The battle of extreme value distributions: A global survey on the extreme daily rainfall

Simon Michael Papalexiou and Demetris Koutsoyiannis

Department of Water Resources, Faculty of Civil Engineering, National Technical University of Athens, Heroon Polytechneiou 5, GR-157 80 Zographou, Greece (smp@itia.ntua.gr)

Abstract

Theoretically, if the distribution of daily rainfall is known or justifiably assumed, then one could argue, based on extreme value theory, that the distribution of the annual maxima of daily rainfall would resemble one of the three limiting types: (a) type I, known as Gumbel, type II, known as Fréchet and, type III, known as reversed Weibull. Yet, the parent distribution usually is not known and often only records of annual maxima are available. Thus, the question that naturally arises is which one of the three types better describes the annual maxima of daily rainfall. The question is of great importance as the naïve adoption of a particular type may lead to serious underestimation or overestimation of the return period assigned to specific rainfall amounts. To answer this question, we analyze the annual maximum daily rainfall of 15 137 records from all over the world, with lengths varying from 40 to 163 years. We fit the Generalized Extreme Value (GEV) distribution, which comprises the three limiting types as special cases for specific values of its shape parameter, and analyze the fitting results focusing on the behavior of the shape parameter. The analysis reveals that: (a) the record length strongly affects the estimate of the GEV shape parameter and long records are needed for reliable estimates, (b) when the effect of the record length is corrected the shape parameter varies in a narrow range, (c) the geographical location of the globe may affect the value of the shape parameter, and (d) the winner of this battle is the Fréchet law.

Keywords: Generalized Extreme Value distribution, Gumbel distribution, extreme rainfall, annual maximum daily rainfall

1. Introduction

“*Φύσις κρύπτεσθαι φιλεῖ*”— Heraclitus of Ephesus

Arguably, the statistical behavior of the annual maximum daily rainfall has been the cornerstone of statistical hydrology, as it is directly related to the design of hydraulic infrastructures and to extreme floods. In hydrology, the study of rainfall or flood extremes has been an active research field and a matter of debate for more than half a century dating back to the works of E. J. Gumbel in 1940s; however, the field of extreme value theory seems to have originated more than three centuries ago in the works of Nicolaus Bernoulli [see e.g. *Gumbel*, 1958]. Yet, it was during the 20th century when the theory was rapidly evolved and found applications in astronomy, hydrology and engineering in general.

A detailed historical survey on the subject would be out of the scope of this study. Nevertheless, we mention here some of the milestones of this fascinating field [for a more complete historical note see e.g. *Kotz and Nadarajah*, 2000]. It seems that the first methodical approach was due to von Bortkiewicz [1922] regarding the range of random samples. In the sequel, Fréchet [1927] identified one of the asymptotic distributions of maxima, and, soon after, Fisher and Tippett [1928] showed that there are only three possible limiting distributions for extremes. These findings were strengthened by von Mises [1936] who identified some sufficient conditions for convergence to the three limiting laws. Yet, it was Gnedenko [1943] who set the solid foundations of the asymptotic theory of extremes providing the precise conditions for the weak convergence to the limiting laws. All these initial theoretical results were refined and generalized later in the works of Juncosa [1949], Smirnov [1949], Watson [1954], Jenkinson [1955], Barndorff-Nielsen [1963], Berman [1964], de Haan [1971], Balkema and de Haan [1972], Galambos [1972] and Pickands III [1975] to mention some of them. Numerous real-world applications followed this theoretical progress not only in flood and rainfall analysis. It is worth noting in this respect Gumbel's [1958] celebrated book who was one of the pioneers promoting and applying the formal theory into engineering practice.

Accordingly, the central question in extreme rainfall analysis is: which one of the three extreme value distributions, i.e., the Gumbel, the Fréchet or the reversed Weibull, should we choose to describe extreme rainfall? Its answer is not only of academic interest, but mainly constitutes a practical matter of eminent significance as the wrong choice may severely underestimate the design rainfall of hydraulic infrastructures leading thus to infrastructure failures and other negative consequences. Overestimation can also be a possibility, which again has negative consequences in terms of the infrastructure cost. During the last decades, accumulation of observations and advances in computers facilitated the analysis of extreme rainfall and literally thousands of studies or technical reports have been published using, or arguing for or against, a particular extreme value distribution. Yet, most of these studies are of "local" character, e.g., case studies analyzing extreme rainfall in particular areas. As an exception, the study by Koutsoyiannis [2004a,b] used records from several sites in the globe but the number of records was small (169 rainfall records worldwide each having 100-154 years of data). Here, we aim to investigate the behavior of the annual maximum daily rainfall at a global scale, using more than 15 000 rainfall records distributed across the globe, and to provide a better answer to the question we address.

2. Theoretical issues of extreme analysis

2.1 The three limiting laws

It is well known that if a random variable (RV) X follows the distribution $F_X(x)$ then according to the classical extreme value theory the distribution function of the maximum of n independent and identically distributed (iid) RV's, i.e., $Y_n = \max(X_1,\ldots,X_n)$ is given by

$$G_{Y_n}(x) = \left(F_X(x)\right)^n \tag{1}$$

Now, loosely speaking, if $n \to \infty$ three limiting laws can emerge from Eq. (1). Actually, as $\lim_{n\to\infty}\left(F(x)\right)^n$ results in a degenerate distribution, the limiting laws are obtained from $\lim_{n\to\infty}\left(F(a_n x + b_n)\right)^n$ for appropriate constants $a_n > 0$ and b_n [*Fisher and Tippett*, 1928]. In addition, these limiting laws emerge not only for iid RV's as Juncosa [1949] extended these results to the case of non-iid random variables and Leadbetter [1974] proved that the limiting distributions hold also for dependent random variables, given that there is no long range dependence of high level exceedences.

The three limiting laws are the type I or Gumbel (G), the type II or Fréchet (F) and the type III or reversed Weibull (RW) with distribution functions respectively given by

$$G_{\mathrm{G}}(x) = \exp\left(-\exp\left(-\frac{x-\alpha}{\beta}\right)\right) \qquad x \in \mathbb{R} \tag{2}$$

$$G_{\mathrm{F}}(x) = \exp\left(-\left(\frac{x-\alpha}{\beta}\right)^{-1/\gamma}\right) \qquad x \geq \alpha \tag{3}$$

$$G_{\mathrm{RW}}(x) = \exp\left(-\left(-\frac{x-\alpha}{\beta}\right)^{1/\gamma}\right) \qquad x \leq \alpha \tag{4}$$

All three distributions comprise a location parameter $\alpha \in \mathbb{R}$ and a scale parameter $\beta > 0$, with the Fréchet and the reversed Weibull distributions having the additional shape parameter $\gamma > 0$. Although the expressions of the Fréchet and the reversed Weibull distributions look very similar, i.e., they differ in a couple of signs, the distributions behave completely differently as the first is bounded from below while the second is bounded from above. Noteworthy, the exponential form of the Fréchet distribution does not imply an exponential right tail, i.e., the Fréchet distribution behaves like a power-type distribution as it can be easily proved that for $\gamma > 0$ the function $1 - \exp(-x^{-1/\gamma})$ is asymptotically equivalent to $x^{-1/\gamma}$ (it is reminded that two functions $f(x)$ and $g(x)$ are asymptotically equivalent if $\lim_{x\to\infty} f(x)/g(x) = 1$). Likewise, the double exponential form of the Gumbel distribution does not imply a double exponential tail, as its right tail is asymptotically equivalent with the exponential tail, i.e., $\exp(-x)$.

Now, any specific parent distribution $F_X(x)$ belongs to the domain of attraction of one the aforementioned limiting laws. To which one depends mainly on the form of its right tail. Several formal mathematical conditions determine the distribution's domain of attraction (formed originally by von Mises [1936] and Gnedenko [1943] and extended by several other authors [for a complete account see e.g. *Embrechts et al.*, 1997; *Reiss and Thomas*, 2007]). Generally speaking, distributions with right tail regularly varying in infinity or, equivalently, not having all of their moments finite, belong to the domain of attraction of the Fréchet law. These include power-type distributions like the Pareto, the Burr type XII and III, the Log-Gamma, the Cauchy and others. In contrast, in the domain of attraction of the Gumbel law belong all distributions with right tail tending to zero faster than any power-type tail, or equivalently distributions having all of their moments finite, e.g., Normal, Lognormal, Gamma, Weibull and others. Finally, in the domain of attraction of the reversed Weibull law belong distributions bounded from above [see e.g. *Kotz and Nadarajah*, 2000].

The afore mentioned three limiting distribution laws can be unified into a single expression known as the Generalized Extreme Value (GEV) distribution (also known as the Fisher-Tippet) with probability distribution function given by

$$G_{\mathrm{GEV}}(x) = \exp\left(-\left(1+\gamma\frac{x-\alpha}{\beta}\right)^{-1/\gamma}\right) \qquad 1+\gamma\frac{x-\alpha}{\beta} \geq 0 \tag{5}$$

This parameterization was proposed by von Mises [1936], although it is commonly attributed to Jenkinson [1955]. The distribution comprises the location parameter $\alpha \in \mathbb{R}$ the scale parameter $\beta > 0$ and the shape parameter $\gamma \in \mathbb{R}$. It can be easily seen that for $\gamma > 0$ it is bounded from below, ($x \geq \alpha - \beta / \gamma$) while for $\gamma < 0$ it is bounded from above ($x \leq \alpha - \beta / \gamma$) (notice that here positive γ means a GEV bounded from below, while some texts use opposite sign convention). Essentially, the GEV distribution formula can be seen as a simple reparameterization of the Fréchet formula as the Fréchet parameters (indexed with F in Eq. (3)) are related with the GEV parameters, i.e., $\alpha_{\mathrm{F}} = \alpha - \beta / \gamma$, $\beta_{\mathrm{F}} = \beta / \gamma$ and $\gamma_{\mathrm{F}} = \gamma$. This simple reparameterization exploits the limiting definition of the exponential function, i.e., $\lim_{\gamma \to 0}\left(1+\gamma x\right)^{-1/\gamma} = \exp(-x)$ so that the Gumbel distribution emerges for $\gamma \to 0$.

2.2 Convergence to the limiting laws

The distribution of the maximum value, given in Eq. (1), converges to one of the three liming laws (depending on the parent distribution) given that the maximum value is selected from a number of variables which tends to infinity. In real world, convergence practically holds if this number is very large. However, in daily rainfall it seems that this number is not even large as in the best case it would equal the number of the year's days, i.e., 365 or 366 values. Actually, the number of rainy days N_{R} that depends on the probability dry is always smaller than the number of year's days and varies from year to year. Thus, whether or not the annual maximum can actually be modeled by one the three limiting laws should not be taken for granted [see also *Koutsoyiannis*, 2004a].

