ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO DI
INGEGNERIA CIVILE, CHIMICA,
AMBIENTALE E DEI MATERIALI - DICAM

K-moments and their application to ombrian modelling

Demetris Koutsoyiannis, Theano Iliopoulou and G.-Fivos Sargentis
School of Civil Engineering, National Technical University of Athens, Greece
(dk@itia.ntua.gr, http://www.itia.ntua.gr/dk/)

Presentation available online: http://www.itia.ntua.gr/2607/

Knowable moments (K-moments): blending properties of different statistical characterizations

- Like **classical moments**, K-moments are expectations of properly defined stochastic variables.
  - The difference is that K-moments are knowable given a sample, while classical moments are unknowable beyond order 2-3.
- Like **order statistics**, K-moments are ordered, inheriting the order from the sample.
  - The difference is that the K-moment estimates combine many sample points, rather than relying on a single one.
- Like **probability weighted moments** and **L-moments**, their estimates are linear combinations of the sample elements.
  - One difference is that they can also be generalized for a nonlinear setting (for example, the variance can be precisely recovered as a K-moment).
- Like **all above**, they can provide summary statistics for the first four orders, which are similar to those of L-moments and much more reliable than those of classical moments.
  - The main difference is that the K-moment framework uses high orders (>>4), which are particularly informative for extremes, and in fitting of distribution tails.

K-moments vs. other statistics: similarities and differences

1. They have a simple, clear, intuitive and rigorous **definition** as expectations of maxima or minima in a sample.
2. Their statistical **estimators** are simple, intuitive, fast, **unbiased** and reliable, enabling the moment estimation for orders as high as the sample size.
    - It is this characteristic that makes them **knowable**.
    - When the datasets are large, e.g., with millions or billions of values, there are techniques to accelerate the calculations.
3. Further to estimation, they enable **entropy**-based derivation of probabilistic models.
4. Both the theoretical definition and the estimators of K-moments apply, with complete correspondence, to both **continuous and discrete stochastic variables**.
    - This is similar to classical moments, but not necessarily shared with other types of moments.
5. Not only can the K-moments be reliably estimated from samples, but they can also provide information about what a classical moment estimator determines.
    - For orders beyond 2–3, the **classical moment** estimates substantially differ from their true values (which for this reason is **unknowable**.)

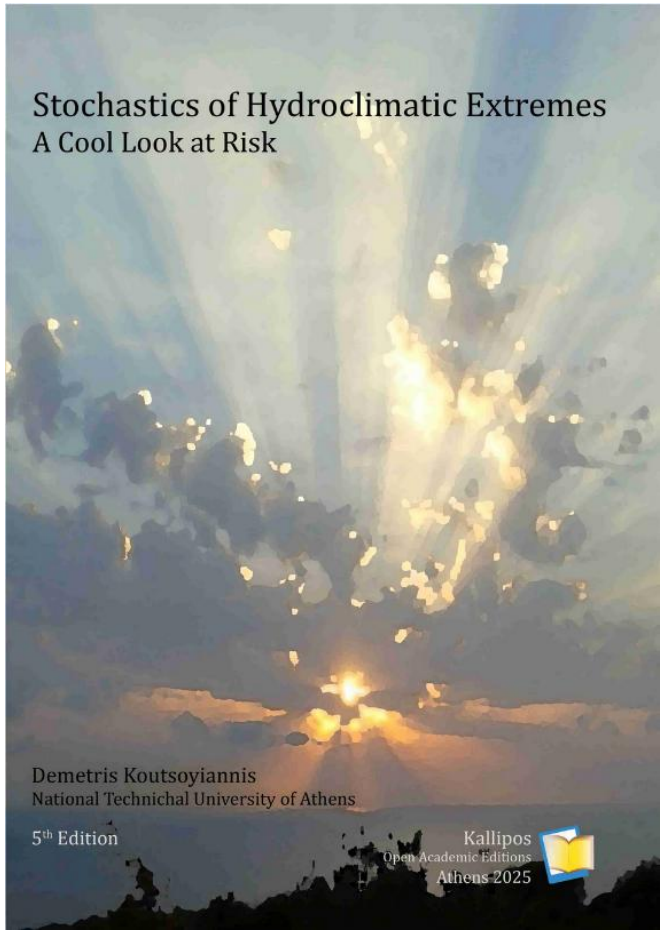
K-moments vs. other statistics: similarities and differences (2)

6. The K-moment framework includes the ability to readily **assign to each K-moment**, in a simple manner, a value of the distribution function or **return period**.
7. K-moments provide a sound and flexible framework for model fitting, making optimal use of the entire dataset, rather than relying on a few moments (as in classical moments and L-moments) or assigning probabilities to single data values (as in order statistics).
   - This enables **utilization of the highest possible moment orders**, which are particularly useful in modelling extreme highs or lows, which are closely associated with high-order moments.
   - The model fitting concept using K-moments is a **new strategy not shared with any other types** of methods of moments and includes **visualization of the goodness of fit**.
8. The K-moment estimators offer the ability to estimate the **probability density function** from a sample.
   - This is another unique feature of the K-moment framework.
9. They can take into account the **estimation bias** when the data are not an independent sample but a time series from a process with **dependence, even of long range** type.
   - This is yet another unique advantage of K-moments.

K-moments vs. other statistics: similarities and differences (3)

10. The K-moment framework offers the possibility of **modelling extremes using merely the parent distribution**, without reference to asymptotic extreme value distributions.
    - This is a more reliable choice as the convergence to the asymptotic extreme value distributions can be extraordinarily slow, while the non-asymptotic distributions of extremes can be quite difficult to determine, particularly when there is time dependence.
    - Yet there is a theoretical connection of the K-moment framework to the asymptotic extreme value theory, related to assigning values of the distribution function to K-moments (point 6) at the distribution tails.
11. Even when the available data are **block maxima** (e.g., annual maximum rainfall data), the K-moment framework can directly **estimate the K-moments of the parent distribution**.
    - This enables the fitting of the parent distribution from block maxima and makes unnecessary any reference to the extreme value distributions.

See details in Koutsoyiannis (2025).



Definition of K-moments

- Given a sample $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ of a stochastic variable $\underline{x}$, we can arrange it in increasing order:

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

- The stochastic variable $\underline{x}_{(i:n)}$ is called the ith **order statistic**, with special cases:

$$\underline{x}_{(1:n)} = \min\left(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n\right), \qquad \underline{x}_{(n)} := \underline{x}_{(n:n)} = \max\left(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n\right)$$

- The **upper knowable moment (K-moment) of order *p*** is the expectation of the largest order statistic $\underline{x}_{(p)}$:

$$K'_p := \mathrm{E}[\underline{x}_{(p)}] = \mathrm{E}\left[\max\left(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p\right)\right]$$

- The **lower knowable moment (K-moment) of order *p*** is the expectation of the smallest order statistic $\underline{x}_{(1:p)}$:

$$\overline{K}'_p := \mathrm{E}\left[\underline{x}_{(1:p)}\right] = \mathrm{E}\left[\min\left(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p\right)\right]$$

- Apparently, the K-moments are ordered as follows:

$$\overline{K}'_p \le \cdots \le \overline{K}'_2 \le \overline{K}'_1 = K'_1 = \mu \le K'_2 \le \cdots \le K'_p$$

- These moments are **noncentral**, and we can also define **central moments** as

$$K_p := K'_p - K'_1, \qquad \overline{K}_p := \overline{K}'_1 - \overline{K}'_p, \qquad K_p, \overline{K}_p \ge 0$$

Basic calculus of K-moments

- For a stochastic variable $\underline{x}$ of continuous type, the upper K-moment of order p of $\underline{x}$ is theoretically calculated as follows:

$$K'_p = p\mathrm{E}\left[\left(F(\underline{x})\right)^{p-1}\underline{x}\right] = p\int_{-\infty}^{\infty}\left(F(x)\right)^{p-1}x\,f(x)\,\mathrm{d}x = p\int_{0}^{1}x(F)\,F^{p-1}\mathrm{d}F$$

- Likewise, the lower K-moment of order p is theoretically calculated as follows:

$$\overline{K}'_p = p\mathrm{E}\left[\left(\overline{F}(\underline{x})\right)^{p-1}\underline{x}\right] = p\int_{-\infty}^{\infty}\left(\overline{F}(x)\right)^{p-1}x\,f(x)\,\mathrm{d}x = p\int_{0}^{1}x(\overline{F})\,\overline{F}^{p-1}\mathrm{d}F$$

- These equations allow for the extension of the evaluation of K-moments for non-integer order p for a stochastic variable $\underline{x}$ of **continuous type.**
- For **discrete-type variables the related equations are different.**
- For this case as well as for generalizations of K-moments, the interested reader is referred to Koutsoyiannis (2023, 2025).

Estimation

- The **unbiased estimators** of K-moments, upper ($\underline{K}'_p$) and lower ($\overline{\underline{K}}'_p$), from a sample of size n are respectively:

$$\underline{\widehat{K}}'_p = \sum_{i=1}^{n} b_{inp}\, \underline{x}_{(i:n)}\,, \qquad \overline{\underline{\widehat{K}}}'_p = \sum_{i=1}^{n} b_{inp}\, \underline{x}_{(n-i+1:n)} = \sum_{i=1}^{n} b_{n-i+1,n,p}\, \underline{x}_{(i:n)}$$

- The weights in the linear estimators are:

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

  while $\Gamma(\,)$ is the gamma function and $\underline{x}_{(i:n)}$ is the *i*th *order statistic*.

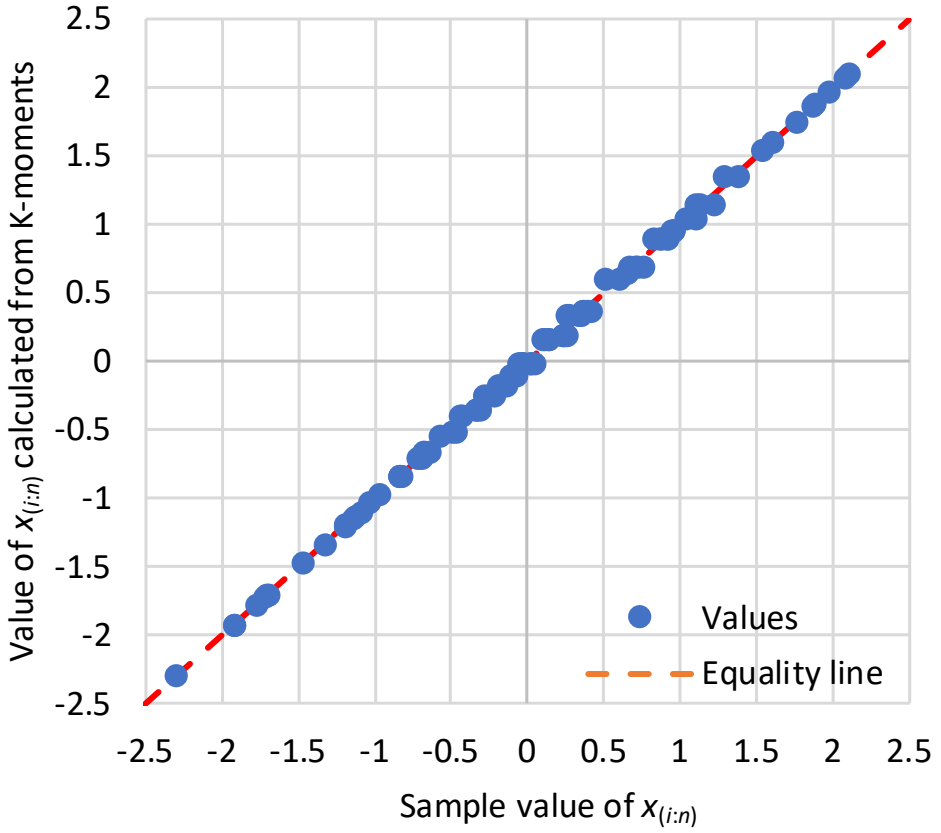
- Useful observations:
  - For integer moment order p and $i \geq p \geq 0$, this simplifies to

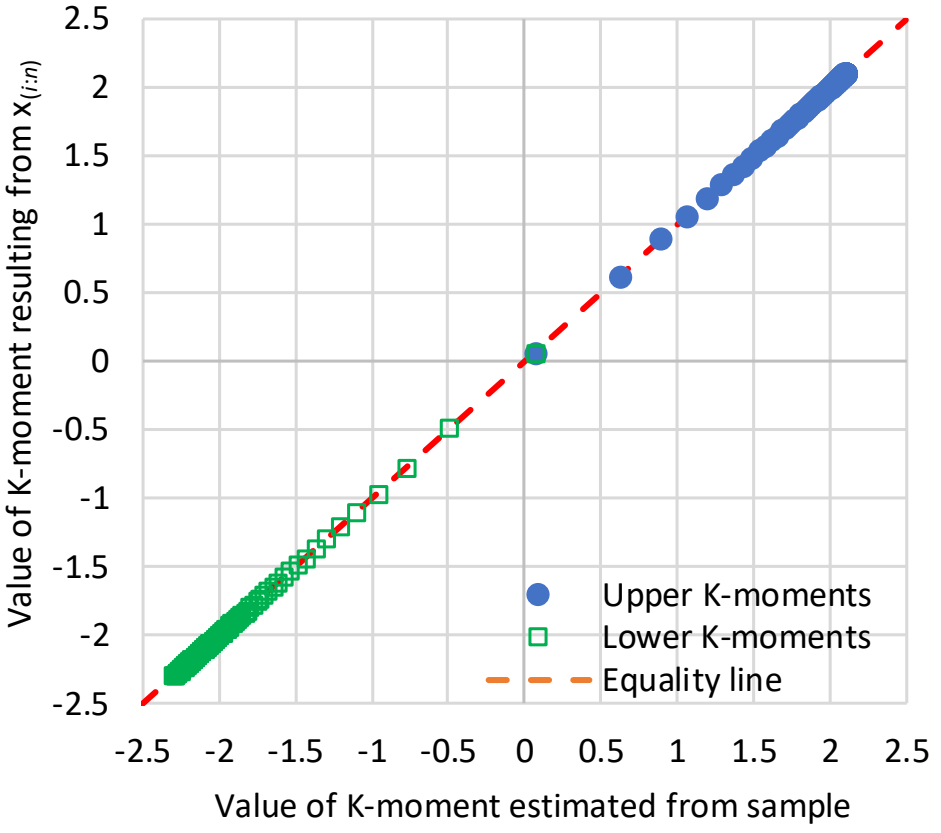
$$b_{inp} = \binom{i-1}{p-1} \Big/ \binom{n}{p}$$

  - For $p = 1$, $b_{in1} = 1/n$ and $\underline{\widehat{K}}'_1$ is identical to the estimator of the mean.
  - for $p = n$, $b_{inn} = 0$ for $i < n$ and $b_{nnn} = 1$, so that $\underline{\widehat{K}}'_n = \max\left(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n\right)$.

