STATISTICA, anno LXXIII, n. 1, 2013

CAN A REGIONAL CLIMATE MODEL REPRODUCE OBSERVED EXTREME TEMPERATURES?

Peter F. Craigmile
School of Mathematics and Statistics, University of Glasgow, Glasgow, United Kingdom
Department of Statistics, The Ohio State University, Columbus, USA

Peter Guttorp
Department of Statistics, University of Washington, Seattle, USA
Norwegian Computing Center, Oslo, Norway

1. INTRODUCTION

In order to adapt to a changing climate, policymakers need information about what to expect for the climate system. In particular, they need local information about changes in extreme weather. Such information typically comes from regional climate models (Rummukainen, 2010). A regional climate model is a downscaled global climate model, a mathematical model that describes, using partial differential equations, the temporal evolution of climate, oceans, atmosphere, ice, and land-use processes over a gridded spatial domain of interest. An important issue is how well the regional models reproduce observed extreme climate variables. There are several statistical problems associated with such comparisons. Models typically produce output on grids in space and time, while data are collected at various time scales at synoptic weather stations. Standard geostatistical methodology can be used to upscale station data to grid squares (Berrocal *et al.*, 2012), but there are no geostatistical tools that can be directly used to upscale station extremes to grid square extremes.

From a statistical point of view it is natural to view climate as the statistical distribution of the variables of interest. In this paper we will consider minimum temperatures in south central Sweden. Most climate models indicate that we should expect increases in minima in this region (Solomon *et al.*, 2007). Such increases can be critical for forests when the intensities and areas of occurrence of pest species change (e.g., Kocmánková *et al.*, 2010; Marini *et al.*, 2012).

In this article we will compare the distribution of seasonal minimum daily mean temperatures as simulated by the Swedish regional model RCA3, driven by a European reanalysis, to the distribution of seasonal minimum temperatures at 17 synoptic stations in south central Sweden consisting of 12–48 years of data. We use the minima of the daily means (rather than the daily minima), because this is more comparable with the regional climate model output. (The regional model produces separate calculations of daily minimum temperature, based on the time step used for the spatial resolution, but for this analysis we did not have access to these model data.) Zwiers *et al.* (2011) used

the annual maxima and minima of daily mean temperatures from climate models, but compared them to annual maxima and minima of measured daily maxima and minima, averaged over grid squares which were substantially larger than ours. The comparisons between models and data in this paper are based on the same variable, and carried out on a much smaller scale, in time (we compare seasonal temperatures) and in space (we have smaller grid boxes). We emphasize that our comparisons are made upon the parameters of the marginal extreme value distribution of the minima. In the discussion, we discuss extensions to multivariate extremes.

We describe the details of the data sets used in Section 2. The methodology is presented in Section 3, while results can be found in Section 4. Finally, we discuss some technical issues, our findings, and some of their consequences in Section 5.

2. DATA

The observations are downloaded from the Swedish Meteorological and Hydrological Institute (SMHI) website, and consists of synoptic observations from 17 sites in an area of south central Sweden, delineated by 57 to 63°N and 12 to 19°E. The longest of the station series have data from March 1, 1961, through May 31, 2008. In this analysis we will use daily mean temperature, as calculated by the Ekholm-Modén formula (see Ma and Guttorp, 2013). We calculate seasonal minima of these values for each station (In sequence, the minimum in December, January, February (DJF); March, April, May (MAM); June, July, August (JJA); and September, October, November (SON)). A summary of the locations and the time series of the seasonal minima are shown in Figure 1. We can see that there are 10 long records, a single medium length record, and 6 short records. As expected, the time series of the seasonal minima have a strong yearly periodicity, and by inspection correlate strongly over space. We further discuss the distribution of the observed seasonal minima series below.

The regional model output consists of a control model of the SMHI regional model RCA3 (Samuelsson *et al.*, 2011), run using boundary conditions given by the ERA40 reanalysis (Uppala *et al.*, 2005) in the earlier years, and the ERA-INTERIM reanalysis (Uppala *et al.*, 2008) in the later years. Based on a model temporal resolution of 7.5 minutes, the 2 meter temperature model values are calculated daily on a 12.5×12.5 km^2 grid over the same region and the same time period as the longest station records. To be comparable with the station temperatures, we restrict our analysis to the 2 meter temperature given for the open land and snow land covers. (Other land types such as forest, lakes and sea water, lakes and sea ice are available from the computer model output but were not used in this analysis.) A map of the centroids of the 1989 grid squares containing open land or snow are shown in Figure 2. For the climate model output we calculate seasonal minima for each grid square in the same fashion as for the station data.

3. METHODS

We use extreme value theory to analyze and model the seasonal minima that we have calculated for both the station data and the regional climate model output. Below we present the model for the observed seasonal minima. We fit the same type of model to the seasonal minima from the regional climate model.

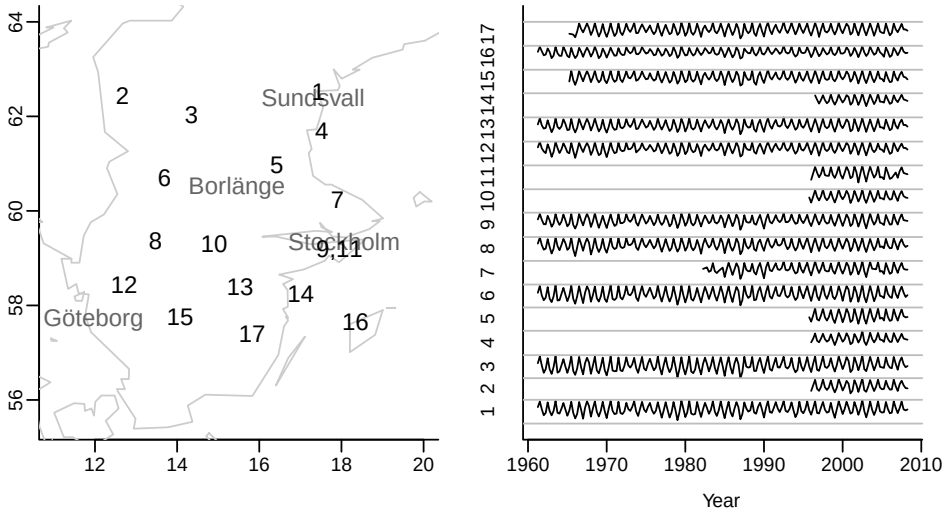

Figure 1 – The left panel shows the location of the 17 synoptic observation sites in south central Sweden. The right panel shows time series of the seasonal minima for these 17 locations.

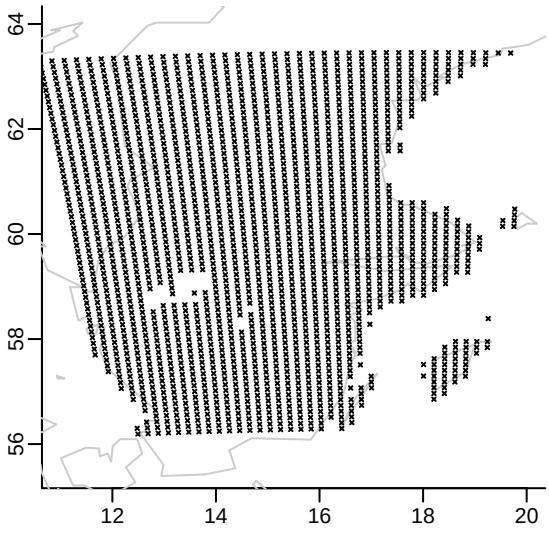

Figure 2 – The location of the centroids of the grid squares for open land and snow in the control model run of the SMHI regional model RCA3.

