Multifractal downscaling models: A crash test

Federico Lombardo, Elena Volpi
Department of Civil Engineering Sciences,
University of Rome "Roma Tre", Italy
(flombardo@uniroma3.it, evolpi@uniroma3.it)

Simon Michael Papalexiou, Demetris Koutsoyiannis
Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Greece

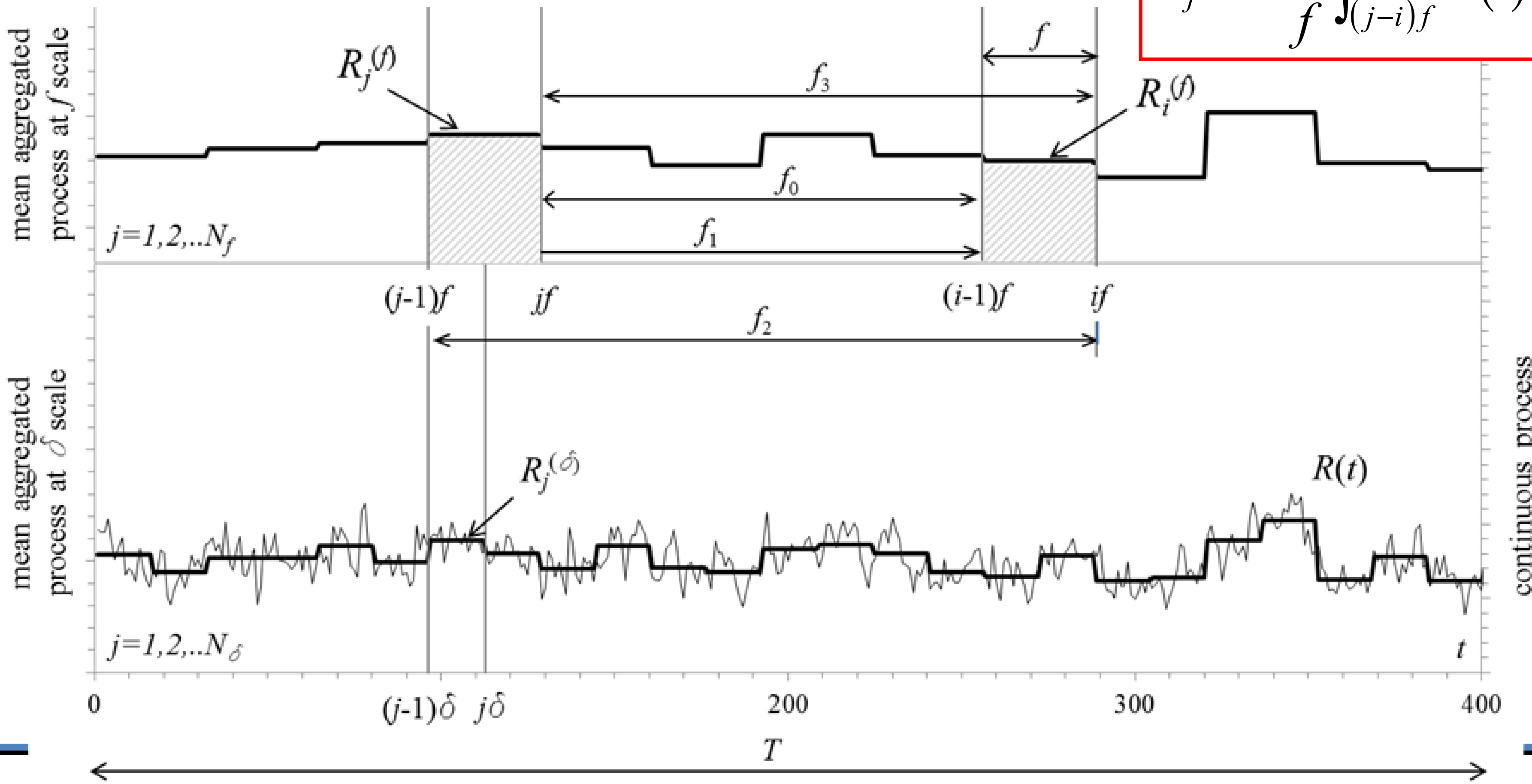
Outline

- Multifractal rainfall models have been widely used to reproduce several statistical properties of actual rainfall fields in finite but practically important ranges of scales (f).
- These properties include the scaling of the moments of different orders (q) which is used in model identification and fitting.
- Sample estimates of q-th order moments (from a single sample) can be very uncertain in the case of a process characterized by temporal dependence; thus, results may lead to false conclusions (e.g. Papalexiou *et al.*, 2010; Houdalaki *et al.*, 2012).
- Therefore, we aim at analysing this uncertainty as a function of the statistical properties of the process, the scale f and the moment order q.

Analytical framework: mean aggregated process at scale *f*

- Let $R(t)$ be a stochastic process in continuous time t with μ=E[R], σ^2=var[R] and $\rho(t-t')$=cov[$R(t)$, $R(t')$]/σ^2.

$$R_j^{(f)} = \frac{1}{f}\int_{(j-i)f}^{jf} R(t)dt$$



Multifractal analysis

- Multifractal analysis is based on detecting power law behaviours for the *q*-order moments estimated over a range of aggregation scales *f*:

$$\left\langle \left[R_j^{(f)}\right]^q \right\rangle \propto f^{K(q)}$$

where $K(q)$ is the moment scaling exponent.

- Calibration of multifractal models is usually based on the estimation of the moment scaling exponents $K(q)$, which relies on the estimate of the sample *q*-th order moments at the scales *f* and their linear regressions in log-log diagrams.

- The sample *q*-th moments of the mean aggregated process are given by:

$$m^{(f)}(q) = \frac{1}{N_f} \sum_{j=1}^{N_f} \left(R_j^{(f)}\right)^q$$

where $N_f = T/f$ is the length of the mean aggregated series.

Sample q-moment estimation

- The mean and the variance of the sample q-th moments are given respectively by:

$$\mathrm{E}\left[m^{(f)}(q)\right] = \frac{1}{N_f}\sum_{j=1}^{N_f}\mathrm{E}\left[\left(R_j^{(f)}\right)^q\right] = \mu_q$$

$$\mathrm{var}\left[m^{(f)}(q)\right] = \frac{1}{N_f^2}\sum_{i=1}^{N_f}\sum_{j=1}^{N_f}\mathrm{E}\left[\left(R_j^{(f)}\right)^q\left(R_i^{(f)}\right)^q\right] - \mathrm{E}\left[m^{(f)}(q)\right]^2$$

- The sample moment is an unbiased estimator of its theoretical value μ_q

- The variance can be assumed as a measure of uncertainty in the moment estimation, and it is expected to depend on:
  - statistical properties of the underlying stochastic process $R(t)$, mean μ, variance σ^2 and autocorrelation function $\rho(t-t')$;
  - aggregation scale f;
  - sample size T;
  - moment order q.

Variance of the mean (q=1)