Correspondence between order statistics and K-moments

- Given a sample with n observations, we can estimate n upper K-moments and n lower K-moments (with the first upper and lower moment being identical, equal to the average).
- Each of the **three families is equivalent** (and transformable) to the two others.
- The graphs show that doing the transformations twice we recover the original family.





K-moments-based summary statistics

Based on the K-moments, we define:

- the **K-centre** of order p, C_p, measuring central tendency,

$$C_p := \frac{K'_p + \overline{K}'_p}{2}$$

- the **K-spread** of order p, D_p, measuring dispersion, as:

$$D_p := \frac{K'_p - \overline{K}'_p}{2} \geq 0$$

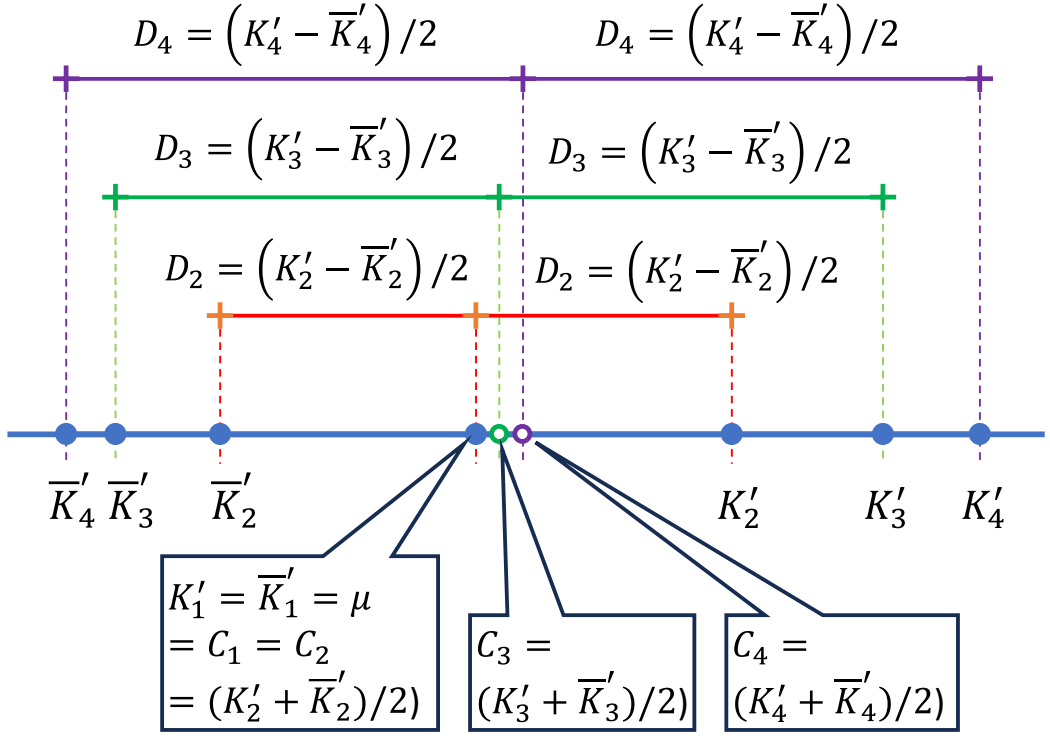
- Characteristics for low orders:

$p = 1$: $C_1 = K'_1 = \overline{K}'_1 = \mu, \qquad D_1 = 0$

$p = 2$: $C_2 = \frac{K'_2 + \overline{K}'_2}{2} = K'_1 = C_1 = \mu, \qquad D_2 = \frac{K'_2 - \overline{K}'_2}{2} = K_2 = \overline{K}_2$

$p = 3$: K-skewness $:= (C_3 - \mu)/K_2$ (dimensionless)

$p = 4$: K-kurtosis $:= D_4/K_2$ (dimensionless)



An additional characteristic dimensionless quantity (meaningful for nonnegative stochastic variables only) [+SOCIOSTOCHASTICS]

- **K-spread coefficient** $\coloneqq K_2/\mu$.
- This is identical to the famous **Gini index** (after Gini, 1936, who introduced it in a very different way, which is still in common use today).
- This is used in economics for quantifying inequality in income and wealth.
- Its **lowest value is 0**, meaning no spread (perfect equality).
- Its **highest value is 1** meaning the largest possible spread (a single individual has all the income while all others have none).
- This is of particular interest when used as a constraint for entropy maximization.
- Both **extreme cases are extremely unlikely as they are associated with extremely low entropy** (theoretically $-\infty$).

Corrado Gini (1884 – 1965)

Italian statistician, demographer and sociologist.

Graduate of the University of Bologna.

A eugenicist, and an advocate of Italian Fascism.

He also had the curious idea that Italy should be annexed by the USA. ☺

Maximum entropy under constrained mean and K-spread

Entropy maximizing distribution

Distribution→	Logistic ($x \in \mathbb{R}$)	Generalized half-logistic ($x \in \mathbb{R}^+$)
Distribution function, $F(x)$	$1 - \frac{1}{1 + \mathrm{e}^{-\varsigma + x/\lambda}}$	$1 - \frac{1}{1 - \mathrm{e}^{-\varsigma} + \mathrm{e}^{-\varsigma + x/\lambda}}$
Mean, μ	$\lambda\varsigma$	$\frac{\lambda\varsigma}{1 - \mathrm{e}^{-\varsigma}}$
Standard deviation, σ	$\lambda\pi/\sqrt{3}$	$\lambda\sqrt{\mathrm{e}^{s}(2(\mathrm{e}^{s} - 1)\mathrm{Li}_2(1 - \mathrm{e}^{-s}) - s^2)}$
K-Moment of order p, K'_p	$\lambda\left(H_{p-1} + \varsigma\right)$	$\lambda\left(H_p + \varsigma + \frac{\mathrm{B}_{1-e^{\varsigma}}(p+1,0)}{(1 - \mathrm{e}^{\varsigma})^p}\right)$
K-spread coefficient, K_2/μ	$\frac{1}{\varsigma}$	$\frac{1}{\varsigma} - \frac{1}{\mathrm{e}^{\varsigma} - 1}$
Standardized Entropy, Φ_μ	$2 - \ln(\varsigma)$	$2 - \frac{\varsigma}{1 - e^{-\varsigma}} - \ln\left(\frac{\varsigma}{e^{\varsigma} - 1}\right)$
Lower-tail index, ζ (ζ' for logistic)	0	1
Upper-tail index, ξ	0	0

Clarifications of symbols: λ is a scale parameter; ς is a shape parameter; $B_z(a, b)$ is the incomplete beta function; H_p is the pth harmonic number; $\mathrm{Li}_2(a)$ is the dilogarithm.

Entropy definition:

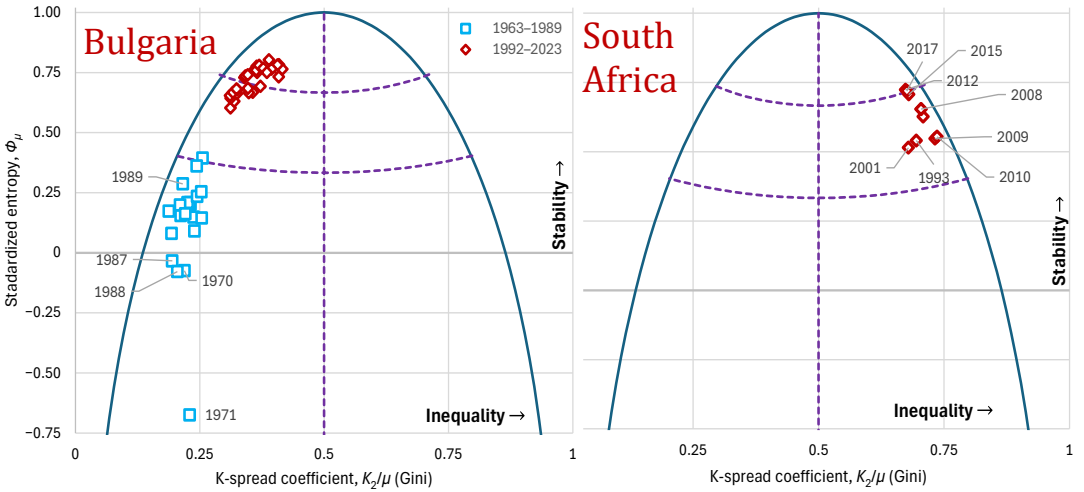
$$\Phi[\underline{x}] := \mathrm{E}\left[-\ln\frac{f(\underline{x})}{\beta(\underline{x})}\right]$$

Standardized entropy:

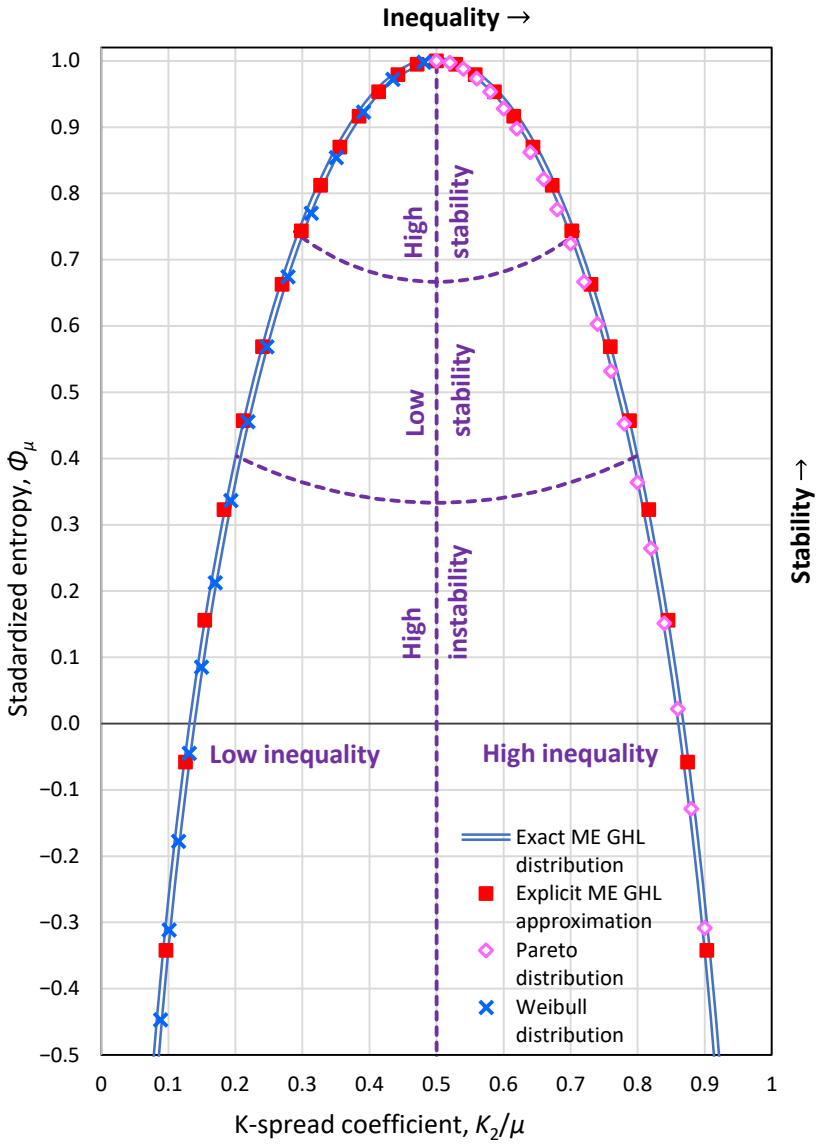
$$\Phi_\mu[\underline{x}] := \mathrm{E}[-\ln(\mu f(\underline{x}))]$$

Standardized entropy vs. K-spread in entropy-maximizing (Generalized half-logistic) distribution

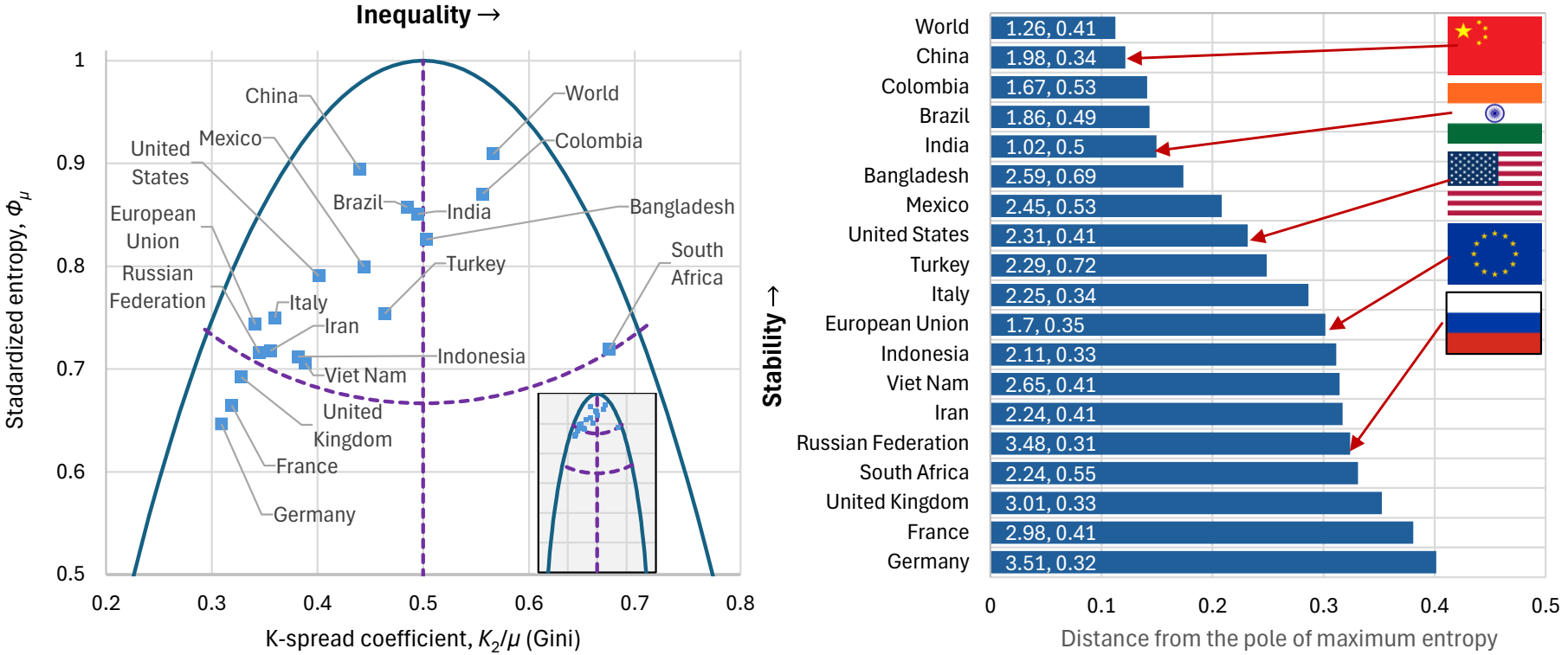
- The curve on the right shows the **trade-off** between standardized quantities:
  - **entropy**, Φ_μ, quantifying (in)stability,
  - **K-spread**, K_2/μ, quantifying (in)equality.
- It also shows that the entropy maximizing distribution can be approximated by the Pareto distribution for high K_2/μ and by the Weibull distribution for low K_2/μ.
- Points above the curve are infeasible.
- Points below the curve are feasible and actually occur.



