10 Dynamic and spatial modelling of block maxima extremes

GABRIEL HUERTA AND
GLENN A. STARK

10.1 Introduction

Figure 10.1 shows the monthly maxima of precipitation near the Maíquetia-Simon Bolivar airport near Caracas, Venezuela within the period 1960–1999. The data only considers measurements at one site and has critical importance in extreme value analysis due to the catastrophic events that occur near this area at the end of the year 1999. For example [5] looked carefully at a Bayesian analysis of the annual maxima rainfall values at the same site, based on models that fully account for parameter uncertainties and non-stationarity. Also [16], analysed the monthly rainfall maxima via dynamic regressions as in [29], [28] and in connection with a climatological index known as the *North Atlantic Oscillation* (NAO). Here we consider these rainfall observations as a starting point to study extreme events via the Generalized Extreme Value (GEV) distribution. Two key questions that arise from Figure 10.1 are: (1) How do we characterize these rainfall events? and (2) How do we assess for the non-stationary behaviour that is shown in the data?

10.1.1 GEV distribution and likelihood

Let $y_{m,1}, y_{m,2}, \ldots, y_{m,n}$ be samples of extremes from m independent observations (i.e. block maxima), where m is the number of observations in each block and n is the number of blocks. For the precipitation maxima of Figure 10.1, a uniform block size across time was considered to produce the observations. In applications of block-maxima values it is typical to assume that $y_{m,i}, i = 1, 2, \ldots, n$, are independent and arise from a common GEV distribution as in [4] and [1]. If $y \sim GEV(\mu, \sigma, \xi)$, the cumulative distribution function for y is:

$$H(y) = \exp\left\{-\left[1 + \xi\left(\frac{y-\mu}{\sigma}\right)\right]_+^{-1/\xi}\right\} \tag{10.1}$$

where $-\infty < \mu < \infty$ is a location parameter, $\sigma > 0$ is a scale parameter and $-\infty < \xi < \infty$ is a shape parameter. The $+$ sign denotes the positive part of the argument so the support for $H(y)$

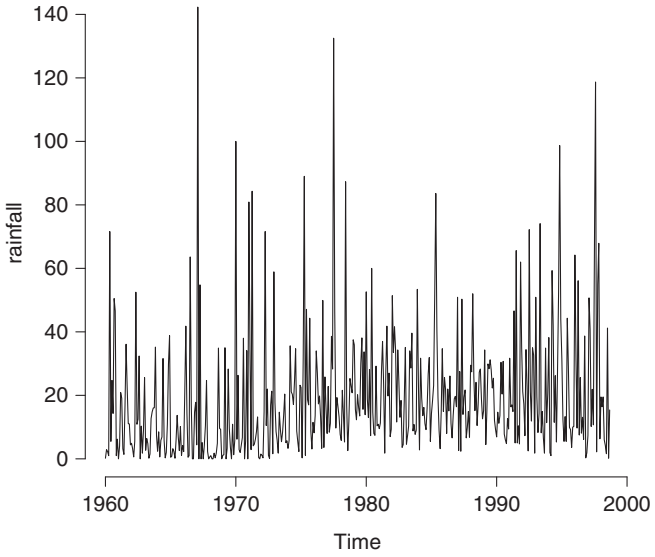

Figure 10.1 Monthly maxima of precipitation in Venezuela.

is given by the set $\{y : 1 + \xi(\frac{y-\mu}{\sigma}) > 0\}$. It is well known that different values of ξ imply different tail behaviours or *domains of attraction* for $H(y)$, namely Gumbel ($\xi \to 0$), Fréchet ($\xi > 0$) and Weibull ($\xi < 0$). $H(y)$ arises through a limiting argument for block maxima in the *Extremal Type* or *Fisher-Tippet* theorem as in [10] and also presented in [4] and [1], so if the block size is large enough, one may assume that the observations $y_{m,i}$ follow a GEV distribution. In particular, if we drop the dependence on the block size so that $y_i = y_{m,i}$, the log-likelihood function based on n observations is

$$l(\mu, \sigma, \xi | \{y_t\}_{t=1}^n) = -n \log \sigma - (1 + 1/\xi) \sum_{i=1}^{n} \log\{1 + \xi(y_i - \mu)/\sigma\}$$
$$- \sum_{i=1}^{n} \{1 + \xi(y_i - \mu)/\sigma\}^{-1/\xi}. \quad (10.2)$$

10.1.2 Bayesian inference on the GEV

From a Bayesian point-of-view inferences on (μ, σ, ξ) can be directly obtained with Markov chain Monte Carlo (MCMC) methods based on a Gibbs sampling approach as in [13] with embedded Metropolis–Hastings (M-H) steps as in [15]. For instance, a prior distribution $p(\mu, \sigma, \xi)$ can be induced through a *trivariate* normal distribution on $(\mu, log(\sigma), \xi)$ which includes the case of an independence prior.

As described in [6], beta distributions for probability ratios or gamma distributions for quantile difference could alternatively be used to elicit $p(\mu, \sigma, \xi)$. In particular, the quantile-difference priors are interesting since the $1 - p$ quantile of a GEV distribution has a closed form in terms of the three parameters, $z_p = \mu - \frac{\sigma}{\xi}[1 - \{-log(1 - p)\}^{-\xi}]$.

The posterior distribution for α can be sampled with random walk proposals on each of the parameters as presented in [4] and [1],

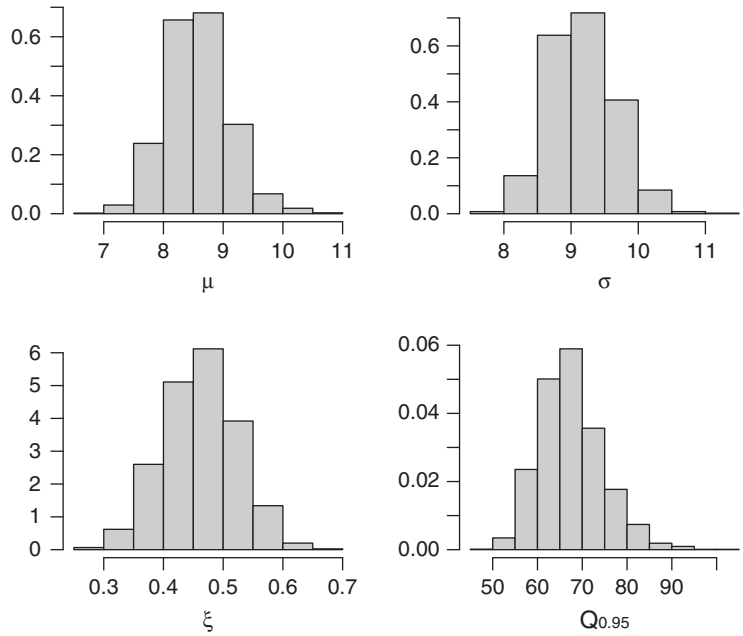

Figure 10.2 Marginal posterior distribution for (μ, σ, ξ) and 95% quantile of the GEV distribution.

$$log(\sigma^*) = log(\sigma^{(i)}) + \nu_\sigma \epsilon_1 \tag{10.3}$$

$$\mu^* = \mu^{(i)} + \nu_\mu \epsilon_2 \tag{10.4}$$

$$\xi^* = \xi^{(i)} + \nu_\xi \epsilon_3 \tag{10.5}$$

where $\epsilon_j \sim N(0,1); j = 1,2,3$ and ν_σ, ν_μ, ν_ξ denote the proposal tuning parameters. The proposed values are then accepted or rejected as iterations of the MCMC are performed. To illustrate this method, Figure 10.2 shows histograms of posterior samples for μ, σ, ξ and $z_{0.05}$ corresponding to the data of Figure 10.1 and via this hybrid Gibbs sampling/Metropolis–Hastings method.

