Lecture notes on stochastics – Part 3

Rome, 3-4 May 2018

Knowable moments and K-climacogram

Demetris Koutsoyiannis
Department of Water Resources and Environmental Engineering
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/1835/

Introduction

Classical moments, raw or central, express important theoretical properties of probability distributions but cannot be estimated from typical samples for order beyond 2—cf. Lombardo et al. (2014): "*Just two moments!*".

L-moments are better estimated but they are all of first order in terms of the random variable of interest. They are good to characterize independent series or to infer the marginal distribution of stochastic processes, but they cannot characterize even second order dependence of processes.

Picking from both categories, we introduce K-moments, which combine advantages of both classical and L moments. They enable reliable estimation from samples (in some cases even more reliable than L moments) and effective description of high order statistics, useful for marginal and joint distributions of stochastic processes.

High-order joint statistics of stochastic properties involve multivariate functions expressing joint high-order moments. Here, by extending the notion of climacogram (Koutsoyiannis, 2010, 2016) and climacospectrum (Koutsoyiannis, 2017) we introduce the K-climacogram and the K-climacospectrum, which enable characterization of high-order properties of stochastic processes, as well as preservation thereof in simulations, in terms of univariate functions.

A note on classical moments

The classical definitions of raw and central moments of order p are:

$$\mu'_p \coloneqq \mathrm{E}\left[\underline{x}^p\right], \qquad \mu_p \coloneqq \mathrm{E}\left[\left(\underline{x} - \mu\right)^p\right] \tag{1}$$

respectively, where $\mu \coloneqq \mu'_1 = \mathrm{E}\left[\underline{x}\right]$ is the mean of the random variable $\underline{x}$. Their standard estimators from a sample $\underline{x}_i$, $i = 1, \dots, n$, are

$$\underline{\hat{\mu}}'_p = \frac{1}{n}\sum_{i=1}^{n} \underline{x}_i^p\,,\ \ \underline{\hat{\mu}}_p = \frac{b(n,p)}{n}\sum_{i=1}^{n}\left(\underline{x}_i - \hat{\mu}\right)^p \tag{2}$$

where $a(n, p)$ is a bias correction factor (e.g. for the variance $\mu_2 =: \sigma^2$, $b(n, 2) = n/(n-1)$). The estimators of the raw moments $\underline{\hat{\mu}}'_p$ are in theory unbiased, but it is practically impossible to use them in estimation if $p > 2$—cf. Lombardo et al. (2014), "Just two moments".

In fact, because for large p, it holds that $\left(\frac{1}{n}\sum_{i=1}^{n} x_i^p\right)^{1/p} \approx \max_{1\le i\le n}(x_i)$ *, we can conclude that, for an unbounded variable $\underline{x}$, asymptotically $\underline{\hat{\mu}}'_p$ **is not an estimator of $\boldsymbol{\mu'_p}$** but one of an extreme quantity, i.e., the nth order statistic raised to power p. Thus, unless p is very small, **$\boldsymbol{\mu'_p}$ is not a *knowable* quantity**: we cannot infer its value from a sample. **This is the case even if *n* is very large**!

---

* This is precise if x_i are positive; see also p. 5.

Definition of K-moments

To derive *knowable* moments for high orders p, in the expectation defining the pth moment we raise $\underline{x} - \mu$ to a lower power $q < p$ and for the remaining $p - q$ terms we replace $\underline{x} - \mu$ with $2F(\underline{x}) - 1$, where $F(x)$ is the distribution function. This leads to the following (central) *K-moment* definition:

$$K_{pq} \coloneqq (p - q + 1)\mathrm{E}\left[\left(2F\left(\underline{x}\right) - 1\right)^{p-q}\left(\underline{x} - \mu\right)^{q}\right] \tag{3}$$

Likewise, we define non-central K-moments as:

$$K'_{pq} \coloneqq (p - q + 1)\mathrm{E}\left[\left(F\left(\underline{x}\right)\right)^{p-q}\underline{x}^{q}\right] \tag{4}$$

The quantity $\left(2F\left(\underline{x}\right) - 1\right)^{p-q}$ is estimated from a sample without using powers of $\underline{x}$. Specifically, for the ith element of a sample $x_{(i)}$ of size n, sorted in ascending order, $F(x_{(i)})$, is estimated as $\hat{F}\left(x_{(i)}\right) = (i - 1)/(n - 1)$, thus taking values from 0 to 1 precisely and irrespective of the values $x_{(i)}$; likewise, $2F(x_{(i)}) - 1$ is estimated as $2\hat{F}\left(x_{(i)}\right) - 1 = (2i - n + 1)/(n - 1)$, taking values from –1 to 1 precisely and irrespective of the values $x_{(i)}$. Hence, the estimators are:

$$\underline{\hat{K}}'_{pq} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)^{p-q}\underline{x}_{(i)}^{q}, \quad \underline{\hat{K}}_{pq} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{2i-n+1}{n-1}\right)^{p-q}\left(\underline{x}_{(i)} - \hat{\mu}\right)^{q} \tag{5}$$

