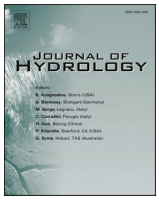
Journal of Hydrology 626 (2023) 130295

Contents lists available at ScienceDirect

Journal of Hydrology

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



Research papers

A method to derive satellite-based extreme precipitation return levels in poorly gauged areas

Matteo Siena [a,b,*], Vincenzo Levizzani [b], Francesco Marra [c,b]

[a] Department of Physics and Astronomy, University of Bologna, Bologna, Italy
[b] Institute of Atmospheric Sciences and Climate, National Research Council of Italy, Bologna, Italy
[c] Departmentof Geosciences, University of Padua, Padua, Italy

ARTICLE INFO

Keywords:
Extreme precipitation
Return levels
Satellite precipitation estimation
SMEV

ABSTRACT

Precipitation frequency analysis based on satellite products is still limited by estimation errors and by the use of statistical methods inadequate for these products. However, when it comes to poorly gauged areas of the world, satellite products can be a vital source of information. We present here a new method to derive satellite-based estimates of extreme precipitation quantiles with long return period in poorly gauged areas. The method relies on the identification of relations between statistics of the satellite estimation error and errors in the parameters of a non-asymptotic extreme value distribution. We show an application of the method in three areas with diverse climatic conditions in Austria and in the South-eastern Mediterranean, showcasing results for different scenarios of rain gauge density. We find that simple linear relations can explain 35–90% of the variance in the error of the parameters of the non-asymptotic extreme value distribution. Using these relations, we derive estimates of extreme return levels with drastically reduced bias and dispersion with respect to the ones directly obtained from the satellite estimates.

1. Introduction

Extreme precipitation has severe impacts on society, ecology, and economy, as it is the primary cause of natural disasters such as floods and landslides, while at the same time replenishing freshwater storages. Monitoring and quantifying the probability of occurrence of extreme precipitation events is thus crucial for hydraulic structure design, weather-related risk forecasting, and water resource management (Katz et al., 2002; Mianabadi, 2023). Hydrologists are interested in computing extreme return levels, which are high precipitation quantiles corresponding to low yearly exceedance probabilities. Usually, the Extreme Value Theory (Fischer and Tippet, 1928) represents the basic framework for precipitation frequency analyses. Extreme Value Theory approaches are based on the observed extremes, either annual maxima or peaks exceeding a high threshold. Long and accurate precipitation records are thus a fundamental prerequisite (Coles, 2001) and are typically provided by rain gauges. However, the Earth is covered by rain gauges in an uneven manner (Kidd et al., 2017; Becker et al., 2013), so that such needs are often not met (Libertino et al., 2018). In fact, vast portions of the globe remain poorly gauged or even ungauged.

Satellite observations may help us overcoming the limitations related to the insufficient rain gauge sampling, as they provide information at spatial resolutions of the order of tens of kilometres and have quasi-global coverage. The temporal coverage of satellite products is currently of the order of 20–30 years (exceeding 40 years in the case of climatological datasets), depending on the product, thus allowing us to attempt the estimation of extreme precipitation quantiles of interest. Despite some promising results (e.g., Gado et al., 2017; Demirdjian et al., 2018), precipitation frequency analyses based on satellite records are still limited by important knowledge gaps, mainly related to (i) estimation errors caused by the indirect nature of satellite retrievals, which are enhanced for extremes (Kidd and Levizzani, 2011; Miao et al., 2015; Prakash et al., 2015), and to (ii) the use of statistical methods proven inadequate for the characteristics of satellite products (Marra et al., 2019a).

When we want to carry on precipitation frequency analysis, high return period quantiles are typically investigated, despite the shortness of the available datasets (Wright et al., 2013). This often leads to important uncertainties in the analysis of extremes and derivation of high return period quantiles (Papalexiou and Koutsoyiannis, 2013). In addition, conditional bias affecting the satellite products (Berne and Krajewski, 2013) increases the magnitude of the uncertainty that

* Corresponding author at: Department of Physics and Astronomy, University of Bologna, Bologna, Italy.
E-mail address: matteo.siena4@unibo.it (M. Siena).

https://doi.org/10.1016/j.jhydrol.2023.130295
Received 16 June 2023; Received in revised form 20 September 2023; Accepted 22 September 2023
Available online 12 October 2023

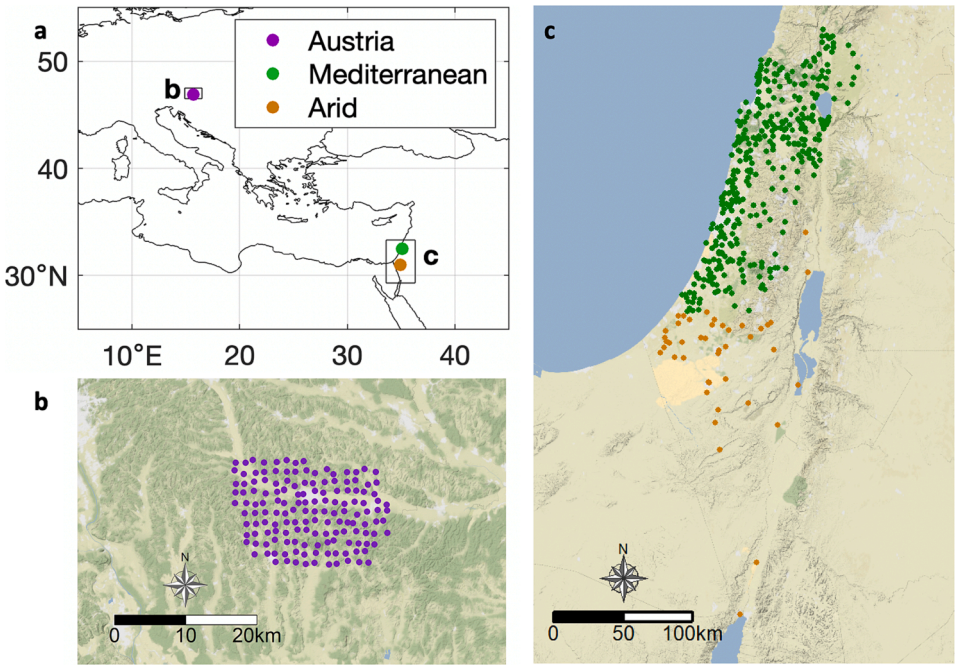


Fig. 1. (a) Location of the three study areas. (b) The dense WegenerNet network (WegenerNet DataPortal, 2023) in the surroundings of Feldbach (Austria). It consists of 151 stations in an area of roughly 300 km^2. (c) The 408 South-eastern Mediterranean rain gauges. Orange dots represent the stations in the ARID area, green dots those in the MED area. Distinction between the two is done based on a 400 mm mean annual precipitation isoline. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

already affects rain gauges when they are used to validate quantiles estimated with statistical approaches based on annual maxima (Pombo and de Oliveira, 2015) or peaks over a threshold (Demirdjian et al., 2018).

In this work, we make a step forward towards the use of satellite products for precipitation frequency analyses. Specifically, we propose a method for correcting satellite-based estimates of extreme quantiles over poorly gauged areas based on the satellite errors in the estimation of the distribution of daily precipitation amounts derived from the available gauge information. Hereafter, we use the term "error statistics" to refer to the satellite error in the estimation of linear moments of the distribution of daily precipitation amounts. This application is made possible by resorting to non-asymptotic statistics for extreme value analysis. Satellite error statistics are then computed from small amounts of rain gauge data and used to adjust extreme quantiles over poorly gauged locations.

2. Data and methods

We focus here on three study areas with diverse climatic conditions and rain gauge densities, and on a state-of-the-art satellite precipitation product. The methodology can be directly extended to other regions and/or applied to any other satellite precipitation product.

2.1. Study areas and rain gauge data

The study areas were selected considering three aspects: (i) availability of relatively long ground-based observation records to be used as a benchmark; (ii) independence of the satellite dataset from the gauge networks; (iii) possibility to explore different climatic conditions. We focused on two rain gauge networks (Fig. 1a): the dense WegenerNet network (Fuchsberger et al., 2021b; Fuchsberger et al., 2021a) covering the surroundings of Feldbach (Austria); and a sparser network that covers two diverse climatic areas in the South-eastern Mediterranean. Following the Köppen and Geiger (1936) climatic classification, the Austrian network falls into the warm-summer humid continental climate Dfb, while the South-eastern Mediterranean network covers three different climatic zones: Csa (Warm Mediterranean), BSh (Warm semi-arid) and BWh (Warm desert) (see Peel et al., 2007). Since the differences between the last two areas are minor and the rain gauge network is here less dense, in this work the BSh and BWh stations were grouped together. The stations belonging to the "Mediterranean" (Csa) area were labelled as MED, and the ones belonging to the "Arid" or "Semi-Arid" (both BSh and BWh) areas, were labelled as ARID. For simplicity MED and ARID stations are separated using the 400 mm mean annual precipitation isoline, as it was found to well reproduce the Köppen-Geiger classes in this area.

As per Austria, 151 rain gauges were selected from the WegenerNet portal (Fig. 1b). The detection threshold is 0.11 mm and precipitation amounts are made available every day at 12:00 UTC. The overall record spans from January 2007 to the end of 2020. No missing values were present, due to the application of interpolated estimates. Since some of the stations were characterized by the absence of one or more yearly record, they were excluded from the analysis, leading to a total of 128 rain gauges. This resulted in an average gauge density of roughly 0.4 stations per km^2 (about 1 station per 2.5 km^2).

