Determination of Precipitation Return Values in Complex Terrain and Their Evaluation

Barbara Früh,* Hendrik Feldmann, Hans-Jürgen Panitz, and Gerd Schädler

Institute for Meteorology and Climate Research, KIT Karlsruhe, Karlsruhe, Germany

Daniela Jacob and Philip Lorenz

Max Planck Institute for Meteorology, Hamburg, Germany

Klaus Keuler

Department of Environmental Meteorology, Brandenburg University of Technology, Cottbus, Brandenburg, Germany

ABSTRACT

To determine return values at various return periods for extreme daily precipitation events over complex orography, an appropriate threshold value and distribution function are required. The return values are calculated using the peak-over-threshold approach in which only a reduced sample of precipitation events exceeding a predefined threshold is analyzed. To fit the distribution function to the sample, the L-moment method is used. It is found that the deviation between the fitted return values and the plotting positions of the ranked precipitation events is smaller for the kappa distribution than for the generalized Pareto distribution.

As a second focus, the ability of regional climate models to realistically simulate extreme daily precipitation events is assessed. For this purpose the return values are derived using precipitation events exceeding the 90th percentile of the precipitation time series and a fit of a kappa distribution. The results of climate simulations with two different regional climate models are analyzed for the 30-yr period 1971–2000: the so-called consortium runs performed with the climate version of the Lokal Modell (referred to as the CLM-CR) at 18-km resolution and the Regional Model (REMO)–Umweltbundesamt (UBA) simulations at 10-km resolution. It was found that generally the return values are overestimated by both models. Averaged across the region the overestimation is higher for REMO–UBA compared to CLM-CR.

1. Introduction

Although climate change occurs on the global scale, its impact varies substantially on local and regional scales (Good and Lowe 2006). Climate projections studying the effects of rising greenhouse gas concentrations are typically provided by global circulation models (GCMs). Since the resolution of the GCMs is rather coarse, the results need to be scaled down to a smaller scale with a higher degree of detail (Giorgi and Mearns 1999). This transfer can be achieved, for example, by dynamical downscaling using regional (i.e., limited area) climate models (RCMs) that are driven by GCMs or by analyzed meteorological fields at their boundaries. The use of the latter reduces errors in the RCM simulations induced by biased boundary conditions from the driving data. However, for future projections GCM-derived boundary conditions are indispensable.

During the last decade, studying climate change on the regional scale was a focal point of several international projects, for example, the Prediction of Regional Scenarios and Uncertainties for Defining European Climate Change Risks and Effects (PRUDENCE; information online at http://prudence.dmi.dk) or the European Union's ENSEMBLES project (information online at http://www.ensembles-eu.org). Thus, the number of regional climate change experiments has grown significantly (Solomon et al. 2007). The typical grid size of such regional simulations is on the order of 50 km (e.g., Frei et al. 2003, 2006; Kunkel et al. 2002; Bülow 2010). A validation of such model

* Current affiliation: Deutscher Wetterdienst, Offenbach am Main, Germany.

Corresponding author address: Barbara Früh, Deutscher Wetterdienst, Frankfurter Str. 135, D-63067 Offenbach am Main, Germany.
E-mail: barbara.frueh@dwd.de

results by Beniston et al. (2007) for southern Germany shows an adequate representation of the 5-yr return level of daily precipitation but an underestimation of the spatial variability, which was probably caused by an unrealistic representation of orography due to the coarse model resolution. The hypothesis that increased resolution improves model results can now be tested since recently two RCM simulations for Europe have become available with grid resolutions below 20 km: the so-called consortium runs (CR) with the climate version of the Lokal Modell (CLM) (Böhm et al. 2006) and the Regional Model (REMO)–Umweltbundesamt (UBA) simulations (Jacob et al. 2008), commissioned by the German Federal Office for Environment Protection, using the REMO regional climate model (Jacob 2001; Jacob et al. 2007). We use both RCM simulations for our analysis to assess the spread of the results. However, due to several differences in the configuration of the simulations (domain, resolution, nesting levels, hydrostatic and nonhydrostatic, etc.), the analysis does not allow one to address the effects of specific configuration characteristics.

One main concern with regard to the future climate is related to the frequency of occurrence and the intensity of extreme events since very often they have a large impact on society and considerable potential to cause damage (van den Brink et al. 2005). An extreme weather event is defined as an event that is rare, that is, has long return periods, at a particular place and time of year and falls below the 10th or exceeds the 90th percentile of the observed probability density function (Solomon et al. 2007). Investigating the regional impact of extreme precipitation is of particular importance because of its high spatial variability as well as implications concerning droughts, floods, or drainage overflows.

Different methods to derive return values (RVs) for extreme daily precipitation events exist (e.g., Maraun et al. 2008; Park and Jung 2002; Casas et al. 2007). An adequate method for describing these rare events is offered by the peak-over-threshold approach, with the advantage of analyzing a reduced sample with precipitation events exceeding a certain threshold only (Wilks 2006). For this approach the generalized Pareto distribution (GPD), as well as the kappa distribution (KD), was designed. Both distributions are fitted using the L-moments method (Hosking and Wallis 1987). We investigated the impact of GPD and KD and varying threshold values on the RV. GPD was also employed by Paeth and Hense (2005) for rainfall extremes in the Mediterranean and by Li et al. (2005) for events in western Australia. Park and Jung (2002) and Parida (1999) used KD to estimate rainfall quantiles in Korea and India, respectively. In addition we assessed the ability of the two RCMs to realistically simulate regional-scale precipitation extremes

at various return periods. Our analysis is intentionally restricted to the 30-yr period 1971–2000 and to southwest Germany because of the complex orography, which should particularly benefit from the high resolution of the RCM and the high data availability for that region. An analysis of the mean precipitation of both RCMs for the same time period and region can be found in Feldmann et al. (2008).

In section 2 a brief description of the RCMs used is given. The high-resolution precipitation climatology is introduced in section 3. Section 4 describes the statistical methods applied. Results with respect to the appropriate threshold and distribution function are presented and discussed in section 5. Section 6 shows the evaluation of the return values. We close the paper with a short summary and some conclusions in section 7.

2. Description of models and experiments

Within this study we use the results of two RCM simulations: the REMO–UBA simulations (Jacob et al. 2008) (grid resolution: $0.088° \approx 10$ km) commissioned by the German Federal Environmental Agency with the hydrostatic RCM REMO (Jacob 2001) and the so-called consortium runs (CLM-CR) (Hollweg et al. 2008) (resolution: $0.165° \approx 18$ km) performed with the nonhydrostatic RCM CLM (Böhm et al. 2006). Both models were derived from routine weather prediction models adapted for climate applications. A brief overview of the models and the simulation setup is given in Table 1.

The model experiments analyzed are taken from simulations with forcing for the twentieth century and cover the period from 1971 to 2000. To estimate the natural variability of precipitation extremes we compared the three realizations of CLM-CR, which only differ in the initial conditions of the large-scale internal variability provided by the driving GCM. We found that the difference of the 10-yr return values between realization 1 and the average of all realizations ranges between −12% and +7% in our domain. Averaged across the region the difference is −0.3% only. Thus, we conclude that the 30-yr period chosen represents a sufficiently long time span for the comparison of the simulated precipitation statistics with observations.

Both models were driven with data from the global climate model ECHAM5 (Roeckner et al. 2006a,b; Hagemann et al. 2006). The ECHAM5 simulation (twentieth-century simulation, realization 1) uses observed anthropogenic forcing for CO_2, CH_4, N_2O, CFCs, O_3, and sulfate initialized by a preindustrial control simulation. This simulation neglects natural forcing from volcanoes and changes of solar activity. The grid resolution is T63 ($1.87° \approx 200$ km) with 31 vertical layers. ECHAM5

TABLE 1. Description of the CLM-CR and REMO–UBA model setups.

	CLM (nonhydrostatic)	REMO (hydrostatic)
Reference	Doms and Schättler (2002)	Jacob (2001)
	Doms et al. (2007)	
	Böhm et al. (2006)	
Based on	Lokal Modell (LM)	Europa Modell (EM) with parameterizations from ECHAM4
Expt	Consortium runs (CLM-CR)	UBA simulations (REMO–UBA)
Resolution	0.165° (≈18 km)	0.088° (≈10 km) nested in 0.44° (≈50 km) coarse grid
Grid points	257 × 271	109 × 121
Layers	32 (11 layers below 2000 m)	27 (8 layers below 2000 m)
Model domain	Polar stereographic grid with rotated	Polar stereographic grid with
corners	pole at 39.25°N, 162°W	rotated pole at 39.25°N, 162°W
		inner grid
lower left	24.24°N, 7.46°W	44.93°N, 4.87°E,
upper left	62.81°N, 40.97°W	55.29°N, 1.61°E
lower right	26.49°N, 35.81°E	45.76°N, 18.31°E
upper right	67.65°N, 63.95°E	56.32°N, 18.38°E
		Coarse grid
lower left		19.82°N, 7.53°W
upper left		64.22°N, 50.51°W
lower right		21.69°N, 37.87°E
upper right		68.58°N, 77.70°E
Dataset citation	CLM-CR	REMO-UBA
	Lautenschlager et al. (2008)	Jacob (2008)

is run in a coupled mode with the Max Planck Institute Ocean Model (MPI-OM). The CLM-CR simulations were nested directly into the ECHAM5 fields. For REMO–UBA, a two-step nesting was applied. A coarser REMO simulation at 0.44° resolution is driven by ECHAM5 and provides the boundary values for the high-resolution REMO–UBA simulation at 0.088°.

Our study area encompasses the region from 47.5° to 49.8°N, 7.5° to 10.5°E, and covers the federal state of Baden-Württemberg, Germany (Fig. 1). The horizontal extent is about 225 km × 255 km. This area is characterized by complex topography and includes orographic features like, for example, the Rhine Valley, the Black Forest, the Swabian Jura, and the valleys of the Neckar and the upper Danube. The Black Forest extends approximately in the north–south direction. Its highest elevation is the Feldberg at nearly 1500 m AMSL, which is represented by a grid cell elevation of 1041 m for the REMO–UBA and 952 m for the CLM-CR resolution. Freiburg (278 m AMSL) with a grid cell height 406 m for the REMO–UBA and 370 m for the CLM-CR resolution is situated west of the Black Forest, that is, to the windward side of the mountain range with regard to the prevailing wind direction. On the other side, Stuttgart (245 m AMSL with grid cell height 340 m for the REMO–UBA and 286 m for the CLM-CR resolution) is situated in the lee of the Black Forest. Adjacent to the

Black Forest is the Rhine River to the south and west, the Danube River to the southeast, and the Neckar River to the northeast. The Swabian Jura is a high plateau between the Neckar and the upper Danube with an altitude up to 1000 m AMSL. South of the Danube are the northern foothills of the European Alps.

