Journal of Hydrology: Regional Studies 33 (2021) 100771

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

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

Probability distributions of daily rainfall extremes in Lazio and Sicily, Italy, and design rainfall inferences

Benedetta Moccia [a,*], Claudio Mineo [a], Elena Ridolfi [b,c], Fabio Russo [a], Francesco Napolitano [a]

[a] Università degli Studi di Roma La Sapienza, Dipartimento di Ingegneria Civile, Edile e Ambientale, Roma, Italy
[b] Uppsala University, Department of Earth Sciences, 75236 Uppsala, Sweden
[c] CNDS, Centre of Natural Hazards and Disaster Science, 75236 Uppsala, Sweden

A R T I C L E I N F O

Keywords:
Daily precipitation
Extreme rainfall
Tail behaviour
Fitting method
Kolmogorov-Smirnov test

A B S T R A C T

Study region: We investigate samples from two Italian regions, i.e. Lazio and Sicily, located in central and south Italy, respectively, and characterized by two diverse climates.
Study focus: In engineering practice, the study of maxima daily rainfall values is commonly dealt with light-tailed probability distribution functions, such as the Gumbel. The choice of a distribution rather than another may cause estimation errors of rainfall values associated to specific return periods. Recently, several studies demonstrate that heavy-tailed distributions are preferable for extreme events modelling. Here, we opt for six theoretical probability distribution functions and evaluate their performance in fitting extreme precipitation samples. We select the samples with two common methods, i.e. the Peak-Over-Threshold and the Annual Maxima. We assess the best fitting distribution to the empirical samples of extreme values through the Ratio Mean Square Error Method and the Kolmogorov-Smirnov test.
New hydrological insights for the region: The assessment of the best fitting distribution to daily rainfall of the two different areas investigated here leads to interesting remarks. Despite the diversity of their climate, results suggest that heavy-tailed distributions describe more accurately empirical data rather than light-tailed ones. Therefore, extreme events may have been largely underestimated in the past in both areas. The proposed investigation can prompt the choice of the best fitting probability distribution to evaluate the design hydrological quantities supporting common engineering practice.

1. Introduction

It is widely recognised that heavy rainfall at a daily time-scale has a crucial role in flood risk estimation and consequently in design and management of flood protection works (Koutsoyiannis, 2007). Due to the availability of daily rainfall data across the world, precipitation at daily scale represents one of the most investigated variables in hydrology (De Michele, 2019). Moreover, daily rainfall data measured by precipitation gauges are much more stable and reliable than those measured with hourly or sub-hourly scales. Not least, studies on long-term series of hydrological phenomena need daily precipitation as input (World Meteorological Organization, 1988). Such data are fundamental in both scientific research and engineering practice, as for instance, the estimation of long series

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

https://doi.org/10.1016/j.ejrh.2020.100771
Received 20 August 2020; Received in revised form 2 November 2020; Accepted 8 December 2020
Available online 18 December 2020

analysis (Alpert et al., 2002; Arnone et al., 2013; Bonaccorso and Aronica, 2016), the study and forecasting of drought (Cancelliere et al., 2007; Mineo et al., 2019), the evaluation of the return period for design hydraulic works (De Michele and Avanzi, 2018; Mineo et al., 2018), the assessment of rainfall erosivity (Petkovšek and Mikoš, 2004; USDA-Agricultural Research Service, 2013), the domestic rainwater harvesting (Campisano and Modica, 2012) and the estimation of sub-daily precipitation, through disaggregation techniques (Koutsoyiannis, 2003).

Among the most common practices that require the analysis of daily records, there is the identification of the theoretical probability distribution that allows the best fit to empirical samples. A poor fit to empirical data can lead to under- or over-sizing hydraulic structures (De Michele and Avanzi, 2018; Młyński et al., 2019). Hence defining a proper method to model the extreme rainfall is still challenging for hydrologists and engineers (Cavanaugh et al., 2015).

Traditionally, in the hydrological practice, Annual Maximum (AM) and Peak-Over-Threshold (POT) are the two main approaches to select extreme values from historical time series. The AM method consists in selecting maximum rainfall values, one for each year of observation. The POT sample is achieved by selecting all extreme values that exceed an arbitrarily fixed threshold (Chow, 1964). Several authors provide a comparative survey on the adaptation of different theoretical probability distributions to hydrological empirical samples selected by these two approaches. Madsen et al. (1997a, 1997b) propose a regional estimation scheme for extreme events modelling by selecting AM and POT samples for discharge data recorded in New Zealand. Their findings reveal that the POT-Generalized Pareto Distribution (GPD) model is generally more efficient than the AM-Generalized Extreme Value (GEV) model. Bezak et al. (2014) show that the POT method provides better results than the AM for a flood frequency analysis in Slovenia. Moreover, the POT approach provides a better adaptation to heavy-tailed distributions compared to the AM (Madsen et al., 1997b), and uses more information since it includes much more extreme events (Beguería, 2005). Although the POT method seems to be a valid alternative to the AM method, the impact of threshold selection can greatly influence the analysis of the extreme values. As a matter of facts, the threshold value directly influences both the length of the sample and the statistical independence of its values. These matters are widely investigated in literature and a general accepted methodology for the identification of the optimal threshold value is still an open issue (Beguería, 2005; Mailhot et al., 2013; Serinaldi and Kilsby, 2014). In order to avoid the choice of an *a priori* threshold value, a commonly accepted approach in literature is known as Annual Exceedance Series (AES, Chow, 1964; Koutsoyiannis, 2004a; Gupta, 2011; Papalexiou et al., 2013). This method works by selecting the first *N* extreme values, where *N* is the number of years of record. Even though the AES method avoids evaluating the threshold value, it still does not imply that the selected events are independent. In the past, many scholars investigated dependent samples behaviour (e.g., Hurst, 1951; Mandelbrot and Wallis, 1968; Leadbetter, 1983; Koutsoyiannis, 2008; Volpi et al., 2015). Dependence may introduce some estimation bias if the dependence is of long-range type with high Hurst coefficient (i.e., $H > 0.8$); however, this problem can be dealt with, by taking dependence into account. Also, the formula for return period does not change in presence of dependence of any type (Koutsoyiannis, 2008; Volpi et al., 2015). Nevertheless, we select independent events using the concept of inter-event time definition (IETD; Adams et al., 1986) which is the minimum inter-arrival time between two events and thus aims at obtaining individual events. Independently on the sample selection method, it is crucial to evaluate which probability distribution fits the empirical samples more accurately. Recently, many authors carried out comparisons between different probability distributions to identify the one providing the best fitting to the empirical sample of extreme rainfall events (e.g., Deidda and Puliga, 2006; Papalexiou and Koutsoyiannis, 2013; De Michele and Avanzi, 2018; Młyński et al., 2019). Generally, the most common probability distributions used in engineering practice are the Gumbel and the Fréchet, both belonging to the GEV distributions family (Papalexiou and Koutsoyiannis, 2013). The Gumbel distribution is often used to describe AM series (Mailhot et al., 2013; Serinaldi and Kilsby, 2014), and more generally, is the most common probabilistic distribution used for the study of hydrological extremes (Koutsoyiannis, 2007; De Michele, 2019). At the same time, however, Koutsoyiannis (2004a, 2004b) highlight that the Gumbel distribution is not appropriate for studying hydrological extremes, while the Fréchet distribution (or EV2) results a better alternative. Specifically, the empirical analyses performed by Koutsoyiannis (2007) highlight that the Gumbel distribution may significantly underestimate the largest extreme rainfall amounts, although its predictions for small return periods are satisfactory.

