Extreme Precipitation Return Levels for Multiple Durations on a Global Scale

Gaby J. Gründemann[1], Enrico Zorzetto[2], Hylke E. Beck[3], Marc Schleiss[4], Nick van de Giesen[1], Marco Marani[2,5], Ruud J. van der Ent[1,6]

[1]Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, Netherlands
[2]Division of Earth and Ocean Sciences, Duke University, Durham, NC, USA
[3]Department of Civil and Environmental Engineering, Princeton University, Princeton, USA
[4]Department of Geoscience and Remote Sensing, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, Netherlands
[5]Dipartimento di Ingegneria Civile, Edile ed Ambientale, Universita' degli Studi di Padova, Padova, Italy
[6]Department of Physical Geography, Faculty of Geosciences, Utrecht University, Utrecht, Netherlands

Key Points:

- Three methods are used to analyze global precipitation return levels for 3-hourly to 10-day durations
- The Metastatistical Extreme Value distribution shows the most coherent spatiotemporal patterns
- A globally consistent change from heavy to thin tails is observed with increasing duration

Corresponding author: Gaby J. Gründemann, `g.j.gruendemann@tudelft.nl`

Abstract

Quantifying the magnitude and frequency of extreme precipitation events is key in translating climate observations to planning and engineering design. Current efforts have mostly focused on the estimation of daily extremes using gauge observations. Due to the recent development of high-resolution global precipitation products it is now also possible to estimate extremes globally. This research aims to calculate extreme precipitation return levels for multiple durations on a global scale, using the global Multi-Source Weighted-Ensemble Precipitation (MSWEP) V2.2 dataset spanning the period 1979–2017 at 0.1° resolution. Both classical and novel extreme value distributions are used. The application of these different methods at the global scale provides valuable insight into the spatial patterns of extreme precipitation frequency. Our results show that estimates based on the traditional Generalized Extreme Value (GEV) distribution and Peak-Over-Threshold (POT) methods, which only use the largest events to estimate precipitation extremes, exhibit some spatial incoherence. The recently developed Metastatistical Extreme Value (MEV) distribution, on the other hand, includes all precipitation events and leads to smoother and more reliable spatial patterns of local extremes. Furthermore, for all three extreme value methods, a consistent shift from heavy to thin tails with increasing duration is observed and quantified. The generated extreme precipitation return levels and corresponding parameters are released as the Global Precipitation EXtremes (GPEX) dataset. We expect the dataset to be particularly relevant to serve as a benchmark, and to provide extreme precipitation estimates in otherwise data-sparse regions for local engineers and planners.

1 Introduction

Extreme precipitation events are a major contributor to natural disasters (CRED, 2019). Accurate estimates of the severity of such 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. Such spatiotemporal estimates of extreme precipitation, based on extreme value distributions (EVDs), are available, 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 (van de Giesen et al., 2014; Kidd et al., 2017; Gründemann et al., 2018), 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 scale and their scaling with different event durations using reanalysis data and the Generalized Extreme Value (GEV) distribution with fixed tail behavior. Donat et al. (2013) produced the HadEx-2 dataset, which contains 29 generic precipitation and temperature indices, although these indices are not based on EVDs. Furthermore, this dataset has a coarse 2.50° latitudinal × 3.75° longitudinal resolution, with data-gaps due to insufficient available gauge data. Other global studies mostly focused on examining which type of distribution is most suitable to capture the tail behavior of extreme precipitation (Papalexiou et al., 2013; Cavanaugh & Gershunov, 2015; Cavanaugh et al., 2015). In addition, the spatial patterns of the parameter controlling the tail decay have been studied for the GEV distribution (Papalexiou & Koutsoyiannis, 2013; Ragulina & Reitan, 2017), and the Generalized Pareto (GP) distribution (Serinaldi & Kilsby, 2014). However, several issues remain to be addressed in order to obtain a comprehensive global scale comparative benchmark of extreme precipitation return levels: 1) the choice of dataset, 2) the focus on daily durations, 3) the choice of the time blocks over which block-maxima are determined, and

4) the exploration of possible alternatives to the classical EVDs and the associated uncertainty with respect to the tail behavior.

1. Several (quasi-)global gridded precipitation datasets have been developed in the recent years, each with strengths and weaknesses. See Sun et al. (2018) and Beck, Pan, et al. (2019) 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 (Schneider et al., 2014; Kidd et al., 2017), and the number of active gauges has been declining in recent decades (Mishra & 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 systematic biases in the magnitude and frequency of precipitation (Decker et al., 2012; Ménégoz et al., 2013; Liu et al., 2018).
2. Global-scale analyses of precipitation extremes are generally based on daily precipitation records (Koutsoyiannis, 2004b, 2004a; Papalexiou & Koutsoyiannis, 2013; Papalexiou et al., 2013; Serinaldi & Kilsby, 2014; Cavanaugh et al., 2015; Ragulina & Reitan, 2017). In practice, however, multiple durations are needed for the design of infrastructure (e.g., Nissen & Ulbrich, 2017) or urban drainage networks (e.g., Mailhot & Duchesne, 2009). It is known that precipitation extremes of different durations scale differently with temperature (Wasko et al., 2015), but little is known about the variation of EVD properties (tail behavior) 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 (McGraw et al., 2019; Nissen & 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., Villarini et al., 2011; Papalexiou & Koutsoyiannis, 2013; Marani & Zanetti, 2015; Ragulina & Reitan, 2017; De Paola et al., 2018). When the variable of interest is river discharge instead of precipitation, however, hydrological years are typically used instead of calendar years to determine the annual maxima (Ward et al., 2016). For discharge values this is important, since peak discharge and flooding could occur during 31 December to 1 January transition and one event would be included in two calendar years. Although not often considered, this could also happen for precipitation. 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 (Rasmusson & Arkin, 1993; Allan & Soden, 2008).
4. The Generalized Extreme Value distribution (GEV), 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 (Marani &

Ignaccolo, 2015; Zorzetto et al., 2016; Marra et al., 2018; Zorzetto & Marani, 2019; Miniussi & Marani, 2020). This is particularly advantageous when analyzing short remote sensing precipitation products, as the commonly applied GEV requires many years of data to accurately estimate the tail of the distribution (Papalexiou & Koutsoyiannis, 2013). 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 & Marani, 2020). The GEV tail behavior is mostly 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 (2004b, 2004a), or a shape parameter value within a narrow range between exponential and heavy-tail behavior (Papalexiou & Koutsoyiannis, 2013), or one shape parameter per region, that is similar within climate types and elevation ranges (Ragulina & Reitan, 2017). The estimation of the shape parameter is particularly difficult with short data series, though crucial for the accurate estimation of extremes.

In this study we contribute to overcome these issues by 1) using a dataset that merges all three main sources of precipitation 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 exploring the global scale behavior of extreme precipitation and the spatiotemporal variation of extreme value distributional tails. Moreover, this study aims to provide a comparative benchmark of extreme precipitation return levels based on a state-of-the-art global precipitation data (MSWEP) for further research and application.

2 Materials and Methods

2.1 Data

For the present study we selected the global gridded Multi-Source Weighted-Ensemble Precipitation (MSWEP-V2.2) dataset, due to its global coverage, long temporal span (1979–2017), high spatial (0.1°) and temporal (3-hourly) resolution, and robust performance compared to other precipitation datasets in numerous evaluations (e.g., Alijanian et al., 2017; Beck et al., 2017; Sahlu et al., 2017; Bai & Liu, 2018; Casson et al., 2018; Beck, Pan, et al., 2019; Satgé et al., 2019; Zhang et al., 2019). We refer to Beck, Wood, et al. (2019) for a comprehensive description of the dataset. 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. MSWEP was derived by merging five different satellite- and reanalysis-based precipitation datasets globally. The dataset is one of the few with daily (as opposed to monthly) gauge corrections, applied using a scheme that accounts for gauge reporting times (Beck, Wood, et al., 2019). We selected all land-cells between 90°N and 58°S for our analysis.