3. Observed climatological precipitation

In this paper we use observed daily precipitation data for the period 1971–2000 interpolated onto the model grids of both CLM-CR and REMO–UBA to compare both simulations with observations at their specific spatial resolution. We call the observed daily precipitation fields OBS10 at the REMO–UBA resolution and OBS18 at the CLM-CR resolution.

To derive OBS10 and OBS18 we make use of the high-resolution (1.25′ ≈ 2 km) climatology for the Alpine region (Schwarb 2001; Schwarb et al. 2001; Frei and Schär 1998), explicitly taking into account topography. It is therefore especially suitable for the complex orography in the area of interest. In the northernmost part of the region, these data are combined with the precipitation climatology of the German Meteorological Service (Müller-Westermeier 1995) to derive a comprehensive observation dataset for the evaluation of the RCMs. Processing of the climatological and daily precipitation
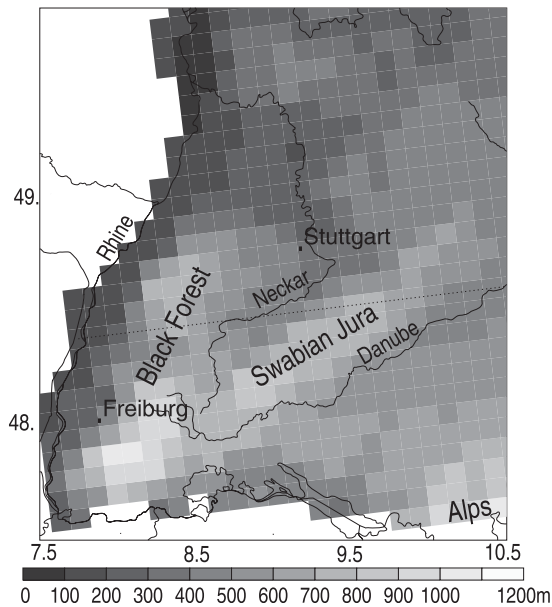
FIG. 1. Map of the orography of southwest Germany at 10-km resolution.

fields is described in detail in Früh et al. (2006) and Früh et al. (2007), respectively.

The interpolation of point measurements onto a regular grid generally leads to a smoothing and thus to a reduction of the variance in the time series. Specifically, the averaged annual standard deviation of the observed precipitation exceeding, for example, the 90th percentile is 17% and 19% lower at 10- and 18-km resolution, respectively, compared to the point measurements. The reduction of the variance narrows the probability density function and, thus, results in lower values for the extreme precipitation events in the interpolated field compared to the station measurements.

Uncertainties and errors of the observed climatology are addressed in Frei et al. (2003), who found that the bias is slightly lower for high precipitation rates than for lower rates because the rain gauge undercatch depends on hydrometeor size and hence on precipitation intensity. In winter and at altitudes between 600 and 1500 m (with a high percentage of snowfall), the undercatch can be up to −12%, depending on the exposure of the measurement site. The precipitation data employed (OBS10 and OBS18) were not corrected for these errors due to the lack of information on the exposure of the stations within the network.

The observed annual cycle of mean precipitation in southwest Germany (Feldmann et al. 2008) exhibits higher precipitation in summer than in winter except in the Black Forest where two maxima exists: one in summer and a higher one in winter. This enhanced winter precipitation is caused by orographic lifting mainly on the western side, which is exposed to the dominant southwesterly flow regime during this season. The annual precipitation amount ranges from ~550 mm in the Rhine Valley to 2000 mm in the Black Forest. There is also a pronounced lee effect to the east of the Black Forest with reduced precipitation. For the Rhine, Danube, and Neckar region the typical annual precipitation amount range between 600 and 1000 mm and for the Swabian Jura between 900 and 1400 mm. The seasonal precipitation sums averaged over the region of interest are 226 mm in March–May (MAM), 285 mm in June–August (JJA), 228 mm in September–November (SON), and 214 mm in December–February (DJF).

4. Analysis method

a. Peak-over-threshold approach

For the analysis of extreme values, statistical distributions need to be derived that closely fit the climate parameter in the upper tail of the distribution. For heavy precipitation events the peak-over-threshold approach (Wilks 2006) (partial-duration series in hydrology), which analyzes precipitation events exceeding a sufficiently high threshold only, seems to be most suitable (Brabson and Palutikof 2000). This approach is an alternative to the classical approach that typically uses annual or seasonal maximum values.

To fit daily extreme precipitation events a number of specific theoretical distributions have been proposed. While the Wakeby and generalized extreme value distributions (Gumbel, Fréchet or Weibull distribution) (Wilks (2006) refer to the classical extreme value theory (Kharin and Zwiers 2000), the three-parameter generalized Pareto distribution (GPD) is especially designed for the peak-over-threshold approach (Hosking and Wallis 1987). The cumulative distribution function of the GPD for the sample x is given by

$$F_{\mathrm{GPD}}(x) = 1 - \left(1 - \frac{k_{\mathrm{GPD}}(x - \xi_{\mathrm{GPD}})}{\alpha_{\mathrm{GPD}}}\right)^{1/k_{\mathrm{GPD}}}. \quad (1)$$

Its quantile function ($x_{\mathrm{GPD}}(F)$), which is the inverse of $F_{\mathrm{GPD}}(x)$, is given by

$$x_{\mathrm{GPD}}(F) = \xi_{\mathrm{GPD}} + \alpha_{\mathrm{GPD}} \frac{1 - (1 - F_{\mathrm{GPD}})^{k_{\mathrm{GPD}}}}{k_{\mathrm{GPD}}}. \quad (2)$$

The three parameters of the GPD describe the location (ξ_{GPD}) that coincides with the threshold chosen, the scale (α_{GPD}) as a measure for the dispersion, and the shape (k_{GPD}).
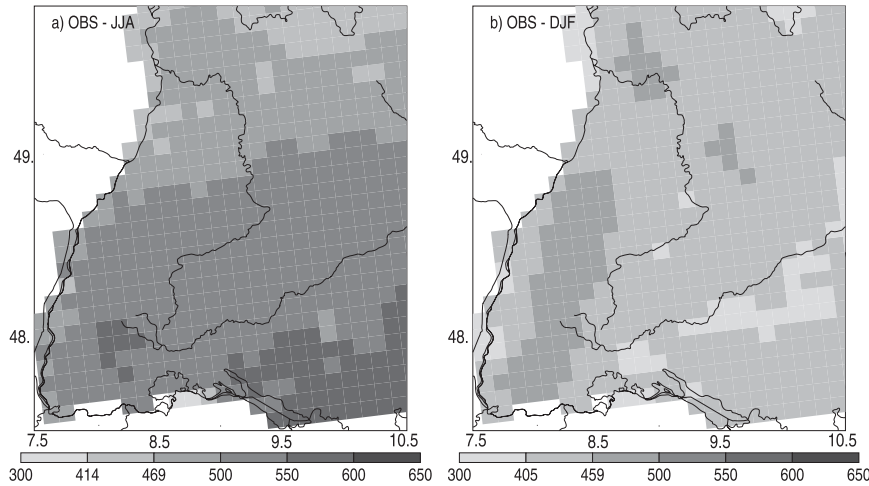
FIG. 2. Map of the number of independent precipitation events exceeding 1 mm day^{-1} for (a) JJA and (b) DJF.

The GPD may be regarded as a special case of the four-parameter kappa distribution (KD) (Hosking 1994). The cumulative distribution function of the KD is given by

$$F_{KD}(x) = \left[1 - h_{KD}\left(1 - \frac{k_{KD}(x - \xi_{KD})}{\alpha_{KD}} \right)^{1/k_{KD}} \right]^{1/h_{KD}} \quad (3)$$

with the quantile function ($x_{KD}(F)$),

$$x_{KD}(F) = \xi_{KD} + \frac{\alpha_{KD}}{k_{KD}}\left[1 - \left(\frac{1 - F^{h_{KD}}}{h_{KD}} \right)^{k_{KD}} \right]. \quad (4)$$

The KD is specified by four parameters: ξ_{KD} describes the location, α_{KD} the scale, and k_{KD} and h_{KD} the shape of the distribution. For $h_{KD} = 1$, KD becomes the GPD. A detailed discussion on the characteristics of GPD and KD can be found in Hosking and Wallis (1987) and Hosking (1994).

b. Percentiles

To investigate the impact of different thresholds on the return values we extract precipitation events exceeding the 83th, 85th, 90th, 95th, and 97th percentiles for all days (including the dry ones) and, thus, obtain a fixed sample size (N_{thr}) for each grid point and percentile. The different samples are called P_{thr} = P83, P85, P90, P95, and P97. To make sure that the exceedances in our sample are mutually independent, that is, that they do not belong to the same precipitation event, we require that the precipitation events are separated by at least two dry days. Dry days are defined as days with less than 1 mm precipitation. This requirement leads to a

smaller total sample size (n_{indep}) and, thus, a smaller threshold for the particular percentile. For the sample that is finally used to fit the distribution functions we selected the largest independent precipitation events at each grid point where the sample size (N_{thr}) only depends on the season (90 for DJF and 92 for JJA) and percentile (P_{thr}) considered.

Figure 2 shows the number of independent precipitation events for JJA and DJF. In JJA short-term convective precipitation events dominate, especially over the northern foothills of the Alps. Thus, the number of independent precipitation events exceeding 1 mm day^{-1} (n_{indep}) is high, especially over the southeast of the domain (Fig. 2a). The maximum and minimum n_{indep} yields 620 and 410 events in JJA of 1971–2000. In DJF n_{indep} is distinctly lower with a maximum and minimum of 500 and 326 since synoptic-scale precipitation events usually tend to last several days (Fig. 2b). The spatial pattern of n_{indep} mainly reflects the orography. The maximum n_{indep} can be found over the ridge of the Black Forest.

