European Geosciences Union General Assembly 2008

Vienna, Austria, 13-18 April 2008

Session IS24: Precipitation: from measurement to modelling and application in catchment hydrology

Ombrian curves in a maximum entropy framework

S.M. Papalexiou and D. Koutsoyiannis

Department of Water Resources and Environmental Engineering
National Technical University of Athens

(www.itia.ntua.gr)

1. Abstract

Ombrian curves (from the Greek ombros, meaning rainfall) are most widely known as rainfall intensity-duration-frequency (IDF) curves or relationships. However, the former term may be preferable as the later is inaccurate. Namely, "frequency" is meant to be "return period" whereas "duration" is in fact the "time scale" on which the rainfall process is averaged. Thus, ombrian relationships are nothing more than multiple time scale expressions of the rainfall probability. Three important issues regarding the mathematical form of the ombrian relationships are examined: (a) whether or not the effects of time scale and return period are separable so that the relationship could be written as the product of two scalar functions; (b) whether or not the rainfall intensity is a power function of return period and (c) whether or not the rainfall is a power function of time scale. All these questions are investigated using the principle of maximum entropy as a theoretical basis and a long rainfall data set as an empirical basis. It turns out that none of the above questions has a precisely positive answer, which makes the theoretical derivation of ombrian curves a complicated task. For this reason, consistent approximations are sought, which eventually do not depart significantly from commonly used forms in engineering practice.

2. Ombrian relationships

- Ombrian relationships (or curves), mostly known as rainfall intensity-duration-frequency (IDF) relationships, determine the rainfall intensity $i(k,T)$ averaged over time scale (not duration) k and exceeded on a return period T.
- These relationships are rather empirical and generally have the form $i(k,T) = f(T)\, g(k)$, where $f(T)$ and $g(k)$ are mathematical functions whose simplest and most common forms are $f(T) = \lambda\, T^{\kappa}$ and $g(k) = k^{-\eta}$. The parameters λ, κ, η are determined from the data.
- Clearly, these formulae imply: (a) a separable functional dependence of i on T and k, (b) a power function of i *vs.* T, and (c) a power function of i *vs.* k.
- In this study we try to investigate the validity of these three assumptions based on probabilistic considerations and using the principle of maximum entropy as a solid theoretical background.
- Additionally, we seek a statistical distribution based on theoretical principles capable of describing rainfall on all time scales. This would be the first step in constructing consistent ombrian relationships.

3. The data set

We have used a 70-year long hourly rainfall data set from the station of National Observatory of Athens, Greece. The table presents statistics of rainfall intensity averaged over several time scales.