As each season can be considered to contain roughly a block of 91 days, we use a block minima approach (see, e.g., Coles, 2001, for a summary of modeling block minima, as well as other statistical methods for extremes). Suppose we are modeling the station data. At each spatial location $s \in D$ in our spatial domain of interest D (south central Sweden), let $Z_t(s)$ denote the minima in year $y_t(s)$ at season $d_t(s)$ (taking on values 1: DJF; 2: MAM; 3: JJA; 4: SON), for time index $t = 1,\ldots,N(s)$. Here, $N(s)$ denotes the number of values at each spatial location. Then, under certain regularity conditions, the negative of the minima series at each spatial location asymptotically follow a generalized extreme value (GEV) distribution (Coles, 2001, Section 3.2). For each time t and spatial location s, let $M_t(s) = -Z_t(s)$ denote the negative of the minima. Then

$$M_t(s) \quad \sim \quad \text{GEV}(\widetilde{\mu}_t(s), \sigma_t(s), \xi_t(s)), \tag{1}$$

conditionally independent over s and t. In this GEV distribution $\widetilde{\mu}_t(s) \in \mathbb{R}$ is the location parameter indicating values which the distribution of the negative minima are concentrated around. Correspondingly let $\mu_t(s) = -\widetilde{\mu}_t(s)$ be the location parameter for the minima. In (1), $\sigma_t(s) > 0$ is the scale parameter defining the spread of the distribution, and $\xi_t(s)$ is the shape parameter. The tails of the GEV distributions are heavier for higher values of the shape parameter (A negative shape parameter leads to bounded tails; otherwise the tails of the distribution are unbounded.)

We think of the parameters of the GEV distribution as describing the climate, with the changes in the parameters indicating seasonal differences and possible trends. Given the climate, the model technically assumes that weather at different stations is conditionally independent. This is an oversimplification, since typically events of extremely cold air arise from arctic air moving south during a high pressure situation. So given that one station is extremely cold, it is more likely that another is. We comment further on this in the Discussion section.

In our model for the station data we assume the following spatio-temporal model for the location parameters, $\{\mu_t(s)\}$:

$$\mu_t(s) = \beta_0(s) + \beta_1(s)(y_t(s) - 1960) + \sum_{d=2}^{4} \beta_d(s) I(d_t(s) = d), \tag{2}$$

where $s \in D$, $t = 1,\ldots,N(s)$ and $I(\cdot)$ is the indicator function. This model is an extension of a model defined in Mannshardt *et al.* (2013) to allow for seasonally varying effects. In the model (2), $\beta_0(s)$ is a spatially varying intercept term for the DJF season, $\beta_1(s)$ is a spatially-varying yearly trend term, and $\beta_d(s)$ $(d = 2,\ldots,4)$ defines the spatially-varying seasonal effects for MAM, JJA, and SON, relative to DJF. To allow for heterogeneity over space, we assume for each $j = 0,\ldots,4$ that each $\{\beta_j(s)\}$ is a Gaussian geostatistical process with mean λ_j and isotropic covariance, $\text{cov}(\beta_j(s), \beta_j(s+h)) = \tau_j \exp(-||h||/\phi_j)$, where $\tau_j > 0$ is the (partial) sill parameter, $\phi_j > 0$ is the range parameter, and $||\cdot||$ denotes the Euclidean norm. We assume mutual independence of these Gaussian processes $\{\beta_j(s)\}$ over j.

In our model we further assume that the scale parameters vary over space, but are constant in time; i.e., $\sigma_t(s) = \sigma(s)$ for all t and s, and that the shape parameter, ξ,

TABLE 1
The prior distributions for the hyperparameters in the model for the minima for the station data. The index j ranges over 0 to 4.

Parameter	Prior distribution
λ_j	Normal with mean 0 and variance 10
τ_j	Inverse gamma with shape 0.01 and rate 0.01
ϕ_j	Gamma with shape 10 and rate 1
λ_σ	Normal with mean 0 and variance 10
τ_σ	Inverse gamma with shape 0.01 and rate 0.01
ϕ_σ	Gamma with shape 10 and rate 1
ξ	Normal with mean 0 and variance 0.5

is constant. This is an approximation, but simplifies the computation somewhat. Indeed, the assumption of a constant shape parameter is reasonable in a number of different applications (e.g., Cooley *et al.*, 2007; Sang and Gelfand, 2010). We suppose $\{\log\sigma(\boldsymbol{s})\}$ is a Gaussian geostatistical process with mean λ_σ and isotropic covariance $\text{cov}(\log\sigma(\boldsymbol{s}),\log\sigma(\boldsymbol{s}+\boldsymbol{h}))=\tau_\sigma\exp(-||\boldsymbol{h}||/\phi_\sigma)$. We assess this assumption in Section 4.

Let $\{\boldsymbol{s}_k : k=1,\ldots,K\}$ denote the locations of the $K=17$ synoptic observation sites in south central Sweden. Our parameters of interest are

$$\begin{aligned}\theta &= \Big(\{\beta_j(\boldsymbol{s}):\boldsymbol{s}\in D, j=0,\ldots,4\},\{\log\sigma(\boldsymbol{s}):\boldsymbol{s}\in D\},\xi\\ &\qquad \{\lambda_j : j=0,\ldots,4\},\{\tau_j\},\{\phi_j\},\lambda_\sigma,\tau_\sigma,\phi_\sigma\Big)^T.\end{aligned}$$

Assuming independence of the hyperparameters, Table 1 summarizes the marginal prior distributions used for this model. With the exception of the hyperparameters for the shape parameter ξ and the spatial range parameters, we assume vague priors. For the range parameter we use the gamma prior choice of Craigmile and Guttorp (2011), who modeled daily mean temperature from the same synoptic stations. The results are robust to minor changes in this choice of prior distribution. Let $\boldsymbol{z}$ denote the realized $(Z_t(\boldsymbol{s}_k): k=1,\ldots,K, t=1,\ldots,N(\boldsymbol{s}_k))^T$ vector. Let $\boldsymbol{R}(\phi)$ denote the correlation matrix with (k,k') element $\exp(-||\boldsymbol{s}_k-\boldsymbol{s}_{k'}||/\phi)$, $\beta_j=(\beta_j(\boldsymbol{s}_k):k=1,\ldots,K)^T$, and $\boldsymbol{\varkappa}=(\log\sigma(\boldsymbol{s}_k): k=1,\ldots,K)^T$. Then the posterior distribution is

$$\begin{aligned}\pi(\theta|\boldsymbol{z}) \propto &\left[\prod_{k=1}^{K}\prod_{t=1}^{N(\boldsymbol{s}_k)} f(z_t(\boldsymbol{s})|\tilde{\mu}_t(\boldsymbol{s}_k),\sigma(\boldsymbol{s}_k),\xi)\right]\times\\ &\left[\prod_{j=0}^{4} n_K(\beta_j|\lambda_j,\tau_j\boldsymbol{R}(\phi_j))\pi(\lambda_j)\pi(\tau_j)\pi(\phi_j)\right]\times\\ &\left[|J|\, n_K(\boldsymbol{\varkappa}|\lambda_\sigma,\tau_\sigma\boldsymbol{R}(\phi_\sigma))\,\pi(\lambda_\sigma)\pi(\tau_\sigma)\pi(\phi_\sigma)\right]\pi(\xi). \qquad (3)\end{aligned}$$

