Signal Processing 105 (2014) 243–257

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Scaling range automated selection for wavelet leader multifractal analysis

Roberto F. Leonarduzzi [a,b,*,1], María E. Torres [a,b], Patrice Abry [c]

[a] Laboratorio de Señales y Dinámicas No Lineales, Universidad Nacional de Entre Ríos, Entre Ríos, Argentina
[b] CONICET, Argentina
[c] Physics Dept. (CNRS UMR 5672), at École Normale Supérieure de Lyon, 69364 Lyon Cedex 07, France

ARTICLE INFO

Article history:
Received 22 December 2013
Received in revised form
26 April 2014
Accepted 4 June 2014
Available online 12 June 2014

Keywords:
Automated scaling range selection
Bootstrap
Wavelet leaders
Multifractal analysis

ABSTRACT

Scale invariance and multifractal analysis constitute paradigms nowadays widely used for real-world data characterization. In essence, they amount to assuming power law behaviors of well-chosen multiresolution quantities as functions of the analysis scale. The exponents of these power laws, the scaling exponents, are then measured and involved in classical signal processing tasks. Yet, the practical estimation of such exponents implies the selection of a range of scales where the power law behaviors hold, a difficult task with yet crucial impact on performance. In the present contribution, a nonparametric bootstrap based procedure is devised to achieve scaling range automated selection. It is shown to be effective and relevant in practice. Its performance, benefits and computational costs are assessed by means of Monte Carlo simulations. It is applied to synthetic multifractal processes and shown to yield robust and accurate estimation of multifractal parameters, despite various difficulties such as noise corruption or inter-subject variability. Finally, its potential is illustrated at work for the analysis of adult heart rate variability on a large database.

1. Motivation, context and contribution

1.1. Scale invariance and multifractal analysis

After Mandelbrot's seminal intuitions and contributions [1,2], the paradigm of *scale invariance*, also referred to as *scaling*, or sometimes *fractal*, has been used to model and/or analyze the temporal dynamics of many different real-world data sets produced by applications of very different natures, including biomedical [3], internet [4], physics [5], geophysics [6], finance [7], etc.

Irrespective of the details of the considered applications, the scale invariance concept amounts to assuming that no specific scale plays a dominant role in the temporal dynamics of the data, and that instead time scales, spread within a large range, are all equally contributing to data temporal dynamics. For such situations, data analysis should no longer consist in identifying preferred scales, but instead in essentially quantifying mechanisms that relate scales ones to the others. Assessing scale invariance requires the use of multiresolution quantities, $T_X(a,t)$, i.e., quantities that depend jointly on time (or space) and scale.

* Corresponding author. Postal address: Facultad de Ingeniería – UNER, C.C. 47 – Suc. 3 – 3100 Paraná, Entre Ríos, Argentina.
Tel.: +54 343 4975100x122.
E-mail addresses: rleonarduzzi@bioingenieria.edu.ar (R.F. Leonarduzzi), metorres@santafe-conicet.gov.ar (M.E. Torres), patrice.abry@ens-lyon.fr (P. Abry).

[1] Work supported by the French-Argentine CNRS-CONICET Grant # 112432 2011-2012, the ANPCyT-UNER PICT Grant # 433 2008, ANPCyT-CONICET PICT Grant # 2954 2012, and the ANR AMATIS Grant # 112432, 2010-2014.

http://dx.doi.org/10.1016/j.sigpro.2014.06.002
0165-1684/

Classical choices are increments, oscillations, or the nowadays commonly used wavelet coefficients. In practice, scale invariance can be evidenced and measured via a power-law (or algebraic) behavior of the time averages of $T_X(a,t)$, with respect to analysis scales a, over a large range of scales,

$$\frac{1}{n_a}\sum_k |T_X(a,k)|^q \simeq C_q a^{\zeta(q)}, \quad a_m \le a \le a_M, \ \frac{a_M}{a_m} \gg 1, \tag{1}$$

where n_a denote the number of such $T_X(a,k)$ available at scale a and the $\zeta(q)$ are usually termed the scaling exponents. These scaling exponents are classically used to analyze, characterize and classify signals or images (cf. e.g., [8,9]). More recently, multifractal analysis (cf. e.g., [10,11]), that aims at quantifying the fluctuations of regularity along time (or space) via the so-called multifractal spectrum, has received considerable interest in signal and image processing and is now considered a standard analysis tool. A recent and powerful formulation of multifractal analysis relies on a choice of specific multiresolution quantities, referred to as wavelet leaders, and is largely used in the sequel.