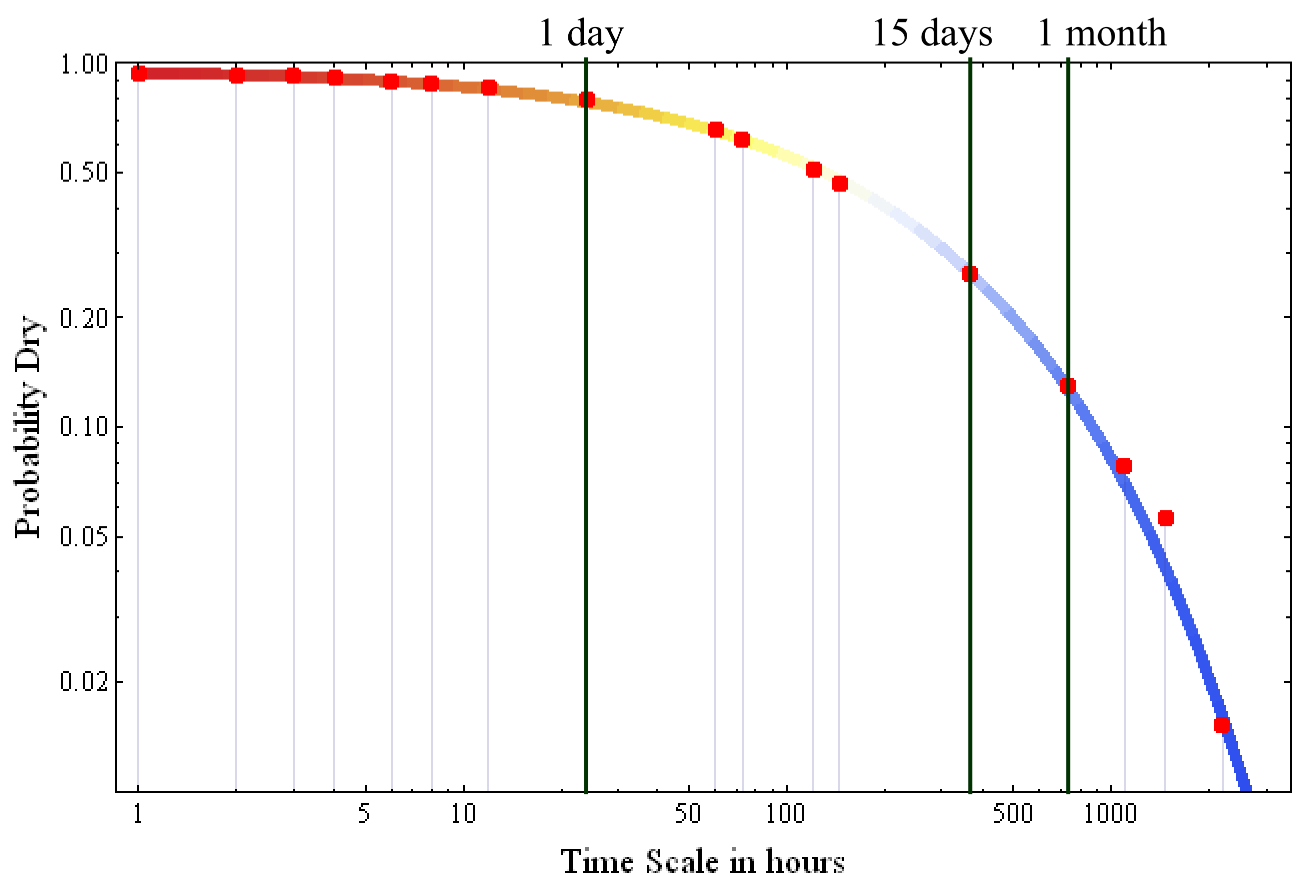
Scale	Length	Missing	Zeros	Positive	P_{Dry}	Mean	StDev	C_v	C_s	C_k	Max
1 h	613632	29197	551718	32717	0.94	0.70	1.71	2.43	7.36	106.92	58.56
2 h	306816	14677	272376	19763	0.93	0.58	1.30	2.24	6.53	87.51	38.25
3 h	204544	9785	179907	14852	0.92	0.51	1.08	2.10	6.34	96.71	33.98
4 h	153408	7427	133313	12668	0.91	0.45	0.92	2.05	5.89	82.08	26.71
6 h	102272	4977	87226	10069	0.90	0.38	0.74	1.96	5.47	69.00	19.25
8 h	76704	3806	64267	8631	0.88	0.33	0.62	1.89	4.73	47.19	13.36
12 h	51136	2531	41774	6831	0.86	0.28	0.50	1.80	5.25	67.89	12.15
24 h	25568	1335	19194	5039	0.79	0.18	0.31	1.69	4.24	40.41	6.08
2.5 d	10028	592	6190	3246	0.66	0.11	0.17	1.51	3.61	25.67	2.39
3 d	8488	512	4955	3021	0.62	0.10	0.15	1.47	3.37	22.75	2.03
5 d	4988	343	2371	2274	0.51	0.08	0.11	1.38	3.24	20.17	1.20
6 d	4218	308	1825	2085	0.47	0.07	0.09	1.31	2.83	15.62	1.01
15 d	1628	169	385	1074	0.26	0.05	0.06	1.15	2.56	15.03	0.60
1 month	840	129	93	618	0.13	0.04	0.04	0.98	1.55	6.43	0.32
1.5 months	560	103	37	420	0.08	0.04	0.04	0.93	1.39	5.52	0.25
2 months	420	98	18	304	0.06	0.04	0.03	0.83	1.03	4.08	0.18
3 months	280	83	3	194	0.02	0.04	0.03	0.84	0.97	3.51	0.16
4 months	210	73	0	137	0.00	0.04	0.03	0.71	0.57	3.03	0.13
6 months	140	62	0	78	0.00	0.04	0.01	0.36	0.64	4.10	0.09
1 year	70	44	0	26	0.00	0.04	0.01	0.27	0.38	3.05	0.07

