Natural Hazards
and Earth System
Sciences
Discussions

# Using high-resolution regional climate models to estimate return levels of daily extreme precipitation over Bavaria

Benjamin Poschlod[1]

[1]Department of Geography, Ludwig-Maximilians-Universität München, Munich, 80333, Germany

5    *Correspondence to*: Benjamin Poschlod (Benjamin.Poschlod@lmu.de)

**Abstract.** Extreme daily rainfall is an important trigger for floods in Bavaria. The dimensioning of water management structures as well as building codes are based on observational rainfall return levels. In this study, three high-resolution regional climate models (RCMs) are employed to produce 10-year daily rainfall return levels and their performance is evaluated by comparison to observational return levels. The study area is governed by different types of precipitation (stratiform, orographic,

10    convectional) and a complex terrain, with convective precipitation also contributing to daily rainfall levels. The Canadian Regional Climate Model version 5 (CRCM5) at 12 km spatial resolution and the Weather and Forecasting Research model (WRF) at 5 km resolution both driven by ERA-Interim reanalysis data use parametrization schemes to simulate convection. The WRF at 1.5 km resolution driven by ERA5 reanalysis data explicitly resolves convectional processes. Applying the Generalized Extreme Value (GEV) distribution, all three model setups can reproduce the observational return levels with an

15    areal average bias of +6.6 % or less and a spatial Spearman rank correlation of ρ > 0.72. The increase of spatial resolution between the 12 km CRCM5 and the 5 km WRF setup is found to improve the performance in terms of bias (+6.6 % and +3.2 %) and spatial correlation (ρ = 0.72 and ρ = 0.82). However, the finer topographic details of the WRF-ERA5 return levels cannot be evaluated with the observation data because their spatial resolution is too low. Hence, this comparison shows no great further improvement (bias = +1.1 %, ρ = 0.82) of the overall performance compared to the 5 km resolution setup.

20    Uncertainties due to extreme value theory are explored by employing three different approaches for the highest-resolution WRF-ERA5 setup. The GEV distribution with fixed shape parameter (bias = +0.9 %, ρ = 0.79) and the Generalized Pareto (GP: bias = +1.3 %, ρ = 0.81) show almost equivalent results for the 10-year return period, whereas the Metastatistical Extreme Value (MEV) distribution leads to a slight underestimation (bias = -6.2 %, ρ = 0.86).

From these results, it follows that high-resolution regional climate models are suitable for generating spatially homogeneous

25    rainfall return level products. In regions with a sparse rain gauge density or low spatial representativeness of the stations due to complex topography, RCMs can support the observational data. Further, RCMs driven by global climate models with emission scenarios can project climate change-induced alterations in rainfall return levels at regional to local scales. This would allow adjustment of structural design and, therefore, adaption to future precipitation conditions.

## 1 Introduction

30  Extreme rainfall is an important driver for different kinds of hydrometeorological hazards, such as flooding and mass movements. The state of Bavaria is exposed to the highest daily rainfall intensities in Germany. Due to the complex topography and a dense river network the area is prone to riverine flooding and landslides (Grieser et al., 2006; Wiedenmann et al., 2016). Furthermore, urban areas are at risk of urban flooding due to the dense population and a large fraction of impervious areas (Chen and Leandro, 2019). To assess the risk of heavy precipitation events and to dimension adaptation measures, engineers

35  and public authorities often use the concept of rainfall return levels. In Germany, a rainfall return level database ("Coordinated heavy precipitation regionalization evaluation"; KOSTRA; Junghänel et al., 2017; Malitz and Ertel, 2015) is supplied by the German weather service, which is based on rain gauge observations. A similar product is available for Austria (Kainz et al., 2007). MeteoSwiss provides mapped return levels and pointwise data (MeteoSwiss, 2021). These products are included in building standards and are, therefore, widely used. Even though the coverage of rain gauges in Germany, Austria, and

40  Switzerland is relatively high, there are uncertainties due to the spatial representativeness of the measuring stations to generate an area-wide rainfall return level product. This problem applies even more on a continental scale as the rain gauge density is distributed heterogeneously over different European countries, where the available time series might be too short to capture a sufficient number of extreme events (Lewis et al., 2019).

Instead of using point-wise measurements, areal precipitation products (e.g. radar, satellite, or reanalysis products) could be

45  used as the basis for return level calculations. However, each of these areal precipitation products shows different limitations, which lead to uncertain or unrealistic return level estimations. Radar data (RADOLAN for Germany; Kreklow et al., 2020) and satellite products (e.g. CMORPH; Joyce et al., 2004 or PERSIANN; Hong et al., 2004) would provide the necessary temporal and spatial resolutions to capture extreme rainfall events. Yet, the temporal coverage of these products extends only to the early 2000s, which is why the sampling of extreme rainfall events is not sufficient for extreme value analysis.

50  Furthermore, radar estimates (Goudenhoofdt and Delobbe, 2016; Kreklow et al., 2020) as well as satellite products (Stampoulis and Anagnostou, 2012) reveal biases compared to rain gauges. Reanalysis data (e.g. E-OBS; Haylock et al., 2008; ERA-Interim; Dee et al., 2011; ERA5; Hersbach et al., 2020) would have the necessary temporal coverage, but they show systematic underestimation of the intensity of extreme precipitation events (Hu and Franzke, 2020; own calculations, not shown).

Since the frequency and intensity of heavy precipitation will change due to climate change (Myhre et al., 2019; Westra et al.,

55  2014), the use of climate models would provide the advantage of being able to estimate climate change-induced alterations in rainfall return levels on a physical basis. However, this application requires careful validation of climate model results for historical conditions.

Regional climate models (RCMs) at 12 km spatial resolution have proven to deliver appropriate rainfall return level estimations for 3-hourly to daily duration (Berg et al., 2019; Poschlod et al., 2021; Ulbrich and Nissen, 2017). Although the results show

60  a high spatial correlation to observational products and a low bias averaged over the area, local deviations are evident,

especially in regions with complex topography (Poschlod et al., 2021). Also, the intensity of short-duration hourly rainfall extremes could not be reproduced at 12 km spatial resolution.

When communicating the results of climate model projections to local or regional stakeholders, insurance companies, and governmental authorities in the field of flood prevention, hydrological modelling, dimensioning of reservoirs, buildings, and

65   water infrastructure, these aforementioned biases may prevent the results from being accepted and implemented (Benjamin and Budescu, 2018).

For shorter durations, many studies have shown that higher-resolution RCMs, so-called convection-permitting models (CPMs), improve the reproduction of high-intensity short-duration convectional precipitation events (Brisson et al., 2016; Coppola et al., 2018; Fosser et al., 2014; Kendon et al., 2014). A spatial resolution of a few kilometres is considered necessary by the

70   RCM community to explicitly resolve convection (Langhans et al., 2012, Panosetti et al., 2020; Prein et al., 2015), whereas at broader-resolutions parametrization schemes are applied to represent convection. However, also long-duration rainfall return levels can be influenced by convectional precipitation. In Germany, convectional rainfall contributes to the 24-hourly return level for roughly 50 % of the area (Malitz and Ertel, 2015). Therefore, CPMs are expected to improve the estimations of these return levels as well. Additionally, the higher spatial resolution enhances the representation of complex terrain (Karki et al.,

75   2017; Langhans et al., 2012; Poschlod et al., 2018).

Hence, in this study, three different high-resolution RCMs featuring 12 km, 5 km, and 1.5 km spatial resolution and driven by 30-year reanalysis data are applied to reproduce daily 10-year rainfall return levels over the complex terrain of the northern Pre-Alps and Alps. Based on interviews with stakeholders from the infrastructure sector and on legislative guidelines, Nissen and Ulbrich (2017) identified the 10-year return level as relevant threshold. Following this recommendation but also to avoid

80   large extreme value statistical uncertainties based on the 30-year time series the 10-year return level is chosen in this study as well representing "moderate extremes". The daily duration is relevant for the generation of riverine floods in the study area (Berghuijs et al., 2019; Keller et al., 2017; Merz and Blöschl, 2003), such as the two extreme flooding events in May 1999 and August 2005 in southern Bavaria, Austria, and Switzerland (BLFW, 2003; Grieser et al., 2006; LfU, 2007; Stucki et al., 2020) induced by high daily precipitation sums. The daily 10-year return levels based on the three RCM setups are evaluated by

85   means of an observational return level product using national datasets from Germany, Austria, and Switzerland. In a second step, different extreme value distributions and sampling strategies are applied to the highest-resolution climate model dataset to explore uncertainties due to extreme value theory and to investigate possible improvements.

It is studied if RCMs can bridge the data gap of spatial homogeneous rainfall return levels and if higher spatial resolution can decrease the biases over areas with complex topography. Therefore, the study aims to evaluate the added value due to higher

90   spatial resolution in terms of biases and spatial correlation between the climate model products and the observational product.

## 2 Study area and data

### 2.1 Description of the study area

The area of investigation is given by the domain of the highest-resolution RCM, which is centred over the state of Bavaria, and the available observational rainfall return level data (see Fig. 1). It covers south-eastern Germany, north-western Austria, north-eastern Switzerland and Liechtenstein. The area shows altitude levels below 100 m in the northwest in the Rhine plain up to altitudes above 2500 m in the Alps. It covers various low mountain ranges, including the Ore Mountains, Odenwald, Swabian Jura and Bavarian Forest. The patterns of annual mean precipitation are governed by the complex topography (see Fig. 2). Different rainfall types (convectional, orographic, stratiform) contribute to this precipitation climatology (Malitz and Ertel, 2015). The lowest annual precipitation sums amount to 500 – 700 mm in the north of the study area. The low mountain ranges induce orographic lifting leading to precipitation sums of 1000 to 1500 mm per year. The highest precipitation sums of more than 2000 mm are found in the Alps, with dry valleys, such as the Inn valley having totals below 1000 mm. Annual average temperatures range from less than 0°C in the Alps to 10°C in northern Bavaria (DWD 2021, ZAMG 2021).
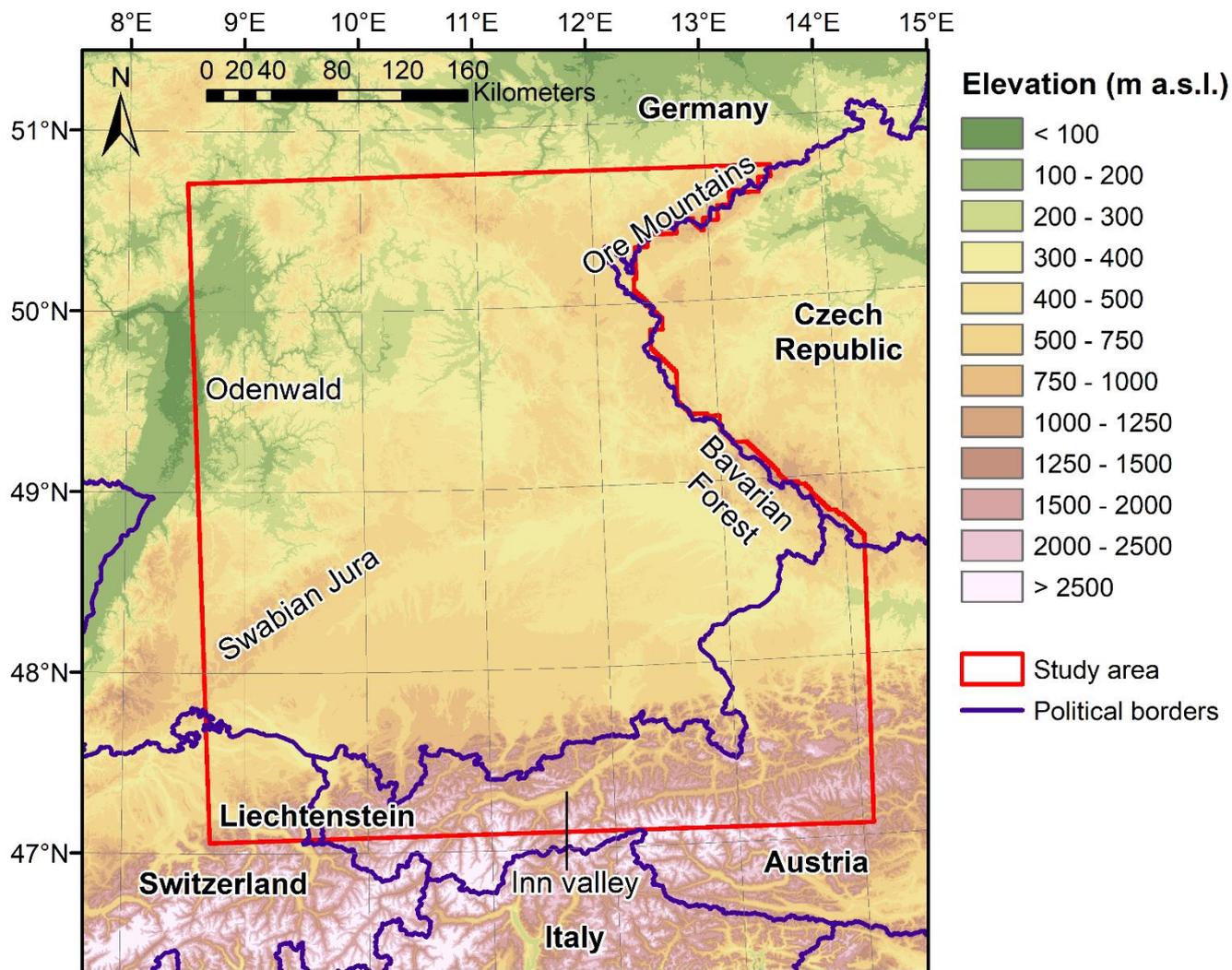
**Figure 1: Topography of the investigated area. The elevation is based on the SRTM at 90 m resolution (Jarvis et al., 2008).**
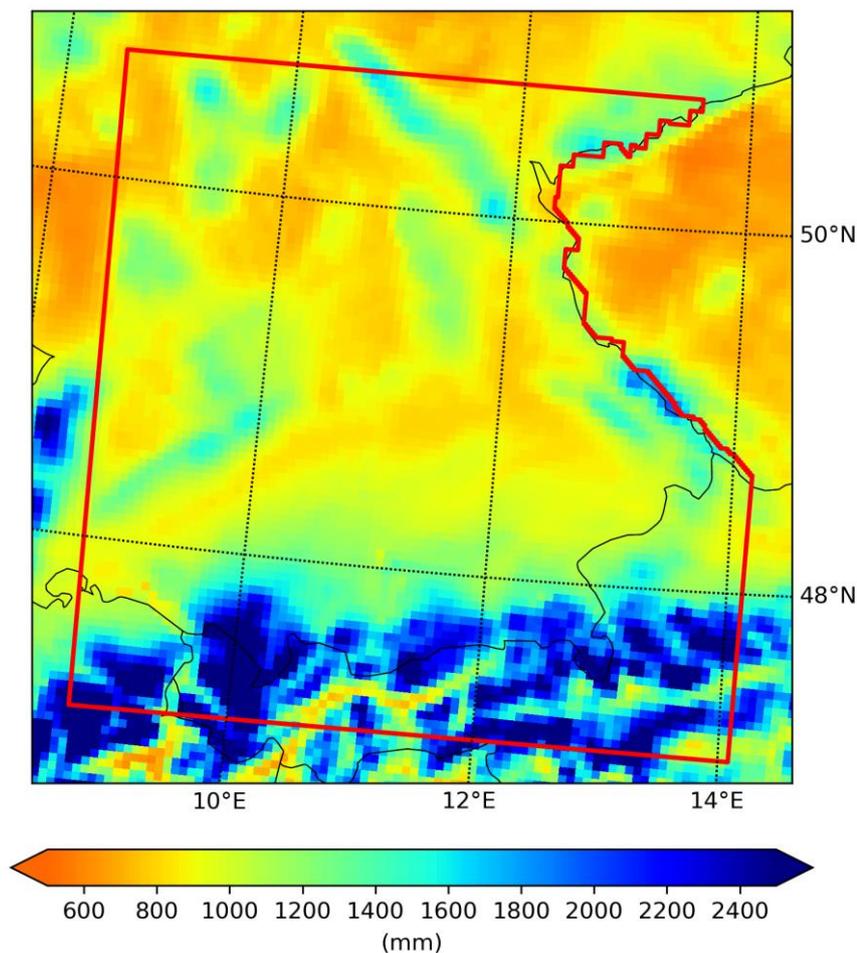
**Figure 2: Annual mean precipitation for the period 1980 – 2009 from the RCM setup of Warscher et al. (2019).**