The 408 MED and ARID stations are located more heterogeneously in space (Fig. 1c) and cover diverse periods between 1948 and 2018 (median record length is 56 years), with average gauge densities of approximately 0.03 MED stations per km^2 (about 1 station per 30 km^2) and 0.003 ARID stations per km^2 (about 1 station per 300 km^2). These rain gauges record daily precipitation values grater or equal to the minimum detection threshold of 0.1 mm every day at 06:00 UTC. Available data were already quality controlled and filtered for missing values (see details in Marra et al., 2021), so that the additional quality check done in this study only consisted in removing years with more than 10 % of missing values.

In all cases satellite data can be considered independent (Austria, no gauge is provided for the satellite product adjustment) or almost independent (MED, ARID, only 8 stations – less than 2 % – are made available for satellite operational adjustment) from the rain gauge networks.

2.2. Satellite data

Satellite data were taken from version 6 of the Integrated Multi-satellitE Retrievals algorithm (IMERG) of the Global Precipitation Measurement (GPM) mission (Huffman et al., 2020) (downloaded from NASA GES DISC service, https://gpm.nasa.gov/node/3328, last access Dec 2021). IMERG (NASA, IMERGv06) combines precipitation estimates from microwave and infrared observations from low-orbiting and geosynchronous satellites with those of the Dual-frequency Precipitation Radar (DPR) on the GPM core satellite and of rain gauges (Hou et al., 2014; Huffman et al., 2007). The precipitation estimates cover portions of the globe between 90° N-S in latitude on a 0.1° grid (on the order of 10 km) and with a temporal resolution of 30 minutes. In this work, the IMERG Final Run is used, which contains research-quality multi-satellite estimates with quasi-Lagrangian time interpolation, gauge data and climatological adjustment, as this represents the ideal version for precipitation frequency analysis.

Four variables can be identified in the original netCDF files: precipitation estimates with 30-min resolution (in mm h^{-1}), time step (in seconds), latitude and longitude of the centre of each 0.1° × 0.1° grid cell. Each station is associated to the co-located satellite pixel based on the minimum distance from the pixel centre. Half-hourly satellite precipitation intensity estimates are converted into daily precipitation accumulations (mm) assuming uniform precipitation rates during the 30-min time intervals and considering the recording time of daily precipitation in each of the rain gauge networks.

2.3. Statistical framework

2.3.1. Ordinary and extreme events

Non-asymptotic statistics relies on the concept of ordinary events, which are all the independent realizations of the examined process (e.g., Marani and Ignaccolo, 2015; Zorzetto et al., 2016). For the case of precipitation, they are usually considered as rainfall amounts of the duration of interest observed during statistically independent rainy periods (Marra et al., 2020). Extremes emerge from finite samples of these ordinary events. In this non-asymptotic perspective, extreme value distributions can be derived following ordinary statistics and explicitly accounting for (i) the distribution of the ordinary events, and (ii) the number of annual occurrences of the ordinary events.

2.3.2. Metastatistical extreme value (MEV) and simplified MEV (SMEV)

To study the yearly exceedance probability of extreme precipitation events, Marani and Ignaccolo (2015) introduced a new statistical method termed Metastatistical Extreme Value (MEV). This method avoids the asymptotic assumption of the Extreme Value Theory (Gumbel, 1958) and derives the distribution of extremes explicitly accounting for the yearly number of occurrences of the ordinary events (see also Zorzetto et al., 2016). In MEV, the parameter defining the number of events and the parameters of the distribution describing the magnitude of the ordinary events are considered as random realizations of stochastic variables. For the case of daily precipitation, the Weibull distribution (Weibull, 1951) is typically used, following the theoretical reasoning by Wilson and Toumi (2005), which showed that the exceedance probability of heavy daily precipitation decreases as a powered exponential, and can thus be described using a two parameter Weibull distribution (e.g., Marani and Ignaccolo, 2015; Zorzetto et al., 2016; Schellander et al., 2019).

The MEV method with the Weibull model has already been used to derive extreme return levels from satellite precipitation products with encouraging results, due to a reduced estimation variance (Zorzetto and Marani, 2019; Hu et al., 2020; Zorzetto and Marani, 2020). However, systematic biases were sometimes reported (Wang et al., 2020; Miniussi and Marra, 2021; Poschlod, 2021). These biases were found to be related to the inexact assumption of having all the ordinary events described by a Weibull distribution (see Marra et al. (2018) and following literature). Indeed, the results by Wilson and Toumi (2005) are asymptotic and only the tail of the ordinary events distribution is expected to be described by a two-parameter Weibull. In fact, this Weibull tail model was recently confirmed to well represent extremes globally when proper left-censoring thresholds are used (Marra et al., 2023).

Following Marra et al. (2020), the ordinary events' tail can be defined as the portion of ordinary events that share the statistical properties with the annual maxima. The above-mentioned bias can thus be reduced by left-censoring the ordinary events, as described in Marra et al. (2020). In this study, we use two-parameter Weibull distributions to describe the tail of the daily precipitation ordinary events. Identification of these 'tails' is described in detail below. However, due to the limited proportion of yearly ordinary events that belong to the Weibull tails (Marra et al., 2023), estimation of the parameters on a yearly basis as required by MEV could lead to uncertain estimates of the precipitation quantiles.

The Simplified MEV (SMEV) framework proposed by Marra et al. (2019b) allows to overcome this practical problem. In SMEV, the inter-annual variability is neglected, and the Weibull tail parameters are estimated based on the entire data record. This provides a much larger data sample. When an identical distribution of the wet days is assumed, the SMEV formulation for daily precipitation becomes (Marra et al., 2019b):

$$G_{\mathrm{SMEV}}(x) = [F(x;\vartheta,\lambda)]^{n} \tag{1}$$

where n is the mean number of wet days per year and F the cumulative distribution function describing their tail, that is the two-parameter Weibull distribution characterized by the shape and scale parameters, respectively ϑ and λ:

$$F(x;\vartheta,\lambda) = 1 - \exp\left[-\left(\frac{x}{\lambda}\right)^{\vartheta}\right] \tag{2}$$

2.4. Precipitation frequency analysis

2.4.1. Identification of the ordinary events

To derive extreme quantiles from ground-based and satellite data using the SMEV non-asymptotic framework, one needs to define the ordinary events. Here, we assume independence of the wet days, as previously done in many studies focusing on extreme precipitation (e.g., see the global study by Marra et al. 2023, and note that Papalexiou 2022 shows that even hourly precipitation has only moderate autocorrelation). We define as "wet" all the days in which at least ω mm of precipitation are reported. Since satellite precipitation products show large variability in their sensitivity (e.g., Kidd and Levizzani, 2011; Sun et al., 2018; Levizzani et al., 2020a; Levizzani et al., 2020b), we optimize the definition of ω for the satellite product in each study area. To do so, we examine the number of wet days identified by IMERG and the rain gauges for different ω. In Austria, the number of wet days is well represented by IMERG already for $\omega = 0.1$ mm day^{-1}, while a larger threshold ($\omega = 1$ mm day^{-1}) is needed in MED and ARID areas. Following these observations, we use $\omega = 0.1$ mm day^{-1} in Austria and $\omega = 1$ mm day^{-1} in the MED and ARID areas. These values are in line with previous studies based on similar statistical approaches for the definition of ordinary events (Zorzetto et al., 2016; Marra et al., 2019b). In generic applications, one can select the most appropriate value based on the available rain gauge observations.

2.4.2. Tail of the ordinary events' distribution

Next, the tail of the ordinary events' distribution must be identified. This is the portion of ordinary events from which annual maxima are sampled and for which our assumption of Weibull tails holds. To this end, we apply the test devised by Marra et al. (2020). The test, described in detail in Marra et al. (2023), ensures that the used assumption is not contradicted by the available observations. The test examines several left-censoring thresholds and compares the observed annual maxima

with synthetic maxima drawn from ordinary events belonging to the corresponding Weibull distribution in a Montecarlo framework. A threshold t^* is "optimal" when the observed annual maxima are likely samples from ordinary events with a Weibull tail $\forall t \geq t^*$. This implies that any threshold $t \geq t^*$ will lead to indistinguishable estimates of the parameters. We apply the test to the rain gauge data generating 1000 Montecarlo iterations and using $p_{\text{test}} = 0.2$. This means that we reject our model when more than 20 % of the annual maxima lie outside of the 80 % confidence interval of the Montecarlo samples. The codes for the test are publicly available in Marra (2022); here we use a version of the code translated to R.

2.4.3. Estimation of return levels with SMEV

Parameters of the Weibull distribution λ and ϑ are computed by left-censoring the ordinary events not belonging to the above defined tail and using a least squares regression in Weibull-transformed coordinates (Marani and Ignaccolo, 2015; codes available in Marra, 2020). Return levels q associated with return periods $T = 2$, 5, 10 and 50 years are computed inverting the SMEV cumulative distribution function in Eq. (1):

$$q = \lambda\left[-ln\left(1 - p^{\frac{1}{n}}\right)\right]^{\frac{1}{\vartheta}} \tag{3}$$

where p is the desired yearly non-exceedance probability $\left(p = 1 - \frac{1}{T}\right)$.