1.2. Scaling range selection

The practical use of the concept of scale invariance essentially amounts to estimating the scaling exponents $\zeta(q)$. Whatever the estimation procedure, Eq. (1) above clearly shows that practical estimation strongly relies on the choice of the range of scales, from now on referred to as the *scaling range*, where the power-law behavior holds. While the theoretical assumption that data X are exactly self-similar would imply an infinite scaling range (i.e., $a_m \to 0$ and $a_M \to +\infty$), in practice, the scaling range must often be considered limited, which may stem from many different causes. At the theoretical level, models used to describe data often yield asymptotic only power law behaviors, as in Eq. (1). For instance, multiplicative construction (underlying multifractal models) implies $a_m \to 0$ [10], while Long Memory models implies $a_M \to +\infty$ [12]. At the practical level, finite scaling range may result from physical (or physiological) mechanisms whose dynamics involve a large yet bounded range of scales, while other competing mechanisms may become dominant at finer or coarser scales (e.g., dissipation in turbulence [5], beat-to-beat nature in heart rate variability [13]). Also, data are digitalized at a given rate, with sampling devices necessarily destroying the power law behaviors at fine scales. Along the same line, the necessarily finite duration of recordings is likely to introduce power-law cut-off at coarse scales. Additionally, noise can be superimposed to truly scaling data, often leading to a substantial narrowing of the range of scales where scale invariance can actually be observed. These different mechanisms thus imply that scale invariance in practice necessarily holds only within a potentially large but finite range of scales, bounded below and above by lower and upper cutoffs. Further, purely from a performance perspective, estimation of the scaling exponents $\zeta(q)$ requires a careful selection of the range of scales, where estimation should be performed. Essentially, that selection is driven by a classical bias-variance trade-off: A large scaling range yields a low variance at the risk of bias, due to the often asymptotic nature of scale invariance; a narrow range, centered over scales where power law holds, reduces bias but at the price of an increased variance.

1.3. Related works

Though most practitioners are perfectly aware both of the crucial impact of scaling range selection on estimation performance and of the difficulties an objective and automated selection raise, this issue remains barely addressed. Essentially, scaling range can be selected either from *fundamental* arguments related to the physics (the physiology, etc.) underlying the data at hand (e.g., Kolmogorov dissipation scale in hydrodynamic turbulence [5], sympathetic–parasympathetic frequency band split in heart rate variability [14]), or from empirical data analysis. In this latter perspective, visual inspection and empirical experience remain the dominant practice amongst practitioners. This is however obviously a tedious and error prone procedure, notably for large databases, where each signal needs to be inspected and where noise level corruption is likely to vary individually from one signal to another, a very common observation in biomedical data notably. Amongst the rare attempts to address the issue in an automated way, χ^2-statistics and F-statistics based procedures were devised and studied in [15] and [16] respectively, to obtain the lower cutoff scale in the analysis of Gaussian long memory processes. In [17], attempts were made to relax the Gaussian assumption, relying on the use of Theil's inequality coefficient. Yet, those approaches do not straightforwardly nor relevantly extend to multifractal analysis, toward which we concentrate in the present contribution, where Gaussian assumptions do not a priori constitute valid approximations.

