Extreme precipitation return levels for multiple durations on a global scale

Gründemann, Gaby J.; Zorzetto, Enrico; Beck, Hylke E.; Schleiss, Marc; van de Giesen, Nick; Marani, Marco; van der Ent, Ruud J.

DOI
10.1016/j.jhydrol.2023.129558

Publication date
2023

Document Version
Final published version

Published in
Journal of Hydrology

Journal of Hydrology 621 (2023) 129558

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Extreme precipitation return levels for multiple durations on a global scale

Gaby J. Gründemann [a,b,*], Enrico Zorzetto [c], Hylke E. Beck [d], Marc Schleiss [e], Nick van de Giesen [a], Marco Marani [f], Ruud J. van der Ent [a,g]

[a] Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, Netherlands
[b] Centre for Hydrology, University of Saskatchewan, Canmore, Alberta, Canada
[c] Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA
[d] King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia
[e] Department of Geoscience and Remote Sensing, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, Netherlands
[f] Dipartimento di Ingegneria Civile, Edile ed Ambientale, Universita' degli Studi di Padova, Padova, Italy
[g] Water Research Centre, School of Civil and Environmental Engineering, University of New South Wales, Sydney, Australia

ARTICLE INFO

This manuscript was handled by Emmanouil Anagnostou, Editor-in-Chief, with the assistance of Shuo Wang, Associate Editor.

Dataset link: https://data.4tu.nl/articles/dataset/GPEX_Global_Precipitation_EXtremes/12764429/4

MSC:
86A05

Keywords:
Precipitation extremes
MSWEP
Metastatistical extreme value distribution
Generalized extreme value distribution
Peaks-over-threshold
Global domain

ABSTRACT

Quantifying the magnitude and frequency of extreme precipitation events is key in translating climate observations to planning and engineering design. Past efforts have mostly focused on the estimation of daily extremes using gauge observations. Recent development of high-resolution global precipitation products, now allow estimation of global extremes. This research aims to quantitatively characterize the spatiotemporal behavior of precipitation extremes, by calculating extreme precipitation return levels for multiple durations on the global domain using the Multi-Source Weighted-Ensemble Precipitation (MSWEP) dataset. Both classical and novel extreme value distributions are used to provide insight into the spatial patterns of precipitation extremes. Our results show that the traditional Generalized Extreme Value (GEV) distribution and Peak-Over-Threshold (POT) methods, which only use the largest events to estimate precipitation extremes, are not spatially coherent. The recently developed Metastatistical Extreme Value (MEV) distribution, that includes all precipitation events, leads to smoother spatial patterns of local extremes. For durations of 5 and 10 days, however, there are less events per year to fit the distribution (37 and 22 on average, respectively), leading to larger inter-annual variability and possible overestimation of the extremes. While the GEV and POT methods predict a consistent shift from heavy to thin tails with increasing duration, the MEV method predicts a relatively constant heaviness of the tail for any precipitation duration, opening up an important research question on what is the 'correct' tail behavior of extreme precipitation for different durations. The generated extreme precipitation return levels and corresponding parameters are provided as the Global Precipitation EXtremes (GPEX) dataset. These data can be useful for studying the underlying physical processes causing the spatiotemporal variations of the heaviness of extreme precipitation distributions.

1. Introduction

Extreme precipitation events are a major contributor to natural disasters (CRED, 2019). Accurate estimates of the severity of intense precipitation events are needed for an enhanced disaster risk understanding, such as that of floods and landslides. The urgency of this is indicated as the first priority of the Sendai Framework for Disaster Risk Reduction (UNSIDR, 2015). The accurate quantification of extremes is also necessary for infrastructure planning and design. Some countries already provide spatiotemporal estimates of extreme precipitation based on extreme value distributions (EVDs), for example, for Australia (Ball et al., 2019), the Netherlands (Beersma et al., 2018), and the US (e.g., Perica et al., 2015, 2018). However, many countries and regions do not have sufficient local data available (Gründemann et al., 2018; Kidd et al., 2017; van de Giesen et al., 2014), such that spatially-distributed extreme precipitation estimates are not possible.

Several previous studies have developed global-scale datasets of extreme precipitation. Courty et al. (2019) calculated intensity-duration-frequency curves at the global domain and their scaling with different event durations using reanalysis data and the Generalized Extreme Value (GEV) distribution with fixed tail parameter. Dunn et al. (2020) produced the HadEX3 dataset, which contains 29 generic precipitation and temperature indices, although these indices are not based on EVDs. Furthermore, this dataset has a coarse 1.25° latitudinal ×1.875° longitudinal resolution, with data-gaps due to insufficient available

* Corresponding author at: Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, Netherlands.
E-mail address: g.j.gruendemann@tudelft.nl (G.J. Gründemann).

https://doi.org/10.1016/j.jhydrol.2023.129558
Received 22 July 2022; Received in revised form 12 October 2022; Accepted 19 April 2023
Available online 26 April 2023

gauge data. Other global studies mostly focused on examining which type of distribution is most suitable to capture the tail behavior of extreme precipitation (Cavanaugh and Gershunov, 2015; Cavanaugh et al., 2015; Papalexiou et al., 2013). In addition, the spatial patterns of the parameter that controls the tail decay have been studied for the GEV distribution (Papalexiou and Koutsoyiannis, 2013; Ragulina and Reitan, 2017), and the Generalized Pareto (GP) distribution (Serinaldi and Kilsby, 2014). However, several issues remain to be addressed in order to obtain global-domain extreme precipitation return levels: (1) the choice of the dataset, (2) the focus on daily extremes without considering sub-daily extremes, (3) the choice of the time blocks over which block-maxima are determined, and (4) the exploration of possible alternatives to the classical EVDs, especially with respect to the tail behavior.

1. Several (quasi-)global gridded precipitation datasets have been developed in recent years, each with strengths, weaknesses, and uncertainties. See Sun et al. (2018), Beck et al. (2019a) and Rajulapati et al. (2020) for recent overviews of available datasets. Most of these datasets are based on gauge, reanalysis, or satellite sensor data. Notable examples of gauge-based datasets include GPCC-FDR (Becker et al., 2013; Schneider et al., 2011) and REGEN (Contractor et al., 2020). However, gauges are extremely unevenly distributed across the globe (Kidd et al., 2017; Schneider et al., 2014), and the number of active gauges has been declining in recent decades (Mishra and Coulibaly, 2009). Satellite-based products such as CMORPH (Joyce et al., 2004), GSMaP (Ushio et al., 2009), IMERG (Huffman et al., 2015), and PERSIANN (Hong et al., 2004) have a relatively high spatio-temporal resolution. However, they do not cover regions outside of 60°N/S, and are only available from 2000 onwards, which significantly hinders their use for extreme value analyses. Precipitation products with a true global coverage and long records are reanalyses, such as ERA-5 (Hersbach et al., 2020), JRA-55 (Kobayashi et al., 2015), and MERRA-2 (Gelaro et al., 2017). However, reanalysis products tend to exhibit strong local and systematic biases in the magnitude and frequency of precipitation (Decker et al., 2012; Liu et al., 2018; Ménégoz et al., 2013).
2. Global-scale analyses of precipitation extremes are generally based on daily precipitation records (Cavanaugh et al., 2015; Koutsoyiannis, 2004a,b; Nerantzaki and Papalexiou, 2019; Papalexiou and Koutsoyiannis, 2013; Papalexiou et al., 2013; Ragulina and Reitan, 2017; Serinaldi and Kilsby, 2014). In practice, however, multiple durations are needed for the design of infrastructure (e.g., Nissen and Ulbrich, 2017) or urban drainage networks (e.g., Mailhot and Duchesne, 2009). It is known that precipitation extremes of different durations scale differently with temperature (Wasko et al., 2015; Schleiss, 2018), but little is known about the variation of EVD location, scale and tail parameters for different temporal resolutions. Studies that did derive extreme precipitation statistics for durations ranging from minutes to a few days have mostly focused on small regions (Zhao et al., 2022a,b; McGraw et al., 2019; Nissen and Ulbrich, 2017; Overeem et al., 2008).
3. Studies estimating return levels of extreme precipitation by using annual maxima typically use calendar years to delineate the annual periods from which maxima values are extracted (e.g., De Paola et al., 2018; Marani and Zanetti, 2015; Papalexiou and Koutsoyiannis, 2013; Ragulina and Reitan, 2017; Villarini et al., 2011). When the variable of interest is river discharge instead of precipitation, however, hydrological years are typically used instead of calendar years (Ward et al., 2016). For discharge values this is important, since peak discharge and flooding could occur during the transition from 31 December to 1 January and one event would be included in two calendar years. Although not often considered, this could also happen for precipitation, especially for longer duration extremes. The annual maxima method could pick multiple values from a single rainy season that may, for example, be highly influenced by the El Niño-/Southern Oscillation, which is known to impact precipitation extremes (Allan and Soden, 2008; Rasmusson and Arkin, 1993).
4. The Generalized Extreme Value (GEV) distribution, the most widely used EVD, is typically fitted through one of two approaches: (a) using annual maximum precipitation series and maximum likelihood (Coles, 2001) or L-moment (Hosking, 1990) estimation approaches, or (b) using a Peak-Over-Threshold (POT) method to fit a Generalized Pareto Distribution to excesses above the threshold and a Poisson process to the sequence of threshold exceedances (Coles, 2001). In contrast to GEV and POT, the recently developed Metastatistical Extreme Value (MEV) distribution is fitted using all events with recorded precipitation instead of only the most severe. The inclusion of more events reduces the uncertainty due to sampling effects, which is important when dealing with short time series (Hu et al., 2020; Marani and Ignaccolo, 2015; Marra et al., 2018, 2019a; Miniussi and Marani, 2020; Zorzetto et al., 2016; Zorzetto and Marani, 2019). Additionally, GEV parameter estimation depends heavily on a few large values, which makes it very sensitive to the possible presence of outliers, a relatively common occurrence in remote sensing estimates of precipitation amounts (Zorzetto and Marani, 2020). The GEV tail behavior is controlled by its shape parameter, which is very sensitive to sampling effects and the choice of the method used for estimation. To overcome these problems, some studies have suggested to use one universal value of the shape parameter that is applicable to the whole world Koutsoyiannis (2004a,b), or a shape parameter value within a narrow range between exponential and heavy-tail behavior (Papalexiou and Koutsoyiannis, 2013), or one shape parameter per region, that is similar within climate types and elevation ranges (Ragulina and Reitan, 2017).

