9

OMBRIAN CURVES ADVANCED TO STOCHASTIC MODELLING OF RAINFALL INTENSITY

Demetris Koutsoyiannis and Theano Iliopoulou

Department of Water Resources and Environmental Engineering, National Technical University of Athens, Heroon Polytechneiou 5, 157 80 Zographou, Greece

Abstract:

Ombrian curves, i.e. mathematic relationships linking average rainfall intensity to time scale of averaging and return period, also known as IDF (intensity-duration-frequency) curves, are essential tools in hydrology and engineering. Their use is supported by long-term hydrological experience, yet related formulas remain mostly empirical and lack a theoretical basis. As such, they entail several theoretical inconsistencies, particularly over large scales, while they cannot be applied in simulation. This Chapter reviews the typical form of ombrian curves along with its merits and limitations, and presents a modelling framework to overcome the latter by advancing curves to stochastic models of rainfall intensity. This is achieved through stochastic modelling of the joint second-order and marginal higher-order properties of the parent process. Two variants of the ombrian model are presented; a full version valid over time scales spanning multiple orders of magnitude, and a simplified relationship applicable over fine scales of the order of common applications, i.e. sub-hourly to daily. Specific emphasis is given to the fitting procedure combining multiple data sources and addressing bias in the estimation induced by temporal dependence. A detailed application of the ombrian model is performed for the rainfall station in Bologna (Italy), highlighting the efficiency of the resulting curves over multiple scales.

Key Words: ombrian curves; rainfall extremes: stochastic rainfall modelling; multi-scale model; IDF curves; Pareto distribution

9.1 Introduction

Ombrian relationships, from the Greek word 'όμβρος', i.e. rainfall, are mathematical relationships linking the average rainfall intensity, x, to time scale of averaging, k, and return period, T. These relationships are widely known as intensity-duration-frequency (IDF) curves, even though the term is a misnomer; 'duration' is a misplaced name for time scale of averaging, and 'frequency' reflects an old tradition of confusing frequency with return period. Nonetheless, ombrian curves are an established tool of most hydrological and engineering operations requiring *design storm* estimates, ranging from flood protection and urban drainage design, to construction of highways, bridges, etc. Related

methodologies are part of standard hydrological practice (Eagleson 1970; Chow et al. 1988), and can be traced back in literature as early as in the works of Sherman (1931) and Bernard (1932). Most countries have performed regionalization analysis of ombrian curves, producing maps for operational uses; such maps are available for the US since 1961 (Hershfield, 1961) and now are available or being updated for most parts of the world (Hailegeorgis et al., 2013; Koutsoyiannis et al., 1998).

The standard modelling procedure consists of fitting a parametric relationship to estimates of (x, k, T), often of the simple power law (Koutsoyiannis, 2021):

$$x = \frac{\lambda T^{\xi}}{k^{\eta}} \tag{1}$$

where λ, ξ, η are positive parameters with $\xi \leq \eta \leq 1$. Formulas of this type are often empirical and sometimes supported by fractal representations of rainfall intensity (e.g. Veneziano and Furcolo 2002; Langousis and Veneziano 2007). Even though the derivation of common formulas is mostly empirical, they have proven useful in practice and been validated by long-term and worldwide hydrological experience. However, there are several reasons that now dictate the need of a solid theoretical basis for their derivation.

First off, the standard empirical approach lacks rigor and entails several theoretical inconsistencies (even dimensional inconsistency, as evident from Equation (1); see details in Koutsoyiannis, 2021), which become important when one is interested in large return periods. A majority of these shortcomings has been exposed and rectified by Koutsoyiannis et al. (1998) who first connected the derivation of ombrian curves to the theoretical properties of the underlying rainfall process, namely its marginal distribution function and, in particular, its tail index. Linking the properties of the curves to their natural basis which is the parent process is also the only way to comprehend issues of bias and uncertainty in the estimation. Both issues are essential considering the presence of temporal dependence in the rainfall process, which induces bias in the estimation of its quantiles. Estimation uncertainty is further worsened by the usual lack of long-term reliable rainfall data at fine scales. Data scarcity is still a great challenge for many parts of the world (Ayman et al., 2011) and underlines the need to support ombrian curves estimation by a more powerful theoretical approach.

Even without considering the issue of data scarcity, a solid theoretical basis is essential when one is interested in applying simulation. The information conveyed by typical ombrian curves on rainfall frequency is usually not sufficient for studying complex hydrological problems beyond the 'design storm' applications. For instance, hydrosystems in which rainfall is only one of the uncertain components of the system, require thorough modelling of the involved processes in order to determine the overall probability of failure. As a matter of fact, as simulation is increasingly adopted in hydrological problems and the traditional ombrian curves have not kept up with such advances, the framework has even, unjustly, been seen as outdated.

The obvious alternative is to derive the ombrian relationships from simulations by models of the rainfall process, so-called rainfall generators. Yet the challenges involved in the multi-scale rainfall generation often beat the purposes of simplicity and practicality of the traditional curves. Moreover, it is not a given that the synthetic series will preserve the empirical ombrian curves, unless if included in the calibration scheme (Willems, 2000).

To this end, this Chapter presents a two-in-one approach, developed by Koutsoyiannis (2021), by which curves are themselves advanced to stochastic models of the all-scale rainfall intensity, i.e. *ombrian models*. The approach allows theoretically consistent derivation of ombrian curves, with provision for estimation bias, validity over extended range of scales and capability to be directly used for simulation. It is shown that these advances can be achieved on the basis of simple stochastic characterizations of the parent process, namely of its joint second-order and marginal higher-order properties.

The remaining of the Chapter is structured in 4 sub-sections. In particular, the following sub-section (9.2) sets the requirements for an ombrian model and is devoted to the presentation of the essential stochastic tools for the characterization of the parent process. Section 9.3 introduces two variants of the ombrian model, a full version covering all the range of time scales and a simplified relationship valid for fine timescales. Particular focus is given to the fitting procedure, outlined in Section 9.4, where issues of dependence-induced bias are also addressed. The entire methodology is illustrated in detail by the case study of rainfall in Bologna, in Section 9.5, which stands as a proof of concept of the ombrian model's power. The possibility of taking advantage of possible existence of multi-source data is also highlighted. Further aspects of the ombrian curves are discussed in section 9.6.

9.2 A stochastic framework for building ombrian models

In this Section, the stochastic concepts used in the ombrian modelling are presented. A stochastic process is an arbitrarily large family of random variables $\underline{x}(t)$ (Papoulis, 1991). To distinguish random variables from regular variables, the former are underlined following the Dutch convention. These variables are indexed by t, which represents time, either discrete (from the set of integers $\mathbb{Z}$, referring to a discrete-time stochastic process), or continuous (from the set of real numbers $\mathbb{R}$, resulting in a continuous-time stochastic process). A continuous-time stochastic process is herein denoted by $\underline{x}(t)$, and a discrete-time one by $\underline{x}_\tau$. A realization of stochastic process $\underline{x}(t)$, i.e. a timeseries, necessarily referring to discrete time and denoted by x_τ.

9.2.1 Basic requirements for an ombrian model

The basic premise of the Chapter is that an ombrian model can be an advance of the classic tool of ombrian curves if it achieves greater modelling power and theoretical consistency, but preserves the practical and simple character of the classic curves. Below we outline the basic requirements for such a model as identified by Koutsoyiannis (2021).

1. A critical prerequisite is that the ombrian model should be applicable in engineering applications without a necessity to perform simulation. Its application should be preferably simple as in traditional IDF curves even if the need for theoretical consistency is compromised to some extent.

2. It is straightforward that, as in every stochastic model, the first and second order properties of the process of interest, i.e. the temporal average of rainfall intensity $\underline{x}^{(k)}$ over any time scale k, should be preserved. Clearly, a constant mean should be preserved for all

time-scales, although this is often violated in common expressions. An effective methodology to preserve the second-order properties at any scale based on the scaling properties of the variance (climacogram) is outlined in Section 9.2.2.