2.5. Correction of satellite-based return levels

The idea behind our approach is that using non-asymptotic statistics it is possible to establish quantitative relations linking satellite estimation errors (i.e., errors affecting the statistics of the ordinary events which depend on the retrieval setup of a given product) to the errors in the estimated parameters of the ordinary events distribution, and then, through the SMEV equation above, the quantiles. Moreover, our idea relies on the assumption that these relations are region-specific and depend on the regional characteristics of the satellite estimation errors and on the local climate. Since estimation errors of satellite precipitation products are often provided by evaluation studies (e.g., Chen et al., 2020; Maggioni et al., 2021) or can be quantified using the entire set of ordinary events from data records much shorter than the ones required for extreme value analyses, the above-mentioned relations could be derived in well-sampled regions and then used to correct satellite-based estimates of extreme quantiles over climatologically similar regions where few or short-recording gauges are available.

2.5.1. Relations between satellite estimation errors and errors in the SMEV parameters

Correcting the parameter n of the SMEV framework is rather trivial, because it accounts for the average yearly number of wet days. The multiplicative error (bias) in the average number of wet days can be computed as:

$$\text{bias}_n = n^{\text{IMERG}}/n^{\text{stations}} \tag{4}$$

This is usually a basic information made available by studies evaluating the accuracy of satellite-based precipitation products all over the globe (e.g., Smith et al., 2006; Liu et al., 2019; Müller and Thomson, 2013).

The problem is thus reduced to the identification of relations between the errors in the satellite precipitation estimates and the errors in the estimated parameters of the intensity distribution. L-moments (Hosking, 1990) can be estimated robustly from limited data and represent a widely used tool in the statistical hydrology community (Vogel and Fennessey, 1993). We propose to use relations linking errors in the L-moments of the ordinary events to errors in the scale and shape parameters of the Weibull distribution describing their tail in the SMEV framework of Eqs. (1) and (2). Each relation is estimated once for each given area, leading to 5 parameters per area. It is important to point out that these relations are in principle non trivial, because the L-moments are computed from the entire set of ordinary events (i.e., all the wet days), while the parameters of the Weibull distribution only describe a small fraction of these events, in our case only the largest 10 % or 25 % of the values defined as wet in a proper manner depending on the area (see below), which is not necessarily known when gauge data availability is poor because the above described test requires the availability of many years of data.

We expect the scale parameter to be closely related to the first L-moment (l_1, i.e., the mean, whose error statistics is also typically provided in validation studies), due to the defining property of scale parameters in distributions with no location parameter, such as the two-parameter Weibull: $F(x;\lambda,\vartheta) = F\left(\frac{x}{\lambda};1,\vartheta\right)$. The relations linking multiplicative errors (bias) in the scale parameter

$$\text{bias}_\lambda = \lambda^{\text{IMERG}}/\lambda^{\text{stations}} \tag{5}$$

to multiplicative errors (bias) in the estimated mean (l_1)

$$\text{bias}_{l_1} = l_1^{\text{IMERG}}/l_1^{\text{stations}} \tag{6}$$

should thus be linear and with intercept equal to 0 and slope equal to 1. However, we expect deviations from this relation, since, as mentioned above, the model parameters λ and ϑ are computed from the distribution tail. For simplicity, we look for general linear relations in the form

$$\text{bias}_\lambda = a_\lambda + b_\lambda \cdot \text{bias}_{l_1} \tag{7}$$

The shape parameter is closely related to the third L-moment ratio (i.e., the L-skewness τ_3). Indeed, τ_3 can be obtained combining the two L-moments l_3 and l_2, which for the case of a two-parameter Weibull distribution can be written as (Akram and Hayat, 2014; Goda et al., 2011):

$$\tau_3 = \frac{l_3}{l_2} = \frac{1 - \frac{3}{2^{\frac{1}{\vartheta}}} + \frac{2}{3^{\frac{1}{\vartheta}}}}{1 - \frac{1}{2^{\frac{1}{\vartheta}}}} \tag{8}$$

We can see that the L-skewness τ_3 only depends on the shape parameter ϑ, and vice versa. Errors in its estimation are thus expected to be directly related to errors in the estimation of ϑ. We look for relations linking the error in the parameters of the Weibull distribution describing the tail of the ordinary events

$$\varepsilon_\vartheta = \vartheta^{\text{IMERG}} - \vartheta^{\text{stations}} \tag{9}$$

with the errors in the L-moments computed from all the ordinary events

$$\varepsilon_{\tau_3} = \tau_3^{\text{IMERG}} - \tau_3^{\text{stations}} \tag{10}$$

Again, L-skewness is computed from all the data and the shape parameter from the tail only, but given the properties of the L-skewness as a descriptor of the distribution tail, it is reasonable to expect a relatively strong relation between the L-skewness and our shape parameter. As a first approximation, we look for linear relations in the form:

$$\varepsilon_\vartheta = a_\vartheta + b_\vartheta \cdot \varepsilon_{\tau_3} \tag{11}$$

2.5.2. Correction of SMEV model parameters and return levels

The correction of n is based on information about the median satellite error in representing the number of wet days in the region of interest, as in Eq. (4). Indeed, since the median is less affected by the presence of outliers caused by biased data, it is better suited for such kind of corrections, especially for areas with few wet days (e.g., in our case, the Southern Mediterranean areas). The correction of the SMEV-Weibull parameters derived from satellite observations over a region of interest is done using the linear relations derived above and exploiting information on the satellite error statistics from the available station data, as follows:

1. The satellite error statistics bias_{l_1} and ε_{τ_3} are computed from Eqs.

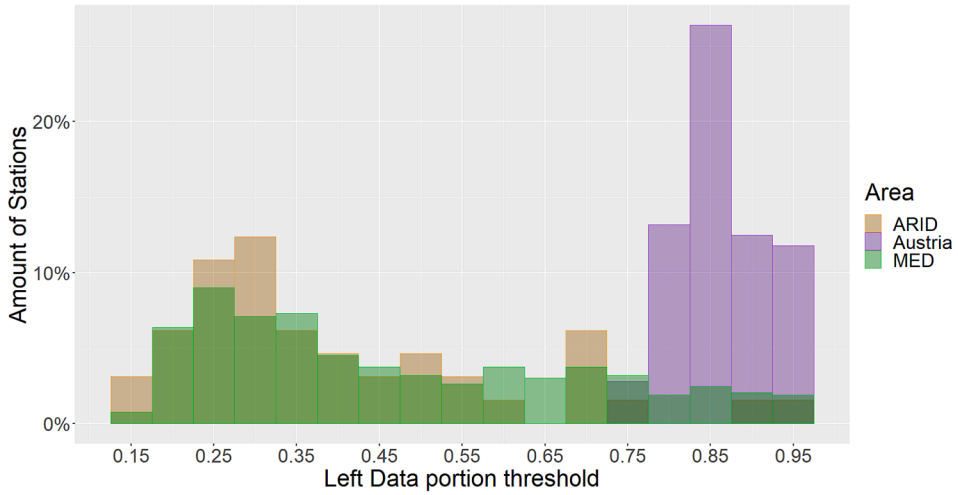


Fig. 2. Optimal left-censoring thresholds identified for MED (green), ARID (orange) and Austria (purple) stations: each bin in the x-axis represents the lowest threshold for which the assumption of having annual maxima sampled from Weibull tails could not be rejected for any threshold equal or larger. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(6) and (8) using the available rain gauge data (e.g., short-recording gauges, nearby gauges, etc) and the representative satellite pixels;

2. The linear relations in Eqs. (7) and (11) are used to compute the errors in the SMEV parameters bias_λ and ε_ϑ from the satellite error statistics at point 1;

3. The corrected parameters n^*, λ^* and ϑ^* are computed inverting Eqs. (5) and (9):

$$\begin{cases} n^* = n^{\text{IMERG}}/\text{median}(\text{bias}_n) \\ \lambda^* = \lambda/\text{bias}_\lambda = \lambda\ /(a_\lambda + b_\lambda \cdot \text{bias}_{l_1}) \\ \vartheta^* = \vartheta - \varepsilon_\vartheta = \vartheta - (a_\vartheta + b_\vartheta \cdot \varepsilon_{\tau_3}) \end{cases} \tag{12}$$

4 The adjusted return levels q^* corresponding to the yearly non-exceedance probability p are then computed from the corrected parameters at point 3 using Eq. (3):

$$q^* = \lambda^* \left[-ln\left(1 - p^{\frac{1}{n^*}}\right) \right]^{\frac{1}{\vartheta^*}} \tag{13}$$

The sensitivity of the resulting return levels to the uncertainty in the linear relations between L-moments and parameters (not addressed in this study) can be quantified using error propagation. To this end, we report in Appendix A the analytical derivation of the propagation of the error from the SMEV parameters to the resulting return levels using Taylor series to the second order. It should be noted that the proposed procedure implicitly corrects possible biases due to the scale mismatch between the satellite pixel (areal estimate) and the rain gauge (point), such as the ones that are handled analytically in the method by Zorzetto and Marani (2019). The obtained return levels are thus to be considered representative of the point scale. Thus, this method is not appropriate for studies interested in the analysis of precipitation extremes at large spatial scales.

2.5.3. Validation of the estimated return levels

The effectiveness of the applied corrections is first evaluated by comparing the new corrected satellite distribution parameters (n^*, ϑ^*, λ^*) to the original parameters of all the available rain gauges. Afterwards, the accuracy of the proposed approach is quantified using a leave-out cross validation. Each study area is treated independently, and different scenarios of available rain gauge densities are explored. The experiment is repeated 100 times for each area and density to derive proper statistics of the performance of the method.