In this study we aim to overcome these issues partly by (1) using a precipitation dataset that merges gauge, reanalysis, and satellite data, (2) estimating extremes for several event durations, (3) using hydrological years in our analyses, and (4) comparing results from three different extreme value methods (GEV, POT and MEV). Specifically, we are interested in quantitatively characterizing the behavior of extreme precipitation and the spatiotemporal variation of extreme value distributional tails at the global domain.

2. Material and methods

2.1. Data

The global precipitation product used in this study is the Multi-Source Weighted-Ensemble Precipitation (MSWEP-V2.2) dataset. MSWEP is particularly suited for our purpose due to its global coverage, long temporal span, high spatial and temporal resolution. We used data from 1 January 1979 to 31 October 2017 at a 0.1° latitude ×0.1° longitude resolution at 3-hourly time steps. We selected all land-cells between 90°N and 58°S for our analysis. MSWEP precipitation estimates are derived by merging five different satellite- and reanalysis-based global precipitation datasets. The dataset is one of the few precipitation products with daily (as opposed to monthly) gauge corrections, applied using a scheme that accounts for gauge reporting times (Beck et al., 2019b). MSWEP has shown robust performance compared to other widely used precipitation datasets (e.g., Alijanian et al., 2017; Bai and Liu, 2018; Beck et al., 2017, 2019a; Casson et al., 2018; Hu et al., 2020; Sahlu et al., 2017; Satgé et al., 2019; Zhang et al., 2019), thus underlying its potential for improving the characterization of extreme precipitation worldwide. We refer to Beck et al. (2019b) for a comprehensive description of the dataset.

2.1.1. Quality control

The integration of erroneous gauge observations into MSWEP-V2.2 can occasionally result in implausible precipitation values. Therefore, we implemented a three-step quality control procedure on the 3-hourly data prior to the analysis. We first discarded negative values, which are physically impossible. The second step was to discard outliers, which we defined as values deviating from the mean by more than 30 standard deviations. The number of 3-hourly blocks containing outliers per grid cell are included in Supplementary Material Section 1 Fig. S1. We also discarded data surrounding the outliers for the same time step using a 11×11 grid-cell window, as erroneous gauge observations may have influenced surrounding cells in the production of the MSWEP dataset. The 11×11 grid-cell window was chosen based on the procedure used in MSWEP to merge gauges and other rainfall products (Beck et al., 2019b). In the third step, we removed all years with > 30 discarded days or < 5 'wet' 3-hourly periods, identified using a threshold of 0.2 mm 3 h^{-1} following Wasko et al. (2015). Finally, we only included in the analysis data from grid cells with at least 30 years of data remaining, as a minimum record length of 30 years is customary and recommended for analyzing extremes (Arguez and Vose, 2011; Kendon et al., 2018; Westra et al., 2013).

2.1.2. Durations and identification of independent events

The durations we selected for our analysis are 3, 6, 12 and 24 h, and 2, 3, 5 and 10 days. In order to create statistically-independent precipitation events for multiple durations, we first calculated the running parameter, which is the minimum distance between two independent events (Fukutome et al., 2015). To separate 3-hourly events and ensure independence for each precipitation event at each grid-cell, we followed the declustering method to limit the autocorrelation of the samples of Marra et al. (2018). In order to do so, we calculated the temporal autocorrelation of the time series for each grid-cell, for time lags up to 10 days. 10 days was deemed sufficient to allow the autocorrelation to drop to very low values and remove the correlation between two events. The long-lag noise for each time lag is the 75th quantile of the autocorrelation as in Marra et al. (2018). The resulting running parameter equals the first time lag at which the temporal autocorrelation is comparable to the long-lag noise. The running parameter is calculated for each grid-cell. For the 3-hourly duration, we removed the blocks containing non-zero rainfall within a correlation window and only kept the highest value. For longer durations, independent events satisfy two conditions: (1) events are separated at least by the running parameter (the length of the independence that was determined by declustering); and (2) 3-hourly blocks are only included once. To ensure this, 3-hourly blocks are summed together using moving windows to create intensities for longer durations. Then, one of two cases arises: (a) the running parameter is smaller than the duration: then the event is already independent (condition 1). We simply take the highest intensity in a moving window, and remove all overlapping ones to satisfy condition 2; (b) the running parameter is greater than the duration: we only take the highest value in the window of the running parameter (condition 1), and remove all overlapping blocks (condition 2).

2.1.3. Hydrological year

A common challenge in global-scale assessments is the delineation of the hydrological year, given the regional variability in the climatological precipitation seasonality. We therefore developed an uniform way to define the hydrological year. To avoid splitting one rainy season over two different years, we computed the median of the monthly precipitation for each grid-cell, and defined the start of the hydrological year to be the first day of the driest month. Supplementary Material Section 2 Fig. S2a shows the starting month of the hydrological year as determined by this method. These data are also available in the GPEX dataset (Gründemann et al., 2021). As MSWEP-V2.2 spans the interval from 1 January 1979 to 31 October 2017, we discarded the data prior to the start of the first hydrological year, thus keeping 38 complete years. Only where the hydrological year starts in December there are just 37 complete years, which occurs in 5.8% of the grid cells.

We also investigated whether there is a significant difference between the use of calendar and hydrological years for the estimated daily extremes for GEV and MEV. The POT method is based on the values over a high threshold, irrespective of when they occurred. Therefore, there is by definition no difference in calculating the extremes using hydrological or calendar years for the POT method. To determine the difference for GEV and MEV, we first calculated the daily return levels for normal calendar years, using the MSWEP data from 1979 to 2016. Then, we calculated the return levels for the same distributions and the same years, by removing the months before the start of the hydrological year from the year 1979 and adding them to the year 2016. We did this in order to use the exact same data and ensure that the differences in the return level estimates are solely due to a different starting month.

2.2. Extreme value distributions

Three different extreme value distributions were fitted to the MSWEP data to calculate extreme precipitation return levels, and to provide an indication of dependence of the spread in return levels as a function of the distribution used. These three extreme value distributions are the GEV, POT and MEV. Annual (hydrological year) maxima were used to estimate the three parameters of the GEV using the L-moments approach, because of its robust performance for small samples (Hosking, 1990). The GEV cumulative distribution function (CDF) is given by:

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

with location parameter $\mu \in (-\infty, \infty)$, scale parameter $\sigma > 0$, and shape parameter $\xi \in (-\infty, \infty)$. The annual extremes estimated by GEV are translated into those of the parent distribution, following Koutsoyiannis (2004a, equation 3).

As a second EV model we used a Peaks Over Threshold approach, describing precipitation accumulations exceeding a high threshold using a GP distribution, while modeling the frequency of threshold exceedances using a Poisson point process (Coles, 2001; Davison and Smith, 1990). This framework also yields GEV as the resulting extreme value distribution, which is then used to determine the quantile corresponding to a given return period. The GP CDF is given by:

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

where $y > 0$ are precipitation excesses over the threshold, with $\beta > 0$ and $\xi \in (-\infty, \infty)$ the GP scale and shape parameters respectively. A relevant aspect in applying the POT model is a suitable choice of the threshold used to define precipitation exceedances. Our global-scale application requires studying the distribution of precipitation extremes across markedly different climatic regions, thus excluding the adoption of a constant threshold value. We studied the effect of the threshold choice using multiple selection methods on a global sample of grid cells (see Supplementary Material Section 3 and Fig. S4). Our results showed that the choice of the method had a limited effect on the estimated return levels (Fig. S4a). We therefore chose to perform our global analysis by selecting for each cell a threshold value such that it is exceeded on average 3 times each hydrological year. Because of this choice, the sample size available for fitting the GP distribution remains constant across different precipitation durations. The method used to fit the parameters of the GP distributions was the Probability Weighted Moments (PWM; e.g., see Hosking and Wallis, 1987).

The third model we used is the MEV distribution (Hosseini et al., 2020; Hu et al., 2020; Marani and Ignaccolo, 2015; Miniussi et al.,

2020a,b; Zorzetto et al., 2016). In the MEV framework, all 'ordinary' precipitation events, i.e. all events above a small threshold, are used to infer the EV distribution. The threshold we applied is 0.2 mm 3 h^{-1}, coinciding with the earlier defined 'wet event'. Weibull parameters were estimated for each hydrological year separately, based on all wet events using the PWM method (Greenwood et al., 1979) as done in Zorzetto et al. (2016). The MEV-Weibull CDF is given by:

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

where j is the hydrological year ($j = 1, 2, \ldots, M$), $C_j > 0$ is the Weibull scale parameter, $w_j > 0$ is the Weibull shape parameter, and n_j is the number of wet events observed in hydrological year j (Marani and Ignaccolo, 2015).

2.2.1. Observed return periods

The MSWEP dataset analyzed here has 38 complete years of data. Therefore, the empirical return period associated with the maximum value on record computed according to the Weibull empirical frequency estimate is $T_{\text{observed}} = 39$ years. However, only 91% of all cells had 38 complete years of data, so the maximum observed return period is sometimes lower: for 7% of the cells only 37 complete years were available, and for 2% of the cells 36 years or less were available. However, for simplicity we still refer to the corresponding maximum return level as T39 in the results.

2.2.2. Study areas

In order to compare the three extreme value distributions, we selected fourteen case study areas. They collectively cover a wide range of climates and domain sizes, the locations of which can be found in Fig. 3a. Within a single case study area, we expect the precipitation estimates to be statistically homogeneous because of their precipitation generating mechanisms (Cavanaugh and Gershunov, 2015; Cavanaugh et al., 2015), elevation (Ragulina and Reitan, 2017), or average annual rainfall.

