# Spatial variability of precipitation extremes over Italy using a fine-resolution gridded product

Benedetta Moccia [a],[*], Simon Michael Papalexiou [b],[c], Fabio Russo [a], Francesco Napolitano [a]

[a] *Dipartimento di Ingegneria Civile, Edile e Ambientale, Sapienza Università di Roma, Rome, Italy*
[b] *Department of Civil Engineering, University of Calgary, Canada*
[c] *Global Institute for Water Security, University of Saskatchewan, Canada*

A R T I C L E   I N F O

A B S T R A C T

*Study region:* Italy.

*Study focus:* Knowing magnitude and frequency of extreme precipitation is necessary to reduce their impact on vulnerable areas. Here we investigate the performance of the Generalized Extreme Value ($\mathscr{GEV}$) distribution, using a fine-resolution satellite-based gridded product, to analyze 13,247 daily rainfall annual maxima samples. A non-extreme value distribution with a power-type behavior, that is, the Burr Type XII ($\mathscr{Br}$-XII), is also evaluated and used to test the reliability of the $\mathscr{GEV}$ in describing extreme rainfall.

*New hydrological insights for the region:* (1) in 44.9 % of the analyzed samples the $\mathscr{GEV}$ predicts an upper rainfall limit; we deem this is an artifact due to sample variations; (2) we suggest the $\mathscr{GEV}^+$ distribution, that is, the $\mathscr{GEV}$ with shape parameters restricted only to positive values as a more consistent model complying with the nature of extreme precipitation; (3) $\mathscr{GEV}$, $\mathscr{GEV}^+$, and $\mathscr{Br}$XII performed equally well in describing the observed annual precipitation, yet all distributions underestimate the observed sample maximum; (4) the $\mathscr{Br}$XII, for large return periods, predicts larger rainfall amounts compared to $\mathscr{GEV}$ indicating that $\mathscr{GEV}$ estimates could underestimate the risk of extremes; and (5) the correlation between the predicted rainfall and the elevation is investigated. Based on the results of this study, we suggest instead of using the classical $\mathscr{GEV}$ to use the $\mathscr{GEV}^+$ and non-extreme value distributions such as the $\mathscr{Br}$XII to describe precipitation extremes.

## 1. Introduction

Designing flood protection and hydraulic infrastructure requires the use of high return periods (or low probability levels) associated with extreme rainfall values. Describing the statistical behavior of extreme events is crucial to limit the risk of failure of civil engineering works and, consequently, to limit the loss of human lives. Deterministic long-term predictions of rainfall are not feasible, thus, using probabilistic methods is a necessity (Papalexiou et al., 2013); however, finding a suitable probability distribution to describe rainfall extremes can be challenging. Choosing a wrong distribution or misfitting it may lead to estimation errors of rainfall quantiles and, consequently, underestimate (or even overestimate) flood risk.

---

* Corresponding author.
  *E-mail address:* benedetta.moccia@uniroma1.it (B. Moccia).

According to the extreme value theory (EVT), the largest values extracted from a set of independent identically distributed random variables tends to an asymptotic distribution that depends on the tail of the distribution of the parent variable (Fisher and Tippett, 1928; Gnedenko, 1943; Gumbel, 1958). EVI (Gumbel), EVII (Fréchet) and EVIII (reversed Weibull) are the three possible asymptotic laws that can be unified into the single expression known as the Generalized Extreme Value ($\mathscr{GEV}$) distribution (von Mises, 1936). The $\mathscr{GEV}$ is widely used worldwide, even though some authors showed a certain skepticism to use it for modelling daily rainfall (De Michele and Avanzi, 2018; Zorzetto et al., 2016).

Coles et al. (2003) suggested the use of the $\mathscr{GEV}$ model, over the Gumbel distribution, for accounting of the uncertainty in the analysis of daily rainfall in Maiqueta (Venezuela). Lee and Maeng (2003) selected the $\mathscr{GEV}$ as the appropriate probability distribution to describe 38 annual maximum daily rainfall samples in Korea. Chu et al. (2009) suggested the use of the Gumbel distribution to describe annual maximum daily rainfall series since, by fitting the $\mathscr{GEV}$ to 158 Hawaiian samples, the shape parameter resulted close to zero. Smith et al. (2011) fitted the $\mathscr{GEV}$ distribution to annual maximum daily rainfall series recorded in the central section of the eastern United States, showing a positive shape parameter, and thus heavy tails for extreme rainfall. Villarini (2012) analyzed annual and seasonal rainfall samples recorded by 44 rain gauges located in East Europe (Ukraine, Moldova, and Romania), using the $\mathscr{GEV}$ distribution to examine the upper tail properties at both time scales. The author shows that the Fréchet distribution performed better than the other two distributions, highlighting a heavy-tail behavior. In a global analysis of 15,137 records Papalexiou and Koutsoyiannis (2013) investigated which of the three limiting types of the $\mathscr{GEV}$ best describes the annual daily maxima rainfall. The study revealed that the Fréchet distribution performs best, while bounded from above distributions (e.g., the reverse Weibull) emerge as artifact from sample variation and are justified. Yang et al. (2013) found a correlation between extreme rainfall and elevation in northern China, by fitting the $\mathscr{GEV}$ to seasonal extreme rainfall samples. Nguyen and Nguyen (2016) considered the $\mathscr{GEV}$ as the most suitable model to describe annual maximum rainfall series for 21 stations located in Ontario (Canada).

The previous studies use rain gauge records. However, many regions around the globe remain ungauged or have a sparce gauge network (Bertini et al., 2020; Mishra and Coulibaly, 2009); also, in some countries, retrieving such data may be difficult due to strict regulations and bureaucracy. For example, the Italian peninsula has a dense rain-gauge network comprising 5265 station over 301,000 km$^2$, yet an open dataset is not available (Libertino et al., 2018; Mazzoglio et al., 2020). Over the last decades, to overcome the limited data availability, the use of satellite precipitation datasets offering global coverage, high resolution and accessibility, has become popular (Rajulapati et al., 2020; Sun et al., 2018).

In this paper we analyze extreme daily rainfall in Italy using a satellite-based fine-resolution gridded product, that is, the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) dataset (Funk et al., 2015b, a). This dataset has been already tested and evaluated in Italy (Duan et al., 2016) and in many other countries such as Argentina (Rivera et al., 2018), Brazil (Paredes-Trejo et al., 2017), China (Bai et al., 2018; Tang et al., 2020), India (Prakash, 2019), and Cyprus (Katsanos et al., 2016). Such studies corroborate the reliability of the CHIRPS dataset to reproduce different characteristics of precipitation, depending on the applied region. For instance, Duan et al. (2016) compared daily data recorded by 101 rain gauges in Italy with eight gridded products in the Adige Basin (North-East). Despite inconsistencies found in all the eight products in the frequency of occurrence of daily precipitation, especially in
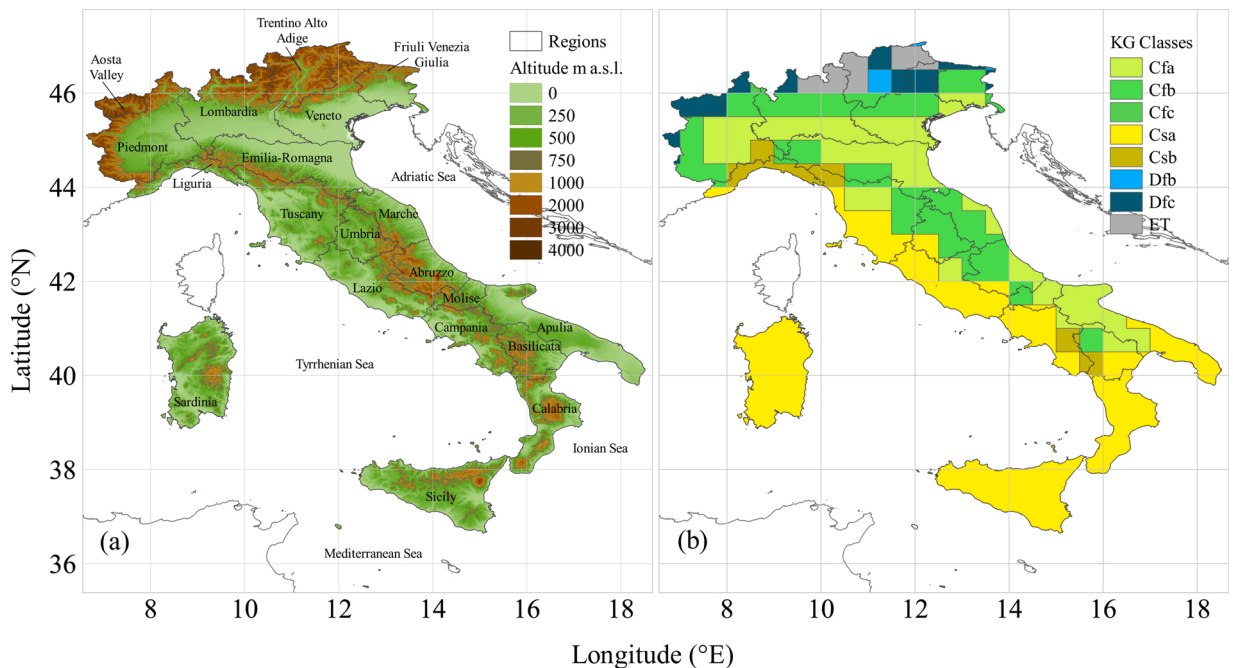


**Fig. 1.** Elevation map of Italy (a), and the Köppen-Geiger classification based on the 1976-2000 period observations (b); the grey lines identify the 20 Italian Regions.

the winter months, CHIRPS, TRMM 3B42 (Huffman et al., 2006), and CMORPH_BLD (Xie and Xiong, 2011) provided the best performances in terms of overall statistics metrics. Also, Caroletti et al. (2019) compared gridded data with ground data for the Calabria Region (South-West of Italy) using monthly rainfall data collected by rain gauges. Their findings confirm that the satellite-based CHIRPS product provides the best performances in terms of Pearson correlation, error metrics and extreme events.

Here, we use the CHIRPS product to investigate the spatial variation of the $\mathscr{GEV}$ distribution and its performance in describing annual maxima over Italy. We compare its performance with a non-extreme value distribution, that is, the Burr type XII ($\mathscr{Br}$XII) (Burr, 1942) to show that the $\mathscr{GEV}$ distribution should not be used blindly. Specifically, we investigate the exceedance probability related to the upper limit predicted by the $\mathscr{GEV}$ distribution when it converges to the reverse Weibull. Using a distribution bounded from above stands contradicts the nature of many hydrological variables and might lead to severe risk underestimation. Instead, we propose fitting the $\mathscr{GEV}$ distribution excluding the reverse Weibull case. The fitting results show that the shape parameter of the $\mathscr{GEV}$ distribution forms coherent spatial patterns. We use these results to build rainfall depth maps for different return periods, that could be useful both for research and engineering practice.

## 2. The data

### 2.1. Study area

Italy is in southern Europe, between latitudes 35.475 °N and 47.125 °N and longitudes 6.625 °E and 18.525 °E. The Mediterranean peninsula is characterized by a complex orography that affects the climate (Fig. 1).

