

Improved moment scaling estimation for multifractal signals

D. Veneziano¹ and P. Furcolo²

¹Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

²Dipartimento di Ingegneria Civile, Università degli Studi di Salerno, Fisciano (SA) 84084, Italy

Received: 12 February 2009 – Revised: 27 September 2009 – Accepted: 28 September 2009 – Published: 4 November 2009

Abstract. A fundamental problem in the analysis of multifractal processes is to estimate the scaling exponent $K(q)$ of moments of different order q from data. Conventional estimators use the empirical moments $\hat{\mu}_r^q = \langle |\varepsilon_r(\tau)|^q \rangle$ of wavelet coefficients $\varepsilon_r(\tau)$, where τ is location and r is resolution. For stationary measures one usually considers “wavelets of order 0” (averages), whereas for functions with multifractal increments one must use wavelets of order at least 1. One obtains $\hat{K}(q)$ as the slope of $\log(\hat{\mu}_r^q)$ against $\log(r)$ over a range of r . Negative moments are sensitive to measurement noise and quantization. For them, one typically uses only the local maxima of $|\varepsilon_r(\tau)|$ (modulus maxima methods). For the positive moments, we modify the standard estimator $\hat{K}(q)$ to significantly reduce its variance at the expense of a modest increase in the bias. This is done by separately estimating $K(q)$ from sub-records and averaging the results. For the negative moments, we show that the standard modulus maxima estimator is biased and, in the case of additive noise or quantization, is not applicable with wavelets of order 1 or higher. For these cases we propose alternative estimators. We also consider the fitting of parametric models of $K(q)$ and show how, by splitting the record into sub-records as indicated above, the accuracy of standard methods can be significantly improved.

1 Introduction

The property of multifractality may be viewed as a scale invariance condition whereby, when the support of a random process $X(t)$ is contracted by a factor $r \geq 1$ and its amplitude is multiplied by some random variable A_r , the resulting process is equivalent to $X(t)$. Hence $X(t)$ is multifractal if

$$X(t) = A_r X(rt), r \geq 1 \quad (1)$$

where equality is in the sense of distributions (Gupta and Waymire, 1990; Veneziano, 1999). This definition of multi-

fractality can be extended to random measures, random fields in spaces of higher dimension, and vector-valued processes (e.g. Meneveau et al., 1990; Schertzer and Lovejoy, 1991; Pflug et al., 1993; Veneziano, 1999). For example, for a random measure on the real line, Eq. (1) becomes

$$\varepsilon(\Omega) = A_r \varepsilon(r\Omega), r \geq 1 \quad (2)$$

where $\varepsilon(\Omega)$ is the average measure density in $\Omega \subset R$.

An important extension of Eqs. (1) and (2) (Veneziano, 1999) considers generalized random processes $X(\varphi)$, where $\varphi(t)$ is a test or kernel function (for example a wavelet). On generalized random functions, see Yaglom (1986). The process $X(\varphi)$ needs not have point values $X(t)$; when it does, $X(\varphi) = \int_{-\infty}^{\infty} \varphi(t) X(t) dt$. For $X(\varphi)$, the multifractal property is expressed as

$$X(\varphi_r) = A_r X(\varphi), r \geq 1 \quad (3)$$

where $\varphi_r(t) = r\varphi(rt)$ is a scaled and contracted version of $\varphi(t)$. What makes Eq. (3) important is that, by variously constraining the class of functions φ under which it holds (for example wavelets of certain orders), Eq. (3) includes a wide variety of multiscaling processes, such as processes with multifractal generalized increments of various orders and so-called multi-affine functions such as those considered by Vicsek and Barabasi (1991) and Benzi et al. (1993).

Multifractality is characterized by the distribution of the scaling factor A_r , but for inference it is more convenient to work with the moments of A_r or equivalently the moment scaling function $K(q) = \log_r E[A_r^q]$, which does not depend on r . The reason is that A_r is not observable, but for q such that the moments $\mu_t^q = E[|X(t)|^q]$, $\mu_\Omega^q = E[|\varepsilon(\Omega)|^q]$, or $\mu_\varphi^q = E[|X(\varphi)|^q]$ exist, one may estimate $K(q)$ using the relationships

$$\mu_{t/r}^q \propto r^{K(q)}, \quad \mu_{\Omega/r}^q \propto r^{K(q)}, \quad \mu_{\varphi_r}^q \propto r^{K(q)} \quad (4)$$

which hold for any given t , Ω or φ .



Correspondence to: P. Furcolo
(p.furcolo@unisa.it)

Equation (4) follows directly from Eqs. (1)–(3) and the scaling of the moments of A_r , but its use is constrained by three factors:

1. For some values of q , say $q \geq q^*$, $E[A_r^q]$ may be finite but the moments $\mu_{r,t}^q$, $\mu_{r,\Omega}^q$ or μ_φ^q may not exist. This is a well-known phenomenon (e.g. Mandelbrot, 1974; Schertzer and Lovejoy, 1987). The critical moment order q^* depends on the $K(q)$ function (Kahane and Peyriere, 1976). Similar problems of moment divergence are often encountered for $q < 0$;
2. Equation (4) holds for the theoretical moments. The empirical moments, even at very high resolutions r , may differ significantly from the theoretical moments and not satisfy Eq. (4). This is associated with the so-called linearization phenomenon, whereby the $K(q)$ function extracted from the empirical moments becomes linear in q above some moment order $q^+ < q^*$ (e.g. Ossiander and Waymire, 2000, 2002; Lashermes et al., 2004). In Sect. 2, we describe in some detail this phenomenon and the calculation of q^+ from $K(q)$;
3. The empirical moments may be heavily distorted by noise or other data artifacts (e.g. Harris et al., 1997).

Our objective is to evaluate and improve current estimators of $K(q)$ for positive or negative q , including cases when the data are exact, corrupted by noise, or degraded by quantization. For numerical illustration and performance assessment, we consider stationary non-negative multifractal measures in the unit interval. These include discrete multiplicative cascades (Mandelbrot's cascades) as well as stationary continuous-scaling measures (conserved universal multifractal processes). In either case, A_r is assumed to have lognormal distribution with unit mean. We use ε_n to denote the average measure density in a generic sub-interval of length 2^{-n} where n is the resolution level and $r=2^n$ is the resolution. The non-diverging moments of ε_n , $\mu_n^q = E[\varepsilon_n^q]$, satisfy $\mu_n^q \propto 2^{nK(q)}$.

We have obtained results also for some processes with multifractal increments, but results are similar to those for the discrete cascades and stationary measures mentioned above and are not presented. It should be noted that, although discrete multiplicative cascades have many limitations as models of physical phenomena, they are pedagogically very attractive due to their simplicity of construction and simulation and the fact that, unlike “continuous scaling” processes, they satisfy the scale invariance condition in Eq. (2) exactly (over the dyadic partition of the unit interval).

The function $K(q)$ is usually estimated in two steps (Harris et al., 1997; Lashermes et al., 2004). In the first step one calculates the empirical moments $\hat{\mu}_n^q = \langle \varepsilon_n^q \rangle$ for a range of n and q and for each q one estimates $\hat{K}(q)$ as the slope of the linear regression of $\log_2(\hat{\mu}_n^q)$ against n over a range $[n_{\min}, n_{\max}]$ of resolution levels. The maximum resolution

level n_{\max} is generally taken as the resolution level of the data, n_{dat} . This operation is performed without assuming any parametric form of $K(q)$; hence we refer to $\hat{K}(q)$ as a nonparametric estimator. One may generalize this basic estimator by replacing the above local averages with the absolute wavelet coefficients $\varepsilon_n = 2^n |\int \varphi(2^n t) X(dt)|$, where φ is a wavelet of order 0 or higher (the order is the lowest integer q such that $\int \varphi^q(t) dt \neq 0$). Hence a wavelet of order 0 has non-zero integral. Wavelets of order 0 are appropriate for the multifractal measures considered here. The use of higher-order wavelets (generalized differences of different order) poses problems for the negative moments, because the probability density of ε_n at zero becomes nonzero and at least the moments of order $q \leq -1$ diverge.

In the second step, one may use various criteria to fit a parametric model $K_{\text{par}}(q)$ to $\hat{K}(q)$. For example, one may fit $K_{\text{par}}(q)$ by least squares over a range $[q_{\min}, q_{\max}]$ of moment orders. In this case the estimator $\hat{K}(q)$ depends on $[n_{\min}, n_{\text{dat}}]$ and the estimator $\hat{K}_{\text{par}}(q)$ depends additionally on $[q_{\min}, q_{\max}]$.

The estimation of $K(q)$ is generally more challenging for negative than for positive moments. The reason is that the negative moments are sensitive to values of ε_n close to zero and therefore are strongly affected by additive noise, data quantization and other measurement artifacts. As mentioned above, for wavelets of order 1 there are problems with the estimation of $K(q)$ for $q \leq -1$, also when the data are not corrupted. In all these cases one must use special techniques that avoid low values of ε_n (e.g. Muzy et al., 1991; Bacry et al., 1993), but in general the accuracy of estimation remains lower than for the positive moments. This is why parametric models $K_{\text{par}}(q)$ are typically fitted to $\hat{K}(q)$ over a range of positive moments.

