Review

# Advances in Deriving the Exact Distribution of Maximum Annual Daily Precipitation 

Carlo De Michele ${ }^{\text {(1) }}$<br>Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133 Milano, Italy; carlo.demichele@polimi.it

Received: 22 May 2019; Accepted: 30 October 2019; Published: 6 November 2019


#### Abstract

Maximum annual daily precipitation does not attain asymptotic conditions. Consequently, the results of classical extreme value theory do not apply to this variable. This issue has raised concerns about the frequent use of asymptotic distributions to model the maximum annual daily precipitation and, at the same time, has rekindled interest in deriving and testing its exact (or non-asymptotic) distribution. In this review, we summarize and discuss results to date about the derivation of the exact distribution of maximum annual daily precipitation, with attention on compound/superstatistical distributions.


Keywords: maximum annual daily precipitation; exact distribution; asymptotic distribution; compound/superstatistical distribution

## 1. Introduction

Classical methods for modeling the probability distribution of maximum annual daily precipitation use an asymptotic extreme-value distribution. Fréchet [1] identified one possible asymptotic extreme value distribution, which received the author's name (Fréchet distribution). Fisher and Tippet [2] and, later Gnedenko [3] showed that, asymptotically, there exist only three types of extreme value distributions, referred to by Gumbel [4] as the "extreme value asymptotes". These are EVI (or Gumbel distribution), EVII (or Fréchet distribution), and EVIII (or reversed Weibull distribution). These three asymptotic distributions were combined by von Mises [5] into the Generalized Extreme Value distribution (GEV). Sometimes, the GEV is also referred to as the von-Mises-type extreme value distribution, or the von Mises-Jenkinson-type distribution.

The first studies in which asymptotic extreme value distributions were used to model the occurrence of maximum annual daily precipitation were those by Jenkinson [6] and then Gumbel [4]. The main advantage of asymptotic laws is that they do not require knowledge of the parent distribution (that is, the distribution of the underlying variable). The tendency to reduce the dimensionality of the extreme-value models and to proceed on the basis of simplified models has also supported the use of asymptotic approaches [7]. In this context, the Gumbel (EVI) distribution has been the most popular model for extremes, especially in the analysis of precipitation maxima [8]. However, the Gumbel distribution was found to underestimate the extreme precipitation amounts [8,9]. It has been suggested that the GEV should always be used instead of the Gumbel distribution, unless enough information supporting the latter is available [7].

A (non-exhaustive) list of papers published in the period 2000-2015 that exclusively used the GEV, or its three asymptotic laws, to describe the behavior of maximum annual daily precipitation is reported in [10]. Some studies have also investigated which of these three asymptotic laws is the most appropriate to represent the statistical variability of maximum annual daily precipitation [8,11]. Using a worldwide database, it has been shown that the EVII is the most suitable asymptotic distribution to describe the maximum annual daily precipitation [12].

The applicability of the asymptotic extreme value theory to annual maxima of precipitation, both for daily and for shorter durations, was recently questioned in Veneziano et al. [13]. The motivation is the comparatively slow convergence of the distribution of annual maxima to the asymptotic laws (see also [8]). Consequently, the annual maxima depend on a range of the parent distribution that is well below its upper tail. For the case of maximum annual daily precipitation, for example, the number of precipitation days in any given year is theoretically bounded to 365 , and so it is far from asymptotic conditions, which assume this number to be much greater than 365 . The number of precipitation days is even smaller than 365, because precipitation is an intermittent process [14-16].

Even if results by Veneziano et al. [13] prevent the asymptotic theory from being strictly applicable to the case of maximum annual daily precipitation, the GEV distribution can still be used to fit extreme precipitation amounts similarly to using any other probability distribution. In this case, however, one should bear in mind that the estimation of GEV parameters will be the result of numerical fit and, as such, it will have no direct link with the statistical properties of the parent distribution. In particular, the shape parameter of the GEV will not represent the shape parameter of the parent distribution. The violation of the asymptotic theory also raises doubts about the extrapolation capability of the fitted GEV model to values (quantiles) associated to high values of probability or return period [10,17]. From this standpoint, an increasing number of studies have recently considered a pool of distributions to model maximum precipitation, including several families of asymptotic- and standard-type (see Table S2 of [10]). Results show that the asymptotic laws often do not yield the best fit according to goodness-of-fit tests and/or statistical indicators.

In order to overcome the asymptotic (or ultimate) assumption in extreme variables, penultimate approximations have been developed [18-21]. For example, Marani and Ignaccolo [22] proposed a penultimate approximation to model the maximum annual daily precipitation, using the Weibull distribution as parent distribution, according to [20,21]. Another promising option would be to analytically derive the exact distribution of annual maxima of daily precipitation, that is, the one that would theoretically be derived from the parent distribution as the distribution of the largest-order statistic (see the next section for details). The word "exact" is used here following the terminology reported in [4], to remark that it is a non-asymptotic distribution.