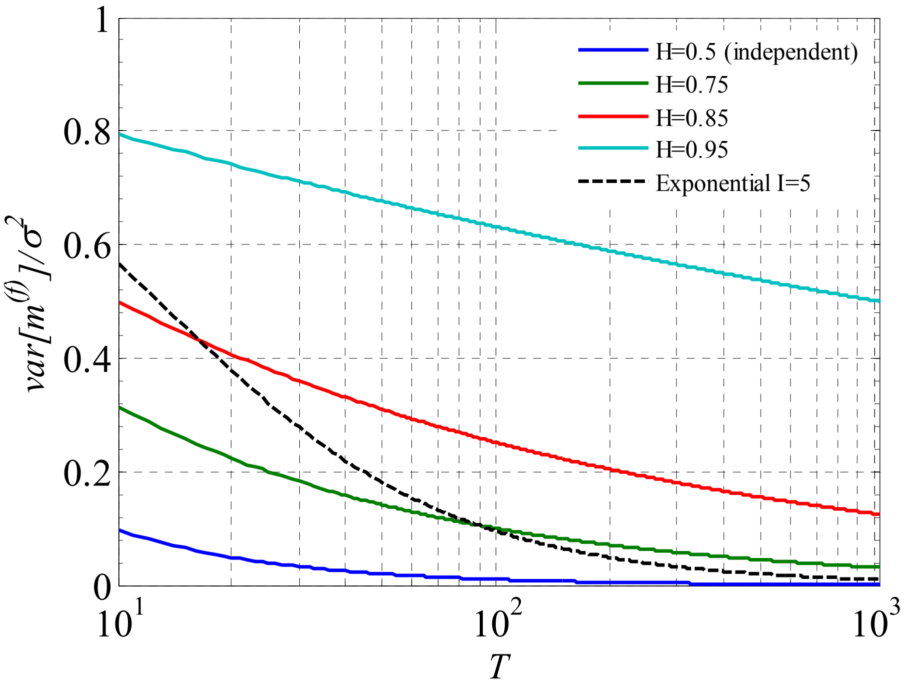
- Following Vanmarcke (1983), the variance of the sample mean is given by:

$$\operatorname{var}\left[m^{(f)}\right] = \frac{2\sigma^2}{T^2}\int_0^T (T-r)\rho(r)dr$$

where r is the time-lag.

- $\operatorname{var}[m^{(f)}]/\sigma^2$ is independent of the scale f, but