To demonstrate this issue, we use results from a previous study [*Papalexiou and Koutsoyiannis*, 2012] where we analyzed more than ten thousand daily rainfall records and we found that the Burr type XII distribution (BrXII) and the Generalized Gamma distribution (GG), are both very good models for describing the non-zero daily rainfall. Their probability density functions are given, respectively, by

$$f_{\mathrm{BrXII}}(x) = \frac{1}{\beta}\left(\frac{x}{\beta}\right)^{\gamma_1 - 1}\left(1 + \gamma_2\left(\frac{x}{\beta}\right)^{\gamma_1}\right)^{-\frac{1}{\gamma_1\gamma_2}-1} \qquad x \geq 0 \tag{6}$$

$$f_{\mathrm{GG}}(x) = \frac{\gamma_2}{\beta\,\Gamma(\gamma_1/\gamma_2)}\left(\frac{x}{\beta}\right)^{\gamma_1 - 1}\exp\left(-\left(\frac{x}{\beta}\right)^{\gamma_2}\right) \qquad x \geq 0 \tag{7}$$

Hence, if we assume that both of these distributions can serve as parent distributions, then for a constant number of rainy days N_{R} we could form the exact distribution of the annual maximum that would respectively be $G_{\mathrm{BrXII}}(x) = \left(F_{\mathrm{BrXII}}(x)\right)^{N_{\mathrm{R}}}$ and $G_{\mathrm{GG}}(x) = \left(F_{\mathrm{GG}}(x)\right)^{N_{\mathrm{R}}}$. It is noted that the BrXII distribution as a power type distribution belongs to the domain of attraction of the Fréchet law; in contrast, the GG distribution is of exponential type, having all of its moments finite and thus belonging to the domain of attraction of the Gumbel law. So, theoretically speaking the first is expected to converge to the Fréchet law and the second to the Gumbel law.

The different daily rainfall records analyzed in the aforementioned study had different statistical characteristics, yet, in order to illustrate the convergence rate based on real world evidence we proceed as follows. First we consider as representative statistics of the nonzero daily rainfall the median (closer to the mode than the mean value) of the sample estimates of the first L-moment λ_1 (mean), of L-variation τ_2 and of L-skewness τ_3; their numerical estimates are $\lambda_1 = 9.86$, $\tau_2 = 0.58$, $\tau_3 = 0.45$ (all parameters with dimensions, e.g., λ_1 or scale parameters, are expressed in mm). Additionally, the median of probability dry was 76.3% corresponding approximately to $N_R = 87$ rainy days. These statistics can be reproduced by a BrXII distribution with parameters $\beta = 8.47$, $\gamma_1 = 0.91$, $\gamma_2 = 0.18$, and a GG distribution with parameters $\beta = 1.83$, $\gamma_1 = 1.16$, $\gamma_2 = 0.54$. For these parent distributions and for $N_R = 87$ we calculated (numerically) the parameters of the exact distribution of the annual maximum for each case. Namely, the G_{BrXII} would have $\lambda_1 = 77.62$, $\tau_2 = 0.23$, $\tau_3 = 0.30$ and the G_{GG} would have $\lambda_1 = 73.71$, $\tau_2 = 0.20$, $\tau_3 = 0.24$. Next we found the corresponding GEV and Gumbel distributions to these parameters, i.e., for the G_{BrXII} parameters the GEV will have $\alpha = 60.71$, $\beta = 20.85$, $\gamma = 0.19$, and the Gumbel will have $\alpha = 62.72$, $\beta = 25.80$. Likewise, for the G_{GG} parameters the GEV will have $\alpha = 60.48$, $\beta = 19.15$, $\gamma = 0.10$, and the Gumbel will have $\alpha = 61.43$, $\beta = 21.28$.

This analysis is graphically depicted in Figure 1 where the fitted distributions are formed in a Rainfall *vs*. Return period plot. It can be easily shown that the exact annual maximum laws, i.e., the G_{BrXII} and the G_{GG} are given by the relationship $x(T) = Q_{X|X>0}\left(\left(1-1/T\right)^{1/N_R}\right)$, where T denotes the return period in years and $Q_{X|X>0}$ the quantile function of the representative BrXII or GG distribution describing the nonzero daily rainfall. The graph reveals that the exact annual maximum law, assuming as a parent distribution the BrXII, quickly converges to the anticipated Fréchet law or GEV with positive γ. Noteworthy, the tail index of the representative BrXII, expressed by the shape parameter γ_2, and the shape parameter γ of the GEV distribution, theoretically should be the same. In reality, while they are not exactly the same, they are very close, i.e., $\gamma_2 = 0.19$ and $\gamma = 0.18$, verifying thus a satisfactory convergence. On the other hand, assuming the GG as a parent distribution, we see that not only does the exact law G_{GG} not converge to the Gumbel law as theoretically expected, but it is better described by the Fréchet law. In this case the GEV overestimates the rainfall for large return periods, yet, it is on the safe side, whereas it is clear that the Gumbel distribution severely underestimates it.

This analysis indicates that even if the parent distribution of daily rainfall is of exponential type, belonging thus theoretically to the domain of attraction of the Gumbel law, the annual maximum is better described by the Fréchet law [see also *Koutsoyiannis*, 2004a]. Is this a paradox? The answer is no. The reason is that the convergence to the Gumbel law is very slow; actually, it does not converge satisfactorily even for $n = 10^7$ as our tests showed. On the contrary, the additional shape parameter of the Fréchet law or of the GEV distribution, adds the required flexibility to this distribution to "imitate" the shape characteristics annual maxima even if the parent distribution does not belong to its domain of attraction. Thus, although the Fréchet law has a power type tail, its flexibility enables it to better describe, compared to Gumbel law, other heavy-type tails like the stretched exponential or the lognormal. Noteworthy, a recent study [*Papalexiou et al.*, 2012] where more than 15 000 daily records were analyzed focusing on the tail behavior of the parent distribution, revealed that the daily rainfall tail is better described by heavy tails. This offers a theoretical argument favoring the use of the Fréchet law in any case instead of Gumbel.

3. The original dataset

In this study we use more than 15 000 rainfall records distributed across the globe. The original data were daily rainfall records obtained from the Global Historical Climatology Network-Daily database (version 2.60, www.ncdc.noaa.gov/oa/climate/ghcn-daily) which includes thousands of records worldwide. We mention though, that many records of this database have a large percentage of missing values, are short in length, e.g., just a few years, or, contain suspicious values in terms of quality (for the quality flags used refer to the aforementioned website).

Thus, among the several thousands of records we studied only those satisfying the following criteria: (a) record length greater or equal than 50 years, (b) percentage of missing values per record less than 20%, and (c) percentage of values assigned with "quality flags" per record less than 0.1%. Special attention was given to values assigned with quality flags "G" (failed gap check) or "X" (failed bounds check) as these values are suspiciously large, e.g., could be orders of magnitude larger compared to the record's second larger value. These extremely large values (probably resulting from recording or registering errors), could alter the record's statistics, and thus we had to identify and delete them (yet, only 594 records contained such values and typically one or two values at each record had to be deleted). The resulted number of records after screening with these criteria is 15 137. The locations of those records are depicted in the map given in Figure 2.

4. A method for extracting the maxima

4.1 Selection procedure

The original dataset comprises daily rainfall records, thus, in order to study the annual maximum daily rainfall we must form the time series of annual maxima. If the original records did not contain any missing-values then forming the annual maximum time series would be trivial. Yet, missing-values occur commonly, and specifically, in the dataset analyzed here records may contain up to 20% of missing-values. Usually, within a record only some years are incomplete, (contain missing-values); hence, the problem is how we can extract the maximum value of incomplete years. Evidently, the recorded maximum value of an incomplete year may not be the real one, as it is likely for a larger value to have occurred in days of missing data. Moreover, as the percentage of missing values gets higher the more probable it becomes that the real maximum has been recorded. Thus, years with missing values, if not treated appropriately, could result in significant errors that may affect the conclusions drawn from the data analysis.

Basically, one could think of three different methods to extract the annual maxima from a daily time series containing missing values: (a) in the first method (M1), specific criteria are used to assess the validity of the annual maxima, e.g., the annual maximum value could be considered valid only if the missing-values percentage is small, (b) in the second method (M2), only the maxima of complete years are accepted as valid while those of incomplete years are assumed unknown, and (c), in the third method (M3), the annual maxima are extracted irrespective of the years' missing-values percentage. Clearly, the method M3 is not safe because, if the missing-values percentage is high, it will result in underestimated maxima. Method M2 is safe and we could be sure that the extracted maxima are the real ones, yet it does not fully utilize the available information. For example, a record may contain many years with just a few missing values per year; according to method M2 all these years would be excluded, thus leading to an unjustifiably small sample. So, it is clear that the most reasonable choice is to set some criteria that need to be fulfilled in order to accept an extracted annual maximum as valid.

It is reasonable to assume that it is safe to extract the annual maximum of those years with small missing-values percentage. Nevertheless, two problems arise. First, the definition of "small" would be subjective, e.g., 1% or 10% could be considered small, and second and most important, maxima of incomplete years may be much greater compared to those of complete years. For example, a year with 90% of missing values may contain the record's maximum; would it be rational to exclude this value? Of course, larger values may have occurred within an incomplete year but this would be unlikely. For these reasons we deem that the acceptance or not of a value extracted from an incomplete year, as the annual maximum, should be based on two criteria; first, on the missing-values percentage, and second, on the value's rank, i.e., its relative position in the extracted sample of maxima after it has been sorted in ascending order (the smallest rank is given to the smallest value).

Accordingly, the annual maxima time series are formed in two steps: (a) the maximum of each year is extracted irrespective of the year's missing-values percentage and, (b) the values of this initial series are tested according to the criteria set and those not fulfilling them are deleted from the time series, i.e., they are assumed unknown. Namely, two criteria, whose validity is justified in section 4.3, were set to justify deletion of a value whenever both hold: (a) the rank is smaller or equal than $40\% \times N$ (where N is the sample size) which means that the particular value belongs to the 40% of the lowest values, and (b) the missing-values percentage within a year is larger than or equal to 1/3 which means that in the particular year approximately the values of more than four months are missing. The method is graphically explained in Figure 3 which depicts along with the annual maxima time series the corresponding percentages and ranks of missing values. Essentially, the method's rationale is simple; if an incomplete year has a high percentage of missing values and its maximum is small compared to the maxima of the other years, then there is a high probability for larger values to have occurred within this year and thus this value should not be accepted as the real annual maximum.

4.2 Validation of the method

One could argue that the criteria defined previously are subjective and different values could be set as thresholds both for the rank and percentage of the missing values. Yet, these thresholds where not selected unjustifiably, but rather emerged after extended Monte Carlo simulations. Particularly, a Monte Carlo scheme was planned and performed in order to validate the method performance and specify the appropriate criteria values. The Monte Carlo scheme could be summarized in four basic steps: (a) a subset of complete daily records is selected and the annual maxima series are created, (b) this daily-records subset is modified to contain missing values, (c) annual maxima series are extracted from the modified daily-records subset by utilizing the maxima extraction method for various criteria values, and (d) the real maxima series created in step (a) are compared with those created in step (c). In other words, the basic idea is to find, if possible, those threshold values resulting in maxima series with statistical characteristics similar to the real ones.

Obviously, to validate the method we need daily time series that are complete. Yet, only few records of the dataset are totally complete, hence, for start we selected those with very small missing-values percentage, i.e., less than 0.1%, and we deleted, if existed, the few incomplete years per record in order to be absolutely certain for the resulting annual maxima series. The result was 1 003 daily rainfall records with lengths varying from 38 to 155 years.