2.2.3. Tail behavior

Both the GEV and MEV distributions are flexible and can describe different tail behaviors. The interpretation of the tail parameter of the two distributions differs, as illustrated in Supplementary Material Section 4 Fig. S5 for different combinations of scale and shape parameters. The shape parameter ξ of the GEV distribution, obtained either through the annual maxima or POT approach, encodes the nature of the tail of the distribution. Based on the value of ξ, the GEV can take one of three forms: a positive GEV shape parameter ($\xi > 0$, *"Fréchet"*) corresponds to a power-law tail, i.e., to a slowly-decaying probability of large events. This heavy-tail behavior contrasts with the case of an exponential tail ($\xi = 0$, *"Gumbel"*), and with the case of a distribution with an upper end point, which corresponds to negative values of the shape parameter ($\xi < 0$, *"inverse Weibull"*).

The MEV distribution assumes that precipitation events are Weibull-distributed. The tail decay of this distribution is controlled by its shape parameter: for $w < 1$ its tail behavior is "sub-exponential", i.e., heavier than that of an exponential (recovered for $w = 1$), albeit with a characteristic scale (Laherrere and Sornette, 1998; Wilson and Toumi, 2005). For $w > 1$ the Weibull tail is super-exponential, with a fast decaying tail, while still retaining an infinite upper end point. Hence, the shape parameter of the Weibull distribution encodes the propensity of a site to be subjected to large extreme events (Wilson and Toumi, 2005; Zorzetto et al., 2016). However, the tail decay of the MEV distribution is not only dependent on that of ordinary values (through w) but is also affected by the yearly number of events (Marra et al., 2018) and by the inter-annual variations of C_j, w_j and n_j.

Whereas the tail behavior of GEV and POT can easily be inferred from the shape parameter of their distributions, for MEV it depends on multiple parameters, making direct comparison between MEV and GEV/POT based on their parameters alone cumbersome. In an effort to nonetheless compare these methods, we have come up with a measure of heaviness that is based on the return levels themselves (Fig. 1). It quantifies how much distributions differ from an exponential one. For an exponential distribution, the 100-year return level (E100) is as follows: $\text{E100} = \text{T1} + b + b$. Where b (blue arrows Fig. 1) is the difference between the 10-year (T10) and 1-year (T1) return level, i.e.: $b = \text{T10} - T1$. For any distribution that differs from a purely exponential one, the difference between the 100-year return level (T100) and the 1-year return level (T1) can be described as:

$$\text{T100} = \text{T1} + b + b + a \quad (4)$$

In this equation a (red arrow in Fig. 1) is the additional increase caused by the heaviness of the tail, $a = \text{T100} - \text{T1} - 2b$ (Fig. 1). For any given extreme value distribution (red dashed line in Fig. 1), a positive a is indicative of heavy tails, and a negative a of thin tails. For pure exponential tails (blue dotted line in Fig. 1) it holds that $a = 0$. The value for a is highly dependent on the local precipitation systems, so we defined the heaviness amplification factor $h_{\text{T1–T10–T100}}$ to be a normalization of a:

$$h_{T1-T10-T100} = \frac{a}{b} = \frac{T100 - T1 - 2\times(T10 - T1)}{T10 - T1} = \frac{T100 - 2\times T10 + T1}{T10 - T1} = \frac{T100 - T10}{T10 - T1} - 1 \quad (5)$$

In words, the meaning of $h_{\text{T1–T10–T100}}$ is the fractional additional increase between T100 and T10 that is more than the increase that could be expected from a pure exponentially tailed distribution. A distribution has a heavy tail for $h > 0$ and a thin tail for $h < 0$. Here, we chose a range for the heaviness metric over return periods from 1 to 100 years, since these return levels are known to be mostly influenced by the underlying data (Rajulapati et al., 2020). Yet, it should be noted that this metric may easily be adjusted to other return periods and other factors between the return periods. For GEV and POT the heaviness metric is independent of the return period range as long as the return periods are a factor 10 apart, because it is solely determined by the shape parameter. Although for MEV this heaviness metric is only valid for the return period range over which it is computed, using other ranges (T2-T20-T200, T5-T50-T500, and T10-T100-T1000) did not yield significant differences (Supplementary Material Section 5+6, Fig. S7+8).

3. Results and discussion

3.1. Hydrological year

Fig. 2 shows the frequency distribution of 1000-year return levels estimated using calendar and hydrological years for GEV and MEV. The spatial distribution of the T1000 differences is presented in Supplementary Material Fig. S2b for GEV and Fig. S2c for MEV, and Fig. S3 presents the frequency distributions of all analyzed return levels. The difference between the return levels estimated using calendar or hydrological years is greatest when the hydrological year starts around April–September, as in the Mediterranean, Middle-East, Southern Africa, Brazil, Indonesia and Western US (see Supplementary Material Fig. S2). This is the case because there are many different events included in hydrological compared to calendar years, resulting in different events and annual maxima and therefore differences in the estimated extremes. For MEV the overall sensitivity in T1000 estimates remains lower than that of GEV, suggesting that regional sensitivity to the definition of block maxima can be quite significant for the GEV approach.

On the other hand, we found that in the case of GEV quantiles the fraction of sites characterized by differences within ±0.5% is larger than for MEV. When the hydrological year starts around November-February, it is only shifted by a few months so the annual maxima mostly stay the same between the calendar and hydrological years. For GEV this means that for many cells there is almost no difference in the T1000 estimates, whereas for MEV the difference is small.

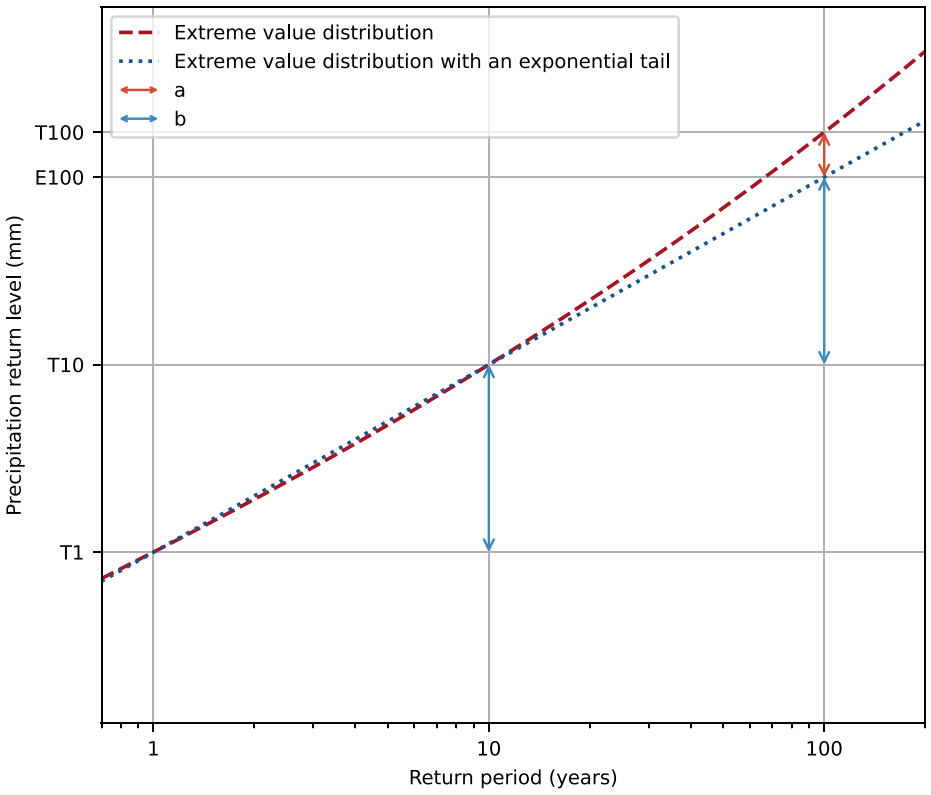

Fig. 1. Illustration of our method to measure the tail heaviness for any distribution based on return levels only. E100 is the 100-year return level if the extreme value distribution has an exponential tail. T100 is the 100-year return level for any extreme value distribution that differs from an exponential one. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

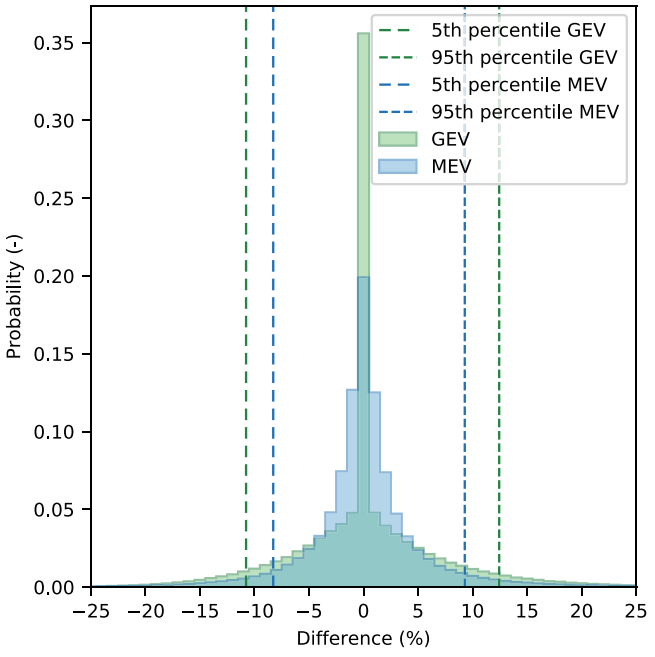

Fig. 2. Weighted histogram showing the percentage difference in the values of T1000 quantiles calculated using calendar years and hydrological years. Included in the figure are all cells where the start of the hydrological year is different than the calendar year (i.e., the hydrological year does not start in January, see Supplementary Material Fig. 2a). A negative difference indicates that the T1000 estimate is larger using hydrological years, whereas a positive difference indicates that the T1000 estimate is larger using calendar years.

