Weather and Climate Extremes 13 (2016) 44–53

Contents lists available at ScienceDirect

Weather and Climate Extremes

journal homepage: www.elsevier.com/locate/wace

A max-stable process model for rainfall extremes at different accumulation durations

Alec G. Stephenson [a,*], Eric A. Lehmann [b], Aloke Phatak [c]

[a] Data Analytics, CSIRO, Clayton South, Victoria, Australia
[b] Data Analytics, CSIRO, Acton, ACT, Australia
[c] Department of Mathematics & Statistics, Curtin University, Bentley, Western Australia, Australia

ARTICLE INFO

Article history:
Received 27 July 2015
Received in revised form
23 June 2016
Accepted 26 July 2016
Available online 27 July 2016

Keywords:
Bayesian hierarchical model
Extreme value theory
Max-stable process
Markov chain Monte Carlo
Intensity-duration-frequency curve
Spatial modeling

ABSTRACT

A common existing approach to modeling rainfall extremes employs a spatial Bayesian hierarchical model, where latent Gaussian processes are specified on distributional parameters in order to pool spatial information. The data are taken to be conditionally independent given the latent processes, and so spatial dependence arises only through this conditional structure. This methodology can be extended by incorporating an explicit max-stable dependence structure, which therefore produces more realistic spatial surfaces, and removes the assumption of conditional independence. We further extend the max-stable methodology to incorporate the joint modeling of rainfall data at different accumulation durations. We therefore pool information across both space and accumulation duration within a broad framework which includes an explicit specification of spatial dependence. Our model can be used to derive inferences at ungauged sites, and easily incorporates missing values.

Our methodology is applied to a dataset of pluviometer records recorded at 182 meteorological stations located on the Central Coast of New South Wales, Australia. For each station, rainfall data are accumulated over 16 different durations ranging from 5 min to 7 days. The model is fitted using Markov chain Monte Carlo simulation, employing auxiliary variables so that exact Bayesian inference can be performed. We present estimated parameters and posterior inferences for isohyetal maps and intensity-duration-frequency curves at selected sites of interest, and compare our inferences with those derived from the standard latent variable model.

1. Introduction

The occurrence, extent and intensity of rainfall has a significant impact on ecosystems, population health, food production, insurance costs, and infrastructure. There is therefore a need for accurate analysis of these extremes and their consequences. The estimation of rainfall intensity-duration-frequency (IDF) relationships is of central importance for impacts assessment when considering extreme rainfall at a particular site. Regional frequency analysis (Hosking and Wallis, 1995) is a relatively simple method for estimation of IDF curves, but it is difficult to expand these procedures to more general settings, such as temporal variability or the incorporation of covariates. Moreover, regional frequency analysis typically makes very simple assumptions regarding the pooling of spatial information in the local area around the site of interest, and it does not explicitly model the relationship between rainfall at different accumulation durations.

An alternative approach to regional frequency analysis is to apply Bayesian hierarchical modeling by specifying latent Gaussian spatial processes on distributional parameters. Bayesian hierarchical modeling represents a approach that allows for the integration of multiple sources of uncertainty. Examples of this approach include Gaetan and Grigoletto (2007) and Sang and Gelfand (2008), who both analyze rainfall maxima using the generalized extreme value (GEV) distribution, where the parameters of that distribution are defined using latent Gaussian processes. Recent work by Lehmann et al. (2013) and Lehmann et al. (2016) has applied this latent variable model to rainfall of different accumulation durations using a relationship given by Koutsoyiannis et al. (1998). Their model therefore pools both spatial information and information across different durations within a formal framework. The disadvantage of the latent variable model is the assumption that the data at different sites are conditionally independent, which means that the spatial dependence arises only through the latent processes. Because of this

* Corresponding author.
E-mail address: alec.stephenson@csiro.au (A.G. Stephenson).

http://dx.doi.org/10.1016/j.wace.2016.07.002

assumption, the model cannot produce plausible extreme rainfall surfaces (Sang and Gelfand, 2010; Davison et al., 2012). The aim of this article is to incorporate the max-stable process framework of Reich and Shaby (2012) and Stephenson et al. (2015) into the latent modeling approach of Lehmann et al. (2013). This max-stable process gives an explicit spatial dependence structure which does not rely on assumptions of conditional dependence. Our model can be used to spatially extrapolate to ungauged sites (see Fig. 4.2 and Appendix B), and easily incorporates missing values (see Section 3).

The contribution of this article is to incorporate the modeling of different rainfall accumulation durations within a max-stable and latent variable process. The methodology we present is novel: as far as we are aware, this is the first published application of spatial max-stable processes applied to rainfall data at different accumulation durations. Return level estimates and other inferences can then be made at any duration and at any site, gauged or ungauged. The addition of a max-stable dependence structure leads to more realistic spatial surfaces than the standard latent variable framework. Additionally, our model is an extension of the latent variable model in the sense that it contains the latent variable model as a special case. For additional applications of latent variable models to spatial extremes see e.g. Cooley et al. (2007), Schliep et al. (2010) and Apputhurai and Stephenson (2013). For applications of max-stable process models to weather data using composite likelihoods see e.g. Smith and Stephenson (2009), Padoan et al. (2010) and Ribatet et al. (2012).

This article is structured as follows. In Section 2 we discuss the derivation of the rainfall maxima data calculated at different accumulation durations, recorded at 182 weather stations located in the Australian state of New South Wales. In Section 3 we describe the model that we apply to our data, and in Section 4 we present the model fit and results, including posterior mean parameter estimates and IDF curves, together with corresponding estimates of uncertainty. Section 5 provides a brief discussion.

2. Rainfall data

The dataset of rainfall maxima used in this work was extracted from pluviometer records acquired at $N=182$ weather stations located around the Sydney and Wollongong metropolitan areas in New South Wales, Australia, as shown in Fig. 2.1. Station records consist of rainfall depths (in mm) registered over 5 min intervals, with different record lengths ranging from 7 to 40 years of measurements during the $T=40$ year period 1961–2000. The pluviometer data at 5 min intervals were subsequently accumulated over $D=16$ different durations, namely 5, 10, 15 and 30 min, and 1, 2, 3, 6, 12, 24, 48, 72, 96, 120, 144 and 168 h. The resulting time series were then used to determine the largest annual rainfall accumulation for each corresponding station, year, and accumulation duration.

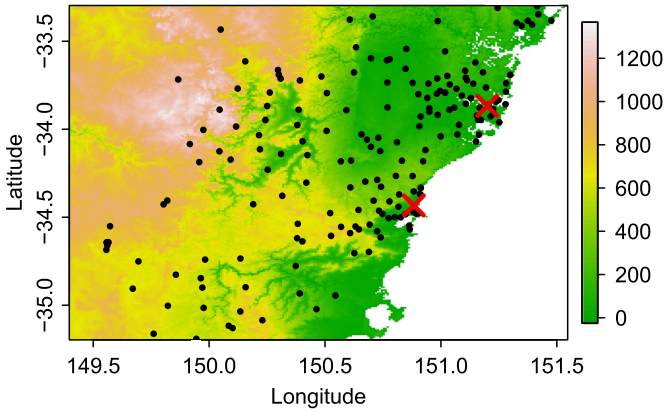

Fig. 2.1. The 182 weather station locations for the dataset used in this article. The two red crosses give the approximate locations of the cities of Sydney (upper) and Wollongong (lower). The legend gives the elevation in meters. The extent of the study area is roughly 200 km by 200 km.

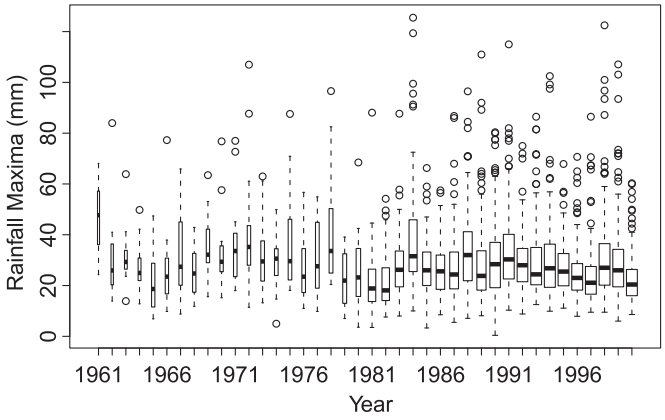

Fig. 2.2. Distributions of annual rainfall maxima at a one hour accumulation duration, plotted against time. Each boxplot shows the distribution of rainfall maxima as recorded at the sites given in Fig. 2.1. The width of each boxplot is proportional to the square-root of the total number of observations in that year.

The extent of the study area in Fig. 2.1 is roughly 200 km by 200 km. For each accumulation duration, the dataset contains a total of 3109 annual rainfall maxima distributed across the 182 weather stations. Within this region, rainfall accumulations tend to be larger during the February to June period, where the monthly average (i.e. the average accumulation for the calendar month) is around 120 mm. The monthly average drops to around 80 mm during the August to December period, and therefore the annual rainfall maxima are more likely to occur during the southern hemisphere autumn period. The largest monthly accumulation at the Sydney Observatory Hill site was 630 mm, recorded in February 1990. Boxplots of distributions of annual rainfall maxima within each year, for a one hour accumulation duration, are given in Fig. 2.2. There was less available data prior to 1980, as can be seen from the width of the boxplots. Fig. 2.2 shows little evidence of any temporal trend for the overall study area. A small number of annual rainfall maxima at a one hour duration exceed 100 mm.