Italy's climate, based on the Köppen-Geiger (KG) classification (Beck et al., 2018; Kottek et al., 2006; Peel et al., 2007), is mainly temperate (C), and cold (D) or polar (E) at higher altitudes. The climate of west coast and of the two major islands, Sicily and Sardinia, is the so-called Mediterranean (Csa-KG), while parts of the inland have dry and warm summers and rainy winters (Csb-KG). Most of the
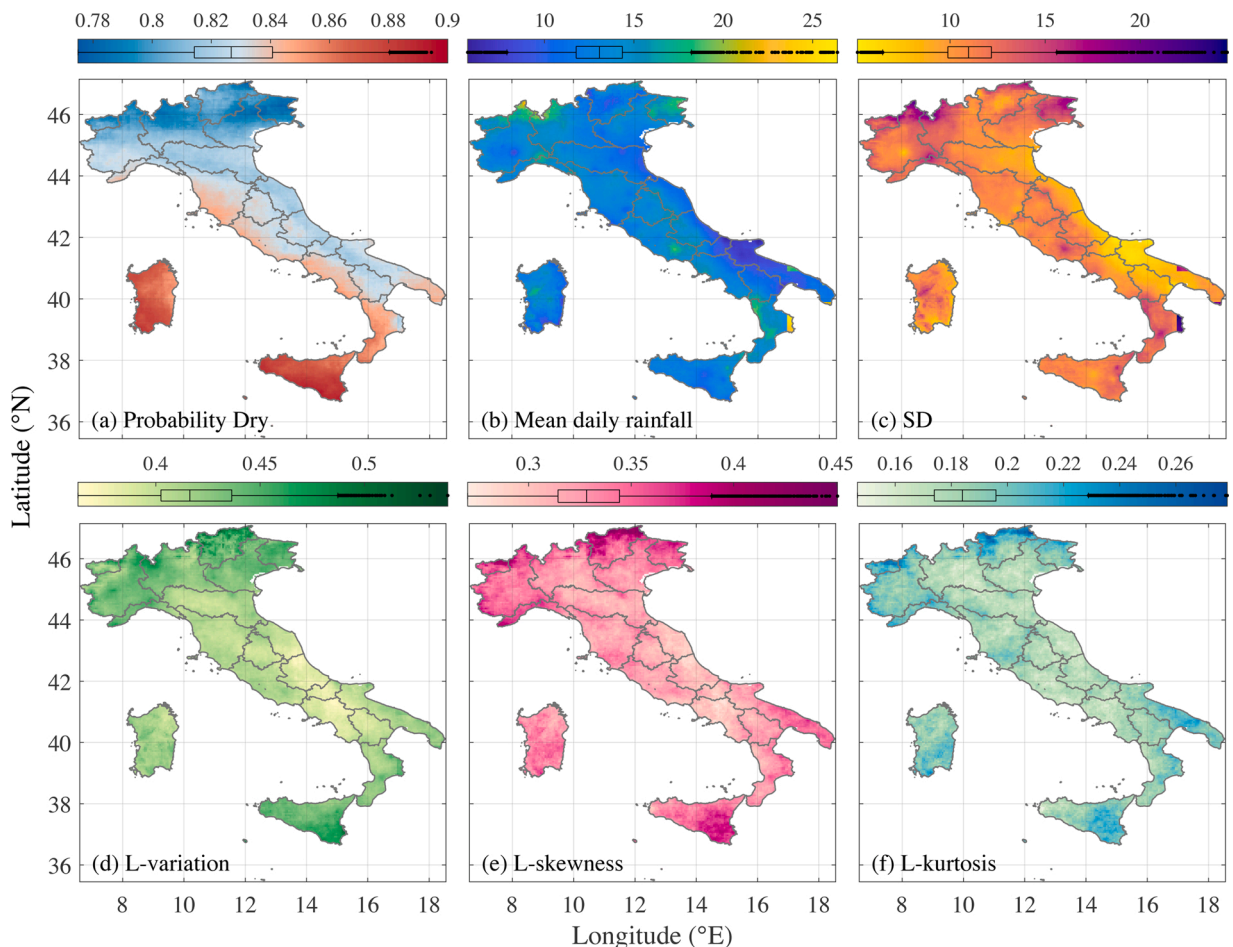


**Fig. 2.** Spatial representation of the estimated statistics calculated for the 13,247 time series. Except for the probability dry, all these statistics were calculated on the continuous part of the marginal distribution of daily rainfall. The boxplot in each panel represents the statistic's variability. The box lines represent the 25th, 50 (or median) and 75th percentile, while the fences are 1.5 times the interquartile range.

Adriatic coast, the Po valley and parts of the central Apennines have humid sub-littoral climate (Cfa-KG) with no dry season, which is characterized by two peaks of rainfall (fall and spring), hot summers and moderately cold winters. The sub-continental climate (Cfb-KG) characterizes the Apennine and the foothills of the Alps, with wet and warm summers. At higher altitudes, the climate is boreal, with no dry season and warm (Dfb-KG) and cold (Dfc-KG) summers, very cold winters, and snow between November and March. Along the Alps there is also the tundra climate (ET-KG) with low temperatures (below 10 °C) all the year.

The various climates of Italy indicate diverse monthly and annual precipitation but do not characterize differences in extremes. It is rational, however, to assume that such climatic conditions can affect the behavior of extremes. Here, we focus on the evaluation of extreme precipitation in Italy which often occurs as thunderstorms. For example, Telesca et al. (2008) identify three thunderstorm types in Italy: large fronts, generally coming from the North-West; local convective cells, typical of the summer season, and orographic thunderstorms. However, to classify the thunderstorms is out of the scope of this study and needs fine temporal resolution data (e.g., hourly); here we focus on the extremes at the daily scale.

### 2.2. Dataset

The Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) is a quasi-global (50 °S-50 °N and all longitudes) precipitation dataset (Funk et al., 2015b, a). CHIRPS provides fine spatial resolution daily rainfall at 0.05 ° (approximately 5 km) in the 1981-present period; it calibrates infrared cold cloud duration observations with the Tropical Rainfall Measuring Mission Multi-Satellite Precipitation Analysis version 7 (TMPA 3B42 v7) and then incorporates ground data. Literature studies show that CHIRPS provides a fine-resolution precipitation measurements and is reliable for Italy (Caroletti et al., 2019; Duan et al., 2016). At the 5 km spatial resolution there are 13,247 grid cells covering Italy, and thus, we extracted and analyzed 13,247 time series in the period 1981−2019.
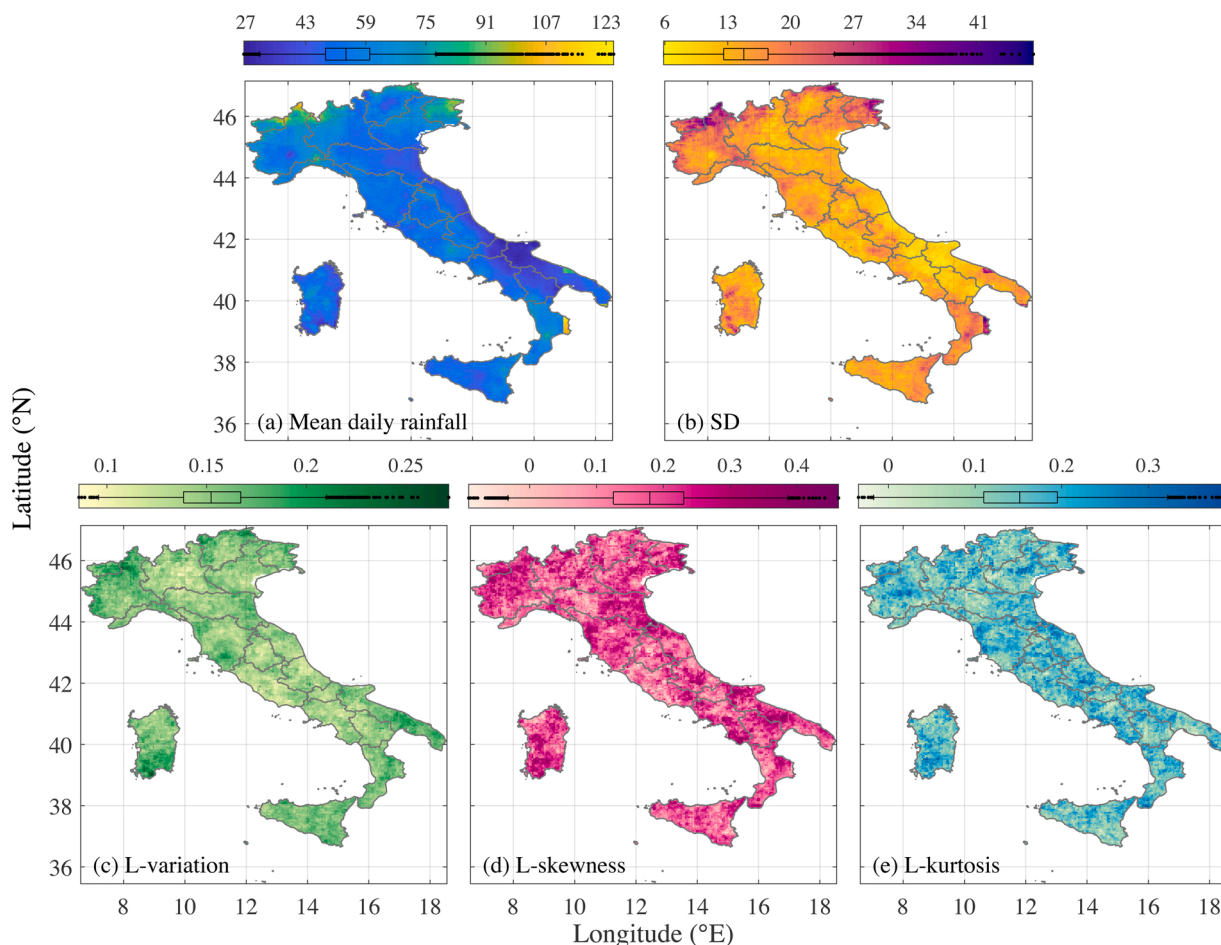


**Fig. 3.** Spatial representation of the estimated statistics calculated for the 13,247 annual maxima samples. The boxplot in each panel represents the statistic's variability. The box lines represent the 25th, 50th (or median) and 75th percentile, while the fences are 1.5 times the interquartile range.

