European Geosciences Union General Assembly 2005 — Vienna, Austria, 24 - 29 April 2005 — Session P3.01 Scaling and nonlinearity in the hydrological cycle

THE SCALING PROPERTIES IN THE DISTRIBUTION OF HYDROLOGICAL VARIABLES AS A RESULT OF THE MAXIMUM ENTROPY PRINCIPLE

D. Koutsoyiannis, Department of Water Resources, National Technical University of Athens

1. Abstract

It is well known that the principle of maximum entropy (ME), when applied to the probability distribution of a random variable with known mean and variance, results in the normal distribution. If the variable is non-negative, as happens with hydrological variables such as rainfall and runoff, the same principle results in the truncated normal distribution. Mathematically, this distribution can have a coefficient of variation ranging from zero to unity, with the upper bound corresponding to the exponential distribution. At fine time scales, rainfall and runoff have coefficients of variations higher than one, so the classical entropy approach, constrained by known mean and variance, is not applicable. However, a generalization of entropy (specifically the use of the concept of nonextensive entropy) allows the application of the ME principle even in such cases and results in power-type distributions, which for low probabilities of exceedence have scaling properties. Thus, the ME principle can be used to infer the type of the distribution of a hydrological variable, i.e. whether it has scaling properties or not, and to quantify the scaling exponent using a simple indicator such as the coefficient of variation. This theoretical framework is validated with several real world examples concerning rainfall, runoff and temperature data at several time scales. Given that entropy is a measure of uncertainty, the applicability of the ME principle to the distribution of hydrological variables emphasizes the dominance of uncertainty in hydrological processes.

2. Type-"which?" versus type-"why?" questions

The typical questions in hydrological statistics are type-"which?": Which is the most appropriate theoretical distribution for a hydrological quantity such as rainfall and runoff? The selection is done from a repertoire that contains distributions such as normal, lognormal, Pearson, log-Pearson, Weibull, Extreme Value of maxima and minima of types I, and II (EV1, EV2, EV3) and is based on comparisons with empirical distributions.

Generally, hydrological statistics avoids questions of the type "why?", which would provide an explanation of the appropriateness or inappropriateness of a certain distribution and thus would also help choose the most appropriate distribution.

The importance of type-"why?" questions becomes more evident when studying extreme events. Obviously, observations of extreme events cannot be as numerous as those of regular events. Therefore, a theoretical reasoning of the appropriateness of a distribution, in addition to the empirical study of the data, would help for a more justified and correct choice.

3. Gaussian versus scaling distributional behaviour

In few cases, explanations of the distributional behaviour of hydrological quantities exist. For example, the Central Limit Theorem (CLT) explains the emerging of the normal distribution, e.g. in annual rainfall in a wet area.

The normal distribution applies also in a lot of cases not explained by CLT (e.g. in daily temperatures).

Recently, evidence has been accumulated in geophysics (including hydrology, e.g. in distribution of rainfall and runoff at daily or finer time scales), in economics and even in humanities, of another distributional behaviour, equally common to, and simultaneously different from normal. This is the state scaling behaviour, characterized by an asymmetric J-shaped density function $f(x)$ and a scaling property of its survival function $F^*(x)$:

$$F^*(l\,x) = l^{-1/\kappa}\,F^*(x)$$

for any $l > 0$ and a specified parameter $\kappa > 0$. This results in

$$F^*(x) = [\lambda/(\kappa\,x)]^{1/\kappa},\ f(x) = (1/\lambda)\,[\lambda/(\kappa\,x)]^{1/\kappa},\ x = (\lambda/\kappa)\,(T/\delta)^{\kappa}$$

where T the return period, δ the interarrival time of an event and λ a parameter.

4. The enrolment of entropy and its definition

The explanation for the two different limiting distributional behaviours is sought upon the **entropy** concept and the **principle of maximum entropy**.

