Joint Estimation of Extreme Spatially Aggregated Precipitation at Different Scales through Mixture Modelling

Jordan Richards[A,B,*], Jonathan A. Tawn[B], and Simon Brown[C]

[A]Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Saudi Arabia

[B]Department of Mathematics and Statistics, Lancaster University, UK

[C]Hadley Centre for Climate Prediction and Research, Met Office, UK

[*]Correspondence to: jordan.richards@kaust.edu.sa

Abstract

Although most models for rainfall extremes focus on pointwise values, it is aggregated precipitation over areas up to river catchment scale that is of the most interest. To capture the joint behaviour of precipitation aggregates evaluated at different spatial scales, parsimonious and effective models must be built with knowledge of the underlying spatial process. Precipitation is driven by a mixture of processes acting at different scales and intensities, e.g., convective and frontal, with extremes of aggregates for typical catchment sizes arising from extremes of only one of these processes, rather than a combination of them. High-intensity convective events cause extreme spatial aggregates at small scales but the contribution of lower-intensity large-scale fronts is likely to increase as the area aggregated increases. Thus, to capture small to large scale spatial aggregates within a

single approach requires a model that can accurately capture the extremal properties of both convective and frontal events. Previous extreme value methods have ignored this mixture structure; we propose a spatial extreme value model which is a mixture of two components with different marginal and dependence models that are able to capture the extremal behaviour of convective and frontal rainfall and more faithfully reproduces spatial aggregates for a wide range of scales. Modelling extremes of the frontal component raises new challenges due to it exhibiting strong long-range extremal spatial dependence. Our modelling approach is applied to fine-scale, high-dimensional, gridded precipitation data. We show that accounting for the mixture structure improves the joint inference on extremes of spatial aggregates over regions of different sizes.

Keywords— extreme precipitation; mixture modelling; spatial aggregates; spatial conditional extremes

1 Introduction

1.1 Motivation

In light of climate change, there is a real need to assess the resilience of national infrastructure to extreme rainfall, both now and under future warming scenarios. This requires knowledge of extreme rainfall behaviour at a range of scales and durations, from the very small and short (about 1km and 10 minutes) for building design and urban drainage, up to thousands of square kilometres and days for river catchments and flooding defence. Acquiring such knowledge is very difficult as observations tend to be sparse spatially, with time series too short in length to reliably inform practitioners about future risk. Climate models can provide spatially complete data for both the present and the future, but when run at resolutions that are too coarse to resolve convection they struggle to reproduce realistic extreme rainfall (Weller et al., 2013). With the advent of convection permitting models (CPMs) run at resolutions that can explicitly represent convection, albeit only at the larger scales, the realism of extreme rainfall is much improved particularly in summer when convection is the dominant mechanism (Kendon et al., 2012; Chan et al., 2014; Kendon et al., 2021). However, such CPMs are very expensive to run and thus currently produce data records which are of insufficient length to characterise extreme rainfall with high precision particularly if considering the area affected by extreme rainfall.

Building a parsimonious model for extreme precipitation at multiple spatial scales is difficult, as the important aspects of the underlying processes that drive the extremal behaviour of precipitation change with spatial resolution. Extreme precipitation is caused by a mixture of known and well-understood processes. The primary drivers are convection, which produces spatially localised, high-intensity precipitation events over short time scales (Schroeer et al., 2018), and processes that cause large scale extreme events, e.g., frontal systems producing stratiform precipitation of lower intensity but affecting a much larger area for a longer duration (Berg et al., 2013; Catto and Pfahl, 2013; Gregersen et al., 2013); Figure 1 shows observations of these two classes of process. The distinction between these two types of events and their space-time characteristics is an important consideration when modelling the upper-tail behaviour of aggregates of rain over different spatial scales; we highlight major differences in the marginal, and extremal dependence, structures of both processes in our application, and show that the relative contribution of frontal events to extreme spatial aggregates of precipitation increases with the size of the region over which aggregates are made. In addition, future changes due to global warming may not affect these two mechanisms equally.

Whilst inference on precipitation extremes can be conducted at different spatial resolutions separately (Institute of Hydrology, 1999), ideally we would adopt an approach that can be used for joint estimation of these variables at multiple spatial scales simultaneously; the benefits include a framework for modelling dependence between aggregates over different regions and physically consistent inference, i.e., a natural ordering for quantiles of aggregates over nested regions (Richards et al., 2022a). Moreover, this approach has the potential to increase the reliability of inference on aggregates at lower-resolution spatial scales as more data will be used for modelling.

In this work we focus on the spatial aggregation of short duration (hourly) extreme rainfall and the modelling of the mixture between convective and non-convective rainfall, showing how that mixture depends on aggregation scale. We restrict the spatial scales considered to the native resolution of the CPM used (2.2km) up to that of UK river catchments, the largest being the Severn at $11,000\text{km}^2$, in anticipation of future work that will consider the issues of aggregation over time and how future extreme rainfall might change in the future.
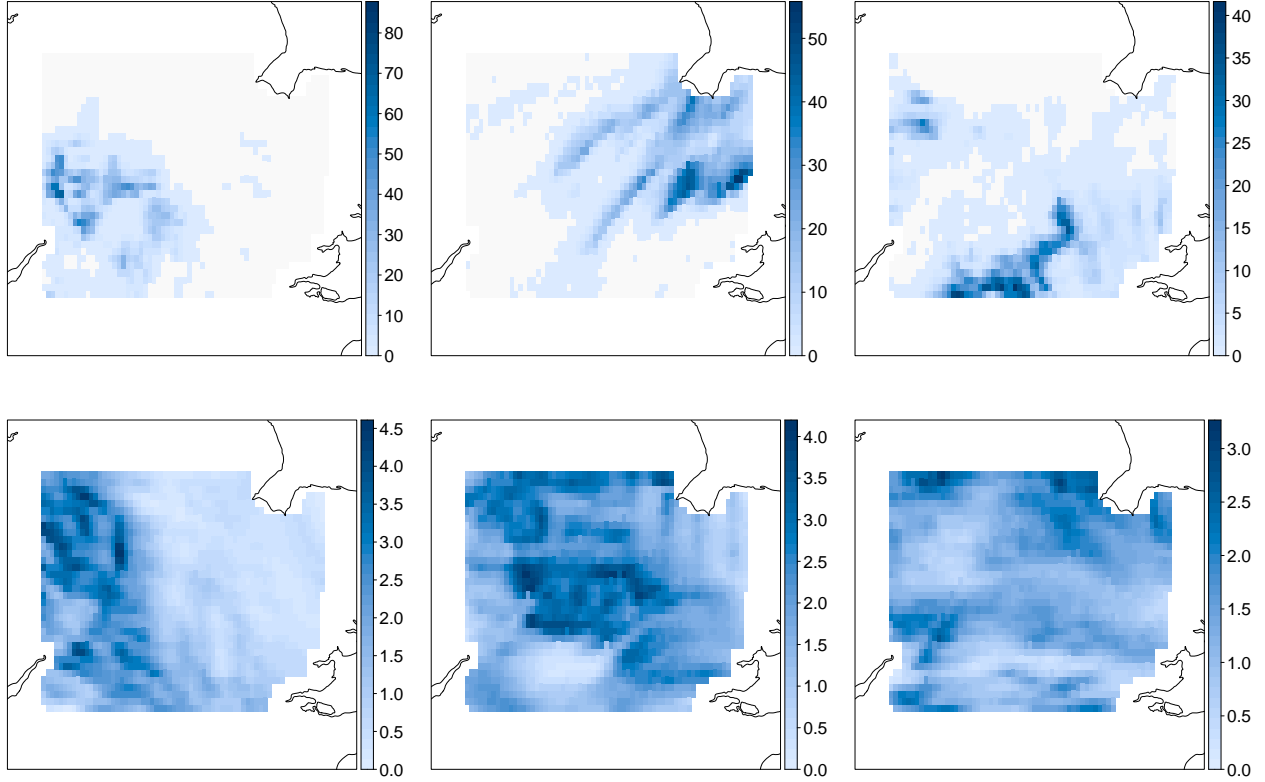
Figure 1: Observed extreme fields identified as convective (top) and non-convective (bottom) (mm/hr).

1.2 Data

We analyse average hourly precipitation rates (mm/hour) taken from the UKCP18 CPM projections (Kendon et al., 2019, 2021), 1981–2000. Richards et al. (2022a) consider the same data on the coarser resolution 5km × 5km grid, but data we use are on the finer resolution native climate model grid over a much larger spatial domain and the number of sampling locations increases from 934 to 7526. The sampling locations are 2.2×2.2 km^2 grid boxes and the spatial domain $\mathcal{S}$ of interest is the southern U.K., approximately centred at Northampton and covering an area of 198×231 km^2 with data only sampled over land. The study domain $\mathcal{S}$ is illustrated in Figure 2.

Each observation corresponds to the average precipitation over a grid-box. As extreme hourly precipitation is more likely to occur in summer, we use June–August observations only, yielding 43200 fields and eliminating the need to include seasonality effects; Figure 1 illustrates six such fields with the top, and bottom, rows of panels displaying events that produce extreme spatial aggregates over
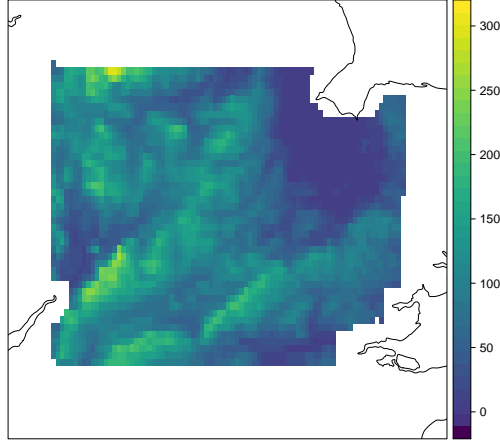
Figure 2: A map of elevation (m) for the spatial domain $\mathcal{S}$ of interest.

small, and large, scales respectively. We follow Richards et al. (2022a) and treat the centre of each grid box as a sampling location, and use the great-circle distance metric. We also set all values of the data less than 1×10^{-5}mm/hour to zero, since this level of precipitation would be recorded as zero by a rain gauge.

1.3   Overview

Richards et al. (2022a) develop methodology for joint inference on precipitation extremes at different spatial scales; they consider a spatial process $\{Y(s) : s \in \mathcal{S}\}$ for some spatial domain $\mathcal{S} \subset \mathbb{R}^2$, and let data used for modelling be observations $\mathbf{y}_t = (y_t(s_1), \ldots, y_t(s_d))^T$ at times $t \in \mathbf{T} = \{1, \ldots, n\}$ and sampling locations $\boldsymbol{S} = (s_1, \ldots, s_d) \subset \mathcal{S}$. Specifically, they perform inference on the upper-tail of the aggregate variable

$$\bar{R}_{\mathcal{A}} = \frac{1}{|\mathcal{A}|} \int_{\mathcal{A}} Y(s)\mathrm{d}s, \tag{1}$$

where $\mathcal{A} \subset \mathcal{S}$ is a sub-region of interest. To achieve this, they propose a model for the extremal behaviour of $\{Y(s) : s \in \mathcal{S}\}$ by adapting the spatial conditional extremes approach, first proposed by Wadsworth and Tawn (2022) and with extensions by Tawn et al. (2018); Shooter et al. (2019, 2021a,b); Huser and Wadsworth (2020); Simpson and Wadsworth (2021) and Simpson et al. (2022). Realisations taken from this model are used to conduct inference on $\bar{R}_{\mathcal{A}}$ for numerous regions $\mathcal{A} \subset \mathcal{S}$ simultaneously. Whilst the approach of Richards et al. (2022a) improved on previous approaches by ensuring self-consistent inference across different sub-regions of interest and flexible extremal dependence modelling,

they make the restrictive assumption that both the marginal and dependence behaviour of the data generating process is constant over time which, as we illustrate in Section 2, is a poor assumption for precipitation; we adapt their approach and propose separate models for convective and non-convective precipitation, and illustrate throughout the paper that this distinction is necessary due to the disparity between both the marginal and dependence behaviour of extreme events arising from the two classes of process. Alongside new parametric forms for the dependence parameters of the Wadsworth and Tawn (2022) model, further extensions of the approach of Richards et al. (2022a) include an approach for simulating events from processes which suffer particularly badly from boundary effects due to their long-range dependence, e.g., non-convective events.

Although we employ a conditional approach to model extremal dependence within the process $Y(s)$, other models have been applied in this context, e.g., max-stable processes (Coles, 1993; Padoan et al., 2010; Westra and Sisson, 2011; Reich and Shaby, 2012), Pareto processes (Palacios-Rodríguez et al., 2020; de Fondeville and Davison, 2022) and censored Gaussian copulas (Sang and Gelfand, 2010; Thibaud et al., 2013). These models make restrictive assumptions about extremal dependence which are not satisfied by our data, described in Section 2.4; moreover, Richards et al. (2022a) illustrate that modelling $Y(s)$ with the same extremal dependence as a max-stable or Pareto process leads to severe bias in estimates of return-levels for $\bar{R}_{\mathcal{A}}$.