*2.3. Main statistics of the CHIRPS product*

The marginal distribution of daily rainfall is mixed having thus a discrete and a continuous part describing, respectively, the probability dry and the non-zero values in wet days (Papalexiou, 2018). For each of the 13,247 grid cells we calculate the probability dry (ratio of dry to total days), the mean value, the standard deviation (SD), and the three L-ratios (L-variation, L-Skewness and L-kurtosis) (Hosking, 1990). The mean daily value and the standard deviation represent, respectively, measures of central tendency and dispersion, while the three L-ratios are dimensionless and represent measures of distributional shape. The west coastal areas and the two major islands are dry, with a probability dry ranging from 0.83 up to 0.89, while the greatest values of the mean daily rainfall, that is, greater than the 75th percentile of 14.3 mm, are spread at higher altitude, with some spots on the west coast (Fig. 2). A high amount of daily precipitation in dry areas could be due to unstable weather conditions, especially during the summer, which cause several convective storms (Davini et al., 2012; Telesca et al., 2008). The east coast, the Po valley and some parts of Trentino, Piedmont, Sicily and of the east coast of Sardinia register the lowest values of the mean rainfall (smaller than the 25th percentile of about 11.7 mm per day). The standard deviation presents the same spatial patterns observed for the mean rainfall (with a Pearson correlation coefficient of 0.93). The L-variation, which is a standardized measure of variance given by the ratio of the L-scale and the central tendency, shows high values (greater than the median of 0.416) in all the Alpine and sub-alpine zones and in the south, especially in Sicily and in the south of Apulia. Same patterns are observed for the L-Skewness and L-kurtosis, with high values (greater than the median) observed also in Tuscany and Lazio.

To summarize and assess the statistical behavior of extreme precipitation, we evaluate those statistics also for daily annual maxima samples (AM) extracted from each 13,247 grid-cells (Fig. 3). The mean daily rainfall exhibits the lowest values (less than the median of about 53.7 mm) on the east coast and in the two major islands, and in some spots of Piedmont, Trentino Alto Adige, and central Italy. The resulting standard deviation shows the minimum values (less than the median of about 14.7 mm) in the south-east areas, and in some parts of Lombardia, Lazio, Tuscany and north of Sardinia. Essentially, at lower values of mean daily precipitation are associated low values of the standard deviation. The spatial variability of the three L-ratios provides a first view of the statistical behavior of the extreme events, because both the magnitude and the frequency of these events are controlled by the shape characteristics of precipitation distribution (Papalexiou and Koutsoyiannis, 2016). The variability of the L-variation presents the lower values (less than the 25th percentile of about 0.13) in the centre of Italy and in Lombardia, while the heavier values are spread in Piedmont, in south Apulia and in south Sardinia. L-Skewness shows a wide variability; indeed, we notice both negative and positive values. The variability of these main statistics evaluated on the annual maxima samples all over Italy may imply the presence of spatial patterns of precipitation extremes.

## 3. Methodology

### 3.1. Extreme value theory and generalized extreme value distribution

The $\mathscr{GEV}$ distribution, firstly introduced by von Mises (1936), is widely adopted to study extreme events; its cumulative distribution function (cdf) is given by

$$F_{\mathscr{GEV}}(x; \alpha, \beta, \gamma) = \exp\left( -\left( 1 + \gamma\frac{x - \alpha}{\beta} \right)^{-1/\gamma} \right), \; 1 + \gamma\frac{x - \alpha}{\beta} \geq 0 \tag{1}$$

where $\alpha \in \mathbb{R}, \beta > 0$, and $\gamma \in \mathbb{R}$ are, respectively, location, scale, and shape parameters. Depending on the value of the shape parameter, the $\mathscr{GEV}$ distribution encompasses the three limiting distributions of the Extreme Value Theory (EVT): the type I or Gumbel ($\mathscr{G}$), the type II or Fréchet ($\mathscr{F}$), and the type III or reverse Weibull ($\mathscr{rW}$). Their cdfs are given by

$$F_{\mathscr{G}}(x; \alpha, \beta) = \exp\left( -\exp\left( -\frac{x - \alpha}{\beta} \right) \right), \; x \in \mathbb{R} \tag{2}$$

$$F_{\mathscr{F}}(x; \alpha, \beta, \gamma) = \exp\left( -\left( \frac{x - \alpha}{\beta} \right)^{-1/\gamma} \right), \; x \geq \alpha, \tag{3}$$

$$F_{\mathscr{rW}}(x; \alpha, \beta, \gamma) = \exp\left( -\left( -\frac{x - \alpha}{\beta} \right)^{1/\gamma} \right), \; x \leq \alpha. \tag{4}$$

The three distributions have a location parameter $\alpha \in \mathbb{R}$ and a scale parameter $\beta > 0$, and the Fréchet and the reverse Weibull have also a shape parameter $\gamma > 0$. The Gumbel distribution is defined for $x \in \mathbb{R}$, Fréchet distribution is bounded from below and reverse Weibull is bounded from above. These three distributions are special cases of the $\mathscr{GEV}$ depending on the values of its shape parameter $\gamma$, that is, for $\gamma$ tending to zero, positive and negative, the Gumbel, Fréchet, and reverse Weibull laws emerge, respectively.

Note that the previous extreme value laws are limiting laws emerging under specific assumptions. Particularly, given a set of independent and identically distributed random variables $\{X_1, .., X_n\}$ described by the distribution $F_X(x)$, the distribution of the maximum $\max\{X_1, .., X_n\}$ is given by $F_{X_{\max}}(x) = (F_X(x))^n$. The three EV laws emerge for $n \to \infty$. However, the convergence to one of these three asymptotic distributions should not be taken for granted, because the length $n$ of observed samples is typically small

(Papalexiou and Koutsoyiannis, 2013). The type of tail of $F_X(x)$ (parent distribution) dictates to which one of the three limiting laws the distribution of the maximum converges. The right tail represents the upper part of a probability distribution and governs the behavior of extreme events (Moccia et al., 2021; Nerantzaki and Papalexiou, 2019); many tail classifications exist (El Adlouni et al., 2008; Ouarda et al., 1994; Werner and Upper, 2002). In general, power-type tails converge to the Fréchet law, exponential ones to the Gumbel, and bounded from above tails to the reversed Weibull. Thus, the Fréchet is a power-type distribution and all distributions with right tail regularly varying in infinity (i.e., power type distributions) belong to its domain of attraction. In contrast, the Gumbel distribution has exponential right tail that approaches to zero more rapidly than any power-type distribution.

Although $\mathscr{GEV}$ is widely used to describe extreme precipitation it has limitations; sample variations can lead to negative shape parameter estimates resulting in an upper bound distribution. An upper bounded distribution, such as the $\mathscr{rW}$, assumes a maximum precipitation value which cannot be exceeded while it is well documented that even probable maximum precipitation (PMP) estimates have been surpassed (Salas et al., 2020). Precipitation is a natural process limited at zero but an upper bound cannot be justified since there is always a probability to rain more.

Recently, several non-extreme value distributions were used in stochastic modelling to preserve the behaviour of extremes in hydroclimatic variables (Papalexiou, 2018; Papalexiou and Serinaldi, 2020). In this study we propose the employment of the Burr type XII ($\mathscr{Br}$XII) distribution, from the Burr system (Burr, 1942), in the formulation of Papalexiou and Koutsoyiannis (2012), with cdf given by

$$F_{\mathscr{Br}\text{XII}}(x;\beta,\gamma_1,\gamma_2) = 1 - \left(1 + \gamma_2\left(\frac{x}{\beta}\right)^{\gamma_1}\right)^{-\frac{1}{\gamma_1\gamma_2}} \tag{5}$$

where $\beta > 0$ is a scale parameter, and $\gamma_1 > 0$ and $\gamma_2 \geq 0$ are shape parameters controlling the left and right tails, respectively. Equivalently to the $\mathscr{GEV}$ for $\gamma > 0$, the asymptotic behavior of $\mathscr{Br}$XII is $\cong x^{-1/\gamma_2}$, therefore it has the potential to model heavy tails. The $\mathscr{Br}$XII includes as special cases the Pareto type II (for $\gamma_1 = 1$), the Weibull (for $\gamma_2 \to 0$), and the exponential distribution (for $\gamma_1 = 1$ and $\gamma_2 \to 0$). We explore the potential of this distribution to describe annual maxima of daily rainfall.

### 3.2. L-moments fitting

To estimate the $\mathscr{GEV}$ and $\mathscr{Br}$XII parameters, we use the method of L-moments (Greenwood et al., 1979; Hosking, 1990). Hosking et al. (1985) showed that L-moments estimators, especially for the $\mathscr{GEV}$ distribution, have small bias and a lower root-mean-square error than the maximum likelihood method. Moreover, Vogel and Fennessey (1993) documented the advantages of L-moments over the conventional product moments. Note that L-moments, since they are defined as linear combination of order statistics, have finite values for all heavy-tailed (power-type) distributions with asymptotic behaviour $x^{-1/\gamma}$ and $\gamma < 1$. This is in contrast to product moments where the $p$-th moment exists if $\gamma < p$. In practice this is crucial as the parameters of a three-parameter power-type distribution cannot be estimated for tails heavier than $x^{-1/3}$ while it can be easily done with L-moments. Therefore, the use of L-moments is highly suggested in estimating parameters of heavy-tailed distributions. Here we fit both distributions to each of the 13,247 annual maxima (AM) samples using L-moments. The spatial variability of the empirical estimates of the first L-moment $\widehat{\lambda}_1$, L-variation $\widehat{\tau}_2$ and L-skewness $\widehat{\tau}_3$ are reported in Fig. 3. The $\mathscr{GEV}$ can be easily fitted by using the first three sample L-moments, that are $\widehat{\lambda}_1, \widehat{\lambda}_2, \widehat{\lambda}_3$ (or L-variation $\widehat{\tau}_2 = \widehat{\lambda}_2/\widehat{\lambda}_1$ and L-skewness $\widehat{\tau}_3 = \widehat{\lambda}_3/\widehat{\lambda}_2$) (Hosking, 1990). Specifically, the shape parameter can be numerically estimated by $\widehat{\tau}_3 = -3 + 2(1 - 3^\gamma)/1 - 2^\gamma$, and the scale and the location parameters by $\alpha = \widehat{\lambda}_1 - \widehat{\lambda}_2(\Gamma(1-\gamma)-1)/(2^\gamma-1)\Gamma(1-\gamma)$ and $\beta = -\widehat{\lambda}_2/(2^\gamma-1)\Gamma(-\gamma)$, where $\Gamma(\cdot)$ is the gamma function.

The $\mathscr{Br}$XII distribution can also be fitted with L-moments method. Here, we follow the approach of Papalexiou and Koutsoyiannis (2012, 2016) which was later generalized in Zaghloul et al. (2020b) allowing to fit any distribution with two-shape parameters defined for positive values. This method relies in estimating the two shape parameters by numerically minimizing the squared difference between theoretical ($\tau_i$) and empirical ($\widehat{\tau}_i$) L-ratios based on the following expressions

$$\tau_2(\gamma_1,\gamma_2) := \frac{\lambda_2}{\lambda_1} = \frac{\int_0^1 Q(u;1,\gamma_1,\gamma_2)(2u-1)du}{\int_0^1 Q(u;1,\gamma_1,\gamma_2)du} \tag{6}$$

$$\tau_3(\gamma_1,\gamma_2) := \frac{\lambda_3}{\lambda_2} = \frac{\int_0^1 Q(u;1,\gamma_1,\gamma_2)(6u^2-6u+1)du}{\int_0^1 Q(u;1,\gamma_1,\gamma_2)(2u-1)du} \tag{7}$$

where $Q(u;1,\gamma_1,\gamma_2)$ is the quantile function of the $\mathscr{Br}$XII with a scale parameter $\beta = 1$. Then, the two shape parameters can be estimated with:

$$(\widehat{\gamma}_1,\widehat{\gamma}_2) = \underset{\gamma_1\gamma_2}{\operatorname{argmin}} \sum_{i=2}^{3} (\tau_i(\gamma_1,\gamma_2) - \widehat{\tau}_i)^2 \tag{8}$$

Given the shape parameters, the scale parameter is estimated by

$$\widehat{\beta} = \underset{\beta}{\mathrm{argmin}}\left(\lambda_1(\beta) - \widehat{\lambda}_1\right)^2 \tag{9}$$

with $\lambda_1(\beta) = \int_0^1 Q(u;\beta,\widehat{\gamma}_1,\widehat{\gamma}_2)du$.