3. The process's asymptotic variance at $k \to 0$ should be finite; the contrary would imply that the process requires infinite energy to materialize which is absurd for physical processes. In addition, the process's asymptotic variance at $k \to \infty$ should be zero, in order for the process to be ergodic.

4. The model should deal with the intermittence of rainfall occurrences at fine time scales, describing both the probability dry $P_0^{(k)} := P\{\underline{x}^{(k)} = 0\}$, and the probability wet, $P_1^{(k)} := \overline{F}^{(k)}(0) = 1 - P_0^{(k)}$ for any time scale k, including for $k \to 0$.

5. The principle modelling focus is on rainfall maxima, and hence it is important to preserve the high-order properties of the process.

6. The tail index of the rainfall intensity distribution should be constant for all time scales. Theoretical justification of this requirement can be found in Koutsoyiannis (2021).

7. The Pareto distribution constitutes an optimal choice for small time scales due to its simplicity and explicit relationship between the time-averaged intensity and return period. Prevalence of the Pareto distribution for rainfall intensities is also supported by worldwide empirical evidence (Koutsoyiannis, 2004a; Koutsoyiannis and Papalexiou, 2016).

9.2.2 Characterization of second-order properties through climacogram

A comprehensive characterization of a process's second-order properties can be achieved by inspecting the properties of its variance when the process is averaged (or aggregated) over multiple scales. The function of the variance of the averaged process versus the scale of averaging is called the climacogram, while the function of the cumulative process versus the scale is called the cumulative climacogram (Koutsoyiannis, 2010). The climacogram of a process $\underline{x}(t)$ is defined as:

$$\gamma(k) := \mathrm{var}\left[\frac{\underline{X}(k)}{k}\right] = \frac{\Gamma(k)}{k^2} \tag{2}$$

where $\Gamma(k)$ is the cumulative climacogram, and $\underline{X}(k)$ is the process $\underline{x}(t)$ aggregated at timescale k:

$$\underline{X}(k) := \int_0^k \underline{x}(t)\,\mathrm{d}t \tag{3}$$

or for a discrete-time process, with climacogram γ_κ:

$$\underline{X}_\kappa := \underline{x}_1 + \underline{x}_2 + \cdots + \underline{x}_\kappa \tag{4}$$

The discrete time scale κ (integer) is related to the continuous-time one k (real number) by $k = \kappa D$, with D denoting the length of the time step.

The climacogram estimator is the same for discrete- and continuous-time processes and is given as:

$$\underline{\hat{\gamma}}_\kappa \equiv \underline{\hat{\gamma}}(k) := \frac{1}{n} \sum_{\tau=1}^{n} \left(\underline{x}_\tau^{(\kappa)} - \underline{\hat{\mu}} \right)^2 \tag{5}$$

where $\underline{\hat{\mu}} := (1/n) \sum_{i=1}^{n} \underline{x}_i$ is the estimator of the true mean.

The climacogram is theoretically equivalent to other second-order properties, namely the autocovariance, autocorrelation and the power-spectrum, but it has superior estimation properties in terms of bias, discretization errors, and sampling uncertainty (Dimitriadis and Koutsoyiannis, 2015). Therefore, it is the basic tool employed here for second-order characterization.

The theoretical climacogram differs among processes with different second-order dependence structure. In case of an independent white-noise process in continuous time, the climacogram is inversely proportional to the time scale:

$$\gamma(k) = \frac{\sigma^2 D}{k} \tag{6}$$

where σ^2 is the variance of the process for $k = D$. Notice that the variance of the instantaneous process is infinite. An extension of the white-noise process, again having infinite variance as $k \to 0$ but now exhibiting dependence in time, is the Hurst-Kolmogorov process, which can be defined through its climacogram as:

$$\gamma(k) = \lambda \left(\frac{a}{k} \right)^{2-2H} \tag{7}$$

where α and λ are scale parameters, with dimensions $[t]$ and $[x^2]$ and H is the so-called Hurst parameter ranging in the interval $(0,1)$. In the case of $H = 0.5$ the white noise is recovered. For $0.5 < H < 1$ the process is persistent and for $0 < H < 0.5$ antipersistent.

The infinite variance of these processes when the scale tends to zero makes them inappropriate for natural processes, as discussed before. In order to remedy this shortcoming, and improve flexibility of the model to describe the dependence in shorter time scales, the Filtered Hurst-Kolmogorov (FHK) process is developed with several climacogram types (Koutsoyiannis, 2017). The generalized Cauchy-type climacogram (FHK-C) is:

$$\gamma(k) = \lambda_1 \left(1 + \left(\frac{k}{a} \right)^{2M} \right)^{\frac{H-1}{M}} \tag{8}$$

where M is an added dimensionless parameter which controls the local scaling of the process (fractal behaviour), denoted as M in honor of Mandelbrot (Koutsoyiannis et al., 2018). Values of $M < 1/2$ indicate a rough process, while $M > 1/2$ indicate a smooth process.

An alternative flexible type is the composite Cauchy-Dagum-type (FHK-CD) climacogram, which for a rough and persistent process, and for the special case $M = 1 - H$, can be written as:

$$\gamma(k) = \lambda_1 \left(1 + \frac{k}{a}\right)^{2H-2} + \lambda_2 \left(1 - \left(1 + \frac{a}{k}\right)^{2H-2}\right) \tag{9}$$

Both these climacogram models have four parameters and thus great flexibility in capturing the scaling properties of the variance at all scales. More information on the bounds of scaling and on other climacogram-type models are provided in Koutsoyiannis (2017).

Therefore, the empirical climacogram is given by estimating the variance over scales by Equation (5), whereas Equations (8)-(9) provide different types of the theoretical climacograms. Because presence of dependence induces downward bias in the estimation of the variance, the two are not directly comparable. To compare them, the bias need to be considered, based on the following equation (Koutsoyiannis, 2003, 2021):

$$\mathrm{E}\left[\hat{\underline{\gamma}}(k)\right] = \gamma(k) - \gamma(L) \tag{10}$$

where L is the length of the observation period.

9.2.3 Assigning empirical return periods using order statistics

Order statistics are a standard tool for dealing with extremes. Below the procedure to apply them for assigning return periods to the data is presented.

Let $\underline{x}$ be a stochastic variable and $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ be IID copies of it, forming a sample. If we rearrange them by increasing order of magnitude such that $\underline{x}_{(i:n)}$ is the ith smallest of the n, i.e.:

$$\underline{x}_{(1:n)} \le \underline{x}_{(2:n)} \le \cdots \le \underline{x}_{(n:n)} \tag{11}$$

then the stochastic variable $\underline{x}_{(i:n)}$ is called the ith order statistic. The minimum and maximum values of a sample are then given respectively by the lowest order statistic, $\underline{x}_{(1:n)} = \min(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n)$ and the highest order statistic $\underline{x}_{(n:n)} = \max(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n)$. It is then well-known (Papoulis 1990) that if we define the stochastic variable $\underline{u} := F(\underline{y}) = F(\underline{x}_{(i:n)})$, then its distribution function is the Beta distribution, whose mean is:

$$\mathrm{E}[\underline{u}] = \mathrm{E}[F(\underline{x}_{(i:n)})] = \frac{i}{n+1} \tag{12}$$

Then an estimate of the return period (in time units D) for order statistics $\underline{T}_{(i:n)} := T(\underline{x}_{(i:n)})$ is:

$$\frac{T_{(i:n)}}{D} = \frac{1}{1 - \mathrm{E}\left[F\left(\underline{x}_{(i:n)}\right)\right]} = \frac{n+1}{n+1-i} \tag{13}$$

which is the well-known *Weibull plotting position*. This, however, is not recommended for use as it results in high bias in the estimation of the return period of the highest events. There are several other formulae for return period all of which are of the form:

$$\frac{T_{(i:n)}}{D} = \frac{n+B}{n-i+A} \tag{14}$$

The parameters A and B depend on the parent distribution of the data. Koutsoyiannis (2021) developed various parameterizations as the best approximations for a number of distributions and properties of interest. In particular, for distributions belonging to the domain of attraction of EV2, such as the Pareto, the parameters A, B are theoretically proved to be:

$$A = \left(\Gamma(1-\xi)\right)^{-1/\xi}, \qquad B = \left(\Gamma(2-\xi)\right)^{-1/\xi} - 1 \tag{15}$$

These can be suggested for use (replacing empirical formulae such as the Weibull plotting positions) for assigning return periods to order statistics, assuming independence of the data. Yet since independence is mostly an untenable assumption for natural processes, there is bias involved in the estimation in this case as well.