3. Modeling rainfall extremes

The general concepts described here are based on the construction of Stephenson et al. (2015), but we extend these concepts to account for the different rainfall accumulation durations, based on the relationship suggested by Koutsoyiannis et al. (1998). We model the annual maximum rainfall as a max-stable spatial process. Full computational details are deferred to Appendix B. A theoretical justification for using max-stable processes to model spatial maxima is given by e.g. Schlather (2002). This justification is an extension of standard asymptotic arguments for the componentwise maxima of random vectors (e.g. Tawn, 1990). These asymptotic arguments assume that the underlying process at each site is stationary and subject to weak dependence assumptions (Leadbetter et al., 1983). For statistical purposes these results are often applied to annual maxima of seasonal meteorological data on the basis that e.g. annual rainfall maxima typically occur during a particular season or period within the year, and stationarity is assumed over that period. The asymptotic results therefore remain valid, but the quality of the approximation deteriorates because the effective block size over which the maxima are taken reduces.

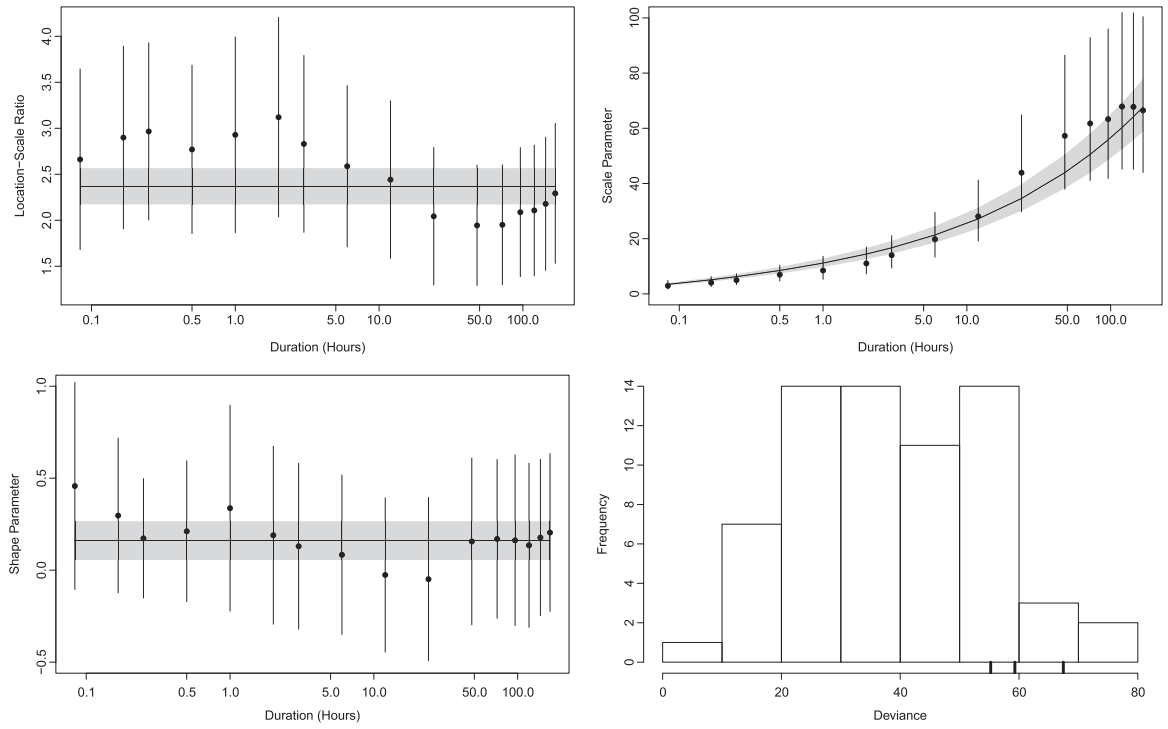

Fig. 3.1. Estimates of generalized extreme value parameters $\tilde{\mu}_d$ (top left), σ_d (top right) and ξ_d (bottom left) for the Cronulla Sewage Treatment Plant pluviometer station, located at 151.16° E, 34.03° S. Black dots with error bars are unconstrained maximum likelihood estimates and 95% confidence intervals for each duration $d = 1, \ldots, D$. The black lines and shaded region give maximum likelihood estimates under the constraints $\tilde{\mu}_d = \tilde{\mu}$, $\xi_d = \xi$, and σ_d defined by Eq. (3.1), for each duration $d = 1, \ldots, D$. The histogram gives deviance differences (i.e. twice the log-likelihood ratio) for the comparison of the two models across all sites with 20 or more observations, with each site considered independently. The three marks above the x-axis denote the individual chi-squared 10%, 5% and 1% significance quantiles for the corresponding likelihood ratio test.

For ease of exposition we focus only on the sites $\boldsymbol{s}_i$ for $i = 1, \ldots, N$ for which we have data, however inferences can also be made at any arbitrary site $\boldsymbol{s}$ (see Appendix B). Let $Y_{t,d}(\boldsymbol{s}_i)$ be the annual maximum rainfall at the data site $\boldsymbol{s}_i$ for year t and at accumulation duration $d = 1, \ldots, D$. The marginal distribution of $Y_{t,d}(\boldsymbol{s}_i)$ is taken to be Generalized Extreme Value (e.g. Coles, 2001) with location, scale and shape parameters given by $\mu_d(\boldsymbol{s}_i)$, $\sigma_d(\boldsymbol{s}_i) > 0$ and $\xi_d(\boldsymbol{s}_i)$ respectively. It can be reasonably assumed (Koutsoyiannis et al., 1998) that both $\tilde{\mu}_d(\boldsymbol{s}_i) = \mu_d(\boldsymbol{s}_i)/\sigma_d(\boldsymbol{s}_i)$ and $\xi_d(\boldsymbol{s}_i)$ do not depend on the accumulation duration, and therefore $\tilde{\mu}_d(\boldsymbol{s}_i) = \tilde{\mu}(\boldsymbol{s}_i)$ and $\xi_d(\boldsymbol{s}_i) = \xi(\boldsymbol{s}_i)$ for each $d = 1, \ldots, D$ (see also Fig. 3.1). Let ρ_d denote the dth accumulation duration in hours, for each $d = 1, \ldots, D$. Then for the scale parameter we employ the relationship given by Koutsoyiannis et al. (1998), with

$$\sigma_d(\boldsymbol{s}_i) = \frac{\rho_d\, \sigma(\boldsymbol{s}_i)}{\{\rho_d + \kappa(\boldsymbol{s}_i)\}^{\eta(\boldsymbol{s}_i)}}, \tag{3.1}$$

which involves a 'duration offset' term $\kappa(\boldsymbol{s}_i) > 0$ and a 'duration exponent' term $0 < \eta(\boldsymbol{s}_i) \leq 1$. We denote this distribution by $Y_{t,d}(\boldsymbol{s}_i) \sim \mathrm{GEV}[\tilde{\mu}(\boldsymbol{s}_i), \sigma_d(\boldsymbol{s}_i), \xi(\boldsymbol{s}_i)]$, and we assume independence across years $t = 1, \ldots, T$ and accumulation durations $d = 1, \ldots, D$.

The relationship in Eq. (3.1) was derived empirically by Koutsoyiannis et al. (1998), who show that it is consistent with several rainfall studies. It has been used successfully with other datasets (Muller et al., 2008; Van de Vyver, 2015). Alternative empirical relationships have been investigated by Garcia-Bartual and Schneider (2001). Nadarajah et al. (1998) used an explicit multivariate model (in a non-spatial setting) for the maxima at different accumulation durations, and showed that the ordering of rainfall depths at increasing durations implies restrictions on the GEV parameters. In particular, they showed that when $\tilde{\mu}$ and ξ are constant across durations, then $\sigma_{d_1} \leq \sigma_{d_2}$ whenever $d_1 < d_2$, which is a relationship that is enforced by Eq. (3.1).

For our precipitation data, we performed an exploratory maximum likelihood analysis at each individual site in order to investigate the Koutsoyiannis et al. (1998) relationship. Fig. 3.1 shows parameter estimates at one particular pluviometer station; we performed the same analysis independently for all sites that had at least 20 years of observed annual maxima. We compared the model fit using unconstrained parameters $(\tilde{\mu}_d, \sigma_d, \xi_d)$ for $d = 1, \ldots, D$ with the model fit under the constraints $\tilde{\mu}_d = \tilde{\mu}$, $\xi_d = \xi$, and σ_d defined by Eq. (3.1), for each duration $d = 1, \ldots, D$. Under this constrained model there are five parameters in total, namely $(\tilde{\mu}, \sigma, \xi, \kappa, \eta)$. The histogram in Fig. 3.1 gives the deviance differences for each model comparison, and the associated chi-squared quantiles for the corresponding likelihood ratio test. At the 5% significance level, only about 10% of our sites show individual significance, which is strong support for the use of Eq. (3.1) for our data.

To specify the max-stable spatial dependence structure, we first transform to residuals with a common marginal distribution. Define

$$X_{t,d}(\boldsymbol{s}_i) = \left\{1 + \xi(\boldsymbol{s}_i)\left[\frac{Y_{t,d}(\boldsymbol{s}_i)}{\sigma_d(\boldsymbol{s}_i)} - \tilde{\mu}(\boldsymbol{s}_i)\right]\right\}^{1/\xi(\boldsymbol{s}_i)}, \tag{3.2}$$

so that the marginal distribution of $X_{t,d}(\boldsymbol{s}_i)$ is given by the standard Fréchet distribution, where $\Pr(X_{t,d}(\boldsymbol{s}_i) < x) = \exp(-1/x)$. The joint distribution of the $X_{t,d}(\boldsymbol{s}_i)$ over the sites $i = 1, \ldots, N$ is then modeled as

$$\Pr(X_{t,d}(\boldsymbol{s}_i) < x_i,\ i = 1, \ldots, N) = \exp\left\{-\sum_{k=1}^{K}\left[\sum_{i=1}^{N}\left(\frac{w_k(\boldsymbol{s}_i)}{x_i}\right)^{1/\alpha}\right]^{\alpha}\right\}, \tag{3.3}$$

for each year t and duration d, where $\alpha \in (0, 1]$ is a spatial dependence parameter and $w_k(\cdot)$ are kernel basis functions defined

by

$$w_k(\boldsymbol{s}_i) = \frac{\Lambda(\| \boldsymbol{s}_i - \boldsymbol{v}_k \|/\tau)}{\sum_{j=1}^{N} \Lambda(\| \boldsymbol{s}_i - \boldsymbol{v}_j \|/\tau)} \tag{3.4}$$

for $k = 1, \ldots, K$, where $\Lambda(\cdot)$ is a kernel, $\tau > 0$ is the kernel bandwidth, and $\|\cdot\|$ is the Euclidean norm. The $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_K$ are a fixed set of K locations that we specify on a regularly-spaced grid over the area of interest. Using the above equation, it follows that $w_k(\boldsymbol{s}_i) \geq 0$ and $\sum_{k=1}^{N} w_k(\boldsymbol{s}_i) = 1$ for all sites $\boldsymbol{s}_i$. We employ the Gaussian kernel $\Lambda(u) \propto \exp(-u^2/2)$.