Rationale of the definition

1. Assuming that the distribution mean is close to the median, so that $F(\mu) \approx 1/2$ (this is precisely true for a symmetric distribution), the quantity whose expectation is taken in (3) is $A(x) \coloneqq \left(2F(\underline{x}) - 1\right)^{p-q} (\underline{x} - \mu)^q$ and its Taylor expansion is
$$A(x) = \left(2f(\mu)\right)^{p-q} (\underline{x} - \mu)^p + (p-q)\left(2f(\mu)\right)^{p-q-1} f'(\mu)(\underline{x} - \mu)^{p+1} + O((\underline{x} - \mu)^{p+2}) \tag{6}$$
where $f(x)$ is the probability density function of $\underline{x}$. Clearly then, K_{pq} depends on μ_p as well as classical moments of $\underline{x}$ of order higher than p. The independence of K_{pq} from classical moments of order < p makes it a good knowable surrogate of the unknowable μ_p.
2. As p becomes large, by virtue of the multiplicative term $(p - q + 1)$ in definition (3), K_{pq} shares similar asymptotic properties with $\hat{\mu}_p^{q/p}$ (the estimate, not the true $\mu_p^{q/p}$). To illustrate this for q = 1, we consider the variable $\underline{z} \coloneqq \max_{1\le i\le p} \underline{x}_i$ and denote $f(\ )$ and $h(\ )$ the probability densities of $\underline{x}_i$ and $\underline{z}$, respectively. Then (Papoulis, 1990, p. 209):
$$h(z) = pf(z)\left(F(z)\right)^{p-1} \tag{7}$$
and thus, by virtue of (4),
$$\mathrm{E}[\underline{z}] = p\mathrm{E}\left[\left(F(\underline{x})\right)^{p-1} \underline{x}\right] = K'_{p1} \tag{8}$$
On the other hand, as seen in p. 2, for positive $\underline{x}$ and large $p \to n$,
$$E[\underline{\hat{\mu}}'_p{}^{1/p}] = E\left[\left(\frac{1}{n}\sum_{i=1}^{n} \underline{x}_i^p\right)^{1/p}\right] \approx E\left[\max_{1\le i\le n} x_i\right] = \mathrm{E}[\underline{z}] = K'_{p1} \tag{9}$$
Note also that the multiplicative term $(p - q + 1)$ in definition (3) and (4) makes K-moments increasing functions of p.

Asymptotic properties of moment estimates

Generally, as p becomes large (approaching n), estimates of both classical and K moments, central or non-central, become estimates of expressions involving extremes such as $\left(\max_{1\le i\le p} x_i\right)^q$ or $\max_{1\le i\le p}(x_i-\mu)^q$. For negatively skewed distributions these quantities can also involve minimum, instead of maximum quantities.

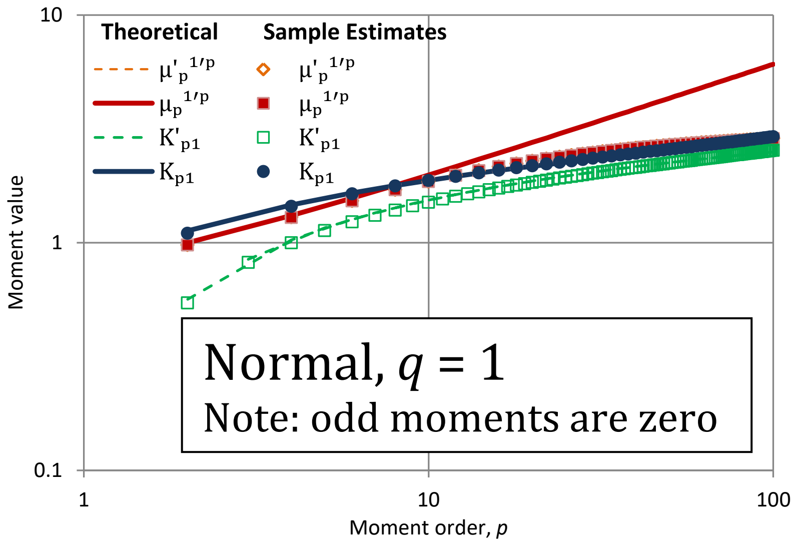
For the K-moments this is consistent with their theoretical definition. For the classical moments this is an inconsistency.