In the above expression $f(z|\tilde{\mu},\sigma,\xi)$ is the GEV density function evaluated at z with

location $\tilde{\mu}$, scale σ, and shape ξ, $n_K(\cdot|\lambda_j \mathbf{1}, \tau_j \boldsymbol{R}_j(\phi_j))$ is the density function for a K-variate normal distribution with mean $\lambda_j \mathbf{1}$ and covariance matrix $\tau_j \boldsymbol{R}_j(\phi_j)$ (here $\mathbf{1}$ is a vector of ones), and J is the Jacobian of the transformation in transforming from $\sigma(\boldsymbol{s}_k)$ to $\log \sigma(\boldsymbol{s}_k)$ at all locations $\boldsymbol{s}_k$, $k = 1, \ldots, K$.

The posterior distribution described by equation (3) is not available in closed form. In order to fit this model to the station data we use a Markov chain Monte Carlo (MCMC) algorithm to sample from the posterior distribution. The algorithm used in this paper is adapted from the algorithm given in the supplemental material of Mannshardt *et al.* (2013), which allows for sampling more than two spatially varying Gaussian processes.

In order to compare distributions, we use Doksum's shift function (Doksum, 1974). The idea is to find a shift function $\Delta(x)$ such that for random variables X and Y we have $X + \Delta(X) \sim Y$. Empirically, if G_m is the empirical distribution function of a sample of size m from Y and F_n is the empirical distribution function of a sample of size n from X we estimate Δ by

$$\hat{\Delta}(x) = G_m^{-1}(F_n(x)) - x.$$

Using Kolmogorov-Smirnov results enables construction of simultaneous confidence bands for Δ (Doksum and Sievers, 1976).

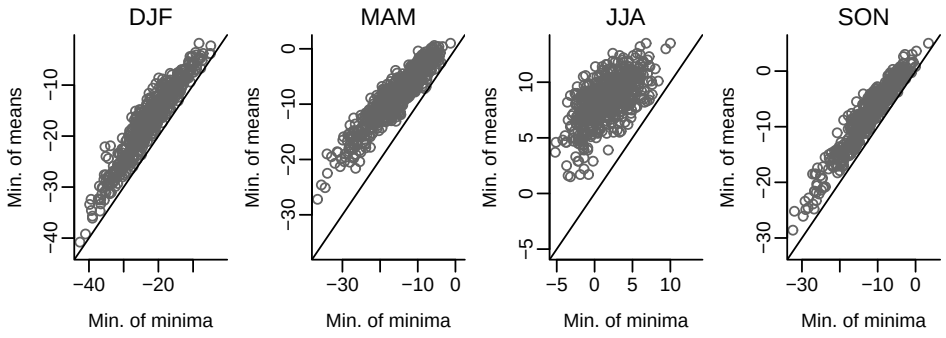

Figure 3 – Seasonal minima of the observed daily means plotted against the season minima of observed daily minima. The solid line, with an intercept of zero and a unit slope, is used as a guide to compare these two distributions.

4. RESULTS

The seasonal minimum daily mean temperature, which is what we use in this paper, is of course somewhat higher than the seasonal daily minimum temperature. Figure 3 demonstrates this effect by season, for the stations used in our analysis. For remainder of the paper we consider the seasonal minimum daily mean temperature.

We compare the seasonal minima of the regional model output to the seasonal minima of the daily mean temperatures at the 17 locations in south central Sweden in two different ways. First we look at the distribution of minimum temperature over the grid square containing the station, comparing the distributions using GEV models and

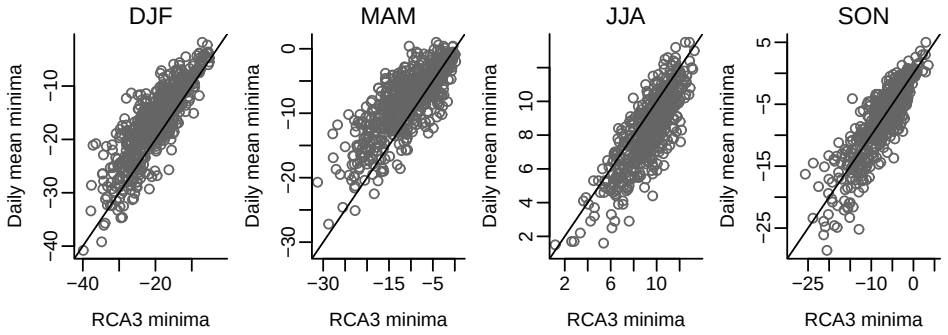

Figure 4 – Seasonal minima calculated for the regional climate model RCA3 plotted against the seasonal minima for the observed daily mean temperatures. The solid line, with an intercept of zero and a unit slope, is used as a guide to compare these two distributions.

Doksum's shift function. Second, we fit our hierarchical Bayesian GEV model to the observed minima and the regional climate model minima, to make comparisons with the regional climate model over the entire spatial region. All data management and computing was carried out in R (R Development Core Team, 2007). Maximum likelihood based estimation of the GEV model parameters was obtained using the `ismev` R library (`http://www.ral.ucar.edu/~ericg/softextreme.php`).

4.1. Comparing the seasonal minima for the climate model with station data

We compare the observations at each station with the regional climate model observations in the grid cell containing that station. We make these comparisons without regard to spatial or temporal effects that may be present. Figure 4 compares, by season, the seasonal minima for the station data with the seasonal minima for the RCA3 regional climate model. We use the solid line on each panel, with an intercept of zero and a unit slope, as a guide to compare these two distributions. The seasonal minima for the station data and the climate model agree best in the autumn (SON). In the winter (DJF) and spring (MAM), the observed minima tend to be slightly higher than is observed for the regional climate model. For the summer (JJA), this relationship is reversed – we observe cooler minima than is predicted by the model.

For each season we fit a GEV model to the negative of the observed seasonal minima using a model for the location parameter that is a linear function of the years since 1960. For each season we assume the scale and shape parameters are constant over the stations. The results are shown in Table 2. We summarize the negative of the intercept and slope parameters to make our comparisons on the original scale of the observed seasonal minima. Assuming independence, the table shows a significant intercept effect, and a positive slope in most seasons, indicating that the distribution of minima are strongly seasonal and tend to be warmer as the years progress. There is a suggestion that the scale parameter may be smaller in the summer (we discuss this point later in the paper), and that the shape parameters may differ (although there is strong correlation between the estimates of the scale and shape).

TABLE 2
The estimated GEV model parameter estimates fit to the observed seasonal minima, using the year as a predictor. A different model is fit to each season. The numbers in parentheses are the standard errors, assuming independence.