Since a general rule to choose the best probability distribution does not exist, in this paper we propose two coupled methods for assessing their goodness: the Ratio Mean Square Error Method (RMSEM; Papalexiou et al., 2013) and the Kolmogorov-Smirnov test (KS; Keutelian, 1991). Specifically, we test six theoretical probability distributions (i.e. Gamma (G), Lognormal (LN), Weibull (W), Gumbel (Gu), Fréchet (F) and Pareto type II (P)) to fit the empirical daily rainfall data recorded in two Italian regions. Furthermore, we perform the analysis for both the two sample selection methods: AM and AES.

The original contribution of this work is to define the most suitable distribution to describe extreme events selected with the AM and AES, respectively and to assess the effects of an incorrect estimation.

For both regional datasets, results show that heavy-tailed distributions are the most suitable to describe the analysed empirical samples. The investigation shows that the use of different distributions rather than the best fitting one can lead to a considerable under- or over-estimation of the rainfall value associated to a specific return period. This difference increases consistently with the return period. Results backup the analysis performed by Papalexiou et al. (2013) who analysed a global data set. However, in their study, Italy was not assessed thoroughly, and Sicily and Lazio regions remained unexplored. With this work we aim at filling the gap left. The two regions are characterized by the Mediterranean climate; however, their topographic and morphological characteristics make their climate peculiar and diversified in the inland. For the first time, the extension of the dataset allows a thoroughly analysis of daily rainfall extremes and their best fitting distributions.

The paper is organised as follow: in Section 2 we describe the datasets; in Section 3 we present the methodology, while in Section 4 we discuss the results of the fitting performance. Finally, in Section 5, we state the conclusions of the study.

2. Dataset

In this study, we use daily rainfall records observed in two Italian regions: Lazio and Sicily. The data, provided by the two Regional Hydrographic Service networks, consist of 345 and 568 rain gauges, respectively for Lazio and Sicily. Lazio is located in central Italy, while Sicily is in the south (see Fig. 1). The two regions have a very different climate: Lazio is characterized by a mild temperate climate along the coast; while the winter is colder in the inner hilly area; the climate becomes continental in the Apennine area. On the other hand, Sicily is characterized by a Mediterranean climate with long and warm summer and mild and wet winter. Along the southern coast, summer is sweltering because of the African air flow. Because of the diversity of their climate, the assessment of the best fitting distribution to daily rainfall in their two different areas leads to interesting remarks.

To be statistically consistent, in our analysis we select only the rain gauges that fulfilled the following criteria: (i) missing data less than 10 % in each year and (ii) record length greater than or equal to 50 years (World Meteorological Organization, 1988). In Table 1 we show the data samples of rain gauges that satisfied these criteria for both regions. In Fig. 1, we report locations and record lengths of each selected rain gauge.

3. Methods

The paper involves the following aspects: (i) preliminary samples selection starting from daily rainfall records, (ii) definition of the parameters of the different theoretical probability distributions, (iii) description of the theoretical probability distributions, and (iv) statistical validation of the best distribution with a goodness-of-fit test.

3.1. Sample definition

The extreme values theory provides the tools to assess the frequency and severity of extremes rainfall events, by fitting theoretical probability distributions to sample of empirical values (Fisher and Tippett, 1928; Gnedenko, 1943; Gumbel, 1958). Starting with the observed daily data, the first step of our analysis consists in selecting and defining two samples via AM and AES methods, respectively. The goal is to understand which are the consequences, of these two sample selection methods, in terms of rainfall predicted values variability for a fixed return period. It is important to underline that in this paper the term “tail” refers to the upper part of a probability distribution function, that governs, at the same time, both magnitude and frequency of extreme events (Papalexiou et al., 2013). The

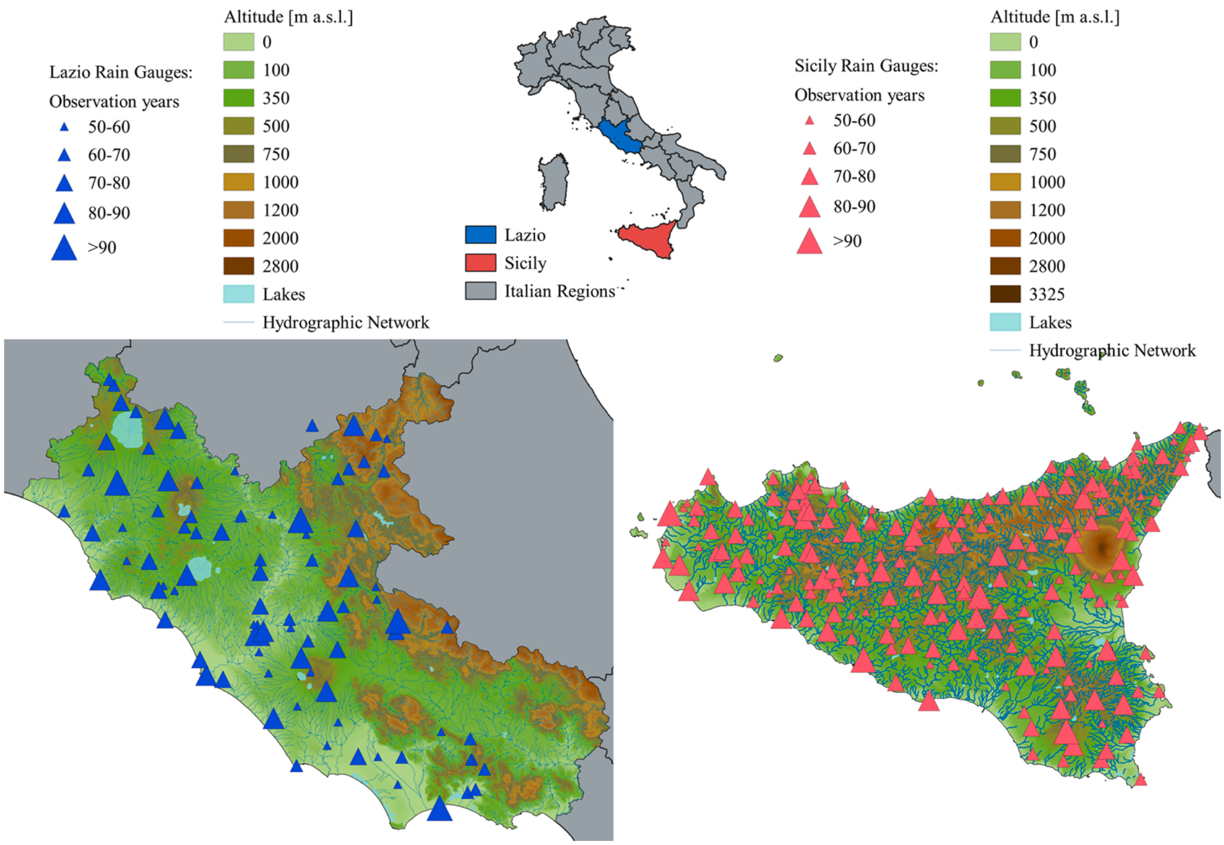

Fig. 1. Location of the two study areas in Italy (upper panel): Lazio (lower left panel); Sicily (lower right panel).

Table 1
Main features of the two dataset.