This option has received little attention so far, as noted by [8]. A first, recurring argument against the exact distribution is that it may not be strictly necessary since a truncated approximation could be enough. A second, and more substantial argument, is that the parent distribution is either not known or difficult to determine in practical cases, as demonstrated by the fact that several distributions have been used in literature to describe the statistical behavior of the non-zero daily precipitation amount. These include both canonical (Exponential, Gamma, Lognormal, generalized Pareto) and more complex ones (mixtures of Exponentials or Lognormals-see [23] for a review). Among these candidate distributions, Gamma has been the most used, especially in weather generators [24].

The argument that the distribution of daily non-zero precipitation is difficult to determine has been recently confuted by increasing evidence, which opened the door to a growing body of literature about the exact distribution. First, Wilson and Toumi [25] provided some physical argumentations in favor of the probability distribution of daily non-zero precipitation being stretched exponential, also known as Weibull distribution. This seminal result is coherent with later findings by [10,17,22,26-28] about both the distribution of the daily amount and the distribution of daily values above a given threshold. Also, Porporato et al. [29] obtained a distribution of daily non-zero precipitation with a stretched exponential tail by compounding an exponential distribution with a parameter having a Gamma distribution. Porporato et al. [29] further motivated the emergence of a Weibull-like tail of daily non-zero precipitation amount based on the interannual variability of the parameter of daily precipitation. More recently, mixture distributions have been used to reproduce daily non-zero precipitation amounts when these amounts are the result of multiple sources/causes [30]. For example, in $[31,32]$ a mixture of a Gamma distribution (for low and moderate values) and a generalized Pareto (for high values) was considered; in [33,34] and then, [25], a mixture of two exponential distributions
was considered; a mixture of an exponential distribution for low to moderate values and a generalized Pareto for high values was used in [35]. Mixtures of two distributions (namely, two Gumbel, or a Gamma and a Gumbel) for the representation of maximum annual value of daily amounts were used in [30].

In this review, we summarize recent advances in the field of extreme daily precipitation distribution, with a focus on recent efforts regarding exact distributions derived as compound-superstatistical distributions. We start by providing some general context about the exact distribution as the largest-order statistic in stationary and non-stationary conditions. We then discuss existing analytical solutions for this exact distribution. We conclude with some directions of future research.

## 2. Some General Results

The exact distribution of the maximum $M=\max \left\{\mathrm{X}_{1}, \mathrm{X}_{2}, \ldots, \mathrm{X}_{n}\right\}$ of $n$ random variables $\left\{\mathrm{X}_{1}, \mathrm{X}_{2}\right.$, $\left.\ldots, X_{n}\right\}$, can be formally derived as the distribution of the largest-order statistic. If $\left\{X_{1}, X_{2}, \ldots, X_{n}\right\}$ are independent and identically distributed (i.i.d.) variables with a common distribution function $F(x)$, also known as a parent distribution, the exact distribution of $M$ is written as

$$
\begin{equation*}
F_{M}(x)=[F(x)]^{n} \tag{1}
\end{equation*}
$$

In the case of maximum annual daily precipitation, $\left\{\mathrm{X}_{1}, \mathrm{X}_{2}, \ldots, \mathrm{X}_{\mathrm{n}}\right\}$ are non-zero daily precipitation variables, $n$ is the counting number of wet days in the year, and $M$ is the maximum annual value (see [4], p. 75, Equation (1); [36], Equation (3.3.45); [37], p. 45, Equation (3.1); [38], p. 10, Equation (1.6); [39], p. 4, Equation (1.16); [40], p. 406, Equation (7.1.1), among others).

A cornerstone of statistical literature, Equation (1) has, however, been obtained under the hypotheses that $n$ is fixed and that $F(x)$ is a-priori known. These hypotheses are not satisfied in the case of maximum annual daily precipitation. First, the distribution $F(x)$ is a-priori unknown. To solve this problem, statistical techniques are used to estimate $F(x)$ from data and substitute it into Equation (1); some examples of these fitting results for daily precipitation are mentioned in the Introduction above. Second, $n$ is not fixed, but it is in fact a realization of a counting random variable $N$. In this case, using the theorem of total probability, the exact distribution can be written as

$$
\begin{equation*}
F_{M}(x)=\sum_{n=0}^{+\infty}[F(x)]^{n} p_{N}(n) \tag{2}
\end{equation*}
$$

where $p_{N}(n)$ is the probability distribution of $N$, denominated also counting distribution-see [4], p. 78, Equation (9); [40], p. 406, Equation (7.1.12); [39], p. 50, Equation (1.107), and [38], p. 29, Equation (1.54).