The structure of the paper is as follows. Section 2 describes an algorithm for classifying observed fields into convective or non-convective precipitation events, and gives an overview of our mixture model setup. Details for modelling the extremes of mixture components are provided in Section 3, and the inference framework is described in Section 3.3.3. We simulate from the two fitted models for both classes of data using a procedure described in Section 4, and then combine realisations from both models to estimate the upper-tail behaviour of $\bar{R}_{\mathcal{A}}$. To illustrate the improvements to modelling $\bar{R}_{\mathcal{A}}$ that our approach provides to that detailed by Richards et al. (2022a), we apply both methods to the precipitation data described in Section 1.2.

6

2 Mixture process

2.1 Mixture component classification

We classify observation times $t \in \mathbf{T}$ for the entire observed fields $\{y_t(s) : s \in \mathcal{S}, t = 1, \ldots, n\}$ into two sets, denoted $\mathcal{C}$ and $\mathcal{N}$, which correspond to "convective" and "non-convective" times, respectively; that is, if $t \in \mathcal{C}$, then the observed field $\{y_t(s) : s \in \mathcal{S}\}$ is a convective event, and similarly for non-convective events. For the classification, we use Algorithm 1, developed at the Met Office Hadley Centre, U.K., for summer precipitation data with gridded sampling locations (Kendon et al., 2012). At time t, the algorithm makes its classification based on the gradient of the surface of $y_t(s)$ in a neighbourhood around s_i for all sites $s_i \in \mathbf{S}$, as convective rainfall are highly localised and so large gradients are expected if convection is present. Gradients are calculated using the difference between precipitation rates at adjacent grid-boxes in a $n_g \times n_g$ grid centred at s_i, denoted $\mathcal{L}_i$, as the data are observed on a regular grid. If it is not possible to create $\mathcal{L}_i$, e.g., at the edges of $\mathcal{S}$, then s_i is removed. Any field not identified as convective is automatically classified as non-convective; this includes any fields with no rain at any $s \in \mathcal{S}$.

Algorithm 1 Identify convective fields

[**Hyperparameters**]: gradient thresholds $g_l > 0, g_u > g_l$, proportion of large rainfall $p^* \in [0, 1]$, neighbourhood size $n_g \in \{2d^* - 1, d^* \in \mathbb{N}\}$
[**Data**]: Observations $\{y_t(s) : s \in \mathbf{S}, t \in \mathbf{T}\}$
For all $t = 1, \ldots n$:
    1. For all $i = 1, \ldots, d$:
        (a) Identify the $n_g \times n_g$ local neighbourhood $\mathcal{L}_i$ for s_i, defined in Section 2.
        (b) Evaluate all differences $\mathcal{G}_i = \{y_t(s_j) - y_t(s_k) : s_j, s_k \in \mathcal{L}_i\}$.
        (c) Calculate the proportion $p_{g,i} = |\{g \in \mathcal{G}_i : g \geq g_u\}|/|\{g \in \mathcal{G}_i : g \geq g_l\}|$.
        (d) If $p_{g,i} \geq p^*$, then label $y_t(s_i)$ as convective.
    2. If any elements of $\{y_t(s) : s \in \mathbf{S}\}$ are labelled as convective, then the entire field is classified as convective, i.e., $t \in \mathcal{C}$. Otherwise, $t \in \mathcal{N}$.

Algorithm 1 involves four hyperparameters, which we take to be $g_l = 0.01, g_u = 1$ and $p^* = 0.2$ and $n_g = 9$; these were tuned by the Met Office through visual inspection, and specifically for the climate model we consider (Roberts, N. and Kendon, E., 2020, personal communication). The most critical hyperparameters are g_u and p^*; if g_u is set too low or p^* too high, then stratiform, or frontal, precipitation may be classified as convective; the lower threshold g_l simply removes the effect of any grid-cells with very little rainfall, i.e., drizzle, and n_g determines the size of the region around site s_i for

which the surface of $y_t(s)$ is considered. With these hyperparameters, the algorithm gives $|\mathcal{C}| = 17724$ convective, and $|\mathcal{N}| = 25976$ non-convective, fields of our data, with examples of both classes presented in Figure 1; these selected fields were randomly sampled from fields which gave site-wise maxima for their respective processes. Note that the 1481 fully-dry fields, where there is zero rainfall at all sites, are classified as non-convective. Fields identified as being non-convective appear smoother over space and with much lower marginal magnitude than the convective fields; note the difference in the scales of Figure 1.

2.2   Mixture model

We assume that, for each $t \in \mathbf{T}$, the entirety of the process $\{Y_t(s) : s \in \mathcal{S}\}$ can be described by one of two mixture components, denoted by $\{Y_{\mathcal{C},t}(s) : t \in \mathcal{C}\}$ and $\{Y_{\mathcal{N},t}(s) : t \in \mathcal{N}\}$, that describe convective and non-convective rainfall, respectively. If $t \in \mathcal{C}$, then Y_t can be described solely by $Y_{\mathcal{C},t}$, rather than a mixture of both components, and similarly for $t \in \mathcal{N}$ and $Y_{\mathcal{N},t}$. Differences in the marginal and dependence behaviour of the two mixture components can be observed in Figures 1 and 3, respectively, and so we propose different marginal and dependence models for $Y_{\mathcal{C},t}$ and $Y_{\mathcal{N},t}$. We assume that, for each $\mathcal{C}$ and $\mathcal{N}$ process, both the marginal behaviour and dependence structures are stationary with respect to time; that is, the process $\{Y_{\mathcal{C},k}(s)\}$ is identically distributed to $\{Y_{\mathcal{C},l}(s)\}$ for all $l, k \in \mathcal{C}$, and similarly for the non-convective process $Y_{\mathcal{N},t}$. We then define the mixture process $\{Y_{\mathcal{M},t}(s)\}$ by

$$\{Y_{\mathcal{M},t}(s) : s \in \mathcal{S}\} = \begin{cases} \{Y_{\mathcal{C},t}(s) : s \in \mathcal{S}\}, & \text{with probability } p_{\mathcal{C}}, \\ \{Y_{\mathcal{N},t}(s) : s \in \mathcal{S}\}, & \text{with probability } 1 - p_{\mathcal{C}}, \end{cases} \tag{2}$$

where $p_{\mathcal{C}} \in [0, 1]$. We estimate $p_{\mathcal{C}}$ empirically using $|\mathcal{C}|/(|\mathcal{C}|+|\mathcal{N}|)$ and Algorithm 1 gives $\hat{p}_{\mathcal{C}} \approx 39.9\%$. As our interest lies only in spatial aggregates, we do not model temporal dependence in either component of $Y_{\mathcal{M},t}$ but account for this dependence through use of a block bootstrap in Section 5; henceforth we drop the time index from notation and use the shorthand $Y_{\mathcal{M}} := \{Y_{\mathcal{M}}(s) : s \in \mathcal{S}\}$ (and similarly for $Y_{\mathcal{C}}$ and $Y_{\mathcal{N}}$).

We compare two modelling approaches for $\{Y(s)\}$: an approach using the existing model developed by Richards et al. (2022a), which we denote throughout by $\{Y_{\mathcal{E}}(s)\}$ or the shorthand $Y_{\mathcal{E}}$, which

ignores any mixture structure in Y; and our improved approach which uses model (2). Note that $Y_{\mathcal{E}}$ and $Y_{\mathcal{M}}$ are equivalent when $p_{\mathcal{C}} = 0$ or $p_{\mathcal{C}} = 1$, or if $Y_{\mathcal{C}}(s) = Y_{\mathcal{N}}(s)$ for all $s \in \mathcal{S}$. Although observations of the process $Y(s)$ are assumed to arise from some probabilistic mixture, we are able to pre-classify observations from each mixture component; in this regard, our approach differs from standard methods using mixture models. Whilst there has been some recent focus on incorporating mixture structures into multivariate extremal dependence models (Simpson et al., 2020; Tendijck et al., 2021), very few methods has been proposed for mixture modelling of spatial extremes; the available models are typically restricted to a single extremal dependence class (Hazra and Huser, 2021), which, as we illustrate in Section 5.2, make them inappropriate for our data. Although model (2) is mathematically simple, from an applied perspective it represents an important and novel step towards flexible mixture modelling of spatial extremes. In Section 5.3, we illustrate that the inclusion of the "rudimentary" mixture structure in eq. (2) leads to a better model fit and improved inference on the extremes of $\bar{R}_{\mathcal{A}}$.

2.3 Model justification

Model (2) introduces separate mixture components that describe both convective and non-convective events over the entire spatial domain $\mathcal{S}$. The assumption that an entire spatial field, i.e., any instantaneous precipitation event observed over $\mathcal{S}$, can be neatly classified into distinct categories is unlikely to be valid in general, particularly as $|\mathcal{S}|$ increases; for very large $|\mathcal{S}|$, we may observe convective and non-convective events occurring simultaneously within $\mathcal{S}$ or even multiple independent convective events. We aim to counter these concerns by focusing on only events leading to extreme precipitation aggregates and by restricting the scale of the domain $\mathcal{S}$ to that of river catchments, which in the UK has an area of less than $11,000 \text{km}^2$. In Section 2.4, we highlight differences in the empirical extremal dependence structure of the two classes of data determined using Algorithm 1.

For an observed field, Algorithm 1 first identifies individual grid-boxes, i.e., sites, as convective (Step 1d of Algorithm 1); it then classifies the entirety of the field as convective (Step 2 of Algorithm 1), if at least one grid-cell is convective. Theoretically, this can lead to situations where entire fields are classified as convective courtesy of a single site (although we did not find this in practice). We investigate the proportion of convective grid-cells within a field to determine if the assumption of model (2), that a field at time t is either wholly convective or non-convective but never a mixture of both, is appropriate for the data. As our interest lies in the extremes of $\bar{R}_{\mathcal{A}}$, we calculate the

proportion of large values of empirical spatial aggregates (of different sizes) that are contributed to by convective sites, whenever any convective rainfall is present in the field; this is achieved via the heuristic described in Algorithm 2. If the estimated proportion is significantly less than one, then this would be an indication of there being mixing of the two rainfall types within extreme spatial aggregates and evidence to suggest that the model (2) cannot be reasonably assumed for the data.

Algorithm 2 Empirical proportion of convective rainfall leading to extreme spatial aggregates

---

[**Hyperparameters**]: radius $r > 0$ [$r \in \{30, 85\}$km], number of samples [$M = 50$]
[**Data**]: All replications with $t \in \mathcal{C}$ temporal replications
  1. For all $t \in \mathcal{C}$:
      (a) Repeat for M samples:
          i. Sample a spatial location uniformly over the domain $s_c \in \mathcal{S}$.
          ii. Compute sample total of hourly rainfall rate over all sites within a radius r.
          iii. Compute sample total of hourly rainfall rate over all sites identified as convective (using Step 1d of Algorithm 1) within a radius r.
  2. Keep only values where samples from Step 1(a)ii exceed the empirical 90% quantile over all corresponding $M \times \mathcal{C}$ samples.
  3. Compute proportions of extreme total hourly rainfall contributed to by convective sites, i.e., divide samples in Step 1(a)iii by those from Step 1(a)ii.

Ideally, for Algorithm 2 we would create samples of extreme aggregates over a specific aggregate region; however, this would leave very few samples available for estimation of the required proportion. We instead make an assumption of stationarity, i.e., that the distribution of the proportion of interest is constant over space and time; this allows us to pool information across sites and observations. Then to create a sample of spatial aggregates which is independent of the position of the aggregate region, we randomly sample a centre and evaluate the spatial average over a circular region with radius $r > 0$ and repeat this process M times for each convective field. We then consider only the largest values within the sample of the aggregates, i.e., those that exceed the sample 90% quantile, and compute the proportion of their total value contributed to by observed values at convective grid-cells. We chose $M = 50$, large enough that reasonably-sized samples were created, with larger values not providing a significant change in the results. We found strong evidence that, for a range of radii ($r = 30$km, $r = 85$km and the whole study domain), the majority of extreme aggregates are driven by mostly convective rainfall, with the median estimate of the proportion for small and large aggregate regions being 0.92, 0.81 and 0.78, respectively. Hence we proceed under the assumption that model (2) is appropriate for the data.