	Period of observation	Number of rain gauges	Minimum sample size [years]	Average sample size [years]	Maximum sample size [years]
Lazio	1916 - 2018	86	50	68	99
Sicily	1916 - 2014	211	50	69	95

AM method may distort the tail behaviour as it takes into account only the largest value among those occurring in one year. On the other hand, using the AES method, the selected events may not be independent. While dependency has a very low probability to occur in the AM samples, since AM method selects a single extreme event per year. As discussed by Volpi et al. (2015), the sample may be populated by dependent events. Nevertheless, here we select independent events considering an inter-event time (IETD; Adams et al., 1986) of at least 24 h (Jun et al., 2017). We assume that the selected N largest daily values of the record, with N equal to the number of observed years, are representative of the distributions' tail and capable to provide information for the tail behaviour.

From a statistical perspective, an important issue to consider is the presence of ties. Ties can influence the probability distributions' fitting performance to the observed sample. As proposed by Salvadori et al. (2014), using a physically based strategy, ties can be dealt with by introducing randomised precipitation values $\widetilde{R}_j$:

$$\widetilde{R}_j = R_j + \varphi \cdot U_j \tag{1}$$

where R_j is the *j-th* repeated value in a sorted sample of daily precipitation values, U_j is an independent identically distributed random variable ranging from -0.5 to 0.5 and $\varphi = 0.1$ mm is the rain gauges resolution. The choice of φ value depends on it being small enough with respect to the range of R_j, in order to obtain a "statistically equivalent value" of $\widetilde{R}_j$. By applying this procedure, we remove the ties by replacing them with statistically equivalent values.

3.2. Parameters estimation and fitting method

Several works established that using heavy-tailed distributions is more appropriate to describe the behaviour of extreme rainfall events (De Michele and Avanzi, 2018; Deidda and Puliga, 2006; Koutsoyiannis and Papalexiou, 2016; Papalexiou et al., 2013). The implication, due to the use of heavy-tailed distributions, is that more frequent and severe rain events are more likely to be predicted than those characterised by light-tails. This means that the general approach of adopting light-tailed distributions would result in a significant underestimation of the associated risk, with potential implications for human lives.

The identification of the best fitting distribution can be strongly affected (i) by the parameter estimation method, (ii) by a proper interpretation of the goodness of fit test and (iii) finally, by the choice of the distribution that is candidate to fit the empirical sample.

In this study we use a straightforward methodology proposed by Papalexiou et al. (2013) to statistically describe extreme rainfall

Table 2
Main statistical indicators of the datasets: (a) samples selected with the AES method; (b) samples selected with the AM method. The letter N indicates the number of observations of the sample, the postscripts 25, 50 and 75 represent the percentiles values, SD stands for standard deviation.

	Lazio						Sicily					
	Sample Length	Min [mm]	Median [mm]	Mean [mm]	SD [mm]	Max [mm]	Sample Length	Min [mm]	Median [mm]	Mean [mm]	SD [mm]	Max [mm]
Min	50	45,82	54	62,73	11,32	100	50	34	41,98	45,18	9,32	76
$N_{AES,25}$	59	53	65,3	72,37	18,79	145,5	61	42,8	53,64	60,92	19,07	143,4
$N_{AES,50}$	67	57	70	76,74	21,37	169,5	70	47,2	61,2	68,17	24,74	180
$N_{AES,75}$	76	65,4	78,85	86,06	26,58	202,6	76	56,85	72,12	81,12	32,99	229,58
Max	99	94	112	121,06	50,58	395,2	95	123	157	181,33	89,52	702
Mean	67,99	59,71	72,6	79,77	22,94	182,84	68,74	51,47	65,8	74,65	27,32	199,69
SD	11,85	9,12	10,45	10,61	6,11	51,59	10,63	14,1	18,75	21,92	11,63	82,32
Skew	0,52	1,16	1,09	1,15	1,41	1,51	0,08	2,35	2,21	2,22	1,71	1,91

(a)

	Lazio						Sicily					
	Sample Length	Min [mm]	Median [mm]	Mean [mm]	SD [mm]	Max [mm]	Sample Length	Min [mm]	Median [mm]	Mean [mm]	SD [mm]	Max [mm]
Min	50	8,6	49,2	56,84	14,9	100	50	6,6	32,1	39,33	14,81	76
$N_{AM,25}$	59	28	60	65,24	23,46	145,5	61	19,2	48,05	53,86	23,4	143,4
$N_{AM,50}$	67	32,9	63,1	68,58	26,48	169,5	70	22,6	53,1	59,71	29,5	180
$N_{AM,75}$	76	36,6	70	76,29	31,25	202,6	76	27,15	62,85	70,62	37,72	229,58
Max	99	49	103,8	107,85	54,51	395,2	95	50,5	149,8	161,2	98,34	702
Mean	67,99	32,23	65,96	71,67	27,65	182,84	68,74	23,57	57,67	65,5	32,51	199,69
SD	11,85	7,73	9,59	9,34	6,16	51,59	10,63	6,65	16,16	18,94	12,94	82,32
Skew	0,52	−0,37	1,29	1,25	1,25	1,51	0,08	0,66	2,39	2,26	1,76	1,91

(b)

events for both the selected samples (AM and AES as described in Section 3.1). The empirical probability of non-exceedance $F_N(x_i)$, according to the Weibull plotting position, is given by:

$$F_N(x_i) = \frac{r(x_i)}{N+1} \tag{2}$$

where N is the length of the sample (equal to the number of recording years) and $r(x_i)$ is the rank of the rainfall value x_i, sorted in ascending order, with $i = 1,\ldots, N$.

The fitting method consists in the numerical minimisation of a modified mean square error norm RMSE (Młyński et al., 2019; Papalexiou et al., 2018, 2013), here proposed in term of non-exceedance probability:

$$RMSE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{F_N(x_i) - F(x_i)}{1 - F_N(x_i)}\right)^2 \tag{3}$$

where $F(x_i)$ is the theoretical cumulative distribution function (CDF) while $F_N(x_i)$ is the empirical distribution function (EDF).

The RMSE method (RMSEM) is chosen since it allows an almost unbiased estimation of the parameters and it has a smaller variance when compared to other norms for specific distribution (e.g. Pareto and Lognormal) as also found by Papalexiou et al. (2013). Moreover, its ease of use allows practitioners a straightforward application. Therefore, in light of the aforementioned properties, when compared to other widespread norms, RMSE presents a high efficiency and performance.

The RMSE value enables for the unambiguous identification of the fit performance: the distribution with the smallest RMSE represents the best one to describe the empirical analysed sample (Papalexiou et al., 2013). In Table 2 we report the main statistical indicators of the four datasets.

Comparing the basic statistics related to the two regions, the precipitations of Sicily is characterized by more accentuated extremes, both for high and for low records. According to the extreme variability of rainfall values during the year, Sicily is characterised by higher values of the standard deviation (SD) than Lazio. Moreover, the lowest values of the events selected with the AM method are smaller than those of AES, while, obviously, the highest values are selected equally by both methods. The AES samples are therefore characterised by events belonging to a range of values smaller than those of AM.