4. Probability dry

- Ombrian relationships are closely related to the probability of a time interval being dry, as this probability affects the average rainfall intensity at different time scales *k* (particularly the moderate ones).
- The variation of probability dry with time scale has been investigated elsewhere (Koutsoyiannis, 2006) based on maximum entropy theoretical considerations.
- The theoretical relationship that describes the probability dry in any time scale (k) is given by

$$p^{(k)} = p^{\left[1+\left(\xi^{-1/\eta}-1\right)(k-1)\right]^{\eta}}$$

where $p = p^{(1)}$ and ξ and η are parameters.



5. Entropy maximization and marginal distributions

The Boltzmann-Gibbs-Shannon (BGS) entropy for a continuous random variable X with density function $f(x)$ is by definition (e.g. Shannon, 1949; Papoulis, 1991)

$$\varphi = E[-\ln f(x)] = -\int_{-\infty}^{\infty} f(x) \ln f(x) dx$$

Maximization of BGS entropy with simple constraints of known mean μ and variance σ^2 and the non-negativity constraint results in

$$f(x) = \exp(-\lambda_0 - \lambda_1 x - \lambda_2 x^2), \quad x \geq 0 \quad (1)$$

where λ_0, λ_1 and λ_2 are parameters depending on the known mean and variance. Inspection of (1) shows that it is none other than the truncated normal density function.

A generalization of the BGS entropy, effectively used in numerous scientific disciplines and also valuable in hydrology has been proposed by Tsallis (1988, 2004):

$$\varphi_q = \frac{1 - \int_{-\infty}^{\infty} [f(x)]^q}{q - 1}$$

with $q = 1$ corresponding to the BGS entropy. Maximization of Tsallis entropy with known μ and σ^2 yields an over-exponential (power-type) distribution,

$$f(x) = [1 + \xi\,(\lambda_0 + \lambda_1 x + \lambda_2 x^2)]^{-1-1/\xi}, \quad x \geq 0 \quad (2)$$

where λ_0, λ_1, λ_2 parameters and $\xi := (1 - q)/q$. The truncated normal distribution fails to describe cases in which the variation $\sigma/\mu > 1$, and, as high variation is common in hydrological variables at fine time scales, this is an indication of the applicability of Tsallis entropy in hydrology (Koutsoyiannis, 2005).

6. A stepwise entropy maximization approach

- Let X_i ($X_i \geq 0$), denote the rainfall rate at time i discretized at a fine time scale (tending to zero) and let us assume that it has a specified mean μ.
- Maximization of BGS entropy with constraints $X_i \geq 0$ and $E[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx = \mu$ results in the exponential distribution: $f(x) = \exp(-x/\mu)/\mu$.
- In addition, let us assume that there is some time dependence of X_i, quantified by $E[X_i\, X_{i+1}] = \gamma$; this will introduce an additional constraint for the multivariate distribution

$$E[X_i\, X_{i+1}] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_i\, x_{i+1}\, f(x_i, x_{i+1})\, dx_i\, dx_{i+1} = \gamma = \rho\, \sigma^2 + \mu^2$$

  Here ρ is the correlation coefficient ($\rho > 0$) and σ is the standard deviation (for the exponential distribution $\sigma = \mu$ and thus $\gamma = \rho\, \sigma^2 + \mu^2 = (\rho + 1)\, \mu^2 > \mu^2$).
- The constant mean constraint in rainfall modelling does not result from a natural principle and although it is reasonable to assume a specific mean rainfall, we can allow this to vary in time.
- In this case we can assume that the mean at time i is the realization of a random process M_i which has mean μ and lag 1 autocorrelation $\rho^M > \rho$.
- Application of the ME principle will produce that M_i is Markovian with exponential distribution.
- Then application of conditional distribution algebra results in

$$f(x) = 2\, K_0(2\, (x/\mu)^{1/2})/\mu, \quad F(x) = 1 - 2\, (x/\mu)^{1/2}\, K_1(2\, (x/\mu)^{1/2})/\mu$$

  where $K_n(x)$ is the modified Bessel function of the second kind (important observation: $f(0) = \infty$, whereas in the exponential distribution $f(0) = \mu < \infty$).