1.4. Goals, contributions and outline

In this context, the present contribution aims at proposing and assessing a practically effective procedure for the automated selection of the scaling range in the general wavelet leader framework for empirical multifractal analysis, thus not assuming a priori Gaussianity of the multiresolution quantities actually used [8,11]. To that end, a short introduction for empirical and practical multifractal analysis is recalled in Section 2. Inspired from the use of the χ^2-statistics in [15] and of the bootstrap framework developed in [8], a bootstrap-based procedure is motivated and constructed in Section 3 that permits an automated scaling range selection. Its performance is assessed by means of Monte Carlo simulations based on synthetic multifractal processes according to the protocol detailed in Section 4.1. Performance is reported and discussed with respect to optimal mean square error, data length and trade-off between estimation and analysis, for perfect multifractal processes in Section 4.2. Performance and robustness are further evaluated against noise corrupted multifractal processes (cf. Section 4.3), or multifractal processes with upper cutoff (cf. Section 4.4), or multifractal processes suffering from both corruptions at fine and

coarse scales, with in addition random amplitude effects (cf. Section 4.5). To finish with, the procedure is illustrated at work for adult heart rate variability analysis, and applied to a large database (cf. Section 5). A MATLAB procedure, designed by ourselves, will be made publicly available from our webpages at the time of publication.

2. Scale invariance and multifractal analysis

2.1. Wavelet leaders

Let $\psi_0(t)$ denote a compact support mother wavelet, characterized by its number of *vanishing moments*, a positive integer N_ψ such that $\int_{\mathbb{R}} t^k \psi_0(t)\,dt = 0$, $\forall k = 0, 1, \ldots, N_\psi - 1$ and $\int_{\mathbb{R}} t^{N_\psi}\,\psi_0(t)\,dt \neq 0$. $\psi_0(t)$ is chosen such that the set $\{\psi_{j,k}(t) = 2^{-j/2}\psi_0(2^{-j}t - k),\ j \in \mathbb{N}, k \in \mathbb{N}\}$ forms an orthonormal basis of $L^2(\mathbb{R})$. The wavelet coefficients of a signal X are defined as $d_X(j,k) = 2^{-j/2} \int_{\mathbb{R}} X(t)\psi_{j,k}(t)\,dt$ (note the use of the nonconventional L^1 norm that better matches multifractal analysis). For a detailed introduction to wavelet analysis, the reader is referred to e.g., [18].

Let further $\lambda_{j,k} = [k2^j, (k+1)2^j)$ denote *dyadic intervals* and $3\lambda_{j,k} = \bigcup_{m \in \{-1,0,1\}} \lambda_{j,k+m}$, the concatenation of 3 such intervals. The wavelet Leader $L_X(j,k)$ is defined from wavelet coefficients as a local supremum, taken within a narrow time neighborhood of $t = 2^j k$ at any finer scales $2^{j'} < 2^j$, [11,8]: $L_X(j,k) = \sup_{\lambda' \subset 3\lambda} |d_{\lambda'}|$.

2.2. Empirical multifractal analysis

In practice, multifractal analysis amounts to forming the so-called structure functions, corresponding to time-averages at given scales, and assuming that they show power-law behaviors against scales 2^j:

$$S(j,q) = 1/n_j \sum_k L_X^q(j,k) \simeq C_L(q) 2^{j\zeta(q)}. \tag{2}$$

It is now well known that a Legendre transform, applied to the function $\zeta(q)$, for both positive and negative q's, yields (an upper bound of) the multifractal spectrum $D(h)$, cf. e.g., [10,11]. The function $D(h)$ consists of a global and geometrical description of the fluctuations along time of local regularity h of the sample path of X.

To avoid the computation of the function $\zeta(q)$ for all q's, it has been proposed to make use of a polynomial expansion [8,19]: $\zeta(q) = \sum_{p \geq 1} c_p (q^p/p!)$. Interestingly, it was shown that the coefficients c_p can be related to the cumulants of order p, $C_p(j)$, of the log-leader $\ln L_X(j,k)$ according to

$$C_p(j) = c_{0,p} + c_p \ln 2^j, \quad \forall p \geq 1. \tag{3}$$

Often, using only the first two cumulants yields a satisfactory approximation for both $\zeta(q) \simeq c_1 q + c_2 q^2/2$ and $D(h) \simeq 1 + (h - c_1)^2/(2c_2)$.

