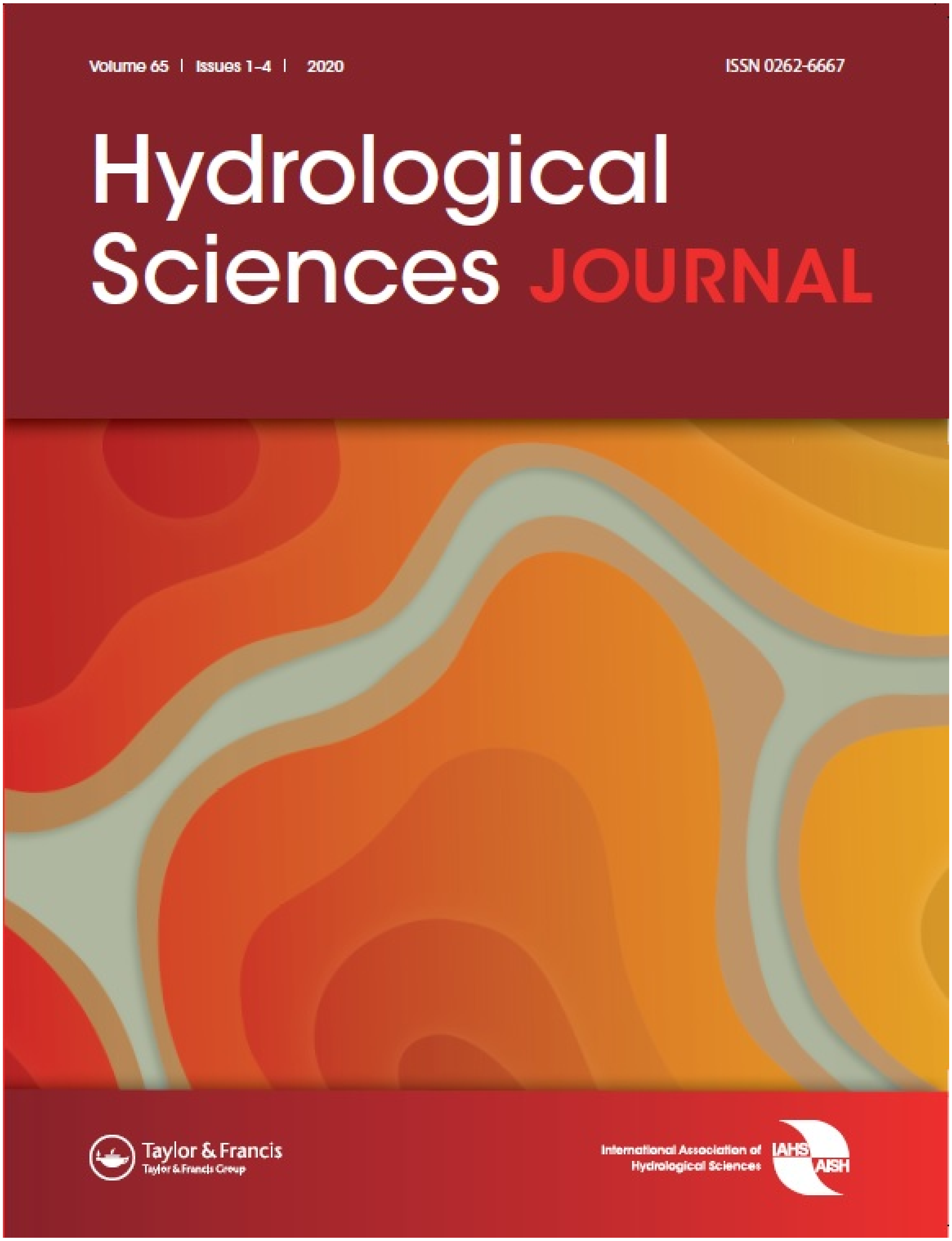

A stochastic framework for rainfall intensity–timescale–return period relationships. Part I: Theory and estimation

URL: http://mc.manuscriptcentral.com/hsj

strategies

Journal:	*Hydrological Sciences Journal*
Manuscript ID	HSJ-2023-0523.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	18-Mar-2024
Complete List of Authors:	Koutsoyiannis, Demetris; National Technical University of Athens, Iliopoulou, Theano; National Technical University of Athens, Water Resources and Environmental Engineering Koukouvinos, Antonis; National Technical University of Athens Malamos, Nikolaos; University of Patras, Department of Agriculture
Keywords:	

A stochastic framework for rainfall intensity–timescale–return period relationships. Part I: Theory and estimation strategies

[1]Demetris Koutsoyiannis, [1*]Theano Iliopoulou, [1]Antonis Koukouvinos, [2]Nikolaos Malamos

[1]Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Heroon Polytechneiou 5, GR-157 72 Zographou, Greece

[2]Department of Agriculture, University of Patras, GR-30200 Messolonghi, Greece; nmalamos@upatras.gr

* Corresponding author. *E-mail address:* tiliopoulou@hydro.ntua.gr

Abstract

This work presents a stochastic framework for the construction of rainfall intensity–timescale–return period relationships, which was applied in the recent regionalization of design rainfall curves over the Greek territory, described in a companion paper. The methodology outlined herein builds upon the Koutsoyiannis' et al. 1998 framework, which has been recently revisited and upgraded, and incorporates two different versions: (a) a theoretically consistent stochastic model applicable for rainfall intensity over any scale of interest and (b) a simplified version valid over small scales, which makes parameter estimation easier. Special focus is given to the presentation of the simplified version, which suffices for most engineering tasks. Parameter estimation approaches are presented in detail including the K-moments framework that allows for reliable high-order moment estimation and handling of bias due to spatiotemporal dependence.

Keywords: rainfall modelling; extreme rainfall; design rainfall; intensity–duration–frequency curves; tail-index; K-moments; stochastic modelling

1. Introduction

Accurate rainfall estimates for various timescales and return periods are one of the most important prerequisites for hydrological and hydraulic design. Engineers typically have access to this information using mathematical relationships that relate the time-averaged rainfall intensity (x) over a given time scale (k) to the return period (T). The latter terms are usually referred to as "duration" and "frequency" in the literature, i.e., forming the term 'intensity–duration–frequency' curves, which however, induces ambiguity in terminology as "duration" can be confounded with actual rainfall duration, while "frequency" is related to but not synonymous to the return period. Here, to avoid confusion, we use the term "ombrian relationships" (or curves), from the Greek "όμβρος" (meaning rainfall), that has been also used in several past studies (Koutsoyiannis et al. 2023a, Koutsoyiannis and Iliopoulou 2022; Iliopoulou et al. 2022).

There are several different types of ombrian relationships in the literature (cf. Svensson and Jones, 2010, Lanciotti et al. 2023 for extensive reviews), ranging from early empirical approaches (Sherman 1931, Bernard, 1932) to generalized parametric approaches, such as the probabilistic approach formulated by Koutsoyiannis et al. (1998) and the approach based on the rainfall simple- or multi-scaling assumption (Burlando and Rosso 1996, Langoussis and Veneziano 2007, Veneziano and Fucolo 2002), as well as other regression-type ('data-driven') approaches (Overeem 2008, Haruna et al. 2023). Most of these methods are either empirical or at least include some empirical relationships that have been established through long-term hydrological experience. Despite certain theoretical shortcomings owing to the empirical derivations (see Koutsoyiannis 2023a, p.290), most approaches address the problem for design for small scales, albeit with different performance being reported worldwide (Shehu et al. 2023, Lanciotti et al. 2023). Meanwhile, however, attempts to provide ombrian relationships with a theoretical foundation have often employed inappropriate assumptions, leading to relationships that are oversimplified and unsuitable for engineering application.

