## Statistics of extremes and estimation of extreme rainfall

## 1. Theoretical investigation

Demetris Koutsoyiannis

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

**Abstract** The Gumbel distribution has been the prevailing model for quantifying risk associated with extreme rainfall. Several arguments including theoretical reasoning and empirical evidence are supposed to support the appropriateness of the Gumbel distribution. These arguments are examined thoroughly in this work and are put into question. Specifically, theoretical analyses show that the Gumbel distribution is quite unlikely to apply to hydrological extremes and its application may misjudge the risk as it underestimates seriously the largest extreme rainfall amounts. Besides, it is shown that hydrological records of typical length (some decades) may display a distorted picture of the actual distribution suggesting that the Gumbel distribution is an appropriate model for rainfall extremes while it is not. In addition, it is shown that the extreme value distribution of type II (EV2) is a more consistent alternative. Based on the theoretical analysis, in the second part of this study an extensive empirical investigation is performed using a collection of 169 of the longest available rainfall records worldwide, each having 100-154 years of data. This verifies the inappropriateness of the Gumbel distribution and the appropriateness of EV2 distribution for rainfall extremes.

**Keywords** design rainfall; extreme rainfall; generalized extreme value distribution; Gumbel distribution; hydrological design; hydrological extremes; probable maximum precipitation; risk.

## Statistique de valeurs extrêmes et estimation de précipitations extrêmes

## 1. Recherche théorique

**Résumé** La distribution de Gumbel a longtemps été le modèle régnant pour la quantification du risque associé aux précipitations extrêmes. Plusieurs arguments comprenant à la fois un raisonnement théorique et des faits empiriques sont censés soutenir la convenance de la distribution de Gumbel. Ces arguments sont examinés exhaustivement dans ce travail et sont mis en question. Spécifiquement, des analyses théoriques montrent que la distribution de Gumbel est peu susceptible de s'appliquer aux valeurs extrêmes de variables hydrologiques et son application peut conduire à une sous-estimation sérieuse du risque des plus grandes valeurs dans la série des valeurs extrêmes de précipitations. En outre, il est montré que les séries hydrologiques de longueur typique (quelques décennies) peuvent montrer une image déformée de la distribution réelle ce qui suggère que la distribution de Gumbel est réellement le modèle approprié pour les précipitations extrêmes alors qu'elle ne l'est pas. En outre, on montre que la distribution de valeurs extrêmes du type II (EV2) est une alternative plus cohérente. Dans la deuxième partie de cette étude, en et se basant sur l'analyse théorique, une recherche empirique étendue est effectuée; celle-ci utilise une collection de 169 parmi les séries de précipitations les plus longues qui sont disponibles dans le monde entier, chacun comportant 100-154 ans de données. Ceci vérifie l'inapproprieté de la distribution de Gumbel et la convenance de la distribution EV2 pour les précipitations extrêmes.

**Mots-clés** Précipitations de projet; précipitations extrêmes; distribution généralisée de valeurs extrêmes; Distribution de Gumbel; conception hydrologique; valeurs extrêmes en hydrologie; précipitation maximum probable; risque.

## 1. Introduction

Almost a century after the empirical foundation of hydrological frequency curves known as “duration curves” (Hazen, 1914) and the theoretical foundation of probabilities of extreme values (von Bortkiewicz, 1922; von Mises, 1923), and half a century after the convergence of empirical and theoretical approaches (Gumbel, 1958) the estimation of hydrological extremes continues to be highly uncertain. This has been vividly expressed by Klemeš (2000), who argues that

> “... the increased mathematization of hydrological frequency analysis over the past 50 years has not increased the validity of the estimates of frequencies of high extremes and

> thus has not improved our ability to assess the safety of structures whose design characteristics are based on them. The distribution models used now, though disguised in rigorous mathematical garb, are no more, and quite likely less, valid for estimating the probabilities of rare events than were the extensions 'by eye' of duration curves employed 50 years ago."

Twenty years earlier, a similar critique was done be Willeke (1980; see also Dooge, 1986), who, among several common myths in hydrology, included the "Myth of the Tails", which reads

> "Statistical distributions applied to hydrometeorological events that fit through the range of observed data are applicable in the tails",

and emphasizes the fact that the tails of distributions fitted to real data are highly uncertain.

Obviously, however, the probabilistic approach to extreme values of hydrological processes signifies a major progress in hydrological science and engineering as it quantifies risk and disputes arbitrary and rather irrational concepts and approaches like the probable maximum precipitation (PMP) and flood (PMF). The latter that essentially assume an upper limit to precipitation and river flow have never been supported by concrete reasoning and data and have been criticized on both ethical grounds, for the promise that design values are risk free, and technical grounds for logical inconsistency and methodological gaps (Benson, 1973). For example, the hydrometeorological approaches to PMP are based on storm maximization assuming a high dew point, equal to the maximum observed values during a period of at least 50 years (World Meteorological Organization, 1986, p. 11). Obviously, the maximum of a 50-year period does not represent a physical limit and had this period been 100 or 200 years the observed dew point would be higher. The statistical approach to PMP, based on the studies of Hershield (1961, 1965) has been revisited recently (Koutsoyiannis, 1999) and it was concluded that the data used by Hershfield do not suggest the existence of an upper limit. It must be recognized however, that the PMP/PMF concepts are still in wide use and are regarded by many as concepts more physically based than the probabilistic approach.

If one is exempted from the concept of an upper limit to a hydrological quantity and adopts a probabilistic approach, one will accept that the quantity may grow without any upper limit but the probability of exceedence decreases as the quantity grows. In this case, as probability of exceedance tends to zero, there exists a lower limit to the rate of growth which is proven mathematically. This lower limit is represented by the Gumbel distribution, which has the thinnest possible tail. So, abandoning the PMP concept and adopting the Gumbel distribution can be thought of as a step from a finite upper limit to infinity, but with the slowest possible growth rate towards infinity. Does nature follow the slowest path to infinity? This question is not a philosophical one but has strong engineering implications. If the answer is positive, the design values for flood protection structures or measures will be the smallest possible ones (among those obtained by the probabilistic approach), otherwise they will be higher.

The fact that the Gumbel distribution has been the most common probabilistic model used in modelling hydrological extremes, especially rainfall extremes, may be interpreted as a positive answer to the above question. It is well known that the estimation of rainfall extremes is very important for major hydraulic structures, given that design floods are generally estimated from appropriately synthesized design storms (e.g. U.S. Department of the Interior, Bureau of Reclamation, 1977, 1987; Sutcliffe, 1978). Recently, several studies have shown that floods seem to have heavier tails than a Gumbel distribution (Farquharson et al., 1992; Turcotte, 1994; Turcotte and Malamud, 2003). Other studies (Wilks, 1993; Koutsoyiannis and Baloutsos, 2000; Coles et al. 2003; Coles and Pericchi, 2003; Sisson et al., unpublished) have extended the scepticism for the Gumbel distribution to the case of rainfall extremes, showing that it underestimates seriously the largest extreme rainfall amounts.