In Fig. 2 we also see that, when a threshold of P83 or P85 is chosen, n_{indep} is lower than the necessary number N_{thr} for the peak-over-threshold approach at some grid points. These grid cells can be identified in Fig. 2 by the light gray (first) color for P85 and, additionally, the slightly darker gray (second) color for P83. In DJF n_{indep} falls below N_{thr} very often. In these cases the sample is filled up to the fixed N_{thr} with precipitation days with very low intensities.

c. L-moments

Usually, sample moment statistics is used to fit an observed sample to a postulated distribution. The

disadvantages of moment statistics are that it is bounded with limits depending on the sample size and that, in many small or moderate samples, skewness and kurtosis rarely take values anywhere near the population values (Hosking and Wallis 1997). Thus, the estimated parameters of the distributions fitted by the method of moments are often markedly less accurate than those obtainable by other estimation procedures such as the method of maximum likelihood (Hosking 1990), which picks the values of the model parameters that make the data most likely for a fixed set of data and the selected distribution function. As a further approach, we use the method of L-moments because the required computational effort is quite small compared to other traditional techniques, such as maximum likelihood, and because L-moment estimators have better sampling properties than the method of maximum likelihood or the method of conventional moments (Zwiers and Kharin 1998). Hosking and Wallis (1997) showed that L-moments are efficient in estimating parameters of a wide range of distributions from small samples. The main advantage of L-moments is that, being a linear combination of data, they are less influenced by outliers.

The L-moments are in some way analogous to the traditional moments of a distribution describing the mean, variance, skewness, and kurtosis of a statistical distribution. The estimation of the sample L-moments is based on a sample of size N_{thr}, arranged in ascending order. Let $x_1 \leq x_2 \leq \cdots \leq x_{N_{thr}}$ be the ordered sample time series. At first an estimator for the probability weighted moment is determined via

$$b_0 = \frac{1}{N_{thr}} \sum_{k=1}^{N_{thr}} x_k;$$

$$b_r = \frac{1}{N_{thr}} \sum_{k=r+1}^{N_{thr}} \frac{(k-1)(k-2)\ldots(k-r)}{(N_{thr}-1)(N_{thr}-2)\ldots(N_{thr}-r)} x_k,$$

$$r > 0. \tag{5}$$

The first four sample L-moments are defined by

$$l_1 = b_0$$
$$l_2 = 2b_1 - b_0$$
$$l_3 = 6b_2 - 6b_1 + b_0$$
$$l_4 = 20b_3 - 30b_2 + 12b_1 - b_0. \tag{6}$$

The parameter estimates for the distribution functions are obtained by equating the first p sample L-moments to the corresponding distribution quantity. Such expressions have been obtained for many standard distributions (Hosking and Wallis 1997) including the GPD and KD.

TABLE 2. Percentiles $F(RV)(\%)$ assigned to the return periods (T) for P90 in JJA/DJF.

JJA/DJF	T(yr)
89.13/88.89	1.00
97.83/97.78	5.00
98.91/98.89	10.00
99.78/99.78	50.00
99.89/99.89	100.00

To ensure the existence [Eqs. (7a) and (7b)] and uniqueness [Eqs. (7c) and (7d)] of the L-moments for KD, the following requirements must be fulfilled (Hosking 1994):

$$
\begin{aligned}
&\text{(a) } k_{KD} > -1 \\
&\text{(b) } h_{KD}k_{KD} < -1 \quad \text{if} \quad h < 0 \\
&\text{(c) } h_{KD} > -1 \\
&\text{(d) } k_{KD} + 0.725h_{KD} > -1.
\end{aligned}
\tag{7a–d}
$$

The L-moments for GPD are defined for $k_{GPD} > -1$ (Hosking and Wallis 1997).

d. Return values

To determine the return value (RV) of the time series (x_k, $k = 1, \ldots, N_{thr}$) for the return period (T), we first define the crossing rate (λ) as the expected number of events exceeding the threshold per year or season (Brabson and Palutikof 2000):

$$\lambda = \frac{N_{thr}}{n_{years}}. \tag{8}$$

The number of exceedances in t years can then be calculated via

$$\lambda_x = \lambda t(1 - F(x)), \tag{9}$$

with $F(x)$ from Eq. (1) or Eq. (3). The RVs are defined as the precipitation intensity that is exceeded by one annual or seasonal extreme with a probability of $1/T$. Therefore, λ_x in Eq. (9) can be set to unity, x to RV, and t to T:

$$1 - F(RV) = \frac{1}{\lambda T}. \tag{10}$$

Note that the relation between $F(RV)$ and T depends on N_{thr}. Inserting Eq. (10) into Eq. (2), the quantile function for GPD becomes

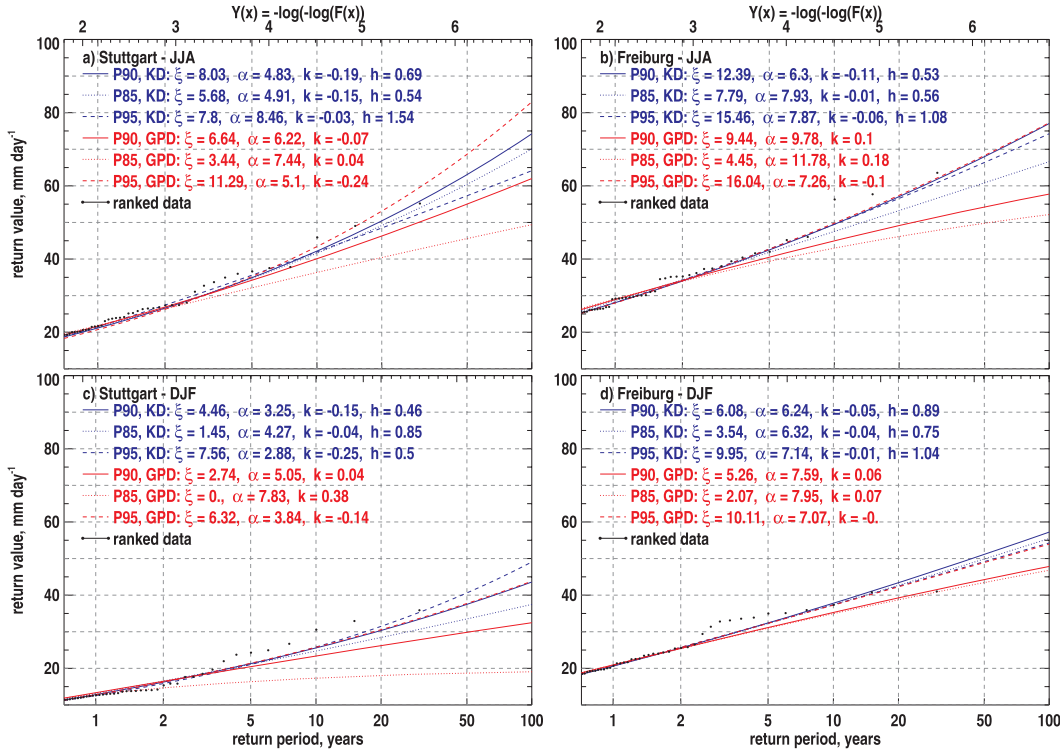$$RV_{GPD} = \xi_{GPD} + \alpha_{GPD} \frac{1 - (\lambda T)^{-k_{GPD}}}{k_{GPD}} \tag{11}$$

FIG. 3. Gumbel plots for two different grid cells close to the cities of (a),(c) Stuttgart and (b),(d) Freiburg for (a),(b) JJA and (c),(d) DJF. The blue (red) curves depict the return values of the KD (GPD). The solid (dotted, dashed) lines refer to the P90 (P85, P95) threshold. The black dots display the plotting positions of the ranked data.

and for KD:

$$\mathrm{RV}_{\mathrm{KD}} = \xi_{\mathrm{KD}} + \frac{\alpha_{\mathrm{KD}}}{k_{\mathrm{KD}}} \left\{ 1 - \left[\frac{1 - (1 - 1/\lambda T)^{h_{\mathrm{KD}}}}{h_{\mathrm{KD}}} \right]^{k_{\mathrm{KD}}} \right\}. \tag{12}$$

The cumulative probability values for a given return period (T) calculated for P90 of the JJA and DJF samples are listed in Table 2. Inserting Eq. (8) into Eq. (10) allows one to compute the percentiles as the cumulative distribution function ($F(\mathrm{RV})$). Since the number of exceedances (λ) is higher for JJA than for DJF, the percentiles belonging to the respective return periods are also higher.

Here $F(\mathrm{RV})$ can be estimated empirically for each of the observed precipitation time series (x_k, $k = 1, \ldots, N_{\mathrm{thr}}$) by simply ordering the data from the smallest (x_1) to the largest ($x_{N_{\mathrm{thr}}}$), and calculating an empirical value of $F(x_k)$ from the ranked position of x_k (Palutikof et al. 1999). These estimates are known as the plotting positions, which can be calculated (Gumbel 1958; Makkonen 2007) via

$$F(x_k) = \frac{\mathrm{rank}(x_k)}{N_{\mathrm{thr}} + 1}. \tag{13}$$

Since the sample x_k is in ascending order, $\mathrm{rank}(x_k) = k$. The corresponding return period, $\hat{T}_k$, can be calculated from Eq. (13) using Eq. (10):

$$\hat{T}_k = \frac{1}{\lambda(1 - F(x_k))} = \frac{N_{\mathrm{thr}} + 1}{\lambda(N_{\mathrm{thr}} + 1 - k)}. \tag{14}$$
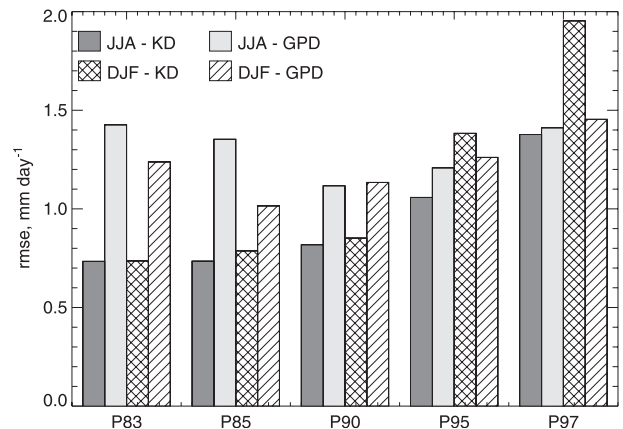

