Earth and Space Science

RESEARCH ARTICLE
10.1029/2019EA000557

Key Points:
• We compare the performance of the metastatistical and generalized extreme value distributions (MEV, GEV)
• The MEV uncertainty is significantly increased when used for smooth spatial modeling of extremes
• The MEV should be preferred over the GEV for point estimates and smooth modeling

Correspondence to:
H. Schellander, harald.schellander@zamg.ac.at

Citation:
Schellander, H., Lieb, A., & Hell, T. (2019). Error structure of metastatistical and generalized extreme value distributions for modeling extreme rainfall in Austria. Earth and Space Science, 6, 1616–1632. https://doi.org/10.1029/2019EA000557

1. Introduction

Accurate estimation of daily rainfall extremes is essential for solving some hydrological problems, like the design of hydraulic infrastructures or the damage assessment of extreme floods. Classical extreme value theory offers the possibility of using maxima within a certain temporal block (often 1 year), demanding the number of events of the underlying process in that block to be sufficiently large. The cumulative distribution function of the block maxima then tends to an asymptotic distribution or generalized extreme value distribution (GEV). A drawback of the commonly used block maxima approach for modeling daily rainfall extremes is that the asymptotic conditions underlying extreme value theory (Fisher & Tippett, 1928) do not apply, unless the number of events per year is large enough. For analysis of annual maximum daily rainfall this implies that the number of wet days tends to infinity, or is at least large, while in reality, the number of wet days within a year is limited to 365 (366 in leap years), and might be rather small, especially in dry years, so that the asymptotic assumption of the GEV is violated. The usual procedure to incorporate more data in an extreme value analysis is the peaks-over-threshold (POT) approach. It uses all values above a suitably high threshold to model the excess values over that threshold using a generalized Pareto distribution (GPD; Davidon & Smith, 1990). This second approach also discards a large portion of the information in the bulk of the parent distribution.

To overcome this problem, Marani and Ignaccolo (2015) proposed the Metastatistical extreme value distribution (MEV), assuming that daily rainfall extremes are block maxima over a finite and stochastically variable number of “ordinary events,” which are defined as samples from the underlying distribution (Zorzetto et al., 2016). The ordinary events are considered to be Weibull right-tail equivalent (Wilson & Toumi, 2005), describing the intensity of daily rainfall. By additionally considering the number of wet days in a year as random variable, daily rainfall occurrence can be taken into account simultaneously. By using the 275 years of rainfall data of Padova (Italy), Marani and Ignaccolo (2015) showed that the error of the GEV is larger than for the MEV. By fitting a POT model to rainfall exceedances, Zorzetto et al. (2016) found that the MEV outperforms the GEV on average, when the estimated return period is larger than the number of calibration years. Similar results were found by Marra et al. (2018) for subdaily rainfall, when a smaller fraction of available years was used for fitting.

Abstract Incorrect estimation of extreme values of daily precipitation can have severe consequences in hydrological and engineering applications. Recent advances in the study of extreme precipitation have shown that the Metastatistical Extreme Value Distribution (MEV) is superior to the Generalized Extreme Value Distribution (GEV) whenever the length of the available record is small compared to the average recurrence time. This paper provides a detailed examination of the relative performance of MEV and GEV for both point estimates and spatial modeling. An analysis for a large number of sample years and return periods for daily precipitation in Austria shows that the MEV exceeds the GEV if the number of sample years is smaller, and the estimated return period is larger than 35 years. This advantage disappears almost entirely if the MEV is used for spatially smooth extreme value modeling instead of the GEV. However, the computational effort is drastically reduced in comparison to spatial modeling with the GEV if a simplified version of the MEV is used.
The current literature convincingly shows that the MEV exceeds the GEV if the length of the available data set is small compared to the average recurrence time. Automatic precipitation stations in Austria have typically been in operation for less than 25 years, motivating the investigation of when exactly the MEV should be used instead of the GEV. In addition to point estimates, maps of daily precipitation extremes are relevant for hydrological planning. There are several concepts for spatial estimation of daily precipitation extremes. Max-stable processes (de Haan, 1984) using block maxima have been very popular in recent years (Blanchet et al., 2016; Fawcett & Walshaw, 2014; Padoan et al., 2010; Reich & Shaby, 2012; Sebille et al., 2017; Westra & Sisson, 2011), also taking into account the anisotropy of extreme values (Blanchet & Cretu-in, 2017). Peak-over-threshold approaches were used with max-stable processes as well (Huser & Davison, 2014; Thibaud et al., 2013). If margins are more important than spatial extremal dependencies, a simpler concept for spatial interpolation of extremes which is termed smooth spatial modeling (Blanchet & Lehning, 2010, for snow depth) can be used. In most of these approaches, the GEV or GPD have been used as a basis for spatial extreme value modeling, while the MEV has never been adopted for this purpose.

Two questions arise in the context of the literature discussed above: What is the relative performance of the MEV distribution with respect to the GEV distribution in estimating extreme values when applied to daily point rainfall observations in Austria? Second, what is the predictive performance of the MEV distribution used within the framework of a spatial model for extreme rainfall?

2. Theoretical Framework

Let $X$ be a random variable on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with cumulative distribution $F$. Assume that $X : \Omega \rightarrow (0, \infty)$ models the amount of nonzero precipitation within 24 hr at a certain location. With $n \in \mathbb{N}$ denoting the number of wet days within 1 year, let $X_1, X_2, \ldots, X_n$ be a sequence of copies of $X$, that is, random variables describing precipitation in that time block. Extreme value theory (EVT) describes the distribution of the “block maximum,” that is, the distribution of the random variable

$$Z_n := \max\{X_1, X_2, \ldots, X_n\}.$$  

Assuming the $n$ events occurring in the block to be independent, the block maximum has cumulative distribution $H_n(x) := F(x)^n, x \in \mathbb{R}$. Note that in practice, $n$ is not fixed, but a realization of a random variable.

2.1. The GEV

Classical EVT assumes the number of events per block “large enough,” such that the asymptotic cumulative distribution function $H_n$ obtained when taking $n$ to infinity, is a good approximation for $H$. This limiting distribution takes the form of a GEV (e.g., Coles, 2001), which we denote by $\text{GEV}(\mu, \sigma, \xi)$. Thereby, $\mu \in \mathbb{R}$ refers to the location parameter, $\sigma > 0$ to the scale parameter and $\xi \in \mathbb{R}$ to the shape parameter. The distribution function of the GEV distribution is given by

$$F_{\text{GEV}}(s; \mu, \sigma, \xi) = \exp \left( -\left(1 + \xi \cdot \frac{s - \mu}{\sigma}\right)^{-1/\xi} \right)$$

for $s \in \mathbb{R}$ and $y_j = \max(y; 0)$ for $y = 1 + \xi \cdot \frac{x - \mu}{\sigma} \in \mathbb{R}$. Given the precipitation values of $T \in \mathbb{N}$ fully recorded years, that is, observations $x_{i_1}, x_{i_2}, \ldots, x_{i_T}$ of years $j \in \{1, \ldots, T\}$, we can calculate the block maxima $z_j := \max\{x_{i_1}, x_{i_2}, \ldots, x_{i_T}\}$ for $j \in \{1, \ldots, T\}$ and fit the GEV to the values $z_1, z_2, \ldots, z_T$.