Equation (2) is the "compound" distribution of the extreme value distribution, $[F(x)]^{n}$, or compound extreme value distribution, because it is the "composition" of $[F(x)]^{n}$ and $p_{N}(n)$. In the literature, Equation (2) is frequently indicated by first reporting the name of the counting distribution after the term "compound", and then the name of the original distribution-here, the extreme value distribution, $[F(x)]^{n}$ ). For example, if the counting distribution is a Poisson distribution, then the name of the compound will be Compound Poisson extreme value distribution (see also Equation (6) below). Usually, the name "compound distributions" is used when the parameter(s) $\theta$ of the distribution is(are) itself(themselves) random variable(s)—indicated with $\boldsymbol{\Theta}$ (see [36], Section 3.5.3); [39], p. 52, Equation (1.113); and [40], Section 3.4.4, p. 151). Equation (2) is also, occasionally, referred to as a "contagious" distribution. The notion of compound, or contagious distribution, was originally proposed by Feller [41]. The term contagious is due to the field of first applications, i.e., entomology and bacteriology.

If the parameter(s) $\theta$ of $F(x)$ is(are) not constant, but instead, is(are) statistically variable, then $F(x)$ becomes $F(x \mid \theta)$, where the symbol | indicates the conditional distribution given a value of parameter(s)
$\Theta=\theta$. Equation (2) will thus provide $F_{M}(x \mid \theta)$, i.e., the conditional distribution of $M$ given $\Theta=\theta$. The unconditional distribution of $M$ is

$$
\begin{gather*}
F_{M}(x)=\sum_{A_{\Theta}} F_{M}(x \mid \theta) p_{\Theta}(\theta)  \tag{3a}\\
F_{M}(x)=\int_{A_{\Theta}} F_{M}(x \mid \theta) f_{\Theta}(\theta) d \theta \tag{3b}
\end{gather*}
$$

depending on whether $\Theta$ is discrete (3a) or continuous (3b), $\boldsymbol{p}_{\boldsymbol{\Theta}}(\theta)$ is the (joint) mass function of $\boldsymbol{\Theta}, f_{\Theta}(\theta)$ is the (joint) density function of $\boldsymbol{\Theta}$, and $A_{\boldsymbol{\Theta}}$ is the domain of $\boldsymbol{\Theta}$. Equation (3) is reported for a generic distribution function $F$ in [36], see Section 3.5.3; [42], vol II, p. 53, Equations (5.1) and (5.2), and [40], p. 152, Equations (3.4.41) and (3.4.42), see also [39], p. 52, Equation (1.113). Note that Equation (2) is a particular case of Equation (3a) where $F_{M}(x \mid \theta)$ is given by Equation (1) and $\Theta=N$, as reported in [40], Example 3.61. Equation (3) can also be written in a compact form as

$$
\begin{equation*}
F_{M}(x)=E\left[F_{M}(x \mid \theta)\right] \tag{4}
\end{equation*}
$$

where $E[$.$] is the expected value calculated with respect to the (joint) distribution of \boldsymbol{\Theta}$ [43].
In physics, the statistical variability of parameters has been referred to as superstatistics [44,45], meaning "statistics of statistics" or superposition of different statistics. From the physical point of view, this characterizes phenomena having variabilities at multiple scales (small and large). For example, the distribution of daily precipitation can have a day-by-day variability (small scales), while the parameters of this distribution can have a year-by-year variability (large scales). Recently, Equation (3) has been referred to as "Metastatistical Extreme Value distribution" [17,22,46]. Since a large amount of literature has referred to Equation (3) as "compound/superstatistical extreme value distributions", we adopt this nomenclature in this review.

Under non-stationary conditions, i.e., when parameters vary with time $t, \boldsymbol{\Theta}(t)$, Equation (3) can be written as:

$$
\begin{gather*}
F_{M}(x \mid t)=\sum_{A_{\Theta}} F_{M}(x \mid \theta(t)) p_{\Theta(t)}(\theta(t))  \tag{5a}\\
F_{M}(x \mid t)=\int_{A_{\Theta}} F_{M}(x \mid \theta(t)) f_{\Theta(t)}(\theta(t)) d \theta(t) \tag{5b}
\end{gather*}
$$

where the distribution of $M$ depends on time. Equation (5) has recently been used by [47] for analyzing the non-stationary behavior of annual maxima of rainfall at different durations.

## 3. Some Analytical Solutions

The practical issue with using Equations (2)-(4), as well as Equation (5), is that analytical solutions of these equations are rarely available. Recently, $[17,22,46]$ tackled this issue via a numerical approach. A downside of numerical approaches is that calculating the inverse distribution and, therefore, quantiles, is not straightforward, which makes analytical solutions particularly appealing.