Now, the records of the dataset analyzed here contain missing-values up to 20%, and these values are distributed among some of the record's years, i.e., only a percentage of the record's years are incomplete. To identify how the percentage of incomplete years per record is distributed we studied all 15 137 records. The empirical distribution is presented in Figure 4, as well as a fitted Beta(α,β) distribution, that will be valuable in the sequel, with estimated parameters $\alpha = 1.32$ and $\beta = 2.41$.

In order to construct time series with missing values distributed similar to the real ones we modified each one of the aforementioned daily records by the following procedure: (a) we generated a random number p_{MV} less than 20% that represents the missing-values percentage of the record, (b) we set the record's total missing-values number then as $n_{\mathrm{MV}} = p_{\mathrm{MV}} \times 365 \times N$, where N is the record's length in years, (c) we distributed the n_{MV} missing values to $N_{\mathrm{MV}} = p_{\mathrm{Y}} \times N \geq n_{\mathrm{MV}} / 365$ years, where p_{Y} is the percentage of incomplete years and is randomly generated from the fitted Beta distribution depicted in Figure 4, (d) we randomly split the number n_{MV} into N_{MV} parts in order to define the number of missing values for each incomplete year, and (e) we selected N_{MV} years randomly from the record and we deleted the number of values previously defined randomly from each year.

Finally, the annual maxima series extracted by the modified records were compared to the corresponding real ones based on four basic statistics, i.e., the mean as a measure of central tendency, the L-variation as a measure of dispersion, and the L-skewness and L-kurtosis as measures of shape characteristics. We applied the maxima extraction method (M1) repeatedly by altering the criteria values until the resulting series were statistically similar to the real ones; this led to the aforementioned threshold values. We also compared the maxima series extracted by methods M2 and M3 to the real ones. Figure 5 presents the box plots formed by the 1 003 differences between the statistics of the real annual maxima series and the ones extracted from the daily series modified to contain missing values.

As expected, method M3 (the one in which maxima are extracted irrespective of the percentage of missing-values) is inappropriate because it significantly alters the statistical character of the extracted maxima series while method M2 does not. Interestingly, not only does method M1 preserve the statistical characteristics (the median is zero and approximately equals the mean as the box plots are almost symmetric) but performs better than method M2. The explanation is that method M1 generates time series with larger length, compared to those of method M2, as fewer values are deleted. Apparently, larger time series means more information and thus more accurate sample estimates. Finally, it is worth noting that the overall range of the differences, taking into account that sample estimates of shape characteristics are usually very uncertain, is very small.

5. Analysis and results

5.1 Fitting results

The application of the maxima extraction method (it is noted that the annual maximum value is determined per calendar year, which is a more appropriate time basis for a study of global rainfall) produced 15 137 annual maximum daily rainfall time series with length varying from 40 to 163 years. To obtain a general idea of the statistical behavior of the annual maximum daily rainfall we calculated basic summary statistics for all records of maxima. The results are given in Table 1. Noteworthy, all statistical characteristics (mean, standard deviation, skewness, L-skewness, L-kurtosis) vary significantly; for example, the mean ranges, from 9.1 mm to 863.7 mm and the standard deviation from 3.9 mm to 430.7 mm. In particular, the large variation of shape characteristics, indicates that any distribution with fixed shape will be inadequate for describing the annual maximum daily rainfall. Consequently, this portends the Gumbel distribution’s inability as a universal model as its shape characteristics are fixed.

We can expect that in some cases the Gumbel distribution suits better, while in other cases the Fréchet, or, even the reversed Weibull are more appropriate; in fact all three distributions have been used in the literature. Theoretically, the estimated shape parameter of a fitted GEV distribution reveals which one of the three distributions performs better, as all of them emerge for specific values of γ. Yet, the Gumbel distribution arises for $\gamma \to 0$, and thus, even if the sample is indeed drawn from a Gumbel distribution the estimated GEV shape parameter (irrespective of the fitting method used) will never be exactly zero. In the literature more than thirteen tests can be found for testing whether the estimated GEV shape parameter can be assumed zero [*Hosking*, 1984]. Nevertheless, all these tests examine whether the null hypothesis H_0: $\gamma = 0$ can be rejected or not. Clearly, a sample not rejecting the null hypothesis does not imply that $\gamma = 0$, or equally, that the underlying distribution is the Gumbel. It is highly probable for a null hypothesis with small values of γ, e.g., H_0: $\gamma = -0$ .01, or, H_0: $\gamma = 0.01$, not to be rejected. Hence, we deem that it is not possible to conclude with certainty applying statistical tests whether the underlying distribution is Gumbel or GEV with γ close to zero.

Nevertheless, apart from the aforementioned tests, graphical tools exist that are especially useful when dealing with a large number of records, which can help to make inference about the underlying distribution. A graphical tool that has gained popularity over the last decade, introduced by Hosking [1990], is provided by the L-moments ratio diagrams. L-ratio plots have superseded classical moments ratio plots as they are superior in many aspects [see e.g., *Hosking and Wallis*, 1993; *Hosking*, 1992; *Peel et al.*, 2001; *Vogel and Fennessey*, 1993]. Essentially, this tool provides a graphical comparison between observed L-ratio values and points or lines or even areas formed by the theoretical formulas of parametric distributions. Figure 6 depicts in an L-kurtosis *vs*. L-skewness plot the 15 137 observed points as well as the theoretical point and line corresponding to the Gumbel and the GEV distributions, respectively. Interestingly, only 20% of points lie on the left of the Gumbel distribution (corresponding to a GEV distribution with $\gamma < 0$;

reversed Weibull law), while 80% of points lie on the right (corresponding to a GEV distribution with $\gamma > 0$; Fréchet law). Also it is worth noting that the average point lies almost exactly on the GEV line and corresponds to $\gamma \approx 0.1$. Figure 6 may not reveal the percentage of points that could be described by a Gumbel distribution, yet, it offers a clear indication that the Fréchet law prevails.

As mentioned before, the GEV shape parameter value indicates the type of the limiting law, a fact that emphasizes the importance to study in depth the behavior of this parameter. To this aim, we fitted the GEV distribution to all available records, and for the completeness of the analysis we also fitted the Gumbel distribution. Both distributions were fitted using the method of L-moments [see e.g., *Hosking*, 1990], as especially for the GEV distribution it has been shown [*Hosking et al.*, 1985] that L-moments estimators are even better than maximum likelihood estimators in terms of bias and variance for samples up to 100 values. The fitting results are shown in Table 2 where various summary statistics of the estimated parameters are given. The table shows the large variation of the estimated GEV shape parameter, which ranges from −0.59 to 0.76 with mean value 0.093; the 90% empirical confidence interval is evidently much smaller, i.e., form -0.11 to 0.28 . The empirical distribution of the GEV shape parameter is depicted on Figure 7 along with a fitted normal distribution with mean 0.093 and standard deviation 0.12.

5.2 GEV shape parameter *vs.* record length

Larger samples offer more accurate estimates because, obviously, the variance of an estimator decreases as the sample size gets larger. Unambiguously thus, the estimate of the GEV shape parameter is expected to be more accurate if based for example on a 100-year record rather than on a ten-year record. In this respect, we study the estimated GEV shape parameter in relationship with the record length as our records vary in length from 40 to 163 years. First, we grouped the

15 137 estimated shape parameter values into nine groups based on the length of the record that were estimated; and second, we estimated various statistics for each group. The summary statistics of each group are given in Table 3, while the mean value and the percentage of records with positive shape parameter in each group are depicted in Figure 8. Clearly, Figure 8 indicates an upward "trend" in the mean shape parameter value over record length, e.g., for the 40-50 years group the mean value of γ is 0.077 while for the last group (with ≥ 121 years) it is markedly larger, i.e., 0.116. Additionally, as the values of Table 3 attest, the standard deviation, as expected, decreases over the record length, e.g., for the 40-50 years group it is 0.141 while for the one with ≥ 121 years it is 0.088. Obviously the smaller the standard deviation the smaller the parameter range, yet we note the drastic decrease, e.g., in the 90% empirical confidence interval (ECI) of γ, which for the 40-50 years group is [−0.152, 0.312] while for the one with ≥ 121 years it is [−0.029, 0.263]. Another key issue to emphasize is the upward "trend" of the percentage of positive γ over record length. This percentage is large (71.8%) even in the 40-50 years and for the group with ≥ 121 years it gets as high as 91.0%, providing a clear indication that the Fréchet law prevails.

The previous analysis gave a clear indication that a relationship between the estimated GEV shape parameter and the record length exists, yet, this relationship is not exactly revealed as the variation in the mean value, as shown in Figure 8, does not suggest a precise law. Nevertheless, if such a law exists, we should conclude that the previous grouping technique fails to reveal its exact form because the record length is not uniformly distributed within the groups (e.g., the 51-60 years group contains 3610 records but this does not imply that there are 361 records of 51 years, 361 records of 52 years, etc.). Thus, in order to create records with exactly the same length, we modified the existing ones by partitioning or cutting off a number of values.

Specifically, we selected records with length greater or equal than 80 years (5 049 records; it would be extremely laborious to use all records), and we partitioned each one into lengths ranging from ten to 115 years increased by a step of five years. The 115-year “upper limit” emerged by demanding at least 1000 records at each record length, a number we deem is large enough to offer a robust analysis (there are 1046 records with length ≥ 115 years and only 540 with length ≥ 120 years). For instance, applying this technique, a 112-year record is partitioned into eleven 10-year records or yields only one 90-year record and obviously none 115-year record. In total the 5 049 selected records generated, for example, 49 270 ten-year records and 1046 115-year records. For all these records at each record length we estimated the GEV shape parameter using the L-moments method.

Figure 9a depicts the observed mean and the 95% confidence interval (CI) values of the GEV shape parameter for the various record lengths as well as the corresponding fitted theoretical functions. The fitted curves have the form $g(L) = a + b\,L^{-c}$, with $c > 0$, L denoting the record length and a, b, c parameters estimated here with a least square error fitting. This formula was figured out so as to have two desiderata: The first stems from the fact that the observed values indicate clearly that the mean and the CI values do not increase or decrease linearly over the record length. Rather, it is reasonable to assume that they tend asymptotically to a fixed value. Clearly, as $L \to \infty$ the function $g(x) \to a$ with a thus expressing the limiting value. The second desideratum is this function to be simple and flexible. Indeed, for $b < 0$ it is concave and for $b > 0$ it is convex, thus being suitable to describe both upward and downward “trends” that converge to a liming value. The estimated parameters for the fitted curves are as follows: (a) for the lower CI curve, $a = 0.021$, $b = -3.90$, $c = 0.80$, (b) for the mean value curve, $a = 0.114$, $b = -0.69$, $c = 0.98$, and (c) for the upper CI curve, $a = 0.195$, $b = 1.29$, $c = 0.55$. Undoubtedly,

Figure 9a indicates a perfect match of the fitted functions to the observed values, unveiling thus the underlying laws. Noteworthy, the 95% limiting CI is very narrow (0.021, 0.195) with the lower bound positive, while the mean value of γ converges to $\mu_\gamma \simeq 0.114$.