2.1.1 Quality Control

The integration of erroneous gauge observations into MSWEP-V2.2 has sometimes resulted in implausible precipitation values. Therefore, we implemented a three-step quality control procedure of 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 the values deviating from the mean more than 30 standard deviations. 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 third step was to discard years with $>$ 30 missing days or $<$ 5 'wet' 3-hourly periods, identified using a threshold of

0.2 mm $3h^{-1}$ following Wasko et al. (2015), and to discard cells with < 30 years of data remaining, as a minimum record length of 30 years is customary and recommended to obtain reliable results (Arguez & Vose, 2011; Kendon et al., 2018; Westra et al., 2013).

2.1.2 Durations

We aggregated the 3-hourly data to create longer duration time series. The durations we selected for our analysis are 3, 6, 12, and 24 hours, and 2, 3, 5, and 10 days. To avoid any overlap in the aggregated data, we took the n-hour window that contains the highest precipitation volume among all n-hour running windows that overlap with it, discarding the overlapping windows with lower precipitation volumes.

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. Supplemental Figure S1a shows the starting month of the hydrological year as determined by this method. As MSWEP-V2.2 spans 1 January 1979 to 31 October 2017, we discarded the data prior to the start of the hydrological year and kept the 38 complete years. Only if the hydrological year starts in December there are just 37 complete years, which occurs in 5.8 % of the 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. Second, 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, so the differences in the return level estimates are solely due to a different starting month.

2.2 Extreme Value Distributions

Three extreme value distributions were fitted to the MSWEP data to calculate extreme precipitation return levels: 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:

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

with location parameter $\mu \in (-\infty, \infty)$, scale parameter $\sigma > 0$, and shape parameter $\xi \in (-\infty, \infty)$.

Another common approach for precipitation extremes is the POT method. Instead of annual maxima, values exceeding a high threshold are described using a GP distribution, while the frequency of threshold exceedances is modelled as a Poisson point process (Davison & Smith, 1990; Coles, 2001). 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:

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

where the precipitation excess over threshold $y > 0$, and is defined by scale parameter $\beta > 0$ and shape parameter $\xi \in (-\infty, \infty)$. The POT-threshold could be a fixed precipitation magnitude, or a fixed percentile to account for local variations in precipitation climatologies. As we are interested in studying the distribution of precipitation extremes across many different climate regions, a fixed number is not suitable and the 98th percentile of all 'wet' 3-hourly periods at a given location was chosen as a threshold. For dry areas and long aggregation time periods, the number of observations above the 98th percentile can be very small. For such instances, we lowered the threshold to equal the number of events to the number of years. The method used to fit the parameters of the GP distributions was the Probability Weighted Moments (PWM; Hosking & Wallis, 1987).

The MEV distribution is a recently developed statistical distribution (Marani & Ignaccolo, 2015; Zorzetto et al., 2016; Hosseini et al., 2020; Miniussi, Marani, & Villarini, 2020; Miniussi, Villarini, & Marani, 2020). All "ordinary" (as opposed to just extreme) precipitation events, i.e. all events above a small threshold, are used to infer this extreme value distribution. The threshold we applied is 0.2 mm $3\mathrm{h}^{-1}$, coinciding with the earlier defined threshold for a wet event. Weibull parameters were estimated for each year based on all wet events using the PWM method (Greenwood et al., 1979), equivalent to the use of L-moments, following the same approach as Zorzetto et al. (2016). The MEV-Weibull CDF is given by:

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

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 in hydrological year j (Marani & Ignaccolo, 2015).

2.2.1 Observed Return Period

The MSWEP dataset analyzed here has 38 complete years of data. Therefore, the return period associated with the maximum value on record computed according to the Weibull empirical frequency estimate is $T_{\mathrm{observed}} = 39$ 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 years were available, and for 2 % of the cells 36 years or less were available. However, for simplicity we still refer to this maximum return period as $T39$ in the results.

2.2.2 Tail Behavior

Both the GEV and MEV distributions are flexible and can describe different tail behaviors and can, therefore, be used to investigate how different locations may be more or less sensitive to large extremes. The tail behavior of the two distributions differs, as illustrated in Figure 1 for different combinations of scale and shape parameters. The shape parameter of the GEV distribution, ξ, 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"*),

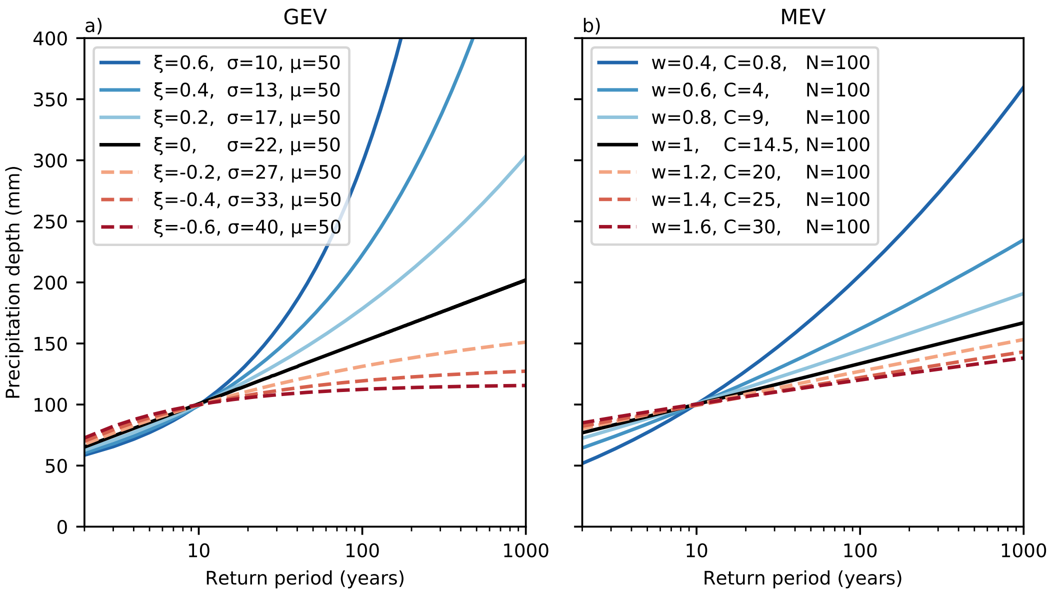

Figure 1. (a) Behavior of the shape (ξ) and scale (σ) parameters of the GEV distribution, with a constant location parameter (μ), and (b) behavior of the shape (w) and scale (C) parameters of the MEV-Weibull distribution, with a constant number of events (N). The results for MEV have been obtained with constant w, C and N parameters for each year. The values of the shape and scale parameter pairs have been chosen such that they all have a precipitation depth of approximately 100 mm for a 10-year return period.

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 (Figure 1b) used here is based on the assumption of a Weibull distribution, which is not heavy-tailed, according to the traditional definition. However, when the shape parameter of the Weibull distribution w is less than one, it does induce a tail behavior which is intermediate between an exponential and a heavy-tail (algebraic decay), also known as sub-exponential, albeit with a characteristic scale (Laherrere & Sornette, 1998; Wilson & Toumi, 2005). When the Weibull shape parameter equals one, the exponential case arises. For a shape larger than one, the distribution becomes hyper-exponential with a fast decreasing tail, while still retaining an infinite upper end point. Hence, the shape parameter of the Weibull distribution can also be used to characterize the propensity of a site to be subjected to large extreme events (Wilson & Toumi, 2005; Zorzetto et al., 2016).

