A non-stationary index-flood model for precipitation extremes in transient Regional Climate Model simulations

Martin Hanel,[1] T. Adri Buishand,[1] and Christopher A. T. Ferro[2]

[1] Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands

[2] School of Engineering, Computing and Mathematics, University of Exeter, Exeter, UK

Abstract

The Generalized Extreme Value (GEV) distribution has often been used to describe the distribution of daily maximum precipitation in observed and climate model data. The model developed in this paper allows the GEV location parameter to vary over the region, while the dispersion coefficient (the ratio of the GEV scale and location parameters) and the GEV shape parameter are assumed to be constant over the region. This corresponds with the index-flood assumption in hydrology. It is further assumed that all three GEV parameters vary with time such that the relative change in a quantile of the distribution is constant over the region. This non-stationary model is fitted to the 1-day summer and 5-day winter precipitation maxima in the river Rhine basin in a simulation of the RACMO regional climate model for the period 1950–2099 and the results are compared with gridded observations. Except for an underestimation of the dispersion coefficient of the 5-day winter maxima by about 35% the GEV parameters obtained from the observations are reasonably well reproduced by RACMO. A positive trend in the dispersion coefficient is found in the summer season, which implies that the relative increase of a quantile increases with increasing return period. In the winter season there is a positive trend in the location parameter and a negative trend in the shape parameter. For large quantiles the latter counterbalances the effect of the increase of the location parameter. It is shown that the standard errors of the parameter estimates are significantly reduced in the regional approach compared to those of the estimated parameters from individual grid box values, especially for the summer maxima.

1. Introduction

Regional climate models (RCMs) nested inside a global climate model provide useful information about potential local climate change. Precipitation extremes in RCM simulations have been analyzed in different ways. One method is to consider the change in a large empirical quantile of the daily precipitation amounts (e.g., the 99th percentile) or the properties of the exceedances of such a quantile [e.g., *Durman et al*., 2001; *Christensen and Christensen*, 2004]. An alternative is to fit an extreme-value distribution to the largest daily precipitation amount in a season [e.g., *Frei et al*., 2006; *Beniston et al*., 2007; *Goubanova and Li*, 2007] or year [e.g., *Huntingford et al*., 2003; *Fowler et al.*, 2005; *Ekström et al*., 2005]. Maxima of multi-day precipitation amounts are treated similarly in several of these studies.

A problem with extreme precipitation is that the likelihood of detecting a systematic change at a single grid box is generally small due to the large year-to-year variability. *Frei and Schär* [2001] mention, for instance, that a frequency change by a factor of 1.5 in daily events with an average return period of 100 days can be detected with a probability of only 0.2 in a 100-year record, assuming a smoothly varying trend component and temporal independence of extreme events. The decrease of this probability with increasing event magnitude limits the detection of systematic changes in extreme events at a single grid box.

Spatial pooling has been used to detect meaningful changes in extremes. *Frei et al*. [2006] and *Goubanova and Li* [2007] averaged an estimated quantile of the extreme-value distribution over large regions. *Kendon et al.* [2008] studied the effectiveness of spatial pooling for the detection of changes in the 95th percentile of wet-day precipitation. An alternative is to assume that the most uncertain parameters of the extreme-value distribution are constant over some region. The estimates of these parameters based on the pooled data across the region are then generally more precise than those from the data of an individual grid box, leading to a reduction of the standard errors of the estimated quantiles of the distribution. This approach has its origin in hydrology where it is known as regional frequency analysis. Although biases will be introduced when the homogeneity assumptions are not met, simulation studies [e.g., *Lettenmaier et al.*, 1987; *Hosking and Wallis*, 1997] show that even in regions with moderate amounts of heterogeneity, a regional frequency analysis is more accurate than the at-site analysis.

The most popular method of regional frequency analysis is the index-flood method. *Fowler et al*. [2005] and *Ekström et al*. [2005] applied this method to the 1-, 2-, 5-, and 10-day annual maximum precipitation amounts across the UK in two RCM simulations. Apart from a change in the distribution parameters between the control and future climate, these parameters do not vary over time in their application.

The purpose of this paper is to introduce an index-flood model with time-varying parameters as a tool to summarize changes of extreme precipitation in transient RCM simulations. The model is applied to daily precipitation in the river Rhine basin in the RACMO-ECHAM5 simulation. In this part of Europe, short-period convective storms may cause local flooding in summer, whereas in winter multi-day episodes may have adverse impacts over large areas. As in *Frei et al.* [2006], we analyze the 1-day precipitation maxima in summer and the 5-day precipitation maxima in winter.

The index-flood model is described in section 2. Section 3 provides some information about the river Rhine basin, the RACMO-ECHAM5 simulation, and the observational data sets that were used for validation. The results for the summer maxima are presented in section 4 and those for the winter maxima in section 5. Section 6 presents the conclusions.

2. Regional modeling of non-stationary precipitation extremes

2.1. Index-flood model

The idea behind the index-flood method is that the variables within a homogeneous region are identically distributed after scaling with a site-specific factor, the index flood. The T-year quantile $Q_T(s)$ of the distribution of the variable $X(s)$ at any given site s, i.e., the value that is exceeded with probability $1/T$, can then be written as

$$Q_T(s) = \mu(s) q_T, \qquad (1)$$

where $\mu(s)$ is the index flood and q_T is a regional, dimensionless quantile function, in this context often called the growth curve. The mean or median of the distribution of $X(s)$ is usually chosen as the index flood.

A consequence of the index-flood assumption is that the coefficient of variation of $X(s)$ should be constant over the region of interest. This property is useful for identifying homogeneous regions. A number of authors have found that the coefficient of variation of the observed annual maximum precipitation is relatively large in dry areas and small in wet, mountainous regions [see *Brath et al.*, 2003]. Nevertheless, the spatial variation in the coefficient of variation of precipitation maxima is generally much less than that in the mean.

The index-flood method has been used with different probability models for the distribution of $X(s)$. For seasonal and annual precipitation maxima the generalized extreme value (GEV) distribution is popular. This is a three-parameter distribution that combines the three possible types of extreme value distributions (i.e., Gumbel, Fréchet, and reverse Weibull distributions). Its distribution function is given by

$$F(x) = \exp\left\{-\left[1+\kappa\left(\frac{x-\xi}{\alpha}\right)\right]^{-\frac{1}{\kappa}}\right\}, \qquad \kappa \neq 0,$$

$$\tag{2}$$

$$F(x) = \exp\left\{-\exp\left[-\left(\frac{x-\xi}{\alpha}\right)\right]\right\}, \qquad \kappa = 0,$$

with ξ, α, and κ the location, scale, and shape parameters, respectively. The shape parameter controls the behavior of the tails of the distribution – positive values imply a heavy upper tail (Fréchet distribution).

Apart from support from extreme value theory [e.g., *Coles*, 2001], the GEV distribution has often been found to describe the distribution of observed or simulated precipitation maxima well. For annual precipitation maxima of various durations *Schaefer* [1990], *Alila* [1999], and *Kyselý and Picek* [2007], using L-moment ratio diagrams, observed that the GEV distribution is generally superior to other candidate distributions. In addition, *Alila* [1999] and *Kyselý and Picek* [2007] found that a goodness of fit test based on the L-kurtosis did not reject the GEV distribution. *Buonomo et al.* [2007] and *Goubanova and Li* [2007] used the Kolmogorov-Smirnov goodness of fit test and concluded that the GEV distribution is appropriate for modeling precipitation extremes in RCM projections for most parts of Europe, although problems were met in dry areas where most of the seasonal maxima were zero.

For the development of our non-stationary GEV model it is convenient to use the location parameter as the index flood, i.e., $\mu(s) = \xi(s)$, rather than the mean or the median. If the seasonal maximum $X(s)$ at site s follows a GEV distribution with parameters $\xi(s)$, $\alpha(s)$, and $\kappa(s)$, then the scaled seasonal maximum $X(s)/\xi(s)$ has a GEV distribution with location parameter 1, scale parameter $\gamma(s) = \alpha(s)/\xi(s)$, and shape parameter $\kappa(s)$. The index-flood method applies if $\gamma(s)$ and $\kappa(s)$ do not vary over the region, i.e., $\gamma(s) = \gamma$ and $\kappa(s) = \kappa$. The dispersion coefficient γ is analogous to the coefficient of variation.

The T-year quantile of the scaled seasonal maximum $X(s)/\xi(s)$ follows from equation (2) by setting $F(q_T) = 1–1/T$, $\xi = 1$, and $\alpha = \gamma$:

$$q_T = 1 - \frac{\gamma}{\kappa}\left\{1 - \left[-\log\left(1 - \frac{1}{T}\right)\right]^{-\kappa}\right\}, \qquad \kappa \neq 0, \tag{3}$$

$$q_T = 1 - \gamma \log\left[-\log\left(1-\frac{1}{T}\right)\right], \qquad \kappa = 0 .$$

Note that $q_T = 1$ and $Q_T(s) = \xi(s)$ when $T = 1/(1–1/e) = 1.58$ years, the return period corresponding to the location parameter. The growth curve is determined by γ and κ. This is also the case if $X(s)$ is scaled by the mean [*Buishand*, 1991; *Sveinsson et al.*, 2001] or the median [*Northrop*, 2004]. However, the index flood then depends on γ and κ, which is inconvenient in the case of temporal trends in these parameters.

2.2. Non-stationary index-flood model

A few studies in the hydrological literature deal with non-stationarity in regional frequency analysis. *Cunderlik and Burn* [2003] assume temporal and spatial variation in both the location and scale parameter of the distribution. Linear trends in these parameters were estimated with a distribution-free method due to *Sen* [1968]. In a subsequent paper [*Cunderlik and Ouarda*, 2006] the scale parameter was assumed to be constant over the region of interest but still time-varying. The regional scale parameter was estimated as a weighted average of the at-site scale parameters. *Renard et al.* [2006] used a regional non-stationary GEV model to describe trends in annual maximum discharges. In that model the shape parameter was constant but the scale and location parameters varied over the region and there was a common linear trend in the location parameter. Statistical inference was based on a Bayesian analysis using Markov chain Monte Carlo methods. Other authors have successfully used a GEV distribution with time-varying parameters, e.g., *Kharin and Zwiers* [2005], *Adlouni et al.* [2007], *García et al.* [2007], and *Brown et al.* [2008], although not in the framework of regional frequency analysis.

Let $X(s, t)$ be the seasonal maximum at site s in year t. Using the location parameter of the GEV distribution as the index flood, the T-year quantile $Q_T(s, t)$ can be represented as

$$Q_T(s,t) = \xi(s,t) q_T(t), \tag{4}$$

where $q_T(t)$ is given by equation (3) but with time-dependent dispersion coefficient $\gamma(t)$ and shape parameter $\kappa(t)$. The location parameter $\xi(s, t)$ varies both in time and space. As in the non-stationary GEV model of *Renard et al.* [2006], the temporal trend in the location parameter is assumed to be constant over the region of interest. A motivation for this is that changes in extreme precipitation are mainly associated with large-scale changes in the atmospheric conditions (changes of the amount of precipitable water due to temperature change and changes of the atmospheric circulation). However, in regions with strong orography the changes in precipitation may be altitude-dependent [*Giorgi et al.*, 1997]. The altitude-dependence of the trend in the location parameter will be examined for the mountainous southern part of the Rhine basin.