It is interesting to note that quantile estimation can also be achieved through the predictive distribution,

$$p(y^f|y) = \int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{-\infty}^{\infty} f(y^f|\mu,\sigma,\xi)p(\mu,\sigma,\xi|y)d\mu\, d\sigma\, d\xi \tag{10.6}$$

where y^f denotes a future observation. Here $f(y^f|\mu,\sigma,\xi)$ represents the probability density function for this future observation based on the GEV distribution. Through the *Method of Composition* it is possible to obtain samples of y^f from the predictive distribution. An empirical quantile based on these samples approximates the value y^* such that $P[Y^f \leq y^*|y] = 1 - p$, which fully takes into account all of the model parametric uncertainties. In fact, for the maximum monthly rainfall data of Figure 10.1, the 99% quantile based on the predictive distribution $(p = 0.01)$ is 162.87. On the other hand, if the GEV distribution is fitted with Maximum Likelihood Estimation (MLE), the 99%

quantile is estimated as 152.3. Furthermore, the posterior mean of the GEV distribution quantile, $z_{0.01}$, is 157.35.

Bayesian approaches for extreme values had been extensively studied from the beginning of Bayesian computation with MCMC and other numerical approximations. The paper by [25] is one of the first to illustrate the Bayesian modelling of extremes within the GEV framework. Also [7] offers some developments of Bayesian approaches to extreme value theory with applications to modelling areal rainfall extremes. A thorough review of the topic is offered by [6] where the authors studied the Bayesian approach from a variety of aspects, including how priors can be better elicited for extreme value data and how well a complete Bayesian analysis, that includes predictive-quantile estimation, performs compared to a pure likelihood based analysis as briefly illustrated here with the rainfall data at the Maíquetia/Simon Bolivar airport. The R-package *evdbayes* described in [26] and [27] implements the techniques in [6] including quantile-based prior specifications and peaks-over-threshold analysis. However, the modelling described in these references is mainly restricted to *constant* parameters or to parameters that may have deterministic trends in time.

10.2 Time-varying models for the GEV distribution

10.2.1 Introduction

To account for the type of non-stationarities that are present in the data of Figure 10.1, we can impose a time-dependent structure on any set of parameters of the GEV distribution. Here, we emphasize on a GEV distribution with a time-varying location parameter,

$$H(y_t) = \exp\left\{-\left[1+\xi\left(\frac{y_t-\mu_t}{\sigma}\right)\right]_+^{-1/\xi}\right\}, \quad t = 1, 2, \ldots n \tag{10.7}$$

If a deterministic function is chosen to model time changes in μ_t, this can be expressed in a generalized form as

$$\mu_t = g(X^T\beta) \tag{10.8}$$

where g is a specific link function, β is a vector of parameters and X^T is a vector of covariates that involves time. In particular μ_t can follow a linear trend, $\mu_t = \beta_0 + \beta_1 t$, a higher degree polynomial such as $\mu_t = \beta_0 + \beta_1 t + \beta_2 t^2$ or a trend/seasonal model, $\mu_t = \beta_0 + \beta_1 t + \beta_2 S(t)$, where $S(t)$ represents the seasonal component. For example, [4] discusses extensively these various models. On the other hand, μ_t can be treated as a stochastic process that depends on time through a hierarchical specification.

10.2.2 Dynamic linear model

We consider *Dynamic Linear Models* (DLMs) or state-space models as in [28] and [29] as our main choice for non-stationary modelling of the location parameter of the GEV distribution. We focus on DLMs to assess whether any short-term changes can occur in the extremal-type distribution. If $z_t, t = 1, 2, \ldots$ represents a vector of observations of dimension r at time t, following the notation in [29], a DLM for z_t is specified as,

$$z_t = F_t'\theta_t + v_t, \quad v_t \sim N(0, V_t) \tag{10.9}$$

$$\theta_t = G_t\,\theta_{t-1} + w_t, \quad w_t \sim N(0, W_t) \tag{10.10}$$

$$\theta_0 | D_0 \sim N(m_0, C_0) \tag{10.11}$$

where F_t' is assumed to be a known $(r \times n)$ regression matrix and G_t is assumed as a known $(n \times n)$ state matrix. Equation 10.9 defines the observation equation of the DLM and the variance of the observation error is given by the $r \times r$ matrix V_t. Equation 10.10 is called the system or evolution equation and W_t is the $n \times n$ variance–covariance matrix of the evolution error. The errors ν_t and w_t are generally assumed to be mutually independent. To obtain Bayesian inference on the state vector of the DLM, a prior distribution on the initial state is needed and is given by equation 10.11, where m_0 defines the forecaster's initial belief about the level θ_0 and C_0 is the associated measure of uncertainty. At the lack of any true prior information, we will adopt a non-informative prior where the mean level is fixed to an arbitrary point and C_0 assumes a large value. The quadruple (F_t, G_t, V_t, W_t) characterizes completely the DLM, so different quadruples define different subsets of the general class of DLMs.

10.2.3 DLMs and the GEV distribution

A special case of the quadruple is given by $(1, 1, V, W)$, which is referred to as a *first-order polynomial* DLM in [29]. DLMs had been used by [12] and [16] for time-varying extreme value models. [12] outlines a semi parametric approach for smoothing extremes with applications to athletic records and temperature data. On the other hand, [16] consider DLMs for assessing dynamic trends and space–time structures for extreme values of ozone levels. Here we follow [16] and focus our presentation for the case of a time-varying location parameter.

Let $\{y_1, y_2, \ldots, y_n\}$ be independent realizations from a GEV distribution (μ_t, σ, ξ) conditional on model parameters. Assume that the temporal dependency on the location parameter is modelled through a DLM as just described above. For this case, the likelihood function is:

$$L(\{\mu_t\}_{t=1}^n\,\xi, \sigma\,|\,\{y_t\}_{t=1}^n) = \prod_{t=1}^{n} \frac{1}{\sigma}\left[1 + \xi\left(\frac{y_t - \mu_t}{\sigma}\right)\right]_+^{-(1+\frac{1}{\xi})} \exp\left\{-\left[1 + \xi\left(\frac{y_t - \mu_t}{\sigma}\right)\right]_+^{-\frac{1}{\xi}}\right\}. \tag{10.12}$$

Assuming a *first-order polynomial* DLM $(1, 1, V, W)$ on μ_t, we have

$$\mu_t = \theta_t + \nu_t\,, \quad \nu_t \sim N(0, V) \tag{10.13}$$

$$\theta_t = \theta_{t-1} + w_t\,, \quad w_t \sim N(0, W) \tag{10.14}$$

In [7] the authors argue that prior eliciting in terms of quantiles or quantile differences is to be preferred over priors on GEV parameters for constant modelling situations. Because of the complexities of a DLM-time-varying GEV model, we choose priors on the GEV parameters. The observational equation in 10.13 defines a prior for each μ_t conditional on θ_t and V, $\mu_t \sim N(\theta_t, V)$, $t = 1, \ldots, n$. We adopt independent normal priors for $\log\sigma$ and ξ,

$$\log\sigma \sim N(M_\sigma, V_\sigma), \quad \xi \sim N(M_\xi, V_\xi) \tag{10.15}$$

For V we adopt an inverse gamma prior, $V \sim IG(a, b)$, which provides a conditionally conjugate structure for this parameter. We adopt *discount factors* as in [29] to deal with the evolution

variance W. Briefly, we can say that the discount factor represents the change of information on state parameters from time $t-1$ to time t, which in [16] has proven its use for temporal and state-space modelling of extremes within the DLM framework.