Graphs from Koutsoyiannis and Sargentis (2026).



Comparison of major countries (w/ population > 50 million) in terms of stability and (in)equality (2022)



- Source of graph: Koutsoyiannis and Sargentis (2026).
- The latter study contains detailed data on the income distributions and additional comparisons.
- For example, USA and China have close similarity in distribution, with the most visible difference being the smaller upper tail index of China (tantamount to discouraging extreme wealth).

Assigning return periods to K-moments

- Like the order statistics, K-moments offer a means to estimate probabilities from samples.
- To determine the theoretical return period $T(K'_p)$ we introduce the ratio Λ_p which happens to vary only slightly with p. Assuming a time unit D we have:

$$T\left(K'_{p1}\right) = \frac{D}{1 - F\left(K'_{p1}\right)}, \qquad \Lambda_p := \frac{T(K'_{p1})}{D\,p} = \frac{1}{p\left(1 - F\left(K'_{p1}\right)\right)}$$

- Any symmetric distribution will have $\Lambda_1 = 2$ exactly because K'_1 is the mean, which equals the median and thus has a return period of $2D$. Thus, an approximation is the rule of thumb:

$$\Lambda_p \approx 2$$

- Generally, the exact value Λ_1 is easy to determine, as it is the return period of the mean (depending on the tail indices ξ, ζ):

$$\Lambda_1 = \frac{1}{1 - F(\mu)} = \frac{T(\mu)}{D}$$

- The exact value of Λ_∞ depends only on the tail index ξ of the distribution (with $\gamma = 0.577$ is the Euler's constant):

$$\Lambda_\infty = \begin{cases} \Gamma(1-\xi)^{1/\xi}, & \xi \neq 0 \\ \mathrm{e}^{\gamma}, & \xi = 0 \end{cases}$$

- These enable the **simple approximation of** Λ_p and hence of the **return period**:

$$\Lambda_p \approx \Lambda_\infty + (\Lambda_1 - \Lambda_\infty)(1/p), \qquad T\left(K'_p\right)/D = p\Lambda_p \approx \Lambda_\infty p + (\Lambda_1 - \Lambda_\infty)$$

Very high order K-moments relationships between *p* and *T*

- From the equation:

$$p = \frac{T}{\Lambda_\infty D} - \frac{\Lambda_1}{\Lambda_\infty} + 1,$$

  we can easily find the **relationship between K-moment order *p* and corresponding return period *T*.** An example is given in the table.

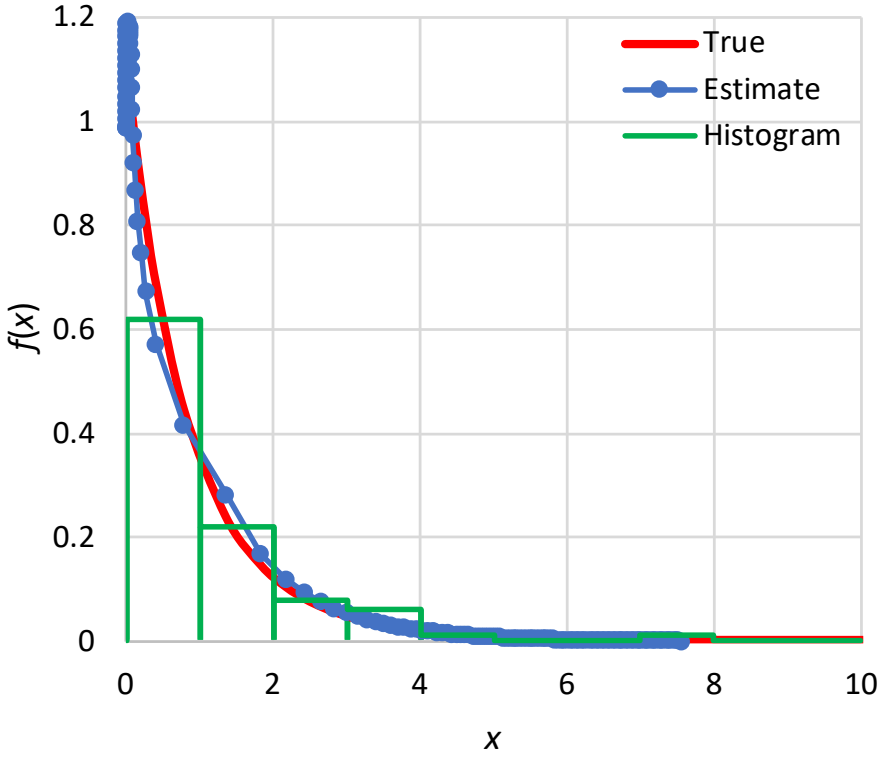
- K-moments of such high order as shown in the table can reliably be estimated.
- In a stochastic process with long-range dependence, the values in the table need to be properly adapted.
- Detailed information for the adaptation is given in the book by Koutsoyiannis (2025).
- The same book contains information for estimating return periods for **lower K-moments** and for **more accurate relationships between moment order and return period.**

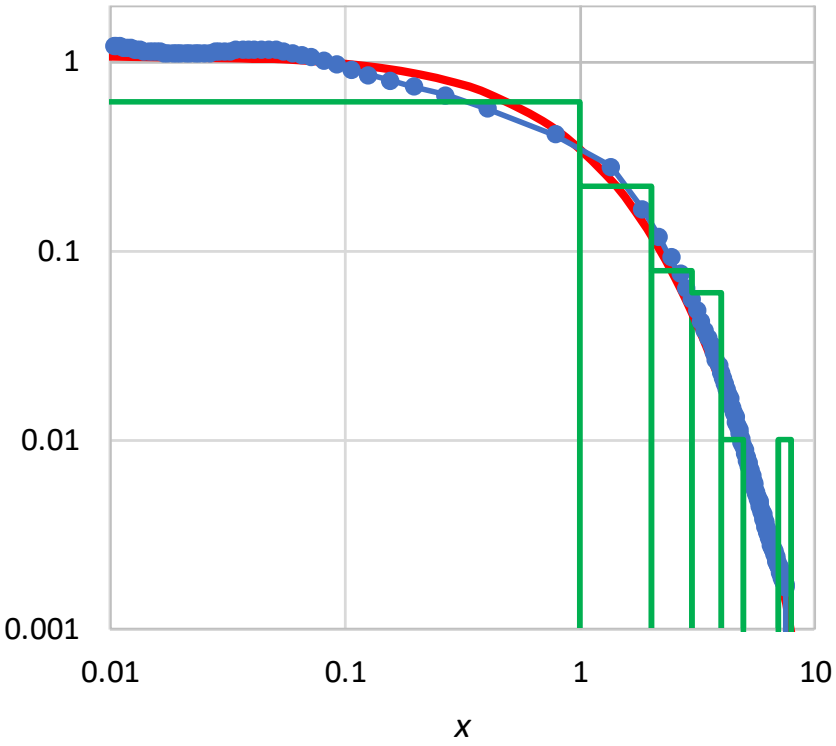
Example of the K-moment order p corresponding to the specified return period for the Pareto distribution with shape parameter ξ = 0.15.

	D = 1 d	D= 1 h	D = 10 min
T = 2 months	29	717	4 307
T = 1 year	179	4 307	25 842
T = 2 years	358	8 614	51 684
T = 100 years	17 945	430 702	2 584 212

Estimation of probability density function

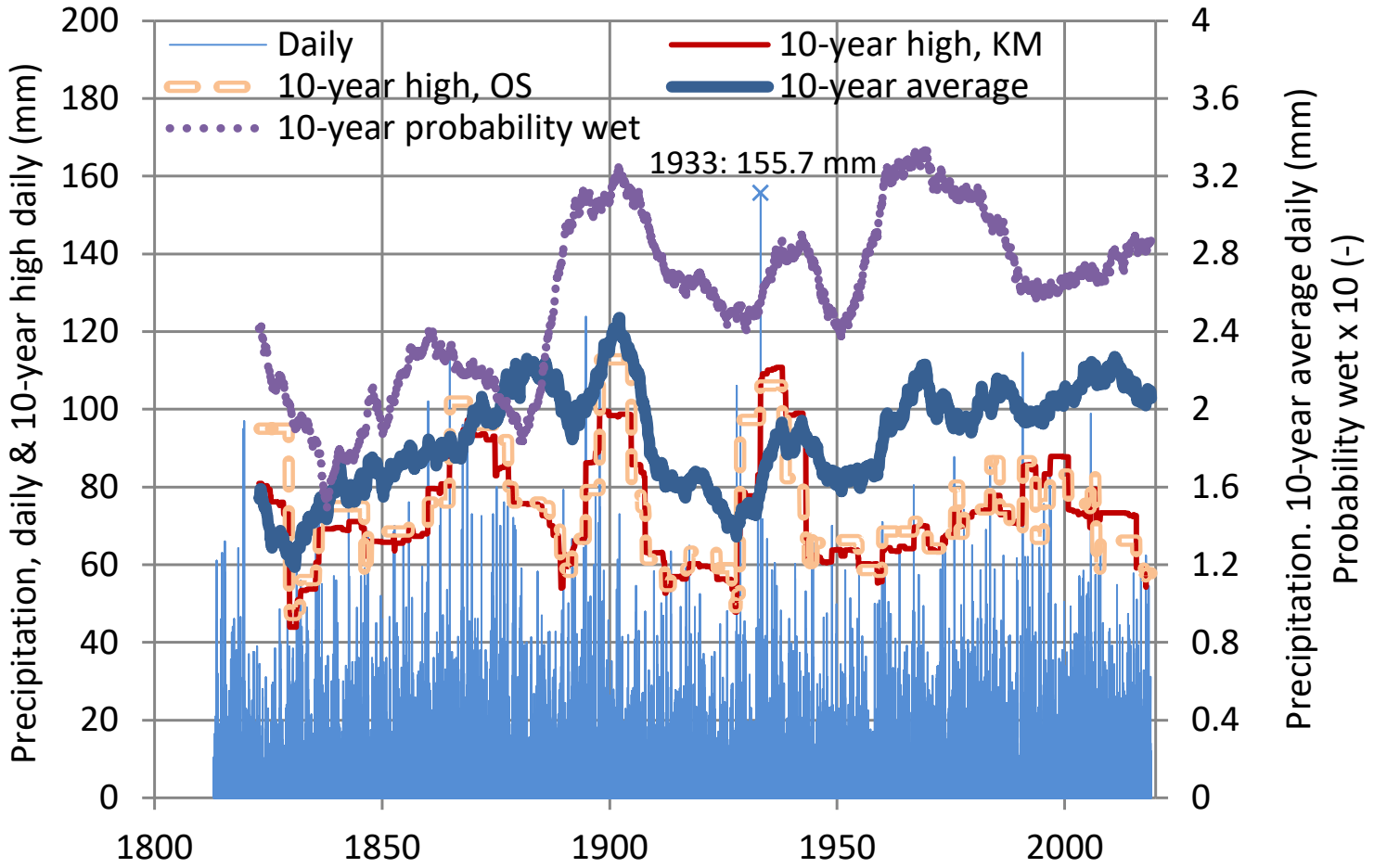
- As K-moments can give the $2n-1$ sample estimates of K-moments $\widehat{K}'_p$ and $\widehat{\overline{K}}'_p$ (the ordered values $\widehat{\overline{K}}'_n \le \widehat{\overline{K}}'_{n-1} \le \cdots \le \widehat{\overline{K}}'_1 \equiv \widehat{K}'_1 \le \widehat{K}'_2 \le \cdots \widehat{K}'_{n-1} \le \widehat{K}'_n$) and the estimates of the distribution function $\widehat{F}\left(K'_p\right)$ and $\widehat{F}\left(\overline{K}'_p\right)$, it is then straightforward to approximate the density function $f(x) = \mathrm{d}F(x)/\mathrm{d}x$ by the discrete derivative.
- This is the only method that can effectively replace the empirical histogram with a decent estimate of probability density.
- See details in Koutsoyiannis (2022, 2025).