The specification of the number of grid locations K presents a trade-off between computational burden and the accuracy of the fit. The computational burden within the fitting algorithm increases for increasing K, because if there are fewer grid locations then there are fewer variables to update (see Appendix B). In our application we define our grid so that $K = 14^2 = 196$ is approximately equal to the number of meteorological stations.

We assign Gaussian spatial processes to the parameters of the Generalized Extreme Value (GEV) distribution. Let $\tilde{\boldsymbol{\mu}}$, $\boldsymbol{\sigma}$, $\boldsymbol{\xi}$, $\boldsymbol{\kappa}$ and $\boldsymbol{\eta}$ denote vectors containing the respective GEV parameters at each site. For example, $\tilde{\boldsymbol{\mu}} = [\tilde{\mu}(\boldsymbol{s}_1), \ldots, \tilde{\mu}(\boldsymbol{s}_N)]$. For each $\chi \in \{\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta}\}$ we therefore take

$$h_\chi(\chi) \sim \mathrm{MVN}(X_\chi \boldsymbol{\beta}_\chi, \Sigma_\chi), \tag{3.5}$$

where $\boldsymbol{\beta}_\chi = (\beta_{\chi,0}, \beta_{\chi,1}, \ldots, \beta_{\chi,n_\chi})^T$ is a column vector of parameters, and X_χ is a N by $n_\chi + 1$ design matrix. The function $h_\chi(\chi)$ represents a componentwise transformation and is used to ensure that the GEV parameters are transformed to a suitable range of values. Thus, we take $h_\chi(\chi) = \chi$ for $\chi \in \{\tilde{\boldsymbol{\mu}}, \boldsymbol{\xi}\}$, $h_\chi(\chi) = \log(\chi)$ for $\chi \in \{\boldsymbol{\sigma}, \boldsymbol{\kappa}\}$ and $h_\chi(\chi) = \log(\chi/(1-\chi))$ for $\chi = \boldsymbol{\eta}$. We use an exponential correlation function, which gives covariance matrix entries

$$[\Sigma_\chi]_{ij} = \delta_\chi \exp(-\| \boldsymbol{s}_i - \boldsymbol{s}_j \|/\lambda_\chi), \quad i, j = 1, \ldots, N, \tag{3.6}$$

where $\delta_\chi > 0$ and $\lambda_\chi > 0$ are the sill and the range respectively. The range parameter provides an indication of the distance beyond which the correlation between the values of a given GEV parameter at different spatial locations drops to a negligible level.

The above framework defines the density $\pi(\chi|\boldsymbol{\phi}_\chi)$ for $\chi \in \{\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta}\}$, with $\boldsymbol{\phi}_\chi = (\boldsymbol{\beta}_\chi, \delta_\chi, \lambda_\chi)$. The spatial dependence in the model defined by Eqs. (3.2) and (3.3) is therefore derived from two sources. Firstly, spatial dependence is derived through the parameters α and τ. Secondly, spatial dependence is induced through the latent Gaussian process specifications for the GEV parameters, as given in Eqs. (3.5) and (3.6). When $\alpha = 1$ we obtain the standard latent variable model, where the $X_{t,d}(\boldsymbol{s}_i)$ are independent, and the $Y_{t,d}(\boldsymbol{s}_i)$ are conditionally independent given the vectors $\tilde{\boldsymbol{\mu}}$, $\boldsymbol{\sigma}$, $\boldsymbol{\xi}$, $\boldsymbol{\kappa}$ and $\boldsymbol{\eta}$. In the $\alpha = 1$ case, the only spatial dependence is that which is induced by integration over $\tilde{\boldsymbol{\mu}}$, $\boldsymbol{\sigma}$, $\boldsymbol{\xi}$, $\boldsymbol{\kappa}$ and $\boldsymbol{\eta}$. The full model defined by Eqs. (3.2) and (3.3) includes both forms of spatial dependence, and it permits straightforward prediction at ungauged sites (see Appendix B).

The model given above permits an exact Bayesian analysis using the following construction, based on Stephenson (2009). Let $\boldsymbol{A} = \{A_{t,k}: t = 1, \ldots, T, \ k = 1, \ldots, K\}$ be independent random variables distributed according to the positive stable distribution (see Appendix A) with index equal to the spatial dependence parameter α. Also, define

$$\theta_t(\boldsymbol{s}_i) = \left[\sum_{k=1}^{K} A_{t,k} w_k(\boldsymbol{s}_i)^{1/\alpha} \right]^{\alpha}. \tag{3.7}$$

Reich and Shaby (2012) show that $Y_{t,d}(\boldsymbol{s}_i)$ are then conditionally independent, with

$$Y_{t,d}(\boldsymbol{s}_i)|\boldsymbol{A} \overset{ind}{\sim} \mathrm{GEV}[\tilde{\mu}_t^*(\boldsymbol{s}_i), \sigma_{t,d}^*(\boldsymbol{s}_i), \xi^*(\boldsymbol{s}_i)] \tag{3.8}$$

where the GEV parameters of the conditional distribution are defined by

$$\tilde{\mu}_t^*(\boldsymbol{s}_i) = \frac{\tilde{\mu}(\boldsymbol{s}_i) + (\theta_t(\boldsymbol{s}_i)^{\xi(\boldsymbol{s}_i)} - 1)/\xi(\boldsymbol{s}_i)}{\alpha \theta_t(\boldsymbol{s}_i)^{\xi(\boldsymbol{s}_i)}} \tag{3.9}$$

$$\sigma_{t,d}^*(\boldsymbol{s}_i) = \alpha \sigma_d(\boldsymbol{s}_i) \theta_t(\boldsymbol{s}_i)^{\xi(\boldsymbol{s}_i)} = \frac{\alpha \rho_d \, \sigma(\boldsymbol{s}_i) \theta_t(\boldsymbol{s}_i)^{\xi(\boldsymbol{s}_i)}}{\{\rho_d + \kappa(\boldsymbol{s}_i)\}^{\eta(\boldsymbol{s}_i)}} \tag{3.10}$$

$$\xi^*(\boldsymbol{s}_i) = \alpha \xi(\boldsymbol{s}_i). \tag{3.11}$$

This formulation permits Bayesian inference using standard Markov chain Monte Carlo techniques (Hastings, 1970) to simulate from the posterior distribution of the parameters and to perform spatial prediction at observed (gauged) or unobserved (ungauged) sites. It also allows us to easily incorporate missing values; if $Y_{t,d}(\boldsymbol{s}_i)$ is missing we simply omit the corresponding term from the likelihood function. Let $\boldsymbol{Y}$ represent the dataset of all rainfall maxima. The posterior density is then proportional to

$$L(\boldsymbol{Y}|\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta}, \boldsymbol{A}, \alpha, \tau) \pi(\boldsymbol{A}|\alpha) \left\{ \prod_{\chi} \pi(\chi|\boldsymbol{\phi}_\chi) \pi(\boldsymbol{\phi}_\chi) \right\} \pi(\alpha, \tau), \tag{3.12}$$

where $L(\cdot)$ is the likelihood function and $\pi(\boldsymbol{A}|\alpha)$ is the density function for $\boldsymbol{A}$ (see Appendix B). The product in Eq. (3.12) is over $\chi \in \{\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta}\}$. The functions $\pi(\boldsymbol{\phi}_\chi)$ and $\pi(\alpha, \tau)$ are independent prior density functions as defined in Appendix B. The density $\pi(\boldsymbol{A}|\alpha)$, cannot be computed in closed form, but we can introduce a further set of variables $\boldsymbol{B} = \{B_{t,k}: t = 1, \ldots, T, \ k = 1, \ldots, K\}$ such that $\pi(\boldsymbol{A}|\alpha)$ is equal to $\int \pi(\boldsymbol{A}, \boldsymbol{B}|\alpha) \, d\boldsymbol{B}$, and where the joint density $\pi(\boldsymbol{A}, \boldsymbol{B}|\alpha)$ can be easily computed. We can then replace $\pi(\boldsymbol{A}|\alpha)$ by $\pi(\boldsymbol{A}, \boldsymbol{B}|\alpha)$ within Eq. (3.12). See Appendix B for details.

4. Results

We initially conducted a series of small Markov chain Monte Carlo simulations in order to assist with the selection of covariates within the linear terms $X_\chi \boldsymbol{\beta}_\chi$ for each $\chi \in \{\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta}\}$, as given by Eq. (3.5). Table 4.1 displays the Deviance Information Criteria (DIC) of Spiegelhalter et al. (2002) for several different models. The exact derivation of these DIC values is given in Appendix B. The largest

Table 4.1
DIC values for several models (East = Easting, North = Northing, Elev = Elevation). The rows represent the covariates included in the expression $X_\chi \boldsymbol{\beta}_\chi$ for each $\chi \in \{\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta}\}$, as given in Eq. (3.5). The column headings give the vectors containing the respective GEV parameters for each site. If a GEV parameter is not contained in the column heading, then all covariates are included. For clarity, a fixed additive constant has been removed from the DIC values, and the minimum within each column is highlighted in bold.