Another less wasteful approach to extreme value modeling would be to not only use block maxima but consider rain events greater than some high threshold, termed as POT approach. Provided the threshold and the number of observations above that threshold are large enough, the natural distribution for such exceedances may be approximated by the GPD (Davison & Smith, 1990), given by

$$F_{\text{GPD}}(s; \sigma, \xi) = 1 - \left(1 + \xi \cdot \frac{s}{\sigma}\right)^{-1/\xi}$$

where $\sigma > 0$ is the scale parameter and $\xi \in \mathbb{R}$ the shape parameter; the range of $s$ is $0 < s < \infty$ if $\xi \leq 0$ and $0 < s < \sigma/\xi$ if $\xi > 0$. Note that the parameters of the GPD of threshold exceedces are equal to that of the corresponding GEV of block maxima. POT approaches are frequently used in modeling extreme rainfall (e.g., Cooley et al., 2007). However, even when more than just the block maxima is used to fit extremes over a threshold, a large part of available information from the bulk of the parent distribution is still neglected.
2.2. The MEV

With the aim of weakening the requirement of an asymptotic assumption, which has been shown to be invalid in many practical cases (Cook & Harris, 2004; Koutsoyiannis, 2004), a metastatistical approach is proposed by Marani and Ignaccolo (2015). This is defined in terms of the distribution of the statistical parameters describing “ordinary” daily rainfall occurrence and intensity. Extremes then emerge from repeated sampling of those ordinary events (Zorzetto et al., 2016). This approach naturally accounts for the influence of the bulk of the distribution of ordinary events on the distribution of annual maximum daily rainfall and is called MEV. The MEV accounts for the random process of event occurrence in each block and the possibly changing probability distribution of event magnitudes across different blocks, by recognizing the number of events in each block, \( n \), and the values of the parameters, \( \tilde{\theta} \), of the parent distribution \( F \) to be realizations of stochastic variables \( N \) and \( \tilde{\Theta} \). The cumulative distribution function of the distribution of block maxima is obtained by the use of the total probability theorem:

\[
\zeta(x) = \sum_{n=1}^{\infty} \int_{\Omega_\tilde{\Theta}} F(x; \tilde{\theta})^n g(n, \tilde{\theta}) d\tilde{\theta} \quad \text{for} \ x \in \mathbb{R} \tag{3}
\]

where \( g(n, \tilde{\theta}) \) is the joint probability distribution of \( N \) and \( \tilde{\Theta} \) and \( \Omega_{\tilde{\Theta}} \) is the set of all possible values of the parameters of the parent distribution. Following Zorzetto et al. (2016) and Wilson and Toumi (2005), we adopt the Weibull distribution \( F(x; C_j, w_j) = 1 - \exp \left( -\frac{x}{C_j} \right)^{w_j} \) to model the ordinary rainfall events, that is, the nonzero daily rainfall in each year \( j = 1, \ldots, T \), whereby, \( C_j > 0 \) is the scale and \( w_j > 0 \) the shape parameter, and compute a sample average approximation to (3) based on \( T \) realizations of \( (\tilde{\Theta}, N) \) by

\[
F_{\text{MEV}}(x; \bar{C}, \bar{w}, \bar{n}) = \frac{1}{T} \sum_{j=1}^{T} \left[ 1 - \exp \left( -\frac{x}{C_j} \right)^{w_j} \right]^{n_j} \tag{4}
\]

for \( x \in \mathbb{R} \), where \( \bar{C} = (C_1, \ldots, C_T) \) and the same applies to \( \bar{w} \) and \( \bar{n} \). With \( T \) fully recorded years, \( C_j \) and \( w_j \) can be estimated by fitting a Weibull distribution to the values \( x_{j1}, x_{j2}, \ldots, x_{jn_j} \) and \( n_j \) is the number of ordinary events for every year \( j \in \{1, \ldots, T\} \).

Particularly for the extension of the smooth modeling approach to the MEV in its original form (4), the Weibull parameters \( C \) and \( w \), as well as the number of wet days would have to be estimated separately for each year, leading to an optimization problem that is difficult to solve. Therefore, the probability distribution of daily rainfall is assumed to be time-invariant, that is, \( C_j = C \) and \( w_j = w \) for all \( j \in \{1, \ldots, T\} \), and \( n_j \) is replaced by the mean \( \bar{n} \). Thus, (4) becomes

\[
F_{\text{SMEV}}(x) \approx F_{\text{SMEV}}(x; C, w, n) = \left[ 1 - \exp \left( -\frac{x}{C} \right)^{w} \right]^{n} \tag{5}
\]

for \( x \in \mathbb{R} \), being a simple straightforward estimate of \( H_n = F^n \). In this case \( C \) and \( w \) are estimated based on all precipitation values from all years, and not by estimating parameters annually and taking the mean. Just recently Marra et al. (2019) proposed a simplified MEV (SMEV) formulation in which interannual variations of the occurrence and intensity of daily rainfall events are neglected and extremes are described as functions of the average properties of multiple underlying processes. As the above formulated, time-invariant version of the MEV (5) and that of Marra et al. (2019) are the same, we stick to the term SMEV.

2.3. PWM

Probability weighted moments (PWM; Greenwood et al., 1979) are potentially useful in expressing parameters of distributions, for which inverse forms allow an explicit notation. Defining \( M_{(k)} \) := \( \int_{\frac{1}{k}}^{1} x (1 - F(x)) \, dx \) for \( k \in \mathbb{R} \) they show in the case of a Weibull distribution \( M_{(0)} = C \cdot \Gamma \left( 1 + \frac{1}{w} \right) \) and \( M_{(1)} = C \cdot \Gamma \left( 1 + \frac{1}{2w} \right) \) and state the following representation for their parameters:

\[
C = \frac{M_{(0)}}{\Gamma \left( \frac{\log M_{(0)}}{\log \left( \frac{1}{2} M_{(1)} \right)} \right)} \quad w = \frac{\log(2)}{\log \left( \frac{M_{(0)}}{2M_{(1)}} \right)} \tag{6}
\]

Given realizations \( x_1 \leq x_2 \leq \ldots \leq x_n \), the sample moments \( \bar{M}_{(0)} := \frac{1}{n} \sum_{i=1}^{n} x_i \) respectively \( \bar{M}_{(1)} := \frac{1}{n(n-1)} \sum_{i=1}^{n} (n-i) x_i \) define unbiased estimators for the Weibull moments \( M_{(0)} \) respectively \( M_{(1)} \), and deliver estimators for the Weibull parameters by exchanging the Weibull moments in (6) by the sample moments (Heo et al., 2001).
3. Data and Methodology

In this study a total of 1,257 stations with daily rainfall observations of the National Weather Service (ZAMG) and the Hydrological Service (HD) of Austria were used (Figure 1). The data set was subject to standard quality control by the maintaining institutions (e.g., Lipa & Jurkovic, 2012) and covers different Austrian climatological regimes, from wet Stau-regions along the northern and southern side of the Austrian Alps to drier inner-alpine areas.