FIG. 4. Domain-averaged RMSE between the plotting positions of the ranked data and the fitted RV for P83, P85, P90, P95, and P97 for KD (dark gray or cross hatched) and GPD (light gray or hatched) in JJA (solid) and DJF (hatched).
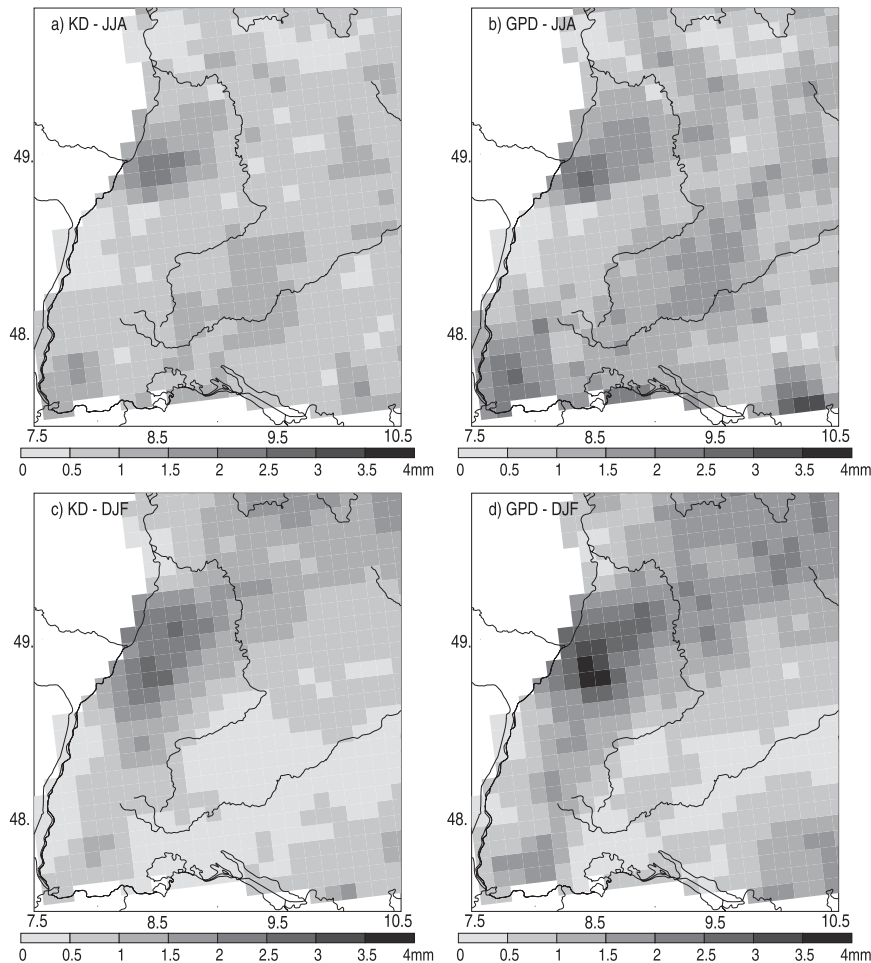
FIG. 5. Map of the RMSE between the plotting position and the fitted RV for P90 for (a),(c) KD and (b),(d) GPD in (a),(d) JJA and (c),(d) DJF.

To estimate the sampling uncertainty with respect to the RV, we follow a practical approach using a parametric bootstrap method (Zwiers and Kharin 1998) with 500 synthetic samples. The 5th and 95th percentiles of the resulting collection of RV estimates are then used as lower and upper 90% confidence bounds for the true RV. This procedure accounts for the uncertainty due to the short-term natural variability (Kendon et al. 2008). The presence of variability on longer time scales (e.g., multidecadal) could lead to an additional uncertainty.

In addition, we compute the signal-to-noise ratio, which is the ratio of the synthetical mean return value divided by the 90% confidence interval. If the ratio is small, then the estimation of the RV is uncertain.

5. Results for an appropriate threshold and distribution function

In this section we investigate the dependence of the return value on the choice of the threshold and distri-

bution function (see section 4). For this purpose we compare the fitted RV for several thresholds and the two distribution functions with the plotting positions of the ranked precipitation events. These investigations are performed employing the OBS10 dataset.

a. Gumbel plots for Stuttgart and Freiburg

Figure 3 displays the Gumbel plots for two grid points close to Stuttgart and Freiburg for JJA and DJF. The geographic positions of these cities are indicated in Fig. 1. The percentiles and probabilities assigned to the return period (T) on the abscissa are listed in Table 2 for the P90 sample.

The fitted RV for the thresholds displayed are very similar for T shorter than about 5 yr. Only for longer T do the thresholds have a noticeable effect on the resulting RV. The fitted RV and their variability due to the different thresholds and distribution functions are higher in JJA than in DJF. The variability of the RV is larger in

Stuttgart, in the lee of the Black Forest, than in Freiburg in its luff. In Stuttgart, especially, the P95 threshold leads to very high RV for long T from the GPD in JJA. This heavy tail is caused by the very low value for the shape parameter ($k_{\mathrm{GPD}} = -0.24$) for this sample fit.

The fitted RV for P90 are higher for KD than for GPD for T longer than about 2 yr. The difference between the return values of KD and GPD increases with increasing T. Averaged across the region, the deviation between the RV resulting from KD and GPD is low, especially for T shorter than about 10 yr ($<\pm 5\%$). The difference in the spatial distribution of the RV between both distributions is also very low (not shown).

b. Root-mean-square deviation of the return values to the plotting positions

However, from these figures it is difficult to decide which threshold and distribution function best fits the plotting positions. To investigate this deviation systematically we calculated the root-mean-square error (RMSE) between the plotting position of the ranked data and the return values resulting from the fitted distribution functions for each threshold and distribution function at each grid cell. For this purpose the RMSE is calculated for the 3% (≈ 80) largest precipitation events to ensure the comparability between the different thresholds chosen. Since the RMSE is based on the square of the difference, it is very sensitive to large deviations.

Figure 4 shows the RMSEs between the plotting position of the ranked data and the fitted RV for P83, P85, P90, P95, and P97 averaged across the region for KD and GPD in JJA and DJF.

In general, the RMSE is relatively low (<2 mm day^{-1}). Here KD yields a lower RMSE than GPD except for the high thresholds P95 and P97 in DJF. In JJA, KD yields the lowest RMSE for P83 and GPD for P85.

The higher RMSE for higher thresholds can be explained by the smaller sample size and thus a higher weight for a single event fitting the distribution function. A small change in the number of events per class has a large impact on the resulting shape of the distribution function and, thus, a considerable influence on the difference between the fitted RV and the ranked observations. In addition, at the far end of the distribution very often only one event per class exists. This discrete character of the binned dataset cannot be represented by any distribution function. By chance, in case of a small sample the fourth parameter of KD can even worsen the fit since the weight of a single event increases with the number of free distribution parameters.

On the other hand, a low threshold leads to a dataset with many moderate events affecting the fit of the distribution function. This fit, which also includes more
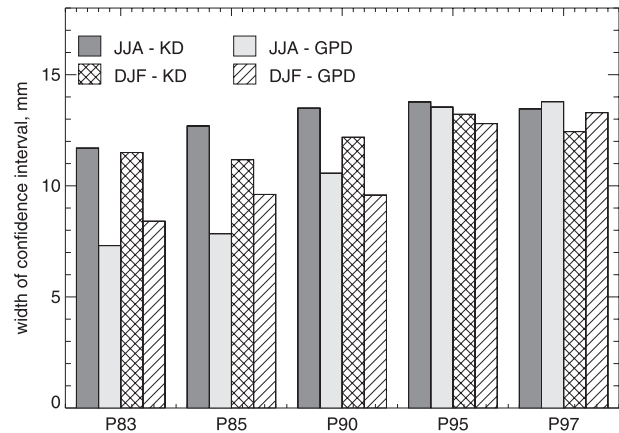

FIG. 6. Domain-averaged width of the confidence interval at the 90% significance level for the $\overline{\mathrm{RV}}_{T=10}$ for P83, P85, P90, P95, and P97 for KD (dark gray or cross hatched) and GPD (light gray or hatched) in JJA (solid) and DJF (hatched).

moderate events, leads to an increased RMSE in the case of GPD on the one hand but a decreased RMSE in the case of KD on the other since the additional parameter of KD enables a better fit to the binned data. This higher flexibility exhibits an advantage of KD over GPD.

The RMSE is smaller in JJA than in DJF except for GPD at P83 and P85.

The spatial variability of the RMSE for P90 can be inferred from Fig. 5. The maximum RMSE is situated in the region around Karlsruhe and in the southeast of the domain. The lowest RMSE can be found in the Neckar region in DJF. The largest spatial variability occurs for GPD with 0.21–3.7 mm day^{-1} in DJF (Fig. 5d). The maximum RMSE for KD amounts to 2.7 mm day^{-1} in DJF (Fig. 5c). In JJA the maximum RMSE is lower for both distribution functions with values of 3.3 mm day^{-1} for GPD and 2.2 mm day^{-1} for KD.

c. Confidence intervals for the 10-yr return value

Figure 6 shows the width of the confidence interval at the 90% significance level for the 10-yr return value ($\mathrm{RV}_{T=10}$) for P83, P85, P90, P95, and P97 averaged across the region for KD and GPD in JJA and DJF. As could be expected, the width of the confidence interval and thus the uncertainty increases with an increasing threshold (decreasing sample size) except for KD at P97.