We propose the following model for the GEV parameters:

$$\xi(s,t) = \xi_0(s) \exp[\xi_1 I(t)] \tag{5}$$

$$\gamma(t) = \exp[\gamma_0 + \gamma_1 I(t)] \tag{6}$$

$$\kappa(t) = \kappa_0 + \kappa_1 I(t) \tag{7}$$

where $I(t)$ is a time indicator or time-dependent covariate, the choice of which is discussed in section 3. Different forms of trends can be considered, but our choices have the following advantages. The dispersion coefficient cannot become negative because of the exponential expression in equation (6). The exponential function in equation (5) ensures that the relative changes in the quantiles are constant over the region of interest, as follows. From equations (4) and (5), the relative change of the T-year quantile between years t_1 and t_2 at site s can be written as

$$\frac{Q_T(s,t_2)}{Q_T(s,t_1)} = \frac{\xi(s,t_2)}{\xi(s,t_1)}\frac{q_T(t_2)}{q_T(t_1)} = \exp\{\xi_1[I(t_2)-I(t_1)]\}\frac{q_T(t_2)}{q_T(t_1)}, \qquad t_2 \geq t_1, \tag{8}$$

which does not depend on s. Apart from the common usage of percentages for changes in extreme precipitation, a reason to assume constant relative changes rather than absolute changes is that specific humidity and hence atmospheric moisture would increase roughly exponentially with temperature (about 6.5% per degree) according to the Clausius-Clapeyron relation [e.g., *Pall et al.*, 2007].

The parameters $\xi_0(s)$, ξ_1, γ_0, γ_1, κ_0, and κ_1 of the model were estimated by maximizing the log-likelihood

$$L = \sum_{s=1}^{S}\sum_{t=1}^{N} L_{s,t}\left(\xi_0(s), \xi_1, \gamma_0, \gamma_1, \kappa_0, \kappa_1\right) \tag{9}$$

where $L_{s,t}(\xi_0(s), \xi_1, \gamma_0, \gamma_1, \kappa_0, \kappa_1)$ is the log-likelihood for the seasonal maxima at grid box s in year t, S is the number of grid boxes in the region and N is the number of years in the record. The number of parameters that has to be determined is thus S+5. Dealing usually with more than 50 grid boxes in one region it was difficult to estimate all parameters simultaneously. Therefore, a two-step procedure was applied [*Arnell and Gabriele*, 1988; *Buishand*, 1991]. Initial values of the parameters were based on L-moments estimates [*Hosking and Wallis*, 1997]. For the parameters $\xi_0(s)$ the individual grid box estimates were used, and the parameters γ_0 and κ_0 were set to the regional average of the grid-box estimates. The trend parameters ξ_1, γ_1, and κ_1 were set initially to zero. In the first step, all the site-specific location parameters $\xi_0(s)$ were estimated by maximum likelihood, keeping the regional parameters ξ_1, γ_0, γ_1, κ_0, and κ_1 fixed. In the second step, the values of $\xi_0(s)$ were fixed at their estimates from the previous step and the regional parameters were estimated by maximum likelihood. These two steps were repeated until convergence. The number of iterations needed for the procedure to converge was usually not more than 5 for the summer and not more than 10 for the winter maxima. The CPU time needed to fit the index-flood model was on average 10% larger in summer and 70% larger in winter than the time needed to fit the model to each of the corresponding grid boxes individually.

2.3. Uncertainty and model checking

The log-likelihood in equation (9) assumes independence between years and between grid boxes within the region. In particular, the latter assumption is not satisfied because the seasonal maxima at adjacent grid boxes are often associated with the same meteorological event. As a consequence, the standard errors of the estimates can no longer be obtained from the second derivatives of the log-likelihood. The bootstrap can be used to assess the uncertainty of the parameters and quantiles of the distribution in the case of spatial dependence. Rather than bootstrapping the data of the grid boxes individually, the data for a certain year are bootstrapped simultaneously in order to preserve the spatial dependence [cf. *Faulkner and Jones*, 1999; *Kharin et al.*, 2007]. Since resampling requires that the data come from the same distribution, the trend is removed from the maxima $X(s, t)$ by the transformation [*Coles*, 2001]

$$\widetilde{X}(s,t) = \frac{1}{\hat{\kappa}(t)} \log\left[1 + \frac{\hat{\kappa}(t)}{\hat{\gamma}(t)}\left(\frac{X(s,t)}{\hat{\xi}(s,t)} - 1\right)\right], \tag{10}$$

where $\widetilde{X}(s,t)$ are the detrended seasonal maxima and $\hat{\xi}(s,t)$, $\hat{\gamma}(t)$, and $\hat{\kappa}(t)$ are the maximum likelihood estimates of the GEV parameters (these are obtained by replacing $\xi_0(s)$, ξ_1, γ_0, γ_1, κ_0, and κ_1 in equations (5)–(7) by their maximum likelihood estimates $\hat{\xi}_0(s), \hat{\xi}_1, \hat{\gamma}_0, \hat{\gamma}_1, \hat{\kappa}_0$, and $\hat{\kappa}_1$). Then a sample $t_1,\ldots,t_u,\ldots,t_N$ is drawn with replacement from the years $1,\ldots,N$. A bootstrap sample of detrended seasonal maxima is obtained by taking the vector $\left(\widetilde{X}(1,t_u),\ldots,\widetilde{X}(s,t_u),\ldots,\widetilde{X}(S,t_u)\right)$ for each resampled year t_u. Finally, the sample is transformed back to the original scale according to

$$X(s,u) = \hat{\xi}(s,u)\left\{1 + \hat{\gamma}(u)\frac{\exp[\hat{\kappa}(u)\widetilde{X}(s,t_u)] - 1}{\hat{\kappa}(u)}\right\} \tag{11}$$

and the parameters are re-estimated.

The transformed maxima $\widetilde{X}(s,t)$ should have a standard Gumbel distribution if the model is correct (we refer to them as standard Gumbel residuals hereafter), which is tested in this study by calculating the Anderson-Darling statistic for each grid box. The Anderson-Darling statistic A^2 is defined as [*Anderson and Darling*, 1952]

$$A^2 = N\int_{-\infty}^{\infty}\frac{[F_N(x) - F(x)]^2}{F(x)[1 - F(x)]}\,dF(x)\,, \tag{12}$$

where $F_N(x)$ is the empirical distribution of the $\widetilde{X}(s,t)$ for the grid box of interest and $F(x)$ is the standard Gumbel distribution function, $F(x) = \exp[-\exp(-x)]$. The A^2 statistic summarizes the mean square distance between the two distributions, putting more weight on the tails of the distribution through the function $1/\{F(x)\,[1-F(x)]\}$. For testing goodness of fit of extreme value distributions it has been shown [e.g., *Shimokawa and Liao*, 1999; *Laio*, 2004] that this statistic is more powerful than the Kolmogorov-Smirnov and Cramer-von Mises statistics and the probability plot correlation coefficient. Here A^2 also tests the adequacy of assumptions about the GEV parameters (the index-flood assumption, constant trends over the region of interest and (log-)linearity with the time indicator $I(t)$). Separate tests for these assumptions can be designed but these are not considered in the present paper. The definition of the region should be re-examined or a different model for the GEV parameters should be used if the fit is not acceptable.

The procedures used to assess uncertainty and goodness of fit assume independence between years. This assumption has been checked by exploring the temporal pattern of residuals. For this purpose, it is convenient to work with residuals that have a symmetric distribution, in particular the normal distribution. Standard normal residuals $\widetilde{X}_{\mathrm{norm}}(s,t)$ are obtained by the transformation

$$\widetilde{X}_{\mathrm{norm}}(s,t) = \Phi^{-1}\left\{\exp\left[-\exp\left(-\widetilde{X}(s,t)\right)\right]\right\}, \qquad (13)$$

with Φ^{-1} the quantile function of the standard normal distribution.

3. Rhine basin and data used

The river Rhine basin has an area of 185,000 km^2 and is situated in the territory of nine European countries (Figure 1a). The basin stretches from the Alps in the south with mountain peaks higher than 4000 m to a flat delta in the Netherlands in the north. Mean annual precipitation is quite variable – the wettest part is the Alpine region with more than 3000 mm of precipitation per year in some areas, the driest part is the area around

Mainz in the center of the Rhine basin where mean annual precipitation is about 400 mm. The overall mean annual precipitation is 910 mm.

The precipitation maxima in the output of the KNMI regional climate model RACMO [*van Meijgaard et al.*, 2008] driven by the ECHAM5 global climate model [*Jungclaus et al.*, 2006] under the SRES A1B emission scenario [*Nakićenović and Swart*, 2000] for the period 1950–2099 were studied. The horizontal resolution of the RACMO model is ≈ 25 km on a rotated longitude-latitude grid. There are 316 grid boxes whose centers lie within the Rhine basin (Figure 1b).

To use the index-flood model homogeneous regions have to be identified. *Hosking and Wallis* [1997] mention several methods for choosing the regions ranging from subjective partitioning to using geographical units and objective partitioning. The latter still requires subjective choices at several stages. We split the Rhine basin into regions subjectively: we estimated the GEV parameters at each grid box for the 1-day summer (JJA) and 5-day winter (DJF) maxima for two time slices (1950–1989 and 2070–2099) using the stationary model, i.e., with $I(t) = 0$ in equations (5)–(7). Since the grid box estimates of the shape parameter are not very reliable, we based the division of the Rhine basin on the spatial pattern of the dispersion coefficient. Spatial heterogeneity of the dispersion coefficient turned out to be stronger for the summer maxima (Figure 1b–c) than for the winter maxima and therefore has more influence on the delimitation of the regions. On the basis of Figure 1b–c we divided the Rhine basin into 5 regions (Figure 1d), each including 48 to 97 grid boxes. Region 1 corresponds roughly to the Swiss part of the basin and region 5 to the Dutch part. The sensitivity of the results to the boundaries of the regions was briefly checked by moving a few grid boxes from one region to another region and refitting the model. There was little change in the estimated parameters and the goodness of fit.

Figure 2 shows the change of the mean seasonal and annual precipitation between the periods 1950–1989 and 2070–2099. In the model output mean annual precipitation increases by about 5% over the whole basin, mean winter precipitation increases by more than 20% over most of the basin and mean summer precipitation decreases by 10–20%.

The model for the GEV parameters defined in equations (5)–(7) requires the choice of the time indicator $I(t)$. The most straightforward approach is to use $I(t) = t$. Since the enhanced greenhouse effect is small during the first decades of the RCM simulation, a more complicated function of the year t is needed to allow the GEV parameters to stay constant or almost constant in this period. Such a function usually contains one or more unknown parameters which generally leads to more uncertain trend estimates. An alternative time indicator which is representative of the enhanced greenhouse effect is the global temperature. In our application a seasonal global temperature anomaly from the driving ECHAM5 model is used. This anomaly is calculated with respect to the overall 1950–2099 mean temperature so that the parameters $\xi_0(s)$ are approximately orthogonal to the regional parameters ξ_1, γ_0, γ_1, κ_0, and κ_1. This significantly speeds up the two-stage estimation procedure. Using temperature anomalies with respect to some historical period such as 1960–1989 (or temperature itself) leads to a significant correlation between

$\hat{\xi}_0(s)$ and $\hat{\xi}_1$. For example, if the historical period 1960–1989 is considered, the average correlation between these parameters is –0.87. This correlation is only 0.14 if the anomalies are calculated with respect to the overall mean. The summer and winter global temperature anomaly is given in Figure 3. The increase between the periods 1950–1989 and 2070–2099 is ≈ 3 °C in the summer and ≈ 3.5 °C in the winter season in the ECHAM5 simulation. The increase of the temperature over the Rhine basin is 3.3 °C in summer and 2.8 °C in winter in the RACMO-ECHAM5 simulation. In the summer season there is, however, a considerable gradient in the warming over the Rhine basin (from 2.5 °C in region 5 in the north to 4.3 °C in region 1 in the south).