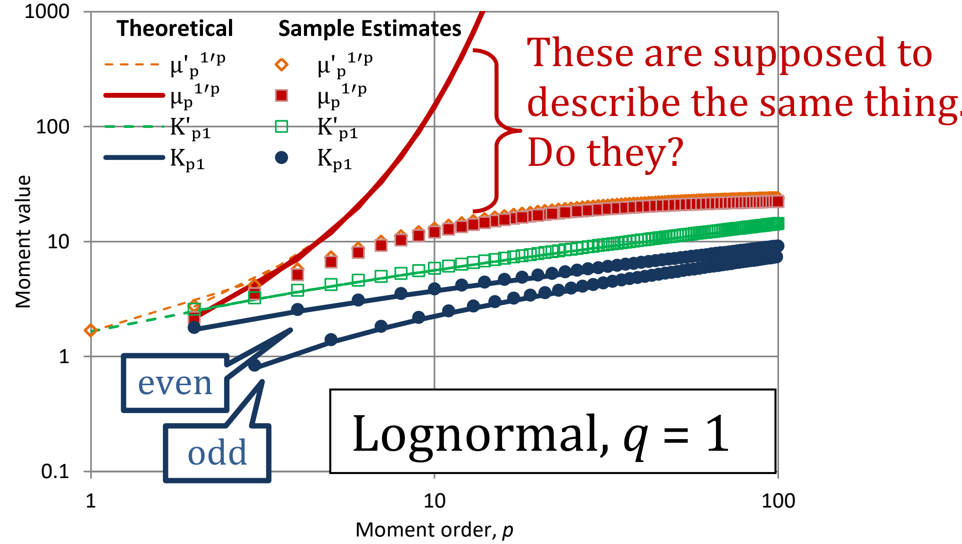
A common property of both classical and K moments is that symmetrical distributions have all their odd moments equal to zero.

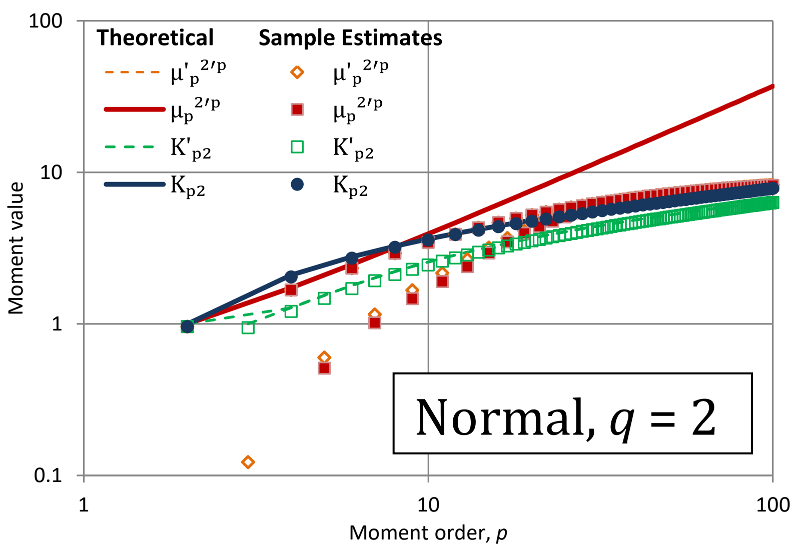
Both classical and K moments are non-decreasing functions of p, separately for odd and even p.

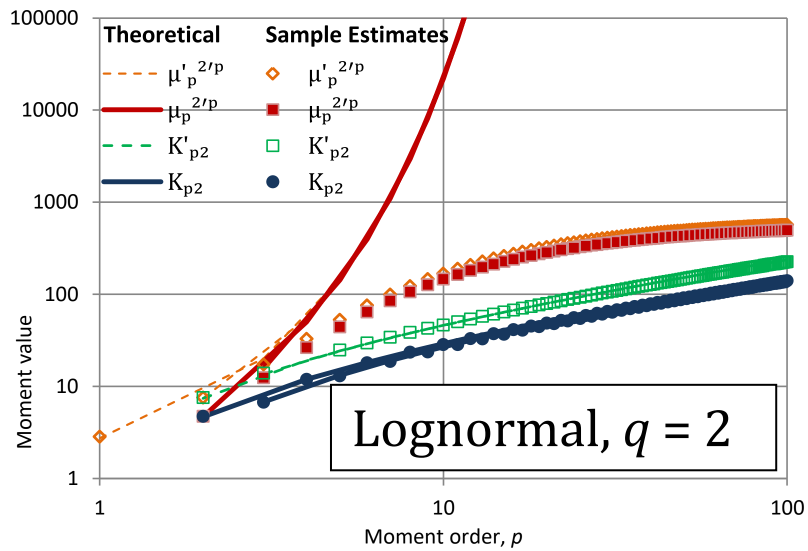
In geophysical processes we can justifiably assume that the variance $\mu_2 \equiv \sigma^2 \equiv K_{22}$ is finite (an infinite variance would presuppose infinite energy to materialize, which is absurd). Hence, high order K-moments K_{p2} will be finite too, even if classical moments μ_p diverge to infinity beyond a certain p (i.e., in heavy tailed distributions).

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.