The bias and variance of the nonparametric estimator $\hat{K}(q)$ depend on the true $K(q)$ function, the length and resolution of the record, the accuracy of the data, the wavelet order used, the parameters $[n_{\min}, n_{\text{dat}}]$, and the moment order q . The same is true for the parametric estimator, with q replaced by the moment range $[q_{\min}, q_{\max}]$. We find how the bias and variance of various estimators vary with the above parameters when A_r in Eq. (2) has lognormal distribution, but our qualitative conclusions hold also for different distributions of A_r . In the lognormal case the variance of $\ln(A_r)$ depends on r as $\text{Var}[\ln(A_r)] = 2C_1 \ln(r)$ where $0 < C_1 < 1$ controls the variability of the measure and the function $K(q)$ is given by (Schertzer and Lovejoy, 1987)

$$K(q) = C_1(q^2 - q) \quad (5)$$

The paper is organized as follows. Section 2 reviews what is known about the performance of the nonparametric estimator $\hat{K}(q)$. Section 3 introduces and evaluates a broader class of estimators that aim at reducing the error variance and the RMS error. The idea is to increase the effective number of independent realizations by partitioning the available

record into a number of sub-records, then separately estimate $K(q)$ for each sub-record, and finally average the results. As we show in Sect. 3, this operation is effective in reducing the mean square error of the estimator and works well also for negative moments, if the data are error-free and one uses zero-order wavelets (e.g. local averages).

Section 4 focuses on the estimation of $K(q)$ for negative moments when the data are corrupted by errors or one uses wavelets of order 1. We show that techniques based on the wavelet transform modulus maxima (WTMM) are biased and suggest alternatives. The alternative methods have parameters that can be adjusted to focus on values of ε_n at different distances from zero, a feature that is useful to deal with data corrupted by different noise levels or with different degrees of quantization.

Issues of parametric estimation are examined in Sect. 5. We show how by using the nonparametric estimators introduced in Sect. 3 one can significantly improve the accuracy of standard parametric estimators. The main conclusions are summarized in Sect. 6.

2 The standard nonparametric estimator $\hat{K}(q)$

An interesting property of the nonparametric estimator $\hat{K}(q)$ is that with probability 1, as the resolution level of the observations $n_{\text{dat}} \rightarrow \infty$, $\hat{K}(q)$ becomes linear with some positive slope γ^+ above some order $q^+ > 1$ and with some negative slope γ^- below some order $q^- < 0$ (Ossiander and Waymire, 2000, 2002; Lashermes et al., 2004). The moment orders q^\pm and slopes γ^\pm are found geometrically by drawing the tangents to $K(q)$ with Y -intercept equal to -1 . γ^+ and γ^- are the slopes of these tangent lines and q^+ and q^- are the values of q at the points of tangency; see Fig. 1. Hence q^\pm and γ^\pm can be calculated analytically by solving

$$\begin{cases} q^\pm K'(q^\pm) - K(q^\pm) = -1 \\ \gamma^\pm = K'(q^\pm) \end{cases} \quad (6)$$

where K' is the derivative of K . For $K(q)$ in Eq. (5), one finds

$$q^\pm = \pm 1/\sqrt{C_1}, \quad \gamma^\pm = \pm 2\sqrt{C_1} - C_1 \quad (7)$$

Next we elaborate on the “straightening” of $\hat{K}(q)$ for $q > q^+$, but similar considerations apply to the negative moments for $q < q^-$.

The straightening of $\hat{K}(q)$ above q^+ is due to the non-observability of extreme singularities in a sample. As the resolution level $n \rightarrow \infty$, the ensemble moments $\mu_n^q = E[\varepsilon_n^q]$ are dominated by values of ε_n on the order of $2^{nK'(q)}$, but in a finite number s of realizations, as the level $n \rightarrow \infty$ these values occur with probability one for $q < q^+$ and with probability zero for $q > q^+$ (Ossiander and Waymire, 2000, 2002; Lashermes et al., 2004; Veneziano et al., 2006). It follows

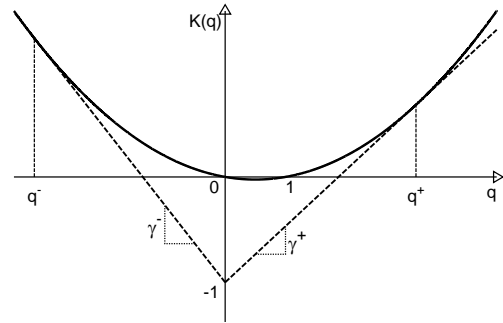


Fig. 1. Illustration of linearization of the estimator $\hat{K}(q)$. As $n_{\text{dat}} \rightarrow \infty$, $\hat{K}(q)$ becomes linear with probability 1 above q^+ and below q^- .

that with probability 1 the sample maximum of ε_n , $\varepsilon_{n,\text{max}}$, satisfies

$$\lim_{n \rightarrow \infty} \frac{\log_2(\varepsilon_{n,\text{max}})}{n} = \gamma^+ \quad (8)$$

These sample maxima control the empirical moments $\hat{\mu}_n^q$ for $q > q^+$, making them behave like $\hat{\mu}_n^q \sim 2^{n(q\gamma^+ - 1)}$ and causing the estimator $\hat{K}(q)$ to become linear with slope γ^+ for $q > q^+$ (with probability 1). As mentioned in the Introduction, this linearization phenomenon occurs at moment orders q^+ below the order q^* of moment divergence. Therefore, what is of concern in practice is linearization due to non-observability of high-order singularities not linearization from divergence of the moments.

The above applies as the maximum resolution level $n_{\text{dat}} \rightarrow \infty$. For finite n_{dat} , $\hat{K}(q)$ displays gradual straightening as q increases, becoming asymptotically linear as $q \rightarrow \infty$. The asymptotic slope for finite n_{dat} , say $\hat{\gamma}^+$, differs from γ^+ in Eq. (8) and is obtained from the sample maxima $\varepsilon_{n,\text{max}}$ as the slope of the linear regression of $\log_2(\varepsilon_{n,\text{max}})$ against n in the range $[n_{\text{min}}, n_{\text{dat}}]$. Transition to linearity might be considered to occur at the moment order \hat{q}^+ at which the sample maxima $\varepsilon_{n,\text{max}}$ contribute some fixed fraction of the empirical moments $\hat{\mu}_n^q$ (for alternative but essentially similar criteria, see Harris et al., 1997). The maxima $\varepsilon_{n,\text{max}}$ vary significantly from realization to realization (Harris et al., 1997; Lashermes et al., 2004); hence also the transition order \hat{q}^+ and the asymptotic slope $\hat{\gamma}^+$ are realization-dependent.

For n_{dat} finite, theoretical evaluation of the bias and variance of $\hat{K}(q)$ is difficult. Wendt and Abry (2006, 2007) and Wendt et al. (2007) estimated these quantities through asymptotic expansion or bootstrap, whereas Harris et al. (1997) used numerical simulation of discrete cascades to show how the bias and variance depend on n_{dat} (in Harris et al., 1997, n_{min} is close to zero and what is effectively varied is n_{dat}). The results confirm that for $q > 1$ the estimator $\hat{K}(q)$ is biased towards low values and the bias and variance increase as q increases or n_{dat} decreases.

The bias is contributed by two sources: one is the finiteness of the sample and the other is the dependence among the empirical moments estimated from the same sample at different scales. The first source of bias may be understood as follows. The empirical moments $\hat{\mu}_n^q$ are unbiased with nonzero variance. Since the logarithmic function is concave, the log moments $\log_2(\hat{\mu}_n^q)$, which are the quantities used in the regression that produces $\hat{K}(q)$, have negative bias. The bias tends to increase as the moment order q increases, causing the slope of the regression to be negatively biased. For $q < q^+$, $\hat{K}(q)$ becomes asymptotically unbiased as the number of realizations $s \rightarrow \infty$.

To understand the second source of bias of $\hat{K}(q)$, consider the maximally biased case when a single realization ($s=1$) is observed at level $n_{\text{dat}}=1$. In this case the data consist of just two values, $\varepsilon_{1,1}$ and $\varepsilon_{1,2}$, being the average measure densities in the first half and second half of the unit interval, and the estimator becomes

$$\hat{K}(q) = \log_2 \left(\frac{\varepsilon_{1,1}^q + \varepsilon_{1,2}^q}{2\varepsilon_0^q} \right) = \log_2 \left(\frac{\varepsilon'_{1,1}{}^q + \varepsilon'_{1,2}{}^q}{2} \right) = \log_2(\varepsilon_1^q) \quad (9)$$

where $\varepsilon_0 = 0.5(\varepsilon_{1,1} + \varepsilon_{1,2})$ is the average measure density in the unit interval and the prime sign indicates normalization by ε_0 . Normalization transforms the measure densities $\varepsilon_{1,1}$ and $\varepsilon_{1,2}$ into the partition coefficients $\varepsilon'_{1,1}$ and $\varepsilon'_{1,2}$ from level 0 to level 1. These partition coefficients are always between 0 and 2 and therefore have finite moments of all positive orders q . For A_r lognormal with parameter $C_1=0.1$, Fig. 2 compares the theoretical function $K(q)$ in Eq. (5) with the \log_2 of the moments of the one-step partition coefficient $\varepsilon'_1 = \varepsilon_1/\varepsilon_0$. For the latter, we first calculated the distribution of ε'_1 using a numerical procedure of the type in Veneziano and Furcolo (2003) and then obtained the moments $E[\varepsilon_1^q]$ numerically. Figure 2 shows that, for $q > 1$, $\log_2(E[\varepsilon_1^q])$ is significantly smaller than $K(q)$. The case $n_{\text{dat}}=1$ produces maximally biased estimators. As the next section shows, the bias decreases as n_{dat} increases.