## 2.2 Observational rainfall return level data

110   To evaluate the RCMs, an observation-based product is generated from the three national datasets described below. As these datasets extend to the national borders and a little beyond, the arithmetic mean is calculated in the overlapping areas. To compare gridded precipitation from the RCMs and point measurements from the observations, Breinl et al. (2020) suggest an areal reduction of 5 % for pointwise 24-hourly 10-year return levels in Austria. However, to be consistent over the study area, no areal reduction factor is applied for the daily duration following Berg et al. (2019) and Poschlod et al. (2021).

### 2.2.1 Germany

The German weather service offers gridded return level data derived from rain gauge measurements (Malitz and Ertel, 2015). The observations cover a period of maximum 1951 – 2010, where only May – September are analysed as the highest rainfall amounts occur during these months. A peak-over threshold (POT) sampling strategy was applied for 2231 rain gauges, where the threshold corresponds to the available time period. A maximum of 2.718 events per year on average was considered. For these samples, an exponential distribution was fitted. The resulting rainfall return levels were spatially interpolated over Germany at roughly 8 x 8 km² resolution. An uncertainty range of 15 % is assumed for the 10-year return levels, which is induced by measurement errors, uncertainties of the extreme value statistics and regionalization, and the internal variability of the climate system (Junghänel et al., 2017). Data are accessed from DWD (2020). As running window 24-hourly return periods are provided, the rainfall intensities are reduced by 14 % to transfer them to daily estimates. This relation between daily fixed windows and 24-hourly moving windows has been applied by Poschlod et al. (2021) following Barbero et al. (2019) and Boughton and Jakob (2008).

### 2.2.2 Austria

The Austrian dataset follows a similar approach as the German dataset also applying POT sampling at 141 ombrographs and 843 ombrometers spatially interpolated to gridded return levels at 6 x 6 km² resolution (BMLRT, 2018). As the rain gauges are distributed inhomogeneously yielding too low return level estimations, the "orographic convective model" OKM (Lorenz and Skoda, 2001) was employed to support the observations (Kainz et al., 2007). The resulting design rainfall is based on a combination of the observational data and the weather model simulations. Further details can be found in Kainz et al. (2007) and BMLRT (2006; 2018). Data are accessed from BMLRT (2020). The 24-hourly return levels are adjusted to daily values using the reduction of 14 %.

### 2.2.3 Switzerland

MeteoSwiss (2021) provides pointwise daily rainfall return levels at 336 rain gauges. The observations cover the time period from 1966 to 2015. To increase the sample size, seasonal maxima were extracted and assumed to follow a Generalized Extreme Value (GEV) distribution. The GEV distribution is fitted via Bayesian estimation and the according return levels are generated (Fukutome et al., 2015). Since an areal comparison product is to be produced in this study, these point return levels are regionalised by means of ordinary kriging.

### 2.3 Climate model data

Three different RCM setups are used. The Canadian Regional Climate Model version 5 (CRCM5) driven by ERA-Interim, the Weather and Research Forecasting Model (WRF; Skamarock et al., 2008) driven by ERA-Interim, and the WRF driven by ERA-5.

### 2.3.1 CRCM5 ERA-INTERIM

The CRCM5 at 0.11° resolution equalling roughly 12 km is driven by ERA-Interim reanalysis data. Convectional processes are parametrized due to the spatial resolution. Processes related to deep convection are calculated with the parametrization scheme by Kain and Fritsch (1990). The Kuo transient scheme (Bélair et al., 2005; Kuo, 1965) is applied to represent shallow convection. A more detailed documentation of the model setup and options used is given by Hernández-Díaz et al. (2012) and Martynov et al. (2012). Daily rainfall sums of 30-year time period of 1980 – 2009 are extracted for this study.

### 2.3.2 WRF ERA-INTERIM

The WRF version 3.6.1 is set up in nested domains of 45 x 45 km², 15 x 15 km² and 5 x 5 km² spatial resolution in its non-hydrostatic mode and driven by ERA-Interim reanalysis data at 75 x 75 km² spatial resolution and 6-hourly temporal resolution (Warscher et al., 2019). Spectral nudging is applied to reduce deviations from the large-scale forcing patterns in the reanalysis data (Wagner et al., 2018). Convection is parametrized with the Grell-Freitas scheme (Grell and Freitas, 2014). The detailed model setup as well as an evaluation of different climate variables is given in Warscher et al. (2019). Here, daily rainfall data of the highest-resolution domain are used for the time period of 1980 – 2009. Data are accessed from Warscher (2019).

### 2.3.3 WRF ERA5

The WRF model version 4.1 is configured with two one-way nested domains of 7.5 x 7.5 km² and 1.5 x 1.5 km² grid spacing centred over Bavaria (Collier and Mölg, 2020). The model is forced at the outer lateral boundaries by ERA5 reanalysis data at 30 x 30 km² spatial resolution and 3-hourly temporal resolution applying spectral nudging. The higher-resolution 1.5 km setup is assumed to explicitly resolve convection, and therefore no parametrization scheme is applied. The 30-year simulation was divided into 30 annual slices starting at 1 September of each year. A detailed description of the model setup and evaluation of various climate variables is provided in Collier and Mölg (2020). However, the authors emphasize that the applied schemes and the model configuration has not been optimized for the study area due to the high computational expenses of the high-resolution run. The physics and dynamics options used in the simulations are based on former convection-permitting WRF applications (e.g. Collier et al., 2019). In this study, daily rainfall sums from 1988 – 2017 are extracted from the climate model data accessed from Collier (2020).

## 3 Extreme value approach

### 3.1 Sampling strategies

Extreme value theory (EVT) is applied to quantify the stochastic behaviour of a process at unusually large or small values. It is commonly used to calculate return levels for different rainfall durations. There, two classical approaches exist to sample these unusual ("extreme") rainfall intensities (Coles, 2001). For the block maxima (BM) approach, a single value is extracted

from a typically seasonal or annual block. This strategy ensures that the samples are distant from each other leading to very
175    low serial dependence. However, not all sampled values might be extreme. Also, the information of more than one extreme
value per block is lost as these values are discarded.

The second approach peak-over threshold (POT) tries to overcome these drawbacks as all values $s$ above a threshold $u$ are
sampled as extreme values (Balkema and de Haan, 1974; Picklands, 1975). Therefore, multiple values per block are allowed.
However, additional restrictions have to be introduced to ensure approximately independent samples. To prevent successive
180    data points from being sampled that originate from one persistent rainfall event, the time series has to be de-clustered.
Therefore, a temporal threshold $t_{decluster}$ is chosen and all values within the duration of $t_{decluster}$ around the sampled extreme
value are discarded (Coles, 2001).

For both classical approaches only a limited number of samples contributes to the database of extreme values. A newer
approach by Marani and Ignaccolo (2015) samples all "wet" events assuming that the information of these "ordinary" values
185    can be used to estimate the distribution of extreme values. Thereby, wet events are defined by a threshold $t_{wet}$. It has been
successfully applied for extreme daily precipitation by Zorzetto et al. (2016).

## 3.2 Extreme value distributions

According to the sampling strategy, different theoretical extreme value distributions (EVDs) are typically found to represent
190    the distributions of the samples.

The Fisher-Tippett-Gnedenko theorem (Fisher and Tippett, 1928; Gnedenko, 1943) states that the distribution of the block
maxima    samples    tends    to    follow    the    GEV    distribution    $G$    with    the    sample    size    $n \rightarrow \infty$:

$$G(x;\ \xi) = \begin{cases} \exp(-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}), \xi \neq 0 \\ \exp(-\exp(-\frac{x-\mu}{\sigma})), \xi = 0 \end{cases} \tag{1}$$

195

The location parameter $\mu$ governs the center, and the scale parameter $\sigma$ governs the spread of the GEV distribution. The tail
behaviour of $G$ is defined by the shape parameter $\xi$ determining whether the GEV follows the Weibull ($\xi < 0$), Gumbel ($\xi =$
0), or Fréchet ($\xi > 0$) distribution (Gilleland et al., 2017).  Hence, the GEV is a very flexible distribution. The drawback of this
flexibility shows up in a high estimation variance of $\xi$ resulting in an unstable quantile estimate (Bücher et al., 2020).
200    For the POT approach, the exceedances $y = s - u$ are sampled for the threshold $u$ and samples $s > u$. Thereby, the number of
exceedances per year is assumed to follow a Poisson distribution (Davison & Smith, 1990). The exceedances $y$ of the POT
threshold $u$ are described by the two-parameter Generalized Pareto (GP) distribution (Davison and Smith, 1990, Martins and
Stedinger, 2001). The corresponding cumulative density function (CDF) is given by

205 $$H(y;\ \xi) = \begin{cases} 1 - (1 + \frac{\xi y}{\beta})^{-1/\xi}, \xi \neq 0, \beta > 0, y > 0 \\ 1 - \exp\left(\frac{-y}{\beta}\right), \xi = 0, \beta > 0, y > 0 \end{cases}, \qquad (2)$$

where $y$ defines the precipitation excess over the threshold $u$ of the POT sampling. The scale parameter β and shape parameter ξ describe the spread and tail behaviour of the GP distribution (Coles, 2001).

