Modelling the clustering of extreme events for short-term risk assessment

Ross Towe[*1], Jonathan Tawn[2], Emma Eastoe[2], and Rob Lamb[3,4]

[1]School of Computing and Communications, Lancaster University, United Kingdom
[2]Department of Mathematics and Statistics, Lancaster University, United Kingdom
[3]JBA Trust, Skipton, United Kingdom
[4]Lancaster Environment Centre, Lancaster University, United Kingdom

August 6, 2019

Abstract

Reliable estimates of the occurrence rates of extreme events are highly important for insurance companies, government agencies and the general public. The rarity of an extreme event is typically expressed through its return period, i.e., the expected waiting time between events of the observed size if the extreme events of the processes are independent and identically distributed. A major limitation with this measure is when an unexpectedly high number of events occur within the next few months immediately after a T year event, with T large. Such instances undermine the trust in the quality of risk estimates. The clustering of apparently independent extreme events can occur as a result of local non-stationarity of the process, which can be explained by covariates or random effects. We show how accounting for these covariates and random effects provides more accurate estimates of return levels and aids short-term risk assessment through the use of a complementary new risk measure.

Keywords: Clustering, covariate modelling, extreme events, flood risk assessment, local non-stationarity and random effects.

1 Introduction

Floods, and other extreme weather-related hazards are often described in terms of their return period; i.e., the expected waiting time between events if the processes being described are assumed to be stationary. In recent years, the annual exceedance probability (AEP) has also been used as an alternative metric to communicate the severity of an extreme event, because it emphasises that the risk is present in any year. When the process is changing over time, such as with a trend, e.g., due to climate change, the return period is not a particularly helpful measure of risk, but the AEP is still useful, it simply changes from year to year. When the process has such non-stationarity a number of alternative ways

---

*r.towe@lancaster.ac.uk

of best presenting the associated risk over time have been proposed (Rootzén and Katz, 2013; Cheng et al., 2014). Here, we examine deficiencies in the return period and AEP descriptions for stationary processes when separate independent flood events cluster in time, and we propose an additional risk measure to address the consequent difficulty in properly communicating this risk.

During the winter of 2015/2016, the north west of the United Kingdom experienced a sequence of storm events known as Desmond, Eva and Frank. Sustained heavy rainfall caused approximately £1.6billion of economic damages, flooded 21,000 properties and severed major transport routes including the west coast train line being closed for a number of months (Agency, 2018). For some communities, such as in Carlisle, which had flooded badly in 1979, 1980, 1984 and 2005, it was the latest in a series of repeated flood events. Similar experiences have been felt elsewhere in the UK e.g., for Harbertonford on the river Harbourne in Devon. This is because Harbertonford was flooded 21 times in the last 60 years and from 1998 to 2000 was flooded on six individual occasions (Bradley, 2005; Cresswell, 2012). Consequently, a number of Environment Agency studies have tried to determine how best to alleviate the risk of flooding.

This reoccurrence of extreme events illustrates long-term problems with the miscommunication of risk to the general public and decision makers (Vogel et al., 2011; Cooley, 2013; Olseon, 2015), which stems from a historical confusion of long- and short-term risks through a single risk measure, namely the return period. The return level z_T is associated with the return period T, through the distribution function G of the annual maximum by $G(z_T) = 1 - 1/T$ and the AEP being T^{-1}. Strictly T is the return period of the annual maxima, with $T_{ALL} = -1/\log(1 - T)$ being the expected waiting time between exceedances of z_T for all of the data, where $T_{ALL} \approx T + 0.5$ for $T > 20$. A major limitation with the return period as a risk measure is that there are regular occurrences when within the few months following a T year event ($T > 50$ year) another event of similar or greater severity occurs. This undermines the reputation of statisticians and flood risk managers.

The problem is that the current method of communicating the severity of an event focusses on the process being stationary, whereas this is not often the case (Black and Werritty, 1997; Hannaford and Buys, 2012; Gilleland et al., 2017). Clustering of apparently independent extreme events exists as a result of local non-stationarity of the process. Flood risk managers need methods to assess the short-term risk of the reoccurrence of flooding after large events have been observed, which the return period is unable to capture if G changes from year-to-year.

In some environmental applications the standard convention is to model the extremes of the data as being identically distributed and to ignore the potential effects of covariates. There are two different univariate approaches: either a block maxima based approach or the method considered here, which looks at exceedances above a predetermined threshold. We adopt the latter approach due to the efficiency gains in maximising the available information in the data as well as it providing an analysis at a temporal scale for which detecting covariate-response relationships is easier (Coles, 2001). There are two alternative threshold approaches, the generalised Pareto distribution (Davison and Smith, 1990) and the non-homogeneous Poisson process (Smith, 1989). We adopt the latter approach as it is most easy to parametrise non-stationarity. For example, a linear trend is accounted for in the location parameter of the Poisson process approach unlike for the generalised Pareto approach which instead needs both the rate of exceedance and the scale parameters to change in a non-linear fashion (Coles, 2001).

The threshold based approaches typically assume that observations are independent and identically
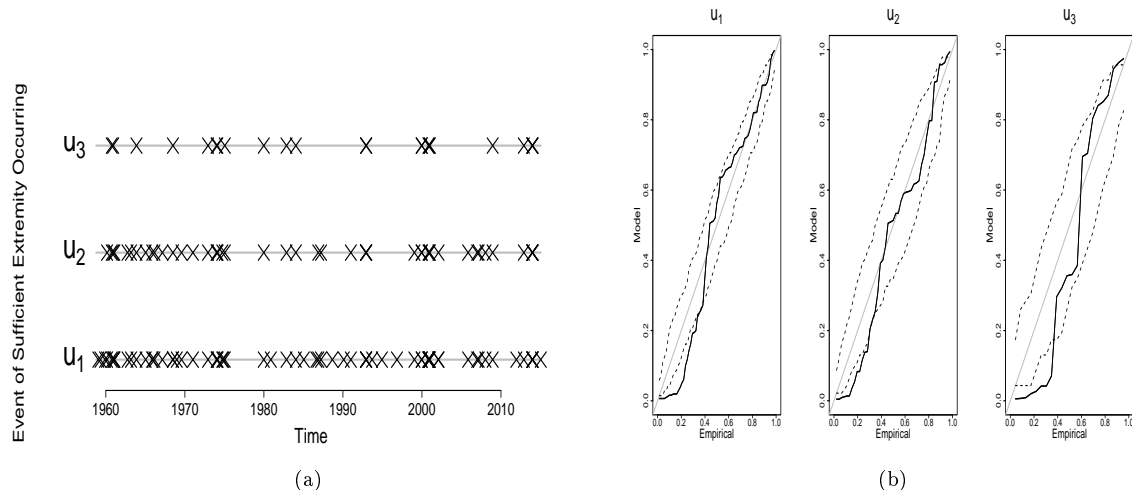
Figure 1: Exceedance and inter-arrival times for independent storm season events extracted from a south Devon river flow time series above thresholds u_1, u_2 and u_3 corresponding to 0.7, 0.8 and 0.9 quantiles. (a): Clustering of extreme events with the crosses representing the set of event times. (b): Probability-Probability plots of inter-arrival times for exceedances of u_1, u_2 and u_3 respectively, 95% tolerance bands on each plot are derived under the assumption that inter-arrivals are exponentially distributed. The grey line shows the line of equality.

distributed (iid). Consequently declustering of extreme values is often performed to obtain independent extreme events for modelling and inference. The most widely used method to decluster extreme observations is the runs method of Ferro and Segers (2003). Using a predetermined event window, w, periods with at least w consecutive observations below the threshold are deemed to define separate independent events. Only the cluster maximum is extracted from each event, to produce a set of independent extreme observations. However, for a declustered river flow time series for a south Devon catchment focussing only on the storm season data, Figure 1a shows that, over a range of thresholds, clusters of extreme events still exist even after declustering has taken place. If the events were iid, the occurrence times would be well approximated by a Poisson process with exponentially distributed inter-arrival times. The P-P plots in Figure 1b show that the inter-arrival times do not follow an exponential distribution, instead the data have a greater probability of short waiting times than is expected under this distribution, and this finding arises for three different thresholds choices.

We will show that this clustering of independent events can be described by local non-stationarity, a local change in the marginal distribution of the process. Ignoring this feature leads to biased return period estimates and an over-optimistic assessment of risk following an extreme event. In many cases, local non-stationarity might be linked to changes in climatic covariate values. Previous studies have shown that metrics such as the North Atlantic Oscillation (NAO) or the Southern Oscillation Index (SOI) influence the weather conditions in their respective spatial regions (Trigo et al., 2002; Hurrell et al., 2003). Recent research has also focused on the influence of atmospheric rivers (defined as a concentrated narrow channel of heavy vapour) and their influence on winter flooding in the United Kingdom (Lavers et al., 2013), which has the potential to identify different covariate effects by spatial location.