This scepticism on the Gumbel distribution for hydrological extremes, with emphasis on rainfall extremes, is the central theme of this study. After a brief review of basic concepts of extreme value distributions, theoretical arguments are provided that show that the Gumbel distribution is quite unlikely to apply to hydrological extremes. Besides, it is shown that hydrological records of inadequate length may display a distorted picture of the actual distribution suggesting that the Gumbel distribution is an appropriate model for rainfall extremes while it is not. Apparently, as record length grows, the picture drawn by hydrological records becomes clearer. Therefore, a collection of 169 of the longest available rainfall records worldwide, each having 100-154 years of data was formed. Based on the theoretical analysis, in the second part of this study an extensive empirical investigation is performed. This verifies the fact that the Gumbel distribution is inappropriate for rainfall extremes and suggests that the three-parameter extreme value distribution of type II is a choice closer to reality and easy to use even with short rainfall records.

## 2. Basic concepts of extreme value distributions

It is recalled from probability theory that the largest of a number $n$ of independent identically distributed random variables, i.e.,

$$X := \max \{Y_1, Y_2, \ldots, Y_n\} \tag{1}$$

has probability distribution function

$$H_n(x) = [F(x)]^n \tag{2}$$

where $F(x) := P\{Y_i \le x\}$ is the common probability distribution function (referred to as parent distribution) of each $Y_i$. If $n$ is not constant but rather can be regarded as a realization of a Poisson distributed random variable with mean $v$, then the distribution of $X$ becomes (e.g. Todorovic and Zelenhasic, 1970; Rossi et al., 1984),

$$H'_v(x) = \exp\{-v[1 - F(x)]\} \tag{3}$$

Since $\ln [F(x)]^n = n \ln \{1 - [1 - F(x)]\} = n \{-[1 - F(x)] - [1 - F(x)]^2 - \ldots\} \approx -n [1 - F(x)]$, it turns out that for large $n$ or large $F(x)$, $H_n(x) \approx H'_n(x)$. Numerical investigation shows that even for relatively small $n$, the difference between $H_n(x)$ and $H'_n(x)$ is not significant (e.g., for $n$ = 10, the relative error in estimating the exceedence probability $1 - H_n(x)$ from (3) rather than from (2) is about 3% at most; for $F(x) = 0.95$, even for $n = 1$, the error does not exceed 2.5%).

In hydrological applications concerning the distribution of annual maximum rainfall or flood, it may be assumed that the number of values of $Y_i$ (e.g., the number of storms or floods per year), whose maximum is the variable of interest $X$ (e.g. the maximum rainfall depth or flood discharge), is not constant. The Poisson model can be regarded as appropriate for such applications. Given also the small difference between (3) and (2), it can be concluded that (3) should be regarded as an appropriate model for every practical hydrological application.

However, the exact distributions (2) or (3), whose evaluation requires the parent distribution to be known, have not been used in hydrological statistics. Instead, hydrological applications have made wide use of asymptotes or limiting extreme value distributions, which are obtained from the exact distributions when $n$ tends to infinity. Gumbel (1958), following the pioneering works by Fréchet (1927), Fisher and Tippet (1928) and Gnedenco (1941) developed a comprehensive theory of extreme value distributions. According to this, as $n$ tends to infinity $H_n(x)$ converges to one of three possible asymptotes, depending on the mathematical form of $F(x)$ (Gumbel, 1958, p. 157). The same limiting distributions may also result from $H'_v(x)$ as $v$ tends to infinity. All three asymptotes can be described by a single mathematical expression introduced by Jenkinson (1955, 1969) and become known as the Generalized Extreme Value (GEV) distribution. This expression is

$$H(x) = \exp\left\{-\left[1 + \kappa\left(\frac{x}{\lambda} - \psi\right)\right]^{-1/\kappa}\right\}, \qquad \kappa\, x \geq \kappa\, \lambda\, (\psi - 1/\kappa) \tag{4}$$

where $\psi$, $\lambda > 0$ and $\kappa$ are location, scale and shape parameters, respectively. Leadbetter (1974) showed that this holds not only for maxima of independent random variables but for dependent random variables, as well, provided that there is no long-range dependence of high-level exceedences. It is noted that the sign convention of $\kappa$ in (4) is opposite to that most commonly used in hydrological texts and the location parameter is dimensionless whereas in most texts a dimensional parameter $\xi = \lambda\, \psi$ is used.

When $\kappa > 0$, $H(x)$ represents the extreme value distribution of maxima of type II (EV2). In this case the variable is bounded from below and unbounded from above ($\lambda\, \psi - \lambda/\kappa \leq x < +\infty$). A special case is obtained when the lower bound becomes zero ($\psi = 1/\kappa$). This special two-parameter distribution has the simplified form

$$H(x) = \exp\left\{-\left(\frac{\lambda}{\kappa\, x}\right)^{1/\kappa}\right\}, \qquad x \geq 0 \tag{5}$$

In some texts, (5) is referred to as the EV2 distribution. Here, as in Gumbel (1958), the name EV2 distribution is used for the complete three-parameter form (4) with $\kappa > 0$. Distribution (5) is referred to as the Fréchet distribution.

The limiting case $\kappa = 0$ represents the type I distribution of maxima (EV1 or Gumbel distribution). Using simple calculus it is found that in this case, (4) takes the form

$$H(x) = \exp[-\exp(-x/\lambda + \psi)] \tag{6}$$

which is unbounded from both below and above ($-\infty < x < +\infty$).

When $\kappa < 0$, $H(x)$ represents the type III (EV3) distribution of maxima. This, however, should be of no practical interest in hydrology as it refers to random variables bounded from above ($-\infty < x \le \lambda\,\psi - \lambda/\kappa$). As discussed in the introduction, there is no general consensus on this and many regard an upper bound in natural quantities as reasonable. Even Jenkinson (1955) regards the EV3 distribution as "the most frequently found in nature, since it is reasonable to expect the maximum values to have an upper bound". However, he leaves out rainfall from this conjecture saying "to a considerable extend rainfall amounts are 'uncontrolled' and high falls may be recorded". In fact, he proposes the EV2 distribution for rainfall (note that he uses a different convention, calling EV2 as type I). In a recent study, Sisson et al. (unpublished), even though detect EV2 behaviour of rainfall maxima, they attempt to incorporate the idea of a PMP upper bound within an EV2 modelling framework.

Furthermore, it is noted that if the distribution of minima is of interest, the roles of types II and III reverse, e.g. the type III distribution is not bounded from above and thus it is a reasonable model for the study of droughts.