	$\tilde{\boldsymbol{\mu}}$	$\boldsymbol{\sigma}, \boldsymbol{\kappa}, \boldsymbol{\eta}$	$\boldsymbol{\xi}$
East-North-Elev	**24.9**	24.9	24.9
East-North	56.8	**22.7**	54.9
East-Elev	74.2	45.6	**21.9**
North-Elev	84.3	27.6	49.7
East	31.8	51.7	50.7
North	56.1	34.7	24.3
Elev	94.1	41.3	33.0
Intercept Only	46.7	55.0	76.9

model under consideration was the use of Easting, Northing (UTM Zone 56 Southern; EPSG:32756) and Elevation covariates for each $\chi \in \{\tilde{\mu}, \sigma, \xi, \kappa, \eta\}$. We examined possible simplifications for μ, ξ and (σ, κ, η); the parameters σ, κ and η were taken as a unit since they all appear in the Koutsoyiannis et al. (1998) relationship in Eq. (3.1).

Table 4.1 shows that Easting, Northing and Elevation are certainly all needed for μ. There is some suggestion that a simpler model would be sufficient for other parameters. For the shape parameter ξ, the Elevation covariate may not be needed, and for σ, κ and η, the Northing covariate may not be needed. However the evidence is not strong, with DIC differences of only 3.0 and 2.2 respectively from the full model, and therefore we continue with the full model for our main analysis.

The posterior means of the model parameters and their 95% credible intervals are provided in Table 4.2. Previous literature on Bayesian hierarchical models for spatial extremes (Sang and Gelfand, 2010; Davison et al., 2012) suggests that it is not possible to learn from the data simultaneously about the sill δ_χ and range λ_χ parameters, however we found that for $\chi \in \{\tilde{\mu}, \sigma, \xi\}$ our data provided clear information on both. This is probably due to the fact that, unlike previous studies, our model combines rainfall extremes across many accumulation durations. For $\chi \in \{\kappa, \eta\}$ there appears to be some minor identifiability issues. Preliminary analysis with simpler latent variable models suggests that λ_κ and λ_η should be the largest of the range parameters and so we fix these at 15 km. This is consistent with the work of Lehmann et al. (2013). There were also identifiability issues between α and τ. Preliminary analysis suggested that τ should be much smaller than λ_κ or λ_η, and so this was fixed to 5 km.

The regression coefficients associated with Easting, Northing and Elevation generally have credible intervals that either include zero or are close to zero. The main exception is for the scale parameter σ which increases with larger easting values, towards the coast. There is some evidence that η is larger at lower elevations for inland areas, and there is some indication that the shape parameter ξ may be larger at lower elevations. Estimates for the range parameters λ_χ indicate a minimal correlation distance for ξ, with an increasing correlation distance for $\tilde{\mu}$ and σ. The estimate for α indicates a moderate level of spatial dependence in addition to that implied by the latent specifications on the GEV parameters.

A comparison of the estimates in Table 4.2 with those obtained from the simpler (i.e. with $\alpha=1$) latent variable model (not displayed) shows that there is little or no change in the Easting, Northing and Elevation regression coefficient estimates. The range parameters are slightly smaller than those obtained from the latent variable model. The most notable difference is in the intercept terms for the parameters $\tilde{\mu}$ and ξ. The ξ parameter is larger and the $\tilde{\mu}$ parameter is smaller than for the latent variable model. This means that the incorporation of the max-stable process dependence structure has generally led to heavier tailed marginal distributions across the study location.

Fig. 4.1 shows the marginal prior and posterior densities of the parameters λ_σ and λ_ξ. The parameter λ_σ was the largest estimated range parameter, whereas λ_ξ was the smallest. The marginal posterior densities are displayed using a histogram of MCMC samples. The marginal prior densities are the same for both parameters. The data gives clear information on these range parameters, as both posterior densities are far more precise than the prior. The posterior densities display a similar asymmetric shape that is typical of range parameters.

For additional information regarding the Markov chain Monte Carlo procedure and additional model diagnostics, see Appendix B. In particular, Fig. B2 in Appendix B gives posterior estimates of cumulative distribution functions for several sites and several accumulation durations, and compares these estimates with their empirical counterparts.

Fig. 4.2 gives maps of marginal rainfall return levels (i.e. isohyetal maps) for a return period of 50 years, for three different accumulation durations. They are derived using the method given in Appendix B. The maps show that larger return levels are typically estimated near coastal areas. The larger values near the cities of Sydney and Wollongong may also be a consequence of having relatively few coastal observation stations outside of these locations (see Fig. 2.1). The spatial pattern is fairly similar across the different accumulation durations, although at longer durations the return levels appear to be relatively larger in the south of the observation area.

A key model output is the intensity-duration-frequency (IDF) curves that can be derived at any location of interest, either gauged or ungauged. An IDF curve shows intensity plotted against duration for any particular return period frequency, typically displayed on a log-log scale. At gauged locations, the posterior distribution of any return period at a given duration can be calculated using the MCMC samples for the GEV parameters at that site. IDF curves were produced at all $N=182$ sites (see Fig. 2.1) using a 20-year return period, and were compared to those produced using the standard latent variable model. Fig. 4.3 displays the IDF curves for two sites: Blacktown, which is fairly close to Sydney, and Jenolan, which is at a similar latitude but is much further inland. For most sites, the intensity values were similar or larger than those produced from the latent variable model, which is a consequence of the larger intercept on the shape parameter. The Jenolan site in Fig. 4.3 is a typical example of this behavior. However there were some exceptions, such as with Blacktown, where intensities were lower than the latent variable model.

We also examined IDF curves for the model that has no spatial dependence of any type (not shown), but still uses Eq. (3.1) to combine information across accumulation durations. The behavior of these IDF curves relative to the two spatial models can differ for different sites, but they generally give lower intensities than our model. In many cases the model with no spatial dependence yields similar IDF curves to the standard latent variable model.

All IDF curves produced from the models were smooth, which is a consequence of the parametric relationship in Eq. (3.1). IDF curves produced by separate estimation at each accumulation duration generally appear less smooth than those given here. Our explicit model formulation allows us to produce pointwise 95% credible intervals for the IDF curves, as displayed in Fig. 4.3. For gauged sites, the width of these intervals is largely due to the length of the history of observations at that site. For ungauged locations the intervals are typically wider, particularly for those locations that are far away from any gauged site. For most sites the interval widths for our model are slightly larger than for the latent variable model (see Fig. 4.3), and this seems to more genuinely reflect the variability that exists in the rainfall data.

5. Discussion

Our model makes the most of the limited amount of available observations, effectively pooling them together through the spatial process and the combining of rainfall data at different accumulation durations.

We have illustrated the simultaneous modeling of rainfall data over space and accumulation durations, all within a max-stable process framework. This allows the derivations of model inferences, such as IDF curves, at both gauged and ungauged locations. In addition, it can be used to produce spatial inferences such as isohyetal maps (e.g. Fig. 4.2) and depth-area-duration curves. It also provides measures of uncertainty, such as 95% credible intervals, for any feature of interest. Our model estimates a large

Table 4.2
Bayesian estimates of model parameters. Posterior means are given in the center of each list, with 95% credible interval limits given on either side. Parameters without credible interval limits are fixed, namely $\lambda_\kappa = \lambda_\eta = 15$ km and $\tau = 5$ km. Easting, Northing and Elevation covariates were approximately standardized.

	$\bar{\mu}$	$\log(\sigma)$	ξ	$\log(\kappa)$	$\text{logit}(\eta)$
Intercept	1.62, 1.7, 1.8	2.58, 2.68, 2.77	0.41, 0.46, 0.51	−2.58, −2.25, −1.92	0.59, 0.68, 0.78
Easting	−0.31, −0.11, 0.1	0.28, 0.55, 0.81	−0.18, −0.05, 0.08	−0.36, 0.41, 1.2	−0.58, −0.36, −0.14
Northing	0.01, 0.1, 0.19	−0.33, −0.2, −0.08	−0.12, −0.07, −0.01	−0.81, −0.39, 0.02	−0.02, 0.1, 0.21
Elevation	−0.03, 0.1, 0.23	−0.12, 0.05, 0.21	−0.24, −0.16, −0.07	−0.73, −0.21, 0.29	−0.4, −0.27, −0.14
$1/\delta_\chi$	17.9, 24.6, 32.4	16.4, 24.9, 33.7	37, 51.9, 69.7	1.8, 2.5, 3.41	23.1, 30.6, 39.4
λ_χ	2.72, 4.52, 7.12	5.9, 9.76, 16.1	0.5, 2.25, 4.48	15	15
τ=5, α=0.39, 0.42, 0.45					

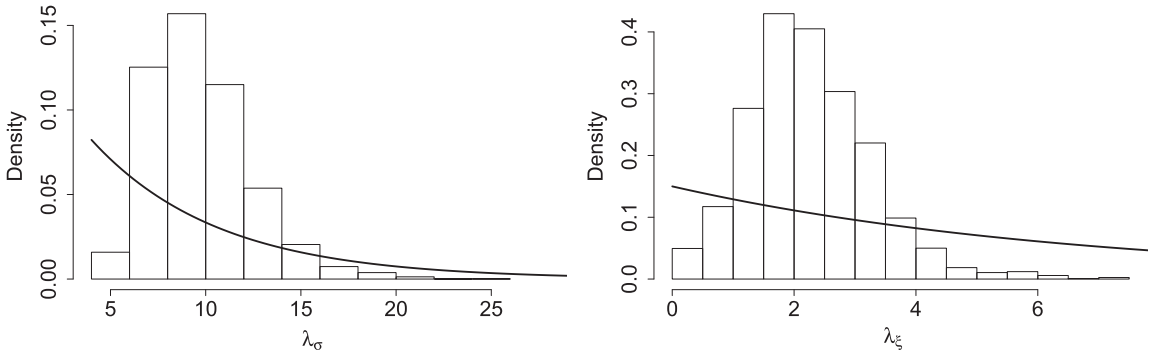

Fig. 4.1. Histograms of MCMC samples for the marginal posterior density of λ_σ (left) and λ_ξ (right). The curves denote the prior density functions.