3.2. Extreme precipitation estimates

Fig. 3 shows the 100-year precipitation return levels for a 24-hour duration. Extreme value estimates for other durations and return periods are featured in the Global Precipitation EXtremes (GPEX) dataset (Gründemann et al., 2021). The spatial patterns of the extremes estimated by GEV and MEV are similar to Zorzetto and Marani (2020, their Figure 9), while the spatial pattern of the underlying GEV parameters are consistent with Courty et al. (2019, their Figure 1). The global spatial pattern of return levels for the three EV methods is similar, although large regional differences can be observed. The GEV and POT results are similar in magnitude and show similar differences when compared to MEV. The estimated precipitation extremes are generally lower for both GEV and POT compared to MEV quantiles. MEV estimates exhibit smooth spatial patterns, whereas the spatial patterns using GEV and POT are more irregular, consistent with the results of Zorzetto and Marani (2020) for the conterminous US. Furthermore, Fig. 3 reveals the presence of a large number of circular areas with heavier extremes, corresponding to the location of gauges used for correcting precipitation estimates in the MSWEP algorithm (Beck et al., 2019b). The effect of these local corrections is much larger for traditional EV models (POT and GEV) than for MEV. The reduced spatial coherence in patterns of extremes for GEV and POT is particularly evident in the Great Plains of North America, and in Northern Russia, Southeast Asia, and Central Africa.

In order to study the ability of the three distributions to capture the spatial coherence of precipitation extremes, we calculated the coefficient of variation (CV) for fourteen study areas, see Fig. 4. The CV is the ratio of the standard deviation to the mean and is used to compare the relative variation between the study areas. The higher the CV, the higher the relative spread of the precipitation estimates within a spatial domain. This figure shows quite similar behavior for GEV and POT, though POT has a slightly lower spread. The CV for

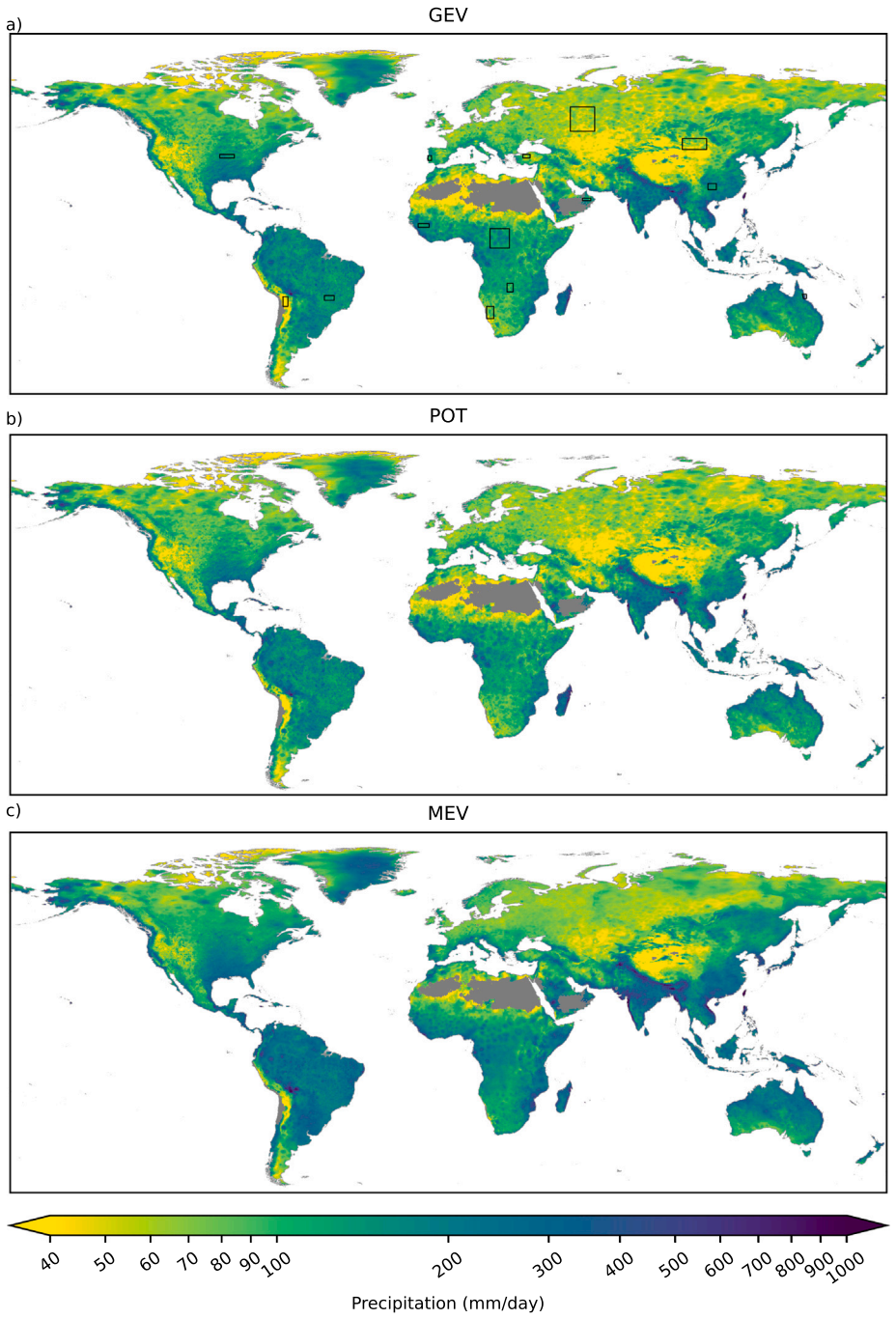

Fig. 3. Precipitation return levels with a duration of 24-hours for a 100-year return period for different extreme value distributions: (a) the Generalized Extreme Value (GEV) distribution, (b) the Peak Over Threshold (POT) method, and (c) the Metastatistical Extreme Value (MEV) distribution. The colorbar has a logarithmic scale, where yellow and purple colors respectively indicate low and very high 100-year precipitation return levels. The black rectangles in panel a are the case studies corresponding to the areas in Fig. 4. Dry areas and cells with too many discarded values are masked in gray following the method in Section 2.1.1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

MEV is generally lower, which points to more spatially coherent T100 precipitation estimates based on single point time series (with 38 years of training data). The CV for all EVDs is higher for areas that are more arid (e.g. Bolivia, Mongolia, Namibia and Oman).

To further investigate the global differences in magnitude between the three methods, we examine the extremes for each distribution using a spatially weighted mean over the global land surface. This is displayed for multiple return periods and durations as depth-duration-frequency curves (Fig. 5). We first compare the maximum precipitation observed (T39 observed, the black dotted line in Fig. 5) in the dataset to the precipitation predicted from each distribution. While locally the empirical T39 estimate could be very different from the true return level, we expect the global average of this value to be representative of the true T39. For GEV and POT, we expected the estimated T39 to be close to the observed value since only the largest values are used to fit these distributions. For MEV, we did not necessarily expect a good agreement for T39, but its performance should be better for return levels greater than the length of the observation time series (Marra et al., 2018, 2019b; Schellander et al., 2019; Zorzetto et al., 2016). The results in Fig. 5 show that for the short duration events, the observed T39 is close to the T39 for all three distributions. For increasing durations, the deviation between empirically observed and EV

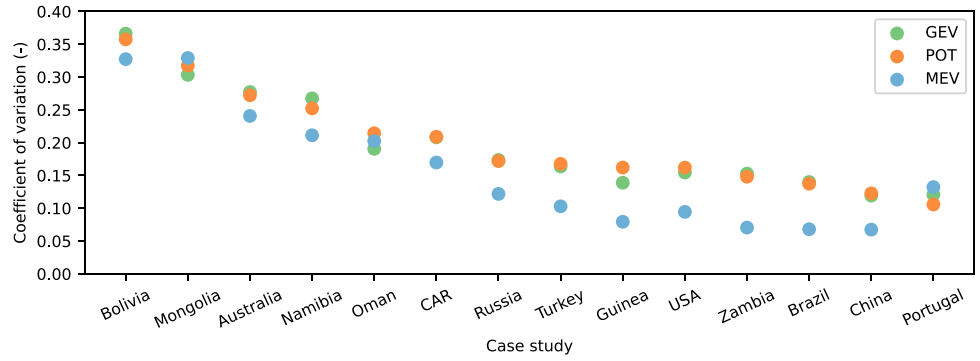

Fig. 4. Coefficient of variation for the difference in estimated T100 quantiles for the three extreme value methods for 24-hour precipitation at selected case study areas. The locations of the case study areas are displayed in Fig. 3a.

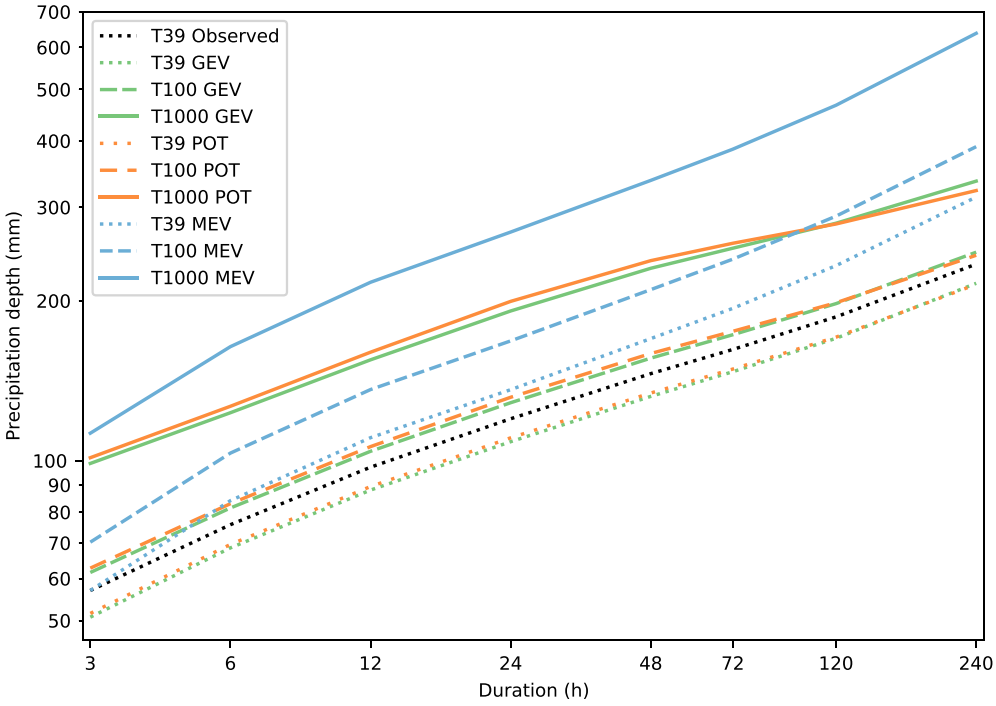

Fig. 5. Area-weighted average depth-duration-frequency curves for the global land surface (log–log scale). T39 Observed is the mean spatially weighted maximum precipitation observed in the MSWEP-V2.2 dataset.

modeled T39 quantiles increases, particularly for MEV. This could be because a smaller number of events per year is used for the fit of MEV-Weibull, whereas the number of events used for the fit of GEV and POT remains constant for all durations. Both GEV and POT show an underestimation and MEV an overestimation. This figure also shows again that the differences between GEV and POT are small. The global average estimated extremes for GEV and POT are notably lower than for MEV, as was already visible from Fig. 3. This difference is more pronounced for larger return periods and longer durations.