Both statistical frameworks can be expressed by the other one as the GP distribution corresponds to the tail distribution of the GEV (Coles, 2001; Goda, 2011; Serinaldi and Kilsby, 2014).

210 The newer approach by Marani and Ignaccolo (2015) features the Metastatistical Extreme Value (MEV) distribution. They propose a framework supposing that the "meta-statistic" of the rainfall sums of wet events per year contains information about the intensity of extreme events. They assume the sampled wet days > $t_{wet}$ to be independent following that the probability distribution of maxima $\zeta_m$ can be expressed as $\zeta_m(x) = F(x)^{n_j}$, where $n_j$ is the number of wet events in a year and $F(x)$ is a distribution describing the rainfall sums of these events. Based on the results of Wilson and Toumi (2005), the distribution of

215 rainfall sums during wet days per year is found to follow a distribution with an exponential tail. They expressed precipitation as the product of mass flux, specific humidity and precipitation efficiency. Following statistical relationships, they concluded that the tail of the distribution of the product of these three random variables is given by a stretched exponential form. Marani and Ignaccolo (2015) and Zorzetto et al. (2016) apply a Weibull distribution to describe this relationship. Hence, Weibull parameters have to be estimated for each year based on all wet events of a year. The MEV-Weibull CDF is given by

220 $$\zeta_m(x) = \frac{1}{M} \sum_{j=1}^{M} \left\{ 1 - \exp\left[ -\left(\frac{x}{C_j}\right)^{w_j} \right] \right\}^{n_j}, C_j > 0, w_j > 0, \qquad (3)$$

where j is the year (j = 1, 2, …, M), and $n_j$ is the number of wet events in year j. $C_j$ and $w_j$ describe the scale and shape of the Weibull distribution (Marani and Ignaccolo, 2015).

## 3.3 Applied approaches

For all three RCM setups, annual maxima of daily precipitation are extracted. Then for all grid cells trends were detected applying the Mann-Kendall test at the significance level of $\alpha = 0.05$. There, the critical p-value is adjusted for multiple testing following Wilks (2016). No significant trends are found for the 30 sampled values at each grid cell of every RCM setup. The parameters of the GEV distribution $G$ are optimized to the BM samples by estimating the L-moments (Hosking et al., 1985).

230 This is carried out applying the R package "extRemes" by Gilleland and Katz (2016). Delicado and Goria (2008) recommend the method of L-moments for sample sizes of $n \leq 50$ as it is robust to outliers in the data. The Anderson-Darling test at the significance level of $\alpha = 0.05$ is applied to ensure the goodness of fit of the estimated GEV distribution at each grid cell. Again, the critical $p$-value is adjusted for multiple testing. Based on these fits, the 10-year return levels are derived. The spatial

distributions      of      the      GEV      parameters      are      mapped      in      Figure      3.
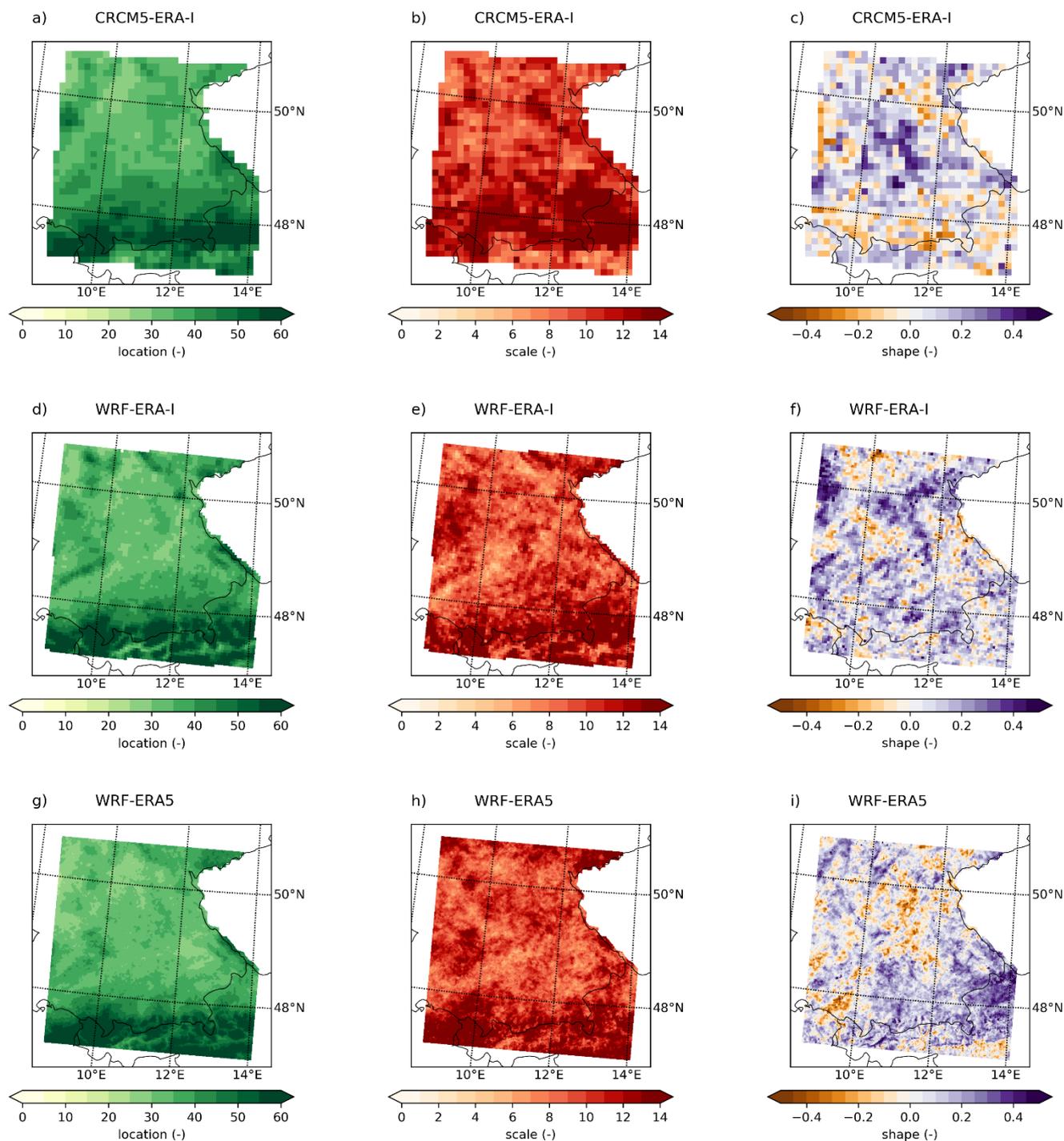


Figure 3: Location, scale, and shape parameters of the GEV-LMOM approach based on the CRCM5-ERA-Interim (a-c), WRF-ERA-Interim (d-f), and WRF-ERA5 (g-i).

235

11

There, the location and scale parameters are governed by the topography, where the spatial distribution of these parameters is

240    similar for all three RCM setups. The fitted shape parameter reveals a chaotic pattern with small patches of positive and negative values differing for the three RCM setups.

The histograms of the parameters are given in the Supplementary Materials (Fig. S1). An exemplary fit for the grid cell of Munich is shown in Figure S2 for all three RCM setups. This EVT approach is referred to as GEV-LMOM.

To assess uncertainties due to the different EVT approaches described in sections 3.1 and 3.2, a modified GEV approach as

245    well as the POT and MEV approaches are explored for the WRF-ERA5 data. As small samples lead to high uncertainties estimating the shape parameter of the GEV distribution, Papalexiou and Koutsoyiannis (2013) recommend using a fixed value of $\xi = 0.114$. This approach is referred to as GEV-FIX.

For the POT approach, the daily rainfall time series is de-clustered applying a conservative threshold $t_{decluster}$ of 5 days. Typical continental cyclones are found to last up to 2.25 days in Bavaria, whereas van Bebber type Vb cyclones can last up to 3 days

250    (Hofstätter et al., 2018; Mittermeier et al., 2019). Hence, the threshold of 5 days ensures independent samples. Precipitation intensities are assumed to be extreme when exceeding the threshold given by 90 events per 30-year period. This threshold has also been selected by Berg et al. (2019). It amounts to 23.4 mm as spatial average for the whole study area, where the lowest threshold is 12.7 mm. Trends are excluded in the same way as for the GEV-LMOM approach. For sample sizes of n > 50, Delicado and Goria (2008) and Madsen et al. (1997) recommend Maximum Likelihood Estimation (MLE) as optimization

255    algorithm to fit an extreme value distribution. Following this recommendation, MLE is applied to fit the GP distribution to the 90 samples per grid cell using the software package by Gilleland and Katz (2016). The goodness-of-fit is assessed in the same way as for the GEV-LMOM approach. An exemplary fit for the grid cell of Munich is shown in Figure S3. This approach is referred to as GP-MLE.

For the MEV approach, I follow the procedure applied by Zorzetto et al. (2016). The Weibull distribution is fitted to the annual

260    wet events by means of the probability weighted moments method (PWM, Greenwood et al., 1979). Wet days are defined by exceedance of the threshold $t_{wet} = 1$ mm d$^{-1}$ in accordance to WMO guidelines (Klein-Tank et al., 2009). This also accounts for the behaviour of RCMs to produce too many very low-intensity precipitation days ("drizzle-effect"; Gutowski et al., 2003). The MEV fitting procedure and the calculation of return levels is carried out using the Python software package mevpy (Zorzetto, 2021). This approach is referred to as MEV-PWM.

265    **4 Evaluation of 10-year return levels**

**4.1 Results**

All approaches and their performance metrics are summarized in Table 1. A mapped comparison of the 10-year return levels calculated via GEV-LMOM based on the three different RCM setups is given in Figure 4. For a better visualization, the observational product is bilinearly interpolated to the respective RCM grid. The following metrics are calculated for the

270  original data (see Fig. S4 for the natively resolved observational product). The observational product shows the rainfall highest intensities above 100 mm d$^{-1}$ at the northern slopes of the Alps. The low mountain ranges of the Bavarian Forest, Swabian Jura, Odenwald and Ore Mountains also induce enhanced intensities between 70 mm d$^{-1}$ and 100 mm d$^{-1}$. The lowest return levels are observed in the north of the study area amounting to intensities below 50 mm d$^{-1}$ (Fig. 4b, e, h). The 12-km resolution CRCM5-ERA-I can reproduce the general spatial pattern with a Spearman rank correlation of $\rho$ = 0.72 (Fig. 4a). The return

275  levels are generally overestimated north of 48° N and underestimated south of 48° N as well as in the Ore Mountains (Fig. 4c). The spatially averaged bias amounts to +6.6 %. The range of simulated rainfall return level intensities is similar to the observations for the whole study area (Fig. 5a) as well as for the southern alpine part (Fig. 5b). However, the histogram also reveals that the bias stems from simulating too few grid cells with return level intensities between 50 mm d$^{-1}$ to 60 mm d$^{-1}$ and too many grid cells with return levels at intensities of 70 mm d$^{-1}$ to 90 mm d$^{-1}$ (Fig. 5a).