For a discrete random variable X taking values x_j with probability mass function $p_j \equiv p(x_j)$, the Boltzmann-Gibbs-Shannon (or extensive) entropy is defined as

$$\phi := E[-\ln p(X)] = -\sum_{j=1}^{w} p_j \ln p_j, \quad \text{where} \quad \sum_{j=1}^{w} p_j = 1$$

For a continuous random variable X with probability density function $f(x)$, the entropy is defined as

$$\phi := E[-\ln f(X)] = -\int_{-\infty}^{\infty} f(x) \ln f(x)\,dx, \quad \text{where} \quad \int_{-\infty}^{\infty} f(x)\,dx = 1$$

In both cases the entropy ϕ is a measure of **uncertainty** about X and equals the information gained when X is observed (Papoulis, 1991).

In other disciplines (statistical mechanics, thermodynamics, dynamical systems, fluid mechanics), entropy is regarded as a measure of **order** or **disorder** and **complexity**.

5. Generalization of entropy

Tsallis (1988) heuristically generalized the Boltzmann-Gibbs-Shannon entropy by postulating the entropic form

$$\phi_q := \frac{1 - \sum_{i=0}^{w} p_i^q}{q-1} \ \text{(for discrete } x\text{)}, \qquad \phi_q := \frac{1 - \int_0^{\infty} [f(x)]^q\,dx}{q-1} \ \text{(for continuous } x\text{)}$$

where q is any real number. This has been called Tsallis entropy or nonextensive entropy and remedies disabilities or inconsistencies in the use of the classical entropy. For $q \to 1$ this precisely reproduces the Boltzmann-Gibbs-Shannon entropy, i.e., $\phi_1 \equiv \phi$.

Other generalized entropic forms have been also proposed, i.e. the Rényi entropy and the normalized entropy; these can be expressed as monotonic functions of Tsallis ϕ_q.

6. The principle of maximum entropy (ME)

In a probabilistic context, the ME principle was introduced by Janes (1957) as a generalization of the "principle of insufficient reason" (PIR) attributed to Bernoulli or to Laplace.

The ME principle is used to infer unknown probabilities from known information.

It is related to the homonymous physical principle that determines thermodynamical states.

It postulates that the entropy of a random variable should be at maximum, under some conditions, formulated as constraints, which incorporate the information that is given about this variable.

For a discrete x taking a finite number of possible states w, in the absence of any information, both ϕ and ϕ_q (for any $q > 0$) achieve their maximum values at equiprobability ($p_i = 1/w$, corresponding to PIR).

For infinite number of possible states, constraints are necessary. Typical constraints used in a probabilistic or physical context are:

Mass — Mean/Momentum

$$\int_{-\infty}^{\infty} f(x)\,dx = 1, \quad E[X] = \int_{-\infty}^{\infty} x\,f(x)\,dx = \mu$$

Variance/Energy — Non-negativity

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \sigma^2 + \mu^2, \quad x \geq 0$$

7. Application of the ME principle to hydrology

We use the above four constraints, which involve two parameters, the mean μ and the standard deviation σ, estimated from the sample.

The non-negativity constraint is essential for hydrological variables. In the case that a variable has a lower bound $\neq 0$, a shift is required to make it 0.

Without loss of generality, before application of ME we can standardize the variable by its mean μ, making it have mean 1 and standard deviation σ/μ, which is the coefficient of variation (CV) of the original variable.

Thus, the ME distribution depends on a single parameter, the CV = σ/μ.

Maximization of the Boltzmann-Gibbs-Shannon entropy with these constraints results in the truncated (for $x \geq 0$) normal distribution:

$$f(x) = \exp(-\lambda_0 - \lambda_1 x - \lambda_2 x^2), \quad \phi = \lambda_0 + \lambda_1 \mu_1 + \lambda_2 \mu_2$$

where λ_0, λ_1, λ_2 are Lagrange multipliers. This tends to the normal distribution as $\sigma/\mu \to 0$ and to the exponential distribution as $\sigma/\mu \to 1$.