Detecting how a covariate affects a response such as river level can help to improve estimates of return levels and reduce the uncertainty in parameter estimates (Coles, 2001). In some cases, there is knowledge of the type of covariates that affect the process of the interest but the data on the covariate(s) are unavailable, or we may have the right covariates but we are unsure of the appropriate functional form in which to include them in the model. Therefore, we adopt similar methods to Eastoe and Tawn (2010) and Eastoe (2019) to account for the unavailable covariates, through the inclusion of random effects in the parameters of the extreme value models. These approaches are particularly beneficial for accounting for presumed large-scale climatic covariates at small environmental scales, where detecting meaningful relationships is particularly difficult. Accounting for unavailable covariate information through random effect modelling can also improve marginal estimates and reduce the uncertainty in estimates of return values through using more physically realistic statistical models to capture the behaviour of river flow data.

This paper will present a novel measure to help inform about short-term risk assessment from local non-stationarity; it will provide additional information to the return period which integrates out the local non-stationarity. The proposed measure characterises the heightened short-term risk of extreme events given the occurrence of a T year event at time t in a season. It is similar in interpretation to the definition of relative risk from epidemiology. The risk measure is $R_T^{(t)}$ with $0 < R_T^{(t)} < \infty$: if $R_T^{(t)} = 1$ there is no change in risk, if $R_T^{(t)} > 1$ ($R_T^{(t)} < 1$) there is an increased (decreased) risk of observing another extreme event for the remainder of the season. Estimates of this new measure can be derived from our models of local non-stationarity of extreme events. The risk measure is illustrated with both simulated and observed data.

The structure of the paper is as follows: Section 2 defines the required underlying univariate extreme value theory used to define the risk measure. Section 3 presents the risk measure methodology and this is illustrated through simulation studies in Section 4. The short-term risk measure is then illustrated in Section 5 in the context of a case study using peak river flow data from Devon. We also provide a practical interpretation of how this risk measure can be used for short-term risk assessment to convey more clearly the reoccurrence chance of extreme events and to help clarify the misinterpretation of an event's return period.

2   Univariate extreme value theory

2.1   Non-homogeneous Poisson process

Consider an iid sequence of random variables $Z_1, \ldots, Z_n$, and let $M_n = \max\{Z_1, \ldots, Z_n\}$. If the distribution of the linearly normalised variable $(M_n - b_n)/a_n$, $a_n > 0$, converges to a non-degenerate distribution as $n \to \infty$, this non-degenerate distribution G must be the generalised extreme value distribution (GEV) (Coles, 2001), i.e.,

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq z\right) \to G(z), \tag{2.1}$$

where $G(z)$ takes the following form,

$$G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]_+^{-\frac{1}{\xi}}\right\}, \tag{2.2}$$

with $[z]_+ = \max\{z, 0\}$. The GEV distribution function (2.2) is defined through three parameters; location $\mu \in \mathbb{R}$, scale $\sigma > 0$ and shape $\xi \in \mathbb{R}$. We may be interested in the maximum up to a scaled time $t \in (0, 1]$, i.e., $M_{1:\lfloor tn \rfloor}$, where here and subsequently $M_{i:j} = \max\{Z_i, \ldots, Z_j\}$ for $i < j$, and $\lfloor x \rfloor$ denotes the integer part of x, for example the index $\lfloor tn \rfloor$ is the number of these n random variables that have been observed up until scaled time t. The distribution of the maximum $M_{1:\lfloor tn \rfloor}$ also converges to a non-degenerate distribution as $n \rightarrow \infty$

$$\mathbb{P}\left(\frac{M_{1:\lfloor tn \rfloor} - b_n}{a_n} \leq z \right) \rightarrow G^t(z), \tag{2.3}$$

with G given by expression (2.2). The point process representation is an extension of limit (2.1) from maxima to all large values and requires the same asymptotic assumptions. Consider a sequence of point processes P_n, defined on $\mathbb{R}^2$, with

$$P_n = \left\{ \left(\frac{i}{(n+1)}, \frac{Z_i - b_n}{a_n} \right) : i = 1, \ldots, n \right\},$$

that converge to a non-homogeneous Poisson process P. The scaling of P_n is chosen so that the first component, time, is scaled to the interval $(0, 1)$ (essentially a normalised temporal index for the random variable Z). The second component of P_n, the maximum of the normalised points, converges in distribution to a non-degenerate limit, given by equation (2.1). The point process $P_n \rightarrow P$ as $n \rightarrow \infty$, on the set $[0, 1] \times (u, \infty)$ for all $u < z$ with u defined such that $u = \sup\{z : G(z) = 0\}$, for $G(z)$ in equation (2.2), where P is a non-homogeneous Poisson process with intensity of the form

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

The integrated intensity for the set $A_u = [t_1, t_2] \times [u, \infty]$ is

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

where $\Lambda(A_u)$ is the expected number of points of P in the set A_u.

In a modelling context, the asymptotic limit for the point process representation is assumed to hold for large n, i.e., above a high threshold u with the normalising constants $a_n > 0$ and b_n absorbed into the location and scale parameters of the non-homogeneous Poisson process. In the stationary case, the likelihood (Coles, 2001) for the points that are above the threshold u, denoted by $\mathbf{z}_{1:n_u} = (z_1, \ldots, z_{n_u})$, is

$$L(\mu, \sigma, \xi; \mathbf{z}_{1:n_u}) \propto \exp\left\{ -n_y \left[1 + \xi \left(\frac{u - \mu}{\sigma} \right) \right]_+^{-1/\xi} \right\} \prod_{i=1}^{n_u} \frac{1}{\sigma} \left[1 + \xi \left(\frac{z_i - \mu}{\sigma} \right) \right]_+^{-\frac{1}{\xi} - 1} \tag{2.4}$$

where n_y is the number of block of data , e.g., the number of years or flooding seasons. With this choice of n_y, the block maximum follows a $\text{GEV}(\mu, \sigma, \xi)$ distribution. Inference is performed by using maximum likelihood estimation.

For standard risk assessments we need to estimate the T year return values denoted by z_T, defined in Section 1. In the stationary case, we estimate z_T by

$$\hat{z}_T = \hat{\mu} + \frac{\hat{\sigma}}{\hat{\xi}} \left\{ \left[-\log\left(1 - \frac{1}{T} \right) \right]^{-\hat{\xi}} - 1 \right\}, \tag{2.5}$$

where $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$ are estimates obtained by maximising likelihood (2.4).

2.2 Observed covariates

In many cases, the distribution of extreme values will be dependent on covariates. This results in the points of the non-homogeneous Poisson process now being independent but non-identically distributed. The covariate space instead of the time space is now considered and as a result points are no longer assumed to be observed uniformly across the space. We initially consider the case where the covariate changes at every observation of Z_t. We focus on a scalar covariate S with corresponding density function $h(s)$ for $s \in (-\infty, \infty)$. The effects of S can be accounted for in all three parameters of the Poisson process by modelling the parameters as $\{\mu(s), \sigma(s), \xi(s)\}$ when $S = s$. The corresponding intensity is

$$\lambda(z, s) = n_y \frac{h(s)}{\sigma(s)} \left[1 + \xi(s) \left(\frac{z - \mu(s)}{\sigma(s)} \right) \right]_+^{-\frac{1}{\xi(s)} - 1},$$

and the integrated intensity on $B_u = [-\infty, \infty] \times (u, \infty)$, is

$$\Lambda(B_u) = n_y \int_{-\infty}^{\infty} \left[1 + \xi(s) \left(\frac{u - \mu(s)}{\sigma(s)} \right) \right]_+^{-\frac{1}{\xi(s)}} h(s)\mathrm{d}s. \tag{2.6}$$

The integrated intensity, stated in equation (2.6), is evaluated through the use of numerical integration as typically closed form expressions cannot be obtained. Further a kernel density estimate using all marginal S data is obtained for $h(s)$ to avoid the need for additional modelling assumptions.

In the presence of non-stationarity due to the covariate S, the likelihood in (2.4) becomes

$$L \propto \exp\left\{ -n_y \int_{-\infty}^{\infty} \left[1 + \xi(s) \left(\frac{u - \mu(s)}{\sigma(s)} \right) \right]_+^{-\frac{1}{\xi(s)}} h(s)\mathrm{d}s \right\} \times \prod_{i=1}^{n_u} \frac{1}{\sigma(s_i)} \left[1 + \xi(s_i) \left(\frac{z_i - \mu(s_i)}{\sigma(s_i)} \right) \right]_+^{-\frac{1}{\xi(s_i)} - 1},$$

where $L = L(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\xi}; \mathbf{z}_{1:n_u}, \mathbf{s}_{1:n_u})$, with $(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\xi})$ being the parameters of $\mu(s), \sigma(s), \xi(s)$ and $\mathbf{s}_{1:n_u} = (s_1, \ldots, s_{n_u})$ the covariates associated with extreme values $\mathbf{z}_{1:n_u}$. Here we can drop the $h(s)$ terms in the product as they do not involve $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$ and $\boldsymbol{\xi}$. The standard method to incorporate covariates is to use linear models, with appropriate link functions (Davison and Smith, 1990), however other more flexible methods such as generalised additive models or splines can also be used (Chavez-Demoulin and Davison, 2005; Yee and Stephenson, 2007). Regardless of the method chosen, the T year return level z_T solves the equation

$$\exp\left\{ -\int_{-\infty}^{\infty} \left[1 + \xi(s) \left(\frac{z_T - \mu(s)}{\sigma(s)} \right) \right]_+^{-\frac{1}{\xi(s)}} h(s)\mathrm{d}s \right\} = 1 - \frac{1}{T}. \tag{2.7}$$

where the effect of the covariate S is integrated out.

