ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ISTITUTO DI STUDI SUPERIORI
INSTITUTE OF ADVANCED STUDIES

Events and Lectures 2019

Sala Rossa, Palazzo Marchesini, via Marsala 26, Bologna, 26 November 2019

Knowable moments for high-order characterization and modelling of hydrological processes for sustainable management of water resources

Theoretical framework and four applications

Demetris Koutsoyiannis

School of Civil Engineering, National Technical University of Athens, Greece
(dk@itia.ntua.gr, http://www.itia.ntua.gr/dk/)

Visiting professor, Dipartimento Ingegneria Civile, Chimica, Ambientale e dei Materiali (DICAM), Università di Bologna

Presentation available online: http://www.itia.ntua.gr/2012/

A reason for pursuing the truth…

καὶ γνώσεσθε τὴν ἀλήθειαν, καὶ ἡ ἀλήθεια ἐλευθερώσει ὑμᾶς (κατὰ Ἰωάννην 8:32)

et cognoscetis veritatem, et veritas liberabit vos (Ioannes 8:32)

e conoscerete la verità, e la verità vi farà liberi (Giovanni 8:32)

Then you will know the truth, and the truth will set you free (John 8:32)

Theoretical framework:
From classical moments to K-moments

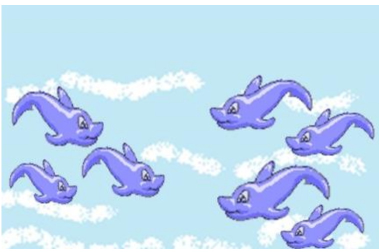
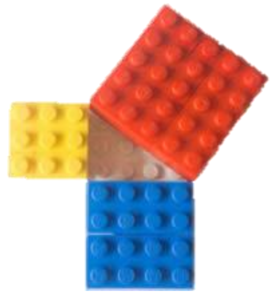
To be followed by four applications

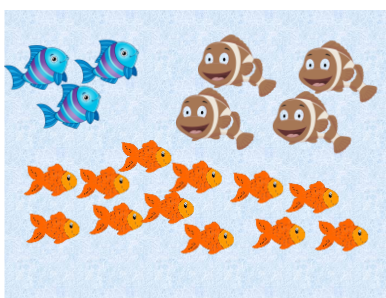
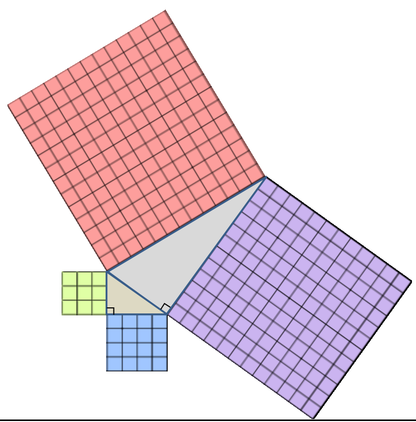
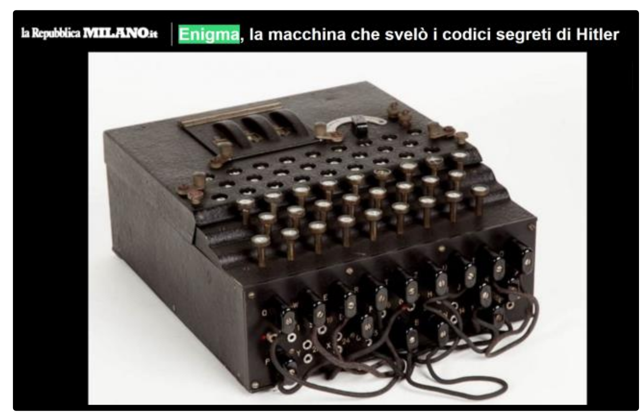
1. **Monitoring and characterization of the hydroclimatic evolution**
2. **Fitting of probabilistic models**
3. **Full stochastic modelling of rainfall (construction of an ombrian model)**
4. **Stochastic (Monte Carlo) simulation**

The applications are based on data from Bologna and Athens

In the beginning was the … sum

What is the result of **raising to a power and adding**, i.e. $\sum_{i=1}^{n} x_i^p$?

<table>
<tr><th>Linear, p = 1</th><th>Pythagorean, p = 2</th><th>Fermatian of high order, e.g., p = 8</th></tr>
<tr>
<td>$3 + 4 = 7$
</td>
<td>$3^2 + 4^2 = 5^2$
</td>
<td>$3^8 + 4^8 = ?^8$

</td>
</tr>
<tr>
<td>$3 + 4 + 12 = 19$
</td>
<td>$3^2 + 4^2 + 12^2 = 13^2$
</td>
<td>$3^8 + 4^8 + 12^8 = ?^8$

</td>
</tr>
</table>

The easy solution of the enigma

- What is the result of **raising to a power and adding**, i.e. $\sum_{i=1}^{n} x_i^p$?

Linear, p = 1	Pythagorean, p = 2	Fermatian of high order, p = 8
3 + 4 = 7	$3^2 + 4^2 = 5^2$	$3^8 + 4^8 \approx 4^8$
3 + 4 +12 = 19	$3^2 + 4^2 + 12^2 = 13^2$	$3^8 + 4^8 + 12^8 \approx 12^8$

- Symbolically, for large (or even modest) p the result is*:

$$\sum_{i=1}^{n} x_i^p \approx \left(\max_{1\le i\le n}(x_i)\right)^p$$

- We recall that this sum, if divided by n,† is the **estimate** of the (noncentral probabilistic) **moment**, μ_p'. Hence:

$$\hat{\mu}_p' = \frac{1}{n}\sum_{i=1}^{n} x_i^p \approx \frac{1}{n}\left(\max_{1\le i\le n}(x_i)\right)^p$$

- Thus, for an unbounded variable and for large p, we conclude that $\underline{\hat{\mu}}_p'$ is **more an estimator of an extreme quantity**, **than an estimator of the moment μ_p'**.
- Thus, unless p is very small, **μ_p' is *unknowable*** (Koutsoyiannis, 2019a): we cannot infer its value from data. **This is the case even if n is extraordinarily large**!
- Also, the various $\underline{\hat{\mu}}_p'$ for different orders p are in fact deformed copies of the same thing: they only differ on the power to which the maximum value is raised.

* This is precise if x_i are positive.
† Note that for large p the term $(1/n)$ could be omitted with a negligible error.

Moments and statistical inference

- We recall that for a **stochastic variable*** $\underline{x}$ the **noncentral** (or raw) and **central moments** of order p are defined as the **expectations**:

$$\mu'_p \coloneqq \mathrm{E}\big[\underline{x}^p\big] = \int_{-\infty}^{\infty} x^p f(x)\mathrm{d}x, \quad \mu_p \coloneqq \mathrm{E}\Big[\big(\underline{x}-\mu\big)^p\Big] = \int_{-\infty}^{\infty} \big(\underline{x}-\mu\big)^p f(x)\mathrm{d}x$$

  respectively, where $f(x)$ is the probability density function and $\mu \coloneqq \mu'_1 = \mathrm{E}\big[\underline{x}\big]$ is the mean.
- Indeed their standard estimators from a sample $\underline{x}_i$, $i = 1, \ldots, n$, are:

$$\underline{\hat{\mu}}'_p = \frac{1}{n}\sum_{i=1}^{n} \underline{x}_i^p, \qquad \underline{\hat{\mu}}_p = \frac{b(n,p)}{n}\sum_{i=1}^{n}\big(\underline{x}_i - \hat{\mu}\big)^p$$

  where $b(n, p)$ is a bias correction factor (e.g. for the variance $\mu_2 \eqqcolon \sigma^2$, $b(n, 2) = n/(n-1)$).
- Statistical inference, which is the **formal probabilistic induction** (the modern version of the **Aristotelian epagoge / επαγωγή**) is based on expectations, and in particular, moments, which are estimated from samples by virtue of **stationarity** and **ergodicity**.
- In theory **the ergodic theorem** enables estimation of moments from data as $n \to \infty$, irrespective of the order p. But as we have seen for finite n and even modest p they are **unknowable**.
- In typical hydrological records it is practically impossible to use the estimators for $p > 2$:

  cf. Lombardo et al. (2014), "**Just two moments**".
- But **high-order moments** are important to characterize a process—particularly its **behaviour in extremes**.

---

* **Stochastic variables**, also known as **random variables** are denoted here by underlined symbols.

A first parenthesis: Stationarity and ergodicity are rigorous scientific concepts—not political ideas

POLICYFORUM

CLIMATE CHANGE

Stationarity Is Dead: Whither Water Management?

P. C. D. Milly,[1*] Julio Betancourt,[2] Malin Falkenmark,[3] Robert M. Hirsch,[4] Zbigniew W. Kundzewicz,[5] Dennis P. Lettenmaier,[6] Ronald J. Stouffer[7]

Climate change undermines a basic assumption that historically has facilitated management of water supplies, demands, and risks.

AGU PUBLICATIONS

Water Resources Research

COMMENTARY

10.1002/2014WR016092

Modeling and mitigating natural hazards: Stationarity is immortal!

Alberto Montanari[1] and Demetris Koutsoyiannis[2]

[1]Department of Civil, Chemical, Environmental, and Materials Engineering, University of Bologna, Bologna, Italy, [2]Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Athens, Greece

Correspondence to:
A. Montanari,
alberto.montanari@unibo.it

Citation:
Montanari, A., and D. Koutsoyiannis (2014), Modeling and mitigating natural hazards: Stationarity is immortal!, *Water Resour. Res.*, *50*, doi:10.1002/2014WR016092.

Abstract Environmental change is a reason of relevant concern as it is occurring at an unprecedented pace and might increase natural hazards. Moreover, it is deemed to imply a reduced representativity of past experience and data on extreme hydroclimatic events. The latter concern has been epitomized by the statement that "stationarity is dead." Setting up policies for mitigating natural hazards, including those triggered

1174 *Hydrological Sciences Journal – Journal des Sciences Hydrologiques*, 60 (7–8) 2015
http://dx.doi.org/10.1080/02626667.2014.959959
Special issue: *Modelling Temporally-variable Catchments*

Negligent killing of scientific concepts: the stationarity case

Demetris Koutsoyiannis[1] and Alberto Montanari[2]

Stationarity and ergodicity constitute the scientific foundation in making inference from data; they enable translation of empirical knowledge into the mathematical language of probability and stochastics.

They are tightly connected to each other.

Without stationarity there cannot be ergodicity.

Without ergodicity inference from data would not be possible.

Ironically, several studies use time series to estimate statistical properties, as if the process was ergodic, while at the same time what they (cursorily) estimate may falsify the ergodicity hypothesis.

Important mathematical concepts **need to be understood** (even if that is not so easy…). **Not to be killed** (which actually is impossible as **they are immortal!**)

Our humble contribution: Seeking theoretical consistency in analysis of geophysical data (Using *stochastics*)

Book in preparation (if the frame of this project):

D. Koutsoyiannis, 2020. *Stochastics of Hydroclimatic Extremes – A Cool Look at Risk*.

Note: Much of the content of this presentation will be included in the book.

The founders of the concepts of stationarity and ergodicity

Ludwig Boltzmann

(1844 –1906, Universities of Graz and Vienna, Austria, and Munich, Germany)

1877 Explanation of the concept of **entropy** in probability theoretic context.

1884/85 Introduction of the term “**ergode**” and the notion of **ergodic*** systems which however he called “isodic”

* The term is etymologized from Greek words but which ones exactly is uncertain (options: (a) έργον + οδός; (b) έργον + είδος; (c) εργώδης; see Mathieu, 1988).

George D. Birkhoff

(1884 – 1944; Princeton, Harvard, USA)

1931 Discovery of the **ergodic (Birkhoff–Khinchin) theorem**

Aleksandr Khinchin

(1894 – 1959; Moscow State University, Russia/Soviet Union)

1933 Purely measure-theoretic proof of the **ergodic (Birkhoff–Khinchin) theorem**

1934 Definition of **stationary stochastic processes** and probabilistic setting of the **Wiener-Khinchin theorem** relating autocovariance and power spectrum

Andrey N. Kolmogorov

(1903 – 1987; Moscow State University, Russia/Soviet Union)

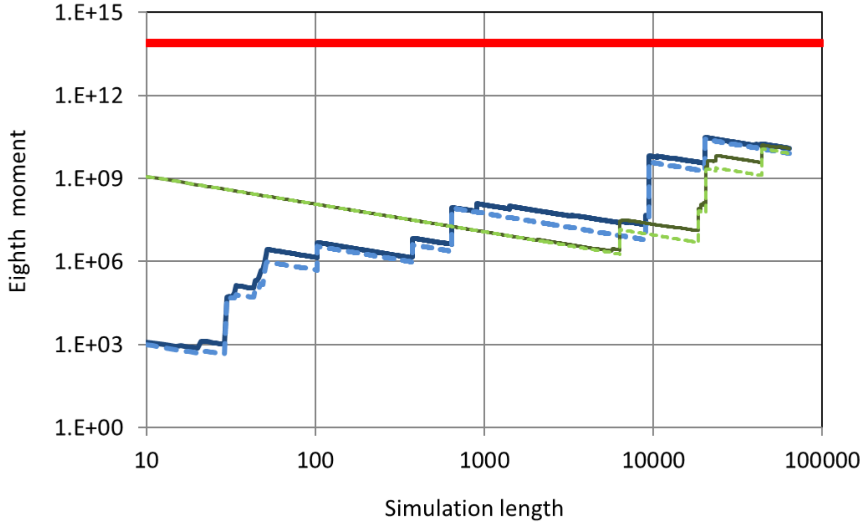
1931 Introduction of the terms **process** to describe change of a certain system and **stationary** to describe a probability density function that is unchanged in time

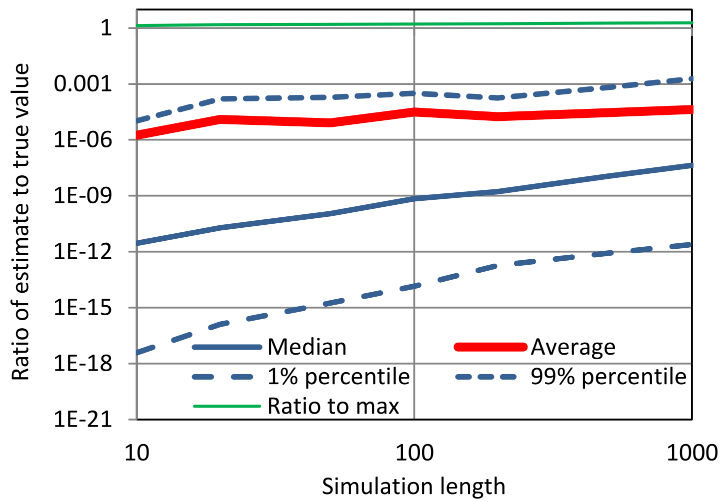
1933 Definition of the concepts of **probability** & **random variable**

1937-1938 Probabilistic exposition of the **ergodic (Birkhoff–Khinchin) theorem** and **stationarity**

1947 Definition of **wide sense stationarity**

Illustration of slow convergence of moment estimators





Convergence of the sample estimate of the eighth non-central moment to its true value (thick horizontal line) corresponding to a lognormal distribution LN(0,1) where the process is an exponentiated Hurst-Kolmogorov process with Hurst parameter H = 0.9. The sample moments ($\sum_{i=1}^{n} x_i^p \,/n$ with p = 8; continuous lines), are estimated from a single simulation of length 64 000, subset to sample size n from 10 to 64 000, with the subsetting being done either from the beginning to the end or from the end to the beginning. Dashed lines represent maximum values $(\max_{1\le i\le n}(x_i))^p/n$.

As in the example on the left but for 200 simulated series of length 1000 each. The sampling distribution of the eighth moment estimator $\sum_{i=1}^{n} \underline{x}_i^8 \,/n$ is visualized by the percentiles, the median and the average, plotted as ratios to the true value. Theoretically, the ratio should be 1, but it is smaller by many orders of magnitude, and the convergence to 1 is very slow. (The convergence of the average could also be achieved if we used millions of simulated series instead of 200). In contrast, the ratio to $(\max_{1\le i\le n}(x_i))^8/n$ is ≈1.

From classical (but unknowable) moments to knowable moments

To derive **knowable** moments for high orders p, in the expectation defining the pth moment:

$$\mu'_p := \mathrm{E}\left[\underline{x}^p\right]$$

we raise $\underline{x}$ to a smaller power $q < p$ (e.g. $q = 1, q = 2$) and for the remaining $(p - q)$ terms in the multiplication $\underline{x}^p = \underbrace{\underline{x} \dots \underline{x}}_{p}$ we replace $\underline{x}$ with the distribution function $F(\underline{x})$:

$$\underline{x}^p \rightarrow \left(F(\underline{x})\right)^{p-q} \underline{x}^q$$

We multiply the latter quantity by $(p - q + 1)$ and take its expected value. This leads to the following definition of **noncentral knowable moment**:

$$K'_{pq} := (p - q + 1)\mathrm{E}\left[\left(F(\underline{x})\right)^{p-q} \underline{x}^q\right], \qquad p \geq q$$

Likewise, we can define **central** and **hypercentral knowable moments** by the following substitutions:

$$(\underline{x} - \mu)^p \rightarrow \left(F(\underline{x})\right)^{p-q} (\underline{x} - \mu)^q \text{ or } (\underline{x} - \mu)^p \rightarrow \left(2F(\underline{x}) - 1\right)^{p-q} (\underline{x} - \mu)^q$$

Knowable moments or **K-moments**, introduced by Koutsoyiannis (2019a, 2020), contain as special cases (or are one-to-one connected to) classical moments, Probability Weighted Moments and L-moments, and are tightly connected to expectations of order statistics.