For $\sigma/\mu > 1$ the Boltzmann-Gibbs-Shannon ME distribution does not exist. In this case the Tsallis entropy can be used which results in:

$$f(x) = [1 + \kappa\,(\lambda_0 + \lambda_1 x + \lambda_2 x^2)]^{-1-1/\kappa}, \quad \phi_q = (\kappa + 1)\,(\lambda_0 + \lambda_1 \mu_1 + \lambda_2 \mu_2)$$

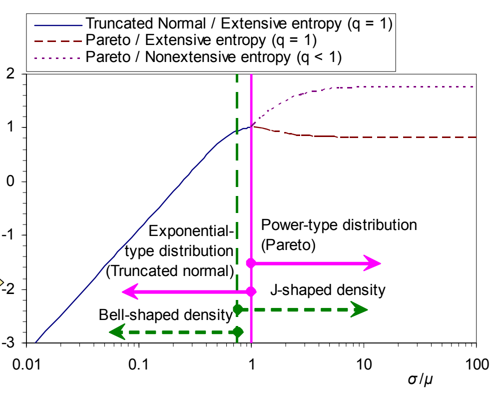
where $\kappa := (1 - q)/q$. In the absence of any information for κ or q, we can set $\lambda_2 = 0$ (Koutsoyiannis, 2005) and obtain the Pareto distribution:

$$f(x) = (1/\lambda)\,(1 + \kappa\,x/\lambda)^{-1-1/\kappa}$$

$$F^*(x) = (1 + \kappa\,x/\lambda)^{-1/\kappa}$$

For large x, this can be approximated by the scaling distribution in panel 3.

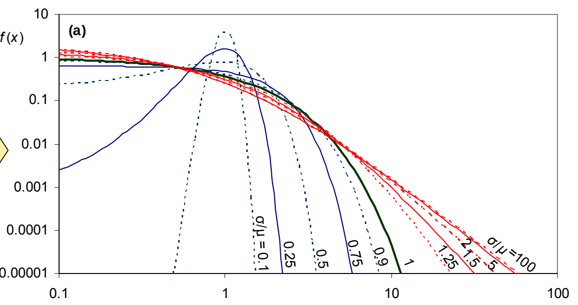
Maximized entropy and maximising distribution versus the coefficient of variation σ/μ



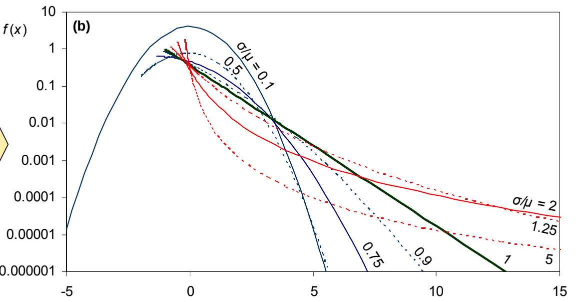
8. Resulting ME distributions

The ME densities are truncated normal for $\sigma/\mu < 1$ (tending to normal as $\sigma/\mu \to 1$), exponential for $\sigma/\mu = 1$ and Pareto for $\sigma/\mu > 1$.

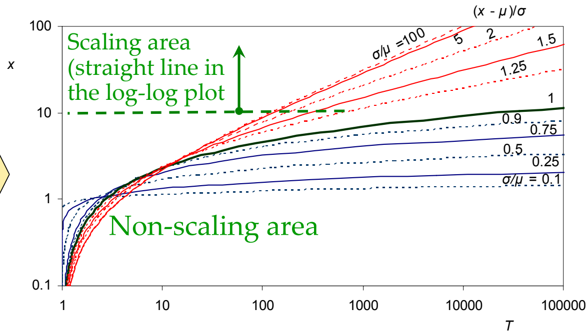
ME probability density functions vs. the random variate x ($\mu = 1$)



ME probability density functions vs. the standardized variate $(x - \mu)/\sigma$ ($\mu = 1$)



ME distribution functions shown as plots of the random variate versus return period ($\mu = 1$, $\delta = 1$)