	DJF		MAM		JJA		SON	
Intercept	-16.90	(0.64)	-6.87	(0.43)	9.10	(0.19)	-5.49	(0.45)
Years after 1960	0.09	(0.02)	0.03	(0.01)	0.00	(0.01)	0.03	(0.01)
Scale	6.12	(0.20)	4.22	(0.15)	1.89	(0.06)	4.46	(0.15)
Shape	-0.16	(0.03)	-0.07	(0.03)	-0.16	(0.02)	-0.02	(0.03)

TABLE 3
As in Table 2, but for GEV models fit to the regional climate model minima, by season, in the grid boxes containing the observed station data

	DJF		MAM		JJA		SON	
Intercept	-20.55	(0.45)	-9.18	(0.45)	10.15	(0.12)	-6.08	(0.30)
Years after 1960	0.10	(0.02)	0.03	(0.02)	0.00	(0.00)	0.04	(0.01)
Scale	6.18	(0.16)	5.78	(0.16)	1.63	(0.05)	4.20	(0.12)
Shape	-0.20	(0.02)	-0.15	(0.03)	-0.01	(0.02)	-0.01	(0.03)

Table 3 shows the corresponding summary of the GEV model fits by season to the regional climate model minima for the grid boxes containing an observed site. With the exception of the estimates of the shape parameters, the direction of effects are quite similar between the two models, indicating preliminary evidence of a concordance between the climate model output minima and the observed minima over south central Sweden. However, we should be cautious in interpreting too much from this non-spatial analysis. In Sections 4.2 and 4.3 we use spatially varying parameter models to more carefully examine spatial effects, choosing to allow location and scale parameters to vary in space, but keeping the shape parameter constant.

Bias correction methods for regional model output relative to data in the literature (e.g., Bordoy and Burlando, 2013) tend to focus on location rather than scale. In order to reduce the variability of model output to that of the data, using Doksum's shift function method (Doksum and Sievers, 1976), we can transform each model output to the distribution of the data. For an example, see Figure 5. This figure demonstrates anomalous behavior of the regional climate model around 0°C. The climate model produces no minimum open air temperatures in an area around freezing temperatures (for example, for the grid square containing Station 1 there are no RCA minima between about -2 and 5°C, but there are observed minima in that range of values). Part of the reason for this is the strong seasonal relationships that we see in the minima, but also there may be clustering effects from the reanalysis model that drives the regional model (see Kharins and Zwiers (2000), p. 3764). Examining Figure 4 again we that minima around 0°C for the regional climate model are most likely in the autumn (SON). They are less likely for the spring (MAM) and summer, and unlikely in the winter (DJF).

The difficulty in extending this approach further is to develop a spatial model for the shift functions. This requires more observational data than we have available. Nikulin *et al.* (2011) use gridded data from E-OBS (Haylock *et al.*, 2008), but these have the disadvantage that the gridding approach smooths the original data, and thus does not preserve

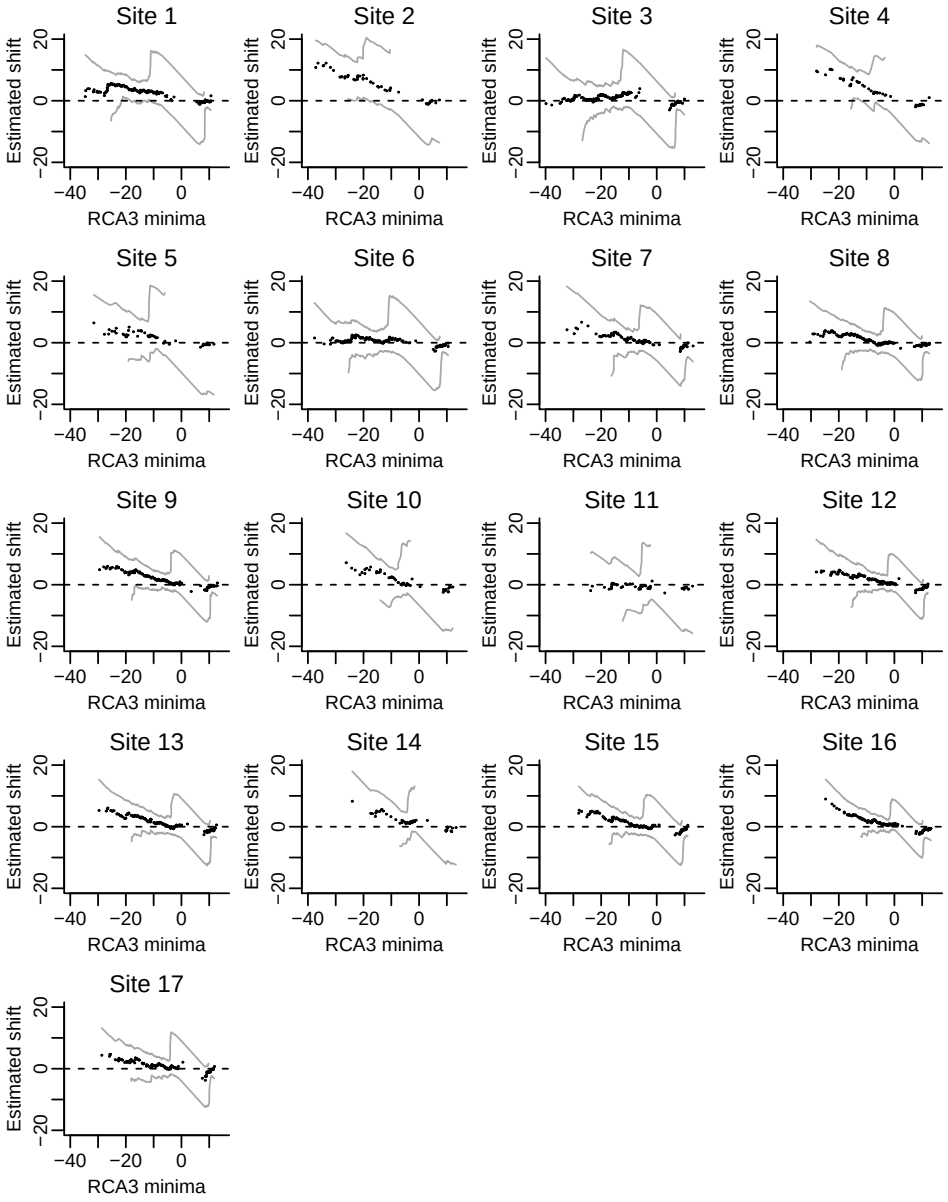

Figure 5 – Estimates of the Doksum shift function, as a function of the seasonal RCA3 minimum value along with simultaneous 95% confidence intervals.

extremes.

4.2. *Fitting the hierarchical Bayesian GEV model to the observed seasonal minima*

We fit the hierarchical Bayesian GEV model described in Section 3 to the observed seasonal minima using Markov chain Monte Carlo. We assessed convergence of the Markov chains by inspecting trace plots from two parallel chains, started from different starting conditions. For each chain we removed the first 2,000 draws as burn-in. Our results are based on running each of the two chains for 50,000 further draws, thinning every 10. This gave us 10,000 samples from the posterior.