To compare the distribution of extremes in the RACMO-ECHAM5 run with that in observations, the gridded observed daily precipitation amounts produced within the EU-funded ENSEMBLES project [*Haylock et al.*, 2008] were used. These data (further denoted as E-OBS) are available on different grids including a rotated longitude-latitude grid with a resolution of ≈ 25 km, which makes the comparison with the RACMO data straightforward. The data cover the period 1950–2006. The density of stations used for gridding varies across the Rhine basin (e.g., Netherlands ≈ 1 station per 400 km^2, Switzerland ≈ 1 station per 1300 km^2, and Germany ≈ 1 station per 3400 km^2). The rather low station density in much of the Rhine basin implies that only a small fraction of grid boxes contains one or more rainfall stations (see Figure 1d). For the gridding of the E-OBS data set, the station data were first interpolated to a 0.1 degree longitude-latitude grid (≈ 10 km by 5 km) using a search radius of 450 km, and then averaged within the grid boxes. The distance between stations that significantly contribute to the interpolated values is relatively large in areas with low station density, resulting in a large amount of spatial smoothing. This questions the representativeness of the extremes in the E-OBS data for these areas. *Hofstra et al*. [2009] compared daily precipitation in the E-OBS data to that in three gridded data sets based on a significantly larger number of rain gauges: one for the UK (1958–2002), the Alpine data set (1971–1995), and the ELDAS data set (October 1999–December 2000) covering central and northern Europe. The upper deciles of the area-average daily rainfall amounts found in these data sets turned out to be larger than those in the E-OBS data set, in particular in the Alpine data set. The latter is also used in our study and will be denoted as ALP from here on. It is available on a regular longitude-latitude grid with a resolution of ≈ 25 km. The density of stations used for gridding was ≈ 1 station per 100–200 km^2 and more high-elevation stations were included than in the E-OBS data. Further details on this data set can be found in *Frei and Schär* [1998].

4. Summer maxima

4.1. Results

Figure 4 shows boxplots of estimated parameters and their trends for the 1-day summer maximum precipitation. These boxplots were obtained from 3000 bootstrap samples. The upper panels (Figures 4a–c) refer to the GEV parameters for the period 1950–1989. The estimated values of $\bar{\xi}$ (average location parameter over the S grid boxes in the region), γ, and κ were derived from equations (5)–(7) using the 1950–1989 average summer global temperature anomaly for $I(t)$.

In the RACMO-ECHAM5 simulation the average location parameter is about 32 mm in the Alpine area and about 21 mm in the rest of the basin. This difference is caused by the high mean seasonal precipitation amounts in the Alps. The dispersion coefficient varies between 0.32 and 0.37 in the RACMO-ECHAM5 simulation. The high value of the dispersion coefficient in region 3 could be related to the low mean precipitation in this region. We do not have any explanation for the high values of the dispersion coefficient in region 5. The shape parameter is positive (Fréchet distribution).

Figures 4a–c also give the estimated parameters from the E-OBS and ALP data sets based on the non-stationary GEV model using the average summer global temperature anomaly from the HadCRUT3 data set of gridded observed temperatures [*Brohan et al.*, 2006] for $I(t)$ in equations (5)–(7). The location parameter in the RACMO-ECHAM5 simulation is on average 10% larger than the location parameter from the E-OBS data. In addition to model error, this difference is caused in part by the low number of stations used for gridding in certain countries (see section 3). This is most pronounced in region 1 where the average estimate of the location parameter from the E-OBS data is 20% lower than that from the ALP data which are based on a substantially larger number of stations. These differences remain large (15%) if the parameters for the E-OBS and ALP data are estimated for the common period 1971–1995. Furthermore, there is little difference between the estimated location parameter from the RACMO-ECHAM5 and E-OBS data in region 5 where the gridding of the E-OBS data was based on a relatively large number of stations. The dispersion coefficient and the shape parameter show a reasonable agreement in the E-OBS and ALP data sets for region 1. These two parameters are in most regions somewhat larger in the RACMO-ECHAM5 simulation than in the E-OBS data.

Figures 4d–f refer to the estimated trends in the GEV parameters $\xi(t)$, $\gamma(t)$, and $\kappa(t)$. The change of $\xi(t)$ and $\gamma(t)$ is given as the ratio of the mean values of these parameters for the periods 2070–2099 and 1950–1989, the change of $\kappa(t)$ is the difference in the mean of $\kappa(t)$ for the same periods. There is a notable positive trend in the dispersion coefficient in all five regions, while the trends in the location and the shape parameters are less clear.

To assess the increase in precision of the parameter estimates due to spatial pooling, the non-stationary GEV model was fitted for each individual grid box (i.e., without spatial pooling) and the 25th and 75th percentiles of the parameter estimates were calculated using 500 bootstrap samples. Then, for each region and each parameter the average interquartile range was obtained as the difference between the average 75th and 25th percentile of the estimates. These average interquartile ranges were compared with those in Figure 4. Table 1 gives the reduction of the interquartile range for the summer season for the RACMO-ECHAM5 data. Note that in the case of no correlation between grid boxes the standard error would be roughly inversely proportional to the square root of the number of grid boxes, which would lead to a reduction by 85–90% of the interquartile range. If the grid boxes were perfectly dependent there would be no reduction at all. The reduction for the RACMO-ECHAM5 data is substantial: 30–80%. Spatial pooling has the largest influence on the uncertainty of the shape parameter and the reduction is larger for parameters describing trends.

The relative changes of quantiles (ratios of the average quantiles in the periods 2070–2099 and 1950–1989) are shown in Figure 5. Despite the decrease of mean summer precipitation, the quantiles of the extremes increase. The change of the 2-year quantile is largely determined by the change of the location parameter. Therefore, there is only a small increase (up to 10%) of the 2-year quantile except for region 2 where a relatively large increase of the location parameter leads to an increase of this quantile of almost 30%. The relative increase of the 50-year quantile is larger in all regions except for region 2 because of the positive trend in the dispersion coefficient. The 50-year quantile increases by 10–30% in regions 1 and 3 and even by 50% in regions 4 and 5 where the positive trend in the dispersion coefficient is enforced by the positive trend in the shape parameter. The relatively small increase of the 50-year quantile in region 2 is caused by the decrease of the shape parameter. The uncertainty of the change of a given quantile is large, in general comparable with its magnitude.

One possible way to reduce the uncertainty of changes of quantiles is to join regions or to assume that certain regions have common parameters. To test for differences between regions the following statistic was used:

$$R = \sum_{i=1}^{n} (\hat{\theta}_i - \overline{\theta})^2 \tag{14}$$

with n the number of regions, $\hat{\theta}_i$ the estimate of the parameter of interest for region i and $\overline{\theta} = \sum_{i=1}^{n} \hat{\theta}_i / n$. The results of the test for the five regions in the Rhine basin are given in Table 2. The p-values were obtained using a bootstrap procedure as described in

Appendix A. The differences between the regions are significant at the 0.1 level for all parameters except the trend parameter γ_1. The differences in the trend parameters ξ_1 and κ_1, however, are mostly due to the results in region 2 only: the trends in the other regions are similar (see Figure 4). Therefore, a restricted model with common trends of the GEV parameters in regions 1, 3, 4, and 5 was also fitted. Regardless of different values of $\hat{\gamma}_0$ and $\hat{\kappa}_0$, the estimated changes of the quantiles for this restricted model are almost identical in these four regions (see Figure 6) and roughly correspond to the mean of the relative changes in these regions assuming no common parameters. The uncertainty is, however, significantly reduced. For the 50-year quantile in Figure 6 a 27% increase is found. This corresponds to a 6.3% increase per degree of summer warming in region 1 and a 10.8% increase per degree in region 5. The latter value is considerably larger than that expected from the Clausius-Clapeyron relation, indicating that other factors than the temperature influence on atmospheric moisture also determine the change in extreme precipitation.

We studied the data further to find an explanation for the deviating trends in the location parameter and the shape parameter for region 2. This region appeared to be part of a larger area east of the Rhine basin exhibiting less summer drying than the rest of the basin in the RACMO-ECHAM5 simulation (not shown). This difference in summer drying might explain why the location parameter increases in region 2 and not in the other regions. The increased soil moisture deficits towards the end of the 21st century limit the increase of summer showers in regions 1, 3, 4, and 5. We further found that the largest values in the last 20–30 years of the simulation for region 2 are not as large as in the rest of the simulation: the trend is different there. This might explain the drop in the shape parameter in this region (Figure 4f).

4.2. Model validation

For the RACMO-ECHAM5 simulation, the goodness of fit was tested using the A^2 statistic. For regions 1 and 3, Figure 7 gives the A^2 value for each grid box together with critical values for a test at the 0.1 significance level. These critical values were determined using a parametric bootstrap procedure (Appendices B and C). The local 0.1 critical values in Figure 7 apply to the goodness of fit test at an individual grid box. The likelihood that all A^2 values fall below these critical values is small. In the case of an adequate fit it is expected that 10% of the A^2 values exceed the local 0.1 critical value. This fraction is higher for regions 1 and 3 (≈ 20%). This does not necessarily imply lack of fit because of spatial dependence. Even if the model provides an adequate fit, clusters of grid boxes may fail the Anderson-Darling test in the case of spatial dependence. In order to evaluate the field significance, the 0.1 global critical values in Figure 7 have to be considered. The chance that some A^2 value exceeds the line of these critical values is 0.1 if the data come from the assumed model. None of the A^2 values for region 1 is above this line, but in region 3 there are five grid boxes for which A^2 exceeds the global 0.1 critical value. Four of these grid boxes are situated near Mainz in the center of the region (Figure 1) where the lowest precipitation in the Rhine basin is found. A separate model fit for these four grid boxes and three adjacent grid boxes with large A^2 values revealed a relatively high dispersion coefficient for this subregion. There was no evidence of lack of fit of the GEV distribution and the trend γ_1 in the dispersion coefficient did not deviate much from that for the rest of region 3. These seven grid boxes in this relatively dry area were excluded. In addition, four grid boxes in region 4 for which A^2 exceeds the global 0.1 critical value were excluded too. One of these grid boxes is located on the western border of the river Rhine basin, whereas the other three are situated in a relatively wet subregion, known as Sauerland, with grid box estimates of γ_0 lower than those for the rest of this region. The GEV model was then fitted again and the A^2 statistics and their critical values were recalculated. The results discussed in section 4.1 refer to the refitted model as well as Figures 4, 5, and 6. Figure 8 shows the location of the excluded grid boxes and summarizes the results of the goodness of fit tests. In region 3 there remains one grid box for which A^2 exceeds the global 0.1 critical value.

Two additional checks were made to assess the presence of temporal dependence: (1) the standard normal residuals were averaged over each of the five regions and smoothed using a locally weighted regression, "loess" [*Cleveland*, 1979], in order to find significant temporal patterns; (2) the average autocorrelation of the standard normal residuals was calculated for each of the five regions. Figures 9 and 10 show the results of these checks for region 1. Both pictures are representative of the other regions as well and both show no evidence of temporal dependence.

5. Winter maxima

5.1. Results

Boxplots of the estimated GEV parameters for the 5-day winter maximum precipitation in the RACMO-ECHAM5 simulation for the period 1950–1989 are given in Figures 11a–

c. As for the summer season the location parameter in the Alpine region is higher than in the rest of the basin. The dispersion coefficient shows a south north gradient. The shape parameter is almost zero in three of the five regions.