For an explicit account of dependence in assigning return periods, one could use knowable moments (Koutsoyiannis, 2019). For simplicity however, the following ad-hoc procedure is proposed as an approximation to account for bias for each time scale.

1. A first estimate of the return period of each nonzero value $x_{(i:n)}$ appearing in a sample sorted in ascending order, is obtained based on the independence assumption from Equation (14). This follows the estimation of the coefficients A and B of Equation (15) based on the tail index of the process.

2. The following approximation for a bias correction factor $\Theta(k, L, H)$ is used (Koutsoyiannis, 2021):

$$\Theta(k, L, H) \approx -\frac{\gamma(L)}{2\gamma(k)} \tag{16}$$

Accordingly, the empirical return periods are corrected as:

$$T' \approx \min\left(\left(2\Theta + (1-2\Theta)\left(\frac{T\hat{P}_1^{(k)}}{2k}\right)^{(1+\Theta)^2}\right)\frac{2k}{\hat{P}_1^{(k)}}, T\right) \tag{17}$$

where $\hat{P}_1^{(k)} = \hat{n}_1/n$ is the ratio of the non-zero values to the total values at each scale, else the probability wet.

3. This procedure is repeated for all nonzero values $x_{(i:n)}$ for each timescale k, yielding a table of empirical values and associated return periods.

9.3 Building a theoretically consistent ombrian model

9.3.1 All-scale version

The first version of the ombrian model refers to a model valid over the whole range of available scales. To achieve the extension of the typical fine-scale curves to large scales an increase in the complexity of the rainfall's intensity distribution is required. This is described by the following assumptions:

1. At small time scales the rainfall intensity follows a mixed type distribution, with a discrete part at the origin described by the probability dry, and a continuous part following the Pareto distribution with a constant tail index ξ and a state scale parameter $\lambda(k)$ as a function of the timescale:

$$F^{(k)}(x) = 1 - P_1^{(k)} \left(1 + \xi \frac{x}{\lambda(k)} \right)^{-1/\xi} \tag{18}$$

2. At larger time-scales the rainfall intensity follows the Pareto-Burr-Feller (PBF) distribution with discontinuity at zero, characterized by an extra parameter $\zeta(k)$ as a function of the timescale:

$$F^{(k)}(x) = 1 - P_1^{(k)} \left(1 + \xi \left(\frac{x}{\lambda(k)} \right)^{\zeta(k)} \right)^{-1/\xi} \tag{19}$$

The Pareto distribution is obtained for $\zeta(k) = 1$. The PBF distribution is chosen because, contrary to the Pareto, it becomes bell-shaped for increasing $\zeta(k)$ which is consistent to the behaviour of the rainfall intensity at large time scales (cf. the central limit theorem).

3. The mean of the time-averaged process is constant across all time-scales:

$$E[\underline{x}^{(k)}] = \mu \tag{20}$$

4. The climacogram follows one of the two four-parameter models introduced in Equations (8)-(9). Clearly, both equations satisfy the asymptotic requirements for the variance set in 9.2.1. As $k \to \infty$, $\gamma(k) \to 0$, whereas for $k = 0$, both variances are finite and equal to $\gamma(0) = \gamma_0 = \lambda_1$ in Equation (8) and $\gamma(0) = \gamma_0 = \lambda_1 + \lambda_2$ in Equation (9).

5. The probability wet $P_1^{(k)} = 1 - P_0^{(k)}$ and dry $P_0^{(k)}$ follow the scaling law:

$$\ln P_0^{(k)} = \ln P_0^{(k^*)} (k/k^*)^\theta, \qquad k \geq k^* \tag{21}$$

where k^* is the transition time scale from Pareto to PBF distribution, for which $P_0^{(k^*)} > 0$ and $\zeta(k^*) = 1$ (for continuity of the transition), and θ is a parameter ($0 \leq \theta \leq 1$). This equation was derived by Koutsoyiannis (2006) from an entropy maximization framework.

The introduction of the two different distributions follows from the need to preserve the shape of the probability of rainfall, which is highly skewed at small time-scales but tends to bell-shape at large scales. However, it is noted in Equation (19) that the tail-index of the PBF distribution is not ξ but $\xi/\zeta(k)$ and tends to zero as $k \to \infty$. Thus, at large time-

scales the constant tail index requirement is violated. An alternative solution would be to replace the PBF with a shape-preserving distribution, yet analytical expressions are too involved and defeat the requirement of practicality. Besides, the violation occurs only at large time-scales which are less of interest in applications.

Having assumed the distribution types, it remains to specify the form of the parameters $\lambda(k)$ and $\zeta(k)$ which are derived from the first- and second-order properties, i.e. the mean and the climacogram of the process. For the PBF distribution these are given by:

$$\frac{1}{\zeta(k)} \approx \sqrt{(1 - 2\xi)\left(P_1^{(k)}\frac{\gamma(k) + \mu^2}{\mu^2} - 1\right)} \tag{22}$$

$$\frac{1}{\lambda(k)} \approx \frac{P_1^{(k)}}{\mu}\left(1 + \frac{1}{(1 - \xi)(\zeta(k))^2} - \frac{1}{(\zeta(k))^{\sqrt{2}}}\right) \tag{23}$$

For the Pareto distribution, $\zeta(k) = 1$, and therefore the probability wet can be explicitly derived from Equation (22) as:

$$P_1^{(k)} = \frac{1 - \xi}{1/2 - \xi}\frac{\mu^2}{\gamma(k) + \mu^2} \tag{24}$$

while in this case Equation (23) can be simplified to:

$$\frac{1}{\lambda(k)} = \frac{P_1^{(k)}}{\mu(1 - \xi)} = \frac{\mu}{(1/2 - \xi)(\gamma(k) + \mu^2)} \tag{25}$$

The special case of $P_1^{(k)} = 1$ denotes the maximum scale till which the Pareto distribution is mathematically feasible, thus $k = k_{max}^*$, and the following hold:

$$P_1^{(k_{max}^*)} = 1, \qquad \frac{\gamma(k_{max}^*)}{\mu^2} = \frac{1}{1 - 2\xi}, \quad \lambda(k_{max}^*) = \mu(1 - \xi) \tag{26}$$

However, in order to preserve the scaling behaviour of the probabilities dry/wet, as specified by Equation (21), the transition scale to the PBF distribution should be chosen much smaller than k_{max}^*, i.e. the Pareto feasibility limit.

On the contrary, the PBF is feasible at any scale, while for large scales in which $P_1^{(k)} = 1$, Equation (22) simplifies to:

$$\frac{1}{\zeta(k)} = \frac{\sqrt{(1 - 2\xi)\gamma(k)}}{\mu} \tag{27}$$

The final version of the ombrian model is obtained by substituting the return period $T = 1/(1 - F^{(k)}(x))$ in the Equation (19) for the PBF:

$$x = \lambda(k) \left(\frac{(P_1^{(k)}T / k)^\xi - 1}{\xi} \right)^{1/\zeta(k)} \tag{28}$$

and for the Pareto ($\zeta(k) = 1$):

$$x = \lambda(k) \frac{(P_1^{(k)}T / k)^\xi - 1}{\xi} \tag{29}$$

For $\xi = 0$, the PBF distribution switches to the Weibull, and the Pareto to the exponential, i.e.:

$$x = \lambda(k) \left(\ln\left(P_1^{(k)}T/k\right) \right)^{1/\zeta(k)}, \qquad x = \lambda(k) \ln(P_1^{(k)}T / k) \tag{30}$$

The final ombrian relationship with its basic properties is summarized in Table 9.1. It is evident that the ombrian relationship is given through the mean, the climacogram, the tail index of the distribution and the probability wet. The full-range model results in a total of seven parameters, depending on the choice of the climacogram model, of four categories: (a) the mean intensity parameter μ with units of $[x]$, i.e. the average of the process, typically mm/h, (b) the intensity scale parameter λ_1, in case of the FHK-C climacogram model, or λ_1, λ_2 in the case of FHK-CD with units of $[x^2]$, (c) the time scale parameter α, in units of time $[t]$, and (d) the dimensionless parameters ξ ($0 < \xi < 0.5$), i.e. the tail index, θ ($0 < \theta < 1$), i.e. the exponent of the expression of probability dry, M ($0 < M < 1$), i.e. the fractal parameter in case of the FHK-C climacogram, and H ($0 < H < 1$), i.e. the Hurst parameter. Note that if the FHK-CD climacogram model (Equation (9)) is used, then the fractal parameter is derived as $M = 1 - H$, and thus it is not an extra parameter.