The close relationship between the distribution of maxima $H(x)$ and the tail of the parent distribution $F(x)$ allows for the determination of the latter if the former is known. The tail of $F(x)$ can be represented by the distribution of $x$ conditional on being greater than a certain threshold $\xi$, i.e., $G_\xi(x) := F(x|x > \xi)$, for which

$$1 - G_\xi(x) = \frac{1 - F(x)}{1 - F(\xi)}, \qquad x \ge \xi \tag{7}$$

If we choose $\xi$ so that the exceedence probability $1 - F(\xi)$ equals $1/\nu$, the reciprocal of the mean number of events in a year (this is implied when the partial duration series is formed from a time series of measurements, by choosing a number of events equal to the number of years of record), and denote $G(x)$ the conditional distribution for this specific value, then

$$1 - G(x) = \nu\,[1 - F(x)] \tag{8}$$

Combining (8) with (3) it is obtained that

$$G(x) = 1 + \ln H'_\nu(x) \tag{9}$$

If $H'_\nu(x)$ is given by the limit distribution $H(x)$ in (4), then it is concluded that for $\kappa > 0$

$$G(x) = 1 - \left[1 + \kappa\left(\frac{x}{\lambda} - \psi\right)\right]^{-1/\kappa}, \qquad x \ge \lambda\,\psi \tag{10}$$

which is the Pareto distribution. Similarly, for $\kappa = 0$

$$G(x) = 1 - \exp\,(-x/\lambda + \psi), \qquad x \ge \lambda\,\psi \tag{11}$$

which is the exponential distribution. For the special case $\psi = 1/\kappa$

$$G(x) = 1 - \left(\frac{\lambda}{\kappa\,x}\right)^{1/\kappa}, \qquad x \ge \lambda/\kappa \tag{12}$$

which is a power law relationship between the distribution quantile $x$ and the return period $T := 1\,/\,[1 - G(x)]$, the mean time interval between exceedences of $x$, expressed in years. Specifically, (12) can be written as $x = (\lambda/\kappa)\,T^\kappa$. Turcotte (1994) used this special power law (also calling it a fractal law) to model flood peaks over threshold in 1200 stations in the United States. In the generalized Pareto case (10), the corresponding relationship is $x = (\lambda/\kappa)\,(T^\kappa - 1 + \kappa\,\psi)$

## 3. The prevailing of the Gumbel distribution

Due to their simplicity and generality, the limiting extreme value distributions $H(x)$ have become very widespread in hydrology. The exact distributions $H_n(x)$ and $H'_\nu(x)$ (equations (2) and (3)) are used rarely in studies of hydrological extremes, as their determination requires the parent distribution $F(x)$ to be known. The determination of $F(x)$ may be too complex and is not necessary as its truncated version $G(x)$ is sufficient for the study of extremes.

In particular, as mentioned in the introduction, the EV1 extreme value distribution has been by far the most popular model of extremes. In hydrological education is so prevailing that most textbooks contain the EV1 distribution only, omitting EV2. In hydrological engineering studies, especially those analysing rainfall maxima, the use of EV1 has become so common that its adoption is almost automatic, without any reasoning or comparing it with other possible models. There are several reasons for this:

a. **Theoretical reasons.** Most types of parent distribution functions that are used in hydrology, such as exponential, gamma, Weibull, normal, and lognormal, (e.g. Kottegoda and Rosso, 1997, p. 431) belong to the domain of attraction of the Gumbel distribution. In contrast, the domain of attraction of the EV2 distribution includes less frequently used parent distributions like Pareto, Cauchy, and log-gamma.
b. **Simplicity.** The mathematical handling of the two-parameter EV1 is much simpler than that of the three-parameter EV2 (see also point **d** below).
c. **Accuracy of estimated parameters.** Obviously, two parameters are more accurately estimated than three. For the former case, mean and standard deviation (or second L-moment) suffice, whereas in the latter case the skewness is also required and its estimation is extremely uncertain for typical small-size hydrological samples.
d. **Practical reasons.** Probability plots are the most common tools used by practitioners, engineers and hydrologists, to choose an appropriate distribution function. EV1 offers a linear probability plot, known as Gumbel probability plot, which is a diagram of $x_H$ versus the so called Gumbel reduced variate, defined as $z_H := -\ln(-\ln H)$. The observed $z_H$ is estimated in terms of plotting positions, i.e. sample estimates of probability of non-exceedence. In contrast, a generalized linear probability plot for the three-parameter EV2 is not possible to construct (unless one of the parameters is fixed). In fact, if $\kappa > 0$, the plot of $x_H$ versus $z_H$ is a convex curve. This may be

regarded as a primary reason of choosing EV1 against EV2 in practice. For the Fréchet distribution, a linear plot is possible (a plot of ln $x_H$ versus $z_H$ that will be referred to as the Fréchet probability plot). However, empirical evidence shows that, in most cases, plots of $x_H$ versus $z_H$ give more straight-line arrangements than plots of ln $x_H$ versus $z_H$. An additional practical reason is the fact that many institutions suggest, or even require, the use of EV1.

As mentioned in the introduction, EV1 has one potential disadvantage, which is very important from the engineering point of view: For small probabilities of exceedence (or large return periods) it yields the smallest possible quantiles $x_H$ in comparison to those of EV2 for any (positive) value of the shape parameter $\kappa$. This means that EV1 results in the highest possible risk for engineering structures. Normally, this would be a sufficient reason to avoid the use of EV1 in engineering studies.

Obviously, this disadvantage of EV1 would be counterbalanced only by strong empirical evidence and theoretical reasoning. In practice, the small size of common hydrological records (e.g. a few tens of years) cannot provide sufficient empirical evidence for preferring EV1 over EV2. This is discussed in section 5 based on a simulation study and in part 2 of the study based on the real-world collection of long records. In addition, the theoretical reasons, exhibited in point **a** above, are not strong enough to justify the automatic adoption of the Gumbel distribution. This is discussed in section 4.

## 4. Theoretical study of the appropriateness of the Gumbel distribution

To begin the theoretical discussion, it will be assumed that the events, whose maximum values are studied, can be represented as independent identically distributed random variables $Y_i$ (Assumption 1). Further, it will be assumed that the (unknown) parent distribution $F(y)$ belongs, with absolute certainty, to the domain of attraction of EV1 (Assumption 2). Are these rather oversimplifying and implausible assumptions sufficient to justify the adoption of EV1? The answer is clearly, No. This answer is demonstrated in Figure 1, which depicts Gumbel probability plots of the exact distribution functions of maxima $H_n(x)$ for $n = 10^3$ and

$10^6$ for two parent distribution functions. The first (left panel) is the standard normal distribution and the second (right panel) is the Weibull distribution ($F(y) = 1 - \exp(-y^k)$) with shape parameter $k = 0.5$. Both parent distributions belong to the domain of attraction of the EV1 limiting distribution, so it is expected that the Gumbel probability plot tends to a straight line as $n \to \infty$. However, the tendency is remarkably slow, and even for $n$ as high as $10^6$ the curvature of the distribution functions is apparent. Obviously, in hydrological applications, such a high number of events within a year, is not possible (it can be expected that the number of storms or floods in a location will not exceed the order of 10-$10^2$). Thus, the limiting distribution for $n \to \infty$ is not useful at all.