For each iteration, a set of calibration stations is randomly selected to match the desired density: the error statistics are computed from the calibration stations (Eqs. 5, 6, 9, 10), and the linear relations between satellite error statistics and errors on the SMEV parameters (Eqs. 7, 11) are derived; the average error on the number of wet days is computed (Eq. 4). The remaining stations are used for validation: local error statistics are computed (Eqs. 6, 10); adjusted parameters are computed (Eq. 12) and the corresponding return levels are derived (Eq. 13). The adjusted return levels are then compared with the ones directly derived from the co-located stations, that serve here as a reference. Finally, we produce boxplots of the multiplicative bias in the estimated return levels and we compute the Fractional Standard Error (FSE) of the estimated parameters and return levels for each study area and rain gauge density. The FSE is the root mean square error normalized over the average reference value, and is computed as

$$FSE = \frac{\sqrt{\frac{1}{N}\sum_i (s_i - g_i)^2}}{\frac{1}{N}\sum_i g_i} \tag{14}$$

where $g_i = 1 \cdots N$ are the original rain gauge estimates and $s_i = 1 \cdots N$ the satellite original or adjusted estimates.

Given the different extent and gauge availability characterizing the study areas, diverse sets of densities were explored in the three areas. For Austria, we studied cases related to 1 station per 5, 10, 25 and 50 km^2; while for MED stations, densities of 1 station per 50, 70, 100 and 200 km^2 were considered. For the ARID area, we studied calibration densities of 1 station per 400, 600 and 1000 km^2.

3. Results and discussion

3.1. Distributions tails

Following the definitions in section 2, we identified the tail of the ordinary events distribution using the method described in section 2.4.2. We report in Fig. 2 the results of the test, which revealed the portion of the ordinary events from which annual maxima are likely sampled. As per Austria, most of the stations showed that annual maxima are likely samples from Weibull tails describing the largest 25 % to 10 % of the ordinary events. In only less than 20 % of the stations a threshold higher than 0.9 was necessary. Given the specificity of the test against alternate hypotheses (Marra et al., 2023), we can safely use a left-censoring threshold of 0.9 for this area, meaning that the Weibull tail of the distribution is here identified as the highest 10 % of the ordinary events. This threshold is consistent with previous studies based in areas with similar climates, such as Germany (Miniussi and Marra, 2021), and with global results which showed that Northern Europe tends to require

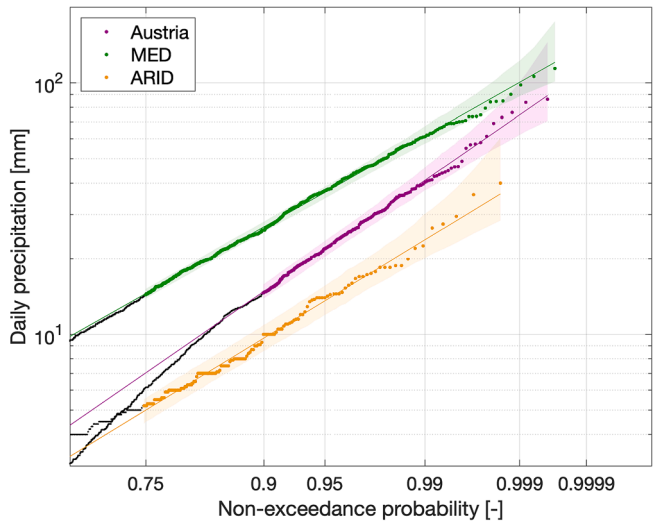


Fig. 3. Tail of the distribution of the ordinary events for three random stations in the Austria (purple), MED (green) and ARID (orange) regions. Dots show the empirical non-exceedance probability of the ordinary events: coloured dots for the events above the left-censoring threshold, black dots for the events below. Solid lines show the Weibull distributions describing the tails estimated as described in Section 2.4.3. Shaded areas show the 90% sampling confidence interval of ordinary events sampled from the estimated distributions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

higher left-censoring thresholds (Marra et al., 2023).

For the South-eastern Mediterranean, the assumption of Weibull tails is not rejected for much lower thresholds. Specifically, 87 % of the 408 analysed stations required left-censoring thresholds lower than 0.75, meaning that here annual maxima are likely samples from Weibull tails that can be described using 25 % of the ordinary events or more. This agrees with previous studies in the same region (Marra et al., 2019b).

The need for different thresholds in the two regions can be associated with the differences in local climatology. As opposed to Austria, in which precipitation events are often associated with stratiform processes, in the South-Mediterranean precipitation is mostly associated to convective processes that more likely satisfy the assumptions of the thermodynamic analyses by Wilson and Toumi (2005) and that were previously found to be more likely associated with stretched exponential tails (e.g., Berg et al., 2013). Fig. 3 shows the cumulative distribution function of the ordinary events for three stations located in the three regions (Austria, MED, ARID). The figure shows how the Weibull tail model is appropriate for all the three cases.

3.2. Satellite estimation of SMEV parameters and return levels

Austrian stations showed a high variability in the number of yearly ordinary events, with values ranging roughly from 150 to 230. Satellite estimates, on the contrary, showed very little dispersion with a mean value centred near 171 days. This is because only 8 satellite pixels cover all the 128 stations. Satellite underestimation of the n parameter was observed for most of the stations. The tipping buckets of the stations collect rain below their detectable threshold, but they tip when the volume of such buckets is full, thus precipitation under the threshold happened in the previous day is recorded as daily values in the next ones; on the other hand, satellite products capture back-scattered radiation corresponding to precipitation signals greater or equal to their detection threshold, ending up in recording less daily events. This suggests that the estimation errors are more likely to be due to the satellite product. MED stations showed overall a compensation between the estimates of n from IMERG and stations, in fact the median value for the satellite is nearly equal to 42, and 39 for the rain gauges. This was not the case of ARID stations, since greater noise was detected and a tendency to IMERG underestimation was caught: in fact, the median value for the satellite estimates amounts to 31 mean rainy days compared to the 40 of the stations. The variability in these two areas is strictly related to the high intrinsic noise in precipitation detection due to the presence of Mediterranean coastal regions, Dead Sea and deserts in the satellite pixels. Furthermore, climatic conditions are such that precipitation events are typically due to convective processes, which could involve just one or few stations during a single event. This last feature becomes more relevant for the ARID stations, since the very few precipitation phenomena in these places are almost entirely convective (Sharon, 1972).

Shape parameter estimates were characterized by overall satellite overestimation for the Austrian dataset, thus leading to an underestimation of extreme quantiles, mainly for the lowest ones (2, 5 and 10 year return period), while a greater compensation was observed for longer

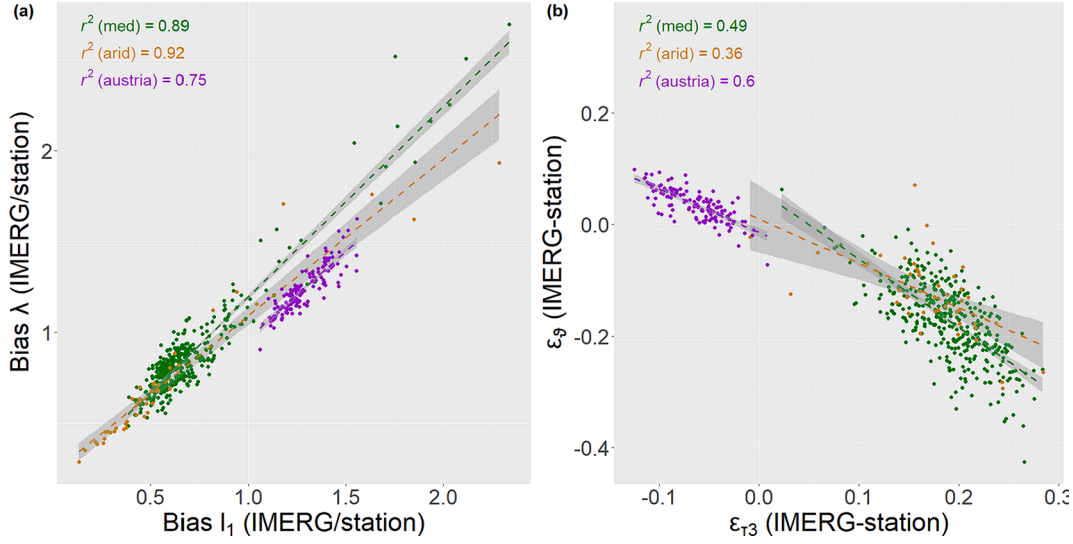


Fig. 4. Linear relations between satellite error statistics and error in the SMEV parameters for the MED (green), ARID (orange), and Austria (purple) areas. (a) Relations between the multiplicative error in the mean daily precipitation bias_{I_1}, and in the SMEV scale parameter bias_λ. (b) Relations between the errors in the L-skewness ε_{τ_3} and errors in the SMEV shape parameter ε_ϑ. In both panels, dashed lines represent the linear regressions and the shaded areas the 90% confidence intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