Relationship among different moment types

The classical moments can be recovered as a special case of K moments: $M_p \equiv K_{pp}$. In particular, in uniform distribution, classical and K moments are proportional to each other:

$$K'_{pq} \coloneqq (p-q+1)\mu'_p, \quad K_{pq} \coloneqq (p-q+1)\mu_p \tag{10}$$

The probability weighted moments (PWM), defined as $\beta_p \coloneqq \mathrm{E}\left[\underline{x}\left(F(\underline{x})\right)^p\right]$, are a special case of K-moments corresponding to q = 1:

$$K'_{p1} = p\beta_{p-1} \tag{11}$$

The L-moments defined as $\lambda_p \coloneqq \frac{1}{p}\sum_{k=0}^{p-1}(-1)^k \binom{p-1}{k} \mathrm{E}\left[\underline{x}_{(p-k):p}\right]$, where $\underline{x}_{k:p}$ denotes the kth order statistic in an independent sample of size p. L-moments are also related to PWM and through them to K moments. In particular, the relationships for the different types of moments for the first four orders are:

$$\begin{aligned}
&K'_{11} = \mu = \beta_0, && K_{11} = 0 \\
&K'_{21} = 2\beta_1, && K_{21} = 2(K'_{21} - \mu) = 4\beta_1 - 2\beta_0 = 2\lambda_2 \\
&K'_{31} = 3\beta_2, && K_{31} = 4(K'_{31} - \mu) - 6(K'_{21} - \mu) = 12\beta_2 - 12\beta_1 + 2\beta_0 = 2\lambda_3 \\
&K'_{41} = 4\beta_3, && K_{41} = 8(K'_{41} - \mu) - 16(K'_{31} - \mu) + 12(K'_{21} - \mu) \\
& && \quad = 32\beta_3 - 48\beta_2 + 24\beta_1 - 4\beta_0 = \frac{8}{5}\lambda_4 + \frac{12}{5}\lambda_2
\end{aligned} \tag{12}$$

Basic characteristics of marginal distribution

Within the framework of K-moments, we can (and should) use "Just two moments" in terms of the power of $\underline{x}$, i.e. q = 1 or 2, but we can obtain knowable statistical characteristics for much higher order p.

In this manner, for $p > 1$ we have two alternative options to define statistical characteristics related to moments of the distribution, as in the table below. (Which of the two is preferable depends on the statistical behaviour and in particular, mean, mode and variance of the estimator.)

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

High order moments for stochastic processes: the K-climacogram and the K-climacospectrum

Second order properties of stochastic properties are typically expressed by the autocovariance function, $c(h) := \mathrm{cov}[\underline{x}(t), \underline{x}(t+h)]$. An equivalent description is by the power spectrum, which is the Fourier transform of the autocovariance, $s(w) \coloneqq 4\int_0^\infty c(h)\cos(2\pi w h)\,\mathrm{d}h$.

Another fully equivalent description with many advantages (Dimitriadis and Koutsoyiannis 2015, Koutsoyiannis 2016) is through the climacogram, the variance of the averaged process, i.e., $\gamma(k) \coloneqq \mathrm{var}[\underline{X}(k)/k]$, where $\underline{X}(t) \coloneqq \int_0^t \underline{x}(\xi)\mathrm{d}\xi$. The climacogram is connected to autocovariance by $\gamma(k) = 2\int_0^1 (1-\chi)c(\chi k)\mathrm{d}\chi$ and $c(h) = \frac{1}{2}\frac{\mathrm{d}^2\left(h^2\gamma(h)\right)}{\mathrm{d}h^2}$. A surrogate of the power spectrum with several advantages over it is the climacospectrum (Koutsoyiannis, 2017) defined as $\zeta(k) \coloneqq \frac{k\left(\gamma(k)-\gamma(2k)\right)}{\ln 2}$.

Full description of the third-order, fourth-order, etc., properties of a stochastic process requires functions of 2, 3, …, variables. For example, the third order properties are expressed in terms of $c_3(h_1, h_2) := \mathrm{E}[(\underline{x}(t) - \mu)\,(\underline{x}(t+h_1) - \mu)\,(\underline{x}(t+h_2) - \mu)]$.

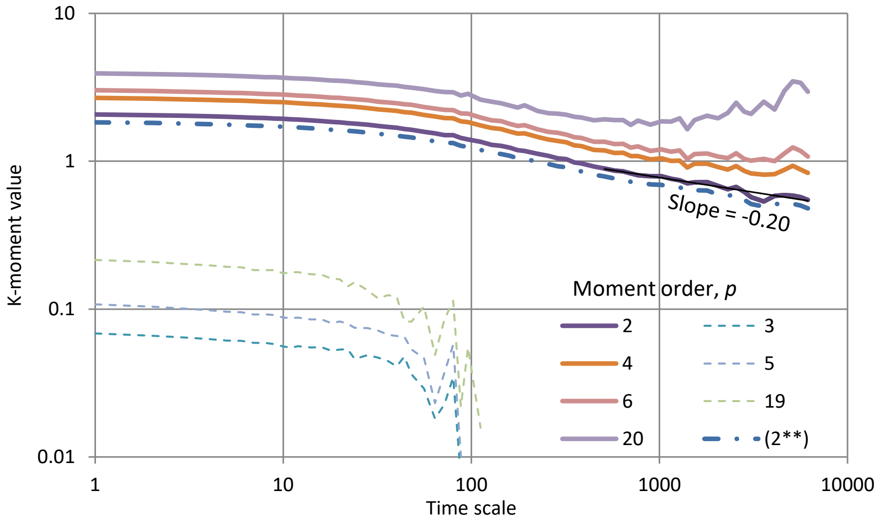
Such a description is not parsimonious and its accuracy holds only in theory, because sample estimates are not reliable. Therefore we introduce single-variable descriptions for any order p, expanding the idea of the climacogram and climacospectrum based on K-moments.