3 Improved nonparametric estimation of $K(q)$ from defect-free data

With the main objective of reducing the error variance, we introduce a class of nonparametric estimators of $K(q)$ that includes the conventional estimator $\hat{K}(q)$ as a special case. The focus here is on the positive moments, but the estimators apply also to the negative moments if the data are defect-free and one uses order 0 wavelets.

Suppose that the available record consists of a single measure realization at some finite resolution level n_{dat} , as in Harris et al. (1997), and that the data are not contaminated by noise or quantized. The variance of $\hat{K}(q)$ is contributed in part by the strong correlation among the measure densities ε_n in different intervals at the same or different resolution

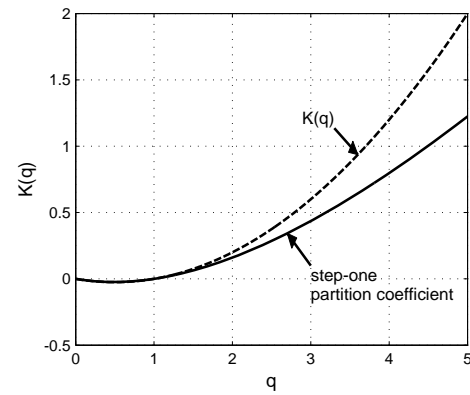


Fig. 2. Comparison of $K(q)$ in Eq. (5) for $C_1=0.1$ with the log-moments $\log_2(E[\varepsilon_1^q])$ of the step-one partition coefficient ε'_1 .

levels n . Therefore the variance would be reduced if one could increase the number of independent samples or reduce the correlations. To achieve this objective, we partition the available record into sub-records, each rooted in a tile at a fixed resolution level $n_o < n_{\text{dat}}$; see Fig. 3 for a schematic illustration. We find $\hat{K}_j(q)$ from each sub-record $j=1, \dots, 2^{n_o}$ using the conventional nonparametric estimator over a range of resolution levels $[n_{\text{min}}, n_{\text{dat}}]$ for some $n_{\text{min}} \geq n_o$, and finally average the results to obtain

$$\bar{K}(q) = 2^{-n_o} \sum_{j=1}^{2^{n_o}} \hat{K}_j(q) \quad (10)$$

For any given n_{dat} , the “average estimator” $\bar{K}(q)$ depends on (n_o, n_{min}) or equivalently $(n_o, \Delta n)$ where $\Delta n = n_{\text{min}} - n_o$. For $n_o=0$, $\bar{K}(q)$ reduces to the standard estimator $\hat{K}(q)$.

A reduction in the error variance is brought about by the fact that the sub-record estimators $\hat{K}_j(q)$ may be considered as mutually independent. To show the independence property for the case of discrete cascades, note that the standard estimator $\hat{K}(q)$ is unaffected by multiplication of the record by a positive constant. Hence $\hat{K}_j(q)$ does not vary if one divides all the average densities ε_n in the j -th sub-record by $\varepsilon_{n_o, j}$, the average measure density for that entire sub-record. Since in discrete cascades the normalized sub-records are independent, the estimators $\hat{K}_j(q)$ are also independent. For continuous multifractal processes, the sub-record estimators $\hat{K}_j(q)$ are serially correlated. However, we have found through simulation that the correlation is at most 0.04 at j -lag 1 and is otherwise very close to zero. Therefore, in practice the estimators $\hat{K}_j(q)$ for different j may be taken as independent.

The price one pays for the increase in the number of independent samples from 1 to 2^{n_o} is that the sub-records have a reduced resolution range $n_{\text{dat}} - n_o$ compared with the range n_{dat} of the original data, and this causes the bias to increase somewhat. This operation of sub-sampling and averaging is formally analogous to a variance-reduction technique used to estimate the power spectrum of stationary signals (Press et

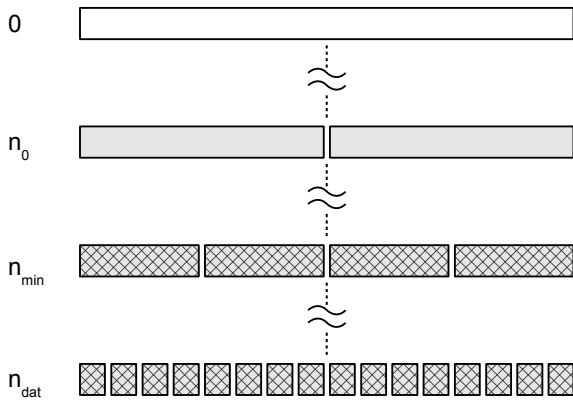


Fig. 3. Illustration of the resolution levels n_o, n_{\min} , and n_{dat} used by the average estimator $\bar{K}(q)$.

al., 1992). In spectral analysis, one may use other variance-reduction methods such as smoothing the empirical spectral density over a range of frequencies, but these alternative methods do not work well with $K(q)$ due to the high correlation of $\hat{K}(q)$ for different q .

Following the approach of Harris et al. (1997) for $\hat{K}(q)$, we estimate the bias and error variance of $\bar{K}(q)$ in Eq. (10) by simulating a large number of discrete cascade realizations in the unit interval. For $K(q)$ in Eq. (5), the bias, variance and root mean square error (RMSE) of $\bar{K}(q)$ depend on the multifractal parameter C_1 , the moment order q , the resolution level of the data n_{dat} and the parameters $(n_o, \Delta n)$ of the estimator. Figure 4 shows results for $C_1=0.1$, which for example is a plausible value for rainfall during rainstorms (Langousis and Veneziano, 2007), $q=-2, 2, 3$ and 4 , $n_{\text{dat}}=14$, and all possible combinations of $n_o=0(2)12$ and $\Delta n=0, \dots, 13-n_o$. The remaining number of levels, $14-n_o-\Delta n$, is the range used to fit the regressions whose slope defines the estimators $\hat{K}_j(q)$. Notice that for $C_1=0.1$, $q=4$ exceeds the moment order $q^+=\sqrt{10}=3.162$ above which $\bar{K}(q)$ tends to be linear. All results in Fig. 4 are based on 1000 fully dressed discrete-cascade simulations. Dressing is accomplished by first calculating the distribution F_Z of the dressing factor Z by the procedure of Veneziano and Furcolo (2003) and then generating independent samples $Z_i=F_Z^{-1}(U_i)$, where the U_i are independent variables with uniform distribution in the unit interval.

Each panel of Fig. 4 includes several plots, one for each value of n_o , whereas Δn varies along the horizontal axis from 0 to $13-n_o$. Therefore, the longer plots are associated with smaller values of n_o . The longest plots, for $n_o=0$, correspond to the traditional estimator $\hat{K}(q)$. In particular, the estimator with $n_o=\Delta n=0$, which is the one most frequently used in practice, is marked with a star.

The bias increases when either n_o increases or Δn decreases. The reason is that either of these changes makes the estimator $\hat{K}_j(q)$ rely more heavily on moments at reso-

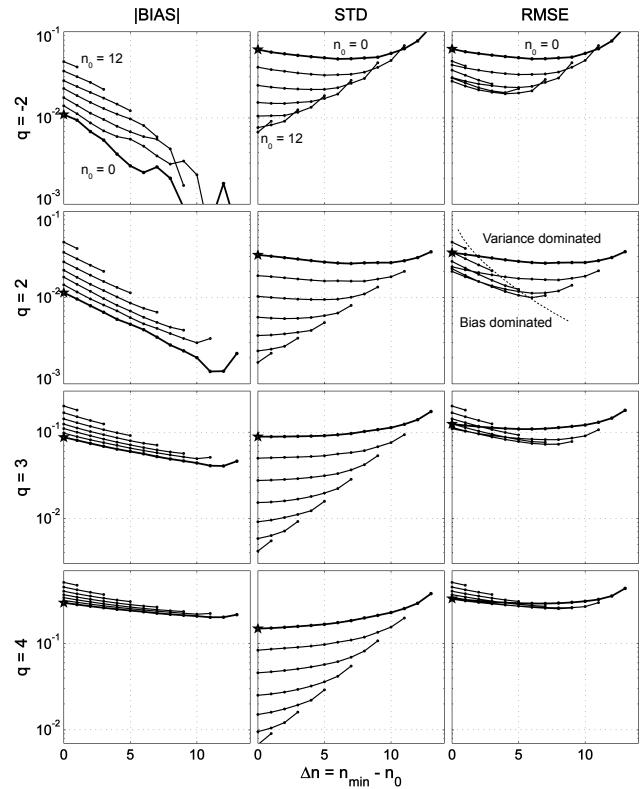


Fig. 4. Absolute bias, standard deviation, and RMS error of $\bar{K}(q)$ for moment orders $q=-2, 2, 3, 4$. A single measure realization with $C_1=0.1$ is observed at resolution level $n_{\text{dat}}=14$. The parameters of the estimator vary as $n_o+0(2)12$ and $\Delta n=0(1)13-n_o$. The different plots in each panel are for different n_o . The longer plots, for $n_o=0$, correspond to the standard estimator $\hat{K}(q)$. Results from 1000 independent simulations.

lution levels near the “root” of the j -th sub-record. As one approaches the root level, the traditional estimator reflects the moment scaling of the partition coefficients and is maximally biased; see Eq. (9) and Fig. 2. The estimator variance decreases with increasing n_o (due to the larger number 2^{n_o} of independent sub-records) but is insensitive to Δn , for the following reason. Lowering Δn increases the range of resolution levels used in the regressions, but the added levels correspond to low resolutions for which the moment estimates are less accurate. Hence it may happen that adding resolution levels increases the variance of the least-squares regression slope. We have investigated the use of weighted rather than ordinary least squares with weights that depend on the accuracy of the log-moment estimates and found that there is some gain in accuracy, but the gain hardly justifies the added complexity of the procedure. These qualitative traits of the bias and variance of $\bar{K}(q)$ hold for all positive and negative moment orders q .