The width of the confidence interval of GPD is smaller than that of KD. This can be explained by the higher degree of freedom in the case of the kappa distribution with four parameters. For the three-parameter GPD only two parameters are fitted to the data. The third parameter is fixed by the threshold chosen. The higher degree of freedom increases the uncertainty for KD.
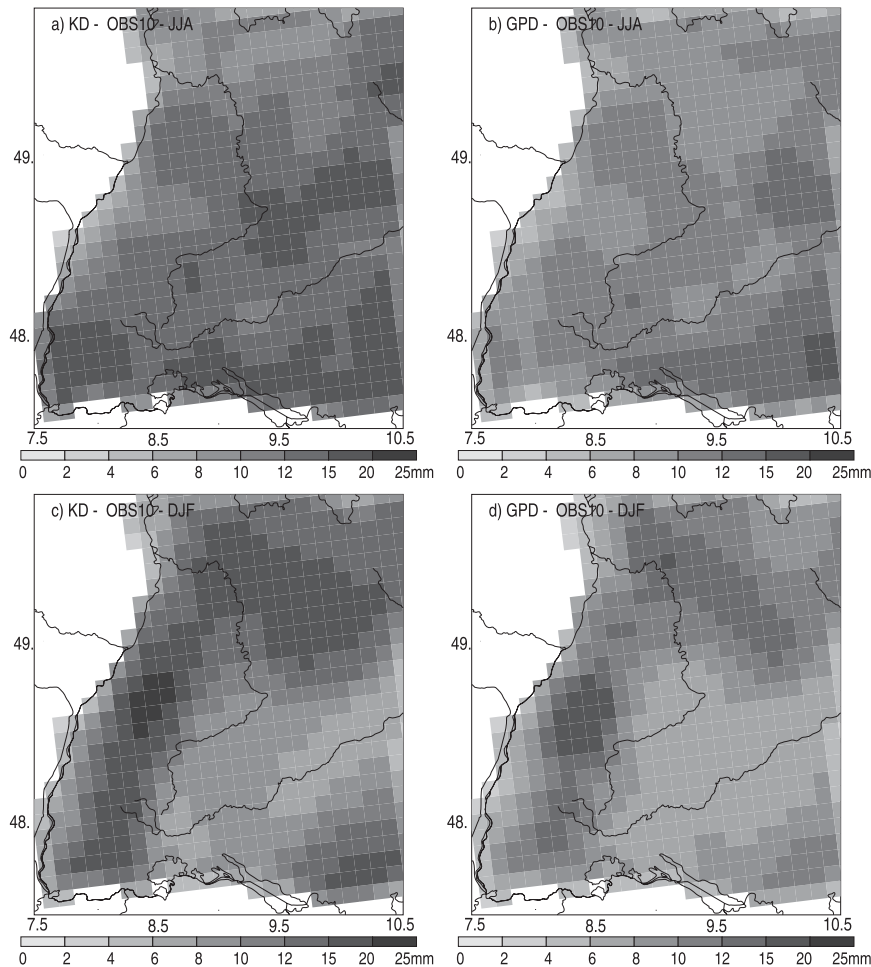
FIG. 7. Map of the width of the confidence interval at the 90% significance level for the $\overline{RV}_{T=10}$ for P90 for (a),(c) KD and (b),(d) GPD in (a),(b) JJA and (c),(d) DJF.

The increase of the width with increasing threshold is stronger for GPD than for KD. For high thresholds (P95 and P97), that is, small sample sizes, the width of the confidence interval of GPD is almost as large as for KD. The width of the confidence interval is smaller for shorter return periods compared to longer ones (not shown).

Figure 7 shows a map of the width of the confidence interval at the 90% significance level for the $RV_{T=10}$ for the P90 sample for KD (panels a and c) and GPD (panels b and d) in JJA (panels a and b) and DJF (panels c and d). The spatial pattern of the width of the confidence interval partly reflects the number of independent precipitation events, which is displayed in Fig. 2.

Comparing Fig. 7 with Fig. 2 (and also Figs. 9 and 10), it becomes apparent that the confidence interval is widest in regions with many independent events (n_{indep}) and large $RV_{T=10}$. Investigating the correlation between n_{indep} and the $RV_{T=10}$ obtained from the sample, we found that the minimum $RV_{T=10}$ of the sample is largest in the mountainous areas of the Black Forest and the alpine foreland in summer as well as winter. In addition, we found that the range of the sample (maximum minus minimum sample $RV_{T=10}$), that is, the spread of the $RV_{T=10}$, is largest in areas where n_{indep} is largest.

To conclude, the RV of both distributions do not differ significantly from the RV of the undistributed, ranked data. Although there is no large difference between the RV of the two distributions, we choose to base our further analysis on the KD with the P90 threshold since the RMSE between the fitted RV and the plotting positions is lower for KD than for GPD and KD is more flexible than GPD. The choice of threshold becomes important only for long return periods.

d. Signal-to-noise ratio

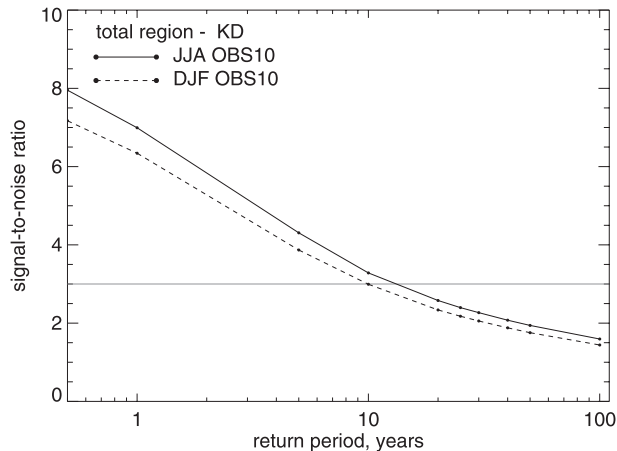Figure 8 shows the regional averaged signal-to-noise ratio of OBS10 for the total region in JJA and DJF. In

FIG. 8. Domain-averaged signal-to-noise ratio as a function of the return period (T) of OBS10 in JJA (solid) and DJF (dashed). The gray horizontal line depicts the signal-to-noise value of three.

general, the signal-to-noise ratio decreases with increasing return period. The higher uncertainty for rarer events is a consequence of their smaller sample size. Because of the high uncertainty of the RV for long T, we analyze return values with return periods up to 10 yr. In this case our samples have an uncertainty less than about one-third of the $RV_{T=10}$, which we consider acceptable.

6. Evaluation of return values

Figure 9 shows the map of the 10-yr return value ($RV_{T=10}$) of OBS10 (panels a and c) and REMO–UBA (panels b and d) for JJA (panels a and b) and DJF (panels d and e). The corresponding maps for the CLM-CR simulation in comparison with OBS18 are given in Fig. 10. For REMO–UBA and CLM-CR a running 3×3 average is applied to the $RV_{T=10}$ to account for the effective model resolution, which is known to be coarser than the numerical grid size.

The spatial pattern of $RV_{T=10}$ is very similar for both models and the observations. REMO–UBA simulates a more distinct rise of $RV_{T=10}$ on the western side of the Black Forest and the northern foothills of the Alps compared to OBS10 or the coarser resolved fields of CLM-CR and OBS18. This pattern can partly be explained by the higher resolution of REMO–UBA compared to CLM-CR. Another part has to be attributed to the formulation of those processes in REMO that affect the redistribution of the atmospheric water. It originates from excess precipitation on the windward side of orographic slopes and a comparable deficit in their lee. This characteristic of the REMO–UBA simulation persists throughout the seasons and return periods. It was also found for the mean precipitation (Feldmann et al. 2008).

TABLE 3. Regional averaged (avg), minimum (min) and maximum (max) $RV_{T=10}$ (mm day^{-1}) in JJA and DJF.

	JJA			DJF		
	Avg	Min	Max	Avg	Min	Max
REMO-UBA	63	41	179	45	17	188
CLM-CR	54	32	88	41	18	118
OBS10	44	27	78	36	17	82
OBS18	44	28	78	36	18	77

This orographically induced misplacement becomes particularly apparent when considering the cross section through the map of $RV_{T=10}$, for example, at 48.5°N, displayed in Fig. 11. Figure 11a refers to the REMO–UBA resolution. The black columns, corresponding to the right ordinate, show the elevation of the grid cells. The simulated $RV_{T=10}$ is indicated by the blue lines (JJA: solid, DJF: dashed) and the observed $RV_{T=10}$ by the green lines. Obviously, REMO–UBA simulates the maximum $RV_{T=10}$ west of the mountain ridge, whereas it is observed on its top. This pattern is also present, though attenuated, at the Swabian Jura especially in DJF. The same pattern evolves for the CLM-CR simulation (Fig. 11b, red lines) but with a distinctly weaker amplitude than for REMO–UBA.

The high $RV_{T=10}$ over the mountainous regions leads to a broader simulated spreading of $RV_{T=10}$ in the REMO–UBA simulation than in the CLM-CR simulation (Table 3). The observed $RV_{T=10}$ of OBS10 and OBS18 is very similar and its variation throughout the region is even lower than for CLM-CR. The intensity of the simulated $RV_{T=10}$ is distinctly higher due to a general overestimation of precipitation by both regional models. The overestimation of the domain-averaged mean precipitation amounts to +9% and +5% in JJA and to +20% and +48% in DJF, respectively (Feldmann et al. 2008). On average, $RV_{T=10}$ is lower in DJF than in JJA but the simulated spreading within the region is much larger in DJF than in JJA. This is due to the very high $RV_{T=10}$ in the Black Forest.

The difference between OBS10 and OBS18 is small. Estimating the areal reduction factor (ARF) relating $RV_{T=10}$ at station observations (245 measuring sites) to the corresponding grid cell values for OBS10 and OBS18 leads to an ARF of 0.896 for 100 km^2 and 0.883 for 324 km^2, respectively, that is, a reduction of about 1.5% on average. This result is consistent with the reduction given by NERC (1975) from ARF = 0.94 at 100 km^2 and ARF = 0.92 at 300 km^2 (−2%). The minimum $RV_{T=10}$ within the region is slightly lower for OBS10 compared to OBS18. The regional maxima of OBS10 and OBS18 are almost identical in JJA but higher for OBS10 than for OBS18 in DJF. The small
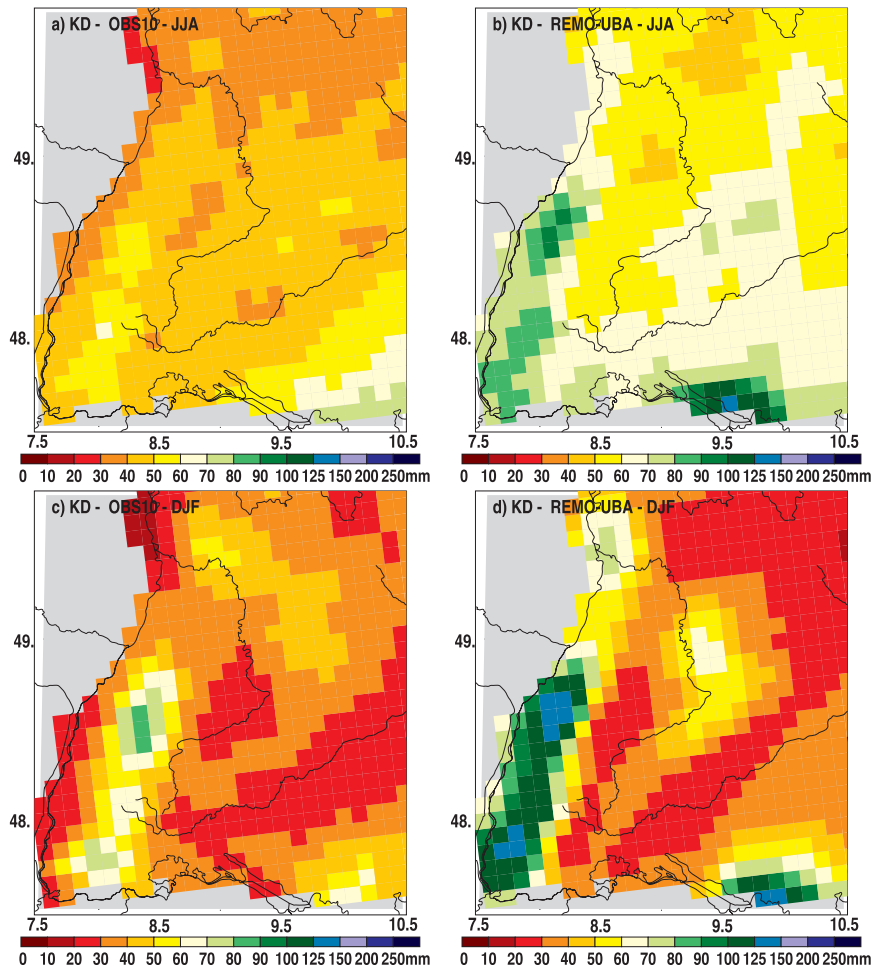
FIG. 9. Map of $\overline{RV}_{T=10}$ for (a),(c) OBS10 and (b),(d) REMO–UBA for (a),(b) JJA and (c),(d) DJF.