This work presents a stochastic framework of ombrian relationships that forms the theoretical background for the recent regionalization of ombrian relationships in Greece (Koutsoyiannis, 2023c), which is the focus of a companion paper (Iliopoulou et al. 2023). The stochastic framework presented herein constitutes an advance of the approach formulated by Koutsoyiannis et al. (1998) which is one of the standard approaches in hydrological practice (e.g., Sane et al. 2018, Shehu et al. 2022, Lanciotti et al. 2022). The new upgraded framework has been developed in two variants (Koutsoyiannis, 2023a; Chapter 8): (a) a theoretically consistent stochastic modelling framework of rainfall intensity, valid over any time scale of interest, and (b) a simplified version applicable over small time scales, e.g., of the order of minutes to a few days.

A key motivation for deriving the ombrian relationships under a theoretically consistent stochastic framework is the ability to retain multi-scale validity and thus, achieve increased modelling efficiency. In particular, the ombrian curves are typically constructed for time intervals ranging from a few minutes to several hours, as this range of timescales is most relevant for common engineering applications. However, having a multi-scale model of rainfall intensity is preferrable since larger temporal scales are important for advanced engineering operations, such as reservoir operation and water management plans, while the model can also be applied in simulation (Koutsoyiannis and Dimitriadis 2021). The problem with this version is that it requires long and complete fine-scale timeseries of the parent process for its determination and such rainfall data are often not available at large regional scales. Also, the parameter estimation procedure becomes complicated if the regionalization of parameter values is sought. In this respect, it is also important to obtain simplified versions of such a model that can be applied in practice with data requirements as common as the ones employed in traditional modelling of ombrian curves. Indeed, in many regions, including Greece, long and complete timeseries of fine-scale rainfall data are sparse and thus, advanced modelling based on the first version cannot be applied at the regional scale (Iliopoulou et al. 2023). For such cases, the simplified version is developed and detailed herein. It is important to note that the simplified version is obtained from the full version under stated assumptions, and therefore it inherits the theoretical consistency of the full model, except in the aspects in which the simplifying assumptions apply.

A second reason for seeking theoretical consistency in modelling of ombrian curves is to make estimation from data with awareness of the involved bias and uncertainty, the quantification of which requires a theoretical model. Even more, it is well-established that in a stochastic process characterized by persistence, both bias and uncertainty increase significantly (Dimitriadis and Koutsoyiannis 2015, Koutsoyiannis 2023a, Iliopoulou and Koutsoyiannis 2019), and this is the case for the rainfall process as well, including its extremes (Iliopoulou et al. 2018, Dimitriadis et al. 2021, Koutsoyiannis 2023a, Iliopoulou and Koutsoyiannis 2019, O’Connell et al. 2023).

To guide empirical estimation, the paper presents in detail efficient estimation methods for all involved parameters. Emphasis is given to a new estimation framework for the estimation of the distribution parameters, including the highly uncertain tail-index, namely the knowable moments (K-moments) approach (Koutsoyiannis 2019, 2023b). This set of moments presents various advantages over classical approaches (Koutsoyiannis 2023b) the most relevant of which for our study are that high-moment orders, related to the tail-behaviour, can be reliably estimated and assigned a return period, whereas estimation can also be adapted for bias due to spatio-temporal dependence, which is common in the analysis of rainfall data. Along with the theoretical development of the simplified model, these estimation approaches complement the theoretical background on which the point modelling and subsequent regionalization of ombrian curves for Greece is based (Iliopoulou et al. 2023).

The remaining of this work is structured in three sections. Section 2 deals with the presentation of the stochastic model detailing the requirements for theoretical consistency and introducing the two variants, with emphasis placed on the simplified version for small scales. Section 3 describes the approaches proposed for parameter estimation, including the K-moments return period estimation framework. Summary and conclusions are reported in Section 4.

2. A stochastic framework for multi-scale rainfall intensity modelling

2.1 Requirements for theoretical consistency

All ombrian relationships are based on some assumptions, which herein are intended to be theoretically consistent with key stochastic properties for the rainfall process, under a stationary framework, which is consistent with long-term stochastic changes (Koutsoyiannis and Montanari 2015). Yet as previously explained, due to the practical/engineering nature of ombrian curves, it is reasonable to sacrifice perfect theoretical consistency if it results in too involved expressions. A set of assumptions for theoretical consistency that are both practical and consistent is proposed by Koutsoyiannis (2023a, p.290), as follows:

1. A basic requirement of any stochastic model is to preserve first and second-order characteristics of a process of interest, which in this case, is the temporal average intensity $\underline{x}^{(k)}$ over any time scale k. For handling the second-order properties, it is convenient to employ the variance of the averaged process over timescale of averaging $\operatorname{var}\left[\underline{x}^{(k)}\right]$, i.e., the climacogram, the preservation of which ensures preservation of any other second-order

characteristic (Koutsoyiannis 2023a, Dimitriadis and Koutsoyiannis 2015). The need for preserving a constant mean is obvious, even though this requirement is violated in common expressions of ombrian curves.

2. The process variance should be finite for $k \to 0$, for physical consistency relating to required energy (which would otherwise be infinite), and zero for $k \to \infty$, in order for the process to be ergodic.
3. The model should account for the fact that the probability dry, $P_0^{(k)} := P\{\underline{x}^{(k)} = 0\}$ is nonzero for small time scales, and equivalently, that the probability wet, $P_1^{(k)} := P\{\underline{x}^{(k)} > 0\} = 1 - P_0^{(k)}$ is smaller than 1 for small k, including for $k \to 0$.
4. Moments of order greater than two are significant to consider since an ombrian model places a strong emphasis on the rainfall extremes.
5. The upper-tail index of the distribution should be constant for all time scales. Theoretical justification of this requirement is given in Koutsoyiannis (2023a, p.318).
6. The model should handle the all-scale rainfall distribution. In this respect, the Pareto distribution is an optimal choice for short time scales due to its simplicity and clear relationship between the time-averaged intensity and return period (the reader is referred to the Appendix for a detailed presentation). However, as the time scale extends to several days or longer, the Pareto distribution becomes insufficient, and a bell-shaped type of probability density, as the Pareto-Burr-Feller distribution (Koutsoyiannis 2023a, p.291) is more appropriate.