3 Results and Discussion

3.1 Hydrological Year

Figure 2 shows the frequency distribution of the difference between the values of $T1000$ estimated using calendar and hydrological years for GEV and MEV. The figure shows more values centered around $\pm 0.5\,\%$ difference for GEV than for MEV, indicating that for GEV there are more cells that have the same $T1000$ estimate irrespective of which type of years is used. However, the figure also shows a larger spread of the 5th and 95th percentile as well as a wider decay of both tails for GEV compared to MEV. This implies that larger differences between hydrological and calendar years occur for

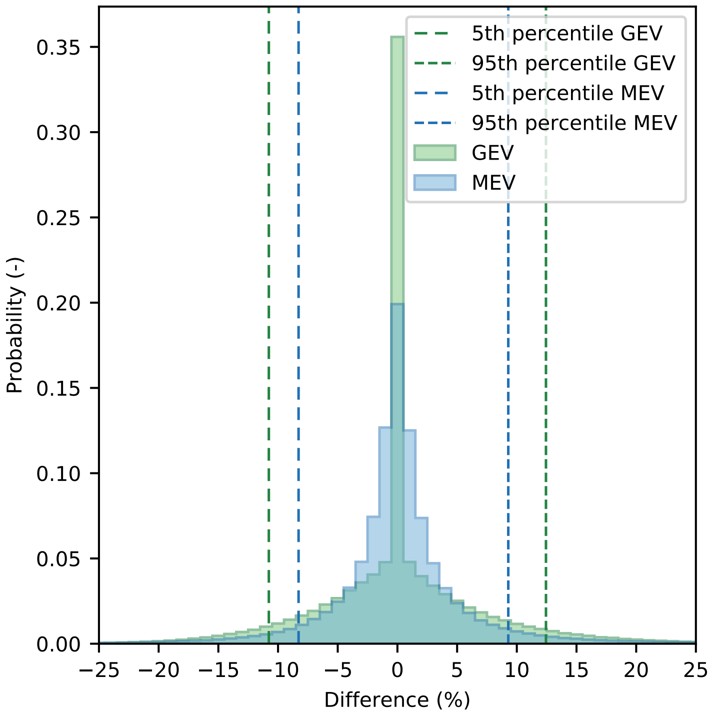

Figure 2. Weighted histogram showing the percentage difference in the values of $T1000$ calculated using calendar years and hydrological years. Included in the figure are all cells where the start of the hydrological year is different than the calendar year (i.e., the hydrological year does not start in January, see Supplement Figure 1a). 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.

GEV. The spatial distribution of these differences is presented in Supplement Figure S1b for GEV and Figure S1c for MEV. The differences are most pronounced in the Southern hemisphere, and at locations where the start of the hydrological year is around June (around the Mediterranean, in the Middle-East, in Brazil, around Indonesia, and in the western US).

When the hydrological year starts around June, the hydrological year is offset by approximately six months compared to a calendar year. As a result, there could be many different events included in the hydrological years compared to the calendar years, resulting in different annual maxima. GEV only uses annual maxima of this relatively short time series, so when there is a large difference in the annual maxima of the hydrological and calendar years, the difference in their estimated extremes are large. When the hydrological year starts in the winter months, the hydrological year is only shifted by a few months. In such instances, the annual maxima mostly stay the same between the calendar and hydrological years. For GEV this means that for many cells there is little difference in the $T1000$ estimates.

MEV, on the other hand, is fitted to all precipitation events in a year. Therefore, even if the hydrological year is only shifted by a couple months, there will be small differences in $T1000$ estimates. However, another implication of MEV being fitted to all precipitation events, is that it is not as sensitive to the inclusion or exclusion of very extreme events. When the hydrological year is offset by approximately six months, there are differences between the hydrological and calendar year estimates, but they are smaller than for GEV.

3.2 Extreme Precipitation Estimates

Figure 3 shows the 100-year precipitation return levels for a 24-hour duration. Other durations and return periods can be viewed by accessing the Global Precipitation EXtremes (GPEX) dataset (Gründemann, 2020). The global 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 overall spatial pattern of return levels of all three 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. MEV exhibits smooth spatial patterns, whereas the spatial patterns using GEV and POT seem more irregular, consistent with the results of Zorzetto and Marani (2020) for the conterminous US. The lack of 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. Furthermore, the figure shows a large number of distinct circular areas with heavier extremes, particularly for GEV and POT, reflecting the locations of gauges used in MSWEP for precipitation correction (Beck, Wood, et al., 2019). MEV has a reduced sensitivity to the presence of outliers and the most extreme events, which may dominate the fitting of GEV.

In order to study the ability of the three distributions to capture the spatial coherence of the estimated extremes, we selected several case studies. They collectively cover several climates, the locations of which can be found in Figure 3a. Within a single case study area, we expect the precipitation estimates to be similar because of their precipitation generating mechanisms (Cavanaugh & Gershunov, 2015; Cavanaugh et al., 2015) or elevation (Ragulina & Reitan, 2017). Figure 4a shows the coefficient of variation (CV) for these case studies. 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 spread of the precipitation estimates within an area. This figure shows quite similar behavior for GEV and POT, though POT is slightly higher overall. The

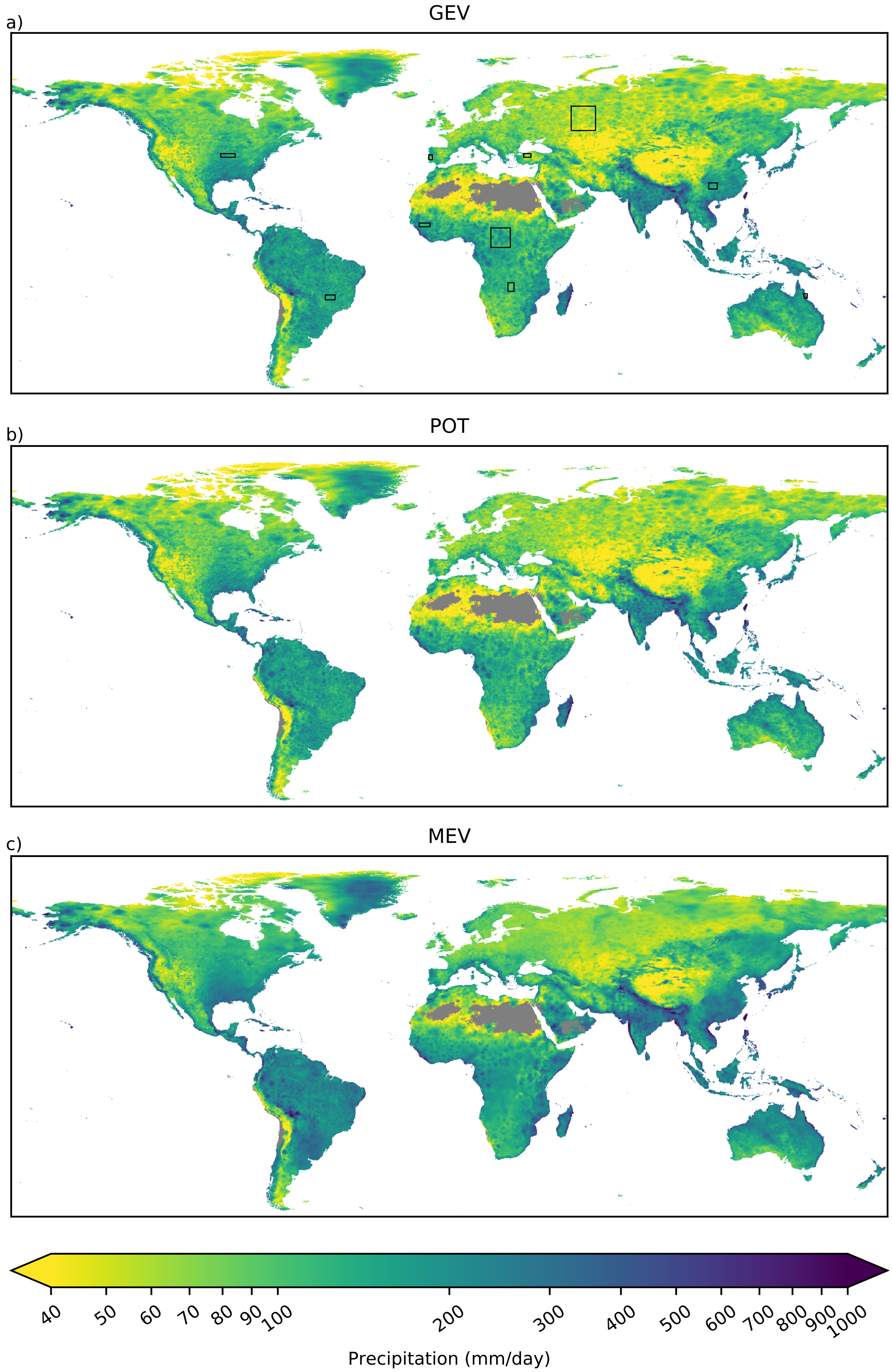

Figure 3. Precipitation return levels with a duration of 24-hours for a 100-year return period for different extreme value distributions: (a) the Generalized Extreme Value (GEV) distribution, (b) the Peak Over Threshold (POT) method, and (c) the Metastatistical Extreme Value (MEV) distribution. The black rectangles in panel a are the case studies corresponding to the areas in Figure 4.