7. A rainfall distribution for all time scales

- Extensive application of the Tsallis distribution (in this study) showed that it fails to simultaneously describe both tails of the empirical probability distribution. It seems that the density $f_X(x)$ of a distribution appropriate to describe rainfall on small time scales, should tend to infinity as x tends to 0, thus achieving a good fit in the region near 0. This basic characteristic is theoretically consistent with the stepwise entropy maximization approach (see panel 6 and Koutsoyiannis, 2008).
- We propose here as a rainfall distribution for all time scales a generalization of the beta prime distribution; this has been called in Koutsoyiannis (2005) the Power-transformed Beta Prime distribution. Here it is referred to as the JH distribution. While the JH distribution does not result from entropy maximization with simple constrains, it is consistent with two essential features theoretically derived from entropy maximization: (a) it is a power type distribution and (b) $f_X(x) \rightarrow \infty$ as $x \rightarrow 0$.
- The probability density function and the distribution function of the JH distribution are:

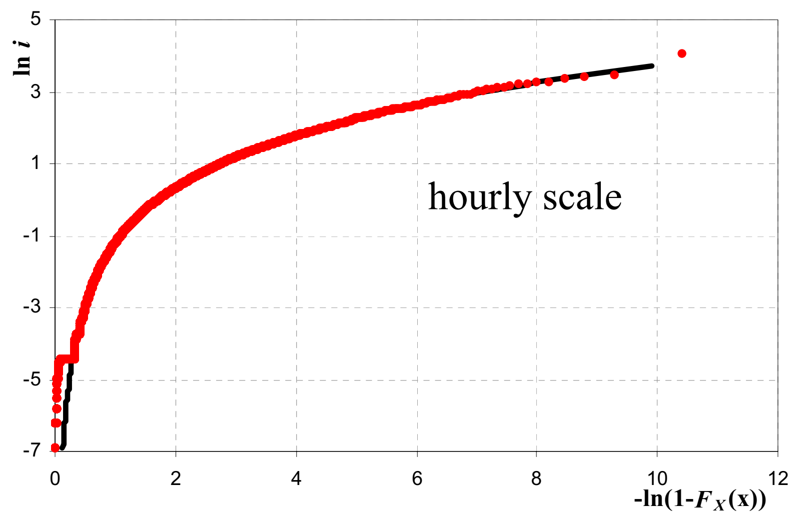
$$f_X(x) = \nu_2 \, a^{\frac{1-\nu_1}{\nu_2}} \, \mathrm{B}\left(\nu_3 - \frac{1-\nu_1}{\nu_2}, -\frac{1-\nu_1}{\nu_2}\right)^{-1} x^{-\nu_1} \left(1 + a\, x^{\nu_2}\right)^{-\nu_3}$$

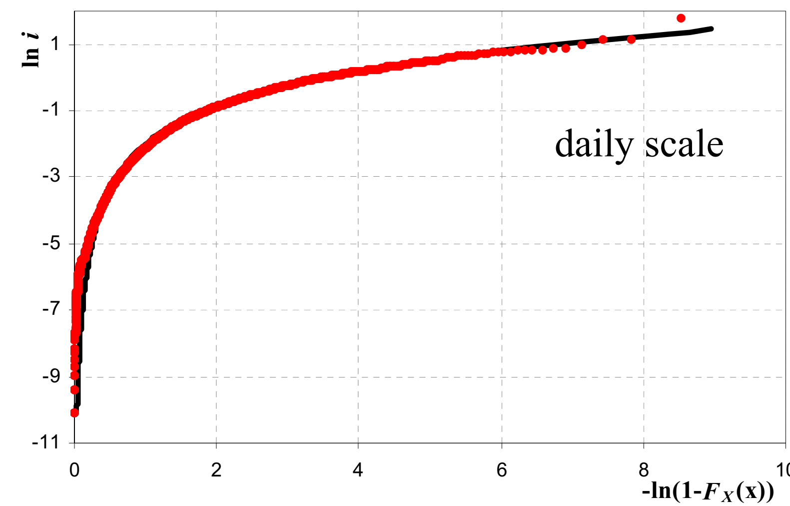
$$F_X(x) = \nu_2 \, a^{\frac{1-\nu_1}{\nu_2}} \left[(1-\nu_1)\, \mathrm{B}\left(\nu_3 - \frac{1-\nu_1}{\nu_2}, -\frac{1-\nu_1}{\nu_2}\right)\right]^{-1} x^{1-\nu_1} \, {}_2F_1\left(\frac{1-\nu_1}{\nu_2}, \nu_3; 1 + \frac{1-\nu_1}{\nu_2}; -a\, x^{\nu_2}\right)$$