The RACMO-ECHAM5 simulation overestimates the location parameter by 10–30% and underestimates the dispersion coefficient by 35% with respect to the E-OBS data. For the 5-day winter maxima the reduction of variability in the E-OBS data due to the gridding of insufficient station data is smaller than for the 1-day summer maxima because of the stronger spatial correlation between the 5-day winter maxima. The low number of stations used for gridding cannot explain the observed differences between the parameter estimates from the RACMO-ECHAM5 and E-OBS data. In contrast to the 1-day summer maxima the differences between the estimated location parameters from the ALP and E-OBS data are small for region 1. There is also a significant difference between the estimated location parameters from the RACMO-ECHAM5 and E-OBS data for the well-gauged region 5. The overestimation of the location parameter in the RACMO-ECHAM5 data is strongly related to the positive model bias in the mean (36%) and the standard deviation (11%) of daily winter precipitation. Part of this bias is caused by the systematic undercatch inherent to rain gauges for which neither the E-OBS nor the ALP data were corrected. For instance, *Frei et al.* [2003] mention for the winter season an average bias of 11% due to undercatch. This bias is expected to be somewhat lower in other parts of the Rhine basin because of a smaller fraction of snowfall. Since the overestimation of the standard deviation is smaller than that of the mean, the coefficient of variation is underestimated (19%). The low relative variability of the daily values in the RACMO-

ECHAM5 simulation partly accounts for the underestimation of the dispersion coefficient in the GEV model for the 5-day maxima across the basin.

The estimated trends of the GEV parameters in the RACMO-ECHAM5 simulation are shown in Figures 11d–f. The location parameter increases and the shape parameter decreases significantly over the whole basin, while there is almost no change in the dispersion coefficient. The relative changes of the quantiles are given in Figure 12. Due to the increase of the location parameter the 2-year quantiles increase over the whole basin by 10–20%. The relative increase of these quantiles is, however, smaller than the relative increase of mean winter precipitation (Figure 2). For the 50-year quantiles the effect of the increase of the location parameter is counterbalanced by the decrease of the shape parameter resulting in only a slight and non-significant change of this quantile. The physical causes of the relatively small change at large quantiles are unknown and need further investigation. The 5-day winter precipitation extremes result from intense large-scale events which are strongly influenced by the atmospheric circulation. A detailed study of the changes in circulation characteristics would therefore be of interest.

In the model fitted to the E-OBS data there is a rather strong positive trend in the location parameter for all five regions (not shown). This trend is in line with the strong increase found by *Hundecha and Bárdossy* [2005] in the 5-day winter maximum precipitation during the period 1958–2001 at rainfall stations in the German part of the Rhine basin. This upward trend is much stronger than that in the RACMO-ECHAM5 simulation. Moreover, the Gumbel residuals for regions 2, 3, and 4 show a small, but statistically significant lag 1 autocorrelation for the E-OBS data. This points to some unknown factor (or factors) causing long-term variability in extreme 5-day winter precipitation. *Hundecha and Bárdossy* [2005] did not find a significant increase in the frequency of circulation patterns associated with wet days over their study period. The presence of this long-term variability makes difficult the interpretation of the differences between the estimated GEV parameters in the RACMO-ECHAM5 simulation and the E-OBS data set. Further investigation is required to understand fully the disparities.

For the 5-day winter precipitation maxima in the RACMO-ECHAM5 data the reduction of the interquartile ranges of parameter estimates due to spatial pooling is 17–53%, where the lower limit applies to the parameter ξ_1 and the upper limit to the parameter κ_1. This reduction is lower than that for the 1-day summer maxima, due to the stronger spatial correlation between the 5-day winter precipitation maxima. In contrast to the summer maxima, the test for differences between regions indicates that for the 5-day winter precipitation maxima the trends in the GEV parameters can be assumed the same for the whole Rhine basin. However, the reduction of the uncertainty of the quantiles by fitting a model with common trend parameters ξ_1, γ_1, and κ_1 is not as large as that for the 1-day summer maxima. This is partly due to the larger correlation between the estimated parameters of different regions in winter and partly due to the fact that the uncertainty of the changes in quantiles is smaller in winter (compare the widths of the confidence bands in Figures 5 and 12).

5.2. Model validation

For the RACMO-ECHAM5 simulation, Figure 13 gives a summary of the goodness of fit testing for the winter season. As for the summer season the model was initially fitted to all grid boxes. Fifteen grid boxes with high values of A^2 were excluded. Most of these grid boxes are located on the border of region 1 or close to it, some of them at high altitude. Two excluded grid boxes are found on the border of region 4. After the exclusion of these grid boxes the model was refitted and the A^2 values were recalculated. The results discussed in section 5.1 refer to this refitted model. After refitting there remains one grid box with an A^2 value exceeding the global 0.1 critical value in region 2. In contrast to the observed data, no signs of persistence or low-frequency variability were found in the standard normal residuals of the RACMO-ECHAM5 data (not shown). This points to a failure of the driving ECHAM5 global model to reproduce long-term variability. There is, however, a strong indication that the magnitude of the trend parameter ξ_1 decreases with increasing altitude in the Swiss part of the Rhine basin (see Figure 14). The relative increase in the GEV location parameter is therefore smaller at high altitude. This is also found for the change in mean winter precipitation in the RACMO-ECHAM5 simulation. Though the relative increase in mean winter precipitation is smaller at high altitude, the absolute increase is larger. The latter is in agreement with the RCM simulation of *Giorgi et al.* [1997]. The physical cause of this altitude-dependence is not clear.

6. Conclusions

In the present study a non-stationary regional GEV model was introduced and applied to the 1-day summer and 5-day winter precipitation maxima in the transient RACMO-

ECHAM5 run for the river Rhine basin in order to evaluate the changes in the properties of simulated precipitation extremes. The capability of the climate model to reproduce observed precipitation extremes was also assessed. The river Rhine basin was subdivided into 5 regions and the GEV model was applied to each of these regions. The model allows the location parameter to vary over the region of interest with common trend in time. The dispersion coefficient and the shape parameter are assumed constant over the region but varying with time.

The regional GEV model provides an informative summary of the differences between observed and simulated precipitation maxima as well as of the changes in the distribution of extremes. Looking at the parameters of the GEV distribution gives a better insight into the differences in distribution than looking at a single quantile only. In addition, the standard errors of the estimated common parameters are significantly reduced compared to the estimates based on the data of an individual grid box.

The choice of regions is a difficult point in the application of the regional GEV model. The size of a region is limited by spatial heterogeneities in the GEV parameters γ and κ as well as spatial heterogeneities in the trends of these parameters. Maps of grid box estimates of γ for the periods 1950–1989 and 2070–2099 proved to be useful for the partitioning of the Rhine basin in this study. Instead of defining certain regions, one could pool the data from the grid box of interest and a fixed number of neighboring grid boxes [e.g., *Zwiers and Kharin*, 1998; *Coelho et al*., 2008]. This is convenient if identifying large, homogeneous regions is difficult or if one wishes to show how the model parameters vary over the entire RCM domain. The size of such neighborhoods is typically much smaller than the regions used in regional frequency analyses, and therefore results in less spatial pooling. Moreover, the use of a fixed number of grid boxes will not be optimal if the degree of spatial heterogeneity varies over the domain.

The values of estimated parameters in the period 1950–1989 for the 1-day summer precipitation extremes are reasonably well reproduced in the RACMO-ECHAM5 simulation. Part of the differences between the values from the E-OBS data can be ascribed to the low density of stations used for gridding. The distribution of the 5-day winter precipitation extremes is affected by strong positive biases in the mean and standard deviation of daily winter precipitation. In particular, the dispersion coefficient of the GEV distribution is severely underestimated across the whole Rhine basin.

The changes of the distribution of the 1-day summer precipitation maxima are primarily related to the positive trend in the dispersion coefficient. Since there is almost no change in the location parameter, the changes in distribution are mainly found at large quantiles (e.g., the 50-year quantile) whereas there are only minor changes in quantiles close to the median (i.e., the 2-year quantile). For the 5-day winter maxima the low quantiles (e.g., 2-year quantile) are increasing due to the increase of the location parameter. As the return period gets longer the effect of the positive trend in the location parameter is counterbalanced by the decrease of the shape parameter resulting in only minor positive or negative changes of large quantiles (e.g., the 50-year quantile).

The opposite direction of the changes in mean and 1-day maximum precipitation in summer is in agreement with earlier findings of *Christensen and Christensen* [2004] and *Frei et al.* [2006]. A relatively small change of the quantiles of extreme multi-day winter precipitation was also found by *Leander et al.* [2008] for the adjacent Meuse basin in a simulation of the RACMO model driven by the HadAM3H atmospheric model of the Hadley Centre. Despite a considerable increase in mean winter precipitation in this experiment there was little change in the distribution of the 10-day winter precipitation maxima and extreme river flows. The differences between changes in mean and extremes indicate that proportional adjustment of observed data can be very misleading.

Despite the reduction of standard errors due to spatial pooling of data, the changes in the quantiles of the extreme-value distributions are often not statistically significant. For the 2-year quantile of the 1-day summer maxima this can be attributed to the fact that the change in the location parameter is small. The estimates of the relative changes of the 50-year quantiles are strongly affected by the estimates of the dispersion coefficient and the shape parameter, which have large standard errors. For the summer season the uncertainty of the change in this quantile for regions 1, 3, 4, and 5 could be reduced considerably by assuming common trend parameters ξ_1, γ_1, and κ_1. The use of an ensemble of RACMO simulations driven by different simulations of the ECHAM5 global climate model is an option to improve the estimates of the changes in extreme value properties of this RCM-GCM configuration further. Apart from the uncertainty in the extreme value properties for a particular RCM-GCM configuration, there are large differences between the estimated changes for different RCM-GCM combinations.

The Anderson-Darling test shows that the model fits well for much of the Rhine basin. In the summer the model fails to fit in a relatively dry subregion with a relatively high dispersion coefficient and in a small relatively wet subregion. In the winter season the model did not fit well at a number of grid boxes on, or close to, the border of the Rhine basin, in particular in the Swiss part of the basin. As a consequence, a small number of grid boxes were excluded. A separate model fit using part of the excluded grid boxes suggests that formation of different, smaller regions could improve the goodness of fit, however, at the cost of increased uncertainty. Another possibility is the reformulation of the statistical model to allow the dispersion coefficient to vary over the region of interest. In addition, for regions with strong orography it may be necessary to incorporate altitude-dependence of the trend in the location parameter.

Appendix A: Test for differences between regions

Let θ_i be one of the parameters ξ_1, γ_0, γ_1, κ_0, or κ_1 in the non-stationary GEV model for region i and let $\boldsymbol{\tau}$ be the vector of the other parameters. We want to test the hypothesis $\mathrm{H}_0 : \theta_1 = \theta_2 = \ldots = \theta_n$ using the statistic R in equation (14). The test consists of the following steps:

1. Calculate the value of the test statistic using equation (14) and denote this value r.
2. Calculate the standard Gumbel residuals using the $\hat{\theta}_i$ and the estimated values of the other parameters.
3. Re-estimate the other parameters $\hat{\boldsymbol{\tau}}_0$ given $\theta_1 = \theta_2 = \ldots = \theta_n = \bar{\theta}$.

4. Draw a bootstrap sample from the standard Gumbel residuals using resampling of years to preserve the spatial dependence structure (see section 2.3) and transform this sample back to the original scale using the parameter estimates $\overline{\theta}$ and $\hat{\boldsymbol{\tau}}_0$.
5. Re-estimate all parameters and re-calculate the test statistic as

$$r_b^* = \sum_{i=1}^{n} (\hat{\theta}_{b,i}^* - \overline{\theta}_b^*)^2 \,, \tag{A1}$$

with $\hat{\theta}_{b,i}^*$ the estimate of θ_i from bootstrap sample b and $\overline{\theta}_b^* = \sum_{i=1}^{n} \hat{\theta}_{b,i}^* / n$.
6. Repeat steps 4–5 until the desired number of bootstrap samples is obtained.

The p-value is the fraction of r_b^* values larger than r. The p-values in Table 2 are based on 500 bootstrap samples.