Formal definition of K-moments

Noncentral knowable moment of order $(p, 1)$ [**analogous to Probability Weighted Moments**]

$$K'_p \coloneqq p\mathrm{E}\left[\left(F(\underline{x})\right)^{p-1} \underline{x}\right], \qquad p \geq 1$$

Noncentral knowable moment (or *noncentral K-moment*) *of order* (p, q) [**recovering classical noncentral moments for $\boldsymbol{p = q}$**]:

$$K'_{pq} \coloneqq (p - q + 1)\mathrm{E}\left[\left(F(\underline{x})\right)^{p-q} \underline{x}^q\right], \qquad p \geq q$$

Central knowable moment of order (p, q) [**recovering classical central moments for $\boldsymbol{p = q}$**]

$$K_{pq} \coloneqq (p - q + 1)\mathrm{E}\left[\left(F(\underline{x})\right)^{p-q} (\underline{x} - \mu)^q\right], \qquad p \geq q$$

where μ is the mean of $\underline{x}$, i.e., $\mu \coloneqq \mathrm{E}[\underline{x}_{(p)}] \equiv K'_1$.

Hypercentral knowable moment (or *central K-moment*) *of order* (p, q) [**analogous to L-moments**]

$$K^+_{pq} \coloneqq (p - q + 1)\mathrm{E}\left[\left(2F(\underline{x}) - 1\right)^{p-q} (\underline{x} - \mu)^q\right], \qquad p \geq q$$

K-moments and order statistics

A **sample** of a stochastic variable $\underline{x}$ is by definition a set $\{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n\}$ of independent copies of $\underline{x}$. We may arrange the sample in increasing order of magnitude such that $\underline{x}_{(i:n)}$ be the ith smallest of the n, i.e.:

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

The stochastic variable $\underline{x}_{(i:n)}$ is termed the ith **order statistic**. The minimum and maximum are, respectively,

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

and represent special cases of the order statistics, the lowest and the highest.

The distribution of the order statistic $\underline{x}_{(i:n)}$ is given in terms of the Beta distribution function as:

$$F_{(i:n)}(x) = P\left\{\underline{x}_{(i:n)} \le x\right\} = P\left\{\underline{u} \le F(x)\right\} = \frac{\mathrm{B}_{F(x)}(i, n-i+1)}{\mathrm{B}(i, n-i+1)}$$

For the special cases of the minimum and maximum we have, respectively,

$$F_{(1:n)}(x) = \frac{\mathrm{B}_{F(x)}(1,n)}{\mathrm{B}(1,n)} = 1 - \left(1 - F(x)\right)^n, \qquad F_{(n:n)}(x) = \frac{\mathrm{B}_{F(x)}(n,1)}{\mathrm{B}(n,1)} = \left(F(x)\right)^n$$

It is shown that the expectation of the order statistic $\underline{x}_{(i:n)}$ is related to the noncentral K-moments by:

$$\frac{\mathrm{E}\left[\underline{x}_{(i:n)}^q\right]}{i} = \binom{n}{i} \sum_{j=0}^{n-i} \binom{n-i}{j} (-1)^j \frac{K'_{i+j-1+q,q}}{i+j}, \qquad \binom{n}{p-q+1} K'_{pq} = \sum_{i=p-q+1}^{n} \binom{i-1}{p-q} \mathrm{E}\left[\underline{x}_{(i:n)}^q\right]$$

K-moments and expectation of extremes

Based on the definition of K-moments it is readily seen that

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

More generally, K-moments of all categories represent expected values of extremes. Thus, for odd q or for nonnegative $\underline{x}$ (so that $\underline{x}_1^q$ be monotonic function of $\underline{x}$):

$$K'_{pq} = \mathrm{E}\left[\max\left(\underline{x}_1^q, \underline{x}_2^q, \dots, \underline{x}_{p-q+1}^q\right)\right] = \mathrm{E}\left[\underline{x}_{(p-q+1)}^q\right]$$

Furthermore, for odd q we have,

$$K_{pq} = \mathrm{E}\left[\max\left(\left(\underline{x}_1 - \mu\right)^q, \left(\underline{x}_2 - \mu\right)^q, \dots, \left(\underline{x}_{p-q+1} - \mu\right)^q\right)\right]$$

which means that the central K-moment K_{pq} of $\underline{x}$ is identical to the expected maximum of order $p - q + 1$ of $\underline{z} = \left(\underline{x} - \mu\right)^q$.

The above properties also hold asymptotically, for large p, also for even q in any case of positively skewed distribution.

For a symmetric distribution, an analogous property holds for the hypercentral moments with even q:

$$K_{pq}^{+} = \mathrm{E}\left[\max\left(\left(\underline{x}_1 - \mu\right)^q, \left(\underline{x}_2 - \mu\right)^q, \dots, \left(\underline{x}_{p-q+1} - \mu\right)^q\right)\right]$$

which means that the hypercentral K-moment K_{pq}^{+} of a stochastic variable $\underline{x}$ with symmetrical distribution for q even is identical to the expected maximum of order $p - q + 1$ of $\underline{z}$. In contrast, for q odd, the hypercentral K-moment K_{pq}^{+} will obviously be zero.

Are those high-order K-moments knowable?

Yes, because we can construct estimators with good properties such as unbiasedness, small variance and fast convergence to the true value.

The unbiased estimator of the noncentral moment K'_{p1} and its extension for $q > 1$ are (Koutsoyiannis, 2020):

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

where for any positive number p (usually, but not necessarily, integer):

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

It can be verified that

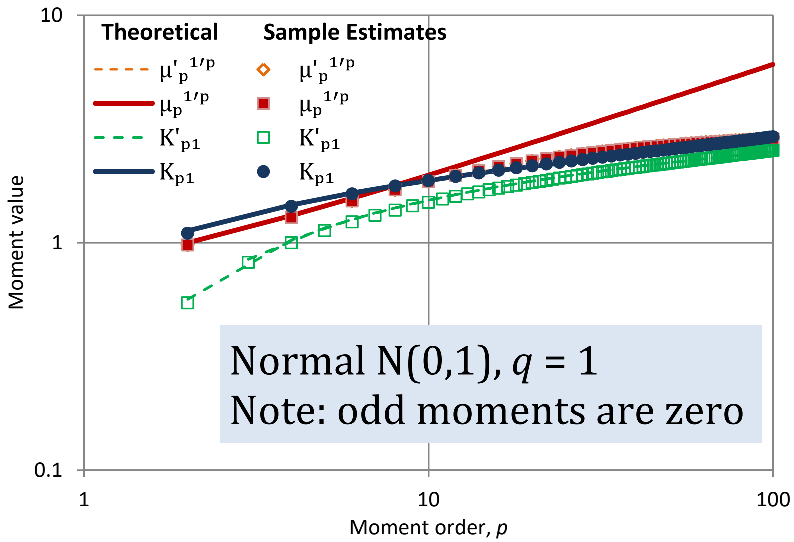
$$\textstyle\sum_{i=1}^{n} b_{inp} = 1$$

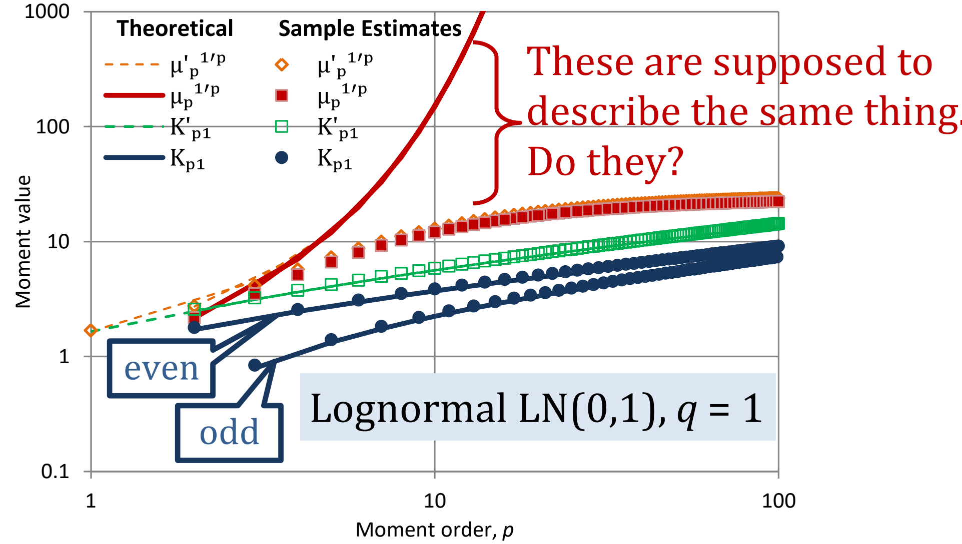
which is a necessary condition for unbiasedness. Furthermore, for $p = 1$, $b_{in1} = 1/n$ and thus we recover the estimator of the mean. For $p = 2$, the quantity $(n/2)b_{in2}$ is the estimator $\hat{F}\left(x_{(i)}\right)$, i.e.,

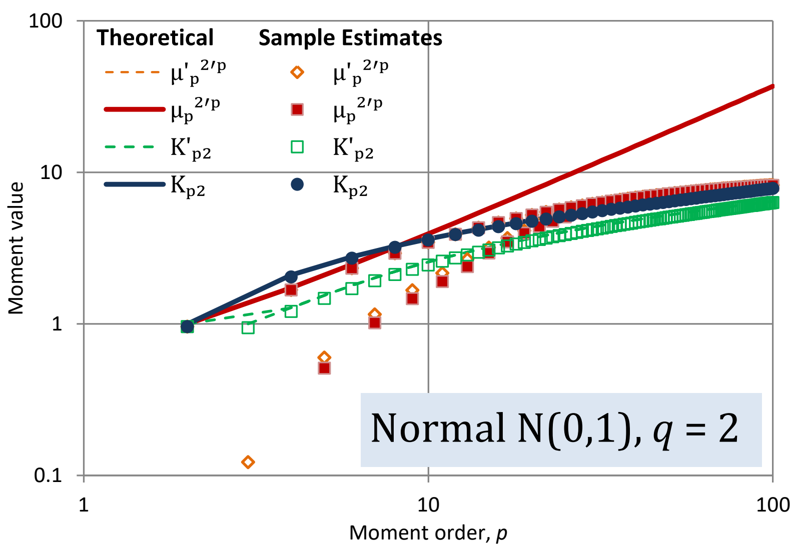
$$\hat{F}\left(x_{(i)}\right) = \frac{i-1}{n-1}$$

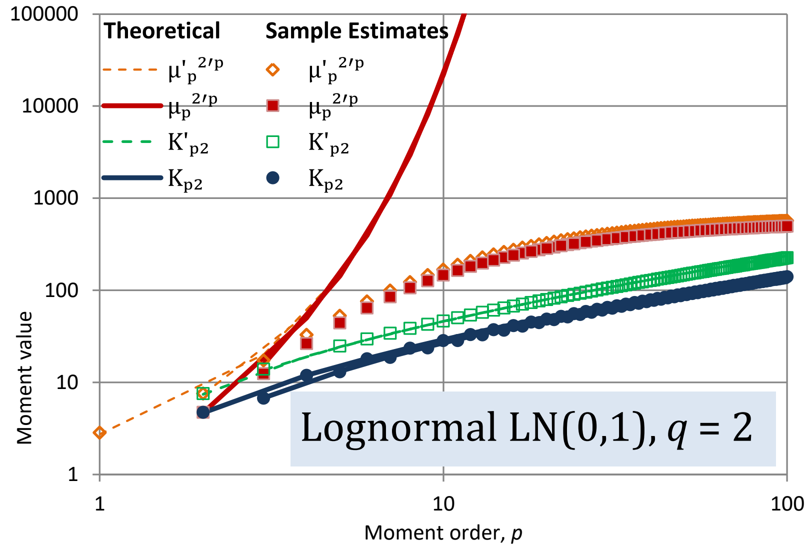
Because $b_{inp} = 0$ for $i < p$, as the moment order increases, progressively, fewer data values determine the moment estimate, until it remains only one, the maximum, when $p = n$, with $b_{nnn} = 1$. Furthermore, if $p > n$ then $b_{inp} = 0$ for all i, $1 \leq i \leq n$, and therefore estimation becomes impossible.

Justification of the notion of unknowable vs. knowable









Note: Sample sizes are ten times higher than the maximum p shown in graphs, i.e., 1000.

What happens if instead of a sample we have a stochastic process? Are the estimators still unbiased?

A first important distinction:

Real world	Theories and models

A second important distinction, regarding models:

Classical statistics (independence)	Statistics within stochastics (dependence)
Sample: identical and independent copies of a random variable	**Stochastic process** (discrete or discretized): families of dependent random variables
Observed sample: observations of the variables that constitute the sample	**Time series**: observations of the same process in consecutive times

In geophysics (including hydroclimatic processes) we **cannot make samples**.

We have always to deal with **time series**.

In geophysical time series, **time dependence and persistence are marked**.

Unbiased estimators of classical statistics are no longer unbiased in stochastics (with the exception of mean—and of noncentral moments, which however suffer from unknowability); thus, they **need to be adapted**.

A second parenthesis: Kolmogorov, Hurst and the Nilometer

Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS
1940. Volume XXVI, № 2

MATHEMATIK

WIENERSCHE SPIRALEN UND EINIGE ANDERE INTERESSANTE KURVEN IM HILBERTSCHEN RAUM

Von A. N. KOLMOGOROFF, Mitglied der Akademie

Wir werden hier einige Sonderfälle von Kurven betrachten, denen meine vorhergehende Note «Kurven im Hilbertschen Raum, die gegenüber einer einparametrigen Gruppe von Bewegungen invariant sind» (¹) gewidmet ist.

Unter einer Ähnlichkeitstransformation im Hil...
Raum H werden wir ...

... $y' \neq x'$ der Punkte, die auf derselben Kurve liegen, übergeht.

Satz 6. *Die Funktion* $B_\xi(\tau_1, \tau_2)$, *die der Funktion* $\xi(t)$ *der Klasse* $\mathfrak{A}$ *entspricht, kann in der Form*

$$B_\xi(\tau_1, \tau_2) = c\,[\,|\tau_1|^\gamma + |\tau_2|^\gamma - |\tau_1 - \tau_2|^\gamma\,]$$

115

AMERICAN SOCIETY OF CIVIL ENGINEERS
Founded November 5, 1852
TRANSACTIONS

1951

Paper No. 2447

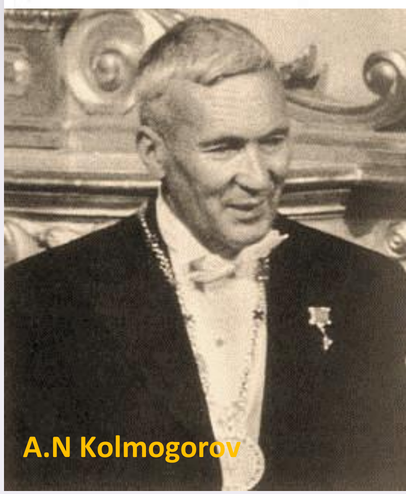
LONG-TERM STORAGE CAPACITY OF RESERVOIRS

By H. E. Hurst[1]

With Discussion by Ven Te Chow, Henri Milleret, Louis M. Laushey, and H. E. Hurst.

Synopsis

A solution of the problem of determining the reservoir storage required on a given stream, to guarantee a given draft, is presented in this paper. For example, if a long-time record of annual total discharges from the stream is available, the storage required to yield the average flow, each year, is obtained by computing the cumulative sums of the departures of the annual totals from ... The range from the maximum to the mini-... taken as the required storage.