It is worth noting that typical ombrian curves involve five parameters, yet the gains of including the extra parameters are manifold. In addition to the recovered mathematical and physical consistency, the model yields a better representation of fine scales (through fractal M parameter) and arbitrarily large scales (through H parameter), while precisely preserving the mean, climacogram and probability dry/wet of the process. In principle, this version of the model has the advantage of being valid over all timescales. Yet if only fine time-scales are of interest, then a less parameterized and simpler version can be used instead. This is discussed next.

Table 9.1 Ombrian models for the full range of scales and the small scales and their basic properties, i.e. mean, climacogram, probability wet, shape scale parameter and state scale parameter.

	All-scale ombrian model		Simplified model
	Small scales (Pareto) $k \leq k^* \ll k^*_{\max}$	Large scales (PBF) $k \geq k^*$	Small scales (Pareto) $k \ll \beta$
x for $\xi > 0$	$\lambda(k)\dfrac{(P_1^{(k)}T \ / \ k)^\xi - 1}{\xi}$	$\lambda(k)\left(\dfrac{(P_1^{(k)}T \ / \ k)^\xi - 1}{\xi}\right)^{1/\zeta(k)}$	$\lambda\dfrac{\left((T/\beta)^\xi - 1\right)}{(1 + k/\alpha)^\eta}$
x for $\xi = 0$	$\lambda(k)\ln(P_1^{(k)}T \ / \ k)$	$\lambda(k)\left(\ln\left(P_1^{(k)}T/k\right)\right)^{1/\zeta(k)}$	$\lambda\dfrac{\ln(T/\beta)}{(1 + k/\alpha)^\eta}$
Properties			
$E[\underline{x}^{(k)}]$	μ		(inconsistent – not constant)
$\gamma(k)$	$\lambda_1(1 + (k/\alpha)^{2M})^{\frac{H-1}{M}}$ or $\lambda_1(1 + k/\alpha)^{2H-2} + \lambda_2(1 - (1 + a/k)^{2H-2})$		$\lambda(1 + k/\alpha)^{2H-2}$
$P_1^{(k)}$	$\dfrac{1-\xi}{1/2 - \xi}\dfrac{\mu^2}{\gamma(k) + \mu^2}$	$1 - (1 - P_1^{(k^*)})\left(\dfrac{k}{k^*}\right)^\theta$	$\dfrac{k}{\beta}$
$\dfrac{1}{\zeta(k)}$	1	$\sqrt{(1 - 2\xi)\left(P_1^{(k)}\dfrac{\gamma(k) + \mu^2}{\mu^2} - 1\right)}$	(not applicable)
$\dfrac{1}{\lambda(k)}$	$\dfrac{P_1^{(k)}}{\mu(1 - \xi)}$	$\dfrac{P_1^{(k)}}{\mu}\left(1 + \dfrac{1}{(1 - \xi)(\zeta(k))^2} - \dfrac{1}{(\zeta(k))^{\sqrt{2}}}\right)$	(not applicable)

9.3.2 Simplified model for small scales

It is possible to simplify the Pareto ombrian relationship (Equation (29)) which is valid for small scales, $k \leq k^*_{\max}$, and can be written as:

$$x = \lambda(k)\frac{(T \ / \ \beta(k))^\xi - 1}{\xi} \tag{31}$$

where $\beta(k)$ is a function of the time scale with units of time, $\beta(k) := k/P_1^{(k)}$. Then by virtue of Equation (25), the ombrian relationship yields:

$$x = \frac{(1/2 - \xi)(\gamma(k) + \mu^2)}{\xi\mu}\left(\left(\frac{T}{\beta(k)}\right)^\xi - 1\right) \tag{32}$$

For small scales we may introduce a series of simplifying assumptions:

1) We assume that $P_1^{(k)} \propto k$, which is an acceptable approximation for small scales, thus that $\beta(k) = \beta = $ constant.
2) Noting that for small scales $\gamma(k) \gg \mu^2$, we neglect the second term in the sum of Equation (32).
3) We assume an FHK-C climacogram (Equation (8)) with a neutral value for $M = 0.5$

By application of these three assumptions to Equation (32), we get:

$$x = \lambda_1 \frac{(1/2 - \xi)}{\xi\mu} \left(1 + \frac{k}{\alpha}\right)^{2H-2} \left(\left(\frac{T}{\beta}\right)^{\xi} - 1\right) \tag{33}$$

which is an ombrian relationship of the form:

$$x = \lambda \frac{b(T)}{a(k)} \tag{34}$$

Thus, obtaining the rainfall intensity as a quotient of a function of the time scale and the return period. This facilitates estimation. The function $a(k)$ is given as:

$$a(k) = \left(1 + \frac{k}{\alpha}\right)^{\eta}, \qquad \eta := 2 - 2H \tag{35}$$

while the parameter λ and the function $b(T)$ are dependent on ξ as follows. For $\xi > 0$:

$$\lambda = \frac{(1/2 - \xi)\lambda_1}{\xi\mu}, \qquad b(T) = \left(\frac{T}{\beta}\right)^{\xi} - 1 \tag{36}$$

and for $\xi = 0$:

$$\lambda = \frac{\lambda_1}{2\mu}, \qquad b(T) = \ln\left(\frac{T}{\beta}\right) \tag{37}$$

This simplification results in a total of five parameters, of three categories: (a) λ with units of x, i.e. typically mm/h, (b) the scale parameters α and β with units of time, and (c) the dimensionless parameters ξ ($0 < \xi < 0.5$), i.e. the tail index, and η ($0 < \eta < 1$), which is related to the Hurst parameter of the process. Any value of $\eta < 1$ corresponds to a process with persistence, $H > 0.5$, while $\eta = 1$ (which in reality is never the case) would correspond to $H = 0.5$, a process with purely random behaviour. The model is also summarized in Table 9.1 along with its basic properties.

It can be seen that the simplified version sacrifices some of the requirements set at the beginning, mostly importantly instead of a constant mean, it yields a mean increasing with time-scale (Table 9.1). However, the inconsistencies are negligible if one restricts the range of timescales using as a lower bound the smallest value of the observed data and choosing an upper bound sufficiently below β. Therefore, the model is applicable over this range of observed scales, but if simulation is of interest then the all-scale version should be used instead.

9.4 Model fitting procedure

9.4.1 All-scale version

By assuming an initial parameterization of the all-scale ombrian model we can obtain a theoretical estimate of: (a) the climacogram $\gamma(k)$, (b) the probability wet vs the time scale $P_1^{(k)}$ and (c) the rainfall intensity as a function of the timescale and the return period, i.e. the ombrian model $x(k,T)$, using the equations summarized in Table 9.1. From the empirical series, we also may obtain the empirical estimates of these three relationships as follows. To obtain an estimate of the climacogram $\hat{\gamma}(k)$, we use Equation (5), following the procedure in Section 9.2.1. To estimate the probability wet, we estimate the ratio $\hat{P}_1^{(k)} = \hat{n}_1/n$ where $\hat{n}_1$ is the number of nonzero observations, and n the total number of observations in the series. Then to assign a return period to each rainfall observation at each scale, thus estimate $x(k,T)$, we use the order statistics method, as presented in Section 9.2.3.