10

2.4 Exploratory analysis for mixture components

We expect the two mixture components defined in eq. (2) to exhibit different extremal dependence structures, as they describe very different rainfall processes: convective and non-convective rainfall. We quantify extremal dependence for a process $\{Y(s)\}$ using the upper tail index $\chi(s_A, s_B) :=$ $\lim_{q\uparrow 1} \chi_q(s_A, s_B)$ (Joe, 1997) for all $s_A, s_B \in \mathcal{S}$, defined by

$$\chi_q(s_A, s_B) = \Pr\left\{Y(s_B) > F^{-1}_{Y(s_B)}(q) \mid Y(s_A) > F^{-1}_{Y(s_A)}(q)\right\}, \tag{3}$$

where $F^{-1}_{Y(s_A)}$ denotes the quantile function of $Y(s_A)$. The measure $\chi(s_A, s_B)$ determines the extremal dependence class; if $\chi(s_A, s_B) > 0$ or $\chi(s_A, s_B) = 0$, then $Y(s_A)$ and $Y(s_B)$ are asymptotically dependent or asymptotically independent, respectively (Coles, 2001). Richards et al. (2022a) found evidence that convective hourly rainfall exhibits asymptotic independence at the smallest observable spatial distances, i.e., $\chi(s_A, s_B) = 0$ for sites $s_A, s_B \in \mathcal{S}$ separated by at least 2.2km, inferring that extreme events of $Y_{\mathcal{C}}$ become increasingly more localised as the magnitude of the events gets larger. However, we find that non-convective events exhibit asymptotic dependence at short-range and asymptotic independence at long-range. To illustrate the need for different extremal dependence structures for both classes of rainfall, we compute empirical estimates $\hat{\chi}_q(s_A, s_B)$ of eq. (3) for both classes of data; Figure 3 presents these for all pairs of a subset of 400 sites randomly sampled over $\mathcal{S}$ and with $q = 0.98$ and $q = 0.995$. We observe that non-convective rainfall (brown points) exhibits much stronger extremal dependence than convective rainfall (green points), as values of $\chi_q(s_A, s_B)$ decay much slower with distance for the brown points than the green points. We also observe that, as q increases, $\chi_q(s_A, s_B)$ appears to tend to zero at a much faster rate for convective rainfall; our model fits in Section 5.2 suggest that the assumption of short-range asymptotic dependence, i.e., $\chi(s_A, s_B) > 0$ for close s_A and s_B, may be appropriate only for non-convective rainfall. We adapt the approach of Richards et al. (2022a) and, in Section 3, propose slightly different parametric sub-models for the two processes to accommodate the differences in their extremal dependence structure.
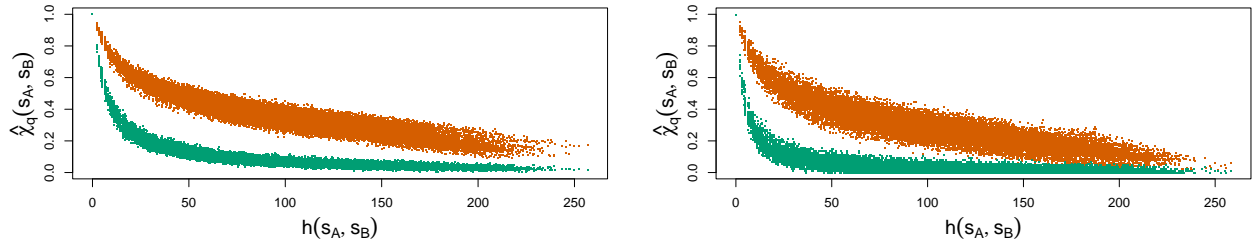
Figure 3: Empirical estimates of pairwise extremal dependence strength $\hat{\chi}_q(s_A, s_B)$ of eq. (3) for $q = 0.98$ (left) and $q = 0.995$ (right) $s_A, s_B \in \mathcal{S}$ against great-circle distances $h(s_A, s_B) := \|s_A - s_B\|$, which are given in km. The green and brown points correspond to estimates for convective and non-convective realisations, respectively.

3 Modelling extremes of each mixture component

3.1 Overview

Following Richards et al. (2022a), we adopt a two-step modelling approach. We first standardise the data using the marginal model described in Section 3.2; this model is presented for a generic stationary process $\{Y(s) : s \in \mathcal{S}\}$, but is fitted separately for each of the processes $Y_\mathcal{C}, Y_\mathcal{N}$ and $Y_\mathcal{E}$. The marginal fits are used to perform site-wise standardisation of data to standard Laplace margins; we denote these standardised processes by $X_\mathcal{C}$, $X_\mathcal{N}$ and $X_\mathcal{E}$.

To characterise extremal dependence, we use the spatial conditional extremes framework (Wadsworth and Tawn, 2022); this is detailed for a generic standardised process $\{X(s) : s \in \mathcal{S}\}$ in Section 3.3.1, but is applied, separately, to each of $X_\mathcal{C}$, $X_\mathcal{N}$ and $X_\mathcal{E}$. Slight differences in modelling choices are made for the three processes, i.e., choices of parametric forms for some dependence functions and hyper-parameters for simulation; discussions of these differences are provided in Sections 3.3.2 and 4.2, respectively.

3.2 Marginal model

We model the site-wise marginals of $\{Y(s) : s \in \mathcal{S}\}$ using three components: generalised Pareto tails (Coles, 2001, p. 75) above some high threshold $q_\lambda(s)$, the $(1 - \lambda)-$quantile of $Y(s)$, the empirical distribution for the bulk and a discrete mass at the lower tail, denoted $p(s)$, which corresponds to the

probability that no rain occurs at site $s \in \mathcal{S}$. The distribution function of $Y(s)$, for all $s \in \mathcal{S}$, is

$$
F_{Y(s)}(y) = \begin{cases} p(s), & \text{if } y = 0, \\ \{1 - \lambda - p(s)\}\frac{F_{Y_+(s)}(y)}{F_{Y_+(s)}\{q_\lambda(s)\}} + p(s), & \text{if } 0 < y \le q_\lambda(s), \\ 1 - \lambda \left[1 + \frac{\xi\{y - q_\lambda(s)\}}{\upsilon(s)}\right]_+^{-1/\xi}, & \text{if } y > q_\lambda(s), \end{cases} \tag{4}
$$

where $[x]_+ = \max\{x, 0\}$, and where $\xi \in \mathbb{R}$, $\upsilon(s) > 0$ and $p(s) \ge 0$, $0 < \lambda < 1$ and $F_{Y_+(s)}(y)$ denotes the empirical distribution function of strictly positive values of $Y(s)$; we further require that $p(s) + \lambda < 1$, suggesting that our marginal model is unlikely to perform well if $p(s)$ is close to 1, as only a small number of observations would then be available to estimate the generalised Pareto distribution parameters.

The marginal parameter functions $p(s), q_\lambda(s)$ and $\upsilon(s)$ are represented through a basis of thin-plate splines. We estimate $p(s)$ using a logistic generalised additive model (Wood, 2006) and estimate $q_\lambda(s)$ for all $s \in \mathcal{S}$ using additive quantile regression (Fasiolo et al., 2021); the latter technique is computationally expensive and so we use only a subset of sites for estimation. Richards et al. (2022a) propose a technique for estimating $q_\lambda(s)$ by fitting a thin-plate spline through point-wise estimates of $q_\lambda(s)$ for each $s \in (s_1, \ldots, s_d)$; however, our approach better accounts for uncertainty in the spline estimates as more data are used in their estimation, i.e., all observations at a single site rather than a single quantile estimate. The scale parameter $\upsilon(s)$ is estimated by fitting a generalised Pareto, with parameters represented as additive functions, to the exceedances of $q_\lambda(s)$, i.e., $Y(s) - q_\lambda(s)$, with a fixed shape parameter ξ for all $s \in \mathcal{S}$ (Youngman, 2019). Our choice to fix ξ over space was fully supported by our data and is a common approach taken when modelling spatial characteristics of extreme rainfall as it reduces the risk of having parameter identifiability issues (Thibaud et al., 2013; Zheng et al., 2015; Saunders et al., 2017; Brown, 2018).

Exploratory analysis of the data reveals that elevation is an important covariate in the marginal behaviour of precipitation, an observation supported by Coles and Tawn (1996); Cooley et al. (2007) and Cooley and Sain (2010). Hence, we allow $p(s), q_\lambda(s)$ and $\log \upsilon(s)$ to vary with both location and elevation; separate spline bases are used for both. Each spline consists of as few basis functions as possible, i.e., four, to ensure that the parameter surfaces are relatively smooth; this is to make the

marginal fits physically interpretable and to reduce over-fitting. Our approach for inference differs from that proposed by Richards et al. (2022a) in that we allow the margins to change with elevation.

3.3 Extremal dependence model

3.3.1 Generic model

We model extremal dependence of the standardised process $\{X(s) : s \in \mathcal{S}\}$, given that it is extreme for some $s \in \mathcal{S}$, by conditioning on the process being above some high threshold u at a specified site $s_O \in \mathcal{S}$ for each $s_O \in \mathcal{S}$. The process $\{X(s)\}$ is assumed to be stationary with dependence a function of distance, i.e., $h(s_A, s_B) = \|s_A - s_B\|$ for $s_A, s_B \in \mathcal{S}$, and some distance metric $\|\cdot\|$. Following Wadsworth and Tawn (2022), we assume that there exists normalising functions $\alpha : [0, \infty) \mapsto [0, 1]$, with $\alpha(0) = 1$, and $\beta : [0, \infty) \mapsto [0, 1]$, such that for each $s_O \in \mathcal{S}$, as $u \to \infty$,

$$\left(\left\{ \frac{X(s) - \alpha\{h(s, s_O)\}X(s_O)}{\{X(s_O)\}^{\beta\{h(s, s_O)\}}} : s \in \mathcal{S} \right\}, X(s_O) - u \right) \,\Bigg|\, \left(X(s_O) > u \right) \xrightarrow{d} \left(\left\{ Z(s|s_O), s \in \mathcal{S} \right\}, E \right), \tag{5}$$

where E is a standard exponential variable, the process $\{Z(s|s_O)\}$ is non-degenerate for all $s \in \mathcal{S}$ where $s \neq s_O$ and it is independent of E. That is, there is convergence in distribution of the normalised process to the residual process $Z(s|s_O)$ which satisfies $Z(s_O|s_O) = 0$ almost surely.

We also follow Wadsworth and Tawn (2022) and take the parametric form for $\alpha(h)$ of

$$\alpha(h) = \begin{cases} 1, & h \leq \Delta, \\ \exp(-\{(h - \Delta)/\kappa_{\alpha_1}\}^{\kappa_{\alpha_2}}), & h > \Delta, \end{cases}$$

where $\Delta \geq 0$ and $\kappa_{\alpha_1}, \kappa_{\alpha_2} > 0$. This function determines the strength and class of extremal dependence within $\{X(s)\}$, see eq. (3); allowing asymptotic dependence up to distance Δ from s_O, and asymptotic independence thereafter, with strength of dependence decreasing with h. Setting $\Delta = \infty$ and $\beta(h) = 0$ for all $h \geq 0$ corresponds to X having the same dependence structure as an $r-$Pareto process conditional on exceedances at a single site (de Fondeville and Davison, 2022).

Richards et al. (2022a) impose that the residual process $Z(s|s_O)$ for $s \in \mathcal{S}$ has delta-Laplace margins with location, scale and shape parameters determined by the spatial functions $\mu : [0, \infty) \mapsto \mathbb{R}$, $\sigma : [0, \infty) \mapsto [0, \sqrt{2}]$ and $\delta : [0, \infty) \mapsto [1, \infty)$, respectively. We denote this by $Z(s|s_O) \sim$

14

$\mathrm{DL}(\mu\{h(s, s_O)\}, \sigma\{h(s, s_O)\}, \delta\{h(s, s_O)\})$, where the delta-Laplace density is

$$f(z) = \delta\{2k\sigma\Gamma(1/\delta)\}^{-1} \exp\{-|(z-\mu)/(k\sigma)|^\delta\} \tag{6}$$

for $z \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0$ and $\delta > 0$, and where Γ denotes the standard gamma function and $k^2 = \Gamma(1/\delta)/\Gamma(3/\delta)$. Dependence within $Z(s|s_O)$, for each $s_O \in \mathcal{S}$, is modelled using a standard Gaussian process with a stationary Matérn correlation function, denoted ρ (see Section B.4 of Richards et al. (2022b)), with $Z(s_O|s_O) = 0$ almost surely. We adopt the same parametric forms for β, μ, σ, and δ as Richards et al. (2022a), unless stated otherwise in Section 3.3.2; full details are provided in Section A of Richards et al. (2022b). These parametric forms allow the extremal dependence model for $\{X(s)\}$ to exhibit perfect independence at long-range as $\alpha(h) \to 0$, $\beta(h) \to 0$, $\mu(h) \to 0$, $\sigma(h) \to \sqrt{2}$ and $\delta(h) \to 1$ when $h \to \infty$.

The function $h(s, s_O)$ accounts for potential geometric anisotropy in the extremal dependence structure of $\{X(s)\}$. As Richards et al. (2022a), we define the distance metric $\|s_A - s_B\|$ as an elliptical transformation of $(s_A, s_B) \in \mathcal{S} \times \mathcal{S}$ before finding the great-circle, or spherical, distance between s_A and s_B (Richards et al., 2022b, eq. S.2); the metric is parametrised by $\theta \in [-\pi/2, 0]$ and $L > 0$ which control the rotation and coordinate stretching effect, respectively, with $L = 1$ being isotropy.