The behavior of the RMSE is more complicated because the bias and variance dominate the RMSE over different ranges of n_o and Δn (see schematic illustration in one of

the panels of Fig. 4), which further depend somewhat on q . In general, the bias dominates for large n_o and small Δn , whereas the variance dominates for small n_o and large Δn . The importance of the bias increases as the moment order increases.

In Fig. 4, the multifractal parameter C_1 and the resolution level of the observations n_{dat} are kept fixed. We have made similar analyses with different values of these parameters. The results (not shown) are qualitatively similar, although as expected the bias and variance tend to increase as C_1 increases or n_{dat} decreases. The optimum value of n_o varies mainly with n_{dat} as $n_o \approx 0.5n_{\text{dat}} - 1$ and the optimum of Δn is such that $n_{\text{min}} = n_o + \Delta n \approx n_{\text{dat}} - 2$. We call $\bar{K}(q)$ with these parameter setting the optimal $\bar{K}(q)$ estimator. The solid lines in Fig. 5 show the relative RMS error of the optimal estimator (the RMS error divided by the actual value of $K(q)$) for different q , $C_1=0.05, 0.1, 0.15$ and 0.2 , and $n_{\text{dat}}=8, 10, 12$ and 14 . For comparison, the relative RMS errors of the standard estimator $\hat{K}(q)$ (this is the estimator $\bar{K}(q)$ for $n_{\text{min}}=n_o=0$) are shown as dashed lines. Increasing the multifractal parameter C_1 and decreasing the resolution of the observations have deleterious effects on both estimators. The same is true for increasing the moment order q . Note that the estimators tend to become linear above $q^+ = 1/\sqrt{C_1}$, which in Fig. 5 is marked by a vertical dashed line; hence in these high- q regions the RMS error is large. For $q=2$ and records of length around 4000 (n_{dat} about 12), the relative RMS error of the optimal estimator is 4.7%, 7.2%, 10.9% and 15.1% when $C_1=0.05, 0.1, 0.15$ and 0.2 , respectively. For the standard estimator, these relative errors are much higher (16.2%, 19.3%, 21.6% and 24.0%).

It is emphasized that we have optimized n_o and Δn for the case when only one multifractal realization is observed. In many practical cases, the available record includes many independent realizations. The higher sample size does not affect the bias of $\hat{K}(q)$ and $\bar{K}(q)$, but reduces the error variance and therefore makes small-bias, large-variance estimators like $\hat{K}(q)$ more appealing.

4 Improved nonparametric estimation of $K(q)$ from noisy or quantized data

Next we consider the estimation of $K(q)$ from data affected by imperfections like additive noise, multiplicative noise, or quantization. Additive noise and quantization influence mainly the estimates of $K(q)$ for negative moments; hence the primary focus of this section is on $q < 0$. This problem has been extensively discussed in the literature; see for example (Harris et al., 1997; Muzy et al., 1991; Bacry et al., 1993). The basic strategy is to use values of ε_n away from zero, while being careful that the operation of screening out the lower values does not alter the scaling of the moments.

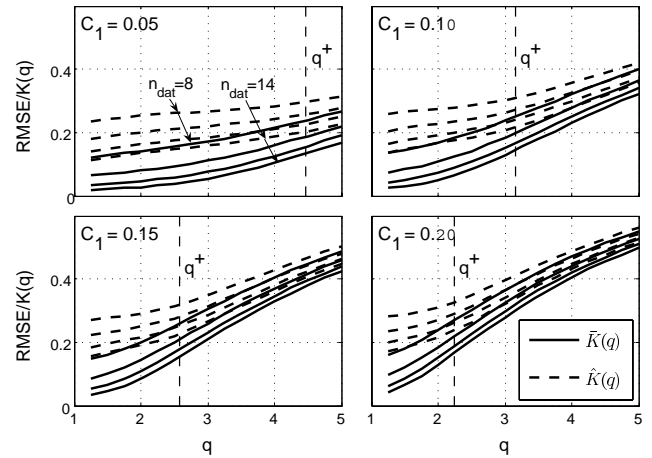


Fig. 5. Comparison of the relative RMS error of the optimal average estimator $\bar{K}(q)$ (solid lines) and the standard estimator $\hat{K}(q)$ (dashed lines). Different lines correspond to observations at different resolution levels $n_{\text{dat}}=8, 10, 12$ and 14 , increasing from above. Results are based on 1000 independent simulations.

4.1 Standard WTMM estimators

A popular technique (Muzy et al., 1991; Bacry et al., 1993; Struzik, 2000) is to first obtain the continuous wavelet transform

$$\varepsilon_n(\tau) = 2^n \int \phi[2^n(t - \tau)]X(dt) \quad (11)$$

where ϕ is a wavelet of order 0 or higher and then calculate the absolute moments using only the local maxima of $|\varepsilon_n(\tau)|$. The use of wavelets of order 1 or higher is problematic, because their absolute coefficients $|\varepsilon_n(\tau)|$ have nonzero probability density at zero; hence there is a higher probability of finding local maxima that are very close to zero than for wavelets of order 0. For stationary multifractal measures one can (and should) use 0-order wavelets, whereas for processes $X(t)$ with stationary multifractal increments this is not an option. Here we consider stationary multifractal measures $X(dt)$ in the unit interval.

In practice, $X(dt)$ is sampled at some finite resolution level n_{dat} and the wavelet coefficients in Eq. (11) are calculated for $2^{n_{\text{dat}}}$ values of τ , equally spaced between 0 and 1. We still refer to this as the continuous wavelet transform (CWT) of the signal. The procedure that derives $\hat{K}(q)$ from the local maxima of CWT is known as wavelet transform modulus maxima (WTMM) method. Several variants of the WTMM method exist, with the general aim to avoid local maxima that are still too close to zero and therefore are affected by data inaccuracies. The most popular method (“WTMM-with-sup”) connects the local maxima at different resolution levels n to form modulus maxima lines. Then each local maximum of $|\varepsilon_n(\tau)|$ is replaced with $\sup_{n \leq n' \leq n_{\text{dat}}} |\varepsilon_{n'}(\tau)|$

where the supremum is along the maxima line that passes through the local maximum point (Muzy et al., 1993, 1994; Arneodo et al., 1995). We consider explicitly only the WTMM method, but the qualitative results should extend to variants of that method, including WTMM-*with-sup*.

Like the ordinary estimator $\hat{K}(q)$, the WTMM estimator $\hat{K}_{\text{WTMM}}(q)$ depends on the true function $K(q)$ and the range of resolution levels $(n_{\min}, n_{\text{dat}})$ used in the regression. The WTMM estimator has two bias components. One component is similar to that of the standard estimator $\hat{K}(q)$ and is due to dependence among the values of ε_n (and hence of the local maxima) for different n . The other component is due to a symmetry breaking of the CWT transform: in order not to affect scaling, the relative amount of overlapping of the wavelet functions used for the calculation of two consecutive coefficients should not vary with the scale (similar concerns have been expressed by Lashermes et al., 2005). This means that the number of values of $\varepsilon_n(\tau)$ from which the local maxima are extracted should vary with n as 2^n , whereas in the WTMM method that number is constant and equal to $2^{n_{\text{dat}}}$. This lack of scale invariance in the identification of the local maxima alters the scaling of the moments and biases the estimate of $K(q)$; see below for a numerical illustration.

As noted above, another problem is that the WTMM procedure is ineffective in filtering out $\varepsilon_n(\tau)$ values that are close to zero. Alternatives like WTMM-*with-sup* are better in this regard, but suffer from other limitations: the additional symmetry breaking caused by the fact that one takes the supremum over a range of resolution levels $[n, n_{\text{dat}}]$ that narrows as n increases, the fact that the method does not apply when $|\varepsilon_{n'}|$ along the maxima lines tends to increase with increasing n' and is not effective in the opposite case, the more complicated implementation etc.

4.2 Unbiased WTMM estimators

In formulating alternatives to the above WTMM methods, one should pursue three objectives: avoid distortions of the $K(q)$ function, retain simplicity, and include tunable parameters to focus on values of $|\varepsilon_n(\tau)|$ that are at different “distances” from zero. This last feature allows one to obtain accurate estimates of $K(q)$ under different levels of data corruption, for example different additive noise levels or degrees of coarseness of the quantization. To avoid distortions, one should subject the wavelet transform to a filtering operation that preserves the scaling of the moments. For example, one can avoid the second source of bias of $\hat{K}_{\text{WTMM}}(q)$ mentioned above by using the discrete wavelet transform whereby one retains only the coefficients $\varepsilon_n(j2^{-n})$ for $j=1, \dots, 2^n$. The estimator of $K(q)$ that uses the local maxima of these coefficients is referred to as the scale-invariant WTMM estimator or simply the “scale-invariant estimator”. Since in this case one can extract the local maxima only at resolution levels $n \geq 2$, one must set $n_{\min} \geq 2$.

To reduce the number of local maxima that are close to zero, one could eliminate a given fraction of lowest local maxima within non-overlapping blocks of j of size 2^{2+k} , where k is a positive integer. This operation produces the “thinned scale-invariant estimator”.