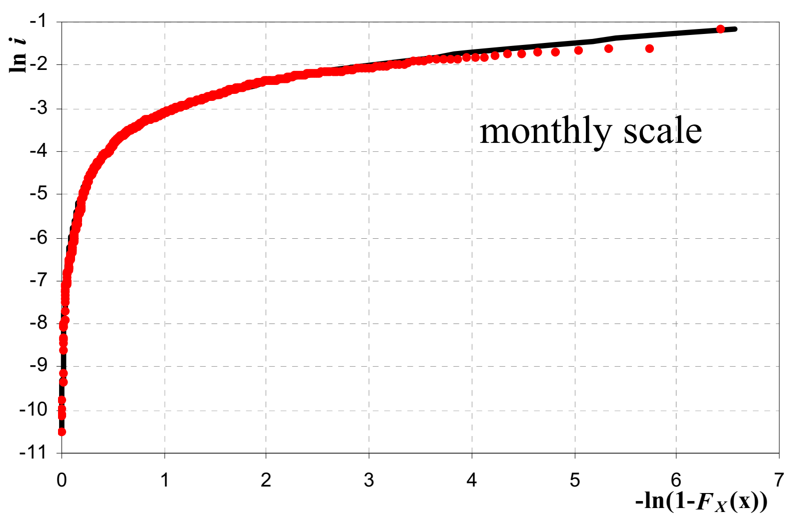
with $\alpha > 0$, $\nu_1 \in$ R, $\nu_2 > 0$, $\nu_3 > 0$ and where ${}_2F_1$ is the Gauss's hypergeometric function 2F1 and B is the Beta function.

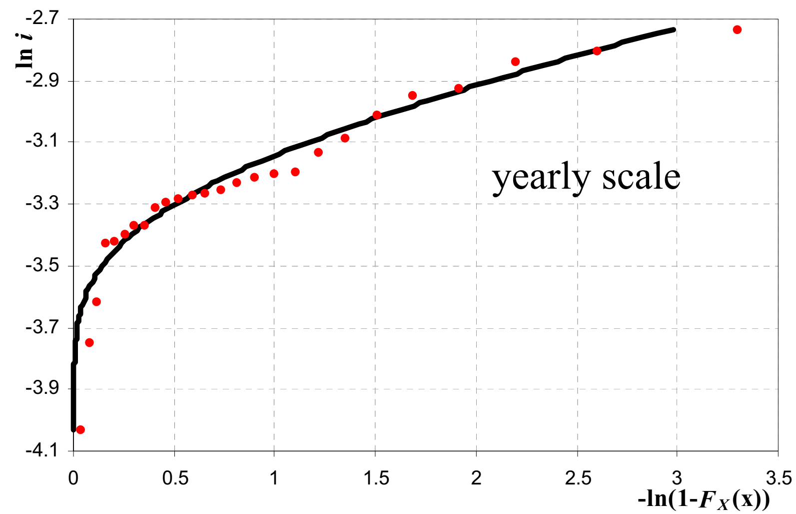
8. The JH distribution in action

The JH distribution performed exceptionally well as the following figures attest for time scales from 1 h to 1 year (~4 orders of magnitude). Asymptotically (for $\nu_3 > 0$) the distribution behaves as $x^{-(\nu_1 + \nu_2 \nu_3)}$ (for our data set $x^{-(\nu_1 + \nu_2 \nu_3)} \approx x^{-7.66}$).