3.3.2 Differences for non-convective component

We apply the model described in Section 3.3.1 to each of $X_\mathcal{C}$, $X_\mathcal{N}$ and $X_\mathcal{E}$. For $X_\mathcal{C}$ and $X_\mathcal{E}$, we use exactly the same parametric forms for the dependence functions as proposed by Richards et al. (2022a), which are provided in Section A of Richards et al. (2022b); for $X_\mathcal{N}$, we adapt the β and σ functions, denoting these new forms $\beta_\mathcal{N}$ and $\sigma_\mathcal{N}$.

The β function advocated by Richards et al. (2022a) for $X_\mathcal{C}$, say $\beta_\mathcal{C}$, satisfies $\beta_\mathcal{C}(0) = 1$ and allows the convective process to exhibit spatial roughness dependent on the magnitude of $X(s_O)$. Whilst this is a desirable property for a process generating convective precipitation events, it is not required for $X_\mathcal{N}$, which generates much smoother spatial events than $X_\mathcal{C}$; this claim is supported by observations of $Y_\mathcal{C}$ and $Y_\mathcal{N}$ in Figure 1. For $\beta_\mathcal{N}$, we adopt the approach of Shooter et al. (2021b) and let

$$\beta_\mathcal{N}(h) = \frac{\kappa_{\beta_1} h^{\kappa_{\beta_2}} \exp(-h/\kappa_{\beta_3})}{\max_{h_* > 0}\{h_*^{\kappa_{\beta_2}} \exp(-h_*/\kappa_{\beta_3})\}}, \qquad \kappa_{\beta_1} \in [0, 1], \kappa_{\beta_2} > 0, \kappa_{\beta_3} > 0, \tag{7}$$

so $0 \leq \beta_{\mathcal{N}}(h) \leq 1$ for all $h \geq 0$ and $\beta_{\mathcal{N}}(h) \to 0$ as $h \to \infty$ with $\beta_{\mathcal{N}}(0) = 0$. In our application, we considered both forms of β for $X_{\mathcal{N}}$, but found that eq. (7) provided a better fitting model.

The scaling function $\sigma_{\mathcal{C}}$ gives long-range independence. However, non-convective events can have a very large spatial extent, e.g., up to 1000km (Houze Jr, 1997). Long-range independence is not required for $X_{\mathcal{N}}$ when the domain of interest $\mathcal{S}$ is small enough, as in our application in Section 5. To ensure a better fitting model for $X_{\mathcal{N}}$, we let the scale function $\sigma_{\mathcal{N}}$ be

$$\sigma_{\mathcal{N}}(h) = \kappa_{\sigma_3}\left[1 - \exp\{-(h/\kappa_{\sigma_1})^{\kappa_{\sigma_2}}\}\right], \qquad \kappa_{\sigma_1} > 0, \kappa_{\sigma_2} > 0, \kappa_{\sigma_3} > 0,$$

which is equivalent to $\sigma_{\mathcal{C}}$ when $\kappa_{\sigma_3} = \sqrt{2}$.

3.3.3 Inference

Inference for each of the three extremal dependence models is conducted using the pseudo-likelihood procedure proposed by Richards et al. (2022a), which they illustrate works well for models of this type and for application to precipitation data. Full details of the inference procedure are given in Richards et al. (2022b, Section B). This technique requires selection of a sub-sample of d_s triples of sampling locations, subject to the maximum pairwise distances subceeding a value of h_{max} km. Maximisation of a pseudo-likelihood is then conducted using observations of the process via a triple-wise censored likelihood; the latter is required to handle point masses in the marginals of $X(s)$ caused by zeroes in the marginals of $Y(s)$. The censoring threshold for $X(s)$, denoted $c(s)$, is found by transforming $p(s)$ to the Laplace scale, i.e., $c(s) = F_L^{-1}\{p(s)\}$, where F_L denotes the standard Laplace distribution. We apply this approach and use different values of h_{max} and d_s for each of the three processes; further details are provided in Section 5.2. Inference is conducted for model (5) by fixing a high threshold $u > 0$ and assuming that the limiting relation of eq. (5) holds exactly; the value of u differs between the three processes. Finding a threshold such that $Z(s|s_O)$ and E, defined in eq. (5), are independent may be infeasible with mixtures of different dependence present in $X(s)$, e.g., for $X_{\mathcal{E}}(s)$. This provides further support for our use of separate extremal dependence models for the mixture components, as the assumptions made in eq. (5) are more likely to hold when applied to $X_{\mathcal{N}}$ and $X_{\mathcal{C}}$, rather than $X_{\mathcal{E}}$, and at lower thresholds u (and thus allows us to use more data for inference); we find the latter to be the case in our analysis.

4 Simulating events

4.1 Efficient simulation

We now consider simulation of $\{Y(s) : s \in \boldsymbol{S}\}$ on a dense discrete set of sites, denoted $\boldsymbol{S} \subseteq \mathcal{S}$; we further distinguish between the continuous aggregate region $\mathcal{A}$ and a discretized version $\boldsymbol{A} \subseteq \mathcal{A}$, which must satisfy $\boldsymbol{A} \subset \boldsymbol{S}$. Typically, we would take $\boldsymbol{S}$ to be the set of sampling locations, such that $\boldsymbol{S} := \mathbf{s}$, and we do so hereafter. When $|\boldsymbol{S}|$ is large, e.g., $|\boldsymbol{S}| > 5,000$ as with our data, it may be too computationally expensive to simulate a large number of replicates of $\{Y(s) : s \in \boldsymbol{S}\}$ in order to do reliable inference on the upper-tail behaviour of $\bar{R}_\mathcal{A}$. It may not be entirely necessary to simulate Y over all of $\boldsymbol{S}$ if our interest lies in aggregates over sufficiently small regions; we can instead simulate fields of Y on a sub-region $\boldsymbol{S}_\tau \subset \boldsymbol{S}$; we define this set for $\tau > 0$ by

$$
\boldsymbol{S}_\tau = \left\{ s \in \boldsymbol{S} : \min_{s_{\boldsymbol{A}} \in \boldsymbol{A}} \{\|s - s_{\boldsymbol{A}}\|\} \leq \tau \right\}, \tag{8}
$$

i.e., so $\boldsymbol{A} \subset \boldsymbol{S}_\tau$. Richards et al. (2022a) chose τ so that there was a predetermined small probability of observing a large event at any $s \in \boldsymbol{A}$ given that there is an extreme event at any single site outside of $\boldsymbol{S}_\tau$. We also introduce a set of conditioning sites $\boldsymbol{A}_c = \boldsymbol{S}_\tau \cup \boldsymbol{S}^+$, where $\boldsymbol{S}^+ \cap \boldsymbol{S} = \emptyset$ for a set $\boldsymbol{S}^+$ to be defined, and create realisations of $\{Y(s) : s \in \boldsymbol{S}_\tau\}$ given that the process is extreme at $s_O \in \boldsymbol{A}_c$. Illustrations of $\boldsymbol{A}$, $\boldsymbol{S}$, $\boldsymbol{S}_\tau$, $\boldsymbol{S}^+$ and $\boldsymbol{A}_c$ for our application are presented in Figure 4 and a heuristic for choosing both $\boldsymbol{S}_\tau$ and $\boldsymbol{A}_c$ is given in Section 4.2; these sets vary over the three processes.

When considering aggregates over a fixed region $\boldsymbol{A} \subset \boldsymbol{S}_\tau$, the simulation procedure proposed by Richards et al. (2022a) never generates events for which the conditioning site, s_O in $\boldsymbol{A}_c$, lies outside of the boundaries of $\boldsymbol{S}_\tau$. When the process exhibits particularly strong extremal spatial dependence, such as for non-convective rainfall, we may bias the upper-quantiles of the distribution of $\bar{R}_\mathcal{A}$ if we set $\boldsymbol{A}_c = \boldsymbol{S}_\tau$, i.e., $\boldsymbol{S}^+ = \emptyset$, since we are ignoring events which could be large somewhere outside of $\boldsymbol{S}_\tau$ but still impact the upper-tails of the aggregate over $\boldsymbol{A}$. So we need to account for the possibility of an extreme at conditioning sites $s_O \notin \boldsymbol{S}$ that are far from $\boldsymbol{A}$ to avoid this problem. When we require that $\boldsymbol{S}^+ \neq \emptyset$ we construct $\boldsymbol{S}^+$ to ensure that $\boldsymbol{A}_c$ contains additional, randomly selected, synthetic conditioning sites located outside of $\boldsymbol{S}$.

To simulate from the fitted models for $Y_{\mathcal{C}}$, $Y_{\mathcal{N}}$ and $Y_{\mathcal{E}}$, we need to extend the procedure proposed by Richards et al. (2022a) due to long-range dependence features of frontal rain; we present this technique for a generic Y and X, but this approach can be applied to any of the three processes considered. To simulate a realisation of $Y(s)$ at a site $s \in \boldsymbol{S}_\tau$, we draw a realisation of

$$Y(s) \,\Big|\, \left\{ \max_{s \in \boldsymbol{S}} \left(F_L^{-1}[F_{Y(s)}\{Y(s)\}] \right) > v \right\} \equiv F_{Y(s)}^{-1}[F_L\{X(s)\}] \,\Big|\, \left[\max_{s \in \boldsymbol{S}} \{X(s)\} > v \right], \tag{9}$$

using models (4) and (5) and for $v \geq u$, with probability

$$\Pr\left\{ \max_{s \in \boldsymbol{S}} \left(F_L^{-1}[F_{Y(s)}\{Y(s)\}] \right) > v \right\},$$

and otherwise draw a realisation from the observed $Y(s)$ subject to $\max_{s \in \boldsymbol{S}} \left(F_L^{-1}[F_{Y(s)}\{Y(s)\}] \right) < v$. To draw $b \in \mathbb{N}$ realisations of process (9), we use Algorithm 3, which adjusts Algorithm 1 of Richards et al. (2022a) to account for the introduction of $\boldsymbol{A}_c$ and $\boldsymbol{S}_\tau$. We use importance sampling to approximately draw realistions from the distribution conditional on $\max_{s \in \boldsymbol{S}} \{X(s)\}$ being large; we first simulate a population of b' replicates and then sub-sample $b < b'$ replicates using an importance sampling scheme.

Algorithm 3 Simulating conditional spatial fields corresponding to formulation (9)

1. For $i = 1, \ldots, b'$ with $b' > b$ denoting the replicate population size:
   (a) Draw a conditioning location $s_O^{(i)}$ from $\boldsymbol{A}_c$ with equal probability $1/|\boldsymbol{A}_c|$.
   (b) Simulate $E^{(i)} \sim \mathrm{Exp}(1)$ and set $x_i(s_O^{(i)}) = v + E^{(i)}$.
   (c) Simulate a field $\{z_i(s|s_O^{(i)}) : s \in \boldsymbol{S}_\tau\}$ from the residual process model.
   (d) Set $x_i(s) = x_i(s_O^{(i)})\alpha\{h(s, s_O^{(i)})\} + \{x_i(s_O^{(i)})\}^{\beta\{h(s,s_O^{(i)})\}}\{z_i(s|s_O^{(i)})\}$ for each $s \in \boldsymbol{S}_\tau$.
2. Assign each simulated field $\{x_i(s) : s \in \boldsymbol{S}_\tau\}$ an importance weight of

$$\begin{cases} \left[\sum_{s \in \boldsymbol{S}_\tau} \mathbb{1}\{x_i(s) > v\}\right]^{-1}, & \text{if } \sum_{s \in \boldsymbol{S}_\tau} \mathbb{1}\{x_i(s) > v\} > 0, \\ 0, & \text{otherwise}, \end{cases}$$

   for $i = 1, \ldots, b'$.
3. Sub-sample b realisations from the collection of replicates with probabilities proportional to their assigned importance weights.
4. For all $s \in \boldsymbol{S}_\tau$, transform $x_i(s)$ to $y_i(s)$ using the marginal transformation in eq. (4). Note that if $x_i(s) \leq c(s)$, we set $y_i(s) = 0$; however, $y_i(s')$ is above its $F_L(v)$-th quantile for some $s' \in \boldsymbol{S}_\tau$.