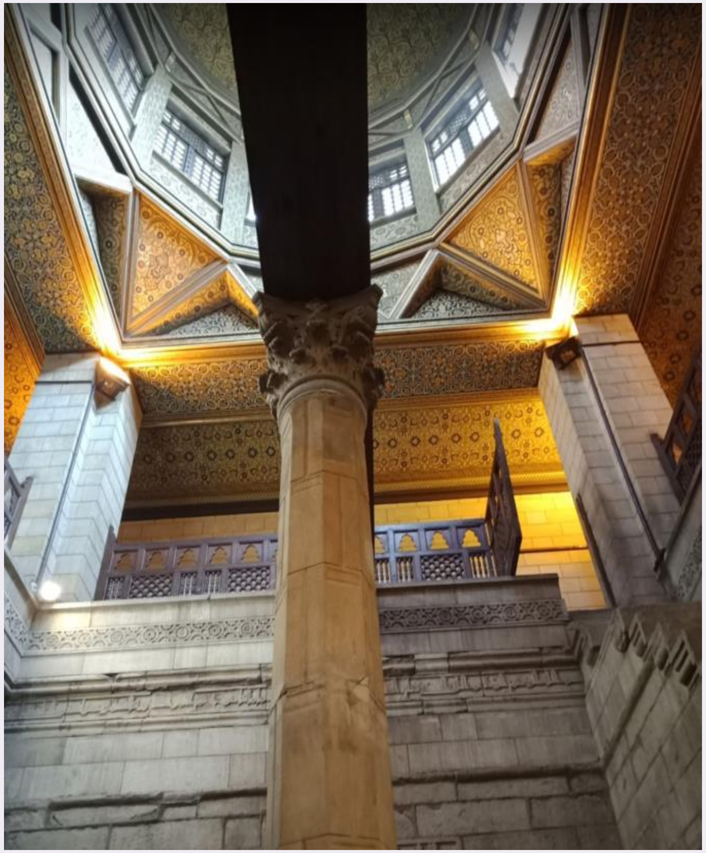
A.N Kolmogorov

Kolmogorov proposed a mathematical (stochastic) process that describes a behaviour unknown at that time. It was discovered a decade later in geophysics by Hurst.

H.E. Hurst
(Courtesy J. Sutcliffe, 2013)

"Although in random events groups of high or low values do occur, their tendency to occur in natural events is greater. This is the main difference between natural and random events."

The Roda Nilometer and the longest instrumental record on Earth



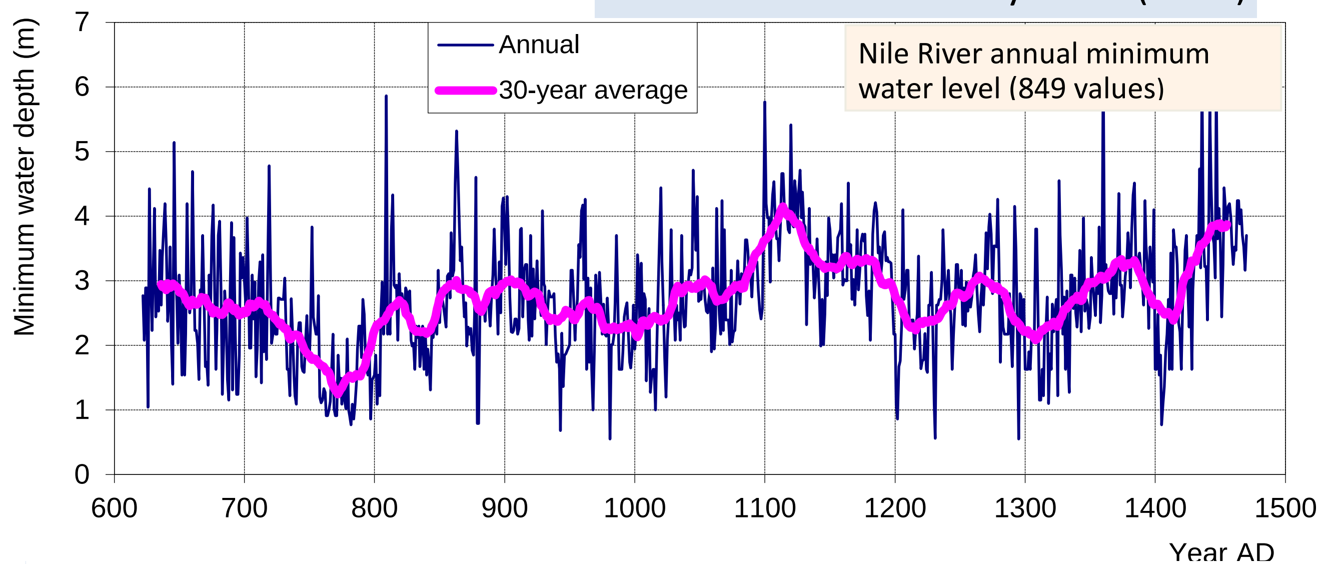
Photos by Loai Samen and Mohamd Mubarak; Google maps, https://goo.gl/maps/T8NUgoDAorK2 and https://goo.gl/maps/dsdJHJYVv572).

The Roda Nilometer, near Cairo. Water entered through three tunnels and filled the Nilometer chamber up to river level. The measurements were taken on the marble octagonal column (with a Corinthian crown) standing in the centre of the chamber; the column is graded and divided into 19 cubits (each slightly more than 0.5 m) and could measure floods up to about 9.2 m. A maximum level below the 16th mark could portend drought and famine and a level above the 19th mark meant catastrophic flood.

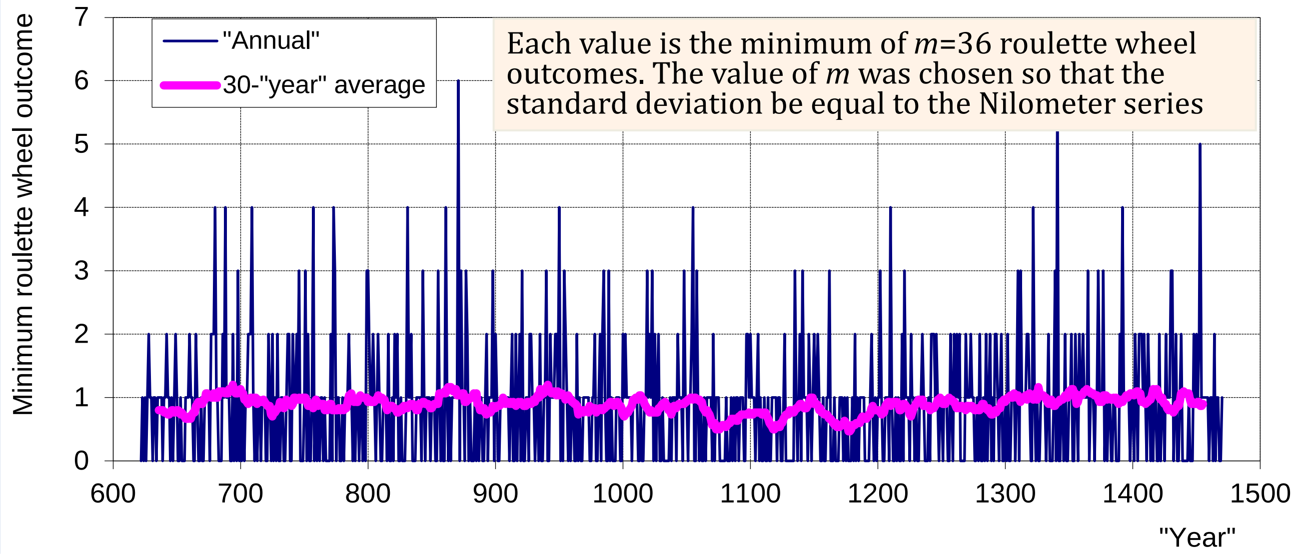
Stationary description of Earth's perpetual change: Hurst-Kolmogorov dynamics

Structured randomness

Nilometer data: Koutsoyiannis (2013)



Pure randomness



The climacogram: A simple statistical tool to quantify change across time scales

- Take the Nilometer time series, $x_1, x_2, ..., x_{849}$, and calculate the sample estimate of variance $\gamma(1)$, where the argument (1) indicates time scale (1 year)
- Form a time series at time scale 2 (years):

$$x_1^{(2)} := \frac{x_1 + x_2}{2}, x_2^{(2)} := \frac{x_3 + x_4}{2}, \dots, x_{424}^{(2)} := \frac{x_{847} + x_{848}}{2}$$

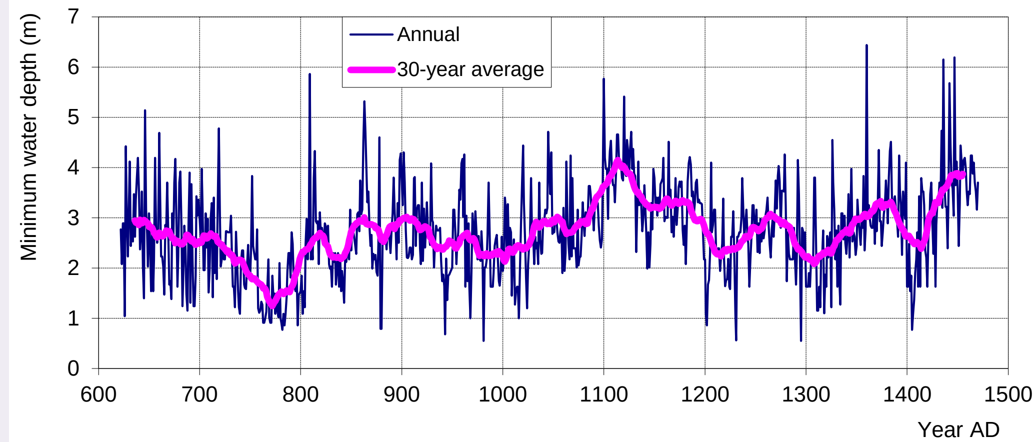
  and calculate the sample estimate of the variance $\gamma(2)$.
- Repeat the same procedure and form a time series at time scale 3, 4, … (years), up to scale 84 (1/10 of the record length) and calculate the variances $\gamma(3), \gamma(4), \dots \gamma(84)$.
- The **climacogram** is the variance $\gamma(\kappa)$ as a function of scale κ; it is visualized as a double logarithmic plot of $\gamma(\kappa)$ vs. κ.
- If the time series x_τ represented a pure random process, the climacogram would be a straight line with slope –1 (the proof is very easy).
- In real world processes, the slope is different from –1, designated as $2H - 2$, where H is the so-called Hurst coefficient ($0 < H < 1$).
- The scaling law $\gamma(\kappa) = \gamma(1) / \kappa^{2-2H}$ defines the **Hurst-Kolmogorov (HK) process**.
- High values of H (> 0.5) indicate **enhanced change** at large scales, else known as (long-term) **persistence**, or strong **clustering** (grouping) of similar values.

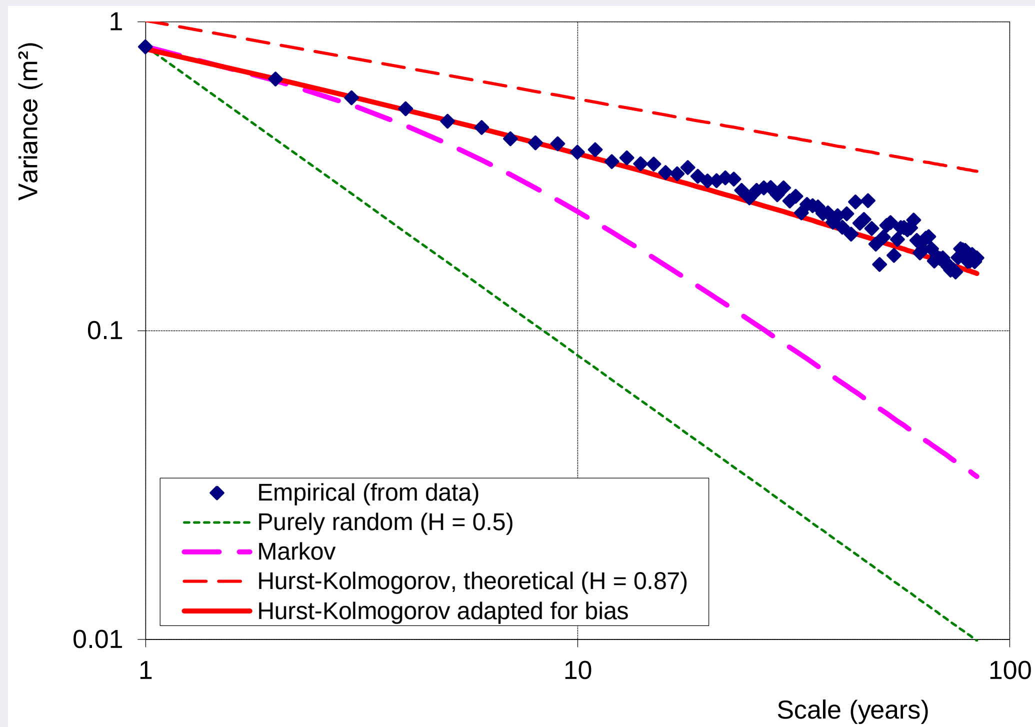
The climacogram of the Nilometer time series

- The Hurst-Kolmogorov process seems consistent with reality.
- The Hurst coefficient is $H = 0.87$ (Similar H values are estimated from the simultaneous record of maximum water levels and from the modern, 131-year, flow record of the Nile flows at Aswan).
- The Hurst-Kolmogorov behaviour, seen in the climacogram, indicates that:

    (a) long-term changes are more frequent and intense than commonly perceived, and

    (b) future states are much more uncertain and unpredictable on long time horizons than implied by pure randomness.





Effect of persistence on K-moment estimates

A K-moment is a characteristic of the marginal, first order, distribution of a process and therefore it is not affected by the dependence structure. However, its estimator is: time dependence induces bias to estimators of K-moments. Thus, the unbiasedness ceases to hold in stochastic processes.

For a **Markov** process the effect of autocorrelation is **negligible**, unless n is low and r high (e.g. > 0.90).

However, for an **HK process**, as shown in Koutsoyiannis (2020), the effect can be **substantial**:

$$\Theta(n, H) = \frac{K_p^{\mathrm{d}} - K_p}{K_p} \approx \frac{2H(1-H)}{n-1} - \frac{1}{2(n-1)^{2-2H}}$$

$$K_{p'} = K_p^{\mathrm{d}} = (1+\Theta)K_p, p' \approx 2\Theta + (1-2\Theta)p^{\left((1+\Theta)^2\right)}$$