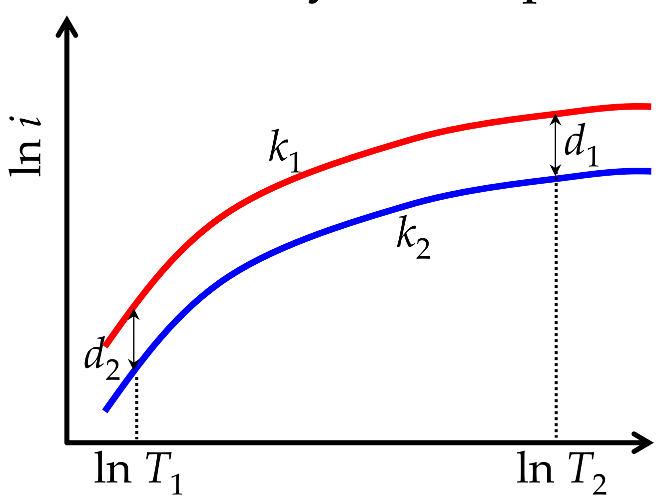


9. Are the effects of time scale and return period separable?

- If $i(k,T) = f(T)\, g(k)$ then for two different time scales k_1, k_2:

  $i(k_1,T) = f(T)\, g(k_1)$, $i(k_2,T) = f(T)\, g(k_2) \rightarrow i(k_1,T)/\, i(k_2,T) = g(k_1) \,/\, g(k_2)$, i.e. independent of T, or

  $\ln i(k_1,T) - \ln i(k_2,T) = \ln g(k_1) - \ln g(k_2) = h(k_1, k_2)$, i.e. independent of T.
- The final equation represents a simple translation in terms of time scale k, the same for any return period T.



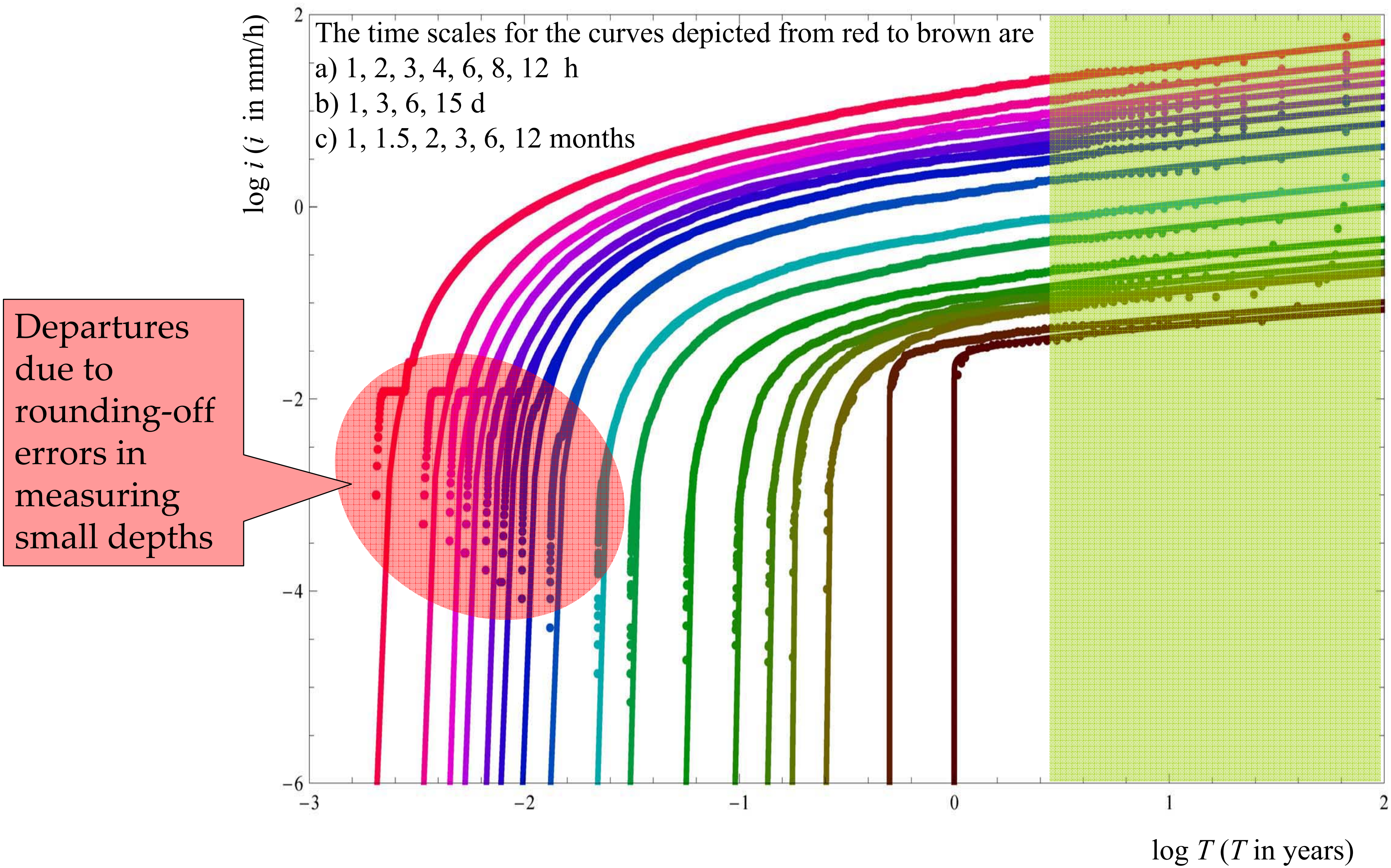
In a translation (as above):