Using Algorithm 3, we can draw realisations of $\{Y_{\mathcal{C}}(s) : s \in \boldsymbol{S}_\tau\}$ and $\{Y_{\mathcal{N}}(s) : s \in \boldsymbol{S}_\tau\}$ and then use these to derive samples of $\{Y_{\mathcal{M}}(s) : s \in \boldsymbol{S}_\tau\}$ by drawing from the samples of $\{Y_{\mathcal{C}}(s) : s \in \boldsymbol{S}_\tau\}$

and $\{Y_{\mathcal{N}}(s) : s \in \boldsymbol{S}_{\tau}\}$ with probabilities $p_{\mathcal{C}}$ and $1 - p_{\mathcal{C}}$, respectively, see eq. (2), which are estimated empirically. Approximate realistions of $\bar{R}_{\mathcal{A}}$ are derived by taking replicates of Y and computing the average of $\{Y(s) : s \in \boldsymbol{A}\}$; we denote estimates of $\bar{R}_{\mathcal{A}}$ derived using $Y_{\mathcal{M}}$ and $Y_{\mathcal{E}}$ by $\bar{R}_{\mathcal{M},\mathcal{A}}$ and $\bar{R}_{\mathcal{E},\mathcal{A}}$, respectively. Although the approach we have described is tailored to a specific aggregate region $\boldsymbol{A} \subset \boldsymbol{S}_{\tau}$, our interest may lie in aggregates over multiple regions $\boldsymbol{A}_1, \dots, \boldsymbol{A}_m$ for $m \in \mathbb{N}$. To ensure that we can perform joint inference on aggregates at multiple scales, we can simulate replicates of Y with $\boldsymbol{A} \supseteq \cup_{i=1}^{m} \boldsymbol{A}_i$; replicates of aggregates of Y over the sub-regions $\boldsymbol{A}_1, \dots, \boldsymbol{A}_m$ are then derived from a common collection of replicates.

4.2   Choosing $\boldsymbol{S}_{\tau}$ and $\boldsymbol{A}_c$

We begin by considering the dense sub-region $\boldsymbol{S}_{\tau}$ of locations at which realisations of $Y(s)$ are to be simulated, ensuring that $\boldsymbol{A} \subset \boldsymbol{S}_{\tau} \subseteq \boldsymbol{S}$, and then subsequently choose $\boldsymbol{S}^+$ to build $\boldsymbol{A}_c \supseteq \boldsymbol{S}_{\tau}$; recall that $\boldsymbol{A}_c = \boldsymbol{S}_{\tau} \cup \boldsymbol{S}^+$ is the set of all possible conditioning sites used in simulation. Note that, in both cases, $\boldsymbol{S}_{\tau}$ and $\boldsymbol{A}_c$ should be chosen as large as is computationally feasible. For Step 2 of Algorithm 3, we require that

$$\sum_{s \in \boldsymbol{S}_{\tau}} \mathbb{1}\{x_i(s) > v\} \approx \sum_{s \in \boldsymbol{S}} \mathbb{1}\{x_i(s) > v\}, \tag{10}$$

to account for the conditioning event in eq. (9); we can improve the accuracy of this approximation by ensuring that sites in $\boldsymbol{S}_{\tau}$ are sufficiently spread out across $\boldsymbol{S}$. We first set $\boldsymbol{S}_{\tau}$ according to eq. (8), as we expect extreme events within $\boldsymbol{A}$ to have the largest impact on the tail behaviour of $\bar{R}_{\mathcal{A}}$. If this choice of $\boldsymbol{S}_{\tau}$ does not give a good approximation in eq. (10), then we additionally sample n_s sites uniformly at random across $\boldsymbol{S} \setminus \boldsymbol{S}_{\tau}$ and add these to $\boldsymbol{S}_{\tau}$.

In our application, we found that setting $\boldsymbol{S}^+ := \emptyset \Rightarrow \boldsymbol{A}_c := \boldsymbol{S}_{\tau}$ was a reasonable choice to make for $Y_{\mathcal{C}}$ and $Y_{\mathcal{E}}$, as no improvement in our inference for $\bar{R}_{\mathcal{A}}$ was achieved by using $\boldsymbol{A}_c \supset \boldsymbol{S}_{\tau}$. However, this was not true for $Y_{\mathcal{N}}$, which exhibits strong extremal dependence at even the greatest distances within $\boldsymbol{S}_{\tau}$. For such processes, we need to artificially increase the maximum possible distance between sites in $\boldsymbol{A}_c$ and $\boldsymbol{A}$, to ensure that any possible conditioning site s_O that could contribute to an extreme event of $\bar{R}_{\mathcal{A}}$ is available in $\boldsymbol{A}_c$. To this end, we construct $\boldsymbol{A}_c$ by taking $\boldsymbol{S}^+$ to be a set of n_c sites, sufficiently far from $\boldsymbol{A}$, where data are not observed. We create $\boldsymbol{S}^+$ via a brute-force approach; we add independent Gaussian noise to the coordinates of a site in the centre of $\boldsymbol{A}$, rejecting any new points
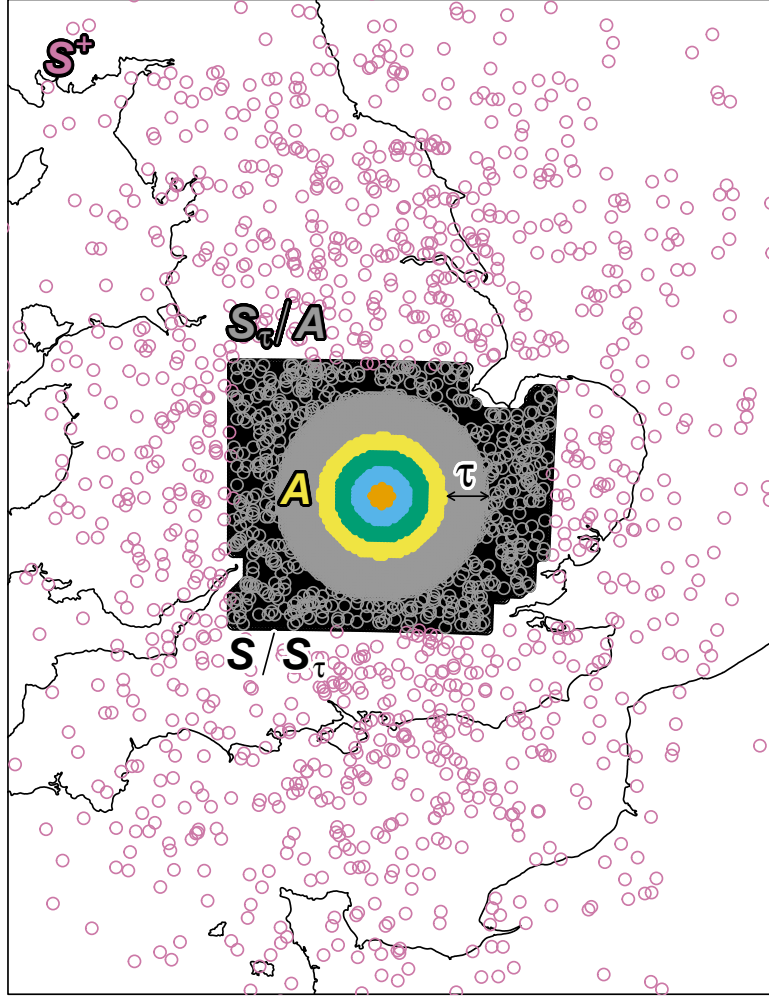
Figure 4: Regions $\boldsymbol{A}$ and $\boldsymbol{S}$ that correspond to discretized versions of the aggregate region $\mathcal{A}$ and study domain $\mathcal{S}$, respectively; regions $\boldsymbol{S}_\tau$, $\boldsymbol{A}_c$ and $\boldsymbol{S}^+ = \boldsymbol{A}_c \setminus \boldsymbol{S}$ are used in Algorithm 3 for simulating model realisations. The aggregate region $\boldsymbol{A}$ encompasses four sub-regions, with corresponding areas $(179, 1263, 3257, 6200)$ km^2, that are coloured orange, blue, green, yellow; sub-regions include both the coloured and interior points and are numbered 1 to 4 in Figures 8 and S9. The grey and black points denote $\boldsymbol{S}_\tau \setminus \boldsymbol{A}$ and $\boldsymbol{S} \setminus \boldsymbol{S}_\tau$, respectively. The $n_c = 1250$ purple points, outside of the boundaries of $\boldsymbol{S}$, correspond to $\boldsymbol{S}^+$, and the region $\boldsymbol{S}_\tau$ encompasses all points within distance $\tau = 27.5$km outside of $\boldsymbol{A}$ and $n_s = 500$ additional randomly-sampled points within $\boldsymbol{S}$; these values mean that $|\boldsymbol{S}_\tau| = 3625$ and $|\boldsymbol{A}_c| = 4875$.

that lie within the boundaries of $\mathcal{S}$, see Figure 4.

Suitable values for the constants n_s, τ and n_c, and the variance of the noise term used to create $\boldsymbol{S}^+$, can be chosen through validation techniques, such as model fit diagnostics for the aggregate variable (Figures 8 and S9); if the fits look poor, one can repeatedly increase the value of these constants until no significant changes in fit are observed.

5   Application

5.1   Marginal analysis

Marginal analysis is conducted by fitting the model described in Section 3.2 to each of the three datasets, i.e., convective, non-convective, and pooled. We set $\lambda = 0.005$ in eq. (4) for all three processes and use a subset of 500 sites sampled randomly over $\mathcal{S}$ to estimate $q_\lambda(s)$; more sites were considered, but no significant difference was observed in the inference. The same sites are used for all processes and are illustrated in Figure S1 (Richards et al., 2022b); we ensure that sampled sites proportionately cover the distribution of elevation values by sampling sites with probability according to the approximate empirical distribution of elevation measurements. We use all sampling locations to estimate $p(s)$, $\upsilon(s)$ and ξ. Richards et al. (2022b, Figures S2–S4) give estimates for the parameters of the marginal models and 20-year marginal return level estimates for convective, non-convective and all rainfall, respectively, with estimation of the return levels accounting for the differing length of the observation record for each process. We observe similar patterns in estimates of $p(s)$ and $q_\lambda(s)$ for each of the three processes, with $p(s)$ and $q_\lambda(s)$ decreasing and increasing with elevation, respectively.

There are differences between the estimates of $\upsilon(s)$ and the 20-year return levels for the three processes; whereas for $Y_\mathcal{N}$ both increase with elevation, for $Y_\mathcal{C}$ and $Y_\mathcal{E}$, we observe spatially smooth estimates of both, with larger values in the east of the domain (see Figures S2 and S4), indicating that the physical processes driving convection are only weakly affected by orography, unlike frontal rain (see Figure S3). The 20-year return levels for $Y_\mathcal{C}$ and $Y_\mathcal{E}$ are much higher than for $Y_\mathcal{N}$, which is consistent with their different physical properties. This property also holds for all larger return levels since the median shape parameter estimates (95% confidence interval), across all bootstrap samples, for convective, non-convective and all rainfall are $\hat{\xi}_\mathcal{C} = 0.226$ (0.201,0.245), $\hat{\xi}_\mathcal{N} = -0.074$ (-0.108, -0.03) and $\hat{\xi}_\mathcal{E} = 0.286$ (0.266,0.307), respectively. Although Richards et al. (2022a) allow ξ to vary

with location, our inference for the marginal tails of all rainfall agree with theirs as our estimate, $\hat{\xi}_{\mathcal{C}} = 0.226$, falls within the range of their estimates; similar agreement is found with results by Chan et al. (2014) and Hosseinzadehtalaei et al. (2020), and the estimate is not inconsistent with the range of feasible values for extreme precipitation proposed by Martins and Stedinger (2000). Note that the change in the shape estimate from 0.286 for all rainfall to 0.226 for convective rainfall supports our assumptions that, marginally, $Y_{\mathcal{E}}$ is a mixture of processes.

To assess the fit of the generalised Pareto additive models, we present Q-Q plots of the marginal fits for each of $Y_{\mathcal{C}}$, $Y_{\mathcal{N}}$ and $Y_{\mathcal{E}}$ at five randomly sampled locations in the Supplementary Material (Richards et al., 2022b, Figures S5–S7). These figures show good individual fits for each of the processes at most locations; some slight biases in return level estimates are found at certain locations near the boundaries of the domain, but we found that this did not compromise the overall marginal fit or the inference on spatial aggregates. To evaluate the fit over all locations, we use a pooled Q-Q plot (Heffernan and Tawn, 2001), transforming all data onto standard exponential margins using the fitted model; these plots are also given in these figures. Again, we observe good fits for each process. Confidence intervals for the Q-Q plots are estimated using the following bootstrap procedure: we create 250 bootstrap samples of the data using the stationary bootstrap approach of Politis and Romano (1994) with expected block size of 48 hours. With $q_\lambda(s)$ treated as fixed across all samples, we then estimate $\upsilon(s)$ and ξ for each bootstrap sample. For the pooled diagnostic plot, we apply the marginal transformation to the original data using the 250 estimated marginal parameter sets.

5.2 Extremal dependence modelling

We proceed by fitting the extremal dependence models described in Section 3.3 separately to each process. We use a different exceedance threshold, i.e., u in eq. (5), for each process; the 96% quantile for $X_{\mathcal{C}}$ and the 99% quantile for $X_{\mathcal{E}}$ and $X_{\mathcal{N}}$. The thresholds were chosen so that the limiting assumptions in eq. (5) were applicable; other choices were considered but these led to poorer inference on the extremal behaviour of spatial aggregates, which we quantified using the measures and diagnostics described in Section 5.3. Ideally, for $X_{\mathcal{E}}$ we would set u such that the number of observations used for inference was identical to the mixture approach to provide a fair comparison. However, we found that using such a threshold led to poorer inference for $X_{\mathcal{E}}$, with significant evidence that assumptions in eq. (5) fail for any lower threshold choice.