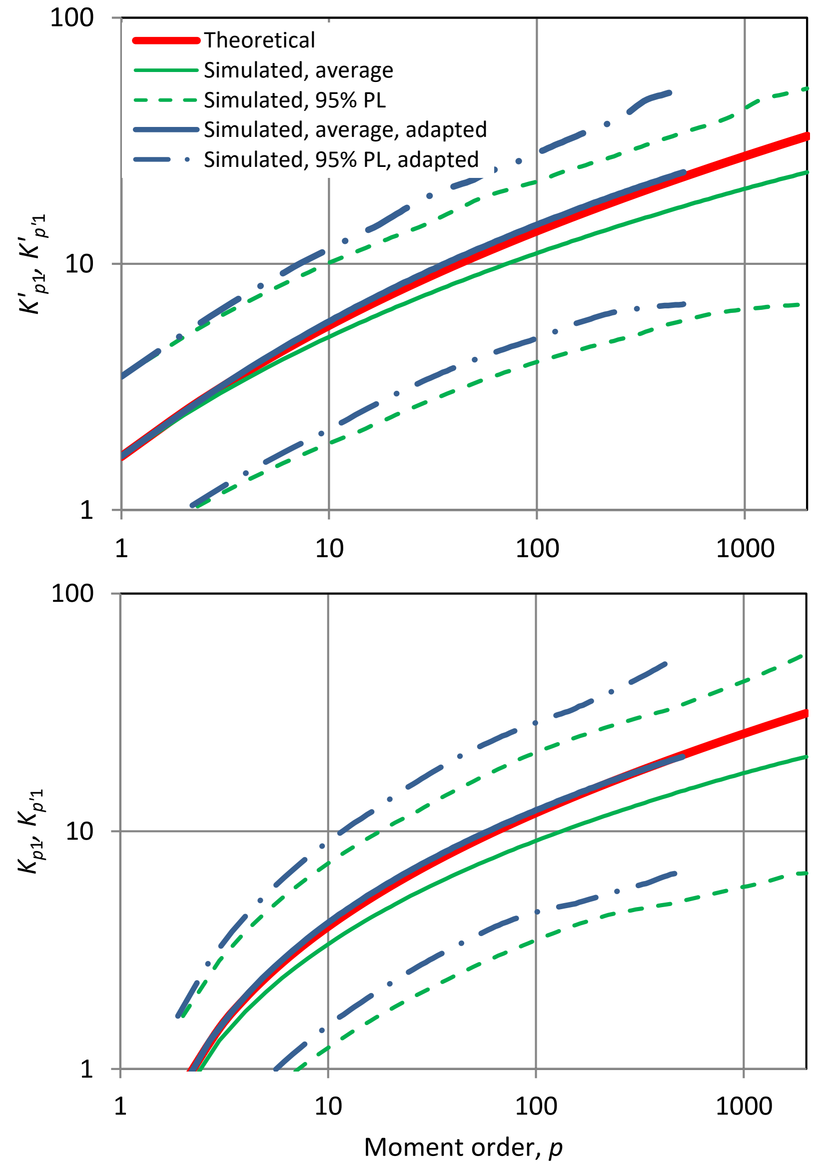


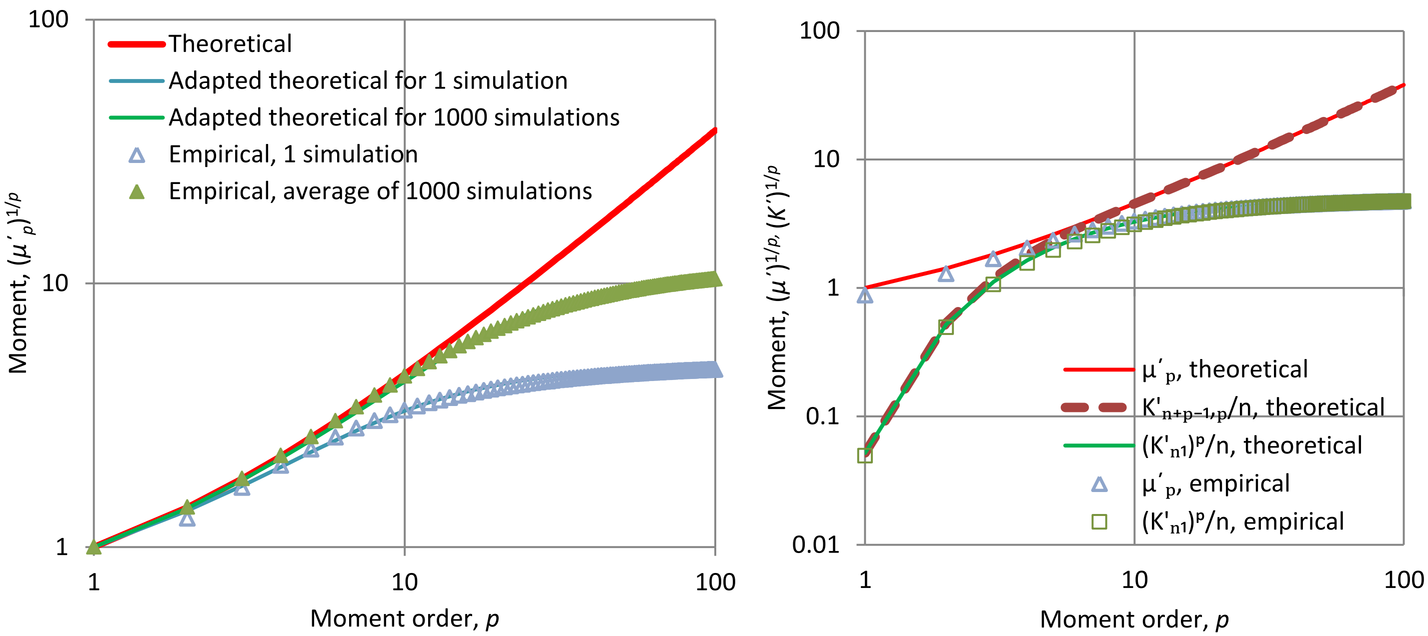
Illustration of the performance of the adaptation of K-moment estimation for an HK process with Hurst parameter 0.9 and lognormal marginal distribution (LN(0,1)). Shown are noncentral and central moments, for $q = 1$. The estimates are averages of 200 simulations each with $n = 2000$ and are almost indistinguishable from the theoretical values. The 95% prediction limits (PL) are also shown. The maximum $p = 2000$ reduces $p' \approx 500$, i.e. to one fourth, with an analogous reduction to the return period.

K-moments and classical moments

Not only are K-moments knowable but they can also predict the value that a classical moment estimator will give, through equation:

$$\hat{\mu}'_p \approx \frac{\left(K'_{mn,1}\right)^p}{K'_{mn+p-1,p}} \mu'_p$$

where n is the sample size and m is the number of samples for the case where more than one sample are available to make the estimate.



(Left) Comparison of the estimates of classical noncentral moments from 1 and 1000 independent samples from the exponential distribution to (a) the theoretical moments and (b) to the values determined by the above equation (adapted theoretical). (Right) Additional information of the terms appearing in the above equation.

K-moments and L-moments

Hypercentral K-moments are virtually equivalent to L-moments for small orders. In addition, the framework of K-moments provides alternative options to define summary statistical characteristics of the distribution, including the classical ones, as in the table below. (Which option is preferable depends on the statistical behaviour, and in particular, the mean, mode and variance, of the estimator.)

Characteristic	Order p	Option 1	Option 2	Option 3*
Location	1	$K'_{11} = \mu$ (the classical mean)		
Variability	2	$K^+_{21} = 2K_{21} = 2(K'_{21} - \mu) = 2\lambda_2$	$K^+_{22} = K_{22} = \mu_2 = \sigma^2$ (the classical variance)	
Skewness (dimensionless)	3	$\dfrac{K^+_{31}}{K^+_{21}} = 2\dfrac{K_{31}}{K_{21}} - 3 = \dfrac{\lambda_3}{\lambda_2}$	$\dfrac{K^+_{32}}{K^+_{22}} = 2\dfrac{K_{32}}{K_{22}} - 2$	$\dfrac{K_{33}}{K_{22}^{3/2}} = \dfrac{\mu_3}{\sigma^3}$
Kurtosis (dimensionless)	4	$\dfrac{K^+_{41}}{K^+_{21}} = 4\dfrac{K_{41}}{K_{21}} - 8\dfrac{K_{31}}{K_{21}} + 6 = \dfrac{4}{5}\dfrac{\lambda_4}{\lambda_2} + \dfrac{6}{5}$	$\dfrac{K^+_{42}}{K^+_{22}} = 4\dfrac{K_{42}}{K_{22}} - 6\dfrac{K_{32}}{K_{22}} + 3$	$\dfrac{K_{44}}{K_{22}^{2}} = \dfrac{\mu_4}{\sigma^4}$

However, the real power of K-Moments is in their determination and use for very high orders p, up to the sample size.

Assigning return periods to K-moments of any order

- The non-central K-moment for q = 1 is $K'_p = p\mathrm{E}\left[\left(F(\underline{x})\right)^{p-1}\underline{x}\right]$
- By definition, it represents the expected value of the maximum of p copies of $\underline{x}$.
- 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:

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

- Any symmetric distribution will give exactly Λ_1= 2 because K'_1 is the mean, which equals the median and thus has a return period of 2D. Thus, a rough 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:

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

- The exact value of Λ_∞ depends only on the tail index ξ of the distribution:

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

  where γ = 0.577 is the Euler's constant.
- These enable the simple approximation of Λ_p and hence of the return period:

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

Better approximation of the relationship between *p* and *T*

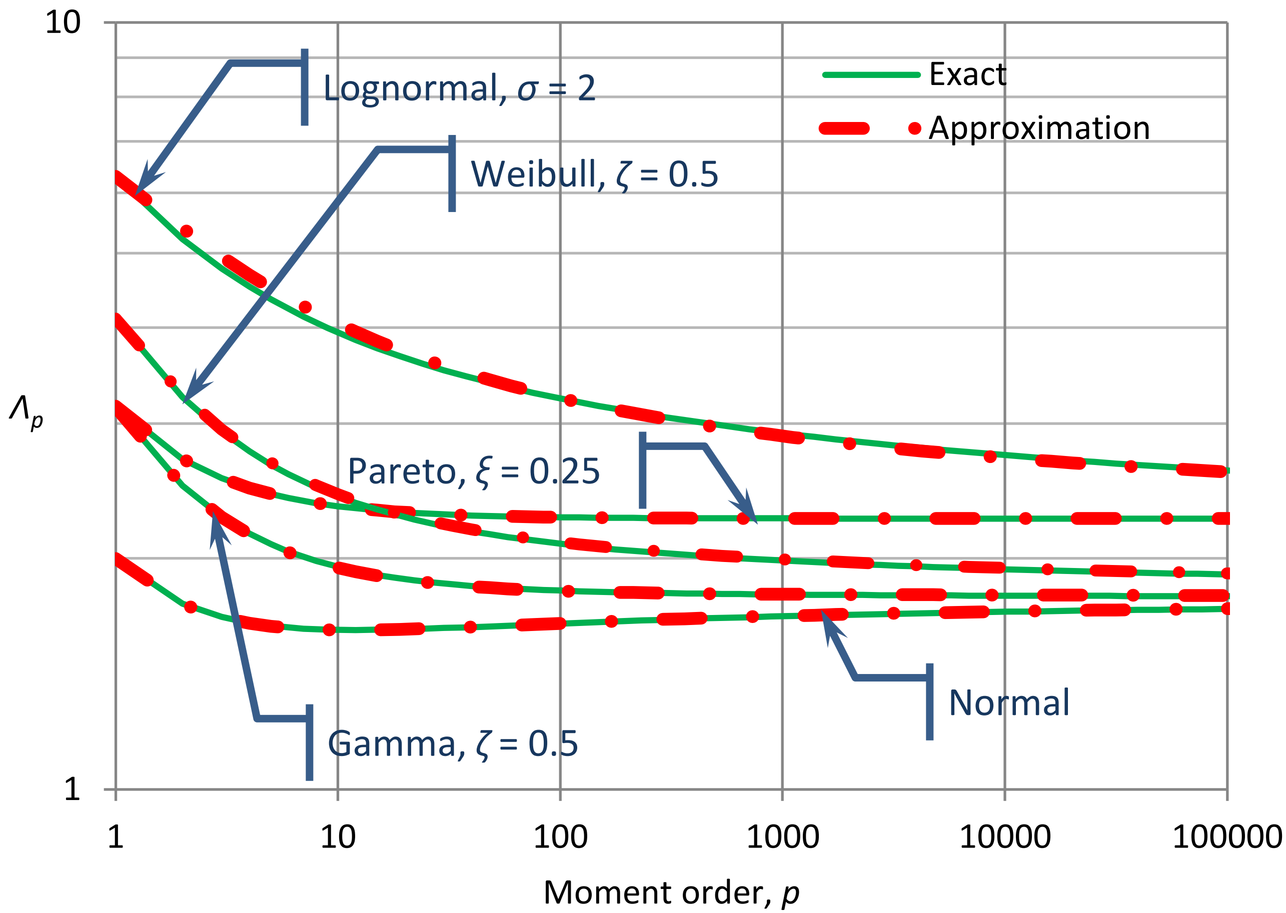
- An almost perfect approximation of Λ_p for any distribution function is:

$$\Lambda_p \approx \Lambda_\infty + \left(\Lambda_1 - \Lambda_\infty - B\ln\left(1 + \frac{\beta}{2^\beta - 1}\right)\right)\frac{1}{p} + B\ln\left(1 + \frac{\beta}{(p+1)^\beta - 1}\right)$$

- This involves two constants β and B, which depend on the distribution function.
- For example, in the Pareto distribution, $\beta = 1$ and

$$B = \frac{(3-\xi)\Lambda_\infty - 2\Lambda_1}{2(1 - \ln 2)}$$

- For parameters of other distributions see Koutsoyiannis (2020).
- The graph indicates the perfect agreement of the approximation to the exact values for several distributions.



Assigning return periods to order statistics (plotting positions)

The classical formula for assigning a return period to $x_{(i:n)}$, i.e., the ith smallest value in a sample of size n is:

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

where A and B are constants. For an unbiased estimate of the distribution quantile these constants are $A = 1/\Lambda_\infty$, $B = \Lambda_1/\Lambda_\infty - 1$ (Koutsoyiannis, 2020) and thus:

$$\frac{T_{(i:n)}}{D} = \frac{\Lambda_\infty(n-1) + \Lambda_1}{\Lambda_\infty(n-i) + 1}$$

For the highest value $x_{(n)} \equiv x_{(n:n)}$ both approaches, K-moments and order statistics, result in precisely the same value, $T(K'_n)/D = T_{(n:n)}/D = \Lambda_\infty p + (\Lambda_1 - \Lambda_\infty)$.

For smaller i, the pth K-moment should be equivalent, in terms of the corresponding return period, if

$$p = \frac{n - (\Lambda_1 - \Lambda_\infty)(n-i)}{\Lambda_\infty(n-i) + 1}$$

This means that:

(a) $x_{(i:n)}$ can be used as a quick-and-dirty (QAD) estimate of K'_p, provided that p is given as a function of i from the above equation.

(b) The return period estimate based on the typical estimator $\underline{\widehat{K}'_p} = \sum_{i=1}^{n} b_{inp}\, \underline{x}_{(i:n)}$ is better than that based on a single $x_{(i:n)}$ because it is derived from many data points (except for the maximum value, when $i = n$, where the two approaches are precisely identical).

Comparison of return periods assigned by the two approaches

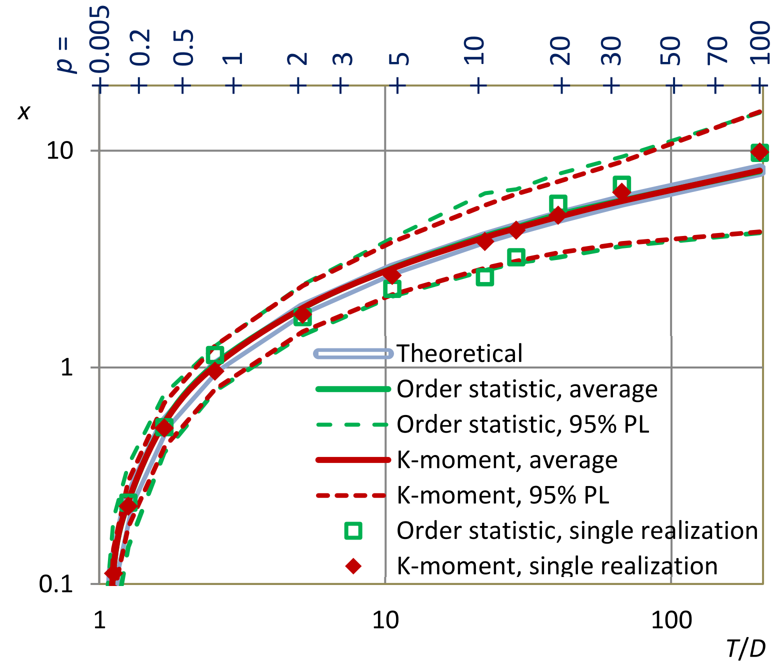
We assume a Pareto distribution with $\xi = 0.15$, in which $\Lambda_\infty = 2.035$ and $\Lambda_1 = 2.955$, and a sample of n = 100. We will have the following values of T and p for the highest and the second highest values:

i	$T_{(i:n)}$	p
100	204.4	100
99	67.4	32.6

Hence the estimate of the x quantile corresponding to a return period of 67.4, will be equal to:

- $x_{(99:100)}$ according to the order-statistics approach (the estimate is based on one data point);
- $\hat{K}'_{32.6}$ according to the K-moments approach (the estimate is based on 68 data points and will be a weighted average of $x_{(i:100)}$ for $i = 33$ to $i = 100$).

Simulation results of empirical return periods assigned to Pareto quantiles (for tail index ξ = 0.15, scale parameter λ = 1 and lower bound zero). Averages and prediction limits (PL) were calculated from 200 simulations each with n = 100. The curves of averages for both the order statistics and the K-moment approaches are indistinguishable from the theoretical curves. The return periods were assigned for the unbiased quantile option. The correspondence between the K-moment of order p and the return period T is also shown through the upper horizontal axis. The plots of a single realization are also shown (but for part of the empirical points to avoid an overcrowded graph).



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

From the relationship:

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

we can easily find the K-moment order p corresponding to the return period T. An example is given in the following table:

Example of the K-moment order p corresponding to the specified return period for the Pareto distribution with shape parameter ξ = 0.15.			
	D = 10 min	D= 1 h	D = 1 d
T = 2 months	4 307	717	29
T = 1 year	25 842	4 307	179
T = 2 years	51 684	8 614	358
T = 100 years	2 584 212	430 702	17 945

In a stochastic process with dependence, what is given in the table is the adapted moment order p' while p should be estimated from p' based on the relationship in p. 22.

Alternatively, the more accurate approximation of p. 26 or even the exact relationships could be used but the resulting p will not differ substantially from the above values.

K-moments of such high order are reliably estimated.

Application 1: Monitoring and characterization of the hydroclimatic evolution

A third parenthesis: What is climate?

Aristotle in his **Meteorologica** describes the climates on Earth in connection with latitude but he does not use the term climate. Instead, he uses the term ***crasis*** (***κρᾶσις***, literally meaning mixing, blending of things which form a compound, temperament).

The term ***climate*** (***κλίμα***, plural ***κλίματα***) was coined as a geographical term by the astronomer Hipparchus (190 –120 BC; founder of trigonometry but most famous for his discovery of precession of the equinoxes by averaging measurements on several stars). It originates from the verb *κλίνειν*, meaning 'to incline' and denoted the angle of inclination of the celestial sphere and the terrestrial latitude characterized by this angle.

A modern definition by the American Meteorological Society is*:

*Climate – The slowly varying aspects of the atmosphere–hydrosphere–land surface system. It is typically characterized in terms of **suitable averages of the climate system over periods of a month or more**, taking into consideration the **variability in time of these averaged quantities**.*

In turn, the **climate system** is defined as:

The system, consisting of the atmosphere, hydrosphere, lithosphere, and biosphere, determining the earth's climate as the result of mutual interactions and responses to external influences (forcing).

In this presentation the term ***hydroclimatic*** refers to

multiscale stochastic characterization of meteorological and hydrological processes

As the averages are not sufficient in monitoring and understanding climate, the high-order K-moments offer a convenient means to characterize its variability.