$d_1 = d_2$

- The figure in panel 10 shows that this does not hold true (the distances of two curves for specified k_1, k_2 increase with the decrease of T).
- Therefore the answer is negative, except in high return periods (green area).

10. Is rainfall a power function of return period?

No, except in high return periods (green area).



11. Is rainfall is a power function of time scale?

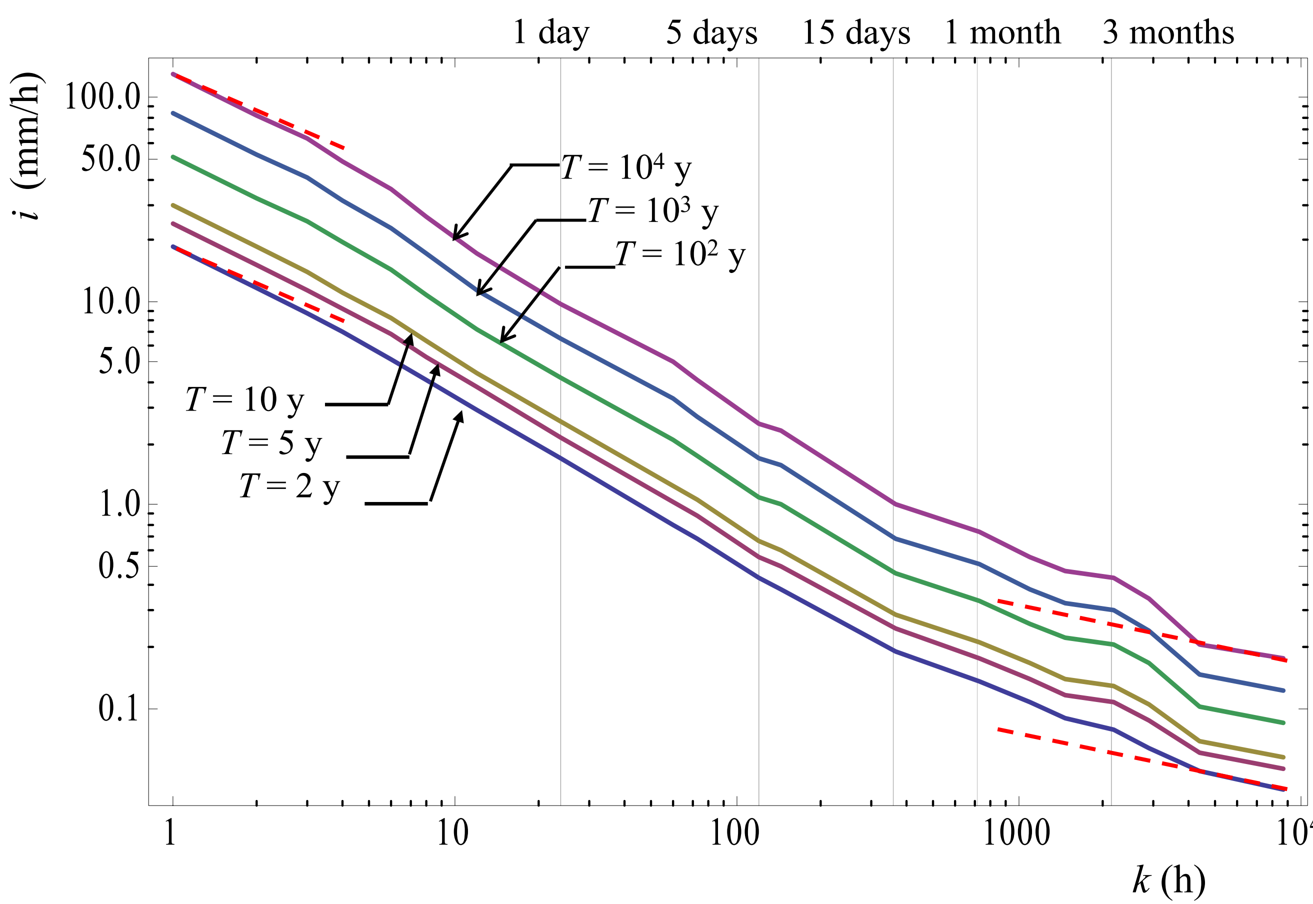
- To study this question, we have constructed double logarithmic plots of rainfall intensity *i vs.* time scale *k* for several values of return period *T*.
- The diagram shows that a power law is a good approximation – but not perfect in either of the two ends.
- A general relationship with better fit, produced by maximum entropy considerations (Koutsoyiannis, 2006) is:

$$g(k) = \frac{\alpha}{1 - p^{\left(\frac{k+\theta}{1+\theta}\right)^{\eta}}}$$

where α, θ, η and $p < 1$ are parameters

- For small time scales this can be approximated as:

$$g(k) = \frac{\alpha'}{(k+\theta)^{\eta}}$$



12. Conclusions and discussion

- Maximum entropy considerations help to construct a probabilistic model (the JH distribution) with an impressively good fit to rainfall intensity data for times scales ranging from hourly to yearly (a range of 4 orders of magnitude).
- The model and the data show that common assumptions in empirical ombrian relationships, i.e., (a) a separable functional dependence of the rainfall intensity i on the return period T and time scale k, (b) a power function of i *vs.* T, and (c) a power function of i *vs.* k, are not verified theoretically.
- However, for moderate and large return periods, these assumptions provide a very good approximation and justify the wide use of these assumptions in the engineering practice ($i = \alpha\ T^{\kappa} / k^{\eta}$).
- Better yet simple ombrian relationships can be derived with approximating the consistent distribution function with the Pareto distribution and also exploiting the generalized dependence of intensity with duration (slide 11); in this case a better approximation is $i = \alpha\ (T^{\kappa} - \psi\ ) / (k + \theta)^{\eta}$.

References

Koutsoyiannis, D. (2005), Uncertainty, entropy, scaling and hydrological stochastics, 2, Time dependence of hydrological processes and time scaling, *Hydrological Sciences Journal*, 50(3), 405-426.

Koutsoyiannis, D. (2006), An entropic-stochastic representation of rainfall intermittency: The origin of clustering and persistence, *Water Resources Research*, 42 (1), W01401.

D. Koutsoyiannis (2008), Long tails of marginal distribution and autocorrelation function of rainfall produced by the maximum entropy principle, European Geosciences Union General Assembly 2008, *Geophysical Research Abstracts*, Vol. 10, Vienna, 00702, European Geosciences Union.

Papoulis, A. (1991), *Probability, Random Variables, and Stochastic Processes* (third edn.), McGraw-Hill, New York.

Shannon, C. E. and Weaver, W. (1949), *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana, Illinois.

Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics, *J. Stat. Phys.*, 52, 479-487.

Tsallis, C. (2004). Nonextensive statistical mechanics: construction and physical interpretation, in *Nonextensive Entropy, Interdisciplinary Applications* (edited by M. Gell-Mann & C. Tsallis), Oxford University Press, New York, NY.