as the temperatures here are moderated by their proximity to the sea. Examining the standard deviation of the excesses at each point (Figure 4(b)) shows a similar trend. Not only are values of the threshold higher as you move away from the coast - the excesses above this threshold are also more variable, indicating more extreme temperature anomalies inland.

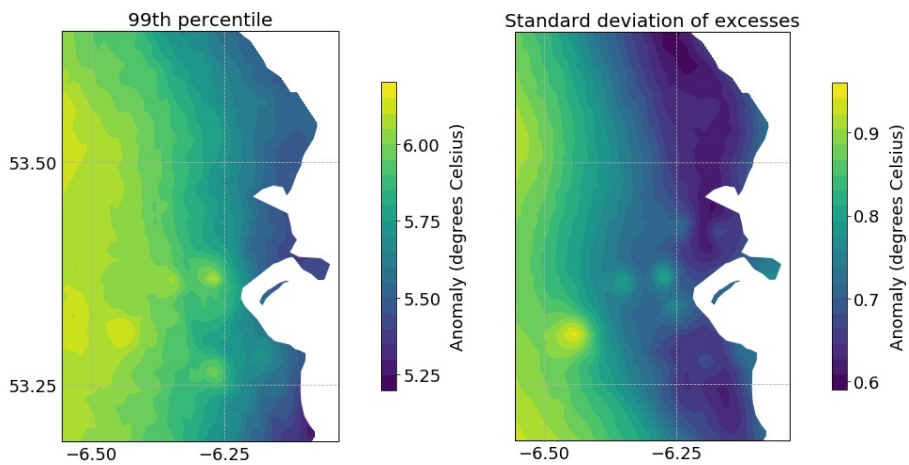

Figure 4: The figure on the left (a) is the threshold anomaly (the 99th percentile) - that is, the value for each location above which only 1% of observations lie; the figure on the right (b) illustrates the standard deviation of the excesses of this threshold. Both scales are in degrees Celsius.

4 Methodology

4.1 Extreme Value Theory

A comprehensive introduction to the field of extreme value theory (EVT) may be found in Coles et al. (2001). One common approach is to model the block maxima. We consider a sequence of independent and identically-distributed random variables, $Z_1, Z_2, \ldots$, and let $M_n = \max(Z_1, \ldots, Z_n)$ be the maximum over a block of n values; for example, we may take M_n to be the annual maxima in a multi-year set of daily maximum temperature data. The extremal types theorem states that, under certain regularity conditions, the distribution function of the M_n will converge to a specific three-parameter distribution, known

as the generalised extreme value (GEV) distribution. A major disadvantage to this approach is the fact that, by using only the maxima from a given block size, the data selected may not fully capture all extreme events (Brown et al., 2008). For example, the two most extreme events in a dataset may occur in the same year - with an annual maxima approach, only one of these will be retained.

Here we consider our dataset of daily temperature maximum values, T_{max} (the fact that these are anomalies is omitted from the notation, but should be remembered). Choosing to model this dataset with, for example, annual maxima would be quite inefficient, leading to posterior parameter estimates with large variance. An alternative is to model the excesses over a given threshold, often called a peaks-over-threshold (POT) approach (Pickands III, 1975). For this, we assume that our sequence of independent random variables, $Z_1, Z_2, \ldots$, satisfies the extremal types theorem described above. For a large enough threshold u, the distribution function of the excesses $Y = Z - u$, conditional on $Z > u$, is described approximately by the generalised Pareto distribution (GPD) with cumulative distribution function:

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

defined on the set $\{y : y > 0 \text{ and } (1 + \xi y/\sigma) > 0\}$. Here, ξ and σ are known as the shape and scale parameters, respectively, and have ranges $-\infty < \xi < \infty$ and $\sigma > 0$.

Both the block-maxima and the POT approach to EVT have been applied extensively to maximum temperatures from different sources. Examples of the use of GEV models for maximum temperatures are to be found in Plavcová and Kyselỳ (2011), Kharin et al. (2013), and Wang et al. (2016). Examples of the use of GPD models for maximum temperatures can be found in Laurent and Parey (2007), Brown et al. (2008), and Osman et al. (2015). In addition, a number of papers have compared the two approaches; see, for example, Unkašević and Tošić (2009), Parey et al. (2010), and Kioutsioukis et al. (2010).

Given the parameters of the GPD distribution, we can compute the N-year return level. For the GPD in (1), we have

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

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

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

Letting $m = N\, n_y$, where n_y is the number of observations per year, we arrive

at the following expression for the N-year return level:

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

4.2 Spatial methods

Typically when dealing with random variables recorded as point-referenced (or geostatistical) data, the location index $\mathbf{s}$ is assumed to vary continuously over D, a fixed subset of $\mathbb{R}^d$ (Banerjee et al., 2014). Let $T(\mathbf{s})$ be a vector of random variables (a random vector) at locations $\mathbf{s}$. This could be, for example, measurements of daily maximum temperatures at locations $\mathbf{s}$. While it is sensible to conceptually assume such values exist at all possible sites in the spatial domain, in practice the data is a partial realisation of this continuous spatial process. Given this partial realisation, the problem then becomes inference about this spatial process $T(\mathbf{s})$ as well as prediction at new locations. To this end, it is assumed that the covariance between the random variables at two locations depends on the distance between these locations. That is, $Cov(T(s_i), T(s_j)) = C(s_i, s_j) = C(d_{ij})$ is a function of d_{ij} where d_{ij} is the distance between locations s_i and s_j (for brevity, the dependence on sites i and j is dropped below). The method of calculating this distance must be specified (with Euclidean distance the most common approach). There are many choices of covariance functions (see Gelfand et al. (2010) for a description of several parametric models, and their relative merits). In this work, we use the Matérn class of covariance functions (Matérn, 2013) with univariate form:

$$C(d) = \begin{cases} \frac{\varsigma^2}{2^{\nu-1}\Gamma(\nu)}(\phi d)^{\nu} K_{\nu}(\phi d) & \text{if } d \neq 0 \\ \tau^2 + \varsigma^2 & \text{if } d = 0. \end{cases}$$

Here, ς^2 is the partial sill (variance of the spatial effect), τ^2 is the nugget (variance of the non-spatial effect), ν is a parameter controlling the smoothness of the spatial field, ϕ is a spatial decay parameter controlling how quickly the covariance decreases with distance, $\Gamma()$ is the gamma function, and K_{ν} is the modified Bessel function of the second kind of order ν.