---

* http://glossary.ametsoc.org/wiki/Climate

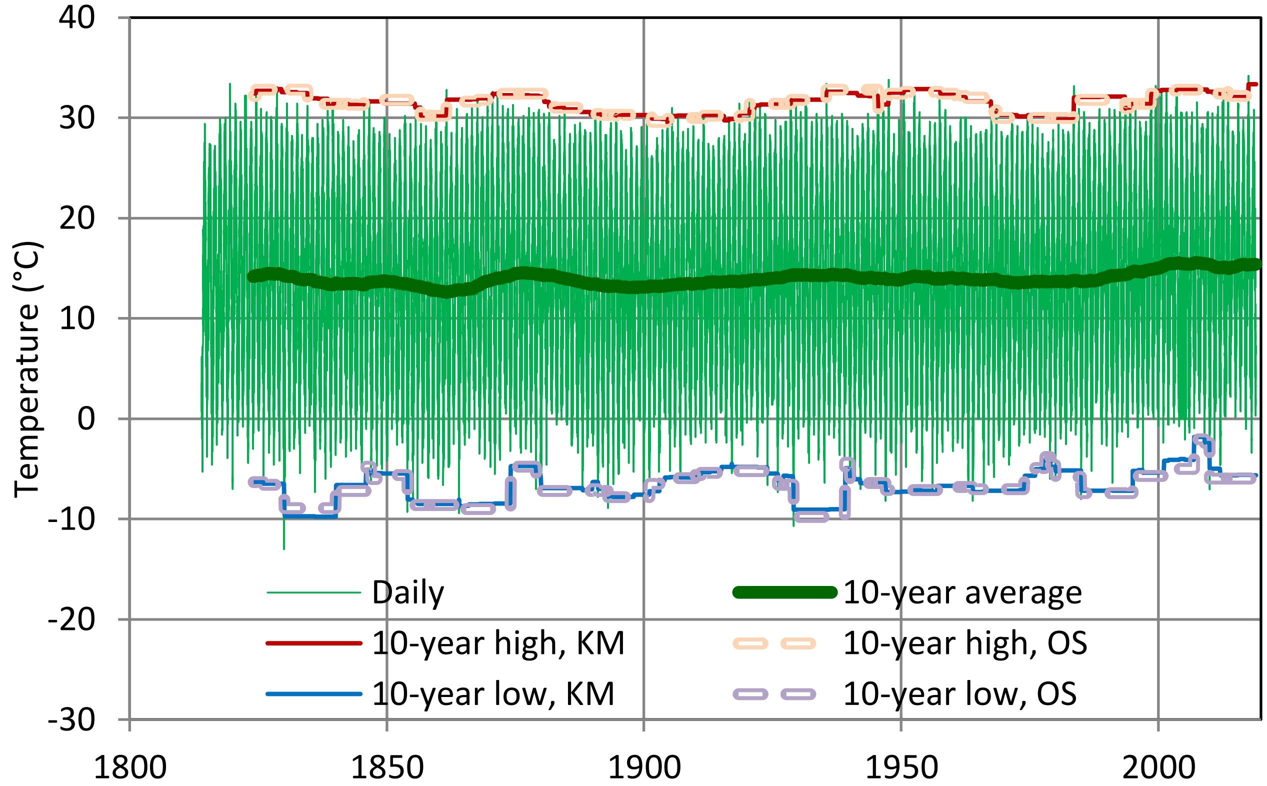
Data for illustration 1: Daily temperature in Bologna

Bologna, Italy (44.50°N, 11.35°E, +53.0 m).

Average daily temperature, available online in the frame of the European Climate Assessment & Dataset.*

Uninterrupted for the period 1814-2003, 190 years in total.

For the most recent period, 2004-2018, daily data are provided by the online data repository Dext3r.† With the additional data, the **record length becomes 205 years**.



---

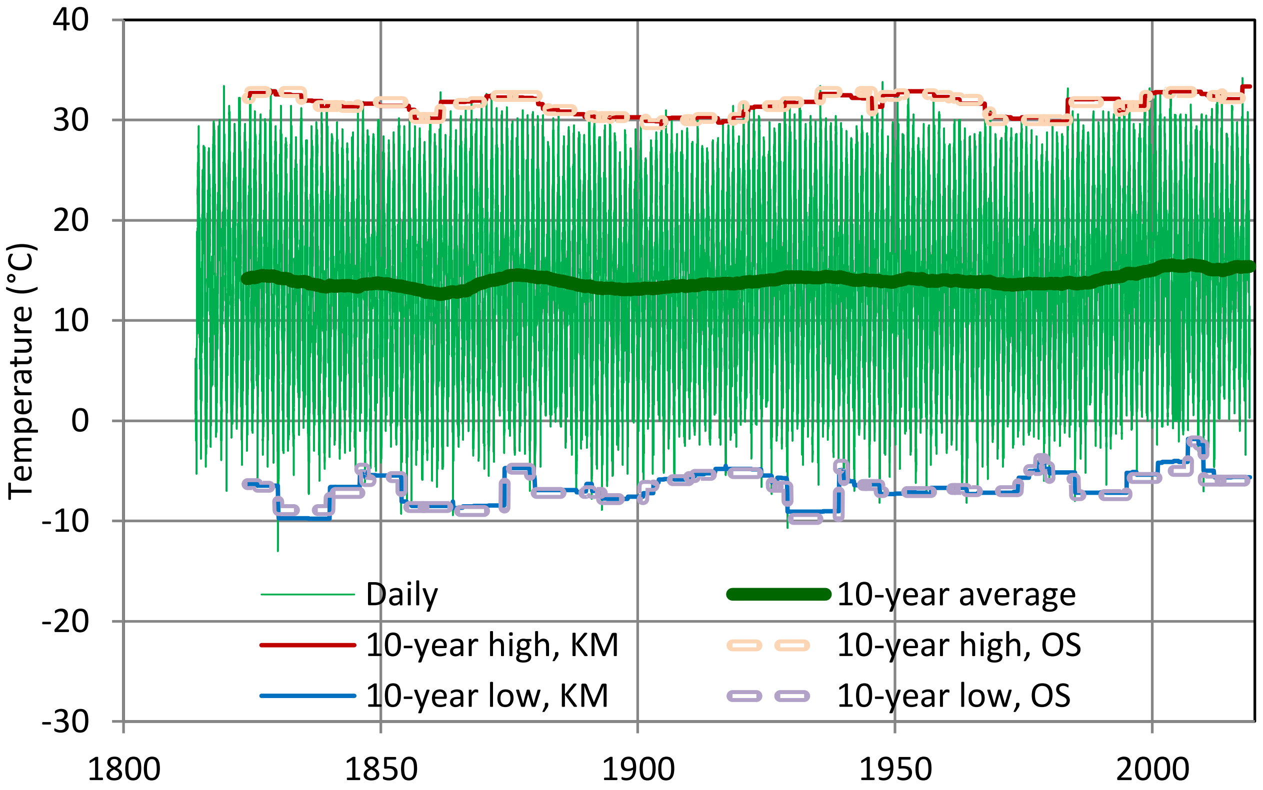
* ECAD; Klein Tank et al., 2002; data retrieved on 2019-02-17 from https://climexp.knmi.nl/ecatemp.cgi?WMO=169.

† http://www.smr.arpa.emr.it/dext3r/. In particular, the average daily temperature values of the station Bologna Urbana (44.500754°N, 11.328789°E, +78.0 m) were used. The data at Bologna Urbana were adjusted by adding a constant temperature difference of 0.19 °C to become consistent with those of the ECAD station. To find this adjustment, as there is no common period of observation between the ECAD station and Bologna Urbana, a third station whose observations have common periods with both, namely the Bologna Meteo station (44.501223°N, 11.328197°E, +80.0 m) was used.

Climate monitoring in Bologna: A diagnostic analysis of the evolution of temperature climatic extremes

In addition to the plot of daily values, the graph contains the climatic evolution in terms of 10-year averages and extremes.

For the extremes, the 10-year high and low values are calculated using a sliding window of 10 years. They correspond to the second highest and second lowest value out of 3652 daily values in the 10-year period, estimated both with the K-moments (KM) and the order statistics (OS) approaches (see next page about the value of p and the corresponding return period).



Main observation: There exist upward and downward fluctuations. The climatic range, measured as the difference of the high and low extremes, had its highest value 42.3 °C in 1860s (arguably, the worst conditions), deceased to 33.8 °C in 1978, and increased to 39.0 °C in 2018.

Assessment of return periods of each of high and low values

Assuming normal distribution, in which $\Lambda_\infty = 1.781, \Lambda_1 = 2$, the return period of the second highest value in ten years (3652 days) is:

$$\frac{T_{(3651:3652)}}{1\text{ d}} = \frac{\Lambda_\infty(3652-1)+\Lambda_1}{\Lambda_\infty(3652-3651)+1} = 2339 \;\Rightarrow T = 6.4\text{ years}$$

This corresponds to K-moment order:

$$p = \frac{n-(\Lambda_1-\Lambda_\infty)(n-i)}{\Lambda_\infty(n-i)+1} = 1313$$

However, the time series suggests Hurst-Kolmogorov behaviour, with Hurst parameter $H = 0.94$ for the annual average (even excluding the 21st century data).

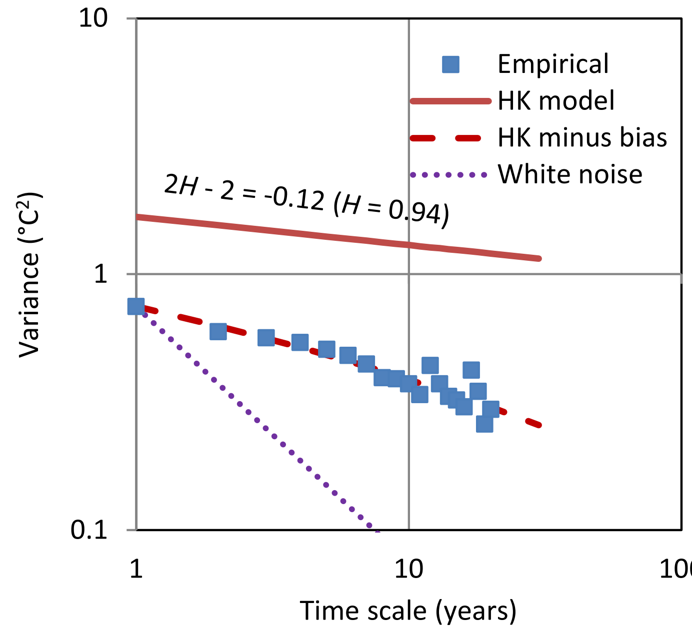
For large n the equation on p. 22 can be written as:

$$\Theta^{\mathrm{HK}}(n,H) \approx -\frac{1}{2n^{2-2H}} = -\frac{1}{2n'} = -\frac{1}{2}\frac{\gamma_n}{\gamma_1}$$

where n' is the effective sample size; for the Bologna daily temperature and n = 3652 (the number of days for 10 years), $n' = 73.5/1.3 = 56$ and $\Theta = -0.018$.

In turn, this results in $p' \approx 2\Theta + (1-2\Theta)p^{\left((1+\Theta)^2\right)} = 1176$ and $T \approx \Lambda_\infty p + (\Lambda_1 - \Lambda_\infty) = 2095\text{ d} = 5.7\text{ years}$.

This is a rough approximation as the climacograms are of HK type only asymptotically (for large k); stochastic simulation is required for a better approximation.



Climacogram of annual average temperature in Bologna.

Data for illustration 2: Daily precipitation in Bologna

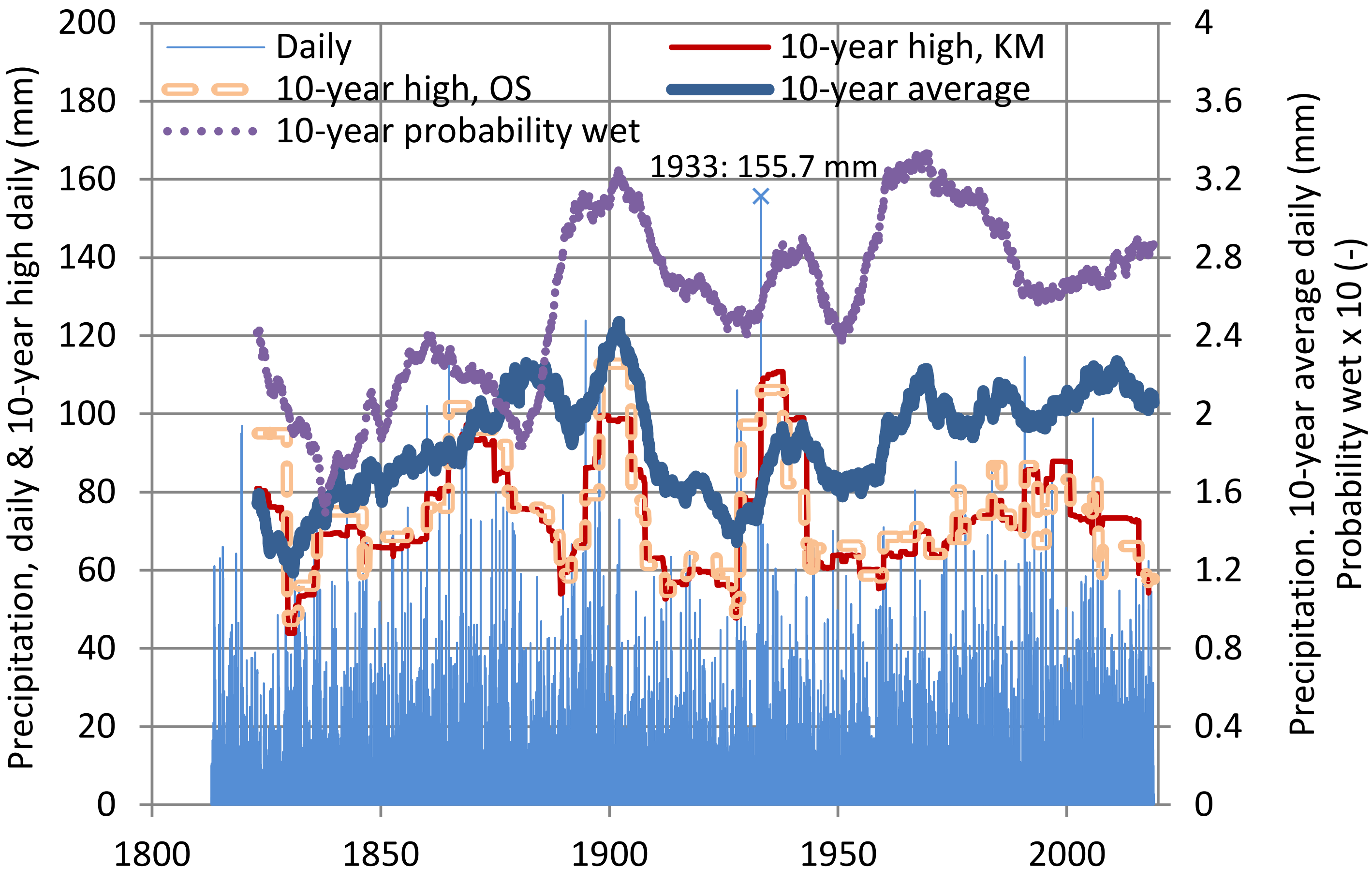
Bologna, Italy (44.50°N, 11.35°E, +53.0 m).

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.