280

**Table 1: Summary of the applied RCM setups and EVT approaches. Performance metrics of the comparison to observational return levels are given in terms of spatially averaged bias and spatial correlation (Spearman).**

| RCM | Reanalysis data | Spatial resolution | Convection | Sampling | EVD | EVD parameter optimization | Bias | Spatial correlation |
|---|---|---|---|---|---|---|---|---|
| CRCM5 | ERA-Interim | 12 km | Parametrized | Block maxima | GEV | L-Moments | +6.6 % | 0.72 |
| WRF | ERA-Interim | 5 km | Parametrized | Block maxima | GEV | L-Moments | +3.2 % | 0.82 |
| WRF | ERA5 | 1.5 km | Explicitely calculated | Block maxima | GEV | L-Moments | +1.1 % | 0.82 |
| WRF | ERA5 | 1.5 km | Explicitely calculated | Block maxima | GEV | Maximum likelihood estimation, fixed shape parameter | + 0.8 % | 0.79 |
| WRF | ERA5 | 1.5 km | Explicitely calculated | Peak-over-threshold | GP | Maximum likelihood estimation | +1.3 % | 0.81 |
| WRF | ERA5 | 1.5 km | Explicitely calculated | All wet events | MEV | Probability weighted moments | - 6.2 % | 0.86 |
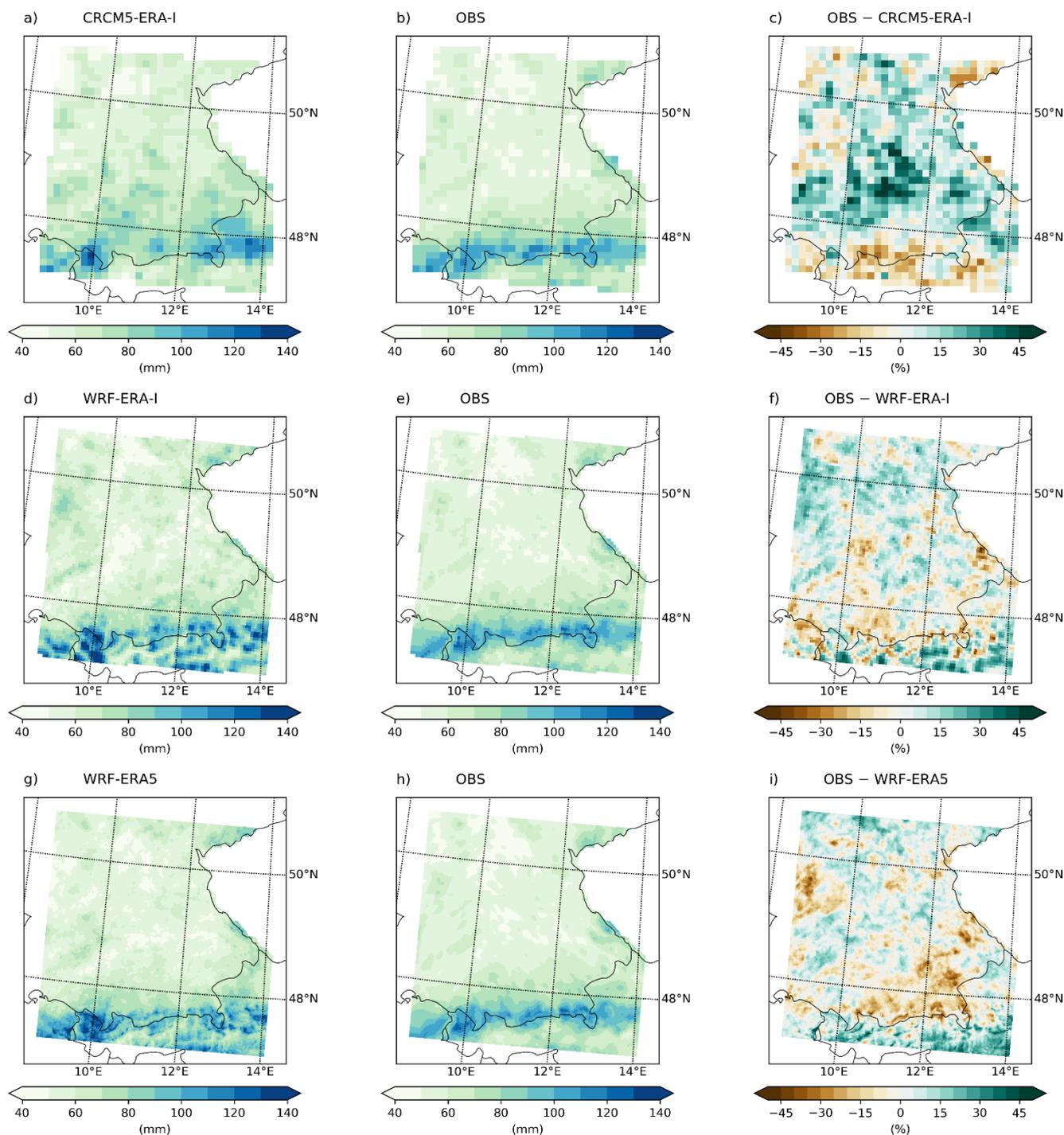
285 **Figure 4: 10-year rainfall return levels applying GEV-LMOM based on the CRCM5-ERA-Interim (a), WRF-ERA-Interim (d), WRF-ERA5 (g). The middle column (b, e, h) shows the observational product bilinearly interpolated to the respective climate model grid. The right column (c, f, i) provides the difference calculated as climate model return level minus observational return level.**

The 10-year return levels based on the WRF-ERA-I at 5 km resolution can recreate the spatial pattern of the observations with

290    a Spearman correlation of ρ = 0.82 (Fig. 4d). The higher intensities due to the orographic precipitation at the lower mountain ranges and their spatial patterns are reproduced, though the intensity around the Bavarian Forest is underestimated. In the alpine area, the WRF-ERA-I simulates higher intensities than observed, especially in the Alps southeast of the Inn valley. The overall bias amounts to +3.2 %. The histogram of simulated return levels is similar to the observed histogram (Fig. 5a), however, the very-high intensities above 110 mm d$^{-1}$ in the alpine area are overrepresented. Also the range of simulated return

295    levels extends to over 140 mm d$^{-1}$ (Fig. 5b).
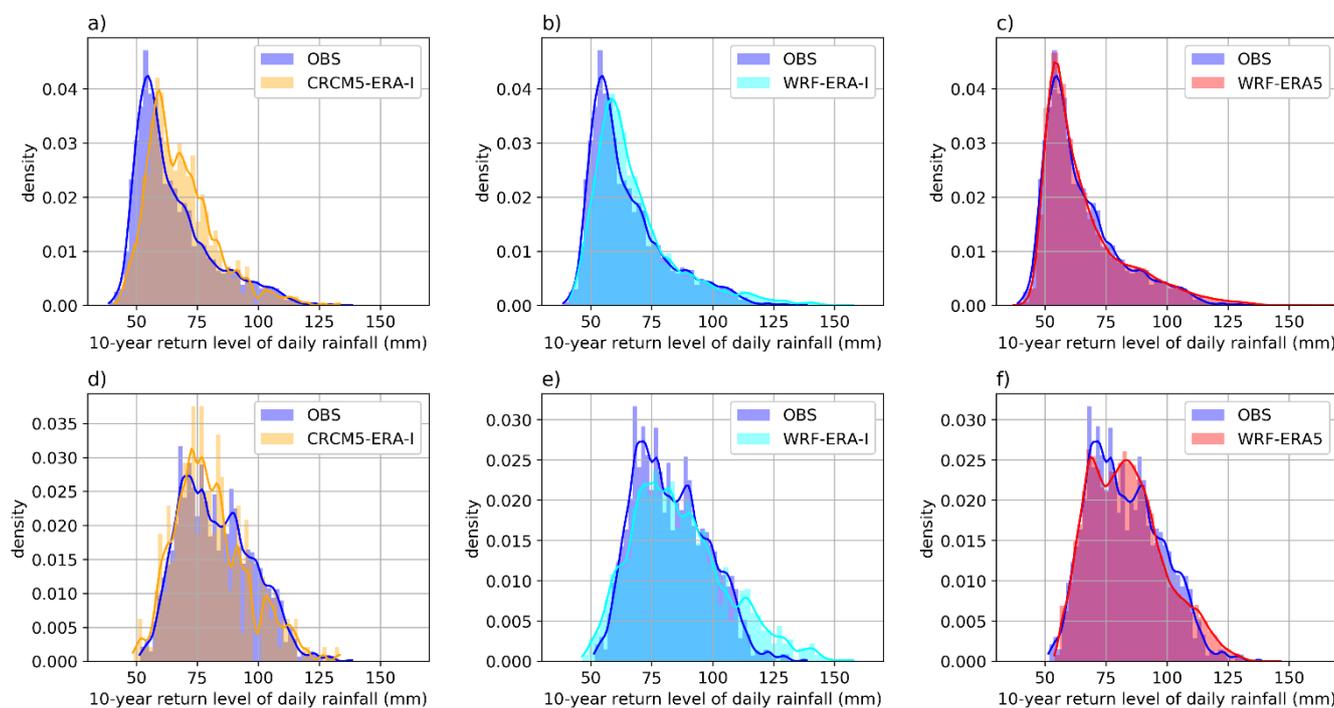


**Figure 5: Histograms of the resulting 10-year return levels in the whole study area (a-c) and the alpine area south of 48° N (d-f). Gaussian kernel density estimates are plotted to enhance the readability.**

300    The 10-year return levels based on the WRF-ERA5 show a generally similar spatial pattern to the WRF-ERA-I (Figs. 4g and 3d). The spatial pattern of orographic precipitation around the low mountain ranges is recreated, whereby the intensities at the Bavarian Forest and the Odenwald are underestimated. The return levels in south-eastern Bavaria are underestimated as well. As the WRF-ERA-I, also the WRF-ERA5 simulates high return levels above 100 mm d$^{-1}$ in the Alps southeast of the Inn valley. This results in a Spearman correlation of ρ = 0.82. The spatial average of the bias amounts to +1.1 %. The range and

305    distribution of the simulated return levels is very close to the observations for the whole study area (Fig. 5a) as well as south of 48° N (Fig. 5b).
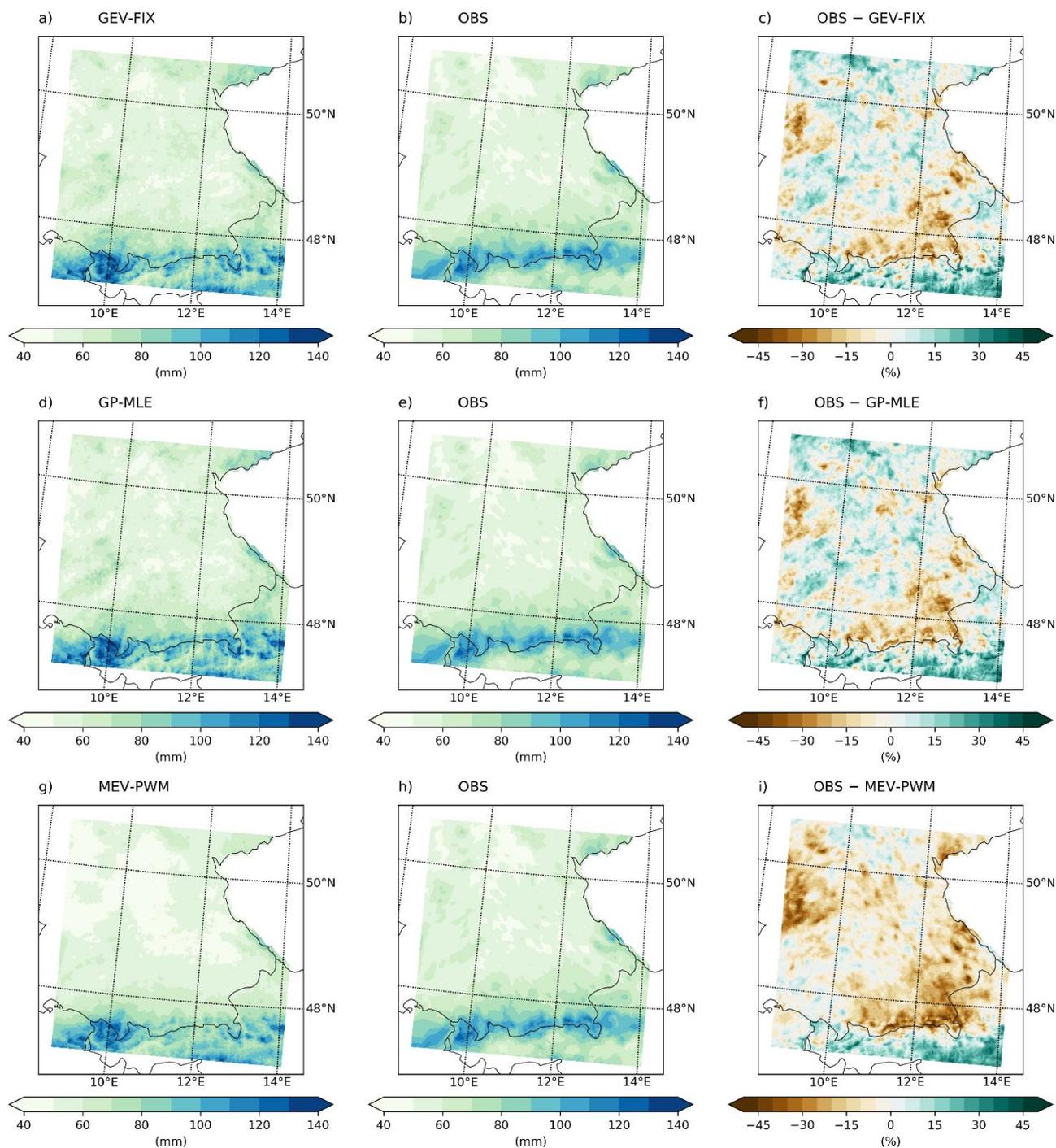
**Figure 6: 10-year rainfall return levels based on the WRF-ERA5 featuring GEV-FIX (a), GP- MLE (d), and featuring MEV-PWM (g). The middle column (b, e, h) shows the observational product interpolated to the WRF-ERA5 grid. The right column (c, f, i) provides the difference calculated as climate model return level minus observational return level.**

Figure 6 compares the three different EVT approaches GEV-FIX, GP-MLE, and MEV-PWM. The intensities as well as the resulting spatial distribution of the GEV-FIX and the GP-MLE are very similar to the GEV-LMOM (Fig 4a, 6a and 6d). The spatial correlation between GEV-FIX and the observations amounts to $\rho = 0.79$ and the overall bias to $+0.8$ %. For the GP-MLE, the spatial correlation is $\rho = 0.81$ and the overall bias is $+1.3$ %. The MEV-PWM method also shows a similar spatial

315 pattern (Fig. 6g), which is slightly more homogeneous than the GEV-LMOM or GP-MLE. The 10-year return levels based on the MEV-PWM are estimated generally lower than by the classical approaches. The spatial correlation between MEV-PWM and the observations amounts to $\rho = 0.86$ and the overall bias to $-6.2$ %.