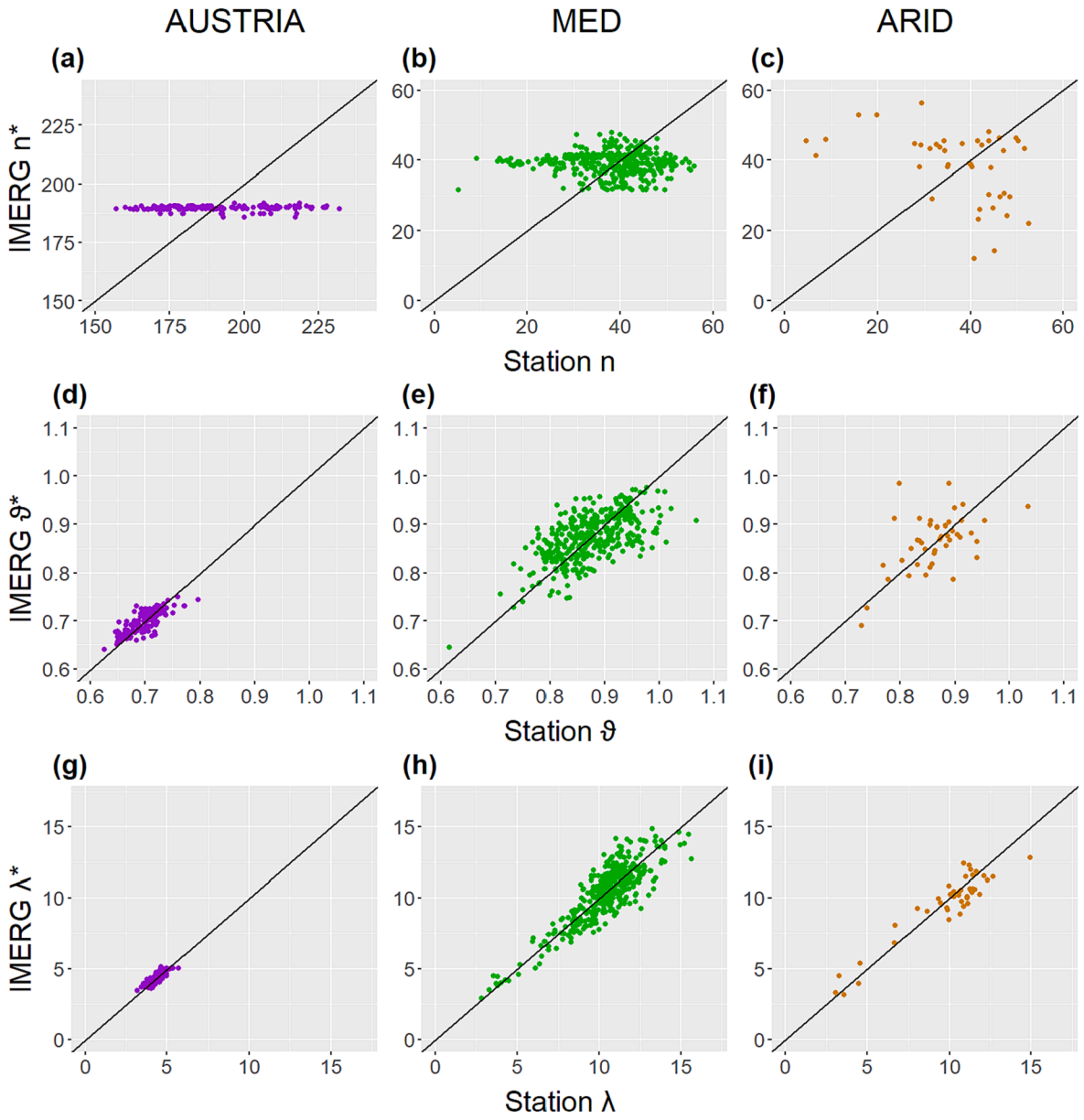


Fig. 5. (a-c) Scatterplots of the adjusted mean number of yearly ordinary events (n) for the three climatic areas, where the 1:1 line (solid black) is plotted to better detect the cases of satellite over-/under-estimation of the parameters compared to the stations. Along the x axis, the rain gauge values are reported, while satellite-based estimates after correction are shown in the y axis. Please note that the y- and x-axis limits of the AUSTRIA n parameter scatterplot are different from the ones for the MED and ARID areas. (d-f) Scatterplots of the adjusted shape parameter for the three climatic areas. (g-i) Scatterplots of the adjusted scale parameter for the three climatic areas.

return periods. Both the MED and the ARID areas showed clear IMERG underestimation of the shape parameter. Scale estimates showed a non-negligible portion of stations characterized by overestimation. Anyway, a general overestimation affects all the computed quantiles due to the combination of two effects: 1) the impact of the underestimation of the shape parameter is stronger than the one of the scale; 2) the shape parameter weighs more than the scale in determining extreme return levels. The last aspect derives from the fact that the shape parameter appears as root index in the quantile inversion formula, while the scale is a simple multiplicative factor. This is corroborated by Appendix A. These analyses helped us assessing the absence of a proper correlation between satellite and ground-based estimates of the model parameters.

3.3. Relations between satellite estimation errors and errors in the SMEV parameters

We found robust relations between the multiplicative error in the first L-moment $bias_{l1}$ (Eq. 6) and the corresponding multiplicative error in the estimation of the scale parameter $bias_{\lambda}$ (Eq. 5). Similarly, we reported an important inverse correlation between the additive error in the L-skewness $\varepsilon_{\tau 3}$ (Eq. 10) and the corresponding additive error in the shape parameter ε_{υ} (Eq. 9). In Fig. 4, the dashed lines represent the linear relations linking the error in the SMEV parameters to the error statistics in the L-moments of the observations for the three climatic areas. The grey shading highlights the 90 % confidence interval in these relations. The figure also reports the coefficient of determination r^2. As expected, we reported a rather strong correlation between $bias_{l1}$ and $bias_{\lambda}$, with about 90 % of the variance in $bias_{\lambda}$ explained by $bias_{l1}$ in the MED and ARID areas, respectively ($r^2 = 0.89$; $r^2 = 0.92$), and about 75 %

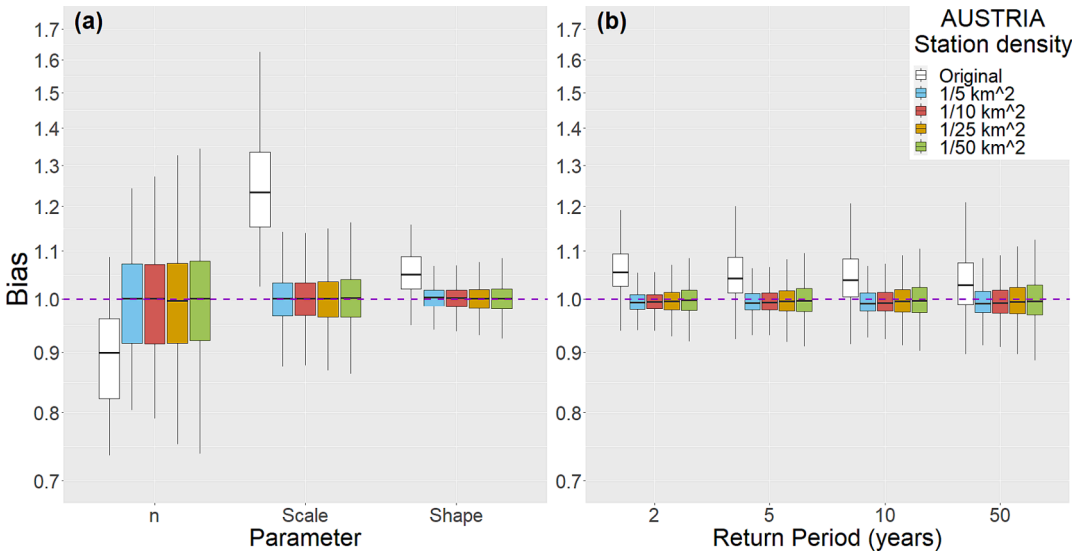


Fig. 6. Boxplots showing the bias in the estimated parameters (a) and return levels (b) for the Austria study case. The white boxes represent the bias in the non-adjusted estimates, while the coloured boxes report the bias in the estimates adjusted with varying densities of the calibration stations.

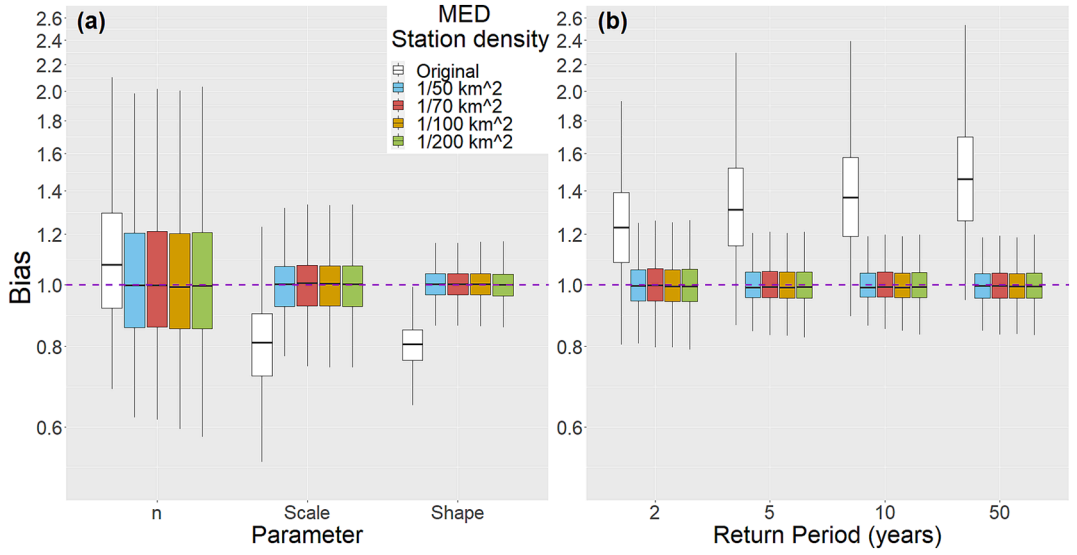


Fig. 7. Same as in Fig. 6 for the MED area. Please note that the y-axis limits are different from the ones in Fig. 6.

in Austria ($r^2 = 0.75$). The relations between ε_0 and $\varepsilon_{\tau 3}$ are generally weaker, but still retain an appreciable explanatory power, with 60 % of explained variance over Austria and about 49 % and 36 % in MED and ARID areas, respectively ($r^2 = 0.36$; $r^2 = 0.49$). These relationships can be used to adjust the satellite quantile estimates with respect to the rain gauges.