To test if the $\mathscr{B}r$XII is a good candidate to describe the AM samples over Italy we use L-moments ratio diagrams (Papalexiou and Koutsoyiannis, 2016; Vogel and Fennessey, 1993). By comparing the sample L-points ($\widehat{\tau}_2$, $\widehat{\tau}_3$) with the $\mathscr{B}r$XII theoretical L-moments space we assess if the tested distribution described well the observations. Results show that almost all empirical L-points lie inside the theoretical L-space (Fig. 4), hence $\mathscr{B}r$XII is a competent model to describe the data. In the supplementary material we also depict the observed L-points and the theoretical $\mathscr{GEV}$ line in the L-skewness and L-kurtosis space (Figure S1). Note that since the $\mathscr{GEV}$ has a location parameter it cannot be represented in an L-skewness vs. L-variation space (L-variation is a function of all of its parameters).

### 3.3. Quantifying the fitting error

In order to perform a comparison between the fitted $\mathscr{GEV}$ and $\mathscr{B}r$XII distributions we use four different error measures (Papalexiou and Koutsoyiannis, 2016; Zaghloul et al., 2020b):

$$\mathrm{ER_I} = \frac{1}{n}\sum_{i=1}^{n}|\Delta x_i| \tag{10}$$

$$\mathrm{ER_{II}} = \frac{1}{m}\sum_{i=n-m+1}^{n}|\Delta x_i| \tag{11}$$

$$\mathrm{ER_{III}} = \max(|\Delta x_1|, \ldots, |\Delta x_n|) \tag{12}$$

$$\mathrm{ER_{IV}} = \frac{\Delta x_n}{\widehat{x}_n}100 \tag{13}$$

where $i$ is the rank of $x_i$ in the ordered sample $\{x_1 \leq \ldots \leq x_n\}$, and $\Delta x_i = x_i - \widehat{x}_i$ is the difference between the predicted and the observed values. Predicted values are estimated by using the quantile function of the fitted distribution with non-exceedance probability $i/(n+1)$. $\mathrm{ER_I}$ is the absolute error; $\mathrm{ER_{II}}$ quantifies the error in the highest $m$ extreme values (here $m = 10$); $\mathrm{ER_{III}}$ estimates the maximum error between the fitted and observed values; and $\mathrm{ER_{IV}}$ quantifies over- or underestimation percentage of the maximum observed value by the fitted distribution.

## 4. Results and discussion

### 4.1. $\mathcal{GEV}$ Fitting artifacts

In the previous section, we showed that the $\mathscr{B}r$XII can describe almost all the 13,247 samples of daily rainfall annual maxima recorded in Italy. Since for $\gamma > 0$ the $\mathscr{GEV}$ distribution presents the same right-tail asymptotic behavior of the $\mathscr{B}r$XII, we expect that
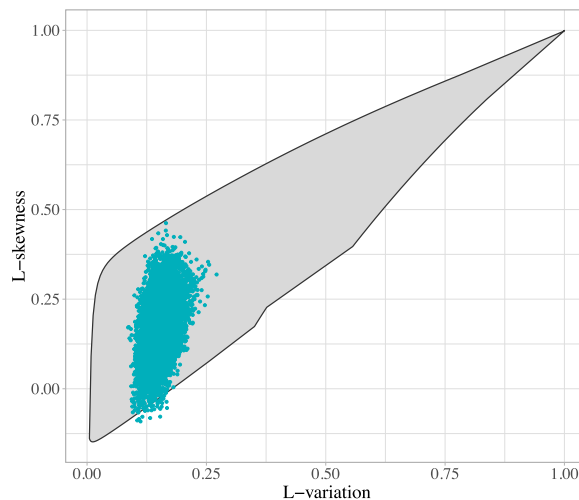


**Fig. 4.** L-ratio diagram of the $\mathscr{B}r$XII distribution: mostly of the 13,247 observed L-points (99.8 %) fall inside the L-area. The theoretical L-moments space was built with LMoFit R-package available at CRAN (Zaghloul et al., 2020a).

our data would also lead to $\mathscr{GEV}$ distributions with a positive shape parameter. To compare the two distributions in terms of tail heaviness, we focus on the estimated values of the shape parameters (i.e., $\gamma$ for the $\mathscr{GEV}$, and $\gamma_2$ for the $\mathscr{Br}$XII). For the $\mathscr{GEV}$, $\gamma$ shows a wide variation, ranging from $-0.45$ to $0.41$ with a median value of $0.02$ (Table 1); while, $\gamma_2$ varies between 0 and 0.33, with a median value of 0.17 (Table 1).

In general, a remarkable difference emerges from the comparison of the tail behavior as quantified by the shape parameters (Fig. 5). The $\mathscr{Br}$XII tails are typically heavier compared to the $\mathscr{GEV}$ tails, with most of the points arranged under the $\gamma_{\mathscr{GEV}} = \gamma_2$ line (see dashed line in Fig. 5). For $\gamma_2 > 0.2$, the two distributions present a similar asymptotic behavior, thus the points scatter in correspondence of the $\gamma_{\mathscr{GEV}} = \gamma_2$ line.

The fitting reveals that 44.9 % of the samples result in $\gamma < 0$ for the $\mathscr{GEV}$ corresponding, thus, to a bounded from above distribution. We deem that the use of a distribution with an upper bound to describe extreme precipitation is physically inconsistent and the good performance of the $\mathscr{Br}$XII distribution strengthens this argument. The $\mathscr{Br}$XII is a power-type distribution with the same tail asymptotic behavior as the $\mathscr{GEV}$, and yet the tail heaviness estimates for the two distributions differ profoundly with almost half of the $\mathscr{GEV}$ tails being bounded from above. We deem that this is an artifact resulting by the mathematical formula of the $\mathscr{GEV}$ distribution and sample variations. To highlight our statements, we perform a Monte Carlo experiment generating 10,000 samples, 40 values each, from a $\mathscr{GEV}$ distribution with a positive shape parameter equal to 0.05. Then fitting the $\mathscr{GEV}$ to those samples shows that almost 38 % of them have a negative shape parameter, with a mean value of 0.04. This result confirms that allowing $\mathscr{GEV}$ to take negative shape parameter is not justifiable as negative values emerge due to sampling variation.

The spatial variability of both the estimated shape parameters (i.e., $\gamma$ and $\gamma_2$ respectively for $\mathscr{GEV}$ and $\mathscr{Br}$XII) for the 44.9 % of grid-cells with $\gamma < 0$ is spread all over Italy (Fig. 6). For all these samples the $\mathscr{GEV}$ distribution implies an upper daily rainfall limit (RL), given by $\alpha - \beta/\gamma$ (Zaghloul et al., 2020b). We use the $\mathscr{Br}$XII distribution, that is unbounded, to estimate the return period related to these upper limits assumed by the $\mathscr{GEV}$, with the expression $T_{\text{RL}} = (1 - F_{\mathscr{Br}\text{XII}}(\text{RL}))^{-1}$. Results reveal that 8% of the AM samples presents an RL with a return period less than 1000 years (Fig. 7b). As these return periods are commonly used in engineering practice, for example, in spillway design, choosing the $\mathscr{GEV}$ distribution might lead to risk underestimation.

We test the previous findings using daily ground data provided by NOAA (https://www.ncdc.noaa.gov/cdo-web/datasets). We select 13 Italian stations with more than 40 years of observation (locations are depicted in Figure S2 of the supplementary material). Then, we fit the $\mathscr{GEV}$ distribution to the extracted AM samples using the L-moments method (summary statistics are provided in Table S1 of the supplementary material). For the selected stations, only one returns a negative $\mathscr{GEV}$ shape parameter, that is Brindisi with 60 years of observation. Comparing between these results and those in Table 1, we observe a slight reduction in the variability of the shape parameters between the two datasets; also, the mean value in ground stations (0.14) is slightly larger from the CHIRPS mean (0.09). Even though the two datasets are not directly comparable, because point observation can have heavier tails which are smoothed out in gridded products due to areal averaging ($\sim$25 km$^2$ for the CHIRPS data), the results confirm the mathematical artifact affecting the $\mathscr{GEV}$ distribution.

## 4.2. The $\mathscr{GEV}^+$ approach

We argued that the reverse Weibull is not a reasonable choice to describe the upper tails, thus, we suggest fitting the $\mathscr{GEV}$ distribution by constraining its shape parameter only to positive values. This essentially implies using a Gumbel distribution instead of a $\mathscr{GEV}$ with negative shape. Hereafter, we denote this version of the $\mathscr{GEV}$ as $\mathscr{GEV}^+$. The Gumbel distribution does not have a shape parameter, and thus, its L-skewness is fixed and equal to $\tau_3 = \ln(9/8)/\ln2$. Consequently, if an AM sample has $\hat{\tau}_3 \leq \ln(9/8)/\ln2$ we fit the Gumbel distribution (Eq. (2)), with scale and location parameters, respectively, given by $\beta = \hat{\lambda}_2/\ln2$ and $\alpha = \hat{\lambda}_1 - \bar{\gamma}\beta$, where $\bar{\gamma}$ is the Euler-Mascheroni constant; otherwise we fit the $\mathscr{GEV}$ distribution. Clearly, the $\mathscr{GEV}^+$ approach results in fitting the Gumbel distribution for 44.9 % of samples with $\gamma < 0$ (Fig. 6b) and fitting the $\mathscr{GEV}$ to the rest 55.1 % (Fig. 8).