## 4.2 Discussion

320 Generally, the high values of the Spearman's rank correlation as well as the visual comparison to the observational product (Fig. 4) prove that all three RCM setups are able to capture the topographic and climatic differences within the study area. Furthermore, the overall low bias of the return level indicates that the complex climate of heavy daily precipitation is reproduced by the climate models.

Both, the overall bias, and the spatial correlation imply that the WRF setups at 5 km and 1.5 km spatial resolution can better

325 reproduce the observed return levels than the broader-resolution CRCM5-ERA-I. There, both WRF setups show similar performance metrics. However, this equivalent performance may be caused by the spatial resolution and spatial representativeness of the observational data (see Fig. S4 for the native resolution). The German dataset is natively resolved at roughly 8 km, and the Austrian dataset at 6 km, whereas the Swiss data are given by single gauges interpolated via ordinary kriging. Hence, small-scale spatial features below such resolution cannot be evaluated by comparison to this observational

330 product. Comparing the WRF-ERA-I and WRF-ERA5 (Figs. 4d and 4g) reveals a similar spatial pattern, where the higher-resolved WRF-ERA5 can especially add more topographically driven spatial variability in the Alps.

### 4.2.1 Uncertainties of the observational datasets

As the German, Austrian, and Swiss data are based on rain gauge measurements, these data are subject to the usual

335 measurement inaccuracies leading to an underestimation of rainfall (Westra et al., 2014). For flat areas in Germany, this deviation is estimated about 5 % during summer (Richter, 1995). In mountainous areas, this deviation is expected to increase due to higher wind speeds. According to Sevruk (1981) it amounts to 7 % for Switzerland during summer. In addition to these systematic underestimations, different rain gauge types yield varying rainfall measurements inducing additional uncertainty (Vuerich et al., 2009). This applies for the different meteorological networks in the study area (Kainz et al., 2007; Frei and

340 Schär, 1998; Rauthe et al., 2013, Zolina et al., 2008).

Apart from these measurement errors, the gridded return level products suffer from a limited number of rain gauges (see Section 2.2), which also differ within their temporal coverage (Isotta et al., 2014). However, not only the number of stations,

Natural Hazards
and Earth System
Sciences
Discussions
EGU
Open Access

but also their spatial representativeness is important for an appropriate interpolation from point-wise measurements to gridded estimations (Ahrens, 2006). In mountainous areas, the spatial representativeness of a station is even more limited due to the

345 heterogeneous topography. In addition, the station distribution with elevation is not representative as well. Due to easier maintenance conditions more stations are located in valleys than on the tops of the mountains (Ahrens, 2006; Sevruk, 1997) leading to an underestimation for spatially interpolated rainfall in these areas (Isotta et al., 2014). Although the monitoring network density in the Alps makes this one of the best-monitored regions with complex topography, Isotta et al. (2014) estimate the "real" spatial resolution of the observations to be $10-25$ km. The regionalization of these point-wise measurements induces

350 additional uncertainties. For the German dataset, the orography is employed as additional variable to interpolate the return levels (Malitz and Ertel, 2015 following Bartels, 1992). Due to the limited spatial representativeness of the rain gauges in the Alps, the weather model OKM at 1.5 km resolution (Lorenz and Skoda, 2001) was used to support the spatial interpolation of the Austrian return level data (BMLRT, 2018; Kainz et al., 2007). Thereby, not only the spatial distribution of return levels was supported by the weather model simulations, but also the intensity of the resulting design rainfall return levels. The return

355 levels based on observations only are classified as "probably too low" due to the spatial distribution of the rain gauges, whereas the weather model return levels are estimated to be "probably too high" (BMLRT, 2006; 2018). Hence, the resulting design rainfall return level is a weighted averaging combination of the measured rainfall intensities and the intensities simulated by the weather model (BMLRT, 2006). This leads to the conclusion that the deviations of the 10-year return levels between the WRF setups (see Figs. 4f and 4i) and the observational data in the Austrian Alps may be caused by the limited spatial

360 representativeness of the measurement stations.

For the Swiss data, ordinary kriging is applied to regionalize the available pointwise 10-year return levels. As different interpolation methods yield differing results (Hu et al., 2019), this processing step induces additional uncertainty.

In summary, it can be stated that the study area offers a good temporal and spatial coverage of measurements, especially compared to other regions in Europe (Poschlod et al., 2021), which are, however, subject to the uncertainties mentioned above.

365 Additionally, uncertainties due to the application of different EVT approaches contribute to the overall uncertainty, which are discussed in Section 4.2.3 as they apply for both observations and climate model data.

Hence, the 10-year return levels (Fig. 4b, e, h) provide the best guess based on observations, but are still matter to substantial uncertainties, especially in the Alps.

### 4.2.2 Uncertainties of the RCM datasets

370 Generally, climate model simulations of historical conditions are subject to two major uncertainty factors (Hawkins and Sutton, 2009). Due to the chaotic nature of atmospheric processes the climate system is governed by internal variability. These non-linear dynamics lead to the behaviour of the system that slightly differing starting conditions may result in significantly differing climate realizations (Deser et al., 2012). However, in this study the degree of internal variability is constrained as the RCMs are forced by reanalysis data. The large-scale atmospheric flows are imposed by the lateral boundary conditions,

375 and, therefore this source of internal variability is not present in these three RCM setups (Christensen et al., 2001). Still, RCMs

are governed by smaller-scale atmospheric variability. Alexandru et al. (2007) have shown that a 20-member RCM ensemble of the CRCM driven by the same lateral boundary conditions with slightly perturbed starting conditions leads to a reasonable spread of simulated precipitation. Even seasonal weather model forecast simulations, which are initialized every month, still show variability, especially for precipitation extremes (Kelder et al., 2020; Thompson et al., 2017). Hence, internal variability
380     cannot be excluded as uncertainty source.

Since models can only represent a simplified image of reality, the structure of climate models leads to the second major uncertainty factor. Even though mainly physically-based, RCMs make use of parametrizations with a differing degree of complexity (Jerez et al., 2013). Model uncertainty includes all limitations of the climate model setup such as model-inherent simplifications, parameterizations and schemes, the lateral boundary conditions, nesting, nudging, spin-up times, and spatial
385     resolution.

Multi-model experiments using the same boundary and starting conditions yield deviating simulations of the climate (Holtanová et al., 2019; Solman et al., 2013). Yet, also the same model applying differing physics options and parametrization schemes can lead to significant variability in the model results (Laux et al., 2019). Hence, climate model setups can be optimized by choosing different model options and schemes and comparing the simulations to observed climate conditions.
390     For the WRF-ERA-I setup, this has been carried out for the whole domain covering central Europe following Wagner et al. (2018; Warscher et al., 2019). The CRCM5-ERA-I and the WRF-ERA5 setups are based on former applications of the respective climate model in different domains. Adapting the applied options to the study area could potentially improve the model performance (Collier and Mölg, 2020).

Furthermore, two different reanalysis datasets at 75 km and 30 km spatial resolution covering differing time periods are used
395     to drive the RCMs. Stucki et al. (2020) argue that the difference of the driving conditions regarding the spatial resolution can alter the simulation results, especially over complex terrain.

The different time windows (1980 – 2009 and 1988 – 2017) of the three model setups lead to different events being sampled. Due to the small sample size, this variance can also lead to deviations in the resulting return levels.

The overall differences between the three RCM setups indicating model uncertainty are less apparent in the resulting 10-year
400     return levels than in the evaluation of individual extreme events. For the close reproduction of extreme rainfall events, Stucki et al. (2020) have shown that initialization of the RCM briefly before the respective events improves the performance at recreating rainfall intensities. Here, the RCMs are run in climate mode featuring transient 30-year simulations (CRCM5-ERA-I, WRF-ERA-I) or annual initialization (WRF-ERA5). It cannot be expected that single extreme events are closely reproduced due to the internal variability. Hence, such comparison is not appropriate to evaluate the skill of the model, but to visualize the
405     differences due to internal variability and model uncertainties. Therefore, the daily rainfall intensities of the two extreme events in May 1999 and August 2005 are given in the Supplementary Materials (Figs. S5 and S6). Furthermore, such a comparison makes it clear that the compared setups are climate model setups and not weather model setups, despite the high spatial resolution (Kelder et al., 2020). While the simulation of individual extremes can differ greatly, the 10-year return levels as a climatic indicator for extreme precipitation show a high degree of agreement (see Fig. 4). This suggests that despite all the

410   simplifications and differences leading to model uncertainty, the models can reproduce the climatic character of extreme precipitation in the study area.

### 4.2.3 Uncertainties due to EVT

The concept of classical EVT (see Sections 3.1 and 3.2) holds under rather restrictive assumptions (Papalexiou et al., 2013)
415   and each step featuring the choice of the distribution and fitting the distribution parameters induces additional uncertainty (Miniussi and Marani, 2020). For the GEV approach, Eq. (1) holds for a large number of samples $n$ (ideally the sample size $n \to \infty$). In practice, the limited available time series make it very difficult to determine whether the distribution of extreme samples is close to its asymptotic GEV limit (Cook and Harris, 2004; Koutsoyiannis, 2004; Miniussi and Marani, 2020).

The POT approach partly overcomes the limitation of very low sample sizes by using the threshold $u$ to define extreme events.
420   However, the choice of this threshold is crucial as the assumptions of a Poisson arrival of exceedances $y$ as well as the GP distribution of these exceedances only hold for a threshold $u$ ensuring both the sampled events to be "extreme" and a large number of samples $n$ (Picklands, 1975, Miniussi and Marani, 2020).

Furthermore, uncertainty is induced by the parameter optimization of the respective EVD to adapt the theoretical EVD to the extreme precipitation samples, even though appropriate methods are chosen (see Section 3.3; Muller et al., 2009). Assessing
425   the goodness of fit by quality measures or statistical testing (e.g. the Anderson-Darling test) can lower the uncertainty due to the aforementioned assumptions. However, the goodness of fit can only assess the quality of the fit between the theoretical EVD and the empirical distribution of the samples. It cannot evaluate if the samples are a "good representation" of possible extreme rainfall events within the boundaries of internal variability of the climate system.

Uncertainty is therefore apparent as the different sampling approaches, EVDs, and fitting methods may lead to differing
430   estimations of rainfall return levels (Lazoglou and Anagnostopoulou, 2017). For the GEV-LMOM (Fig. 4g) and GP-MLE (Fig. 6d) the mean absolute deviation (MAD) between the 10-year return levels based on both approaches amounts to a spatial average of 1.7 %. Hence, despite different sampling, distributions, and fitting methods, the results are almost equivalent.

However, both classical approaches still suffer from drawbacks. Papalexiou and Koutsoyiannis (2013) as well as Serinaldi and Kilsby (2014) argue that producing stable fits of the shape parameter of the GEV and GP distributions needs larger sample
435   sizes than typically available. They have shown that the estimation of the shape parameter of the GEV and GP distributions is dependent on the sample size, whereby it is also influenced by the geographical location. Papalexiou and Koutsoyiannis (2013) propose to restrict the shape parameter of the GEV to a window described by a normal distribution around the mean value of 0.114 or to apply a fixed shape parameter of $\xi = 0.114$. Indeed, the shape parameter of GEV-LMOM shows a heterogeneous spatial distribution with small patches of positive and negative values for all three RCM setups (Fig. 3c, f, i). This spatial
440   distribution can be interpreted as "noise" due to too low sample sizes, where the 30 annual maximum precipitation events do not fully represent the range of possible extreme precipitation within the boundaries of internal climate variability. The distribution of the shape parameter $\xi$ based on all three RCM setups is centred around a value close to 0.114 (see Fig. S1c).

However, Papalexiou and Koutsoyiannis (2013) suggest that 99 % of the distribution should be between 0 and 0.225, whereas the distribution of all three RCM setups reveals a larger spread. When restricting the shape parameter to $\xi = 0.114$, however,

445 the 10-year return levels only differ very slightly from the GEV-LMOM resulting in an average MAD of 3.4 % (see Fig. 6a). As $\xi$ only defines the tail of the distribution it is more relevant for longer return periods.