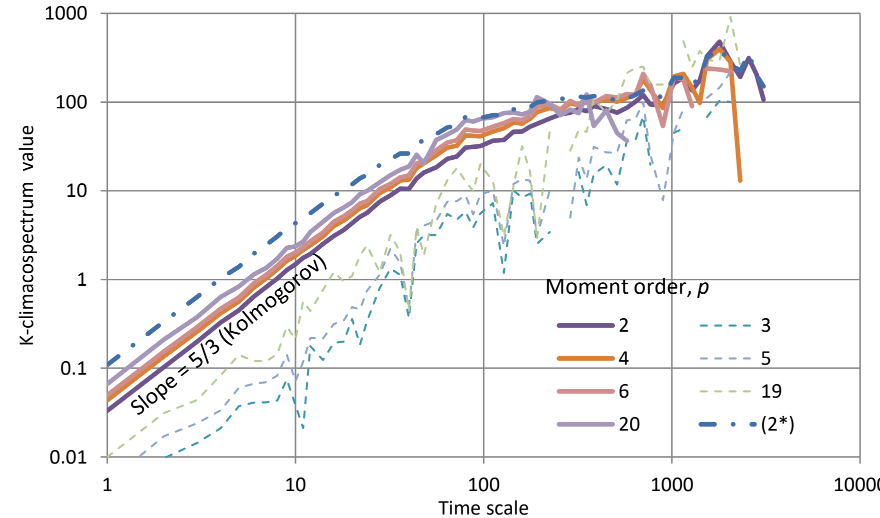
K-climacogram: $$\gamma_{pq}(k) = (p-q+1)\mathrm{E}\left[\left(2F\left(\underline{X}(k)/k\right)-1\right)^{p-q}\left(\underline{X}(k)/k-\mu\right)^q\right] \tag{13}$$

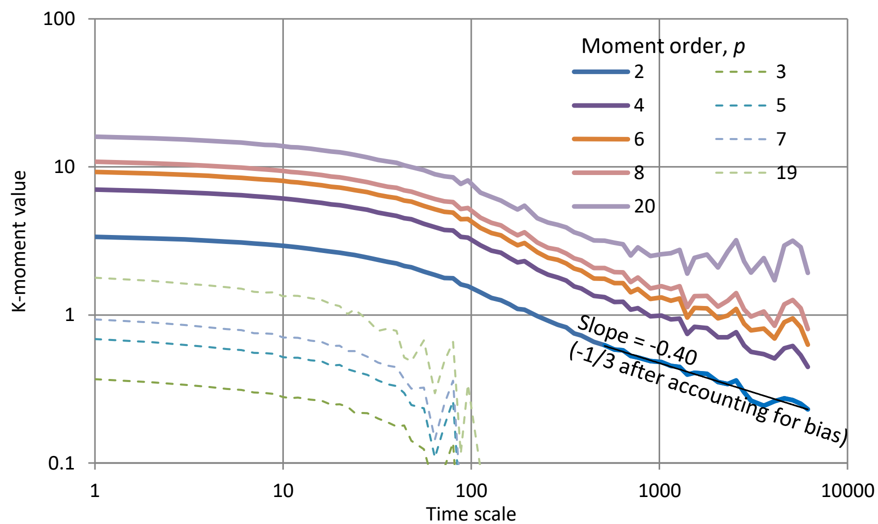
K-climacospectrum: $$\zeta_{pq}(k) = \frac{k\left(\gamma_{pq}(k)-\gamma_{pq}(2k)\right)}{\ln 2} \tag{14}$$