When studying storms and floods at a fine time scale, the parent distribution has typically a positively skewed, J-shaped density function. Thus, the normal distribution is not relevant in this case, but the Weibull distribution with shape parameter smaller than 1 (e.g. $k = 0.5$ as in the example of Figure 1) can be a plausible parent distribution. In this case, it is observed in Figure 1, that the probability plots are convex curves, which indicates that, for a specified $n$, an EV2 distribution may approximate sufficiently the exact distribution. Thus, even if the parent distribution belongs to the domain of attraction of the Gumbel distribution, an EV2 distribution can be a choice closer to the exact distribution of maxima in comparison to EV1.

Now, the Assumption 1 set above will be relaxed, forming the more plausible Assumption 1A. According to this, the events whose maximum values are studied are independent random variables $Y_i$ but not identically distributed ones. Instead, it is assumed that all $Y_i$ have the same type of distribution function $F_i(y)$ but with different parameters. This distribution function belongs to the domain of attraction of the Gumbel distribution, i.e., Assumption 2 is valid for each $F_i(y)$.

The relaxed assumption 1A is more consistent with hydrological reality. The statistical characteristics (e.g., averages, standard deviations etc.) and, consequently, the parameters of distribution functions exhibit temporal (e.g. seasonal) variation. In addition, evidence from long geophysical records shows that there appear fluctuations of local statistical properties on large time scales (e.g., tens of years, hundreds of years, etc.). It has been proposed that such fluctuations, either periodical or irregular, occurring either on a single time scales or on multiple time scales simultaneously, constitute the physical basis of the well-known Hurst phenomenon (Klemes, 1974; Montanari et al., 1999; Koutsoyiannis, 2002). Here it should be noted that such fluctuations may not be detected in series of maxima, which typically satisfy Leadbetter's (1974) condition of the absence long-range dependence, but surely affect the parent distribution of rainfall at low- and intermediate-level exceedences.

The consequences of Assumption 1A are demonstrated by examples in which the parent distribution is specified to be the gamma distribution (which belongs to the domain of attraction of EV1) with varying scale parameter. (Additionally, the shape parameter could be regarded as a varying one, but the study of the effect of the variation of the scale parameter is mathematically more convenient and sufficient for the demonstration attempted.) More specifically, it may be assumed that during some 'epoch' (e.g. a specific month of a year through one or more years) the scale parameter is fixed to some value $\alpha_i > 0$. In this case, the probability density function of $Y_i$, conditional on $\alpha_i$, is

$$f_i(y|\alpha_i) = \alpha_i^{\theta}\, y^{\theta - 1}\, \mathrm{e}^{-\alpha_i y} / \Gamma(\theta) \tag{13}$$

where the shape parameter $\theta > 0$ was kept constant for all 'epochs'. In the first example it will be assumed that $\alpha_i$ varies randomly following a gamma distribution itself with scale parameter $\beta > 0$ and shape parameter $\tau > 0$, so that its density is

$$g(\alpha_i) = \beta^{\tau}\, \alpha_i^{\tau - 1}\, \mathrm{e}^{-\beta\, \alpha_i} / \Gamma(\tau) \tag{14}$$

If one is interested on the unconditional distribution of the variable $Y$, that is valid over all epochs, instead of a specified epoch, then one should use (13) and (14) to determine the marginal density of $Y$, which is

$$f(y) = \int_0^{\infty} f_i(y|\alpha_i)\, g(\alpha_i)\, d\alpha_i = \{\beta^{\tau}\, y^{\theta - 1} / [\Gamma(\theta)\, \Gamma(\tau)]\} \int_0^{\infty} \alpha_i^{\theta + \tau - 1}\, \mathrm{e}^{-(y + \beta)\, \alpha_i}\, d\alpha_i \tag{15}$$

After algebraic manipulations it is obtained that

$$f(y) = \frac{1}{\beta\, B(\theta, \tau)} \frac{(y/\beta)^{\theta - 1}}{(1 + y/\beta)^{\tau + \theta}} \tag{16}$$

which shows that the marginal distribution of $Y/\beta$ is beta of the second kind (Kendal and Stuart, 1963, p. 151; Yevjevich, 1972, p. 149). Consequently, the marginal probability distribution function of $Y$ is

$$F(y) = B_{y/(y+\beta)}(\theta, \tau) / B(\theta, \tau) \tag{17}$$

where $B_z(\theta, \tau)$ and $B(\theta, \tau)$ denote respectively the incomplete beta function and the Euler (complete) beta function, i.e.,

$$B_z(\theta, \tau) := \int_0^z t^{\theta - 1} (1 - t)^{\tau - 1}\, dt, \quad B(\theta, \tau) := \int_0^\infty t^{\theta - 1} (1 - t)^{\tau - 1}\, dt \tag{18}$$

Thus, the exact distribution of maxima for constant and variable $n$ is respectively

$$H_n(x) = [B_{x/(x+\beta)}(\theta, \tau) / B(\theta, \tau)]^n, \quad H'_\nu(x) = \exp\{-\nu[1 - B_{x/(x+\beta)}(\theta, \tau) / B(\theta, \tau)]\} \tag{19}$$

For $\theta = 1$, the parent distribution (17) simplifies to

$$F(y) = 1 - (1 + y/\beta)^{-\tau} \tag{20}$$

which is the Pareto distribution. Clearly, this belongs to the domain of attraction of EV2 with zero lower bound, i.e., the limiting distribution of maxima $H(x)$ is the Fréchet distribution. In the general case, it can be shown that

$$\lim_{y \to \infty} \frac{y\, f(y)}{1 - F(y)} = \tau > 0 \tag{21}$$

which is a sufficient condition for convergence of $H_n(x)$ to the EV2 distribution (e.g. Kottegoda and Rosso, 1997, p. 430).

In Figure 2 it is demonstrated how the exact distribution tends to the Fréchet distribution as $n$ increases. In this case the shape parameter $\theta$ was assumed 0.5 and the exact distribution was calculated from (19). For $n$ as high as 1000 the Fréchet probability plot becomes almost a straight line. However, as in the cases of Figure 1, for smaller values of $n$, which are more relevant in hydrological applications, the Fréchet plot of the exact distribution appears to be curved, so an EV2 distribution would yield a better approximation to the exact distribution than the Fréchet distribution. (It is noted that the concave curvature appearing in the Fréchet plot of Figure 2 would be convex in a Gumbel plot).

A more specific numerical experiment is depicted in Figure 3. Here the exact distributions of maxima $H_5(x)$ (for $n$ = 5), based on assumptions 1 and 1A, are compared. In case 1A, a variable parameter gamma distribution was assumed, with parameters $\theta = 0.5$, $\tau = 5$ and $\beta = 1$. In case 1, the variable scale parameter is replaced by a constant parameter $\alpha = \tau/\beta = 5$ (equal to the mean of the scale parameter of case 1A). The exact distribution of maxima for case 1 is almost a straight line on the Gumbel probability plot whereas that of case 1A is a convex curve. In addition to the theoretical distribution functions, empirical ones were also plotted, based on 4000 synthetic maxima. To these synthetic data series the EV1 and EV2 distributions were fitted and were also plotted in Figure 3. As expected, EV1 is in good agreement with the exact distribution of case 1 but departs significantly from the exact distribution in case 1A, especially in the tail that corresponds to large return periods. In contrast, the EV2 distribution (estimated $\kappa = 0.20$) is almost indistinguishable from the exact distribution.