In addition to unstable parameters fits, the sampling strategies of both classical EVT approaches only use a small fraction of available data. In this study, only 0.3 % (GEV) and 0.8 % (GP) of the daily precipitation sums from the RCMs are sampled. Especially with respect to short available observation time series, but also with respect to the extensive computational power

450 and related costs of such high-resolution climate models, this sampling is a waste of valuable information (Miniussi and Marani, 2020). The sampling of the MEV approach overcomes this limitation and uses the information of rainfall intensities of all wet days as well as the frequency of these days. This is found to result in more stable fits (Zorzetto et al., 2016). Zorzetto et al. (2016) concluded that the MEV outperforms the classical GEV approach due to the more stable parameter fits if the return period exceeds the length of the available samples. However, this is not the case for the 10-year return levels based on

455 30 years of RCM data in this study. As also found by Zorzetto et al. (2016), the performance of the MEV-PWM at reproducing the return levels for such a combination of return period and available sample size is slightly worse than of the GEV-LMOM (see Figs. 4 and 6).

## 5 Conclusion

Various combinations of high-resolution regional climate models driven by reanalysis data and state-of-the-art EVT

460 approaches have been explored to reproduce 10-year return levels or daily rainfall. The increase in spatial resolution comparing the 12 km CRCM5-ERA-I and the 5 km WRF-ERA-I setups to observations reveals added value in terms of spatial correlation and bias. The further increase in spatial resolution featuring the 1.5 km WRF-ERA5, accompanied by an explicit simulation of convective processes, cannot greatly improve the performance metrics. This is possibly since the observational product is resolved natively between 6 km and 8 km. Hence finer-scale spatial features cannot be evaluated by such comparison.

465 The resulting 10-year return levels based on the different applied EVT approaches show good agreement to each other and to the observational product. This suggests that the methodological uncertainty for return levels of moderate extremes is relatively low. However, if return periods outside the sample size are to be extrapolated, the estimation uncertainty of the shape parameter governing the tail of the GEV and GP distributions becomes more important. Two approaches are studied to overcome this uncertainty. The GEV with fixed shape parameter shows 10-year return levels, which are almost equivalent to the three-

470 parameter GEV optimized via L-moments. The rather new EVT approach by Marani and Ignaccolo (2015) featuring the MEV distributions leads to a slight underestimation of 10-year return levels. Nevertheless, this methodology shows great potential for extrapolation of longer return periods due to the larger sampling and, therefore, increased stability of the fits (Zorzetto et al., 2016).

The question remains to be answered as to what the findings of this study can contribute to in practice.

Natural Hazards
and Earth System
Sciences
Discussions

Open Access

EGU

475 First, in regions with a low density of rain gauges, such RCM setups can contribute to a homogeneous spatial estimation of return levels. Even in regions, where the rain gauges cannot represent the spatial heterogeneity, RCMs can be applied to support observational products. This is already being done in Austria using a convective-permitting weather model, and the results of this study reinforce such use of regional climate models. It is also conceivable to use the high-resolution spatial patterns of CPMs as an auxiliary variable for the interpolation of the return levels based on measured data (e.g. via kriging with external

480 drift; Haberlandt, 2007 or spatial GEV models; Davison et al., 2012). A visualization of a simple combination approach for such a subsequent enhancement is provided in Figure 7. Therefore, the differences at each grid cell between the observational product and the WRF-ERA5 GEV-LMOM are smoothed with a Gaussian filter and again added to the climate model return levels. However, this rather naive approach only serves to provide a visual impression of a possible enhancement.
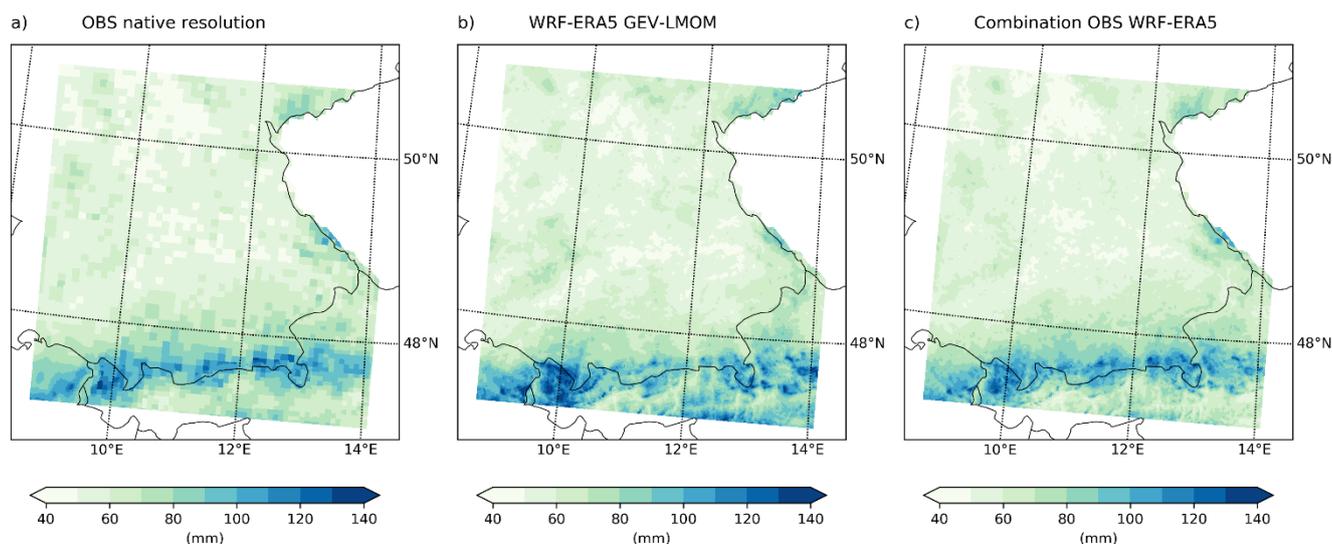


485

**Figure 7: (a) Observational 10-year return levels based on the German (8 km), Austrian (6 km), and Swiss (interpolated via ordinary kriging) data at original resolution. (b) 10-year return levels based on the WRF-ERA5 applying the GEV-LMOM approach. (c) Combination of (a) and (b) by applying a Gaussian filter on the differences.**

490 Second, large ensembles of RCMs can be set up to increase the sample size within the boundaries of the internal climate variability. On the one hand, increased sample sizes lower the uncertainty related to EVT, on the other hand large ensembles enable to quantify uncertainties due to internal variability (Poschlod et al., 2021).

Third, RCMs driven by global climate models following different emission scenarios allow to simulate climate change induced alterations of return levels (Ban et al., 2020). Even though an increase in extreme precipitation intensities is known for decades

495 (Trenberth et al., 2003), there is a lack of operational implementation and adaptation. In 2004, a climate change surcharge of a flat +15 % on top of the 100-year flood return level was introduced in Bavaria for the planning of flood protection facilities (LfU, 2021). However, such an adaptation for extreme rainfall is missing so far.

Despite all model-specific uncertainties, the evaluation of RCMs in this study proved that they are suitable to reproduce daily extreme precipitation intensities over complex terrain.

500

**Data availability**

The observational rainfall return level data are available at the German weather service (DWD, 2020), the Federal Ministry of Agriculture, Regions and Tourism Austria (BMLRT, 2020), and MeteoSwiss (MeteoSwiss, 2021). The daily precipitation of the WRF-ERA-I and WRF-ERA5 are publicly available at Warscher (2019) and Collier (2020),

505 which is cordially acknowledged. The CRCM5-ERA-I data are available on request from the author.

**Competing interests**

The author declares that he has no conflict of interest.

510 **Financial support**

**References**

515 Ahrens, B.: Distance in spatial interpolation of daily rain gauge data, Hydrol. Earth Syst. Sci., 10, 197–208, doi:10.5194/hess-10-197-2006, 2006.

Alexandru, A., Elia, R. D., and Laprisé, R.: Internal Variability in Regional Climate Downscaling at the Seasonal Scale, Mon. Weather Rev., 135, 3221–3238, doi:10.1175/MWR3456.1, 2007.

Balkema, A. A. and de Haan, L.: Residual life time at great age, Ann. Probab., 2, 5, 792–804, 1974.

520 Ban, N., Rajczak, J., Schmidli, J., and Schär, C.: Analysis of Alpine precipitation extremes using generalized extreme value theory in convection-resolving climate simulations, Clim. Dyn., 55, 61–75, doi:10.1007/s00382-018-4339-4, 2020.

Barbero, R., Fowler, H. J., Blenkinsop, S., Westra, S., Moron, V., Lewis, E., Chan, S., Lenderink, G., Kendon, E., Guerreiro, S., Li, X.-F., Villalobos, R., Ali, H., and Mishra, V.: A synthesis of hourly and daily precipitation extremes in different climatic regions, Weather. Clim. Extremes, 26, 100219, doi:10.1016/j.wace.2019.100219, 2019.

525     Bartels, H.: Regionalisierung am Beispiel der flächendeckenden Starkniederschlagsauswertung für die Bundesrepublik
        Deutschland. In: Regionalisierung in der Hydrologie, DFG, Mitteilung XI der Senatskommission für Wasserforschung,
        Weinheim, Germany, 1992.

        Bélair, S., Mailhot, J., Girard, C., and Vaillancourt, P.: Boundary layer and shallow cumulus clouds in a medium-range forecast
        of a large-scale weather system, Mon. Wea. Rev., 133, 7, 1938–1960, doi:10.1175/MWR2958.1, 2005.

530     Benjamin. D. M. and Budescu, D. V.: The Role of Type and Source of Uncertainty on the Processing of Climate Models
        Projections, Front. Psychol., 9, 403, doi:10.3389/fpsyg.2018.00403, 2018.

        Berg, P., Christensen, O. B., Klehmet, K., Lenderink, G., Olsson, J., Teichmann, C., and Yang, W.: Summertime precipitation
        extremes in a EURO-CORDEX 0.11° ensemble at an hourly resolution, Nat. Hazards Earth Syst. Sci., 19, 957–971,
        doi:10.5194/nhess-19-957-2019, 2019.

535     Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., and Kirchner, J. W.: The Relative Importance of Different Flood-
        Generating Mechanisms Across Europe, Water Resour. Res., 55, 4582–4593, https://doi.org/10.1029/2019WR024841, 2019.

        BLFW (Bayerisches Landesamt für Wasserwirtschaft): Hochwasser Mai 1999 Gewässerkundliche Beschreibung. Tech. Rep.,
        Munich, Germany, 44pp., 2003.

        BMLRT (Bundesministerium für Landwirtschaft, Regionen und Tourismus): Forschungsprojekt "Bemessungsniederschläge
540     in der Siedlungswasserwirtschaft", Tech. Rep., Vienna, Austria, 79pp., 2006.

        BMLRT (Bundesministerium für Landwirtschaft, Regionen und Tourismus): eHyd Auswertungen Karte
        Bemessungsniederschlag, Tech. Rep., Vienna, Austria, 10pp., 2018.

        BMLRT (Bundesministerium für Landwirtschaft, Regionen und Tourismus): eHYD, https://ehyd.gv.at/, last access: 22
        January 2020.

545     Boughton, W. and Jakob, D. Adjustment factors for restricted rainfall, Aust. J. Water Resour., 12, 37–47,
        doi:10.1080/13241583.2008.11465332, 2008.

        Breinl, K., Müller-Thomy, H., and Blöschl, G.: Space–Time Characteristics of Areal Reduction Factors and Rainfall Processes,
        J. Hydromet., 21, 4, 671–689, doi:10.1175/JHM-D-19-0228.1, 2020.

        Brisson, E., Van Weverberg, K., Demuzere, M., Devis, A., Saeed,S., Stengel, M., and van Lipzig, N. P. M.: How well can a
550     convection-permitting climate model reproduce decadal statistics of precipitation, temperature and cloud characteristics?,
        Clim. Dynam., doi:10.1007/s00382-016-3012-z, 2016.

        Bücher, A., Lilienthal, J., Kinsvater, P. and Fried, R.: Penalized quasi-maximum likelihood estimation for extreme value
        models with application to flood frequency analysis, Extremes, doi:10.1007/s10687-020-00379-y, 2020.

        Chen, K.-F. and Leandro, J.: A Conceptual Time-Varying Flood Resilience Index for Urban Areas: Munich City, Water, 11,
555     830; doi:10.3390/w11040830, 2019.

        Christensen, O. B., Gaertner, M. A., Prego, J. A., and Polcher, J.: Internal variability of regional climate models, Clim. Dyn.,
        17, 875–887, doi:10.1007/s003820100154, 2001.

        Coles, S.: An introduction to statistical modeling of extreme values, Springer, London, UK, 2001.

Collier, E., Sauter, T., Mölg, T., and Hardy, D.: The influence of tropical cyclones on circulation, moisture transport, and snow
560    accumulation at Kilimanjaro during the 2006–2007 season, J. Geophys. Res.-Atmos., 124, 6919–6928,
https://doi.org/10.1029/2019JD030682, 2019.