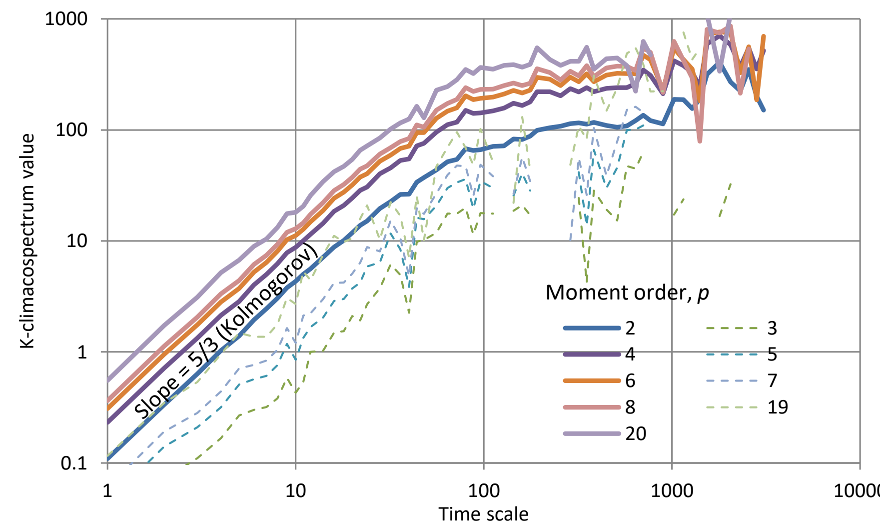
where $\gamma_{22}(k) \equiv \gamma(k)$ and $\zeta_{22}(k) \equiv \zeta(k)$. Even though the K-moment description is not equivalent to the multivariate high-order one, it suffices to define the marginal distribution at any scale k.