Although the median value of the shape parameter is moderate (0.084, see Table 2), in some regions high values are spotted (light

**Table 1**
Main statistics of the fitting results of the estimated parameters of both the $\mathscr{Br}$XII and the $\mathscr{GEV}$ distributions to the 13,247 AM daily rainfall samples.

| | $\mathscr{Br}$XII parameters | | | $\mathscr{GEV}$ parameters | | |
| --- | --- | --- | --- | --- | --- | --- |
| | $\beta$ | $\gamma_1$ | $\gamma_2$ | $\alpha$ | $\beta$ | $\gamma$ |
| Min | 18.95 | 3.41 | 0.00 | 22.54 | 4.51 | $-0.45$ |
| $p_5$ | 27.16 | 4.96 | 0.08 | 33.04 | 7.34 | $-0.20$ |
| $p_{25}$ | 34.73 | 6.11 | 0.13 | 42.10 | 9.80 | $-0.07$ |
| $p_{50}$ | 38.97 | 7.10 | 0.17 | 47.11 | 11.54 | 0.02 |
| $p_{75}$ | 43.29 | 8.35 | 0.20 | 52.36 | 13.58 | 0.09 |
| $p_{95}$ | 54.83 | 10.95 | 0.24 | 66.33 | 18.12 | 0.19 |
| Max | 86.82 | 55.79 | 0.33 | 106.26 | 31.74 | 0.41 |
| Mean | 39.51 | 7.43 | 0.16 | 47.83 | 11.98 | 0.01 |
| SD | 8.23 | 2.03 | 0.05 | 9.95 | 3.38 | 0.12 |
| Skew | 0.90 | 2.91 | $-0.32$ | 0.93 | 1.25 | $-0.29$ |
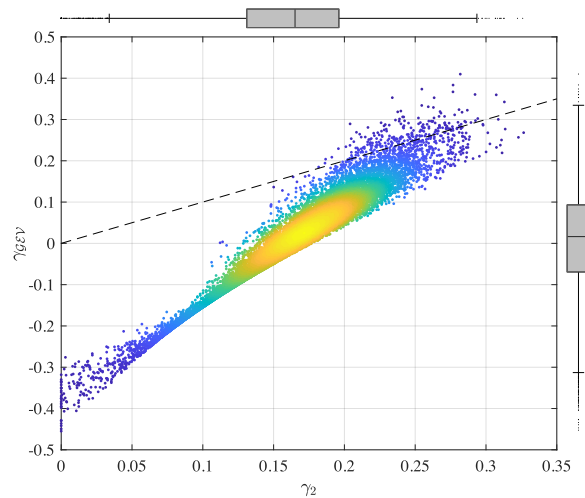
**Fig. 5.** Comparison based on the tail heaviness of $\mathscr{B}r$XII and $\mathscr{GEV}$ distributions when fitted to annual maximum rainfall. Each point is related to an annual maximum sample, while the dashed line represents $\gamma_2 = \gamma_{\mathscr{GEV}}$. The color scale is proportional to the density of points (blue: low, yellow: high) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).
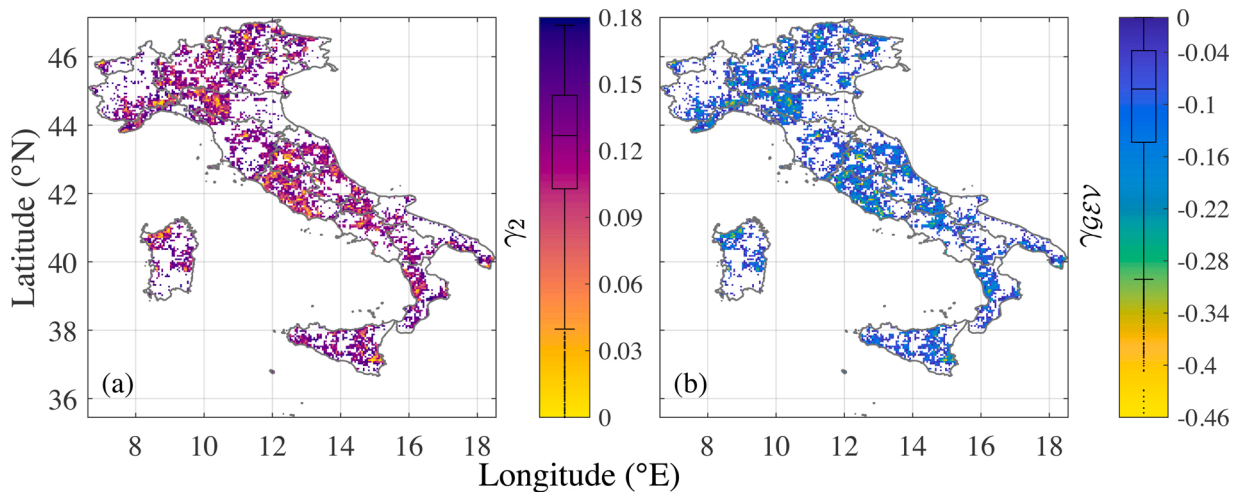


**Fig. 6.** Representation of the grid-cells with a negative $\mathscr{GEV}$ shape parameter. (a) Spatial variability of $\gamma_2$ that governs the heaviness of the right tail of the $\mathscr{B}r$XII distribution. (b) Spatial variability of the $\mathscr{GEV}$'s shape parameter: white cells are representative of $\gamma > 0$, instead colored cells are characterized by $\gamma < 0$.

colours in Fig. 8), denoting heavy-tails. Many regions, for example, the North-West area (Aosta Valley and Piedmont) is characterized by high values of $\gamma$ (especially in the Alpine area). Other clear patterns are observed in Friuli-Venezia Giulia, in the Po valley, in Tuscany, in some parts of the Apennine, in Apulia and Basilicata, in the north part of Calabria and in south Sardinia. In Table 2 we provide summary statistics of the estimated parameters for the $\mathscr{GEV}^+$ distribution.

We compare the potential of the $\mathscr{B}r$XII, $\mathscr{GEV}$, and $\mathscr{GEV}^+$ to describe the observed annual maxima in terms of the four error measures Eq.s (10–13) (Fig. 9). Their performance is almost identical. An important finding is that although are in general heavy-tailed, still they under-estimate the observed rainfall maximum for the 74 %, 78.5 % and 69.3 % of the analyzed samples (Fig. 9d), respectively. Such under-estimation in the observed sample's maximum might imply a more severe underestimation for larger return periods.

The previous findings show that the $\mathscr{B}r$XII and $\mathscr{GEV}$ versions perform equally well but underestimate the observed largest value. A reasonable assumption then could be that the distribution predicting higher values for larger return periods would underestimate less the actual future magnitudes. To quantify the difference between the rainfall predictions, we evaluate the percentage difference PD between the $\mathscr{B}r$XII and $\mathscr{GEV}$ estimates as
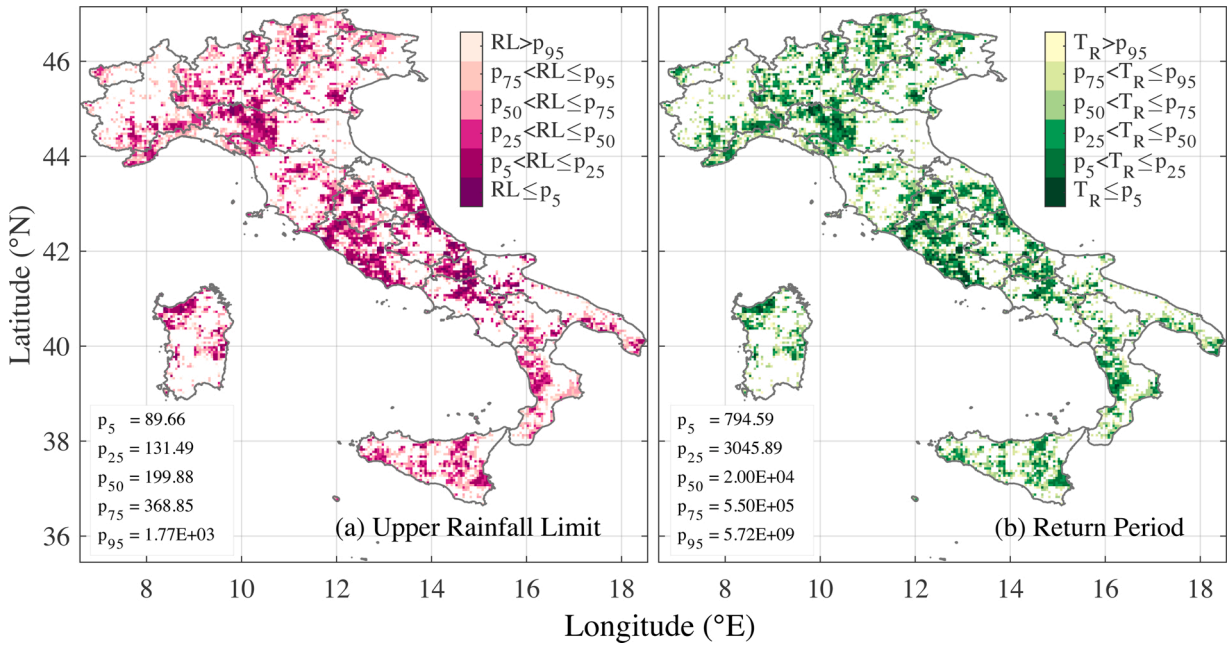
**Fig. 7.** Spatial representation of the upper rainfall limit (a) and of the related return periods (b) evaluated with the quantile function of the $\mathscr{Br}$XII. For a clear representation, we represent the results grouped in percentile classes, whose value are reported in table.
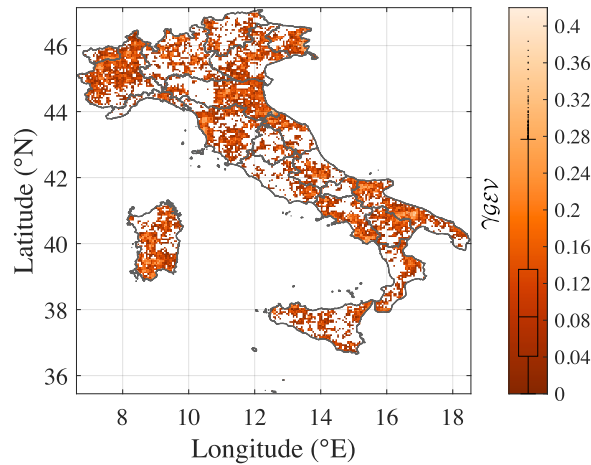


**Fig. 8.** $\mathscr{GEV}$'s shape parameter spatial representation. Contrary to Fig. 6(b), coloured cells depict only the positive $\gamma$; white cells represent samples fitted with Gumbel distribution. Inside the colorbar is presented the variability of $\gamma$.

$$PD = \left( \frac{F^{-1}_{\mathscr{Br}XII}\left(1 - \frac{1}{T}\right)}{F^{-1}_{\mathscr{GEV}}\left(1 - \frac{1}{T}\right)} - 1 \right) \times 100\% \tag{14}$$

and similarly for $\mathscr{GEV}^+$, by applying several return periods $T$ commonly used for engineering purpose (i.e., 50, 100, 200, 500 and 1000 years). Our results show that the $\mathscr{Br}$XII predicts larger values compared to the other two distributions (Fig. 10). Indeed, the median PD values are positive for all return periods. The predicted rainfall values by the $\mathscr{Br}$XII, for the 50-years return period, are comparable with those of the $\mathscr{GEV}$ and $\mathscr{GEV}^+$, yet as return period increases up to 1000-years the PD becomes much larger with median higher than 15 %. Notably, a larger variance is observed in the PD values evaluated for the $\mathscr{GEV}^+$ compared to $\mathscr{GEV}$, which also increases with the return period. This greater variance is explained because when $\mathscr{GEV}$ has $\gamma < 0$ it predicts always lower rainfall values than $\mathscr{Br}$XII and thus the PD is positive; in contrast, it seems that in the $\mathscr{GEV}^+$ when essentially the Gumbel is used to describe the negative shape parameter cases, the $\mathscr{Br}$XII can predict less and this even creates negative PD values and the larger