There are alternatives to using the Matérn class of covariance functions, such as kernel convolution (or moving average) models, and convolutions of covariance models (Gelfand et al., 2004). However, we decided to work with the Matérn class of functions as it is a flexible class, with parameters that have attractive interpretations, and includes as special cases the exponential and Gaussian covariance functions (Banerjee et al., 2014).

4.2.1 Gaussian Processes

The process $T(\mathbf{s})$ is said to be Gaussian if, for any $n \geq 1$ and any set of sites $\{s_1, s_2, ...s_n\}$, $\mathbf{T} = (T(s_1), T(s_2), ...T(s_n))^T$ has a multivariate normal distribution (Banerjee et al., 2014). Gaussian Processes (GPs) can be thought of as

extending the finite multivariate normal distribution to infinitely many random variables; in other words, a GP is an infinite collection of variables such that every finite subset follows a multivariate normal distribution. This is a very flexible framework for modelling spatial data, as the covariance matrix can be specified using any valid covariance function.

Given realisations of the the process $T(\mathbf{s})$, and p spatially-referenced covariates at the same locations $\mathbf{s}$, let $X(\mathbf{s})$ be the $n \times (p+1)$ matrix associated with the spatial regression model:

$$T(\mathbf{s}) = X^T(\mathbf{s})\alpha + w(\mathbf{s}) + \epsilon(\mathbf{s})$$

where $X^T(\mathbf{s})\alpha$ is the mean response, $w(\mathbf{s})$ is a zero-centred GP with covariance function $C(\mathbf{s}, \mathbf{s}')$ and $\epsilon(\mathbf{s}) \overset{iid}{\sim} N(0, \tau^2)$ is an independent measurement error (see, e.g., Finley et al. (2009) and Cressie and Wikle (2015)).

There are alternative spatial methods to GPs such as the integrated nested Laplace approximation (INLA) approach proposed by Rue et al. (2009), and INLA combined with a stochastic partial differential equation approach (INLA-SPDE) proposed by Lindgren et al. (2011). However, we have chosen to use GPs due to the ease with which they fit into a Bayesian hierarchical framework. They are flexible empirical models, which are appropriate for an irregularly fluctuating and real-valued spatial surface (Diggle et al., 2007), as we have here.

Bayesian inference using GPs typically involves the need to invert or factor the covariance matrix. This becomes computationally impractical as the dimension n becomes large (that is, a large number of gridpoints), particularly when using an algorithm such as Markov chain Monte Carlo (MCMC) which involves inverting or factoring this matrix hundreds of thousands of times in one run of a model fitting. For this reason, we decided to use reduced-rank representations focusing on Gaussian predictive process models.

4.2.2 Gaussian Predictive Process Models

A comprehensive overview of hierarchical Gaussian predictive process models is given in Banerjee et al. (2014). We offer a brief summary below. Following the notation from section 4.2.1, we can avoid dealing with the dense covariance matrix induced by the zero-centred GP $w(\mathbf{s})$ by projecting it onto a subspace spanned by its realisation over the n^*-dimensional S^* where $n^* \ll n$. An optimal projection of the process w at location s, based upon its realisation over S^* is given by the kriging equation:

$$\tilde{w}(s) = E(w(s)|w^*) \tag{4}$$

where $w^* = (w(s_1^*), w(s_2^*), ...w(s_n^*))^T$. $\tilde{w}(s)$ is referred to as the predictive process derived from the parent process $w(s)$.

Further expanding the above:

$$\tilde{w}(s) = C(s, S^*)^T \;\; C(S^*, S^*)^{-1} \;\; w(S^*)$$

where $C(s, S^*)$ is the $n^* \times 1$ vector with $C(s, s_j^*)$ as the jth element; $C(S^*, S^*)^{-1}$ is the $n^* \times n^*$ covariance matrix on the subspace S^*; and $w(S^*)$ is the $n^* \times 1$ values of the process w on the subspace S^*. The important point now is that inference on the process w will now involve inverting or factoring an n^*-dimensional matrix rather than an n-dimensional one.

4.3 Hierarchical Model

The aim of this study is to produce $N-$year return levels of anomalies of extreme temperature (T_{max}). The dataset we want to model has already been described. In summary, anomalies were calculated at each location; site-specific thresholds (u_i) of the 99.5th percentile were calculated; only declustered excesses above this threshold were kept for modelling. We require location-specific estimates of the parameters from the GPD (σ_i and ξ_i), as well as the probability of exceeding the threshold (ζ_i) in order to calculate return levels using equation (3).

Following the approach of Cooley et al. (2007), we employ a Bayesian hierarchical model with three layers (see the directed acyclic graph in Figure 5). There are two separate hierarchies - both with three layers. The first hierarchy models the parameters of the GPD; the second models the probability of exceeding the threshold. The first layers in both of these hierarchies consist of modelling the data; the second layers describe the latent spatial process underlying the extremes; while the third layers consist of the prior distributions on the parameters controlling the second.
Let the declustered extreme anomalies of T_{max} be denoted by $z_k(x_i)$ where the indices i and k are such that $z_k(x_i)$ refers to the k-th exceedance ($k = 1 \ldots n_i$) at gridpoint x_i ($i = 1 \ldots n_x$). Let M_i refer to the number of exceedances. The layers in the hierarchies are described in detail below.