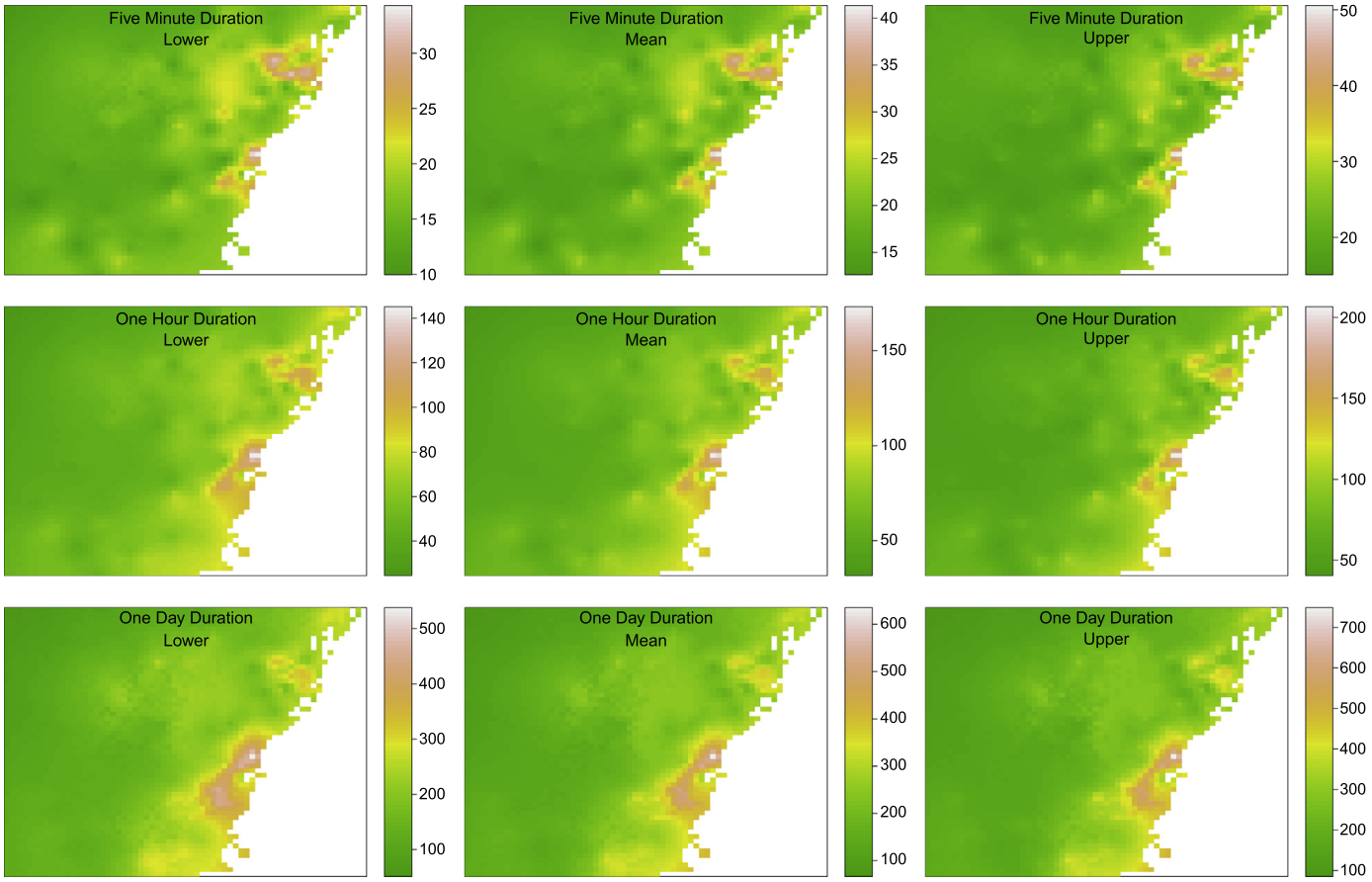

Fig. 4.2. Isohyetal maps of marginal rainfall return levels (mm) for a return period of 50 years (i.e. '1-in-50 year' events), at accumulation durations of five minutes (row one), one hour (row two) and one day (row three). The middle column gives posterior mean estimates. The left and right columns give lower and upper limits of 95% posterior probability intervals. The method of construction is detailed in Appendix B.

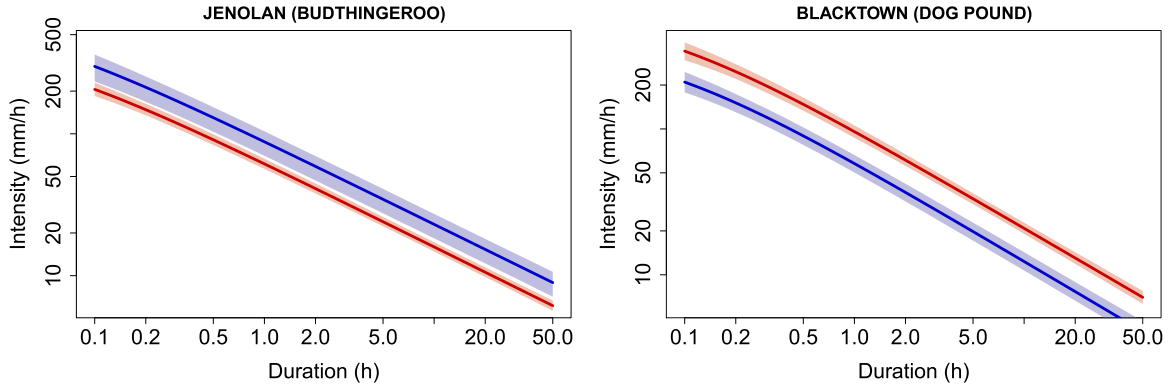

Fig. 4.3. Intensity-Duration-Frequency (IDF) curves on a log-log scale for a 20-year return period, at two different sites. The blue IDF curve is derived from the model of Section 3. The red IDF curve is derived from the standard latent variable model where $\alpha = 1$, so that the spatial dependence arises only through the Gaussian processes on the GEV parameters. The shaded regions represent pointwise 95% credible intervals.

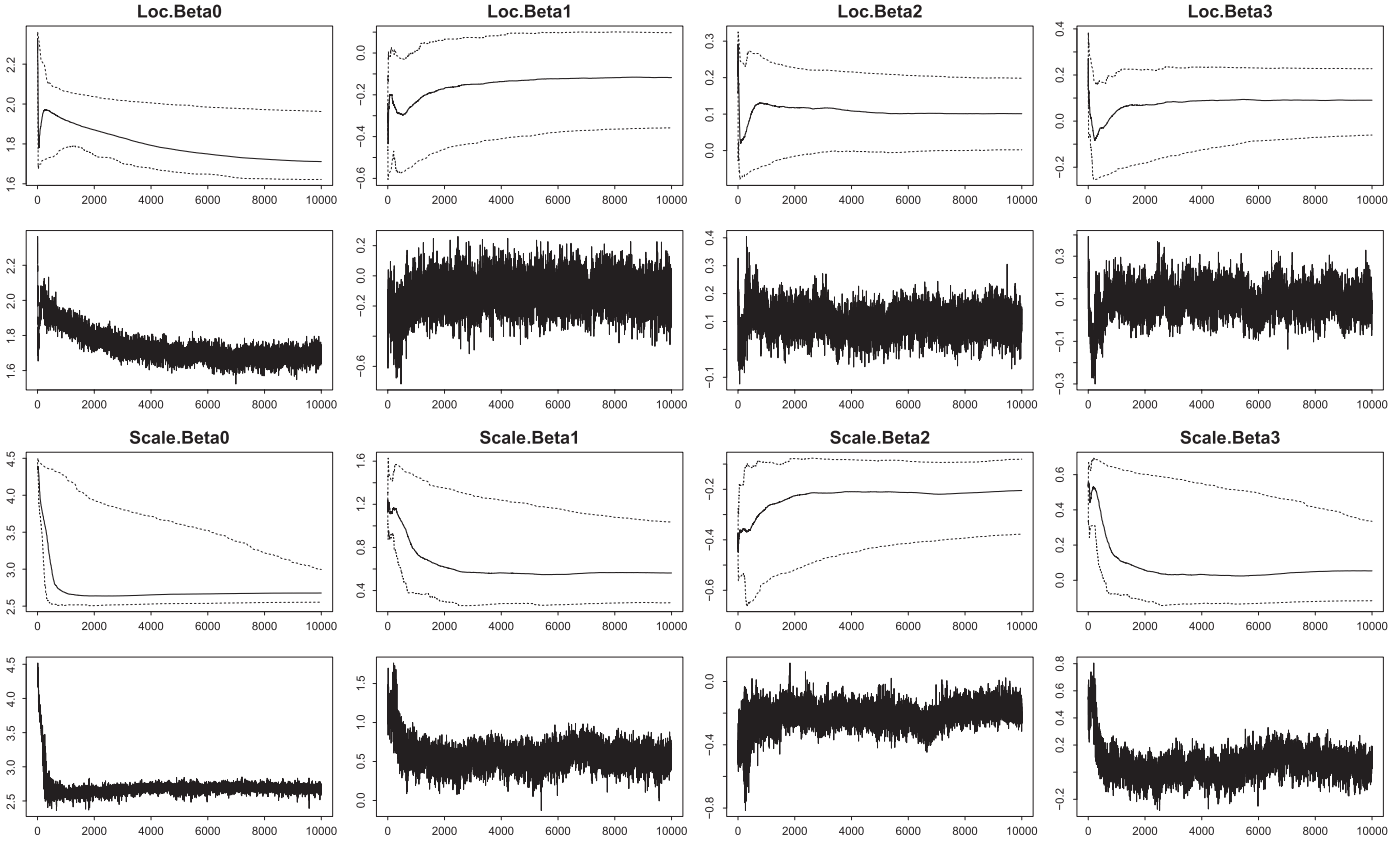

Fig. B1. Cumulative quantile plots (first and third rows) and trace plots (second and forth rows) for the first 10,000 iterations of the Markov chain (including burn-in), for the parameters $\boldsymbol{\beta}_{\tilde{\mu}}$ (first two rows) and $\boldsymbol{\beta}_{\sigma}$ (last two rows). The columns (from left to right) correspond to the intercept and the Easting, Northing and Elevation covariates. The cumulative quantile plots show the quantiles at 2.5%, 50% and 97.5%.

number of parameters from a relatively limited number of observations; although three parameters in the general formulation of Section 3 were fixed, the remainder used relatively uninformative prior distributions.