The data series have lengths between 24 and 161 years and are evenly spaced across Austria. Ten data series start before 1880, with the very first record from 1852, many of the time series last until 2017. Only 0.7% of all available years are missing, whereby the maximum of one station is 15.4% missing years. Three test statistics (Kolmogorov-Smirnov, Cramer-Von Mises, and Anderson-Darling) were used to test the applicability of the Weibull distribution to the ordinary events of daily rainfall at all stations. Gridded temperature (Hiebl & Frei, 2016) and precipitation (Hiebl & Frei, 2017) data sets (SPARTACUS, ZAMG) were used as covariates for the smooth spatial modeling, with horizontal and temporal resolutions of $1 \times 1$ km and 1 day. For the temperature grid, daily minimum and maximum temperature data from 1961 to 2012 at 112 stations were interpolated using nonlinear temperature profiles with altitude and accounting for the non-Euclidean spatial representativity of station observations. The precipitation grid was constructed from Austrian stations with daily rainfall from 1961 to 2014, using separate interpolations for mean monthly precipitation and daily relative anomalies.

3.1. Smooth Modeling

Using the approach of Blanchet and Lehning (2010) as blueprint, daily rainfall extremes are assumed to follow a GEV distribution (1), whose parameters $\mu$, $\sigma$, and $\xi$ can then be modeled in space considering linear relations of the form

$$\eta(x) = \alpha_0 + \sum_{k=1}^{m} \alpha_k y_k(x)$$

at location $x$, where $\eta$ denotes one of the GEV parameters, $y_1, \ldots, y_m$ are the considered covariates as smooth functions of the location, and $\alpha_0, \ldots, \alpha_m \in \mathbb{R}$ are the coefficients. Assuming stations to be spatially independent, the log-likelihood function $l_k(\mu(x_k), \sigma(x_k), \xi(x_k))$ at the $k$th station then reads as

$$l_k = \log l(\mu(x_k), \sigma(x_k), \xi(x_k)).$$
where \( l \) depends only on the coefficients of the linear models for the GEV parameters. This approach is called smooth modeling. Note that Blanchet and Lehning (2010) already showed that extreme snow depths can be considered spatially independent in the Swiss Alps. As daily rainfall extremes are spatially more heterogeneous compared to snow depth, the assumption of spatial independence for rainfall extremes in Austria is also expected to hold. The smooth assumption for the spatial component of daily rainfall might be questionable. However, on average, roughly half of the yearly precipitation falls as snow in large parts of Austria, particularly in the west and center, where high mountains dominate (Hiebl et al., 2011). This justifies daily rainfall in Austria to be modeled as spatially smooth. Moreover, the Weibull scale and shape parameters \( C \) and \( w \) of all available stations computed from all available years (not shown) show a fairly smooth distribution in space, further justifying the smooth spatial modeling approach for daily rainfall in Austria.

The two research questions posed in section 1 are addressed as follows: To investigate whether the GEV or the MEV should be used for point estimations of daily rainfall extremes in Austria, in section 4.1 a set of calibration stations is selected, and the three distributions GEV (1), MEV (4), and SMEV (5) are used to compute a relative root-mean-square (RMS) error for a large number of calibration years and recurrence times. To gain insight in the predictive performance of the MEV distribution used within the framework of a spatial model for rainfall extremes, in section 4.2, the smooth modeling concept is adapted to the SMEV, and both the smooth GEV and SMEV are compared using the same relative RMS error. In addition, absolute errors are computed for both point estimates and smoothly modeled quantiles with the GEV and SMEV distributions.

4. Results

4.1. Point Estimation

As a first step, the 55 stations with at least 100 fully recorded years were selected as the calibration data set \( S_{\text{point}} \) (cf. Figure 1). Let \( m \in \mathcal{M} := \{4, 6, 8, \ldots, 50\} \) describe the number of years used to fit an extreme value distribution and let \( q \in \mathcal{Q} := \{5, 10, 15, \ldots, 100\} \) specify a return period given in years. Then, separately for each calibration station \( s \in S_{\text{point}}, \) we chose \( \max(M) \) arbitrary years \( Y_s \) in a not necessarily chronological order. Next, for each \( s \) and for every \( m \in \mathcal{M}, \) we used the daily rainfall of the first \( m \) years in \( Y_s \) to fit the GEV as first variant. In contrast to Zorzetto et al. (2016), who obtained local estimations at the stations via a POT approach and \( l \) moments, in this study maximum likelihood (MLE) was used for parameter estimation, to stay consistent between the point estimations and the spatial modeling approach discussed in section 4.2.

As second variant, based on the same data, the parameters \( C \) and \( w \) of the MEV (4) in each single year by means of PWMs (see section 2.3) were estimated. PWM attributes a greater weight to the tail of the distribution than MLE. For small samples, MLE is known to be biased, especially for the estimation of \( w, \) whereas PWM works well and is more robust when outliers are present (Greenwood et al., 1979; Sornette, 2006).

As a third variant, we estimated the Weibull parameters using PWM based on the precipitation values of all \( m \) years at once and determined \( \hat{n}, \) that is, the mean number of wet days per year at the station \( s \) to obtain an explicit expression for SMEV.

After estimating these three distributions (GEV/MLE, MEV/PWM, and SMEV/PWM), for every return period \( q \in \mathcal{Q}, \) the corresponding return level \( RL_{\text{est}}(m, q) \) and the empirical return level \( RL_{\text{emp}}(q) \) using the Weibull plotting position were computed. More precisely, we estimated the sample frequency of an annual maximum \( z_i \) by \( F_i := i/(T + 1), \) where \( i \) is the rank of \( z_i \) in a list of the annual maxima sorted in ascending order, that is, \( z_1 \leq z_2 \leq \ldots \leq z_T, \) and the corresponding return period by \( q_i := 1/(1 − F_i). \) To get empirical return levels for arbitrary return periods, we fitted a monotonic smoothing spline through the points \( (q_i, z_i). \)

To compare the performance of GEV (1), MEV (4), and SMEV (5), the nondimensional RMS error was considered:

\[
err(m, q) = \sqrt{\frac{1}{|S_{\text{point}}|} \sum_{s \in S_{\text{point}}} \left( \frac{RL_{\text{est}}(m, q) − RL_{\text{emp}}(q)}{RL_{\text{emp}}(q)} \right)^2},
\]