2.2 Ombrian model for small scales

The mathematical framework for rainfall curves proposed by Koutsoyiannis et al. (1998) is one of the most widely used approaches for design rainfall (e.g., Sane et al. 2018, Shehu et al. 2022, Lanciotti et al. 2022). Recently, Koutsoyiannis (2023a, Chapter 8) revisited this framework developing an approach that generalizes typical ombrian curves applicable over small scales, to stochastic models of rainfall intensity, valid over any scale (arbitrarily large) supported by the data. The revisited model preserves all requirements discussed in Section 2.1. However, for large time scales the mathematics becomes somewhat involved while these scales are less relevant to typical applications. Here, we apply the framework only for small time scales, for which a Pareto distribution for the non-zero rainfall intensity is justified. As already mentioned, for larger scales, this should be replaced by a Pareto-Burr-Feller distribution. For the former case, the Pareto distribution quantile is given as (Koutsoyiannis 2023a, p.294):

$$x = \lambda(k)\frac{\left(P_1^{(k)}T/k\right)^{\xi} - 1}{\xi} \tag{1}$$

where $P_1^{(k)}$ is the probability wet, $\lambda(k)$ is a scale parameter and ξ is the tail-index of the Pareto distribution. Both $P_1^{(k)}$ and $\lambda(k)$ are functions of the timescale obtained as (related derivations are given in the Appendix):

$$P_1^{(k)} = \frac{1-\xi}{1/2-\xi}\frac{\mu^2}{\gamma(k)+\mu^2} \tag{22}$$

$$\lambda(k) = \frac{\mu(1-\xi)}{P_1^{(k)}} = \frac{(1/2-\xi)\left(\gamma(k)+\mu^2\right)}{\mu} \tag{3}$$

where μ is the mean intensity (constant at all time scales) and $\gamma(k)$ the climacogram of the process, i.e., the function of the variance across timescale, which can follow different models.

This model respects the requirements set for small scales, yet its application requires estimation of stochastic properties (e.g., the climacogram) based on the parent process, which in this case should be based on complete fine-scale timeseries of the rainfall intensity. Such series are often sparse, and this may hinder reliable estimation of the ombrian model in regional analyses, as is the case of Greece (Iliopoulou et al. 2023). For such cases, a simplification of the model with less intensive data requirements, i.e., employing solely series of extremes (block maxima or values over threshold), is possible as detailed next.

2.3 Simplified version

Based on the following assumptions which are reasonable for the fine-scale behaviour of rainfall, Equations (1)–(3) may be simplified. For small time scales:

- We assume that $P_1^{(k)} \propto k$, and hence we can set the quantity $\beta(k) := k/P_1^{(k)} = \beta =$ constant in Equation (1). This is a strong assumption, yet necessary to derive the simplified version of the model, and for this reason, it is usually implicitly adopted in most simplified ombrian relationships (cf. Koutsoyiannis 2023a, p.290-291, p.299). The assumption can only hold as long as $k < \beta$ (otherwise $P_1^{(k)} > 1$ would result). If larger timescales are of interest, then the all-scale model version (Koutsoyiannis 2023a, p.291-295) should be used.

- $\gamma(k) \gg \mu^2$, and thus, we can neglect the latter term in their sum.
- The empirical climacogram may be modelled through the generalized Cauchy-type model climacogram (Koutsoyiannis, 2023; p. 113):

$$\gamma(k) = \lambda_1^2 \left(1 + \left(\frac{k}{\alpha}\right)^{2M}\right)^{\frac{H-1}{M}} \tag{4}$$

where α and λ_1 are scale parameters, with dimensions of time $[t]$ and $[x]$, respectively, and H, M are dimensionless parameters in the interval (0,1), controlling the long-range, i.e., Hurst-Kolmogorov (HK) dynamics, and local scaling of the process (fractal behaviour) of the process, respectively. For M we take the neutral value $M = 1/2$ as default.

These simplifying assumptions, result in some violations of a full stochastic consistency, as detailed in Koutsoyiannis (2023a, p.299). However, at small scales, of the order of minutes to a few days the violations are negligible. By virtue of these simplifications, the ombrian relationship is given as:

$$x = \lambda_1^2 \frac{(1/2 - \xi)}{\xi\mu} \left(1 + \frac{k}{\alpha}\right)^{2H-2} \left(\left(\frac{T}{\beta}\right)^{\xi} - 1\right) \tag{5}$$

It is easily observed that Equation (5) can be written concisely as the quotient of two separable functions $b(T)$ and $a(k)$ of the return period and the timescale, respectively, in the form:

$$x = \frac{b(T)}{a(k)} \tag{6}$$

From Equation (5) it follows that the function $a(k)$ has the following general form:

$$a(k) = \left(1 + \frac{k}{\alpha}\right)^{\eta}, \ \eta := 2 - 2H \tag{7}$$

where α and η are parameters to be estimated from the data with $\alpha > 0$ (in units of time, e.g., h) and $0 < \eta < 1$ (dimensionless). Accordingly, assuming $\xi > 0$ and setting $\lambda = (1/2 - \xi)\lambda_1^2 / \xi\mu$ the function $b(T)$ is:

$$b(T) = \lambda\left(\left({}^{T}/\beta\right)^{\xi} - 1\right), \ \xi > 0 \tag{8}$$

Therefore, Equation (6) takes the following final form, for the usual case where $\xi > 0$:

$$x = \lambda \frac{\left({T}/{\beta}\right)^{\xi} - 1}{\left(1 + {k}/{\alpha}\right)^{\eta}} \tag{9}$$

where the following five parameters are involved: λ an intensity scale parameter in units of the rainfall intensity x (e.g. mm/h), β a timescale parameter in units of the return period (e.g. years), α a timescale parameter in units of timescale (e.g. h) with $\alpha > 0$, η a dimensionless parameter with $0 < \eta < 1$, and $\xi > 0$ the upper tail index of the process.

In the case that the return period of the rainfall intensity is empirically determined based on rainfall exceedances extracted from the full series, a Pareto distribution can be generally assumed for modelling the rainfall intensity, as implied by Equation (9). However, if the return period is determined based on series of annual maxima (AM) of rainfall intensity, then both long-term empirical evidence and theoretical arguments support the use of the Extreme Value Type 2 (EV2) distribution from the Generalized Extreme Value (GEV) distribution family:

$$F(y) = \exp\left(-\left(1 + \xi\left(\frac{y}{\nu} - \psi\right)\right)^{-\frac{1}{\xi}}\right), \; y \geq \nu\left(\psi - \frac{1}{\xi}\right) \tag{10}$$

where ψ (dimensionless), $\nu > 0$ (units same as in y) and $\xi > 0$ (dimensionless) are location, scale and shape parameters, respectively. It should be mentioned that the case of $\xi < 0$ is not appropriate for maximum rainfall, since it presumes the existence of an upper limit for the variable, which is inconsistent to the physical reality. Also, the case of ξ=0, i.e., assuming a Gumbel (Extreme Value Type 1—EV1) distribution for the maximum rainfall intensity, is also not supported by worldwide empirical evidence and is to be avoided in general (Koutsoyiannis 2004). Therefore, it is not developed herein, but further details for this case are given in Koutsoyiannis (2023a).

Equivalently, the EV2 distribution as given by Equation (10) can be re-parameterized consistently to Equation (9) as follows:

$$F(y) = \exp\left(-\frac{\Delta}{\beta}\left(\frac{y}{\lambda} + 1\right)^{-\frac{1}{\xi}}\right) \tag{11}$$

where $\Delta = 1$ year, $\beta = (1 - \xi\,\psi)^{1/\xi}\Delta$ and $\lambda = (1 - \xi\,\psi)\,\nu\,/\,\xi$ and $\xi > 0$.