3.4. Validation of the corrected satellite return levels

Finally, we carried out corrections of satellite-based extreme quantile estimates exploiting relations such as the ones identified above.

Scatterplots of the satellite parameters (n^*, ϑ^*, λ^*) adjusted using relations obtained considering all the available rain gauges versus the corresponding rain gauge values are reported in Fig. 5. As expected, n^* values do not show a specific relation, since it was adjusted simply using the median error (see above in section 2.5.2). Conversely, the adjusted shape (ϑ^*) and scale (λ^*) parameters show a clear direct correlation with the rain gauge values.

We then quantified the accuracy of the proposed method using a leave-out approach. New relations were derived for each iteration of the leave-out procedure using the calibration rain gauges and were then evaluated in the adjustment of satellite estimates over the validation rain gauges. Figs. 6–8 show the distribution of bias in the SMEV parameters and in the estimated return levels obtained over the three areas for different densities. The boxplots show the results across the validation stations for 100 sampling iterations. In all cases, the median bias is 1 or very close to 1 (dashed purple line), despite the biases in the parameters and return levels directly estimated from the satellite are sometimes large.

For the case of Austria (Fig. 6), one can see how the tendency to overestimate the scale parameter of the satellite estimates (nearly 1.2 times higher than the rain gauge estimates) was compensating the overestimation of the shape parameter (median bias about 5 %) leading to estimates of the 50-year return levels bias lower than 5 % and a standard error lower than 5 % (Table 1). This is caused by the compensation of the errors (see Appendix A). The corrected parameters are almost unbiased in their median and show a much smaller dispersion. This translates into a more accurate estimation of return levels that

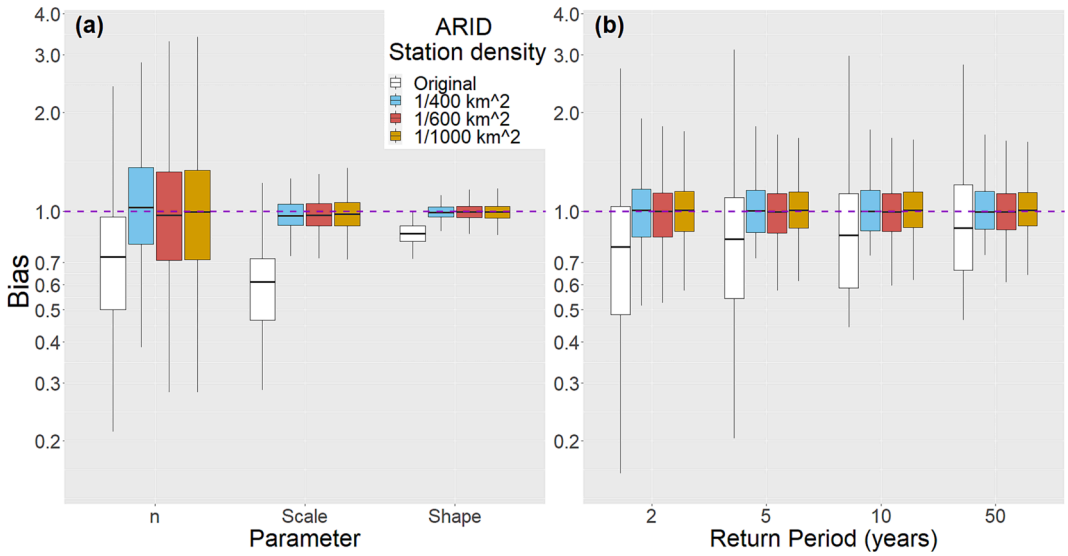


Fig. 8. Same as in Fig. 6 for the ARID area. Please note that the y-axis limits are different from the ones in Figs. 6 and 7.

Table 1
Fractional Standard Error (FSE, See Eq. (14) of the estimated parameters and return levels for each of the three climatic areas. Non-adjusted (original) and adjusted (for different calibration station densities) values are obtained considering the original station values as truth and the estimated satellite values as the prediction.

AUSTRIA	2 yr RP	5 yr RP	10 yr RP	50 yr RP	Shape	Scale	n
Original	7.8 %	7.5 %	7.4 %	7.4 %	6.6 %	2.5 %	1.5 %
1/50 km^2	3.5 %	3.9 %	4.1 %	4.8 %	3.1 %	5.8 %	1.1 %
1/25 km^2	2.9 %	3.2 %	3.6 %	4.3 %	2.9 %	5.4 %	1.0 %
1/10 km^2	2.4 %	2.8 %	3.1 %	3.7 %	2.7 %	5.0 %	9.7 %
1/5 km^2	2.3 %	2.6 %	2.9 %	3.6 %	2.7 %	5.0 %	9.5 %
MED	2 yr RP	5 yr RP	10 yr RP	50 yr RP	Shape	Scale	n
Original	34%	40%	44%	52%	21%	27%	27%
1/200 km^2	9.1%	7.7%	7.3%	7.1%	5.7%	10%	25%
1/100 km^2	8.9%	7.4%	7.0%	6.8%	5.7%	10%	25%
1/70 km^2	8.9%	7.5%	7.0%	6.8%	5.7%	10%	25%
1/50 km^2	8.9%	7.4%	6.9%	6.7%	5.8%	10%	25%
ARID	2 yr RP	5 yr RP	10 yr RP	50 yr RP	Shape	Scale	n
Original	49%	46%	45%	44%	17%	48%	50%
1/1000 km^2	23%	22%	22%	22%	7.4%	12%	49%
1/600 km^2	23%	22%	22%	22%	7.4%	11%	50%
1/400 km^2	23%	22%	21%	21%	7.0%	10%	50%

can now be estimated with biases mostly lower than 5 % (and always lower than 15 %), even for relatively sparse densities of 1 station per 50 km^2. Dispersion of the estimates is nearly symmetric, with little underestimation prevalence of median values, especially for higher calibration densities. It is important to note that these results are obtained using relatively short rain gauge records (14 years), which well mimic situations with recently deployed dense raingauge networks. The small dispersion in the estimates also indirectly supports the use of non-asymptotic methods for these purposes (e.g., Zorzetto et al., 2016).

Corrections in the MED area (Fig. 7) proved to be highly accurate as well, despite the lower density of stations and the higher climate variability of the area. Median values of the corrected parameters are almost unbiased (laying near the exact match line and the dispersion) for all the SMEV parameters. Here, the impact of the proposed correction is dramatic, as the large bias in the estimated return levels (about 40 % overestimation in the 50-year return levels, in median) is adjusted and the dispersion largely reduced. Taking the 50-year return levels, the FSE of the least dense validation precedure (1/200 km^2) is 7.1 %, while the FSE for the orignal not corrected quantile is equal to 52 % (Table 1). It is interesting to note how the large dispersion in the estimation of the average yearly number of wet days does not impact the estimated return levels. This is related to the rather low sensitivity of return levels to this parameter (see Appendix A).

The most difficult case study is the ARID one, due to the lower density of stations, the larger climatic variability, and the less explanatory power of the linear relations between the satellite error statistics and the errors in the SMEV parameters (see above). Despite this, the proposed corrections are reasonably accurate also in this case (Fig. 8). This is true especially for the shape and scale parameters, which show a good match of the median value and an improved dispersion. The average number of wet days, originally characterized by satellite underestimation due to the low number of stations and the high noise of the satellite data in the area, is drastically reduced. Overall, the proposed correction reduces the bias to values generally within 10–15 % even for return levels as high as 50-years and for rain gauge densities as low as 1 station per 1000 km^2, with a reduction in the FSE from 44 % of the original satellite estimates to about 22 % of the adjusted estimates. Interestingly, the accuracy of the adjusted estimates depends only weakly on the rain gauge density. This means that the proposed adjustment method is robust in presence of sparse networks, such as the ones available in poorly gauged regions.

The residual dispersion is related to different factors: i) the use of linear relations, which leads to inevitable approximations, ii) the use of relatively short records (14 years for the Austria rain gauge network, 20 years for IMERG), which entails stochastic uncertainties related to natural variabiliy both in our estimates and in the reference. It is difficult to contrast the results of our method with other approaches in literature. In fact, previous studies (e.g., Pombo and de Oliveira, 2015; Demirdjian et al., 2018) worked with traditional extreme value methods and derived empirical corrections for the annual maxima or peaks over threshold and their distribution without the possibility to connect those to satellite estimation error statistics. Conversely, the method proposed by Marra et al. (2022) for radar estimates, also based on non-asympotic approaches, relies on the interpolation of empirical correction factors for the parameters. As such it requires a rather high-density of rain gauges and does not allow to transfer information across similar regions.

4. Conclusions

We propose a new method to derive accurate extreme daily precipitation return levels from satellite data in poorly gauged regions. The method relies on the idea that satellite estimation errors, which can be derived from poor gauge information, can be used to correct the parameters of a non-asymptotic extreme value distribution. To do so, we use the SMEV non-asymptotic statistical framework, and we derive linear relations between errors in the satellite estimation of the L-moments of the ordinary events distribution and errors in the SMEV parameters. These relations can be obtained from poor rain gauge coverage or from regions with similar climate and satellite estimation error characteristics and with available rain gauge information. Compared to previous studies in which the estimates of extremes (e.g., annual maxima or their distribution) were corrected, we work in the parameter space of the ordinary events distribution, under the assumption that these are related rather directly to satellite estimation errors.