The second and third columns of Table 4 displays marginal summaries of the posterior distributions of the hyperparameters in the model. In agreement with the models presented earlier in this section there is strong evidence that the shape parameter is negative, and hence the distribution of seasonal minima is bounded. The posterior summary for the parameter λ_1 (the mean of the geostatistical process for $\{\beta_1(s)\}$, the spatially-varying process for a yearly trend in the location parameter) indicates weak evidence of warming of the minima. The posterior mean value corresponds to an increase of 0.05 degrees celsius per year. The parameters λ_2, λ_3 and λ_4 are positive indicating that the non-winter minima are warmer. From this parameter alone there is suggestion the SON effect is the larger than the MAM and JJA effect. However this does not account for the larger spatial variability in the parameters for the MAM and JJA seasons.

Figure 6 displays summaries of the spatially varying parameters at each regional climate model grid box centroid, based on producing samples of the process conditional on the posterior distribution of the parameters. For the spatially varying seasonal effects, we choose to summarize the intercepts for each season (DJF: $\beta_0(s)$; MAM: $\beta_0(s)+\beta_2(s)$; JJA: $\beta_0(s)+\beta_3(s)$; SON: $\beta_0(s)+\beta_4(s)$). Given the form of the model, these spatially-varying intercepts correspond to the posterior location parameter for the four seasons in 1960. There are strong spatial trends in these four seasonal spatially-varying intercepts. Regardless of the season, the minima are colder in the north and toward the center (which also tends to be at a higher elevation), and warmer to the south east. Comparing the seasons we can see that the spatial patterns do differ, and that the seasonal effects are asymmetric (the MAM and SON effect are not identical). As expected the minima tend to be warmer in the summer, JJA. The spatially-varying slopes characterizing the year effect (relative to 1960) in the location parameter of the GEV distribution also has a distinctive spatial pattern. It is strongest (the location parameter increases by around 0.10°C per year on average) in the center of the north of the region of interest, and weakest (the location parameter increases around 0.04°C per year on average) in the south and west.

4.2.1. *Model verification*

Quantile plots (see, e.g., Coles, 2001, Section 6.2.3) were used to assess the distributional assumptions made by the GEV model on a site-by-site basis. The quantile plots for the model fit to the observational data, shown in Figure 7, indicate excellent goodness of fit at all locations (the majority of the values are within the simultaneous 95% confidence limits). The goodness of fit of this model is improved over the fit of a model in which the

TABLE 4
Posterior means and associated 95% credible intervals (CIs) for the GEV model parameters fit to the observed seasonal minima, and the regional climate model seasonal minima.

	Observed stations		RCM	
Parameter	Post. mean	95% CI	Post. mean	95% CI
λ_0	-4.09	(-9.98, 2.05)	-17.86	(-20.24, -11.74)
τ_0	62.60	(25.15, 141.67)	9.84	(4.43, 44.81)
ϕ_0	6.67	(3.84, 10.37)	14.92	(9.53, 19.30)
λ_1	0.05	(-0.05, 0.16)	0.04	(0.02, 0.06)
τ_1	0.00	(0.00, 0.01)	0.00	(0.00, 0.00)
ϕ_1	6.09	(3.14, 9.90)	15.88	(1.01, 21.80)
λ_2	5.52	(0.99, 8.95)	9.77	(9.07, 10.40)
τ_2	8.97	(2.58, 23.48)	0.65	(0.32, 1.78)
ϕ_2	5.50	(2.81, 9.15)	7.07	(1.77, 11.06)
λ_3	4.57	(-1.85, 10.87)	26.51	(24.16, 28.56)
τ_3	78.90	(31.74, 171.40)	5.15	(2.89, 19.29)
ϕ_3	7.60	(4.62, 11.41)	12.91	(1.01, 17.50)
λ_4	8.72	(6.75, 9.92)	12.70	(11.59, 13.65)
τ_4	0.84	(0.02, 3.37)	1.49	(0.78, 4.56)
ϕ_4	5.29	(2.58, 8.87)	8.44	(0.92, 12.05)
λ_σ	1.22	(0.90, 1.52)	1.28	(1.23, 1.34)
τ_σ	0.04	(0.01, 0.11)	0.01	(0.00, 0.02)
ϕ_σ	5.18	(2.60, 8.65)	7.76	(0.11, 11.63)
ξ	-0.18	(-0.21, -0.15)	-0.14	(-0.15, -0.14)

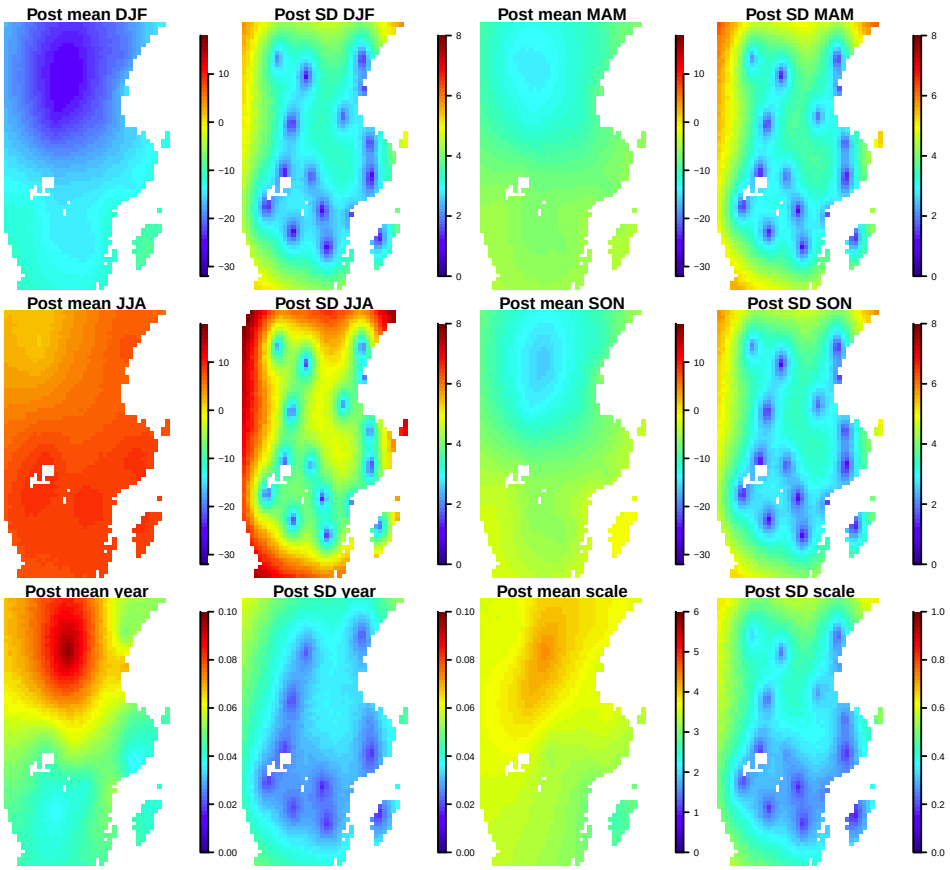

Figure 6 – Posterior means and standard deviations of the spatially varying parameters in the hierarchical Bayesian GEV model for the observed seasonal minima.