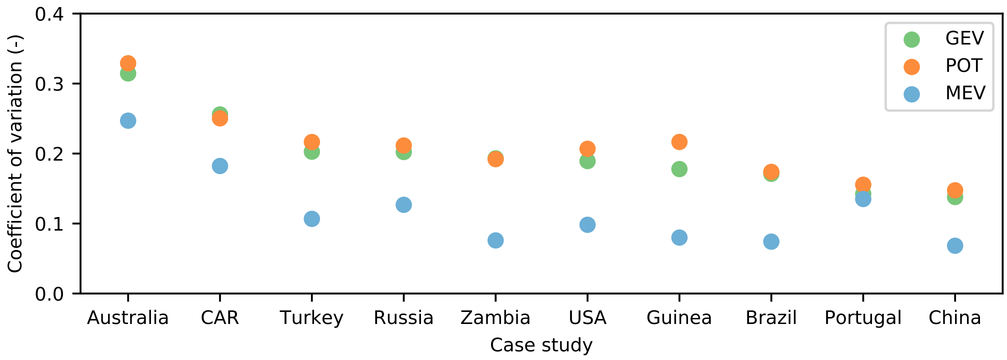

Figure 4. Coefficient of variation for the difference in estimated $T100$ quantiles for the three extreme value methods for 24-hour precipitation at selected case studies. The coefficient of variation is the standard deviation of the precipitation divided by the mean precipitation. The locations of the case studies are displayed in Fig 3a.

CV for MEV is lower. This implies that the internal variability within the study areas is lower using MEV, and that MEV gives a more coherent estimate of the extremes.

To further investigate the 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 several return periods and durations as depth-duration-frequency curves (Figure 5). We first compare the maximum precipitation in the dataset to the predicted precipitation from each distribution. As there are 38 complete years of MSWEP data, the maximum observed return period is 39 years ($T39$ observed, the black dotted line in Figure 5). Doing this for individual grid cells would make little sense as the observed return level may be very different from the true return level, but globally averaged this should be close. For the lower duration events, the observed $T39$ is close to the $T39$ of all three distributions. Both GEV and POT show an underestimation and MEV overestimates in these cases. For increasing durations, the deviation between observed and estimated increases, particularly for MEV. 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 Figure 3. This difference is more pronounced for larger return periods and longer durations.

One reason for the higher estimations of MEV may be that MEV improves significantly for return times that are much larger than the length of the observational time series. Zorzetto et al. (2016) and Marra et al. (2018, 2019) showed that MEV only has a 20 % estimation error for a return period greater than 5 times the length of the time series. Another potential reason for this could be our use of a fixed threshold for defining a precipitation event. Zorzetto et al. (2016), Zorzetto and Marani (2019), and Zorzetto and Marani (2020) use a minimum of 1 mm day^{-1} for a precipitation event, and as we have 3-hour durations instead of solely daily we lowered this threshold. This results in a high number of sub-daily precipitation events per year. We also note that a running parameter based on auto-correlation to distinguish between separate events would have reduced the number of sub-daily events (Marra et al., 2018), but characterizing the temporal correlation of precipitation fields worldwide was outside the scope of our analysis. For a small number of events, occurring for long durations and dry regions, there are few events in each hydrological year to fit the distribution, which results in increased estimation uncertainty. It is, therefore, possible that this leads to an overestimation by MEV for particularly long durations and return periods. To overcome this, windows of two or

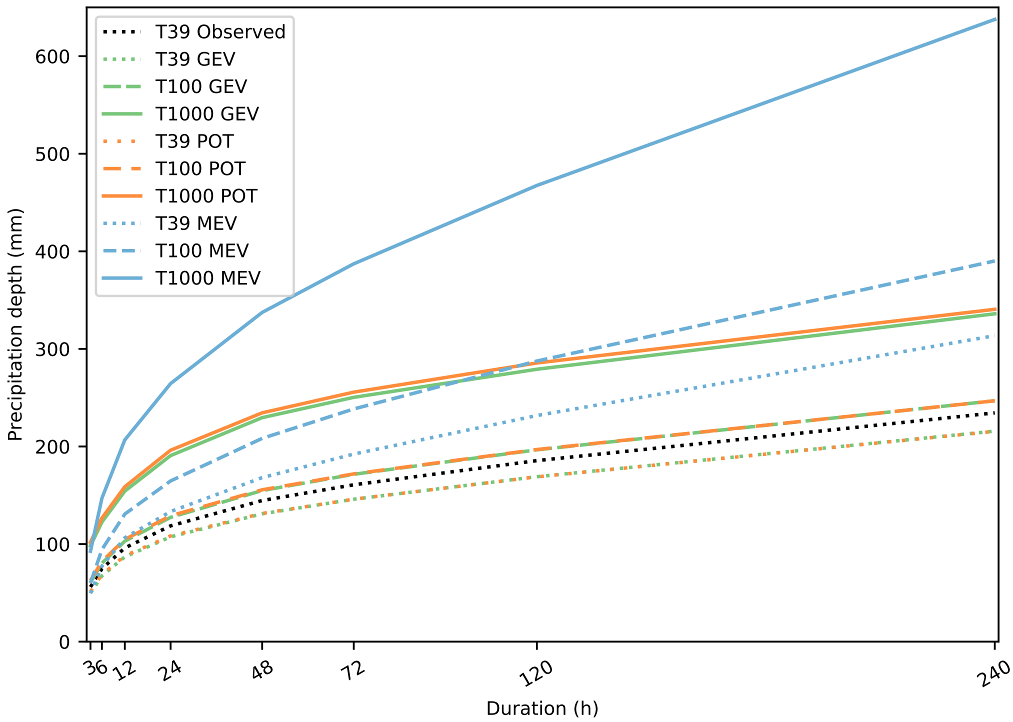

Figure 5. Area-weighted average depth-duration-frequency curves for the global land surface. T39 Observed is the mean spatially weighted maximum precipitation observed in the MSWEP-V2.2 dataset.

more years could result in a better parameter estimation (Miniussi & Marani, 2020). Therefore, the high MEV extreme precipitation estimates are not necessarily a characteristic of the MEV distribution, but rather our application of it.

3.3 Tail Behavior

To better understand the differences between extremes estimated using the three extreme value methods, we analyze their shape parameters in more detail as they provide information on the tail behavior (see section 2.2.2). Figure 6 presents the shape parameter for a 24-hour duration worldwide for each of the three distributions. Both GEV (Figure 6a) and POT (Figure 6b) exhibit a large spatial variability in addition to a low spatial coherence. This makes it difficult to discern spatial patterns, though some can be distinguished. In the Amazon, for instance, the shape parameter is relatively low, suggesting a tail with upper limit, while in Southern Australia it is quite high, denoting heavy tail behavior. This map roughly corresponds to the spatial patterns of the daily GEV shape parameter shown by Papalexiou and Koutsoyiannis (2013, their Figure 13) and GP shape parameter shown by Ragulina and Reitan (2017, their Figure 4), although the actual values differ. We also find that for GEV and POT cells associated with a heavy tail may be adjacent to cells with thin tails. This conflicts with the expected spatial correlation of precipitation regimes, which should lead to spatial coherence (Ragulina & Reitan, 2017). Furthermore, GEV and POT do not always show the same type of tail, heavy or thin, at the same location. In 75 % of the cases the sign of the shape parameter agrees, while in 25 % of the cases the signs are different.

The mean value of the shape parameters underlying the MEV-Weibull distribution (Figure 6c) shows a clear and coherent spatial pattern. At virtually all grid cells the average MEV-Weibull shape parameter indicates a heavy sub-exponential tail (shape parameter less than 1), and the spatial patterns of the shape parameter of MEV-Weibull are much smoother and there is more consistency in different geographic regions. This predominantly heavy-tail behavior for daily precipitation was expected based on previous studies (Papalexiou & Koutsoyiannis, 2013; Papalexiou et al., 2013; Cavanaugh et al., 2015; Ragulina & Reitan, 2017), and is well captured by MEV. There are also clear topographical patterns visible, and Figure 6c has many similarities to a global elevation map. The mean shape values indicate thinner tails at higher elevations, such as the Rocky Mountains and Sierra Madres in North America, the Andes and large areas of the Brazilian Highlands in South America, the Ethiopian Highlands, the Scandinavian Mountains, and the Ural and Tibetan Plateau in Asia. 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 southeast. However, our results correspond well to Ragulina and Reitan (2017, their Figure 4), who showed that the GP shape parameter decreases with elevation.