4.3.1 Layer 1: Extremal Data and Probability of Exceedance

We model the extremal data $z_k(x_i)$ using the GPD. This is the first layer in the first hierarchy. To ensure a positive scale parameter throughout the computations, we reparameterise $\phi = log(\sigma)$. We have two spatially-varying parameters for the distribution. The first layer in the first hierarchy is then given by:

$$z_k(x_i) \sim GPD(u(x_i), \sigma(x_i), \xi(x_i))$$

The first layer in the second hierarchy involves the parameter ζ_i (the probability of exceeding the threshold u_i - or more precisely, following declustering, the probability of being a cluster maximum at gridpoint i). This needs to be modelled in order to calculate return level surfaces using equation (3). Following the methodology of Cooley et al. (2007), we model this as a binomial random

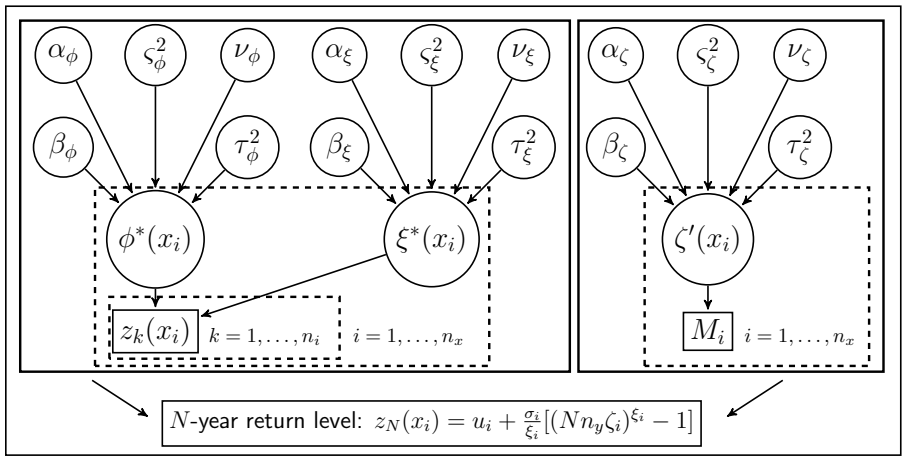

Figure 5: A directed acyclic graph (DAG) of the Bayesian hierarchical model fitted to the spatial dataset. Details of each layer and the parameters involved may be found in the text.

variable. Here, the probability of being a cluster maximum is modelled using the empirical probability as our data. It is assumed that the observed number of cluster maxima M_i at gridpoint i is a binomial random variable with m_i trials (the total number of observations in the period of study), each with a probability $\zeta(x_i)$ of being a cluster maximum:

$$M_i \sim Bin(m_i, \zeta(x_i))$$

From this point, we omit all dependence on location x_i for ease of notation, unless emphasis at a particular point is necessary.

4.3.2 Layer 2: Process

We assume that the GPD parameters vary smoothly over space and thus model the two variables ($\phi = log(\sigma)$ and ξ) as GPs. Following our decision to model these parent processes using the reduced-rank representation of predictive processes, at this layer we thus directly model $\phi^* = log(\sigma^*)$ and ξ^* (where ϕ^* and ξ^* are defined on a smaller grid, and are thus lower-dimensional, than ϕ and ξ, which are defined on the full grid). The second layer for ϕ^* in this hierarchy then is:

$$\phi^* \sim MVN(\mu_\phi, \Sigma_\phi)$$

Here, $\mu_\phi = X\alpha_\phi$ where X is a matrix of covariates for the linear regression component of the model, and α_ϕ is a vector of coefficients. Σ_ϕ is the covariance matrix of the spatial process, and is modelled using the Matérn covariance function described earlier. This has parameters β_ϕ consisting of a matrix of

range parameters (controlling how quickly the correlation drops off in different directions), ς_ϕ is the partial sill, τ_ϕ is the nuggest, and ν_ϕ is the smoothness parameter. The n-dimensional $\tilde{\phi}$ is then calculated using the kriging equation (4) detailed in section 4.2.1. The layer for ξ^* is similar.

In a similar manner (and again following the methodology of Cooley et al. (2007)), we assume ζ to vary smoothly over space. We let ζ^* be a reduced-rank representation of ζ, and then apply the transformation $\zeta' = \text{logit}(\zeta^*)$, and use predictive processes to model this reduced-rank transformed representation. As before, the kriging equation is used to calculate $\tilde{\zeta}$ given ζ'.

4.3.3 Layer 3: Hyperparameters

The third and final layers of the hierarchies consist of the 15 prior distributions on the parameters in the second layers - that is, the distributions of the α, β, ς, τ and ν hyperparameters, for each of the three parameters ϕ^*, ξ^* and ζ'.

Using Bayesian inference allows additional information about a process to be incorporated in the form of prior information. This could be of great benefit in our case, due to the scarcity of extreme data. However, Coles and Powell (1996) argue that with such scarce data, an expert may not be able to independently formulate prior beliefs about this process. With this in mind, we aim to use semi-informative prior distributions which are based on physically plausible values, but with enough flexibility so that the data is not restricted from informing the posterior distributions.

We constructed the regression matrix X to model an intercept parameter in addition to the covariates of latitude, longitude, and elevation. As elevation had a strong positive skew (skewness = 2.36), we transformed it using a log-transform. The three covariates were then scaled to be centred on 0 and with a standard deviation of 1.

For the prior distributions for the corresponding vectors of regression coefficients α_ϕ, α_ξ, and α_ζ, we chose Normal distributions centred on 0, 0 and -6 respectively for the intercept parameters. Remembering that ϕ^* is the log of the scale parameter, this corresponds to a prior distribution for the scale parameter σ which is centred on 1. The prior for the intercept for ξ^* is centred on 0, which assumes the data to have infinite support. The prior for the intercept for ζ' is centred on -6, which is approximately the logit of the probability of an observation selected at random exceeding the chosen threshold. Standard deviations of 2 are used for all three intercepts. This is arguably a little too wide in some cases (e.g., Coles and Powell (1996) point out that a shape of -1 rarely occurs when modelling the maxima of environmental data), but as the shape parameter is particularly difficult to model, we preferred to err on the side of caution, and allow the data play the dominant role in informing the posterior distribution for this parameter.