2.3. Estimation and scaling range selection

As detailed in [8], estimation of the $\zeta(q)$ (resp., c_p) is commonly performed by linear regressions of $\log_2 S(j,k)$ (resp., $(\log_2 e) C_p(j)$) versus $\log_2 2^j = j$ (cf. Eq. (4) below). The crucial step in actually performing these linear regressions consists in selecting the range of scales across which it should be carried out so as to obtain the best performance estimates. This range selection is the core issue of the next section.

3. Automated scaling range selection

3.1. Multifractal parameter estimation by linear regressions

To discuss the estimation of the $\zeta(q)$'s and the c_p's within a single framework, θ will from now on denote either $\zeta(q)$ or c_p, and $M(m,j)$ will represent either $\log_2 S(m,j)$ or $(\log_2 e) C_m(j)$. Eqs. (2) and (3) suggest the generic behavior: $M(m,j) = \theta j + \gamma$. This leads to estimating θ (and γ) by linear regressions written as

$$\hat{\theta}(m,\underline{j}) = \sum_{j \in \underline{j}} w_j M(m,j), \quad \hat{\gamma}(m,\underline{j}) = \sum_{j \in \underline{j}} v_j M(m,j), \tag{4}$$

where $\underline{j}$ denotes the set of octaves over which the regression is performed. The weights w_j and v_j can be selected to achieve weighted or nonweighted linear regressions (cf. e.g., [4,8]).

3.2. Scaling range selection

Motivated by the most often encountered issues in processing real-world data, it is assumed, in the present contribution, that the set $\mathcal{V}$ of valid regression ranges $\underline{j}$, forms a continuum of contiguous octaves: $\underline{j} = [j_1, j_2]$, and $\mathcal{V} = \{(j_1, j_2) : 1 \leq j_1 < j_2 \leq J,\ j_1 < j_2 - 1,\ j_i \in \mathbb{N},\ i = 1, 2\}$, where J is the number of octaves actually available. Further, the restriction $j_1 < j_2 - 1$ ensures that at least three octaves are involved in the regression, thus requiring scale invariance to hold at least across two octaves.

Following the approach developed in [15], the selection of the scaling range where linear regression is to be performed relies on the use of a standard goodness-of-fit measure in a least-square sense:

$$R(m,\underline{j}) = \sum_{j \in \underline{j}} [M(m,j) - (\hat{\theta}(m,\underline{j}) j + \hat{\gamma}(m,\underline{j}))]^2. \tag{5}$$

In [15], it is argued that when properly normalized (by the variances $\sigma^2_{m,j}$ of the $M(m,j)$) and when applied to $m = 2$, for the wavelet coefficients of Gaussian processes, $R(m = 2, \underline{j})$ follows a χ^2 distribution with $j_2 - j_1 + 1$ degrees of freedom. Here, because multifractal processes are considered, that are usually not Gaussian, because R is computed from wavelet leaders, and because m is not restricted to 2, there is no reason why $R(m,\underline{j})$, even if properly normalized by the variances of $M(m,j)$, should follow a χ^2 distribution. Instead, let $F_{m,\underline{j}}$ denote the a priori unknown cumulative distribution functions (CDF) of $R(m,\underline{j})$. Then, the quantity

$$\Lambda(m,\underline{j}) = 1 - F_{m,\underline{j}}(R(m,\underline{j})), \tag{6}$$

provides practitioners with a convenient measure of the goodness-of-fit of $M(m,\underline{j})$ with $\hat{\theta}(m,\underline{j}) j + \hat{\gamma}(m,\underline{j})$ across the range of scales $\underline{j}$.