In order to identify the true underlying distribution of the GEV shape parameter (assuming it is well approximated by a normal distribution), apart from the limiting mean value estimated before, we need to estimate the limiting value of the standard deviation. Figure 9b depicts the estimated standard deviation values versus record length and a fitted curve of the same form used for the mean. The estimated parameters of the fitted curve are $a = 0.045$, $b = 1.27$ and $c = 0.70$, indicating thus that the true standard deviation of γ is $\sigma_\gamma \simeq 0.045$, a value significantly smaller than the smallest observed. Interestingly, assuming that the shape parameter follows the estimated normal distribution, i.e., $\gamma \sim \mathrm{N}(\mu_\gamma, \sigma_\gamma^2)$, the 95% CI of γ would be (0.03, 0.21) which is very close to the limiting CI estimated and depicted in Figure 9a. Furthermore the 99% CI (rounded at the second decimal digit) is estimated at (0, 0.23), and apparently the probability for a negative shape parameter to occur is only 0.005.

Additionally, Figure 9c depicts the percentage of records with negative γ over record length. Evidently, the observed points suggest a quickly non-linear decreasing "trend". The fitted curve has the same simple form as above but with $c < 0$. With estimated parameters $a = 221.3$, $b = -154.1$, $c = -0.067$ it crosses the horizontal axis at $L = (-a/b)^{-1/c} \approx 226$ years, implying that for record length greater than 226 years the percentage of records with negative γ would be zero. Indeed, none of the 16 records available with length greater than 140 years resulted in negative γ. This indicates a deviation from the fitted curve; yet, the number of stations for this record length is very small to take it into account but this is additional evidence that the Fréchet law prevails.

Finally, based on the previous findings, it is possible to create an “unbiased” or record-length-free estimator for the GEV shape parameter that incorporates its relation with the record length. Given that the true distribution of γ is the N(μ_γ, σ_γ^2) while for specific record length n is the N($\mu_\gamma(n)$, $\sigma_\gamma^2(n)$), with $\mu_\gamma(n) = \mu_\gamma - 0.69\, n^{-0.98}$ and $\sigma_\gamma(n) = \sigma_\gamma + 1.27\, n^{-0.70}$ being the functions fitted previously for the mean and the standard deviation, it can be easily proved that an “unbiased” estimator $\tilde{\gamma}(n)$ is the

$$\tilde{\gamma}(n) = \frac{\sigma_\gamma}{\sigma_\gamma(n)}\left(\hat{\gamma} - \mu_\gamma(n)\right) + \mu_\gamma \qquad (8)$$

where n is sample size (number of years), $\hat{\gamma}$ is the L-moments estimate of γ, whereas $\mu_\gamma \simeq 0.114$ and $\sigma_\gamma \simeq 0.045$ are the limiting mean and standard deviation values estimated previously.

5.3 Monte Carlo validation of the results

In order to validate our results regarding the underlying distribution of the GEV shape parameter we performed a Monte Carlo simulation. Specifically, we generated 15 137 random samples, with sizes precisely equal with the original records lengths, from a GEV distribution with the shape parameter being randomly generated from the anticipated normal distribution, i.e., the N(μ_γ, σ_γ^2), and with the location and scale parameter fixed to their mean values given in Table 2 as they do not affect the shape parameter estimates. In sequel, we estimated the shape parameter values of those samples and we formed the empirical distribution shown in Figure 10. We can see that while the prior distribution of γ was the N(μ_γ, σ_γ^2) the estimated posterior is almost identical with the empirical distribution emerged from the real records given in Figure 7. The comparison of the two distributions reveals a very close match, i.e., the empirical distribution emerged from the real records has mean and the standard deviation, respectively, equal to 0.092

and 0.12 while the corresponding values for the empirical distribution emerged from the synthetic records are, respectively, 0.104 and 0.11.

This minor deviation is probably justified by the fact that the L-skewness and the L-kurtosis of the empirical distribution of γ, which are −0.017 and 0.158, respectively, deviate slightly from the theoretical values of a normal distribution which are 0 and 0.123. The small negative skewness may have caused the slight decrease in the mean value while the higher L-kurtosis implies more extremes γ values, both negative and positive, and this obviously leads to higher variance. The fact is that both the empirical evidence and the Monte Carlo simulation suggest that the distribution of the GEV shape parameter is very well approximated by the normal distribution N(μ_γ, σ_γ^2). Even if the shape characteristics between the empirical and the Monte Carlo distributions do not match exactly (mainly the L-kurtosis) this is something anticipated; when a set of 15 137 real-world records is analyzed we should expect that some records may either contain incorrectly recorded values or some extraordinary events occurred, leading thus to unrealistically small or large shape parameter estimates. For example a couple or even one "extremely" extreme event in a relatively small sample, e.g., 40-60 years may alter significantly the value of L-skewness and consequently the estimate of the shape parameter γ resulting thus in a distribution that may not describe realistically the behavior of the rainfall in general. "Errors" of this kind are unavoidable as it is possible for a small sample to contain, e.g., the 1000-year event.

The previous analysis also indicated that the true mean value of the underlying distribution of the GEV shape parameter is $\mu_\gamma = 0.114$, markedly larger than zero, i.e. the value specifying the Gumbel distribution. This consequently leads us to assume that the Gumbel distribution is not a good model in general for annual maximum daily rainfall. Nevertheless, it does not reveal how bad or good the Gumbel model is if compared to the GEV model or more specifically to the Fréchet law. Obviously the GEV and the Gumbel distributions cannot be compared directly in the sense that the first one is a three-parameter model while the second one is a two-parameter model and a special case of the first one. For this reason we compare here the Gumbel distribution with a representative fixed-shape-parameter GEV distribution, i.e., a GEV with shape parameter equal to $\mu_\gamma = 0.114$.

Specifically, we generated 15 137 random samples, with sizes equal to those of the original records using: (a) a Gumbel distribution, and (b) a GEV distribution with $\gamma = 0.114$ (the location and scale parameters were fixed in both distributions as their values do not affect the shape characteristics). Next, we estimated the Monte Carlo (MC) L-kurtosis *vs*. L-skewness points and depicted them in comparison with the observed ones already presented in Figure 6. The idea is to compare the extent of the area formed by the MC points with the area formed by the points of the real records.

The results of this Monte Carlo simulation are presented in Figure 11. For the Gumbel case (upper graph) we note that indeed there is a spread around the theoretical Gumbel point, yet, the area covered by the MC points is significantly smaller than the one formed by the observed points and the cloud of points are placed toward the left. Clearly, the Gumbel distribution fails to generate points with high values of L-skewness. In the GEV case with fixed γ (lower graph) we observe not only the expected shift of the cloud of the MC points toward the right, but also the expansion of this cloud, so that the area formed is much larger compared to that of the Gumbel case. In addition, the MC area better fits the one formed by the empirical points. This reveals that the GEV distribution with fixed γ performs in general much better compared with the Gumbel distribution.

5.4 Geographical variation of the GEV shape parameter

The previous analysis reveals that the GEV shape parameter estimates depend on the record length and that essentially the parameter varies in the interval (0, 0.23). Thus, the question that naturally arises is how the parameter varies over geographical location, as it is reasonable to expect that different areas of the world exhibit different behavior not only in the mean annual rainfall but also the in the shape of distribution of the annual extremes. Yet, we should bear in mind that even if the behavior of extreme rainfall is the same in a big area, in practice the estimated GEV shape parameters in different locations within the area will differ due to sampling effects. As a consequence, the different estimates may lead to false conclusions.

Thus, in order to reduce the sampling effect and to investigate the geographical distribution of the GEV shape parameter seeking to reveal any kind of geographical pattern, we divided the earth's surface into cells and studied the mean value of the GEV shape parameter within the cell; obviously the mean value offers a simple and rational smoothing. Each cell is defined by a latitude difference of $\Delta\varphi = 2.5°$ and longitude difference of $\Delta\lambda = 5°$; as latitude φ ranges from $-90°$ to $90°$ and longitude λ from $-180°$ to $180°$, a total of 5 184 cells emerged. The mean value of the GEV shape parameter of each cell is simply estimated as the average of those shape parameter estimates that correspond to stations lying within the cell, given that the cell contains at least two records, Clearly, the number of stations within each cell is not constant, and most of the cells (notably those in the oceans) do not contain any stations while there are 258 cells containing only one record. Specifically, from the 5184 cells formed, only 792 cells had available records and only 534 had at least two records, while there are 46 cells with more than 100 records each. The results using the typical (record-length dependent) estimates of the GEV shape parameter are depicted in the world map given in Figure 12 where the cell's mean value is expressed by coloring the cell according to the map's legend. It is noted that the values defining the bins in the map's legend are defined by the minimum value, the Q_{10}, Q_{25}, Q_{50}, Q_{75}, and Q_{90} empirical quantile (or percentile) points and the maximum value of the 534 mean shape parameter values after rounding off to the second decimal, e.g., the central 50% of values or the interquartile range is approximately form 0.06 to 0.14. The numbers of cells with mean values at each successive bin (from low to high values) are: 57, 76, 146, 115, 89 and 51, while the number of cells with negative mean values is 52. Clearly, the map reveals that large and discrete areas exist with the same behavior in extreme rainfall manifested by the approximately equal GEV shape parameter values.

Nevertheless, the analysis of the previous section unveiled the clear relationship of the estimated GEV shape parameters with the record length. Consequently, a more accurate map should incorporate these findings as a region contains records of variable length leading thus to a record-length depended estimate of the mean value. Additionally, we showed that the GEV shape parameter estimates can be corrected by Eq. (8) to be record-length free and follow the normal distribution $N(\mu_\gamma, \sigma_\gamma^2)$ which constitutes a very good approximation of the true distribution of the GEV shape parameter. For these reasons, we reconstructed the map by using the unbiased (free of record-length dependence) estimate of the shape parameter values according to Eq. (8). The results are presented in Figure 13. As in the previous map, the bins are defined the same way but obviously the values differ as the range of variation is much smaller. The numbers of cells with values spotted in each successive bin are different from the previous map, i.e., 59, 88, 105, 143, 93 and 46 (due to rounding of the quantile values), while the number of points representing negative values is now zero. Comparing the two maps we note that they look almost the same but in fact they differ. Finally, it is notable that large areas or zones are formed by points representing shape parameter values belonging in a very narrow range. For example, in the US there are two large zones where the shape parameter ranges from 0.10 to 0.11 in the one (green color) and from 0.11 to 0.13 in the other (yellow-green color); additionally, in the entire Atlantic coasts of South America a zone of low values is formed while a large area of high values can be spotted in South-West Australia.

Obviously, the accuracy in the estimation of the shape parameter mean values is not the same for every cell as the number of records per cell is not constant. Thus, in order to provide a measure of uncertainty or a measure of estimation error, we constructed the map given in Figure 14 that presents each cell's standard error (SE) values with respect to the mean values given in the map Figure 13 (unbiased estimates). The SE is defined as $\mathrm{SE} = \sigma / \sqrt{n}$ and in this case σ is the sample standard deviation of the shape parameter values of the cell and n the number of those values. In order for the estimates of SE to be relatively accurate we selected only those cells that contain at least six records (a total of 281 cells), as it is well-known that the estimation of the standard deviation is markedly biased for very small samples. A cell's SE expresses the standard deviation of the cell's shape parameter mean value, and can be used directly to calculate the 95% CI of this estimate as it is well-known that the 95% CI is given by $\overline{\gamma} \pm 1.96\,\mathrm{SE}$, where $\overline{\gamma}$ is the cell's shape parameter mean value. The values defining the bins of SE in the map's legend (Figure 14) are defined by the minimum value, the Q_{25}, Q_{50}, Q_{75} empirical quantile (or percentile) points and the maximum value of the 281 SE values after rounding off to the third decimal, e.g., the 50% of SE values are less than 0.008. The numbers of cells with SE values at each successive bin (from lower to higher values) are: 67, 75, 68, and 71. As expected, areas with high density of stations and large records have very low values of SE.