Climate monitoring in Bologna: A diagnostic analysis of the evolution of precipitation climatic extremes

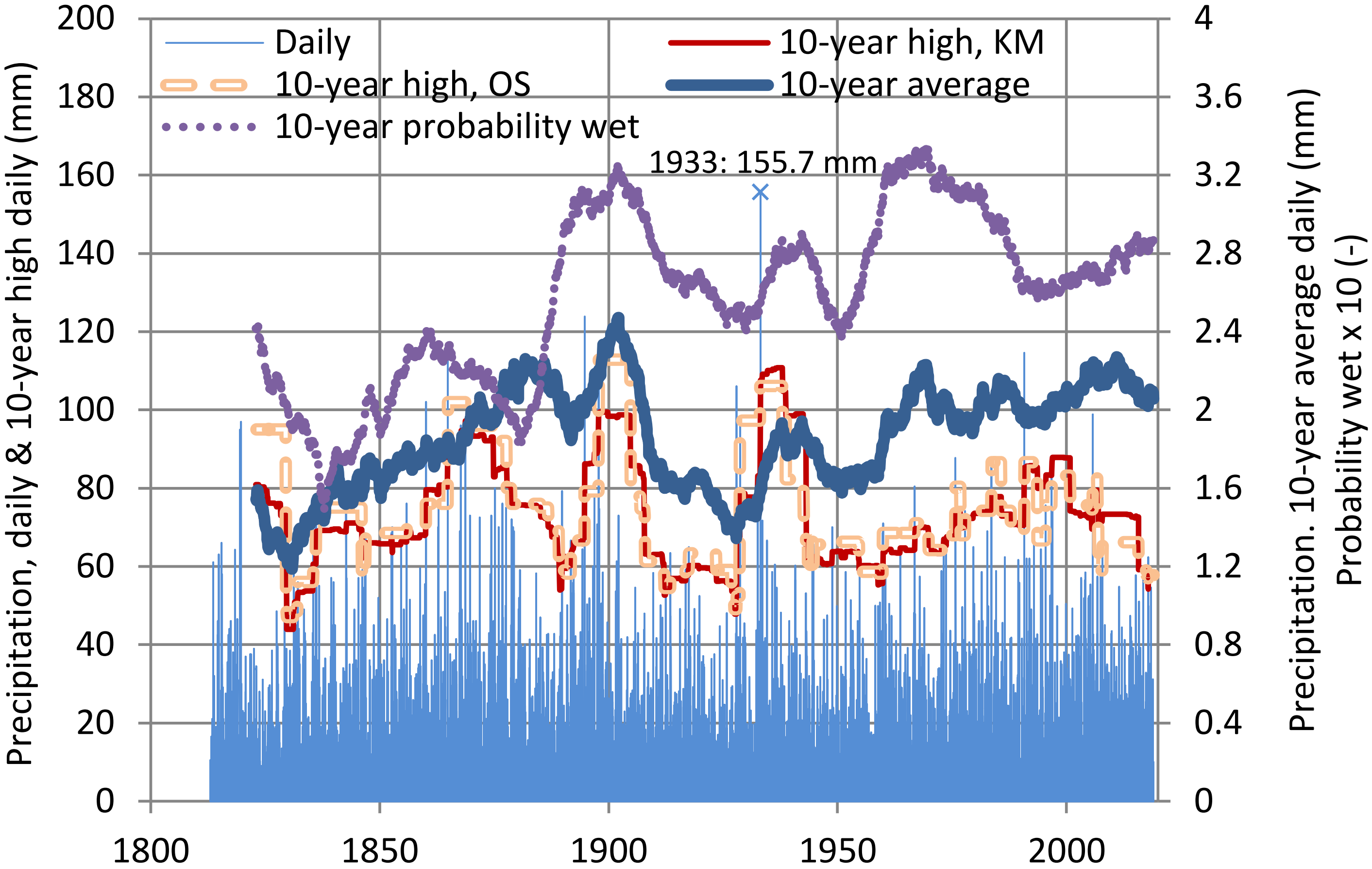
Here instead of the minimum climatic value (= zero) we plot the probability that a day is wet.

Main observations:

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);



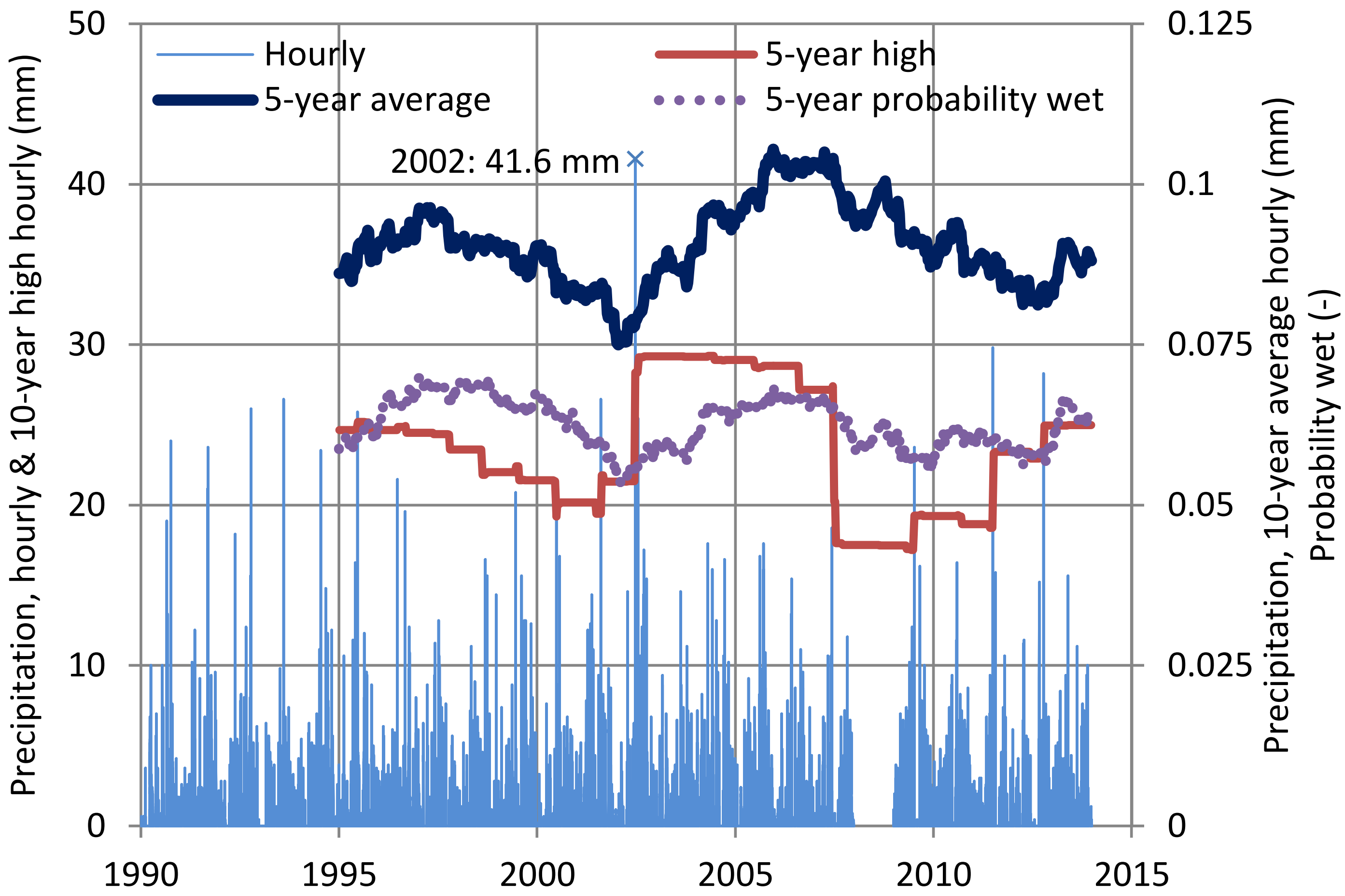
The high daily precipitation by a factor 2.5 (from 44 to 110 mm/d).

Data for illustration 3: Hourly precipitation in Bologna

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.



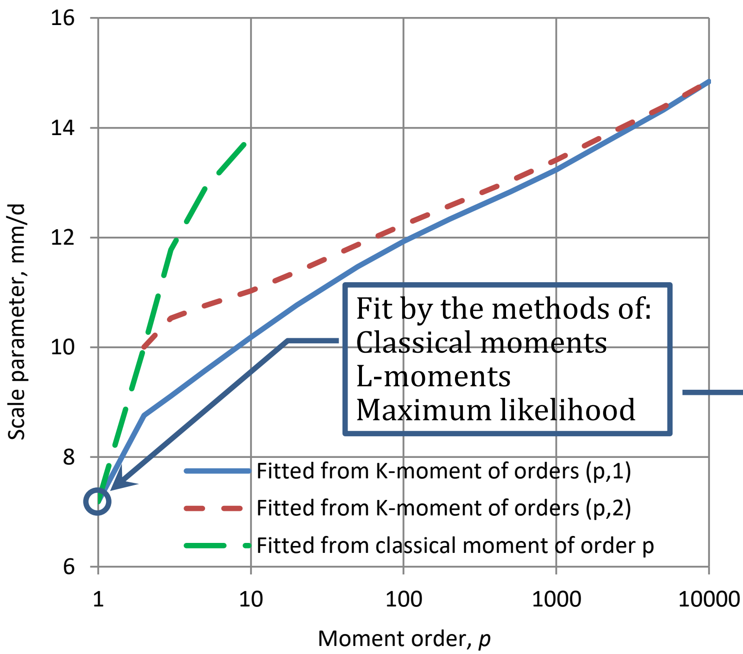
Application 2: Fitting of probabilistic models

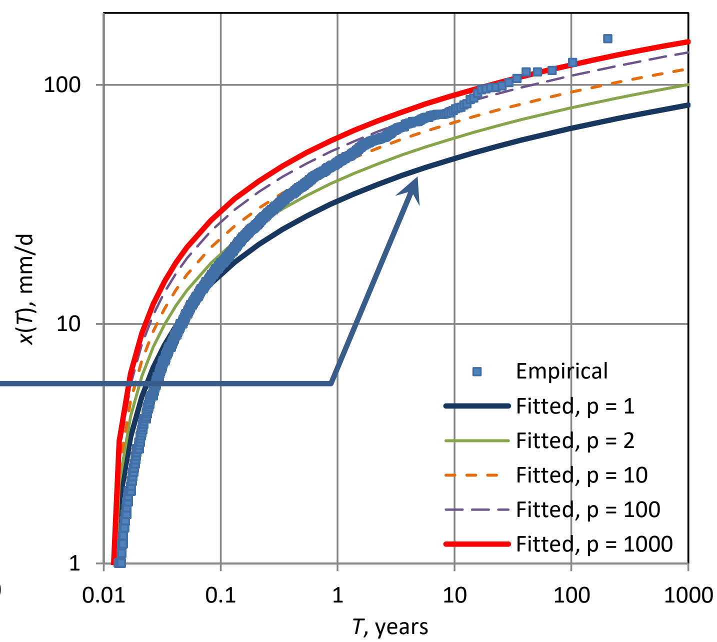
Illustration that high-order K-moments are preferable to low-order moments

For the sake of illustration, for the daily rainfall in Bologna, we intentionally choose the simplest and blatantly unsuitable model, the 1-parameter exponential distribution, $F(x) = 1 - \mathrm{e}^{-x/\lambda}$.

One moment suffices to estimate the single (scale) parameter λ—but which moment to choose?

The exact K-moments are: $K_{p1} = (H_p - 1)\lambda$, $K_{p2} = \left(\left(H_{p-1} - 1\right)^2 + H_{p-1}^{(2)}\right)\lambda^2$, $K_{pp} = \mu_p = (!\,p)\lambda^p$, where H_p is the pth harmonic number and $H_p^{(2)}$ is the pth harmonic number of order 2.



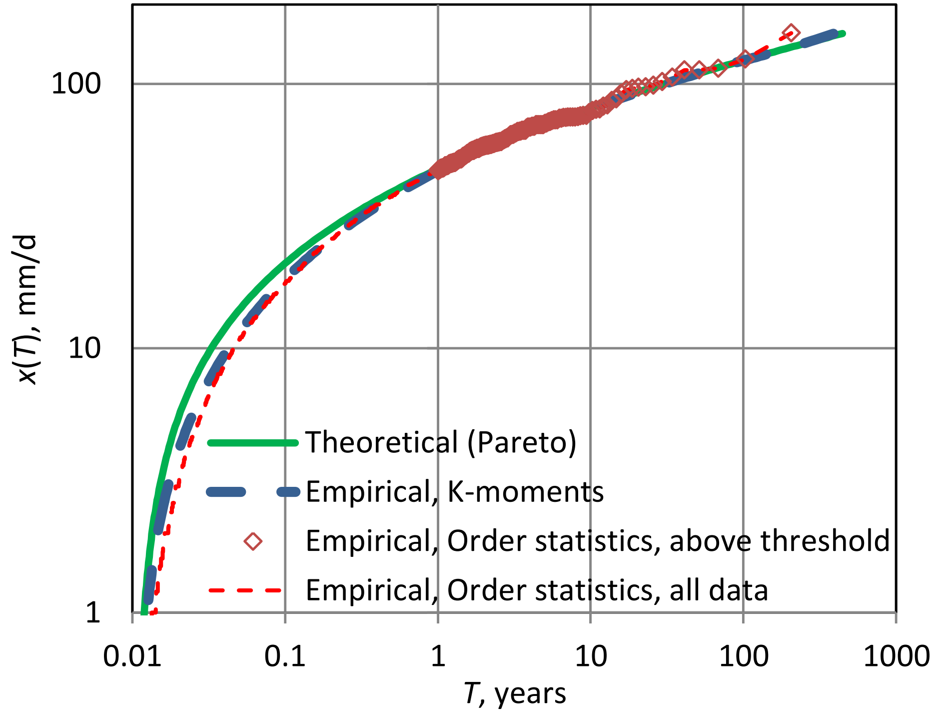


The moment order p affects the fitting dramatically.

The scale parameter λ increases with increasing p, q.

If we wish to model maxima, it is better to fit based on the 1000th K-moment than on the 1st!

Better fitting on K-moments for orders p from ~100 to 10 000 (T = ~2 to 200 years)



We assume Pareto distribution with zero lower bound (for physical consistency):

$$F(x) = 1 - (1 + \xi\, x/\lambda)^{-\frac{1}{\xi}} \text{ or}$$

$$\frac{T(x)}{D} = (1 + \xi\, x/\lambda)^{\frac{1}{\xi}}$$

The exact relationship of K-moments with return period is:

$$\frac{\hat{T}(\hat{K}_p')}{D} = p\Lambda_p = \left((p + 1 - \xi)\, \mathrm{B}(1 - \xi, p + 1)\right)^{\frac{1}{\xi}}$$

We estimate the parameters by minimizing the mean square error of the logarithms of the empirical $\hat{T}(\hat{K}_p')$ from the theoretical $T(\hat{K}_p')$. We calculate the error for a range of T from 2 to 200 years. We utilize as many data as possible (cf. Volpi et al., 2019: "Save hydrological observations"). The fitted parameters are $\xi = 0.096$, $\lambda = 8.37$ mm/d.

The graph shows a perfect fit of theoretical and empirical curves for $T > 1$ year (the two curves are indistinguishable).

For comparison, empirical curves for order statistics are also plotted (Weibull plotting positions).

Note: Minimizing the error of $\hat{K}_p'$ with respect to K_p', without reference to T, is another possibility but presupposes exact relationships for K_p', which in other distributions may be infeasible to derive.

Slight improvement for a global fitting

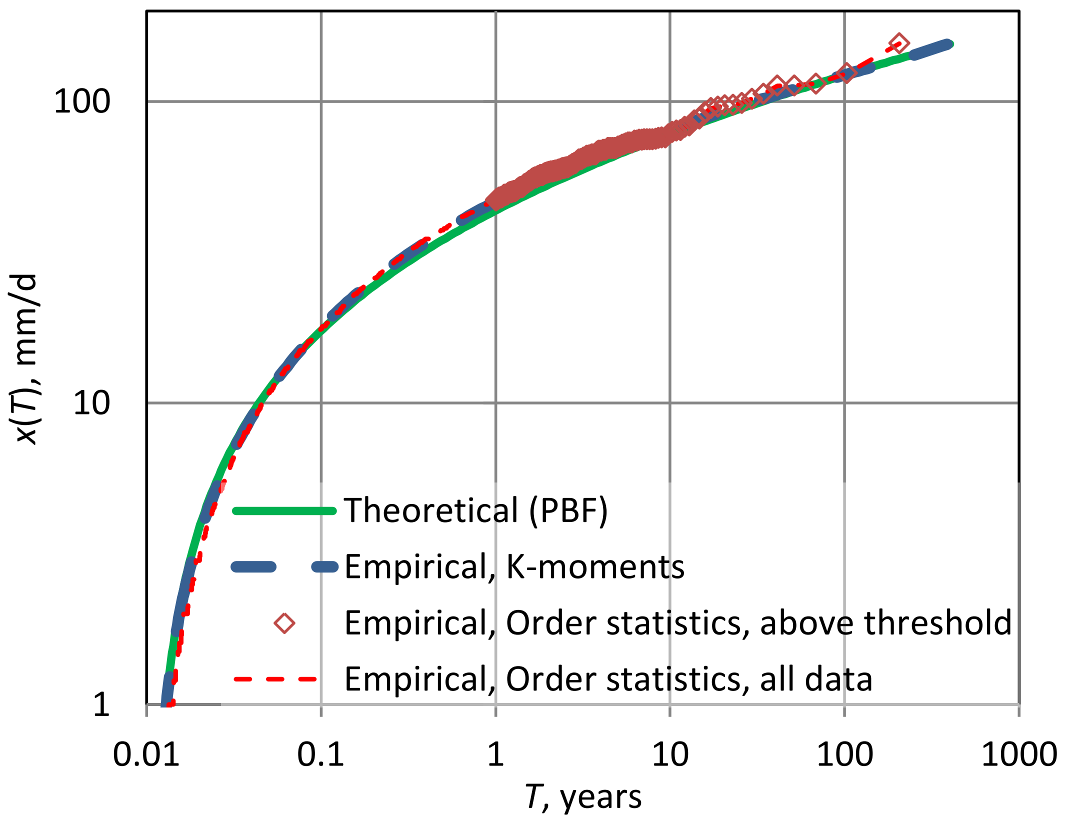
By adding one parameter to the theoretical distribution function we can get a model applicable for the entire range of rainfall depth.

Namely, we use the Pareto-Burr-Fuller (PBF) distribution with zero lower bound (for physical consistency for rainfall):

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

We use the same estimation procedure as above but calculate the error on the entire range of values.

The estimated parameters are: ξ = 0.096, ζ= 0.883, λ = 5.04 mm/d.



A perfect fit of the model (green continuous line) and empirical curve (blue dashed line) is seen for the entire range.

For comparison, empirical curves for order statistics are also plotted (Weibull plotting positions) but not used at any step.