3.3. *Theoretical probability distributions and fitting performance*

The method proposed in the previous section, allows directly fitting and comparing the adaptation to the empirical data of different theoretical distributions. In this paper, we focus on six probability distributions: Gamma (G), Lognormal (LN), Weibull (W), Gumbel (GU), Fréchet (F) and Pareto type II (P). Primarily, we adopt these distributions since they are the most used in the study of extreme rainfall events, secondly as they are all bi-parametric can be considered comparable in terms of degree of freedom. According to different tail classification (Werner and Upper, 2002; El Adlouni et al., 2008), the selected distributions can be ordered, from heavier-tailed to lighter-tailed, as follows: Pareto type II, Fréchet, Lognormal, Weibull (with shape parameter <1), Gumbel, Gamma, Weibull (with shape parameter >1). The Pareto type II distribution is a power-type distribution defined in $[0,\infty)$. Its CDF is provided by Eq. 4 with a scale parameter $\beta > 0$ and a shape parameter $\gamma \geq 0$.

$$F_P(x) = \left(1 + \gamma\frac{x}{\beta}\right)^{-1/\gamma} \tag{4}$$

The Fréchet distribution is a special case of the Generalized Extreme Value Distribution (GEV, type II), defined in $[0,\infty)$. Its CDF is provided by Eq. 5 and it is defined by the scale parameter $\beta > 0$ and the shape parameter $\gamma > 0$.

$$F_F(x) = e^{-\left[\left(\frac{x}{\beta}\right)\right]^{-\gamma}} \tag{5}$$

Koutsoyiannis (2007) shows that the Fréchet distribution, in its form provided by Eq. 5, is characterised by an heavy tail: moreover, following the classification provided by Ouarda et al. (1994), Fréchet belongs to the case in which the tail has the same behaviour of the Pareto distribution.

The Lognormal distribution is characterised by the scale parameter $\beta > 0$ and the shape parameter $\gamma > 0$ and is also considered a heavy-tailed distribution:

$$F_{LN}(x) = 1 - \left[\frac{1}{2}\mathrm{erfc}\left(\ln\left(\frac{x}{\beta}\right)^{\frac{1}{\gamma}}\right)\right];\ \mathrm{erfc} = 2\pi\int_x^{\infty} e^{-t^2}dt \tag{6}$$

Weibull distribution can be considered as a generalisation of the exponential distribution, with CDF expresses as follow:

$$F_W(x) = 1 - e^{-\left[\left(\frac{x}{\beta}\right)\right]^{\gamma}} \tag{7}$$

The Weibull distribution, with both scale β and shape γ parameters greater than zero, can belong to the heavy-tail class if $\gamma < 1$, or to light-tail class if $\gamma > 1$.

The Gumbel distribution represents another particular case of the GEV, with the CDF given as:

$$F_{GU}(x) = e^{-e^{-\beta(x-\theta)}} \tag{8}$$

For the Gumbel distribution the shape parameter is $\gamma = 0$, and both the scale and the location are $\beta, \theta > 0$. Koutsoyiannis (2007) has shown that the Gumbel distribution, in its form provided by Eq. 8, is characterised by a light tail. Therefore, unlike the other distributions where the shape parameter is an unknown quantity, in Eq. 8, γ is defined and equal to zero. This aspect can be significant in the fitting performance, since the shape parameter governs the asymptotic behaviour of the tail. The last candidate distribution is the Gamma, which, like the Weibull, belongs to the exponential family. Its CDF is:

$$F_G(x) = 1 - \frac{\Gamma\left(\gamma, \frac{x}{\beta}\right)}{\Gamma(\gamma)};\ \Gamma(s,x) = \int_x^\infty t^{s-1} e^{-t} dt;\ \Gamma(s) = \int_0^\infty t^{s-1} e^{-t} dt \tag{9}$$

The tail of the Gamma distribution also has different behaviour based on the value assumed by the shape parameter: (a) for $0 < \gamma < 1$, it has a slightly lighter tail than the exponential tail (as it decreases faster); (b) for $\gamma = 1$, the tail degenerates in the exponential one; (c) for $\gamma > 1$ it exhibits a slightly heavier tail (as it decreases more slowly than the exponential).

3.4. Best fitting analysis: Kolmogorov-Smirnov Test

By solving Eq. 3, the RMSE values and the two parameters for the six candidate distributions are estimated. As mentioned, the RMSEM can be considered a straightforward method to determine the best fit performance between CDFs, since the distribution that has the lowest RMSE value is the one that guarantees the best adaptation to the empirical sample. Although we have long data series, estimated parameters are characterized by uncertainty (McNeil, 1997). To assess the reliability of the theoretical distribution with the empirical one, the KS is adopted (Zeng et al., 2015). It allows to compare cumulative frequency $F_N(x)$ with cumulative probability distribution $F(x)$ (Adirosi et al., 2016). In addition to the KS test, in literature several goodness-of-fit test have been developed, among the others we can list the Cramer-von Mises and the Anderson-Darling test conceived to give heavier weighting to tails of a distribution (Kottegoda and Rosso, 2008). In this study the KS test have been preferred as its reliability makes it a widespread test in daily precipitation extremes studies (Cugerone and De Michele, 2015).

We carry out the popular KS test in order to compare the fitting performance provided by the RMSEM.

The KS test statistic is represented by the maximum distance, in absolute value, between cumulative frequency and hypothesized cumulative distribution (Kottegoda and Rosso, 2008). Since it is often misunderstood in practical engineering, it is worth to recall that if the parameters of the distributions to be tested are estimated from the data, the distribution of the KS test statistic is dependent on the distribution under analysis, and the critical value must be calculated ad hoc. In this paper, we follow the procedure proposed by Keutelian (1991) by estimating critical values of the test statistic using Monte Carlo simulations. For each distribution and each rain gauge station, we estimate the parameters with the RMSEM. Then, we use such parameters to generate 100, 500, 1000, 5000 and 10, 000 samples for each rain gauge, with a length N equal to the original one. Subsequently, we re-estimate the parameters for each of the generated samples with the RMSEM, and compute the maximum difference D_{cr} between the CDF and the cumulative frequencies. We conduct the KS test for two significance levels α, equal to 95 % and 99 %, therefore the critical values of the test statistic are at the 95 % and 99 % percentile, respectively $D_{cr,95}$ and $D_{cr,99}$. The KS test is accepted if the maximum difference in absolute value, D, between the

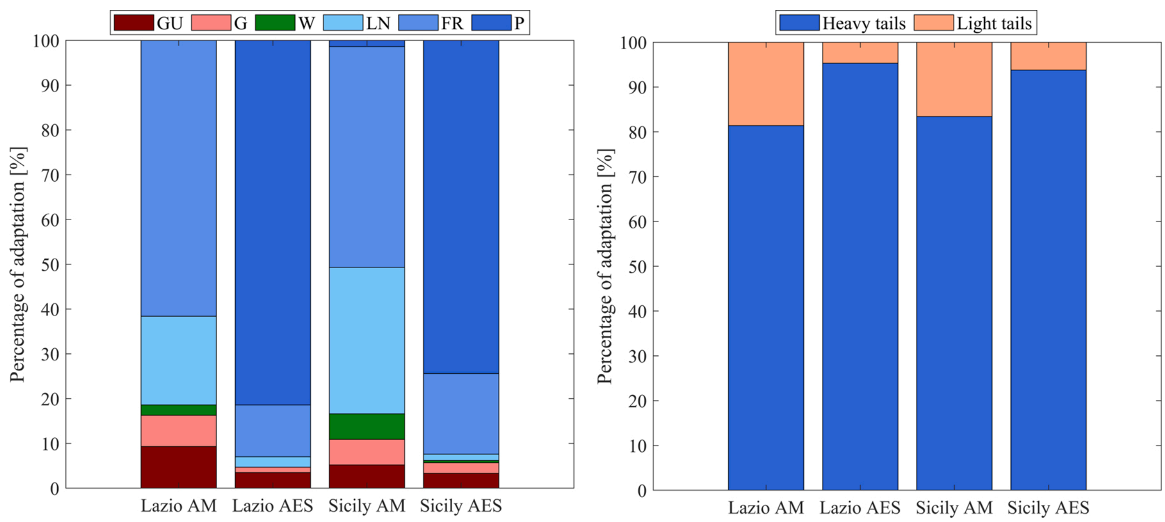

Fig. 2. Percentages of adaptation for the six different probability distributions to the empirical data for the AM and AES selected samples (left panel); percentages of adaptation of heavy and light tails probability distributions for all the AM and AES selected samples (right panel).