6. Summary and conclusions

Extreme value distributions have been extensively used in hydrology for more than half a century as a basic tool for estimating the design rainfall of infrastructures or assessing flood risks; however, selecting the appropriate law is usually based on small samples without guaranteeing the correct choice or the accurate estimate of the law's parameters. Here, we analyze 15 137 rainfall records from all over the world aiming to assess which one of the three limiting distributions better describes the annual maximum daily rainfall. Initially, we formed a method comprising two simple criteria, in order to treat the very common problem of extracting annual maxima of daily rainfall from records containing missing values. The method was successfully validated and applied to form the annual maximum daily rainfall records.

The question, which of the three limiting extreme value distributions to use, is the focus of this study. Starting from the reversed Weibull distribution, we may note that it implies a parent distribution for daily rainfall with an upper bound; we contend that this is physically inconsistent and moreover, to our knowledge distributions bounded from above have never been used for daily rainfall in competent studies. With reference to the Fréchet *vs*. Gumbel "battle", we showed that, as strange it may seem, annual maxima extracted from a parent distribution that belongs to the domain of attraction of the Gumbel law, are better described by the Fréchet law. This occurs for two reasons: first, the convergence rate to the Gumbel law is extremely slow, and second, the shape parameter of the Fréchet law enables the distribution to approximate quite well not only distributions with power-type tails but also other heavy-tailed distributions.

The empirical investigation using 15 137 records started with an L-moments ratio plot which reveals that 80% of observed points are located on the right of the "Gumbel point" providing clear evidence that the Fréchet law prevails. Additionally, the analysis of the estimated

GEV shape parameters unveils a clear relationship between the shape parameter value over the record length, implying that only very large samples can reveal its true distribution or the true behavior of the extreme rainfall. The "asymptotic" analysis performed, based on the fitted functions to the mean and standard deviation of the GEV shape parameter over record length, suggests that the distribution of the GEV shape parameter that would emerge if extremely large samples were available is approximately normal with mean value 0.114 and standard deviation 0.045. The meaning of this finding is that the GEV shape parameter is expected to belong in a narrow range, approximately from 0 to 0.23 with confidence 99%. Essentially, the analysis shows that we cannot trust blindly the data, as small samples may distort the true picture. In this direction, we propose the use of Eq. (8) that corrects the L-moments estimate of the GEV shape parameter removing the bias due to limited sample size.

While originally a small percentage of records have negative shape parameter (reversed Weibull law), the analysis reveals that this percentage rapidly decreases over sample size, while the fitted function indicates that for record length greater than 226 years this percentage would be zero. Interestingly, none of the 16 records available with length greater than 140 years resulted in negative γ. Moreover, the probability for a negative shape parameter to occur, according to the distribution fitted, is only 0.005, and combined with the previous findings suggests that a GEV distribution with negative shape parameter (bounded from above) is completely inappropriate for rainfall. Concerning the geographical distribution of the GEV shape parameter, the constructed maps show that large areas of the world share approximately the same GEV shape parameter, yet different areas of the world exhibit different behavior in extremes.

We believe the "verdict" is clear: the Fréchet law, or else the GEV law with positive shape parameter, should prevail over the Gumbel law and a fortiori over the reversed Weibull law, with latter suggesting a dangerous choice. If we had to form a rule of thumb, we would propose that in the case where data suggest a GEV distribution with negative shape parameter, this should not be used. Instead it is more reasonable to use a Gumbel or, for additional safety, a GEV distribution with a shape parameter value equal to 0.114. The prevailing practice of the past that favored the use of the Gumbel distribution does not suggest a proof of its outperformance over the Fréchet law, as it seems it takes a long time to reveal Nature's "secrets" and its true behavior. As Heraclitus of Ephesus stated more than 2500 years ago in the aphorism given in the introduction (loosely translated) "Nature loves to hide".

Acknowledgment We wish to thank the Editor Praveen Kumar and the Associate Editor Alberto Montanari for the positive evaluation of the paper, as well as the eponymous reviewer Salvatore Grimaldi and two anonymous reviewers for their detailed comments which helped us to improve the presentation of the study.

References

Balkema, A. A., and L. D. Haan (1972), On R. Von Mises' Condition for the Domain of Attraction of exp(-e-x)1, *The Annals of Mathematical Statistics*, *43*(4), 1352–1354.

Barndorff-Nielsen, O. (1963), On the Limit Behaviour of Extreme Order Statistics, *The Annals of Mathematical Statistics*, *34*(3), 992–1002.

Berman, S. M. (1964), Limit Theorems for the Maximum Term in Stationary Sequences, *The Annals of Mathematical Statistics*, *35*(2), 502–516.

von Bortkiewicz, L. (1922), Variationsbreite und mittlerer Fehler, *Sitzungsber. Berli. Math. Ges.*, (21), 3–11.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997), *Modelling extremal events for insurance and finance*, Springer Verlag, Berlin Heidelberg.

Fisher, R. A., and L. H. C. Tippett (1928), Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample, *Mathematical Proceedings of the Cambridge Philosophical Society*, *24*(02), 180–190, doi:10.1017/S0305004100015681.

Fréchet, M. (1927), Sur la loi de probabilité de l'écart maximum, *Ann. soc. polon. math*, *6*, 93.

Galambos, J. (1972), On the Distribution of the Maximum of Random Variables, *The Annals of Mathematical Statistics*, *43*(2), 516–521.

Gnedenko, B. (1943), Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire, *The Annals of Mathematics*, *44*(3), 423–453, doi:10.2307/1968974.

Gumbel, E. J. (1958), *Statistics of Extremes*, Columbia University Press.

De Haan, L. (1971), A form of regular variation and its application to the domain of attraction of the double exponential, *Z. Wahrsch. Geb.*, (17), 241–258.

Hosking, J. R. M. (1984), Testing Whether the Shape Parameter is Zero in the Generalized Extreme- Value Distribution, *Biometrika*, *71*(2), 367–374, doi:10.2307/2336254.

Hosking, J. R. M. (1990), L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics, *Journal of the Royal Statistical Society. Series B (Methodological)*, *52*(1), 105–124.

Hosking, J. R. M. (1992), Moments or L Moments? An Example Comparing Two Measures of Distributional Shape, *The American Statistician*, *46*(3), 186–189, doi:10.2307/2685210.

Hosking, J. R. M., and J. R. Wallis (1993), Some statistics useful in regional frequency analysis, *Water Resour. Res.*, *29*(2), PP. 271–281, doi:199310.1029/92WR01980.

Hosking, J. R. M., J. R. Wallis, and E. F. Wood (1985), Estimation of the Generalized Extreme-Value Distribution by the Method of Probability-Weighted Moments, *Technometrics*, *27*(3), 251–261, doi:10.2307/1269706.

Jenkinson, A. F. (1955), The frequency distribution of the annual maximum (or minimum) values of meteorological elements, *Quarterly Journal of the Royal Meteorological Society*, *81*(348), 158–171, doi:10.1002/qj.49708134804.

Juncosa, M. L. (1949), The asymptotic behavior of the minimum in a sequence of random variables, *Duke Mathematical Journal*, *16*(4), 609–618, doi:10.1215/S0012-7094-49-01658-0.

Kotz, S., and S. Nadarajah (2000), *Extreme value distributions: theory and applications*, Imperial College Press.

Koutsoyiannis, D. (2004a), Statistics of extremes and estimation of extreme rainfall, 1, Theoretical investigation, *Hydrological Sciences Journal*, *49*(4), 575–590.

Koutsoyiannis, D. (2004b), Statistics of extremes and estimation of extreme rainfall, 2, Empirical investigation of long rainfall records, *Hydrological Sciences Journal*, *49*(4), 591–610.

Leadbetter, M. R. (1974), On extreme values in stationary sequences, *Probability theory and related fields*, *28*(4), 289–303.

von Mises, R. (1936), La distribution de la plus grande de n valeurs, *Rev. math. Union interbalcanique*, *1*(1).

Papalexiou, S. M., and D. Koutsoyiannis (2012), Entropy based derivation of probability distributions: A case study to daily rainfall, *Advances in Water Resources*, *45*, 51–57, doi:10.1016/j.advwatres.2011.11.007.

Papalexiou, S. M., D. Koutsoyiannis, and C. Makropoulos (2012), How extreme is extreme? An assessment of daily rainfall distribution tails, *Hydrology and Earth System Sciences Discussions*, *9*(5), 5757–5778, doi:10.5194/hessd-9-5757-2012.

Peel, M. C., Q. Wang, R. M. Vogel, and T. A. McMahon (2001), The utility of L-moment ratio diagrams for selecting a regional probability distribution, *Hydrological sciences journal*, *46*(1), 147–156.

Pickands III, J. (1975), Statistical Inference Using Extreme Order Statistics, *The Annals of Statistics*, *3*(1), 119–131.

Reiss, R. D., and M. Thomas (2007), *Statistical analysis of extreme values: with applications to insurance, finance, hydrology and other fields*, Birkhauser.

Smirnov, N. V. (1949), Limit distributions for the terms of a variational series, *Trudy Matematicheskogo Instituta im. VA Steklova*, *25*, 3–60.

Vogel, R. M., and N. M. Fennessey (1993), L moment diagrams should replace product moment diagrams, *Water Resources Research*, *29*(6), 1745–1752.

Watson, G. S. (1954), Extreme Values in Samples from m-Dependent Stationary Stochastic Processes, *The Annals of Mathematical Statistics*, *25*(4), 798–800.

Tables

Table 1. Basic summary statistics of the 15 137 records; Q indicates the empirical quantile.

	Record Length	Median	Mean	SD	Skew	L-scale λ_2	L-skew τ_3	L-kurtosis τ_4
min	40	7.40	9.10	3.94	-0.71	2.15	-0.16	-0.06
Q_5	49	25.60	28.51	11.00	0.53	5.80	0.10	0.09
Q_{25}	58	39.20	43.13	17.41	0.98	9.06	0.18	0.14
Q_{50} (Median)	68	57.20	62.24	23.73	1.35	12.35	0.23	0.18
Q_{75}	91	77.50	83.96	33.84	1.84	17.43	0.28	0.22
Q_{95}	117	114.80	126.23	57.81	3.03	29.86	0.37	0.30
max	163	864.50	863.69	430.69	9.87	244.66	0.76	0.73

Mean	74.85	61.97	67.73	27.72	1.51	14.40	0.23	0.18
SD	21.84	30.71	33.16	15.38	0.85	7.98	0.08	0.06
Skew	0.80	2.68	2.37	2.72	2.06	3.16	0.15	0.85
L-scale λ_2	12.07	15.97	17.35	7.80	0.43	4.01	0.04	0.03
L-skew τ_3	0.22	0.19	0.20	0.27	0.23	0.28	0.02	0.10

Table 2. Summary statistics of the estimated parameter of the fitted Gumbel and GEV distributions to the 15 137 annual maximum daily rainfall records; the fitting was done by the method of L-moments.