A third possibility is to use the maximum of $|\varepsilon_n(j2^{-n})|$ in each non-overlapping block of 2^Δ values of j , where Δ is a positive integer acting as a tuning parameter (larger values of Δ let one focus on increasingly higher values of $|\varepsilon_n|$, although the number of maxima retained for moment analysis decreases). More in general, one could retain the largest $1 \leq m \leq 2^\Delta$ values of $\varepsilon_n(j2^{-n})$ within each 2^Δ block. Preliminary analyses have shown that the (m, Δ) estimator is rather insensitive to m and Δ and that retaining 1/8 to 1/4 of the original wavelet coefficients is about optimal. We refer to this as the (m, Δ) WTMM estimator or simply the “ (m, Δ) estimator”. In this case it must be $n_{\min} \geq \Delta$. For a schematic illustration of the continuous WTMM estimator, the scale-invariant estimator, the thinned scale-invariant estimator, and the (m, Δ) estimator, see Fig. 6.

4.3 Performance of various WTMM estimators

Below we assess and compare the performance of these four estimators under perfect and imperfect data conditions. Since WTMM estimators operate on the local maxima of $|\varepsilon_n|$ in time, it is desirable that the measure $X(dt)$ be stationary. This excludes discrete cascades. The measures we simulate are stationary lognormal multifractal measures on the unit interval, with $C_1=0.1$ and one-sided log-spectral density function $S_{\ln(X)}(\omega)=0.1\omega^{-1}$ over the scaling range. Simulation of $\ln(X)$ is performed in Fourier space: for each 2-fold increase of ω starting from $\omega_{\min}=\pi \cdot 2^{-2}$ and ending at $\omega_{\max}=\pi \cdot 2^{14}$, we use 2^4 waves with logarithmically spaced frequencies. We have used the simulations at two resolution levels: $n_{\text{dat}}=14$ and $n_{\text{dat}}=8$.

The generation of well-dressed continuous MF processes at high resolution (e.g. $n_{\text{dat}}=14$) can be computationally very demanding. To dress our simulations for $n_{\text{dat}}=14$ we used the following ad-hoc procedure. The signal was first (over)sampled at intervals of length $2^{-(14+3)}$ and then averaged in groups of 8 to achieve the desired resolution of 2^{14} . We call this the “bare” process. To generate a sample of “dressing factors”, we increased the resolution of the simulation in 10 000 intervals of size 2^{-14} by a factor of 2^{10} and took the ratios between the averages of the high resolution signal and the “bare” values as randomly simulated dressing factors. Finally, we obtained the dressed simulation by applying randomly selected “dressing factors” from the pool of 10 000 to the “bare” signal.

The parameters of the estimators have been set as follows. For the (m, Δ) WTMM estimator, in analyses with $n_{\text{dat}}=14$ we have set $m=4$ and $\Delta=5$ (meaning that we select the 4 largest values from each non-overlapping block of 32 coefficients) and in those with $n_{\text{dat}}=8$, we have set $m=1$

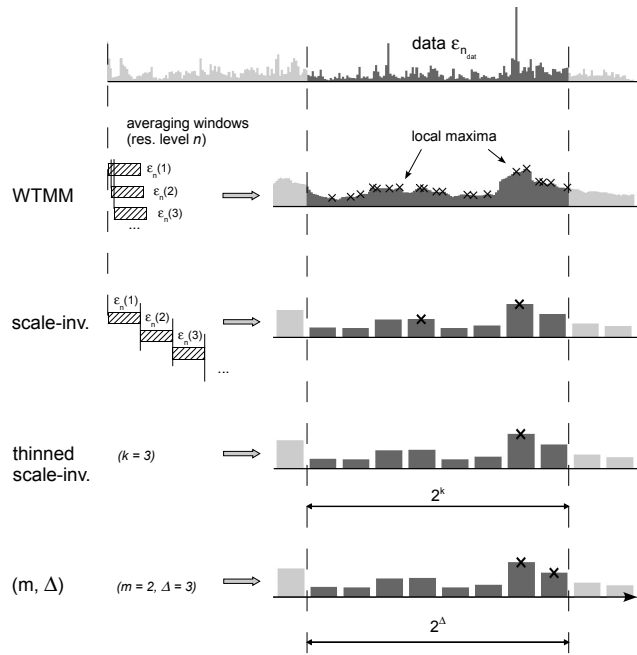


Fig. 6. Schematic illustration of four estimators of $K(q)$. Subsequent figures compare the performance of these estimators in the case of negative moments and either exact observations or variously corrupted data.

and $\Delta=3$ (thus retaining the largest value within each block of 8 coefficients). If the increments of $|\varepsilon_n(j2^{-n})|$ had independent signs, on average one fourth of the wavelet coefficients would be local maxima, whereas the above (m, Δ) estimators retain only one eighth of the values. Due to this higher selectivity, one may anticipate the (m, Δ) estimator to be less affected by data inaccuracies but have larger variance than the scale-invariant estimator. To compare the (m, Δ) estimator with an estimator that uses approximately the same number of wavelet coefficients, in the thinned scale-invariant estimator we eliminate 50% of the local maxima within non-overlapping j blocks of size 32. We set n_{\min} , the lower resolution level for the regression, to 2 for the WTMM and scale-invariant estimators and to 5 for the thinned scale-invariant and (m, Δ) estimators. Finally, we set the upper resolution level for the regression n_{\max} to the upper limit of the moment scaling range. This upper limit coincides in some cases with n_{dat} , the resolution level of the data, but in general depends on the estimator and the type and level of data corruption; see below. The analyzing wavelets are of order 0 (corresponding to local averaging) or order 1 (Haar wavelets).

We apply the above estimators to noiseless data and simulated records with various types and levels of corruption. The uncorrupted process is a continuous-scaling stationary lognormal multifractal measure in the unit interval. Data corruption is always applied at the observation level n_{dat} , as follows. In the case of additive noise, we add independent zero-mean Gaussian variables with standard deviation equal to ei-

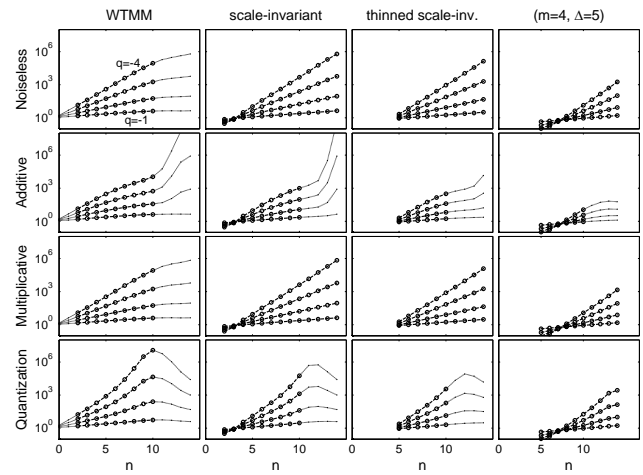


Fig. 7. Plots of the moments of order $q=-1, -2, -3$ and -4 for the estimators in Fig. 6. The plotted values are the averages of the log of the moments over 1000 simulations with $C_1=0.1$, observed at resolution level $n_{\text{dat}}=14$. The bottom three rows include “low levels” of additive noise, multiplicative noise or quantization; see text. The open circles indicate the resolution levels n used to estimate $K(q)$.

ther 5% or 10% of the standard deviation of $\varepsilon_{n_{\text{dat}}}$ (setting to zero any resulting negative value). In the case of multiplicative noise, we multiply the $\varepsilon_{n_{\text{dat}}}$ values by independent log-normal random variables with mean value 1 and log-standard deviation equal to either 20% or 50% of the log-standard deviation of the noiseless values. Finally, quantization is introduced by assuming a resolution of the recording instrument equal to either the 5-th or 20-th percentile of the actual $\varepsilon_{n_{\text{dat}}}$ values.

Figure 7 shows plots of the moments of order $q=-1, -2, -3$ and -4 for the four estimators when $C_1=0.1, n_{\text{dat}}=14$, and the data are noiseless (top row) or variously defective, with defects set at “low levels”. The analyzing wavelet is of order 0. The values plotted in Fig. 7 are the averages of the log of the moments obtained in 1000 independent simulations. Therefore the plots are affected by small statistical variability and can be used to assess systematic deviations from moment scaling. The open dots indicate the range (n_{\min}, n_{\max}) used in each case to estimate the $K(q)$ function. In general, there is a progression towards increased linearity of the plots as one moves to the right, from the continuous WTMM estimator to the (m, Δ) estimator. Of the various forms of data corruption, additive noise and quantization are the most disruptive of moment scaling, especially at the high resolutions (because the high-resolution values can be very close to zero). These qualitative observations hold also for other values of n_{dat} and for the higher levels of data corruption but, as we shall see later, not for analyzing wavelets of order 1.