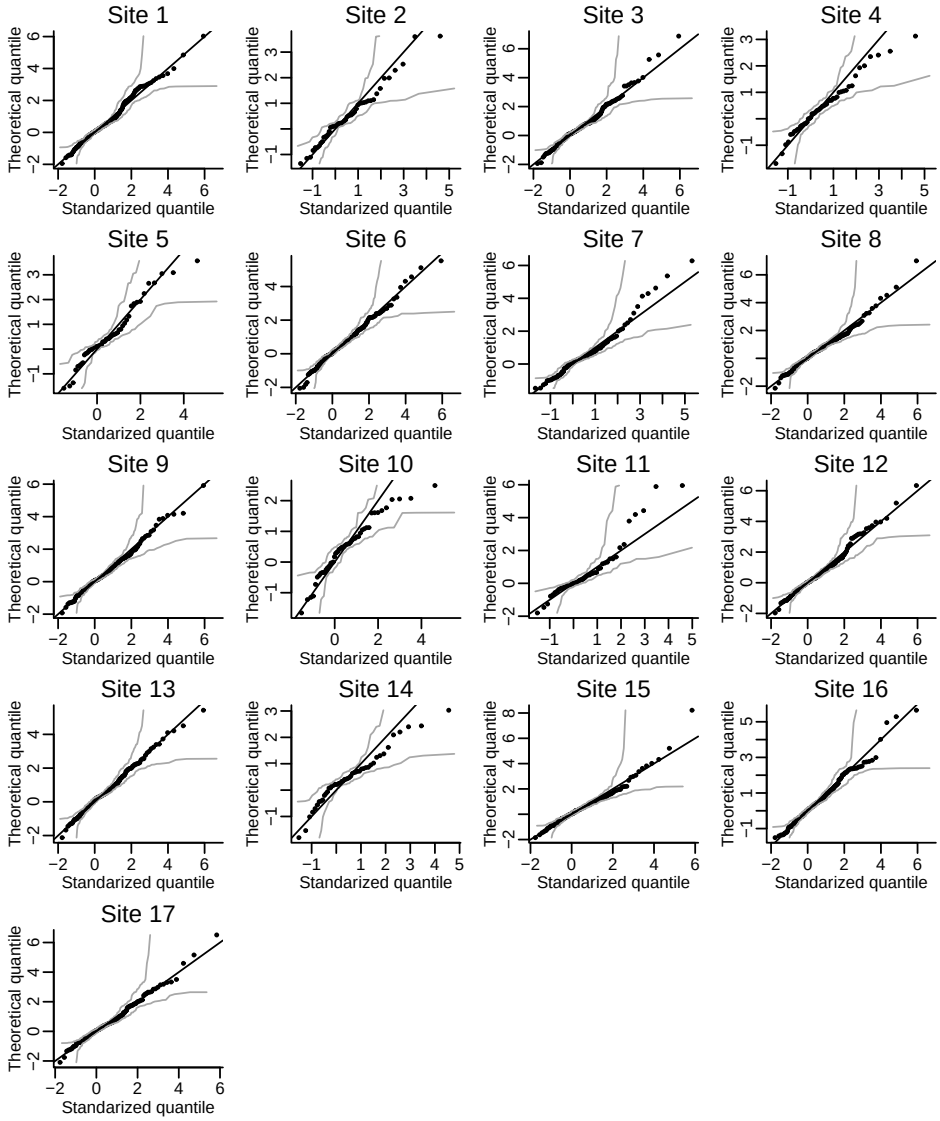

Figure 7 – Site-by-site quantile plots to assess the goodness of fit of the hierarchical Bayesian GEV model fit to observed seasonal minima.

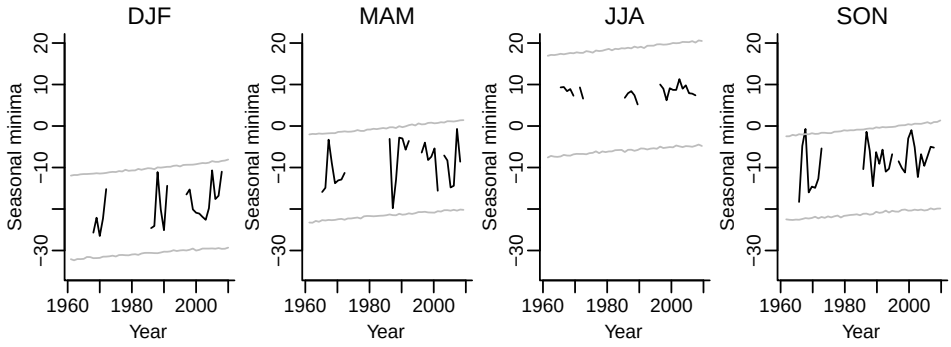

Figure 8 – By season, time series plots of the observed Borlänge seasonal minima are shown in black. The gray limits show pointwise 95% predictive intervals, generated from the hierarchical GEV model fit to the observed seasonal minima.

scale parameter of the GEV model was held fixed over spatial locations. This indicates the utility of using a hierarchical Bayesian model that borrows strength in estimating the GEV parameters over spatial locations.

To further examine the quality of the marginal GEV model fit we use the observed series at Borlänge which is located at 15° 30' E and 60° 25' N, at an elevation of 152 meters. This station has been moved twice, and has long stretches of missing data, and was not included in the GEV model fit. Let $\boldsymbol{s}_B$ denote the spatial location of the Borlänge station. At this location seasonal minima were simulated from a GEV distribution, using the posterior spatially-varying parameters for the location, $\{\beta_j(\boldsymbol{s}_B)\}$, and the scale $\sigma(\boldsymbol{s}_B)$, along with posterior distribution for the scale parameter ξ. Based on calculating pointwise 95% prediction intervals for the seasonal minima, we found that 99% of the 105 observed seasonal minima at this station were included in these prediction intervals. Time series plots of these prediction intervals by season (in gray) are shown in Figure 8. Comparing to the observed time series (shown in black) we can see that the lack of fit is due to underestimating the variation of the seasonal minima in the summer. We discuss the need to investigate this further in Section 5.

4.3. Fitting the hierarchical Bayesian GEV model to the regional climate model seasonal minima; making comparisons

We next fit the hierarchical Bayesian GEV model described in Section 3 to the RCA regional climate model output seasonal minima using Markov chain Monte Carlo. To be able to fit such a large model to the minima series at 1989 spatial locations, we made two approximations. Firstly, in updating the spatially varying GEV parameters, we calculated the acceptance ratios for Metropolis updates at each spatial location using the 15 nearest neighbors, rather than all the spatial locations. Experiments showed very little effect upon the results by increasing this number above 15. In updating the hyperparameters in the spatial models, we broke the spatial field up into 4 independent regions (northwest, northeast, southwest, and southeast). This sped up the inversions required, but again had little effect on the results. To ensure that our chains were mixing we com-

pared the results of fitting the Bayesian model to the entire dataset with results from fitting our model to smaller, randomly selected, subsets of the regional climate output. Except for an increase in the uncertainty estimates when smaller fractions of the data were used, our results were robust.

As for the observational data model, we assessed the goodness of fit of the hierarchical GEV model fit to the seasonal minima in the RCM output using quantile plots (not shown). The goodness of fit was even better than the fit for the observational minima.