	Gumbel parameters		GEV parameters		
	α	β	α	β	γ
min	6.81	3.10	6.00	2.66	-0.587
Q_5	23.21	8.37	22.59	7.36	-0.107
Q_{25}	35.26	13.07	34.67	11.71	0.020
Q_{50} (Median)	51.54	17.82	50.82	16.16	0.093
Q_{75}	70.07	25.15	69.24	22.69	0.169
Q_{95}	102.54	43.09	101.14	38.53	0.283
max	659.96	352.97	688.17	401.68	0.760
Mean	55.74	20.77	54.95	18.71	0.092
SD	27.21	11.51	27.08	10.68	0.120
Skew	2.23	3.16	2.38	4.67	-0.130
L-scale λ_2	14.30	5.78	14.17	5.25	0.067
L-skew τ_3	0.18	0.28	0.18	0.27	-0.017
L-kurt τ_4	0.13	0.18	0.14	0.18	0.158

Table 3. Summary statistics of the estimated GEV shape parameter for various record length categories.

Record length (years)	40 - 50	51 - 60	61 - 70	71 - 80	81 - 90	91 - 100	101 - 110	110 - 120	≥ 121
Records No.	1161	3610	3972	1467	1134	1164	1132	1017	480
Records % ($\gamma > 0$)	71.8	72.9	77.8	83.6	85.0	86.8	88.1	91.1	91.0
Records % ($\gamma \leq 0$)	28.2	27.1	22.2	16.4	15.0	13.2	11.9	8.9	9.0
	GEV shape parameter γ								
min	-0.461	-0.587	-0.493	-0.307	-0.287	-0.283	-0.188	-0.193	-0.204
Q_5	-0.152	-0.156	-0.112	-0.086	-0.068	-0.048	-0.046	-0.035	-0.029
Q_{25}	-0.014	-0.009	0.011	0.030	0.036	0.042	0.049	0.047	0.060
Q_{50} (Median)	0.079	0.082	0.086	0.102	0.100	0.106	0.108	0.102	0.118
Q_{75}	0.172	0.166	0.166	0.176	0.169	0.175	0.169	0.158	0.170
Q_{95}	0.312	0.290	0.291	0.285	0.268	0.271	0.271	0.247	0.263
max	0.541	0.706	0.760	0.567	0.539	0.573	0.750	0.471	0.345
Mean	0.077	0.077	0.089	0.103	0.101	0.108	0.110	0.105	0.116
SD	0.141	0.138	0.124	0.112	0.102	0.100	0.096	0.088	0.088

Skew	-0.135	-0.253	0.120	0.096	-0.029	0.171	0.367	0.220	-0.137
L-scale λ_2	0.079	0.077	0.069	0.063	0.057	0.056	0.053	0.048	0.049
L-skew τ_3	-0.012	-0.034	0.015	0.006	0.002	0.014	0.023	0.024	-0.011
L-kurt τ_4	0.142	0.149	0.153	0.134	0.135	0.137	0.144	0.166	0.156

Figures

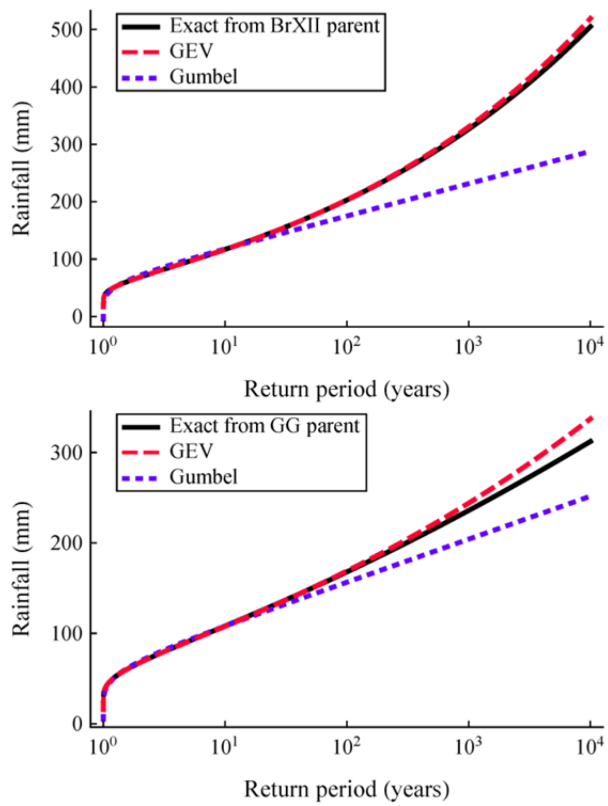

Figure 1. Demonstration of the convergence of the true distribution of maxima to the limiting laws.

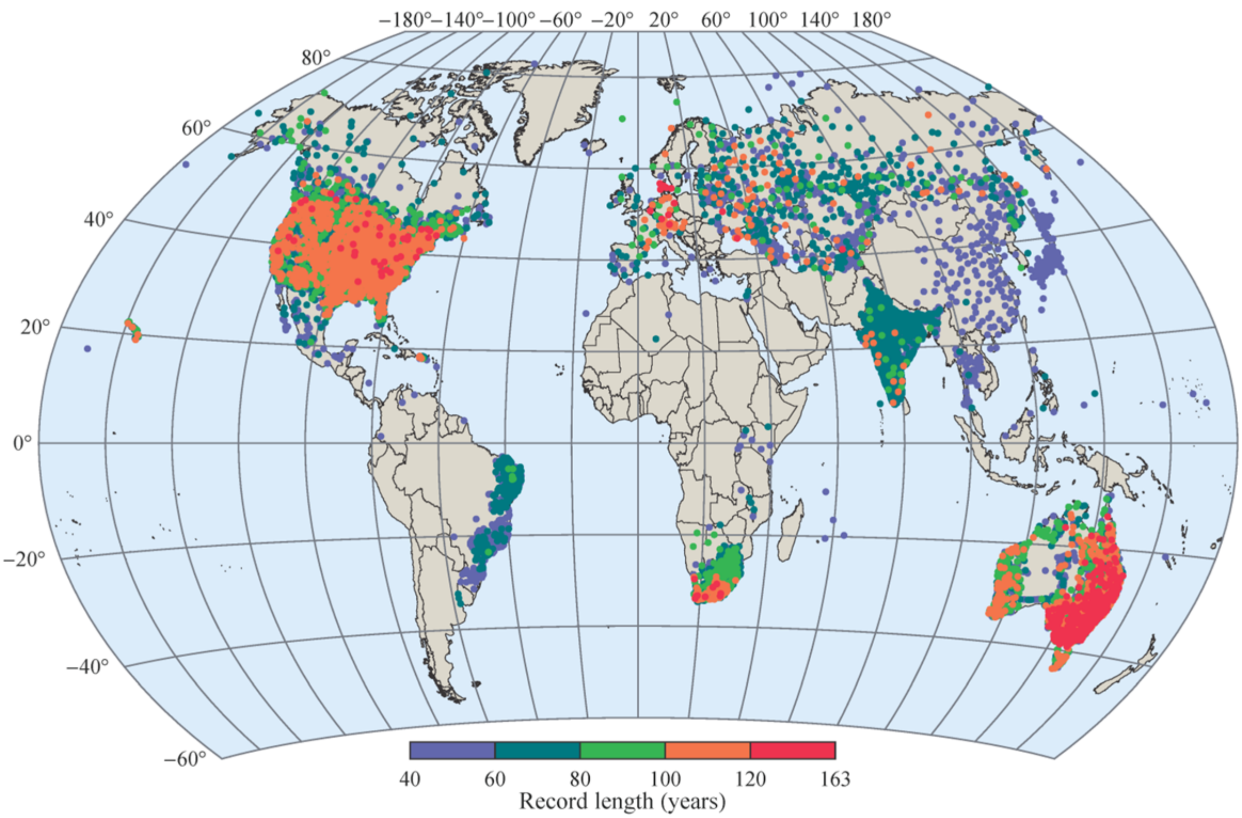

Figure 2. Locations of the 15 137 stations with annual maximum records of daily rainfall analyzed with number of values ranging from 40 to 163 years. Note that there are overlaps with points corresponding to high record lengths shadowing (being plotted in front of) points of lower record lengths.

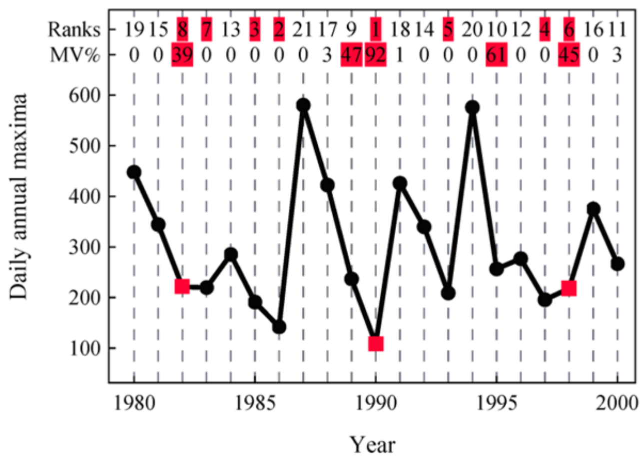

Figure 3. Explanatory plot of the maxima extraction method. The annual maximum daily rainfall is considered unknown (red rectancles) if its rank is in the smaller 40% of ranks (red shaded ranks) and the missing-value percentage (MV%) of the year it belongs is larger than 1/3 (red shaded percentages).

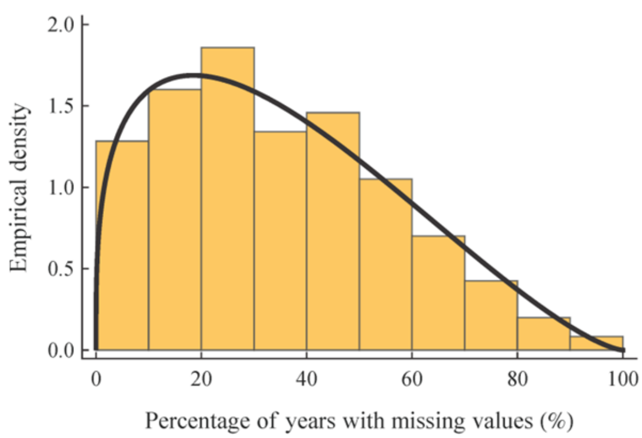

Figure 4. Empirical distribution of the year's percentage per record having missing values as resulted from the analysis of the 15 137 records; the solid line depicts a fitted Beta distribution.