still depends on the time dependence of the process.

- The greater the time dependence (e.g. Hurst coefficient H for Hurst-Kolmogorov processes), the larger samples required in order to obtain estimates of similar quality.



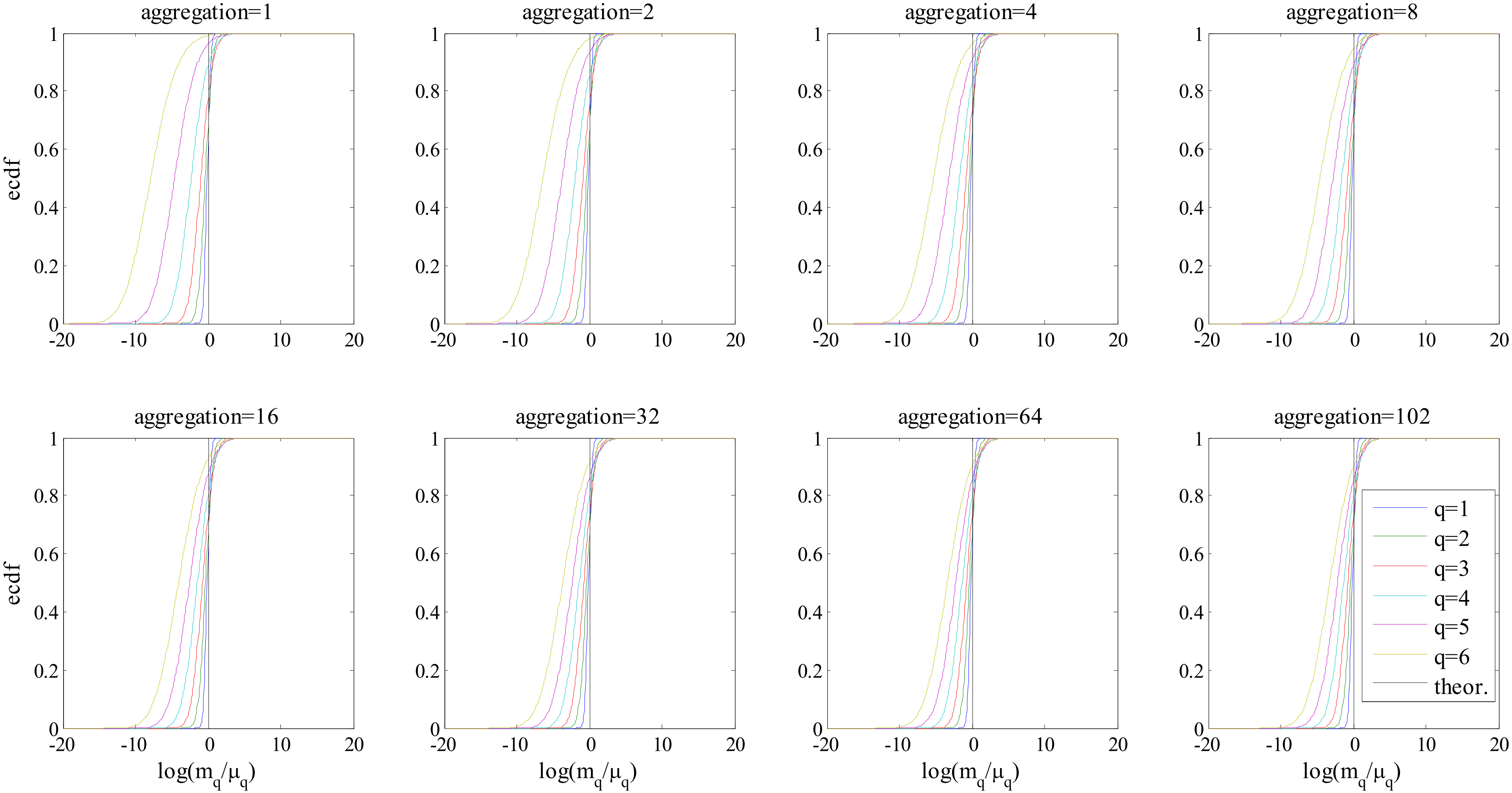
q-order moments: Monte Carlo simulations