The variable y represents either the rainfall intensity x or, equivalently, the product $x\,a(k)$ (Equation (6)). Solving Equation (11) in terms of y and substituting $F(y) = 1 - \varDelta\,/\,T$, where $\varDelta = 1$ year for annual maxima, yields, respectively:

$$x = \lambda \frac{\left(-\left(\beta/\varDelta\right)\ln\left(1 - {\varDelta}/{T}\right)\right)^{-\xi} - 1}{\left(1 + k/\alpha\right)^{\eta}},\ \xi > 0 \tag{12}$$

and therefore, in this case the function $b(T)$ is:

$$b(T) = \lambda\Big(\left(-\left(\beta/\varDelta\right)\ln\left(1 - {\varDelta}/{T}\right)\right)^{-\xi} - 1\Big),\ \xi > 0 \tag{13}$$

It is easily shown that for small return periods, Equation (9) deriving from a Pareto distribution yields higher intensity than Equation (12) whereas for larger return periods ($T > 10$ years) the two are practically indistinguishable given that for small $\varDelta/T$ holds $\ln\left(1 - (\varDelta/T)\right) = -(\varDelta/T) - (\varDelta/T)^2 -$ ☹ $\approx -\varDelta/T$. Therefore, from an engineering perspective, it is safer to express the final model as Equation (9) even when the fitting is based on Equation (14), i.e., in the case that annual maxima are employed. Therefore, Equation (9) is the final design relationship, as it is in full correspondence with the natural rainfall process (without reference to the subjective choice of the yearly time scale for the maxima extraction), has a simpler mathematical description, and its validity covers return periods also smaller than 1 year. Equations (9) and (12) were also presented by Koutsoyiannis et al. (1998) for small scales, albeit with a slightly different parameterization, and without links to the multi-scale stochastic model. Despite this, they are adequate for most engineering applications of ombrian curves, namely those involving flood analyses. A summary of the simplified relationships along with the parameters is provided in Table 1.

An attractive advantage of this simplified version is the separability of functions $a(k)$ and $b(T)$ that allows for an independent, two-step procedure of parameter estimation. This is useful in practice and even more for regional analyses where different data sources may be available. In particular, the estimation of the parameters of the timescale function (of the expression $a(k)$) requires the use of sub-daily or even sub-hourly data, available from tipping-buckets and automated censors. On the other hand, the estimation of the distribution parameters (of the expression $b(T)$) may be performed by also exploiting daily rainfall records which are characterized by longer lengths and a denser spatial resolution while they are usually more reliable in recording heavy rainfall during storm events (e.g., Molini et al. 2005).

2.4 Physical/mathematical basis of the parameters

The simplified Equations (9) and (12) are dimensionally consistent, and the five parameters have physical or logical meaning, as explained below following Koutsoyiannis (2023a, p. 297). It should be noted, however, that the procedure for parameter estimation is based on minimizing an error expression rather than on their actual meaning. Therefore, the connection between the parameter values and their meanings is imperfect, it nonetheless allows us to comprehend the full theoretical framework.

- η [−]: Persistence parameter, where larger values indicate less strong persistence. It is asymptotically connected to the Hurst parameter H, with the relationship $\eta = 2 - 2H$. For a purely random process, $H = 0.5$ and $\eta = 1$, a value that is the upper allowable limit of η, but certainly not supported by empirical evidence. Clearly, any value of $\eta < 1$ results in $H > 0.5$, i.e., a process with persistence. In a fully persistent process, $H = 1$ and $\eta = 0$, a value that is the lower allowable limit of η. For $H = 0.75$, $\eta = 0.5$, which is a typical value of η.
- α [T]: Time scale parameter, expressing the rate of deviation of the term $A \coloneqq 1/(1 + k/\alpha)^\eta$ from the pure power law $B \coloneqq (\alpha/k)^\eta$. For time scale $k \gg \alpha$, A and B are practically identical. For $k = \alpha$, the A term already deviates quite a bit (by 1/3 to 1/2) from the power law. For $k \to 0$ (instantaneous time scale), $A = 1$, while $B \to \infty$. For $\alpha \to 0$, A and B tend to coincide, but the rainfall intensity tends to infinity. For that reason, the value $\alpha = 0$ should be excluded. Typical values are close to (Koutsoyiannis et al. 2023c) $\alpha = 0.2$ h, while, for a set of global rainfall records, Koutsoyiannis and Papalexiou (2017) suggested α =0.07 h.
- ξ [−]: Upper-tail index of the distribution of rainfall depth or intensity. Its minimum value, $\xi = 0$, corresponds to an exponential distribution (or Gumbel distribution for annual maximum rainfall). Values of $\xi > 0$ correspond to a Pareto distribution (or Fréchet distribution for annual maximum rainfall). For better understanding of the meaning of the parameter ξ it is noted that, when $\xi > 0$, the classical moments of the distribution are finite only for order $p < 1/\xi$, while for $p > 1/\xi$ they diverge to infinity. Therefore, values of $\xi \geq 1$ correspond to an infinite mean of the rainfall depth or intensity, which has no physical meaning. Values $\xi \geq 1/2$ are not considered admissible because they make the variance ($p = 2$) infinite. Values $\xi \geq 1/3$ and $\xi \geq 1/4$ result in infinite skewness ($p = 3$) and kurtosis ($p = 4$), respectively. Typical values range

from ξ = 0.1 to 0.2 (Koutsoyiannis et al. 2023c), while global investigations of precipitation extremes have given ξ = 0.13 to 0.15 (Koutsoyiannis 1999, 2004b). All these empirically estimated values suggest finite mean, variance, and classical skewness and kurtosis of the distribution.

- β [T]: Scale parameter for return period, expressing the average temporal distance of two consecutive wet periods (e.g., days). It is recalled that the simplified ombrian expression is based on the assumption that the ratio of the time scale k to the probability wet at scale k, $P_1(k)$ is constant, equal to β, i.e., $\beta = k/P_1(k)$. Considering k = 1 d, we find β = 1 d/P_1(1 d) = N/ν d, where $N \approx 365$ is the number of days in a year and ν is the average number of wet days in a year. Thus, the ratio N/ν is the average time interval between two wet days. If it rains every day, then $\nu \approx 365$ and the average distance between two wet days is β = 1 d. If it rains 20% of the days, then β = 1/0.2 = 5 d =0.0137 years. Since the rain depth has a lower bound of 0, if we set $T = \beta$, then the ombrian equation should yield $x = 0$, which is indeed the case. Values $T < \beta$ are meaningless. Likewise, time scales $k > \beta$ cannot be modelled by the simplified equations.
- λ [LT^{-1}]: Characteristic instantaneous rainfall intensity (scale parameter), roughly corresponding to a one-year return period (T = 1 year). Indeed, for k = 0, for typical values ξ = 0.15, β = 4 d (cf. the explanation of parameters ξ and β above), and for T = 1 year = 365 d, it follows $(T/\beta)^{\xi} = (365/4)^{0.15} \approx 2$ and thus, from equation (9), x(0, 1 year) = λ.