One reason the quantiles estimated using MEV are higher than using GEV and POT is related to the increase in estimation uncertainty of Weibull parameters and higher inter-annual variability when the number of events per hydrological year is low (Miniussi and Marani, 2020). This is especially relevant in arid regions and for long durations (Figs. 3 and 5). For instance, for 5 and 10-day durations the average annual number of events is 37 and 22 events respectively. It is therefore possible that this leads to an overestimation by MEV. To overcome this, windows of two or more years could result in a better parameter estimation (Miniussi and Marani, 2020).

3.3. Tail behavior

To better understand the differences between extremes estimated using the three methods, we analyze their tail behavior using the heaviness amplification factor $h_{T1-T10-T100}$ (Eq. (5)). Fig. 6 presents $h_{T1-T10-T100}$ for a 24-hour duration worldwide for each of the three distributions. We refer to Figs, S9-S15 in Section 7 in the supplementary material for maps of $h_{T1-T10-T100}$ for the other durations. Note that as expected the heaviness metric gives a near-identical pattern for both GEV and POT as compared to using their shape parameter directly, but this is less similar for MEV and its yearly mean shape parameter (compare Fig. 6 to Fig. S6).

Both GEV (Fig. 6a) and POT (Fig. 6b) exhibit a large spatial variability in addition to a low spatial coherence. This makes it difficult to discern clear spatial patterns with the exception of a few notable regions. For instance, in the Amazon, $h_{T1-T10-T100}$ is mostly negative, suggesting a tail with an upper limit, while in Eastern and Southern Australia $h_{T1-T10-T100}$ it is strongly positive, denoting strong heavy tail behavior. This map roughly corresponds to the spatial patterns of the GEV shape parameter for daily precipitation shown by Papalexiou

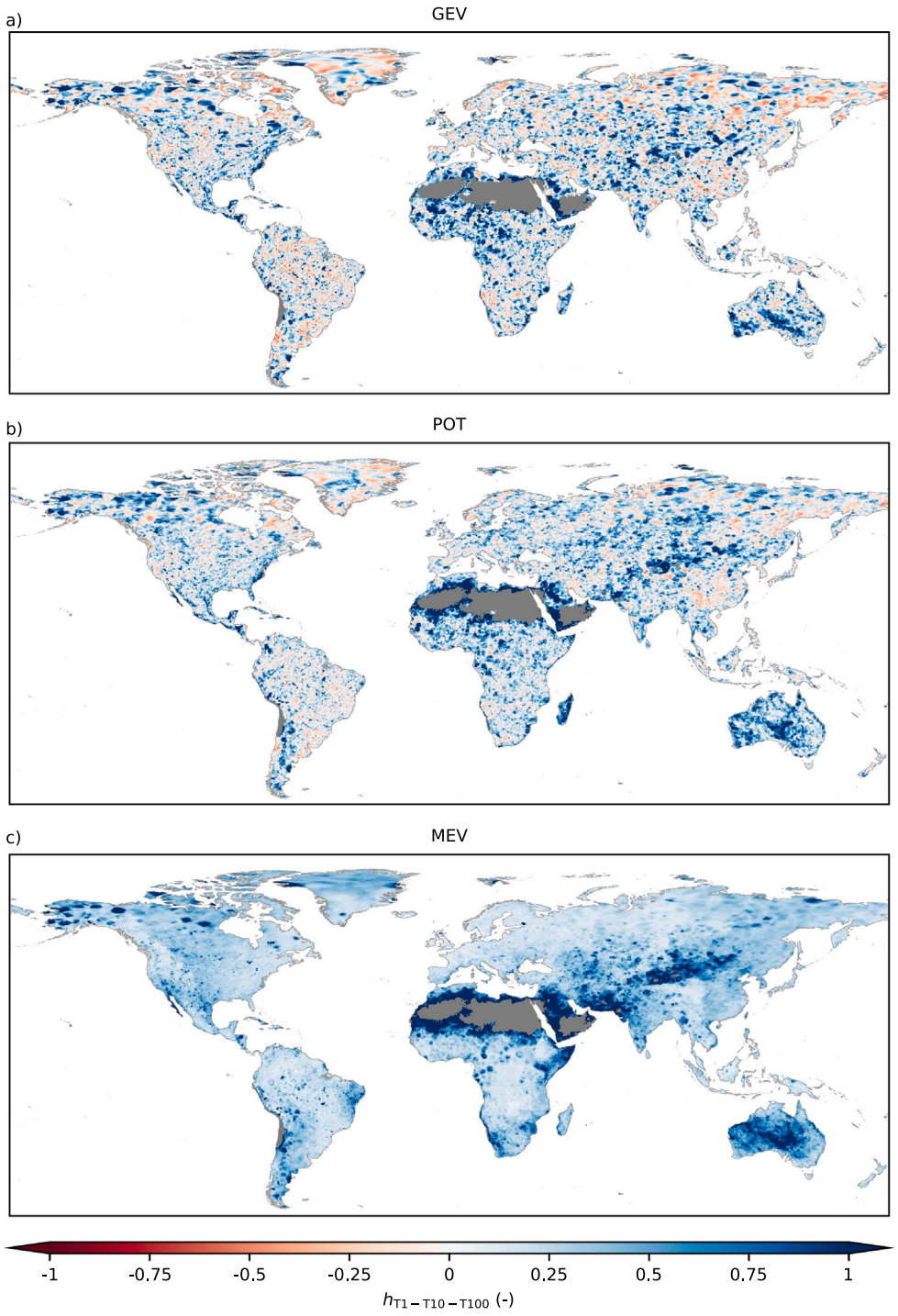

Fig. 6. The heaviness amplification factor $h_{T1-T10-T100}$ (Eq. (5)) for daily precipitation calculated for different extreme value methods: (a) GEV, (b) POT, (c) MEV. Red indicates a thin tail, white an exponential tail, and blue a heavy tail. See Section 2.2.3 for more information on the heaviness metric, and Figures S8-S14 for maps of $h_{T1-T10-T100}$ for the other durations. Dry areas and cells with too many discarded values are masked in gray following the method in Section 2.1.1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and Koutsoyiannis (2013, their Figure 13) and Ragulina and Reitan (2017, their Figure 4) as well as a metric based on the mean excess function Nerantzaki and Papalexiou (2019, their Figure 7a). Given the more similar results between these studies, who used for a large part the same data, we think that the differences with our maps are likely caused by using different underlying data rather than the particular heaviness metric. We also find that for the GEV and POT methods, grid cells associated with heavy tails can be adjacent to cells with thin tails. Furthermore, in 28% of the cells for daily precipitation GEV and POT show a different type of tail, heavy/thin, in the same grid cells. This highlights the large uncertainty associated with estimating reliable tail parameters from short time series and the sensitivity of the GEV and POT methods to sampling effects.

The heaviness of the MEV distribution (Fig. 6c) shows a more coherent spatial pattern. At virtually all grid cells the heaviness amplification factor $h_{T1-T10-T100}$ (Eq. (5)) indicates heavy tail behavior and there is a high consistency within geographic regions and for all durations (Figures S8-S14). Based on previous studies (Cavanaugh et al., 2015;

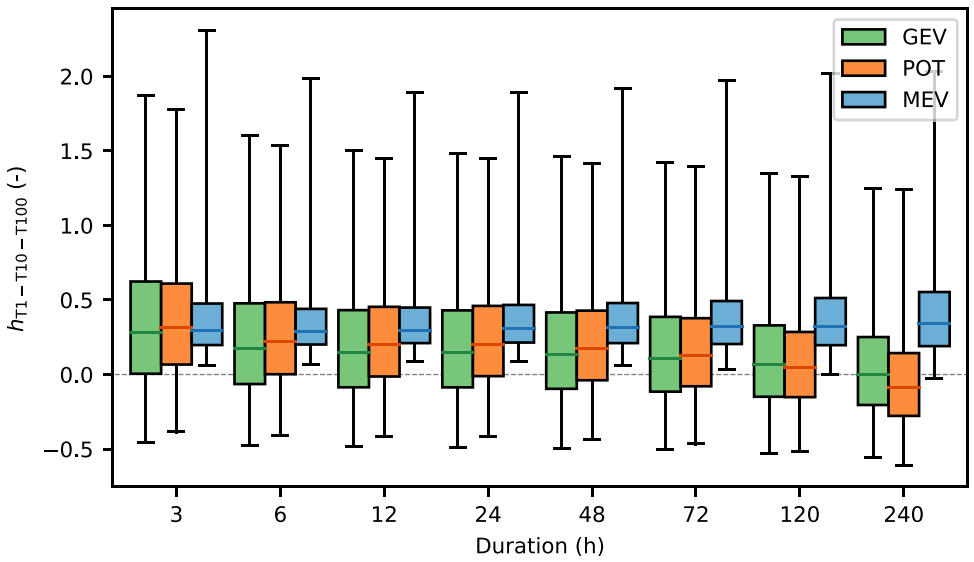

Fig. 7. Boxplot showing the distribution of the heaviness amplification factor $h_{T1-T10-T100}$ (-) for different durations and extreme value methods: GEV, POT and MEV. The whiskers denote the 1st and 99th percentiles. The top and bottom of the boxes represent the 75th and 25th percentiles, respectively. The dashed gray horizontal lines indicate exponential tails. See Section 2.2.3 for more information on the heaviness metric.