CDF and the cumulative frequency of the empirical sample, is smaller than the critical values ($D_{cr,95}$, $D_{cr,99}$). Ultimately, it can be recalled that since both D and D_{cr} are functional for each different distribution and for each empirical sample, the KS test cannot be used to define which one among the candidate distributions best fits a particular sample.

4. Results and discussion

In this paper, we investigate which probability distribution provides the best fit to the samples selected for the two datasets collected in Lazio and Sicily. The first analysis concerns the identification of the parameters of the six theoretical probability distributions with the RMSEM, for both the selected samples (i.e., AM and AES). The results obtained for both datasets, Lazio and Sicily respectively, and for both sample selection methods, are shown in Fig. 2.

It can be argued that for the two regional dataset (i.e. Sicily and Lazio), Fréchet distribution provides the best fit for the AM samples, while AES samples are optimally fitted by the Pareto type II. As expected, the performance of the Gumbel distribution is significantly reduced for the AES samples (Coles et al., 2001; Madsen et al., 1997a). It is interesting to note that for the AES samples, the Weibull distribution never provides the best fit for the Lazio region, and for the Sicilian dataset Weibull distribution results as the best fitting distribution limited to a negligible AES samples percentage (i.e.1%). For the AM samples instead, the Weibull distribution provides the best fit for a low percentage of the analysed stations (2.3 % and 5.7 % respectively for Lazio and Sicily). When the Weibull best fits the empirical data (although for a minimal percentage), it always shows a shape parameter greater than one ($\gamma > 1$): it is therefore representative of a light-tailed distribution. It is noteworthy that the Gumbel distribution presents a much lower percentage in the Sicilian samples compared to the Lazio AM data. These findings backup the results of De Michele and Avanzi (2018) as the Gumbel distribution seems to provide a better fit to a wet pluviometric regime while Sicily is characterized by a diverse climate in its lands.

According to a probabilistic approach, it is possible to ascribe a return period to any rainfall value. In Fig. 3 we show predicted values of extreme rainfall against their corresponding return period. With filled markers, we identify the distribution that provides the best fit to the empirical samples. It is noteworthy to observe that the use of a distribution other than the best fitting one can entail a considerable under- or over-estimation of the rainfall value associated to a specific return period. For the sake of simplicity, in Fig. 3 we represent the results only for two rain gauge stations, one for each region (Subiaco-Scolastica and Caltagirone, respectively for Lazio and Sicily).

In Fig. 3 it is noticeable that for high return periods, there is a remarkable difference in the predicted rainfall values obtained from the six probability distributions. Considering a return period of 1000 years the use of a light-tailed distribution (i.e. Weibull or Gamma in Fig. 3) may significantly under-estimate the rainfall amount. The use of millennial return periods, despite the uncertainty related to their estimation from an empirical sample with limited observations, is commonly required for the design of hydraulic structures (Bertini et al., 2020). Moreover, the choice of the sample selection (i.e. AM and AES) entails different values of predicted rainfall for the same return period. Fig. 4 provides the possible maximum under- and over-estimation (i.e., Δh) of the predicted rainfall value for six different return periods. Particularly, we estimate Δh as the difference of the rainfall depths predicted by the best fitting distribution, and the one predicted by the worst fitting distribution (based on the RMSEM). It is worth noting that for both regions, the AM samples show a greater variability, although the median value is always lower than that obtained for the AES samples. Such aspect is due to the peculiarity of the sample and thus to its selection method: precipitation values selected with the AM method show a wider variability compared to those obtained with the AES method (as also shown in Table 2).

The highlighted variability in Table 2 for the Sicilian datasets is confirmed in Fig. 4(b). Several extreme and catastrophic precipitation events occurred in Sicily during the period of registration: this results in a wider variability in the calculated Δh and makes

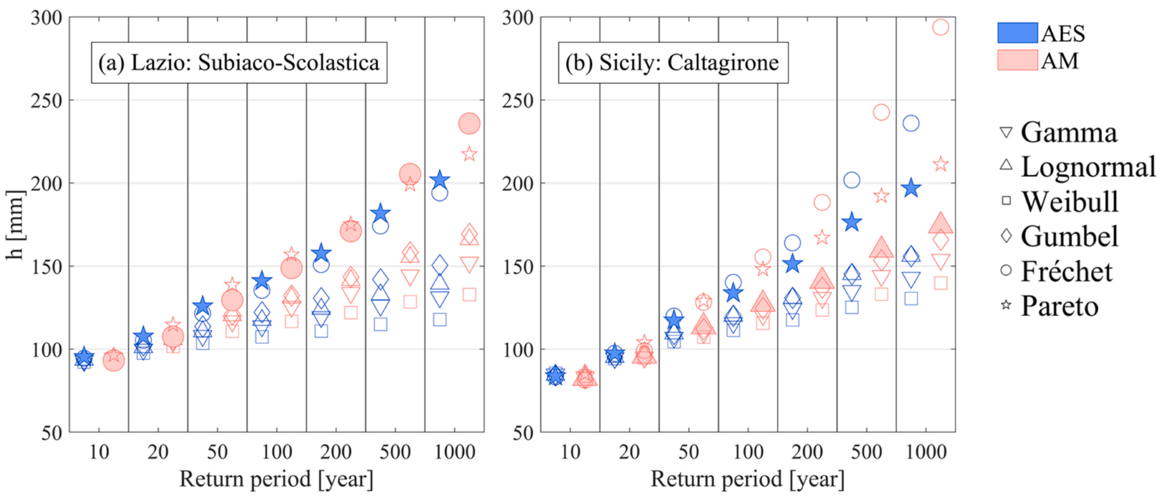

Fig. 3. Variability of the predicted values of extreme rainfall according to the return period, for both methods of sample selection and for both datasets. The best fitting distribution is marked with the filled symbol.

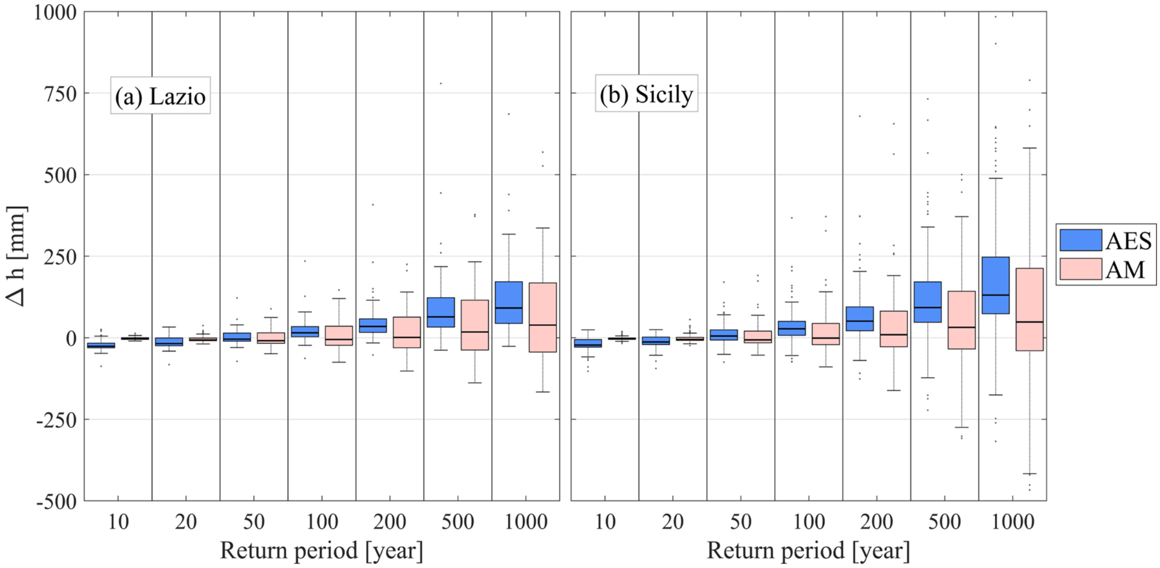

Fig. 4. Differences in the rainfall predicted values for fixed return periods. Δh is evaluated as the difference between the quantile of the best and the worst fitting distributions.

the choice of the best distribution essential.