A second simpler example was based again on gamma parent distribution function with constant shape parameter $\theta = 0.5$ and scale parameter shifting between two values, $\alpha_1 = 2$ and $\alpha_2 = 6$ which are sampled at random with probabilities 0.25 and 0.75, respectively. The two values of the scale parameter could be thought of as representing two epochs, a wet and a dry or even two distinct competing processes as in Walshaw (2000). For comparison, a gamma distribution with constant parameter $\alpha = 5$ (again equal to the mean of $\alpha_1$ and $\alpha_2$) was used. Here, the theoretical distributions were not determined but rather empirical ones were plotted, based on 4000 synthetic maxima. To these synthetic data series, the EV1 and EV2 distributions were fitted and were also plotted in Figure 4. As in Figure 3, EV1 is in good agreement with the empirical distribution of the constant parameter case but departs significantly from the empirical distribution of the variable parameter case. Again, the departure is greatest in the tail, i.e. in large return periods. In contrast, EV2 (with $\kappa = 0.20$) agrees well with the simulated distribution.

All this theoretical discussion and the examples show that the theoretical reasons, which have endorsed the use of the Gumbel distribution for hydrological extremes, are not strong enough to compensate the high risk it implies.

## 5. The hiding of the EV2 distribution

If, according to the previous analysis, an EV2 distribution is more likely to represent hydrological maxima than an EV1 distribution, the question arises, Why this was not manifested in maximum rainfall series, which are typically attributed an EV1 behaviour? The answer to this question is simple: Typical annual maximum rainfall series extend over 20-50 years and such a record length hides the EV2 distribution and displays an EV1 behaviour.

This was demonstrated by Koutsoyiannis and Baloutsos (2000) using an annual series of maximum daily rainfall in Athens, Greece, extending through 1860-1995 (136 years). This series was found to follow the EV2 distribution, but if small parts of the series were analysed, the EV1 distribution seemed to be an appropriate model.

Here a more systematic analysis has been done based on Monte Carlo simulations for different sample sizes $m$ and different shape parameters $\kappa$. For each combination of $m$ and $\kappa$, 200 synthetic records were generated from the EV2 distribution. For each synthetic record, the parameter $\kappa$ was assumed unknown and was estimated from the record, using the methods of moments and L-moments. From the simulation results, a negative bias, defined as estimated $\kappa$ minus true $\kappa$, became apparent, and its expected magnitude $b$ was computed as the average of the 200 samples. As expected, $b$ is found to be a function of both $m$ and $\kappa$, which can be approximated by the following expressions (chosen after several trials and fitted numerically to simulation results):

$$b = -1.7\,\kappa^3 - \frac{55\,\kappa + 2}{m + 200\,\kappa + 20}, \quad b = -\frac{0.44\,e^{4\,\kappa}}{m} \tag{22}$$

for the moments and L-moments estimators, respectively. The expressions are plotted in Figure 5. It can be observed that for $\kappa = 0.15$ (a value that is typical for extreme rainfall as it will be shown in part 2 of the study) and for a record length of 20 years the bias is –0.15, which means that the estimated $\kappa$ will be zero! Even for a record length of 50 years the negative bias is high ($b$ = –0.12), so that $\kappa$ will be estimated at 0.03, a value that will not give good reason for preferring EV2 to EV1.

The situation is improved if L-moments estimators are used as the resulting bias is much lower (Figure 5). However the method of L-moments is relatively new (Hosking et al., 1985; Hosking, 1990) and its use has not been very common so far. In addition, even the method of L-moments will not reject the hypothesis of an EV1 distribution against the hypothesis of EV2 distribution thus making a type II error (no rejection of a false hypothesis) with a high probability. To demonstrate this, the L-moments $\kappa$ test (Hosking et al., 1985; see also Stedinger et al., 1993) was used, which tests whether $\kappa = 0$ (i.e., appropriateness of the EV1 distribution; null hypothesis) or not (alternative hypothesis). To determine the probability of type II error using this test, a simulation experiment was performed similar to the one already described, assuming that the true distribution is EV2 with several values of $\kappa$. With this assumption, for each combination of $\kappa$ and $m$ 10000 synthetic records were generated and for each record the $\kappa$-test was applied with null hypothesis $\kappa = 0$, alternative hypothesis $\kappa > 0$ and significance level 5%. The results of simulations have been plotted in Figure 6. It can be observed that for $\kappa = 0.15$ and $m = 20$ the frequency of not rejecting the EV1 distribution is 80%! Even for $m = 50$ this frequency is high, 62%.

For the method of maximum likelihood, a single simulation experiment was performed corresponding to $\kappa = 0.15$ and $m = 30$. The bias was found to be about half that of the method of L-moments, i.e. not very substantial. Its sample variance, however was equally high as in the other methods, which indicates that the uncertainty in estimating the shape parameter is high even using the more accurate method of maximum likelihood. It is noted that the latter method is not widely used in engineering applications, as, in contrast to the other two methods which are simple, it requires numerical optimization (no analytical solution can be obtained). Consequently, a more detailed demonstration of the behaviour of this method is not relevant to this investigation, whose purpose is to demonstrate whether common engineering practices hide or display an underlying EV2 behaviour of a series of maxima. The reader interested in additional information of the application of the maximum likelihood method in rainfall extremes, including testing of rejecting the hypothesis of EV1 versus EV2, as well as in Baysian analysis in fitting EV2 to rainfall extremes, is referenced to Coles and Pericchi (2003).

The findings of this investigation show that the empirical evidence supporting the wide applicability of the Gumbel distribution may in fact be the result of too small sample sizes and imperfections of parameter fitting methods, rather than a manifestation of the real behaviour of rainfall maxima. To improve the clarity of the real behaviour of extreme rainfall series longer records are needed. Such records are investigated in the second part of the study.

## 6. Synopsis and conclusion

The Gumbel or EV1 distribution has been the prevailing model for rainfall extremes despite of the fact that it results in the highest possible risk for engineering structures, i.e. it yields the smallest possible design rainfall values in comparison to those of EV2 for any value of the shape parameter. The simplicity of the calculations of the EV1 distribution along with its geometrical elucidation through a linear probability plot may have contributed to its popularity in hydrologists and engineers. There is also a theoretical justification, as EV1 is the asymptotic extreme value distribution for a wide range of parent distributions that are common in hydrology.