A special case of a covariate relationship that we will consider is when both the covariate and its effect on the non-homogeneous Poisson process remain constant over a year (or a season). We write s_i for the covariate in year i and $\mathbf{z}_{1:n_u} = \{z_{ij}, \ j = 1, \ldots, n_u(i)\}$, where $n_u(i)$ is the number of exceedances in year i and $\sum_i n_u(i) = n_u$, the total number of exceedances. The likelihood $L = L(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\xi}; \mathbf{z}_{1:n_u}, \mathbf{s}_{1:n_y})$ is then

$$\begin{aligned} L \ &\propto \ \prod_{i=1}^{n_y} \left(\exp\left\{ -\left[1 + \xi(s_i) \left(\frac{u - \mu(s_i)}{\sigma(s_i)} \right) \right]_+^{-\frac{1}{\xi(s_i)}} \right\} \prod_{j=1}^{n_u(i)} \frac{1}{\sigma(s_i)} \left[1 + \xi(s_i) \left(\frac{z_{ij} - \mu(s_i)}{\sigma(s_i)} \right) \right]_+^{-\frac{1}{\xi(s_i)} - 1} \right). \\ &\propto \ \exp\left\{ -\sum_{i=1}^{n_y} \left[1 + \xi(s_i) \left(\frac{u - \mu(s_i)}{\sigma(s_i)} \right) \right]_+^{-\frac{1}{\xi(s_i)}} \right\} \prod_{i=1}^{n_y} \prod_{j=1}^{n_u(i)} \frac{1}{\sigma(s_i)} \left[1 + \xi(s_i) \left(\frac{z_{ij} - \mu(s_i)}{\sigma(s_i)} \right) \right]_+^{-\frac{1}{\xi(s_i)} - 1}. \end{aligned}$$

2.3 Unavailable covariates

It will not always be possible to obtain an appropriate covariate S as required by the models in Section 2.2. In this situation, we want to account for covariates without explicitly stating their value. Unavailable covariates can be included in the model through the use of random effects (Laird and Ware, 1982). Random effects have been adopted to capture unexplained behaviour in a number of previous applications of extreme value theory (Cooley et al., 2007; Cooley and Sain, 2010). Similar methods to those discussed below were presented by Eastoe (2019) for the generalised Pareto distribution, however this is the first time that these models have been presented for the non-homogeneous Poisson process. As with covariates, random effects can be incorporated into all three parameters of the non-homogeneous Poisson process and can be different for all parameters. We assume that these random effects remain constant within each block of time (e.g., years) $i = 1, \ldots, n_y$ and are iid over blocks. The vector of all random effects is denoted by $\mathbf{r} = (\mathbf{r}_\mu, \mathbf{r}_\sigma, \mathbf{r}_\xi)$, with $\mathbf{r}_\mu = (r_{\mu,1}, \ldots, r_{\mu,n_y})$ where $r_{\mu,i}$ is the random effect for μ in year i, and similarly for $\mathbf{r}_\sigma$ and $\mathbf{r}_\xi$. It is assumed that $r_{i,j}$ is independent of $r_{k,l}$ when $(i,j) \neq (k,l)$ for all $i,k \in \{\mu, \sigma, \xi\}$ and $j,l \in \{1, \ldots, n_y\}$. This results in non-homogeneous Poisson process (NHPP) parameters for a given block i

$$\mu(r_{\mu,i}) = \mu_0 + \mu_1 r_{\mu,i}; \;\; \log\left[\sigma(r_{\sigma,i})\right] = \sigma_0 + \sigma_1 r_{\sigma,i}; \;\; \xi(r_{\xi,i}) = \xi_0 + \xi_1 r_{\xi,i}, \tag{2.8}$$

where $(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\xi}) = (\mu_0, \mu_1; \sigma_0, \sigma_1; \xi_0, \xi_1)$. As the parametrisation in equation (2.8) is linear in the random effect, we can assume that the random effects have zero mean and unit variance. For parsimony and computational convenience we assume that they jointly follow the multivariate Normal distribution

$$\begin{bmatrix} r_{\mu_i} \\ r_{\sigma_i} \\ r_{\xi_i} \end{bmatrix} \;\sim\; \mathrm{MVN}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \; \Sigma = \begin{bmatrix} 1 & \rho_{\mu,\sigma} & \rho_{\mu,\xi} \\ \rho_{\mu,\sigma} & 1 & \rho_{\sigma,\xi} \\ \rho_{\mu,\xi} & \rho_{\sigma,\xi} & 1 \end{bmatrix} \right), \tag{2.9}$$

where Σ is the covariance matrix and $\boldsymbol{\rho} = (\rho_{\mu,\sigma}, \rho_{\mu,\xi}, \rho_{\sigma,\xi})$ represents the correlations between the parameter specific random effects. The resulting single site likelihood is then

$$L(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \mathbf{r}_{1:n_y}, \boldsymbol{\rho}; \mathbf{z}_{1:n_u}) \;\propto\; \exp\left\{ -\sum_{i=1}^{n_y} \left[1 + \xi(r_{\xi,i})\left(\frac{u - \mu(r_{\mu,i})}{\sigma(r_{\sigma,i})}\right) \right]_+^{-\frac{1}{\xi(r_{\xi,i})}} \right\}$$

$$\times \prod_{i=1}^{n_y} \left\{ \prod_{j=1}^{n_u(i)} \frac{1}{\sigma(r_{\sigma,i})} \left[1 + \xi(r_{\xi,i})\left(\frac{z_{ij} - \mu(r_{\mu,i})}{\sigma(r_{\sigma,i})}\right) \right]_+^{-\frac{1}{\xi(r_{\xi,i})}-1} \right\} \phi(r_{\mu,i}, r_{\sigma,i}, r_{\xi,i}; \Sigma), \tag{2.10}$$

where ϕ is the function of the standard multivariate Normal distribution given in equation (2.9).

Unlike in Section 2.2, obtaining maximum likelihood estimates for the parameters and random effects in (2.10) is difficult. Following Eastoe (2019) we therefore use a Bayesian framework and estimate posterior distributions for the parameters using Markov chain Monte Carlo (MCMC) methods with a Metropolis-Hastings random walk (Gilks et al., 1995). The adoption of Bayesian methods allows us to incorporate prior information, which can be ascertained from domain experts as described in Coles and Tawn (1996). In order to obtain convergence of the MCMC chains, the distributions from which parameters are proposed are tuned by using an adaptive algorithm until the optimal acceptance rate is reached (Roberts and Rosenthal, 2009). Standard visual assessments, such as trace plots, are performed to assess the convergence of the chains.

For a historical block i for which we have an estimate of the random effects, we can determine a block i specific return level $z_{T,i}$ through solving equation (2.11)

$$\exp\left\{-\left[1+\hat{\xi}(\hat{r}_{\xi,i})\left(\frac{z_{T,i}-\hat{\mu}(\hat{r}_{\mu,i})}{\hat{\sigma}(\hat{r}_{\sigma,i})}\right)\right]_{+}^{-\frac{1}{\hat{\xi}(\hat{r}_{\xi,i})}}\right\}=1-\frac{1}{T}, \tag{2.11}$$

with the parameters and the random effect evaluated at the posterior estimates. If required, we can still obtain a single cross-year return level curve by integrating over the random effects

$$\int_{\mathbb{R}^3}\exp\left\{-\left[1+\hat{\xi}(r_{\xi})\left(\frac{z_T-\hat{\mu}(r_{\mu})}{\hat{\sigma}(r_{\sigma})}\right)\right]_{+}^{-\frac{1}{\hat{\xi}(r_{\xi})}}\right\}\phi(\mathbf{r};\hat{\Sigma})\mathrm{d}\mathbf{r}=1-\frac{1}{T} \tag{2.12}$$

and solving to obtain an estimate of z_T.

3 Short-term risk measure

3.1 Risk measure for threshold exceedance data

We propose a short-term risk measure that is updated for the rest of a block using information on the size of the largest event that has been observed up a time point t in the block. Consider a window of normalised time $[0,1]$ for which the covariate remains constant and there exists of a threshold u above which the models of Section 2 hold.

For a given time series suppose that a T-year event of size z_T is observed at time t. As a result of this event z_T we want to determine for the remainder of the time period $(t,1]$, the probability of observing an event of size greater than $z_{T^*} > u$, corresponding to a T^* year return period event. We denote this by $R^{(t)}_{T,T^*}$. Most often we envisage that interest will be in $T^* = T$, i.e., future events are as rare as the previously largest event denoted by $R^{(t)}_T$. A visualisation of $R^{(t)}_{T,T^*}$ is given in Figure 2a with $Z_t = z_T$, and with $z_{T^*} > z_T$ so the return period $T^* > T$.