The last step of our analysis is represented by the KS test to verify if the probability distribution selected with the RMSEM is actually suitable to describe the population of data. We conduct the KS test on both datasets and for both sample selection methods, with two different significance levels: 95 % and 99 %. Starting with the parameters calculated with the RMSEM, we generated 100, 500, 1000, 5000 and 10,000 different samples for each station, with a length equal to the original one. To evaluate the number of synthetic samples suitable for the analysis, as the number of synthetic samples increases, the critical values of each distribution tend to stabilise towards a constant value. This is because the portion of the population, of which the empirical sample represents a realisation, increases with the number of generated synthetic sample. To perform a rigorous statistical analysis, it would be necessary to refer to the highest number of generated samples. However, we notice that already for 5000 synthetic samples the critical value D_{cr} starts to converge to a constant value. Therefore, apart from a significant computational cost in generating 10,000 synthetic samples, it does not significantly influence the results of the KS test. The subsequent elaborations and results presented below refer to a number of synthetic

Table 3
KS test results for both regions and tails selection methods: Lazio (upper panel) and Sicily (lower panel). All the values are given in percentage.

		Lazio							
		α	Gamma	Lognormal	Weibull	Gumbel	Fréchet	Pareto	None
AES	Acceptance rates	99	4.7	9.3	0.0	22.1	74.4	98.8	
		95	3.5	3.5	0.0	14.0	50.0	93.0	
	Best fitting RMSEM		1.2	2.3	0.0	3.5	11.6	81.4	
	KS validated	99	1.2	2.3	0.0	3.5	12.8	80.2	0
		95	1.2	0.0	0.0	4.7	15.1	75.6	3
AM	Acceptance rates	99	76.7	94.2	39.5	94.2	96.5	12.8	
		95	65.1	82.6	18.6	80.2	83.7	3.5	
	Best fitting RMSEM		7.0	19.8	2.3	9.3	61.6	0.0	
	KS validated	99	7.0	19.8	2.3	9.3	61.6	0.0	0
		95	7.0	19.8	2.3	9.3	61.6	0.0	0
		Sicily							
		α	Gamma	Lognormal	Weibull	Gumbel	Fréchet	Pareto	None
AES	Acceptance rates	99	2.8	10.0	0.0	18.5	75.8	98.6	
		95	0.0	2.4	0.0	6.2	53.1	93.4	
	Best fitting RMSEM		2.4	1.4	0.5	3.3	18.0	74.4	
	KS validated	99	0.9	1.4	0.0	3.8	18.5	75.4	0
		95	0.0	1.4	0.0	0.9	20.4	75.4	0
AM	Acceptance rates	99	82.0	94.3	45.0	90.5	91.9	28.4	
		95	58.3	87.2	21.3	75.8	80.6	9.0	
	Best fitting RMSEM		5.7	32.7	5.7	5.2	49.3	1.4	
	KS validated	99	5.7	33.2	5.7	5.2	48.8	1.4	0
		95	6.2	34.6	5.2	5.2	47.4	1.4	0

samples equal to 5,000.

Once defined the value $D_{cr,95}$ and $D_{cr,99}$ for each sample, we consider two different possibilities to choose the final best fitting distribution:

- if the distribution with the lowest value of RMSE had been validated by the KS test, this would have been accepted as the final distribution;
- if instead the distribution with the lowest value of RMSE had been rejected by the KS test, the final distribution was selected among the distributions accepted by the test and corresponding to the one remaining with the lowest RMSE.

Thus, it should be noted that in case the KS test does not accept any probability distribution, the distribution with the lowest RMSE is chosen. In Table 3, we report the rate of rain gauges that pass the KS test for each distribution. It is necessary to specify the distinction between "acceptance" and "validation" of the distribution: the first term indicates whether, for a given station, the distribution was accepted regardless of the result of RMSEM, while the second concerns only distributions with the best fitting RMSEM.

Overall, the AM method always provides a best fitting distribution for each analysed rain gauge. Obviously, we can note that the use of a 95 % significance level leads to more stringent results than the 99 %. It is interesting to highlight that, for the majority of the cases, the RMSEM and the KS are equivalent. However, the RMSEM seems to be easier and offers two main advantages: it allows the estimation of the parameters of the distributions and compares and provide the best fitting distribution. Finally, in Fig. 5, we show the geographic representation of the results relative to a level of reliability of 99 %. The maps in Fig. 5 can be extremely useful as they provide the best fitting distributions for each site. As shown by results, heavy-tailed distributions have a better performance than light-tailed ones in both regions. This suggests that despite the difference in terms of pluviometric regime, heavy-tailed distributions play a dominant role. As a consequence, extreme events may have been under-estimated in the engineering practice whereby the use of light-tailed distributions is widespread with repercussions in the design of hydraulic infrastructures.

5. Conclusions

Analysis of daily precipitation data is a topic of interest in many hydrologic fields and in engineering practice, being their modelling fundamental to define the project loading for both the design of hydraulic structures and the risk analysis.

In this paper, we analyse daily precipitation values recorded by a dense network of rain gauges deployed in two Italian regions with two very different climates, although they both belong to the Mediterranean climate area: Lazio and Sicily. For the analysis, we select only rain gauges having at least 50 years of daily rainfall data with at most 10 % of missing data per year. We use two different selection methods to extract the empirical samples: Annual Maxima (AM) and Annual Exceedance series (AES). To conduct a proper statistical analysis, we process the ties by replacing them with statistically equivalent values. Then, we assess the performance of six theoretical probability distributions among the most used to describe daily rainfall: Pareto Type II (P), Fréchet (F), Lognormal (LG), Weibull (W), Gamma (G) and Gumbel (GU). To fit these six probability distributions to the empirical samples, we use a modified norm of the mean square error (RMSE) to calculate numerically the parameters and simultaneously to identify the best fitting distribution. Alongside the RMSE, we choose the widespread Kolmogorov-Smirnov test to validate statistically the results obtained from the previous method. We perform the Kolmogorov Smirnoff (KS) test generating several synthetic samples using Monte Carlo simulations. The number of generated synthetic samples does not significantly influence the results of the KS test.