10.2.4 MCMC for DLM-GEV distribution

The joint posterior distribution for all the model parameters, $(\{\mu_t\}_{t=1}^n, \xi, \sigma, \{\theta_t\}_{t=1}^n, V)$ given the data $\{y_t\}_{t=1}^n$ is

$$
\begin{aligned}
&p(\{\mu_t\}_{t=1}^n, \xi, \sigma, \{\theta_t\}_{t=1}^n, V|\{y_t\}_{t=1}^n) \propto L(\{\mu_t\}_{t=1}^n \xi, \sigma|\,\{y_t\}_{t=1}^n) \\
&\prod_{t=1}^{n}\left[\frac{1}{V^{1/2}}\exp\left\{-\frac{(\mu_t-\theta_t)^2}{2\,V}\right\}\exp\left\{-\frac{(\theta_t-\theta_{t-1})^2}{2\,W}\right\}\right] \\
&\exp\left\{-\frac{(\xi-M_\xi)^2}{2\,V_\xi}\right\}\exp\left\{-\frac{(\log\sigma-M_\sigma)^2}{2\,V_\sigma}\right\}V^{-(a+1)}\exp(-b/V)
\end{aligned}
\tag{10.16}
$$

Posterior draws can be obtained through full conditional draws of each parameter based on the Metropolis or Metropolis–Hastings algorithm. For example, to draw the shape parameter ξ at iteration $i+1$,

1. We sample ξ^{i+1} from a normal distribution centred at ξ^i which defines a symmetric proposal distribution based on a *random walk*.
2. We compute $\alpha(\xi^i, \xi^{i+1}) = min\left\{1\,\frac{p(\xi^{i+1})}{p(\xi^i)}\right\}$, where $p(\xi^{i+1})$ denotes the full conditional posterior distribution from Equation 10.16 evaluated at the proposed value and $p(\xi^i)$ is the full conditional posterior evaluated at the previous sampled value.
3. Generate u from a $U(0,1)$ distribution. If $u < \alpha(\xi^i, \xi^{i+1})$, we accept the proposed value ξ^{i+1} as our current point in the chain. Otherwise, we reject ξ^{i+1} and keep the previous point ξ^i.

Sampling σ and μ_t follows similar steps. Specifically, we draw each μ_t individually and use the prior as the proposal distribution. This leads into a Metropolis ratio that exclusively depends on the full conditional distribution of μ_t. More details on this sampling strategy appear in [16]. To draw $\{\theta_t\}_{t=0}^n$ from its full conditional distribution, we use *Forward Filtering and Backward Simulation* (FFBS) as in [2] and [11] which, as mentioned in [28], is central to MCMC implementations of *conditionally linear* normal models. In our case and following Chapter 4 of [29], the state posterior distributions conditional on all other model parameters are computed sequentially in time using the following recursive equations. If δ denotes the DLM discount factor so that $W = \frac{(1-\delta)}{\delta}C_{t-1}$, and D_t represents all the information available up to time t, for $t = 1, 2, \ldots, n$

$$
\begin{aligned}
&(\theta_t|D_{t-1}) = N(m_{t-1}, R_t), \quad R_t = C_{t-1} + W = C_{t-1}/\delta \\
&(\mu_t|\theta_t, D_{t-1}) = N(m_{t-1}, Q_t), \quad Q_t = R_t + V \\
&(\theta_t|D_t) = N(m_t, C_t), \quad m_t = m_{t-1} + A_t(\mu_t - m_{t-1})\,, C_t = A_t V\,, A_t = R_t/Q_t
\end{aligned}
\tag{10.17}
$$

To perform the backward simulation at time $t = n$, we draw θ_n from $(\theta_n|D_n)$ and draw the other θ_t parameters via *retrospective filtering* so for $t = n-1, n-2, \ldots, 0$, θ_t is drawn conditionally on θ_{t+1} from a $N(h_t, H_t)$ where

$$h_t = m_t + \delta(\theta_{t+1} - m_t), \quad H_t = C_t(1-\delta). \tag{10.18}$$

The posterior simulation for the GEV-DLM model proposed in (10.12–10.14) can be summarized as follows. At iteration $i+1$, we

- Draw $\mu_t^{i+1}|y_t, \mu_t^i, \sigma^i, \xi^i, \theta_t^i, V^i$, for $t = 1, \ldots, n$ with individual Metropolis steps.
- Draw $\sigma^{i+1}|\{y_t\}_{t=1}^n, \{\mu_t^{i+1}\}_{t=1}^n, \sigma^i, \xi^i$ with a Metropolis step.
- Draw $\xi^{i+1}|\{y_t\}_{t=1}^n, \{\mu_t^{i+1}\}_{t=1}^n, \sigma^{i+1}, \xi^i$ with a Metropolis step.
- Draw $\{\theta_t^{i+1}\}_{t=1}^n|\{\mu_t^{i+1}\}_{t=1}^n, V^i$ via FFBS as described in (10.17–10.18).
- Draw $V^{i+1}|\{\mu_t^{i+1}\}_{t=1}^n, \{\theta_t^{i+1}\}_{t=1}^n$ from an inverse-gamma distribution.

Figure 10.3 shows the posterior mean of μ_t and θ_t with a 95% probability interval for θ_t based on our MCMC approach for the Maiquetía rainfall time series with a discount factor $\delta = 0.9$ and a burn-in period of 30000 iterations. The estimated parameters are consistent with a notion of more intense extremes in recent years. The changes in location parameter from the DLM-GEV are nonlinear and remarkably different from an estimated trend based only on a deterministic line where the intercept and slope area fitted via MLE. Furthermore, Figure 10.4 shows the posterior mean estimates of predictive GEV quantiles for four different probability levels: 95%, 75%, 50% and 5% based on our modelling approach and with the Maiquetía extreme rainfall measurements. These estimates were more stable than the parameter estimates but clearly exhibit the non-stationary behaviour and skewness that is typically present in block maxima.

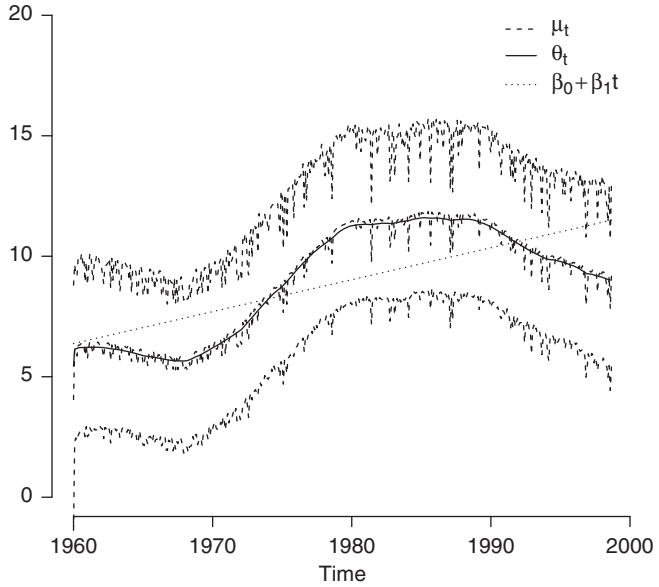

Figure 10.3 Posterior mean of μ_t and θ_t. 95% probability interval for μ_t under the GEV-DLM for Maiquetía rainfall data