- Let us investigate the behaviours of estimators of higher order moments when varying the temporal dependence of the underlying random process; this can be done by Monte Carlo simulation.

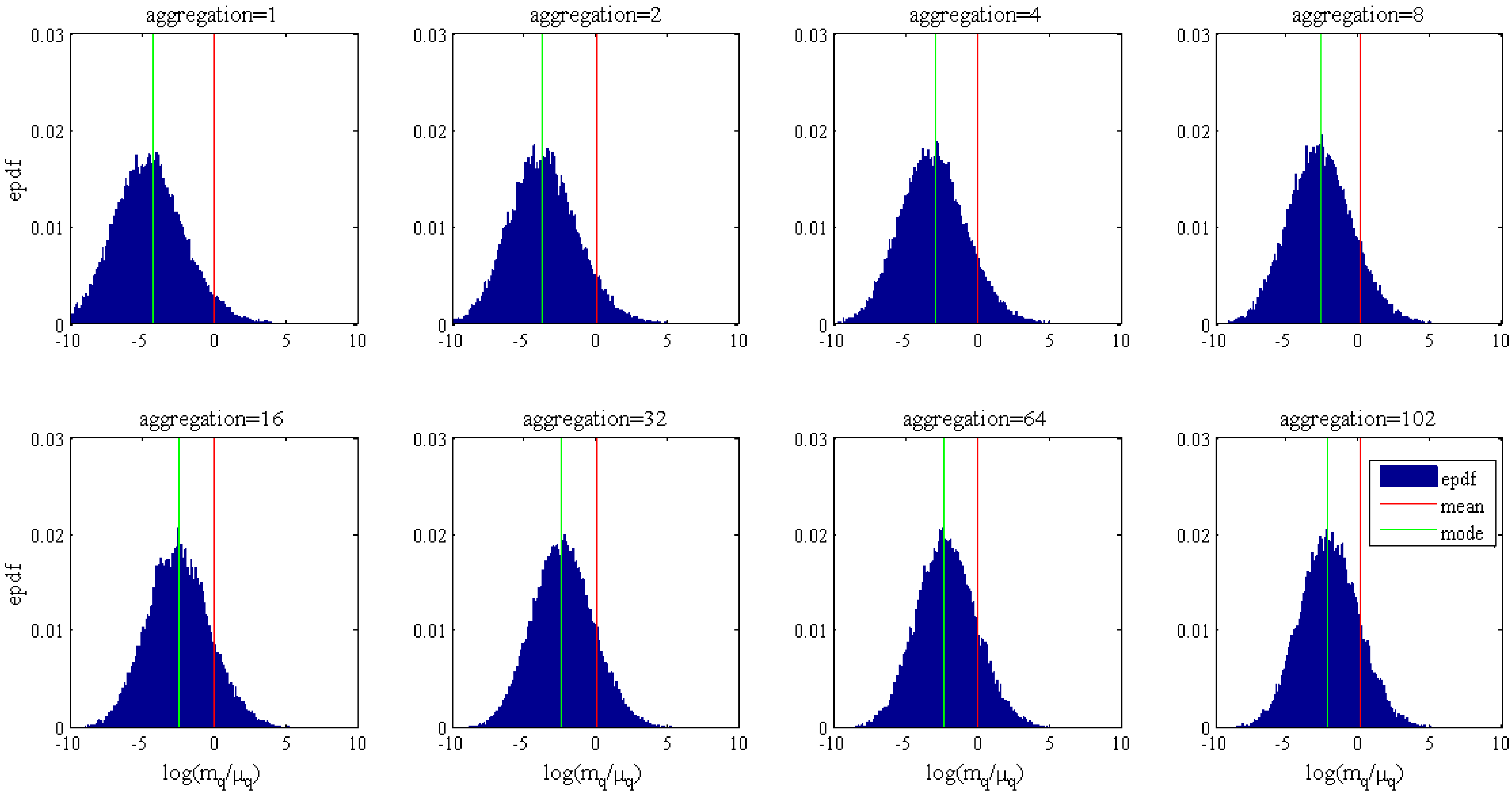
- Moreover, we particularly focus on assessing the influence or relative importance of the data aggregation at different scales f in determining uncertainty in estimation of raw moments.