  The model is fitted to the empirical estimates, after bias is accounted for, by minimizing the error between the two through a nonlinear solver, available even in any computational (e.g. spreadsheet) environment. Specifically, the error that should be minimized to fit the climacogram adjusted for bias is of the form:

$$E_\gamma := \sum_k w_\gamma(k)(\ln(\gamma(k) - \gamma(L)) - \ln\hat{\gamma}(k))^2 \tag{38}$$

where $w_\gamma(k)$ is a weighting function of scale. The logarithm is introduced to account for the different orders of magnitude that the climacogram spans. By minimizing Equation (38) we obtain all the climacogram-related parameters.

  Likewise, the error for the probability wet is defined as:

$$E_P := \sum_k w_P(k)(P_1^{(k)} - \hat{P}_1^{(k)})^2 \tag{39}$$

where $w_P(k)$ is the weight, which can be chosen as a function of scale. Since the expression for the probability wet, involves all parameters of the ombrian model, Equation (39) could be used to specify the full version of the model. However, this would give more weight to the representation of the probability dry, than to the extreme values, which are the ones of interest. To this aim, it is better to obtain the parameters directly from the distribution quantiles $x(k,T)$. The total fitting error in this case is given as:

$$E_x := \sum_k \frac{1}{\gamma(k)} \frac{1}{n_k} \sum_T w_x(T)(x(k,T) - \hat{x}(k,T))^2 \tag{40}$$

where $w_x(T)$ is a weighting function of the return period and n_k is the number of x values at each scale k. The total mean square error over the entire set of return periods is further normalized by the climacogram $\gamma(k)$.

  In so doing, we have determined the full parameterization of the ombrian model also accounting for dependence-induced bias. It is also possible to optimize the parameters

of the model by formulating an objective function that includes, as a weighted sum, all three errors defined above:

$$E := a_\gamma E_\gamma + a_P E_P + a_x E_x \tag{41}$$

where a_γ, a_P, a_x the weights for the three errors.

9.4.2 Simplified version

The simplified version of the ombrian model also allows for a simplified fitting procedure, adjusting the steps previously outlined. In fact, by observing the separability of functions $a(k)$ and $b(T)$ in this version an independent, two-step fitting approach can be used, introduced by Koutsoyiannis (1998). Equation (34) can be expressed as:

$$a(k)x = \lambda b(T) \tag{42}$$

We note that the timescale k is not a stochastic variable as it takes values from a fixed set, depending on data availability, whereas $a(k)$ is a deterministic function thereof. The right-hand side of the equation is in fact an expression of the Pareto distribution, independent of timescale k. Substituting Equations (35)-(36) in the above equation, yields:

$$\left(1 + \frac{k}{\alpha}\right)^\eta x = \lambda\left(\left(\frac{T}{\beta}\right)^\xi - 1\right) \tag{43}$$

Now, it is easy to see that for the different timescales k_j the stochastic variables $\underline{y}_j := a(k_j)\underline{x} = (1 + k/\alpha)^\eta \underline{x}$ have a common distribution function, with the $\underline{y}_j$ for the different k_j being samples of it. Let then, $\underline{y}_{ji} := a(k_j)\underline{x}_{ji}$ of length $n = \sum_j n_j$ denote the merged sample of all sub-samples $\underline{x}_{ji}$ of size n_j corresponding to timescale k_j. Let also $\underline{r}_{ji}$ denote the rank of each sub-sample $\underline{x}_{ji}$ 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 value $\underline{r}_j$ we get a sample of n values, with n_1 equal to r_1, n_2 equal to r_2 etc. Then the mean and variance estimators are, respectively:

$$\overline{\underline{r}} := \frac{1}{n}\sum_j n_j\,\underline{r}_j \tag{44}$$

$$\underline{\gamma}_r := \frac{1}{n}\sum_j n_j\left(\underline{r}_j - \overline{\underline{r}}\right)^2 \tag{45}$$

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 (43), then each $\underline{r}_j$ should be close to the mean $\overline{\underline{r}}$ while the variance should be minimal. Therefore, we can find the parameters α and η as the values that minimize the estimate of the variance γ_r from the observations $\underline{x}_{ji}$. The original values

$\underline{y}_{ji}$ could be used as well instead of the ranks, yet the use of the ranks makes the estimation process more robust to outliers. In order to improve the fit to the higher quantile region, we could also use a part of the data of each sample, belonging to the highest 1/2 or 1/3 of the data (Koutsoyiannis, 1998).

Having estimated the α and η parameters, it remains to specify the parameters of the function $b(T)$. Following the same rationale, i.e. of a single distribution function, we merge all k sub-samples into a single sample and we estimate the parameters of the Pareto distribution, which fully determines the form of $b(T)$.

The two-step fitting procedure has an attractive flexibility in using different sources of data. Namely, a reliable determination of parameters α and η requires sub-hourly and sub-daily data, respectively, whereas, on the contrary, the parameters of the function $b(T)$ are better inferred from daily rain-gauge data. Particularly, the most uncertain and critical parameter is the tail-index of the distribution, which requires long timeseries to be reliably estimated. In the absence of long observational rain-gauge records, the tail index of the Pareto should be estimated from regional analysis or be assumed independently of the data, based on local hydrological experience.

9.5 Development of an ombrian model for Bologna in Italy

The all-scale ombrian model is applied to the rainfall of Bologna in Italy, which has one of the longest daily rainfall records worldwide spanning 206 years. The time series of daily observations is available online in the frame of the Global Historical Climatology Network – Daily (GHCN-Daily). Hourly rainfall data from the Dext3r repository (made available by Lombardo et al. 2019) are also employed, covering the entire period 1990-2013, with the exception of the missing year 2008. To take advantage of the availability of the two data sources, the ombrian model is fitted to both rainfall series simultaneously.

As a first step, the data of both series are aggregated at larger time-scales. Specifically, the hourly rainfall data are aggregated at timescales of 2, 4, 6, 12, 24, 48 and 96 h, thus the modelling scales extend from 1 h to 4 d. The longer daily rainfall data are aggregated at timescales of 2, 4, 8, 16, 32, 64, 128, 182, 365, 730, 1460, and 5840 d, thus in this case, the range of scales extends from 1 d to 16 years. Longer timescales are studied for the daily data due to both their longer length and their higher reliability for the estimation of long-term properties compared to the hourly series. Therefore, the combined series spans from 1 h to 16 years (140 256 h).