As discussed in Section 1, multifractal analysis requires, in theory, that the scaling range should be the same for all

m. Therefore, the optimal scaling range $\underline{j}^D$ can be defined as the one that maximizes $\Lambda(m,\underline{j})$ on average across m:

$$\underline{j}^D = \arg\max_{\underline{j}} \sum_m \Lambda(m,\underline{j}), \tag{7}$$

where summation is performed over the selected statistical orders m. However, in practice and with real-world data, either because of noise issues or because data only approximately follows scale invariance, it can prove fruitful to estimate the scaling ranges independently for c_1 and c_2 and analyze the extent to which they match (cf. Section 5). This can be easily achieved by measuring $\underline{j}^D$ independently for different m's:

$$\underline{j}^D_m = \arg\max_{\underline{j}} \Lambda(m,\underline{j}). \tag{8}$$

The use of Eq. (7) or (8) implies however the explicit calculation of $\Lambda(m,\underline{j})$, as in Eq. (6), which, in turn, implies that distributions $F_{m,\underline{j}}$ must be estimated. The combined use of wavelet leaders, consisting on a nonlinear transform of wavelet coefficients, of various statistical orders m, and the fact that data may a priori be not Gaussian, preclude the theoretical derivation of $F_{m,\underline{j}}$ for all m and $\underline{j}$. Instead, a nonparametric bootstrap procedure in the time-scale domain enables an efficient estimation, as detailed in the next section.

3.3. Bootstrap-based estimation of $F_{m,\underline{j}}$ and $\Lambda(m,\underline{j})$

3.3.1. Nonparametric block-bootstrap procedure

Following [8], where a bootstrap procedure was devised in the time-scale domain to construct confidence intervals for multifractal parameter estimation or for hypothesis testing, a bootstrap based estimation of the $F_{m,\underline{j}}$ is proposed here.

Nonparametric bootstrap consists of a now classical drawing with replacement procedure [20,21]. When data show dependencies, the drawing with replacement procedure must be applied to blocks of data [20,21]. Because the $L_X(j,k)$ are dependent quantities both along time k and scale 2^j, the drawing with replacement procedure must be applied to time-scale strips.

3.3.2. Bootstrap-based estimation

A large number, B, of time-scale block bootstrap surrogates, $L_X^{*,b}(j,k)$, $b=1,\ldots,B$, are drawn with replacement. For each b, from the $L_X^{*,b}(j,k)$, the $M^{*,b}(m,j)$, the $(\hat{\theta}^{*,b}(m,j), \hat{\gamma}^{*,b}(m,j))$ (cf. Eq. (4)) and the $R^{*,b}(m,\underline{j})$ (cf. Eq. (5)) are computed. These bootstrap surrogates are used to produce bootstrap-based estimates for $F_{m,\underline{j}}$ and Λ:

$$\hat{F}_{m,\underline{j}}(x) = \frac{\#\{b: R^{*,b}(m,\underline{j}) < x\}}{B}, \quad \hat{\Lambda}(m,\underline{j}) = 1 - \hat{F}_{m,\underline{j}}(R(m,\underline{j})). \tag{9}$$

The scaling range selection finally stems from the use of Eq. (7) or (8) with $\hat{\Lambda}(m,\underline{j})$.

Table 1
Proposed algorithm: Pseudocode for the proposed algorithm.

Begin algorithm
Compute $L_X(j,k)$.
for j,m **do**
  Compute $M(m,j)$.
end for
for $\underline{j} \in \mathcal{V}$ **do**
  Estimate $\hat{\theta}(m,\underline{j})$, $\hat{\gamma}(m,\underline{j})$ (Eq. (4)).
  Compute $R(m,\underline{j})$ (Eq. (5)).
end for
BEGIN PARALLEL REGION
for b in $1,\ldots,B$ **do**
  Bootstrap $L_X^{*,b} \leftarrow L_X$.
  for j,m **do**
    Compute $M^{*,b}(m,j)$.
  end for
  for $\underline{j} \in \mathcal{V}$ **do**
    Estimate $\hat{\theta}^{*,b}(m,\underline{j})$, $\hat{\gamma}^{*,b}(m,\underline{j})$ (Eq. (4)).
    Compute $R^{*,b}(m,\underline{j})$ (Eq. (5)).
  end for
end for
END PARALLEL REGION
for $\underline{j} \in \mathcal{V}$ **do**
  Estimate $\hat{F}_{m,\underline{j}}$ from $\{R^{*,b}(m,\underline{j})\}_{b=1,\ldots,B}$ (Eq. (9)).
  Compute $\hat{\Lambda}(m,\underline{j})$ (Eq. (9)).
end for
Solve $\underline{j}^D \leftarrow \arg\max_{\underline{j}} \sum_m \Lambda(m,\underline{j})$ (Eq. (7)).
Output $\underline{j}^D$, $\hat{\theta}(m,\underline{j}^D)$, $\hat{\gamma}(m,\underline{j}^D)$
End algorithm