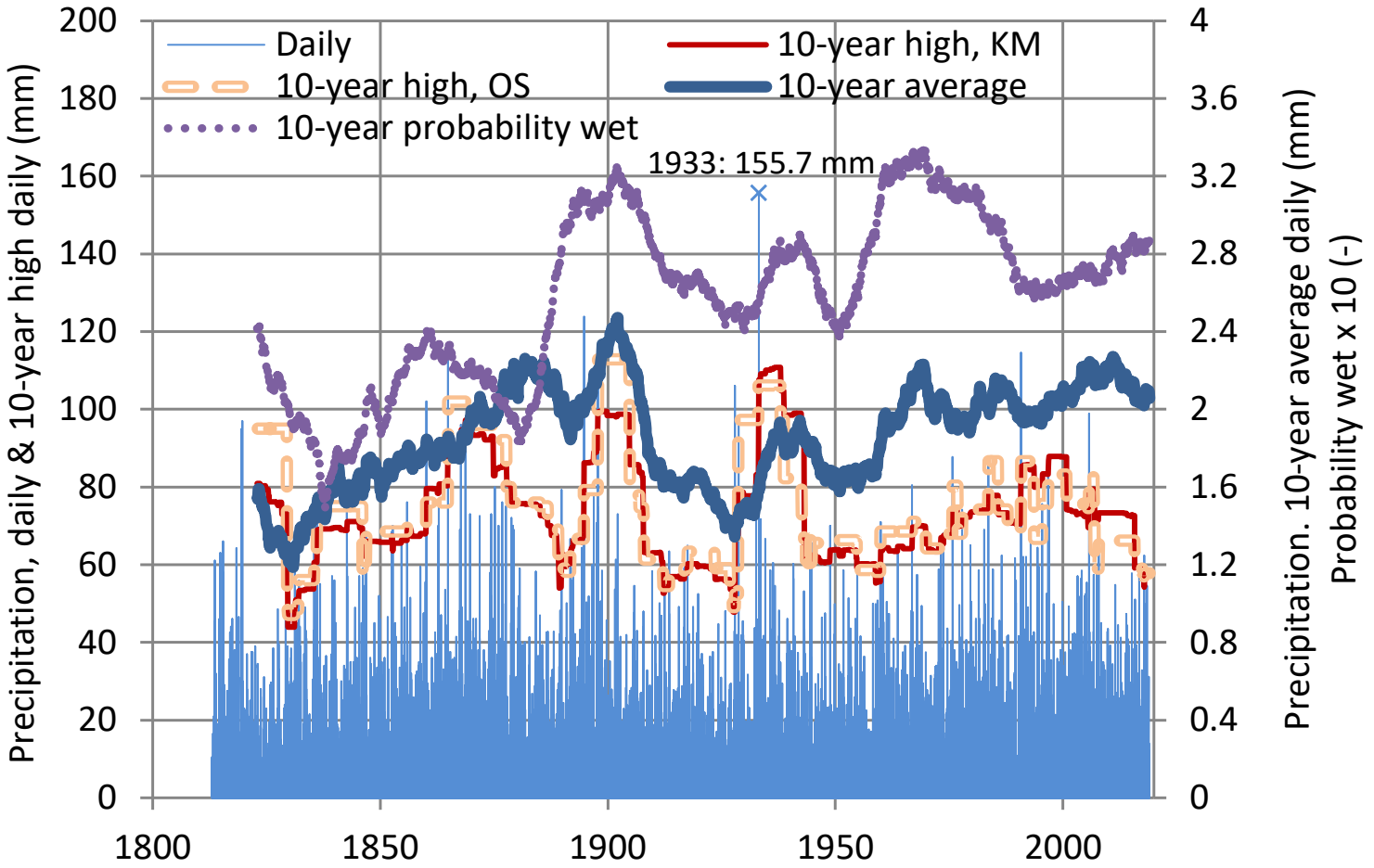
Application to Bologna precipitation: Daily data

- Bologna, Italy (44.50°N, 11.35°E, +53.0 m).
- Data available from the Global Historical Climatology Network (GHCN) – Daily.
- Uninterrupted for the period 1813-2007: 195 years.
- For the period 2008-2018, daily data are provided by the repository Dext3r of ARPA Emilia Romagna.
- **Total record length: 206 years.**
- The plot shows daily data along with 10-year average and maximum depths and probability dry (as a climatic indicator of low rainfall).



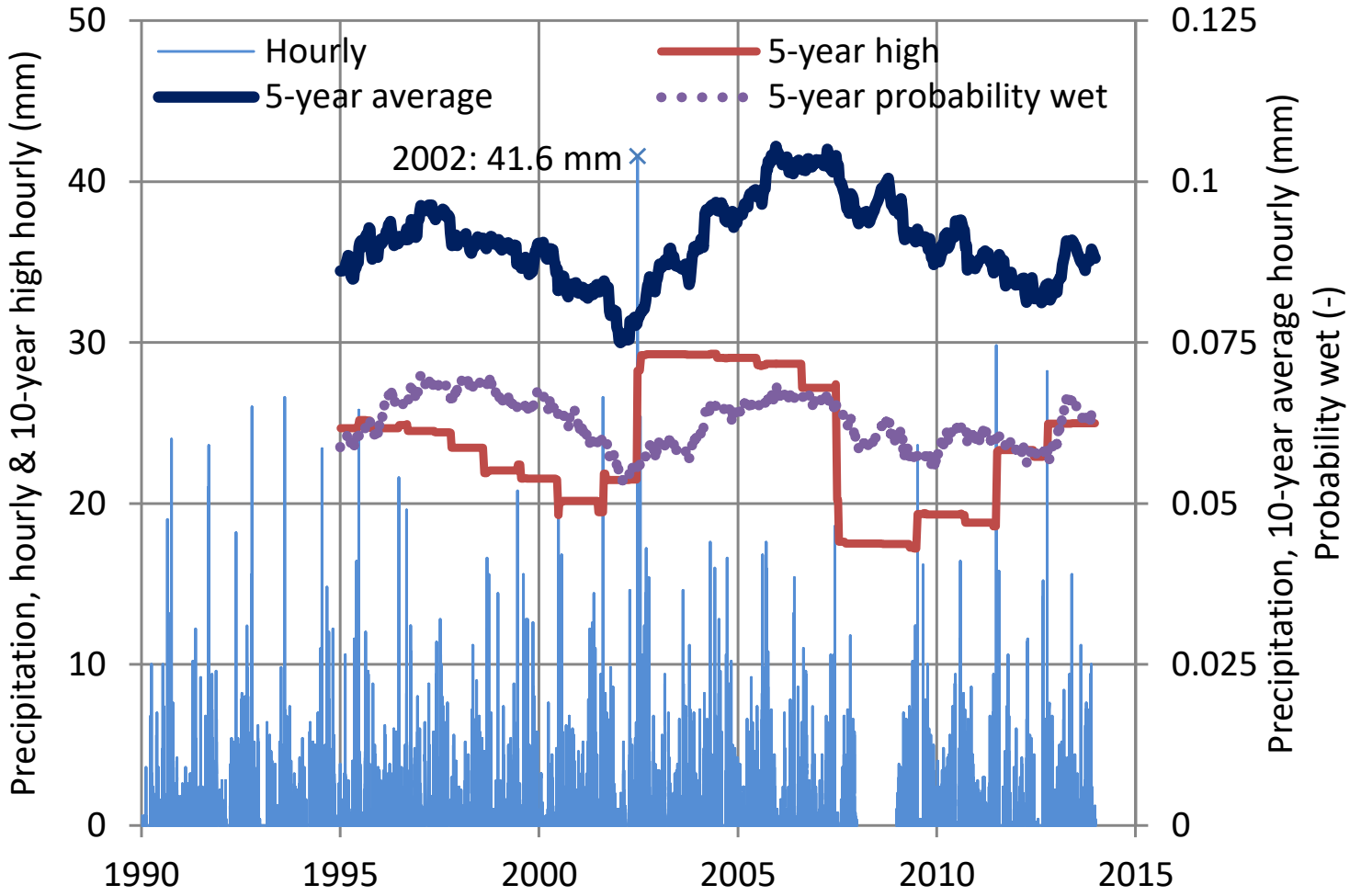
Application to Bologna precipitation: Diagnostic analysis of the evolution of climatic extremes

- **Main observation**:
  All 10-year climatic indices have **varied substantially and irregularly**:
  - The **average by a factor of 2** (from 1.2 to 2.4 mm);
  - The **probability wet by a factor >2** (from 0.15 to 0.33);
  - The **high daily precipitation by a factor 2.5** (from 44 to 110 mm/d).

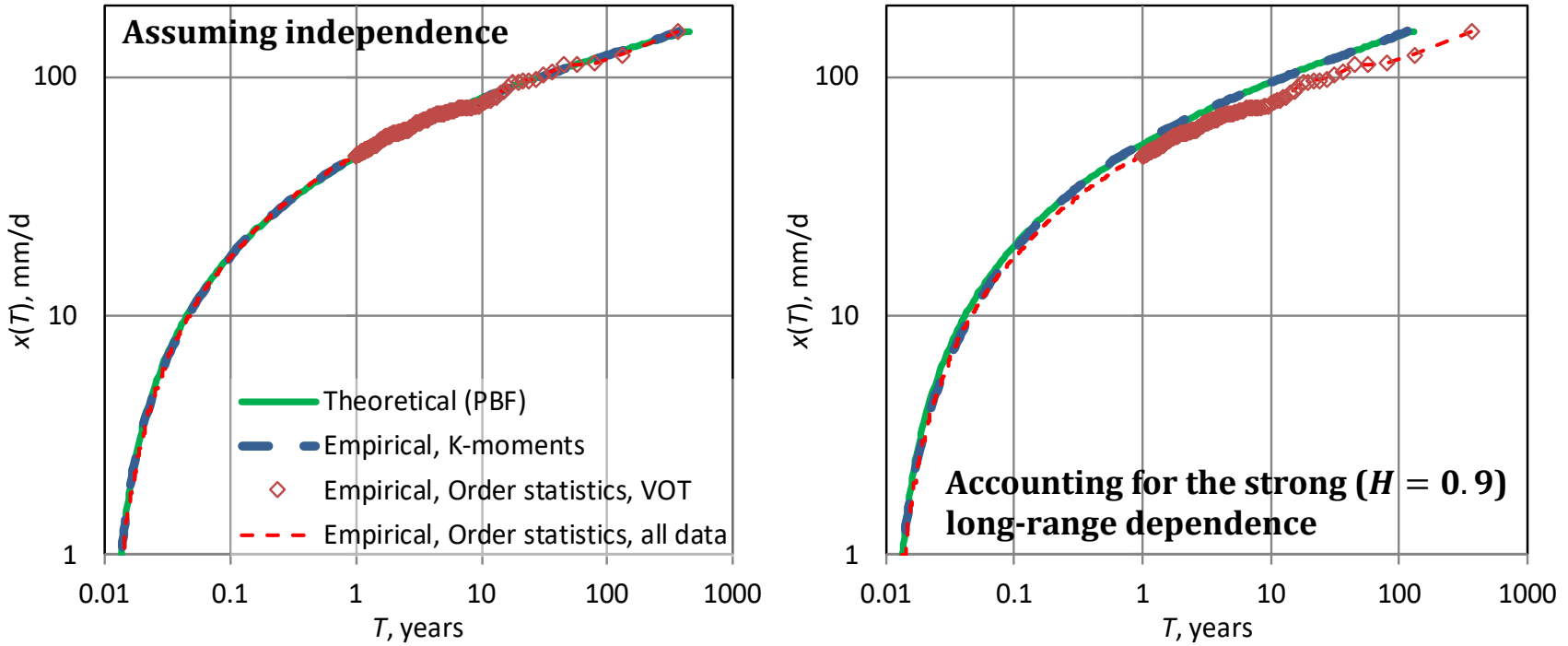


Application to Bologna precipitation: Hourly data

- Hourly rainfall data of the Bologna station for the period 1990-2013 are also available, provided by the Dext3r repository.
- **23 years** full coverage, while the entire 2008 is missing (retrieved and processed by Lombardo et al., 2019).
- **Main observation**: Again we have **fluctuations with upward and downward segments**. However, 23 years are too few (and the 5-year window used too small) to monitor climatic variation of hourly rainfall.



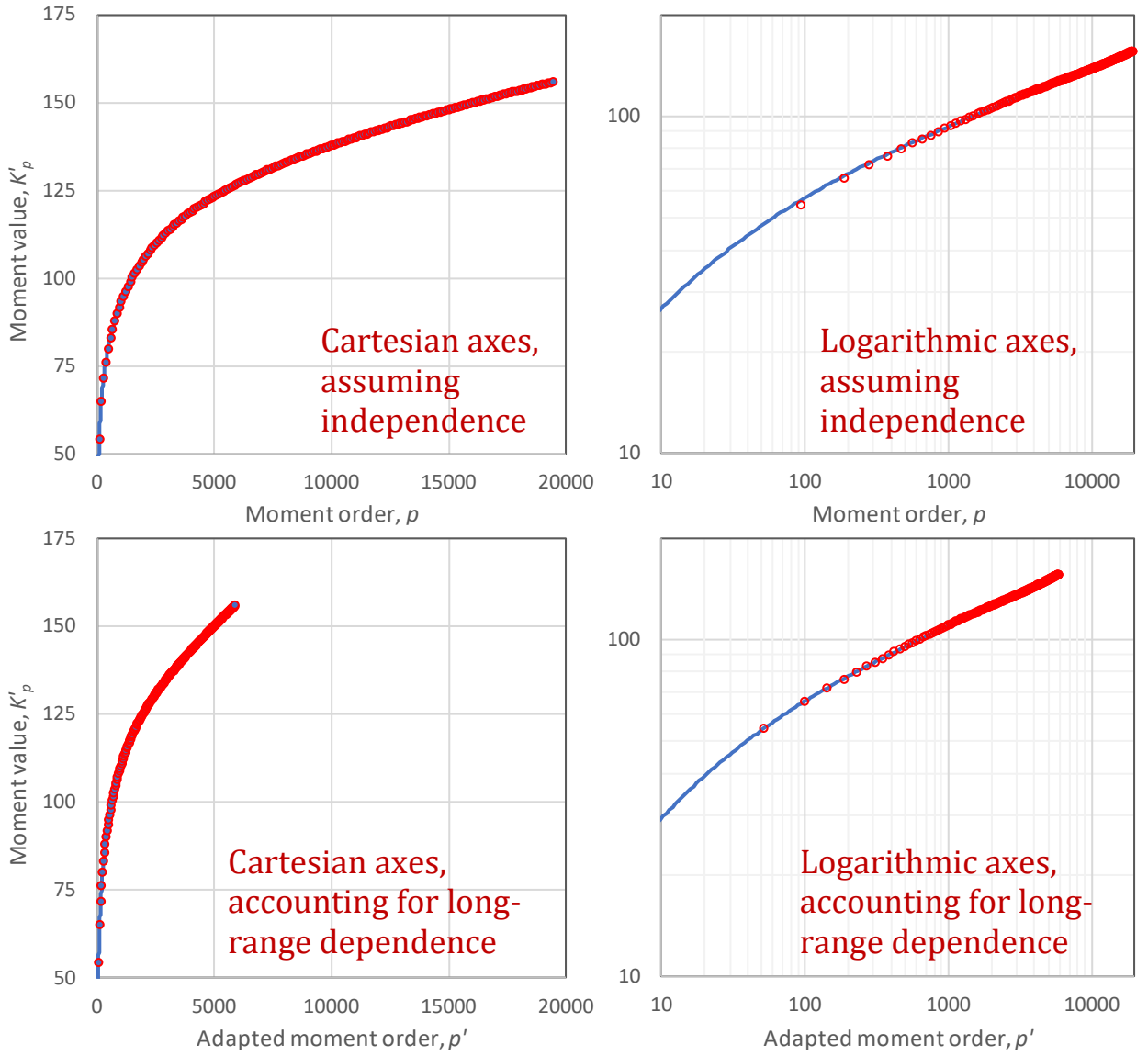
Application to Bologna precipitation: Extreme-oriented estimation taking the Hurst-Kolmogorov behaviour into account



- The Pareto-Burr-Feller (PBF) distribution was fitted on Bologna precipitation data for the entire range of return period.
- Unlike classical statistical methods, the K-moment framework offers the **capability to take persistence into account.**
- The **differences due to persistence are substantial.** (Notice that the logarithmic axes.)