The hierarchical model that we present here can be extended in a number of ways if this is deemed necessary for any given rainfall dataset. For example, if non-stationarity is present in the annual maxima, then time varying parameters can be incorporated into the framework. Additional covariates such as additional spatial and topographic characteristics can be easily incorporated if such information is available for the study area. Higher-order terms in the GEV parameter specifications could also improve the fit.

The method provided here is more computationally complex and computationally time consuming than the standard latent variable model, which assumes conditional independence. The latent variable model requires one N by N positive definite matrix inversion for each update of the parameter λ_{χ}, for each $\chi \in \{\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta}\}$. In addition, our method requires updating the variables $A_{t,k}$ for each $t = 1, \ldots, T$ and each $k = 1, \ldots, K$ (see Appendix B). Together, the updates of λ_{χ} and $A_{t,k}$ take up almost all the computational time. For large N the updates of λ_{χ} are subject to the "big N problem" of spatial statistics, for which there are a number of proposed solutions (Rue and Held, 2005). For example, Stephenson et al. (2015) use an intrinsic Markov random field representation. The updates of $A_{t,k}$ can be made faster by simply reducing the number of grid locations K, but too few locations will lead to a poor fit. The adjustment of the algorithm to improve both

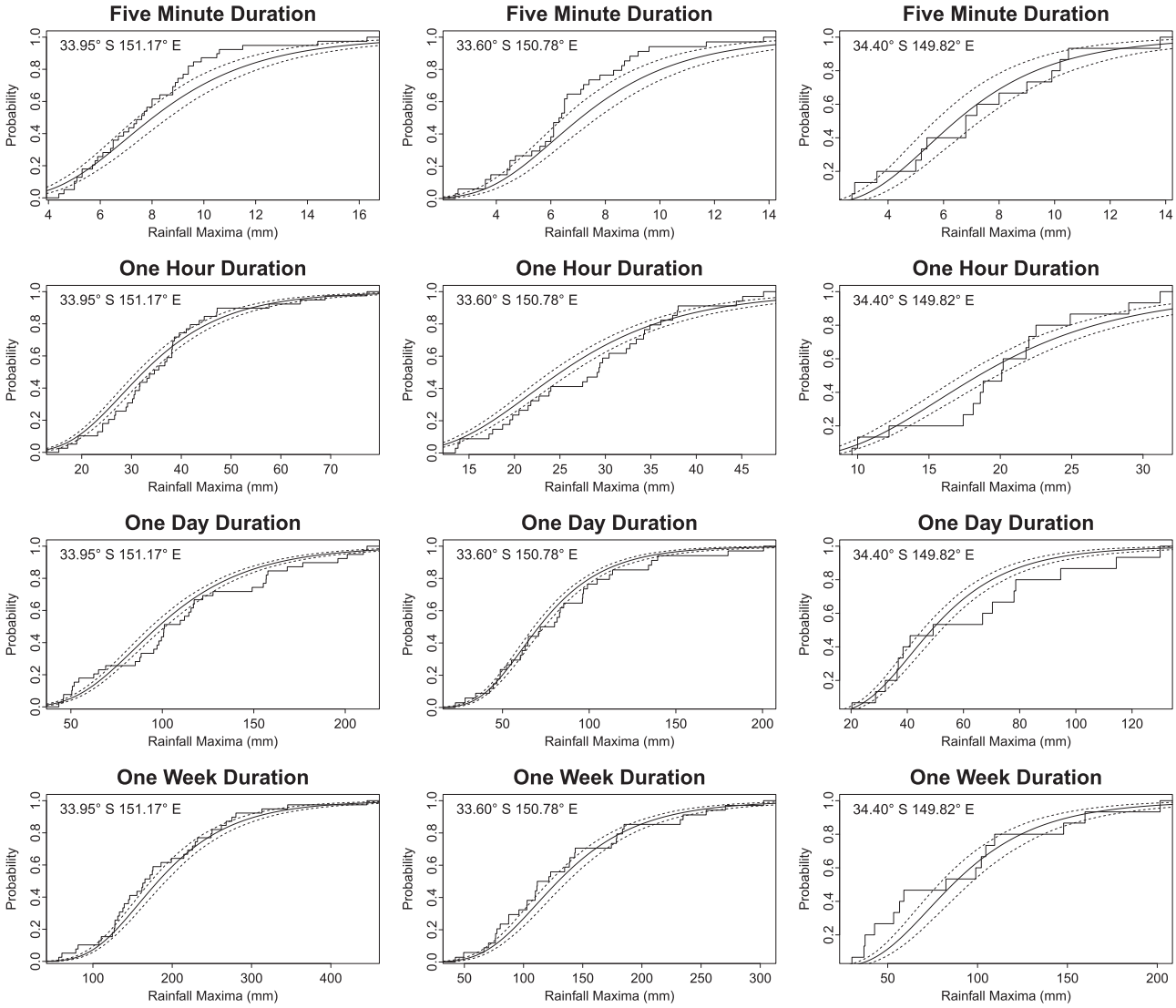

Fig. B2. The smooth curves show fitted cumulative distribution functions for annual rainfall maxima at accumulation durations of five minutes (row one), one hour (row two), one day (row three) and one week (row four), for each of three pluviometer stations. The sites are Sydney Airport AMO (column one), Richmond RAAF (column two) and Taralga Post Office (column three). The locations are as given in the plots. The dashed lines give the limits of pointwise 95% posterior probability intervals. The step functions give the corresponding empirical estimates.

the speed and the mixing properties is an ongoing issue. For further discussion, see Appendix B.

The computational complexity of the proposed method and the standard latent model is far greater than simply assuming independence and fitting each site separately. This independence approach loses the pooling of spatial information, but if inference is only needed on marginal characteristics and the spatial dependence is known to be weak, then simpler methods may be suitable for some applications.

An anonymous referee suggested to us the potential to extend the model so that the dependence structure also depends on the accumulation duration. In particular, we could specify a dependence parameter α_d that depends on the accumulation duration d, in which case we would define $\boldsymbol{A}$ as $\{A_{t,k,d}: t = 1, \ldots, T,\ k = 1, \ldots, K,\ d = 1, \ldots, D\}$, and then Eq. (3.7) would give $\theta_{t,d}(\boldsymbol{s}_i)$ and hence $\tilde{\mu}_t^*(\boldsymbol{s}_i)$ would become $\tilde{\mu}_{t,d}^*(\boldsymbol{s}_i)$. It would also be possible to impose a smoothness constraint on α_d over the duration d. The disadvantage of this approach is that $\boldsymbol{A}$ becomes a three dimensional array, which could lead to impractical computation times for anything other than a small number of sites.

Appendix A. Positive stable distribution

Let X be distributed according to the positive stable distribution, with index $\alpha \in (0, 1]$. When $\alpha = 1$, the positive stable distribution is degenerate at one. When $\alpha \in (0, 1)$, the density function has support on the positive real line and can be expressed as

$$f(x) = \int_0^1 \frac{\alpha x^{-1/(1-\alpha)}}{1-\alpha} h(u) \exp\left[-h(u) x^{-\alpha/(1-\alpha)}\right] \mathrm{d}u \tag{A.1}$$

where

$$h(u) = \left[\frac{\sin(\alpha\pi u)}{\sin(\pi u)}\right]^{1/(1-\alpha)} \frac{\sin[(1-\alpha)\pi u]}{\sin(\alpha\pi u)} \tag{A.2}$$

for $0 < u < 1$, a representation first given with minor typographical errors by Ibragimov and Chernin (1959). It can be shown that the integrand in Eq. (A.1) represents a bivariate density function with positive stable and uniformly distributed marginal distributions. Our use of positive stable random variables derives from the simple form of the Laplace transform of the density function,

which is given by

$$\int_0^\infty f(x)\exp(-xt)\,dx = \exp(-t^\alpha). \tag{A.3}$$

Appendix B. Model estimation and diagnostics

B.1. Model estimation

Model estimation is achieved through standard Markov chain Monte Carlo (MCMC) simulations (Hastings, 1970) applied to the posterior distribution of the parameters. The posterior density is proportional to Eq. (3.12), with $\pi(\boldsymbol{A}, \boldsymbol{B}|\alpha)$ replaced by $\pi(\boldsymbol{A}, \boldsymbol{B}|\alpha)$, as discussed below that equation. This gives the expression

$$L(\boldsymbol{Y}|\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta}, \boldsymbol{A}, \alpha, \tau)\pi(\boldsymbol{A}, \boldsymbol{B}|\alpha)\left\{\prod_{\chi}\pi(\boldsymbol{\chi}|\boldsymbol{\phi}_\chi)\pi(\boldsymbol{\phi}_\chi)\right\}\pi(\alpha, \tau). \tag{B.1}$$