Example 1: Turbulent velocity



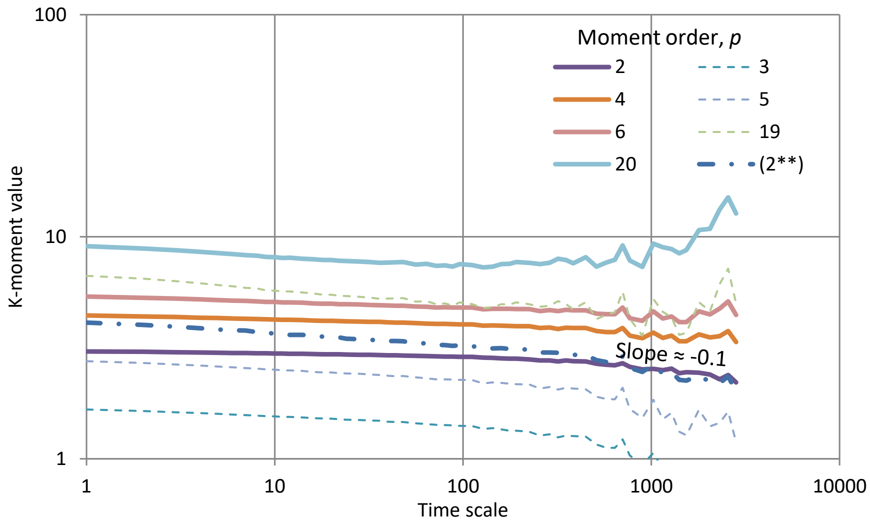


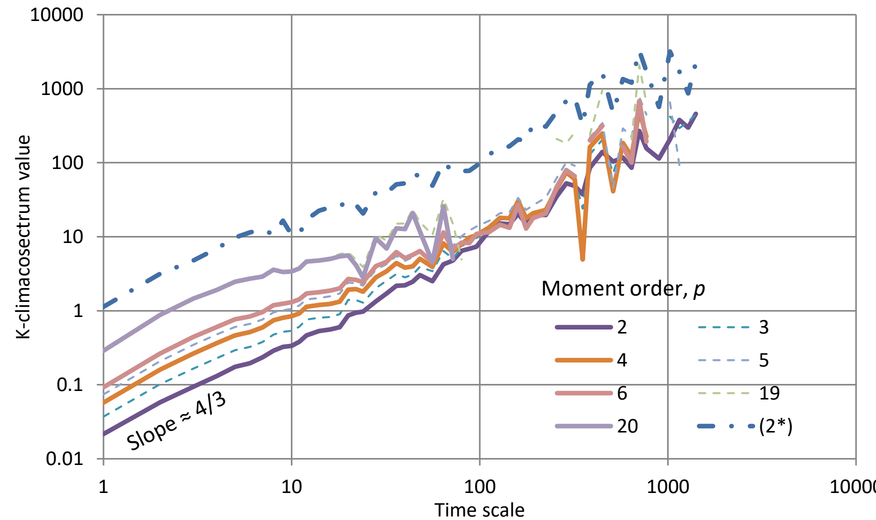


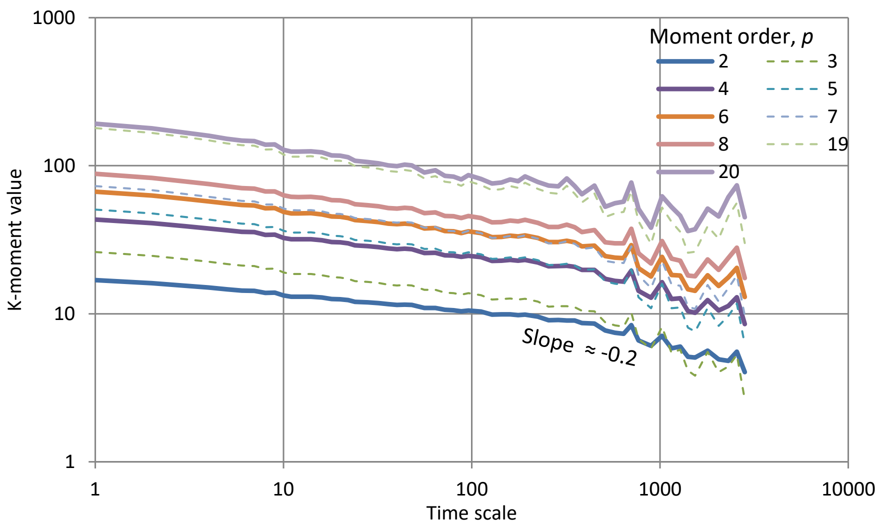


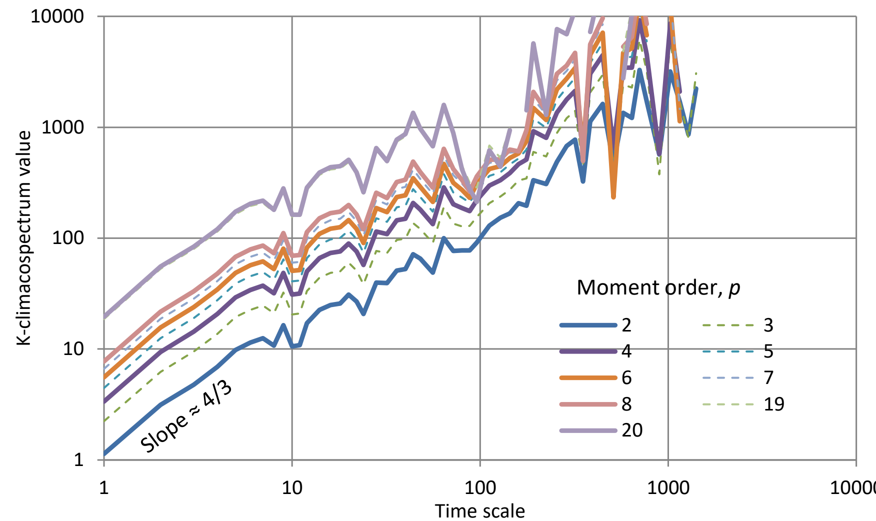
Data: 60 000 values of turbulent velocity along the flow direction (Kang, 2003; Koutsoyiannis 2017, Dimitriadis and Koutsoyiannis, 2018); the original series was averaged so that time scale 1 corresponds to 0.5 s.
Note: Plot (2*) is constructed from the variance and (2**) corresponds to standard deviation.