Genuine distribution function fitting from block-maxima

- Extreme precipitation data are often **available as block maxima** (usually annual maxima).
- Any stochastic model should be fitted to the actual process (same requirement for the marginal distribution).
- The K-moment framework offers the **capability to estimate parameters of the actual process from block maxima**.
- The figure, based on the **Bologna precipitation, shows the excellent performance** of the method in estimating the actual process properties (lines) from block maxima (points).



Ombrian model: Multiscale distribution of rainfall intensity

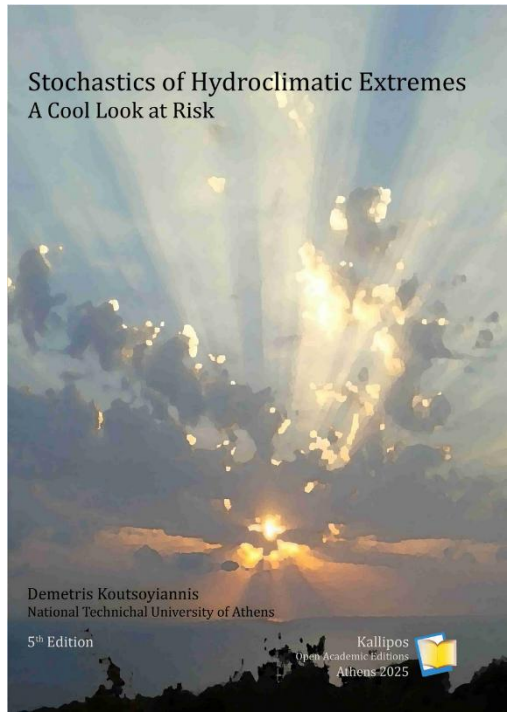
- Estimation of rainfall intensity at various scales and for various return periods is central in hydrological design and management, and in other engineering tasks.
- This is commonly based on empirically constructed curves, the **ombrian curves,** often relying on empirical approaches, which hinder their applicability over large scales, and preclude simulation.

  Note: The 'ombrian curves' (from the Greek ombros, meaning rainfall) are often referred to with the misnomer 'rainfall intensity – duration – frequency' (or IDF).

  'Duration' is meant to be 'time scale' and 'frequency' is meant to be 'return period'.

  Hopefully the misleading acronym IDF is less popular now…

- Recent work has advanced these curves to theoretically-consistent models describing the stochastic properties of the distribution of rainfall of any order at any time scale (see Koutsoyiannis, 2025).
- The model extends applicability to the full range of available scales, e.g. from minutes to decades.
- From an ombrian model, which is simple enough, the ombrian relationships are directly extracted.



Ombrian model: Distribution function

1. Generally we use the three-parameter PBF distribution with discontinuity at the origin. However, for small scales it can be simplified using the two-parameter Pareto distribution:

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

   It is shown by theoretical reasoning (Koutsoyiannis, 2025) that the tail index ξ should be constant, while the probability wet, $P_1^{(k)}$, and the state scale parameter, $\lambda(k)$, are functions of the time scale k.

2. For larger time scales we keep the PBF distribution (with possible discontinuity at zero for intermediate scales), i.e.:

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

   In this case a new parameter $\zeta(k)$ is introduced, which is again a function of time scale. The Pareto distribution is a special case of PFB for $\zeta(k) = 1$. In contrast to the Pareto distribution, whose density is a decreasing function of x, the PBF tends to be bell-shaped for increasing $\zeta(k)$. Here we sacrifice the constancy of tail index ($= \xi/\zeta(k)$) to assure simplicity and ergodicity.

Ombrian model: Mean and climacogram

3. Constant mean of the time averaged process:

$$\mathrm{E}\left[\underline{x}^{(k)}\right] = \mu$$

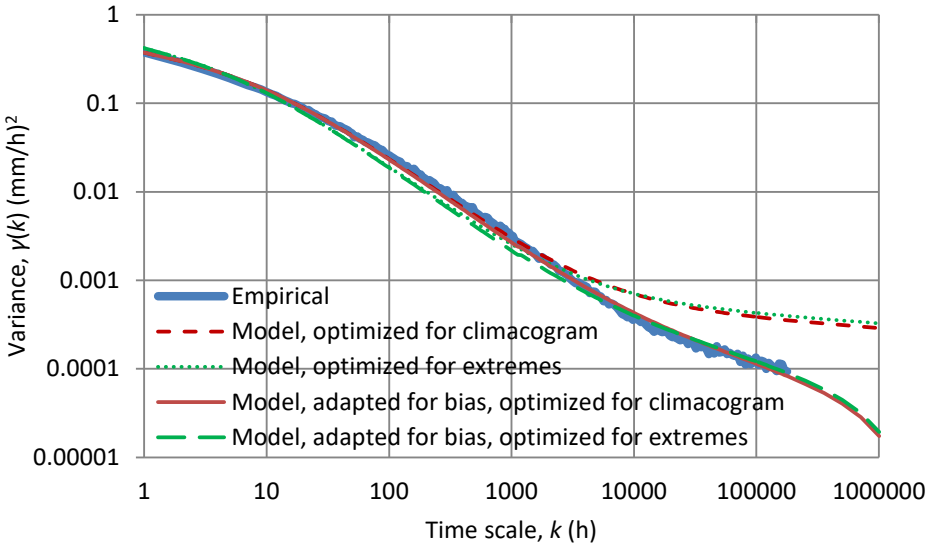
4. Climacogram of Filtered Hurst-Kolmogorov – Cauchy (FHK-C) type, i.e.:

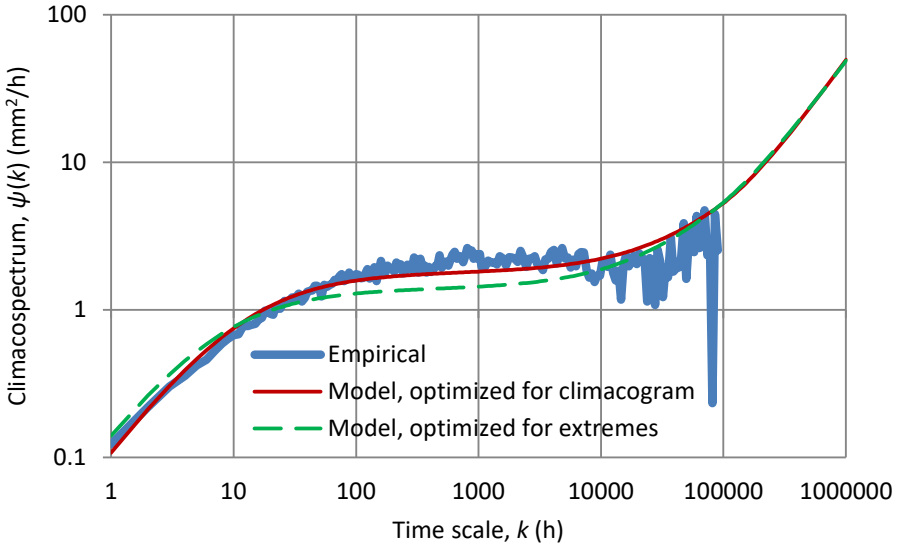
$$\mathrm{var}\left[\underline{x}^{(k)}\right] = \gamma(k) = \lambda_1^2 \left(1 + \left(\frac{k}{\alpha}\right)^{2M}\right)^{\frac{H-1}{M}}$$

or of Filtered Hurst-Kolmogorov – Cauchy-Dagum (FHK-CD) type; in the latter case, to avoid an overparametrized model (and as we expect $H > 0.5$ and $M < 0.5$ due to roughness), we set $M = 1 - H$ and thus we get:

$$\gamma(k) = \lambda_1^2 \left(1 + \frac{k}{\alpha_1}\right)^{2H-2} + \lambda_2^2 \left(1 - \left(1 + \frac{\alpha_2}{k}\right)^{2H-2}\right)$$

Clearly, in both cases, $\gamma(k) \to 0$, as $k \to \infty$, which makes the process ergodic; for $k = 0$, $\gamma(0) = \gamma_0 = \lambda_1^2$ in the FHK-C case and $\gamma(0) = \gamma_0 = \lambda_1^2 + \lambda_2^2$ in the FHK-CD case. In both cases $\gamma(0)$ is finite and the number of parameters is four.





The ombrian model: Probability wet/dry

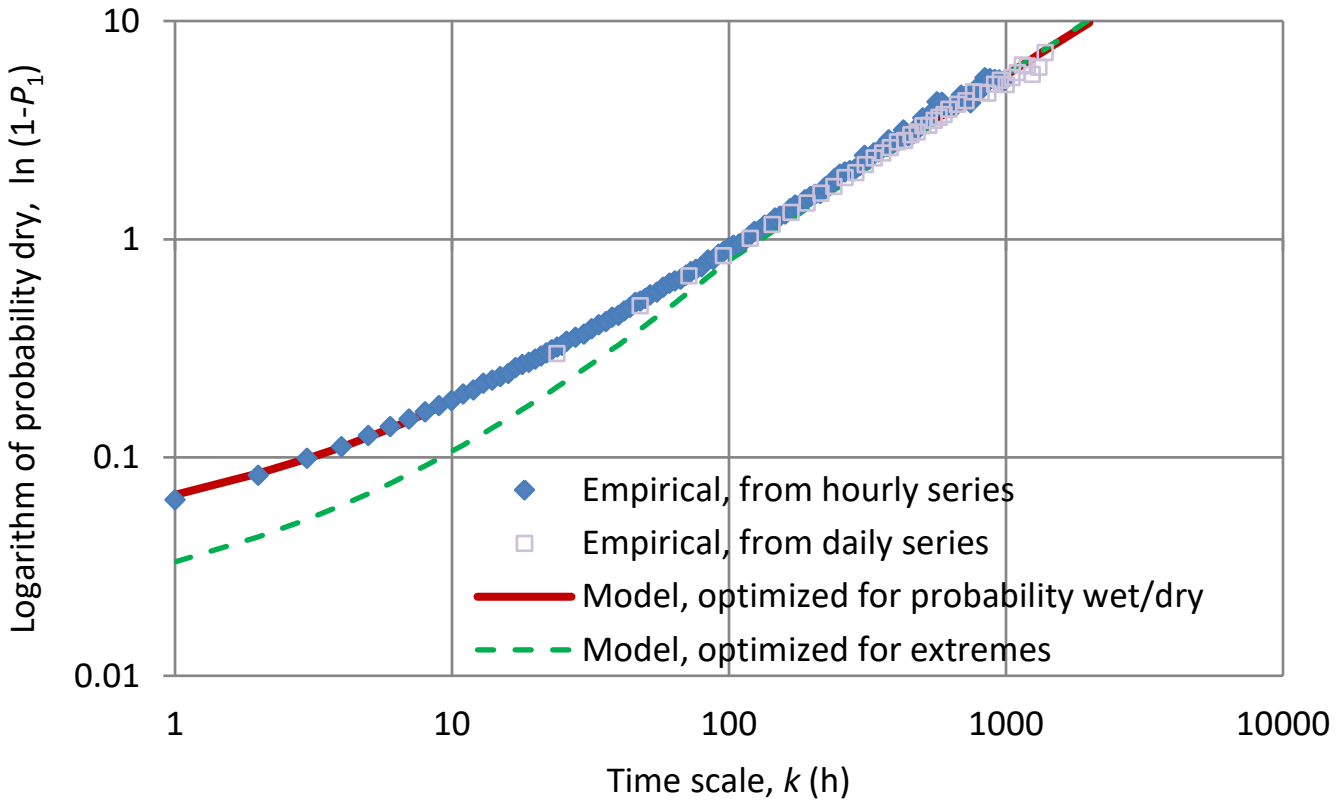
5. Probability wet and dry, $P_1^{(k)} = 1 - P_0^{(k)}$, varying with time scale according to:

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

where k^* is the transition time scale from Pareto to PBF distribution, for which $P_0^{(k^*)} > 0$ and $\zeta(k^*) = 1$ (for continuity in the transition point), and θ is a parameter ($0 \leq \theta \leq 1$). This equation has been derived in Koutsoyiannis (2006) based on maximum entropy considerations.

In the Pareto case, since $\zeta(k) = 1$, the probability wet is fully determined from the other parameters:

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



Mathematical relationships of the ombrian model

Quantity	Small scales, $k \le k^*$ (Pareto)	Large scales, $k \ge k^*$ (PBF)[1]
$\mathrm{E}\left[\underline{x}^{(k)}\right]$	μ	
$\gamma(k)$	$\lambda_1^2(1+(k/\alpha)^{2M})^{\frac{H-1}{M}}$ or $\lambda_1^2\left(1+\frac{k}{\alpha}\right)^{2H-2}+\lambda_2^2\left(1-\left(1+\frac{\alpha}{k}\right)^{2H-2}\right)$	
$P_1^{(k)}$	$\frac{1-\xi}{1/2-\xi}\frac{\mu^2}{\gamma(k)+\mu^2}$	$1-\left(1-P_1^{(k^*)}\right)^{(k/k^*)^\theta}$
$\frac{1}{\zeta(k)}$	1	$\sqrt{(1-2\xi)\left(P_1^{(k)}\frac{\gamma(k)+\mu^2}{\mu^2}-1\right)}$
$\frac{1}{\lambda(k)}$	$\frac{\mu}{(1/2-\xi)(\gamma(k)+\mu^2)}$	$\left(1+\frac{1}{(1-\xi)\left(\zeta(k)\right)^2}-\frac{1}{\left(\zeta(k)\right)^{\sqrt{2}}}\right)\frac{P_1^{(k)}}{\mu}$
x for $\xi > 0$	$\lambda(k)\frac{\left(P_1^{(k)}T^{(k)}/k\right)^\xi-1}{\xi}$	$\lambda(k)\left(\frac{\left(P_1^{(k)}T^{(k)}/k\right)^\xi-1}{\xi}\right)^{\frac{1}{\zeta(k)}}$
x for $\xi = 0$	$x=\lambda(k)\ln\left(P_1^{(k)}T^{(k)}/k\right)$	$x=\lambda(k)\left(\ln\left(P_1^{(k)}T^{(k)}/k\right)\right)^{\frac{1}{\zeta(k)}}$

The ombrian curves per se are given in the last two rows. The transition time scale k^* has a default value of ~100 h.