The following details regarding the non-intercept α coefficients, and the β, ν, ς^2 and τ^2 hyperparameters are identical for the ϕ^*, ξ^* and ζ' parameter surfaces. For the non-intercept α coefficients for the scaled covariates of latitude, longitude and altitude, we chose $N(0,1)$ distributions (as proposed by Dyrrdal et al. (2015)). Although we expect to see a relationship between the covariates and the parameter surfaces (e.g, more extreme excesses are expected further from the sea, and so the scale surface is likely to increase as the distance from the sea increases), we wanted to assume no relationship *a priori* and allow the data to inform the relationship between the parameter surfaces and the covariates. A standard deviation of 1 is, again, arguably a little too wide (with the scaled covariates, a value of 1 would imply a parameter surface which increases by at least 3 across the fields of latitude, longitude or altitude - a slope of this magnitude is unrealistic for all three parameters under consideration in this study). We have aimed to use physical arguments in informing the prior distributions, but again have chosen to err on the side of caution in order to allow the data play the dominant role in informing the posterior distributions.

The β and ν priors needed to be considered together. Looking again at the univariate Matérn function for the covariance between two gridpoints, where d is the distance between the gridpoints s and s':

$$C(s,s') = C(d) = \begin{cases} \frac{\varsigma^2}{2^{\nu-1}\Gamma(\nu)}(\beta d)^\nu K_\nu(\beta d) & \text{if } d \neq 0 \\ \varsigma^2 + \tau^2 & \text{if } d = 0, \end{cases}$$

it can be seen that the shape of the parameter surface depends on the spatial decay parameter β (which controls how quickly the covariance decreases with distance) and ν (which controls the smoothness of the spatial field).

In order to allow the spatial decay parameter β to differ depending on the direction of its two-dimensional coordinates (longitude and latitude), it is necessary to extend the univariate case above to incorporate a 2×2 positive definite matrix β. With this approach, the (1, 1) entry in the matrix represents the spatial decay in the direction of longitude, the (2, 2) entry represents the spatial decay in the direction of latitude, and the (1, 2) entry represents the covariance between the two. The covariance function between two points s and s' now becomes:

$$C(s,s') = C(d) = \begin{cases} \frac{\varsigma^2}{2^{\nu-1}\Gamma(\nu)}\left(\sqrt{d^T\beta^{-1}d}\right)^\nu K_\nu\left(\sqrt{d^T\beta^{-1}d}\right) & \text{if } d \neq 0 \\ \varsigma^2 + \tau^2 & \text{if } d = 0, \end{cases}$$

where d is now a 2-dimensional vector of the distance between the two points.

In order to construct a set of prior matrices β, all possible combinations of matrices were formed using values from the set {0, 0.05, 1, 10}, from which only the positive definite ones were retained. This led to a set of 20 matrices. The smoothness parameter ν is assigned a prior support of {0.5, 2.5}. It is common to use such values for ν, as the data can rarely inform about smoothness of

higher orders (Finley et al., 2009). This means there are 20 x 2 = 40 prior (β, ν) combinations to model the spatial decay across the parameter field and the smoothness of the resulting field.

Though it is possible for the spatial dependence of individual extreme observations to have a short range, the surfaces being modelled here refer to a climatological quantity rather than a weather quantity - and so it is reasonable to assume the climate will be similar at two nearby locations (Cooley et al., 2007). With this in mind, the (β, ν) combinations above allow for either very short or very long effective spatial ranges, which we take to be the distance at which the correlation equals 0.05 (Finley et al., 2013). The extent of these combinations means that correlation can either drop off very quickly with distance (with an effective spatial range of 10 km) or else reduce very slowly (reducing from 1 to 0.67 between the two furthest points on the grid). This latter combination means all points can be well within the effective spatial range of all other points in the domain of the study. We feel that this represents a sufficiently broad selection to allow for great flexibility in modelling the spatial fields of the parameters. Though physically plausible (but conservative approach), we will check that the posteriors for β assign negligible probabilities to the extreme combinations in the specified prior range (Diggle et al., 2007).

The remaining two parameters in the Matérn covariance function, ς^2 and τ^2, represent the partial sill (variance of the spatial effect) and the nugget (variance of the non-spatial effect - essentially representing the measurement error in repeated measurements at any site) respectively. It is difficult to have information on these parameters *a priori*, so we chose relatively uninformative priors. As both of these quantites are positive, we chose to model their log: $\log(\varsigma^2) \sim N(0, 1)$ and $\log(\tau^2) \sim N(-2.3, 1)$ (i.e., we assume ς^2 has a mode of 1, and τ^2 has a mode of ~ 0.1). The standard deviations of 1 (for the priors on $\log(\varsigma^2)$ and $\log(\tau^2)$) are again sufficiently large to allow the data play the dominant role in informing the posterior distributions for these parameters, without allowing for unrealistic larger values to occur.

4.4 Threshold selection

In order to ensure independence of observations in time, we declustered the dataset of extremes anomalies by removing all observations which occurred in clusters except for the maximum of this cluster (a process described by Coles et al. (2001), p.99). That is, if two or more consecutive days at any point exceeded the threshold, only the maximum of these values was retained. The model was run for varying thresholds from the 98th percentile upwards. As detailed in Coles et al. (2001), threshold selection involves a trade-off between bias and variance (higher thresholds lead to reduced bias, but increased variance). In our case, choosing too low a threshold will result in many less-extreme exceedances (i.e., extremes which are marginally above the threshold), which will threaten the asymptotic nature of the GPD model. Choosing too high a thresh-

old will result in too few datapoints, and result in large uncertainties in the posterior estimates. Here, we selected a threshold of the 99.5th percentile - results for simulations below this tended to be overly dominated by the large number of excesses marginally above the threshold, and failed to model the extreme excesses sufficiently well. Following declustering of excesses above this 99.5th percentile threshold, there was a median number of 40 independent excesses retained across the domain (the 2.5th percentile value of the number of independent excesses retained was 35, and the corresponding number for the 97.5th percentile value was 50 (see Table 1)).