**Table 2**
Basic statistics of the estimated parameters for the $\mathscr{GEV}^+$ distribution.

| | $\mathscr{GEV}^+$ parameters | | |
|---|---|---|---|
| | $\alpha$ | $\beta$ | $\gamma$ |
| Min | 22,18 | 4,51 | 0.00001 |
| p$_5$ | 32,94 | 7,10 | 0.008 |
| p$_{25}$ | 41,97 | 9,46 | 0.041 |
| p$_{50}$ | 46,84 | 11,12 | 0.084 |
| p$_{75}$ | 52,02 | 13,05 | 0.135 |
| p$_{95}$ | 65,90 | 17,53 | 0.222 |
| Max | 106,26 | 31,55 | 0.410 |
| Mean | 47,58 | 11,55 | 0.094 |
| SD | 9,87 | 3,25 | 0.067 |
| Skew | 0,94 | 1,28 | 0.804 |
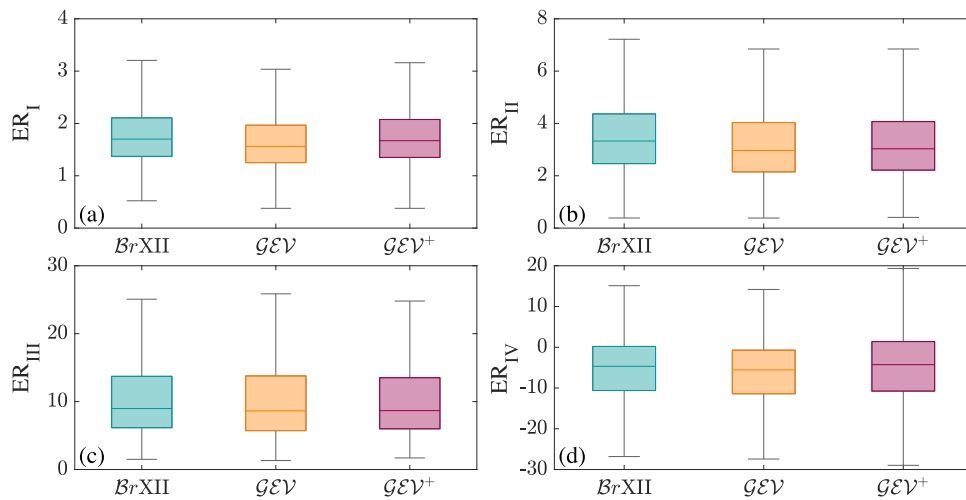


**Fig. 9.** Box plots of the four error measures evaluated for $\mathscr{Br}$XII, $\mathscr{GEV}$ and $\mathscr{GEV}^+$ distributions when fitted to the AM samples.
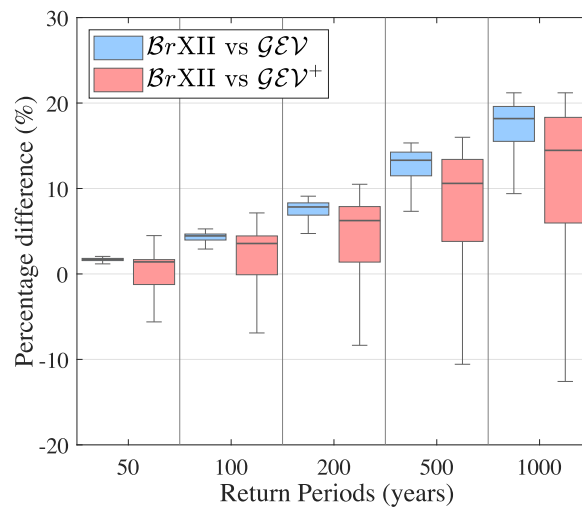


**Fig. 10.** Box plots showing the rainfall prediction differences for several return periods, between the fitted $\mathscr{Br}$XII in comparison with the fitted $\mathscr{GEV}$ and $\mathscr{GEV}^+$.

variance.

### 4.3. Rainfall maps for fixed return periods

Here, we evaluate the spatial variation of the predicted precipitation for different return periods. Indeed, the knowledge of rainfall patterns for particular return periods is a necessity for flood mapping and for designing hydraulic structures (Van De Vyver, 2012). Based on the previous results we estimate the rainfall amount that, on average, is exceeded once every $T$ years. Rainfall depths $h$ are evaluated with the quantile functions of the three tested distributions, that are, $\mathscr{B}r$XII, $\mathscr{GEV}$ and $\mathscr{GEV}^+$. Differences in terms of rainfall values predicted by the three distributions for three representative return periods are shown in Fig. 11. For a clear representation of the differences previously discussed in Fig. 10, we grouped the rainfall values in six percentile classes (see the legend in Fig. 11) that were evaluated for each return period.

As anticipated the three tested distributions exhibit, in general, the same spatial patterns (Fig. 11). The highest rainfall values, greater than the 95th percentile, are in the Alpine zone, in Liguria, in the north of Tuscany, and in specific parts of Calabria and Apulia. The same areas have been previously recognized as the ones with the highest values of mean daily rainfall (Fig. 2b). Most of the Adriatic coast is characterized by predicted values lower than the 25th percentile. The extension of the Apennine plays a crucial role for this behavior: the occidental winds from the Tyrrhenian Sea are blocked by the mountain chain causing orographic rainfall. The same behavior is found nearby the Alps, as the mountain chain blocks the winds from the north and protect the Po valley. In the south and in the two major islands we observe that the highest values of the rainfall depth are characteristics of the highest elevation.

The identified patterns remain roughly constant for smaller return periods, yet careful inspection shows strong differences for the 1000-years period. We further investigated the differences emerged in Fig. 11 by analysing the subdivision of the cells into the six percentile classes (Fig. 12). At 50-years of return period, the behavior is approximately the same for the three distributions, consistently with the results obtained for ER$_{IV}$ (Fig. 9d). The higher variance already emerged in Fig. 10 for the $\mathscr{GEV}^+$ is reflected in a greater number of cells falling in the higher percentile classes. By increasing the return period, the potential large under-estimation by the $\mathscr{GEV}$ distribution is found in many cells belonging to the first two percentile classes. These results indicate that for many regions in Italy the $\mathscr{GEV}$ distribution could potentially severely underestimate the magnitude of extremes for large return periods.

### 4.4. Relationship between rainfall depth and orography

The spatial patterns emerged in the rainfall maps (Fig. 11) seem to follow the orography of the country (Fig. 1a). To investigate this aspect, we show in Fig. 13 a bivariate choropleth map to evaluate the correlation between the elevation above the sea level (hereinafter $q$) and the classes of rainfall depths. To perform this comparison, we choose the $\mathscr{GEV}^+$ to estimate the rainfall depths for several return periods (i.e., 39, 50, 100, 200, 500 and 1000 years).

The highest values of elevation and rainfall depths, whose grid-cells are depicted with darkest purple-blue colours, are mostly located along the Alps and the Apennine. Particularly, the Alpine and pre-Alpine areas present heavier rainfall values with respect to the Apennine chain. At lower altitudes classes (with $q$ close to its median value), a greater variability emerged in the distribution of rainfall: with lower values (light orange colours on the left of the legend) along the Apennine, especially on the Adriatic coast, and higher values in the North (purple values on the right of the legend). The same variability can be observed over the flat areas, mainly located along the coasts and in the Po valley. The lowest rainfall values are located on the Adriatic and Lazio coasts (white cells). Blue and light-blue cells are representative of greater values of precipitation over the plains (see the Po valley in Fig. 13). The relationship between the elevation and the rainfall depth does not change with the return period. To highlight the above aspects, we provide six different maps for each class of elevation in Figure S3 of the supplementary material.

### 5. Summary and conclusions

Catastrophic events in Italy, such as floods and landslides, have caused significant infrastructures damage an economic loss. A robust estimation of rainfall depths for different return periods is crucial to plan and design hydraulic protection infrastructure to alleviate the impact of extremes. The Generalize Extreme Value ($\mathscr{GEV}$) distribution is the most popular model to describe extreme rainfall worldwide; yet it should be used with caution. Theoretical assumptions justifying the use of $\mathscr{GEV}$ are often violated in real-world, and thus, we suggest using $\mathscr{GEV}$ with a priori restricting its shape parameter to positive values; we name this variant the $\mathscr{GEV}^+$. We believe that using a distribution bounded from above, such as the reverse Weibull ($\mathscr{GEV}$ with negative shape parameter), is not consistent with natural processes limited at zero, such as precipitation. To support our statements, we also used the Burr type XII ($\mathscr{B}r$XII) distribution, a non-extreme value distribution defined in $[0, \infty)$, to describe annual maxima and to compare it with $\mathscr{GEV}$. The $\mathscr{B}r$XII has the same asymptotic behavior with the $\mathscr{GEV}$ distribution which makes the comparison of the tail behavior valid.

The fitting performance and the spatial variability of the $\mathscr{GEV}$ are tested on 13,247 daily rainfall time-series over Italy, using L-moments. Our analysis reveals that:

1. The $\mathscr{GEV}$ distribution in 44.9 % of the samples had a negative shape parameter implying an upper daily rainfall limit (Fig. 6). Return periods related to these upper limits, however, are finite when estimated with the $\mathscr{B}r$XII distribution. This indicates that
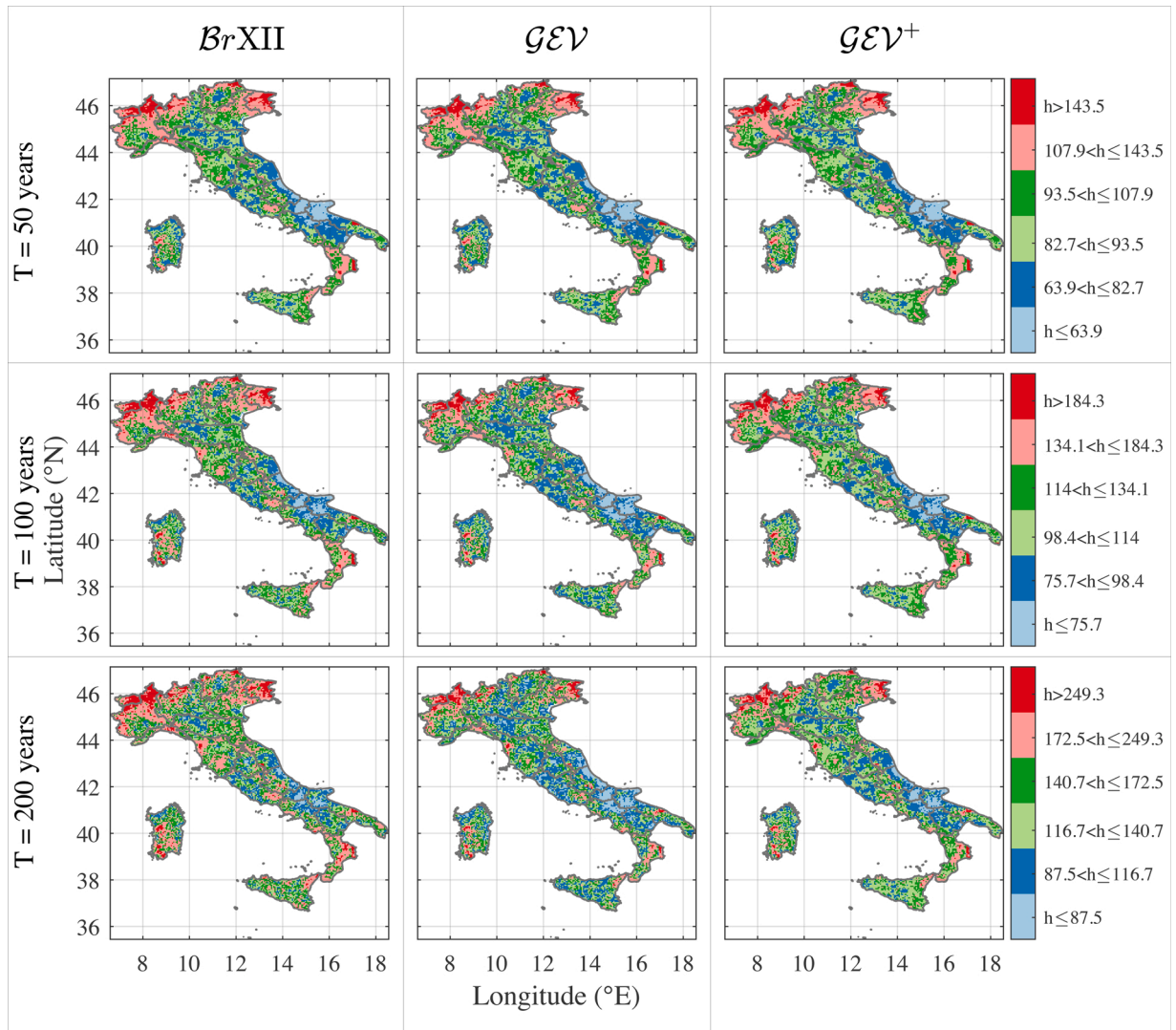
**Fig. 11.** Differences on the predicted rainfall values $h$ (mm) by fitting the three distributions. To provide a clear representation, we divide the rainfall quantiles in six percentile classes (i.e., 5th, 25th, 50th, 75th, and 95th) whose values are evaluated using the values obtained with the three distribution for each return period.