Appendix B: Determination of the critical values of the Anderson-Darling statistic

The critical values of the Anderson-Darling statistic A^2 in the literature usually refer to the situation of independent realizations from a distribution that is entirely specified under the null hypothesis. This does not apply to the standard Gumbel residuals $\widetilde{X}(s,t)$ at a given grid box, which are in fact dependent due to the use of estimated GEV parameters instead of their true but unknown values. It is well-known that parameter estimation has a substantial effect on the distribution of A^2 [e.g., *Laio*, 2004]. This appendix deals with the derivation of the local and global critical values of A^2 from bootstrap samples. The generation of these bootstrap samples is discussed in Appendix C. In our application B = 3000 bootstrap samples were generated.

Let $t(s)$ be the value of A^2 from the climate model data at grid box s (s = 1, …, S) and let $t_b^*(s)$ be the value of A^2 from bootstrap sample b (b = 1, …, B) for this grid box. For a chosen significance level α_{LOC}, the local critical values $c^{\alpha_{LOC}}(s)$ are obtained for each grid box as the kth smallest value $t_{(k)}^*(s)$ of the $t_b^*(s)$, where $k = (1-\alpha_{LOC})(B+1)$.

The determination of the global critical values is based on an approach suggested by *Davison and Hinkley* [1997]. Let $c_{-b}^{\alpha_{LOC}}(s)$ be the local critical values that we get if we exclude bootstrap sample b. Then a bootstrap estimate of the global error rate α_{GLOB} is obtained as:

$$\alpha_{GLOB} = \frac{\#\{b : [t_b^*(s) \geq c_{-b}^{\alpha_{LOC}}(s), \quad \text{for any } s]\}}{B}, \tag{B1}$$

where $\#\{b: A_b\}$ is the number of b for which A_b is true. This error rate can easily be calculated using the fact that bootstrap sample b fulfills the condition $[t_b^*(s) \geq c_{-b}^{\alpha_{LOC}}(s), \quad \text{for any } s]$ if and only if $\text{rank}[t_b^*(s)] \geq k = (1-\alpha_{LOC})(B+1)$ for at least one s. Thus if the values of $t_b^*(s)$ are stored in a matrix with grid boxes in columns and bootstrap samples in rows, then we first calculate the columnwise ranks and subsequently the proportion of rows in which the maximum rank is greater than or equal to k. The value of k is chosen such that α_{GLOB} is as close as possible to the desired global significance level.

Appendix C: Comparison of two bootstrap procedures for goodness of fit testing

The determination of the critical values of the Anderson-Darling statistic A^2 requires simulation from the model under the null hypothesis. In particular, the preservation of spatial dependence is important. The bootstrap procedure outlined in section 2.3 to asses the uncertainty of the parameter estimates and quantiles is not appropriate for testing goodness of fit because the distribution of the $\widetilde{X}(s,t)$ may deviate from the Gumbel distribution due to lack of fit of the GEV model and because of the occurrence of ties in the bootstrap samples. The latter influences the statistical properties of the empirical distribution function $F_N(x)$ in equation (12). In this appendix two alternatives are discussed:

- Replacement of resampled standard Gumbel residuals by samples from the standard Gumbel distribution, preserving the spatial structure of the ranks of the maxima as suggested by *Heffernan and Tawn* [2004]. This approach requires no assumptions about the underlying dependence structure of data.

- Sampling standard normal residuals from the multivariate normal distribution [*Hosking and Wallis*, 1997]. These residuals are assumed to be equicorrelated, i.e., the correlation $\rho_{i,j}$ between the residual at grid box i and the residual at grid box j equals $\rho_{i,j} = \rho$ for $i \neq j$ and $\rho_{i,j} = 1$ for $i = j$. In this case the multivariate normal dependence structure is introduced into the simulated samples.

In the following the procedures are referred to as "HT" and "MVN", respectively, and both are fully described below.

Bootstrap procedure based on the Heffernan and Tawn approach

1. Fit the statistical model to the original sample.
2. Calculate standard Gumbel residuals with the parameter estimates from step 1.
3. Bootstrap the residuals from step 2 (using resampling of years to preserve the spatial dependence as described in section 2.3).
4. Generate S independent samples of size N from the standard Gumbel distribution (S is the number of grid boxes and N the number of years).
5. Rearrange the values in the samples from step 4 such that the dependence structure of the ranks corresponds to that of the bootstrapped residuals from step 3.
6. Transform the rearranged standard Gumbel values from step 5 back to the original scale using the parameter estimates from step 1.
7. Fit the statistical model again.
8. Calculate standard Gumbel residuals with the parameter estimates from step 7 and calculate the A^2 statistics.
9. Repeat steps 3–8 until the desired number of bootstrap samples is obtained.

Parametric bootstrap procedure with sampling from the multivariate normal distribution

1. Fit the statistical model to the original sample.

2. Calculate standard normal residuals (see section 2.3) with the parameter estimates from step 1.
3. Calculate the average correlation $\hat{\rho}$ of the standard normal residuals.
4. Generate a sample of S equicorrelated standard normal variables with correlation $\hat{\rho}$.
5. Transform the sample from step 4 back to the original scale using the parameter estimates from step 1.
6. Fit the statistical model again.
7. Calculate standard Gumbel residuals with the parameter estimates from step 6 and calculate the A^2 statistics.
8. Repeat steps 4–7 until the desired number of bootstrap samples is obtained.

A simulation experiment was conducted to assess the validity of both approaches: 3000 samples of size 150 from an equicorrelated 30-dimensional normal distribution with known correlation were generated (think about 30 grid boxes in the RACMO-ECHAM5 simulation which has a length of 150 years). These samples (further denoted as control samples) were transformed according to the non-stationary GEV model

$$\xi(s,t) = \xi_0(s)\exp[\xi_1(t-40)_+] \tag{C1}$$

$$\gamma(t) = \exp[\gamma_0 + \gamma_1(t-40)_+] \tag{C2}$$

$$\kappa(t) = \kappa_0 + \kappa_1(t-40)_+ \tag{C3}$$

with $s = 1, \ldots, 30$; $t = 1, \ldots, 150$, and $(x)_+ = \max(x, 0)$. The values of the parameters were set to be representative of those obtained for the 1-day summer maximum precipitation in the Rhine basin, i.e., $\xi_0(s)$ ranged between 22 and 38, $\xi_1 = 0.00055$, $\exp(\gamma_0) = 0.37$, $\gamma_1 = 0.0013$, $\kappa_0 = 0.05$, and $\kappa_1 = 0.00015$.

For each sample the parameters of the GEV model were estimated and the values of the A^2 statistics were calculated. The 0.1 critical value from these simulations is denoted the "true" critical value. Further, for one of the control samples two sets of 3000 bootstrap samples were generated using the "HT" and "MVN" approaches, respectively, and the 0.1 local and global critical values of the A^2 statistic were calculated according to Appendix B.

Table C1 gives the local rejection rates of the null hypothesis as obtained from the control samples, i.e., the proportion of the A^2 values of these samples lying above the "HT" and "MVN" critical values. For the "MVN" critical values the rejection rate corresponds quite well with the nominal 0.1 significance level, but for the "HT" critical values the actual rejection rate is lower than 0.1 in the case of correlation and the difference grows with increasing correlation coefficient. Table C1 further shows that the "MVN" critical values resemble the "true" critical values and decrease with increasing correlation. By contrast the "HT" critical values do not depend on correlation. Though Table C1 refers to the local rejection rates and the local critical values, very similar results were obtained for the global test at the 0.1 significance level.

To understand why the critical values of the A^2 statistic are decreasing with increasing correlation, we have to examine how the estimates of the parameters are influenced by the data from a particular grid box. The estimate of $\xi_0(s)$ is largely determined by the maxima of the grid box of interest. If there is no or little correlation, the maxima of this grid box have little influence on the estimates of the other parameters γ_0, κ_0, ξ_1, γ_1, and κ_1. The influence of the maxima of the grid box of interest on the estimates of these parameters grows with increasing spatial correlation. As a result the fitted regional GEV model will describe the local maxima better and therefore the critical value of the A^2 statistic should be smaller than in the case of independence. The "MVN" and "true" critical values for $\rho = 0.99$ are close to the critical value for the case that all six parameters are estimated from the maxima at the grid box of interest only.

The reason of the failure of the "HT" approach in the case of goodness of fit testing is that the test statistic is insensitive to a permutation of the data, i.e., rearranging residuals at a grid box to preserve the spatial dependence of the ranks does not influence the value of the A^2 statistic. Unlike the "MVN" bootstrap samples, the values of the A^2 statistic do not exhibit any spatial correlation in the "HT" bootstrap samples. Although the "HT" approach is not suitable for goodness of fit testing, it can be used for the estimation of standard errors and the construction of confidence intervals, for which it was originally introduced by *Heffernan and Tawn* [2004].

It is not surprising that the "MVN" critical values do quite well because of the underlying multivariate normal dependence structure of the data. To study the robustness to the type of association at extreme levels, 3000 new samples were generated from our non-stationary GEV model but now with a dependence structure of a limiting extreme-value distribution. This was achieved by generating the standard Gumbel residuals from an equicorrelated multivariate Gumbel distribution as described by *Stephenson* [2003]. The results (not shown) are very similar to those presented in Table C1 for a multivariate normal dependence structure from which it may be concluded that the "MVN" critical values are robust to the dependence structure.

Acknowledgments. We acknowledge the ENSEMBLES project, funded by the European Commission's 6th Framework Programme through contract GOCE-CT-2003-505539. The Alpine data set was kindly provided by MeteoSwiss.

References

Adlouni, S. El., T. B. M. J. Ouarda, X. Zhang, R. Roy, and B. Bobée (2007), Generalized maximum likelihood estimators for the nonstationary generalized extreme value model, *Water Resour. Res.*, *43*, W03410, doi:10.1029/2005WR004545.

Alila, Y. (1999), A hierarchical approach for the regionalization of precipitation annual maxima in Canada, *J. Geophys. Res.*, *104*(D24), 31,645–31,655.

Anderson, T. W., and D. A. Darling (1952), Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes, *Ann. Math. Stat.*, *23*(2), 193–212, doi:10.1214/aoms/1177729437.

Arnell, N. W., and S. Gabriele (1988), The performance of the two-component extreme value distribution in regional frequency analysis, *Water Resour. Res.*, *24*(6), 879–887.

Beniston, M., D. B. Stephenson, O. B. Christensen, C. A. T. Ferro, C. Frei, S. Goyette, K. Halsnaes, T. Holt, K. Jylhä, B. Koffi, J. Palutikof, R. Schöll, T. Semmler, and K. Woth (2007), Future extreme events in European climate: an exploration of regional climate model projections, *Climatic Change*, *81*, 71–95, doi:10.1007/s10584-006-9226-z.

Brath A., A. Castellarin, and A. Montanari (2003), Assessing the reliability of regional depth-duration-frequency equations for gaged and ungaged sites, *Water Resour. Res.*, *39*(12), 1367, doi:10.1029/2003WR002399.

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones (2006), Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850, *J. Geophys. Res.*, *111*, D12106, doi:10.1029/2005JD006548.

Brown, S. J., J. Caesar, and C. A. T. Ferro (2008), Global changes in extreme daily temperature since 1950, *J. Geophys. Res.*, *113*, D05115, doi:10.1029/2006JD008091.

Buishand, T. A. (1991), Extreme rainfall estimation by combining data from several sites, *Hydrological Sciences Journal*, *36*(4), 345–365.

Buonomo, E., R. Jones, C. Huntingford, and J. Hannaford (2007), On the robustness of changes in extreme precipitation over Europe from two high resolution climate change simulations, *Q. J. R. Meteorol. Soc.*, *133*, 65–81, doi:10.1002/qj.13.

Christensen, O. B., and J. H. Christensen (2004), Intensification of extreme European summer precipitation in a warmer climate, *Global Planet. Change*, *44*, 107–117, doi:10.1016/j.gloplacha.2004.06.013.