### 3.1. Analytical Solutions for Equation (2)

A first example of an analytical solution of Equation (2) was derived under the assumption that $N$ is Poisson distributed with distribution $p_{N}(n)=\frac{v^{n} e^{-v}}{n!}$, where the parameter $v$ denotes the mean
number of occurrences in a given time period (here, this period is one year). In this case, the exact distribution of $M$ can be written as
$F_{M}(x)=\sum_{n=0}^{+\infty}[F(x)]^{n} \cdot \frac{v^{n} e^{-v}}{n!}=\sum_{n=0}^{+\infty} \frac{[v F(x)]^{n} e^{-v}}{n!}=e^{-v[1-F(x)]} \sum_{n=0}^{+\infty} \frac{[v F(x)]^{n} e^{-v F(x)}}{n!}=\exp [-v(1-F(x))]$
Equation (6) was originally developed by [48] and then used for the first time in hydrology by [49]. Equation (6) was then reported in many textbooks, such as [40], (p. 440, Equation (7.2.69), [39], p. 51, Equation (1.108), and [38], p. 29, Equation (1.54). According to the terminology reported above, Equation (6) is denominated Compound Poisson Extreme Value distribution.

If the distribution $F(x)$ in Equation (6) is a shifted Exponential, then the distribution of $M$ is a Gumbel (EVI); if the distribution $F(x)$ in Equation (6) is a Pareto, then the distribution of $M$ is a Fréchet (EVII); finally, if the distribution $F(x)$ in Equation (6) is a generalized Pareto, then the distribution of $M$ is a GEV, see [40] (pp. 440-441), [50]. This means that Gumbel, Fréchet, and GEV can be viewed as the exact distributions of Compound Poisson Extreme Value distributions with shifted Exponential, Pareto, and generalized Pareto as parent distributions, respectively.

As a generalization of Equation (6), the product of two compound Poisson Extreme Value distributions in order to characterize ordinary and rare components of extremes (assumed independent from each other) were proposed in [51,52]. $F_{M}(x)$ is then:

$$
\begin{equation*}
F_{M}(x)=\exp \left[-v_{1}\left(1-F_{1}(x)\right)\right] \cdot \exp \left[-v_{2}\left(1-F_{2}(x)\right)\right] \tag{7}
\end{equation*}
$$

The two components have both Poisson chronology with parameters $\boldsymbol{v}_{1}$ and $\boldsymbol{\nu}_{2}$, respectively. The parent distributions are $F_{1}(x)$ and $F_{2}(x)$, respectively. Shifted exponential distributions were used for the parent distributions, i.e., $F_{1}(x)=1-\exp \left(-\frac{x-a_{1}}{b_{1}}\right)$ and $F_{2}(x)=1-\exp \left(-\frac{x-a_{2}}{b_{2}}\right)$. The result is that Equation (7) becomes the following:

$$
\begin{equation*}
F_{M}(x)=\exp \left[-\exp \left(-\frac{x-a_{1}^{\prime}}{b_{1}^{\prime}}-\frac{x-a_{2}^{\prime}}{b_{2}^{\prime}}\right)\right] \tag{8}
\end{equation*}
$$

where $b_{1}^{\prime}=b_{1^{\prime}} a_{1}^{\prime}=a_{1}+b_{1} \ln v_{1}, b_{2}^{\prime}=b_{2^{\prime}} a_{2}^{\prime}=a_{2}+b_{2} \ln v_{2}$. Equation (8) is the product of two Gumbel distributions and is known in the literature as the Two-Component Extreme Value distribution-TCEV, see [53], or [40], p. 443, Equation (7.2.77). The TCEV was formulated for the analysis of extremes of daily precipitation and daily flow (among other variables) and is widely used in Italy to describe the behavior of maximum annual daily precipitation in regionalization studies realized under the VAPI (Valutazione delle Piene in Italia; Flood Evaluation in Italy) project, see e.g., [54,55].

In order to account for the seasonal variability or the non-identical distribution of the parent process in the distribution of annual maxima of precipitation, Revfeim [56] proposed another generalization of Equation (6), using Equation (6) to model the monthly maximum of daily precipitation, following the work of Todorovic and Rousselle [57]. In that case, the distribution of $M$ becomes

$$
\begin{equation*}
F_{M}(x)=\prod_{i=1}^{12} \exp \left[-v_{i}\left(1-F_{i}(x)\right)\right]=\prod_{i=1}^{12} \exp \left[-\exp \left(-\frac{x-a_{i}^{\prime}}{b_{i}^{\prime}}\right)\right] \tag{9}
\end{equation*}
$$

where $\boldsymbol{v}_{i}$ and $\boldsymbol{F}_{\boldsymbol{i}}(\boldsymbol{x})$ are the Poisson parameter and the parent distribution for the $i$-th month, respectively. An exponential distribution was used for $F_{i}(x)=\exp \left(-x / b_{i}\right)$, while $a_{i}^{\prime}=b_{i} \ln v_{i}$. Revfeim [56] considered a harmonic variability of both the two parameters $\left(\boldsymbol{v}_{i}, \boldsymbol{b}_{i}\right.$ ) in order to reduce the number of parameters to determine from $2 \times 12$ to $2 \times 2$, which is the same number of parameters of TCEV. Equation (9) is known as Multi-Component Extreme Value distribution, also indicated as MCEV.