difference in the $RV_{T=10}$ may be caused by the relatively small difference between the resolutions for OBS10 and OBS18: 10 and 18 km, respectively. Neither of the resolutions resolve the scale of convection—just the major topographical scales. A stronger dependence of the $RV_{T=10}$ field on the resolution can be expected for much higher or coarser resolutions that can resolve convection or use a broader averaging across the orography. Additionally, the interpolation scheme reduces the standard deviation of the time series at each grid point and, thus, smoothes the RV regionally. Because of the small difference between OBS10 and OBS18, we conclude that the different resolution is responsible for only a minor part of the difference between the two RCM simulations. The larger part is caused by the differences induced by differences in the RCMs themselves.

Figure 12 shows the percentage deviation of the simulated to the observed $RV_{T=10}$ of REMO–UBA to

OBS10 (panels a and c) and CLM-CR to OBS18 (panels b and d) for JJA (panels a and b) and DJF (panels c and d).

In JJA, REMO–UBA and CLM-CR generally overestimate the observed $\overline{RV}_{T=10}$ on average by +45% (range: −6% to +141%) and +24% (range: −5% to +52%), respectively. For coarser resolution extreme precipitation is usually underestimated. This is in part caused by insufficient representation of the orography and interrelated processes. Frei et al. (2003, 2006) found that the extreme rainfall in the alpine region is generally underestimated by models at 50-km resolution. For that study simulation results of both REMO and CLM (there named GKSS for the Forschungszentrum Geesthacht Research Center) are analyzed. Semmler and Jacob (2004) also found that REMO version 5.1, driven by the global climate model HadAM3H at 0.5° resolution, shows a tendency to underestimate the 10-yr and 20-yr RV in southern Germany. In contrast, the overestimation at
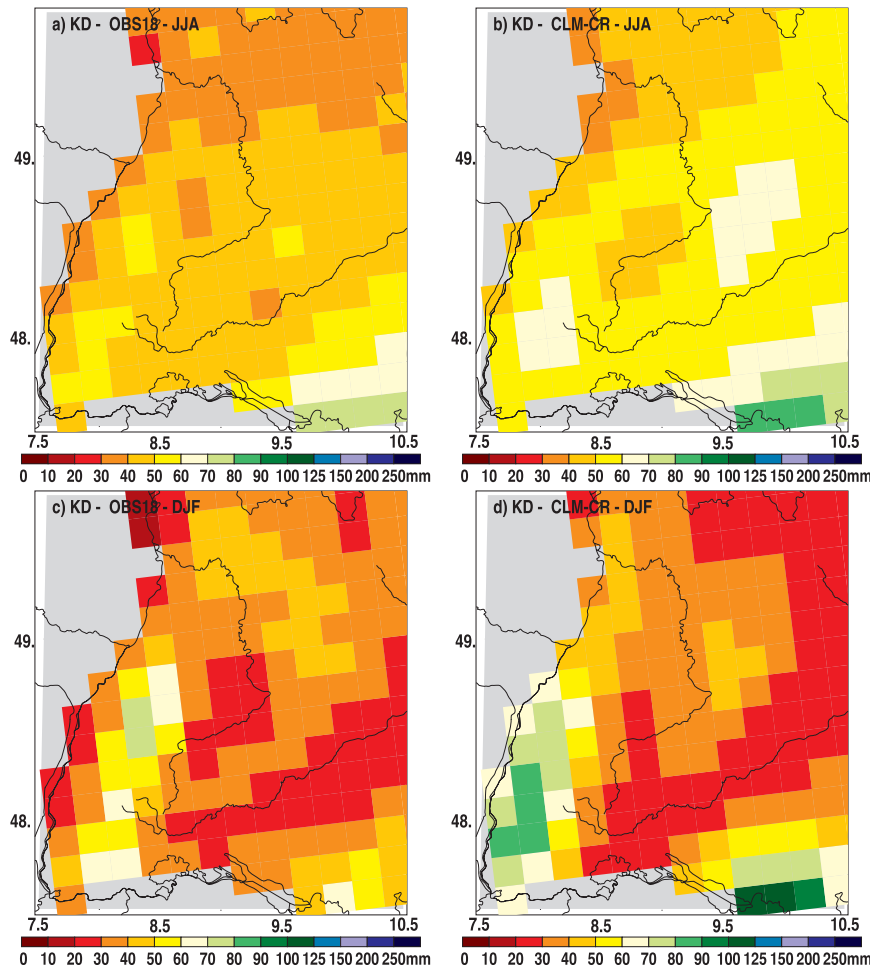
FIG. 10. Map of $\overline{RV}_{T=10}$ for (a),(c) OBS18 and (b),(d) CLM-CR for (a),(b) JJA and (c),(d) DJF.

higher resolutions may be due to inappropriate parameterizations in the RCMs.

The highest deviation of REMO–UBA to OBS10 is located at the western slope of the Black Forest. The deviation pattern of CLM-CR is not as strongly correlated with the orography as REMO–UBA. As visualized in Fig. 11, REMO–UBA overemphasizes the orographic influence in DJF even stronger than in JJA (average deviation +31%, range −68% to +342%). Again, this effect is distinctly weaker for CLM-CR (average deviation +13%, range, −39% to +205%).

The seasonal $\overline{RV}_{T=10}$ of OBS10, REMO–UBA, and CLM-CR averaged across the total region is displayed in Fig. 13a. The percentage deviation of REMO–UBA to OBS10 and of CLM-CR to OBS18 is shown in Fig. 13b.

Averaged across the total region OBS10, REMO–UBA, and CLM-CR simulate the maximum $\overline{RV}_{T=10}$ in JJA, followed by SON and DJF, and the minimum $\overline{RV}_{T=10}$ in MAM. The simulated annual cycle of $\overline{RV}_{T=10}$

has a distinctly larger amplitude than the observations. REMO–UBA and CLM-CR overestimate the seasonal $\overline{RV}_{T=10}$ in all seasons except MAM. The overestimation of REMO–UBA in JJA, SON, and DJF is more distinct than that of CLM-CR. The underestimation in MAM is larger for CLM-CR than for REMO–UBA. The maximum deviation of the simulated $\overline{RV}_{T=10}$ can be found in JJA with an overestimation of more than +40% by REMO–UBA and about +20% by CLM-CR. In the other seasons the deviation is much smaller.

Figure 14 shows the domain-averaged $\overline{RV}$ of OBS10, REMO–UBA, and CLM-CR as a function of T between 0.5 and 100 yr (panel a) and the percentage deviation of the REMO–UBA (blue) and CLM-CR (red) to the observed $\overline{RV}$ Obviously, both REMO–UBA and CLM-CR overestimate the observed $\overline{RV}$ in JJA as well as in DJF.

In JJA the deviation to the observed $\overline{RV}$ increases with increasing T. That means the accordance between REMO–UBA and OBS10 or CLM-CR and OBS18 is
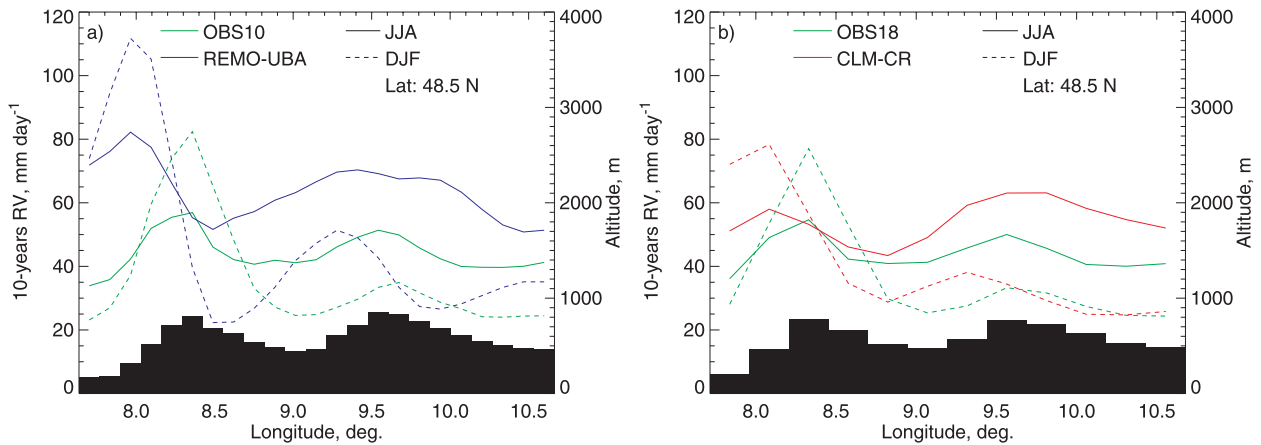
FIG. 11. Cross section through the map of $\overline{RV}_{T=10}$ for (a) REMO–UBA and (b) CLM-CR at 48.5°N.

higher for frequent events with low intensities than for rare events with high intensities. Here, CLM-CR results in better agreement than REMO-UBA. In DJF the increase of the deviation with increasing T is lower than in JJA. Even a decreasing relation is found for CLM-CR.