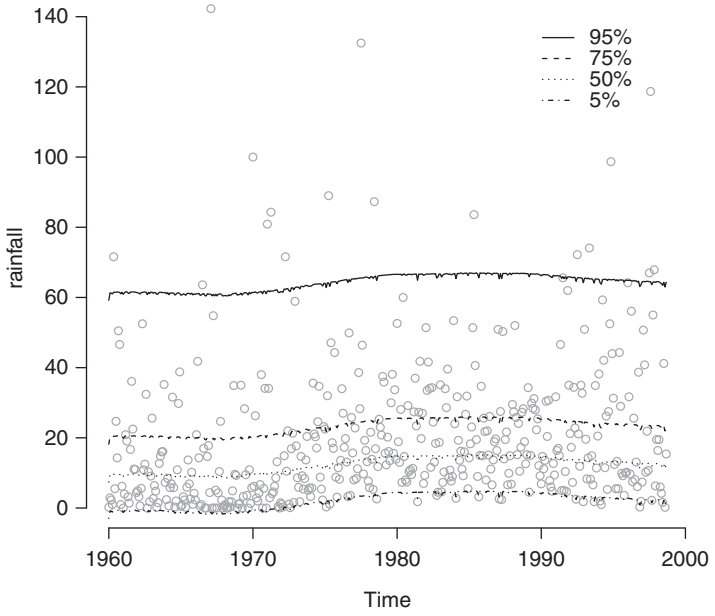

Figure 10.4 Posterior predictive quantiles based on GEV-DLM model for Maiquetía rainfall data

10.3 A spatial GEV distribution

10.3.1 Introduction

The analysis of extremes from a spatial perspective can arise as a natural extension to the time-varying GEV distribution-DLM models described previously. The main scope of these spatial models considers the theory and applications of *Gauss Markov Random Fields* (GMRFs) as described in [21], which also presents the connections of GRMFs to structural time series in the form of DLMs. Here we consider output of the Penn State/NCAR mesoscale (MM5) Regional Climate Model (RCM) which was driven by a NCAR/DOE parallel climate model. A similar output was previously analysed in [8]. The output contains extreme winter (December–January–February) precipitation of a 20-year control run that assumes current levels of greenhouse gases and begins in 1995. The spatial domain of the RCM includes 616 (28×22) grid points covering the western United States and southwestern Canada. Figure 10.5 shows maps of the output precipitation corresponding to two selected years. For year 2004, the extreme values are more intense in regions covering the Pacific coast, southwestern Canada and Arizona. Figure 10.6 shows the grid points over the spatial domain along with twenty-five locations that were held out and treated as missing in precipitation for all 20 years to assess the predictive ability of our models.

10.3.2 Objective and Gauss Markov random fields

The goal of our analysis is to treat the RCM output as data and to develop a hierarchical model around the GEV distribution that permits the characterization of the extremes through predictive quantiles, which could assist climate modellers to evaluate the performance of such as RCM. We are not comparing the results from our RCM analysis to real precipitation measurements arising

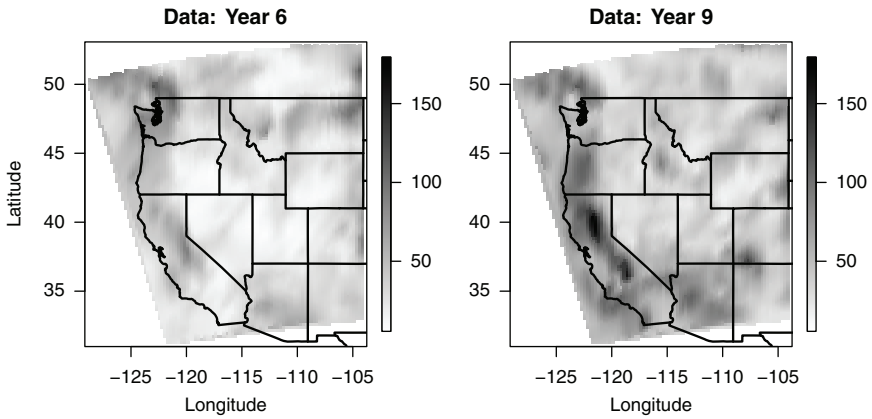

Figure 10.5 Two years of extreme precipitation for the MM5 Regional Climate Model.

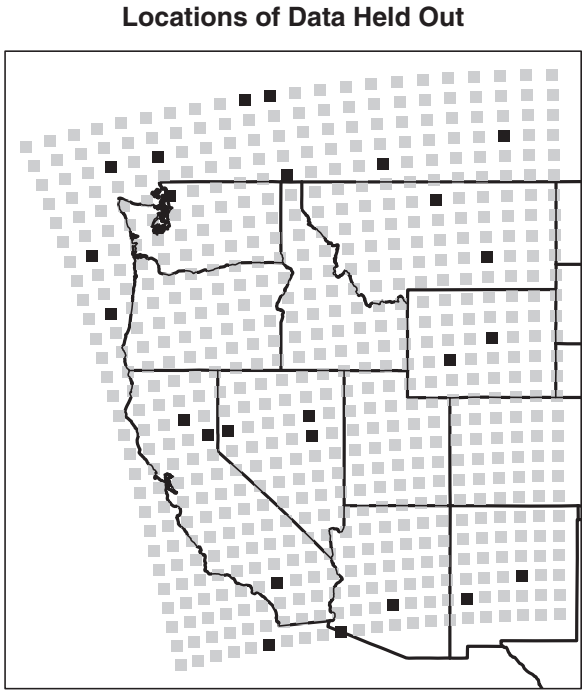

Figure 10.6 Grid points for the Regional Climate Model and holdout locations.

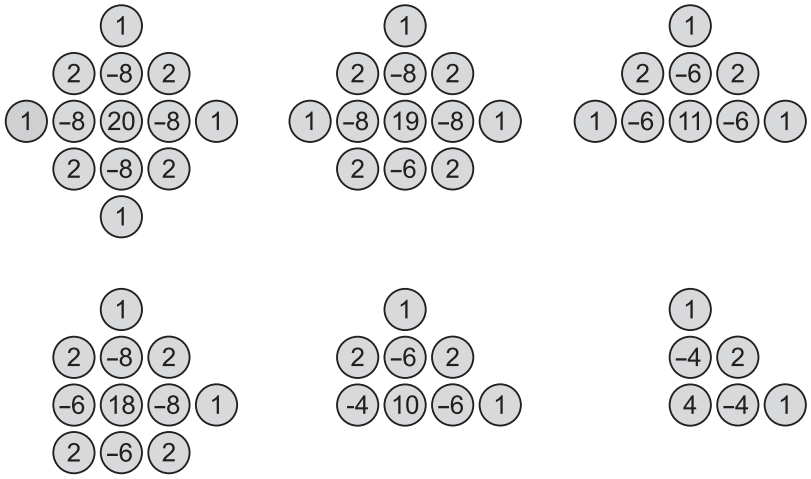

Figure 10.7 Stencils of the precision matrix for a IGMRF second-order neighbourhood structure.

from station or satellite data. From a modelling standpoint, GMRFs provide a structure defined through neighbours and precision matrices that has a graphical representation and is attractive to represent spatial relationships on high-dimensional output from climate models. The *nodes* and *vertices* representing the graph of the GMRF correspond to points on a grid and neighbours. If $\mathbf{Q}$ represents the precision matrix of the GMRF, a point (node) i is connected to a point j or $i \sim j$, if and only if $\mathbf{Q}_{ij} \neq 0$. Different specifications of the matrix $\mathbf{Q}$ provide different spatial structures. In particular, our models consider a precision matrix from a *biharmonic difference operator* which corresponds to an *intrinsic* (improper) GMRF as in Chapter 3 of [21]. The schematic of Figure 10.7 gives this second-order neighbourhood structure where the point of reference is the node with the associated highest value and corrections are imposed for when this node is near a boundary. The hierarchical model we propose to analyse the RCM output depends on a likelihood based on the GEV distribution and a prior level on parameters that involve this second-order IGMRF. For example [8] considered an IGMRF prior based on a *first-order* neighbourhood. A potential advantage of the second-order prior over a first-order structure, is that it adds extra flexibility to represent dependencies for localized climate phenomena like regional storms.