Given that $M_{1:\lfloor tn\rfloor} = z_T > u$ we are interested in determining for the remaining observations $Z_{\lfloor tn\rfloor+1}, \ldots, Z_n$ the probability that $M_{\lfloor tn\rfloor+1:n} = \max\{Z_{\lfloor tn\rfloor+1}, \ldots, Z_n\} > z_{T^*}$ as $n \to \infty$. In terms of the Poisson process this is equivalent to the probability of at least one point in the set $B_{t,z_{T^*}} = [t,1]\times[z_{T^*},\infty)$ given $M_{1:\lfloor tn\rfloor} = z_T$, see Figure 2a. This probability is compared to the marginal probability of $M_{\lfloor tn\rfloor+1:n} > z_{T^*}$ without conditioning on the maximum value $M_{1:\lfloor tn\rfloor}$ being observed up to time t; this allows us to assess the effect of observing an event of size z_T. We define the measure of short-term risk by

$$R^{(t)}_{T,T^*} = \frac{\mathbb{P}(M_{\lfloor tn\rfloor+1:n} > z_{T^*} | M_{1:\lfloor tn\rfloor} = z_T)}{\mathbb{P}(M_{\lfloor tn\rfloor+1:n} > z_{T^*})}, \tag{3.1}$$

where $R^{(t)}_{T,T^*} \in [0,\infty)$. If $R^{(t)}_{T,T^*} = 1$ there is no change in risk, however if $R^{(t)}_{T,T^*} > 1 \ (< 1)$ there is an increase (decrease) in risk of an event of size z_{T^*} occurring.

3.2 Derivation of the risk measure under a covariate effect

Knowing that $M_{1:\lfloor tn\rfloor} = z_T$ provides information about the value of the covariate S and consequently about how the risk has changed for that specific block of time. The denominator in definition (3.1) can be calculated using the the non-homogeneous Poisson process with covariate S common over the block since
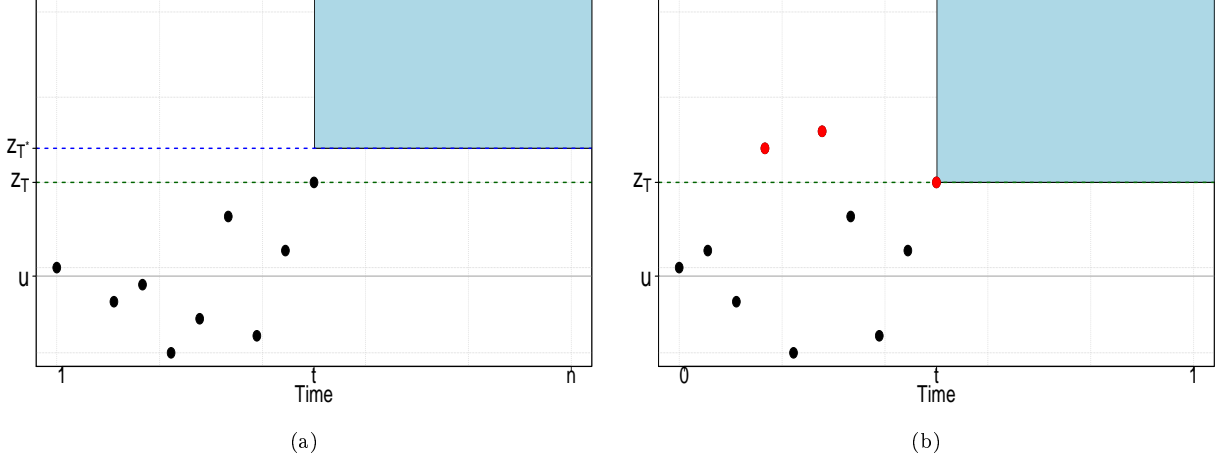
Figure 2: (a): Visualisation of the risk measure for the threshold exceedance modelling approach with $Z_t = z_T$, where z_T has a T year return period. The blue shaded area represents the region $B_{t,z_{T^*}} = [t,1] \times [z_{T^*}, \infty)$. (b): Visualisation of an alternative risk measure (discussed in Section 6), whereby we are interested in the probability of any points being observed in the region $B_{t,z_T} = [t,1] \times [z_T, \infty)$ given that a number of points were above the level z_T before time t.

$$
\begin{aligned}
\mathbb{P}(M_{\lfloor tn \rfloor + 1:n} \leq z) &= \int_{-\infty}^{\infty} \mathbb{P}(M_{\lfloor tn \rfloor + 1:n} \leq z | S = v) h(v) \mathrm{d}v \\
&= \int_{-\infty}^{\infty} \exp\left\{ -(1-t)\left[1 + \xi(v)\left(\frac{z - \mu(v)}{\sigma(v)}\right)\right]_{+}^{-\frac{1}{\xi(v)}} \right\} h(v)\mathrm{d}v,
\end{aligned}
$$

as in Section 2.2, where h is the density of the covariate. The conditional probability of $M_{\lfloor tn \rfloor + 1:n}$ being above a level z given that the maximum of $M_{1:\lfloor tn \rfloor}$ is z corresponds to the numerator of expression (3.1) and is calculated as follows,

$$
\mathbb{P}(M_{\lfloor tn \rfloor + 1:n} > z | M_{1:\lfloor tn \rfloor} = z) = \int_{-\infty}^{\infty} \mathbb{P}(M_{\lfloor tn \rfloor + 1:n} > z | S = v) h_{S|1:\lfloor tn \rfloor}(v|z) \mathrm{d}v,
$$

where

$$
h_{S|1:\lfloor tn \rfloor}(v|z) = \frac{g_{1:\lfloor tn \rfloor|S}(z|v)h(v)}{g_{1:\lfloor tn \rfloor}(z)}, \tag{3.2}
$$

and $g_{1:\lfloor tn \rfloor}(z)$ and $g_{1:\lfloor tn \rfloor|S}(z|v)$ are the marginal and conditional GEV densities of $M_{1:\lfloor tn \rfloor}$. In order to calculate these densities, we use the marginal probability

$$
\mathbb{P}(M_{1:\lfloor tn \rfloor} < z) = \int_{-\infty}^{\infty} \exp\left\{ -t\left[1 + \xi(v)\left(\frac{z - \mu(v)}{\sigma(v)}\right)\right]_{+}^{-\frac{1}{\xi(v)}} \right\} h(v)\mathrm{d}v,
$$

from which it follows that

$$
g_{1:\lfloor tn \rfloor}(z) = \int_{-\infty}^{\infty} g_{1:\lfloor tn \rfloor|S}(z|v)h(v)\mathrm{d}v,
$$

where

$$
g_{1:\lfloor tn \rfloor|S}(z|v) = \frac{t}{\sigma(v)}\left[1 + \xi(v)\left(\frac{z - \mu(v)}{\sigma(v)}\right)\right]_{+}^{-\frac{1}{\xi(v)} - 1} \exp\left\{ -t\left[1 + \xi(v)\left(\frac{z - \mu(v)}{\sigma(v)}\right)\right]_{+}^{-\frac{1}{\xi(v)}} \right\}.
$$

9

3.3 Derivation of the risk measure for an unavailable covariate

The risk measure of equation (3.1) can also be estimated in the presence of random effects. The marginal probability given in equation (3.1) becomes

$$
\begin{aligned}
\mathbb{P}(M_{\lfloor tn \rfloor + 1:n} \leq z_{T^*}) &= \int_{\mathbb{R}^3} \mathbb{P}(M_{\lfloor tn \rfloor + 1:n} \leq z_{T^*} | \mathbf{R} = \mathbf{r}) \phi(\mathbf{r}; \Sigma) \mathrm{d}\mathbf{r} \\
&= \int_{\mathbb{R}^3} \exp \left\{ -(1-t) \left[1 + \xi(r_\xi) \left(\frac{z_{T^*} - \mu(r_\mu)}{\sigma(r_\sigma)} \right) \right]_+^{-\frac{1}{\xi(r_\xi)}} \right\} \phi(\mathbf{r}; \Sigma) \mathrm{d}\mathbf{r},
\end{aligned}
$$

where $\mathbf{r} = (r_\mu, r_\sigma, r_\xi)$ and ϕ is the multivariate Normal density defined in equation (2.9). The numerator of expression (3.1) is calculated as follows,

$$
\mathbb{P}(M_{\lfloor tn \rfloor + 1:n} > z_{T^*} | M_{1:\lfloor tn \rfloor} = z) = \int_{\mathbb{R}^3} \mathbb{P}(M_{\lfloor tn \rfloor + 1:n} > z_{T^*} | \mathbf{R} = \mathbf{r}) \phi_{\mathbf{R}|1:\lfloor tn \rfloor}(\mathbf{r}; \Sigma) \mathrm{d}\mathbf{r},
$$