Collier, E.: BAYWRF, dataset, doi:10.17605/OSF.IO/AQ58B, 2020.

Collier, E. and Mölg, T.: BAYWRF: a high-resolution present-day climatological atmospheric dataset for Bavaria, Earth Syst.
Sci. Data, 12, 3097–3112, doi:10.5194/essd-12-3097-2020, 2020.

565    Cook, N. J. and Harris, R. I.: Exact and general FT1 penultimate distributions of extreme wind speeds drawn from tail-
equivalent Weibull parents, Struct. Saf., 26, 391–420, doi:10.1016/j.strusafe.2004.01.002, 2004.

Coppola, E., Sobolowski, S., Pichelli, E., Raffaele, F., Ahrens, B., Anders, I., Ban, N., Bastin, S., Belda, M., Belušić, D.,
Caldas-Alvarez, A., Cardoso, R. M., Davolio, S., Dobler, A., Fernandez, J., Fita, L., Fumiere, Q., Giorgi, F., Goergen, K.,
Güttler, I., Halenka, T., Heinzeller, D., Hodnebrog, Ø., Jacob, D., Kartsios, S., Katragkou, E., Kendon, E., Khodayar, S.,
570    Kunstmann, H., Knist, S., Lavín-Gullón, A., Lind, P., Lorenz, T., Maraun, D., Marelle, L., van Meijgaard, E., Milovac, J.,
Myhre, G., Panitz, H.-J., Piazza, M., Raffa, M., Raub, T., Rockel, B., Schär, C., Sieck, K., Soares, P. M. M., Somot, S., Srnec,
L., Stocchi, P., Tölle, M. H., Truhetz, H., Vautard, R., de Vries, H., and Warrach-Sagi, K.: A first-of-its-kind multi-model
convection permitting ensemble for investigating convective phenomena over Europe and the Mediterranean, Clim. Dynam.,
55, 3–34, doi:10.1007/s00382-018-4521-8, 2020.

575    Davison, A. C. and Smith, R. L.: Models for exceedances over high thresholds, J. Roy. Stat. Soc., 52, 3, 393–442, 1990.

Davison, A. C., Padoan, S. A., and Ribatet, M.: Statistical modelling of spatial extremes, Stat. Sci., 27, 161–186,
doi:10.1214/11-STS376, 2012.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli,P, Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo,
G.,Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bid-lot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M.,
580    Geer,A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V.,Isaksen, L., Kallberg, P., Köhler, M., Matricardi, M.,
McNally,A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey,C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and
Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc.,
137, 553–597, doi:10.1002/qj.828, 2011.

Delicado, P. and Goria, M. N.: A small sample comparison of maximum likelihood, moments and l-moments methods for the
585    asymmetric exponential power distribution, Comput. Stat. Data Anal., 52, 1661–1673, doi:10.1016/j.csda.2007.05.021, 2008.

Deser, C., Phillips, A., Bourdette, V., and Teng, H.: Uncertainty in climate change projections: the role of internal variability,
Clim. Dynam., 38, 527–546, doi:10.1007/s00382-010-0977-x, 2012.

DWD (Deutscher Wetterdienst): KOSTRA_DWD_2010R,
https://opendata.dwd.de/climate_environment/CDC/grids_germany/return_periods/precipitation/KOSTRA/KOSTRA_DWD
590    _2010R/asc/, last access: 21 October 2020.

DWD (Deutscher Wetterdienst): Klimaatlas, https://www.dwd.de/DE/klimaumwelt/klimaatlas/klimaatlas_node.html, last
access: 10 February 2021.

Fisher, R. A.and Tippett, L. H. C.: Limiting forms of the frequency distribution of the largest or smallest member of a sample, Math. Proc. Cambr. Phil. Soc., 24, 2, 180–190, doi:10.1017/S0305004100015681, 1928.

595 Fosser, G., Khodayar, S., and Berg, P.: Benefit of convection permitting climate model simulations in the representation of convective precipitation, Clim. Dynam., 44, 45–60, doi:10.1007/s00382-014-2242-1, 2014

Frei, C. and Schär, C.: A precipitation climatology of the Alps from high-resolution rain-gauge observations, Int. J. Climatol., 18, 873–900, doi:10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9, 1998.

Fukutome, S., Schindler, A, and Capobianco, A.: MeteoSwiss extreme value analyses: User manual and documentation. 600 Technical Report MeteoSwiss, 255, 3rd Edition, Zurich, Switzerland, 80 pp, 2015.

Gilleland, E., and Katz, R. W.: extRemes 2.0: An Extreme Value Analysis Package in R, J. Stat. Softw., 72, 8, doi:10.18637/jss.v072.i08, 2016.

Gnedenko, B.: Sur la distribution limite du terme maximum d'uneserie aleatoire, Ann. Math., 44, 423–453, 1943.

Goda, Y.: Inherent negative bias of quantile estimates of annual maximum data due to sample size effect: A numerical 605 simulation study, Coastal Eng. J., 53, 4, 397–429, doi:10.1142/S0578563411002409, 2011.

Goudenhoofdt, E. and Delobbe, L.: Generation and Verification of Rainfall Estimates from 10-Yr Volumetric Weather Radar Measurements, J. Hydrometeor., 17, 1223–1242, doi:10.1175/JHM-D-15-0166.1, 2016.

Greenwood, J., Landwehr, J., and Matalas, N.: Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form, Water Resour. Res., 15, 5, 1049–1054, doi:10.1029/WR015i005p01049, 1979.

610 Grell, G.A. and Freitas, S.R.: A scale and aerosol aware stochastic convective parameterization for weather and air quality modelling, Atmos. Chem. Phys, 14, 5233–5250, doi:10.5194/acp-14-5233-2014, 2014.

Grieser, J., Beck, C. and Rudolf, B.: The Summer Flooding 2005 in Southern Bavaria – A Climatological Review. In: Deutscher Wetterdienst (Ed.): Klimastatusbericht 2005, 168–173. Offenbach a. M., Germany, 2006.

Hawkins, E., and Sutton, R.: The potential to narrow uncertainty in regional climate predictions, Bull. Am. Meteorol. Soc., 90, 615 1095–1108, doi:10.1175/ 2009BAMS2607.1, 2009.

Gutowski, W. J., Decker, S. G., Donavon, R. A., Pan, Z. T., Arritt, R. W., and Takle, E. S.: Temporal-spatial scales of observed and simulated precipitation in central U.S. climate, J.Climate, 16, 3841–3847, doi:10.1175/1520-0442(2003)016<3841:tsooas>2.0.co;2, 2003.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D. and New, M.: A European daily high-resolution 620 gridded data set of surface temperature and precipitation for 1950–2006, J. Geophys. Res., 113, D20119, doi:10.1029/2008JD010201, 2008.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in regional climate predictions, B. Am. Meteorol. Soc., 90, 1095–1107, doi:10.1175/2009BAMS2607.1, 2009

Hernández-Díaz L., Laprise, R., Sushama, L., Martynov, A., Winger, K., Dugas, B.: Climate simulation over the CORDEX-625 Africa domain using the fifth-generation Canadian Regional Climate Model (CRCM5), Clim. Dynam., 40, 1415–1433, doi:10.1007/s00382-012-1387-z, 2012.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G. D., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L.,

630    Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P. de, Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 Global Reanalysis, Q. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hofstätter, M., Lexer, A., Homann, M., and Blöschl, G.: Large-scale heavy precipitation over central Europe and the role of atmospheric cyclone track types, Int. J. Climatol., 38, e497–e517, doi:10.1002/joc.5386, 2018.

635    Holtanová, E., Mendlik, T., Koláček, J., Horová, I., and Mikšovský, J.: Similarities within a multi-model ensemble: functional data analysis framework, Geosci. Model Dev., 12, 735–747, https://doi.org/10.5194/gmd-12-735-2019, 2019.

Hong, Y., Hsu, K.-L., Sorooshian, S. and Gao, X.: Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. J. Appl. Met., 43, 12, 1834–1853, doi:662 10.1175/JAM2173.1, 2004.

Hosking, J. R. M., Wallis, J. R., and Wood, E. F.: Estimation of the Generalized Extreme-Value Distribution by the

640    Method of Probability-Weighted Moments, Technometrics, 27, 251–261, doi:10.1080/00401706.1985.10488049, 1985.

Hu, Q., Li, Z., Wang, L., Huang, Y., Wang, Y., and Li, L.: Rainfall Spatial Estimations: A Review from Spatial Interpolation to Multi-Source Data Merging, Water, 11, 579, doi:10.3390/w11030579, 2019.

Hu, G. and Franzke, C. L. E. (2020): Evaluation of daily precipitation extremes in reanalysis and gridded observation-based data sets over Germany, Geophys. Res. Lett., 47, e2020GL089624, doi:10.1029/2020GL089624, 2020.

645    Isotta, F. A., Frei, C., Weilguni, V., Tadić, M. P., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., and Vertačnik, G.: The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, Int. J. Climatol., 34, 1657–1675, doi:10.1002/joc.3794, 2014.

Jarvis, A., Reuter, H. I., Nelson, A., and Guevara, E.: Hole-filled SRTM for the globe Version 4, dataset, 2008.

650    Jerez, S., Montavez, J. P., Gomez-Navarro, J. J., Lorente-Plazas, R.,Garcia-Valero, J. A., and Jimenez-Guerrero, P.: A multi-physics ensemble of regional climate change projections over the Iberian Peninsula, Clim. Dyn., 41, 1749–1768, doi:10.1007/s00382-012-1551-5, 2013.

Joyce, R. J., Janowiak, J. E., Arkin, P. A. and Xie, P.: CMORPH: A 680 method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydromet., 5, 3, 487-503,

655    doi:10.1175/1525-7541(2004)005h0487:CAMTPGi2.0.CO;2, 2004.

Junghänel, T., Ertel, H., and Deutschländer, T.: Bericht zur Revision der koordinierten Starkregenregionalisierung und -auswertung des Deutschen Wetterdienstes in der Version 2010, Tech. Rep., Deutscher Wetterdienst, Offenbach a. M., Germany, 30 pp., 2017.

Kain, J. S., and Fritsch, J. M.: A one-dimensional entraining/detraining plume model and its application in convective

660    parameterization, J. Atmos. Sci., 47, 2784–2802, doi:10.1175/1520-0469(1990)047%3C2784:AODEPM%3E2.0.CO;2, 1990.

Kainz, H., Beutle, K., Ertl, T., Fenz, R., Flamisch, N., Fritsch, E., Fuchsluger, H., Gruber, G., Hackspiel, A., Hohenauer, R., Klager, F., Lesky, U., Nechansky, N., Nipitsch, M., Pfannhauser, G., Posch, A., Rauch, W., Schaar, W., Schranz, J., Sprung, W., Telegdy, T., and Lehner, F.: Niederschlagsdaten zur Anwendung der ÖWAV-Regelblätter 11 und 19, Tech. Rep., ÖWAV, 2007.

665 Karki, R., ul Hasson, S., Gerlitz, L., Schickhoff, U., Scholten, T., and Böhner, J.: Quantifying the added value of convection-permitting climate simulations in complex terrain: a systematic evaluation of WRF over the Himalayas, Earth Syst. Dynam., 8, 507–528, doi:10.5194/esd-8-507-2017, 2017.

Kelder, T., Muller, M., Slater, L., Marjoribanks, T., Wilby, R. L., Prudhomme, C., Bohlinger, P., Ferranti, L., and Nipen, T.: Using UNSEEN trends to detect decadal changes in 100-year precipitation extremes, npj Clim. Atmos. Sci., 1–13, 670 doi:10.1038/s41612-020-00149-4, 2020.

Keller, L., Rössler, O., Martius, O., and Weingartner, R.: Delineation of flood generating processes and their hydrological response. Hydrol. Proc., 228–240, doi:10.1002/hyp.11407, 2017.

Kendon, E. J., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., and Senior, C. A.: Heavier summer downpours with climate change revealed by weather forecast resolution model, Nat. Clim. Change, 4, 7, 570–576, doi:10.1038/nclimate2258, 675 2014.

Klein Tank, A., Zwiers, F. W., and Zhang, X.: Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation, Tech. Rep., World Meteorological Organization, Geneva, Switzerland, 55pp., 2009.

Koutsoyiannis, D.: Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. Hydrol. Sci. J., 49, 4, 575–590, doi:10.1623/hysj.49.4.575.54430, 2004.

680 Kreklow, J., Tetzlaff, B., Burkhard, B. and Kuhnt, G.: Radar-Based Precipitation Climatology in Germany—Developments, Uncertainties and Potentials, Atmosphere, 11, 217, doi:10.3390/atmos11020217, 2020.

Kuo, H.-L.: On formation and intensification of tropical cyclones through latent heat release by cumulus convection, J. Atmos. Sci., 22, 1, 40–63, doi:10.1175/1520-0469(1965)022%3C0040:OFAIOT%3E2.0.CO;2, 1965.