However, the theoretical investigation of this study shows that the convergence of the exact distribution of maxima to the asymptote may be extremely slow, thus making the EV1 asymptotic distribution an inappropriate approximation of the exact distribution of maxima. Besides, the attraction of parent distributions to the EV1 asymptote relies on a stationarity assumption, i.e. the assumption that parameters of the parent distribution are constant in time, which may not be the case in hydrological processes. Slight relaxation of this assumption may result in the EV2 rather than the EV1 asymptote.

On the contrary, the EV2 distribution does not have the theoretical disadvantages of the EV1 distribution. Even though it is still a limiting distribution, it can yield good approximations to the exact distribution of maxima yet away from the limit, and it is not very sensitive to changes of parameters in time.

The simulation experiments of the study show that small sizes of records, e.g. 20-50 years, hide the EV2 distribution and display it as if it were EV1. This allows the conjecture that the broad use of the EV1 distribution worldwide may in fact be related to small sample sizes rather than to the real behaviour of rainfall maxima, which should be better described by the EV2 distribution. This conjecture is investigated in the second part of the study using 169 of the longest available rainfall records worldwide.

**Acknowledgements** The comments by Andras Bardossy, Charles Obled, Kieran O'Connor, Renzo Rosso and Ezio Todini on an early version of this work that was presented at the ESF LESC Exploratory Workshop, Hydrological Risk held at the University of Bologna were encouraging and helpful. With the editor's help, the paper was so fortunate as to bring two very encouraging, constructive and detailed reviews by Alberto Montanari and an anonymous reviewer, which improved the paper substantially.

## References

Benson, M. A., (1973) Thoughts on the design of design floods, in *Floods and Droughts, Proc. 2nd Intern. Symp. in Hydrology*, pp. 27-33, Water Resources Publications, Fort Collins, Colorado.

Coles, S., and Pericchi, L. (2003) Anticipating catastrophes through extreme value modelling, Applied Statistics, 52, 405-416.

Coles, S., Pericchi, L. R., and Sisson, S., (2003) A fully probabilistic approach to extreme rainfall modeling, Journal of Hydrology, 273(1-4), 35-50.

Dooge, J. C. I., (1986) Looking for hydrologic laws, *Water Resour. Res.*, 22(9) pp. 46S-58S.

Farquharson, F. A. K., Meigh, J. R., and Sutcliffe, J. V. (1992), Regional flood frequency analysis in arid and semi-arid areas, *J. Hydrol.*, 138, 487–501.

Fisher, R. A., and Tippet, L. H. C. (1928) Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proc. Cambridge Phil. Soc.*, 24, 180-190.

Fréchet, M. (1927) Sur la loi de probabilité de l'écart maximum, *Ann. de la Soc. Polonaise de Math.*, Cracow, 6, 93-117.

Gnedenco, B.V. (1941) Limit theorems for the maximal term of a variational series, *Doklady Akad. Nauk SSSR*, Moscow, 32, 37 (in Russian).

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

Hazen, A. (1914) Storage to be provided in impounding reservoirs for municipal water supply, *Trans. ASCE*, ASCE, New York, 77, 1539–1640.

Hershfield, D. M. (1961) Estimating the probable maximum precipitation, *Proc. ASCE, J. Hydraul. Div.*, 87(HY5), 99-106.

Hershfield, D. M. (1965) Method for estimating probable maximum precipitation, *J. American Waterworks Association*, 57, 965-972.

Hosking, J. R. M. (1990) L-moments: Analysis and estimation of distributions using linear combinations of order statistics, *J. R. Stat. Soc.*, Ser. B, 52, 105-124.

Hosking, J. R. M., Wallis, J. R., and Wood, E. F. (1985) Estimation of the generalized extreme value distribution by the method of probability weighted moments, *Technometrics*, 27(3), 251-261.

Jenkinson, A. F. (1955) The frequency distribution of the annual maximum (or minimum) value of meteorological elements, *Q. J. Royal Meteorol. Soc.*, 81, 158-171.

Jenkinson, A. F. (1969) Estimation of maximum floods, *World Meteorological Organization, Technical Note No 98*, ch. 5, 183-257.

Kendall, M. G. and Stuart, A. (1963) *The advanced theory of Statistics, Vol.1, Distribution theory*, 2nd edition, C. Griffin & Co., London.

Klemeš, V. (1974) The Hurst phenomenon - a puzzle?", *Water Resour. Res.*, 10, 675-688.

Klemeš, V. (2000) Tall tales about tails of hydrological distributions, *J. Hydrol. Engineering*, 5(3), 227-231 & 232-239.

Kottegoda, N. T., and Rosso, R. (1997) *Statistics, Probability, and Reliability for Civil and Environmental Engineers*, McGraw-Hill, New York.

Koutsoyiannis, D. (1999) A probabilistic view of Hershfield's method for estimating probable maximum precipitation, *Water Resources Research*, 35(4), 1313-1322.

Koutsoyiannis, D. (2002) The Hurst phenomenon and fractional Gaussian noise made easy, *Hydrological Sciences Journal*, 47(4), 573-595.

Koutsoyiannis, D., and G. Baloutsos, G. (2000) Analysis of a long record of annual maximum rainfall in Athens, Greece, and design rainfall inferences, *Natural Hazards,* 22(1), 31-51.

Leadbetter M. R. (1974) On extreme values in stationary sequences, *Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 28, 289-303.

Montanari, A., Taqqu, M.S. and Teverovsky, V. (1999) Estimating long-range dependence in the presence of periodicity: an empirical study, *Mathematical and Computer Modeling*, 29, 217-228.

Rossi, F., Fiorentino, M., and Versace, P. (1984) Two-component extreme value distribution for flood frequency analysis, *WaterResour. Res.*, 20(7), 847-856.

Sisson, S. A., Pericchi, L. R., and Coles, S. (unpublished) A case for a reassessment of the risks of extreme hydrological hazards in the Caribbean (www.maths.unsw.edu.au/~scott/papers/paper_caribbean_abs.html)

Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E. (1993) Frequency analysis of extreme events, ch. 18 in *Handbook of Hydrology,* edited by D. R. Maidment, McGraw-Hill, New York.

Sutcliffe J. V. (1978) *Methods of Flood Estimation, A Guide to Flood Studies Report*, Report No. 49, Institute of Hydrology, UK.

Todorovic, P., and Zelenhasic, E. (1970) A stochastic model for flood analysis , *Water Resour. Res.*, 6(6), 1641-1648.

Turcotte, D. L. (1994) Fractal theory and the estimation of extreme floods, J. Res. Natl. Inst. Stand. Technol., 99(4), 377-389.

Turcotte, D. L., and Malamud, B. D. (2003) Applicability of fractal flood-frequency statistics, *Hydrofractals '03, An international conference on fractals in hydrosciences*, Monte Verita, Ascona, Switzerland, August 2003, ETH Zurich, MIT, Université Pierre et Marie Curie.

U.S. Department of the Interior, Bureau of Reclamation (1977) *Design of Arch Dams*, US Goverment Printing Office, Denver, Co.