Papalexiou and Koutsoyiannis, 2013; Papalexiou et al., 2013; Ragulina and Reitan, 2017; Nerantzaki and Papalexiou, 2019), this predominantly heavy-tail behavior of daily precipitation was expected and is well captured by MEV. There are also topographical patterns visible in the heaviness amplification factor (Fig. 6c), though they are not as clearly distinguishable as for the shape parameter itself (Fig. S6). The heaviness tends to be higher in arid areas, and lower in mountainous areas. Examples of arid areas with high heaviness include the Sahara, the Namib and Kalahari in Africa, the Gobi, Thar and Taklamakan in Asia, the Atacama Desert in South America, large areas of Southwestern Australia, and the Arabian desert and other areas in the Middle East. This same pattern is to a lesser extent also visible for the heaviness of GEV (Fig. 6a) and POT (Fig. 6b).

At high elevations a small $h_{T1-T10-T100}$ is usually found for MEV (Fig. 6c). Examples include the Rocky Mountains and the Sierra Madres in North America, the northern Andes and large areas of the Brazilian Highlands in South America, the Ethiopian Highlands, the Scandinavian Mountains, and the Tibetan Plateau. These spatial patterns are in contrast with what Papalexiou et al. (2018, their Figure 6) found for hourly Weibull tails in the USA, where the heaviest tails are in the mountainous areas, and the thin tails are in the south-east. However, our results correspond well to Ragulina and Reitan (2017, their Figure 4), who showed that heaviness decreases with elevation.

A comparison of the heaviness for different distributions and durations is presented as a boxplot in Fig. 7. For spatial maps of the heaviness for the different durations we refer to Figures S8-S14. For GEV and POT, predominantly heavy tails are observed for short durations and thinner tails for long durations. Furthermore, GEV and POT both show a decreasing variability in the heaviness for longer durations, indicated by both shorter whiskers and boxes. The decrease of the heaviness of the tails for increasing durations is in line with the findings of Cavanaugh and Gershunov (2015), who found that longer duration extremes exhibit thinner tails. For GEV and POT the longer durations largely indicate tails with a finite upper end point, in half of the cases for a duration of 10 days for GEV, and more than half for POT. One implication of this finding is that, when computing return levels for a single location (see Figures S4 and S15), it is possible for the largest return periods that shorter duration quantiles estimate a higher precipitation depth than the longer duration quantiles. This is physically impossible (see Fig. S16a,b,f and g), and we should thus be extremely careful when interpreting such results.

MEV, on the other hand, shows different heaviness patterns than GEV and POT (Fig. 7 and Figures S8-S14). MEV shows almost entirely heavy-tail behavior, which remains consistent across the range of durations examined. Furthermore, also the variability for MEV is constant across durations, though with a slight increase for longer durations. This is in line with Fig. 5, where the difference between the T39 observed and estimated with MEV increases for the longest durations. It could be that MEV is overestimating these return levels of these largest durations. A possible way to improve the estimations would be to group multiple years together, as per Miniussi and Marani (2020).

The MEV distribution produces a spatially and temporally coherent heavy tail behavior based on a 38 years calibration sample and a single grid-cell analysis. This is a promising result, as MEV, in contrast to the traditional methods analyzed, provides a more spatially coherent picture of precipitation extremes without any prior hypothesis on its spatial structure, for example through a spatial clustering scheme (Demirdjian et al., 2018). In fact, the spatial structure of the tail heaviness obtained through the MEV analysis could be used as a measure of statistical homogeneity for regionalization studies.

4. Conclusions

The aim of this research was to quantitatively characterize the spatiotemporal variation of global precipitation extremes and their associated extreme value distribution tails. We have fitted three different extreme value methods (GEV, POT, and MEV) to a global precipitation dataset, MSWEP V2.2, to estimate extreme precipitation return levels for eight durations. In order to compare the tails of the three distributions, we introduced a novel heaviness amplification factor $h_{T1-T10-T100}$ (Eq. (5)). Instead of using calendar years to delineate between different years, we used hydrological years, the start of which we defined as the driest month. We demonstrated that there is a substantial difference in the extremes depending on the definition of yearly blocks used in the extreme value analysis (Fig. 2). Although there is no systematic bias, we still recommend to apply the extreme value analyses for estimating extreme precipitation based on hydrological years in future studies. Our analysis indicates that this can be particularly relevant in the Southern hemisphere and in regions characterized by marked seasonal cycles.

It is well known that the traditional GEV and POT methods require very long data series for accurate estimation of the tail behavior, and our study confirms that there is a low spatial coherence for the tail

properties of both distributions (Fig. 6a and b) using just 38 years of training data. The tail properties of the MEV distribution are spatially more coherent (Fig. 6c) and hence the estimated return levels are more spatially coherent as well (Fig. 3c). This spatially coherent behavior, consistent with previous results obtained over the conterminous US (Zorzetto and Marani, 2020), shows that the MEV distribution is able to capture spatially consistent tail behavior from short time series and by a single grid-cell analysis, without any prior information on the spatial precipitation structures. The analysis of the MEV tail behavior reveals distinct spatial patterns, as the heaviness appears to be controlled by climate zones and orography. Heavier tails are observed in arid areas, and thinner tails in mountainous regions. More in-depth analyses are necessary to draw definite conclusions on what exactly controls the heaviness of extreme value distribution tails. The performance of MEV is promising for regions without long local precipitation records. Furthermore, our study shows that the tail behavior captured by MEV is coherent and heavy both spatially and temporally (Figs. 6, 7 and S8–S14). For GEV and POT, on the other hand, the tail behavior decreases with increasing event duration, resulting in a thin tail with a finite endpoint for about half of the cells for a duration of 10 days. The 'correct' relationship between heaviness and precipitation duration warrants further investigation.

We also conclude that both GEV and POT generally underestimate the observed extremes, whereas MEV overestimates them (Fig. 5). This occurs particularly for long-duration extremes and large return periods. For MEV an explanation could be that there are fewer events per year used for the yearly distribution fits, resulting in larger interannual variability and thus overestimation of extremes. We do consider it likely, however, that the results could be improved, for instance by changing the event threshold or by fitting the Weibull distribution over two or more years for dry areas (Miniussi and Marani, 2020), so as to reduce inter-annual variability of the parameters due to samples of limited length. Our results suggest that this issue is particularly relevant at the longest durations examined. For GEV and POT the spatial consistency of the results could also be improved by adopting spatial extreme models (Davison et al., 2012; Huser and Wadsworth, 2020).

The data generated for this study are openly available as the GPEX dataset (Gründemann et al., 2021). These data include extreme precipitation return levels and extreme value distribution parameters for durations between 3 h and 10 days at a global gridded 0.1° resolution. They could be used by engineers as a reference of precipitation extremes for data-scarce regions in particular. For scientific purposes, all underlying parameters are also available and can be used to answer several outstanding questions, such as: what are the controls on the tail behavior of extremes, and what is driving the different changes in tail heaviness with duration for GEV, POT, and MEV?

Data and code availability

The GPEX dataset is available at the 4TU repository (Gründemann et al., 2021). The data included are the extremes estimated using the different distributions, the observed extremes, and the parameters to estimate the extremes. These data are available for all durations included in this study. The resolution of the dataset is 0.1°, the resolution of the MSWEP-V2.2 dataset. For more information we refer to the Dataset Usage Notes in Section 8 of the supplementary material. The scripts for creating the figures in this manuscript are available at https://doi.org/10.4121/21293760.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data is published as a dataset: https://data.4tu.nl/articles/dataset/GPEX_Global_Precipitation_EXtremes/12764429/4.

Acknowledgments

We acknowledge the use of the `mevpy` Python package (https://github.com/EnricoZorzetto/mevpy) for the extreme value analysis. This work was carried out on the Dutch National e-Infrastructure (DNI) with support from the SURF cooperative. Enrico Zorzetto acknowledges support from the NASA Earth and Space Science Fellowship 80NSSC17K0364. Hylke Beck was supported in part by the U.S. Army Corps of Engineers' International Center for Integrated Water Resources Management (ICIWaRM). Nick van de Giesen acknowledges support of the European Commission's Horizon 2020 Programme under grant agreement number 776691 (TWIGA). Ruud van der Ent acknowledges funding from the Netherlands Organization for Scientific Research (NWO), project number 016.Veni.181.015.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jhydrol.2023.129558.

References

Alijanian, M., Rakhshandehroo, G.R., Mishra, A.K., Dehghani, M., 2017. Evaluation of satellite rainfall climatology using CMORPH, PERSIANN-CDR, PERSIANN, TRMM, MSWEP over Iran. Int. J. Climatol. 37 (14), 4896–4914. http://dx.doi.org/10.1002/joc.5131.

Allan, R.P., Soden, B.J., 2008. Atmospheric warming and the amplification of precipitation extremes. Science 321 (5895), 1481–1484. http://dx.doi.org/10.1126/science.1160787.

Arguez, A., Vose, R.S., 2011. The definition of the standard WMO climate normal: The key to deriving alternative climate normals. Bull. Am. Meteorol. Soc. 92 (6), 699–704. http://dx.doi.org/10.1175/2010BAMS2955.1.

Bai, P., Liu, X., 2018. Evaluation of five satellite-based precipitation products in two gauge-scarce basins on the Tibetan Plateau. Remote Sens. 10 (8), http://dx.doi.org/10.3390/RS10081316.

Ball, J., Babister, M., Nathan, R., Weeks, W., Weinmann, E., Retallick, M., Testoni, I., 2019. Australian Rainfall and Runoff: A Guide to Flood Estimation. Commonwealth of Australia (Geoscience Australia).

Beck, H.E., Pan, M., Roy, T., Weedon, G.P., Pappenberger, F., van Dijk, A.I.J.M., Huffman, G.J., Adler, R.F., Wood, E.F., 2019a. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. Hydrol. Earth Syst. Sci. 23, 207–224. http://dx.doi.org/10.5194/hess-23-207-2019.