(8)
Figure 2. Relative root-mean-square error $ERR$ (9) for the GEV (top) and MEV (bottom). Smooth error surfaces—visualized by color—through $M \times Q \times ERR$ were obtained applying spline interpolation. GEV = generalized extreme value distribution; MEV = metastatistical extreme value distribution.

where $S_{\text{point}}$ is a set of stations to calculate the error, in our case the set of 55 stations with at least $\max(Q) = 100$ years. For each of the three approaches GEV, MEV, and SMEV and each pair $(m, q)$, the above procedure was averaged over $r = 20$ realizations, leading to the following relative RMS error to assess model performance:

$$ERR(m, q) := \frac{1}{r} \sum_{k=1}^{r} err_k(m, q).$$

For every $k$ a new calibration subset of $\max(M) = 50$ years was randomly selected. The small standard deviation of roughly 1% of the 20 realizations of $err_k(m, q)$ strengthens the assumption of time-invariant daily rainfall in section 2.2. For the spatial models, $r = 20$ already requires the optimization of nearly 1,000 different models leading to computation times of several hours only for the GEV/MLE models. The small standard deviation of 1% also shows that more than 20 realizations would bring almost no gain in information. For consistency with the spatial models $r$ was set to 20 also for the point estimations.

Figure 2 compares the relative RMS error measure (9) obtained with the GEV (1) and the MEV (4). For both distributions, not surprisingly, the error increases with increasing return periods. While the MEV error increases only slightly up to 22.5% independent of the sample length $m$, the GEV error quickly reaches large values up to 100% particularly for short samples. Only for sample lengths of roughly $m > 30$ and return periods exceeding 30 years, the GEV slightly outperforms the MEV. For very small samples with lengths below roughly 14 years, the GEV leads to unacceptably large errors of 200%, while the MEV provides useful estimations, even for the largest quantiles. These findings nicely corroborate results of Zorzetto et al. (2016), who showed that the MEV outperforms the GEV even when a POT approach was used for fitting the GEV.

The difference of the error measures $ERR$ of the GEV and MEV distributions in Figure 2 is shown in Figure 3. The blue line separates the $m$-$q$-plane in two different regions. The lower half (bluish colors) indicates clearly when the MEV should be chosen in favor of the GEV for return level estimation of daily rainfall in Austria. For return periods larger than the sample size, the MEV clearly outperforms the GEV until sample lengths of
Figure 3. Difference of the pointwise estimated relative root-mean-square error measures $ERR_{GEV} - ERR_{MEV}$. Colors represent differences in percent of empirical estimates and were obtained applying spline interpolation. The blue line separates the area where the MEV performs better (bottom) from the area where the GEV could be taken instead for estimation (top). The black line indicates where $M/Q = 1$. MEV = metastatistical extreme value distribution; GEV = generalized extreme value distribution.

approximately 30 years. Particularly for high return periods with respect to sample lengths, the GEV shows up to 30% larger errors than the MEV. For samples with more than 35 years, Figure 2 indicates a slightly better performance of the GEV, albeit the improvement is only about 2.5%. Again this was also shown by Zorzetto et al. (2016), where for larger return periods above $\sim$40 years, some evidence can be found that the GEV (fitted via POT) is slightly superior to the MEV.

Heuristically, this may not be a surprise, because the MEV employs the whole information contained in the parent distribution and is therefore estimated on the basis of the bulk of rainfall values even for small numbers of sample years. In contrast, the GEV, whether fitted directly to maxima only or to some values above a high threshold is expected to get closer to the empirical distribution for large samples. The SMEV (5), that is, the MEV under the assumption that the Weibull parameters $C$ and $w$ are time invariant, leads to almost the same errors as the MEV (4). Figure 4 shows that the largest differences are on the order of $10^{-4}$. The SMEV is therefore considered to be a justified simplification of the MEV and a reasonable choice for the smooth modeling approach based on the MEV. Note that the Padova time series described in Marani and Ignaccolo (2015) also exhibits no visible trend for $C$ and $w$, at least for the years 1880–2006, which is the period that most of the stations used here span. Moreover, the relative RMS error (9) proposes the MEV...
Figure 5. Mean relative error, MRE (10) of the GEV (top) and the SMEV (bottom), computed pointwise at the 55 calibration stations introduced as $S_{point}$ in section 4.1. Negative (brown) values refer to underestimation, positive values describe overestimation of empirical quantiles. GEV = generalized extreme value distribution; SMEV = simplified metastatistical extreme value distribution.

(4) and its simplified variant SMEV (5) to be robust, as indicated by a small standard deviation of the 20 randomly selected calibration samples $err_s(m, q)$ (9) (maximal standard deviation $\sim 1\%$).

In order to investigate the underestimation or overestimation of empirical quantiles of the GEV and the SMEV, the mean relative error averaged over $r = 20$ realizations

$$MRE(m, q) := \frac{1}{r} \sum_{k=1}^{r} \frac{1}{|S_{point}|} \sum_{s \in S_{point}} \left( \frac{RL_{est}^s(m, q) - RL_{emp}^s(q)}{RL_{emp}^s(q)} \right)$$

was computed for the GEV and the SMEV for all pairs $(m, q) \in M \times Q$ of the 55 calibration stations $S_{point}$ (Figure 5). When less than $\sim 14$ years are used for calibration, the GEV severely overestimates all quantiles (Figure 5, top). The SMEV only overestimates return periods up to roughly 25 years, then it provides a remarkably stable underestimation of approximately 2% up to the largest return periods, which are much longer than the sample lengths (Figure 5, bottom). When the number of calibration years increases to more than 14 years, the GEV tends to underestimate empirical quantiles for all return periods up to $\sim 4\%$. The underestimation of the SMEV increases only marginally up to roughly 4% for the largest return periods.

4.2. Smooth Modeling

As stated at the end of the last section, and to avoid optimization problems with the standard formulation of the MEV (4) (cf. section 2.2), in the smooth modeling approach only the GEV and the SMEV are compared, termed sGEV and sSMEV hereafter. Similar to section 4.1, first, a set of calibration stations was selected to fit the smooth models described in section 3. As smooth modeling for the whole domain of Austria would not perform well with only the 55 stations selected as calibration stations in section 4.1, a new set $S_{spatial}$ with 264 stations with at least max$(M) = 50$ fully recorded years was selected, thereby neglecting the 55 calibration stations of $S_{point}$, which are then used as verification data set for the smooth models (cf. Figure 1).