Inference is conducted using the stratified sampling regime described in Section 3.3. We use the same value of d_s for all three fitted models; we follow Richards et al. (2022a) and set $d_s = 5000$. Different values of h_{max} were also taken: 35km for $X_{\mathcal{E}}$ and $X_{\mathcal{C}}$ and 250km for $X_{\mathcal{N}}$, with the latter reflecting the long-range extremal dependence of $X_{\mathcal{N}}$. Inference for all three extremal dependence models was conducted using different sub-samples of sampling locations. For $X_{\mathcal{N}}$, we tried both parametric forms of β, i.e., $\beta_{\mathcal{N}}$ in eq. (7) and the form described in Richards et al. (2022b); with $\beta_{\mathcal{N}}$ giving a better fit, we present findings for this form only. Parameter estimates for all models are given in Richards et al. (2022b). For $X_{\mathcal{E}}$ and $X_{\mathcal{C}}$, we found that fixing $\kappa_{\delta_4} = \kappa_{\beta_3} = 1$ and $\Delta = 0$ did not restrict the quality of either fit.

In Figure 5, we fix a conditioning site s_O in the centre of the domain and evaluate all dependence functions for each process at each pairwise anisotropic distance, i.e., $h(s_i, s_O)$ for $s_i \in \boldsymbol{S}$ and $i = 1, \ldots, d$; these are then plotted against the great-circle distances in the original coordinate system, denoted $h_*\{s_i, s_O\}$; as the anisotropy transformations differ for the three processes, we cannot compare estimates of their respective dependence functions in the original space, i.e., as functions of $h\{s_i, s_O\}$. Figure 5 indicates that there are similarities within the extremal dependence structures of $X_{\mathcal{C}}$ and $X_{\mathcal{E}}$, as most of the functions are approximately equal with slight differences for μ and δ. In contrast, there are widely different estimated structures for the α and β functions for $X_{\mathcal{N}}$, with α decaying much slower and independence not being achieved for any pairs of sites within $\mathcal{S}$.

The estimate of $\Delta = 6.81$km given in Table 1 suggests that the process $X_{\mathcal{N}}$ is asymptotically dependent up to this distance in the anisotropic setting, which corresponds to roughly 7.04km in the original coordinate system. Evidence of asymptotic dependence, even at short-range, has not been found previously using the spatial conditional extremes modelling approach (Wadsworth and Tawn, 2022; Huser and Wadsworth, 2020; Simpson and Wadsworth, 2021; Shooter et al., 2021a). As extremal dependence is predominantly exhibited through the α function, this suggests that extreme realisations of $X_{\mathcal{N}}$ are much smoother than extreme events drawn from the other two processes, which agrees with our understanding of convective and non-convective precipitation. Keef et al. (2013) identify additional constraints for the Heffernan and Tawn (2004) framework, which lessen the risk of both parameter identifiability and inconsistency issues, and which in the spatial context correspond to $\beta(h) = 0$ whenever $\alpha(h) = 1$; our fitted model breaks this constraint with $0 < \beta < 0.003$ for $0 < h \leq \Delta$,

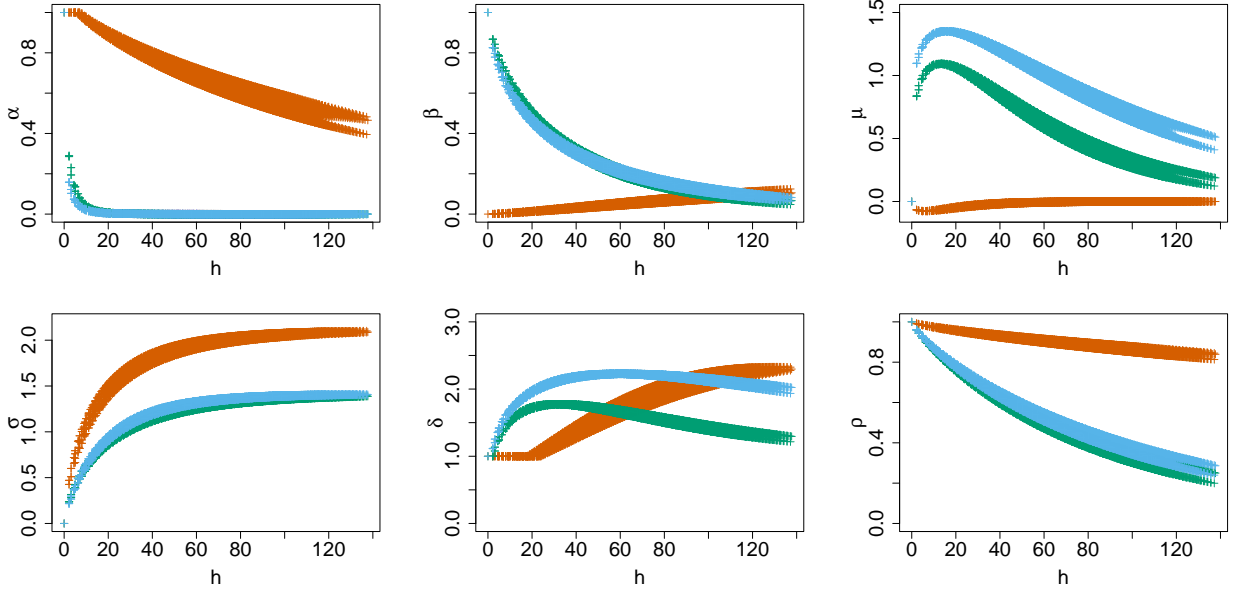but this minor deviation is not problematic.

Figure 5: Estimates of extremal dependence functions evaluated at $h(s_i, s_O)$ for $i = 1, \ldots, d$, i.e., anisotropic distances, against great-circle distances $h := h_*(s_i, s_O)$, which are given in km. The conditioning site s_O is in the centre of the spatial domain $\mathcal{S}$ and the colours correspond to the estimates for the different spatial processes; these are green, orange and blue for convective (green), non-convective (orange) and all rainfall (blue), respectively. Dependence functions α and β determine the strength of extremal dependence within the standardised process $X(s)$, whilst (μ, σ, δ) and ρ control the marginal and dependence properties, respectively, of the residual process $Z(s|s_O)$.

To compare the model fits of $Y_{\mathcal{C}}$, $Y_{\mathcal{N}}$ and $Y_{\mathcal{E}}$, we propose the following diagnostic that we present for a generic process. We investigate how the "conditional process"

$$\{Y(s) : s \in \mathcal{S}\} \mid (Y(s_O) = y^{(l)})$$

changes with distance $h_*(s, s_O)$ for $s, s_O \in \mathcal{S}$, where $y^{(l)}$ denotes the l-year return level for $Y(s_O)$. Figure 6 gives point-wise estimates for the median, and the 2.5% and 97.5% marginal quantiles, for $l = 1$ and 50; these are obtained from $50,000$ realisations from the fitted model. To achieve the largest possible distances, s_O is chosen to be on the boundary of $\mathcal{S}$ and, to ease comparison, we consider a transect of points in $\mathcal{S}$ only. We take into account the respective length of observation periods when evaluating the respective $l-$year return levels, e.g., for $Y_{\mathcal{C}}$ we take $y^{(l)} = F_{Y_{\mathcal{C}}(s_O)}^{-1}\{1 - 1/(l \times |\mathcal{C}|/20)\}$ and similarly for $Y_{\mathcal{N}}$ and $Y_{\mathcal{E}}$. The diagnostics for $Y_{\mathcal{C}}$ and $Y_{\mathcal{E}}$ are almost identical, with the conditional

24

medians of the respective processes both decaying quickly with distance. For $Y_{\mathcal{N}}$, the conditional median decays fairly slowly, with non-zero values at even the largest distances.
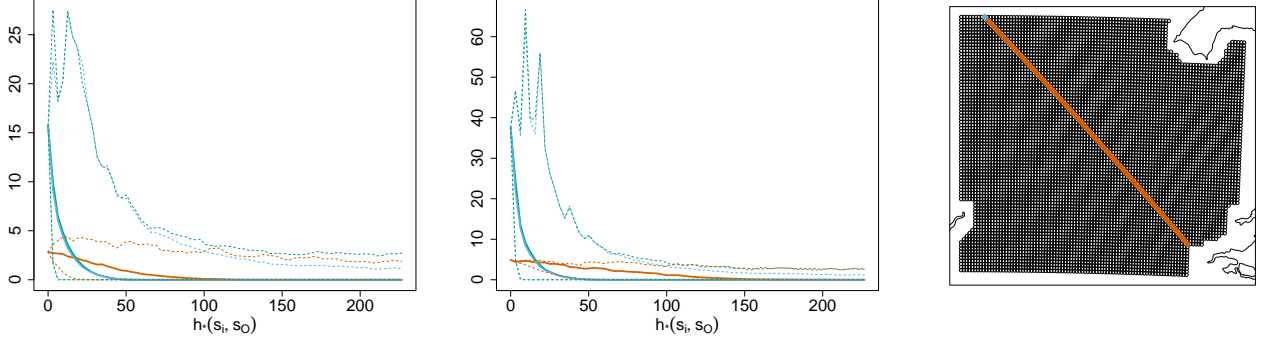

Figure 6: Summary statistics for $\{Y(s) : s \in \mathcal{S}\} \mid \{Y(s_O) = y^{(l)}\}$ (mm/hr) against distance $h_*(s, s_O)$ (km), with $y^{(l)}$ the $l = 1$ and $l = 50$ years return level in the left and centre panels, respectively; the process is considered along a transect of points (right) only. Solid lines correspond to estimates for conditional medians, dashed lines denote 95% confidence intervals. Lines are coloured green, brown and blue for convective, non-convective and all rainfall, respectively. Right: brown and blue points denote the transect and s_O, respectively.

Figure 7 presents realisations of $\{Y(s) : s \in \mathcal{S}\} \mid \{Y(s_O) > F^{-1}_{Y(s_O)}(0.99)\}$ for $Y_{\mathcal{C}}$ and $Y_{\mathcal{N}}$, where each s_O was uniformly sampled over $\mathcal{S}$, with their locations identified; similar plots are given for $Y_{\mathcal{E}}$ in Richards et al. (2022b, Figure S8). There are similarities for $Y_{\mathcal{C}}$ and $Y_{\mathcal{E}}$, with characteristics we anticipated for convective rainfall, i.e., high intensity, spatially localised events with a large proportion of the domain $\mathcal{S}$ being dry. In contrast, $Y_{\mathcal{N}}$ produces events that are much smoother spatially, cover a much large area, and are lower in their intensity.