Application 3:
Full stochastic modelling of rainfall
(construction of an ombrian model)

Ombrian model: Marginal distribution of rainfall intensity

An ombrian model (from the Greek ombros, meaning rainfall) describes the stochastic properties of the distribution of rainfall of any order, or equivalently, at any time scale.

From an ombrian model that is simple enough, the ombrian relationships, also known with the misnomer rainfall intensity – duration [meaning time scale] – frequency [meaning return period] curves are directly extracted. The assumptions of the proposed ombrian model follow.

1. Pareto distribution with discontinuity at the origin for small time scales:

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

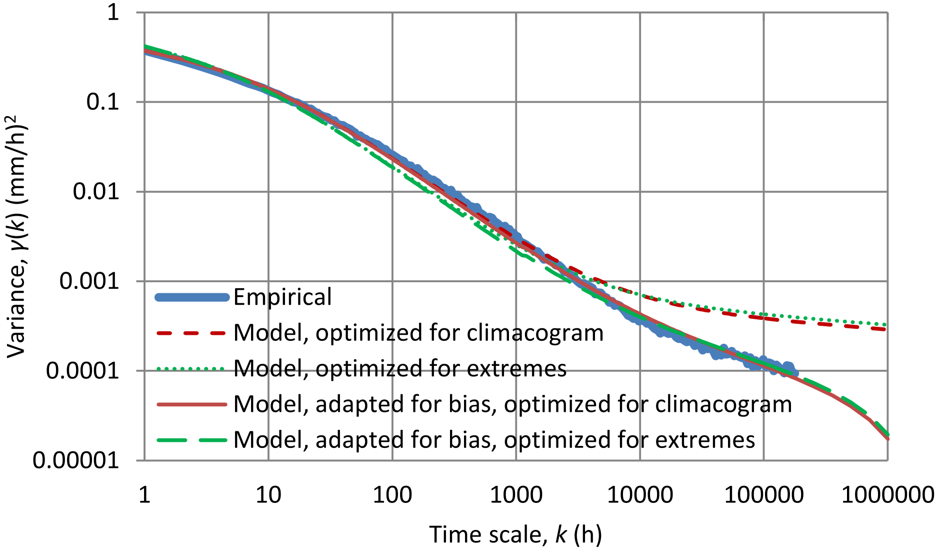
It is shown by theoretical reasoning (Koutsoyiannis, 2020) that the tail index ξ should be constant, while the probability wet, $P_1^{(k)}$, and the state scale parameter, $\lambda(k)$, are functions of the time scale k. Here we sacrifice the exactness of the PBF distribution (see previous page) in order to get simpler ombrian relationships for small scales.

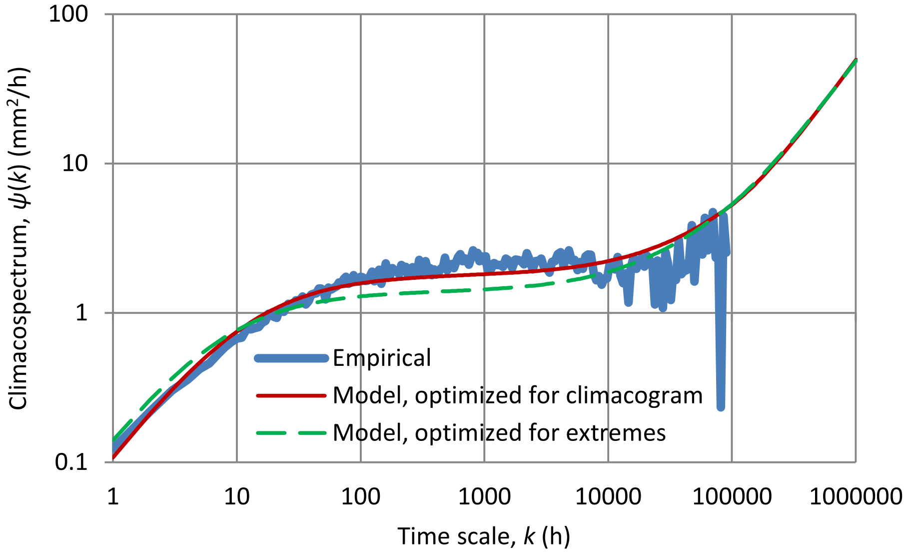
2. Continuous PBF distribution with possible discontinuity at zero for large time scales, i.e.:

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

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.:

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

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:

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

Clearly, in both cases, $\gamma(k) \to 0$, as $k \to \infty$, which makes the process ergodic; for $k = 0$, $\gamma(0) = \gamma_0 = \lambda$ in the FHK-C case and $\gamma(0) = \gamma_0 = \lambda_1 + \lambda_2$ 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

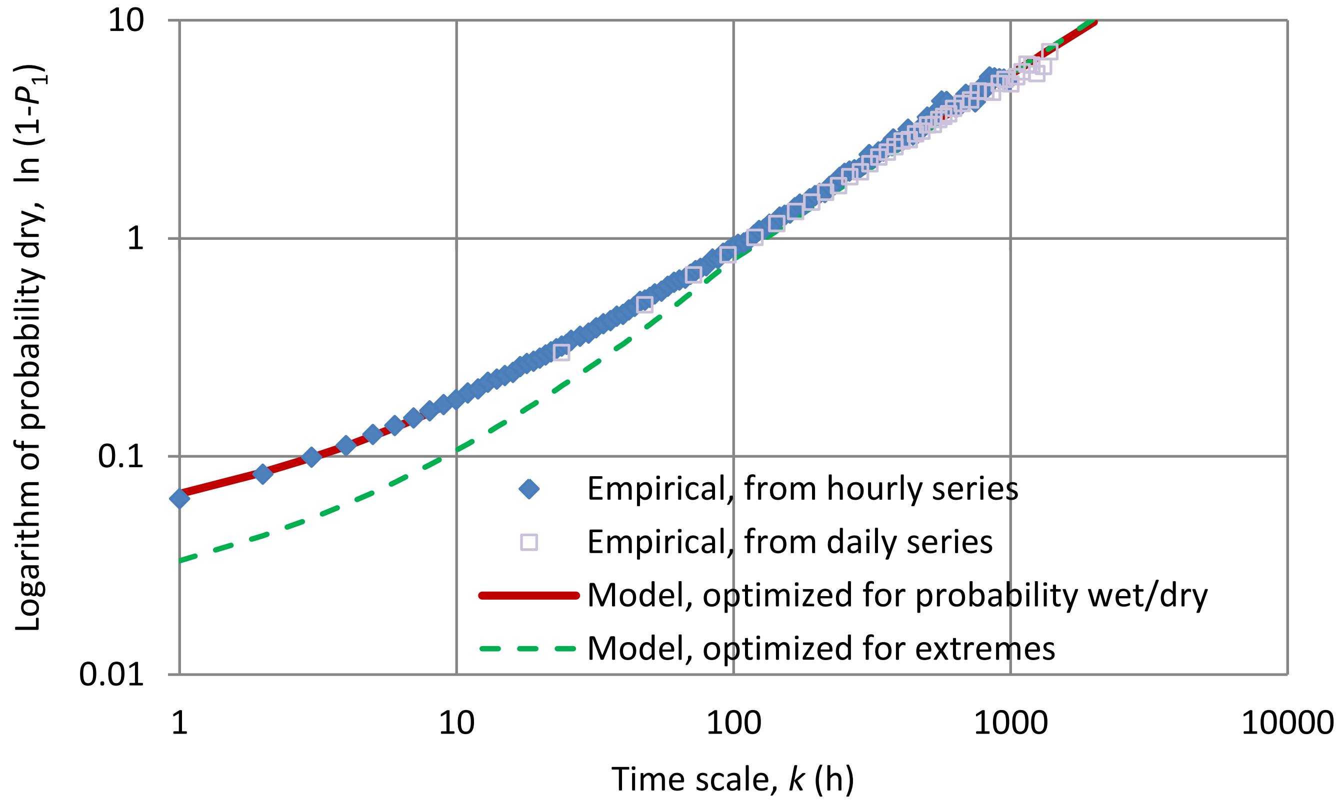
5. Probability wet and dry, $P_1^{(k)} = 1 - P_0^{(k)}$, varying with time scale according to:

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

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:

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



Fitting of the ombrian model to the empirical estimates of probability wet (P_1) or dry ($1 - P_1$).

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	
$\gamma(k)$	$\lambda_1(1+(k/\alpha)^{2M})^{\frac{H-1}{M}}$ or $\lambda_1\left(1+\frac{k}{\alpha}\right)^{2H-2} + \lambda_2\left(1-\left(1+\frac{\alpha}{k}\right)^{2H-2}\right)$	
$P_1^{(k)}$	$\frac{1-\xi}{1/2-\xi}\frac{\mu^2}{\gamma(k)+\mu^2}$	$1-\left(1-P_1^{(k^*)}\right)^{(k/k^*)^\theta}$
$\frac{1}{\zeta(k)}$	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

Parameter	Meaning of parameter	Related tool
μ	Mean intensity	Mean, μ
λ (or 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.

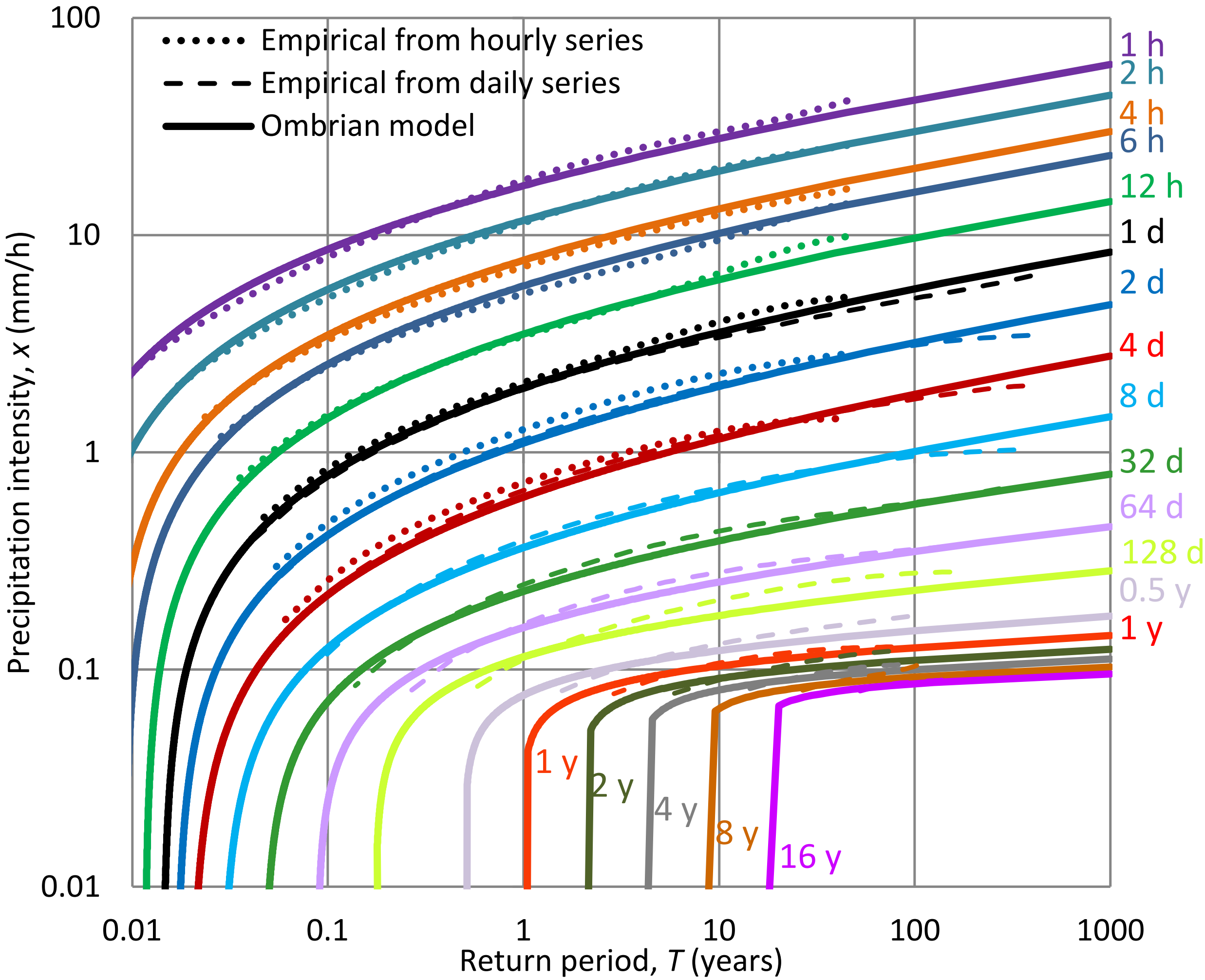
With these seven parameters, the ombrian model achieves:

(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;
(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.

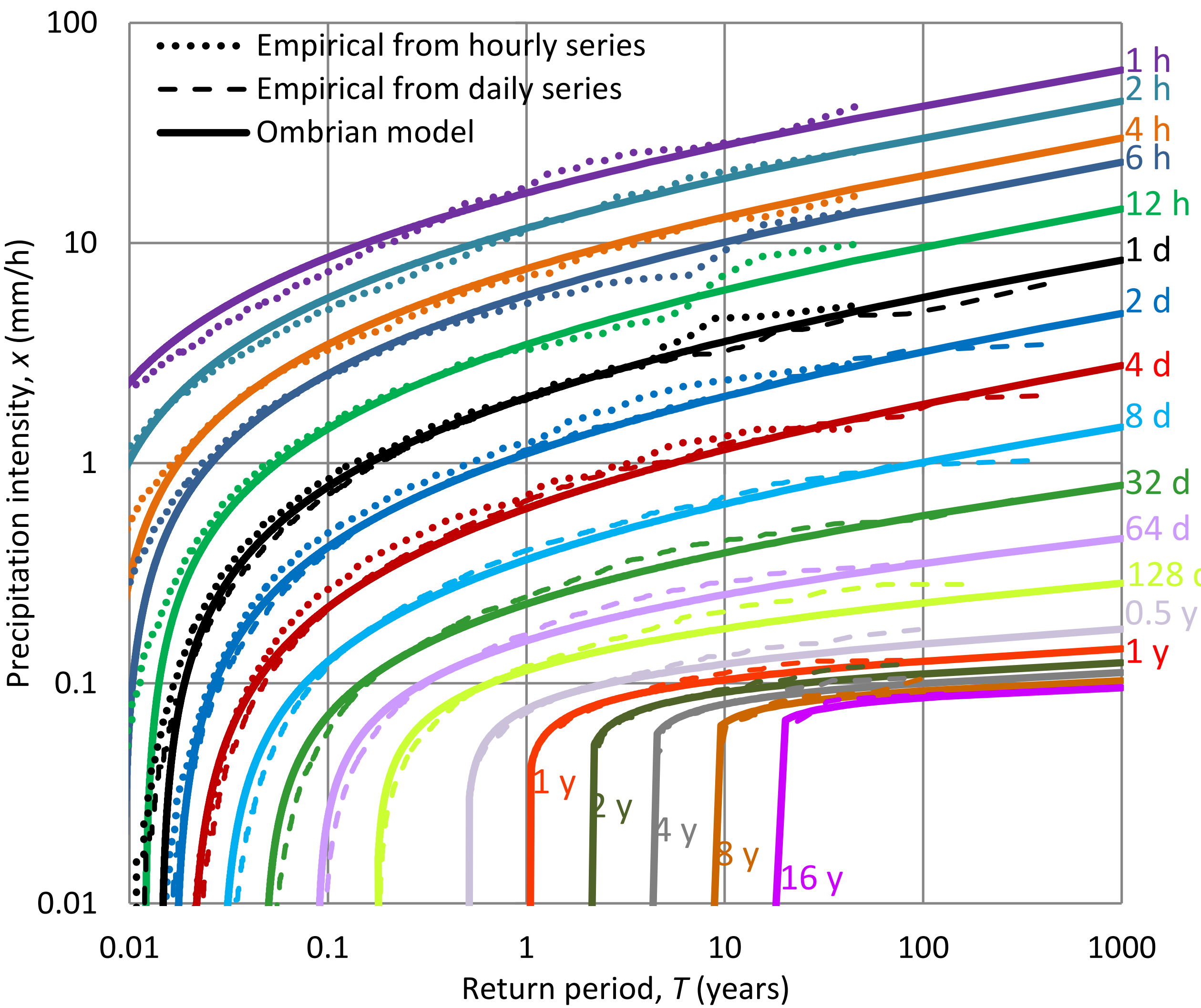
Parameter values	
μ	0.0773 mm/h
λ_1	0.00103 mm²/h²
λ_2	1.978 mm²/h²
α	9.704 mm
H	0.95
ξ	0.120
θ	0.849



Ombrian curves for Bologna: Comparison of model to order-statistics estimates of return period