Figures 8 and 9 compare the performance of the same four estimators, applied to a single measure realization with

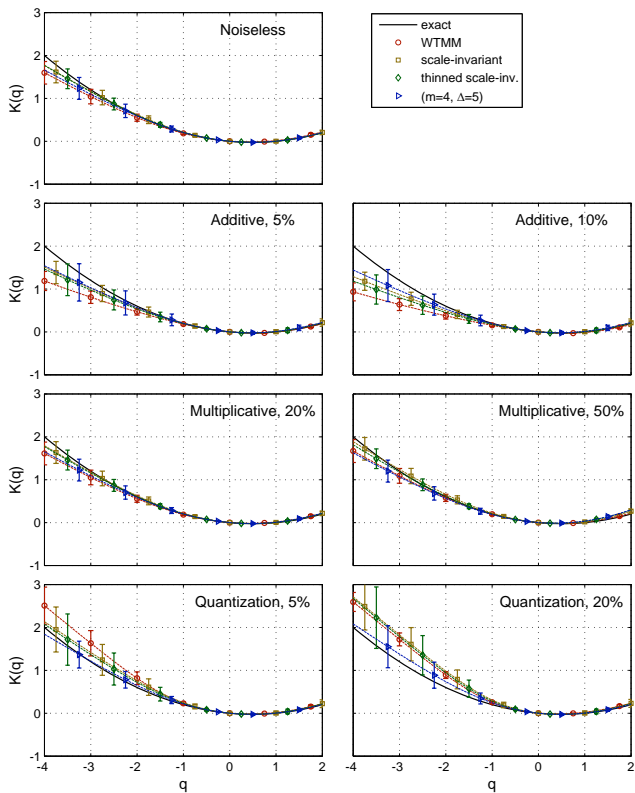


Fig. 8. $m \pm \sigma$ intervals for the estimators in Fig. 6, applied to a single measure realization with $C_1=0.1$ observed at resolution $n_{\text{dat}}=14$. The panel at the top is for noiseless data. The bottom three rows consider variously corrupted data. The type and level of corruption are indicated in each panel; see text for details. All results are based on 1000 measure simulations.

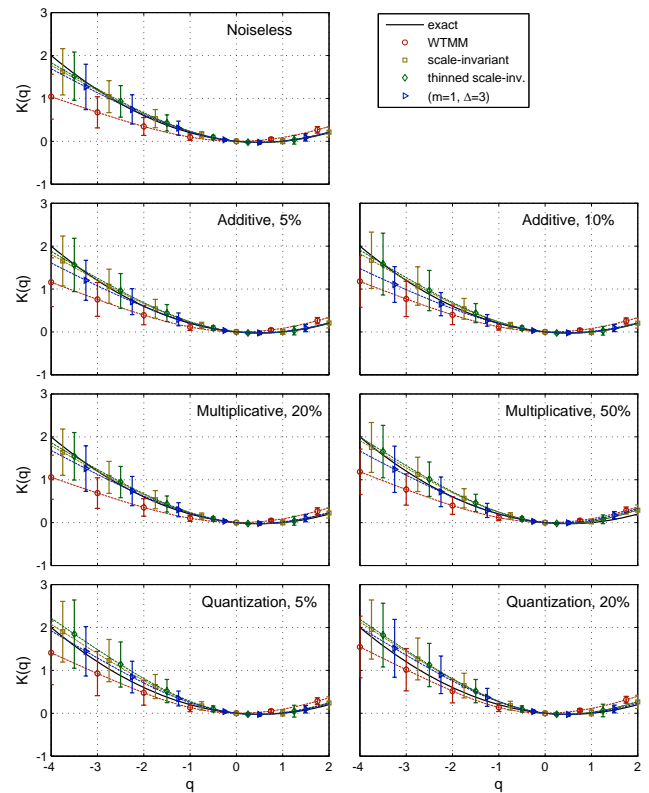


Fig. 9. Same as Fig. 8 for $n_{\text{dat}}=8$.

$C_1=0.1$, when observations are at resolution level $n_{\text{dat}}=14$ (Fig. 8) or resolution level $n_{\text{dat}}=8$ (Fig. 9). Also included are results for low and high levels of data corruption (left and right columns in each figure). In each panel, the continuous black line is the actual $K(q)$ function and the other lines are the mean values of the different estimators, obtained from 1000 independent simulations. The vertical bars are $m \pm \sigma$ intervals for the estimators. The panel at the top, for noiseless data, shows the maximum achievable performance of the estimators. The plots extend mainly over the negative moment orders. Results for positive moments up to order 2 mainly confirm that for $q > 0$ all the estimators perform well and are insensitive to data artifacts. In general, the continuous WTMM estimator has smaller variance but significantly higher bias than the other estimators. This is especially true when the record covers a narrow range of scales; see Fig. 9. The other estimators have generally comparable performance, with the (m, Δ) estimator having lower bias but slightly higher variance than the scale-invariant and thinned scale-invariant estimators. Multiplicative noise, even at the higher level, produces minimal deterioration in the performance of all estimators.

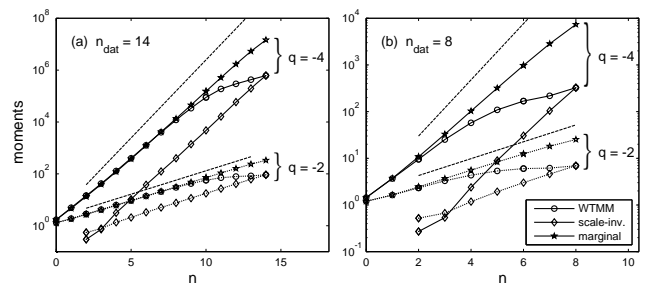


Fig. 10. Sources of bias for the continuous WTMM method. Comparison of the marginal moments of order $q=-2$ and -4 (star symbols) with the same moments of the modulus maxima used by the WTMM estimator (open circles) and scale-invariant estimator (diamonds). The reference straight lines have the theoretical slope $K(q)$ from Eq. (5) with $C_1=0.1$. The measures are observed at resolution level $n_{\text{dat}}=14$ or 8.

To understand the sources of bias of the continuous WTMM estimator especially in the case of low-resolution records, Fig. 10 shows moment plots of the type in Fig. 7 for $q=-2$ and -4 , comparing the theoretical slope $K(q)$ (thin straight lines) with the marginal moments of the ε_n values (stars; these are the moments used by the standard estimator $\hat{K}(q)$), the moments of the local maxima used by the continuous WTMM estimator (open circles), and the moments of the local maxima used by the scale-invariant estimator (dia-

monds). The resolution level of the observations, n_{dat} , is 14 in Fig. 10a and 8 in Fig. 10b. In all cases, the moments are log-averages over 1000 independent simulations with noiseless data. Therefore the moments for the continuous WTMM and scale-invariant estimators are the same as those in the first row of Fig. 7. It is interesting that, over a range of low resolutions, the continuous WTMM moments are practically identical to the marginal moments. This means that the distribution of the local maxima is virtually the same as the marginal distribution (due to the high degree of overlapping of the averaging windows in the continuous wavelet transform). In this scale range the continuous WTMM method displays scaling but is ineffective and retains the bias of the ordinary estimator. At higher resolutions, the WTMM moments do not scale due to the lack of scale invariance of the maximum-extraction procedure that was noted earlier in this section. By contrast, the moments of the scale-invariant method scale over all resolutions, with an exponent that is much closer to the theoretical slope $K(q)$. The lack of scaling of the WTMM method is especially severe when the data are at low resolution (Fig. 10b), explaining the high negative bias of that method in Fig. 9.

Results using Haar wavelets of order 1 are shown in Figs. 11 and 12, in both cases limited to $n_{\text{dat}}=14$. Figure 11 is analogous to Fig. 7 and Fig. 12 is analogous to Fig. 8. The main difference between using wavelets of order 0 or 1 is that in the latter case the absolute wavelet coefficients have nonzero probability density at zero. This makes the estimation of $K(q)$ for negative q more difficult, especially for methods that are ineffective at filtering out the low values. This is why in Figs. 11 and 12 the continuous WTMM method produces nonsensical results. The performance of the other methods is somewhat deteriorated relative to the order 0 case, although these methods do not break down.

5 Improved parametric estimation of $K(q)$

Harris et al. (1997) discuss at length the selection of the positive moment-order range $[q_{\min}, q_{\max}]$ such that the parametric estimator $\hat{K}_{\text{par}}(q)$, fitted by least-squares to $\hat{K}(q)$ inside that range, is minimally affected by the straightening of $\hat{K}(q)$ for large q . They set $q_{\max}=q^+$, the moment order above which $\hat{K}(q)$ is essentially linear, and suggest two specific ways to obtain estimates \hat{q}^+ from a realization. Here we consider the same problem in the context of the lognormal model in Eq. (5) (hence the problem of estimating C_1), with the difference that we replace $\hat{K}(q)$ with the average estimator $\bar{K}(q)$ in Eq. (10) and compare two different estimators of C_1 , one obtained from LS fitting of $\bar{K}(q)$ within some range $[q_{\min}, q_{\max}]$ and the other with C_1 estimated as the slope of $\bar{K}(q)$ at $q=1$. We refer to these as the LS estimator and Slope estimator of C_1 , respectively. Due to the superior performance of $\bar{K}(q)$, one may expect significant gains in accuracy from using $\bar{K}(q)$ instead of $\hat{K}(q)$.

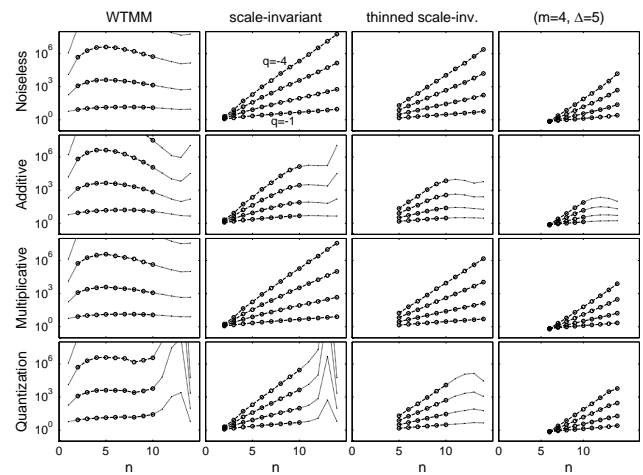


Fig. 11. Same as Fig. 7 for Haar wavelets of order 1.