Parameters of the ombrian model

With seven parameters, the ombrian model achieves:

(a) mathematical and physical consistency;

(b) coverage of all time scales (from zero to infinity);

(c) good respresentation on the very fine time scales, through the fractal parameter M;

Parameter	Meaning of parameter	Related tool
μ	Mean intensity	Mean, μ
λ_1 (& λ_2)	Variance scale parameters[1]	Climacogram, $\gamma(k)$
α	Time scale parameter	Climacogram, $\gamma(k)$
M	Fractal (smoothness) parameter[2]	Climacogram, $\gamma(k)$
H	Hurst parameter	Climacogram, $\gamma(k)$
ξ	Tail index	Probability distribution, $F(x)$
θ	Exponent of the expression of probability dry	Probability wet, $P_1^{(k)}$

[1] One or two parameters for the two cases depending on the choice of the climacogram expression.
[2] The fractal (roughness/smoothness) parameter M is an independent parameter if we have chosen a climacogram expression with one λ; otherwise it is assumed $M = 1 - H$.
[3] The transition time scale k^* but this is not regarded a parameter but a modelling choice.

(d) good respresentation on very large time scales, through the Hurst parameter H and the preserved mean μ whose effect becomes important as time scale increases;

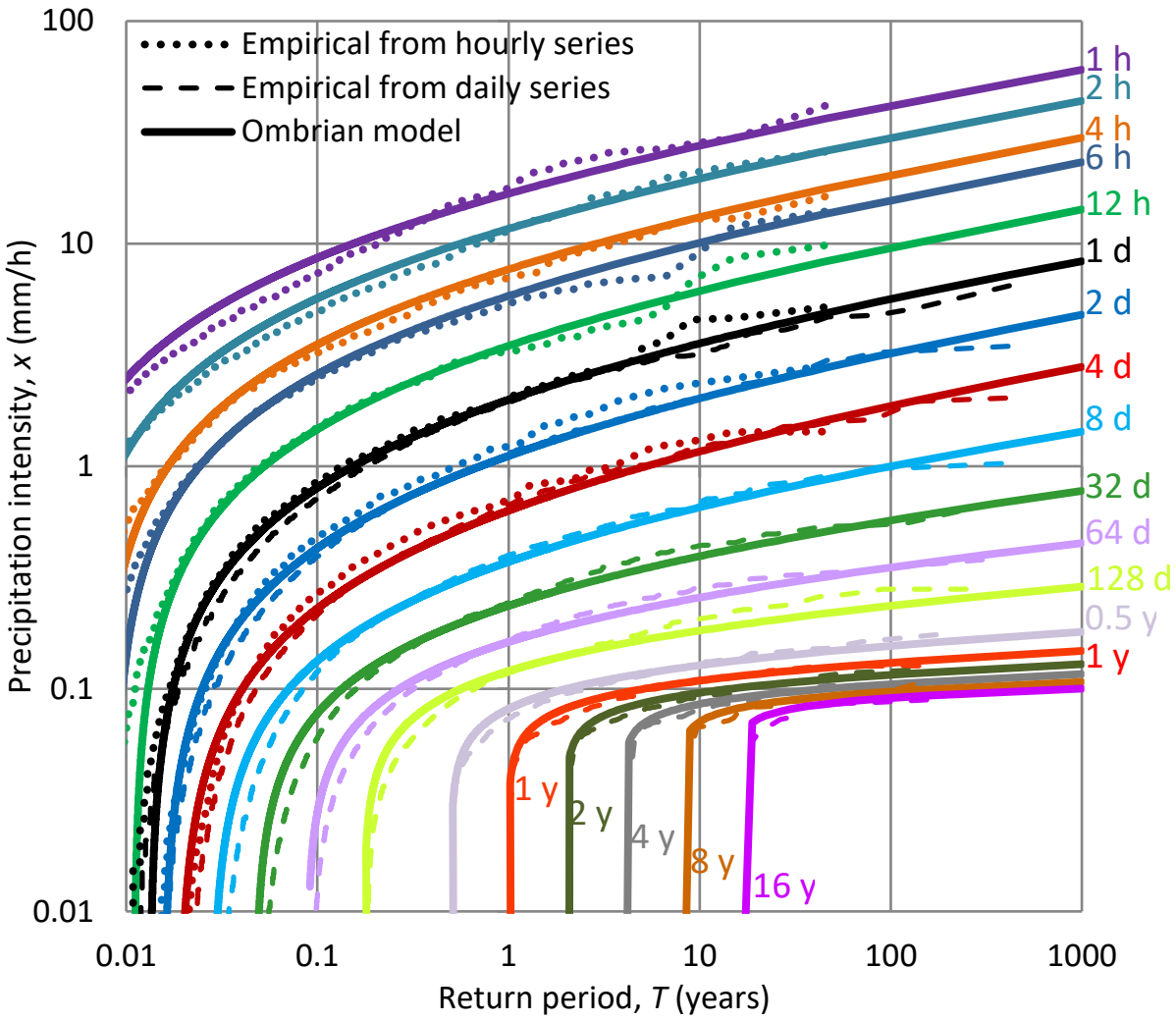
(e) simultaneous treatment and preservation of the climacogram; and

(f) simultaneous treatment and preservation of the probability dry/wet.

Ombrian curves for Bologna: Comparison of model to K-moment estimates of return period

Ombrian curves as resulted from the ombrian model for Bologna for time scales spanning 5 orders of magnitude (1 h to 16 years = 140 256 h). The empirical points are estimated from K-moments. The effect of persistence was taken into account; the model was plotted with bias-adapted variance in order to be comparable with empirical plots.

Parameter values	
μ	0.0823 mm/h
λ_1	0.03317 mm/h
λ_2	1.20 mm/h
α	8.74
H	0.92
ξ	0.121
θ	0.787



A Python implementation of the methodology

- All calculations are easy to make in a spreadsheet software (e.g. Excel).
- A Python implementation is also available online (Iliopoulou and Koutsoyiannis, 2021).
- As seen in the screenshots, the required code for K-moments and climacogram calculations (and all other modules) is very simple.

```
## K-moment for order p
def kmom(p,x): ## Non-central Unbiased for q=1, x must be sorted in increasing order
    x=x[~np.isnan(x)]
    n=len(x)
    b=np.empty(n)
    b[:]=np.NaN
    i=np.arange(1,n+1,1)
    i_0=i[i<p]
    i_1=i[i>=p]
    b[i_0-1]=0
    b[i_1-1]=(p/n)*np.exp(gammaln(n-p+1)-gammaln(n)+gammaln(i_1)-gammaln(i_1-p+1))
    kmom=np.sum(b*x)
    return kmom
```

```
## climacogram models
# Filtered HK Cauchy type
def FHK_C(k,l1,a,M,H):
    g_k=l1*((1+(k/a)**(2*M))**((H-1)/M))
    return g_k

## Filtered HK Cauchy-Dagum-type (FHK-CD) for M=1-H (rough process)
def FHK_CD(k,l1,l2,a,H):
    g_k=l1*(1+k/a)**(2*H-2)+l2*(1-(1+a/k)**(2*H-2))
    return g_k
```

Simplified model version for large scale applications

- For **small time scales** k we make the **simplifying assumptions**:
  - $P_1^{(k)} = \beta\, k$ (probability wet increases in proportionality with time scale).
  - $\gamma(k)/\mu^2 \gg 1$ (large coefficient of variation).
- Under these assumptions the rainfall intensity x for small timescales k (of the order of minutes to a few days) and return period T is given by the following relationships, in which the intensity is given by a separable function $x = b(T)/a(k)$:
  - for return period estimated from a full series or of rainfall exceedances over threshold:

$$x = \frac{b(T)}{a(k)} = \lambda \frac{(\,T/\beta)^{\xi} - 1}{(1 + k/\alpha)^{\eta}}, \qquad \xi > 0$$

  - from series of annual maxima (where Δ = 1 year):

$$x = \lambda \frac{(-(\beta/\Delta)\ln(1 - \Delta/T))^{-\xi} - 1}{(1 + k/\alpha)^{\eta}}, \qquad \xi > 0$$

Theoretically equivalent for all T and for the same parameter values; giving virtually same values for T > 10 years

Comparison of parameter sets

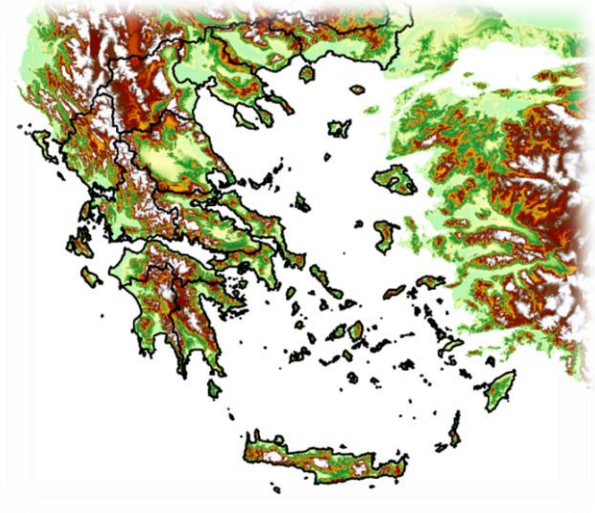
Full ombrian model		Simplified ombrian relationship	
Parameter	Meaning of parameter	Parameter	Meaning of parameter
μ	Mean intensity		
λ_1, λ_2	Intensity scale parameters	λ	Intensity scale parameter
α	Time scale parameter for k	α	Time scale parameter for k
θ	Exponent of the expression of probability dry	β	Ratio of probability wet to time scale (or mean distance of wet periods)
M	Fractal (smoothness) parameter	η	Exponent of the expression of the time scale function $a(k)$
H	Hurst parameter		
ξ	Upper-tail index	ξ	Upper-tail index

- The parameters in the simplifying model are five (fewer by two).
- They **retain their physical/ mathematical meaning.**

Fitting the simplified model

- An attractive feature of this simplification, related to the separable function $x = b(T)/a(k)$, is that it allows the parameters to be estimated **by a convenient, two-step procedure.**
  - The two-step procedure allows using different data sets in each step.
- At the first step the **time scale parameters** (of the expression $a(k)$) are obtained by the optimization procedure introduced in Koutsoyiannis et al. (1998).
  - This step requires **subdaily data.**
- At the second step, the **distributional parameters** (of the expression $b(T)$) are obtained by **the method of K-moments** (Koutsoyiannis, 2025).
  - This step is better performed using **daily datasets**, which are typically available in full series (not only annual maxima), longer and more accurate.
  - The full data series allow inspecting **long-term persistence** (which series of annual maxima tend to hide; Iliopoulou, and Koutsoyiannis, 2019), and the method of K-moments allows taking it into account.
- The fitted parameters can be **generalized for large geographical areas** as was done in the large scale **application for Greece** (see below, as well as in Iliopoulou and Koutsoyiannis, 2021; Iliopoulou et al., 2022, 2024; Koutsoyiannis et al., 2023, 2024; Malamos et al. 2025; Koutsoyiannis, 2025).

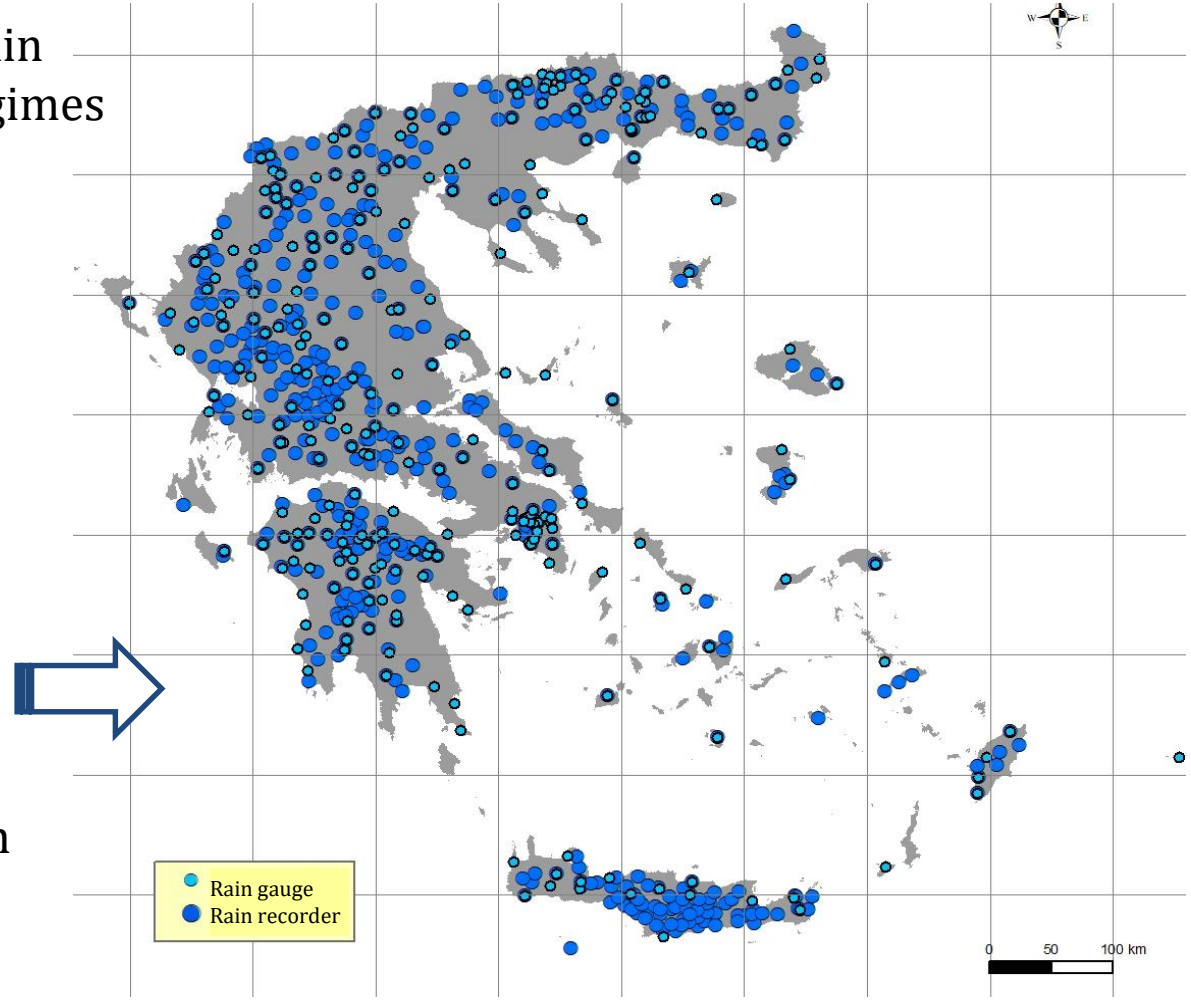
Compiling a dataset of rainfall maxima for Greece