U.S. Department of the Interior, Bureau of Reclamation (1987) *Design of Small Dams*, 3rd edition, US Goverment Printing Office, Denver, Co.

von Bortkiewicz, L. (1922) Variationsbreite und mittlerer Fehler, *Sitzungsberichte d. Berliner Math. Ges.*, 21, 3.

von Mises, R. (1923) Über die Variationsbreite einer Beobachtungsreihe, *Sitzungsber, d. Berliner Math. Ges.*, 22, 3.

Walshaw, D. (2000) Modelling extreme wind speeds in regions prone to hurricanes, *Applied Statistics*, 49, 51-62.

Wilks, D. S. (1993) Comparison of three-parameter probability distributions for representing annual extreme and partial duration precipitation series, *Water Resour. Res.*, 29(10), 3543-3549.

Willeke, G. E. (1980) Myths and uses of hydrometeorology in forecasting, in *Proceedings of March 1979 Engineering Foundation Conference on Improved Hydrological Forecasting – Why and How*, pp. 117-124, American Society of Civil Engineers, New York.

World Meteorological Organization (1986), *Manual for Estimation of Probable Maximum Precipitation*, Operational Hydrology Report 1, 2nd edition, Publication 332, World Meteorological Organization, Geneva.

Yevjevich, V. (1972), *Probability and Statistics in Hydrology*, Water Resources Publications, Fort Collins, Colorado.

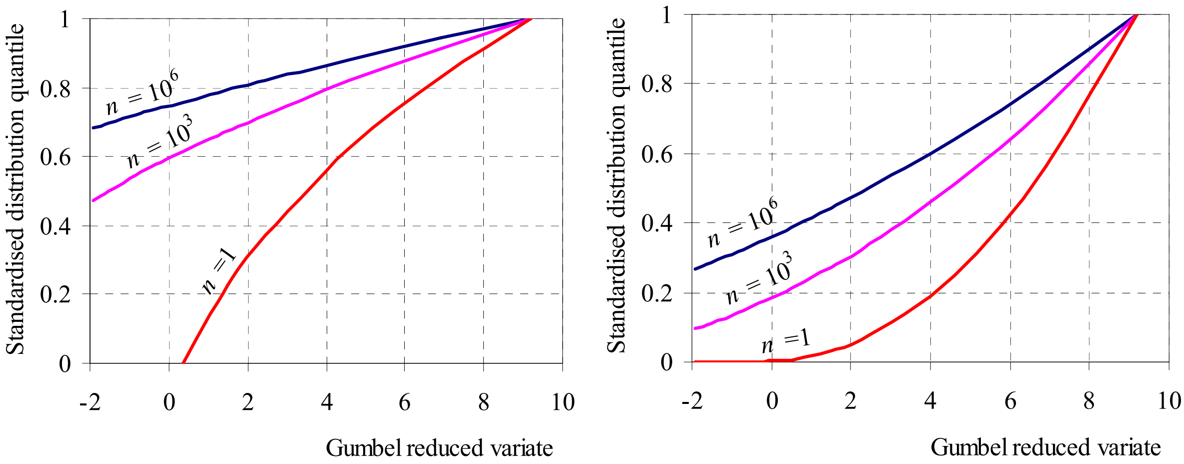


**Figure 1.** Gumbel probability plots of the exact distribution function of maxima $H_n(x)$ for $n = 10^3$ and $10^6$, also in comparison with the parent distribution function $F(y) \equiv H_1(y)$, which in the left panel is standard normal and in the right panel Weibull with shape parameter $k = 0.5$. The distribution quantile has been standardized by $x_{0.9999}$ corresponding to $z_H = 9.21$.

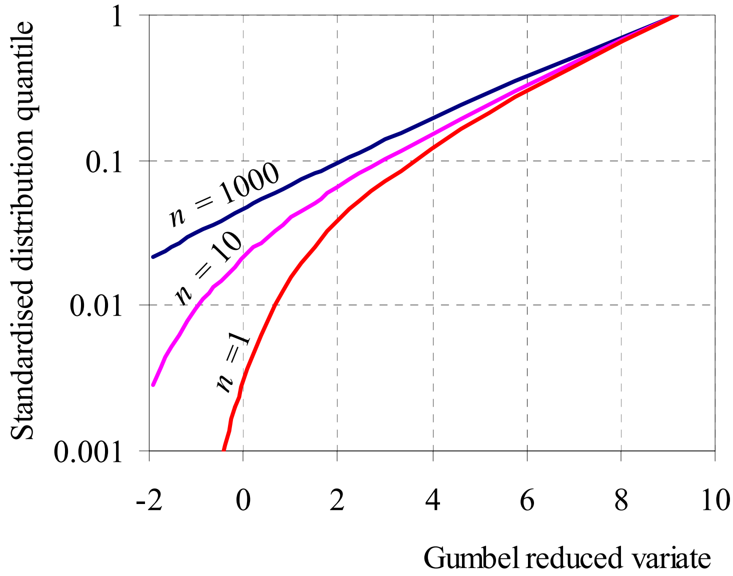


**Figure 2.** Fréchet probability plot of the exact distribution function of maxima $H_n(x)$ for $n$ = 1, 10 and 1000, as this results assuming a gamma parent distribution with shape parameter $\theta$ = 0.5 and scale parameter randomly varying following a second gamma distribution with shape parameter $\tau$ = 3 and scale parameter $\beta = 1$. The distribution quantile has been standardized by $x_{0.9999}$ corresponding to $z_H = 9.21$.

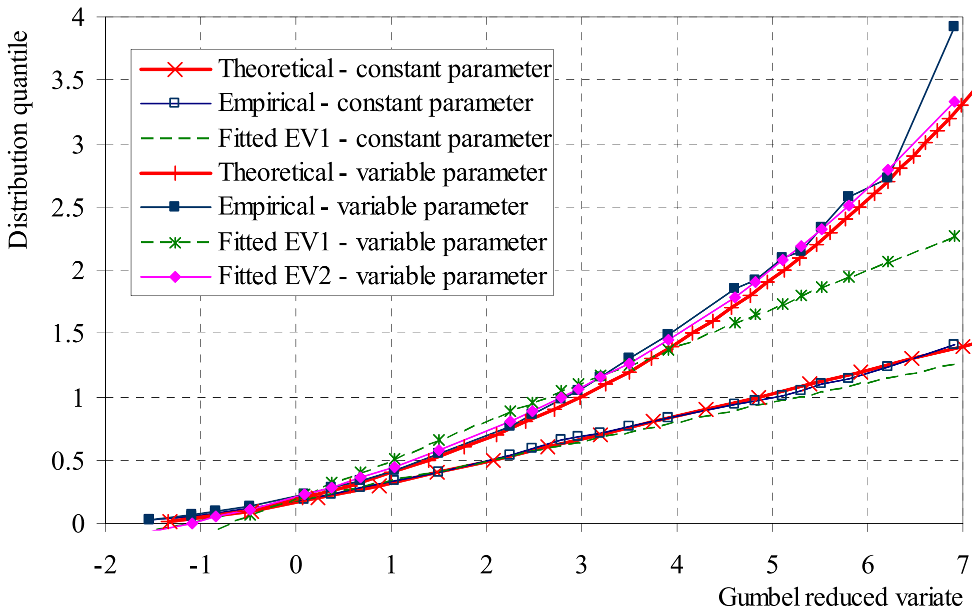


**Figure 3.** Gumbel probability plot of the exact distribution function of maxima $H_5(x)$, as this results assuming a gamma parent distribution with shape parameter $\theta$ = 0.5 and scale parameter either constant $\alpha$ = 5 (case 1) or randomly varying following a second gamma distribution with shape parameter $\tau$ = 5 and scale parameter $\beta$ = 1 (case 1A). The additional plotted curves are empirical distribution functions from synthesized series of length 4000, and fitted to these series EV1 and EV2 distribution functions.