In addition to choosing the range $[q_{\min}, q_{\max}]$ for the LS estimator, one must select the parameters n_o and n_{\min} of $\bar{K}(q)$. Three choices are especially meaningful:

$$\begin{cases} n_o = n_{\text{dat}} - 1 \\ n_{\min} = n_{\text{dat}} - 1 \end{cases}, \quad \text{for minimum variance} \quad (12a)$$

$$\begin{cases} n_o = 0.5n_{\text{dat}} - 1 \\ n_{\min} = n_{\text{dat}} - 2 \end{cases}, \quad \text{for minimum RMS error} \quad (12b)$$

$$\begin{cases} n_o = 0 \\ n_{\min} = 0 \end{cases}, \quad \text{reproducing } \hat{K}(q) \quad (12c)$$

For the range of moment orders used in the LS estimator, we set $q_{\min}=0$ and select q_{\max} such that the slope of the estimator $\bar{K}(q)$ at q_{\max} is a fixed fraction of the estimated asymptotic slope \hat{y}^+ . The latter slope is obtained by regressing the logs of the moments of the maximum measurements $\varepsilon_{n,\text{max}}$ at different resolution levels n , as explained in Sect. 2. In the numerical results shown below, the fraction is 0.5. Therefore our q_{\max} is about one half the value used in Harris et al. (1997).

Figure 13 compares results for different true values of C_1 (which varies along the horizontal axis), different resolutions of the data (different columns), different parameters of $\bar{K}(q)$ in Eq. (11) (different rows), and different ways of fitting the parametric function in Eq. (5) (solid lines for the LS estimator, dotted lines for the Slope estimator). The data are noiseless and the results are based on 1000 independent simulations. The vertical bars are $m \pm \sigma$ intervals. As expected, all estimators perform much better for $n_{\text{dat}}=14$, since the record is $2^6=64$ times longer than for $n_{\text{dat}}=8$. This is especially true for the parameter settings (a) and (b) in Eq. (12), since these settings take advantage of the higher resolution to split the record into a larger number of independent sub-records and by so doing reduce the estimator variance. The gain in accuracy is not as high in the case of the standard estimator

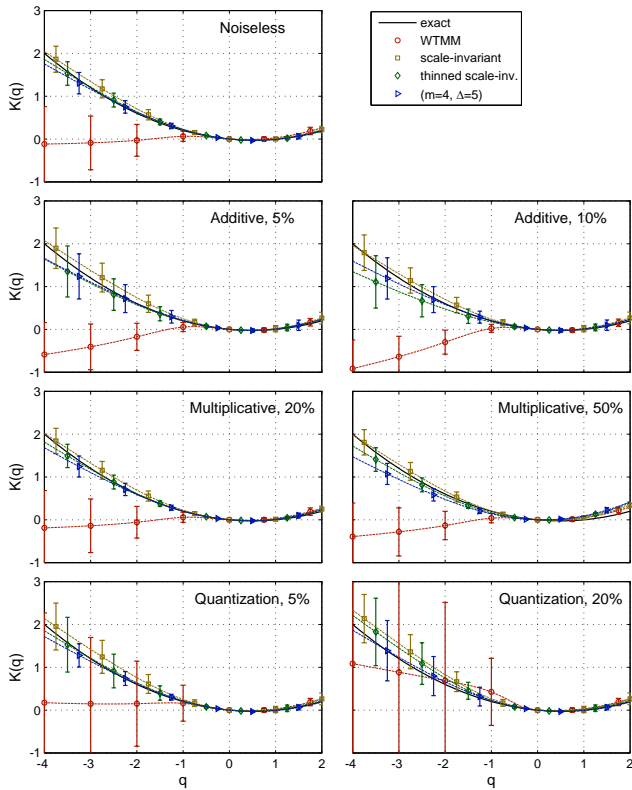


Fig. 12. Same as Fig. 8 for Haar wavelets of order 1.

(Eq. 12c). Also notice that the settings (b) and (c) produce nearly unbiased estimates, whereas there is a large bias in case (a). The LS and Slope estimators have very similar performance levels and in fact are nearly identical (for any given setting in Eq. 12, any C_1 and any n_{dat} , the correlation coefficient between them is about 0.95). Since one of the estimators is local and the other is global, this is an indication that the parametric estimators are highly efficient in using information from $\bar{K}(q)$. Due to its greater simplicity, the Slope estimator is more attractive, but the LS estimator is more general and can be used in cases when $K(q)$ has several unknown parameters.

Overall, in terms of the RMS error, estimator (b) performs the best. However, one can devise an even better estimator by correcting the minimum-variance estimator (a) for bias. In our experiment with lognormal MF processes, this is done by finding the value of C_1 for which the expected estimator $E[\hat{C}_1]$ equals the actual estimate. The top panels in Fig. 13 show that, for estimator (a), \hat{C}_1 underestimates C_1 by approximately 20%, independently of n_{dat} . Therefore, bias-corrected estimators are given by

$$\hat{C}_{1,\text{nobias}} = 1.20\hat{C}_{1(a)} \tag{13}$$

where $\hat{C}_{1(a)}$ is the estimator \hat{C}_1 under setting (a) in Eq. (12). These bias-corrected estimators are far superior to the standard parametric estimators (c).

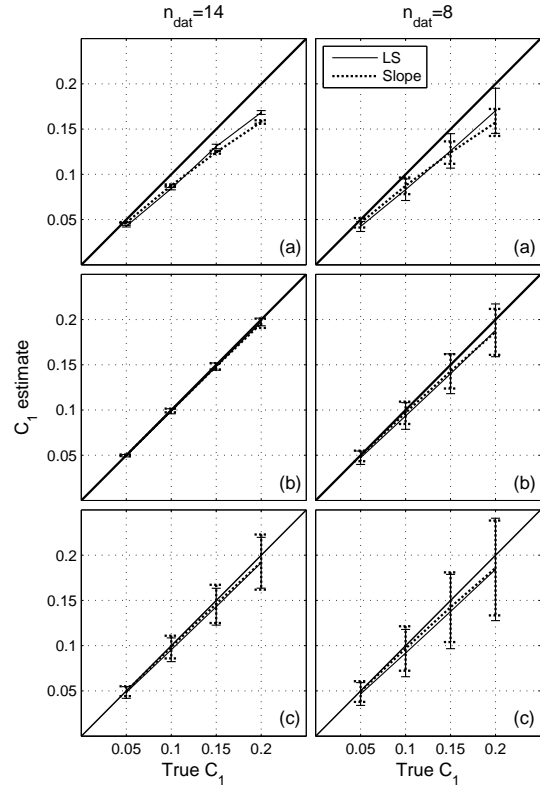


Fig. 13. Performance of various estimators of C_1 in Eq. (5), based on a single measure realization observed at resolution level $n_{\text{dat}}=14$ or 8. The continuous lines are for the LS estimator and the dotted lines are for the Slope estimator. The parameters n_o and n_{min} vary in different rows. The top (a) and middle (b) rows correspond to the estimators $\bar{K}(q)$ with minimum-variance ($n_o=n_{\text{min}}=n_{\text{dat}}-1$) and minimum RMSE ($n_o=0.5n_{\text{dat}}-1; n_{\text{min}}=n_{\text{dat}}-2$), respectively. The bottom row (c) is for the standard estimator $\hat{K}(q)$ ($\bar{K}(q)$ with $n_o=n_{\text{min}}=0$). Results are based on 1000 independent simulations.

6 Conclusions

We have revisited the classical problem of estimating the moment-scaling function $K(q)$ of multifractal signals from data. Specifically, we have focused on stationary multifractal measures (discrete-cascades or continuous-scaling), observed at different resolution levels with or without noise or quantization. For these measures and under these data conditions, we have sought to evaluate and improve existing non-parametric estimators of $K(q)$ for positive and negative moment orders q , as well as existing parametric estimators. All the estimators considered are based on the absolute wavelet coefficients ε_n at different resolution levels n calculated from the given record. We use wavelets of order 0 (in this case ε_n is just a local average) or wavelets of order 1 (of the Haar type). Numerical results have been obtained by simulating “lognormal measures” with $K(q)$ given by Eq. (5). In parametric estimation, we have assumed $K(q)$ to satisfy Eq. (5) with C_1 unknown. Our main results and conclusions

are as follows.