10.3.3 modelling strategy

More specifically, we assume that $Y_{st} \sim \text{GEV}(\mu^*_{st}, \sigma_s, \xi); s = 1, \ldots, n; t = 1, \ldots, 20$ are conditionally independent where Y_{st} represents precipitation from the RCM at the grid location s and at time t. In our case, $n = 616$, the number of grid point locations being considered in our analysis. Therefore the probability distribution for Y_{st} has the form,

$$H(y_{st}|\mu^*_{st}, \sigma_s, \xi) = exp\left\{-\left[1 + \xi\left(\frac{y_{st} - \mu^*_{st}}{\sigma_s}\right)\right]^{-1/\xi}_{+}\right\} \tag{10.19}$$

$$\mu^*_{st} = \mu_s + \phi_t \tag{10.20}$$

so μ^*_{st} has been additively decomposed in space and time. The scale parameter σ_s is allowed to vary in space while the shape parameter ξ is kept fixed across space and time with a prior distribution

$(\xi - 0.5) \sim \text{Beta}(9, 5)$ which guarantees that $-0.5 < \xi < 0.5$. This prior distribution was proposed in [17] and has been used as a penalization term in a GEV likelihood function in low sample size situations. This prior is not overly informative and is restricted to values that are sensible in studies of extreme precipitation.

Furthermore, we introduce vectors to represent a spatial component for the location and scale parameters respectively, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n), \boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ and $\boldsymbol{\eta} = \text{vec}(\boldsymbol{\mu}, log(\boldsymbol{\sigma}))$ is a vector that represents the concatenation of the elements in $\boldsymbol{\mu}$ and the logarithm of the elements in $\boldsymbol{\sigma}$. We model $\boldsymbol{\eta}$ as,

$$\boldsymbol{\eta} = (\boldsymbol{I}_2 \otimes \boldsymbol{X})\boldsymbol{B} + \boldsymbol{U} + \boldsymbol{\epsilon}, \tag{10.21}$$

$$\boldsymbol{\epsilon} \sim \text{N}(\boldsymbol{0}, \boldsymbol{T} \otimes \boldsymbol{I}_n), \tag{10.22}$$

$$\boldsymbol{U} \sim \text{GMRF}(\boldsymbol{0}, \boldsymbol{\theta} \otimes \boldsymbol{Q}), \tag{10.23}$$

$$\boldsymbol{T} \sim \text{Wishart}(n_T, \boldsymbol{V}_T), \tag{10.24}$$

$$\boldsymbol{\theta} \sim \text{Wishart}(n_\theta, \boldsymbol{V}_\theta), \tag{10.25}$$

$$\boldsymbol{B} \sim \text{N}(\boldsymbol{0}, \tau_{\boldsymbol{B}} \boldsymbol{I}_{10}). \tag{10.26}$$

For this specification both $\boldsymbol{T} \otimes \boldsymbol{I}_n$ and $\boldsymbol{\theta} \otimes \boldsymbol{Q}$ are precision matrices, where $\boldsymbol{Q}$ depends on the neighbourhood structure, $\otimes$ represents a *Kronecker* product and $\boldsymbol{I}_k$ is an identity matrix of dimension k. $\boldsymbol{X}$ is a matrix of covariates with $p = 5$ columns including an intercept, the longitudes, latitudes and altitudes of each grid point location and a vector indicating whether a grid point is located over the ocean or land. We expect that these covariates account for the variability as well as some of the spatial patterns in $\boldsymbol{\eta}$. $\boldsymbol{B}$ is a 10 $(2p)$-dimensional vector of regression coefficients where its first five elements are associated to $\boldsymbol{\mu}$ and the second five elements to $\boldsymbol{\sigma}$. The spatial properties of $\boldsymbol{U}$ are defined through a GMRF prior which has a precision matrix $\boldsymbol{\theta} \otimes \boldsymbol{Q}$. Here $\boldsymbol{Q}$ is defined through the construction of a second-order intrinsic GRMF as described in Chapter 3 of [21] and illustrated in Figure 10.7. The second-order precision matrix $\boldsymbol{Q}$ is not a full rank matrix, and so the GMRF prior for $\boldsymbol{U}$ is not a proper probability distribution. Therefore, we constrain $\boldsymbol{U}$ so that $(\boldsymbol{I_2} \otimes \boldsymbol{E})'\boldsymbol{U} = \boldsymbol{0}$ where $\boldsymbol{E}$ is a $n \times 3$ matrix whose columns are the eigenvectors of $\boldsymbol{Q}$ with zero eigenvalues. This constrain guarantees a proper prior on $\boldsymbol{U}$ and allows us to improve on the computational efficiencies of our MCMC simulations.

In addition, $\boldsymbol{\theta}$ is a 2×2 positive definite matrix that we model through a Wishart prior and gives the precision matrix of the blocks $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$. The term $\boldsymbol{\epsilon}$ represents global variability in $\boldsymbol{\eta}$ that is not captured by the model covariates. Each element of $\boldsymbol{\epsilon}$ is modelled independently at each grid location with a bivariate normal random variable with a 2×2 precision matrix $\boldsymbol{T}$, that we also model through a Wishart prior. The time term for μ^*_{st}, ϕ_t, follows a $\text{N}(\beta_1(t - \bar{t}), \nu), t = 1, \ldots, 20$, so β_1 is used to assess for a potential annual shift in the GEV location parameter over the 20-year control run. We assigned flat prior distributions on both β_1 and ν.

10.3.4 MCMC approach

The following steps describe the *ith* iteration of our MCMC method to sample model parameters and imputed values at held out locations in a Gibbs sampling as in [13] and resembling some of our GEV-DLM analysis. $\boldsymbol{y}$ represents the full set of observed and inputed data. $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_{20})$.

- For a held out location s, we impute $Y_{s,t}$ from the GEV distribution, $Y^{(i)}_{st} \sim \text{GEV}(\mu_s^{(i-1)} + \phi_t^{(i-1)}, \sigma_s^{(i-1)}, \xi^{(i-1)})$ for $t = 1, \ldots, 20$.