Cleveland, W. S. (1979), Robust locally weighted regression and smoothing scatterplots, *J. Am. Stat. Assoc.*, *74*, 829–836.

Coelho, C. A., C. A. T. Ferro, D. B. Stephenson, and D. J. Steinskog (2008), Methods for exploring spatial and temporal variability of extreme events in climate data, *J. Clim.*, *21*(10), 2072–2092, doi:10.1175/2007JCLI1781.1.

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

Cunderlik, J. M., and D. H. Burn (2003), Non-stationary pooled flood frequency analysis, *J. Hydrol.*, *276*, 210–223.

Cunderlik, J. M., and T. B. M. J. Ouarda (2006), Regional flood-duration-frequency modeling in the changing environment, *J. Hydrol.*, *318*, 276–291.

Davison, A. C., and D. V. Hinkley (1997), *Bootstrap methods and their application*, Cambridge University Press.

Durman, C. F., J. M. Gregory, D. C. Hassell, R. G. Jones, and J. M. Murphy (2001), A comparison of extreme European daily precipitation simulated by a global and regional climate model for present and future climates, *Q. J. R. Meteorol. Soc.*, *127*, 1005–1015.

Ekström, M., H. J. Fowler, C. G. Kilsby, and P. D. Jones (2005), New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 2. Future estimates and use in impact studies, *J. Hydrol.*, *300*, 234–251, doi:10.1016/j.jhydrol.2004.06.019.

Faulkner, D. S., and D. A. Jones (1999), The FORGEX method of rainfall growth estimation III: Examples and confidence intervals. *Hydrol. Earth Syst. Sci.*, *3*, 205–212.

Fowler, H. J., M. Ekström, C. G. Kilsby, and P. D. Jones (2005), New estimates of future changes in extreme rainfall across the UK using regional climate model integrations, 1. Assessment of control climate, *J. Hydrol.*, *300*, 212–233, doi:10.1016/j.jhydrol.2004.06.017.

Frei, C., and C. Schär (1998), A precipitation climatology of the Alps from high-resolution rain-gauge observations, *Int. J. Climatol.*, *18*(8), 873–900.

Frei, C., and C. Schär (2001), Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region, *J. Clim.*, *14*, 1568–1584.

Frei, C., J. H. Christensen, M. Déqué, D. Jacob, R. G. Jones, and P. L. Vidale (2003), Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps, *J. Geophys. Res.*, *108*(D3), 4124, doi:10.1029/2002JD002287.

Frei, C., R. Schöll, S. Fukutome, J. Schmidli, and P. L. Vidale (2006), Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models, *J. Geophys. Res.*, *111*, D06105, doi:10.1029/2005JD005965.

García, J. A., M. C. Gallego, A. Serrano, and J. M. Vaquero (2007), Trends in block-seasonal extreme rainfall over the Iberian Peninsula in the second half of the twentieth century, *J. Clim.*, *20*, 113–130, doi:10.1175/JCLI3995.1.

Giorgi, F., J. W. Hurrell, M. R. Marinucci, and M. Beniston (1997), Elevation dependency of the surface climate change signal: A model study, *J. Clim.*, *10*, 288–296, doi:10.1175/1520-0442(1997)010.

Goubanova, K., and L. Li (2007), Extremes in temperature and precipitation around the Mediterranean basin in an ensemble of future climate scenario simulations, *Global Planet. Change*, *57*, 27–42, doi:10.1016/j.gloplacha.2004.06.010.

Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New (2008), A European daily high-resolution gridded dataset of surface temperature and precipitation, *J. Geophys. Res.*, *113*, D20119, doi:10.1029/2008JD010201.

Heffernan, J. E., and J. A. Tawn (2004), A conditional approach for multivariate extreme values (with discussion), *J. R. Stat. Soc. B*, *66*(3), 497–546, doi:10.1111/j.1467-9868.2004.02050.x.

Hofstra, N., M. Haylock, M. New, and P. D. Jones (2009), Testing E-OBS European high-resolution gridded dataset of daily precipitation and surface temperature, *J. Geophys. Res.*, submitted.

Hosking, J. R. M., and J. R. Wallis (1997), *Regional frequency analysis*, Cambridge University Press.

Hundecha, Y., and A. Bárdossy (2005), Trends in daily precipitation and temperature extremes across western Germany in the second half of the 20th century, *Int. J. Climatol.*, *25*, 1189–1202, doi:10.1002/joc.1182.

Huntingford, C., R. G. Jones, C. Prudhomme, R. Lamb, J. H. C. Gash, and D. A. Jones (2003), Regional climate model predictions of extreme rainfall for a changing climate, *Q. J. R. Meteorol. Soc.*, *129*, 1607–1621, doi:10.1256/qj.02.97.

Jungclaus, J. H., N. Keenlyside, M. Botzet, H. Haak, J.-J. Luo, M. Latif, J. Marotzke, U. Mikolajewitcz, and E. Roeckner (2006), Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM, *J. Clim.*, *19*, 3952–3972, doi:10.1175/JCLI3827.1.

Kendon, E. J., R. G. Jones, and E. Buonomo (2008), Robustness of future changes in local precipitation extremes, *J. Clim.*, *21*, 4280–4297, doi:10.1175/2008JCLI2082.1.

Kharin, V. V., and F. W. Zwiers (2005), Estimating extremes in transient climate change simulations, *J. Clim.*, *18*, 1156–1173, doi:10.1175/JCLI3320.1.

Kharin, V. V., F. W. Zwiers, X. Zhang, and G. C. Hegerl (2007), Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations, *J. Clim.*, *20*, 1419–1444, doi:10.1175/JCLI4066.1.

Kyselý, J., and J. Picek (2007), Regional growth curves and improved design value estimates of extreme precipitation events in the Czech Republic, *Clim. Res.*, *33*, 243–255.

Laio, F. (2004), Cramer-von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters, *Water Resour. Res.*, *40*, W09308, doi:10.1029/2004WR003204.

Leander, R., T. A. Buishand, B. J. J. M. van den Hurk, and M. J. M. de Wit (2008), Estimated changes in flood quantiles of the river Meuse from resampling of regional climate model output, *J. Hydrol.*, *351*, 331–343, doi:10.1016/j.jhydrol.2007.12.020.

Lettenmaier, D. P., J. R. Wallis, and E. Wood (1987), Effect of regional heterogeneity on flood frequency estimation, *Water Resour. Res.*, *23*(2), 313–323.

van Meijgaard, E., L. H. van Ulft, W. J. van de Berg, F. C. Bosveld, B. J. J. M. van den Hurk, G. Lenderink, and A. P. Siebesma (2008), The KNMI regional atmospheric climate model RACMO, version 2.1, *KNMI Technical Report 302*, Royal Netherlands Meteorological Institute, De Bilt, Netherlands.

Nakićenović, N., and R. Swart (Eds.) (2000), *Special Report on Emissions Scenarios. A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*, Cambridge University Press.

Northrop, P. J. (2004), Likelihood-based approaches to flood frequency estimation, *J. Hydrol.*, *292*, 96–113, doi:10.1016/j.jhydrol.2003.12.031.

Pall, P., M. R. Allen, and D. A. Stone (2007), Testing the Clausius-Clapeyron constraint on changes in extreme precipitation under CO_2 warming, *Clim. Dyn.*, *28*, 351–363, doi:10.1007/s00382-006-0180-2.

Renard, B., V. Garreta, and M. Lang (2006), An application of Bayesian analysis and Markov chain Monte Carlo methods to the estimation of a regional trend in annual maxima, *Water Resour. Res.*, *42*, W12422, doi:10.1029/2005WR004591.

Schaefer, M. G. (1990), Regional analyses of precipitation annual maxima in Washington State, *Water Resour. Res.*, *26*(1), 119–131.

Sen, P. K. (1968), Estimation of the regression coefficient based on Kendall's tau, *J. Am. Stat. Assoc.*, *63*, 1379–1389.

Shimokawa, T., and M. Liao (1999), Goodness of fit tests for type-1 extreme-value and 2-parameter Weibull distributions, *IEEE Trans. Reliab.*, *48*(1), 79–86.

Stephenson, A. (2003), Simulating multivariate extreme value distributions of logistic type, *Extremes*, *6*, 49–59.

Sveinsson, O. G. B., D. C. Boes, and J. D. Salas (2001), Population index flood method for regional frequency analysis, *Water Resour. Res.*, *37*(11), 2733–2748.

Zwiers, F. W., and V. Kharin (1998), Changes in the extremes of the climate simulated by CCC GCM2 under CO_2 doubling, *J. Clim.*, *11*(9), 2200–2222, doi: 10.1175/1520-0442(1998)011.

List of Figures

Figure 1. (a) The river Rhine basin. (b) The dispersion coefficient of the 1-day summer (JJA) maxima for the fit of the stationary GEV model to the RACMO-ECHAM5 simulation for the period 1950–1989. (c) Same as (b) but for the period 2070–2099. (d) Subdivision of the river Rhine basin into five regions. The numbers in subscript give the number of grid boxes included in the region. The rectangles represent the RACMO model grid boxes, the gray dots show the locations of stations that have been used for gridding of the E-OBS data.

Figure 2. Relative change of mean seasonal and annual precipitation between the periods 1950–1989 and 2070–2099 in the RACMO-ECHAM5 simulation for all 5 regions of the Rhine basin.

Figure 3. Summer (JJA) and winter (DJF) global temperature anomalies in the ECHAM5 simulation.

Figure 4. (a–c) Estimates of the GEV parameters for the 1-day summer (JJA) precipitation extremes for the period 1950–1989 for the ALP, E-OBS, and RACMO-ECHAM5 data. The results are averaged over the region in the case of the location parameter. (d–f) The changes of the GEV parameters for the 1-day summer (JJA) precipitation extremes between the periods 1950–1989 and 2070–2099. The boxplots were obtained from 3000 bootstrap samples. The boxes represent the interquartile range, the whiskers extend from the 5th to the 95th percentile of these bootstrap samples.

Figure 5. Relative changes of quantiles of the 1-day summer maximum precipitation between the periods 1950–1989 and 2070–2099 in the RACMO-ECHAM5 simulation for all five regions. The confidence bands were obtained from 3000 bootstrap samples. The 5th, 25th, 50th, 75th, and 95th percentile of these bootstrap samples are shown.

Figure 6. Same as Figure 5 but for the restricted model with common trends over regions 1, 3, 4, and 5. The panel on the right gives the average relative change of the four regions together with the average confidence band.

Figure 7. The values of the Anderson-Darling statistic for (a) region 1 and (b) region 3 for the 1-day summer (JJA) precipitation extremes in the RACMO-ECHAM5 simulation.

Figure 8. Summary of the goodness of fit testing of the non-stationary GEV model for the 1-day summer (JJA) precipitation extremes in the RACMO-ECHAM5 simulation.

Figure 9. Averaged standard normal residuals (gray line) for the 1-day summer (JJA) precipitation extremes in the RACMO-ECHAM5 simulation in region 1. The black line shows residuals smoothed by locally weighted regression "loess".

Figure 10. Average autocorrelation coefficients (ACC) of the standard normal residuals (vertical bars) for the 1-day summer (JJA) precipitation extremes in the RACMO-

ECHAM5 simulation in region 1. The 90% confidence band (shaded area) was obtained from 3000 bootstrap samples.

Figure 11. Same as Figure 4 but for the 5-day winter (DJF) precipitation extremes.

Figure 12. Same as Figure 5 but for the 5-day winter (DJF) precipitation extremes.

Figure 13. Same as Figure 8 but for the 5-day winter (DJF) precipitation extremes.

Figure 14. Grid box estimates of the trend in the location parameter as a function of altitude for the 5-day winter (DJF) precipitation maxima in the RACMO-ECHAM5 simulation. The values for the grid boxes in region 1 (black dots) are smoothed by locally weighted regression "loess" (black line).