- To this aim, we use a rainfall downscaling model based on the Hurst-Kolmogorov (HK) process (Lombardo *et al.*, 2012); the Hurst coefficient H is the only model parameter and it has a simple dyadic cascade structure.

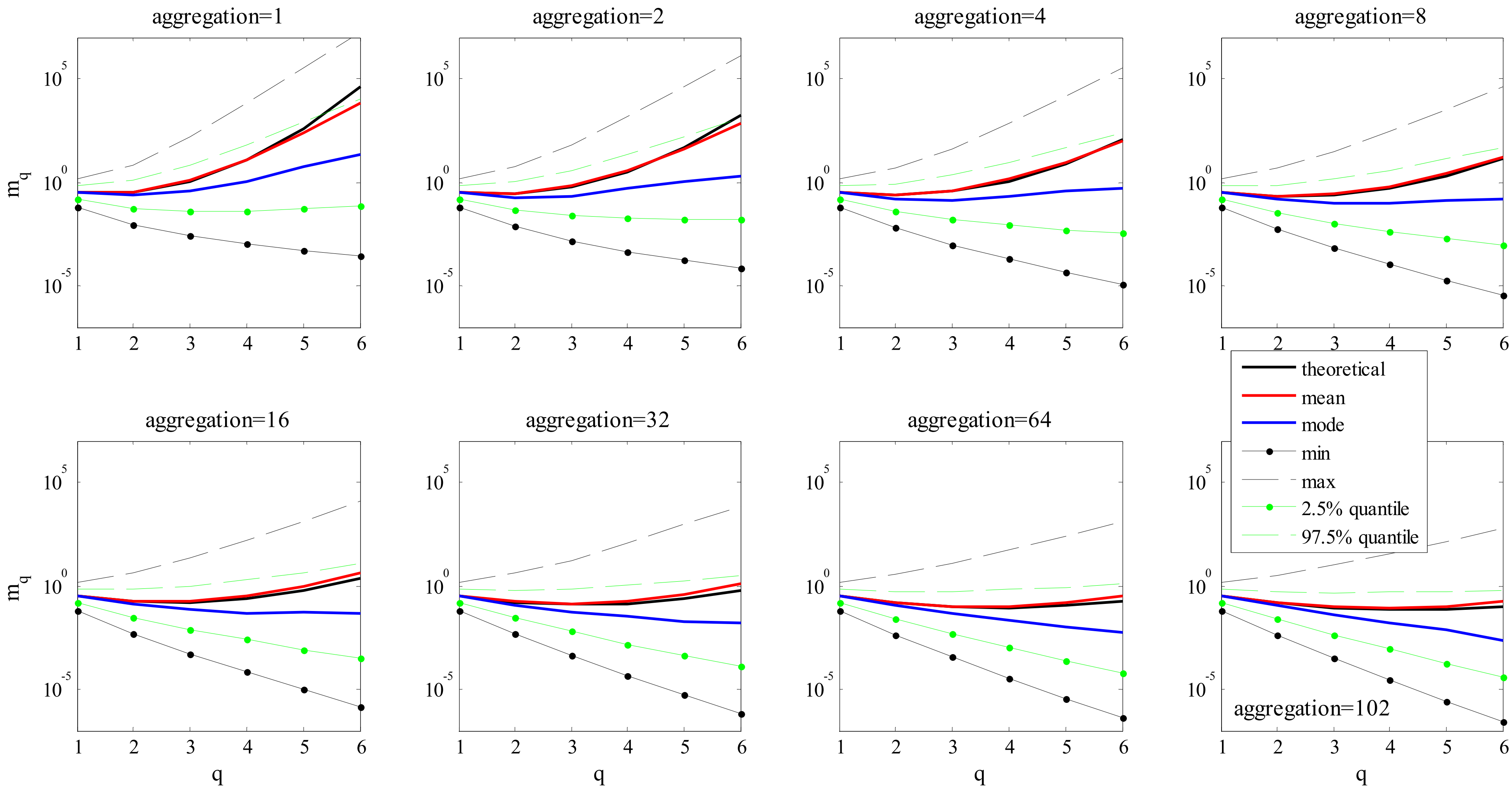
- We generate 30000 log-normal time series with sample size of T=2^{10}=1024, with μ=0.33, σ=0.49 and H=0.85.