### 3.2. Analytical Solutions for Equation (3)

The analytical solution of Equation (3) is even more challenging and less investigated for extreme values than that of Equation (2). We mention here the theoretical result obtained by Dubey [58], who considered $F_{M}(x \mid \theta)$ as a generalized Gumbel distribution with parameters $a$ and $b$ and an extra parameter $\theta$ :

$$
\begin{equation*}
F_{M}(x \mid \theta)=\exp \left[-\theta b \exp \left(-\frac{x-a}{b}\right)\right] \tag{10}
\end{equation*}
$$

In Equation (10), Dubey [58] assumed the parameter $\boldsymbol{\theta}$ to be a random variable $\boldsymbol{\Theta}$ having a Gamma distribution with density function $f_{\Theta}(\theta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)} \theta^{\alpha-1} \exp (-\beta \theta)$, so the compound distribution of $M$ is

$$
\begin{equation*}
F_{M}(x)=\left[1+\frac{\beta}{b} \exp \left(-\frac{x-a}{b}\right)\right]^{-\alpha} \tag{11}
\end{equation*}
$$

Equation (11) is known as generalized Logistic distribution, or type I generalized Logistic distribution, see [59]. This shows how the generalized Logistic distribution can be viewed as an extreme-type distribution, explaining why it has been successfully used in modeling maximum annual daily precipitation, see Table S2 in [10].

### 3.3. Non-Poisson Analytical Solutions

Based on this literature review, we conclude that the exact distribution of annual maxima has been mostly developed around the Compound Poisson Extreme Value distribution (see Equation (6)). However, this distribution has an often-overlooked weak point: even if Equation (6) is an exact distribution, it still asymptotic characteristics because it is the combination of two distributions, a generic distribution $F$ and a Poisson distribution. The latter is the limit case of a Binomial distribution for a number of trials $\rightarrow+\infty$ and a "small" probability of occurrence for wet days (or of a generic event). Thus, the Poisson distribution is an upper unbounded distribution, which is theoretically not coherent with the upper bounded behavior of the counting variable $N$ under study, i.e., the number of wet days during a given year ( $\leq 365$ ). The Poissonian chronology of wet days due to the clusterization of rain events was questioned in [7].

Recently, De Michele and Avanzi [10] derived the analytical solution of Equation (2) under the assumption that $N$ is Binomially distributed, with distribution $p_{N}(n)=\binom{N_{T}}{n} p_{0}^{n}\left(\mathbf{1}-p_{0}\right)^{N_{T}-n}$, where $N_{T}$ and $p_{0}$ are the total number of days in the year (365) and the probability of dry days, respectively. Note that the Binomial distribution is upper-bounded to $N_{T}$, which somehow overcomes the weak point of the Compound Poisson Extreme Value distribution. Accordingly, the exact distribution of $M$ can be written as

$$
\begin{equation*}
F_{M}(x)=\sum_{n=0}^{+\infty}[F(x)]^{n} \cdot\binom{N_{T}}{n} p_{0}^{n}\left(1-p_{0}\right)^{N_{T}-n}=\left[p_{0}+\left(1-p_{0}\right) F(x)\right]^{N_{T}} \tag{12}
\end{equation*}
$$

In De Michele and Avanzi [10], Equation (12) was derived as a particular case (i.e., zero-order Markov chain) of a first-order Markov chain [10], but proofs of Equation (12) can also be obtained using the Binomial theorem [60]. Similarly to Equation (6), Equation (12) is called Compound Binomial Extreme Value distribution.

A shifted Weibull distribution as parent distribution, $F(x)=\mathbf{1 - \operatorname { e x p }}\left[-\left(\frac{x-x_{T}}{\lambda}\right)^{\beta}\right]$, was used in [10], where $\lambda$ and $\beta$ are, respectively, the scale and shape parameter, and $x_{T}$ is a fixed threshold. Thus, Equation (12) becomes the following:

$$
\begin{equation*}
F_{M}(x)=\left\{p_{0}+\left(1-p_{0}\right)\left[1-\exp \left(-\left(\frac{x-x_{T}}{\lambda}\right)^{\beta}\right)\right]\right\}^{N_{T}} \tag{13}
\end{equation*}
$$

De Michele and Avanzi [10] tested the hypothesis that the daily precipitation amount is Weibull distributed over a large dataset ( 20,561 sites across the world), using threshold values in the range $[0,16] \mathrm{mm}$ and showing that an increase in the threshold rapidly increases the percentage of sites where the hypothesis is accepted to $100 \%$ [10].