In a next step, according to (7), appropriate linear models for the parameters $\mu$, $\sigma$, and $\xi$ of the sGEV, and $C$ and $w$ of the sSMEV have to be selected. Note that no model is needed for the mean number of wet days $\bar{n}$,
which can be derived directly from the data series. The sGEV and Weibull parameters at each of the 1,257 stations were estimated using all available rainfall data. Then a generalized linear regression was performed between the parameters and meaningful covariates added in a stepwise manner. Using the Akaike information criterion (AIC; Akaike, 1974) as outcome of the regression, the best linear model between a given parameter of the sGEV distribution and to the scale parameter of the sSMEV distribution. Covariates not considered are marked with “nc.” That means, for example, the sGEV scale and shape parameters are modeled depending on the sGEV location parameter. AIC = Akaike information criterion; MLE = maximum likelihood; PWM = probability weighted moment; GEV = generalized extreme value distribution; MEV = metastatistical extreme value distribution.

Table 1
Optimized Global Coefficients of the Linear Models for Daily Rainfall Extremes After AIC Model Selection for the Smooth SGEV and the SSMEV Models and Their Parameters

| Model               | Parameter estimation | Spatially optimized | Parameter | sGEV/MLE | sSMEV/MLE | sSMEV/PWM |
|---------------------|----------------------|---------------------|-----------|----------|-----------|-----------|
|                     |                      |                     | (1/10 mm) | (1/10 mm) | (1/10 mm) | (1/10 mm) |
| (Intercept)         | 1,167.15             | 1167.15             | -700.28   | -1.96    | 152.71    | 114.05    |
| Longitude           | 5.21                 | 2.19                | -         | -        | -2.62e-3  | -         |
| Latitude            | -26.67               | 12.42               | 3.4e-2    | -2.92    | -2.17     | 1.75      |
| Altitude            | 0.03                 | -                   | -2.33e-5  | 1         | -         | -         |
| Mean Precipitation  | 51.58                | -1.94               | 4.82e-2   | 2.32     | 1.71      | 9.01      |
| Mean Summer Temperature | 3.21             | 2.78                | 3.87e-3   | 8.52e-4  | 0.34      | 0.003     |
| 𝑛̄                   | 0.6                  | 2.19                | 6.14e-4   | -1.32e-4 | -         | -         |
| 𝜇                    | nc                   | 0.32                | -5.04e-4  | nc       | nc        | nc        |
| 𝜎                    | nc                   | nc                  | -5.74e-4  | nc       | nc        | nc        |
| 𝐶                    | nc                   | nc                  | -0.74     | nc       | 0.006     |

Note. The dash symbol “-” indicates parameters excluded by the model selection. The covariates longitude, latitude, altitude, mean precipitation, mean summer temperature, and the mean number of wet days are available on the SPARTACUS grid and referred to in the text. The variables μ, σ, and C in the first column refer to the location and scale parameter of the sGEV distribution and to the scale parameter of the sSMEV distribution.

In (11) the dash symbol “-” indicates parameters excluded by the model selection. The covariates longitude, latitude, altitude, mean precipitation, mean summer temperature, and the mean number of wet days are available on the SPARTACUS grid and referred to in the text. The variables μ, σ, and C in the first column refer to the location and scale parameter of the sGEV distribution and to the scale parameter of the sSMEV distribution.

In accordance to section 3.1, the smooth sGEV modeling approach determines global model coefficients \(\tilde{\alpha}, \tilde{\beta}\), and \(\tilde{\gamma}\) of the linear models for \(\mu, \sigma,\) and \(\xi\) by employing maximum likelihood estimation jointly over all stations in \(S_{\text{spatial}}\), that is, by maximizing

\[
\sum_{s \in S_{\text{spatial}}} \sum_{j=1}^{T_s} \log \left( f \left( z_{sj}^{(s)}, \mu^{(s)}(\tilde{\alpha}), \sigma^{(s)}(\tilde{\beta}), \xi^{(s)}(\tilde{\gamma}) \right) \right)
\]

through \(\tilde{\alpha}, \tilde{\beta},\) and \(\tilde{\gamma}\). In (11) \(T_s\) denotes the number of years available at station \(s\) and \(z_{sj}^{(s)}\) the annual maximum of year \(j\). The parameters \(\mu^{(s)}, \sigma^{(s)},\) and \(\xi^{(s)}\) are determined by \(\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}\), and the covariates, thus, once global model coefficients are obtained, they can be computed easily for an arbitrary site in Austria using the covariates provided on the SPARTACUS grid.

For consistency with the point estimates examined in section 4.1, the smooth sSMEV modeling approach employs PWMs to determine global model coefficients \(\tilde{\delta}\) and \(\tilde{\eta}\) of the linear models for \(C\) and \(w\). At the begin-
Relative root-mean-square error $ERR$ (9) of the models sGEV/MLE (top), sSMEV/MLE (center), and sSMEV/PWM (bottom). Colors depict smooth error surfaces through $M \times Q \times ERR$ and were obtained applying spline interpolation. GEV = generalized extreme value distribution; MLE = maximum likelihood; MEV = metastatistical extreme value distribution; PWM = probability weighted moments.

For every station $s \in S_{spatial}$ and putting all observations $x_{j_1}, x_{j_2}, \ldots, x_{j_n}$ of years $j \in \{1, \ldots, T_s\}$ together, sample moments $\hat{M}^{(0)}_{(s)}$ and $\hat{M}^{(1)}_{(s)}$ are computed once and for all. Subsequently, the global coefficients are chosen by minimizing

$$
\sum_{s \in S_{spatial}} \left[ \frac{\hat{M}^{(0)}_{(s)} - M^{(0)}_{(C^{(s)}, W^{(s)})}}{\hat{M}^{(0)}_{(s)}} \right]^2 + \left[ \frac{\hat{M}^{(1)}_{(s)} - M^{(1)}_{(C^{(s)}, W^{(s)})}}{\hat{M}^{(1)}_{(s)}} \right]^2
$$

through $\delta$ and $\eta$. In every iteration step, the parameters $C^{(s)}$ and $w^{(s)}$ in (12) are determined through the linear models by $\delta$, $\eta$, and the covariates.

To further investigate the influence of the spatial optimization method, three different smooth model configurations were tested. For the sGEV/MLE model, first, the parameters were estimated with MLE, the spatial optimization was then also carried out with MLE, that is, maximizing (11). The sSMEV/MLE model used PWM for local parameter estimations and MLE for the spatial optimization, and sSMEV/PWM applied PWM.
Mean relative error MRE (10) of the smooth models sGEV (top) and sSMEV (center) optimized with MLE, and the smooth sSMEV optimized with PWM (bottom), computed at the same 55 verification stations as for the point estimations. Negative (brown) values refer to underestimation, positive values describe overestimation of empirical quantiles. GEV = generalized extreme value distribution; MLE = maximum likelihood; MEV = metastatistical extreme value distribution; PWM = probability weighted moments.