Example 2: Rainfall rate at Iowa measured every 10 s



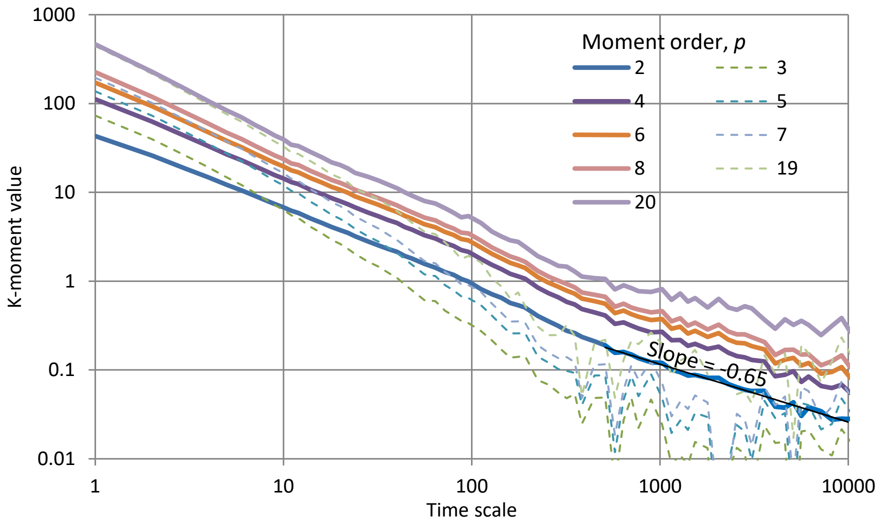


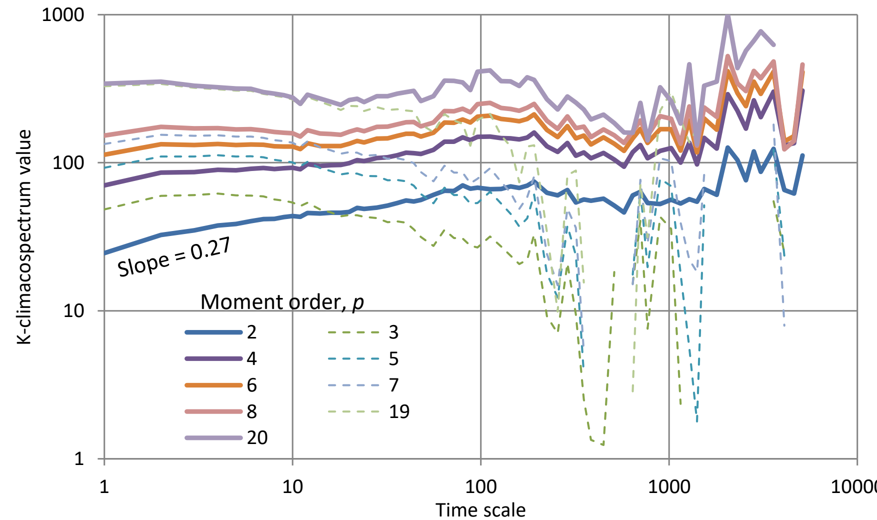


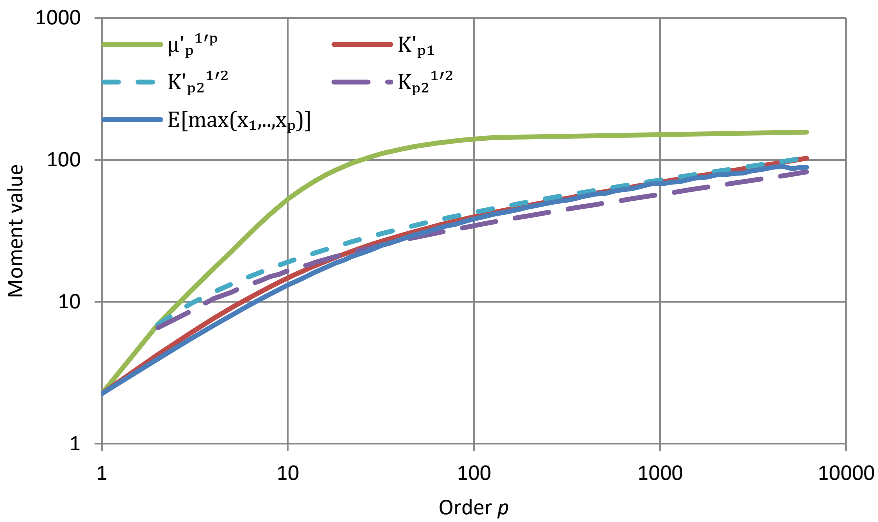


Data: 29542 values of rainfall at Iowa measured at temporal resolution of 10 s (merger of seven events from Georgakakos *et al.* 1994; see also Lombardo et al. 2012). Plot (2*) is constructed from the variance and (2**) corresponds to standard deviation.

Example 3: Daily rainfall at Padova







Data: 100 442 values of daily rainfall at Padova (the longest rainfall record existing worldwide; Marani and Zanetti, 2015).

Note about the graph on the left: Notice that moments are plotted against order p and thus approximately represent maxima for a time window of length p. For independent processes $\mathrm{E}\left[\max\left(\underline{x}_1, \dots, \underline{x}_p\right)\right]$ should be equal to K'_{p1}, but when there is dependence the two quantities slightly differ; the former reflects the joint distribution and the latter the marginal one.

References

Dimitriadis, P., and Koutsoyiannis, D. (2015), Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes, *Stochastic Environmental Research & Risk Assessment*, 29 (6), 1649–1669, doi:10.1007/s00477-015-1023-7.

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

Georgakakos, K. P., Cârsteanu, A. A., Sturdevant, P. L. and Cramer, J. A. (1994), Observation and Analysis of Midwestern Rain Rates, *Journal of Applied Meteorology*, 33, 1433-1444.

Kang, H.S., Chester, S., and Meneveau, C. (2003). Decaying turbulence in an active-grid-generated flow and comparisons with large-eddy simulation, *Journal of Fluid Mechanics*, 480, 129–160.

Koutsoyiannis, D. (2010), A random walk on water, *Hydrology and Earth System Sciences*, 14, 585–601, doi:10.5194/hess-14-585-2010.

Koutsoyiannis, D. (2016), Generic and parsimonious stochastic modelling for hydrology and beyond, *Hydrological Sciences Journal*, 61 (2), 225–244, doi:10.1080/02626667.2015.1016950.

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

Lombardo, F., Volpi, E., and Koutsoyiannis, D. (2012), Rainfall downscaling in time: Theoretical and empirical comparison between multifractal and Hurst-Kolmogorov discrete random cascades, *Hydrological Sciences Journal*, 57 (6), 1052–1066.

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.

Marani, M., and Zanetti, S. (2015), Long-term oscillations in rainfall extremes in a 268 year daily time series, *Water Resources Research*, 51, 639–647.