where

$$
\phi_{\mathbf{R}|1:\lfloor tn \rfloor}(\mathbf{r}; \Sigma | z) = \frac{g_{1:\lfloor tn \rfloor | \mathbf{R}}(z | \mathbf{r}) \phi(\mathbf{r}; \Sigma)}{g_{1:\lfloor tn \rfloor}(z)},
$$

and $g_{1:\lfloor tn \rfloor}(z)$ and $g_{1:\lfloor tn \rfloor | \mathbf{R}}(z | \mathbf{r})$ are again marginal and conditional GEV densities of $M_{1:\lfloor tn \rfloor}$. Finally, using the same strategy as in Section 3.2 we have

$$
g_{1:\lfloor tn \rfloor}(z) = \int_{\mathbb{R}^3} g_{1:\lfloor tn \rfloor | \mathbf{R}}(z | \mathbf{r}) \phi(\mathbf{r}; \Sigma) \mathrm{d}\mathbf{r},
$$

where

$$
g_{1:\lfloor tn \rfloor | \mathbf{R}}(z | \mathbf{r}) = \frac{t}{\sigma(r_\sigma)} \left[1 + \xi(r_\xi) \left(\frac{z - \mu(r_\mu)}{\sigma(r_\sigma)} \right) \right]_+^{-\frac{1}{\xi(r_\xi)} - 1} \exp \left\{ -t \left[1 + \xi(r_\xi) \left(\frac{z - \mu(r_\mu)}{\sigma(r_\sigma)} \right) \right]_+^{-\frac{1}{\xi(r_\xi)}} \right\}.
$$

4 Numerical illustration of the risk measure

To illustrate the ideas behind our proposal for the risk measure, Figure 3 shows a simplified version of the problem where the covariate S takes only three distinct values s_1, s_2 and s_3. Here s_1, s_2 and s_3 correspond to the 2.5, 50 and 97.5% quantiles of a $N(0,1)$ variable and the distribution of the random variable of interest $Z|S$ has a negative shape parameter. By comparing the conditional densities it is clear that $S = s_3$ is the most likely value of S; further Figure 3 clearly shows that z_T cannot occur if $S = s_1$. As this value of S is larger than the average S it follows that $R_{T,T^*}^{(0.4)} > 1$. More generally S will have a continuum of possible values, and an observed maximum value of z_T in a year so far will favour some range of values of S over others and hence give an updated short-term risk for the rest of the year.

Risk measure (3.1) is illustrated in the case of a linear trend in the location parameter of the NHPP($\mu(s) = \alpha + \beta s, \sigma(s) = \sigma_0, \xi(s) = \xi_0$) with 30 years' of simulated data. The following parametrisation is used: $\alpha = 0$, $\beta = 2.5$, $\sigma_0 = 1.5$ and the positive and negative shape parameter cases $\xi_0 = \pm 0.2$, with $S \sim N(0,1)$. The choice of values of the shape parameter is consistent with observed values in the environment. We condition on z_T being equal to a 1 in 100 year event and the event occurring at $t = 0.4$, i.e., 40% of the way into the storm season. We are interested in the occurrence of a value above z_{T^*} after the 1 in T year event and we consider this for $T^* \in [1, 100]$.

Figure 4 shows that the true short-term risk measure $R_{100,T^*}^{(0.4)}$ is above one for both values of ξ. When $\xi > 0$, the distribution is unbounded so there is always a non-zero probability, however small, of
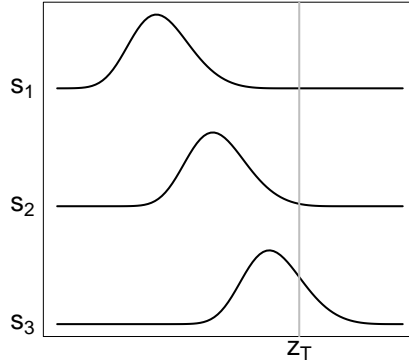
Figure 3: Conditional density for a range of z, showing the impact of the change in value of the covariate s. Three choices of the covariates are shown s_1, s_2 and s_3, which correspond to 2.5, 50 and 97.5% quantile of the covariate density. The vertical line corresponds to the value of z_T.

being above a level z_{T^*} whatever the value of the covariate. A more dramatic effect is seen for $\xi < 0$, with the short-term risk measure increasing with level in Figure 4 (b) with a much larger magnitude than when $\xi > 0$. For example, the chance of observing a 1 in 50 year event is now more like a 1 in 2 year event.

We next estimate the risk measure using the non-homogeneous Poisson process model fitted to the data and treating when the covariate S as either available or unavailable. When the covariate is available the methods in Sections 2.2 and 3.2 are used. When the covariate S is unavailable but there is evidence that the data are not identically distributed, the methods in Sections 2.3 and 3.3 provide us with random effects estimates of the unavailable covariate.

The two estimates of short-term risk in Figure 4 are both slightly lower than the truth but still reflect an increase in risk and capture the pattern of how $R_{100,T^*}^{(0.4)}$ varies with T^*. When the covariates are unavailable the estimates are lower than when they are available and have larger uncertainty intervals; this is expected as there are a larger number of parameters to estimate. The uncertainty intervals show that both estimation methods give intervals which contain the truth. The estimates in Figure 4 show the benefits of using random effect models to account for an unavailable covariate as they give similar risk measure estimates to when the correct covariate is observed. Furthermore, if we had not included random effects the risk measure would have been estimated to be equal to one, a considerable underestimate of the associated short-term risk.

Comparisons of estimates of the return levels of the annual maxima for a random year within the observation period are given in Figure 5 for the same values of the shape parameter as in Figure 4. The estimates for when the covariates are observed agree very with the truth. When the covariates are unavailable the estimates are still good, but naturally perform slightly less well than if the covariate was observed. In contrast, if the covariate is ignored and an iid GEV distribution is wrongly fitted, this estimate is considerably different, in each case giving far too heavy tails and overestimating return levels because the covariate variation across years is interpreted as variation to be modelled by the GEV. This comparison highlights the importance of accounting for covariates to explain the local non-stationarity of extreme events, as lower and more accurate return levels estimates are achieved.
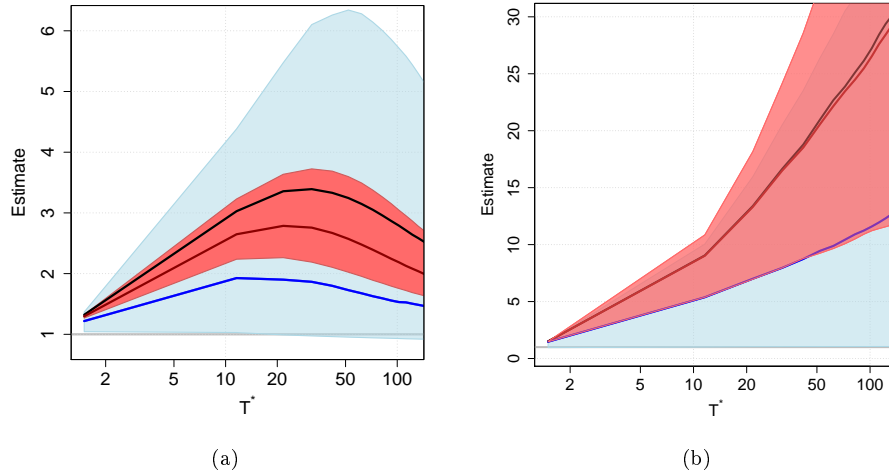
(a)   (b)

Figure 4: Estimates of the short-term risk measure $R_{100,T^*}^{(0.4)}$ when (a) $\xi > 0$ and (b) $\xi < 0$. The true values are shown by the black lines: The point estimate of the risk when the covariate is observed (unavailable) are shown by the red (blue) lines with associated 95% pointwise uncertain intervals given by the shaded sets with identical colour (intervals are overlaid). Uncertainty intervals are calculated from 50 simulated samples of 30 years' of data.

The behaviour of the return level estimates are consistent with the results presented in Carter and Challenor (1981), who showed theoretically that return values estimates from sub-samples of the population are lower than the estimates from sampling of the population. This highlights that the inclusion of covariates improves the marginal estimation of return values.

(a)   (b)
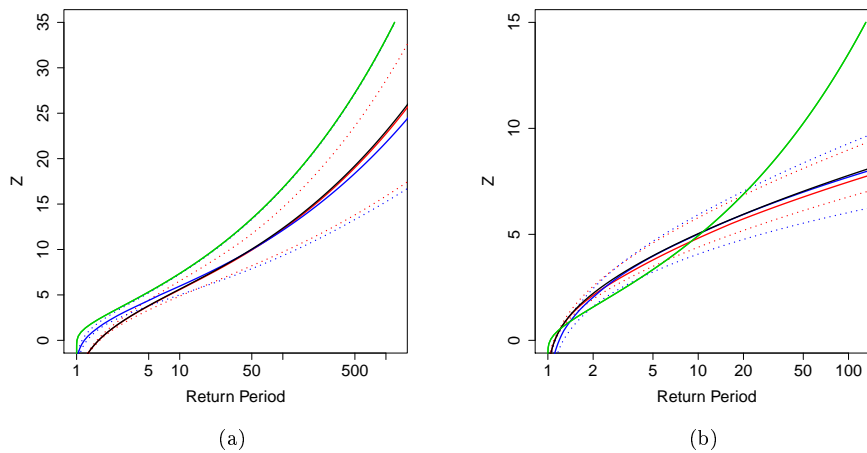
Figure 5: Comparison of the estimated return levels for different methods (a) $\xi > 0$ and (b) $\xi < 0$. The true values are shown by the black lines. The cases where the covariate is observed (unavailable) are shown by the red (blue) lines. Uncertainty intervals are calculated from 50 simulated samples of 30 years' worth of data and are shown as dotted lines in the same colours as the point estimates. The green line shows the estimate if we ignore the existence of the covariate.