upper bounds estimated by $\mathcal{GEV}$ might be artifacts caused by sample variations. Thus, we suggest using the $\mathcal{GEV}$ only with positive shape parameters ($\mathcal{GEV}^+$).

2. Comparison of the right tail between the $\mathcal{GEV}$ and $\mathcal{Br}$XII distributions reveals notable differences. Although they have the same asymptotic behavior when $\gamma > 0$, the $\mathcal{Br}$XII tails resulted heavier than the $\mathcal{GEV}$ tails, especially when the shape parameter $\gamma$ of $\mathcal{GEV}$ is lower than 0.2. Nevertheless, the spatial representation of $\mathcal{GEV}$'s positive shape parameter for the 55.1 % of the samples allowed to identify many regions with heavy-tails (Fig. 8).

3. The $\mathcal{Br}$XII, $\mathcal{GEV}$, and $\mathcal{GEV}^+$ performed almost identically for the observed return period (i.e., 39 years, Fig. 9d). Although these models are heavy tailed, they underestimate the observed rainfall maximum in 74 %, 78.5 % and 69.3 % of the analysed samples, respectively.

4. The $\mathcal{Br}$XII, for large return periods, predicts larger rainfall amounts compared to the $\mathcal{GEV}$ estimates (Fig. 10). This might indicate that $\mathcal{GEV}$ estimates could underestimate the risk of extremes.

5. The rainfall spatial patterns, emerged by fitting the three distributions, are similar for the selected return periods. The estimated rainfall depth tends to follow the orography of the territory, with high rainfall at high altitude and moderate rainfall over the plains (Fig. 13). This result confirmed the important role played by the high mountain chains on the rainfall distribution in Italy. These maps might be useful to update and integrate the estimated rainfall depths for fixed return periods evaluated with the rain gauge network.
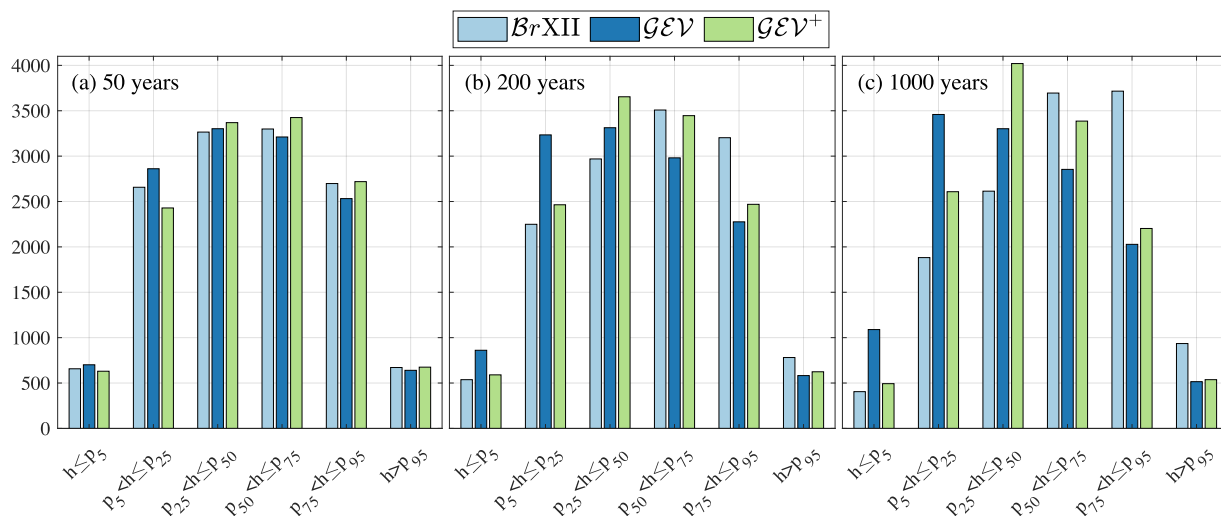
**Fig. 12.** Barplots showing for the $\mathcal{Br}$XII, $\mathcal{GEV}$ and $\mathcal{GEV}^+$ the number of cells falling into the six percentile classes in three different return periods.
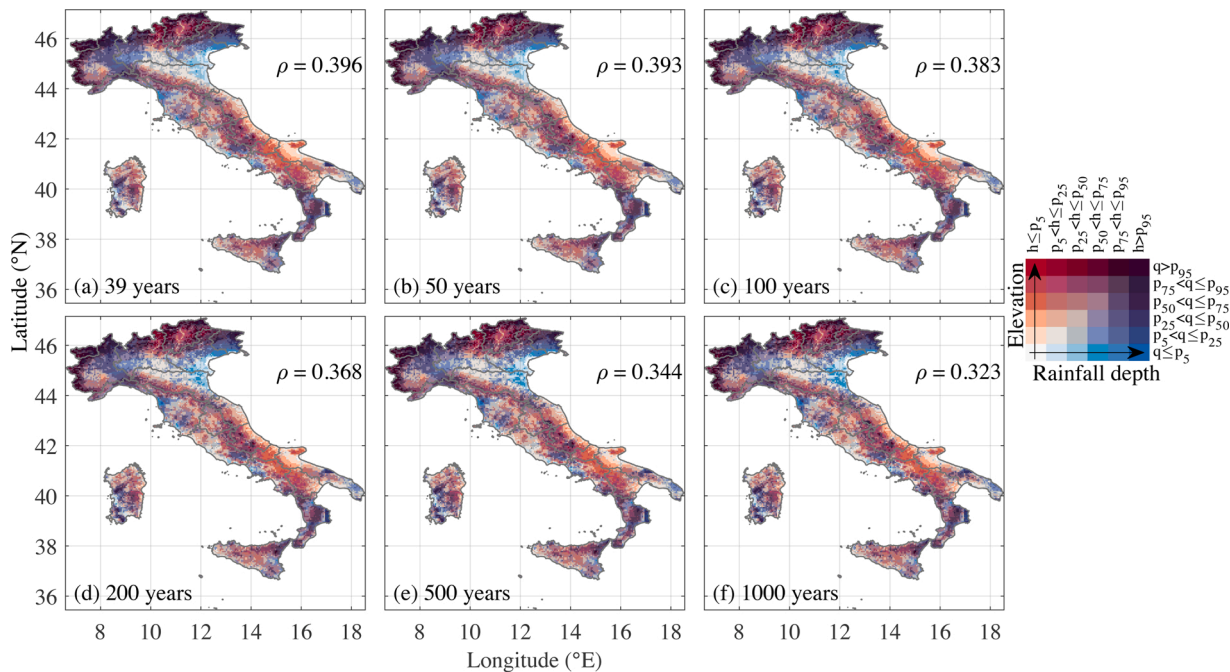


**Fig. 13.** Bivariate choropleth map of the rainfall depth $h$ and the elevation $q$ for different return periods. By grouping the two variables in six percentile classes (5th, 25th, 50th, 75th and 95th) we identify 36 classes for the classification. Legend shows that the horizontal colours are related to the variability of the rainfall depth (from lower values on the left to higher values on the right), while the vertical colours depict the variability of the elevation (from lower values on the bottom to higher values on the top). $\rho$ indicates the Pearson correlation values calculated between the rainfall depth and the elevation of each grid-cell.

Based on the previous findings, we suggest using the $\mathcal{GEV}^+$ instead of the classical $\mathcal{GEV}$ distribution but also include in the analysis non-extreme value distributions such as the $\mathcal{Br}$XII to describe precipitation extremes. It is impossible to verify which model predicts more accurately the extremes for large period; however, since the three models underestimate the observed sample maximum, we could claim that the one underestimating it less might be a better choice.

## Authors statement

**Benedetta Moccia:** Data Curation, Formal Analysis, Methodology, Visualization, Writing – original draft, Writing – Review & Editing. **Simon Michael Papalexiou:** Conceptualization, Formal Analysis, Methodology, Supervision, Writing – original draft, Writing – Review & Editing. **Fabio Russo:** Supervision, Founding acquisition, Writing – Review & Editing. **Francesco Napolitano:** Supervision, Founding acquisition, Writing – Review & Editing.

## Declaration of Competing Interest

The authors declare no competing interest.

## Acknowledgements

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ejrh.2021.100906.

## References

Bai, L., Shi, C., Li, L., Yang, Y., Wu, J., 2018. Accuracy of CHIRPS satellite-rainfall products over mainland China. Remote Sens. 10 (3) https://doi.org/10.3390/rs10030362.

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. Sci. Data 5, 1–12. https://doi.org/10.1038/sdata.2018.214.

Bertini, C., Buonora, L., Ridolfi, E., Russo, F., Napolitano, F., 2020. On the use of satellite rainfall data to design a dam in an ungauged site. Water 12 (3028), 1–20. https://doi.org/10.3390/w12113028.

Burr, I.W., 1942. Cumulative frequency functions. Ann. Math. Stat. 13 (2), 215–232. https://doi.org/10.1214/aoms/1177731607.

Caroletti, G.N., Coscarelli, R., Caloiero, T., 2019. Validation of satellite, reanalysis and RCM data of monthly rainfall in Calabria (Southern Italy). Remote Sens. 11 (13) https://doi.org/10.3390/rs11131625.