3. Estimation of ombrian parameters

3.1 Time-averaging of rainfall intensity

The ombrian relationships describe the probabilistic behaviour of the time-averaged rainfall intensity $x^{(k)}$ over any scale k of interest. Therefore, as in all studies investigating a process at many scales, the first step is to aggregate the available data from several time series to different time scales. The aggregated series are formed with no specific provision for the starting point for aggregation of each original time series. For instance, if the original series is a daily time series x_τ we can construct the 2-day time series $x_\tau^{(2)}$, in two different ways depending on the selection we make for the first term. Namely, the $x_1^{(2)}$ that contains the daily term x_2 could be either $(x_1 + x_2)/2$ or $(x_2 + x_3)/2$. Likewise, if we construct a time series at time scale 10, there are 10 variants (the first term $x_1^{(10)}$ that contains the daily term x_{10} could be anyone among $(x_1 + ... + x_{10})/10$ through $(x_{10} + ... + x_{19})/10)$. These are all numerically distinct time series, yet their statistical properties are same. Following a stochastic approach, all these realizations are equivalent since we are interested in statistical properties and not the time series values per se. As a result, since we are using a stochastic approach, there is no need to apply a sliding window and take the maximum value among the variants for our investigation, although this has been a common practice in the studies of hydrological extremes (e.g., Linsley et al. 1975, p. 357). In fact, by doing so, instead of constructing a time series $x_\tau^{(2)}$ whose first term would be, e.g., $x_1^{(2)}=(x_1 + x_2)/2$, we construct the time series $y_\tau^{(2)}$ whose first term is: $y_1^{(2)}:=\max\{(x_1 + x_2)/2,(x_2 + x_3)/2\} = (x_2 + \max\{x_1,x_3\})/2$, which is a different stochastic process from the one of interest, $\underline{x}_\tau^{(k)}$. For this reason, we do not apply a sliding window but use a fixed time window, with any arbitrary starting time, and without any conversion of the original time series (e.g., by a Hershfield coefficient), except taking temporal averages at several time scales.

3.2 Estimation of the timescale function parameters

The simplified version of the ombrian model utilizing the separability of functions $a(k)$ and $b(T)$ allows for a simplified fitting procedure, in two independent steps, as introduced by Koutsoyiannis et al. (1998). From the expression of Equations (9) and (12) it is easy to see that for the different timescales k_j the stochastic variables:

$$\underline{y}_j := a(k_j)\underline{x} = (1 + k_j/\alpha)^\eta \underline{x} \tag{14}$$

have a common distribution function, with the y_j for the different k_j being samples of it. Let then, $\underline{y}_{ji} := a(k_j)\underline{x}_{ji}$ denote the merged sample with length $n = \sum_j n_j$ where $\underline{x}_{ji}$ is the ith item of the sub-sample of size n_j for timescale k_j. Let also $\underline{r}_{ji}$ denote the rank of $\underline{x}_{ji}$ in the merged sample $\underline{y}_{ji}$ so that the mean rank of each sub-sample is given as $\underline{r}_j = \sum_i \underline{r}_{ji}/n_j$. Replacing all $\underline{r}_{ji}$ with the mean rank $\underline{r}_j$ we get a sample of n values, with $n_1$equal to $\underline{r}_1$, n_2 equal to $\underline{r}_2$ etc. Then the mean and variance estimators are, respectively:

$$\bar{\underline{r}} := \frac{1}{n}\sum_j n_j \underline{r}_j \tag{15}$$

$$\underline{\gamma}_r \coloneqq \frac{1}{n}\sum_j n_j(\underline{r}_j - \overline{\underline{r}})^2 \qquad (16)$$

If no ties are present among the different ranks, then $\overline{\underline{r}} = (n+1)/2$.

Following the assumption that the samples are from the same distribution, given by the right-hand side of Equation (14), then each realization of $\underline{r}_j$ should be close to the mean while that of the variance $\underline{\gamma}_r$ should be minimal. Therefore, the parameters α and η can be identified as the values that minimize the estimate of the variance γ_r from the observations x_{ji}. The original variables $\underline{y}_{ji}$ could be used as well instead of the ranks $\underline{r}_{ji}$, yet the use of the ranks makes the estimation process more robust to outliers. In order to improve the fit in the region of higher intensities, we may use a part of the data, belonging to the highest 1/2 or 1/3 of intensity values for each timescale.

3.3 Estimation of the distribution function parameters using K-moments

After estimating the parameters of the $a(k)$ function, the $b(T)$ function parameters must be specified. The distribution fitting is based on the method of K-moments (Koutsoyiannis 2019, 2023a, Chapter 6). K-moments have been developed with the aims of being knowable for very large orders (depending on the sample size) and interpretable in terms of order statistics. K-moments are a more general type than classical moments, probability-weighted moments, and L-moments, and share some properties with all of them as well as order statistics (for a detailed analysis the reader is referred to Koutsoyiannis 2023b). The distinctive feature of K-moments for the study of extremes though is that they are tailored to perform extreme-oriented analyses, as they enable reliable estimation of very high-order moments. Furthermore, each high-order K-moment estimate can be assigned a return period, which provides a direct means to empirical estimation of probability, alternative to order statistics. In addition, their estimation can be appropriately modified in the case that there is bias due to dependence, as discussed in the following sections. Finally, they have a simple, clear, intuitive and rigorous definition as expectations of maxima (upper K-moments) or minima (lower K-moments) in a sample. The former case which is relevant to our study is presented below.

Koutsoyiannis (2019) has introduced several variants of K-moments, of which here we use the simplest non-central variant, as defined in Koutsoyiannis (2023a, p.183). The simplest non-central upper K-moments K'_p is defined to be the expected value of the maximum of p independent stochastic variables identical to $\underline{x}$:

$$K'_p := \mathrm{E}\left[\max\left(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_p\right)\right] = p\mathrm{E}\left[\left(F(\underline{x})\right)^{p-1}\underline{x}\right] \quad (17)$$

for the moment order $p \geq 1$ and the last part of the equation is valid for continuous stochastic variables $\underline{x}$. The estimators of the non-central K-moments are given by the following formulae (Koutsoyiannis 2023a, p. 189):

$$\hat{\underline{K}}'_p = \sum_{i=1}^{n} b_{inp}\, \underline{x}_{(i:n)} \quad (18)$$

$$b_{inp} = \begin{cases} 0, & i < p \\ p\dfrac{\Gamma(n-p+1)}{\Gamma(n+1)}\dfrac{\Gamma(i)}{\Gamma(i-p+1)}, & i \geq p \geq 0 \end{cases} \quad (19)$$

where $\underline{x}_{(i:n)}$ is the ith smallest variable in a sample $\underline{x}$, of size n, (the ith item of the sample in ascending order) and p is the order of the moment, which can be any positive number $p \leq n$. In addition, the following holds:

$$\sum_{i=1}^{n} b_{inp} = 1 \quad (20)$$

The fact that $b_{inp} = 0$ for $i < p$ means that as the moment order increases, fewer data are used in the estimation, until only one is left, the maximum, when $p = n$, and $b_{nnn} = 1$. For $p > n$, $b_{inp} = 0$ for every i, $1 \leq i \leq n$, and the estimation becomes impossible. The first order non-central K-moment is the mean value of the sample.