Percentile \ **Threshold**	98th	98.5th	99th	99.5th	99.7th	99.9th
2.5%	136	105	73	35	21	8
25%	148	114	76	37	22	9
50%	152	117	79	40	27	10
75%	162	126	86	46	29	11
97.5%	176	137	93	50	31	11

Table 1: This table shows, for six selected threshold surfaces of the 98th percentile and above, the remaining number of declustered excesses across the surface for the 2.5%, 25%, 50%, 75% and 97.5% percentile levels. For example, the 98th percentile threshold has a median number of excesses at a gridpoint of 152, with a 2.5th percentile value of 136 excesses and a 97.5th percentile value of 176 excesses.

4.5 Model implementation

We implemented our model using the programming language R (R Core Team, 2017) and the package Rcpp (Eddelbuettel and François, 2011). A Metropolis-Hastings MCMC algorithm was used to draw samples from the posterior distributions of all parameters in the hierarchy. For both components of the model, three chains were run for 50,000 iterations. A burn-in of 10,000 iterations was discarded from each chain. The remaining chains were then thinned by retaining only every 10th sample to reduce auto-correlation. Convergence was then assessed using the $\hat{R}$ criterion (Gelman and Rubin, 1992), with values below the suggested criterion of 1.2 taken to imply convergence. The resulting simulations are presented in the next section. All code needed to reproduce this analysis is available in a public repository on GitHub at
https://github.com/jackos13/extremes.

5 Results

5.1 Posterior parameter estimates

Posterior median surfaces for the scale and shape parameters of the GPD are shown in Figure 6. The scale parameter (indicating the variance of the distribution) increases with increasing distance from the coast. Along the coastline, it has a median value of between 0.75 and 1. Further inland, it exceeds 1 at most locations, extending upwards of 1.5 in the west and south-west of the study domain. This is consistent with the surface of the standard deviation in Figure 4(b), which showed that the excesses of the threshold farther from the coast had larger variance. In other words, more variable temperature extremes are observed farther from the coast, due to the increased distance from the moderating effect of the Irish Sea. This corresponds to standard meteorological theory where the diurnal range of temperature generally increases with distance from the sea (Rafferty, 2011).

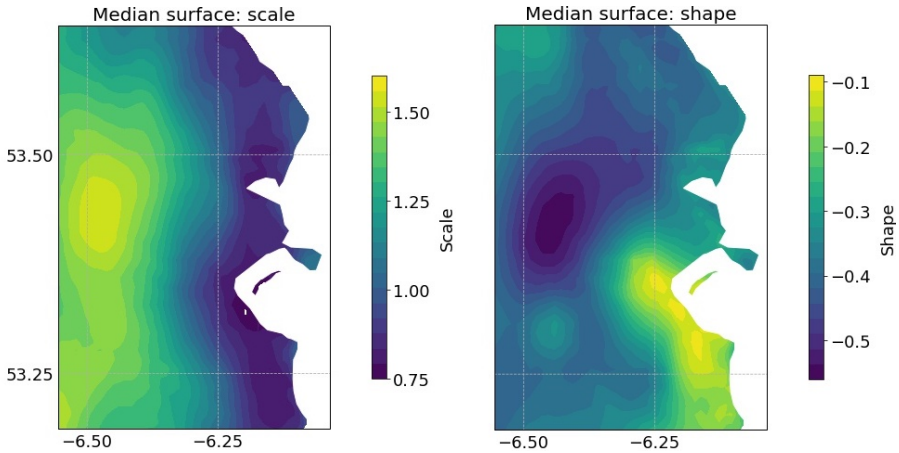

Figure 6: Shown are the posterior estimates for the median surface of the scale parameter on the left (a) and the shape parameter on the right (b).

The shape parameter is slightly more difficult to interpret. A negative shape indicates a finite upper-bound to the corresponding posterior distribution; a shape of 0 indicates that the data has infinite support; while a positive shape indicates a finite lower-bound to the corresponding posterior distribution. The scale and shape are known to be generally negatively correlated (Cooley and Sain, 2010). This is evident here, where the median shape is seen to have the opposite trend to the median scale - the shape surface generally decreases with increasing distance from the coast (Figure 6(b)).

The median surface of the zeta parameter is shown in Figure 7. The relatively simple nature of this parameter (essentially, it models the binomial probabil-

ity that a randomly chosen day is a cluster maximum) led it to converge quite quickly, with very little uncertainty in its posterior distribution. The general pattern here is for higher values nearer the coast. This indicates that extreme temperatures exceeding the threshold are more likely to be isolated incidents here, whereas further inland extremes are more likely to occur in clusters (and therefore there are fewer independent excesses retained here). The higher frequency of isolated excesses near the coast is due to the moderating effect of the sea, which makes it more unlikely for prolonged periods where successive days exceed the threshold to occur.

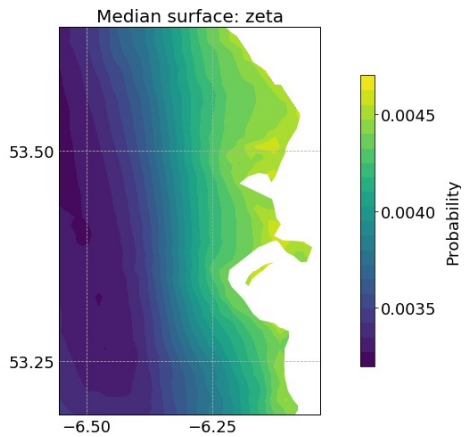

Figure 7: Shown is the posterior estimate for the median surface of the zeta parameter, the probability of a randomly selected day being a cluster maximum.

5.2 Posterior return level estimates

20- and 100-year return-level median surfaces are shown in Figure 8. The 20-year median surface is seen to range from just below 8°C to almost 10°C. The lowest levels are seen along the north-eastern coast of the study domain, while higher return levels are observed inland. The highest return levels are observed on the western side of the domain. The 100-year median surface ranges from 8°C to just under 10.7°C. A similar pattern to the 20-year return level is observed - lowest values appear along the sea in the north-east of the domain, with highest values on the western boundary. The highest part of the mean climate curve calculated for Casement Aerodrome (its location is shown in Figure 1(b)) exceeds 20°C in July. This means that if the temperature anomaly return levels seen in Figure 8 occur at this time of year, daily maximum temperatures will be in excess of 30.5°C. Since observations began at this location in 1944, there have been only two days recorded here with temperatures exceeding 30°C - one in 1975 (before the time period of this study) and one in 2006.