The likelihood function is given by

$$L(\boldsymbol{Y}|\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta}, \boldsymbol{A}, \alpha, \tau) = \prod_{t=1}^{T}\prod_{d=1}^{D}\prod_{i=1}^{N}\left\{\frac{1}{\sigma^*_{t,d}(\boldsymbol{s}_i)}g(Y_{t,d}(\boldsymbol{s}_i))^{\xi^*(\boldsymbol{s}_i)+1}\exp[-g(Y_{t,d}(\boldsymbol{s}_i))]\right\}, \tag{B.2}$$

with

$$g(Y_{t,d}(\boldsymbol{s}_i)) = \left[1 + \xi^*(\boldsymbol{s}_i)\left(\frac{Y_{t,d}(\boldsymbol{s}_i)}{\sigma^*_{t,d}(\boldsymbol{s}_i)} - \tilde{\mu}^*_t(\boldsymbol{s}_i)\right)\right]_+^{-1/\xi^*(\boldsymbol{s}_i)} \tag{B.3}$$

$$= \theta_t(\boldsymbol{s}_i)^{1/\alpha}\left[1 + \xi(\boldsymbol{s}_i)\left(\frac{Y_{t,d}(\boldsymbol{s}_i)}{\sigma_d(\boldsymbol{s}_i)} - \tilde{\mu}(\boldsymbol{s}_i)\right)\right]_+^{-1/\alpha\xi(\boldsymbol{s}_i)} \tag{B.4}$$

for $\xi^*(\boldsymbol{s}_i) \neq 0$, where $\tilde{\mu}^*_t(\boldsymbol{s}_i)$, $\sigma^*_{t,d}(\boldsymbol{s}_i)$ and $\xi^*(\boldsymbol{s}_i)$ are defined in Section 3 and where $\lambda_+ = \max(\lambda, 0)$. If $\xi^*(\boldsymbol{s}_i) = 0$ then Eq. (B.2) is defined in the limit as $\xi^*(\boldsymbol{s}_i) \to 0$.

The densities $\pi(\boldsymbol{\chi}|\boldsymbol{\phi}_\chi)$ for $\boldsymbol{\chi} \in \{\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta}\}$ are multivariate normal densities as defined in Section 3, where $\boldsymbol{\phi}_\chi = (\boldsymbol{\beta}_\chi, \delta_\chi, \lambda_\chi)$. The density of the variables $(\boldsymbol{A}, \boldsymbol{B})$ is given by

$$\pi(\boldsymbol{A}, \boldsymbol{B}|\alpha) = \prod_{t=1}^{T}\prod_{k=1}^{K}\frac{\alpha A_{t,k}^{-1/(1-\alpha)}}{1-\alpha}h(B_{t,k})\exp\left[-h(B_{t,k})A_{t,k}^{-\alpha/(1-\alpha)}\right], \tag{B.5}$$

with $h(\cdot)$ as defined in Eq. (A.2).

We define $\pi(\alpha, \tau)$ using independent prior distributions, taking $\alpha \sim \text{Unif}(0, 1)$ and $\tau \sim \text{Gam}(\psi_\tau, \gamma_\tau)$. For the $\boldsymbol{\phi}_\chi$ parameters we take $\delta_\chi \sim \text{InvGamma}(\psi_{\delta_\chi}, \gamma_{\delta_\chi})$, $\lambda_\chi \sim \text{Gam}(\psi_{\lambda_\chi}, \gamma_{\lambda_\chi})$ and $\boldsymbol{\beta}_\chi \sim \text{MVN}(\boldsymbol{0}, 100I)$ where I is the identity matrix. Following Banerjee et al. (2004), we use weakly informative priors for δ_χ and λ_χ, setting the hyperparameters so that the distributions cover a reasonable range of values as determined by an exploratory analysis.

The Markov chain Monte Carlo simulations are performed using standard Metropolis-Hastings proposals (Hastings, 1970). We individually update the spatial dependence parameters (τ, α), the Generalized Extreme Value parameters $(\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta})$, the range parameters λ_χ and the variables $(\boldsymbol{A}, \boldsymbol{B})$ using either normal, log-normal or logit-normal proposal distributions. For example, if $A^*_{t,k}$ is the log-normal proposal for $A_{t,k}$, then the acceptance ratio simplifies to

$$\frac{A^*_{t,k}}{A_{t,k}} \times \frac{A^{*-1/(1-\alpha)}_{t,k}\exp\left[-h(B_{t,k})A^{*-\alpha/(1-\alpha)}_{t,k}\right]}{A^{-1/(1-\alpha)}_{t,k}\exp\left[-h(B_{t,k})A^{-\alpha/(1-\alpha)}_{t,k}\right]} \times \prod_{d=1}^{D}\prod_{i=1}^{N}\frac{\theta^*_t(\boldsymbol{s}_i)^{1/\alpha}\exp\left[-\theta^*_t(\boldsymbol{s}_i)^{1/\alpha}\tilde{g}(Y_{t,d}(\boldsymbol{s}_i))\right]}{\theta_t(\boldsymbol{s}_i)^{1/\alpha}\exp\left[-\theta_t(\boldsymbol{s}_i)^{1/\alpha}\tilde{g}(Y_{t,d}(\boldsymbol{s}_i))\right]},$$

where the first term is the proposal ratio, the second term derives from Eq. (B.5), and the third term derives from Eq. (B.2), with

$$\tilde{g}(Y_{t,d}(\boldsymbol{s}_i)) = \left[1 + \xi(\boldsymbol{s}_i)\left(\frac{Y_{t,d}(\boldsymbol{s}_i)}{\sigma_d(\boldsymbol{s}_i)} - \tilde{\mu}(\boldsymbol{s}_i)\right)\right]_+^{-1/\alpha\xi(\boldsymbol{s}_i)}, \tag{B.6}$$

and

$$\theta^*_t(\boldsymbol{s}_i)^{1/\alpha} = \theta_t(\boldsymbol{s}_i)^{1/\alpha} + w_k(\boldsymbol{s}_i)^{1/\alpha}(A^*_{t,k} - A_{t,k}). \tag{B.7}$$

The parameters $\boldsymbol{\beta}_\chi$ and δ_χ have closed form conditional posterior distributions, and so our proposals for these parameters simulate from the posterior directly. In particular, the conditional posterior distributions are given by

$$\boldsymbol{\beta}_\chi|\boldsymbol{\chi}, \delta_\chi, \lambda_\chi \sim \text{MVN}(V_\chi X_\chi^T \tilde{\Sigma}_\chi^{-1}\boldsymbol{\chi}, V_\chi)$$

$$\delta_\chi|\boldsymbol{\chi}, \boldsymbol{\beta}_\chi, \lambda_\chi \sim \text{InvGamma}(N/2 + \psi_{\delta_\chi}, S_\chi/2 + \gamma_{\delta_\chi}),$$

where $V_\chi = (X_\chi^T\tilde{\Sigma}_\chi^{-1}X_\chi + I/100)^{-1}$ and $S_\chi = (\boldsymbol{\chi} - X_\chi\boldsymbol{\beta}_\chi)^T\tilde{\Sigma}_\chi^{-1}(\boldsymbol{\chi} - X_\chi\boldsymbol{\beta}_\chi)$, with

$$[\tilde{\Sigma}_\chi]_{ij} = \exp(-\|\boldsymbol{s}_i - \boldsymbol{s}_j\|/\lambda_\chi), \quad i, j = 1, \ldots, N. \tag{B.8}$$

B.2. Inference at ungauged sites

Inferences can be made at any arbitrary ungauged site $\boldsymbol{s}$ in addition to the N gauged sites $\boldsymbol{s}_i$ for $i = 1, \ldots, N$. This is required for e.g. the isohyetal maps in Fig. 4.2, which were produced as follows. Take for example $\boldsymbol{\xi} = [\xi(\boldsymbol{s}_1), \ldots, \xi(\boldsymbol{s}_N)]$. Each iteration of the Markov chain simulation produces a realization of $\boldsymbol{\xi}$ and of $\boldsymbol{\beta}_\xi$. Therefore $\boldsymbol{\xi} - X_\xi\boldsymbol{\beta}_\xi$ gives N values that are modeled as a zero mean Gaussian process. We construct a new matrix X^*_ξ which now contains the Easting, Northing and Elevation covariates for a grid of ungauged locations. Hence X^*_ξ has the same number of columns as X_ξ, but has as many rows as there are points in the grid. We then use the realizations of δ_ξ and λ_ξ to simulate a conditional zero mean Gaussian process over the ungauged locations, conditioning on the N values $\boldsymbol{\xi} - X_\xi\boldsymbol{\beta}_\xi$ at sites $\boldsymbol{s}_i$ for $i = 1, \ldots, N$, and then adding the mean $X^*_\xi\boldsymbol{\beta}_\xi$. This generates $\xi(\boldsymbol{s})$ on a grid of arbitrary sites, where $\xi(\boldsymbol{s}) = \xi(\boldsymbol{s}_i)$ whenever $\boldsymbol{s} = \boldsymbol{s}_i$. A similar process is performed for each parameter vector $(\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta})$ and for each iteration of the Markov chain. Posterior means and posterior probability intervals can then be calculated in the usual manner, either for individual parameters or for any function of the parameters, such as return levels for any return period and for any accumulation duration.