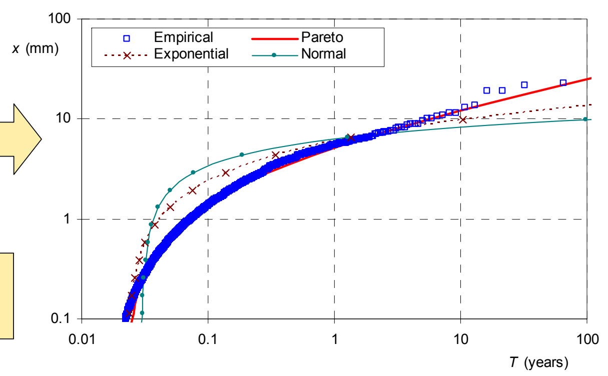


9. Application to hourly rainfall

Data set: Hourly rainfall in Athens, Greece, in January (in mm); period 1927-1996; 70 years of data (65 equivalent years if missing data are not counted); 48397 data values, out of which 2919 are nonzero. (Note: only nonzero values are modelled.)

μ = 1.07 mm
σ/μ = 1.47 > 1
ME distribution: Pareto
κ = 0.269
ϕ_q = 1.32

Conclusion: Scaling for T > ~2 yr

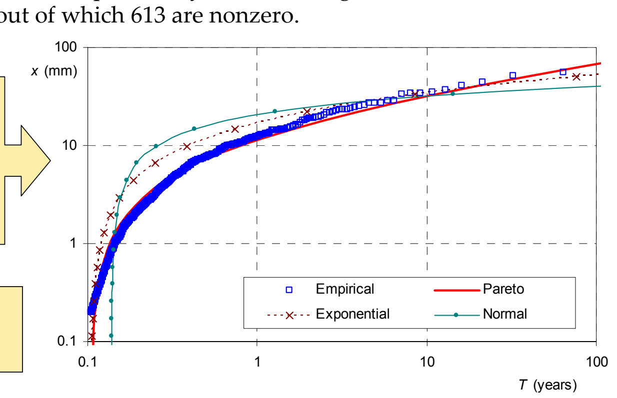


10. Application to daily rainfall

Data set: Daily rainfall in Athens, Greece, in January (in mm); period 1927-1996; 70 years of data (65 equivalent years if missing data are not counted); 1998 data values, out of which 613 are nonzero.

μ = 5.05 mm
σ/μ = 1.39 > 1
ME distribution: Pareto
κ = 0.242
ϕ_q = 1.29

Conclusion: Scaling for T > ~10 yr

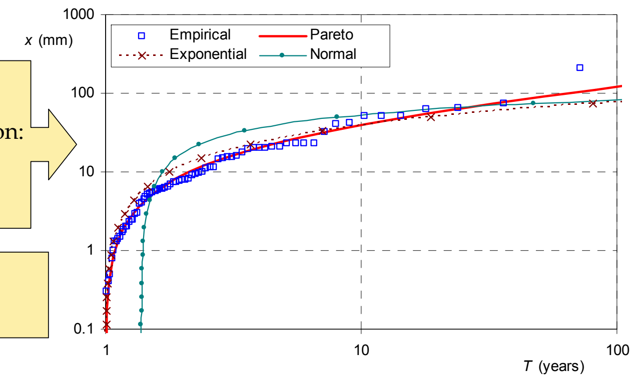


11. Application to monthly rainfall in a dry month

Data set: Monthly rainfall in Aliartos, Greece, in August (in mm); period 1907-2003; 96 years of data; 96 data values, 71 of which are nonzero.

μ = 17.0 mm
σ/μ = 1.67 > 1
ME distribution: Pareto
κ = 0.321
ϕ_q = 1.40

Conclusion: Scaling for T > ~20 yr

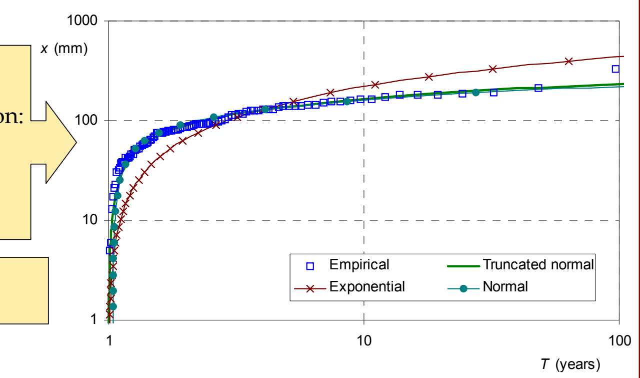


12. Application to monthly rainfall in a wet month

Data set: Monthly rainfall in Aliartos, Greece, in January (in mm); period 1907-2003; 96 years of data; 96 data values, all nonzero.