If the parent distribution $F(x)$ in Equation (12) was written as a mixture of two distributions, $F_{1}(x)$ and $F_{2}(x)$, in order to account for the different behavior of low and high values [31,32], then Equation (12) could be written as $F(x)=a_{1} F_{1}(x)+a_{2} F_{2}(x)$, with $a_{1}$ and $a_{2}$ being the weights of the two components and $\boldsymbol{a}_{\mathbf{1}}+\boldsymbol{a}_{\mathbf{2}}=\mathbf{1}$ [43]. The distribution of $M$ could then be written as

$$
\begin{equation*}
F_{M}(x)=\left[p_{0}+\left(1-p_{0}\right)\left(a_{1} F_{1}(x)+a_{2} F_{2}(x)\right)\right]^{N_{T}} \tag{14}
\end{equation*}
$$

Equation (14) can be denominated Two-Component Compound Binomial Extreme Value distribution. Equation (14) has the same inspiring principle of the TCEV but uses a Binomial distribution instead of Poisson for the variable $N$, which, as mentioned, overcomes the asymptotic characteristics of the latter. The applicability and performance of Equation (14) is not known at this stage and should be the target of further research.

## 4. Conclusions and Outlook

Maximum annual daily precipitation is a fundamental hydrologic variable. The determination of its probability distribution is a key issue in the design and verification of many engineering works. This task has mainly been addressed using asymptotic distributions, and more recently, through selection from a broader pool of distributions. Less attention has been paid to the development and use of a more rigorous, exact distribution that does not assume asymptotic conditions. The violation of asymptotic conditions for the maximum annual daily precipitation has rekindled interest in this exact distribution, for which the compound/superstatistical distributions play a key role.

Here, we have reviewed existing results to date regarding the exact distribution of maximum annual daily precipitation. The results of our review show that non-asymptotic, exact results already exist, but several of them somehow retain an asymptotic flavor as they are based on a Poisson-occurrence assumption. Some recent results based on a Binomial-occurrence law overcome this limitation and suggest that the way toward an analytical and non-asymptotic definition of maximum-precipitation probability distributions is open. While asymptotic distributions are appealing to use compared to the exact distribution because they do not require information about the underling process, results by several authors show that their predictions are less robust than the exact one beyond the observed values, simply because the latter include information about all precipitation events and not only about maxima [10,17].

We identify several possible directions of future research at this stage. First, we note that model uncertainty associated to quantiles of the exact distribution has been generally overlooked compared to the broad body of literature about quantile uncertainty for the GEV and thus, should be the target of future research. Relatedly, the extrapolation power of the exact distribution compared to asymptotic distributions beyond observed values should be more systematically related to climatic conditions and event types. Third, the statistical variability of the parameters of the exact distribution should be better understood in the wake of potential future changes to climate. Fourth, more investigations of the parent distribution are still needed, possibly considering mixtures of distributions [30-35,61-64].

Funding: This research was partially funded by FONDAZIONE CARIPLO through FLORIMAP project, grant number 2017-0708.

Acknowledgments: The Author is indebted to Demetris Koutsoyannis, Francesco Serinaldi, Simon Papalexiou, Francesco Avanzi, and two anonymous Reviewers, for their comments in the first draft of the manuscript.
Conflicts of Interest: The author declares no conflict of interest.