5 River flow extremes for the Harbourne

5.1 Strategy and River Flow Data

The methods and risk measure developed in Sections 2 and 3 are now applied to river flow data from the river Harbourne in a catchment in South Devon in the South West of England. As discussed in Section 1, England's Environment Agency have focused on trying to reduce the risk of flooding to the communities situated along this river. They built a flood defence at Palmer's Dam, upstream of Harbertonford, to create extra capacity and reduce the risk of flooding (Cresswell, 2012). However, this was overtopped in 2012 by a flood estimated to be a 1 in 40 year event, which caused further flooding to properties in the local area (Bott et al., 2013).

Daily mean river flow data from 1998-2017 for the Rolster Bridge gauging station on the river Harbourne are considered. The data were declustered above the 97th percentile with an event window of w=7 days (Keef et al., 2009) to produce on average 3 independent events per year. The methods described in Section 2.2 were used to model the declustered data, and there was no statistically significant evidence to include random effects in the parameters of the Poisson process possibly due to a relatively short record length of 20 years. However, this changes if we examine evidence for random effects by also incorporating data from neighbouring sites with longer record lengths. This longer record length and an assumption of a common random effect across a region allows us to pool information to obtain more reliable estimates of the random effects. The regional random effect, representing a function of large scale weather events, induces common shared behaviour between river flow gauges in a spatially homogeneous region.

The additional data set consists of peak flows from five nearby river flow gauges, for the water years 1958-2013 (year defined from October to September). Data were obtained from the National River Flow Archive (NRFA, 2014), and an exploratory analysis of these data can be found on the National River Flow Archive website (NERC CEH, 2018). The gauges are all situated within the same hydrometric area as the river Harbourne, and so share similar physical characteristics, see Figure 6. However, they do not include any stations that are situated on the river Harbourne. All peak flow data sets were declustered to produce on average 5 independent events ever year.

5.2 Regional Random Effects Model

The random effects model of Section 2.3 can be adopted to include a common regional random effect using the methods outlined in Eastoe (2019), who develops an equivalent model for the generalised Pareto distribution. The model given in equation (2.8) is extended to include both site-specific effects as well as a common regional random effect across sites. Consider a set of sites that are located in a spatially homogeneous region with NHPP parameters for site d and year i given by

$$\mu_d(r_{\mu,i}) \quad = \quad \mu_{d,0} + \mu_1 r_{\mu,i}; \ \log\left[\sigma_d(r_{\sigma,i})\right] = \sigma_{d,0} + \sigma_1 r_{\sigma,i}; \ \xi_d(r_{\xi,i}) = \xi_{d,0} + \xi_1 r_{\xi,i}$$

where for a given year i the random effects $\mathbf{r}_i = (r_{\mu,i}, r_{\sigma,i}, r_{\xi,i})$ have the joint distribution defined in equation (2.9). The parameters $\mu_{d,0}$, $\sigma_{d,0}$ and $\xi_{d,0}$ are site-specific intercept terms that account for differences between the gauging stations, e.g., the catchment size. The parameters μ_1, σ_1 and ξ_1, and the random effects, are common across sites. Under the assumption of spatial independence of observations at different sites in year i, given the random effects $\mathbf{r}_i$ for that year, coupled with independence of $\mathbf{r}_1, \ldots, \mathbf{r}_{n_y}$, the likelihood function over sites is
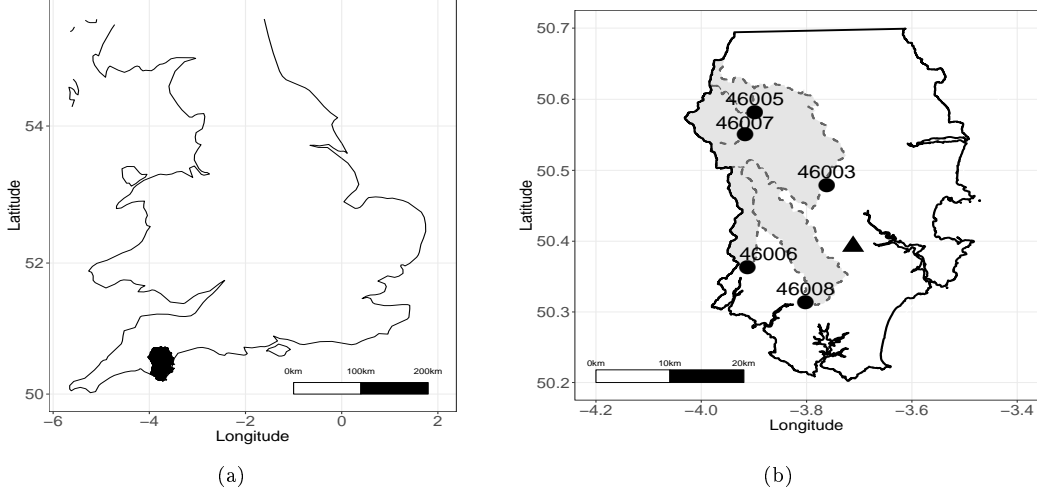
Figure 6: Hydrometric region 46, (a): location within England and (b): region map. The triangle shows the location of the gauge on the river Harbourne and the circles show the locations of NRFA gauges. The grey shaded areas are the catchments of the NRFA gauges. Further details of the gauges can be found by searching their catalogue numbers (shown on the map) at nrfa.ceh.ac.uk.

$$L(\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\xi},\mathbf{r}_{1:n_y},\Sigma;\mathbf{z}_1,\ldots,\mathbf{z}_d) \;\propto\; \exp\left\{-\sum_{d=1}^{n_d}\sum_{i=1}^{n_y}\left[1+\xi_d(r_{\xi,i})\left(\frac{u_d-\mu_d(r_{\mu,i})}{\sigma_d(r_{\sigma,i})}\right)\right]_+^{-\frac{1}{\xi_d(r_{\xi,i})}}\right\}$$

$$\times\left[\prod_{i=1}^{n_y}\left\{\prod_{d=1}^{n_d}\prod_{j=1}^{n_u(i,d)}\;\frac{1}{\sigma_d(r_{\sigma,i})}\left[1+\xi_d(r_{\xi,i})\left(\frac{z_{dij}-\mu_d(r_{\mu,i})}{\sigma_d(r_{\sigma,i})}\right)\right]_+^{-\frac{1}{\xi_d(r_{\xi,i})}-1}\right\}\phi(r_{\mu,i},r_{\sigma,i},r_{\xi,i};\Sigma)\right], (5.1)$$

where $n_u(i,d)$ is the number of exceedances of threshold u_d in year i at site d and $\mathbf{z}_j$ are these $\sum_{i=1}^{n_y}n_u(i,d)$ exceedances and z_{dij} is the jth of $n_u(i,d)$ exceedances of u_d at site d in year i.

Model (5.2) is fitted under a Bayesian framework to the NRFA peak river flow data described in Section 5.1, using an adaptive MCMC algorithm which was run for 200,000 iterations with a burn-in of 50,000. Uninformative priors were used for all parameters of the Poisson process. It was found that there was no evidence to incorporate random effects into the shape parameter, therefore random effects were incorporated only into the location and scale parameters. This is a typical finding with covariates rarely found to be statistically significant for the shape parameter in environmental applications (Eastoe and Tawn, 2009; Eastoe, 2019). As a result of this the random effects model of Section 2.3 is simplified so that $\mathbf{r}_i = (r_{\mu,i}, r_{\sigma,i})$ follows a standard bivariate Normal with correlation parameter ρ and ξ_d is the shape parameter for site d.

The estimated random effects in Figures 7a and 7b show that it is unsuitable to assume that the location and scale parameter of the Poisson process remain constant over years. The location parameter shows much greater variability across years than the scale parameter. In particular the largest values for the location random effect correspond to known flood events such as those in the water years 1990 and 2003 (Bott et al., 2013). Interestingly the random effect in the scale parameter shows a clear positive value in 1979, corresponding to known floods in the region (Bott et al., 2013), whereas there is no significant random effect in the location parameter for this year. This particular year has the largest flow across all river flow gauges and a particularly large value of NAO, which corresponds to

wet winters. It saw several more typical size extreme events as well as the large flooding event, so the large value of the random effect in the scale parameter captures both of these features. A change in the random effect of the location parameter would be insufficient as it could not cover the associated change in the variance of the threshold excesses in this particular year. The correlation of the random effects was estimated to be 0.62 (0.13, 0.88). Within-series dependence of both sequences of random effects was explored to see whether we should extend the model e.g., using an autoregressive process to model the random effects. However, there was no evidence for serial correlation and so we retain the simpler model assuming random effects are independent over years.