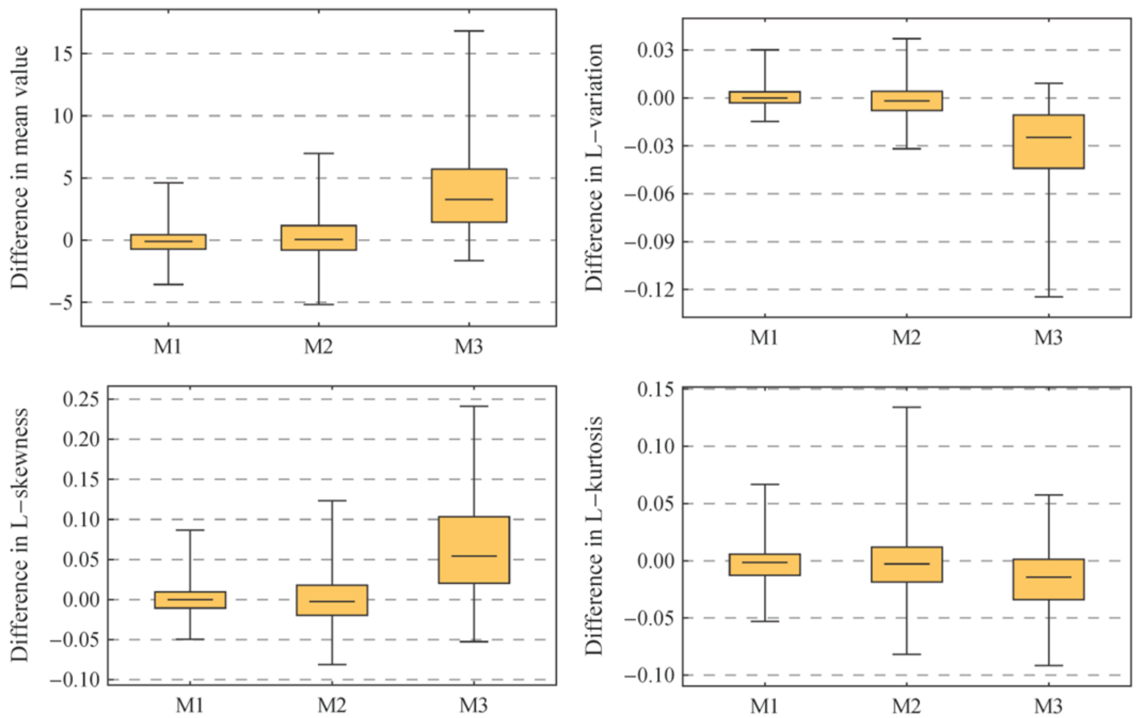

Figure 5. Box plots depicting the resulting sample differences of various statistics between the real annual maxima series and the ones created from the incomplete daily series. The advantage of the first method compared to the others is clearly seen by the smaller range of the box plots. The lower and upper fences of the box plots represent the sample quantiles Q_1 and Q_{99}, respectively.

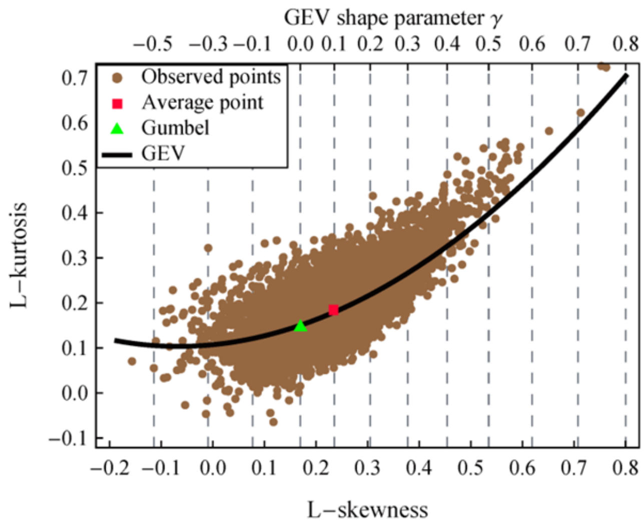

Figure 6. Observed L-kurtosis *vs*. L-skewness points of the 15 137 annual maximum daily rainfall records and the theoretical point and line of the Gumbel and GEV distribution, respectively.

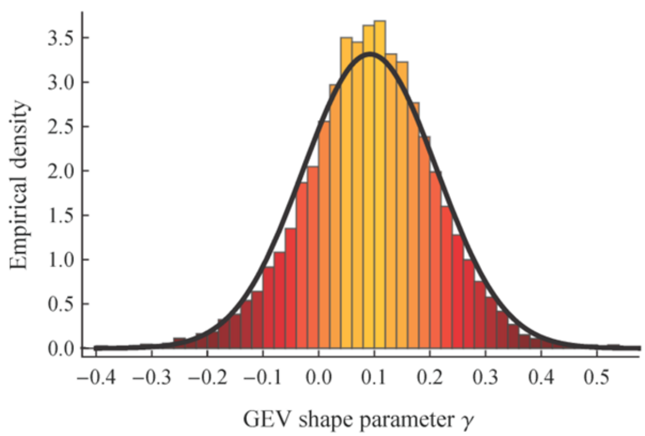

Figure 7. Empirical distribution of the GEV shape parameter as resulted by fitting the GEV distribution to the 15 137 annual maximum daily rainfall records. The solid line depicts a fitted normal distribution.

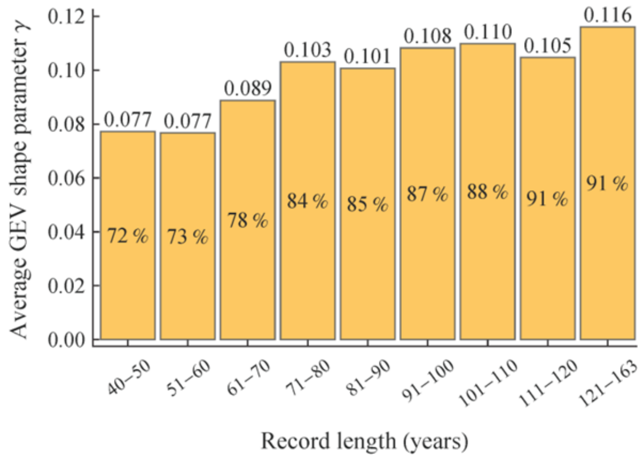

Figure 8. Mean value of the GEV shape parameter for various categories of record length. The numbers in the boxes indicates the percentage of records with positive shape parameter value.

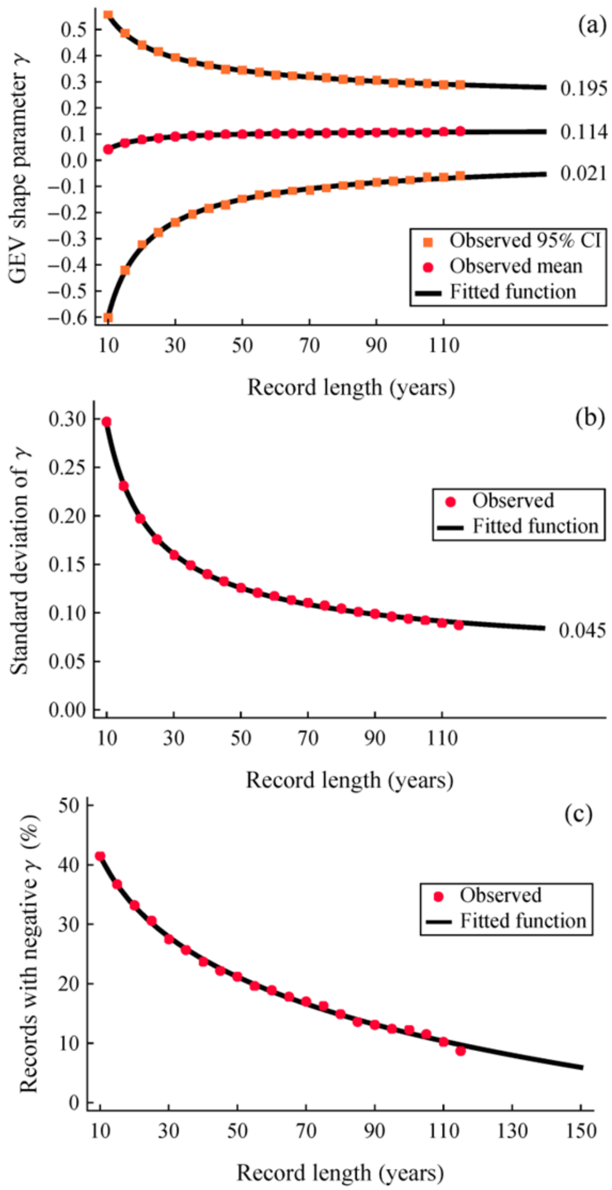

Figure 9. (a) Mean, quantiles Q_5 and Q_{95} as estimated for various records lengths and their fitted asymptotic values; (b) standard deviation; (c) percentage of records with negative shape parameter.

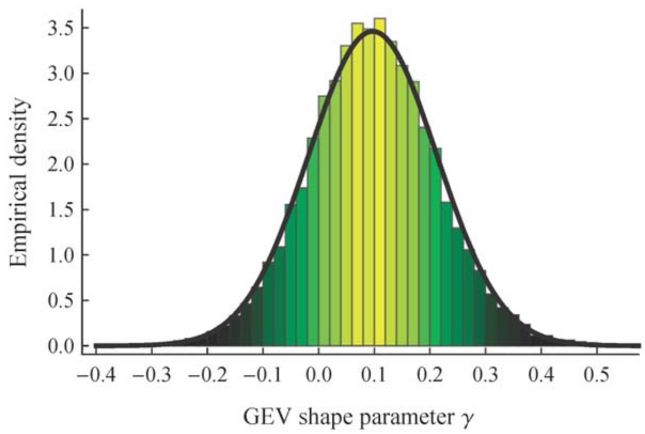

Figure 10. Empirical distribution of the GEV shape parameter as resulted from the Monte Carlo simulation where 15 137 synthetic records generated with the shape parameter being randomly sampled from the N(μ_γ, σ_γ^2). The solid line depicts the fitted normal distribution.

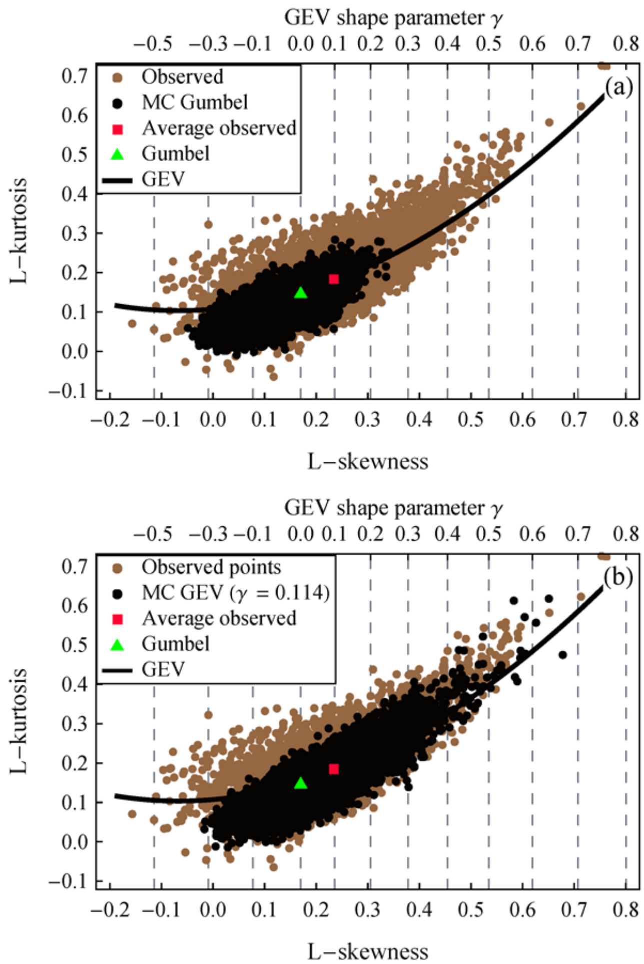

Figure 11. Monte Carlo points estimated (a) for the Gumbel distribution, and (b) for the GEV distribution with fixed shape parameter $\gamma = 0.114$, depicted in comparison to the observed ones.