- We draw $\boldsymbol{\eta}^{(i)}|\boldsymbol{y}, \boldsymbol{B}^{(i-1)}, \boldsymbol{U}^{(i-1)}, \boldsymbol{\phi}^{(i-1)}, \boldsymbol{T}^{(i-1)}, \xi^{(i-1)}, \boldsymbol{X}$ via Metropolis steps for the pair of values $\mu_s^{(i)}$ and $log(\sigma_s^{(i)})$ at each location $s = 1, \ldots, n$ where $n = 616$ is the number of grid point locations.
- We use a random walk Metropolis step to draw $\phi_t^{(i)}|\boldsymbol{y}, \boldsymbol{\eta}^{(i)}, \xi^{(i-1)}, \beta^{(i-1)}, \nu^{(i-1)}$ for $t = 1, \ldots, 20$.
- We directly draw $\beta^{(i)}|\nu^{(i-1)}, \boldsymbol{\phi}^{(i)}$ and $\nu^{(i)}|\beta^{(i)}, \boldsymbol{\phi}^{(i)}$ from normal and inverse-gamma distributions respectively.
- We use random walk Metropolis step to draw $\xi^{(i)}|\boldsymbol{y}, \boldsymbol{\eta}^{(i)}, \boldsymbol{\phi}^{(i)}$.
- We draw $\boldsymbol{B}^{(i-1)}|\boldsymbol{\eta}^{(i)}, \boldsymbol{U}^{(i-1)}, \boldsymbol{T}^{(i-1)}, \boldsymbol{X}$ from a multivariate normal with mean vector $\mu_{\boldsymbol{B}} = \Sigma_{\boldsymbol{B}}(\boldsymbol{T}^{(i-1)} \otimes \boldsymbol{X}')(\boldsymbol{\eta}^{(i)} - \boldsymbol{U}^{(i-1)})$ and covariance matrix $\Sigma_{\boldsymbol{B}} = (\tau_{\boldsymbol{B}}\boldsymbol{I_{10}} + \boldsymbol{T}^{(i-1)} \otimes \boldsymbol{X}'\boldsymbol{X})^{-1}$.
- We draw $\boldsymbol{U}^i|\boldsymbol{\eta}^{(i)}, \boldsymbol{T}^{(i-1)}, \boldsymbol{B}^{(i)}, \boldsymbol{\theta}^{(i-1)}, \boldsymbol{X}$ by first generating $\boldsymbol{U}^*$ from a $N_C(\boldsymbol{b_U}, \boldsymbol{Q_U})$ where $\boldsymbol{b_U} = (\boldsymbol{T}^{(i-1)} \otimes \boldsymbol{I}_n)\boldsymbol{\eta}^{(i)} - (\boldsymbol{T}^{(i-1)} \otimes \boldsymbol{X})\boldsymbol{B}^{(i)}$ and $\boldsymbol{Q_U} = (\boldsymbol{T}^{(i-1)} \otimes \boldsymbol{I}_n) + (\boldsymbol{\theta}^{(i-1)} \otimes \boldsymbol{Q})$. N_C designates the canonical parameterization of a GMRF as in [21]. We then correct for the linear constraint $(\boldsymbol{I}_2 \otimes \boldsymbol{E})'\boldsymbol{U} = \boldsymbol{0}$, via the method of *conditioning by Kriging* as detailed in [21], page 37. This leads to the expression, $\boldsymbol{U}^{(i)} = (\boldsymbol{I}_2 \otimes (\boldsymbol{I}_n - \boldsymbol{E}\boldsymbol{E}'))\boldsymbol{U}^*$. The matrix $(\boldsymbol{I}_2 \otimes (\boldsymbol{I}_n - \boldsymbol{E}\boldsymbol{E}'))$ is the perpendicular projection operator that projects $\boldsymbol{U}^*$ into the column space of $\boldsymbol{Q}$.
- We draw $\boldsymbol{T}^{(i)}|\boldsymbol{\eta}^{(i)}, \boldsymbol{B}^{(i)}, \boldsymbol{U}^{(i)}, \boldsymbol{X}$ from a Wishart$(n_{\boldsymbol{T}} + n, (\boldsymbol{V_T} + \boldsymbol{C}'\boldsymbol{C})^{-1})$ where $\boldsymbol{C}$ is a $n \times 2$ matrix with columns $\boldsymbol{\mu}^{(i)} - \boldsymbol{X}\boldsymbol{B}^{(i)}_{1,\ldots,p} - \boldsymbol{U}^{(i)}_{1,\ldots,n}$ and $log(\boldsymbol{\sigma}^{(i)}) - \boldsymbol{X}\boldsymbol{B}^{(i)}_{p+1,\ldots,2p} - \boldsymbol{U}^{(i)}_{n+1,\ldots,2n}$, and where $n = 616$ is the number of grid points and $p = 5$ is the number of model covariates, including an intercept. The notation $\boldsymbol{A}_{j,\ldots,k}$ with $j \leq k$ represents the vector formed with the consecutive entries of $\boldsymbol{A}$ starting from element j and ending in element k.
- Finally, we draw $\boldsymbol{\theta}^{(i)}|\boldsymbol{D}, \boldsymbol{Q}$ from a Wishart$(n_{\boldsymbol{\theta}} + n, (\boldsymbol{V_\theta} + \boldsymbol{D}'\boldsymbol{Q}\boldsymbol{D})^{-1})$ where $n = 616$ and $\boldsymbol{D}$ is a $n \times 2$ matrix with columns formed by the vectors $\boldsymbol{U}^{(i)}_{1,\ldots,n}$ and $\boldsymbol{U}^{(i)}_{n+1,\ldots,2n}$ respectively.

10.3.5 Analysis of the RCM output

Figure 10.8 shows maps for the posterior mean estimates for the vector of location and scale parameters respectively based on the MCMC described in Section 10.3.4. The posterior estimates $\boldsymbol{\mu}$ were higher at most locations along the Pacific coast, from Canada to central California, as well as locations in central Arizona, suggesting that annual precipitation maxima were generally more extreme in these areas. The estimates of $\boldsymbol{\sigma}$ were higher along the Pacific coast of Washington state and British Columbia, along the Sacramento Valley of northern and central California, the coast of southern California, and in southern Arizona. This suggests that the distribution of annual precipitation maxima in these areas is more variable than in other areas.

Figure 10.9 shows the posterior distributions of ξ, the constant shape parameter, and of β_1, the slope of the location parameter trend. The posterior distribution of ξ has a mean of 0.063 and a standard deviation of 0.0083. This is very different compared to the prior distribution of ξ, which is a shifted beta distribution on $[-0.5, 0.5]$ with a mean of 0.1 and a standard deviation of 0.12. The posterior probability that $\xi > 0$ is close to one, which corresponds to a *Fréchet* case. However the range of values for ξ is lower than that traditionally obtained with GEV distribution fits to real measurements of precipitation. On the other hand, the posterior distribution of β_1 has a posterior mean of -0.016, but a standard deviation of 0.11, not showing any evidence that β_1 differs from zero. This indicates that over the 20 years of control runs for the RCM output, our statistical model is not able to detect any deterministic time changes in the GEV location parameters. Maps of posterior predictive quantiles of the distribution of annual precipitation maxima are shown in Figure

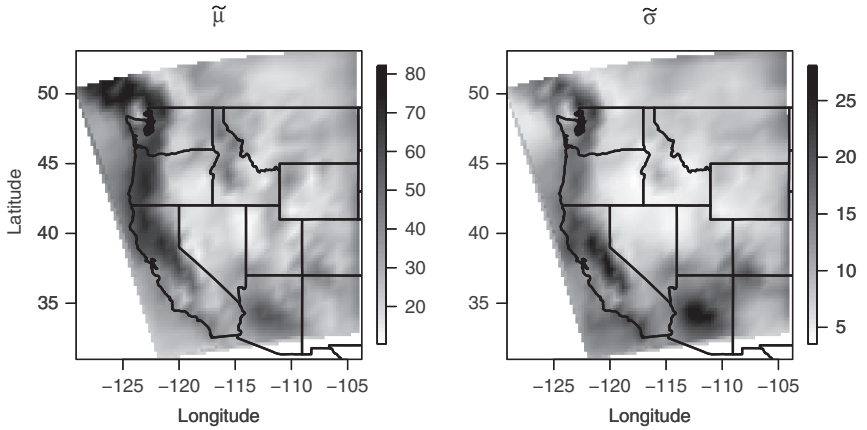

Figure 10.8 Posterior mean estimates for $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ for the RCM output.