μ = 94.2 mm
σ/μ = 0.56 < 1
ME distribution: Truncated normal
κ = 0
ϕ_q = 0.78

Conclusion: No scaling

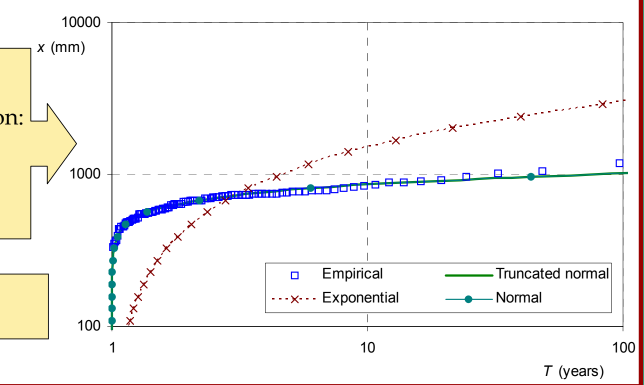


13. Application to annual rainfall

Data set: Annual rainfall in Aliartos, Greece (in mm); period 1907-2003; 96 years of data; 96 data values, all nonzero.

μ = 658.4 mm
σ/μ = 0.24 < 1
ME distribution: Truncated normal
κ = 0
ϕ_q = -0.01

Conclusion: No scaling

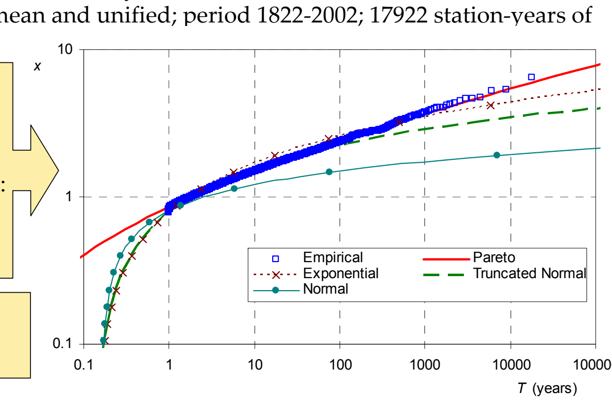


14. Application to extreme daily rainfall worldwide

Data set: Daily rainfall from 168 stations worldwide (Koutsoyiannis, 2004, 2005) each having at least 100 years of measurements; series above threshold, standardized by mean and unified; period 1822-2002; 17922 station-years of data.

μ = 0.28 (mean minus threshold)
σ/μ = 1.19 > 1
ME distribution: Pareto
κ = 0.15
ϕ_q = 1.16

Conclusion: Scaling for T > ~50 yr

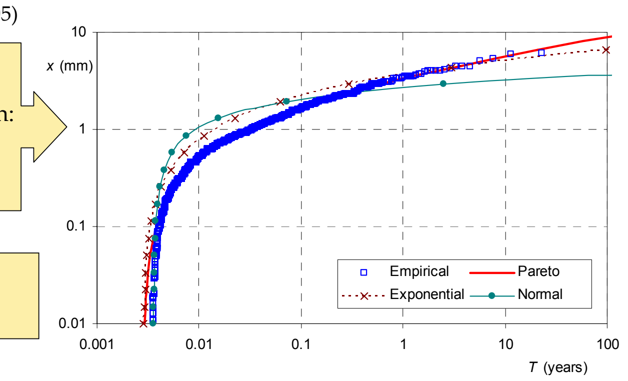


15. Application to daily runoff

Data set: Daily runoff in Boeoticos Kephisos river basin, Greece (in mm); period 1978-2003; 25 years of data (23 equivalent years if missing data are not counted); 8402 data values, out of which 6637 are nonzero. (Note: see details in Koutsoyiannis, 2005)