We find, however, that the relationship of the Weibull shape parameter with elevation is more complicated. Heavier tails are generally observed on the leeward side of large mountain ranges, and thinner tails on the windward side that is dominated by orographically enhanced frontal precipitation (Figure 6c). This is for example visible in the Rocky Mountains, Indonesia, and Norway, and corresponds to findings of Cavanaugh and Gershunov (2015, their Figure 5), who showed that exponential tails are observed in regions where extreme precipitation is predominantly generated by one type of system. There are also some similarities with the main climate types of the Köppen-Geiger classification (Peel et al., 2007; Beck et al., 2018). We observe thinner tails in Western Europe, which has a warm temperate climate. Heavier tails are located in both hot and dry Southern Europe and cold and snowy regions as parts of the Alps. Interestingly, in the USA this pattern is almost reversed, where the heavier tails are in the arid areas of the Great Plains and in the warm temperate climates in the southeast, whereas thinner tails are predominantly observed in the colder mountainous regions.

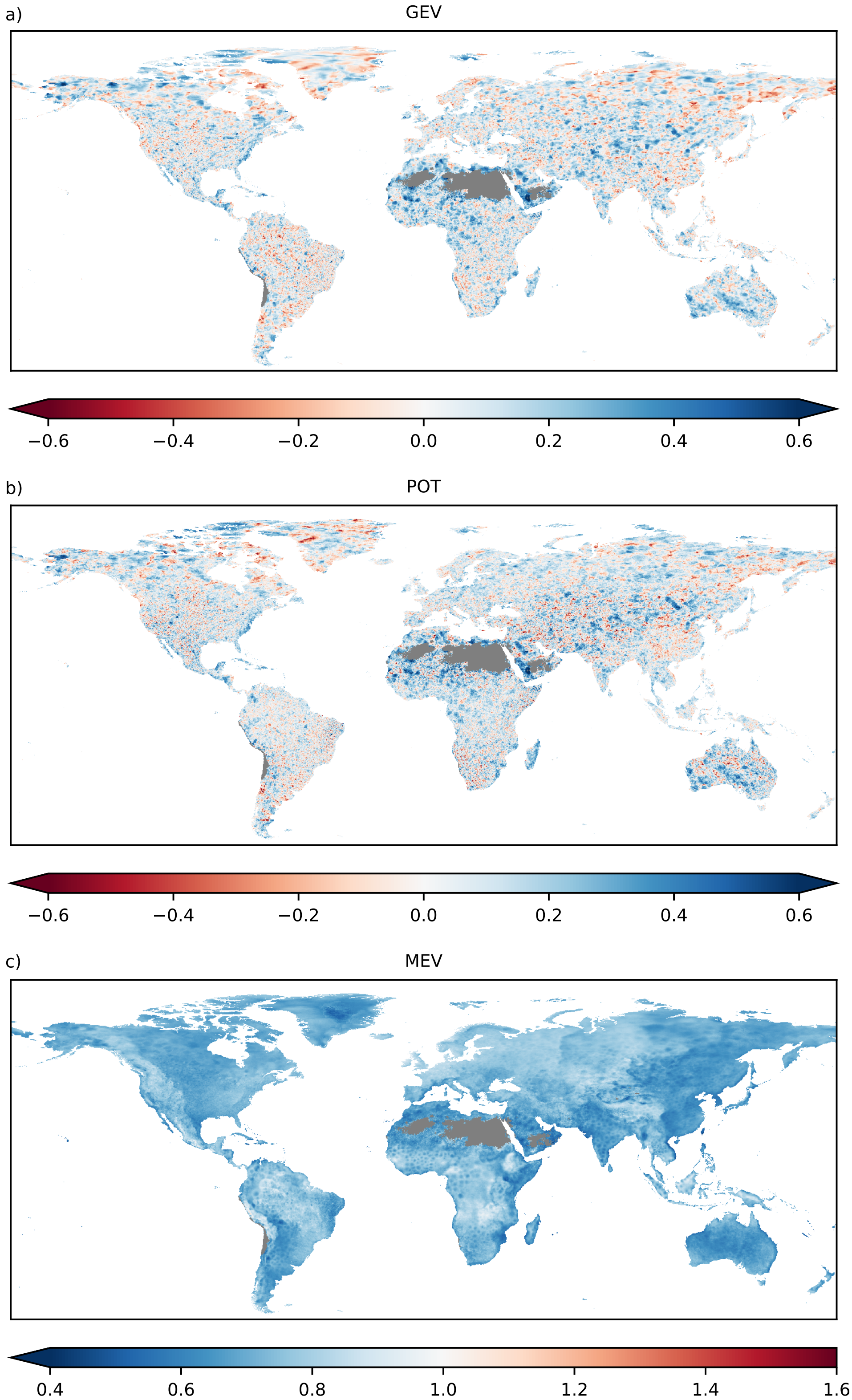

Figure 6. The shape parameter (−) for daily precipitation calculated for different extreme value methods: (a) GEV, equation 1 — ξ_{GEV}, (b) POT, equation 2 — ξ_{GP} and (c) MEV, equation 3 — w. For MEV, the mean shape parameter of all yearly Weibull distributions is displayed. Red indicates a thin tail (for GEV and POT with an upper limit, for MEV hyper-exponential), white an exponential tail, and blue a heavy tail (power-law for GEV and POT, sub-exponential for MEV, see section 2.2.2 and Figure 1).

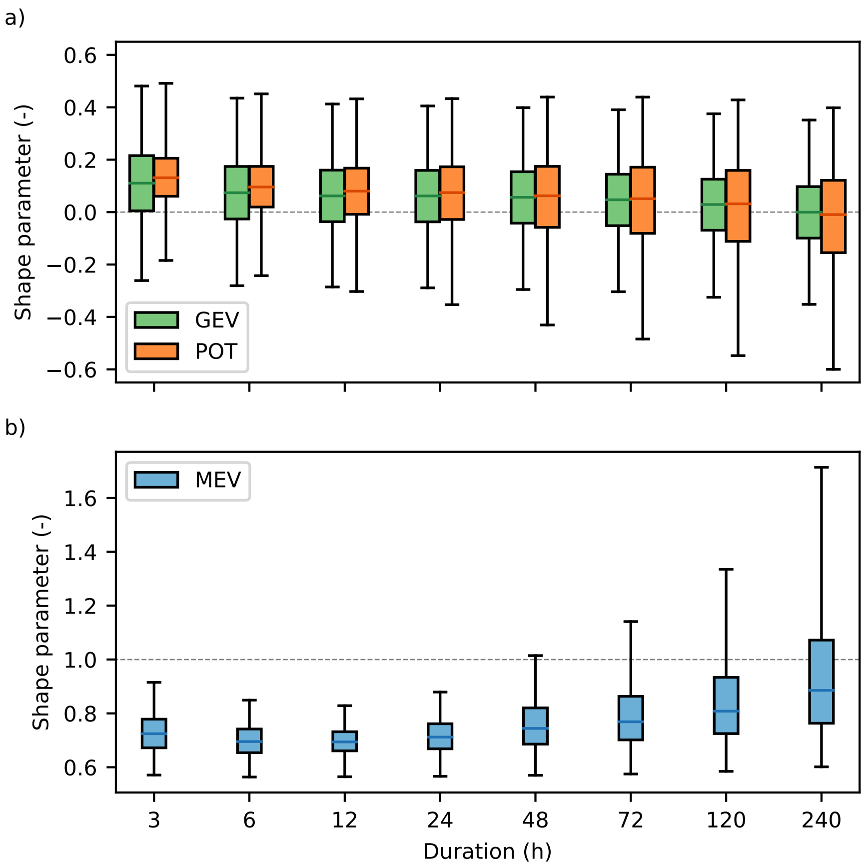

Figure 7. Boxplots showing the distribution of the shape parameter for different durations and extreme value methods: (a) GEV and POT, and (b) MEV. The whiskers denote the 1st and 99th percentiles. The top and bottom of the boxes represent the 75th and 25th percentiles, respectively. The dashed gray horizontal lines indicate exponential tails. Section 2.2.2 and Figure 1 explain the effect of the shape parameter on the tail behavior.