## References

1. Fréchet, M. Sur la loi de probabilité de l'écart maximum. Ann. Soc. Pol. Math. 1927, 6, 93-116.
2. Fisher, R.A.; Tippet, H.C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc. Camb. Philos. Soc. 1928, 24, 180-190. [CrossRef]
3. Gnedenko, B. Sur la distribution limite du terme maximum d'une serie aleatoire. Ann. Math. 1943, 44, 423-453. [CrossRef]
4. Gumbel, E.J. Statistics of Extremes; Columbia University Press: New York, NY, USA, 1958.
5. von Mises, R. La distribution de la plus grande de n valeurs. In Selected Papers; American Mathematical Society: Providence, RI, USA, 1954; Volume II, pp. 271-294.
6. Jenkinson, A.F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. Q. J. R. Meteorol. Soc. 1955, 81, 158-171. [CrossRef]
7. Coles, S.; Pericchi, L.R.; Sisson, S. A fully probabilistic approach to extreme rainfall modeling. J. Hydrol. 2003, 273, 35-50. [CrossRef]
8. Koutsoyiannis, D. Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. Hydrol. Sci. J. 2004, 49, 576-590. [CrossRef]
9. Wilks, D.S. Comparison of three-parameter probability distributions for representing annual extreme and partial duration precipitation series. Water Resour. Res. 1993, 29, 3543-3549. [CrossRef]
10. De Michele, C.; Avanzi, F. Superstatistical distribution of daily precipitation extremes: A worldwide assessment. Sci. Rep. 2018, 8, 14204. [CrossRef]
11. Koutsoyiannis, D. Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. Hydrol. Sci. J. 2004, 49, 591-610. [CrossRef]
12. Papalexiou, S.M.; Koutsoyannis, D. Battle of extreme value distributions: A global survey on extreme daily rainfall. Water Resour. Res. 2013, 49, 187-201. [CrossRef]
13. Veneziano, D.; Langousis, A.; Lepore, C. New asymptotic and preasymptotic results on rainfall maxima from multifractal theory. Water Resour. Res. 2009, 45, W11421. [CrossRef]
14. De Michele, C.; Ignaccolo, M. New perspectives on rainfall from a discrete view. Hydrol. Process. 2013, 27, 2379-2382. [CrossRef]
15. Trenberth, K.E.; Zhang, Y.; Gehne, M. Intermittency in Precipitation: Duration, Frequency, Intensity, and Amounts Using Hourly Data. J. Hydrometeorol. 2017, 18, 1393-1412. [CrossRef]
16. Bernardara, P.; De Michele, C.; Rosso, R. A simple model of rain in time: An alternating renewal process of wet and dry states with a fractional (non-Gaussian) rain intensity. Atmos. Res. 2007, 84, 291-301. [CrossRef]
17. Zorzetto, E.; Botter, G.; Marani, M. On the emergence of rainfall extremes from ordinary events. Geophys. Res. Lett. 2016, 43, 8076-8082. [CrossRef]
18. Cohen, J.P. The penultimate form of approximation to normal extremes. Adv. Appl. Probab. 1982, 14, 324-339. [CrossRef]
19. Gomes, M.I. Penultimate limiting forms in extreme value theory. Ann. Inst. Stat. Math. 1984, 236, 71-85. [CrossRef]
20. Harris, R. Extreme value analysis of epoch maxima-Convergence, and choice of asymptote. J. Wind Eng. Ind. Aerodyn. 2004, 92, 897-918. [CrossRef]
21. Cook, N.J.; Harris, R.I. Exact and general ft1 penultimate distributions of extreme wind speeds drawn from tail-equivalent Weibull parents. Struct. Saf. 2004, 26, 391-420. [CrossRef]
22. Marani, M.; Ignaccolo, M. A metastatistical approach to rainfall extremes. Adv. Water Resour. 2015, 79, 121-126. [CrossRef]
23. Papalexiou, S.M.; Koutsoyannis, D. A global survey on the seasonal variation of the marginal distribution of daily precipitation. Adv. Water Resour. 2016, 94, 131-145. [CrossRef]
24. Wilks, D.S.; Wilby, R.L. The weather generation game: A review of stochastic weather models. Prog. Phys. Geogr. 1999, 23, 329-357. [CrossRef]
25. Wilson, P.S.; Toumi, R. A fundamental probability distribution for heavy rainfall. Geophys. Res. Lett. 2005, 32, L14812-1-4. [CrossRef]
26. Papalexiou, S.M.; Koutsoyiannis, D. Entropy based derivation of probability distributions: A case study to daily rainfall. Adv. Water Resour. 2012, 45, 51-57. [CrossRef]
27. Papalexiou, S.M.; Koutsoyiannis, D.; Makropoulos, C. How extreme is extreme? An assessment of daily rainfall distribution tails. Hydrol. Earth Syst. Sci. 2013, 17, 851-862. [CrossRef]
28. Serinaldi, F.; Kilsby, C.G. Rainfall extremes: Toward reconciliation after the battle of distributions. Water Resour. Res. 2014, 50, 336-352. [CrossRef]
29. Porporato, A.; Vico, G.; Fay, P.A. Superstatistics of hydro-climatic fluctuations and interannual ecosystem productivity. Geophys. Res. Lett. 2006, 33, L15402. [CrossRef]
30. Shin, J.-Y.; Lee, T.; Ouarda, T.B.M. Heterogeneous mixture distributions for modeling multisource extreme rainfalls. J. Hydrometeorol. 2015, 16, 2639-2657. [CrossRef]
31. Furrer, E.M.; Katz, R.W. Improving the simulation of extreme precipitation events by stochastic weather generators. Water Resour. Res. 2008, 44, W12439. [CrossRef]