μ = 0.45 mm
σ/μ = 1.19 > 1
ME distribution: Pareto
κ = 0.15
ϕ_q = 1.16

Conclusion: Scaling for T > ~5 yr

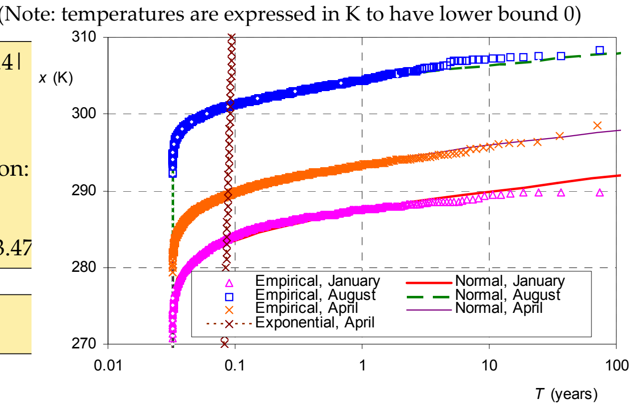


16. Application to mean daily temperature

Data set: Mean daily temperature in Athens, Greece, in three months (in K); period 1930-2003; 74 years of data; 2294|2220|2294 data values for January|August|April. (Note: temperatures are expressed in K to have lower bound 0)

μ = 282.2|288.4|300.2 K
σ/μ = 0.0075|0.0095|0.01
ME distribution: ~Normal
κ = 0
ϕ_q = -3.19 to -3.47

Conclusion: No scaling

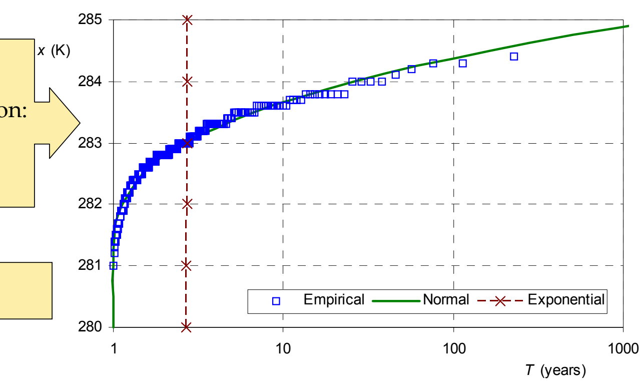


17. Application to mean annual temperature

Data set: Mean annual temperature in Geneva, Switzerland (in K); period 1753-1980 (one of the longest worldwide); 228 data values.

μ = 282.8 K
σ/μ = 0.0024
ME distribution: ~Normal
κ = 0
ϕ_q = -4.62

Conclusion: No scaling



18. Conclusions

1. Maximum entropy + Low variation → Exponential-type (truncated normal) distribution.
2. Maximum entropy + High variation → Power-type (Pareto) distribution.
3. Maximum entropy + High variation + High return periods → State scaling.
4. State scaling is only an approximation, good for high return periods and for variables with high variation.
5. Real world data series validate the applicability of the principle of maximum entropy in hydrometeorological processes.
6. This can be interpreted as dominance of uncertainty in nature.

References

Jaynes, E.T. (1957) Information theory and statistical mechanics, Physical Review 106(4), 620-630.

Koutsoyiannis, D. (2004) Statistics of extremes and estimation of extreme rainfall, 2, Empirical investigation of long rainfall records, *Hydrol. Sci. J.*, 49(4), 591-610.

Koutsoyiannis, D. (2005), Uncertainty, entropy, scaling and hydrological stochastics, 1, Marginal distributional properties of hydrological processes and state scaling, *Hydrol. Sci. J.*, (in press).

Papoulis, A. (1991) Probability, Random Variables, and Stochastic Processes (third edn), 57. McGraw-Hill, New York, USA.

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