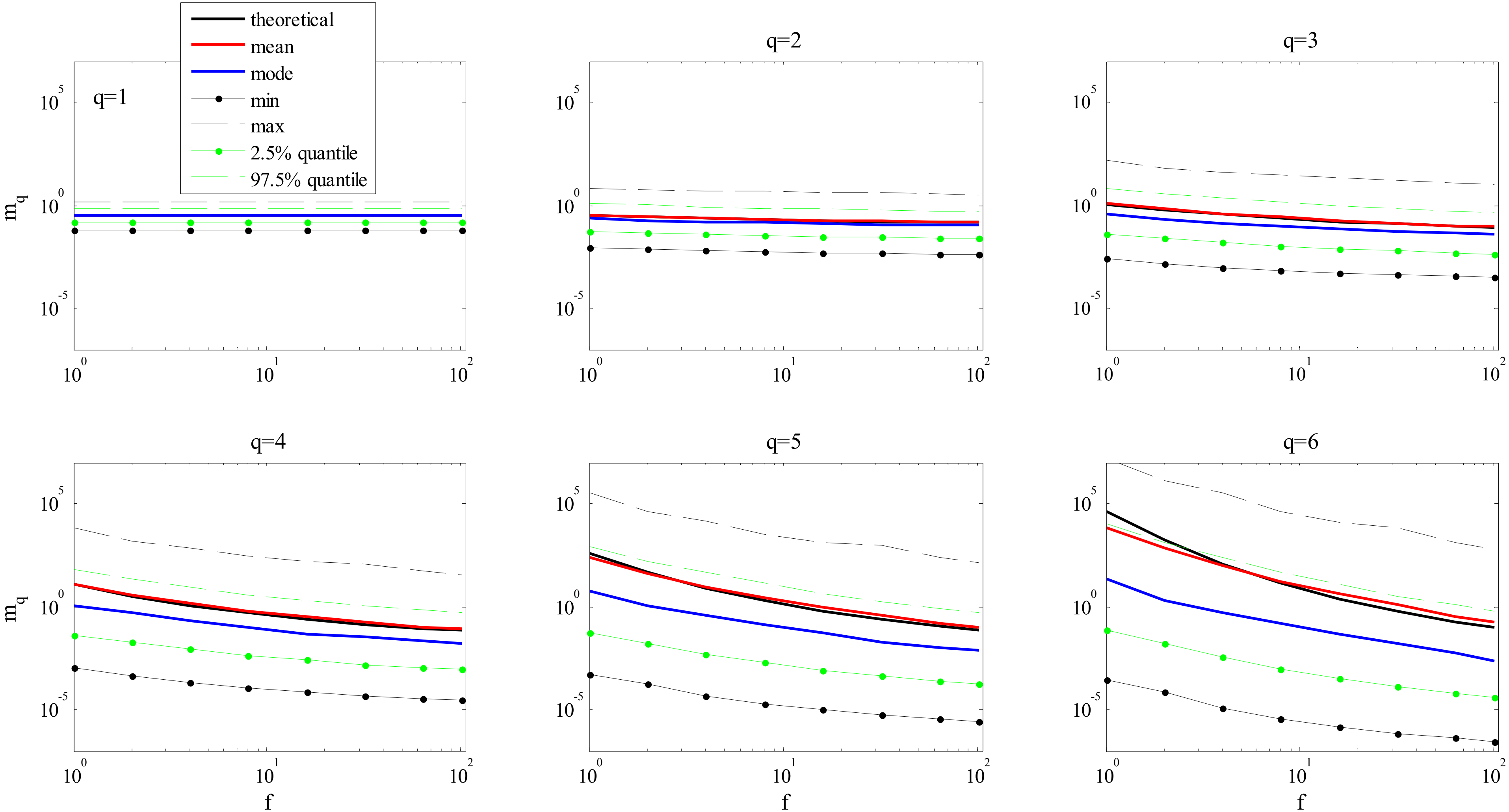
- We show the empirical distribution of the natural logarithm of the ratio of sample q-th moments to their theoretical values.
- The higher the order the less the information content of the empirically estimated moments is (the distribution is less concentrated around 0).
- Increasing the aggregation scale, the moment variability slightly decreases.