The results of the analysis carried out on the empirical samples of both regions, show that heavy-tailed distributions provide a better fit than light-tailed (for about 80 % of stations for both the selection sample methods). In particular, Fréchet distribution is the most performing for the AM samples and the Pareto type II for the AES. This aspect conflicts with what is commonly sensed in engineering practice: in fact, the study of daily rainfall is generally dealt with using a light-tailed probability distribution (i.e. Gumbel). The difference in terms of return period and the related precipitation values is notable, for each of the six candidate distributions, and it increases dramatically for high return periods, especially for the AM samples. As expected, the sample selection method of extreme rainfall values leads to significantly different results. This aspect needs to be further investigated especially considering that, in the engineering practice, the AM is the most widely adopted method.

Despite we consider this study preliminary at this stage, we believe that the proposed investigation can represent a useful occasion to reassess the widespread practice to evaluate design variables as it is of utmost importance the correct estimation of project loading of hydraulic structures.

The RMSEM and the KS test show to be equivalent in terms of results. However, the RMSEM offers two main advantages as it allows the comparison and the definition of the best fitting distribution and of its parameters. This can largely benefit practitioners when dealing with the determination of the best fitting distribution.

The analysis show that heavy-tailed distributions perform better in both regions, despite they have different climatic characteristics. Results are thus in agreement with Papalexiou et al. (2013) who investigated many regions around the globe, except central and south Italy that remained unexplored. In this regard, this paper aims at filling the gap left on these areas. As a result, the project loading of hydraulic structures may have been considerably under- or over-estimated in the past. This difference increases consistently with the return period, as shown in Fig. 4. We thus believe that the proposed investigation can raise awareness to accurately evaluate the design hydrological quantities supporting common engineering practice.

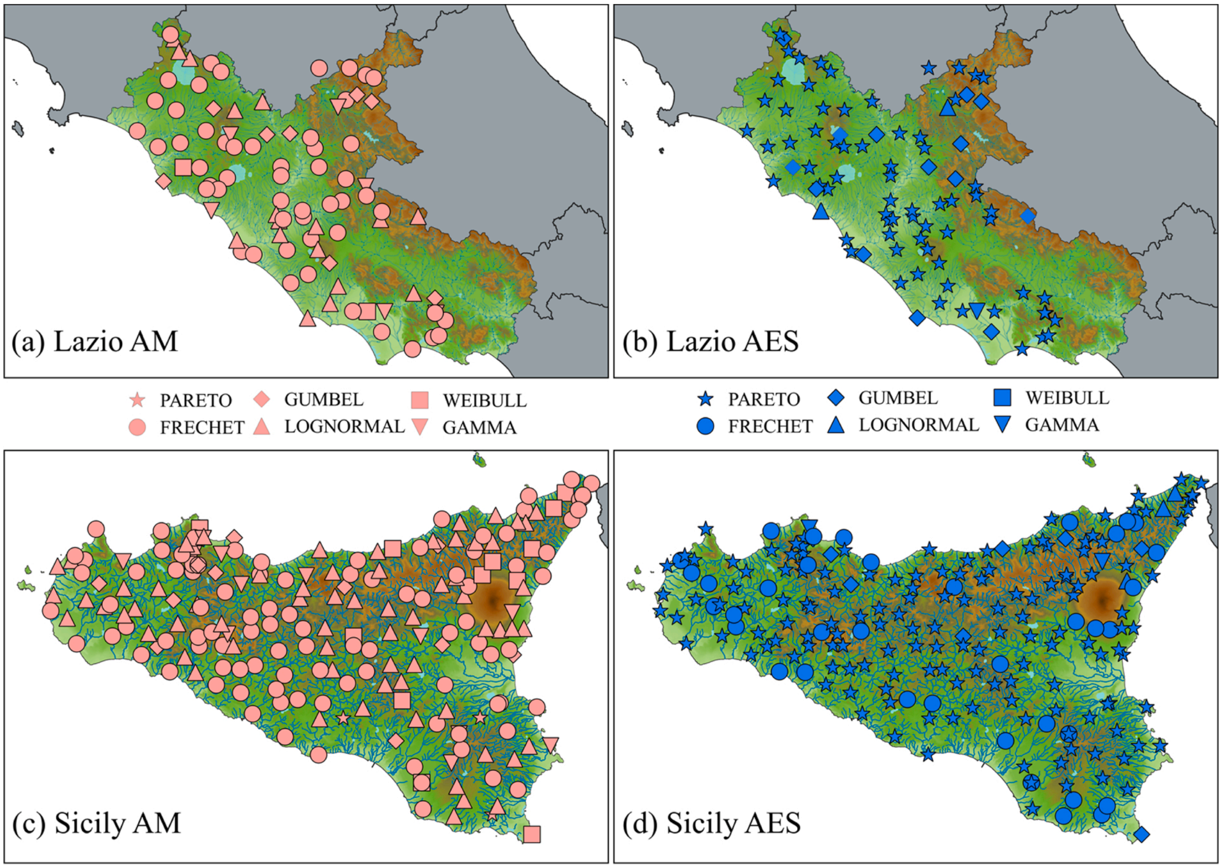

Fig. 5. Best fitting distributions after the application of the KS test for both regions, Lazio (a-b) and Sicily (c-d). The results are relative to a level of reliability of 99 %.

CRediT authorship contribution statement

Benedetta Moccia: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing - original draft, Writing - review & editing. **Claudio Mineo:** Conceptualization, Data curation, Methodology, Writing - original draft, Writing - review & editing. **Elena Ridolfi:** Conceptualization, Supervision, Methodology, Writing - original draft, Writing - review & editing. **Fabio Russo:** Conceptualization, Supervision, Writing - original draft, Writing - review & editing. **Francesco Napolitano:** Conceptualization, Supervision, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgments

The Authors would like to express their sincere gratitude to the Servizio Idrografico e Mareografico of Regione Lazio and Osservatorio delle acque of Regione Sicilia for providing the rainfall data set and to Eleonora Boscariol for her precious contribution in a preliminary analysis. The authors are also thankful for the comments and suggestions provided by Demetris Koutsoyiannis and by an anonymous reviewer that contributed to consistently improve the paper.

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.2020.100771.

References

Adams, B.J., Fraser, H.G., Howard, C.D., Hanafy, M.S., 1986. Meteorological data analysis for drainage system design. J. Environ. Eng. 112, 827–848.

Adirosi, E., Volpi, E., Lombardo, F., Baldini, L., 2016. Raindrop size distribution: fitting performance of common theoretical models. Adv. Water Resour. 96, 290–305. https://doi.org/10.1016/j.advwatres.2016.07.010.

Alpert, P., Ben-Gai, T., Baharad, A., Benjamini, Y., Yekutieli, D., Colacino, M., Diodato, L., Ramis, C., Homar, V., Romero, R., Michaelides, S., Manes, A., 2002. The paradoxical increase of Mediterranean extreme daily rainfall in spite of decrease in total values. Geophys. Res. Lett. 29, 29–32. https://doi.org/10.1029/2001GL013554.

Arnone, E., Pumo, D., Viola, F., Noto, L.V., La Loggia, G., 2013. Rainfall statistics changes in Sicily. Hydrol. Earth Syst. Sci. Discuss. 17, 2449–2458. https://doi.org/10.5194/hess-17-2449-2013.

Beguería, S., 2005. Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value. J. Hydrol. 303, 215–230. https://doi.org/10.1016/j.jhydrol.2004.07.015.

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, 1–20.

Bezak, N., Brilly, M., Šraj, M., 2014. Comparaison entre les méthodes de dépassement de seuil et du maximum annuel pour les analyses de fréquence des crues. Hydrol. Sci. J. 59, 959–977. https://doi.org/10.1080/02626667.2013.831174.