The right two columns of Table 4 displays marginal summaries of the posterior distributions of the hyperparameters in the GEV model for the seasonal RCM minima, and Figure 9 shows posterior summaries of the corresponding spatially varying parameters. While the intensity scales of the posterior standard deviations differ in Figure 9 relative to Figure 6, reflective of the increase in the certainly of the spatially varying parameters in the regional climate output, we have fixed the intensity scales of the posterior means in Figures 9 and 6 to be the same for the two models. This aids in the comparison of the spatial fields.

Before discussing these spatial fields, we outline a number of conclusions made from comparing the posterior distributions of the GEV model hyperparameters for the observed minimal and the RCM minima. The uncertainty of the spatial fields (as measured by $\{\tau_j\}$ and τ_σ) is smaller for the RCM model over the observed station model. This decrease in the variance of the spatial field for the RCM model versus the station model is associated with increased spatial range parameters in the RCM model, as well as spatial mean values that are larger in magnitude for the RCM model. Comparing the 95% posterior credible intervals, there is evidence from Table 4 that the yearly linear slope in the location parameters for the GEV model for the RCM minima is significantly positive (a 95% credible interval of between 0.02 and 0.06 does not contain zero), however this slope in the location parameter may not be as large as that posited for the station model. From the marginal summaries the scale parameters may be quite similar between to the two models, but the shape parameter is larger for the RCM model versus the observed model (—0.14 versus —0.18). The could be an effect of data length, or could be due to the parameterization of the regional climate model. Exploring the spatial distributions of the location parameters and scale parameters in more detail, we can see that except for the yearly slope location effect the spatial patterns are quite similar. The biggest difference spatially is the yearly slope parameter, which the observed minima model estimates to be much higher in the north than for the RCM model. A possible reason could be that a closer comparison of Figure 6 and 9 indicates that the seasonal spatial patterns are more negative in the north for the regional climate model, possibly suggesting that the seasonal patterns are changing over time. Given the limited spatial and temporal coverage of station data in the northwest (site 2 is a very short record) it would be worth obtaining more station data for the northwest region to confirm the warming trend in this region.

5. DISCUSSION

This article demonstrates an extreme value analysis that allows us to compare the distribution of the extremes of the observations with the distribution of extremes of the climate model output. There are number of ways that we can extend our analysis. We can

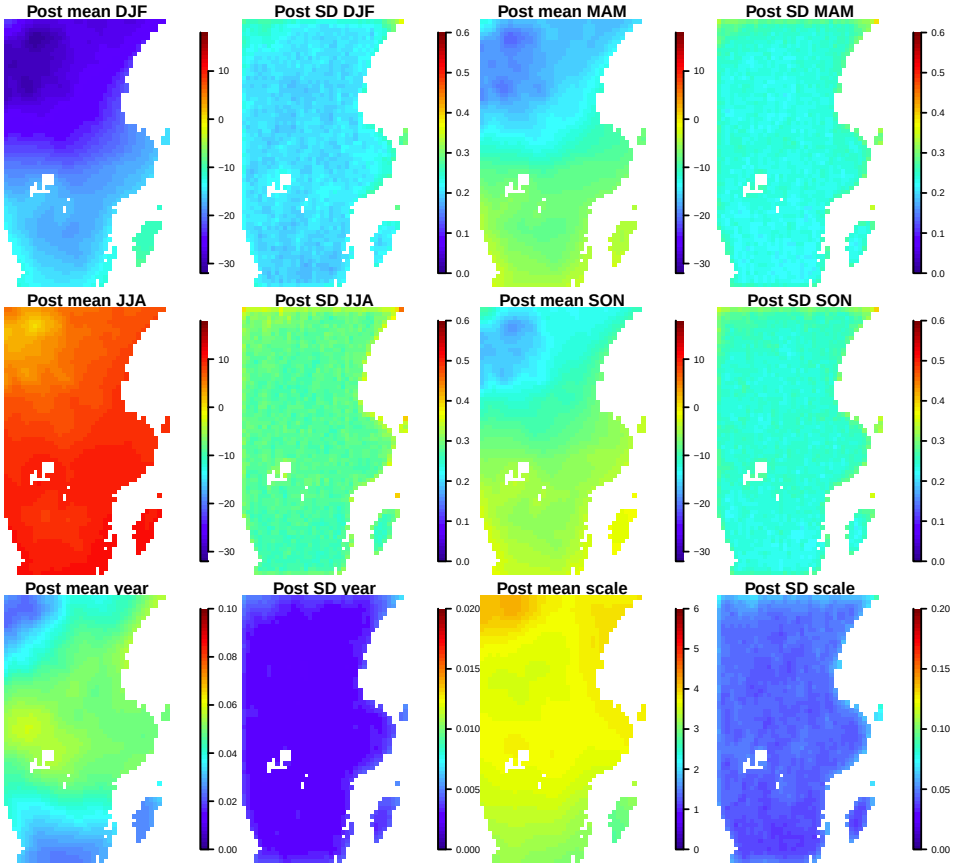

Figure 9 – Posterior means and standard deviations of the spatially varying parameters in the hierarchical Bayesian GEV model for the RCM seasonal minima.

consider more general models to compare the extreme value distributions. For example we could allow for nonlinear trends in the location parameter (see, e.g., the discussion in Gaetan and Grigoletto, 2007), and allow for more general models for the scale and shape parameters. In particular, our analysis points towards allowing the scale parameter in the GEV distribution to additionally vary by season.

It is possible to replace the block maxima approach with exceedance analysis. While this may seem to include more data, the temporal dependence in the daily data would tend to require declustering of exceedances, thereby reducing the potential advantage of exceedance analysis.

Another issue that we have not yet tackled is the change of support question. For the climate model output, what does the seasonal minimum calculated for a grid box represent? Is it the minimum over the grid box region, or does it represent the value of the minimum at the centroid of the grid box? We could answer these questions by simulating minima (based on the observation model) at a finer spatial scale than the regional climate model output, and then changing support by calculating minima at the grid box level. Answering this question, however, may require the use of multivariate models for extremes, which we introduce below.

In this article we have modeled the marginal distribution of extremes, assuming conditional independence of the extremes in space and time. We used a block minima approach. A competing marginal modeling approach would be to compare the distributions of the station data and the regional climate model output below a range of different thresholds. While threshold methods retain more values for analysis, care is required in the presence of clustering in space and in time (see, e.g., Coles, 2001). In future work we will look at modeling the dependent extreme values in space and in time. There are a number of approaches for modeling multivariate extremes. We can either take a conditional modeling approach (e.g., Ledford and Tawn, 1996; Heffernan and Tawn, 2004), a max-stable approach (see, e.g., Davison *et al.*, 2012), or a copula-based approach (e.g., Ghosh and Mallick, 2011). In practice methods of multivariate extreme value inference can be computationally demanding (e.g., Ribatet *et al.*, 2012).

ACKNOWLEDGEMENTS

We are grateful to the Swedish Meteorological and Hydrological Institute (SMHI) for making the synoptic observation network data available online. Gregory Nikulin of SMHI provided the model runs of RCA3 divided up into different land cover types. We would like to thank two anonymous reviewers who greatly improved this article and Daniela Cocchi for organizing this special issue.