Figure 7. Mean relative error MRE (10) of the smooth models sGEV (top) and sSMEV (center) optimized with MLE, and the smooth sSMEV optimized with PWM (bottom), computed at the same 55 verification stations as for the point estimations. Negative (brown) values refer to underestimation, positive values describe overestimation of empirical quantiles. GEV = generalized extreme value distribution; MLE = maximum likelihood; MEV = metastatistical extreme value distribution; PWM = probability weighted moments.

to local parameter estimations and the spatial optimization, that is, minimizing (12). The optimized parameter models for the three smooth model variants are listed in Table 1. Both sSMEV models use the same covariates, but have different coefficients due to different optimization methods.

Next, for every station $s \in S_{\text{spatial}}$ we chose $\max(M) = 50$ arbitrary years $Y_s$, not necessarily in chronological order, and then, for every $m \in M$, we maximized (11) or minimized (12) both with $T_s = m$ for all $s \in S_{\text{spatial}}$ by using the first $m$ years in $Y_s$. To achieve a fair comparison of model performance between different $m$ values, the very same stations are used for both models and all $m \in M$. Once smooth models for every $m \in M$ are optimized, we calculated for every station $s \in S_{\text{spatial}}$ that is, every station with at least $\max(Q) = 100$ fully recorded years, and for every $q \in Q$ the corresponding return level $RL_{est}^s(m, q)$ and as a reference the empirical return level $RL_{emp}^s(q)$, followed by the same error computation used for the point estimates (section 4.1). Recall that the error calculation took place at the 55 calibration stations of section 4.1, which were not used for fitting the smooth models. Analogously to the point estimations, the above exposed procedure
Empirical densities of the relative error \( \text{RE} \) for the three models sGEV/MLE, sSMEV/MLE, and sSMEV/PWM, depicted by colors. Positive \( \text{RE} \) values refer to overestimation, negative values to underestimation of empirical quantiles.

The error of the sSMEV/MLE model is surprisingly large for all return periods, giving rise to the suspicion that MLE is not particularly well suited as spatial optimization method for the smooth sSMEV model. As explained in section 4.1, maximizing the log-likelihood may be inappropriate for estimating the Weibull parameters, regardless of whether the likelihood is determined locally at one station or as the sum over all stations (Sornette, 2006).

The results clearly indicate that the advantage of the SMEV for the point estimations is almost absent for the spatial model sSMEV/PWM, but largely increased for the model sSMEV/MLE (cf. Figure 2). The maximum \( \text{ERR}_{\text{sSMEV/PWM}} \) in Figure 6 is 19.8%, compared to 21.4% for the sGEV/MLE model and 28.4% for the sSMEV/MLE model. Moreover, the maximum difference between the average error of the sSMEV/PWM model and that of the sGEV/MLE model at some point \((m, q)\) is only 1.9%, which may be negligible in practice. Both models seem to provide similar adaptation at the validation points for all \((m, q) \in M \times Q\). The error of the sSMEV/MLE model shows a slight increase for high return periods \(q\), when \(m\) becomes smaller than 10 to 15, demonstrating the large bias of the MLE for estimating GEV parameters from small samples (Martins & Stedinger, 2000). In contrast, the error of the sSMEV/PWM model seems to be almost independent of the number of years used for fitting.

The compensating behavior of the sGEV/MLE in case of smooth modeling may be explained by the following hypothesis, which would, however, need further testing: The uncertainty associated with smooth modeling is much smaller than that associated with local GEV estimates, because by maximizing the sum of the log-likelihood functions over all stations, strength is borrowed from neighboring observation sites. However, as the local SMEV parameter estimations are already much better than those for the GEV, the gain with smooth SMEV modeling might be smaller than for smooth GEV modeling. In addition, already for small samples the number of values available for fitting a sGEV distribution in case of smooth modeling (number of stations \(\times m\)) may be large enough for a good fit of the sGEV.

As can be expected from the analysis of the relative RMS error, the sSMEV/MLE model with MLE as spatial optimization method shows a severe overestimation of all return periods. The overestimation tends to be somewhat smaller for larger samples, which could be related to the problematic Weibull parameter estimation of MLE from small samples.

To gain insight into the spatial error structure of the three model variants, the following cross-validation was performed. All 1,257 available stations described in section 3 were used for calibration, 319 stations (all stations with more than 50 fully recorded years) were selected for verification. The verification stations were randomly split into 29 groups of 11 stations. For each verification group the model was fitted with all calibration stations except those from the verification group, and the 50-year return level of daily rainfall was computed at the 11 stations in the verification group.

Empirical densities of the relative error \( \text{RE} = (\text{RL}_{\text{est}} - \text{RL}_{\text{emp}})/\text{RL}_{\text{emp}} \) where \( \text{RL}_{\text{est}} \) stands for estimated return level, and \( \text{RL}_{\text{emp}} \) for empirical return level at the 319 verification stations for the three models (Figure 8) show that the sSMEV/MLE model mostly overestimates empirical return values, whereas the other models slightly underestimate them. The other two models are comparable in accuracy and it is difficult to identify a clear winner. Figure 9 illustrates different spatial distributions of the RE for the three models. Along,
Figure 9. Spatial structure of the relative error RE of the smooth (top) sGEV/MLE, (center) sSMEV/MLE, and (bottom) sSMEV/PWM models for daily rainfall with return period 50 years. Background colors illustrate elevation (cf. Figure 1). The dashed black line depicts the main alpine crest. Verification stations are marked in red (blue) colors when the distributions overestimate (underestimate) empirical return levels.

and to the south of, the main alpine crest (marked as black dashed line) the smooth sSMEV/PWM model mostly overestimates empirical 50-year quantiles, whereas on the northern side of the Austrian alps empirical quantiles are generally underestimated, albeit to a smaller extent. The error is close to zero, that is, between $-0.05$ and $0.05$ (light gray-colored points in Figure 9), at $\sim 21\%$ of the verification stations ($25\%$ for the sGEV/MLE model). We suspect one of the reasons for this spatial pattern might lie in the fact that daily rainfall extremes in the south and southeast of Austria are more related to storms on a subdaily scale, whereas rainfall in the north is dominated by elongated frontal systems. These are spatially much more homogeneous and therefore better captured by a smooth modeling approach.

While the accuracy between the smooth sGEV/MLE and sSMEV/PWM models is comparable, the latter brings distinct computational advantages as using the sSMEV/PWM requires that only two parameters $(C, w)$ must be estimated instead of three for the sGEV/MLE. Using the original form of the MEV (4) would be even more computationally expensive as it would require distinct models for $(C_j, w_j)$ for each year $j$. When minimizing (12), the parameters of only two instead of three linear models need to be optimized. Further-
more, when building a sSMEV/PWM model, the recorded values are needed only once at the beginning for a cheap computation of the sample moments $\hat{M}(s,m)_{(0)}$ and $\hat{M}(s,m)_{(1)}$ occurring in (12), while in the case of a smooth sGEV/MLM model every evaluation of the fitness function (11) makes use of the observed annual maxima. Therefore, the cost function (12) is just a sum over a set of stations, whereas (11) is a double sum consisting of $m$ times more summands; more precisely every summand of the outer sum is a sum over all recorded years, which are used for the fit, at this location. This led to a temporal performance boost in the optimization with the smooth sSMEV/PWM of nearly 80%.