- We show here the empirical frequency distribution of the sample 5-th moment estimated from generated series aggregated at different scales.
- Despite being a theoretically unbiased estimator, the most probable value of the moment (the mode) can be two orders of magnitude less than its expected value. The situation slightly improves when increasing the scale.



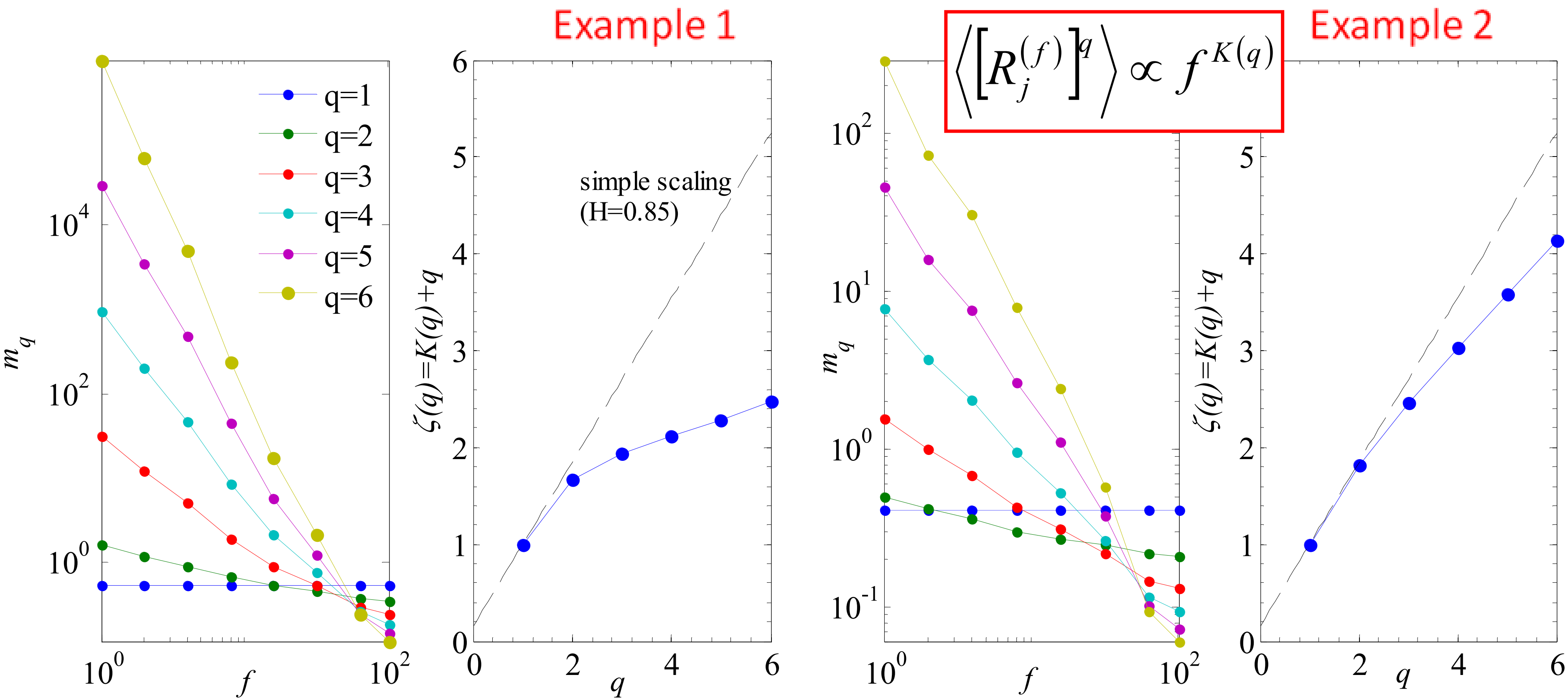
- Semilogarithmic plots of the confidence intervals of the sample moments, which depict the huge variability of estimates when the order increases.
- The discrepancy between the expected value and the mode is also evident.



- log-log diagrams of the confidence intervals of the sample moments against the scale *f*.
- As noted earlier, the scale of aggregation has little influence on the variability of the moments, meaning that the reduction of sample size is somewhat compensated by the time averaging.

Crash test

- Even if the generated process is not multifractal the sample estimates of the *q*-moments from a unique sample can lead to measling results.



$$\left\langle \left[R_j^{(f)} \right]^q \right\rangle \propto f^{K(q)}$$

Linear regressions of sample moments vs scale in log-log diagrams to estimate *K*(*q*)

Conclusions

- The multifractal framework provide parsimonious models to study the space-time variability of several natural processes in geosciences, such as rainfall.
- Models following this approach require the scaling of the moments of different orders (q), which is used in model identification and fitting.
- Using simple Monte Carlo simulations, we form “crash test” conditions and find that the reliability of such methods is questionable.
- Indeed, the estimation of moments is problematic in case of processes exhibiting temporal dependence (such as many geophysical processes). This is especially so in higher orders.
- Despite the sample moment being an unbiased estimator of its theoretical value, Monte Carlo simulations show that the most probable value of the moment (the mode) can strongly differ from its expected value.

References

- Houdalaki, E., M. Basta, N. Boboti, N. Bountas, E. Dodoula, T. Iliopoulou, S. Ioannidou, K. Kassas, S. Nerantzaki, E. Papatriantafyllou, K. Tettas, D. Tsirantonaki, S.M. Papalexiou, and D. Koutsoyiannis, On statistical biases and their common neglect, *European Geosciences Union General Assembly 2012, Geophysical Research Abstracts, Vol. 14*, Vienna, 4388, European Geosciences Union, 2012.

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

- Papalexiou, S.M., D. Koutsoyiannis, and A. Montanari, Mind the bias!, *STAHY Official Workshop: Advances in statistical hydrology*, Taormina, Italy, International Association of Hydrological Sciences, 2010.

- Vanmarcke, E., Random fields: Analysis and synthesis, *MIT Press*, Cambridge, MA., 382 pp., 1983.