5.3   Inference on spatial aggregates

For each process, we draw $b = 5.5 \times 10^5$ realisations using Algorithm 3, with the number of population replicates, b', being increased until the estimated quantiles of $\bar{R}_{\mathcal{A}}$ were stable; we used $b' := 8b$ for $Y_{\mathcal{C}}$ and $Y_{\mathcal{E}}$, and $b' = 8b$ for $Y_{\mathcal{N}}$. Regions $\boldsymbol{S}_\tau, \boldsymbol{A}_c$ and $\boldsymbol{A}$, described in Section 4.2, are illustrated in Figure 4 with $\boldsymbol{S} := \mathbf{s}$; the sets $\boldsymbol{A}$ and $\boldsymbol{S}_\tau$ are not changed over the processes, but we take $\boldsymbol{A}_c = \boldsymbol{S}_\tau$ for $Y_{\mathcal{C}}$ and $Y_{\mathcal{E}}$ and for $Y_{\mathcal{N}}$ we create $\boldsymbol{A}_c$ using the heuristic described in Section 4.2, with $\boldsymbol{S}^+ := \boldsymbol{A}_c \setminus \boldsymbol{S}_\tau$ illustrated by the purple points in Figure 4. Using these regions, we create samples of $\{Y_{\mathcal{M}}(s) : s \in \boldsymbol{S}_\tau\}$ for different regions $\boldsymbol{A}$ (see Figure 4) by drawing from $\{Y_{\mathcal{C}}(s) : s \in \boldsymbol{S}_\tau\}$ and $\{Y_{\mathcal{N}}(s) : s \in \boldsymbol{S}_\tau\}$ with estimated probability $\hat{p}_{\mathcal{C}}$ and $1 - \hat{p}_{\mathcal{C}}$, respectively.
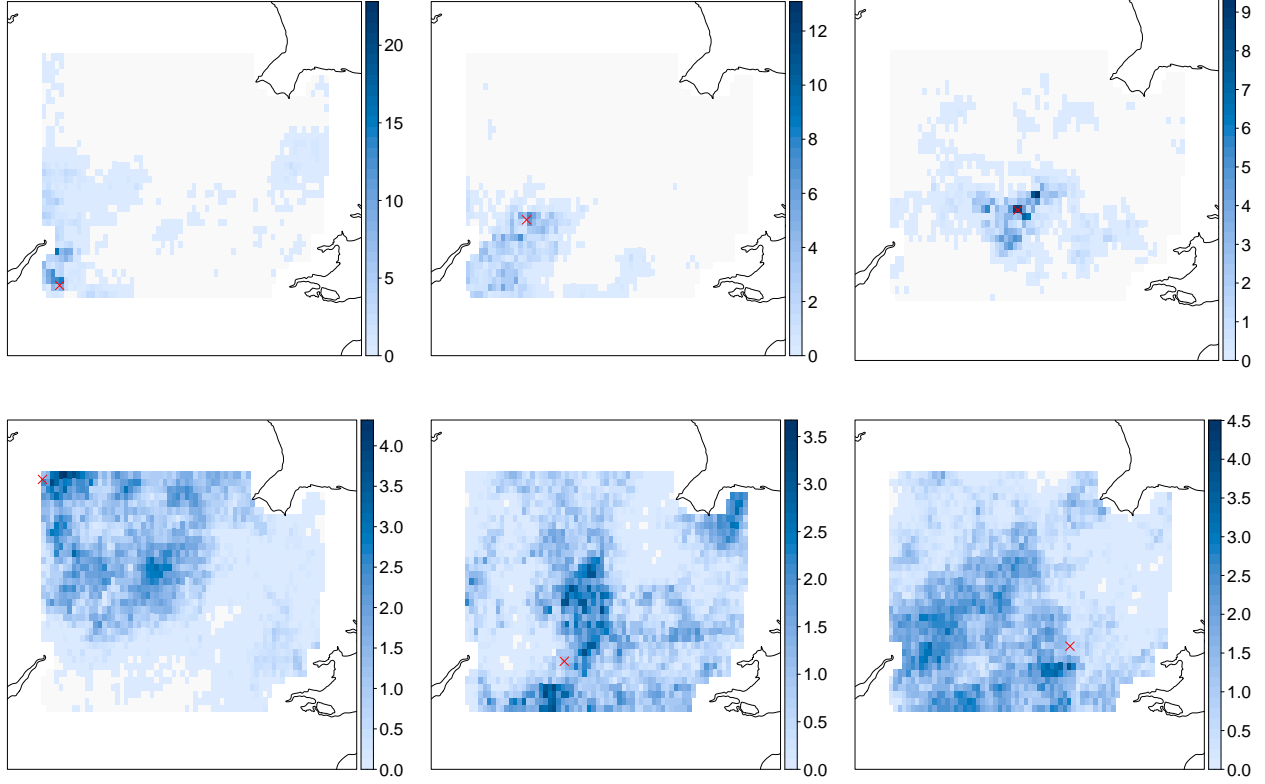
Figure 7: Extreme precipitation fields (mm/hr). Realisations from the fitted models for exceedances over the 0.99 quantile at conditioning site s_O for convective (top) and non-convective (bottom) events. The conditioning sites s_O are given by the red crosses. Scales differ within each panel and row.

We take $\bar{R}_{\mathcal{M},\mathcal{A}}$ to denote samples from $\bar{R}_{\mathcal{A}}$ created using our new modelling approach and $\bar{R}_{\mathcal{E},\mathcal{A}}$ for samples using the single process approach of Richards et al. (2022a). To compare mixture and non-mixture approaches, we present Q-Q plots in Figures 8 and S9 (Richards et al., 2022b) for spatial aggregates. In Figure S9, we assess the fit of individual mixture components $Y_{\mathcal{C}}$ and $Y_{\mathcal{N}}$ by using Q-Q plots of the variables $\bar{R}_{\mathcal{C},\mathcal{A}} := |\mathcal{A}|^{-1} \int_{\mathcal{A}} Y_{\mathcal{C}}(s)\mathrm{d}s \approx |\boldsymbol{A}|^{-1} \sum_{s \in \boldsymbol{A}} Y_{\mathcal{C}}(s)$ and similarly for $\bar{R}_{\mathcal{N},\mathcal{A}}$. For Figure 8, we compare Q-Q plots for both $\bar{R}_{\mathcal{M},\mathcal{A}}$ and $\bar{R}_{\mathcal{E},\mathcal{A}}$. Point-wise confidence interval estimates for the quantiles on these figures are created using a bootstrap technique. Specifically, we obtain a sample of 50 parameter estimates for each of model $X_{\mathcal{C}}$ and $X_{\mathcal{N}}$ by first applying a stationary bootstrap (Politis and Romano, 1994) to observations of the process with expected block size of 48 hours, and then fitting the extremal dependence model of Section 3.3, i.e., with the same h_{max} and d_s, but using different triples of sampling locations. For each of the 50 parameter estimates for $X_{\mathcal{C}}$ and $X_{\mathcal{N}}$, we draw 5.5×10^5 realisations of $\{Y_{\mathcal{M}}(s) : s \in \boldsymbol{A}_c\}$, with $\boldsymbol{A}_c$ and $\boldsymbol{S}_\tau$ consistent across parameter estimates. Our

implementation does not account for the classification uncertainty as we believe that incorporating its uncertainty does not justify the computational expense given the precision with which $p_\mathcal{C}$ is estimated.

Figures 8 and S9 show good fits for all components of both modelling approaches, with slight underestimation for the smaller quantiles of $\bar{R}_{\mathcal{N},\mathcal{A}}$. It appears that $\bar{R}_{\mathcal{M},\mathcal{A}}$ provides slightly better fits than the corresponding $\bar{R}_{\mathcal{E},\mathcal{A}}$ for most of the tail, particularly for the largest regions (3 and 4); indicating an improvement over the approach of Richards et al. (2022a). To help understand these observed differences in fit, we estimate the proportion $p_\mathcal{M}(r^{(l)}, \mathcal{A})$ of non-convective events that contribute to $\bar{R}_{\mathcal{M},\mathcal{A}} > r^{(l)}$, where here $r^{(l)}$ denotes the 99% and 99.5% quantile of the simulated $\bar{R}_{\mathcal{M},\mathcal{A}}$. The proportion $p_\mathcal{M}(r^{(l)}, \mathcal{A})$ increases with the size of $\mathcal{A}$; for the 99% quantile, $p_\mathcal{M}(r^{(l)}, \mathcal{A})$ increases from 0.029 to 0.052 for the smallest and largest regions, respectively, and for the 99.5% quantile, the values increase from 0.004 to 0.040. Neither the models for $\bar{R}_{\mathcal{M},\mathcal{A}}$ nor $\bar{R}_{\mathcal{E},\mathcal{A}}$ are able to capture the very largest empirical quantiles when $\mathcal{A}$ is the smallest region, region 1. This discrepancy was caused by two potentially spurious events that produced large values for the average over region 1, at consecutive hours (Whitall, M., 2021, personal communication).

As the differences in Figure 8 are difficult to observe by eye, we better quantify our assertion that model $\mathcal{M}$ fits better than model $\mathcal{E}$, through a comparison using a metric adapted from a measure of Varty et al. (2021). Let $Q(p; W) : [0, 1] \mapsto \mathbb{R}_+$ denote the sample p-th quantile of random variable W and $\tilde{R}_\mathcal{A}$ denote the observed spatial aggregate; we define measures of the expected deviance in the Q-Q plot for W against $\tilde{R}_\mathcal{A}$ from the line $y = x$,

$$\Lambda_i(W) = \frac{1}{m} \sum_{j=1}^{m} |Q(p_j; W) - Q(p_j; \tilde{R}_\mathcal{A})|^i, \ \ \text{for } i = 1, 2, \tag{11}$$

described through the mean absolute and mean squared distances Λ_1 and Λ_2, respectively, where $\{0 < p_1 < \cdots < p_m < 1\}$ is a grid of m ($m \geq 2$) equally spaced values with $p_m = 1 - (p_2 - p_1)$. As our interest lines in the upper-tails of $\bar{R}_\mathcal{A}$, we set p_1 close to one. We evaluate Λ_1 and Λ_2 for $W := \bar{R}_{\mathcal{M},\mathcal{A}}$ and $W := \bar{R}_{\mathcal{E},\mathcal{A}}$, and for each region $\mathcal{A}$ illustrated in Figure 4, with $p_1 = 0.99$ and $m = 432$, i.e., the number of observations of $\tilde{R}_\mathcal{A}$ that exceed the 0.99-quantile; this is done for each sample of bootstrapped parameter estimates, see above. These goodness-of-fit diagnostics are given in Table 1, with smaller values indicating a best fit under that metric; our approach provides better fits
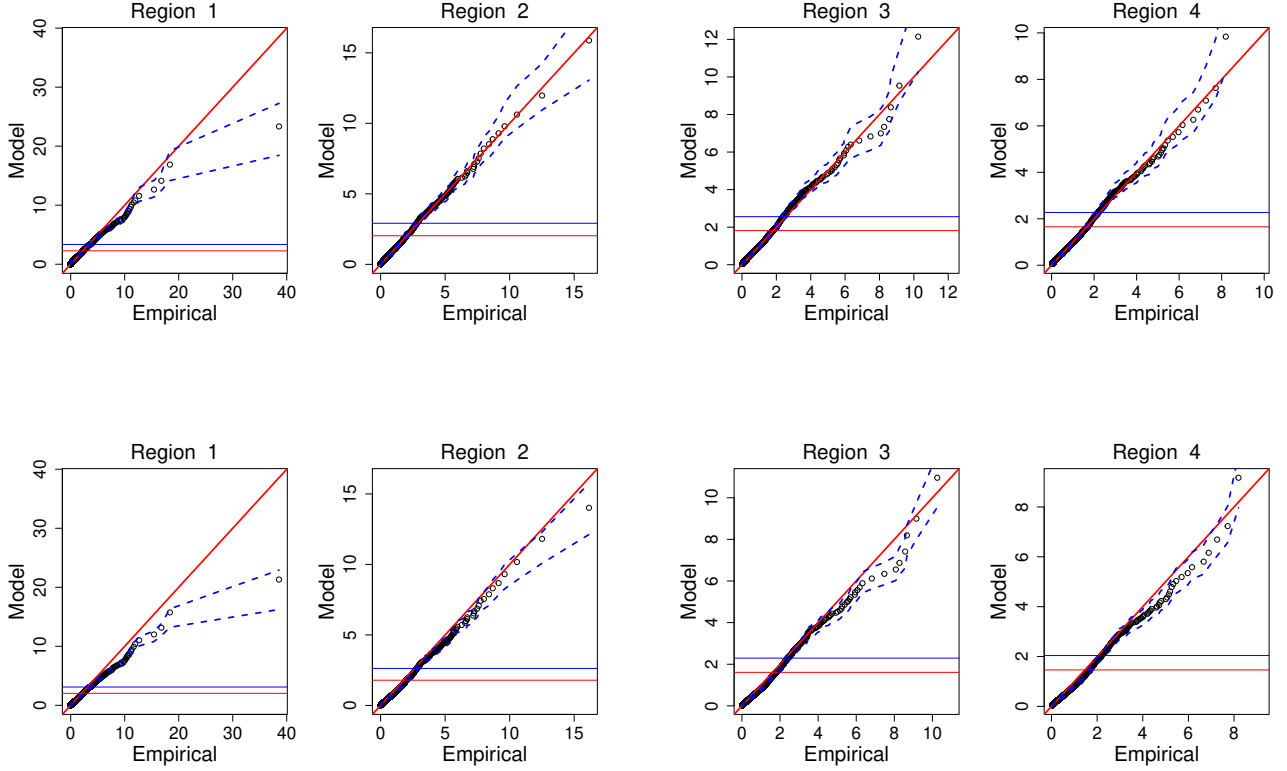
Figure 8: Q-Q plots for the model, and empirical, of average rainfall rate (mm/hr) over regions of increasing size. Top: new approach with $\bar{R}_{\mathcal{M},\mathcal{A}}$, bottom: existing approach with $\bar{R}_{\mathcal{E},\mathcal{A}}$. The considered quantiles range from the 0.8-quantile to a value corresponding to the respective 20 year return level. Blue dashed lines denote point-wise 95% quantile estimates. Regions 1–4 correspond to those illustrated in Figure 4. The blue and red horizontal lines denote the 99% and 99.5% quantiles of the respective simulated aggregates.

for all four regions.

6   Discussion

We have presented a simple but effective mixture model extension of the approach proposed by Richards et al. (2022a) for modelling the extremes of spatial aggregates of precipitation over a wide range of spatial scales. We classify observed fields as being either convective or non-convective, fit separate spatial models to each dataset and then simulate from both models and combine samples to explore the upper-tail behaviour of $\bar{R}_{\mathcal{A}}$. Fits of our spatial models reveal major differences between the extremal dependence structures of convective and non-convective precipitation events, with evidence to suggest that the latter process may exhibit short-range asymptotic dependence but long-range

Table 1: Median estimates (0.025 and 0.975 quantile estimates) of aggregate diagnostics Λ_1 and Λ_2, defined in eq. (11), over all bootstrap samples (2 decimal places).