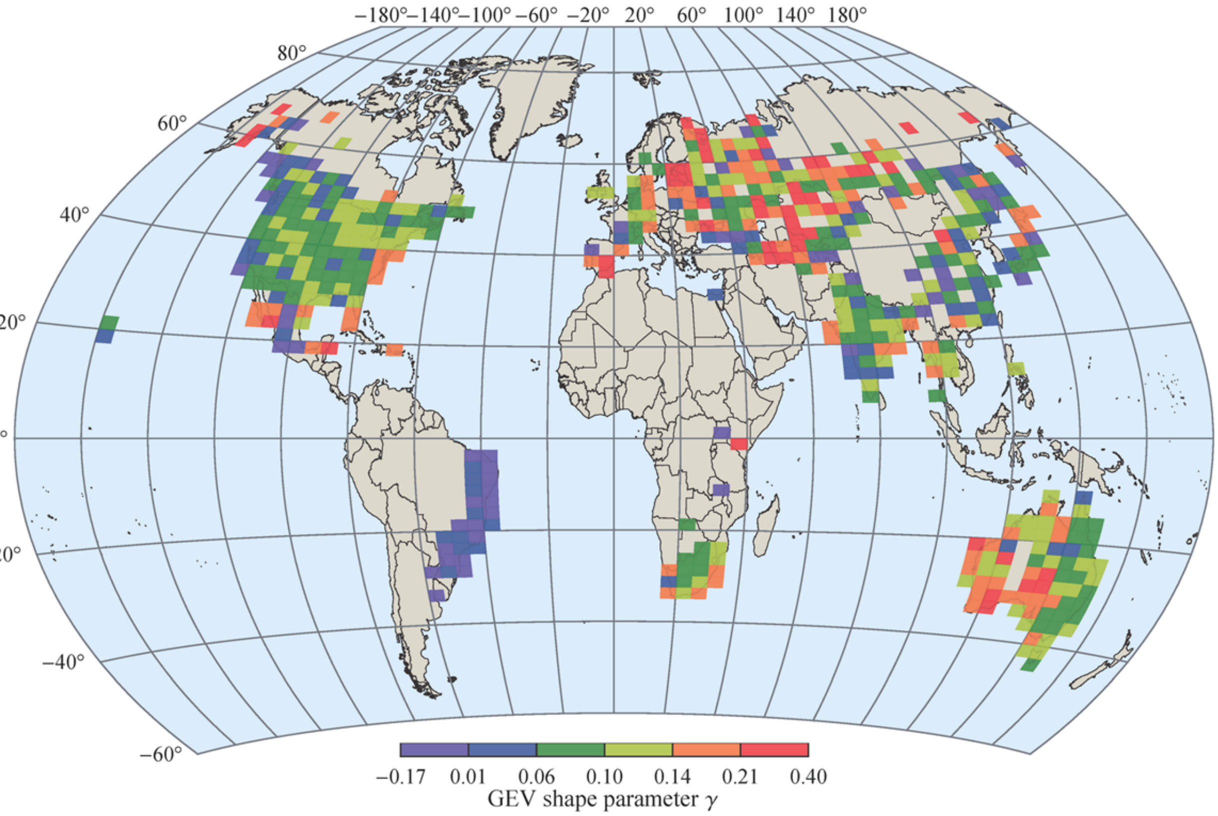

Figure 12. Geographical distribution of the mean value of the GEV shape parameter (estimated by the standard L-moment estimator) in regions of latitude difference $\Delta\varphi = 2.5°$ and longitude difference $\Delta\lambda = 5°$.

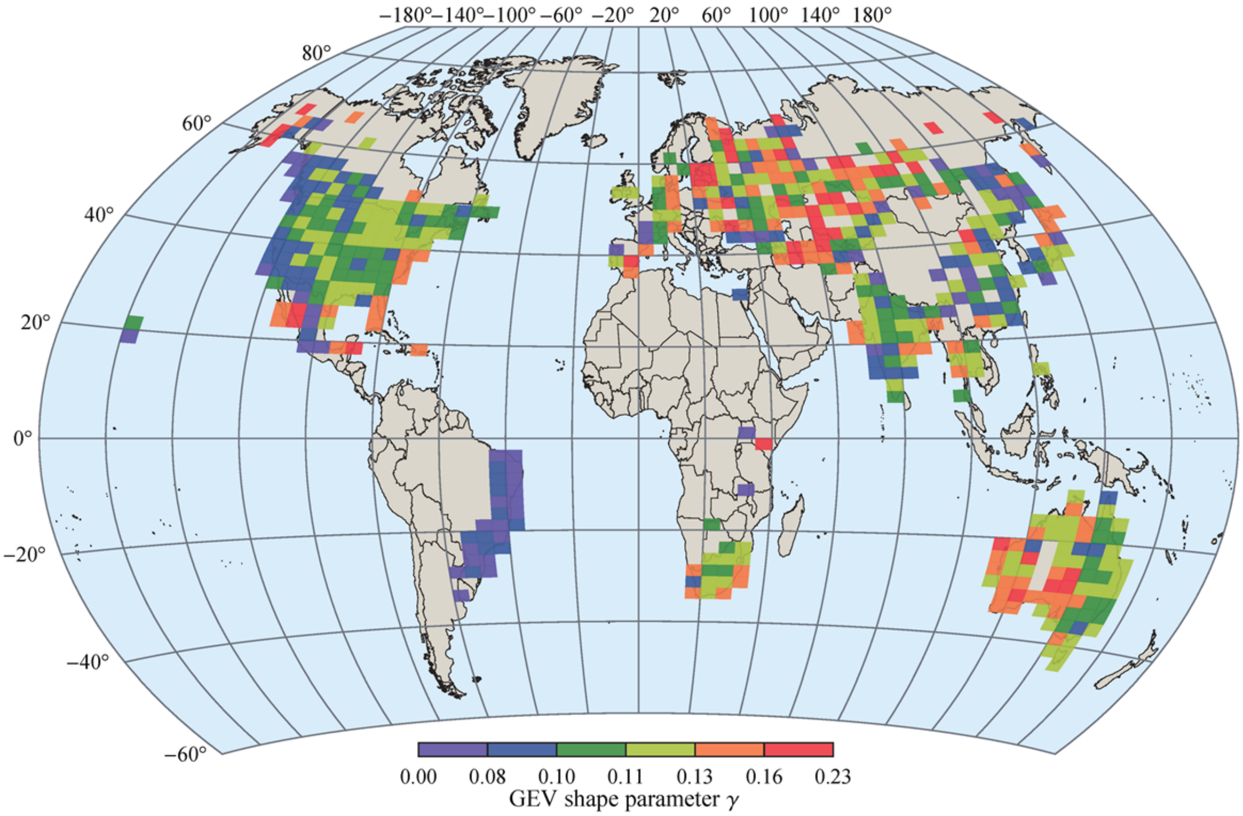

Figure 13. Geographical distribution of the mean value of the GEV shape parameters estimated by the unbiased estimator of Eq. (8) that corrects the sample-size effect; notice the difference in the values of the legend with the legend of Figure 12.

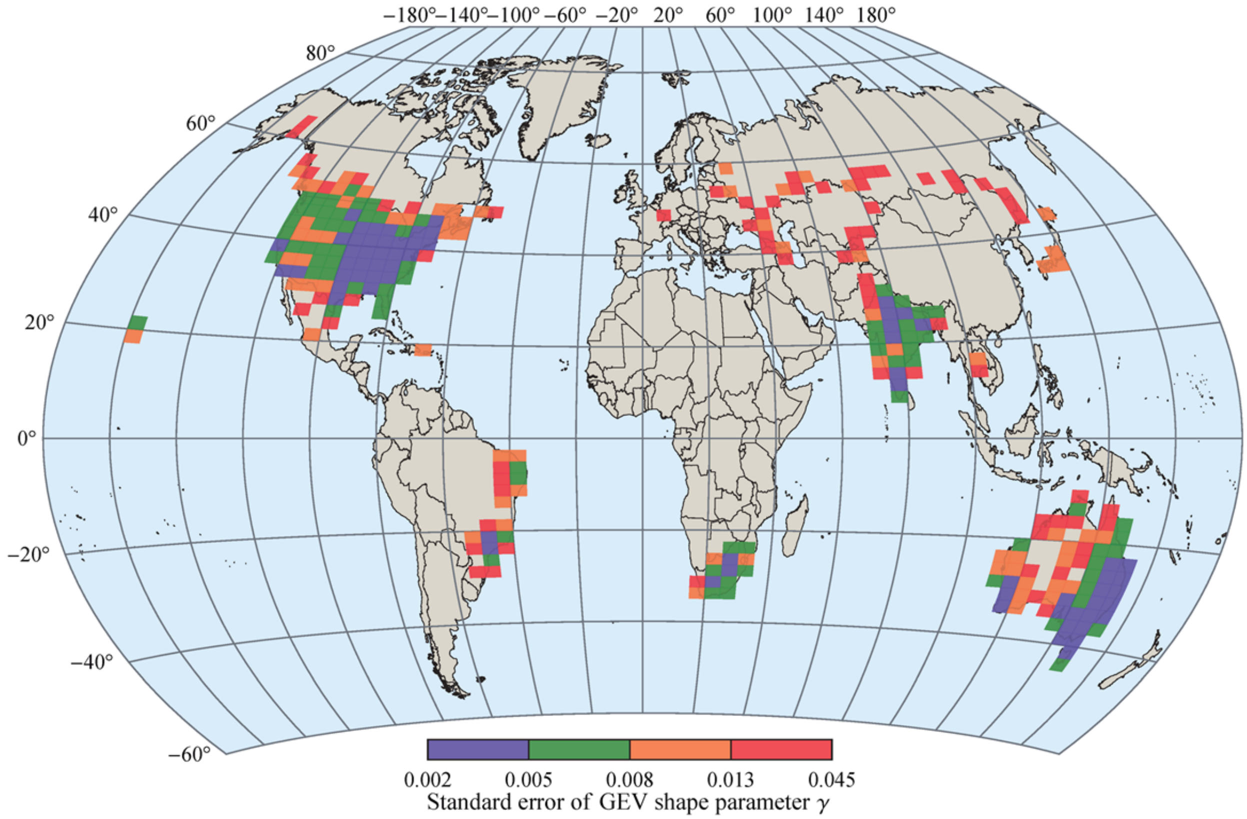

Figure 14. Standard error values of the GEV shape parameter mean values that are given in the map of Figure 13.