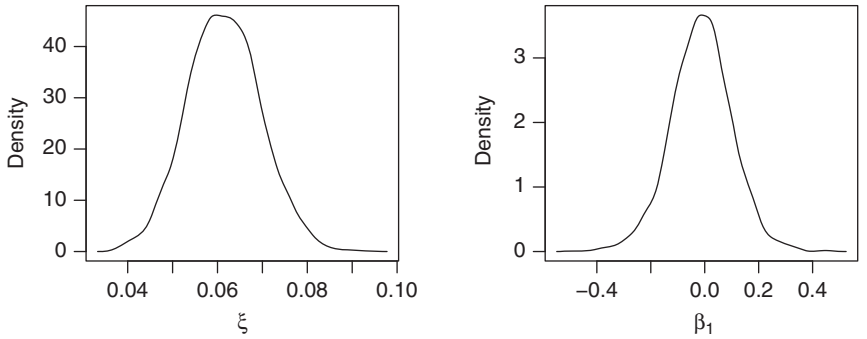

Figure 10.9 Posterior distributions for shape and slope parameters. RCM output.

10.10. These maps are based on posterior predictive samples and not on GEV quantiles or GEV *return levels*. All of these quantiles are relatively high along the Pacific coast from British Columbia through northern California, through northern and central California, and in central Arizona. The annual precipitation maxima in these areas are higher than in other areas of the spatial domain, but rarely, very extreme. In terms of posterior predictive evaluations, Figure 10.11 shows histograms of samples from the posterior predictive distribution corresponding to the held-out values of y_{st} at four different locations. The four locations were selected from the 25 originally held out locations and to be roughly representative of the northwest, northeast, southwest, and southeast regions of the study area. Vertical bars show the posterior predictive median and the 0.95 posterior predictive quantile based on the predictive samples. The small vertical lines along the x-axis show the actual 20 observed (output) values that were held out at each of the four locations. The observations are in accordance with our predictive distribution and were computed under the assumption of a zero-trend parameter. Similar figures were obtained for other held out locations. Figure 10.12 shows

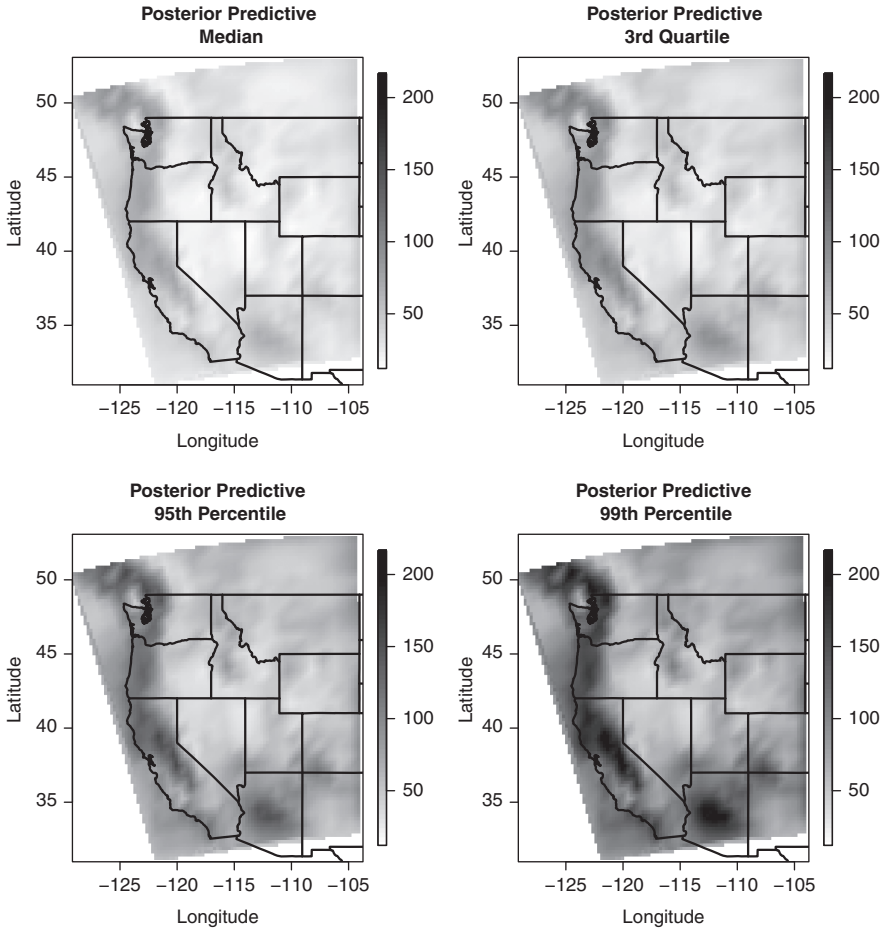

Figure 10.10 Posterior predictive percentiles at four levels: 50%,75%, 90% and 99% for the RCM output.

the posterior Distributions for each of the elements in $\boldsymbol{B}$. The first row of histograms corresponds to the parameters associated with $\boldsymbol{\mu}$ while the second row corresponds to $\boldsymbol{\sigma}$. The covariates that appear more relevant are those associated with longitude, latitude and relative position to ocean. The latitude coefficients show a negative change for both $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$, indicating that annual precipitation maxima become both generally smaller and less variable with increasing latitude. The longitude coefficients also show a negative change with both $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$, indicating that annual precipitation maxima are both generally smaller and less variable in the eastern portion of the study area than in the western portion. The elevation coefficients indicate a positive change on $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$, indicating that in general annual precipitation maxima are both more extreme and more variable at higher elevations, however, this change is anticipated to be rather small. The coefficients for the ocean indicator variable induce a negative change for $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$, so that annual precipitation maxima are both less extreme and less variable over the ocean than over land.

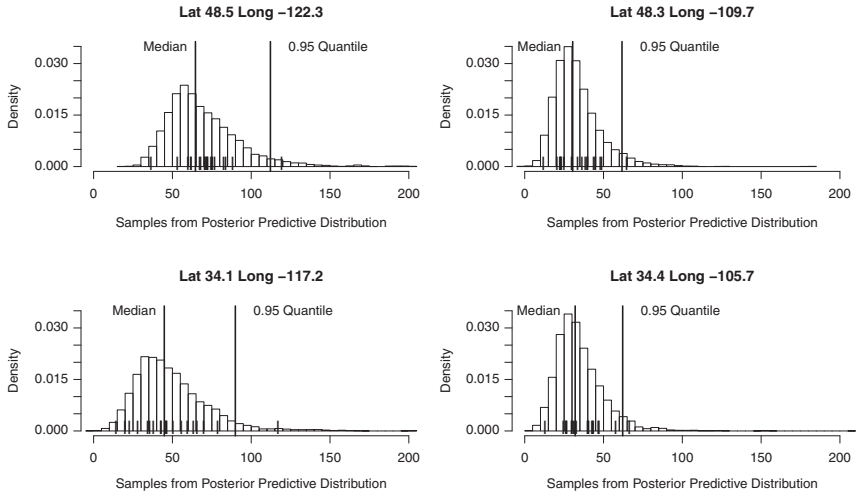

Figure 10.11 RCM output analysis. Predictive posterior distribution at four held out grid points.