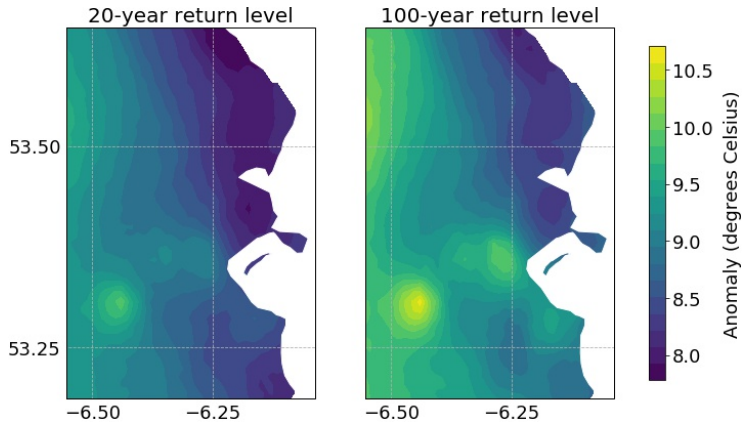

Figure 8: Shown are the 20- and 100-year return level median surfaces of the temperature anomaly.

5.3 Comparison with recent observations

Figure 9 shows the return level curve at Casement Aerodrome (location shown in Figure 1), with the observations from the synoptic station overlaid on the plot. The return period (measured in years) is displayed on a log scale for ease of interpretation. The median return level curve is shown in black; a 95% credible interval is contained within the upper and lower bounds in grey. Observations are included by plotting their empirical return period against their return level anomaly. This is found by ordering the observed excesses in a vector y such that $y_1 \leq y_2 \leq \cdots \leq y_n$ for the $i = 1 \ldots n$ excesses at that location, and then calculating the corresponding vector of return periods x with entries given by:

$$x_i = \frac{1}{1 - {}^{i}/_{(n+1)}} \times \frac{1}{npy}$$

Here npy refers to the number of observations per year. Observations from the time period of the study (1981-2010) are plotted with blue circles. Observations from the most recent eight years (2011-2018) are plotted with orange circles. There were more excesses per year in the period 2011-2018 (12 in total - 1.5 per year) than there were in the period of the study (34 in total - 1.33 per year). However, this increase in the frequency of threshold excesses does not appear to be due to an increase in the severity of threshold excesses: as can be seen in the plot, all of the more recent excesses appear in the lower region of the return level curve. For reasons of clarity, the graph only includes those points with a return period of 2 years or greater. 11 points from 1981-2010 are included in this set, and 7 points from 2011-2018. The greatest empirical return period calculated for an anomaly from 1981-2010 is almost 39 years, whereas the greatest empirical return period calculated for an anomaly for the 2011-2018 period is just under 4 years. This again demonstrates the increase in the frequency of threshold

excesses at this location, but shows that there is no corresponding increase in their severity.

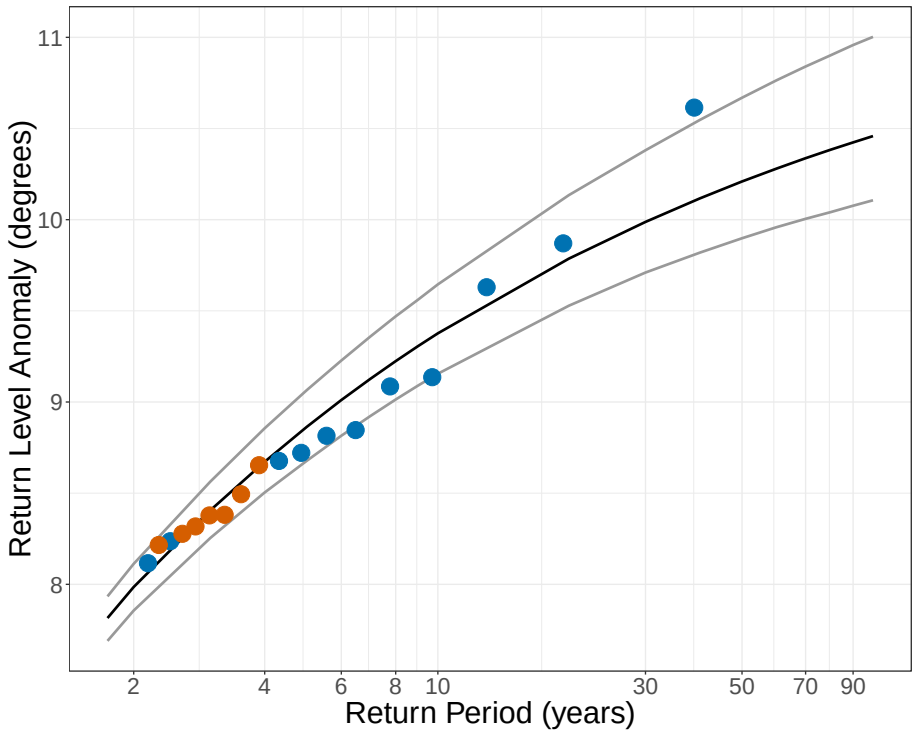

Figure 9: Shown is the return level curve for Casement Aerodrome showing the median (black) and the 95% credible interval (grey) curves, with observations superimposed for 1981-2010 (blue dots) and 2011-2018 (orange dots).

This increase in the frequency of threshold excesses from recent observational data is seen at more than one location. Figure 10 shows posterior 95% credible intervals (black line segments, with a filled circle indicating the median) of the probability of a particular day being a cluster maximum at four synoptic stations across the domain of the study: Casement Aerodrome, Dublin Airport, Dun Laoghaire, and the Phoenix Park. Superimposed on these credible intervals is the data used to fit the model - the observed site-specific probability of being a cluster maximum for the period 1981-2010 (blue diamonds). Also superimposed is the more recent (2011-2018) observed site-specific probability of being a cluster maximum (red diamonds). From this plot, it can be seen that the greatest increase in the frequency of threshold excesses is at Casement Aerodrome (where the 2011-2018 observed probability is 20% greater than the upperbound of the credible interval at that site). The next largest increase is seen at the Phoenix Park station, where the 2011-2018 observed probability is 7% greater than the upperbound of the corresponding credible interval. Of the

remaining two stations, the 2011-2018 observed probability at Dun Laoghaire is marginally outside the credible interval bounds (it is 2% greater than the upperbound at this location), while the corresponding probability for Dublin Airport is within the credible interval bounds.