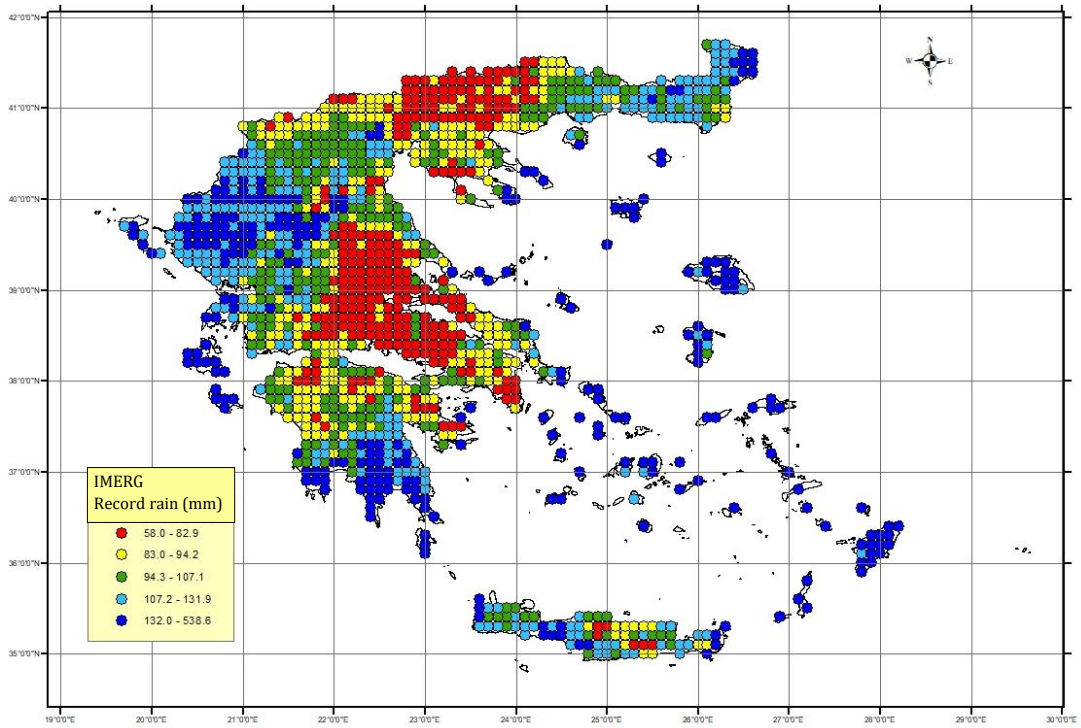
- Highly varied terrain
- Diverse rainfall regimes

After extensive data collection efforts and quality control, we compiled a dataset of **783 stations**, comprising:

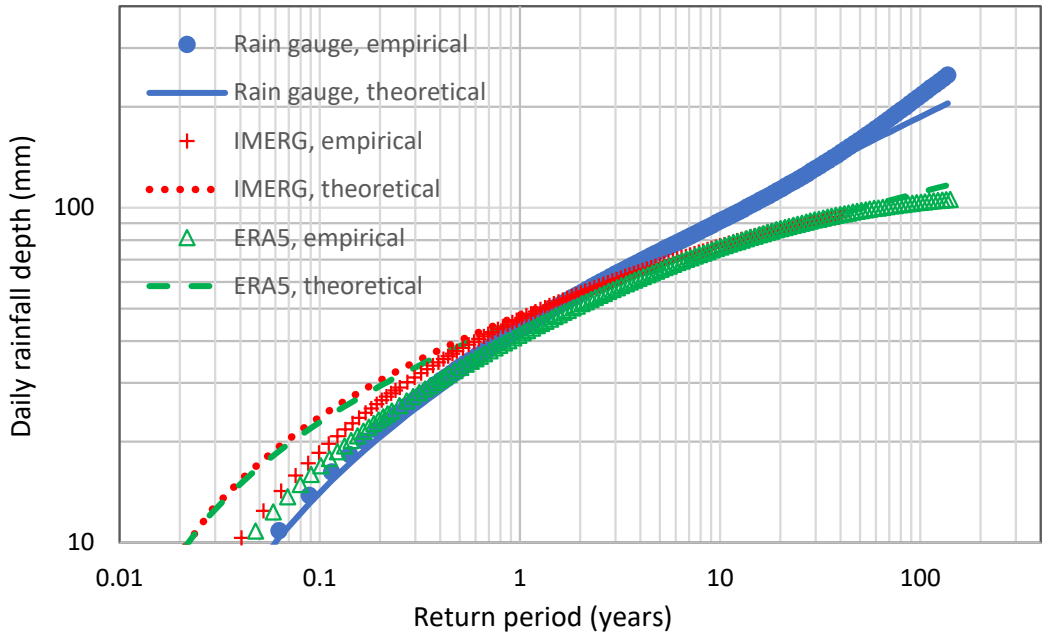
- 503 **daily** rain gauges,
- 280 rain gauges with **sub-daily** resolution (rain recorders).



Non-conventional rainfall data



- From satellite-based information, we investigated the usefulness of the IMERG data set (half hourly time step at 0.1° spatial resolution, period 2000-today),
- From the reanalysis information we investigated the usefulness of the ERA5 data set (daily time step at 0.25° spatial resolution; period 1950-today).

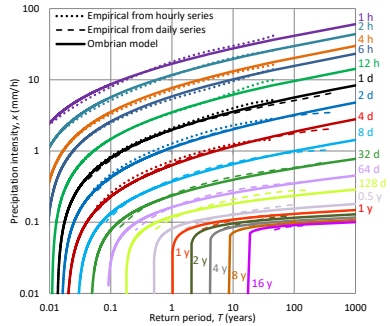


- Both data sets (especially the IMERG) underestimate the highest rainfall depths (as seen in the example for the station of Karditsa) and proved not appropriate for the construction of ombrian curves.

Modelling approach

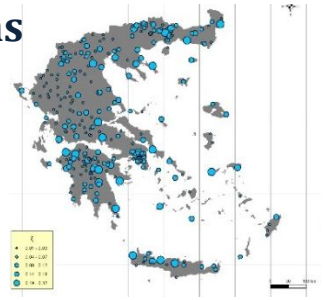
1 At-site independent estimation of parameters

using the **new** framework for **ombrian** relationships

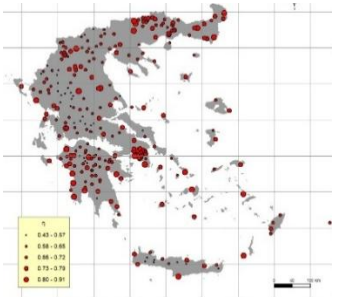


2 Evaluation of spatial variability of parameters:

- ❖ mostly **random patterns**

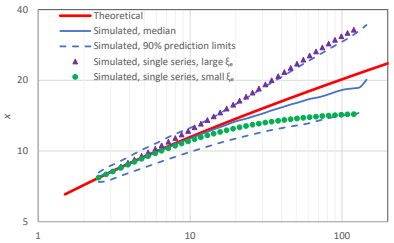


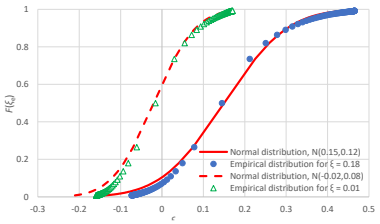
- ❖ mostly **systematic patterns**



3 Identification of **common** parameter values

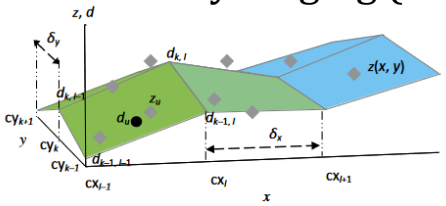
- using simultaneous optimization
- and stochastic simulations

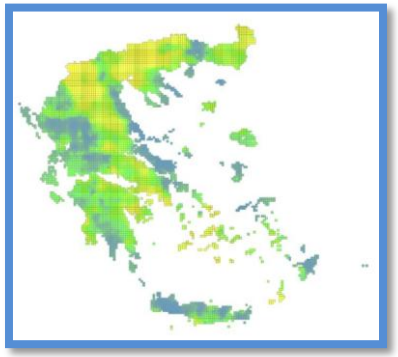




4 Regionalization using **spatial models**

- Inverse Distance Weighted (IDW)
- Bilinear surface smoothing (BSS)
- Ordinary Kriging (OK)





Final product

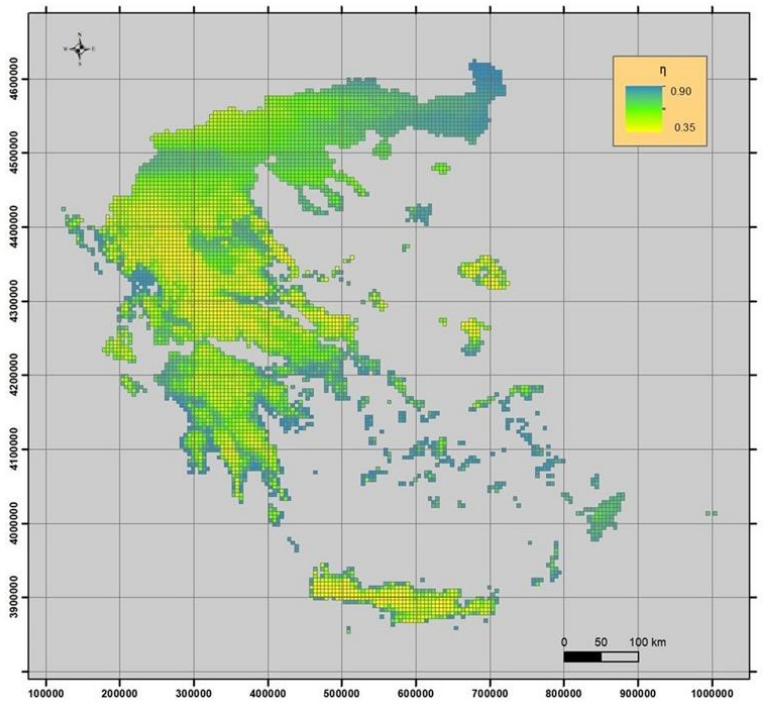
The following generalized form of ombrian curves is derived for rainfall intensity x (mm/h), return period T (years) and temporal scale k (h):

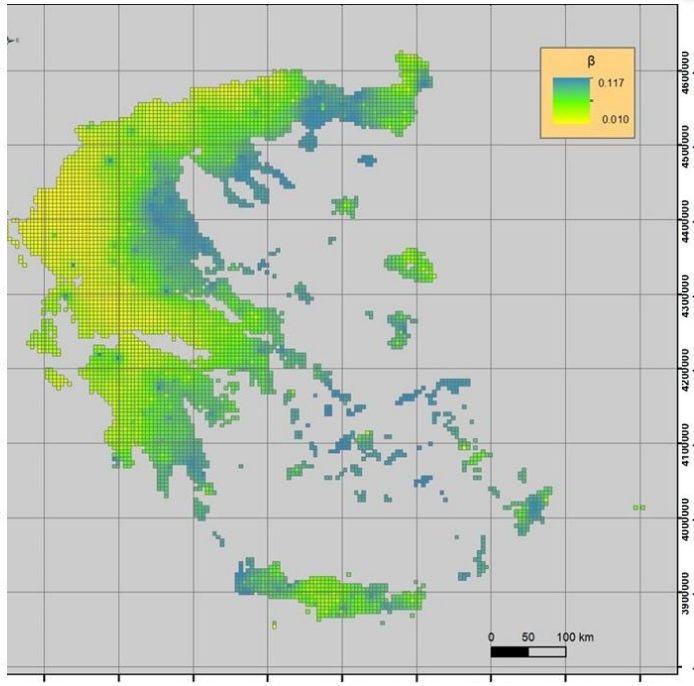
$$x = \lambda_* \frac{(T/\beta_*)^{\xi} - 1}{(1 + k/\alpha)^{\eta_*}}$$

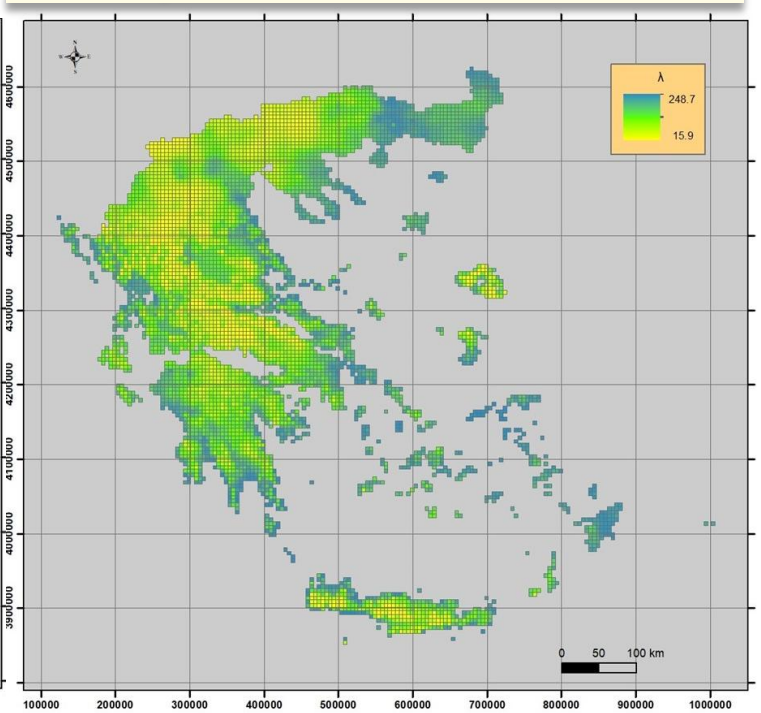
with the following **five** parameters:

- characteristic timescale α = 0.18 h
- tail index ξ = 0.18,
- three spatially varying parameters $\eta_*[-]$, β_* (years) and λ_* (mm/h):