This demonstrates that the MEV method is able to produce spatially coherent tail behavior of precipitation extremes using a time series of just 38 years. This smoothness of MEV was obtained by doing a single-cell analysis, as opposed to the need to identify homogeneous regions using a clustering scheme (Demirdjian et al., 2018). This is a promising result as MEV can thus be applied on a single grid-cell-basis using much shorter time series than GEV, enabling extreme value analysis in those regions of the world where only short precipitation time series are available.

A comparison of shape parameters for different durations is presented in Figure 7. Heavier tails are observed for short durations, and thinner tails for long durations. This 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 an upper limit. This occurs for instance in half of the cases for a duration of 10 days. MEV shows some thin tails for long durations, but not as many as for GEV and POT. Furthermore, there are clear patterns for the variability of the shape parameter. POT shows an increasing spatial variability of the shape for an increasing duration, indicated by longer whiskers. For GEV, on the other hand, the variability remains

Table 1. Variables included in the GPEX dataset.

Variable	Description
GEV estimate	Extreme precipitation return levels estimated using GEV (mm)
POT estimate	Extreme precipitation return levels estimated using POT (mm)
MEV estimate	Extreme precipitation return levels estimated using MEV (mm)
Observed estimate	Observed extreme precipitation return levels (mm)
GEV location parameter	Location parameter of the GEV distribution
GEV scale parameter	Scale parameter of the GEV distribution
GEV shape parameter	Shape parameter of the GEV distribution
POT location parameter	Location parameter for a GEV distribution estimated by fitting the GP distribution
POT scale parameter	Scale parameter for a GEV distribution estimated by fitting the GP distribution
POT shape parameter	Shape parameter for a GEV distribution estimated by fitting the GP distribution
MEV scale parameter	Scale parameter of the MEV distribution for each hydrological year
MEV shape parameter	Shape parameter of the MEV distribution for each hydrological year
MEV number of events	Number of events per hydrological year, n parameter of the MEV distribution
Annual maxima	Annual maximum precipitation for each hydrological year (mm)
Start hydrological year	Number indicating the month in which the hydrological year starts
Land mask	Mask used for this study to indicate land cells and ocean cells

constant for different durations. For MEV, an increasing duration shows first a slight decrease of the variability with a heavier tails, and then after 12 hours a steady increase of the spread with thinner tails.

4 Dataset Usage Notes

The GPEX dataset created in this study is available at the 4TU repository (Gründemann, 2020). It provides openly accessible and readily available hydrologically relevant return levels of extreme precipitation estimates worldwide. It contains the precipitation estimates of the three extreme values distributions, the observed estimates, the parameters of the three distributions, as well as a few other variables used in this study (Table 1). In this section we provide some possible uses of the dataset, and instructions and disclaimers for proper use, both for large or regional-scale usage as well as for a single cell or point-scale.

4.1 Large-Scale Applications

The GPEX dataset contains global scale extreme precipitation estimates and its parameters at a spatial resolution of 0.1°, covering 3-hourly to 10-day durations. The dataset contains information about precipitation extremes for the entire Earth's land surface except Antarctica. The estimates of three distributions as well as the return levels as observed are included in the dataset. Among the three distributions, the traditional GEV and POT provide comparable large-scale average extremes, although differences can be substantial at smaller scales. When using the dataset at regional scales, we advise taking the average of the precipitation estimates, as neighboring cells could differ. Note that since only 38 years of data were available, the fitted model parameters and associated return values are subject to considerable uncertainty. Furthermore, we acknowledge that the use of just one dataset does not represent the true uncertainty in the generation of the dataset created. We do not think this affects our results for observed global spatial patterns significantly, but in a practical setting we recommend verifying the estimates with local observations if available, and to reproduce the precipitation return level estimates with a full uncertainty range estimation.

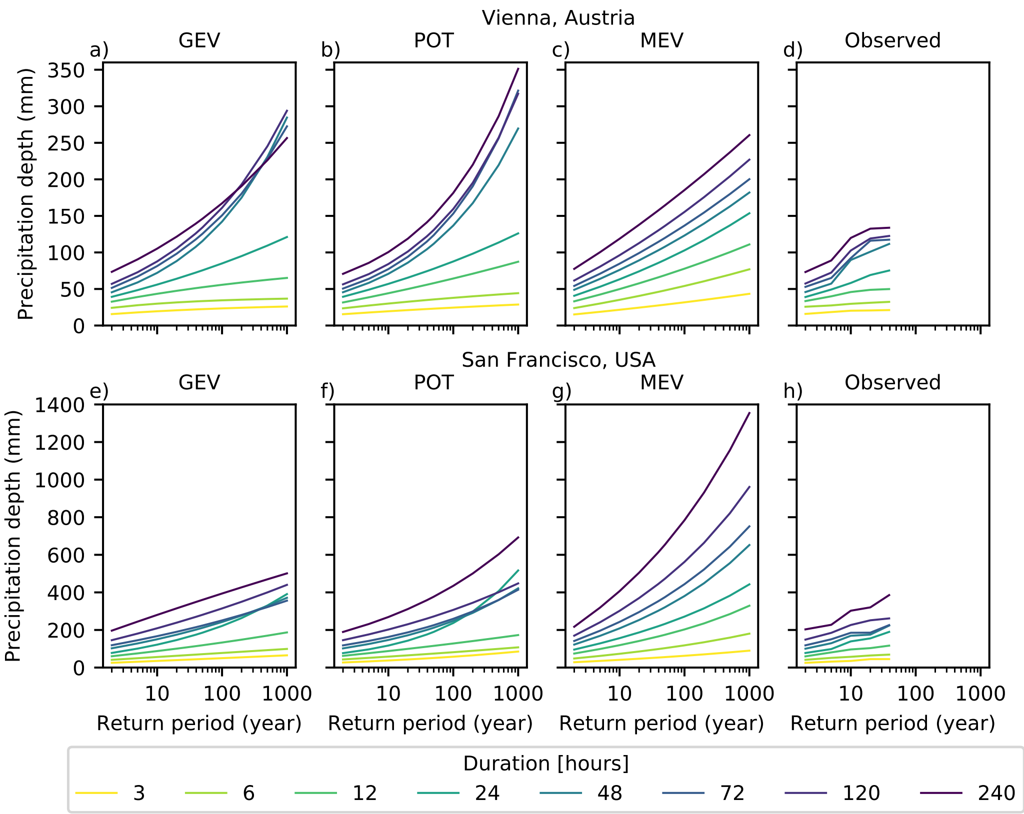

Figure 8. Return level plots for specific locations and different distributions. a–d) Vienna, Austria (48.234°N, 16.415°E) and e–g) San Francisco, California, USA (37.784°N, 122.400°W). Observed (d & h) are the annual maxima converted to return periods.

The novel MEV distribution provides more spatially coherent patterns of the extremes. Its mean shape parameter for daily events captures the (heavy-)tail behavior, and follows orographic patterns. The extremes estimated by MEV are higher than those estimated by GEV and POT. However, for large return periods and long durations, MEV can overestimate the extremes, due to the small number of events available for the fitting. We, therefore, recommend analyzing the extremes of all three distributions to obtain an indication of the uncertainty.

4.2 Small-Scale Applications

The dataset is also suitable for small-scale applications either in comparative studies or for direct use in data sparse regions, but one should be aware of the different statistical characteristics of point-scale and grid-scale. Due to averaging effects in gridded datasets, precipitation extremes of point-scale observations are higher (Sivapalan & Blöschl, 1998; De Michele et al., 2001; Ensor & Robeson, 2008; Cavanaugh & Gershunov, 2015; Zorzetto & Marani, 2019). Illustrative examples of two locations, Vienna and San Francisco, are included in Figure 8. Analysis of the return level plots shows the estimates of the three distributions compared to the observed ones. We converted the annual maximum precipitation to 'observed' return levels (Figure 8d+h). It should be kept in mind though that these 'observed' return levels are also different from the 'true' return levels. For (sub-)daily durations and low return periods, there is generally a good agreement between the observed return levels and the estimates of the three EVDs. For longer durations and return periods, however, the estimated extremes deviate from the observed extremes. This is seen in San Fransisco (Figure 8e-h) where MEV overestimates and GEV and POT underestimate the extremes.