REFERENCES

V. BERROCAL, P. CRAIGMILE, P. GUTTORP (2012). *Regional climate model assessment using statistical upscaling and downscaling techniques*. Environmetrics, 23, pp. 482–492.

R. BORDOY, P. BURLANDO (2013). *Bias correction of regional climate model simulations in a region of complex orography*. Journal of Applied Meteorology and Climatology, 52, pp. 82–101.

S. COLES (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag.

D. COOLEY, D. NYCHKA, P. NAVEAU (2007). *Bayesian spatial modeling of extreme precipitation return levels*. Journal of the American Statistical Association, 102, pp. 824–840.

P. F. CRAIGMILE, P. GUTTORP (2011). *Space-time modelling of trends in temperature series*. Journal of Time Series Analysis, 32, pp. 378–395.

A. C. DAVISON, S. A. PADOAN, M. RIBATET (2012). *Statistical modelling of spatial extremes*. Statistical Science, 27, pp. 161–186.

K. A. DOKSUM (1974). *Empirical probability plots and statistical inference for nonlinear models in the two sample case*. Annals of Statistics, 2, pp. 267–277.

K. A. DOKSUM, G. L. SIEVERS (1976). *Plotting with confidence: Graphical comparisons of two populations*. Biometrika, 63, pp. 421–434.

C. GAETAN, M. GRIGOLETTO (2007). *A hierarchical model for the analysis of spatial rainfall extremes*. The Journal of Agricultural, Biological and Environmental Statistics, 12, pp. 434–449.

S. GHOSH, B. K. MALLICK (2011). *A hierarchical Bayesian spatio-temporal model for extreme precipitation events*. Enivronmetrics, 22, pp. 192–204.

M. R. HAYLOCK, N. HOFSTRA, A. M. G. KLEIN TANK, E. J. KLOK, P. D. JONES, CO AUTHORS (2008). *European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006*. Journal of Geophysical Researh, 113.

J. E. HEFFERNAN, J. A. TAWN (2004). *A conditional approach for multivariate extreme values*. Journal of the Royal Statistical Society, Series B, 66, pp. 497–546.

V. V. KHARINS, F. W. ZWIERS (2000). *Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM*. Journal of Climate, 13, pp. 3760–3788.

E. KOCMÁNKOVÁ, M. TRNKA, J. EITZINGER, H. FORMAYER, M. DUBROVSKÝ, D. SEMERDOVÁ, Z. ŽALUD, J. JUROCH, M. MOŽNÝ (2010). *Estimating the impact of climate change on the occurrence of selected pests in the Central European region*. Climate Research, 44, pp. 95–105.

A. W. LEDFORD, J. A. TAWN (1996). *Statistics for near independence in multivariate extreme values*. Biometrika, 83, pp. 169–187.

Y. MA, P. GUTTORP (2013). *Estimating daily mean temperature from synoptic climate observations*. International Journal of Climatology, 33, pp. 1264–1269.

E. MANNSHARDT, P. F. CRAIGMILE, M. P. TINGLEY (2013). *Statistical modeling of extreme value behavior in North American tree-ring density series*. Climatic Change, 117, pp. 843–858.

L. MARINI, M. P. AYRES, A. BATTISTI, M. FACCOLI (2012). *Climate affects severity and altitudinal distribution of outbreaks in an eruptive bark beetle*. Climatic Change, 115, pp. 327–341.

G. NIKULIN, E. KJELLSTRÖM, U. HANSSON, G. STRANDBERG, A. ULLERSTIG (2011). *Evaluation and future projections of temperature, precipitation and wind extremes over Europe in an ensemble of regional climate simulations*. Tellus A, 63, pp. 41–55.

R DEVELOPMENT CORE TEAM (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL `http://www.R-project.org`. ISBN 3-900051-07-0.

M. RIBATET, D. COOLEY, A. C. DAVISON (2012). *Bayesian inference from composite likelihoods, with an application to spatial extremes*. Statistica Sinica, 22, pp. 813–845.

M. RUMMUKAINEN (2010). *State-of-the-art with regional climate models*. WIREs Climate Change, 1, pp. 82–96.

P. SAMUELSSON, C. G. JONES, U. WILLÉN, A. ULLERSTIG, S. GOLLVIK, U. HANSSON, C. JANSSON, E. KJELLSTRÖM, G. NIKULIN, K. WYSER (2011). *The Rossby Centre regional climate model RCA3: model description and performance*. Tellus A, 63, pp. 4–23.

H. SANG, A. E. GELFAND (2010). *Continuous Spatial Process Models for Spatial Extreme Values*. Journal of Agricultural, Biological and Environmental Statistics, 15, pp. 49–65.

S. SOLOMON, D. QIN, M. MANNING, *et al.* (eds.) (2007). *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, Cambridge.

S. UPPALA, D. DEE, S. KOBAYASHI, P. BERRISFORD, A. SIMMONS (2008). *Towards a climate data assimilation system: status update of ERA-Interim*. ECMWF Newsletter, pp. 12–18.

S. UPPALA, P. KÖLBERG, A. SIMMONS, U. ANDRAE, V. DA COSTA BECHTOLD, M. FIORINO, J. GIBSON, J. HASELER, A. HERNANDEZ, G. KELLY, X. LI, K. ONOGI, S. SAARINEN, N. SOKKA, R. ALLAN, E. ANDERSSON, K. ARPE, M. BALMASEDA, A. BELJAARS, L. VAN DE BERG, J. BIDLOT, N. BORMANN, S. CAIRES, F. CHEVALLIER, A. DETHOF, M. DRAGOSAVAC, M. FISHER, M. FUENTES, S. HAGEMANN, E. HÓLM, B. HOSKINS, L. ISAKSEN, P. JANSSEN, R. JENNE, A. MCNALLY, J.-F. MAHFOUF, J.-J. MORCRETTE, N. RAYNER, R. SAUNDERS, P. SIMON, A. STERL, K. TRENBERTH, A. UNTCH, D. VASILJEVIC, P. VITERBO, J. WOOLLEN (2005). *The ERA-40 re-analysis*. Quarterly Journal of the Royal Meteorological Society, 131, pp. 2961–3012.

F. W. ZWIERS, Z. ZHANG, Y. FENG (2011). *Anthropogenic influence on long return period daily temperature extremes at regional scales*. Journal of Climate, 24, pp. 881–892.

SUMMARY

Can a regional climate model reproduce observed extreme temperatures?

Using output from a regional Swedish climate model and observations from the Swedish synoptic observational network, we compare seasonal minimum temperatures from model output and observations using marginal extreme value modeling techniques. We make seasonal comparisons using generalized extreme value models and empirically estimate the shift in the distribution as a function of the regional climate model values, using the Doksum shift function. Spatial and temporal comparisons over south central Sweden are made by building hierarchical Bayesian generalized extreme value models for the observed minima and regional climate model output. Generally speaking the regional model is surprisingly well calibrated for minimum temperatures. We do detect a problem in the regional model to produce minimum temperatures close to 0°C. The seasonal spatial effects are quite similar between data and regional model. The observations indicate relatively strong warming, especially in the northern region. This signal is present in the regional model, but is not as strong.