The reasons for this overestimation might be manifold. 1) The mean precipitation is also distinctly overestimated by both REMO–UBA and CLM-CR (Feldmann et al. 2008). 2) One part of the distinctly high deviation in the Rhine Valley especially for REMO–UBA is induced by the misplacement of the orographically induced precipitation in the western Black Forest (Fig. 11). 3) The systematic undercatch of the rain gauge measurements, especially in DJF, results in a lower observed precipitation amount than really fell. This causes a seeming overestimation of the simulated precipitation. 4) The spatial interpolation of the observations also leads to a smoothing of the rainfall, especially for the extreme events, and thus to possibly lower values in the interpolated field compared to the "true" intensities. 5) The stronger overestimation of REMO–UBA compared to CLM-CR may, to some extent, be caused by the higher resolution of REMO–UBA since a coarser resolution smoothes extreme events.

7. Summary and conclusions

We discussed the ability of the peak-over-threshold approach to derive return values (RVs) for extreme daily precipitation events. The advantage of this approach is the focus of the analysis on heavy precipitation events since the sample is reduced to events exceeding a certain threshold. To analyze the same number of independent precipitation events across all grid points we applied a relative threshold. However, in some regions this leads to a sample not only containing heavy but also moderate

precipitation events, particularly in winter. For the peak-over-threshold approach several distribution functions like the generalized Pareto distribution (GPD) and the more general kappa distribution (KD) were designed to represent the frequency of occurrence of the sample. To fit these distribution functions to the sample we used the L-moment method. Because of the complex orography, which particularly benefits from the high resolution of the RCM and the high data availability, we performed our analysis for southwest Germany for the 30-yr period 1971–2000.

We investigated the sensitivity of the RVs at various return periods on several threshold values and the distribution functions GPD and KD. We found that 1) the number of independent precipitation events larger than 1 mm day^{-1} is at some grid points too low to allow a threshold below the 90th percentile (e.g., P83 and P85), 2) the deviation between the fitted return values and the plotting position of the ranked precipitation events is smaller for the KD than for the GPD, 3) with increasing threshold (i.e., decreasing sample size) the deviation increases, 4) the width of the confidence interval is smaller for the GPD than for the KD due to the lower number of free parameters at GPD, and 5) the width of the confidence interval increases with increasing threshold. This means that the KD yields RVs closer to the ranked data at the expense of a higher uncertainty. This especially holds for low threshold values up to the 90th percentile. We conclude that it is not necessary for the KD to reduce the sample to the very small amount of the extreme events as required for the GPD to yield high accuracy (in terms of RMSE). This advantage can be ascribed to the additional parameter of the KD. Thus, we decided that the KD with a 90th percentile threshold is most appropriate to derive extreme daily precipitation properties in orographically structured regions like southwest Germany.
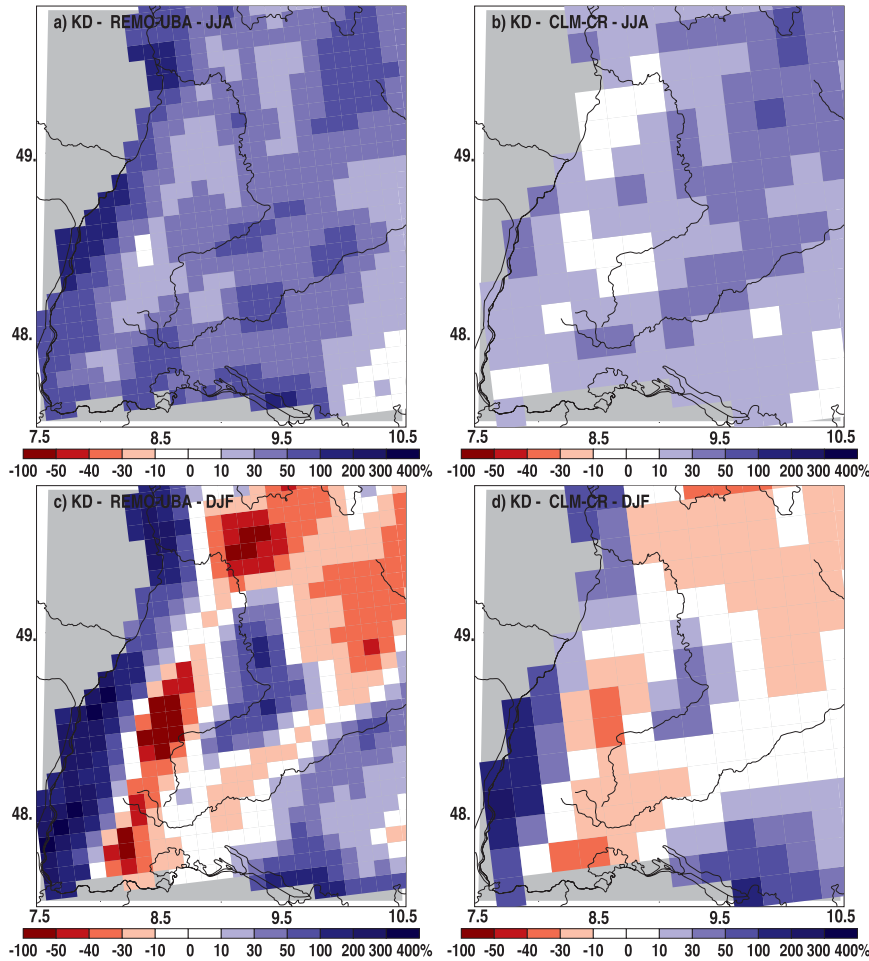
FIG. 12. Map of the percentage deviation of the simulated to the observed $\overline{\mathrm{RV}}_{T=10}$ of (a),(c) REMO–UBA and (b),(d) CLM-CR for (a),(b) JJA and (c),(d) DJF.

The second focus of this paper is the evaluation of simulated return values. We assessed the ability of regional climate models to realistically simulate the extreme daily precipitation. We compared the results of

present-day climate simulations for 1971–2000 resulting from two different RCMs with observations. The RCMs employed are the so-called consortium runs performed with the CLM at 18-km horizontal resolution and the
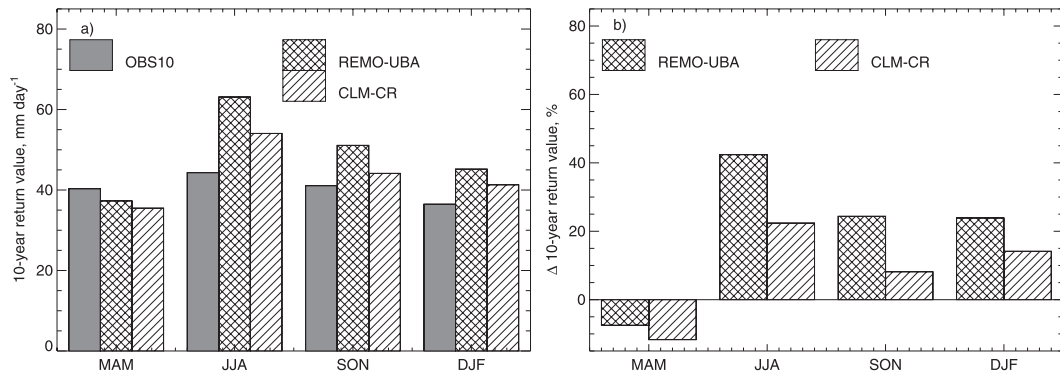

FIG. 13. (a) Domain-averaged seasonal $\overline{\mathrm{RV}}_{T=10}$ of OBS10 (gray), REMO–UBA (cross hatched), and CLM-CR (hatched) and (b) the percentage deviations of REMO–UBA to OBS10 and CLM-CR to OBS18.
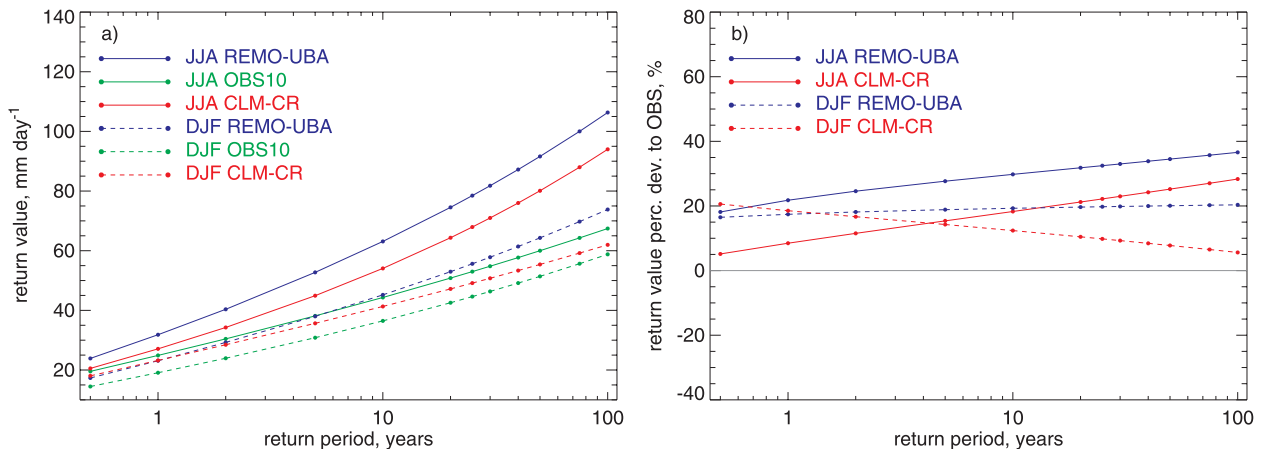
FIG. 14. (a) Domain-averaged $\overline{RV}$ of OBS10 (green), REMO–UBA (blue), and CLM-CR (red) as a function of the return period T for JJA (solid) and DJF (dashed). (b) The percentage deviations of the REMO–UBA (blue) and CLM-CR (red) to the observed $\overline{RV}$.

REMO–UBA simulations at 10-km horizontal resolution. We found that the 10-yr RVs are generally overestimated by both models. The highest deviation for REMO–UBA is located at the western slope of the Black Forest due to a misplacement of the orographic precipitation. This pattern exists also to a lesser extent in the more coarsely resolved CLM-CR simulations.

Averaged across the total region both model results agree with the observations and we find the annual maximum of the 10-yr return value $\overline{RV}_{T=10}$ in JJA and the minimum in MAM. The simulated annual cycle of $\overline{RV}_{T=10}$ has a distinctly larger amplitude than the observed one. Both models overestimate the seasonal $\overline{RV}_{T=10}$ in all seasons except MAM. The overestimation of REMO–UBA is more pronounced than that of CLM-CR. The underestimation in MAM is larger for CLM-CR than for REMO–UBA.