The K-moment values, being closely related to order statistics, can also be assigned a return period, as follows (Koutsoyiannis 2023a, p.225-227):

$$\frac{T\left(K'_p\right)}{D} = p\Lambda_p \approx \Lambda_\infty p + \left(\Lambda_1 - \Lambda_\infty\right) \quad (21)$$

where Λ_1, Λ_∞ are coefficients depending on the distribution function and D is the time step or, more generally, a time period reference for the specification of return period. For the EV2 distribution it is shown (Koutsoyiannis 2023a, p. 229) that the Λ coefficients are functions of the shape parameter ξ:

$$\Lambda_1 = \frac{1}{1 - \exp\left(-(\Gamma(1-\xi))^{-\frac{1}{\xi}}\right)} \tag{22}$$

$$\Lambda_\infty = \Gamma(1-\xi)^{\frac{1}{\xi}} \tag{23}$$

For validation purposes, the following relationship of empirical return periods based on order statistics is also used, which is shown to provide an unbiased estimate of the logarithm of the return period (Koutsoyiannis 2023a, p.170):

$$\frac{T_{(i:n)}}{D} = \frac{n + e^{1-\gamma} - 1}{n - i + e^{-\gamma}} = \frac{n + 0.526}{n - i + 0.561} \tag{24}$$

The procedure outlined above could be directly applied for assigning return periods to the K-moments of any sample and the parameters of the EV2 distribution could be obtained by minimizing an error metric (e.g., the root mean square error) between the theoretical quantiles and the empirical K-moment values, or between the corresponding return periods. To take advantage of the large number of reliably estimated moments but also to check the behaviour of the model, it is also possible to use only some moments for calibration of the model, and to use the higher ranks for comparison/verification purposes (as a validation set). In so doing, the moments used in the calibration are still much more than the ones used in regular moment fitting procedures (typically up to 3 or 4 orders) while a second set of higher moments is also available for validation. It is also noted that the estimation method can be applied either to the rainfall intensities or directly to the values of the annual maximum rainfall depths or their standardized values, with an appropriate adjustment of the final parameters (e.g., Iliopoulou et al. 2022). It is also possible to use either the merged intensities of all time scales, provided that they are first suitably adjusted with the time scale function $a(k)$ (via Equation (14)), or to base the fit only on one time scale, e.g. 24 h.

3.4 Effect of temporal dependence on return period estimation

Since a K-moment is a property of the process's marginal, first-order distribution, its definition is unaffected by the dependence structure. On the other hand, temporal dependence introduces bias into K-moment estimators and thus, in the case of stochastic processes, the estimator's unbiasedness, asserted for the case of independent samples, is no longer valid. However, it is possible to quantify and offset the effect of dependence by appropriately modifying the moment order p used for the estimation, as follows.

For the case of temporal dependence of the general HK type as represented by the estimated H parameter, the following first-order approximation of the bias $-\Theta^{\mathrm{HK}}$, suitable for positively skewed processes, can be used (Koutsoyiannis 2023a, p. 216):

$$-\Theta^{\mathrm{HK}}(n,H) = \frac{1}{2n^{2-2H}} - \frac{2H(1-H)}{n} \tag{25}$$

Based on this, and assuming that the same adjustment applies approximately to all orders p (cf. simulation experiments in Koutsoyiannis 2023a, p.263), we can obtain the modified orders p' using the corresponding bias-correction factor Θ^{HK} (for simplicity denoted Θ) as:

$$p' \approx 2\Theta + (1-2\Theta)p^{\left((1+\Theta)^2\right)} \tag{26}$$

For $p = 1$, Equation (26) results to no modification, consistent to the fact that the estimator of the mean (i.e., the 1st order upper non-central moment) is unbiased. Based on the modified orders, the empirical return periods are adjusted according to Equation (21).

3.5 Effect of spatial dependence on return period estimation from merged records

In the case that we have many observation records, representing the same stochastic process, we often use them simultaneously by merging the samples to increase reliability of our estimations. It is well known however that in the case the records are not independent of each other, the estimation uncertainty for the merged record depends on an equivalent sample length that is reduced compared to that for the case of independence (in which the equivalent sample length is the sum of the individual record lengths) but remains greater than that of an individual record. The framework of K-moments allows for the effect of spatial dependence to be explicitly accounted for in the estimation of the return period. This is achieved through proper modification of the order of the moments of the unified sample, p', which in turn modifies the estimation of the return period. Let n_1 denote the sample length of each station, m denote the number of stations, and $n = m\, n_1$ denote the size of the merged sample, then the following methodology is applied (Koutsoyiannis 2023a, section 6.18):

- For $p \le n_1$ we set $p' = p$, therefore no modification is performed.

- For $p > n_1$ the following approximation is used. We estimate the equivalent Hurst parameter H, based on the spatial correlation of the stations ρ:

$$H = \frac{1}{2} + \frac{\ln(1+\rho)}{2\ln 2} \tag{27}$$

Then the modified orders of the moments are obtained as:

$$p' \approx 2\Theta + (1-2\Theta)(p - n_1 + 1)^{\left((1+\Theta)^2\right)} + n_1 - 1 \tag{28}$$

where Θ is obtained from Equation (25), and their corresponding return periods are adjusted according to Equation (21).

In the case that the records to be merged cannot be regarded as random samples but time series with time dependence (of general HK form) it is possible to apply the following approximate methodology for modifying their return periods (Koutsoyiannis 2023a; p.224). A representative, single value of the dependence parameter, which captures the effect of both temporal and spatial dependence, denoted as H_b ('bulk' H), is determined as follows:

$$H_b = \left(1 - \frac{\ln m}{\ln n}\right)H + \frac{\ln\left(\left(1 + \rho(m-1)\right)\right) + \ln m}{2\ln n} \tag{29}$$

where m, n, ρ, parameters as before, and H the long-term dependence parameter characterizing the individual records (see also Equation (4)). Then, the bias factor Θ^{HK} is estimated as before (Equation (25)), and the moment orders and corresponding return periods are modified. It is noted that unlike the previous case, in this case all the orders of the moments and the corresponding return periods are modified.

Therefore, these procedures provide a means to compensate for bias due to spatiotemporal dependence when multiple records are used simultaneously for the estimation of the distribution function parameters.

4. Conclusions

This work outlines the methodological framework for the construction of rainfall intensity–timescale–return period (also, called ombrian) curves that was used in the recent regionalization of the ombrian curves throughout the Greek territory, which is described in a companion paper (Iliopoulou et al. 2023). It introduces a stochastic framework for ombrian relationships based on a set of pre-requisites for theoretical consistency as set by Koutsoyiannis (2023a) that may be relaxed depending on practical needs. In this respect, the framework has two different levels of validity and complexity, (a) a full-scale ombrian model, covering any scale of interest, that requires the complete parent series for its determination, and (b) a simpler version applicable only over the usual small scales of hydrological design, that can be determined based on series of extremes. Special focus is given to the second case, since its more economic data requirements are most often satisfied for regional analyses of extremes, as was also the case in Greece. In addition, the simpler framework has an added flexibility related to involving two separable functions, a timescale function and a distribution function, the parameters of which can be estimated by different data sources. This enables, for instance, the straightforward use of daily raingauges for the fitting of the distribution function.