Bonaccorso, B., Aronica, G.T., 2016. Estimating temporal changes in extreme rainfall in Sicily Region (Italy). Water Resour. Manag. 30, 5651–5670. https://doi.org/10.1007/s11269-016-1442-3.

Campisano, A., Modica, C., 2012. Optimal sizing of storage tanks for domestic rainwater harvesting in Sicily. Resour. Conserv. Recycl. 63, 9–16. https://doi.org/10.1016/j.resconrec.2012.03.007.

Cancelliere, A., Di Mauro, G., Bonaccorso, B., Rossi, G., 2007. Drought forecasting using the standardized precipitation index. Water Resour. Manag. 21, 801–819. https://doi.org/10.1007/s11269-006-9062-y.

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

Chow, V.Te, 1964. Handbook of Applied Hydrology. McGraw Hill.

Coles, S., Bawa, J., Trenner, L., Dorazio, P., 2001. An Introduction to Statistical Modeling of Extreme Values, Vol. 208. Springer, London.

Cugerone, K., De Michele, C., 2015. Johnson SB as general functional form for raindrop size distribution. Water Resour. Res. 51, 6276–6289. https://doi.org/10.1002/2014WR016259.

De Michele, C., 2019. Advances in deriving the exact distribution of maximum annual daily precipitation. Water 11. https://doi.org/10.3390/w11112322.

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

Deidda, R., Puliga, M., 2006. Sensitivity of goodness-of-fit statistics to rainfall data rounding off. Phys. Chem. Earth 31, 1240–1251. https://doi.org/10.1016/j.pce.2006.04.041.

El Adlouni, S., Bobée, B., Ouarda, T.B.M.J., 2008. On the tails of extreme event distributions in hydrology. J. Hydrol. 355, 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, 180–190.

Gnedenko, B., 1943. Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire Author. Ann. Math. 423–453.

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

Gupta, S.K., 2011. Modern Hydrology and Sustainable Water Development. John Wiley & Sons. https://doi.org/10.1002/9781444323962.

Hurst, E.H., 1951. Long-term storage capacity of reservoirs. Trans. Amer. Soc. Civ. Eng. 116, 770–799.

Jun, C., Qin, X., Gan, T.Y., Tung, Y.K., De Michele, C., 2017. Bivariate frequency analysis of rainfall intensity and duration for urban stormwater infrastructure design. J. Hydrol. 553, 374–383. https://doi.org/10.1016/j.jhydrol.2017.08.004.

Keutelian, H., 1991. The Kolmogorov-Smirnov test when parameters are estimated from data. CDF Notes.

Kottegoda, N.T., Rosso, R., 2008. Applied Statistics for Civil and Environmental Engineers. Blackwell, Malden, MA.

Koutsoyiannis, D., 2003. Rainfall disaggregation methods: theory and applications. Work. Stat. Math. Methods Hydrol. Anal., pp. 1–23. https://doi.org/10.13140/RG.2.1.2840.8564.

Koutsoyiannis, D., 2004a. Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. Hydrol. Sci. J. 49, 591–610. https://doi.org/10.1623/hysj.49.4.591.54424.

Koutsoyiannis, D., 2004b. Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. Hydrol. Sci. J. 49, 575–590. https://doi.org/10.1623/hysj.49.4.575.54430.

Koutsoyiannis, D., 2007. A critical review of probability of extreme rainfall: principles and models. Adv. Urban Flood Manag. 139–166. https://doi.org/10.1201/9780203945988 ch7.

Koutsoyiannis, D., 2008. Probability and Statistics for Geophysical Processes. Natl. Tech. Univ. Athens. https://doi.org/10.13140/RG.2.1.2300.1849/1.

Koutsoyiannis, D., Papalexiou, S.M., 2016. Extreme rainfall: global perspective. Chow's Handbook of Applied Hydrology. McGraw Hill.

Leadbetter, M., 1983. Extremes and local dependence in stationary sequences. Z. wahrscheinlichkeit 65, 291–306.

Madsen, H., Pearson, C.P., Rosbjerg, D., 1997a. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events 2. Regional modeling. Water Resour. Res. 33, 759–769. https://doi.org/10.1029/96WR03849.

Madsen, H., Rasmussen, P.F., Rosbjerg, D., 1997b. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At site modelling. Water Resour. Res. 33, 747–757.

Mailhot, A., Lachance-Cloutier, S., Talbot, G., Favre, A.C., 2013. Regional estimates of intense rainfall based on the Peak-Over-Threshold (POT) approach. J. Hydrol. 476, 188–199. https://doi.org/10.1016/j.jhydrol.2012.10.036.

Mandelbrot, B.B., Wallis, J.R., 1968. Noah, Joseph, and operational hydrology. Water Resour. Res. 4, 909–918.

McNeil, A.J., 1997. Estimating the tails of loss severity distributions using extreme value theory. ASTIN Bull. J. IAA 27, 117–137. https://doi.org/10.2143/ast.27.1.563210.

Mineo, C., Ridolfi, E., Napolitano, F., Russo, F., 2018. The areal reduction factor: a new analytical expression for the Lazio Region in central Italy. J. Hydrol. 560, 471–479. https://doi.org/10.1016/j.jhydrol.2018.03.033.

Mineo, C., Moccia, B., Lombardo, F., Russo, F., Napolitano, F., 2019. Preliminary analysis about the effects on the SPI values computed from different best-fit probability models in two Italian regions. New Trends Urban Drain. Model. Green Energy Technol., pp. 958–962. https://doi.org/10.1007/978-3-319-99867-1_165.

Młyński, D., Wałęga, A., Petroselli, A., Tauro, F., Cebulska, M., 2019. Estimating maximum daily precipitation in the Upper Vistula Basin, Poland. Atmosphere (Basel) 10, 43. https://doi.org/10.3390/atmos10020043.

Ouarda, T.B.M.J., Ashkar, F., Bensaid, E., Hourani, I., 1994. Statistical distributions used in hydrology. Transformations and asymptotic properties. Sci. Report. Dep. Math. Univ. Moncton, New Brunswick, 31 pp.

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

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

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

Petkovšek, G., Mikoš, M., 2004. Estimating the R factor from daily rainfall data in the sub-Mediterranean climate of southwest Slovenia / Estimation du facteur R à partir de données journalières de pluie dans le climat sub- méditerranéen du Sud-Ouest de la Slovénie the sub-Mediterranean. Hydrol. Sci. J. 49 https://doi.org/10.1623/hysj.49.5.869.55134.

Salvadori, G., Tomasicchio, G.R., D'Alessandro, F., 2014. Practical guidelines for multivariate analysis and design in coastal and off-shore engineering. Coast. Eng. 88, 1–14. https://doi.org/10.1016/j.coastaleng.2014.01.011.

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

USDA-Agricultural Research Service, 2013. Science Documentation Revised Universal Soil Loss Equation Version 2.

Volpi, E., Fiori, A., Grimaldi, S., Lombardo, F., Koutsoyiannis, D., 2015. One hundred years of return period: strengths and limitations. Water Resour. Res. 51, 8570–8585. https://doi.org/10.1002/2015WR017820.

Werner, T., Upper, C., 2002. Time Variation in the Tail Behaviour of Bund Futures Returns, European Central Bank, Working Paper Series, 199.

World Meteorogical Organization, 1988. Analyzing Long Time Series of Hydrological Data with Respect To Climate Variability.

Zeng, X., Wang, D., Wu, J., 2015. Evaluating the three methods of goodness of fit test for frequency analysis. J. Risk Anal. Cris. Response 5, 178. https://doi.org/10.2991/jrarc.2015.5.3.5.