Furthermore, increasing event durations result in lower shape parameters, which was seen for all three distributions (Figure 7). An implication of this is that for long-durations the shape parameter indicates a finite endpoint (GEV and POT), or a very thin tail (MEV), while heavier tails are generally observed for short-durations. When estimating very large return periods (e.g., $T500$), it is therefore possible for shorter duration estimates to be more intense than the corresponding quantiles computed for longer durations, which is physically impossible (see also Figure 8a,b,e and f). We should thus be careful when calculating very large return periods for multiple durations based on short records and interpreting the estimated extremes.

To get a better understanding of the range and uncertainty of a single cell location, we recommend to look at return level plots of the three distributions at the cell of interest in combination with its neighboring cells. This is particularly important for GEV and POT, due to the absence of coherent spatial patterns and the erratic manifestation of the tail behaviors. Previous results (Zorzetto et al., 2016) show that the benefits of MEV over GEV are greater for large return periods relative to the sample size available for the fit. Hence, for the estimation of large quantiles, MEV may be presumed to be more accurate. Depending on the practical application one could then choose to use the most extreme value, use the MEV value, or a spatial average of the GEV and POT estimates.

5 Conclusions

In this study 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 several durations. The estimated precipitation extremes for the three approaches as well as their parameters have been published in the openly available GPEX dataset (Gründemann, 2020). Instead of using calendar years to delineate between different years, we used hydrological years. We demonstrated that there is a substantial difference in the extremes depending on the definition of yearly blocks used in the extreme value analysis (Figure 2). These differences were most notable in the Southern hemisphere, and in locations where the driest month is around June (Figure S1). 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 issue 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 dataseries for accurate estimation of the tail parameter, and our study confirms that there is a low spatial coherence for the tail properties of both distributions (Figure 6a and b). The tail properties of the MEV distribution are spatially more coherent (Figure 6c) and hence the estimated return levels are spatially coherent as well (Figure 3c). In particular, the analysis of the MEV-Weibull shape parameter reveals distinct spatial patterns. The shape parameter appears to be significantly controlled by orographic patterns, with the Weibull tail becoming thinner at higher elevations, as well as on the windward side of large mountain ranges. This behavior, consistent with previous results obtained over the conterminous US (Zorzetto & Marani, 2020), shows that the MEV distribution is able to capture spatially consistent tail behavior from short time series, a promising result for regions without long local precipitation records. Furthermore, our study shows that for all three distributions, the tail behavior decreases with increased event duration (Figure 7). For GEV and POT for instance, about half of the cells have a thin tail with a finite endpoint for a duration of 10 days. We also conclude that both GEV and POT generally underestimate the observed extremes, whereas MEV overestimates them (Figure 5). This occurs particularly for long-duration extremes and large return periods. We do consider it likely, however, that the results of the MEV estimation could be improved, for instance by changing the event threshold, by using a running parameter to separate events (Marra et al., 2018), or by fitting the Weibull distribution over two or more years for dry areas and long durations.

The global gridded 0.1° GPEX dataset we created covers extreme precipitation return levels and its parameters for durations between 3 hours and 10 days. The dataset is freely available and could be used as a benchmark of precipitation extremes for data-scarce regions in particular. The dataset can be used for a single location, to obtain return level plots for planning and engineering design. For scientific purposes, not only the extreme precipitation estimates can be useful, but also the underlying parameters, allowing a more in-depth analysis of the local or regional circumstances. GPEX thus serves as a benchmark of global precipitation extremes with applications in global- and local-scale research as well as a reference for engineering practitioners in data sparse environments.

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 19-EARTH19R-27. Hylke Beck was supported in part by the U.S. Army Corps of Engineers' International Center for Integrated Water Resources Management (ICIWaRM). Nick van der 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.

Upon acceptance the GPEX dataset will be available at the 4TU repository (Gründemann, 2020). 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.

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. *International Journal of Climatology*, *37*(14), 4896–4914. doi: 10.1002/joc.5131

Allan, R. P., & Soden, B. J. (2008). Atmospheric warming and the amplification of precipitation extremes. *Science*, *321*(5895), 1481–1484. doi: 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. *Bulletin of the American Meteorological Society*, *92*(6), 699–704. doi: 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 Sensing*, *10*(8). doi: 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., … Wood, E. F. (2019). Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrology and Earth System Sciences*, *23*, 207–224. doi: 10.5194/hess-23-207-2019

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. M., Weedon, G. P., … Wood, E. F. (2017). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, *21*(12), 6201–6217. doi: 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., … Adler, R. F. (2019). MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bulletin of the American Meteorological Society*, *100*(3), 473–500. doi: 10.1175/BAMS-D-17-0138.1

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future köppen-geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*(180214). doi: 10.1038/sdata.2018.214

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 System Science Data*, *5*(1), 71. doi: 10.5194/essd-5-71-2013

Beersma, J., Versteeg, R., & Hakvoort, H. (2018). *Neerslagstatistieken voor korte duren actualisatie 2018* (Tech. Rep.). Amersfoort: STOWA.

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. *Hydrology and Earth System Sciences*, *22*(9), 4685–4697. doi: 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. *Climate Dynamics*, *45*(11-12), 2965–2975. doi: 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. *Geophysical Research Letters*, *42*, 1560–1567. doi: 10.1002/2015GL063238

Coles, S. (2001). *An introduction to statistical modeling of extreme values.* Springer Series in Statistics; Springer: London, UK.

Contractor, S., Donat, M., Alexandre, L. V., Ziese, M., Meyer-Christoffer, A., Schneider, U., … Vose, R. S. (2020). Rainfall estimates on a gridded network (regen)–a global land-based gridded dataset of daily precipitation from 1950 to 2016. *Hydrology and Earth System Sciences*, *24*(2), 919–943. doi: 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. *Environmental Research Letters*, *14*, 084045. doi: 10.1088/1748-9326/ab370a

CRED. (2019). *Natural disasters 2018* (Tech. Rep.). Brussels, Belgium: Institute Health and Society UClouvain. Retrieved from `https://www.cred.be/sites/default/files/CREDNaturalDisaster2018.pdf`

Davison, A. C., & Smith, R. L. (1990). Models for Exceedances over High Thresholds. *Journal of the Royal Statistical Society. Series B (Methodological)*, *52*(3), 393–442. Retrieved from `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). doi: 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. *Journal of Climate*, *25*(6), 1916-1944. doi: 10.1175/JCLI-D-11-00004.1

De Michele, C., Kottegoda, N. T., & Rosso, R. (2001). The derivation of areal reduction factor of storm rainfall from its scaling properties. *Water Resources Research*, *37*(12), 3247–3252. doi: 10.1029/2001WR000346

Demirdjian, L., Zhou, Y., & Huffman, G. J. (2018). Statistical modeling of extreme precipitation with trmm data. *Journal of applied meteorology and climatology*, *57*(1), 15–30. doi: 10.1175/JAMC-D-17-0023.1

Donat, M. G., Alexander, L. V., Yang, H., Durre, I., Vose, R., Dunn, R. J., ... Kitching, S. (2013). Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. *Journal of Geophysical Research: Atmospheres*, *118*, 2098–2118. doi: 10.1002/jgrd.50150

Ensor, L. A., & Robeson, S. M. (2008). Statistical characteristics of daily precipitation: comparisons of gridded and point datasets. *Journal of Applied Meteorology and Climatology*, *47*(9), 2468–2476. doi: 10.1175/2008JAMC1757.1

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., ... Zhao, B. (2017). The modern-era retrospective analysis for research and applications, version 2 (merra-2). *Journal of Climate*, *30*(14), 5419–5454. doi: 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 Resources Research*, *15*(5), 1049–1054. doi: 10.1029/WR015i005p01049

Gründemann, G. J. (2020). *Global Precipitation EXtremes dataset.* doi: 10.4121/uuid:12b5c941-cd54-45db-8d7b-fefaacecaa69

Gründemann, G. J., Werner, M., & Veldkamp, T. I. E. (2018). The potential of global reanalysis datasets in identifying flood events in southern africa. *Hydrology and Earth System Sciences*, *22*(9), 4667–4683. doi: 10.5194/hess-22-4667-2018

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horanyi, A., Munoz-Sabater, J., ... Thepaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 1-51. doi: 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. *Journal of Applied Meteorology*, *43*(12), 1834–1853. doi: 10.1175/JAM2173.1

Hosking, J. R. M. (1990). L-Moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, *52*(1), 105–124.