The model is fitted according to the procedure outlined in Section 9.4.1. First, the climacogram is graphically inspected in order to choose the most suitable form of the climacogram-models given (Table 9.1). This is shown in Figure 9.1. It is evident that the behaviour of the variances switches over larger scales, which makes a type FHK-CD climacogram (Equation (9)) more suitable. Next, in order to account for estimation bias, initial theoretical values of the 7 parameters are assumed, namely of μ (mm/h), λ_1 (mm^2/h^2), λ_2 (mm^2/h^2), α (h), H, θ, and ξ. Once the parameters are assumed, the theoretical values of the variance, the probability wet and the empirical quantiles are known for all scales by virtue of equations shown in Table 9.1. The empirical properties at each timescale, i.e. the variance, the probability wet and the return periods of the non-zero rainfall values, are also estimated. The transition time scale k^* is chosen as 96 h (= 4 d) by inspection of the probability wet, in Figure 9.2.
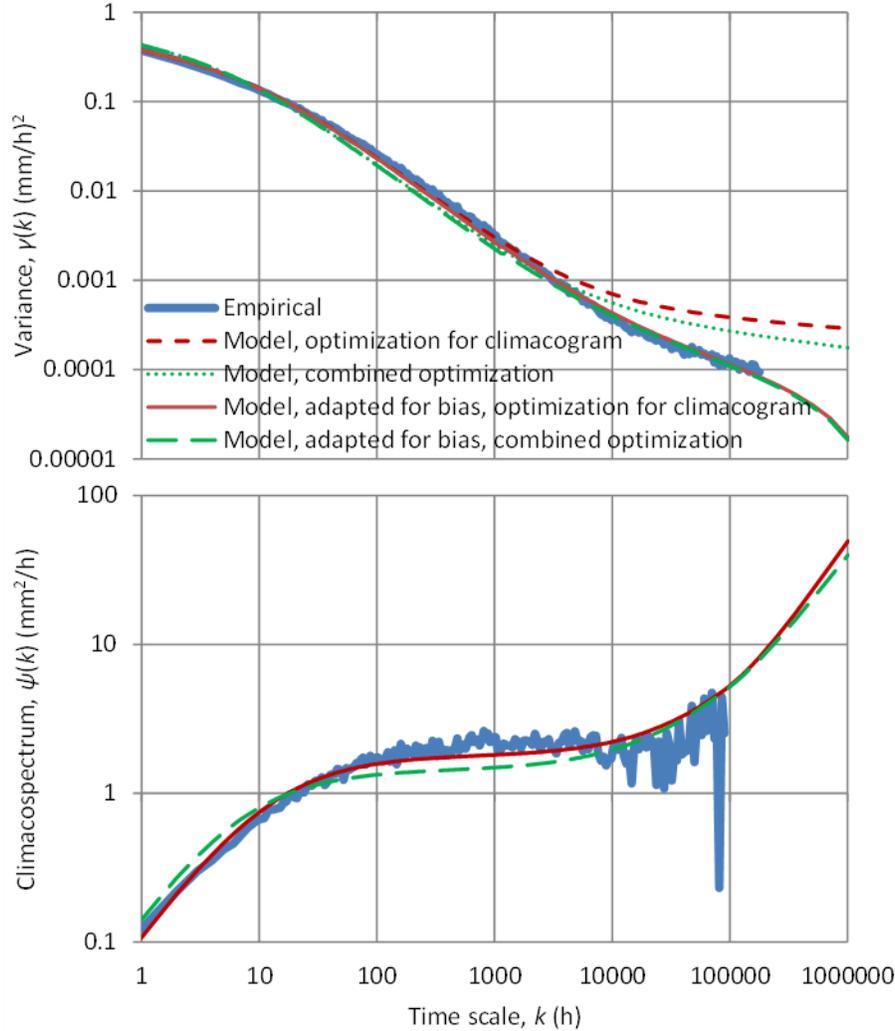
Figure 9.1 Fitting of the ombrian model to the empirical estimates of the climacogram (Equation (9)) and the climacospectrum for Bologna. The empirical estimates for time scales smaller than or greater than 1000 h (~42 d) are taken from the hourly and daily series, respectively. Note that the climacospectrum is defined through the climacogram as $\psi(k) \coloneqq k(\gamma(k) - \gamma(2k))/\ln 2$ and is most appropriate for visualization of the process behaviour at small time scales. *Source*: Koutsoyiannis (2021).

The model is fitted employing four different optimization procedures with the latter three including all the model parameters, i.e. with optimization targeting in minimizing the (a) error in the climacogram, (b) error in the probability wet, (c) error in the rainfall quantiles and (c) combined error in all previous. Optimization scheme (a), yielding the four parameters of the climacogram model, uses Equation (38) with all scales given equal weight, $w_\gamma(k) = 1$. Optimization scheme (b) is based on Equation (39), assuming equal weights for all scales, $w_P(k) = 1$, whereas optimization scheme (c) is based on Equation (40). In this case, using equal weights would result to a model fit biased in favour of lower return periods, which are more frequent in our dataset. To improve the fit to the higher return periods, which are typically the ones of interest, we use a weighting function increasing with return period, i.e. $w_x(k) \propto \sqrt{T}$. Optimization scheme (d) is done using

Equation (41) with weights $a_\gamma = 0.1$, $a_P = 100$, $a_x = 1$. (Note that the chosen high value of a_P counterbalances the fact that E_P is much smaller than the other error components.)
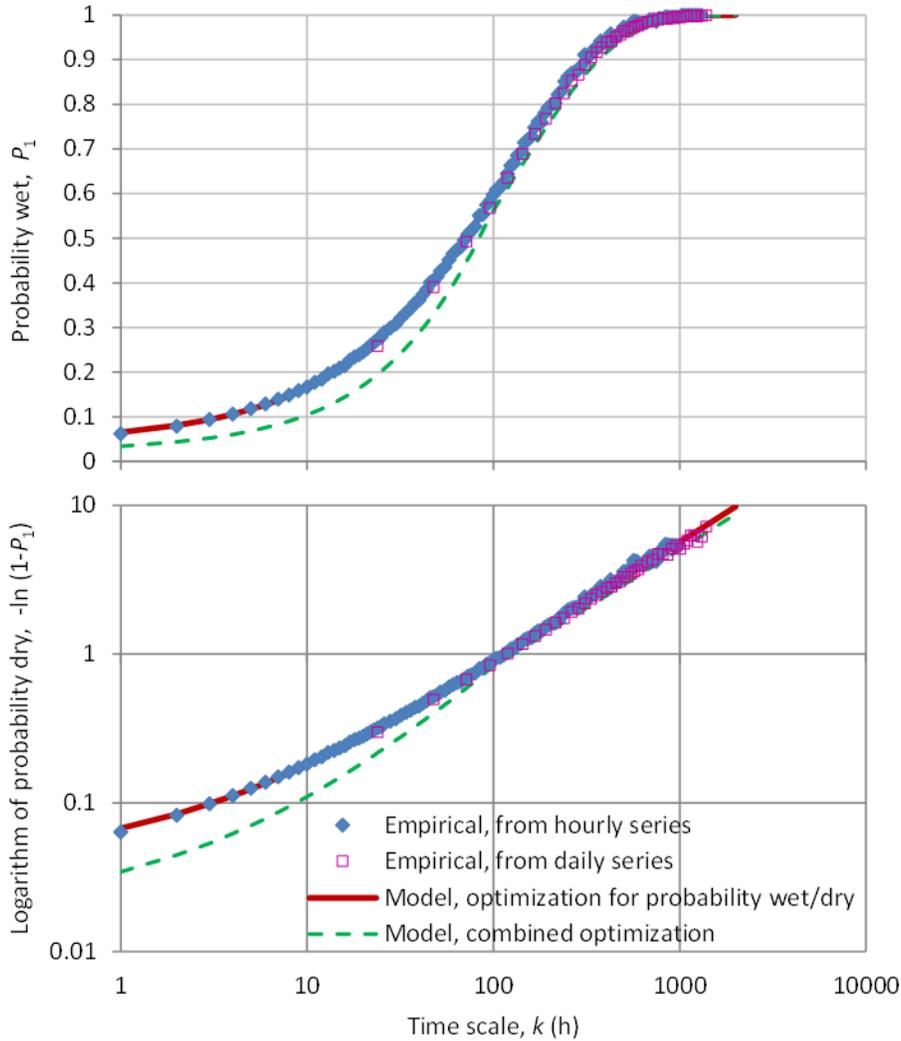

Figure 9.2 Fitting of the ombrian model (Equations (21) and (24)) to the empirical estimates of probability wet and dry of Bologna. *Source*: Koutsoyiannis (2021).

Results from the optimization are shown in Table 9.2. It is interesting to note the high H parameters resulting from all optimization schemes, which yield considerable bias in the climacogram estimation (Figure 9.1). The ombrian model resulting from the combined optimization is shown in Figure 9.3, along with the hourly and daily empirical estimates. In this plot, the bias-adjusted results are plotted in order to be comparable to the empirical estimates; therefore, the true theoretical intensity is higher for scales $k > 1000$ h or about 40 d. Overall, the power of the ombrian model is impressive over the whole range of scales spanning 5 orders of magnitude, i.e. from 1 h to 16 years (Figure 9.3).

Table 9.2 Parameters of the ombrian model of Bologna from the four optimization schemes (table adapted from Koutsoyiannis, 2021).