(a)   (b)
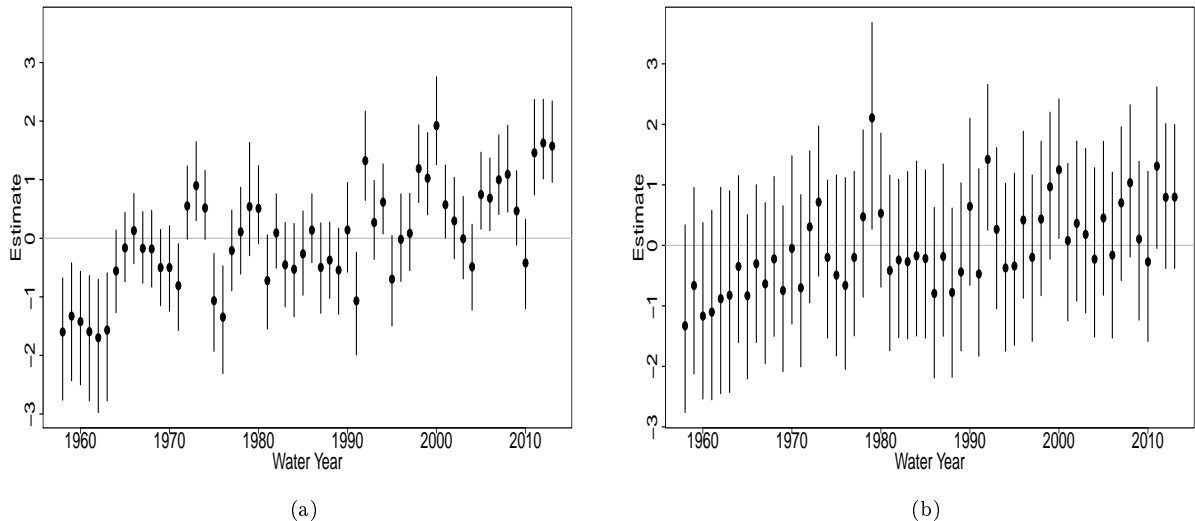
Figure 7: Random effect estimates for the location (a) and scale (b) parameters plotted against water years. The vertical lines represent 95% credible intervals.

5.3   Short-term Risk Measure Estimates

We now estimate the short-term risk measure derived in Section 3 for two of the five peak flow river gauges discussed in Section 5.2. Figure 8 presents how $R^{(t)}_{T,T^*}$ changes depending on t, T and T^* for each gauge. Two different time points t, December and June, are considered with $T = 10, 100, 1000$. The short-term risk measure estimates are larger when conditioning on more severe events. For each gauge, there is a significant increase in risk for the 100 and 1000 year events as shown in Figure 8. We can see the trade-off between the size (return period T) and timing, t, of the event in influencing the chance of observing extreme event of return period T^*, with $R^{(t)}_{T,T^*}$ appearing to decrease as a function of t. As expected the confidence interval is widest when $T = 1000$ years. By conditioning on an extreme event occurring later in the water year, in this case June, the short-term risk measure decreases due to the shorter amount of time remaining in the season in which to observe another extreme event. Although the regional random effects are common across the gauges, the risk estimates are different due to the site-specific parameters of the non-homogeneous Poisson process: so the risk measure captures the different effects of extreme events at each site.
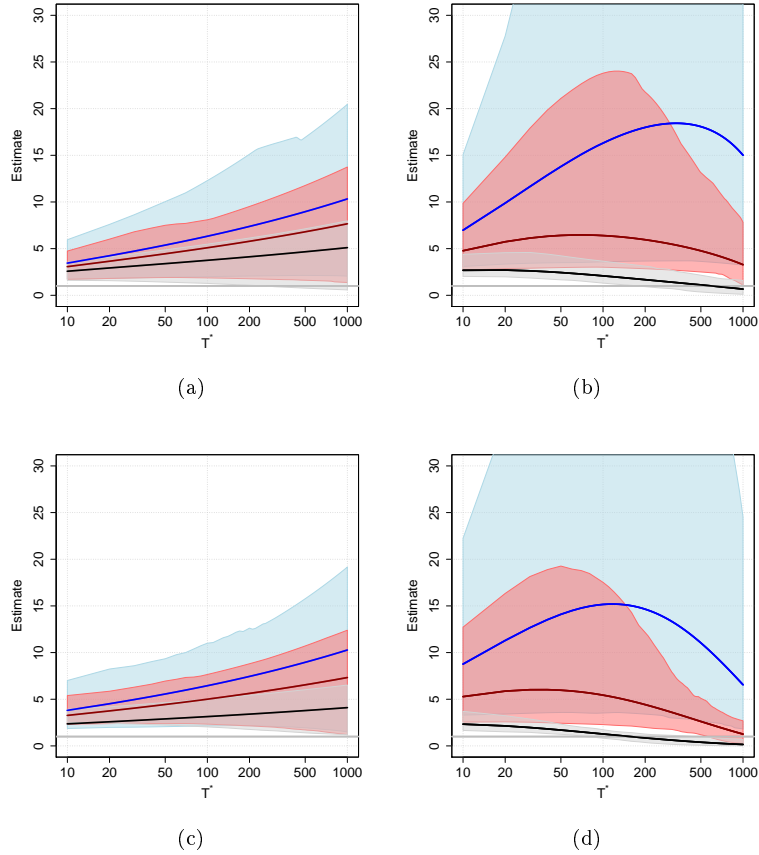
15

Figure 8: Short-term risk measure $R_{T,T^*}^{(t)}$ for gauges 46006 (a,c) and 46008 (b,d) if an extreme event occurred in December ($t = 0.20$) or June ($t = 0.80$) (top and bottom row respectively). The black, red and blue lines correspond to already observing a 1 in T=10, 100 and 1000 year event with the corresponding shaded areas providing 95% credible intervals.

16

5.4 Relating the results to the River Harbourne

To assess the value of the annual regional random effect estimates $(r_{\mu,i}, r_{\sigma,i})$, we compare them with the associated annual number of threshold exceedances for the river Harbourne in Figure 9. For each random effect there does seem to be an association, which is stronger for the location parameter.

In order to assess the validity of the random effects shown in Figure 7, the random effect estimates are taken as fixed and the model (2.8) is fitted to the data from the gauge on the river Harbourne, denoted by site H. The Poisson process parameter estimates of $\mu_{1,H}$ and $\sigma_{1,H}$ are significantly different from zero, suggesting we have good covariates in the form of $(r_{\mu,i}, r_{\sigma,i})$. However, the estimate of the local location parameter $\mu_{1,H}$ for the spatial random effect was statistically significantly less than μ_1 (the regional location parameter estimate for the spatial random effect) so not all aspects of the regional model translate to the river Harbourne.
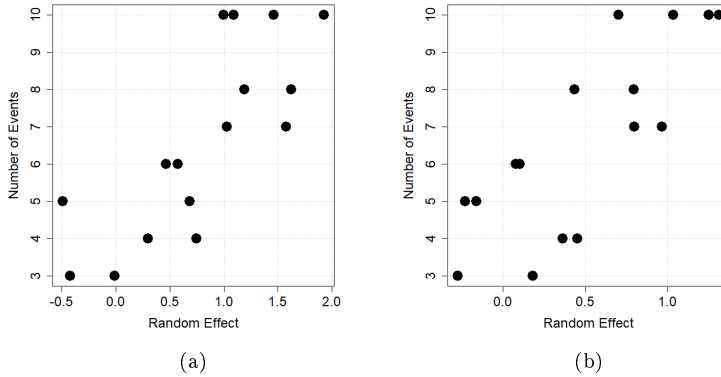

(a) (b)

Figure 9: Comparison of the regional random effects estimate for the (a) location and (b) scale parameters with the number of events for the Rolster Bridge gauging station from the river Harbourne.