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. doi: https://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. *Geophysical Research Letters*, *47*(1), 1–9. doi: 10.1029/2019GL086138

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

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. *Journal of Hydrometeorology*, *5*(3), 487–503. doi: 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? *Journal of Climate*, *31*(7), 2945–2964. doi: 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? *Bulletin of the American Meteorological Society*, *98*, 69–78. doi: 10.1175/BAMS-D-14-00283.1

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., ... others (2015). The jra-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan*, *93*(1), 5–48. doi: 10.2151/jmsj.2015-001

Koutsoyiannis, D. (2004a). Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. *Hydrological Sciences Journal*, *49*(4), 591–610. doi: 10.1623/hysj.49.4.591.54424

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

Laherrere, J., & Sornette, D. (1998). Stretched exponential distributions in nature and economy:"fat tails" with characteristic scales. *The European Physical Journal B-Condensed Matter and Complex Systems*, *2*(4), 525–539.

Liu, Z., Liu, Y., Wang, S., Yang, X., Wang, L., Baig, M. H. A., ... Wang, Z. (2018). Evaluation of Spatial and Temporal Performances of ERA-Interim Precipitation and Temperature in Mainland China. *Journal of Climate*, *31*(11), 4347-4365. doi: 10.1175/JCLI-D-17-0212.1

Mailhot, A., & Duchesne, S. (2009). Design criteria of urban drainage infrastructures under climate change. *Journal of Water Resources Planning and Management*, *136*(2), 201–208. doi: 10.1061/(asce)wr.1943-5452.0000023

Marani, M., & Ignaccolo, M. (2015). A metastatistical approach to rainfall extremes. *Advances in Water Resources*, *79*, 121–126. doi: 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 Resources Research*, *51*, 639–647. doi: 10.1002/2014WR015885

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. *Advances in Water Resources*, *117*, 27–39. doi: 10.1016/j.advwatres.2018.05.001

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*, 280–290. doi: 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. *Journal of Hydrology*, *573*, 299–310. doi: 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. *Hydrology and Earth System Sciences*, *17*(10), 3921–3936. doi: 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 Resources Research*, *56*(7), e2019WR026535. doi: 10.1029/2019WR026535

Miniussi, A., Marani, M., & Villarini, G. (2020). Metastatistical Extreme Value Distribution applied to floods across the continental United States. *Advances* *in Water Resources*, *136*, 103498. doi: 10.1016/j.advwatres.2019.103498

Miniussi, A., Villarini, G., & Marani, M. (2020). Analyses Through the Metastatistical Extreme Value Distribution Identify Contributions of Tropical Cyclones to Rainfall Extremes in the Eastern United States Geophysical Research Letters. *Geophysical Research Letters*, *47*, e2020GL087238. doi: 10.1029/2020GL087238

Mishra, A. K., & Coulibaly, P. (2009). Developments in hydrometric network design: A review. *Reviews of Geophysics*, *47*(2). doi: 10.1029/2007RG000243

Nissen, K. M., & Ulbrich, U. (2017). Increasing frequencies and changing characteristics of heavy precipitation events threatening infrastructure in Europe under climate change. *Natural Hazards and Earth System Sciences*, *17*(7), 1177–1190. doi: 10.5194/nhess-17-1177-2017

Overeem, A., Buishand, A., & Holleman, I. (2008). Rainfall depth-duration-frequency curves and their uncertainties. *Journal of Hydrology*, *348*, 124–134. doi: 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 Resources Research*, *54*(9), 6725–6738. doi: 10.1029/2018WR022732

Papalexiou, S. M., & Koutsoyiannis, D. (2013). Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, *49*, 187–201. doi: 10.1029/2012WR012557

Papalexiou, S. M., Koutsoyiannis, D., & Makropoulos, C. (2013). How extreme is extreme? An assessment of daily rainfall distribution tails. *Hydrology and Earth System Sciences*, *17*, 851–862. doi: 10.5194/hess-17-851-2013

Peel, M., Finlayson, B., & McMahon, T. (2007). Updated world map of the Köppen–Geiger climate classification. *Hydrology and Earth System Sciences*, *11*(5), 1633–1644. doi: 10.5194/hess-11-1633-2007

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.* (Vol. 11; Tech. Rep.). 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* (Vol. 11; Tech. Rep.). 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. *Hydrological Sciences Journal*, *62*(6), 863–879. doi: 10.1080/02626667.2016.1260134

Rasmusson, E. M., & Arkin, P. A. (1993). A global view of large-scale precipitation variability. *Journal of Climate*, *6*(8), 1495–1522.

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. *Advances in Meteorology*, *2017*. doi: 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. *Hydrology and Earth System Sciences*, *23*(1), 595–619. doi: 10.5194/hess-23-595-2019

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.* doi: 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. *Theoretical and Applied Climatology*, *115*(1-2), 15–40. doi: 10.1007/s00704-013-0860-x

Serinaldi, F., & Kilsby, C. G. (2014). Rainfall extremes: Toward reconciliation after the battle of distributions. *Water Resources Research*, *50*, 336–352. doi: 10.1002/2013WR014211

Sivapalan, M., & Blöschl, G. (1998). Transformation of point rainfall to areal rainfall: Intensity-duration-frequency curves. *Journal of Hydrology*, *204*(1-4), 150–167.

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. *Review of Geophysics*, *56*(1), 79–107. doi: 10.1002/2017RG000574

UNSIDR. (2015). *Sendai Framework for Disaster Risk Reduction 2015 - 2030* (Tech. Rep.). Retrieved from `www.unisdr.org/we/inform/publications/43291`

Ushio, T., Sasashige, K., Kubota, T., Shige, S., Okamoto, K., Aonashi, K., … Kawasaki, Z. I. (2009). A kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. *Journal of the Meteorological Society of Japan*, *87 A*, 137–151. doi: 10.2151/jmsj.87A.137

van de Giesen, N., Hut, R., & Selker, J. (2014). The Trans-African Observatory (TAHMO). *WIREs Water*, *1*, 341–348. doi: 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. *Journal of Hydrology*, *400*(1-2), 103–120. doi: 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. *Journal of Hydrology*, *539*, 358–378. doi: 10.1016/j.jhydrol.2016.05.045

Wasko, C., Sharma, A., & Johnson, F. (2015). Does storm duration modulate the extreme precipitation-temperature scaling relationship? *Geophysical Research Letters*, *42*(20), 8783-8790. doi: 10.1002/2015GL066274

Westra, S., Alexander, L. V., & Zwiers, F. W. (2013). Global increasing trends in annual maximum daily precipitation. *Journal of Climate*, *26*(11), 3904–3918. doi: 10.1175/JCLI-D-12-00502.1

Wilson, P. S., & Toumi, R. (2005). A fundamental probability distribution for heavy rainfall. *Geophysical Research Letters*, *32*(14812). doi: 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 Sensing*, *11*(2). doi: 10.3390/rs11020151

Zorzetto, E., Botter, G., & Marani, M. (2016). On the emergence of rainfall extremes from ordinary events. *Geophysical Research Letters*, *43*(15), 8076–8082. doi: 10.1002/2016GL069445

Zorzetto, E., & Marani, M. (2019). Downscaling of rainfall extremes from satellite observations. *Water Resources Research*, *55*, 156–174. doi: 10.1029/2018WR022950

Zorzetto, E., & Marani, M. (2020). Extreme value metastatistical analysis of remotely sensed rainfall in ungauged areas: Spatial downscaling and error modelling. *Advances in Water Resources*, *135*, 103483. doi: 10.1016/j.advwatres.2019.103483