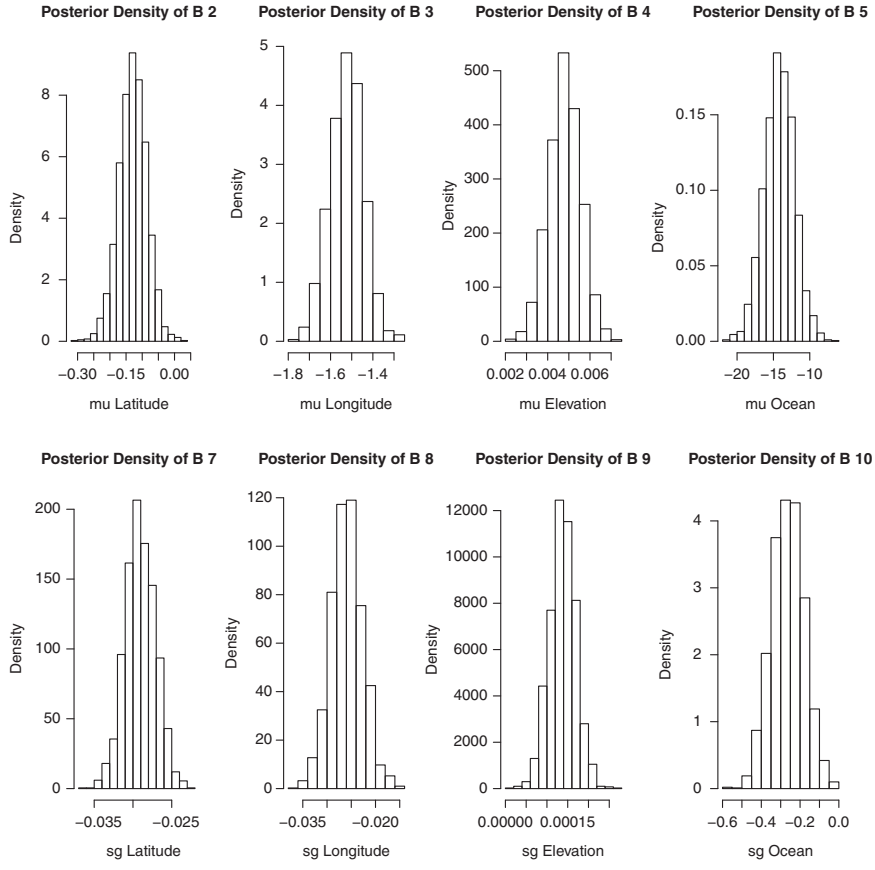

Figure 10.12 RCM analysis. Posterior distributions for regression parameters.

10.4 Conclusions

This chapter presents a general class of models to study extreme values based on the GEV distribution and that rely on time domain and spatial latent components. These models had not been available until recently and are thanks to the developments of MCMC approaches. In particular, dynamic models as described in [29] and [28] provide a flexible approach to deal with time-varying extremes via MCMC algorithms based on FFBS, in contrast with the traditional deterministic parameter regression models. Interesting developments in this area had arisen from the *Particle Filter* perspective starting from the work by [12] and more recently extended by [9], both emphasizing the analysis of the athletics dataset of [20]. In addition, [18] present a different framework where the time dependence in extremes is modelled through AR and MA with innovations arising from the Gumbel distribution and illustrated for extreme returns of daily stock data. From the spatial or spatial temporal perspective, some of the main model developments for the study of extremes arose from the Bayesian hierarchical perspective with MCMC methods initiated by [3] and also considered in for example, [16], [22] and [8]. As shown in this chapter, these modelling approaches had proven their value and flexibility in the representation of extremal phenomena from station data or from climate model output in high-dimensional situations. However, the main drawback of these approaches is that they rely on assumptions of *conditional independence* which may not be adequate to represent the spatial dependencies of extremes. Developments based on *Copulas* as in [23] and [14] and the *Max stable* process as in [24] through *composite likelihood* methods combined with MCMC as in [19], provide some examples of the recent focus for modelling spatial extremes from a Bayesian point-of-view.

Acknowledgements

The rainfall data at Maíquetia was kindly provided by Bruno Sansó. The RCM output data was kindly provided by Steve Sain and Dan Cooley while the first author was visiting NCAR.

References

[1] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statistics of Extremes*. Wiley, Chichester, England.

[2] Carter, C. K. and Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, **81**(3), 541–553.

[3] Casson, E. and Coles, S. (1999). Spatial regression models for extremes. *Extremes*, **1**(4), 449–468.

[4] Coles, S. (2001). *An Introduction to Statistical modelling of Extreme Values*. Springer-Verlag, New York, USA.

[5] Coles, S. and Pericchi, L. (2003). Anticipating catastrophes through extreme value modelling. *Journal of the Royal Statistical Society Series C, Applied Statistics*, **52**, 405–416.

[6] Coles, S. G. and Powell, E. A. (1996). Bayesian methods in extreme value modelling: A review and new developments. *International Statistical Review*, **64**(1), 119–136.

[7] Coles, S. G. and Tawn, J. A. (1996). Modelling extremes of the areal rainfall process. *Journal of Royal Statistics Society*, **B**(58), 329–347.

[8] Cooley, D. and Sain, S. (2010). Spatial hierarchical modelling of precipitation extremes from a regional climate model. *Journal of Agricultural Biological and Environmental Statistics*, **15**(3), 381–402.

[9] Fearnhead, P., Wyncoll, D. and Tawn, J. (2010). A sequential smoothing algorithm with linear computational cost. *Biometrika*, **97**(2), 447–464.

[10] Fisher, R. A. and Tippet, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, **24**, 180–190.

[11] Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, **15**, 183–202.

[12] Gaetan, C. and Grigoletto, M. (2004). Smoothing sample extremes with dynamic models. *Extremes*, **7**, 221–236.

[13] Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of American Statistical Association*, **85**(410), 398–409.

[14] Ghosh, S. and Mallick, B. (2010). A hierarchical Bayesian spatio-temporal model for extreme precipitation events. *Environmetrics*, **22**, 192–204.

[15] Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **87**, 97–109.

[16] Huerta, G. and Sansó, B. (2007). Time-varying models for extreme values. *Environmental and Ecological Statistics*, **14**(3), 285–299.

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

[18] Nakajima, J., Kunihama, T. Omori, T. and Früwirth-Schnatter (2011). Generalized extreme value distribution with time-dependence using the ar and ma models in state space form. *Computational Statistics and Data Analysis* (in press).

[19] Ribatet, M., Cooley, D. and Davison, A. (2011). Bayesian inference for composite likelihood models and an application to spatial extremes. *Statistica Sinica* (accepted).

[20] Robinson, M. E. and Tawn, J. A. (1995). Statistics for exceptional athletics records. *Journal of the Royal Statistical Society Series C, Applied Statistics*, **44**, 499–511.

[21] Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields*. Chapman & Hall/CRC.

[22] Sang, H. and Gelfand, A. E. (2009). Hierarchical modelling for extreme values observed over space and time. *Environmental and Ecological Statistics*, **16**, 407–426.

[23] 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.

[24] Smith, R. (1990). Max-stable processes and spatial extremes. Unpublished manuscript.

[25] Smith, R.L. and Naylor, J.C. (1988). A comparison of maximum likelihood and Bayesian estimators for the three parameter Weibull distribution. *Applied Statistics*, **36**, 358–369.

[26] Stephenson, A.G. (2002, September). *A User's Guide to the Evdbayes Package(Version 1.0)*. http://www.maths.lancs.ac.uk/~stephena/.

[27] Stephenson, A. G. and Gilleland, E. (2005). Software for the analysis of extreme events: the current state and future directions. *Extremes*, **8**(3), 87–109.

[28] West, M. (2012). Bayesian dynamic modelling. In *Bayesian Theory and Applications*, pp. 145–166. Oxford University Press, Oxford.

[29] West, M. and Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models* (Second edn). Springer-Verlag, New York.