Chu, P.S., Zhao, X., Ruan, Y., Grubbs, M., 2009. Extreme rainfall events in the Hawaiian Islands. J. Appl. Meteorol. Climatol. 48 (3), 502–516. https://doi.org/10.1175/2008JAMC1829.1.

Coles, S., Pericchi, L.R., Sisson, S., 2003. A fully probabilistic approach to extreme rainfall modeling. J. Hydrol. 273, 35–50.

Davini, P., Bechini, R., Cremonini, R., Cassardo, C., 2012. Radar-based analysis of convective storms over Northwestern Italy. Atmosphere 3 (1), 33–58. https://doi.org/10.3390/atmos3010033.

De Michele, C., Avanzi, F., 2018. Superstatistical distribution of daily precipitation extremes: a worldwide assessment. Sci. Rep. 8 (1), 1–11. https://doi.org/10.1038/s41598-018-31838-z.

Duan, Z., Liu, J., Tuo, Y., Chiogna, G., Disse, M., 2016. Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. Sci. Total Environ. 573, 1536–1553. https://doi.org/10.1016/j.scitotenv.2016.08.213.

El Adlouni, S., Bobée, B., Ouarda, T.B.M.J., 2008. On the tails of extreme event distributions in hydrology. J. Hydrol. 355 (1–4), 16–33. https://doi.org/10.1016/j.jhydrol.2008.02.011.

Fisher, R.A., Tippett, L.H.C., 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Math. Proc. Cambridge Philos. Soc. 24 (2), 180–190. https://doi.org/10.1017/S0305004100015681.

Funk, C., Verdin, A., Michaelsen, J., Peterson, P., Pedreros, D., Husak, G., 2015a. A global satellite-assisted precipitation climatology. Earth Syst. Sci. Data Discuss. 7, 275–287. https://doi.org/10.5194/essd-7-275-2015.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., et al., 2015b. The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. Sci. Data 2, 1–21. https://doi.org/10.1038/sdata.2015.66.

Gnedenko, B., 1943. Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire Author. Annals of Mathematics, pp. 423–453.

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

Gumbel, E.J., 1958. Statistics of Extremes. Columbia University Press, New York.

Hosking, J.R.M., 1990. L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. J. R. Stat. Soc. Ser. B 52 (1), 105–124.

Hosking, J.R.M., Wallis, J.R., Wood, E.F., 1985. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. Technometrics 27 (3), 251–261. https://doi.org/10.1080/00401706.1985.10488049.

Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., et al., 2006. The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeorol. 8, 38–55. https://doi.org/10.1175/JHM560.1.

Katsanos, D., Retalis, A., Michaelides, S., 2016. Validation of a high-resolution precipitation database (CHIRPS) over Cyprus for a 30-year period. Atmos. Res. 169, 459–464. https://doi.org/10.1016/j.atmosres.2015.05.015.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World map of the Köppen-Geiger climate classification updated. Meteorol. Z. 15 (3), 259–263. https://doi.org/10.1127/0941-2948/2006/0130.

Lee, S.H., Maeng, S.J., 2003. Frequency analysis of extreme rainfall using L-moment. Irrig. Drain. 52 (3), 219–230. https://doi.org/10.1002/ird.90.

Libertino, A., Ganora, D., Claps, P., 2018. Technical note: space-time analysis of rainfall extremes in Italy: clues from a reconciled dataset. Hydrol. Earth Syst. Sci. 22 (5), 2705–2715. https://doi.org/10.5194/hess-22-2705-2018.

Mazzoglio, P., Butera, I., Claps, P., 2020. I2-red: a massive update and quality control of the italian annual extreme rainfall dataset. Water (Switzerland) 12 (12). https://doi.org/10.3390/w12123308.

Mishra, A.K., Coulibaly, P., 2009. Developments in hydrometric network design: a review. Rev. Geophys. 47 https://doi.org/10.1029/2007RG000243.1.
    INTRODUCTION.
Moccia, B., Mineo, C., Ridolfi, E., Russo, F., Napolitano, F., 2021. Probability distributions of daily rainfall extremes in Lazio and Sicily, Italy, and design rainfall
    inferences. J. Hydrol. Reg. Stud. 33, 100771 https://doi.org/10.1016/j.ejrh.2020.100771.
Nerantzaki, S.D., Papalexiou, S.M., 2019. Tails of extremes: advancing a graphical method and harnessing big data to assess precipitation extremes. Adv. Water
    Resour. 134 (May) https://doi.org/10.1016/j.advwatres.2019.103448, 103448.
Nguyen, V.T. Van, Nguyen, T.H., 2016. Statistical modeling of extreme rainfall processes (SMExRain): a decision support tool for extreme rainfall frequency analyses.
    Procedia Eng. 154, 624–630. https://doi.org/10.1016/j.proeng.2016.07.561.
Ouarda, T.B.M.J., Ashkar, F., Bensaid, E., Hourani, I., 1994. Statistical distributions used in hydrology. Transformations and asymptotic properties. Scientific Report.
    Department of Mathematics, Univ. of Moncton, New Brunswick, 31 Pp.
Papalexiou, S.M., 2018. Unified theory for stochastic modelling of hydroclimatic processes: preserving marginal distributions, correlation structures, and
    intermittency. Adv. Water Resour. 115, 234–252. https://doi.org/10.1016/j.advwatres.2018.02.013.
Papalexiou, S.M., Koutsoyiannis, D., 2012. Entropy based derivation of probability distributions: a case study to daily rainfall. Adv. Water Resour. 45, 51–57. https://
    doi.org/10.1016/j.advwatres.2011.11.007.
Papalexiou, S.M., Koutsoyiannis, D., 2013. Battle of extreme value distributions: a global survey on extreme daily rainfall. Water Resour. Res. 49 (1), 187–201.
    https://doi.org/10.1029/2012WR012557.
Papalexiou, S.M., Koutsoyiannis, D., 2016. A global survey on the seasonal variation of the marginal distribution of daily precipitation. Adv. Water Resour. 94,
    131–145. https://doi.org/10.1016/j.advwatres.2016.05.005.
Papalexiou, S.M., Serinaldi, F., 2020. Random fields simplified: preserving marginal distributions, correlations, and intermittency, with applications from rainfall to
    humidity. Water Resour. Res. 56 (2) https://doi.org/10.1029/2019WR026331 e2019WR026331.
Papalexiou, S.M., Koutsoyiannis, D., Makropoulos, C., 2013. How extreme is extreme? An assessment of daily rainfall distribution tails. Hydrol. Earth Syst. Sci. 17 (2),
    851–862. https://doi.org/10.5194/hess-17-851-2013.
Paredes-Trejo, F.J., Barbosa, H.A., Lakshmi Kumar, T.V., 2017. Validating CHIRPS-based satellite precipitation estimates in Northeast Brazil. J. Arid Environ. 139,
    26–40. https://doi.org/10.1016/j.jaridenv.2016.12.009.
Peel, M.C., Finlaysonm, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. Hydrol. Earth Syst. Sci. 11, 1633–1644. https://
    doi.org/10.5194/hess-11-1633-2007, 2007.
Prakash, S., 2019. Performance assessment of CHIRPS, MSWEP, SM2RAIN-CCI, and TMPA precipitation products across India. J. Hydrol. 571 (January), 50–59.
    https://doi.org/10.1016/j.jhydrol.2019.01.036.
Rajulapati, C.R., Papalexiou, S.M., Clark, M.P., Razavi, S., Tang, G., Pomeroy, J.W., 2020. Assessment of extremes in global precipitation products: How reliable are
    they? J. Hydrometeorol. 21 (12), 2855–2873. https://doi.org/10.1175/JHM-D-20-0040.1.
Rivera, J.A., Marianetti, G., Hinrichs, S., 2018. Validation of CHIRPS precipitation dataset along the Central Andes of Argentina. Atmos. Res. 213 (January), 437–449.
    https://doi.org/10.1016/j.atmosres.2018.06.023.
Salas, J.D., Anderson, M.L., Papalexiou, S.M., Frances, F., 2020. PMP and climate variability and change: a review. J. Hydrol. Eng. 25 (12) https://doi.org/10.1061/
    (asce)he.1943-5584.0002003, 03120002.
Smith, J.A., Villarini, G., Baeck, M.L., 2011. Mixture distributions and the hydroclimatology of extreme rainfall and flooding in the Eastern United States.
    J. Hydrometeorol. 12 (2), 294–309. https://doi.org/10.1175/2010JHM1242.1.
Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K.L., 2018. A review of global precipitation data sets: data sources, estimation, and intercomparisons.
    Rev. Geophys. 56 (1), 79–107. https://doi.org/10.1002/2017RG000574.
Tang, G., Clark, M., Papalexiou, S.M., Ma, Z., Hong, Y., 2020. Have satellite precipitation products improved over last two decades? A comprehensive comparison of
    GPM IMERG with nine satellite and reanalysis datasets. Remote Sens. Environ. 240 (September 2019) https://doi.org/10.1016/j.rse.2020.111697, 111697.
Telesca, L., Bernardi, M., Rovelli, C., 2008. Time-scaling analysis of lightning in Italy. Commun. Nonlinear Sci. Numer. Simul. 13 (7), 1384–1396. https://doi.org/
    10.1016/j.cnsns.2007.02.001.
Van De Vyver, H., 2012. Spatial regression models for extreme precipitation in Belgium. Water Resour. Res. 48 (9), 1–17. https://doi.org/10.1029/2011WR011707.
Villarini, G., 2012. Analyses of Annual and Seasonal Maximum Daily raiNfall Accumulations for Ukraine, Moldova, and Romania, 2226, pp. 2213–2226. https://doi.
    org/10.1002/joc.3394.
Vogel, R.M., Fennessey, N.M., 1993. L moment diagrams should replace product moment diagrams. Water Resour. Res. 29 (6), 1745–1752.
von Mises, R., 1936. La distribution de la plus grande de n valeurs. Rev. Math. Union Interbalcanique 1, 141–160.
Werner, T., Upper, C., 2002. Time Variation in the Tail Behaviour of Bund Futures Returns. European Central Bank. Working paper series, 199.
Xie, P., Xiong, A.Y., 2011. A conceptual model for constructing high-resolution gauge-satellite merged precipitation analyses. J. Geophys. Res. 116, 1–14. https://doi.
    org/10.1029/2011JD016118.
Yang, L., Villarini, G., Smith, J.A., Tian, F., Hu, H., 2013. Changes in seasonal maximum daily precipitation in China over the period 1961-2006. Int. J. Climatol. 33
    (7), 1646–1657. https://doi.org/10.1002/joc.3539.
Zaghloul, M., Papalexiou, S.M., Elshorbagy, A., 2020a. LMoFit: Advanced L-Moment Fitting of Distributions. R Package Version 0.1.6.
Zaghloul, M., Papalexiou, S.M., Elshorbagy, A., Coulibaly, P., 2020b. Revisiting flood peak distributions: a pan-Canadian investigation. Adv. Water Resour. 145
    (103720) https://doi.org/10.1016/j.advwatres.2020.103720.
Zorzetto, E., Botter, G., Marani, M., 2016. On the emergence of rainfall extremes from ordinary events. Geophys. Res. Lett. 43 (15), 8076–8082. https://doi.org/
    10.1002/2016GL069445.