Figure 10 shows both return level and risk measure plots for the Harbourne derived from using the regional random efforts and treating them as fixed. Figure 10a shows the return level curve for each annual random effect, illustrating how much the annual maximum distribution varies over time, as well as the time averaged estimate which gives us a best estimate for the future return levels when excluding long-term climate changes. For comparison we also show the estimated return level curve based on a GEV fit to the annual maxima from the Harbourne, similarly to Figure 5. Again we find that, relative to fitting with random effects, this single-site GEV fit overestimates the risk of large events, i.e., events with return periods of more than 100 years. The risk measure estimates, shown in Figure 10b, show that the short-term risk of flooding for the Harbourne increases as the size of the conditioning event increases. For example, for a risk assessment in December (20% of the way into the year) the increased risk of a 100-year event ($T^* = 100$) in the remainder of the year doubles from 3 to 6 if the previously observed event had a return period of $T = 10$ or 1000 years, i.e., these updated odds of these events for the rest of the year are $1 : 33$ to $1 : 17$.
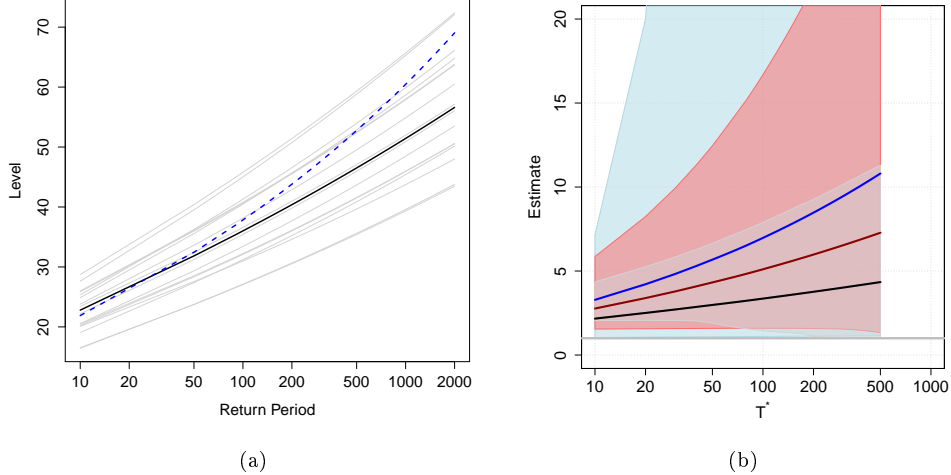
| (a) | (b) |

Figure 10: (a): Return level estimates for the gauge on the river Harbourne. The blue dashed line is the estimate from fitting a GEV distribution to the annual maxima. The grey lines show the return level estimates for each year conditional on its random effect estimate with the black line showing the cross-year posterior averages; (b): Short-term risk measure $R_{T,T^*}^{(t)}$ for the river Harbourne if an extreme event occurred in December ($t = 0.20$). The black, red and blue lines correspond to already observing a 1 in T=10, 100 and 1000 year event with the corresponding shaded areas providing 95% credible intervals.

6   Discussion

There has been much media coverage of the reoccurrence of extreme events with a seemingly long return period at a single location. What is often ignored is the fact the definition of the return periods of such events is only valid when the process is stationary. Clustering of apparently independent extreme events in a short period of time provides clear evidence that such an assumption is no longer valid. As a result, using return periods to communicate the severity of an event is no longer suitable for short-term risk assessment.

We have developed a short-term risk measure to convey the change in the risk of observing extreme events later in a storm season after observing an initial extreme event. This measure captures the local non-stationarity in time through the incorporation of covariate information, either from observations or through the use of random effects. The covariate information allows us to improve the accuracy of the statistical models and also to produce more reliable and informative return value estimates. This novel methodology helps to provide an advancement towards improving the long-term modelling of extreme flooding in the presence of temporal variability, an issue raised by the National Flood Resilience Review (Her Majesty's Government, 2016; Tawn et al., 2018; Towe et al., 2018).

The methodology was applied to river flow data from South Devon. The models fitted in Section 5 showed that there is clear evidence of a similar inter-year non-stationarity over sites in this area. The random effect were compared against the the climate index NAO but no significant relationship was found. These random effects estimates were then used to estimate the short-term risk measure. For a hydrometric area we have shown that there is a clear change in risk of observing an extreme event once one has already been observed. Naturally, the magnitude of this change in risk is a function of

18

the size of the original event and the time at which it was observed.

The short-term risk measure developed in Section 3.1 is one particular method of assessing risk but alternative risk measures could also be considered. We may be interested in the probability of there being an exceedance of z_T for the remainder of the time period given that there have been n exceedances of z_T so far in the year. Figure 2b illustrates this event with $n = 3$. The risk measure could then be constructed as the ratio of this probability with the marginal probability $\mathbb{P}(M_{\lfloor tn \rfloor:n} > z_T)$.

References

Agency, E. (2018). Estimating the economic costs of the 2015 to 2016 winter floods. Technical report, Enviroment Agency, Bristol.

Black, A. R. and Werritty, A. (1997). Seasonality of flooding: a case study of North Britain. *Journal of Hydrology*, 195(1-4):1–25.

Bott, J., Meredith, B., and M., B. (2013). Flood investigation report: Devon floods 19th - 24th December 2012. Technical report, Devon County Council.

Bradley, W. T. (2005). Effective flood alleviation design and construction. *Proceedings of the Institution of Civil Engineers-Municipal Engineer*, 158(2):107–113.

Carter, D. J. T. and Challenor, P. G. (1981). Estimating return values of environmental parameters. *Quarterly Journal of the Royal Meteorological Society*, 107(451):259–266.

Chavez-Demoulin, V. and Davison, A. C. (2005). Generalized additive modelling of sample extremes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(1):207–222.

Cheng, L., AghaKouchak, A., Gilleland, E., and Katz, R. W. (2014). Non-stationary extreme value analysis in a changing climate. *Climatic Change*, 127(2):353–369.

Coles, S. G. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics. Springer London, London.

Coles, S. G. and Tawn, J. A. (1996). A Bayesian analysis of extreme rainfall data. *Applied Statistics*, 45(4):463–478.

Cooley, D. (2013). Return periods and return levels under climate change. *In: AghaKouchak A., Easterling D., Hsu K., Schubert S., Sorooshian S. (eds) Extremes in a Changing Climate*, 65:97–114.

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. (2010). Spatial hierarchical modeling of precipitation extremes from a regional climate model. *Journal of Agricultural, Biological, and Environmental Statistics*, 15(3):381–402.

Cresswell, R. (2012). South Devon catchment flood management plan: Summary report. Technical report, Environment Agency.

Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 52(3):393–442.

Eastoe, E. F. (2019). Non-stationarity in peaks-over-threshold river flows: a regional random effects model. *Environmetrics*.

Eastoe, E. F. and Tawn, J. A. (2009). Modelling non-stationary extremes with application to surface level ozone. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 58(1):25–45.

Eastoe, E. F. and Tawn, J. A. (2010). Statistical models for overdispersion in the frequency of peaks over threshold data for a flow series. *Water Resources Research*, 46(2):W02510.

Ferro, C. A. T. and Segers, J. (2003). Inference for clusters of extreme values. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):545–556.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. (1995). *Markov chain Monte Carlo in Practice.* CRC press.

Gilleland, E., Katz, R. W., and Naveau, P. (2017). Quantifying the risk of extreme events under climate change. *CHANCE*, 30(4):30–36.

Hannaford, J. and Buys, G. (2012). Trends in seasonal river flow regimes in the UK. *Journal of Hydrology*, 475:158–174.

Her Majesty's Government (08 September 2016). National Flood Resilience Review. Technical report.

Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M. (2003). An overview of the North Atlantic Oscillation. *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*, pages 1–35.

Keef, C., Tawn, J. A., and Svensson, C. (2009). Spatial risk assessment for extreme river flows. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 58(5):601–618.

Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 1:963–974.

Lavers, D. A., Allan, R. P., Villarini, G., Lloyd-Hughes, B., Brayshaw, D. J., and Wade, A. J. (2013). Future changes in atmospheric rivers and their implications for winter flooding in Britain. *Environmental Research Letters*, 8(3):034010.

NERC CEH (2018). National River Flow Archive (nrfa). Technical report, NERC CEH, Wallingford.

NRFA (2014). Hydrometric areas for Great Britain and Northern Ireland. *NERC-Environmental Information Data Centre*.

Olseon, T. (2015). The '100-year flood' fallacy: Return periods misleading in communication of flood risk. *Earth Magazine*.

Roberts, G. O. and Rosenthal, J. S. (2009). Examples of adaptive MCMC. *Journal of Computational and Graphical Statistics*, 18(2):349–367.

Rootzén, H. and Katz, R. W. (2013). Design life level: quantifying risk in a changing climate. *Water Resources Research*, 49(9):5964–5972.

Smith, R. L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science*, 4(4):367–377.

Tawn, J. A., Shooter, R., Towe, R. P., and Lamb, R. (2018). Modelling spatial extreme events with environmental applications. *Spatial Statistics*, 28:39–58.

Towe, R. P., Tawn, J. A., and Lamb, R. (2018). Why extreme floods are more common than you might think. *Significance*, 15(6):16–21.

Trigo, R. M., Osborn, T. J., and Corte-Real, J. M. (2002). The North Atlantic Oscillation influence on Europe: climate impacts and associated physical mechanisms. *Climate Research*, 20(1):9–17.

Vogel, R. M., Yaindl, C., and Walter, M. (2011). Nonstationarity: Flood magnification and recurrence reduction factors in the United States. *Journal of the American Water Resources Association*, 47(3):464–474.

Yee, T. W. and Stephenson, A. G. (2007). Vector generalized linear and additive extreme value models. *Extremes*, 10(1-2):1–19.