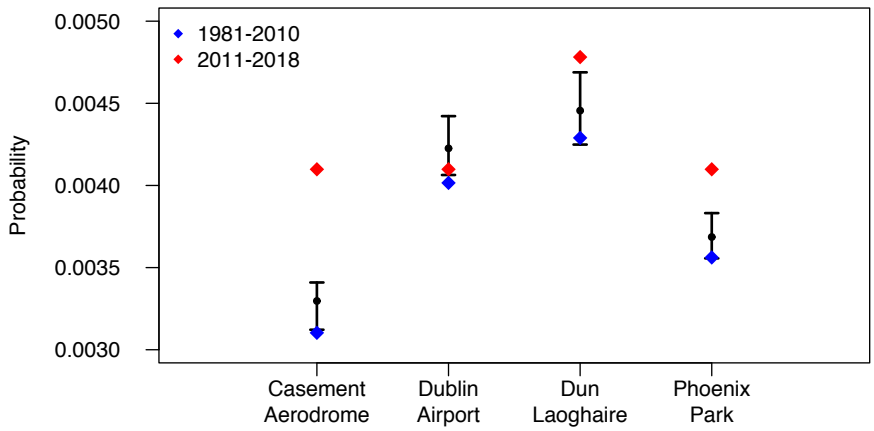

Figure 10: Shown in black is the 95% credible interval of the probability of being a cluster maximum (ζ) at four synoptic stations: Casement Aerodrome, Dublin Airport, Dun Laoghaire, and the Phoenix Park; black dots show the posterior median values; blue diamonds show the observed (1981-2010) probability of being a cluster maximum - that is, the data used to fit the model; red diamonds show the more recent (2011-2018) observed probability of being a cluster maximum.

6 Discussion and conclusion

In this research, we began a comprehensive characterisation of temperature extremes in Ireland for the period 1981-2010. We produced return-level surfaces of daily maximum temperature across County Dublin, the domain of the study. We also produced site-specific return-level curves at synoptic stations, and super-imposed data from 2011-2018 onto these plots. To our knowledge, this is the first study to combine predictive processes and EVT in the manner we have used here.

We modelled a spatial dataset of daily maximum temperatures over the domain of County Dublin, Ireland, in order to better understand the nature of temperature extremes there. We first created a dataset of anomalies to focus attention on temperatures which would be considered extreme relative to the time of year in which they occurred. In order to make the best use of this dataset, we chose a

peaks-over-threshold approach, and declustered the exceedances of this threshold to remove successive days of extremes, retaining only the maximum excess from a cluster. We then used the GPD and the reduced-rank representation of predictive processes in the first component of two three-level Bayesian hierarchical models. This resulted in (samples of) posterior densities for the scale and shape surfaces, the parameters which uniquely determine the distribution and behaviour of the temperature anomaly excesses. The second Bayesian hierarchical model was applied in a similar manner to the surface of the probability that a day selected at random is a cluster maximum which exceeds the threshold. Our chosen Bayesian approach means that uncertainty is accounted for directly in the estimate of the full posterior density, in contrast to approaches which yield only summary statistics from the target density. Values from these posterior densities for the three surfaces were drawn at random in order to calculate estimates of return-level surfaces using equation (3). This direct accounting for uncertainty in the parameter distributions means that the uncertainty of the return-level surfaces can be quantified directly, as the process yields posterior densities for these values too. We used the reduced-rank representation of predictive processes to significantly reduce the computational burden of MCMC by modelling the surfaces of interest directly on a sub-grid of the domain, while still being able to incorporate and make use of the data at every gridpoint in the domain.

Extremes are, by defintion, rare - our approach ensured that valuable information about the observed extremes was not neglected in pursuit of a relatively fast model-fitting algorithm. Modelling the GPD and binomial parameters as GPs which vary continuously over space also allowed us to make the best use of the limited data which we had - parameter values at a gridpoint are not only informed by the data at that gridpoint, but by the data at surrounding gridpoints too, with nearby gridpoints having more influence than those at a greater distance. This reduces the (large) uncertainties which result in parameter estimates from a single-site analysis (e.g, maximum-likelihood estimation at a single point). This is particularly helpful when it comes to the shape parameter, on which it can be very difficult to perform inference.

One of the problems with modelling large spatial datasets with a likelihood-based approach using MCMC algorithms is the need to invert large and dense matrices thousands of times. We tackled this problem with the use of reduced-rank representations of the latent spatial processes, and hence the computational burden was vastly reduced. The problem of modelling extremes (that is, the unavoidable scarcity of data) was approached by using spatial models to make best use of the information contained in the data, and a Bayesian hierarchy in order to set prior distributions for the parameters which were based on physical principles.

Following model fitting and convergence diagnostics, posterior parameter estimates and return level surfaces were produced. These were presented in the

previous section. The median return level surfaces showed that, for example, for both 20- and 100-year time periods, exceedances of these values in July would mean maximum daily temperatures in excess of 30.5°C at Casement Aerodrome. Further site-specific analysis here showed that, for the period 2011-2018, an increase in the frequency of extreme anomalies, but not the severity, was observed. This increased frequency of extreme anomalies at Casement Aerodrome is such that the observed probability of being a cluster maximum was 20% greater here than the upperbound of the credible interval produced by the model using the 1981-2010 data. This increase was also present to a lesser extent at the Phoenix Park and Dun Laoghaire stations (where the observed probability of being a cluster maximum was 7% and 2% respectively above their corresponding upperbounds).