Table 1. Reduction (%) of interquartile ranges of the parameter estimates due to spatial pooling for the summer (JJA) in the case of the RACMO-ECHAM5 data.

	region					
parameter	1	2	3	4	5	mean
ξ_1	37	32	34	31	39	35
γ_0	58	45	48	53	44	50
γ_1	67	60	60	61	58	61
κ_0	73	71	67	66	58	67
κ_1	80	75	75	72	66	74

Table 2. The *p*-values resulting from the test for differences between regions for the summer (JJA).

parameter	*p*-value
ξ_1	0.01
γ_0	0.00
γ_1	0.13
κ_0	0.00
κ_1	0.00

Table C1. Local rejection rates and critical values (nominal significance level of 0.1) for testing goodness of fit using the Anderson-Darling statistic. The "true" critical values are based on 3000 simulated samples from a non-stationary GEV model, the critical values "HT" and "MVN" are based on 3000 bootstrap samples from one of these simulations using respectively the Heffernan and Tawn approach and a multivariate normal distribution to preserve spatial dependence.

	rejection rate		critical value		
correlation	"HT"	"MVN"	"HT"	"MVN"	"true"
0.00	0.098	0.102	0.881	0.870	0.875
0.40	0.077	0.095	0.888	0.837	0.823
0.60	0.050	0.088	0.901	0.778	0.751
0.80	0.025	0.081	0.892	0.686	0.648
0.99	0.000	0.093	0.905	0.523	0.514

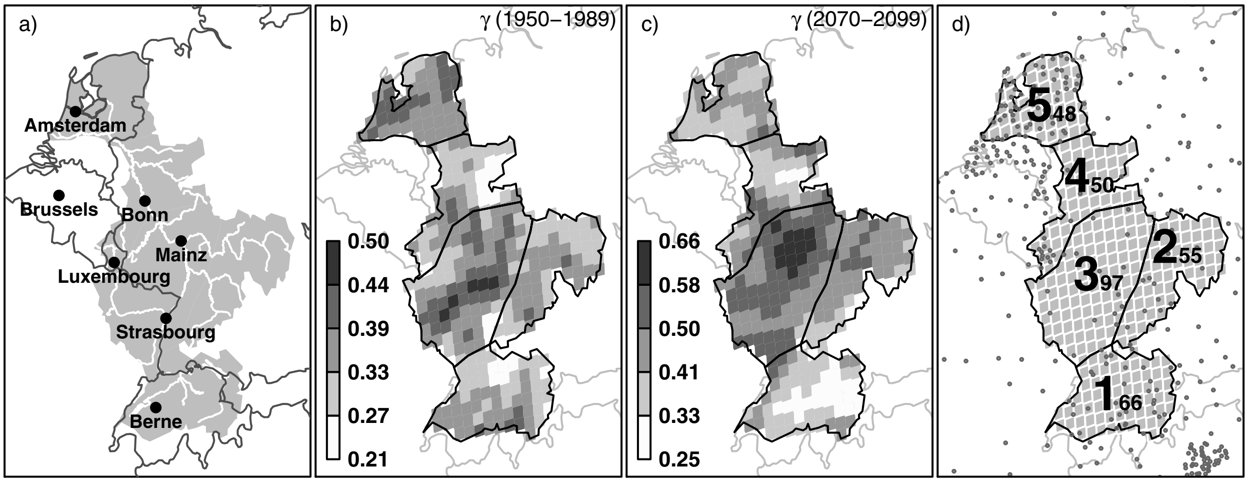

Figure 1. (a) The river Rhine basin. (b) The dispersion coefficient of the 1-day summer (JJA) maxima for the fit of the stationary GEV model to the RACMO-ECHAM5 simulation for the period 1950–1989. (c) Same as (b) but for the period 2070–2099. (d) Subdivision of the river Rhine basin into five regions. The numbers in subscript give the number of grid boxes included in the region. The rectangles represent the RACMO model grid boxes, the gray dots show the locations of stations that have been used for gridding of the E-OBS data.

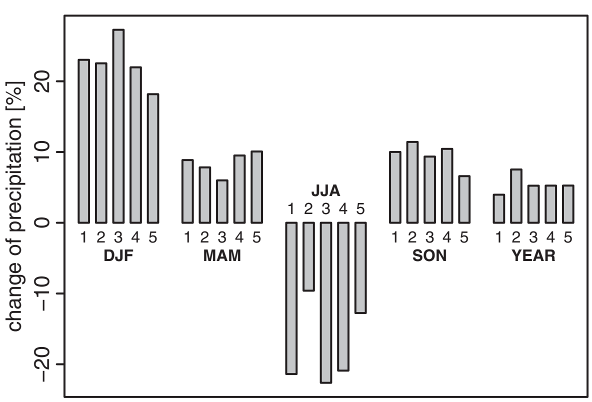

Figure 2. Relative change of mean seasonal and annual precipitation between the periods 1950–1989 and 2070–2099 in the RACMO-ECHAM5 simulation for all 5 regions of the Rhine basin.

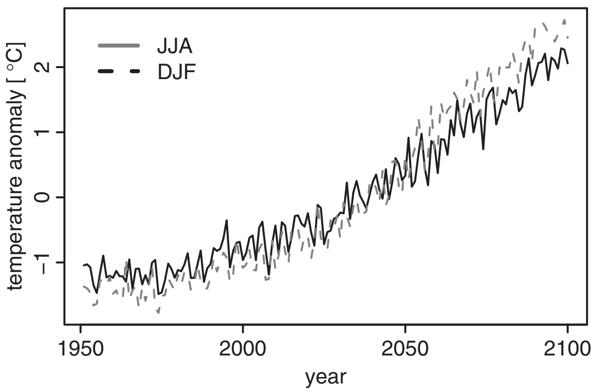

Figure 3. Summer (JJA) and winter (DJF) global temperature anomalies in the ECHAM5 simulation.

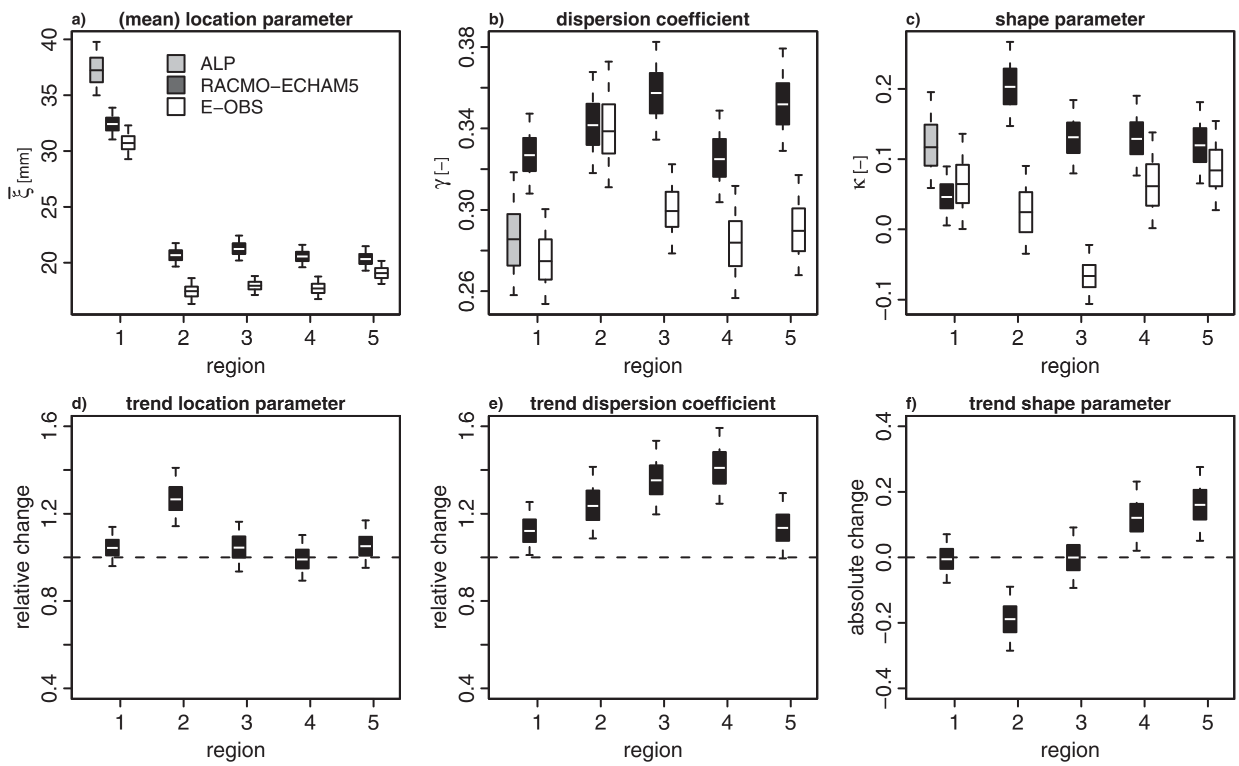

Figure 4. (a–c) Estimates of the GEV parameters for the 1-day summer (JJA) precipitation extremes for the period 1950–1989 for the ALP, E-OBS, and RACMO-ECHAM5 data. The results are averaged over the region in the case of the location

parameter. (d–f) The changes of the GEV parameters for the 1-day summer (JJA) precipitation extremes between the periods 1950–1989 and 2070–2099. The boxplots were obtained from 3000 bootstrap samples. The boxes represent the interquartile range, the whiskers extend from the 5th to the 95th percentile of these bootstrap samples.

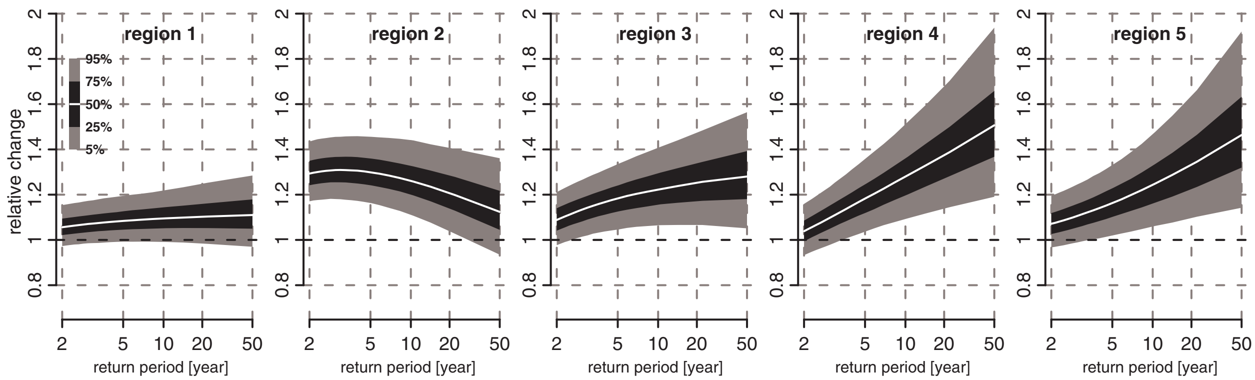

Figure 5. Relative changes of quantiles of the 1-day summer maximum precipitation between the periods 1950–1989 and 2070–2099 in the RACMO-ECHAM5 simulation for all five regions. The confidence bands were obtained from 3000 bootstrap samples. The 5th, 25th, 50th, 75th, and 95th percentile of these bootstrap samples are shown.