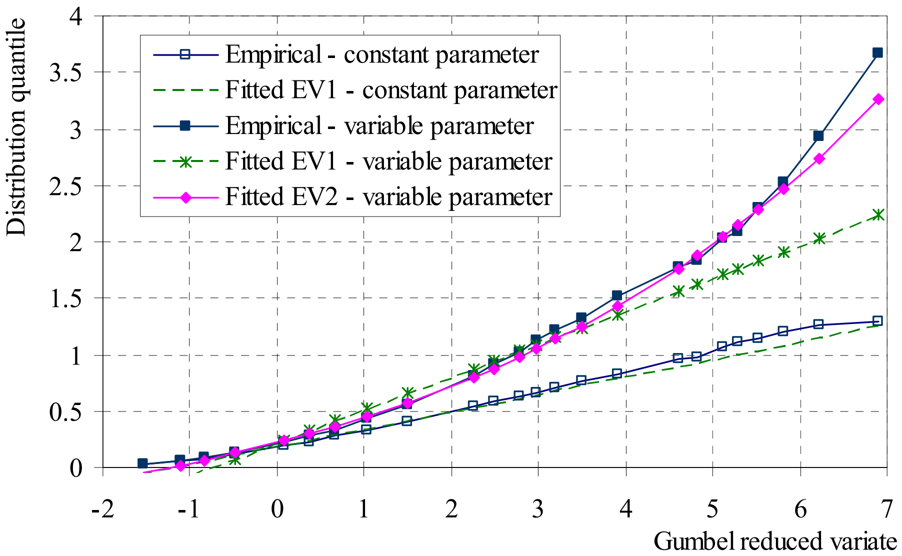


**Figure 4.** Gumbel probability plot of the empirical distribution functions of maxima $H_5(x)$ and fitted EV1 and EV2 distribution functions, as they result from synthesized series of length 4000 assuming gamma parent distribution with shape parameter $\theta$ = 0.5 and scale parameter either constant $\alpha$ = 5, or shifting at random between the values $\alpha_1$ = 2 and $\alpha_2$ = 6 with probabilities 0.25 and 0.75, respectively.

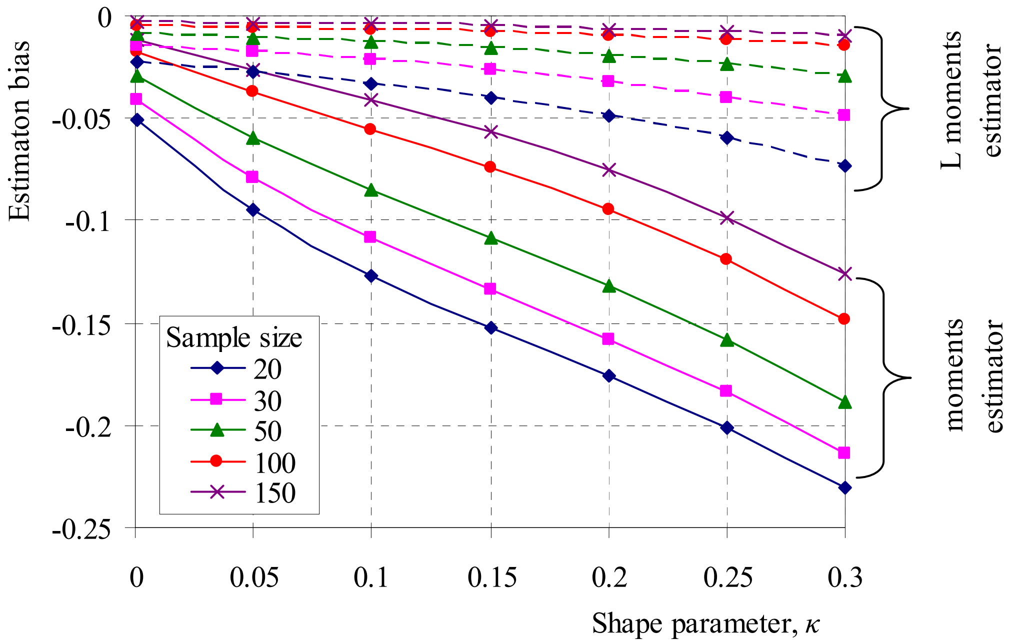


**Figure 5.** Bias in estimating the shape parameter $\kappa$ of the GEV distribution using the methods of moments and L-moments.

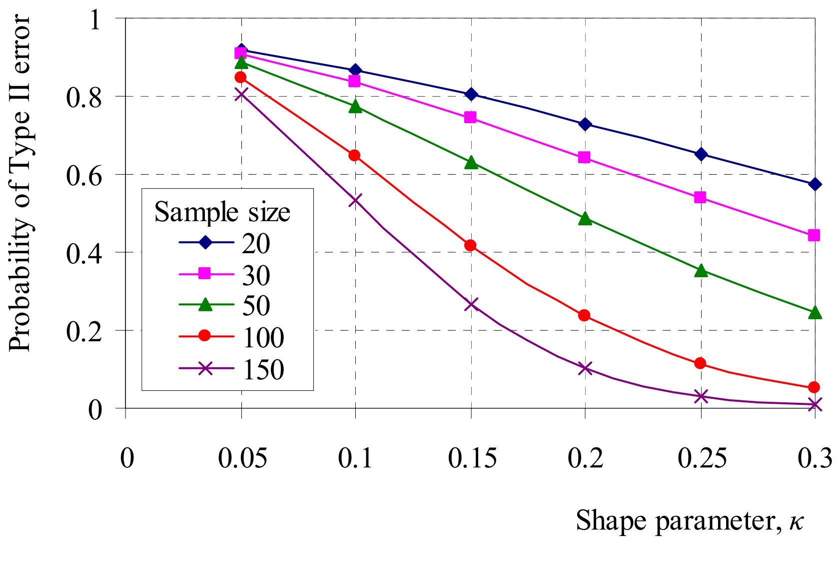


**Figure 6.** Probability of type II error (no rejection of a false hypothesis) in testing the null hypothesis that a series originates from the EV1 distribution against the hypothesis that it originates from the EV2 distribution.