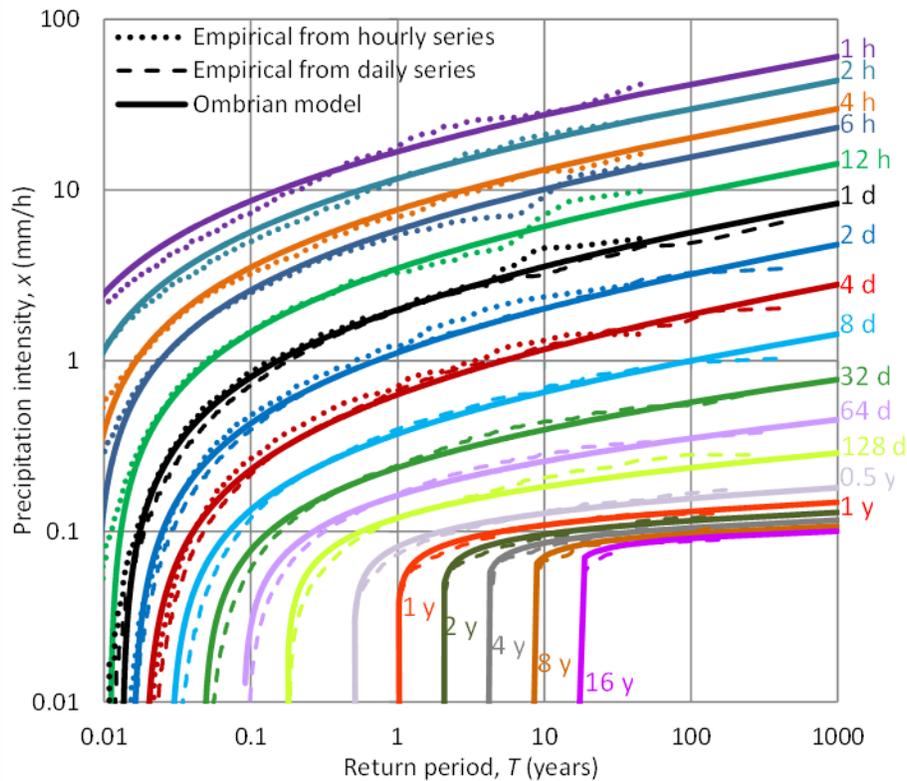| Optimization scheme	μ (mm/h)	λ_1 (mm²/h²)	λ_2 (mm²/h²)	α (h)	H (-)	θ (-)	ξ (-)
(a) Climacogram	-	0.000864	1.51	16.4	0.95	-	-
(b) Probability wet	0.0773	0.00775	0.836	14.15	0.95	0.795	0.121
(c) Quantiles	0.0788	0.00407	1.60	7.70	0.93	0.693	0.125
(d) Combined	0.0823	0.00110	1.43	8.74	0.92	0.787	0.121

Figure 9.3 Ombrian curves derived from the ombrian model for Bologna for time scales spanning 5 orders of magnitude (1 h to 16 years = 140256 h). The empirical points are estimated from order statistics. The ombrian model results are plotted for bias-adapted variance in order to be comparable with empirical plots (thus, for $k > 1000$ h or about 40 d, the true intensity resulting from the model is higher than what is shown in the graph). The abbreviation "y" stands for year. *Source*: Koutsoyiannis (2021).

9.6 Discussion and further aspects

This Section provides an overview of wider research topics within the ombrian modelling framework in light of recent advances in stochastic modelling. Detailed information on these topics can be found in Koutsoyiannis (2021).

9.6.1. On the use of all data

It is strongly suggested within the ombrian model approach to exploit and combine all available sources of information, particularly when these are available at different time scales. This approach was first proposed in Koutsoyiannis et al. (1998) on the basis that daily records are usually of higher reliability compared to the shorter sub-hourly series, which are also subjected to larger measurement uncertainty. The exploitation of information from different metering devices is exemplified by the case study of rainfall in Bologna.

Another basic premise of the proposed methodology is to infer the rainfall properties from the whole parent timeseries. This approach may seem at odds with the common practice of using block maxima or a certain amount of values over threshold, yet it is sounder in terms of retained information on the process properties. Series of extremes tend to hide the persistence of the parent process (Iliopoulou and Koutsoyiannis, 2019), while discarding the body of the distribution, in favour of modelling its tails, has been criticized as wasteful usage of data (Volpi et al., 2019). Current approaches promote the use of the parent process as the natural basis for estimating design quantities in hydrological design. Issues of estimation uncertainty related to classic estimators of higher-order properties (Lombardo et al., 2014) may also be resolved by reliable higher-order estimators of the form of *knowable* moments (Koutsoyiannis, 2019). It is however often the case that only a part of the series is available for some records, for instance the annual maxima. In such cases, the EV2 distribution which corresponds to a Pareto tail, should be used instead for fitting the data. Yet the final model should again be formulated for the Pareto distribution (Koutsoyiannis et al., 1998).

9.6.2 On the estimation of the tail-index

The tail index of the process is one of the most important properties of extremes and also the hardest to estimate from data. In the Bologna case study, the timeseries was long and could support the estimation from the data, yet this is not the case for typical record lengths < 50 years. In such record lengths, the Gumbel distribution may be falsely supported from the data, even when the true distribution is of the EV2 type (Koutsoyiannis, 2004b, 2004a). This may lead to severe underestimation of risk. It is therefore useful to seek longer timeseries from using data from other stations in the region, or even refer to published results from global or large-scale analyses (Koutsoyiannis, 2004b, Papalexiou and Koutsoyiannis, 2013). The latter analyses provided global-scale evidence of the prevalence of a positive shape parameter, with a mean value of $\xi = 0.15$ and 0.114, respectively. It is also stressed that estimates yielding negative shape parameter, which correspond to a process bounded from above, should be discarded as physically unrealistic, and instead replaced by either the Gumbel distribution, or preferably an EV2 type with a regionally-estimated shape parameter.

9.6.3 On the use of a Hershfield coefficient

The study of a statistical property of a timeseries over multiple scales involves some procedure of aggregation of its values for different scales. Typically, the choice of the

starting point for the aggregation is arbitrary, and a change thereof likely results to a different estimate. When studying extremes, it is a common hydrological practice, to either take the maximum estimate resulting from all possible positions of the starting point, or 'inflate' the given estimate by a specific factor, known as the Hershfield coefficient (Hershfield and Wilson, 1957). This practice aims to safer estimates from an engineering point of view. However, it is noted here that when the behaviour of a process is studied in stochastic terms, all realizations are stochastically equivalent and there is no theoretical basis to 'correct' them. In fact, by correcting the series, we distort its stochastic properties by studying, instead of the behaviour of $\underline{x}_\tau^{(k)}$, the behaviour of $\underline{y}_\tau^{(k)} := \max_j\left(\underline{x}_{\tau+j}^{(k)}, j = 0, \dots k - 1\right)$, which is a different stochastic process.

9.6.4 Area-reduction of point ombrian curves

So far, the estimation of ombrian curves has been presented for the case of point rainfall. However, many hydrological applications, especially the estimation of the streamflow process, require estimates of areal rainfall. To account for the spatial variability of point rainfall within a generally homogeneous climatic region, hydrologists have long used the concept of area-reduction factors (ARF). An ARF is defined as the ratio of areally averaged precipitation depth over a certain area A for a specified return period T and time scale k to the precipitation depth over any point of the area (assumed to be climatically homogeneous) for the same return period and time scale (Flammini et al., 2022). To estimate this ratio, samples of both areal precipitation and point precipitation are needed for several scales and return periods. Moreover, according to this definition, samples of areal precipitation should be derived and stochastically analysed for the entire period, rather than calculating ARF values for isolated events, as sometimes performed in the literature. (A comprehensive review on other empirical approaches and definitions of the ARF concept is provided by Svensson and Jones, 2010 and Flammini et al., 2022.)

Extensive investigation on ARF were conducted by the UK by NERC (1975) which resulted to tabulated values of ARF for a wide range of areas (1 to 30 000 km^2) and time scales (1 min to 25 days), ignoring the effect of the return period. Koutsoyiannis and Xanthopoulos (1999, p54) fitted the following relationship to these values:

$$\varphi = \max\left(0.25, 1 - \frac{0.048A^{0.36-0.01\ln A}}{k^{0.35}}\right) \tag{46}$$

where A is the area given in km^2 and k is in h. This relationship has been validated by results from the US by Hershfield and Wilson (1957) for the eastern USA and from the U.S. Weather Bureau (1960) for the western USA. Therefore, it could support ARF estimation in other regions as well.