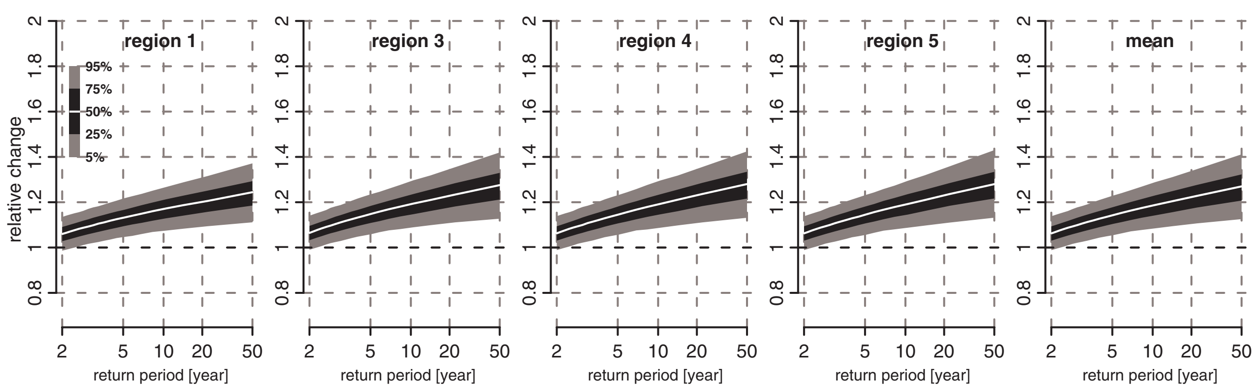

Figure 6. Same as Figure 5 but for the restricted model with common trends over regions 1, 3, 4 and 5. The panel on the right gives the average relative change of the four regions together with the average confidence band.

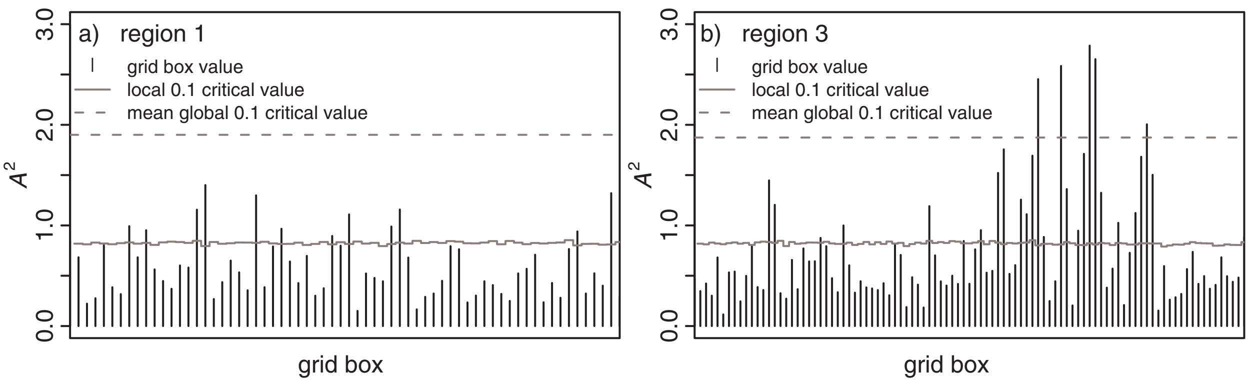

Figure 7. The values of the Anderson-Darling statistic for (a) region 1 and (b) region 3 for the 1-day summer (JJA) precipitation extremes in the RACMO-ECHAM5 simulation.

Figure 8. Summary of the goodness of fit testing of the non-stationary GEV model for the 1-day summer (JJA) precipitation extremes in the RACMO-ECHAM5 simulation.

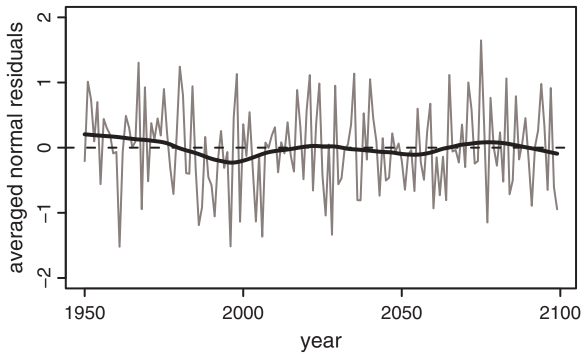

Figure 9. Averaged standard normal residuals (gray line) for the 1-day summer (JJA) precipitation extremes in the RACMO-ECHAM5 simulation in region 1. The black line shows residuals smoothed by locally weighted regression "loess".

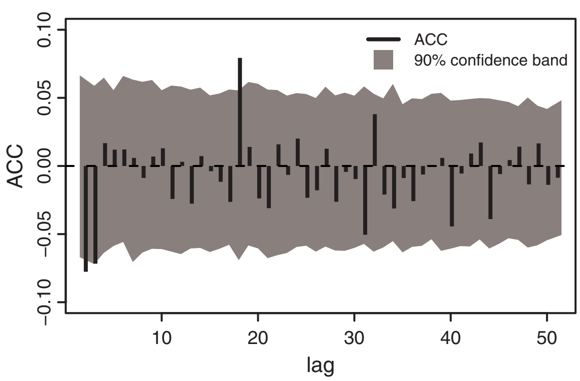

Figure 10. Average autocorrelation coefficients (ACC) of the standard normal residuals (vertical bars) for the 1-day summer (JJA) precipitation extremes in the RACMO-ECHAM5 simulation in region 1. The 90% confidence band (shaded area) was obtained from 3000 bootstrap samples.

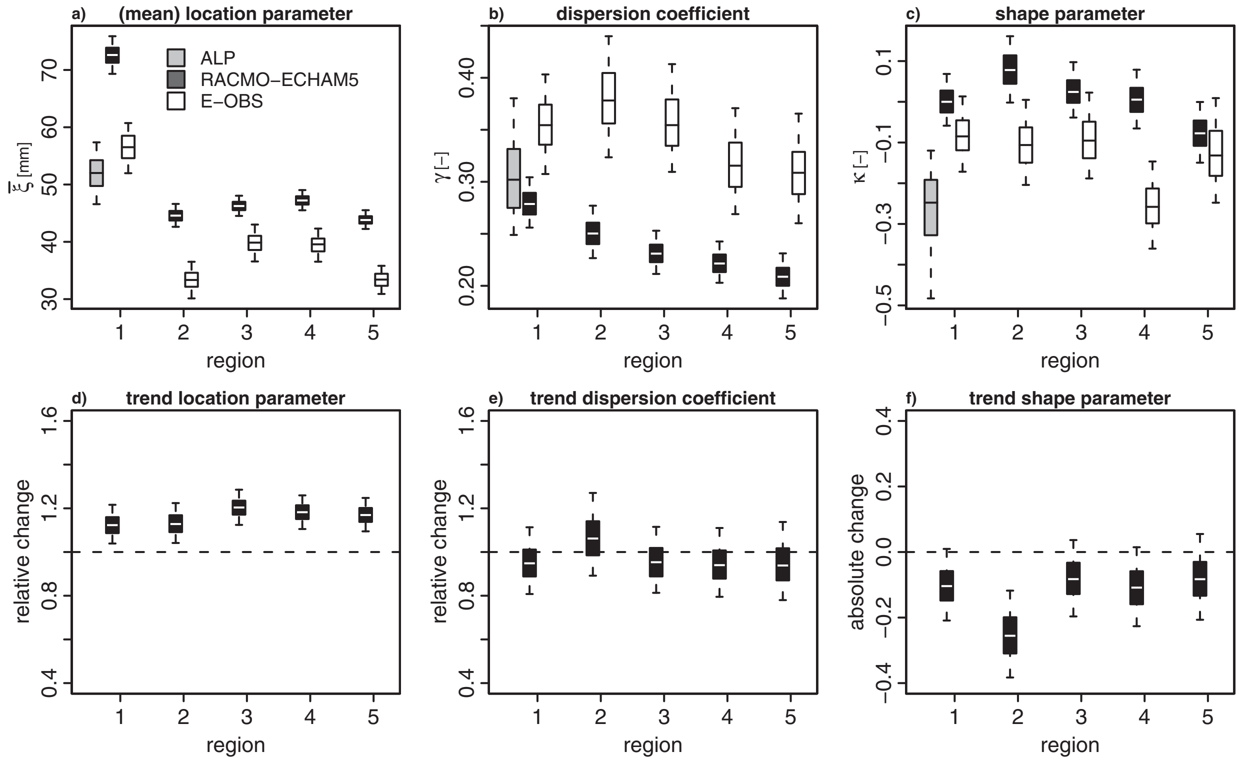

Figure 11. Same as Figure 4 but for the 5-day winter (DJF) precipitation extremes.

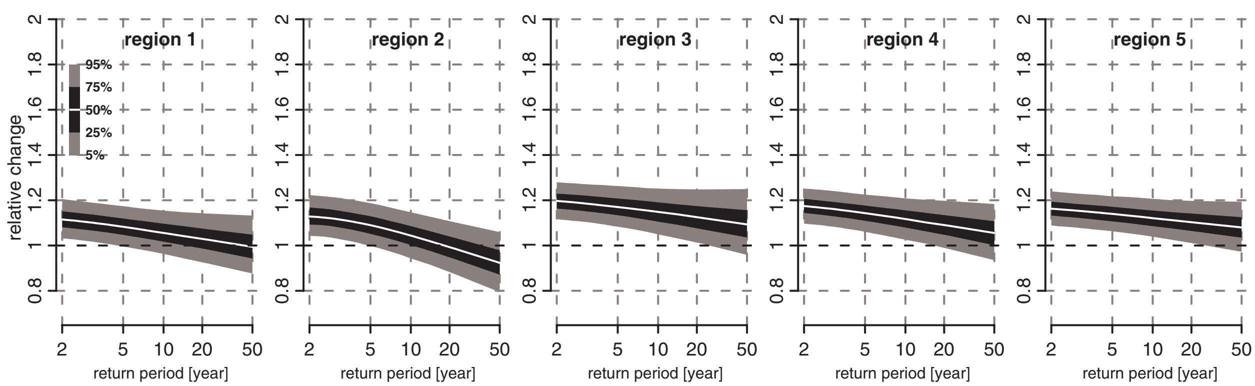

Figure 12. Same as Figure 5 but for the 5-day winter (DJF) precipitation extremes.

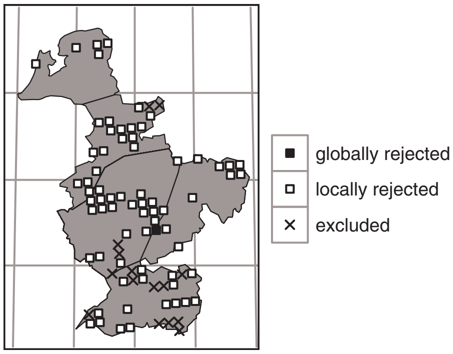

Figure 13. Same as Figure 8 but for the 5-day winter (DJF) precipitation extremes.

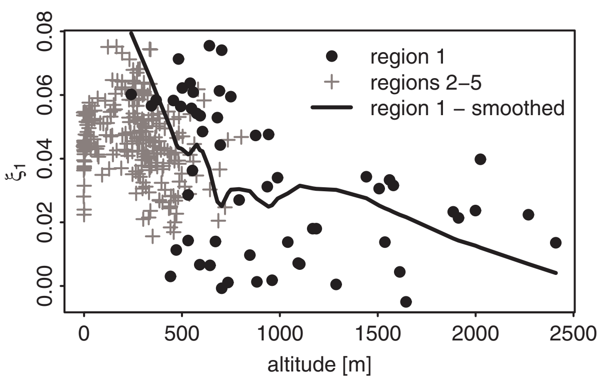

Figure 14. Grid box estimates of the trend in the location parameter as a function of altitude for the 5-day winter (DJF) precipitation maxima in the RACMO-ECHAM5 simulation. The values for the grid boxes in region 1 (black dots) are smoothed by locally weighted regression "loess" (black line).