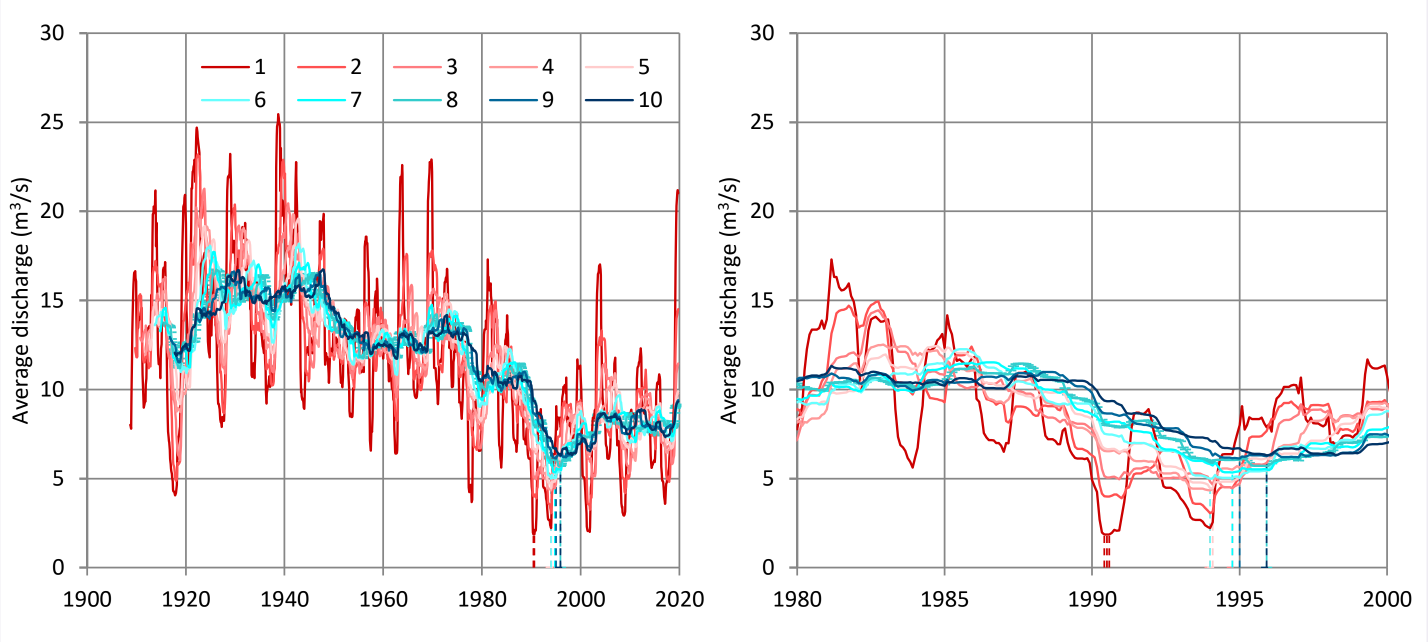
Ombrian curves as resulted from the ombrian model for Bologna for time scales spanning 5 orders of magnitude (1 h to 16 years = 140256 h). The empirical points are estimated from order statistics. The effect of persistence was taken into account; the model was plotted with bias-adapted variance in order to be comparable with empirical plots.

Parameter values as in previous page.



Application 4:
Stochastic (Monte Carlo) simulation

A fourth parenthesis: The Athens drought and the importance of decent stochastic simulation



During the 7-year period 1988-95, a severe drought hit Greece, including the Athens area.

The graph of the discharge of one of the rivers supplying Athens (Boeoticos Kephisos) shows a clustering of low flow values for time scales of 1 to 10 years (see legend), with the 10-year average becoming 1/2 to 1/3 of earlier decadal means.

A stochastic generation method respecting the Hurst-Kolmogorov dynamics was developed and used for the system management.

Despite the drought severity and duration, the management, assisted by stochastic simulation, was effective and we recovered from the crisis without even a day of failure of the water supply system.

The stochastic simulation method has ever been improving since then and what follows is the most recent development (Koutsoyiannis, 2019b).

Reminder of second-order stochastic tools

- **Climacogram**, $\gamma(k)$ for time scale k (variance of the averaged process as a function k):

$$\gamma(k) \coloneqq \operatorname{var}\left[\frac{\underline{X}(k)}{k}\right], \qquad \underline{X}(t) \coloneqq \int_0^t \underline{x}(\xi)\mathrm{d}\xi$$

- **Autocovariance function**, $c(h)$ for time lag h, defined as:

$$c(h) := \operatorname{cov}[x(t), x(t+h)] = \frac{1}{2}\frac{\mathrm{d}^2\left(h^2\gamma(h)\right)}{\mathrm{d}h^2}$$

- **Power spectrum** (also known as *spectral density*), $s(w)$ for frequency w, defined as the Fourier transform of the autocovariance function, i.e.:

$$s(w) \coloneqq 4\int_0^\infty c(h)\cos(2\pi wh)\,\mathrm{d}h\,, c(h) = \int_0^\infty s(w)\cos(2\pi wh)\,\mathrm{d}w\,, \gamma(k) = \int_0^\infty s(w)\ \operatorname{sinc}^2(\pi wk)\,\mathrm{d}w$$

- **Structure function** (also known as *semivariogram* or *variogram*),

$$v(h) \coloneqq \frac{1}{2}\operatorname{var}\left[\underline{x}(t) - \underline{x}(t+h)\right] = \gamma_0 - c(h)$$

- **Climacospectrum**, $\zeta(k)$, for time scale k:

$$\zeta(k) \coloneqq \frac{k\left(\gamma(k) - \gamma(2k)\right)}{\ln 2} = \frac{k\,\gamma_\mathrm{C}(k)}{\varepsilon \ln 2}$$

This resembles the power spectrum and combines the asymptotic behaviours of the climacogram and the structure function.

The generic generator

- Any stationary stochastic process $\underline{x}_\tau$ can be generated by the moving average scheme (Koutsoyiannis 2000):

$$\underline{x}_\tau = \sum_{j=-J}^{J} a_j \underline{v}_{\tau-j}$$

where a_j are weights to be calculated from the autocovariance function, $\underline{v}_j$ is white noise averaged in discrete-time and J is a large integer (theoretically, $J = \infty$).

- The autocovariance is given by the convolution expression:

$$c_\eta = \sum_{j=-J}^{J} a_j a_{\eta+j}$$

- Given the stochastic model, the weights a_η can be calculated from its second-order characteristics by the following explicit relationship (Koutsoyiannis, 2019b):

$$a_\eta = \int_{-1/2}^{1/2} \sqrt{2}\mathrm{e}^{2\pi\mathrm{i}(\theta(\omega)-\eta\omega)}\sqrt{s_\mathrm{d}(\omega)}\,\mathrm{d}\omega$$

where ω denotes frequency, $s_\mathrm{d}(\omega)$ is the power spectrum of the discrete-time representation of the process (see below) and $\theta(\omega)$ is any arbitrary odd real function.

☺ Notice the appearance of 2, $\pm 1/2$, $\sqrt{2}$, e, π, i (imaginary unit) in last equation.

The generic generator (2)

- The equations:

$$\underline{x}_\tau = \sum_{j=-J}^{J} a_j \underline{v}_{\tau-j}, \quad a_\eta = \int_{-1/2}^{1/2} \sqrt{2}\mathrm{e}^{2\pi\mathrm{i}(\theta(\omega)-\eta\omega)}\sqrt{s_\mathrm{d}(\omega)}\,\mathrm{d}\omega$$

  define the **asymmetric moving average (AMA)** scheme, which can be used in any problem of stochastic simulation of time irreversible and reversible processes.
- The sequence of a_η given by the above equation:
  - consists of real numbers, despite the expression involving complex numbers;
  - reproduces precisely the required autocovariance function; and
  - is easy and fast to calculate using the fast Fourier transform (FFT).
- The above equation gives not a single solution, but a variety of infinitely many ones, all of which preserve exactly the second-order characteristics of the process.
  - A particular solution is characterized by the chosen function $\theta(\omega)$.
  - Even assuming $\theta(\omega) = \theta_0 =$ constant, again there are infinitely many solutions.
  - This enables preservation of additional statistics, e.g. those related to time asymmetry.
- In addition, we always have several options related to the distribution of the white noise $\underline{v}_\tau$ (which in general **is not Gaussian**), thus enabling preservation of moments of any order (Koutsoyiannis, 2019a; Dimitriadis and Koutsoyiannis, 2018).
- In particular, moments of order > 2, are dealt with by preserving the **cumulants** $\underline{v}_\tau$ ($\kappa_p^{(v)}$), which are related to those of $\underline{x}_\tau$ (κ_p), by $\kappa_p = \sum_{l=-J}^{J} a_j^p\ \kappa_p^{(v)}$. Cumulants are determined by **theoretical calculations** on the distribution function fitted on the basis of **empirical K-moments**.

Computational algorithm

1. From the continuous-time stochastic model, expressed through its climacogram $\gamma(k)$, we calculate its autocovariance function in discrete time (assuming time step D):

$$c_j = \frac{(j+1)^2\gamma(|j+1|D)+(j-1)^2\gamma\big((|j-1|D\big)}{2} - j^2\gamma(|j|D)$$

(This step is obviously omitted if the model is already expressed in discrete time through its autocovariance function.)

2. We choose an appropriate number of coefficients J that is a power of 2 and use the inverse FFT to calculate the discrete-time power spectrum and the frequency function $A^{\mathrm{R}}(\omega)$ for an array of $\omega_j = j\, w_1, j = 0,1, \dots, J, w_1 \coloneqq 1/JD$:

$$s_{\mathrm{d}}(\omega_j) = 2c_0 + 4\sum_{\eta=1}^{J} c_\eta \cos(2\pi\eta\omega_j), \quad A^{\mathrm{R}}(\omega_j) = \sqrt{2s_{\mathrm{d}}(\omega_j)}$$

3. We choose $\theta(\omega)$ and form the arrays (vectors) $\boldsymbol{A}^{\mathrm{R}}$ and $\boldsymbol{A}^{\mathrm{I}}$, both of size $2J$ indexed as $0, \dots, 2J-1$, with the superscripts R and I standing for a real and an imaginary vector, respectively:

$$[\boldsymbol{A}^{\mathrm{R}}]_j = \begin{cases} A^{\mathrm{R}}(\omega_j)\cos\big(2\pi\theta(\omega_j)\big)/2, & j = 0, \dots, J \\ [\boldsymbol{A}^{\mathrm{R}}]_{2J-j}, & j = J+1, \dots, 2J-1 \end{cases}$$

$$[\boldsymbol{A}^{\mathrm{I}}]_j = \begin{cases} -A^{\mathrm{R}}(\omega_j)\sin\big(2\pi\theta(\omega_j)\big)/2, & j = 0, \dots, J-1 \\ 0 & j = J \\ -[\boldsymbol{A}^{\mathrm{I}}]_{2J-j}. & j = J+1, \dots, 2J-1 \end{cases}$$

4. We perform FFT on vectors $\boldsymbol{A}^{\mathrm{R}}$ and $\boldsymbol{A}^{\mathrm{I}}$, and get the real part of the result for $j = 0, \dots, J$, which is precisely the sequence of a_η.

Conclusions

- In a **stationary and ergodic** framework, the newly introduced knowable moments (K-moments) are powerful tools that unify other statistical moments (classical, L-, probability weighted) and order statistics, offering several advantages.
- In particular, they offer a sound basis for distribution fitting with emphasis on extremes, as well as for climate monitoring, again with emphasis on extremes.
- For **independent** identically distributed variables, 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.
- **Stochastic simulation** of non-Gaussian processes is also assisted by the K-moments, through their unbiased estimators, from which the high-order properties are assessed and then preserved, by combining K-moments with cumulants.

References

Birkhoff, G. D., 1931. Proof of the ergodic theorem. *Proc. Nat. Acad. Sci.*, 17, 656–660.

Boltzmann, L., 1877. Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht. *Wiener Ber.*, 76, 373–435 (in German).

Boltzmann, L., 1884/85. Über die Eigenschaften monocyklischer und anderer damit verwandter Systeme. *Crelles J.*, 98, 68–94.

Dimitriadis, P., and Koutsoyiannis, D., 2018. Stochastic synthesis approximating any process dependence and distribution, *Stochastic Environmental Research & Risk Assessment*, 32 (6), 1493–1515, doi:10.1007/s00477-018-1540-2.

Hurst, H.E., 1951. Long term storage capacities of reservoirs. *Trans. Am. Soc. Civil Engrs.*, 116, 776–808.

Khintchine, A., 1933. Zu Birkhoffs Lösung des Ergodenproblems. *Math. Ann.*, 107, 485-488.

Khintchine, A., 1934. Korrelationstheorie der stationären stochastischen Prozesse. *Mathematische Annalen*, 109 (1), 604–615.

Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A., et al., 2002. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *Intern. J. Climatol.*, 22, 1441-1453, doi:10.1002/joc.773.

Kolmogorov, A.N., 1931. Uber die analytischen Methoden in der Wahrscheinlichkcitsrechnung, *Mathematische Annalen*, 104, 415-458. (English translation: On analytical methods in probability theory, In: Kolmogorov, A.N., 1992. *Selected Works of A. N. Kolmogorov - Volume 2, Probability Theory and Mathematical Statistics*, A. N. Shiryayev, ed., Kluwer, Dordrecht, The Netherlands, pp. 62-108).

Kolmogorov, A. N., 1933. *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Ergebnisse der Mathematik und Ihrer Grenzgebiete), Berlin. (2nd English Edition: *Foundations of the Theory of Probability*, 84 pp. Chelsea Publishing Company, New York, 1956).

Kolmogorov, A. N., 1937. Ein vereinfachter Beweis der Birkhoff-Khintchineschen Ergodensatzes. *Mat. Sb.*, 2, 367-368.

Kolmogorov, A.N., 1938. A simplified proof of the Birkhoff-Khinchin ergodic theorem. *Uspekhi Mat. Nauk*, 5, 52-56. (English edition: Kolmogorov, A.N., 1991, *Selected Works of A. N. Kolmogorov - Volume 1, Mathematics and Mechanics*, Tikhomirov, V. M. ed., Kluwer, Dordrecht, The Netherlands, pp. 271-276).

Kolmogorov, A.N., 1940. Wienersche Spiralen und einige andere interessante Kurven in Hilbertschen, *Raum. Dokl. Akad. Nauk* URSS, 26, 115–118 (In German).

Kolmogorov, A.N., 1947. Statistical theory of oscillations with continuous spectrum. *Collected papers on the 30th anniversary of the Great October Socialist Revolution*, Vol. 1, Akad. Nauk SSSR, Moscow-Leningrad, pp. 242-252. (English edition: Kolmogorov, A.N., 1992. *Selected Works of A. N. Kolmogorov - Volume 2, Probability Theory and Mathematical Statistics*, A. N. Shiryayev ed., Kluwer, Dordrecht, The Netherlands, pp. 321-330).

Koutsoyiannis, D., 2000. A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series. *Water Resources Research*, 36 (6), 1519–1533.

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.

References (2)

Koutsoyiannis, D., 2013. Hydrology and Change. *Hydrological Sciences Journal*, 58 (6), 1177–1197, doi: 10.1080/02626667.2013.804626.

Koutsoyiannis, D., 2019a. Knowable moments for high-order stochastic characterization and modelling of hydrological processes. *Hydrological Sciences Journal*, 64(1), 19–33, doi:10.1080/02626667.2018.1556794.

Koutsoyiannis, D., 2019b. Simple stochastic simulation of time irreversible and reversible processes. *Hydrological Sciences Journal* (in press).

Koutsoyiannis, D., 2020. *Stochastics of Hydroclimatic Extremes – A Cool Look at Risk* (in preparation).

Koutsoyiannis, D., and Montanari, A., 2015. Negligent killing of scientific concepts: the stationarity case. *Hydrological Sciences Journal,* 60(7-8), 1174–1183, doi: 10.1080/02626667.2014.959959.

Lombardo, F., Volpi, E., Koutsoyiannis, D., and Papalexiou, S.M., 2014. Just two moments! A cautionary note against use of high-order moments in multifractal models in hydrology. *Hydrology and Earth System Sciences*, 18, 243–255, doi:10.5194/hess-18-243-2014.

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.

Mathieu. M., 1988. On the origin of the notion 'Ergodic Theory', *Expositiones Mathematicae*, 6, 373-377.

Milly, P.C.D., Betancourt, J. Falkenmark, M. Hirsch, R.M. Kundzewicz, Z.W. Lettenmaier, D.P. and Stouffer, R.J., 2008. Stationarity Is Dead: Whither Water Management?. *Science*, 319, 573-574.

Montanari, A., and Koutsoyiannis, D., 2014. Modeling and mitigating natural hazards: Stationarity is immortal!. *Water Resources Research*, 50 (12), 9748–9756, doi: 10.1002/2014WR016092.

Volpi, E., Fiori, A., Grimaldi, S., Lombardo, F., and Koutsoyiannis, D., 2019. Save hydrological observations! Return period estimation without data decimation. *Journal of Hydrology*, doi:10.1016/j.jhydrol.2019.02.017.

Data sources

GHCN Version 3 data: retrieved on 2019-02-17 from https://climexp.knmi.nl/gdcnprcp.cgi?WMO=ITE00100550

Dext3r data: retrieved on 2019-02-17 from http://www.smr.arpa.emr.it/dext3r/