4.2.1. Return Level Maps

Figure 10 shows return level maps of 50-year daily rainfall in Austria, computed with the three smooth models sGEV/MLM, sSMEV/MLM, and sSMEV/PWM defined in Table 1. All maps basically reflect the findings of a classical alpine precipitation climatology (Frei & Schärf, 1998): The highest return levels are located in the southernmost part of Austria, as well as along the northern border to Germany and the main crest. The driest regions are some inner-alpine valleys in the west and center and the adjacent flatlands in the north. This implies that the daily precipitation regime in Austria is mainly governed by large-scale atmospheric ascent...
Figure 11. Difference of sGEV/MLE and sSMEV/PWM models of 50-year daily rainfall. All available stations with all available daily rainfall data were used for fitting. Positive (red) values indicate larger return values for the sGEV/MLE model. The main alpine crest is marked by the black dashed line.

along the Alps and subsequent enhanced precipitation, as well as shadowing effects for the inner-alpine areas.

As the previous analyses suggest, the different models lead to return levels deviating considerably in some regions. As expected, the largest daily rainfall returns are produced by the sSMEV/PWM model. Despite the comparable accuracy of the sGEV/MLE and sSMEV/PWM models described in the previous section, the latter provides the smallest 50-year return values for daily rainfall. While on average the underestimation of the sGEV/MLE model is somewhat larger than that of the sSMEV/PWM (cf. Figure 7), it provides higher return values north of the main alpine ridge than the sSMEV/PWM model and smaller return values south of it. This can also nicely be seen in Figure 11. Although it is difficult to determine a winner between the sGEV/MLE and sSMEV/PWM models from the error analyses in section 4.2, for practical applications it should be considered that the calculated return values of the two models north and south of the main alpine ridge differ significantly from each other. Thus, the sGEV/MLE model achieves up to 50 mm more daily rainfall in a 50-year event than the sSMEV/PWM model, especially in the Stau-regions in the north. In the south, the sGEV/MLE model produces about 20 mm less precipitation.

The reason for this clear separation of rainfall extremes into two distinct zones north and south of the Austrian Alps can only be suspected. Nevertheless, as already pointed out in the previous section, one explanation might be that the smooth modeling approach is not well suited for modeling daily rainfall extremes. Another reason might be that daily rainfall extremes in Austria originate from different processes on either side of the Alps, namely from large-scaled frontal systems in the north and local thunderstorms in the south. This could however be accounted for with the SMEV concept as already suggested by Marra et al. (2019).

Another way to face this problem would be to use separate smooth models for each side of the Alps.

5. Conclusions

The systematic pointwise investigation of daily rainfall extremes in Austria with the classical GEV and the MEV distributions reveals that in total the MEV outperforms the GEV for point estimations as well as when the distribution parameters are smoothly modeled in space. The pointwise estimated relative RMS error of the MEV remains below 25% of the empirical quantiles, regardless of the sample length $m$ and the return period of interest $q$. For quantiles larger than the sample size, being the classical practical application of such probabilistic models, the MEV clearly outperforms the GEV up to return periods of roughly 30 years. For sample lengths larger than this, the GEV very slightly outperforms the MEV, albeit the difference lies in the order of 2.5% of empirical quantiles. The largest benefit with the MEV is gained when only very small sample lengths are available, and the sought-after return period is much longer. For $m = q \leq 14$ years...
the GEV overestimates empirical quantiles by more than 20%, while the MEV only slightly underestimates empirical quantiles by up to 3%. The presented results require that parameters of the GEV are estimated via maximum likelihood, whereas the Weibull parameters of the MEV have to be estimated with PWMs.

For smooth modeling with almost no loss of accuracy, a simplified MEV version (SMEV) with fixed scale and shape parameter for all years, and the mean number of wet days can be used. The clear advantage for point-wised errors disappears almost completely, when the SMEV is used for modeling rainfall extremes spatially via the smooth modeling approach. The reason is suspected to be (1) better gain of accuracy of the sGEV in case of smooth modeling, than for the MEV, and (2) a strongly increased sample size already for small sample lengths due to the optimization of the sum of all station-wise log-likelihoods. Again these results assume that the sGEV model uses MLE as spatial optimization method and SMEV is spatially optimized by PWM's. As the spatial SMEV optimization with MLE leads to unusable large errors and return levels, it can be argued, that the spatial optimization concept for smooth modeling could be improved for the SMEV. As for the SMEV only two parameters have to be estimated (three for the sGEV), saving approximately 80% of computation time makes the smooth SMEV a better choice also for smooth modeling.

Fifty-year return level maps of Austrian daily rainfall based on the three models sGEV/MLE, sSMEV/MLE, and sSMEV/PWM basically reflect the results of alpine precipitation climatologies, in that the largest rainfall amounts are located along the northern and central crest of the Alps, as well as in the south of Austria. The mean error for the 50-year daily rainfall extremes estimated via a cross-validation procedure at 319 stations show a small underestimation on the northern side of the Austrian Alps for the sGEV/MLE and SMEV/PWM models. Empirical quantiles are overestimated along the main alpine crest and in the south of Austria for both models, where subdaily rainfall events dominate the precipitation climatology. Despite the similar accuracy of the sGEV/MLE and SMEV/PWM models, return levels of the first are significantly larger than those of the second model north of the alpine ridge, whereas they are smaller in the south of Austria. One reason for this ambivalent behavior is suspected to be the violation of the identical-distribution assumption, because extremes in the north emerge from a different process (frontal system) than in the south (thunderstorms).

The results presented extend the current knowledge about when the MEV should be preferred over the GEV distribution for the estimation of daily rainfall extremes in Austria. For every pair of \((m, q)\) the proper distribution can be precisely selected. Nevertheless, regardless of point estimations or smooth spatial modeling, the MEV is more robust and should always be preferred.

Acknowledgments
The authors are grateful to ZAMG for data acquisition and provision. Note that observations used in this study are generally not freely available. However, in compliance with particular regulations they are available on request from ZAMG or the Hydrological Service for research purposes. We appreciate comments of three anonymous referees who made numerous suggestions to improve the manuscript. Sincere thanks go to Lindsey Nicholson for proofreading the manuscript.