Furthermore, we found that all area-averaged return values $\overline{RV}$ are overestimated by both REMO–UBA and CLM-CR in JJA as well as in DJF. This is in part due to the general overestimation of the simulated precipitation. In JJA the deviation from the observed $\overline{RV}$ increases with increasing T, which means the agreement between the simulation results and the observations is higher for frequent events with low intensities than for rare events with high intensities. In DJF the increase of the deviation with increasing T is less than in JJA for REMO–UBA. Even a decreasing relation is found for CLM-CR.

The regional climate models reproduce the characteristics of the annual cycle and the spatial distribution well. However, they have some difficulties in producing realistic return values, especially for longer return periods. In contrast to coarser-resolved simulations (Frei et al. 2003) higher resolution leads to an overestimation of the return values. Thus, more sensitivity studies and

larger ensembles are necessary to reduce uncertainty of the model results and to obtain more reliable climate change information on the regional scale.

REFERENCES

Beniston, M., and Coauthors, 2007: Future extreme events in European climate: An exploration of regional climate model projections. *Climatic Change,* **81,** 71–95.

Böhm, U., M. Kücken, W. Ahrens, A. Block, D. Hauffe, K. Keuler, B. Rockel, and A. Will, 2006: CLM–The climate version of LM: Brief description and long-term applications. *COSMO Newsletter,* No. 6, COSMO, 225–235.

Brabson, B. B., and J. P. Palutikof, 2000: Tests of the generalized Pareto distribution for predicting extreme wind speeds. *J. Appl. Meteor.,* **39,** 1627–1640.

Bülow, K., 2010: Zeitreihenanalyse von regionalen Temperatur- und Niederschlagssimulationen in Deutschland (Time series analysis of simulated temperature and precipitation in Germany). Ph.D. thesis, Universität Hamburg, Berichte zur Erdsystemforschung 75, 148 pp.

Casas, M. C., M. Herrero, M. Ninyerola, X. Pons, R. Rodríguez, A. Rius, and A. Redaño, 2007: Analysis and objective mapping of extreme daily rainfall in Catalonia. *Int. J. Climatol.,* **27,** 399–409.

Doms, G., and U. Schättler, 2002: A description of the non-hydrostatic regional model LM. Part I: Dynamics and numerics. Deutscher Wetterdienst Tech. Rep., 134 pp. [Available online at www.cosmo-model.org.]

——, J. Förstner, E. Heise, H.-J. Herzog, M. Raschendorfer, R. Schrodin, T. Rheinhardt, and G. Vogel, 2007: A description of the non-hydrostatic regional model LM. Part II: Physical parameterisation. Deutscher Wetterdienst Tech. Rep., 139 pp. [Available online at www.cosmo-model.org.]

Feldmann, H., B. Früh, G. Schädler, H.-J. Panitz, K. Keuler, D. Jacob, and P. Lorenz, 2008: Evaluation of the precipitation for south-western Germany from high resolution simulations with regional climate models. *Meteor. Z.,* **17,** 455–465.

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

——, 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,** 4124, doi:10.1029/2002JD002287.

——, 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.

Früh, B., J. W. Schipper, A. Pfeiffer, and V. Wirth, 2006: A pragmatic approach for downscaling precipitation in alpine scale complex terrain. *Meteor. Z.,* **15,** 631–646.

——, J. Bendix, T. Nauss, M. Paulat, A. Pfeiffer, J. W. Schipper, B. Thies, and H. Wernli, 2007: Verification of precipitation from regional climate simulations and remote-sensing observations with respect to ground-based observations in the upper Danube catchment. *Meteor. Z.,* **16,** 275–293.

Giorgi, F., and L. Mearns, 1999: Introduction to special section: Regional climate modeling revisited. *J. Geophys. Res.,* **104** (D6), 6335–6352.

Good, P., and J. Lowe, 2006: Emergent behavior and uncertainty in multimodel climate projections of precipitation trends at small spatial scales. *J. Climate,* **19,** 5554–5569.

Gumbel, E. J., 1958: *Statistics of Extremes.* Columbia University Press, 375 pp.

Hagemann, S., K. Arpe, and E. Roeckner, 2006: Evaluation of the hydrological cycle in the ECHAM5 model. *J. Climate,* **19,** 3810–3827.

Hollweg, H.-D., and Coauthors, 2008: Ensemble simulations over Europe with the regional climate model CLM forced with IPCC AR4 global scenarios. Model and Data Group Tech. Rep. 3, Max Planck Institute for Meteorology, Hamburg, Germany, 150 pp.

Hosking, J. R. M., 1990: L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. Roy. Stat. Soc.,* **52B,** 105–124.

——, 1994: The four parameter kappa distribution. *IBM J. Res. Develop.,* **38,** 251–258.

——, and J. R. Wallis, 1987: Parameter and quantiles estimation for the generalized Pareto distribution. *Technometrics,* **29,** 339–349.

——, and ——, 1997: *Regional Frequency Analysis.* Cambridge University Press, 224 pp.

Jacob, D., 2001: A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin. *Meteor. Atmos. Phys.,* **77,** 61–73.

——, cited 2008: REMO climate of the 20th century run, UBA project, 0.088 degree resolution run no. 006210, 1h data. CERA-DB 'REMO_UBA_C20_1_R006210_1H'. World Data Center for Climate. [Available online at http://cera-www.dkrz.de/WDCC/ui/Compact.jsp.]

——, and Coauthors, 2007: An intercomparison of regional climate models for Europe: Design of the experiments and model performance. *Climatic Change,* **81,** 31–52, doi:10.1007/s10584-006-9213-4.

——, H. Göttel, S. Kotlarski, P. Lorenz, and K. Sieck, 2008: Klimaauswirkungen und Anpassung in Deutschland—Phase 1: Erstellung regionaler Klimaszenarien für Deutschland. Abschlussbericht zum UFOPLAN-Vorhaben Tech. Rep. 204 41 138, 154 pp.

Kendon, E. J., D. P. Rowell, R. G. Jones, and E. Buonomo, 2008: Robustness of future changes in local precipitation extremes. *J. Climate,* **21,** 4280–4297.

Kharin, V. V., and F. W. Zwiers, 2000: Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere–ocean GCM. *J. Climate,* **13,** 3760–3788.

Kunkel, K. E., K. Andsager, X.-Z. Liang, R. W. Arritt, E. S. Takle, W. J. J. Gutowski, and Z. Pan, 2002: Observations and regional climate model simulations of heavy precipitation events and seasonal anomalies: A comparison. *J. Hydrol.,* **3,** 322–334.

Lautenschlager, M., K. Keuler, C. Wunram, E. Keup-Thiel, M. Schubert, A. Will, B. Rockel, and U. Böhm, cited 2008: Climate simulation with CLM, climate of the 20th century run no. 1, data stream 2: European region MPI-M/MaD. World Data Center for Climate. [Available online at http://cera-www.dkrz.de/WDCC/ui/Compact.jsp.]

Li, Y., W. Cai, and E. P. Campbell, 2005: Statistical modeling of extreme rainfall in southwest western Australia. *J. Climate,* **18,** 852–863.

Makkonen, L., 2007: Problems in the extreme value analysis. *Struct. Saf.,* **30,** 405–419.

Maraun, D., T. J. Osborn, and N. P. Gillett, 2008: United Kingdom daily precipitation intensity: Improved early data, error estimates and an update from 2000 to 2006. *Int. J. Climatol.,* **28,** 833–842.

Müller-Westermeier, G., 1995: Numerisches Verfahren zu Erstellung klimatologischer Karten. Deutschen Wetterdienstes Reps., Vol. 193, Deutscher Wetterdienst, Offenbach am Main, Germany, 17 pp.

NERC, 1975: Flood Studies Report. Meteorological Studies, Vol. II, National Environment Research Council, London, United Kingdom, 81 pp.

Paeth, H., and A. Hense, 2005: Mean versus extreme climate in the Mediterranean region and its sensitivity to future global warming conditions. *Meteor. Z.,* **14,** 329–347.

Palutikof, J. P., B. B. Brabson, D. H. Lister, and S. T. Adcock, 1999: A review of methods to calculate extreme wind speeds. *Meteor. Appl.,* **6,** 119–132.

Parida, B. P., 1999: Modelling of Indian summer monsoon rainfall using a four-parameter kappa distribution. *Int. J. Climatol.,* **19,** 1389–1398.

Park, J.-S., and H.-S. Jung, 2002: Modelling Korean extreme rainfall using a kappa distribution and maximum likelihood estimate. *Theor. Appl. Climatol.,* **72,** 55–64.

Roeckner, E., and Coauthors, 2006a: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model. *J. Climate,* **19,** 3771–3791.

——, M. Lautenschlager, and H. Schneider, cited 2006b: IPCC-AR4 MPI-ECHAM5_T63L31 MPI-OM_GR1.5L40 20C3M

run no. 1: Atmosphere monthly mean values MPImet/MaD Germany. World Data Center for Climate, doi:10.1594/WDCC/EH5-T63L31_OM-GR1.5L40_20C_1_MM.

Schwarb, M., 2001: The alpine precipitation climate evaluation of a high-resolution analysis scheme using comprehensive rain-gauge data. Ph.D. dissertation ETH 13911, Swiss Federal Institute of Technology, Zürich, Switzerland, 130 pp.

——, C. Daly, C. Frei, and C. Schär, 2001: Mean annual and seasonal precipitation in the European Alps, 1971–1990. *Hydrological Atlas of Switzerland,* Federal Office for Water and Geology, Bern, Switzerland, Plates 2.6 and 2.7.

Semmler, T., and D. Jacob, 2004: Modeling extreme precipitation events—A climate change simulation for Europe. *Global Planet. Change,* **44,** 119–127.

Solomon, S., D. Qin, M. Manning, M. Marquis, K. Averyt, M. M. B. Tignor, H. L. Miller Jr., and Z. Chen, Eds., 2007: *Climate Change 2007: The Physical Science Basis.* Cambridge University Press, 996 pp.

van den Brink, H. W., G. P. Können, J. D. Opsteegh, G. J. van Oldenborgh, and G. Burgers, 2005: Estimating the return periods of extreme events from ECMWF seasonal forecast ensembles. *Int. J. Climatol.,* **25,** 1345–1354.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* 2nd ed. International Geophysics Series, Vol. 91, Academic Press, 627 pp.

Zwiers, F. W., and V. V. Kharin, 1998: Changes in the extremes of the climate simulated by CCC GCM2 under CO_2 doubling. *J. Climate,* **11,** 2200–2222.