Beck, H.E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A.I.M., Weedon, G.P., Brocca, L., Pappenberger, F., Huffman, G.J., Wood, E.F., 2017. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. Hydrol. Earth Syst. Sci. 21 (12), 6201–6217. http://dx.doi.org/10.5194/hess-21-6201-2017.

Beck, H.E., Wood, E.F., Pan, M., Fisher, C.K., Miralles, D.G., van Dijk, A.I.J.M., McVicar, T.R., Adler, R.F., 2019b. MSWEP V2 global 3-hourly 0.1° precipitation: methodology and quantitative assessment. Bull. Am. Meteorol. Soc. 100 (3), 473–500. http://dx.doi.org/10.1175/BAMS-D-17-0138.1.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., Ziese, M., 2013. A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901-present. Earth Syst. Sci. Data 5 (1), 71. http://dx.doi.org/10.5194/essd-5-71-2013.

Beersma, J., Versteeg, R., Hakvoort, H., 2018. Neerslagstatistieken voor korte duren actualisatie 2018. Technical Report, STOWA, Amersfoort, p. 51.

Casson, D.R., Werner, M., Weerts, A., Solomatine, D., 2018. Global re-analysis datasets to improve hydrological assessment and snow water equivalent estimation in a sub-Arctic watershed. Hydrol. Earth Syst. Sci. 22 (9), 4685–4697. http://dx.doi.org/10.5194/hess-22-4685-2018.

Cavanaugh, N.R., Gershunov, A., 2015. Probabilistic tail dependence of intense precipitation on spatiotemporal scale in observations, reanalyses, and GCMs. Clim. Dynam. 45 (11–12), 2965–2975. http://dx.doi.org/10.1007/s00382-015-2517-1.

Cavanaugh, N.R., Gershunov, A., Panorska, A.K., Kozubowski, T.J., 2015. The probability distribution of intense daily precipitation. Geophys. Res. Lett. 42, 1560–1567. http://dx.doi.org/10.1002/2015GL063238.

Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. In: Springer Series in Statistics, Springer, London, UK, p. 208.

Contractor, S., Donat, M., Alexandre, L.V., Ziese, M., Meyer-Christoffer, A., Schneider, U., Rustemeier, E., Becker, A., Durre, I., Vose, R.S., 2020. Rainfall Estimates on a Gridded Network (REGEN)–a global land-based gridded dataset of daily precipitation from 1950 to 2016. Hydrol. Earth Syst. Sci. 24 (2), 919–943. http://dx.doi.org/10.5194/hess-24-919-2020.

Courty, L.G., Wilby, R.L., Hillier, J.K., Slater, L.J., 2019. Intensity-duration-frequency curves at the global scale. Environ. Res. Lett. 14, 084045. http://dx.doi.org/10.1088/1748-9326/ab370a.

CRED, 2019. Natural Disasters 2018. Technical Report, Institute Health and Society UClouvain, Brussels, Belgium, pp. 1–8, URL: https://www.cred.be/sites/default/files/CREDNaturalDisaster2018.pdf.

Davison, A.C., Padoan, S.A., Ribatet, M., et al., 2012. Statistical modeling of spatial extremes. Statist. Sci. 27 (2), 161–186. http://dx.doi.org/10.1214/11-STS376.

Davison, A.C., Smith, R.L., 1990. Models for exceedances over high thresholds. J. R. Stat. Soc. Ser. B Stat. Methodol. 52 (3), 393–442, URL: https://www.jstor.org/stable/2345667.

De Paola, F., Giugni, M., Pugliese, M., Annis, A., Nardi, F., 2018. GEV parameter estimation and stationary vs. non-stationary analysis of extreme rainfall in African test cities. Hydrology 5 (28), http://dx.doi.org/10.3390/hydrology5020028.

Decker, M., Brunke, M.A., Wang, Z., Sakaguchi, K., Zeng, X., Bosilovich, M.G., 2012. Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux tower observations. J. Clim. 25 (6), 1916–1944. http://dx.doi.org/10.1175/JCLI-D-11-00004.1.

Demirdjian, L., Zhou, Y., Huffman, G.J., 2018. Statistical modeling of extreme precipitation with TRMM data. J. Appl. Meteorol. Climatol. 57 (1), 15–30. http://dx.doi.org/10.1175/JAMC-D-17-0023.1.

Dunn, R.J., Alexander, L.V., Donat, M.G., Zhang, X., Bador, M., Herold, N., Lippmann, T., Allan, R., Aguilar, E., Barry, A.A., et al., 2020. Development of an updated global land in situ-based data set of temperature and precipitation extremes: HadEX3. J. Geophys. Res.: Atmos. 125 (16), http://dx.doi.org/10.1029/2019JD032263, e2019JD032263.

Fukutome, S., Liniger, M., Süveges, M., 2015. Automatic threshold and run parameter selection: a climatology for extreme hourly precipitation in Switzerland. Theor. Appl. Climatol. 120 (3), 403–416. http://dx.doi.org/10.1007/s00704-014-1180-5.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., Da Silva, A.M., Wei, G., Kim, G.K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., Zhao, B., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J. Clim. 30 (14), 5419–5454. http://dx.doi.org/10.1175/JCLI-D-16-0758.1.

Greenwood, J.A., Landwehr, J., Matalas, N., Wallis, J., 1979. Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. Water Resour. Res. 15 (5), 1049–1054. http://dx.doi.org/10.1029/WR015i005p01049.

Gründemann, G.J., Werner, M., Veldkamp, T.I.E., 2018. The potential of global reanalysis datasets in identifying flood events in Southern Africa. Hydrol. Earth Syst. Sci. 22 (9), 4667–4683. http://dx.doi.org/10.5194/hess-22-4667-2018.

Gründemann, G.J., Zorzetto, E., Beck, H.E., Schleiss, M., van de Giesen, N., Marani, M., van der Ent, R.J., 2021. Global precipitation EXtremes dataset. http://dx.doi.org/10.4121/uuid:12b5c941-cd54-45db-8d7b-fefaacecaa69.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horanyi, A., Munoz-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., Healy, S., Hogan, R.J., Holm, E., Janiskova, M., Keeley, S., Laloyaux, P., Lopez, P., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thepaut, J.N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 1–51. http://dx.doi.org/10.1002/qj.3803.

Hong, Y., Hsu, K.L., Sorooshian, S., Gao, X., 2004. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. J. Appl. Meteorol. 43 (12), 1834–1853. http://dx.doi.org/10.1175/JAM2173.1.

Hosking, J.R.M., 1990. L-Moments: Analysis and estimation of distributions using linear combinations of order statistics. J. R. Stat. Soc. Ser. B Stat. Methodol. 52 (1), 105–124, http://www.jstor.org/stable/2345653.

Hosking, J.R.M., Wallis, J.R., 1987. Parameter and quantile estimation for the generalized pareto distribution parameter and quantile estimation generalized pareto distribution. Technometrics 23 (3), 339–349. http://dx.doi.org/10.1080/00401706.1987.10488243.

Hosseini, S.R., Scaioni, M., Marani, M., 2020. Extreme atlantic hurricane probability of occurrence through the metastatistical extreme value distribution. Geophys. Res. Lett. 47 (1), 1–9. http://dx.doi.org/10.1029/2019GL086138.

Hu, L., Nikolopoulos, E.I., Marra, F., Morin, E., Marani, M., Anagnostou, E.N., 2020. Evaluation of MEVD-based precipitation frequency analyses from quasi- global precipitation datasets against dense rain gauge networks. J. Hydrol. 590, 125564. http://dx.doi.org/10.1016/j.jhydrol.2020.125564.

Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E.J., Xie, P., 2015. NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG): Algorithm Theoretical Basis Document (ATBD) Version 4.5. Technical Report, NASA/GSFC, Greenbelt, MD 20771, USA, p. 26.

Huser, R., Wadsworth, J.L., 2020. Advances in statistical modeling of spatial extremes. Wiley Interdiscip. Rev. Comput. Stat. e1537. http://dx.doi.org/10.1002/wics.1537.

Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydrometeorol. 5 (3), 487–503. http://dx.doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.

Kendon, E.J., Blenkinsop, S., Fowler, H.J., 2018. When will we detect changes in short-duration precipitation extremes? J. Clim. 31 (7), 2945–2964. http://dx.doi.org/10.1175/JCLI-D-17-0435.1.

Kidd, C., Becker, A., Huffman, G.J., Muller, C.L., Joe, P., Skofronick-Jackson, G., Kirschbaum, D.B., 2017. So, how much of the Earth's surface is covered by rain gauges? Bull. Am. Meteorol. Soc. 98, 69–78. http://dx.doi.org/10.1175/BAMS-D-14-00283.1.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., et al., 2015. The JRA-55 reanalysis: General specifications and basic characteristics. J. Meteorol. Soc. Japan 93 (1), 5–48. http://dx.doi.org/10.2151/jmsj.2015-001.

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

Koutsoyiannis, D., 2004b. Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. Hydrol. Sci. J. 49 (4), 591–610. http://dx.doi.org/10.1623/hysj.49.4.591.54424.

Laherrere, J., Sornette, D., 1998. Stretched exponential distributions in nature and economy:"fat tails" with characteristic scales. Eur. Phys. J. B 2 (4), 525–539. http://dx.doi.org/10.1007/s100510050276.

Liu, Z., Liu, Y., Wang, S., Yang, X., Wang, L., Baig, M.H.A., Chi, W., Wang, Z., 2018. Evaluation of spatial and temporal performances of ERA-interim precipitation and temperature in mainland China. J. Clim. 31 (11), 4347–4365. http://dx.doi.org/10.1175/JCLI-D-17-0212.1.

Mailhot, A., Duchesne, S., 2009. Design criteria of urban drainage infrastructures under climate change. J. Water Resour. Plann. Manag. 136 (2), 201–208. http://dx.doi.org/10.1061/(asce)wr.1943-5452.0000023.

Marani, M., Ignaccolo, M., 2015. A metastatistical approach to rainfall extremes. Adv. Water Resour. 79, 121–126. http://dx.doi.org/10.1016/j.advwatres.2015.03.001.

Marani, M., Zanetti, S., 2015. Long-term oscillations in rainfall extremes in a 268 year daily time series. Water Resour. Res. 51, 639–647. http://dx.doi.org/10.1002/2014WR015885.