In Austria, 75 % of the variance in the SMEV-Weibull scale parameter bias is explained by the bias in the first L-moment, that is the mean daily precipitation amount of the wet days. This value increases to approximately 90 % for the Southern MED and ARID areas. As much as 35 % (ARID), 40 % (MED) and 60 % (Austria) of the variance in the SMEV-Weibull shape parameter error is explained by errors in the third L-moment ratio (i.e., the L-skewness).

These approximate relations allow us to develop a method to correct the satellite estimates of daily precipitation return levels related to 2, 5, 10 and 50 years return periods. The corrections prove to be highly accurate for each of the three climatic areas, removing the median biases and drastically reducing the dispersion, as indicated by the Fractional Standard Error, although the three areas were characterised by diverse satellite estimation errors. The accuracy of the method is rather independent of the available rain gauge density. This indicates that the proposed approach could help adjusting satellite estimates of extreme precipitation return levels in regions of the globe characterized by poor rain gauge coverage.

CRediT authorship contribution statement

Matteo Siena: Conceptualization, Data curation, Writing – original draft, Writing – review & editing. **Vincenzo Levizzani:** Supervision, Funding acquisition, Writing – review & editing. **Francesco Marra:** Conceptualization, Data curation, Supervision, Writing – review & editing.

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

Links to codes and data are mentioned in the paper

Acknowledgments

The authors thank the developers and maintainers of the IMERG dataset, Jürgen Fuchsberger and the developers and maintainers of the Wegener Network, and Yoav Levi from Israel Meteorological Service and Efrat Morin from the Hebrew University of Jerusalem for creating and providing the rain gauge data used in this study. MS was supported by joint funds of the University of Bologna and Atos Italia SpA. FM was supported by the CARIPARO Foundation through the Excellence Grant 2021 to the “Resilience” Project and by internal projects of CNR-ISAC.

Appendix A. Propagation of the error from the SMEV parameters to the return levels

We provide here an analytical derivation of the error propagation from the parameters of the SMEV formulation to the estimated quantiles (Eq. (3)). Applying a second order Taylor series expansion, the error on the quantiles δq can be written as:

$$\delta q = \left(\frac{\partial q}{\partial \lambda}\right)\delta\lambda + \left(\frac{\partial q}{\partial \vartheta}\right)\delta\vartheta + \left(\frac{\partial q}{\partial n}\right)\delta n + \frac{1}{2}\left[\left(\frac{\partial^2 q}{\partial \lambda^2}\right)\delta\lambda^2 + \left(\frac{\partial^2 q}{\partial \vartheta^2}\right)\delta\theta^2 + \left(\frac{\partial^2 q}{\partial n^2}\right)\delta n^2\right] \tag{A1}$$

where $\delta\lambda, \delta\vartheta$, δn are the errors on the estimated scale and shape parameters and on the average number of wet days, respectively. Because of the approximation, this expansion is well suited for small errors, while greater deviations are expected for larger errors. Specifically, the first derivatives of Eq. (3) can be written as follows:

$$\frac{dq|_\lambda}{q}\delta\lambda = \frac{\left[-ln\left(1-P^{\frac{1}{n}}\right)\right]^{\frac{1}{\vartheta}}\delta\lambda}{q} = \frac{\delta\lambda}{\lambda} \tag{A2}$$

$$\frac{dq|_\vartheta}{q}\delta\vartheta = -\frac{\ln\left[\left(\frac{q}{\lambda}\right)^{\vartheta}\right]}{\vartheta}\frac{\delta\vartheta}{\vartheta} \tag{A3}$$

$$\frac{dq|_n}{q}\delta n = -\frac{\lambda^{\vartheta}}{q^{\vartheta}\vartheta n}\left[\frac{p^{\frac{1}{n}}\ln(p)}{1-p^{\frac{1}{n}}}\right]\frac{\delta n}{n} \underset{p\to 1}{\to} \frac{\lambda^{\vartheta}}{q^{\vartheta}\vartheta}\frac{\delta n}{n} \tag{A4}$$

where the last one is simplified to a more treatable expression in the limit $p\to1$, that is the case of extreme quantiles.

Again, considering the limit $p\to1$, we find that the second order contributions of the scale parameter is null, and the one of n can be neglected:

$$\frac{dq^2|_\lambda}{q} = 0 \tag{A5}$$

$$\frac{dq^2|_n}{q} = -\frac{1}{q}\frac{\lambda p^{\frac{1}{n}} ln(p)\left(\frac{q}{\lambda}\right)^{1-2\vartheta}}{\vartheta^2 n^4\left(p^{\frac{1}{n}}-1\right)^2}\left\{ln(p)\left[(\vartheta-1)p^{\frac{1}{n}} - \vartheta\left(\frac{q}{\lambda}\right)^{\vartheta}\right] + 2\vartheta n\left(p^{\frac{1}{n}}-1\right)\left(\frac{q}{\lambda}\right)^{\vartheta}\right\} \underset{p\to1}{\ll} 1 \tag{A6}$$

Thus, the only second order term which influences the series expansion is

$$\frac{dq^2|_\vartheta}{q} = \frac{\ln\left[\left(\frac{q}{\lambda}\right)^\vartheta\right]\left\{2\vartheta + \ln\left[\left(\frac{q}{\lambda}\right)^\vartheta\right]\right\}}{\vartheta^4}\delta\vartheta^2 \tag{A7}$$

Overall, defining $y = \ln\left[\left(\frac{q}{\lambda}\right)^\vartheta\right]$ the error on the estimated quantiles in presence of small errors in the parameters can be written as:

$$\frac{\delta q}{q} \simeq \frac{\delta\lambda}{\lambda} - \frac{y}{\vartheta}\frac{\delta\vartheta}{\vartheta} - \frac{\lambda^\vartheta}{\vartheta q^\vartheta}\frac{\delta n}{n} + \frac{y(2\vartheta + y)}{2\vartheta^2}\frac{\delta\vartheta^2}{\vartheta^2} \tag{A8}$$

References

Akram, M., Hayat, A., 2014. Comparison of estimators of the Weibull distribution. Journal of Statistical Theory and Practice 8, 238–259. https://doi.org/10.1080/15598608.2014.847771.

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, 71–99. https://doi.org/10.5194/essd-5-71-2013.

Berg, P., Moseley, C., Haerter, J.O., 2013. Strong increase in convective precipitation in response to higher temperatures. Nature Geoscience Letters 6, 181–185. https://doi.org/10.1038/ngeo1731.

Berne, A., Krajewski, W.F., 2013. Radar for hydrology: unfulfilled promise or unrecognized potential? Advances in Water Resources 51, 357–366. https://doi.org/10.1016/j.advwatres.2012.05.005.

Chen, H., Yong, B., Qi, W., Wu, H., Ren, L., 2020. Investigating the evaluation uncertainty for satellite precipitation estimates based on two different ground precipitation observation products. Journal of Hydrometeorology 21, 2595–2606. https://doi.org/10.1175/JHM-D-20-0103.1.

Coles, S., 2001. An introduction to statistical modeling of extreme values. Springer-Verlag, London, p. 209 pp. https://doi.org/10.1007/978-1-4471-3675-0.

Demirdjian, L., Zhou, Y., Huffman, G.J., 2018. Statistical modelling of extreme precipitation with TRMM data. Journal of Applied Meteorology and Climatology 57, 15–30. https://doi.org/10.1175/JAMC-D-17-0023.1.

Fischer, R.A., Tippet, L.H.C., 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proceedings of the Cambridge Philosophical Society 24, 180–190. https://doi.org/10.1017/S0305004100015681.

Fuchsberger, J., Kirchengast, G., Bichler, C., Leuprecht, A., Kabas, T. (2021). WegenerNet climate station network Level 2 data version 7.1 (2007–2020). *Wegener Center for Climate and Global Change*, University of Graz, Austria, https://wegenernet.org/portal/v7.1/2021/1.

Fuchsberger, J., Kirchengast, G., Kabas, T., 2021b. WegenerNet high-resolution weather and climate data from 2007 to 2020. Earth System Science Data 13, 1307–1334. https://doi.org/10.5194/essd-13-1307-2021.

Gado, T.A., Hsu, K., Sorooshian, S., 2017. Rainfall frequency analysis for ungauged sites using satellite precipitation products. Journal of Hydrology 554, 646–655. https://doi.org/10.1016/j.jhydrol.2017.09.043.

Goda, Y., Kudaka, M., Kawai, H., 2011. Incorporation of Weibull distribution in L-moments method for regional frequency analysis of peaks-over-threshold wave heights. Coastal Engineering Proceedings 1 (32). https://doi.org/10.9753/icce.v32.waves.62.

Gumbel, E.J., 1958. Statistics of extremes. Columbia University Press. https://doi.org/10.7312/gumb92958.

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. https://doi.org/10.1111/j.2517-6161.1990.tb01775.x.

Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K., Iguchi, T., 2014. The Global Precipitation Measurement mission. Bulletin of the American Meteorological Society 95, 701–722. https://doi.org/10.1175/BAMS-D-13-00164.1.

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

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Wolff, D.B., Adler, R.F., Gu, G., Hong, Y., Bowman, K.P., Stocker, E.F., 2007. The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. Journal of Hydrometeorology 8, 38–55. https://doi.org/10.1175/JHM560.1.

Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E.J., Sorooshian, S., Tan, J., Xie, P. (2020). NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). *Algorithm Theoretical Basis Document (ATBD) Version 06*. [available at https://gpm.nasa.gov/sites/default/files/2020-05/IMERG_ATBD_V06.3.pdf, last accessed May 2023].

Katz, R., Parlange, M., Naveau, P., 2002. Statistics of extremes in hydrology. Advances in Water Resources 25, 1287–1304. https://doi.org/10.1016/S0309-1708(02)00056-8.

Kidd, C., Levizzani, V., 2011. Status of satellite precipitation retrievals. Hydrology and Earth System Sciences 15, 1109–1116. https://doi.org/10.5194/hess-15-1109-2011.

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 (1), 69–78. https://doi.org/10.1175/BAMS-D-14-00283.1.

Köppen, W., Geiger, R., 1936. Das geographische System der Klimate. Handbuch der Klimatologie vol. 1.

Levizzani, V., Kidd, C., Kirschbaum, D.B., Kummerow, C.D., Nakamura, K., Turk, F.J., Eds. (2020a). Satellite Precipitation Measurement. Vol. 1, Springer Nature, Cham, Switzerland, *Advances in Global Change Research*, 67, 450 pp, doi:10.1007/978-3-030-24568-9.

Levizzani, V., Kidd, C., Kirschbaum, D.B., Kummerow, C.D., Nakamura, K., Turk, F.J., Eds. (2020b). Satellite Precipitation Measurement. Vol. 2, Springer Nature, Cham, Switzerland, *Advances in Global Change Research*, 69, 712 pp, doi:10.1007/978-3-030-35798-6.

Libertino, A., Allamano, P., Laio, F., Claps, P., 2018. Regional-scale analysis of extreme precipitation from short and fragmented records. Advances in Water Resources 112, 147–159. https://doi.org/10.1016/j.advwatres.2017.12.015.

Liu, J., Xia, J., She, D., Li, L., Wang, Q., Zou, L., 2019. Evaluation of six satellite-based precipitation products and their ability for capturing characteristics of extreme precipitation events over a climate transition area in China. Remote Sensing 11, 1477. https://doi.org/10.3390/rs11121477.

Maggioni, V., Massari, C., Kidd, C. (2021). Errors and uncertainties associated with quasiglobal satellite precipitation products. In Michaelides (Ed.), *Precipitation Science. Measurement, Remote Sensing, Microphysics and Modeling*, Elsevier, Amsterdam, 377-390, doi:10.1016/B978-0-12-822973-6.00023-8.

Marani, M., Ignaccolo, M., 2015. A metastatistical approach to rainfall extremes. Advances in Water Resources 79, 121–126. https://doi.org/10.1016/j.advwatres.2015.03.001.

Marra, F., 2020. A unified framework for extreme sub-daily precipitation frequency analyses based on ordinary events - data & codes (Version v1). Zenodo. https://doi.org/10.5281/zenodo.3971558.

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. https://doi.org/10.1016/j.advwatres.2018.05.001.

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

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

Marra, F., Borga, M., Morin, E., 2020. A unified framework for extreme subdaily precipitation frequency analyses based on ordinary events. Geophysical Research Letters 47(18), e2020GL090209. https://doi.org/10.1029/2020GL090209.

Marra, F., Armon, M., Adam, O., Zoccatelli, D., Gazal, O., Garfinkel, C., Rostkier-Edelstein, D., Dayan, U., Enzel, Y., Morin, E., 2021. Toward narrowing uncertainty in future projections of local extreme precipitation. Geophysical Research Letters 48 (5), e2020GL091823. https://doi.org/10.1029/2020GL091823.

Marra, F., Amponsah, W., Papalexiou, S.M., 2023. Non-asymptotic Weibull tails explain the statistics of extreme precipitation. Advances in Water Resources 173 (2), 104388. https://doi.org/10.1016/j.advwatres.2023.104388.

Marra, F., Armon, M., Morin, E. (2022). Coastal and orographic effects on extreme precipitation revealed by weather radar observations. Hydrology and Earth System Sciences, 26, 1439-1458, doi:105194/hess-26-1439-2022.

Marra, F. (2022). A test for the hypothesis: block maxima are samples from a parent distribution with Weibull tail. (Version v1). *Zenodo*, doi:10.5281/zenodo.7234708.

Mianabadi, A., 2023. Evaluation of long-term satellite-based precipitation products for developing intensity-frequency (IF) curves of daily precipitation. Atmospheric Research 286, 106667. https://doi.org/10.1016/j.atmosres.2023.106667.

Miao, C., Ashouri, H., Hsu, K., Sorooshian, S., Duan, Q., 2015. Evaluation of the PERSIANN-CDR daily rainfall estimates in capturing the behaviour of extreme precipitation events over China. Journal of Hydrometeorology 16, 1387–1396. https://doi.org/10.1175/JHM-D-14-0174.1.

Miniussi, A., Marra, F., 2021. Estimation of extreme daily precipitation return levels at-site and in ungauged locations using the simplified MEV approach. Journal of Hydrology 603, 126946. https://doi.org/10.1016/j.jhydrol.2021.126946.

Müller, M.F., Thomson, S.E., 2013. Bias adjustment of satellite rainfall data through stochastic modeling: Methods development and application to Nepal. Advances in Water Resources 60, 121–134. https://doi.org/10.1016/j.advwatres.2013.08.004.
NASA, IMERG V06: Documentation, https://gpm.nasa.gov/data/imerg.
Papalexiou, S.M., 2022. Rainfall generation revisited: introducing cosmos-2 s and advancing copula-based intermittent time series modeling. Water Resources Research 58(6), e2021WR031641. https://doi.org/10.1029/2021WR031641.
Papalexiou, S.M., Koutsoyiannis, D., 2013. Battle of extreme value distributions: a global survey on extreme daily rainfall. Advances in Water Resources 49, 187–201. https://doi.org/10.1029/2012WR012557.
Peel, M. C., Finlayson, B. L., McMahon, T. A. (2007). Updated world map of the Köppen–Geiger climate classification, *Hydrology and Earth System Sciences*, 11, 1633–1644, doi:10.5194/hess-11-1633- 2007.
Pombo, S., de Oliveira, R.P., 2015. Evaluation of extreme precipitation estimates from TRMM in Angola. Journal of Hydrology 523, 663–679. https://doi.org/10.1016/j.jhydrol.2015.02.014.
Poschlod, B., 2021. Using high-resolution regional climate models to estimate return levels of daily extreme precipitation over Bavaria. Natural Hazards and Earth System Sciences 21, 3573–3598. https://doi.org/10.5194/nhess-21-3573-2021.
Prakash, S., Mitra, A.K., Pai, D.S., AghaKouchak, A., 2015. From TRMM to GPM: how well can heavy rainfall be detected from space? Advances in Water Resources 88, 1–7. https://doi.org/10.1016/j.advwatres.2015.11.008.
Schellander, H., Lieb, A., Hell, T., 2019. Error structure of Metastatistical and Generalized Extreme Value distributions for modeling extreme rainfall in Austria. Earth and Space Science 6, 1616–1632. https://doi.org/10.1029/2019EA000557.
Sharon, D., 1972. The spottiness of rainfall in a desert area. Journal of Hydrology 17 (3), 161–175. https://doi.org/10.1016/0022-1694(72)90002-9.
Smith, T.M., Arkin, P., Bates, J.J., Huffman, G.J., 2006. Estimating bias of satellite-based precipitation estimates. Journal of Hydrometeorology 7 (5), 841–856. https://doi.org/10.1175/JHM524.1.
Sun, Q., Miao, C., Duan, Q., Ashouri, A., Sorooshian, S., Hsu, K.-L., 2018. A review of global precipitation data sets: Data sources, estimation and intercomparisons. Reviews of Geophysics 56, 79–107. https://doi.org/10.1002/2017RG000574.
Vogel, R.M., Fennessey, N.M., 1993. *L* moment diagrams should replace product moment diagrams. Water Resources Research 29 (6), 1745–1752. https://doi.org/10.1029/93WR00341.
Wang, L., Marra, F., Onof, C. (2020). Modelling sub-hourly rainfall extremes with short records – a comparison of MEV, Simplified MEV and point process methods. *European Geosciences Union (EGU) General Assembly 2020 (Online)*. [available at https://presentations.copernicus.org/EGU2020/EGU2020-6061_presentation.pdf, last accessed May 2023].
WegenerNet DataPortal. Stations location info and documentation, https://wegenernet.org/portal/(last accessed May 2023).
Weibull, W., 1951. A statistical distribution function of wide applicability. Journal of Applied Mechanics 18, 239–296. https://doi.org/10.1115/1.4010337.
Wilson, P.S., Toumi, R., 2005. A fundamental probability distribution for heavy rainfall. Geophysical Research Letters 32, L14812. https://doi.org/10.1029/2005GL022465.
Wright, D.B., Smith, J.A., Villarini, G., Baeck, M.L., 2013. Estimating the frequency of extreme rainfall using weather radar and stochastic storm transposition. Journal of Hydrology 488, 150–165. https://doi.org/10.1016/j.jhydrol.2013.03.003.
Zorzetto, E., Botter, G., Marani, M., 2016. On the emergence of rainfall extremes from ordinary events. Geophysical Research Letters 43, 8076–8082. https://doi.org/10.1002/2016GL069445.
Zorzetto, E., Marani, M., 2019. Downscaling of rainfall extremes from satellite observations. Water Resources Research 55, 156–174. https://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. Advances in Water Resources 135, 103483. https://doi.org/10.1016/j.advwatres.2019.103483.