Acknowledgements

The authors are grateful to Séamus Walsh and Sandra Spillane for making the gridded dataset available and for their initial input into this research. This work was supported by the Environmental Protection Agency grant number 2012-CCRP-PhD.3. Andrew Parnell's work was supported by a Science Foundation Ireland Career Development Award grant 17/CDA/4695 and by the Insight Centre for Data Analytics SFI/12/RC/2289.

References

Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2014). *Hierarchical Modeling and Analysis for Spatial Data*. CRC Press.

Banerjee, S., Gelfand, A. E., Finley, A. O., and Sang, H. (2008). Gaussian predictive process models for large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(4):825–848.

Brown, S., Caesar, J., and Ferro, C. A. (2008). Global changes in extreme daily temperature since 1950. *Journal of Geophysical Research: Atmospheres*, 113(D5).

Climate Change Advisory Council (2018). *Annual Review 2018*. Climate Change Advisory Council 2018.

Coles, S., Bawa, J., Trenner, L., and Dorazio, P. (2001). *An Introduction to Statistical Modeling of Extreme Values*, volume 208. Springer.

Coles, S. G. and Powell, E. A. (1996). Bayesian methods in extreme value modelling: a review and new developments. *International Statistical Review/Revue Internationale de Statistique*, pages 119–136.

Cooley, D., Nychka, D., and Naveau, P. (2007). Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, 102(479):824–840.

Cooley, D. and Sain, S. R. (2010). Spatial hierarchical modeling of precipitation extremes from a regional climate model. *Journal of agricultural, biological, and environmental statistics*, 15(3):381–402.

Cressie, N. and Wikle, C. K. (2015). *Statistics for Spatio-Temporal Data*. John Wiley & Sons.

Diggle, P., Ribeiro, P., and Geostatistics, M.-b. (2007). Springer series in statistics.

Dyrrdal, A. V., Lenkoski, A., Thorarinsdottir, T. L., and Stordal, F. (2015). Bayesian hierarchical modeling of extreme hourly precipitation in Norway. *Environmetrics*, 26(2):89–106.

Eddelbuettel, D. and François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, 40(8):1–18.

Finley, A. O., Banerjee, S., and Gelfand, A. E. (2012). Bayesian dynamic modeling for large space-time datasets using Gaussian predictive processes. *Journal of geographical systems*, 14(1):29–47.

Finley, A. O., Banerjee, S., and Gelfand, A. E. (2013). spbayes for large univariate and multivariate point-referenced spatio-temporal data models. *arXiv preprint arXiv:1310.8192*.

Finley, A. O., Sang, H., Banerjee, S., and Gelfand, A. E. (2009). Improving the performance of predictive process modeling for large datasets. *Computational statistics & data analysis*, 53(8):2873–2884.

Gelfand, A. E., Diggle, P., Guttorp, P., and Fuentes, M. (2010). *Handbook of Spatial Statistics*. CRC press.

Gelfand, A. E., Schmidt, A. M., Banerjee, S., and Sirmans, C. (2004). Nonstationary multivariate process modeling through spatially varying coregionalization. *Test*, 13(2):263–312.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, pages 457–472.

Gelman, A., Stern, H. S., Carlin, J. B., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis*. Chapman and Hall/CRC.

Ghosh, S. and Mallick, B. K. (2011). A hierarchical Bayesian spatio-temporal model for extreme precipitation events. *Environmetrics*, 22(2):192–204.

Guhaniyogi, R., Finley, A. O., Banerjee, S., and Gelfand, A. E. (2011). Adaptive Gaussian predictive process models for large spatial datasets. *Environmetrics*, 22(8):997–1007.

IPCC (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

IPCC (2014). *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Kharin, V. V., Zwiers, F., Zhang, X., and Wehner, M. (2013). Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic change*, 119(2):345–357.

Kioutsioukis, I., Melas, D., and Zerefos, C. (2010). Statistical assessment of changes in climate extremes over Greece (1955–2002). *International Journal of Climatology*, 30(11):1723–1737.

Laurent, C. and Parey, S. (2007). Estimation of 100-year-return-period temperatures in France in a non-stationary climate: Results from observations and IPCC scenarios. *Global and Planetary Change*, 57(1):177–188.

Lehmann, E. A., Phatak, A., Stephenson, A., and Lau, R. (2016). Spatial modelling framework for the characterisation of rainfall extremes at different durations and under climate change. *Environmetrics*, 27(4):239–251.

Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498.

Matérn, B. (2013). *Spatial Variation*, volume 36. Springer Science & Business Media.

Osman, Y. Z., Fealy, R., and Sweeney, J. C. (2015). Modelling extreme temperatures in Ireland under global warming using a hybrid peak–over–threshold and generalised Pareto distribution approach. *International Journal of Global Warming*, 7(1):21–47.

Parey, S., Hoang, T. T. H., and Dacunha-Castelle, D. (2010). Different ways to compute temperature return levels in the climate change context. *Environmetrics*, 21(7-8):698–718.

Pickands III, J. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics*, pages 119–131.

Plavcová, E. and Kyselỳ, J. (2011). Evaluation of daily temperatures in Central Europe and their links to large-scale circulation in an ensemble of regional climate models. *Tellus A*, 63(4):763–781.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rafferty, J. P. (2011). *Climate and Climate Change (The Living Earth)*. Rosen Educational Services.

Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)*, 71(2):319–392.

Shaby, B. A. and Reich, B. J. (2012). Bayesian spatial extreme value analysis to assess the changing risk of concurrent high temperatures across large portions of European cropland. *Environmetrics*, 23(8):638–648.

Unkašević, M. and Tošić, I. (2009). Changes in extreme daily winter and summer temperatures in Belgrade. *Theoretical and Applied Climatology*, 95(1):27–38.

Walsh, S. (2012). A summary of climate averages for Ireland. *Met Eireann, Dublin*.

Walsh, S. (2017). Met Éireann Climatological Note No . 16 Long-term temperature averages for Ireland 1981-2010. *Met Eireann, Dublin*.

Wang, J., Han, Y., Stein, M. L., Kotamarthi, V. R., and Huang, W. K. (2016). Evaluation of dynamically downscaled extreme temperature using a spatially-aggregated generalized extreme value (GEV) model. *Climate dynamics*, 47(9-10):2833–2849.