Marra, F., Nikolopoulos, E.I., Anagnostou, E.N., Bárdossy, A., Morin, E., 2019a. Precipitation frequency analysis from remotely sensed datasets: A focused review. J. Hydrol. 574, 699–705. http://dx.doi.org/10.1016/j.jhydrol.2019.04.081.

Marra, F., Nikolopoulos, E.I., Anagnostou, E.N., Morin, E., 2018. Metastatistical Extreme Value analysis of hourly rainfall from short records: Estimation of high quantiles and impact of measurement errors. Adv. Water Resour. 117, 27–39. http://dx.doi.org/10.1016/j.advwatres.2018.05.001.

Marra, F., Zoccatelli, D., Armon, M., Morin, E., 2019b. A simplified MEV formulation to model extremes emerging from multiple nonstationary underlying processes. Adv. Water Resour. 127, 280–290. http://dx.doi.org/10.1016/j.advwatres.2019.04.002.

McGraw, D., Nikolopoulos, E.I., Marra, F., Anagnostou, E.N., 2019. Precipitation frequency analyses based on radar estimates: An evaluation over the contiguous United States. J. Hydrol. 573, 299–310. http://dx.doi.org/10.1016/j.jhydrol.2019.03.032.

Ménégoz, M., Gallée, H., Jacobi, H.W., 2013. Precipitation and snow cover in the Himalaya: from reanalysis to regional climate simulations. Hydrol. Earth Syst. Sci. 17 (10), 3921–3936. http://dx.doi.org/10.5194/hess-17-3921-2013.

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

Miniussi, A., Marani, M., Villarini, G., 2020a. Metastatistical Extreme Value Distribution applied to floods across the continental United States. Adv. Water Resour. 136, 103498. http://dx.doi.org/10.1016/j.advwatres.2019.103498.

Miniussi, A., Villarini, G., Marani, M., 2020b. Analyses through the metastatistical extreme value distribution identify contributions of tropical cyclones to rainfall extremes in the eastern United States geophysical research letters. Geophys. Res. Lett. 47, http://dx.doi.org/10.1029/2020GL087238, e2020GL087238.

Mishra, A.K., Coulibaly, P., 2009. Developments in hydrometric network design: A review. Rev. Geophys. 47 (2), http://dx.doi.org/10.1029/2007RG000243.

Nerantzaki, S.D., Papalexiou, S.M., 2019. Advances in Water Resources Tails of extremes: Advancing a graphical method and harnessing big data to assess precipitation extremes. Adv. Water Resour. 134, 103448. http://dx.doi.org/10.1016/j.advwatres.2019.103448.

Nissen, K.M., Ulbrich, U., 2017. Increasing frequencies and changing characteristics of heavy precipitation events threatening infrastructure in Europe under climate change. Nat. Hazards Earth Syst. Sci. 17 (7), 1177–1190. http://dx.doi.org/10.5194/nhess-17-1177-2017.

Overeem, A., Buishand, A., Holleman, I., 2008. Rainfall depth-duration-frequency curves and their uncertainties. J. Hydrol. 348, 124–134. http://dx.doi.org/10.1016/j.jhydrol.2007.09.044.

Papalexiou, S.M., Aghakouchak, A., Foufoula-Georgiou, E., 2018. A diagnostic framework for understanding climatology of tails of hourly precipitation extremes in the United States. Water Resour. Res. 54 (9), 6725–6738. http://dx.doi.org/10.1029/2018WR022732.

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

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

Perica, S., Pavlovic, S., St. Laurent, M., Trypaluk, C., Unruh, D., Martin, D., Wilhite, O., 2015. NOAA Atlas 14 Volume 10, Precipitation-Frequency Atlas of the United States, Northeastern States. Technical Report, 11, NOAA, National Weather Service.

Perica, S., Pavlovic, S., St. Laurent, M., Trypaluk, C., Unruh, D., Wilhite, O., 2018. NOAA Atlas 14: Precipitation-frequency Atlas of the United States, Texas. Technical Report, 11, NOAA, National Weather Service.

Ragulina, G., Reitan, T., 2017. Generalized extreme value shape parameter and its nature for extreme precipitation using long time series and the Bayesian approach precipitation. Hydrol. Sci. J. 62 (6), 863–879. http://dx.doi.org/10.1080/02626667.2016.1260134.

Rajulapati, C.R., Papalexiou, S.M., Clark, M.P., Razavi, S., Tang, G., Pomeroy, J.W., 2020. Assessment of extremes in global precipitation products: How reliable are they? J. Hydrometeorol. 21 (12), 2855–2873. http://dx.doi.org/10.1175/JHM-D-20-0040.1.

Rasmusson, E.M., Arkin, P.A., 1993. A global view of large-scale precipitation variability. J. Clim. 6 (8), 1495–1522. http://dx.doi.org/10.1175/1520-0442(1993)006<1495:AGVOLS>2.0.CO;2.

Sahlu, D., Moges, S.A., Nikolopoulos, E.I., Anagnostou, E.N., Hailu, D., 2017. Evaluation of high-resolution multisatellite and reanalysis rainfall products over East Africa. Adv. Meteorol. 2017, http://dx.doi.org/10.1155/2017/4957960.

Satgé, F., Ruelland, D., Bonnet, M.P., Molina, J., Pillco, R., 2019. Consistency of satellite-based precipitation products in space and over time compared with gauge observations and snow-hydrological modelling in the Lake Titicaca region. Hydrol. Earth Syst. Sci. 23 (1), 595–619. http://dx.doi.org/10.5194/hess-23-595-2019.

Schellander, H., Lieb, A., Hell, T., 2019. Error structure of metastatistical and generalized extreme value distributions for modeling extreme rainfall in Austria. Earth Space Sci. 6 (9), 1616–1632. http://dx.doi.org/10.1029/2019EA000557.

Schleiss, M., 2018. How intermittency affects the rate at which rainfall extremes respond to changes in temperature. Earth System Dyn. 9 (3), 955–968. http://dx.doi.org/10.5194/esd-9-955-2018.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Ziese, M., 2011. GPCC full data reanalysis version 6.0 at 0.5°: monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. http://dx.doi.org/10.5676/DWD_GPCC/FD_M_V6_050.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., Rudolf, B., 2014. GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. Theor. Appl. Climatol. 115 (1–2), 15–40. http://dx.doi.org/10.1007/s00704-013-0860-x.

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

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K., 2018. A review of global precipitation data sets: Data sources, estimation, and intercomparisons. Rev. Geophys. 56 (1), 79–107. http://dx.doi.org/10.1002/2017RG000574.

UNSIDR, 2015. Sendai Framework for Disaster Risk Reduction 2015 - 2030. Technical Report, p. 37, URL: www.unisdr.org/we/inform/publications/43291.

Ushio, T., Sasashige, K., Kubota, T., Shige, S., Okamoto, K., Aonashi, K., Inoue, T., Takahashi, N., Iguchi, T., Kachi, M., Oki, R., Morimoto, T., Kawasaki, Z.I., 2009. A kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. J. Meteorol. Soc. Japan 87 A, 137–151. http://dx.doi.org/10.2151/jmsj.87A.137.

van de Giesen, N., Hut, R., Selker, J., 2014. The trans-african observatory (TAHMO). WIREs Water 1, 341–348. http://dx.doi.org/10.1002/wat2.1034.

Villarini, G., Smith, J.A., Lynn, M., Vitolo, R., Stephenson, D.B., Krajewski, W.F., 2011. On the frequency of heavy rainfall for the Midwest of the United States. J. Hydrol. 400 (1–2), 103–120. http://dx.doi.org/10.1016/j.jhydrol.2011.01.027.

Ward, P.J., Kummu, M., Lall, U., 2016. Flood frequencies and durations and their response to El Niño Southern Oscillation: Global analysis. J. Hydrol. 539, 358–378. http://dx.doi.org/10.1016/j.jhydrol.2016.05.045.

Wasko, C., Sharma, A., Johnson, F., 2015. Does storm duration modulate the extreme precipitation-temperature scaling relationship? Geophys. Res. Lett. 42 (20), 8783–8790. http://dx.doi.org/10.1002/2015GL066274.

Westra, S., Alexander, L.V., Zwiers, F.W., 2013. Global increasing trends in annual maximum daily precipitation. J. Clim. 26 (11), 3904–3918. http://dx.doi.org/10.1175/JCLI-D-12-00502.1.

Wilson, P.S., Toumi, R., 2005. A fundamental probability distribution for heavy rainfall. Geophys. Res. Lett. 32 (14812), http://dx.doi.org/10.1029/2005GL022465.

Zhang, D., Liu, X., Bai, P., Li, X.H., 2019. Suitability of satellite-based precipitation products for water balance simulations using multiple observations in a humid catchment. Remote Sens. 11 (2), http://dx.doi.org/10.3390/rs11020151.

Zhao, W., Kinouchi, T., Ang, R., Zhuang, Q., et al., 2022a. A framework for quantifying climate-informed heavy rainfall change: Implications for adaptation strategies. Sci. Total Environ. 835, 155553. http://dx.doi.org/10.1016/j.scitotenv.2022.155553.

Zhao, W., Kinouchi, T., et al., 2022b. Uncertainty quantification in intensity-duration-frequency curves under climate change: Implications for flood-prone tropical cities. Atmos. Res. 270, 106070. http://dx.doi.org/10.1016/j.atmosres.2022.106070.

Zorzetto, E., Botter, G., Marani, M., 2016. On the emergence of rainfall extremes from ordinary events. Geophys. Res. Lett. 43 (15), 8076–8082. http://dx.doi.org/10.1002/2016GL069445.

Zorzetto, E., Marani, M., 2019. Downscaling of rainfall extremes from satellite observations. Water Resour. Res. 55, 156–174. http://dx.doi.org/10.1029/2018WR022950.

Zorzetto, E., Marani, M., 2020. Extreme value metastatistical analysis of remotely sensed rainfall in ungauged areas: Spatial downscaling and error modelling. Adv. Water Resour. 135, 103483. http://dx.doi.org/10.1016/j.advwatres.2019.103483.