Langhans, W., Schmidli, J., and Schär, C.: Bulk Convergence of Cloud-Resolving Simulations of Moist Convection over 685 Complex Terrain, J. Atmos. Sci., 69, 2207–2228, doi:10.1175/JAS-D-11-0252.1, 2012.

Laux, P., Kerandi, N., Kunstmann, H.: Physics Parameterization Selection in RCM and ESM Simulations Revisited: New Supporting Approach Based on Empirical Copulas, Atmosphere, 10, 150, doi:10.3390/atmos10030150, 2019.

Lazoglou, G., and Anagnostopoulou, C.: An Overview of Statistical Methods for Studying the Extreme Rainfalls in Mediterranean, Proceedings, 1, 681; doi:10.3390/ecas2017-04132, 2017.

690 Lewis, E., Fowler, H., Alexander, L., Dunn, R., McClean, F., Barbero, R., Guerreiro, S., Li, X.-F., and Blenkinsop, S.: GSDR: A Global Sub-Daily Rainfall Dataset, J. Clim., 32, 4715-4729, doi:10.1175/JCLI-D-18-0143.1, 2019.

LfU (Bayerisches Landesamt für Umwelt): Gewässerkundlicher Bericht Hochwasser August 2005, Tech. Rep., Augsburg, Germany, 75pp., 2007.

LfU (Bayerisches Landesamt für Umwelt): Anpassung an Hochwasser,
695 https://www.lfu.bayern.de/wasser/klima_wandel/wawi_anpassung/hochwasser/index.htm, last access: 15 February 2021.

Lorenz, P. and Skoda, G.: Bemessungsniederschläge kurzer Dauerstufen (D ≤ 12 Stunden) mit inadäquaten Daten. Mitt. Hydr. Dienst in Österr., 80, 1–24, Vienna, Austria, 2001.

Madsen, H., Rasmussen, P. F., and Rosbjerg, D.: Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling, Water Resour. Res., 33, 747, doi:10.1029/96WR03848, 1997.

700 Malitz, G., and Ertel, H.: KOSTRA-DWD2010: Starkniederschlagshöhen für Deutschland (Bezugszeitraum 1951 bis 2010), Tech. Rep., Deutscher Wetterdienst, Offenbach a. M., Germany, 40 pp., 2015.

Marani, M. and Ignaccolo, M.: A metastatistical approach to rainfall extremes. Adv. Wat. Res., 79, 121–126, doi:org/10.1016/j.advwatres.2015.03.001, 2015.

Martins, E. S. and Stedinger, J. R.: Generalized maximum likelihood Pareto-Poisson estimators for partial duration series,
705 Water Resour. Res., 37, 2551–2557, doi:10.1029/2001WR000367, 2001.

Martynov, A., Sushama, L., Laprise, R., Winger, K., and Dugas, B.: Interactive lakes in the Canadian Regional Climate Model, version 5: the role of lakes in the regional climate of North America, Tellus A, 64, 16226, doi:10.3402/tellusa.v64i0.16226, 2012.

Merz, R. and Blöschl, G.: A process typology of regional floods, Water Resour. Res., 39, 1340,
710 https://doi.org/10.1029/2002WR001952, 2003.

MeteoSwiss: Standard period 1966 – 2015. https://www.meteoswiss.admin.ch/home/climate/swiss-climate-in-detail/extreme-value-analyses/standard-period.html?station=int, last access: 20 January 2021.

Miniussi, A. and Marani, M.: Estimation of daily rainfall extremes through the metastatistical extreme value distribution: Uncertainty minimization and implications for trend detection, Water Resour. Res., 56, 7, e2019WR026535,
715 doi:10.1029/2019WR026535, 2020.

Mittermeier, M., Braun, M., Hofstätter, M., Wang, Y., and Ludwig, R.: Detecting climate change effects on Vb cyclones in a 50-member single-model ensemble using machine learning, Geophys. Res. Lett., 46, 14653–14661, doi:10.1029/2019GL084969, 2019.

Muller, A., Arnaud, P., Lang, M., and Lavabre, J.: Uncertainties of extreme rainfall quantiles estimated by a stochastic rainfall
720 model and by a generalized Pareto distribution / Incertitudes des quantiles extrêmes de pluie estimés par un modèle stochastique d'averses et par une loi de Pareto généralisée, Hydrol. Sci. J., 54, 3, 417–429, doi:10.1623/hysj.54.3.417, 2009.

Myhre, G., Alterskjær, K., Stjern, C. W., Hodnebrog, Ø., Marelle, L., Samset, B. H., Sillmann, J., Schaller, N., Fischer, E., Schulz, M., and Stohl, A.: Frequency of extreme precipitation increases extensively with event rareness under global warming, Sci. Rep., 9, 16063, doi:10.1038/s41598-019-52277-4, 2019

725 Nissen, K., and Ulbrich, U.: Increasing frequencies and changing characteristics of heavy precipitation events threatening infrastructure in Europe under climate change, Nat. Hazards Earth Syst. Sci., 17, 1177–1190, doi:10.5194/nhess-17-1177-2017, 2017.

Panosetti, D., Schlemmer, L., and Schär, C.: Convergence behavior of idealized convection-resolving simulations of summertime deep moist convection over land, Clim. Dyn., 55, 215–234, doi:10.1007/s00382-018-4229-9, 2020.

730 Papalexiou, S. M. and Koutsoyiannis, D.: Battle of extreme value distributions: A global survey on extreme daily rainfall. Water Resour. Res., 49, 187–201, doi:10.1029/2012WR012557, 2013.

Papalexiou, S. M., Koutsoyiannis, D., and Makropoulos, C.: How extreme is extreme? An assessment of daily rainfall distribution tails, Hydrol. Earth Syst. Sci., 17, 851–862, doi:10.5194/hess-17-851-2013, 2013.

Pickands, J. I. I. I.: Statistical inference using extreme order statistics, Ann. Stat., 3, 1, 119–131, doi:10.1214/aos/1176343003, 735 1975.

Poschlod, B., Hodnebrog, Ø., Wood, R. R., Alterskjær, K., Ludwig, R., Myhre, G., and Sillmann, J.: Comparison and evaluation of statistical rainfall disaggregation and high-resolution dynamical downscaling over complex terrain, J. Hydrometeorol., 19, 1973–1982, doi:10.1175/JHM-D-18-0132.1, 2018.

Poschlod, B., Ludwig, R. and Sillmann, J.: Ten-year return levels of sub-daily extreme precipitation over Europe, Earth Syst. 740 Sci. Data, 13, 983–1003, doi:10.5194/essd-13-983-2021, 2021.

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., van Lipzig, N. P. M., and Leung, R.: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges, Rev. Geophys., 53, 323–361, doi:10.1002/2014RG000475, 2015.

Richter, D.: Ergebnisse methodischer Untersuchungen zur Korrektur des systematischen Meßfehlers des Hellmann-745 Niederschlagsmessers, Tech. Rep., Deutscher Wetterdienst, Offenbach a. M., Germany, 1995.

Rauthe, M., Steiner, H., Riediger, U., Mazurkiewicz, A., and Gratzki, A.: A Central European precipitation climatology – Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS), Meteorol. Z., 22, 3, 235–256, doi:10.1127/0941-2948/2013/0436, 2013.

Serinaldi, F. and Kilsby, C. G.: Rainfall extremes: Toward reconciliation after the battle of distributions, Water Resour. Res., 750 50, 336–352, doi:10.1002/2013WR014211, 2014.

Sevruk, B.: Regional dependency of precipitation-altitude relationship in the Swiss Alps. – Climatic Change, 36, 355–369, doi:10.1023/A:1005302626066, 1997.

Sevruk, B.: Methodische Untersuchungen des systematischen Messfehlers der Hellmann-Regenmesser im Sommerhalbjahr in der Schweiz. Mitteilungen Nr. 52, Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie, ETH Zürich, 299pp., 1981.

755 Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Duda, M., Huang, X., Wang, W. and Powers, J.: A Description of the Advanced Research WRF Version 3; Tech. Rep. NCAR/TN-475+STR, NCAR TECHNICAL NOTE; University Corporation for Atmospheric Research: Boulder, CO, USA, p. 113, 2008.

Solman, S. A., Sanchez, E., Samuelsson, P., da Rocha, R. P., Li, L., Marengo, J., Pessacg, N. L., Remedio, A. R. C., Chou, S. C., Berbery, H., Treut, H. L., de Castro, M., and Jacob, D.: Evaluation of an ensemble of regional climate model simulations 760 over South America driven by the ERA-Interim reanalysis: model performance and uncertainties, Clim. Dynam., 41, 1139–1157, doi:10.1007/s00382-013-1667-2, 2013.

Stampoulis, D. and Anagnostou, E. N.: Evaluation of Global Satellite Rainfall Products over Continental Europe, J. Hydrometeorol., 13, 588–603, doi:10.1175/JHM-D-11-086.1, 2012.

Stucki, P., Froidevaux, P., Zamuriano, M., Isotta, F. A., Messmer, M., and Martynov, A.: Simulations of the 2005, 1910, and 1876 Vb cyclones over the Alps – sensitivity to model physics and cyclonic moisture flux, Nat. Hazards Earth Syst. Sci., 20, 35–57, doi:10.5194/nhess-20-35-2020, 2020.

Thompson, V., Dunstone, N. J., Scaife, A. A., Smith, D. M., Slingo, J. M., Brown, S., and Belcher, S. E.: High risk of unprecedented UK rainfall in the current climate, Nat. comm., 8, 1–6, doi: 10.1038/s41467-017-00275-3, 2017.

Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B.:The changing character of precipitation, B. Am. Meteorol. Soc.,84, 1205–1217, doi:10.1175/BAMS-84-9-1205, 2003.

Vuerich, E., Monesi, C., Lanza, L. G., Stagi, L., and Lanzinger, E.: WMO Field Intercomparison of Rainfall Intensity Gauges, Tech. Rep., WMO, Geneva, Switzerland, 287pp., 2009.

Wagner, A., Heinzeller, D., Wagner, S., Rummler, T., and Kunstmann, H.: Explicit Convection and Scale-Aware Cumulus Parameterizations: High-Resolution Simulations over Areas of Different Topography in Germany, Mon. Weather Rev., 146, 1925–1944, https://doi.org/10.1175/mwr-d-17-0238.1, 2018.

Warscher, M., Wagner, S., Marke, T., Laux, P., Smiatek, G., Strasser, U., and Kunstmann, H.: A 5 km Resolution Regional Climate Simulation for Central Europe: Performance in High Mountain Areas and Seasonal, Regional and Elevation-Dependent Variations, Atmosphere, 10, 682, doi:10.3390/atmos10110682, 2019.

Warscher, M.: High-resolution (5 km) RCM data for Central Europe, 1980-2009 and 2020-2049, WRF 3.6.1 forced by ERA-Interim and MPI-ESM, RCP4.5. , dataset, doi:10.5281/zenodo.2533904, 2019.

Westra, S., Fowler, H. J., Evans, J. P., Alexander, L. V., Berg, P., Johnson, F., Kendon, E. J., Lenderink, G., and Roberts, N. M.: Future changes to the intensity and frequency of short-duration extreme rainfall, Rev. Geophys., 52, 522–555, doi:10.1002/2014RG000464, 2014.

Wiedenmann, J., Rohn, J. and Moser, M.: The relationship between the landslide frequency and hydrogeological aspects: a case study from a hilly region in Northern Bavaria (Germany), Environ. Earth Sci., 75, 609, doi:10.1007/s12665-016-5451-6, 2016.

Wilks, D. S.: "The stippling shows statistically significant grid points": How research results are routinely overstated and overinterpreted, and what to do about it, Bull. Am. Meteorol. Soc., 97, 2263–2273, doi:10.1175/BAMS-D-15-00267.1, 2016.

Wilson, P. S. and Toumi, R.: A fundamental probability distribution for heavy rainfall, Geophys. Res. Lett., 32, L14812, doi:10.1029/2005GL022465, 2005.

ZAMG (Zentralanstalt für Meteorologie und Geodynamik): Klimamittel, https://www.zamg.ac.at/cms/de/klima/informationsportal-klimawandel/daten-download/klimamittel, last access: 10 February 2021.

Zolina, O., Simmer, C., Kapala, A., Bachner, S., Gulev, S., and Maechel, H.: Seasonally dependent changes of precipitation
795   extremes over Germany since 1950 from a very dense observational network, J. Geophys. Res., 113, 1–17,
doi:10.1029/2007JD008393, 2008.

Zorzetto, E., Botter, G., and Marani, M.: On the emergence of rainfall extremes from ordinary events. Geophys. Res. Lett., 43,
8076–8082, doi:10.1002/2016GL069445, 2016.

Zorzetto, E.: mevpy, https://github.com/EnricoZorzetto/mevpy, last access: 1 February 2021.