1. When the data are accurately recorded (no noise or quantization), one can improve the performance of the standard estimator $\hat{K}(q)$, which is given by the slope of the least squares regression of $\log_2\langle\varepsilon_n^q\rangle$ against n over the resolution range of the data. The improvement is obtained by partitioning the record into sub-records, calculating $\hat{K}(q)$ for each sub-record, and averaging the results. The resulting average estimator $\bar{K}(q)$ has two parameters, one that determines the number of sub-records and the other that controls the range of moment orders used to estimate $\hat{K}(q)$ from each sub-record. Based on extensive simulations, we have determined a simple rule to set these parameters to minimize the RMS error of $\bar{K}(q)$. Different settings minimize the variance and the bias. As Fig. 5 shows, $\bar{K}(q)$ is much more accurate than $\hat{K}(q)$.
2. For the positive moments, $\bar{K}(q)$ performs well also when the data are noisy or quantized and with analyzing wavelets of order 0 or 1. By contrast, for $q < 0$, $\bar{K}(q)$ should be used only with wavelets of order 0 and data that are perfectly recorded or are corrupted by multiplicative noise. The reason for these restrictions is that in all other cases the wavelet coefficients ε_n have nonzero probability density at 0 and their absolute moments of order $q \leq -1$ diverge. The standard way to deal with this problem is to calculate the scaling of the moments of the local maxima of $|\varepsilon_n|$ (WTMM method and variants). We have shown that the WTMM estimator has several drawbacks, also in the case of perfect data recording and wavelets of order 0: at high resolution levels n , the moments do not scale because the local maxima are extracted in a way that violates scale invariance. For smaller n , the method is ineffective in filtering out the values of ε_n that are close to zero, so that the negative modulus-maxima moments are virtually identical to the marginal wavelet transform moments. In the case of wavelets of order 1, the latter drawback is even more severe. We have suggested various alternatives to the WTMM method that are essentially unbiased and perform well also with noisy or quantized data and wavelets of either order.
3. For parametric estimation, we follow the standard procedure of least-squares fitting a parametric $K(q)$ function to nonparametric estimates over a range of positive moments. In our case we use nonparametric estimates of the $\bar{K}(q)$ type. For the range of moments, we consider two cases: a range $[0, q_{\max}]$ where q_{\max} is approximately one half of the moment order beyond which $\hat{K}(q)$ becomes linear (see text for details), or a range $[1^-, 1^+]$ around 1. The latter fitting procedure is suggested by the fact that, when $K(q)$ has the form in Eq. (5), the parameter C_1 is the slope of $K(q)$ at $q=1$.

We call the resulting estimators of C_1 the LS estimator and the Slope estimator, respectively. These estimators perform very similarly and, when used with the optimal nonparametric estimator $\bar{K}(q)$ in place of the traditional estimator $\hat{K}(q)$, significantly reduce the RMS error of \hat{C}_1 . Additional accuracy can be gained by using the estimator $\bar{K}(q)$ with minimum error variance (with the parameter settings in Eq. 12a) and then correcting \hat{C}_1 for bias using Eq. (13).

All the numerical work reported here is for stationary multifractal measures. In many cases, one needs to estimate the moment-scaling properties of processes $X(t)$ with multifractal increments (often referred to as multifractal processes). We have done some analyses also with these processes (not reported here). The main difference is that the analyzing wavelet must be of order at least 1. We have found that the qualitative results are similar to those when using the same wavelets with stationary measures.

Another context in which higher-order wavelets are used is the analysis of multifractal measures (or functions with multifractal increments) with additive polynomial trends (Lashermes and Foufoula-Georgiou, 2007). The non-scaling trend can be filtered out by choosing a wavelet of suitable order, but one has to contend with the issues of divergent negative moments and low local maxima mentioned above.

We conclude with a remark on the negative moments. There are many difficulties in accurately estimating $K(q)$ for $q < 0$. Moreover, one is often interested in the large values of ε_n , which are controlled by the scaling of the positive moments. Therefore, it is generally recommended that one limits the inference of $K(q)$ to the positive moments. If needed, scaling of the negative moments may be inferred through parametric extension of $K(q)$ from the positive moments.

Acknowledgements. This work was supported in part by project RISK (High Speed Rail) of the MIT-Portugal Program and in part by ex60% funds of the University of Salerno.

Edited by: T. Chang

Reviewed by: three anonymous referees

References

- Arneodo, A., Bacry, E., and Muzy, J. F.: The Thermodynamics of Fractals Revisited with Wavelets, *Physica A*, 213, 232–275, 1995.
- Bacry, E., Muzy, J. F., and Arneodo, A.: Singularity Spectrum of Fractal Signals from Wavelet Analysis: Exact Results, *J. Stat. Phys.*, 70, 635–674, 1993.
- Benzi, R., Biferale, L., Crisanti, A., Paladin, G., Vergassola, M., and Vulpiani, A.: A Random Process for the Construction of Multiaffine Fields, *Physica D*, 65, 352–358, 1993.
- Gupta, V. K. and Waymire, E. C.: Multiscaling Properties of Spatial Rainfall and River Flow Distributions, *J. Geophys. Res.*, 95(D3), 1999–2009, 1990.

- Harris, D., Seed, A., Menabde, M., and Austin, G.: Factors affecting multiscaling analysis of rainfall time series, *Nonlin. Processes Geophys.*, 4, 137–156, 1997, <http://www.nonlin-processes-geophys.net/4/137/1997/>.
- Kahane, J.-P. and Peyriere, J.: Sur Certaines Martingales de Benoit Mandelbrot, *Adv. Math.*, 22, 131–145, 1976.
- Langousis, A. and Veneziano, D.: Intensity-Duration-Frequency Curves from Scaling Representations of Rainfall, *Water Resour. Res.*, 43(2), W02422, doi:10.1029/2006WR005245, 2007.
- Lashermes, B. and Foufoula-Georgiou, E.: Area and Width Functions of River Networks: New Results on Multifractal Properties, *Water Resour. Res.*, 43, W09405, doi: 10.1029/2006WR005329, 2007.
- Lashermes, B., Jaffard, S., and Abry, P.: Wavelet Leader Based Multifractal Analysis, *IEEE International Conference on Acoustic, Speech and Signal Processing*, Philadelphia, PA, USA, IV, 161–164, 18–23 March 2005.
- Lashermes, B., Abry, P., and Chainais, P.: New Insights into the Estimation of Scaling Exponents, *Int. J. Wavelets, Multi.*, 2, 497–523, 2004.
- Mandelbrot, B. B.: Intermittent Turbulence in Self-similar Cascades: Divergence of High Moments and Dimension of the Carrier, *J. Fluid Mech.*, 62, 331–358, 1974.
- Meneveau, C., Sreenivasan, K. K., Kailasnath, P., and Fan, M. S.: Joint Multifractal Measures: Theory and Application to Turbulence, *Phys. Rev. A*, 41(2), 894–913, 1990.
- Muzy, J. F., Bacry, E., and Arneodo, A.: Wavelets and Multifractal Formalism for Singular Signals: Application to Turbulence Data, *Phys. Rev. Lett.*, 67, 3515–3518, 1991.
- Muzy, J. F., Bacry, E., and Arneodo, A.: Multifractal Formalism for Fractal Signals: The Structure-Function Approach Versus the Wavelet-Transform Modulus-Maxima Method, *Phys. Rev. E*, 47, 875–884, 1993.
- Muzy, J. F., Bacry, E., and Arneodo, A.: The Multifractal Formalism Revisited with Wavelets, *Int. J. Bifurcat. Chaos*, 4, 245–302, 1994.
- Ossiander, M. and Waymire, E.C.: Statistical Estimation for Multiplicative Cascades, *Ann. Stat.*, 28, 1533–1560, 2000.
- Ossiander, M. and Waymire, E. C.: On Estimation Theory for Multiplicative Cascades, *The Indian Journal of Statistics*, 64, 323–343, 2002.
- Pflug, K., Lovejoy, S., and Schertzer, D.: Generalized Scale Invariance, Differential Rotation and Cloud Texture: Analysis and Simulation, *J. Atmos. Sci.*, 50, 538–553, 1993.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edn., Cambridge University Press, 1992.
- Schertzer, D. and Lovejoy, S.: Physical Modeling and Analysis of Rain and Clouds by Anisotropic Scaling Multiplicative Processes, *J. Geophys. Res.*, 92(D8), 9693–9714, 1987.
- Schertzer, D. and Lovejoy, S.: Nonlinear Geodynamical Variability: Multiple Singularities, Universality and Observables, in: *Non-Linear Variability in Geophysics*, edited by: Schertzer, D. and Lovejoy, S., Kluwer, The Netherlands, 41–82, 1991.
- Struzik, Z. R.: Determining Local Singularity Strengths and Their Spectra with the Wavelet Transform, *Fractals*, 8(2), 163–179, 2000.
- Veneziano, D. and Furcolo, P.: Marginal Distribution of Stationary Multifractal Measures and Their Haar Wavelet Coefficients, *Fractals*, 11(3), 253–270, 2003.
- Veneziano, D., Langousis, A., and Furcolo, P.: Multifractality and Rainfall Extremes: A Review, *Water Resour. Res.*, 42, W06D15, doi:10.1029/2005WR004716, 2006.
- Veneziano, D.: Basic Properties and Characterization of Stochastically Self-similar Processes in R^D , *Fractals*, 7(1), 59–78, 1999.
- Vicsek, T. and Barabasi, A.-L.: Multiaffine Model for the Velocity Distribution in Fully Developed Turbulent Flows, *J. Phys. A-Math. Gen.*, 24, L845–851, 1991.
- Wendt, H. and Abry, P.: Bootstrap for Multifractal Analysis, *IEEE International Conference on Acoustic, Speech and Signal Processing*, Toulouse France, III, 812–815, 14–19 May 2006.
- Wendt, H. and Abry, P.: Multifractality Tests Using Bootstrapped Wavelet Leaders, *IEEE T. Signal Proces.*, 55(10), 4811–4820, doi:10.1109/TSP.2007.896269, 2007.
- Wendt, H., Abry, P., and Jaffard, S.: Bootstrap for Empirical Multifractal Analysis, *IEEE Signal Processing Magazine*, 38–48, July 2007.
- Yaglom, A. M.: *Correlation Theory of Stationary and Related Random Functions*, Basic Results, Springer-Verlag, Vol. 1, 526 pp., 1986.