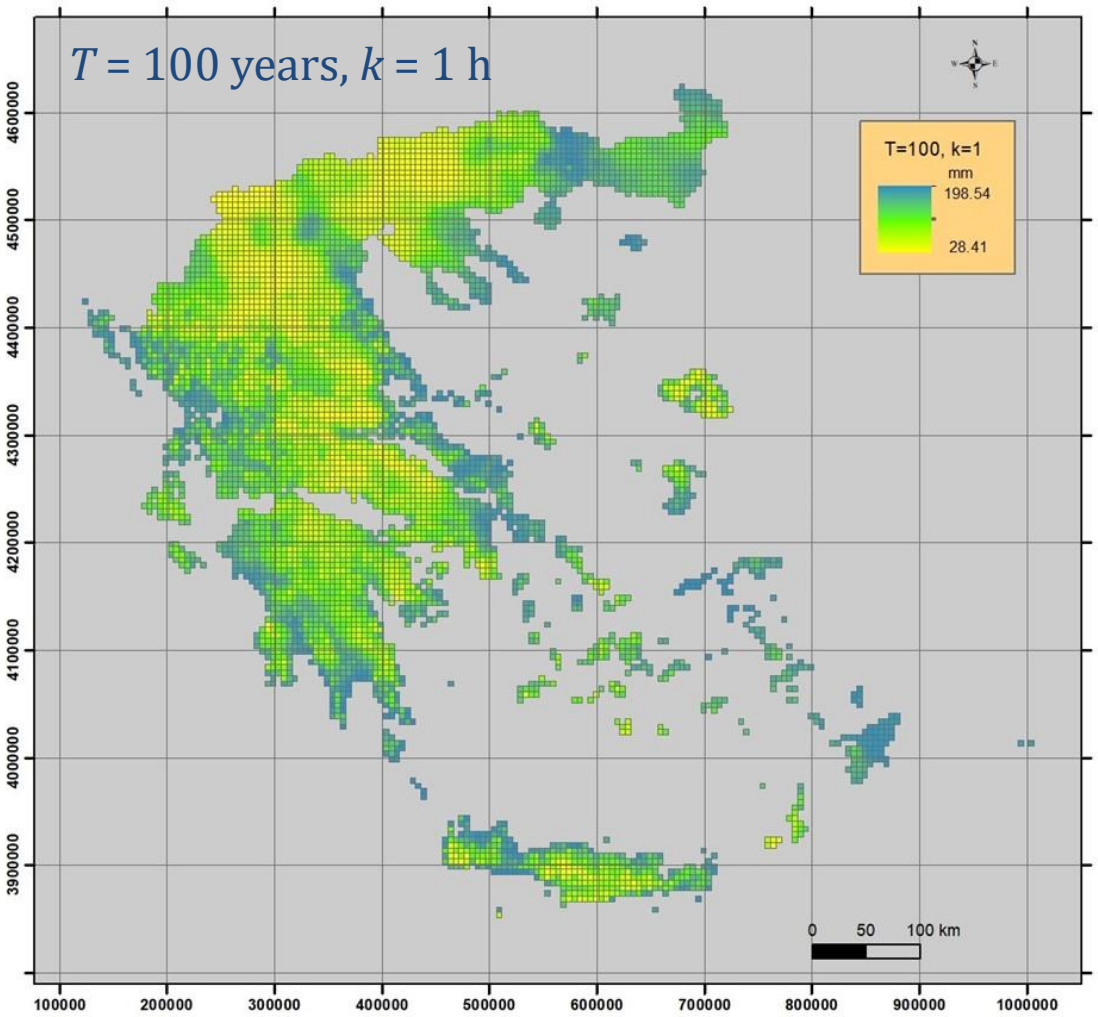
For the first time, we have produced a geographically distributed design rainfall model for the entire Greek territory.

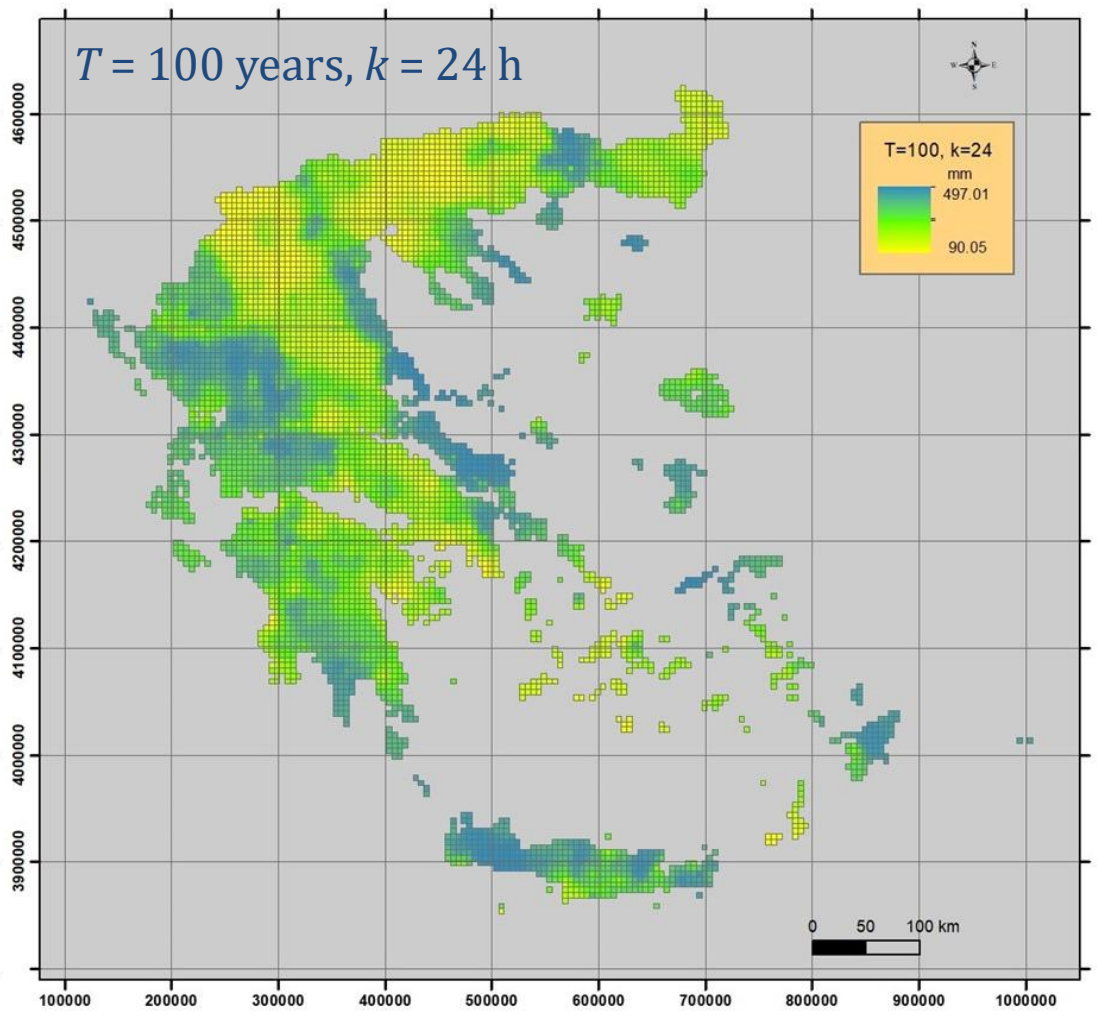




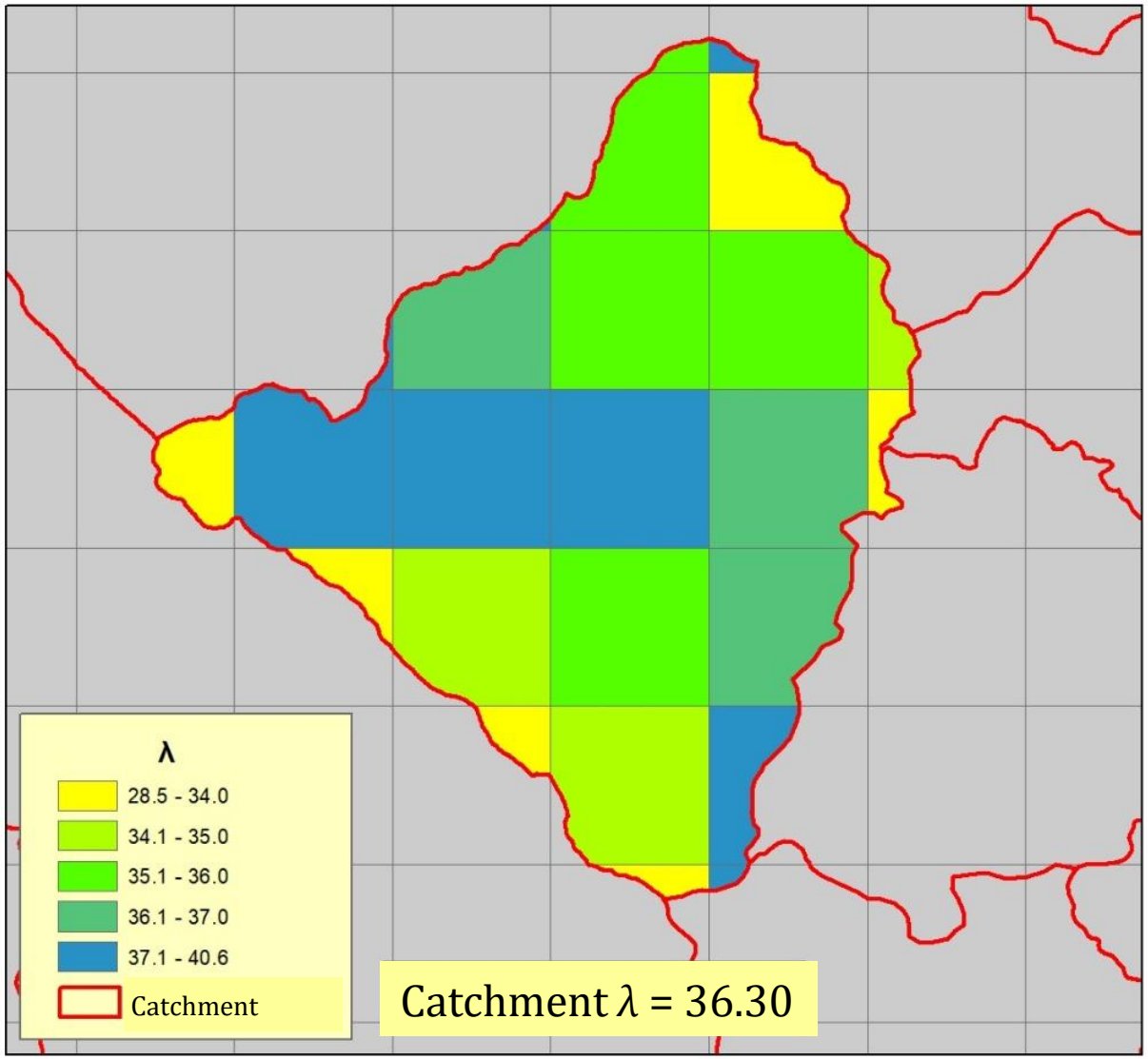


Design rainfall maps





Design rainfall for the catchment scale



The ombrian curves for any region within the Greek territory are derived based on the two constant-value parameters and the three regionally varying parameters which are obtained as a weighted average of the grid points falling within the area.

Concluding remarks

- K-moments are powerful tools that unify other statistical moments (classical, L-, probability weighted) and order statistics, offering several advantages.
- It is important that they have a simple, clear, intuitive and rigorous definition as expectations of maxima and minima.
- They offer a sound basis for distribution fitting with emphasis on extremes.
- For standard (IID) samples, K-moments offer unbiased, reliable and workable estimators for low and high orders p, up to order equal to the sample size n.
- Time dependence influences the estimates, yet K-moments offer a basis to assess that influence and properly adapt the estimates.
- Rainfall extremes can be effectively modelled by a rather simple ombrian model, which, in addition to modelling extremes, provides a good representation the climacogram of the rainfall process and its intermittence.
- A regionalization approach of ombrian relationships enables integration of smoothing and interpolation methods and flexible usage of different data sources.
- A large scale application covering the entire territory of Greece resulted in a powerful tool, easy to apply for engineering tasks.

References

Gini, C., 1936. On the measure of concentration with special reference to income and statistics. Colorado College Publication, General Series, 208, 73.

Iliopoulou, T., and Koutsoyiannis, D., 2019. Revealing hidden persistence in maximum rainfall records,.*Hydrological Sciences Journal*, 64 (14), 1673–1689, doi:10.1080/02626667.2019.1657578.

Iliopoulou, T., and Koutsoyiannis, D., 2021. PythOm: A python toolbox implementing recent advances in rainfall intensity (ombrian) curves. *EGU General Assembly 2021, online*, EGU21-389, doi:10.5194/egusphere-egu21-389, European Geosciences Union, https://www.itia.ntua.gr/2111/.

Iliopoulou, T., Koutsoyiannis, D., Malamos, N., Koukouvinos, A., Dimitriadis, P., Mamassis, N., Tepetidis, N., and Markantonis, D., 2024. A stochastic framework for rainfall intensity–time scale–return period relationships. Part II: point modelling and regionalization over Greece. *Hydrological Sciences Journal*, 69 (8), 1092–1112, doi:10.1080/02626667.2024.2345814.

Iliopoulou, T., Malamos, N., and Koutsoyiannis, D., 2022/ Regional ombrian curves: Design rainfall estimation for a spatially diverse rainfall regime. *Hydrology*, 9 (5), 67, doi: 10.3390/hydrology9050067.

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

Koutsoyiannis, D., 2022. Replacing histogram with smooth empirical probability density function estimated by K-moments. *Sci*, 4 (4), 50, doi:10.3390/sci4040050.

Koutsoyiannis, D., 2023. Knowable moments in stochastics: Knowing their advantages, *Axioms*, 12 (6), 590, doi:10.3390/axioms12060590.

Koutsoyiannis, D., 2025. *Stochastics of Hydroclimatic Extremes - A Cool Look at Risk*. Edition 5, *ISBN: 978-618-85370-0-2*, 420 pages, doi:10.57713/kallipos-1, Kallipos Open Academic Editions, Athens, https://www.itia.ntua.gr/2000/.

Koutsoyiannis, D., and Sargentis, G.-F., 2026. Trade-off between entropy and Gini Index in income distribution. *Entropy*, 28 (1), 35, doi:10.3390/e28010035.

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

Koutsoyiannis, D., Iliopoulou, T., Koukouvinos, A., Malamos, N., Mamassis, N., Dimitriadis, P., Tepetidis, N., and Markantonis, D., 2023. In search of climate crisis in Greece using hydrological data: 404 Not Found, *Water*, 15 (9), 1711, doi:10.3390/w15091711.

Koutsoyiannis, D., Iliopoulou, T., Koukouvinos, A., Malamos, 2024. A stochastic framework for rainfall intensity–time scale–return period relationships. Part I: theory and estimation strategies. *Hydrological Sciences Journal*, 69 (8), 1082–1091, doi:10.1080/02626667.2024.2345813.

Lombardo, F., Napolitano, F., Russo, F., and D. Koutsoyiannis, D., 2019. On the exact distribution of correlated extremes in hydrology. *Water Resources Research*, doi:10.1029/2019WR025547.

Malamos, N., Iliopoulou, T., Dimitriadis, P., and Koutsoyiannis, D., 2025. Comparative analysis of IMERG satellite rainfall and elevation as covariates for regionalizing average and extreme rainfall patterns in Greece by means of Bilinear Surface Smoothing. *Geosciences*, 15, 212, doi:10.3390/geosciences15060212,

Data sources

GHCN Version 3 data: https://climexp.knmi.nl/gdcnprcp.cgi?WMO=ITE00100550

Dext3r data: http://www.smr.arpa.emr.it/dext3r/