The framework is complemented by appropriate estimation strategies for both functions. Emphasis is placed on the efficient estimation of the distribution function using the new method of K-moments, which allow reliable high-order moment estimation and treatment of the extremes, while accounting for the bias induced by temporal and spatial dependence, which is non-negligible when adopting a stochastic approach.

Overall, the methodology provides a framework for theoretical understanding and modelling of ombrian relationships consistent with a stochastic representation of the parent rainfall process, accompanied by estimation procedures that are adjusted for data availability encountered in standard hydrological practice. The framework was applied for the recent construction of ombrian relationships in Greece and may be of use to other regional probabilistic analyses of rainfall extremes.

Funding: This research was funded by the General Directorate of Water of the Greek Ministry of Environment and Energy through the unit "Επιτελική Δομή ΕΣΠΑ Υπουργείου Περιβάλλοντος και Ενέργειας Τομέα Περιβάλλοντος" in the framework of the project "Production of maps with updated parameters of the ombrian curves at country level (implementation of the EU Directive 2007/60/EC in Greece)".

Data and Code Availability Statement: The PythOm package may be used to apply the multi-scale version of the ombrian model (Iliopoulou and Koutsoyiannis 2021).

Acknowledgments: We are grateful to the Associate Editor Elena Volpi and two anonymous reviewers for their comments which helped us to improve the manuscript.

Conflicts of Interest: We declare no conflict of interest.

References

Bernard, M.M., 1932. Formulas for rainfall intensities of long duration. *Transactions of the American Society of Civil Engineers*, *96*(1), pp.592-606.

Burlando, P. and Rosso, R., 1996. Scaling and muitiscaling models of depth-duration-frequency curves for storm precipitation. Journal of Hydrology, 187(1-2), pp.45-64.

Dimitriadis, P. and Koutsoyiannis, D., 2015. Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes. Stochastic environmental research and risk assessment, 29, pp.1649-1669.

Dimitriadis, P., Koutsoyiannis, D., Iliopoulou, T. and Papanicolaou, P., 2021. A global-scale investigation of stochastic similarities in marginal distribution and dependence structure of key hydrological-cycle processes. Hydrology, 8(2), p.59.

Greenwood, J.A., Landwehr, J.M., Matalas, N.C. and Wallis, J.R., 1979. Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. Water resources research, 15(5), pp.1049-1054.

Haruna, A., Blanchet, J. and Favre, A.C., 2023. Modeling Intensity-Duration-Frequency Curves for the Whole Range of Non-Zero Precipitation: A Comparison of Models. Water Resources Research, 59(6), p.e2022WR033362.

Hershfield, D.M. and Wilson, W.T., 1957. Generalizing of rainfall-intensity-frequency data. AIHS. Gen. Ass. Toronto, 1, pp.499-506.

Hershfield, D.M., 1961. Rainfall frequency atlas of the United States. Technical paper, 40, pp.1-61.

Hosking, J.R., 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. Journal of the Royal Statistical Society Series B: Statistical Methodology, 52(1), pp.105-124.

Hosking, J.R.M. and Wallis, J.R., 1988. The effect of intersite dependence on regional flood frequency analysis. Water Resources Research, 24(4), pp.588-600.

Hosking, J.R.M. and Wallis, J.R., 1997. Regional frequency analysis (p. 240).

Iliopoulou, T. and Koutsoyiannis, D., 2019. Revealing hidden persistence in maximum rainfall records. Hydrological Sciences Journal, 64(14), pp.1673-1689.

Iliopoulou, T., Koutsoyiannis, D., Malamos, N., Koukouvinos, A., Dimitriadis, P., Mamassis, N., Tepetidis, N., Markantonis, D. 2023. A stochastic framework for rainfall intensity-timescale-return period relationships: from point modelling to regionalization over Greece, Hydrological Sciences Journal, in review.

Iliopoulou, T., Malamos, N. and Koutsoyiannis, D., 2022. Regional ombrian curves: Design rainfall estimation for a spatially diverse rainfall regime. Hydrology, 9(5), p.67.

Iliopoulou, T., Papalexiou, S.M., Markonis, Y. and Koutsoyiannis, D., 2018. Revisiting long-range dependence in annual precipitation. Journal of Hydrology, 556, pp.891-900.

Iliopoulou, T. and Koutsoyiannis, D., 2021. PythOm: A python toolbox implementing recent advances in rainfall intensity (ombrian) curves. *Eur. Geosci. Union Gen. Assem*.

Koutsoyiannis, D. and Dimitriadis, P., 2021. Towards generic simulation for demanding stochastic processes. Sci, 3(3), p.34.

Koutsoyiannis, D. and Iliopoulou, T., 2022. Ombrian curves advanced to stochastic modeling of rainfall intensity. Rainfall, pp.261-284.

Koutsoyiannis, D. and Papalexiou, S.M., 2017. Extreme rainfall: Global perspective. Handbook of Applied Hydrology; McGraw-Hill: New York, NY, USA, pp.74-1.

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

Koutsoyiannis, D., 2023a.Stochastics of Hydroclimatic Extremes - A Cool Look at Risk, Edition 3, ISBN: 978-618-85370-0-2, 391 pages, doi:10.57713/kallipos-1, Kallipos Open Academic Editions, Athens.

Koutsoyiannis, D., 2023b. Knowable Moments in Stochastics: Knowing Their Advantages. Axioms, 12(6), p.590.

Koutsoyiannis. D., Iliopoulou, T., Koukouvinos, A., Malamos, N., Mamassis, N., Dimitriadis, P., Tepetidis, N., and Markantonis, D., 2023c. Technical Report, Production of maps with updated parameters of the ombrian curves at country level (impementation of the EU Directive 2007/60/EC in Greece), Department of Water Resources and Environmental Engineering – National Technical University of Athens.

Koutsoyiannis, D., 2004. Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: II. Recherche empirique sur de longues séries de précipitations. Hydrological Sciences Journal, 49(4).

Koutsoyiannis, D., 2006. An entropic-stochastic representation of rainfall intermittency: The origin of clustering and persistence. Water Resources Research, 42(1).

Koutsoyiannis, D., 2019. Knowable moments for high-order stochastic characterization and modelling of hydrological processes. Hydrological sciences journal, 64(1), pp.19-33.

Koutsoyiannis, D., Kozonis, D. and Manetas, A., 1998. A mathematical framework for studying rainfall intensity-duration-frequency relationships. Journal of hydrology, 206(1-2), pp.118-135.

Koutsoyiannis, D. and Montanari, A., 2015. Negligent killing of scientific concepts: the stationarity case. Hydrological Sciences Journal, 60(7-8), pp.1174-1183.

Lanciotti, S., Ridolfi, E., Russo, F. and Napolitano, F., 2022. Intensity–Duration–Frequency Curves in a Data-Rich Era: A Review. Water, 14(22), p.3705.

Langousis, A. and Veneziano, D., 2007. Intensity-duration-frequency curves from scaling representations of rainfall. Water Resources Research, 43(2).