	$\mathcal{A}$			
	1	2	3	4
$\Lambda_1(\bar{R}_{\mathcal{M},\mathcal{A}})$	0.46 (0.23, 0.73)	0.12 (0.07, 0.31)	0.15 (0.09, 0.37)	0.12 (0.07, 0.32)
$\Lambda_1(\bar{R}_{\mathcal{E},\mathcal{A}})$	0.78 (0.52, 1.02)	0.32 (0.11, 0.53)	0.22 (0.11, 0.41)	0.23 (0.10, 0.40)
$\Lambda_2(\bar{R}_{\mathcal{M},\mathcal{A}})$	1.12 (0.45, 1.79)	0.04 (0.01, 0.15)	0.05 (0.03, 0.21)	0.03 (0.01, 0.18)
$\Lambda_2(\bar{R}_{\mathcal{E},\mathcal{A}})$	1.70 (1.17, 2.80)	0.13 (0.02, 0.43)	0.08 (0.03, 0.26)	0.08 (0.02, 0.21)

asymptotic independence. Our mixture modelling approach was compared against an approach where the mixture structure is ignored and we found that the former was able to better capture the extremal behaviour of aggregates over very large spatial regions. We now discuss some further extensions that can be made to improve the model.

In Section 5.3, we identified two consecutive convective events that provided the two largest values of the empirical average rainfall $\bar{R}_{\mathcal{A}}$ over the smallest aggregate region $\mathcal{A}$, denoted region 1; neither the new nor old model for $\bar{R}_{\mathcal{A}}$ were able to capture the characteristics of these two events. The consecutive nature of the values suggests that there may be a temporal aspect to the data that we cannot currently capture with the model. The data themselves are aggregates of continuous-in-time precipitation over a temporal interval of one hour and a spatial grid-box. Hence, two spatially identical storms can produce different hourly values simply by moving at different speeds through a grid-box. Potential model extensions would incorporate the speed of a storm as a covariate in the model or build a conditional spatial-temporal model (Simpson and Wadsworth, 2021) adapted for precipitation.

We classified entire observed fields as convective or not regardless of the amount of convective rainfall properties across $\mathcal{S}$, or the presence/proportion of non-convective rainfall within the same spatial domain. The classification of the whole field into one of two processes appears to work well as we observe distinctly different structures for the two identified components. A more realistic approach is to accommodate spatial mixing of convective and non-convective rainfall across $\mathcal{S}$ within a single field. This may require modelling of extremal dependence between rain types and a framework for simulating different types of event within the same region.

The classification technique is deterministic and so improvements may be achieved by adopting a more probabilistic approach with labelling of mixture components conducted within inference. In

this vein, Tendijck et al. (2021) have developed mixture models in a conditional extremes setting by allowing the α and β functions to change with the mixture component, but only in a bivariate context. For the residual process $Z(s|s_O)$, a Dirichlet mixture of Gaussian processes (Duan et al., 2007) could be used, rather than a single process. This would have the added benefit of the model not being limited to only two mixture components. Instead, a number of mixture components could be used, each with their own dependence structure; that is, we could model more than two "types" of extreme rainfall. In the context of modelling spatial extremes, Hazra and Huser (2021) advocate the use of Dirichlet mixtures of Student's-t processes.

This work is an important step towards protecting society from the impacts of extreme rainfall through the facilitation of more precise estimation of rainfall hazards from short-run climate model simulations. For this work to be appropriate in the future, we must accommodate future climate change into both the marginal and extremal dependence structures of the underlying process, which could be achieved through the use of linear or additive regression models. Further non-stationarity in the extremal dependence structure may be handled through the inclusion of other meteorological covariates, e.g., wind-speed, which is likely to have a large impact on the anisotropy in $Y_{\mathcal{M}}$. Finally, we recognise that it is not only the extremes of hourly-spatial aggregates that contribute to river flooding, but also the extremes of aggregates taken over temporal windows of varying length. In order to adapt our approach to allow for modelling of spatio-temporal aggregates, we would require a conditional spatial-temporal dependence model (Simpson and Wadsworth, 2021).

Acknowledgments

References

Berg, P., Moseley, C., and Haerter, J. O. (2013). Strong increase in convective precipitation in response to higher temperatures. *Nature Geoscience*, 6(3):181–185.

Brown, S. J. (2018). The drivers of variability in UK extreme rainfall. *International Journal of Climatology*, 38(S1):e119–e130.

Catto, J. L. and Pfahl, S. (2013). The importance of fronts for extreme precipitation. *Journal of Geophysical Research: Atmospheres*, 118(19):10–791.

Chan, S. C., Kendon, E. J., Fowler, H. J., Blenkinsop, S., Roberts, N. M., and Ferro, C. A. (2014). The value of high-resolution Met Office regional climate models in the simulation of multihourly precipitation extremes. *Journal of Climate*, 27(16):6155–6174.

Coles, S. G. (1993). Regional modelling of extreme storms via max-stable processes. *Journal of the Royal Statistical Society: Series B (Methodological)*, 55(4):797–816.

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

Coles, S. G. and Tawn, J. A. (1996). Modelling extremes of the areal rainfall process. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(2):329–347.

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.

de Fondeville, R. and Davison, A. C. (2022). Functional peaks-over-threshold analysis. *Journal of the Royal Statistical Society: Series B (Methodology)*, 84(4):1392–1422.

Duan, J. A., Guindani, M., and Gelfand, A. E. (2007). Generalized spatial Dirichlet process models. *Biometrika*, 94(4):809–825.

Fasiolo, M., Wood, S. N., Zaffran, M., Nedellec, R., and Goude, Y. (2021). Fast calibrated additive quantile regression. *Journal of the American Statistical Association*, 116(535):1402–1412.

Gregersen, I. B., Sørup, H. J. D., Madsen, H., Rosbjerg, D., Mikkelsen, P. S., and Arnbjerg-Nielsen, K. (2013). Assessing future climatic changes of rainfall extremes at small spatio-temporal scales. *Climatic Change*, 118(3):783–797.

Hazra, A. and Huser, R. (2021). Estimating high-resolution Red Sea surface temperature hotspots, using a low-rank semiparametric spatial model. *The Annals of Applied Statistics*, 15(2):572 – 596.

Heffernan, J. E. and Tawn, J. A. (2001). Extreme value analysis of a large designed experiment: A case study in bulk carrier safety. *Extremes*, 4(4):359–378.

Heffernan, J. E. and Tawn, J. A. (2004). A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society: Series B (Methodology)*, 66(3):497–546.

Hosseinzadehtalaei, P., Tabari, H., and Willems, P. (2020). Climate change impact on short-duration extreme precipitation and intensity–duration–frequency curves over Europe. *Journal of Hydrology*, 590:125249.

Houze Jr, R. A. (1997). Stratiform precipitation in regions of convection: A meteorological paradox? *Bulletin of the American Meteorological Society*, 78(10):2179–2196.

Huser, R. and Wadsworth, J. L. (2020). Advances in statistical modeling of spatial extremes. *Wiley Interdisciplinary Reviews: Computational Statistics*, 14:e1537.

Institute of Hydrology (1999). *Flood Estimation Handbook*. Wallingford, UK.

Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. CRC Press, New York.

Keef, C., Papastathopoulos, I., and Tawn, J. A. (2013). Estimation of the conditional distribution of a multivariate variable given that one of its components is large: Additional constraints for the Heffernan and Tawn model. *Journal of Multivariate Analysis*, 115:396–404.

Kendon, E., Fosser, G., Murphy, J., Chan, S., Clark, R., Harris, G., Lock, A., Lowe, J., Martin, G., Pirret, J., Roberts, N., Sanderson, M., and Tucker, S. (2019). UKCP Convection-permitting model projections: Science report. Technical report, The Met Office.

Kendon, E., Short, C., Pope, J., Chan, S., Wilkinson, J., Tucker, S., Bett, P., and Harris, G. (2021). Update to UKCP Local (2.2km) projections. Technical report, The Met Office.

Kendon, E. J., Roberts, N. M., Senior, C. A., and Roberts, M. J. (2012). Realism of rainfall in a very high-resolution regional climate model. *Journal of Climate*, 25(17):5791–5806.

Martins, E. S. and Stedinger, J. R. (2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, 36(3):737–744.

Met Office Hadley Centre (2019). UKCP Convection-Permitting Model Projections for the UK at 2.2km resolution. NERC EDS Centre for Environmental Data Analysis. `https://catalogue.ceda.ac.uk/uuid/ad2ac0ddd3f34210b0d6e19bfc335539`. Last accessed 14/09/2022.

Padoan, S. A., Ribatet, M., and Sisson, S. A. (2010). Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association*, 105(489):263–277.

Palacios-Rodríguez, F., Toulemonde, G., Carreau, J., and Opitz, T. (2020). Generalized Pareto processes for simulating space-time extreme events: an application to precipitation reanalyses. *Stochastic Environmental Research and Risk Assessment*, 34(12):2033–2052.

Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313.

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

Richards, J., Tawn, J. A., and Brown, S. (2022a). Modelling extremes of spatial aggregates of precipitation using conditional methods. *The Annals of Applied Statistics*, 16(4):2693–2713.

Richards, J., Tawn, J. A., and Brown, S. (2022b). Supplement to "Joint estimation of extreme spatially aggregated precipitation at different scales through mixture modelling".

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

Saunders, K., Stephenson, A. G., Taylor, P. G., and Karoly, D. (2017). The spatial distribution of

rainfall extremes and the influence of El Niño Southern Oscillation. *Weather and Climate Extremes*, 18:17–28.

Schroeer, K., Kirchengast, G., and O, S. (2018). Strong dependence of extreme convective precipitation intensities on gauge network density. *Geophysical Research Letters*, 45(16):8253–8263.

Shooter, R., Ross, E., Ribal, A., Young, I., and Jonathan, P. (2021a). Spatial dependence of extreme seas in the North East Atlantic from satellite altimeter measurements. *Environmetrics*, 32(4):e2674.

Shooter, R., Ross, E., Tawn, J. A., and Jonathan, P. (2019). On spatial conditional extremes for ocean storm severity. *Environmetrics*, 30(6):e2562.

Shooter, R., Tawn, J. A., Ross, E., and Jonathan, P. (2021b). Basin-wide spatial conditional extremes for severe ocean storms. *Extremes*, 24(2):241–265.

Simpson, E. S., Opitz, T., and Wadsworth, J. L. (2022). High-dimensional modeling of spatial and spatio-temporal conditional extremes using INLA and Gaussian Markov random fields. *arXiv e-prints*, arXiv:2011.04486.

Simpson, E. S. and Wadsworth, J. L. (2021). Conditional modelling of spatio-temporal extremes for Red Sea surface temperatures. *Spatial Statistics*, 41:100482.

Simpson, E. S., Wadsworth, J. L., and Tawn, J. A. (2020). Determining the dependence structure of multivariate extremes. *Biometrika*, 107(3):513–532.

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

Tendijck, S., Eastoe, E., Tawn, J., Randell, D., and Jonathan, P. (2021). Modeling the extremes of bivariate mixture distributions with application to oceanographic data. *Journal of the American Statistical Association*, doi:10.1080/01621459.2021.1996379.

Thibaud, E., Mutzner, R., and Davison, A. C. (2013). Threshold modeling of extreme spatial rainfall. *Water Resources Research*, 49(8):4633–4644.

Varty, Z., Tawn, J. A., Atkinson, P. M., and Bierman, S. (2021). Inference for extreme earthquake magnitudes accounting for a time-varying measurement process. *arXiv preprint arXiv:2102.00884*.

Wadsworth, J. and Tawn, J. (2022). Higher-dimensional spatial extremes via single-site conditioning. *Spatial Statistics*, 51:100677.

Weller, G. B., Cooley, D., Sain, S. R., Bukovsky, M. S., and Mearns, L. O. (2013). Two case studies on NARCCAP precipitation extremes. *Journal of Geophysical Research: Atmospheres*, 118(18):10–475.

Westra, S. and Sisson, S. A. (2011). Detection of non-stationarity in precipitation extremes using a max-stable process model. *Journal of Hydrology*, 406(1-2):119–128.

Wood, S. (2006). *Generalized Additive Models: An Introduction with R.* Chapman & Hall/CRC, New York. 2nd edition.

Youngman, B. D. (2019). Generalized additive models for exceedances of high thresholds with an application to return level estimation for U.S. wind gusts. *Journal of the American Statistical Association*, 114(528):1865–1879.

Zheng, F., Thibaud, E., Leonard, M., and Westra, S. (2015). Assessing the performance of the independence method in modeling spatial extreme rainfall. *Water Resources Research*, 51(9):7744–7758.