B.3. MCMC output and model diagnostics

The DIC values given in Table 4.1 of Section 4 are given by

$$2\log L(\boldsymbol{Y}|\bar{\tilde{\boldsymbol{\mu}}}, \bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\xi}}, \bar{\boldsymbol{\kappa}}, \bar{\boldsymbol{\eta}}, \bar{\boldsymbol{A}}, \bar{\alpha}, \bar{\tau}) - \frac{4}{M}\sum_{j=1}^{M}\log L(\boldsymbol{Y}|\tilde{\boldsymbol{\mu}}_j, \boldsymbol{\sigma}_j, \boldsymbol{\xi}_j, \boldsymbol{\kappa}_j, \boldsymbol{\eta}_j, \boldsymbol{A}_j, \alpha_j, \tau_j), \tag{B.9}$$

where the sum is over $M = 5\,000$ Markov chain Monte Carlo iterations obtained following a burn-in of 1000 iterations, and where e.g. $\bar{\boldsymbol{\sigma}}$ is the componentwise mean $\sum_{j=1}^{M}\boldsymbol{\sigma}_j/M$. This is the standard DIC definition for hierarchical models, where $\boldsymbol{A}$ is not

integrated out and is therefore included in the effective number of parameters. The terms $\pi(\boldsymbol{A}, \boldsymbol{B}|\alpha)$ and $\prod_{\chi} \pi(\chi|\boldsymbol{\phi}_{\chi})\pi(\boldsymbol{\phi}_{\chi})$ in Eq. (B.1) correspond to lower levels of the hierarchy and are therefore excluded.

The results presented in Section 4 were obtained from a Markov chain simulated for a total of 25,000 iterations, with the first 5000 iterations discarded. No thinning was used. In the most general case, each iteration of the Markov chain provides realizations of the site specific parameters $(\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta})$, the latent matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, the dependence parameters (α, τ) and finally the Gaussian process parameters $(\boldsymbol{\beta}_{\chi}, \delta_{\chi}, \lambda_{\chi})$ for each $\chi \in \{\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\kappa}, \boldsymbol{\eta}\}$. Generally speaking, the site specific parameters and the latent matrices are well-behaved, although it can be difficult to obtain reasonable acceptance rates for all the parameters in $\boldsymbol{A}$. The parameters $(\boldsymbol{\beta}_{\chi}, \delta_{\chi}, \lambda_{\chi})$ and (α, τ) are more problematic. Even though $(\boldsymbol{\beta}_{\chi}, \delta_{\chi})$ are simulated directly and therefore acceptance rates are not an issue, some of the intercept parameters and $(\delta_{\chi}, \lambda_{\chi})$ were seen to have high auto-correlations and to be slow to converge. This is shown in Fig. B1, which displays cumulative quantile plots and trace plots for $\boldsymbol{\beta}_{\tilde{\mu}}$ and $\boldsymbol{\beta}_{\sigma}$, for the first 10,000 iterations (including burn-in). The intercept parameter for $\tilde{\mu}$ is particularly slow to converge, and similar issues arise for (α, τ). This resulted in our fixing of the parameters λ_{κ}, λ_{η} and τ in Section 4. The application of the Markov chain convergence diagnostic of Heidelberger and Welch (1983) showed that 85.8% of site specific parameters passed the test, and 93.1% of the remaining parameters passed the test. Almost all parameters that passed the convergence diagnostic test also passed the Heidelberger and Welch (1983) run length diagnostic. For more information on output diagnostics for Markov chains, see Brooks and Roberts (1998) and Cowles and Carlin (1996). For Markov chain output diagnostics and Bayesian model diagnostics in the context of extremes modeling, see Stephenson (2016).

Fig. B2 is a model diagnostic which compares empirical cumulative distribution estimates against fitted cumulative distribution functions derived from the model for four different accumulation durations, for each of three different sites. The durations are five minutes, one hour, one day and one week. The pluviometer stations are Sydney Airport AMO (33.95°S, 151.17°E), Richmond RAAF (33.60°S, 150.78°E) and Taralga Post Office (34.40°S, 149.82°E). Sydney Airport AMO is near the upper red cross in Fig. 2.1, and Richmond RAAF is approximately 60 km to the north-west of this. Taralga Post Office is at a higher elevation, about 125 km to the west of Wollongong. There are less available data at Taralga, and there are fewer neighboring sites, and hence the confidence intervals here are wider. The plots indicate that the model provides a satisfactory fit to the data. Similar results are achieved for other pluviometer stations and other durations.

References

Apputhurai, P., Stephenson, A.G., 2013. Spatiotemporal hierarchical modelling of extreme precipitation in Western Australia using anisotropic Gaussian random fields. Environ. Ecol. Stat. 20 (4), 667–677.

Banerjee, S., Carlin, B.P., Gelfand, A.E., 2004. Hierarchical Modelling and Analysis for Spatial Data. Chapman & Hall, London.

Brooks, S.P., Roberts, G.O., 1998. Convergence assessment techniques for Markov chain Monte Carlo. Stat. Comput. 8, 319–335.

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

Cooley, D., Nychka, D., Naveau, P., 2007. Bayesian spatial modeling of extreme precipitation return levels. J. Am. Stat. Assoc. 102, 824–840.

Cowles, M.K., Carlin, B.P., 1996. Markov chain Monte Carlo convergence diagnostics: a comparative review. J. Am. Stat. Assoc. 91, 883–904.

Davison, A.C., Padoan, S.A., Ribatet, M., 2012. Statistical modelling of spatial extremes. Stat. Sci. 27, 161–186.

Gaetan, C., Grigoletto, M., 2007. A hierarchical model for the analysis of spatial rainfall extremes. J. Agric. Biol. Environ. Stat. 12, 434–449.

Garcia-Bartual, R., Schneider, M., 2001. Estimating maximum expected short-duration rainfall intensities from extreme convective storms. Phys. Chem. Earth PartB: Hydrol. Oceans Atmos. 26 (9), 675–681.

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57, 97–109.

Heidelberger, P., Welch, P.D., 1983. Simulation run length control in the presence of an initial transient. Oper. Res. 31, 1109–1144.

Hosking, J.R.M., Wallis, J.R., 1995. Regional Frequency Analysis: An Approach Based on L-Moments. Cambridge University Press, Cambridge.

Ibragimov, I.A., Chernin, K.E., 1959. On the unimodality of stable laws. Theory of Probability and Its Applications, 4, pp. 417–419.

Koutsoyiannis, D., Kozonis, D., Manetas, A., 1998. A mathematical framework for studying rainfall-intensity-duration-frequency relationships. J. Hydrol. 206, 118–135.

Leadbetter, M.R., Lindgren, G., Rootzén, H., 1983. Extremes and Related Properties of Random Sequences and Series. Springer-Verlag, New York.

Lehmann, E.A., Phatak, A., Soltyk, S., Chia, J., Lau, R., Palmer, M., 2013. Bayesian hierarchical modelling of rainfall extremes. In: Piantadosi, J., Anderssen, R., Boland, J. (Eds.), MODSIM 2013. Modelling & Simulation Society of Australia & New Zealand Inc., Adelaide, Australia, pp. 2806–2812.

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

Muller, A.H., Barco, J.N., Lang, M., 2008. Bayesian comparison of different rainfall depth-duration-frequency relationships. Stoch. Environ. Res. Risk Assess. 22 (1), 33–46.

Nadarajah, S., Anderson, C.W., Tawn, J.A., 1998. Ordered multivariate extremes. R. Stat. Soc. B 60, 473–496.

Padoan, S., Ribatet, M., Sisson, S., 2010. Likelihood-based inference for max-stable processes. J. Am. Stat. Assoc. 105, 263–277.

Reich, B.J., Shaby, B.J., 2012. A hierarchical max-stable spatial model for extreme precipitation. Ann. Appl. Stat. 6 (4), 1430–1451.

Ribatet, M., Cooley, D., Davison, A., 2012. Bayesian inference from composite likelihood, with an application to spatial extremes. Stat. Sin. 22, 813–845.

Rue, H., Held, L., 2005. Gaussian Markov Random Fields: Theory and Applications. Chapman & Hall/CRC Press, London.

Sang, H., Gelfand, A.E., 2008. Hierarchical modeling for extreme values observed over space and time. Environ. Ecol. Stat. 16, 407–426.

Sang, H., Gelfand, A.E., 2010. Continuous spatial process models for spatial extreme values. J. Agric. Biol. Environ. Stat. 15, 49–65.

Schlather, M., 2002. Models for stationary max-stable random fields. Extremes 5, 33–44.

Schliep, E.M., Cooley, D., Sain, S.R., Hoeting, J.A., 2010. A comparison study of extreme precipitation from six different regional climate models via spatial hierarchical modeling. Extremes 13, 219–239.

Smith, E.L., Stephenson, A.G., 2009. An extended Gaussian max-stable process model for spatial extremes. J. Stat. Plan. Inference 139, 1266–1275.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measures of model complexity and fit. J. R. Stat. Soc. B 64, 583–639 (with discussion).

Stephenson, A.G., 2009. High-dimensional parametric modelling of multivariate extreme events. Aust. N. Z. J. Stat. 51, 77–88.

Stephenson, A.G., 2016. Bayesian inference for extreme value modelling. In: Dey, D., Yan, J. (Eds.), Extreme Value Modeling and Risk Analysis: Methods and Applications. Chapman & Hall/CRC, Boca Raton, FL, pp. 257–280.

Stephenson, A.G., Shaby, B.A., Reich, B.J., Sullivan, A.L., 2015. Estimating spatially varying severity thresholds of the forest fire danger rating system using max-stable extreme event modelling. J. Appl. Meteorol. Climatol. 54, 395–407.

Tawn, J.A., 1990. Modelling multivariate extreme value distributions. Biometrika 77, 245–253.

Van de Vyver, H., 2015. Bayesian estimation of rainfall intensity-duration-frequency relationships. J. Hydrol. 529 (3), 1451–1463.