Linsley Jr, R.K., Kohler, M.A. and Paulhus, J.L., 1975. Hydrology for engineers.

Molini, A., Lanza, L.G. and La Barbera, P., 2005. The impact of tipping-bucket raingauge measurement errors on design rainfall for urban-scale applications. Hydrological Processes: An International Journal, 19(5), pp.1073-1088.

Overeem, A., Buishand, A. and Holleman, I., 2008. Rainfall depth-duration-frequency curves and their uncertainties. Journal of Hydrology, 348(1-2), pp.124-134.

O'Connell, E., O'Donnell, G. and Koutsoyiannis, D., 2023. On the Spatial Scale Dependence of Long-Term Persistence in Global Annual Precipitation Data and the Hurst Phenomenon. Water Resources Research, 59(4), p.e2022WR033133.

Sane, Y., Panthou, G., Bodian, A., Vischel, T., Lebel, T., Dacosta, H., Quantin, G., Wilcox, C., Ndiaye, O., Diongue-Niang, A. and Diop Kane, M., 2018. Intensity–duration–frequency (IDF) rainfall curves in Senegal. Natural Hazards and Earth System Sciences, 18(7), pp.1849-1866.

Shehu, B., Willems, W., Stockel, H., Thiele, L.B. and Haberlandt, U., 2023. Regionalisation of rainfall depth–duration–frequency curves with different data types in Germany. Hydrology and Earth System Sciences, 27(5), pp.1109-1132.

Sherman, C.W., 1931. Frequency and intensity of excessive rainfalls at Boston, Massachusetts. Transactions of the American Society of Civil Engineers, 95(1), pp.951-960.
Svensson, C. and Jones, D.A., 2010. Review of rainfall frequency estimation methods. Journal of Flood Risk Management, 3(4), pp.296-313.
Veneziano, D. and Furcolo, P., 2002. Multifractality of rainfall and scaling of intensity-duration-frequency curves. Water resources research, 38(12), pp.42-1.

Appendix

The Pareto distribution as an ombrian model for small scales

The Pareto distribution for a continuous stochastic variable $\underline{x} \geq 0$ can be expressed as (Koutsoyiannis 2023a, p. 45):

$$F(x) = 1 - \left(1 + \xi\frac{x}{\lambda}\right)^{-\frac{1}{\xi}}, \ \xi > 0, \ \lambda > 0 \tag{A1}$$

where ξ is the upper-tail index, λ is a scale parameter and the lower bound is assumed zero. The corresponding return period is easily obtained as:

$$\frac{T(x)}{D} = \frac{1}{1 - F(x)} = \left(1 + \xi\frac{x}{\lambda}\right)^{\frac{1}{\xi}} \tag{A2}$$

For the chosen Pareto distribution of the parent stochastic variable $\underline{x}$, its multi-scale version with discontinuity at the origin for small time scales, which in the case of the rainfall process is equal to the probability dry, i.e., $P_0^{(k)} := P\{\underline{x}^{(k)} = 0\} = 1 - P_1^{(k)}$, where $P_1^{(k)} := P\{\underline{x}^{(k)} > 0\}$ is the probability wet, is obtained as:

$$F^{(k)}(x) = 1 - P_1^{(k)}\left(1 + \xi\frac{x}{\lambda(k)}\right)^{-1/\xi} \tag{A3}$$

The upper-tail index ξ should be scale-invariant (see proof in Koutsoyiannis 2023a, p.318), while the probability wet, $P_1^{(k)}$, and the state scale parameter, $\lambda(k)$, are functions of the time scale k. By setting $T = {}^{1}/(1 - F^{(k)}(x))$ in Equation (A3), the resulting rainfall quantile for scale k and return period T, i.e., the ombrian model, is:

$$x = \lambda(k)\frac{\left(P_1^{(k)}T/k\right)^{\xi} - 1}{\xi} \tag{A4}$$

To fully specify the model it suffices to determine the functions $\lambda(k)$ and $P_1^{(k)}$ which can be derived from the mean μ and the climacogram $\gamma(k)$ of the process, as follows. By standard algebra on equation **Error! Reference source not found.**, we find that the pth moment of $\underline{x}^{(k)}$ is:

$$\mathrm{E}\left[\left(\underline{x}^{(k)}\right)^p\right] = \mu'_p = \frac{P_1^{(k)}(\lambda(k))^p p}{\xi^p}\mathrm{B}\left(p, \frac{1}{\xi} - p\right) \tag{A5}$$

where B(,) denotes the beta function. Hence, the mean is:

$$\mathrm{E}\big[\underline{x}^{(k)}\big] = \mu = \frac{P_1^{(k)}\lambda(k)}{1-\xi} \tag{A6}$$

and the squared coefficient of variation is:

$$C_\mathrm{v}^2\big[\underline{x}^{(k)}\big] = \frac{\gamma(k)}{\mu^2} = \frac{2(1-\xi)}{(1-2\xi)P_1^{(k)}} - 1 \tag{A7}$$

which can be solved for $P_1^{(k)}$and $\lambda(k)$, to derive Equations (2) and (3).

It is noted that the special case $P_1^{(k)} = 1$ signifies the maximum time scale $k_{\max}^*$, at which the Pareto distribution is mathematically feasible, at which:

$$P_1^{(k_{\max}^*)} = 1,\ \frac{\gamma\big(k_{\max}^*\big)}{\mu^2} = \frac{1}{1-2\xi},\ \lambda\big(k_{\max}^*\big) = \mu(1-\xi) \tag{A8}$$

However, if we are interested in preserving the probabilities dry/wet, we should choose the time scale k^* (of transition from Pareto to a bell-shaped type of probability density, as the Pareto-Burr-Feller) smaller enough than $k_{\max}^*$, at a point where the deviation of probability dry derived from the Pareto model from the empirical one is marginally acceptable.

Table 1. Summary of the simplified ombrian relationships (for application over small scales) and their parameters, for rainfall intensity x, time scale k and return period T. Note that the equations are dimensionally consistent, so if, as usual, rainfall intensity is expressed in mm/h, the temporal scale in hours (h) and the return period in years (years), the parameters λ,α,β must be expressed in the same units, respectively.

Validity	**Mathematical relationship**
▪ For return period that is defined with reference to series above a threshold and therefore can also take values less than 1 year; it is the final relationship used for design	$x = \lambda \frac{\left({}^{T}/_{\beta}\right)^{\xi} - 1}{\left(1 + {}^{k}/_{\alpha}\right)^{\eta}}$
▪ For return period that refers to annual maxima series and thus takes values greater than Δ = 1 year; it is an intermediate relationship used for parameter estimation when annual maxima series are used	$x = \lambda \frac{\left(-\left({}^{\beta}/_{\Delta}\right)\ln\left(1 - {}^{\Delta}/_{T}\right)\right)^{-\xi} - 1}{\left(1 + {}^{k}/_{\alpha}\right)^{\eta}}$
Parameter	**Symbol (usual units)**
▪ Rainfall intensity scale parameter	λ (mm/h)
▪ Shape parameter (upper tail-index)	ξ (-)
▪ Timescale parameter for return period	β (years)
▪ Timescale parameter	α (h)
▪ Persistence parameter	η (-)