References
Akaïke, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. Retrieved from http://ieeexplore.ieee.org/document/1100705/, https://doi.org/10.1109/TAC.1974.1100705
Blanchet, J., & Creutin, J.-D. (2017). Co-occurrence of extreme daily rainfall in the French Mediterranean region. *Water Resources Research, 53*, 9330–9349. https://doi.org/10.1002/2017WR020717
Blanchet, J., & Lehning, M. (2010). Mapping snow depth return levels: Smooth spatial modeling versus station interpolation. *Hydrology and Earth System Sciences, 14*(12), 2527–2544. Retrieved from https://www.hydrol-earth-syst-sci.net/14/2527/2010/, https://doi.org/10.5194/hess-14-2527-2010
Blanchet, J., Molinié, G., & Touati, J. (2016). Spatial analysis of trend in extreme daily rainfall in southern France. *Climate Dynamics, 51*, 799–812. https://doi.org/10.1007/s00382-016-3122-7
Coles, S. (2001). *An introduction to statistical modeling of extreme values*. London: Springer.
Cook, N. J., & Harris, R. (2004). Exact and general FTI penultimate distributions of extreme wind speeds drawn from tail-equivalent Weibull parents. *Structural Safety, 26*(4), 391–420. Retrieved from http://www.sciencedirect.com/science/article/pii/S0167473004000037, https://doi.org/10.1016/j.strusafe.2004.01.00
Cooley, D., Nychka, D., & Naveau, P. (2007). Bayesian spatial modeling of extreme precipita-tion return levels. *Journal of the American Statistical Association, 102*(479), 824–840. https://doi.org/10.1198/016214506000000780
Davison, A. C., & Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B (Methodological), 52*, 393–442. WileyRoyal Statistical Society. Retrieved from https://www.jstor.org/stable/2345667, https://doi.org/10.2307/2345667
de Haan, L. (1981). A spectral representation for Max-stable processes. *The Annals of Probability, 12*(4), 1194–1204. Retrieved from http://www.jstor.org/stable/2243357, https://doi.org/10.1214/aop/1176993148
Fawcett, L., & Walshaw, D. (2014). Estimating the probability of simultaneous rainfall extremes within a region: a spatial approach. *Journal of Applied Statistics, 41*(5), 955–976. https://doi.org/10.1080/02664763.2013.856872
Fisher, R. A., & Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society, 24*(02), 180. Retrieved from http://www.journals.cambridge.org/abstract_S0305004100015681, https://doi.org/10.1017/S0305004100015681
Frei, C., & Schär, C. (1998). A precipitation climatology of the Alps from high-resolution rain-gauge observations. *International Journal of Climatology, 18*(8), 873–900. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9/abstract, https://doi.org/10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9

Greenwood, J. A., Landwehr, J. M., Matalas, N. C., & Wallis, J. R. (1979). Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research, 15*(5), 1049–1054. https://doi.org/10.1029/WR015i005p01049

Heo, J.-H., Salas, J. D., & Kim, K.-D. (2001). Estimation of confidence intervals of quantiles for the Weibull distribution. *Stochastic Environmental Research and Risk Assessment, 15*(4), 284–309. https://doi.org/10.1007/s007040100071

Hiebl, J., & Frei, C. (2016). Daily temperature grids for Austria since 1961—Concept, creation and applicability. *Theoretical and Applied Climatology, 124*(1-2), 161–178. https://doi.org/10.1007/s00704-015-1411-4

Hiebl, J., & Frei, C. (2017). Daily precipitation grids for Austria since 1961—Development and evaluation of a spatial dataset for hydroclimatic monitoring and modelling. *Theoretical and Applied Climatology, 132*(1-2), 327–345. https://doi.org/10.1007/s00704-017-2093-x

Hiebl, J., Reisenhofer, S., Auer, I., Böhm, R., & Schöner, W. (2011). Multi-methodical realisation of Austrian climate maps for 1971–2000. *Advances in Science and Research, 6*, 19–26. Retrieved from http://www.adv-sci-res.net/6/19/2011/, https://doi.org/10.5194/asr-6-19-2011

Huser, R., & Davison, A. C. (2014). Space–time modelling of extreme events. *Journal of the Royal Statistical Society Series B, 76*, 439–461.

Koutsoyiannis, D. (2004). Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. *Hydrological Sciences Journal, 49*(4), 575–590. https://doi.org/10.1623/hyos.49.4.575.54430

Lipa, W., & Jurkovic, A. (2012). *Arbeitsanweisung OFFLINE Klimadatenaufarbeitung: Dokumentation für internen Gebrauch*. Vienna: ZAMG.

Lippe, Q., Fougères, A.-L., & Mercadier, C. (2017). Modeling extreme rainfall. A comparative study of spatial extreme value models. *Spatial Statistics, 22*(Part A), 187–208. Retrieved from www.sciencedirect.com/science/article/pii/S2211675316300975, https://doi.org/10.1016/j.spasta.2017.06.009

Marra, F., Zoccatelli, D., Armon, M., & Morin, E. (2019). A simplified MEV formulation to model extremes emerging from multiple nonstationary underlying processes. *Advances in Water Resources, 127*(April), 280–290. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S0309170818309011, https://doi.org/10.1016/j.advwatres.2019.04.002

Marani, M., & Ignaccolo, M. (2015). A metastatistical approach to rainfall extremes. *Hydrological Sciences Journal, 60*(3), 737–744. https://doi.org/10.1002/hysj.999WR015i005p01049

Padoan, S. A., Ribatet, M., & Sisson, S. A. (2010). Likelihood-based inference for max-stable processes. *Annals of Applied Statistics, 4*(8), 263–277. https://doi.org/10.1198/jasa.2009.tm08577

Reich, B. J., & Shaby, B. A. (2012). A hierarchical max-stable spatial model for extreme precipitation. *Annals of Applied Statistics, 6*(4), 1430–1451. https://doi.org/10.1214/12-AOAS591

Sebille, E., & Mercadier, C. (2017). Detection of non-stationarity in precipitation extremes using a max-stable process model. *Journal of Hydrology, 406*(1-2), 119–128. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022169410041112, https://doi.org/10.1016/j.jhydrol.2011.06.014

Sornette, D. (2006). *Critical phenomena in natural sciences*. Berlin/Heidelberg: Springer-Verlag.

Thibaud, E., Mutzner, R., & Davison, A. C. (2013). Threshold modeling of extreme spatial rainfall. *Water Resources Research, 49*, 4633–4644. https://doi.org/10.1002/wrcr.20329

Westra, S., & Sisson, S. A. (2011). Detection of non-stationarity in precipitation extremes using a max-stable process model. *Journal of Hydrology, 406*(1-2), 119–128. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022169410041112, https://doi.org/10.1016/j.jhydrol.2011.06.014

Wilson, P. S., & Toumi, R. (2005). A fundamental probability distribution for heavy rainfall. *Geophysical Research Letters, 32*, L14812. https://doi.org/10.1029/2005GL022465

Zorzetto, E., Botton, G., & Marani, M. (2016). On the emergence of rainfall extremes from ordinary events. *Geophysical Research Letters, 43*, 8076–8082. https://doi.org/10.1002/2016GL069445