32. Hundecha, Y.; Pahlow, M.; Schumann, A. Modeling of daily precipitation at multiple locations using a mixture of distributions to characterize the extremes. Water Resour. Res. 2009, 45, W12412. [CrossRef]
33. Johnson, G.L.; Hanson, C.L.; Hardegree, S.P.; Ballard, E.B. Stochastic weather simulation: Overview and analysis of two commonly used models. J. Appl. Meteorol. 1996, 35, 1878-1896. [CrossRef]
34. Wilks, D.S. Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. Agric. For. Meteorol. 1999, 93, 153-169. [CrossRef]
35. Cameron, D.; Beven, K.; Tawn, J. Modelling extreme rainfalls using a modified pulse Barlett-Leweis stochastic rainfall model (with uncertainty). Adv. Water Resour. 2001, 24, 203-211. [CrossRef]
36. Benjamin, J.R.; Cornell, C.A. Probability, Statistics, and Decision for Civil Engineers; McGraw-Hill: New York, NY, USA, 1970.
37. Coles, S. An Introduction to Statistical Modeling of Extreme Values; Springer: London, UK, 2001.
38. Reiss, R.D.; Thomas, M. Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and other Fields, 3rd ed.; Birkhauser: Berlin, Germany, 2007.
39. Salvadori, G.; De Michele, C.; Kottegoda, N.T.; Rosso, R. Extremes in Nature: An Approach Using Copulas; Springer: Berlin, Germany, 2007.
40. Kottegoda, N.T.; Rosso, R. Applied Statistics for Civil and Environmental Engineers; Blackwell: Oxford, UK, 2008.
41. Feller, W. On a general class of "contagious" distributions. Ann. Math. Stat. 1943, 14, 389-400. [CrossRef]
42. Feller, W. An Introduction to Probability Theory and Its Applications; Wiley: New York, NY, USA, 1971; Volume II.
43. Johnson, N.L.; Kotz, S.; Kemp, A.W. Univariate Discrete Distributions, 2nd ed.; Wiley: New York, NY, USA, 1992.
44. Beck, C.; Cohen, E.G.D. Superstatistics. Physica A 2003, 322, 267-275. [CrossRef]
45. Beck, C. Superstatistics: Theory and applications. Contin. Mech. Thermodyn. 2004, 16, 293-304. [CrossRef]
46. Marra, F.; Nikolopoulos, E.I.; Anagnostou, E.N.; Morin, E. Metastatistical extreme value analysis of hourly rainfall from short records: Estimation of high quantiles and impact of measurement errors. Adv. Water Resour. 2018, 117, 27-39. [CrossRef]
47. De Luca, D.L.; Galasso, L. Stationary and non-stationary frameworks for extreme rainfall time series in southern Italy. Water 2018, 10, 1477. [CrossRef]
48. Epstein, B. A modified extreme value problem. Ann. Math. Stat. 1949, 20, 99-103. [CrossRef]
49. Todorovic, P.; Zelenhasic, E. A stochastic model for flood analysis. Water Resour. Res. 1970, 6, 1641-1648. [CrossRef]
50. van Monfort, M.A.J.; van Putten, B. A comment on modelling extremes: Links between multi-component extreme value and generalized extreme value distribution. J. Hydrol. 2002, 41, 197-202.
51. Rossi, F. Legami di connessione tra le distribuzioni di probabilità delle massime piogge e delle massime piene in una regione. In Proceedings of the Seminar on Extreme Hydrological Events: Floods and Droughts, Erice, Italy, 22-25 March 1980.
52. Rossi, F.; Versace, P. Criteri e Metodi per L'analisi Statistica delle Piene; CNR pubs 165; CNR: Roma, Italy, 1982.
53. Rossi, F.; Fiorentino, M.; Versace, P. Two-component extreme value distribution for flood frequency analysis. Water Resour. Res. 1984, 20, 847-856. [CrossRef]
54. Cannarozzo, M.; D'Asaro, F.; Ferro, V. Regional rainfall and flood frequency analysis for Sicily using the two-component extreme value distribution. Hydrol. Sci. J. 1995, 40, 19-42. [CrossRef]
55. Ferro, V.; Porto, P. Regional analysis of rainfall-depth duration equation for south Italy. J. Hydrol. Eng. 1997, 4, 326-336. [CrossRef]
56. Revfeim, K.J.A. Annual maxima and totals of seasonally varying processes. Stoch. Hydrol. Hydraul. 1991, 5, 147-153. [CrossRef]
57. Todorovic, P.; Rousselle, J. Some Problems of Flood Analysis. Water Resour. Res. 1971, 7, 1144-1150. [CrossRef]
58. Dubey, S.D. A new derivation of the logistic distribution. Nav. Res. Logist. Q. 1969, 16, 37-40. [CrossRef]
59. Johnson, N.L.; Kotz, S.; Balakrishnan, N. Continuous Univariate Distributions, 2nd ed.; Wiley: New York, NY, USA, 1995; Volume 2.
60. Marani, M.; Zorzetto, E. Doubly stochastic distributions of extreme events. arXiv 2019, arXiv:1902.09862.
61. Vrac, M.; Naveau, P. Stochastic downscaling of precipitation: From dry events to heavy rainfalls. Water Resour. Res. 2007, 43, W07402. [CrossRef]
62. Carreau, J.; Bengio, Y. A hybrid Pareto model for asymmetric fat-tailed data: The univariate case. Extremes 2009, 12, 53-76. [CrossRef]
63. Carreau, J.; Vrac, M. Stochastic downscaling of precipitation with neural network conditional mixture models. Water Resour. Res. 2011, 47, W10502. [CrossRef]
64. Li, C.;Singh, V.P.; Mishra, A.K. Simulation of the entire range of daily precipitation using a hybrid probability distribution. Water Resour. Res. 2012, 48, W03521. [CrossRef]
© 2019 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access