9.7 Conclusions

Ombrian curves widely known under the misnomer IDF curves are central design tools for a majority of hydrological and engineering tasks. Most of their applications have been based on empirical evidence and hydrological experience. Yet empirically-derived curves

entail prominent theoretical inconsistencies and cannot support simulation beyond the range of observed scales. This Chapter presented the traditional tool of ombrian curves and outlined a methodology to traverse their limitations toward building consistent and more powerful stochastic models of rainfall intensity, i.e. *ombrian models*. Going from ombrian curves to models requires understanding the assumptions that are implicit in traditional curves and revisiting thereof through stochastic modelling of the parent process.

Two modelling versions are provided; a simplified relationship valid for small-scales, and an all-scale ombrian model, i.e., covering all the range of available scales. Particular emphasis is devoted to the fitting procedure in which issues of bias and data uncertainty are discussed and addressed. It is shown how to account for the effect of dependence-induced bias and further how to combine information from multiple data sources. The entire methodology is illustrated by the case study of rainfall in Bologna, which also stands as a proof of concept of the exceptional performance of the ombrian model from hourly to 16-years scale.

Advancing empirically-derived ombrian curves to theoretically-consistent ombrian models, allows the user to address bias and estimation uncertainty, extrapolate results to longer timescales and perform simulation for complex hydrological systems. These theoretical and practical gains are manifold, while the operational character of traditional ombrian curves is preserved.

References

Ayman, G.A., Mohamed, E., Ashraf, E., Hesham, E., 2011. Developing intensity-duration frequency curves in scarce data region: an approach using regional analysis and satellite data. Engineering 3 (3), 215–226. doi:10.4236/eng.2011.33025.

Bernard, M.M., 1932. Formulas for rainfall intensities of long duration. Trans. Am. Soc. Civil Eng. 96 (1), 592–606. doi:10.1061/TACEAT.0004323.

Chow, V.T., Maidment, D.R., Mays, L.W., 1988. Applied Hydrology. McGraw-Hill, New York.

Dimitriadis, P., Koutsoyiannis, D., 2015. Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes. Stoch. Environ. Res. Risk. Assess. 29 (6), 1649–1669. doi:10.1007/s00477-015-1023-7.

Eagleson, P.S.: Dynamic hydrology,1970.

Flammini, A., Dari, J., Corradini, C., Saltalippi, C., Morbidelli, R., 2022. Areal Reduction Factor Estimate for Extreme Rainfall Events. In: Morbidelli, R. (Ed.), Rainfall. Modelling, Measurement and Applications. Elsevier, Amsterdam, pp. 285–306. doi:10.1016/C2019-0-04937-0.

Hailegeorgis, T.T., Thorolfsson, S.T., Alfredsen, K., 2013. Regional frequency analysis of extreme precipitation with consideration of uncertainties to update IDF curves for the city of Trondheim. J. Hydrol. 498, 305–318. doi: 10.1016/j.jhydrol.2013.06.019.

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

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

Iliopoulou, T., Koutsoyiannis, D., 2019. Revealing hidden persistence in maximum rainfall records. Hydrol. Sci. J. 64 (14), 1673–1689. doi:10.1080/02626667.2019.1657578.

Koutsoyiannis, D., 2003. Climate change, the Hurst phenomenon, and hydrological statistics. Hydrol. Sci. J. 48 (1), 3–24. doi:10.1623/hysj.48.1.3.43481.

Koutsoyiannis, D., 2004a. Statistics of extremes and estimation of extreme rainfall, 2, empirical investigation of long rainfall records. Hydrol. Sci. J. 49 (4), 591–610, doi:10.1623/hysj.49.4.575.54430.

Koutsoyiannis, D., 2004b. Statistics of extremes and estimation of extreme rainfall, 1, theoretical investigation. Hydrol. Sci. J. 49 (4), 575–590, doi:10.1623/hysj.49.4.591.54424.

Koutsoyiannis, D., 2006. An entropic-stochastic representation of rainfall intermittency: The origin of clustering and persistence. Water Resour. Res. 42 (1), W01401. doi:10.1029/2005WR004175.

Koutsoyiannis, D., 2019. Knowable moments for high-order stochastic characterization and modelling of hydrological processes. Hydrol. Sci. J. 64 (1), 19–33. doi:10.1080/02626667.2018.1556794.

Koutsoyiannis, D., 2021. Stochastics of Hydroclimatic Extremes - A Cool Look at Risk, 1st ed. Kallipos, Athens, p. 333, ISBN: 978-618-85370-0-2. http://www.itia.ntua.gr/en/docinfo/2000/ (Accessed 27 May 2021).

Koutsoyiannis, D., 2017. Entropy production in stochastics. Entropy 19 (11), 581. doi:10.3390/e19110581.

Koutsoyiannis, D., Dimitriadis, P., Lombardo, F., Stevens, S., 2018. From fractals to stochastics: Seeking theoretical consistency in analysis of geophysical data. Advances in Nonlinear Geosciences. Springer, Cham, Switzerland, pp. 237–278.

Koutsoyiannis, D., Kozonis, D., Manetas, A., 1998. A mathematical framework for studying rainfall intensity-duration-frequency relationships. J. Hydrol. 206 (1–2), 118–135. doi:10.1016/S0022-1694(98)00097-3.

Koutsoyiannis, D., Papalexiou, S.M., 2016. Extreme rainfall: Global perspective. Chow's Handbook of Applied Hydrology, 2nd edition. McGraw-Hill, New York.

Koutsoyiannis, D., Xanthopoulos, T., 1999. Engineering Hydrology, 3rd edition. National Technical University of Athens, Athens (in Greek), p. 418. doi:10.13140/RG.2.1.4856.0888.

Langousis, A., Veneziano, D., 2007. Intensity-duration-frequency curves from scaling representations of rainfall. Water Resour. Res. 43 (2). doi:10.1029/2006WR005245.

Lombardo, F., Volpi, E., Koutsoyiannis, D., Papalexiou, S.M., 2014. Just two moments! A cautionary note against use of high-order moments in multifractal models in hydrology. Hydrol. Earth Syst. Sci. 18 (1), 243–255. doi:10.5194/hess-18-243-2014.

NERC (Natural Environment Research Council), 1975. Flood Studies Report. Institute of Hydrology, Wallingford, UK.

Papalexiou, S.M., Koutsoyiannis, D., 2013. Battle of extreme value distributions: a global survey on extreme daily rainfall. Water Resour. Res. 49 (1), 187–201. doi:10.1029/2012 WR012557.

Papoulis, A., 1991. Probability, Random Variables, and Stochastic Processes, 3rd edition. McGraw-Hill, New York.

Sherman, C.W., 1931. Frequency and intensity of excessive rainfalls at Boston, Massachusetts.Trans. Am. Soc. Civil Eng. 95 (1), 951–960. doi:10.1061/TACEAT.0004286.

Svensson, C., Jones, D.A., 2010. Review of methods for deriving areal reduction factors. J.Flood Risk Manag. 3 (4), 232–245. doi:10.1111/j.1753-318X.2010.01079.x.

U.S. Weather Bureau, 1960. Generalized Estimates of Probable Maximum Precipitation West of the 105th Meridian, Technical Paper No 38. U.S. Department of Commerce, Washington, DC. https://www.nws.noaa.gov/oh/hdsc/Technical_papers/TP38.pdf.

Veneziano, D., Furcolo, P., 2002. Multifractality of rainfall and scaling of intensity-duration-frequency curves. Water Resour. Res. 38 (6), 42–1 doi:10.1029/2005WR004716.

Volpi, E., Fiori, A., Grimaldi, S., Lombardo, F., Koutsoyiannis, D., 2019. Save hydrological observations! Return period estimation without data decimation. J. Hydrol. 571, 782–792. doi:10.1016/j.jhydrol.2019.02.017.

Willems, P., 2000. Compound intensity/duration/frequency-relationships of extreme precipitation for two seasons and two storm types. J. Hydrol. 233 (1–4), 189–205. doi:10.1016/S0022-1694(00)00233-X.