3.4. Algorithm for scaling range selection and estimation

The overall procedure for the automated scaling range selection, and multifractal parameter estimation, can be summarized in the algorithm in Table 1.

The computational cost of the sole procedure producing estimates of θ is essentially that of the computation of a Discrete Wavelet Transform, in $O(N)$, plus the computation of leaders (negligible) and of linear regressions, in $O(J)$, where $J=\log_2(N)$. For the scaling range selection procedure, there are three extra costs. First, linear regressions yielding estimates need to be computed for each bootstrap resample, i.e., B times. Second, linear regressions need to be repeated for all $\underline{j} \in \mathcal{V}$, whose size is of the order of J^2, yielding a partial cost in $O(J^3)$. Third, the optimization problem in Eq. (7) or (8) can be solved by exhaustive search across all $\underline{j} \in \mathcal{V}$, at a cost in $O(J^2)$. The total cost is thus $O((B+1)(N+J^3)+J^2) \simeq O(B(N+\log(N)^3)) \simeq O(BN)$, which, thus, remains perfectly acceptable in practice. Moreover, the bootstrap procedure can be easily run in parallel, thus yielding a time complexity of $O((B/P)N)$, where P is the number of processing units that are used.

4. Performance assessment

To assess the performance of the proposed *automated scaling range selection* procedure, to show its adaptivity and benefits, and to illustrate its robustness against noise, Monte Carlo simulations are conducted. They rely on the use of a large number of independent copies of synthetic

stochastic processes, whose characteristics are aiming at reproducing a variety of realistic situations, ranging from pure scale invariance processes to data corrupted by noise, either at fine or coarse scales (or both). Simulation set-up and statistical performance evaluation are reported below in this section.

4.1. Methodology and numerical experiments

4.1.1. Performance evaluation

A large number N_{MC} of realizations of a process with perfect and controlled *scale invariance* and *multifractal* properties are synthesized numerically. For each realization, the scaling range is automatically selected by the proposed procedure and used to estimate various multifractal parameters, $\zeta(q)$ or c_p, arbitrarily referred to as θ. To assess the quality of the scaling range selection procedure with respect to multifractal parameter estimation, performance is systematically compared against that obtained from the *MSE-optimal* scaling range, i.e., the scaling range that yields the minimum mean-squared error (MSE) estimates of θ. That *MSE-optimal* scaling range is found by computing, exhaustively across all possible scaling ranges $\underline{j} \in \mathcal{V}$, averages over Monte Carlo realizations of MSE($\underline{j}$), thus used as an ensemble average for $\mathbb{E}[(\theta-\hat{\theta})^2]$.

A second set of simulations involve the more realistic case where scaling is not perfect but corrupted either at fine or coarse scales. This is simulated by either adding Gaussian noise, which pollutes fine scales, or by high-pass filtering, which breaks the scaling at coarse scales. These simulations permit to address the issue of whether the proposed algorithm is able to select a scaling range that excludes the corrupted scales.

To report results, the following notations are used:

- $\underline{j}^M = [j_1^M, j_2^M] = \arg\min_{\underline{j} \in \mathcal{V}} \mathrm{MSE}(\underline{j})$: scaling range that produces the minimum MSE for $\hat{\theta}$, obtained from independent copies of the process, and thus regarded as the target to reach for the proposed procedure.
- $\underline{j}^D = [j_1^D, j_2^D]$: automatically selected scaling range.
- $\langle \underline{j}^D \rangle = [\mathrm{median}(j_1^D), \mathrm{median}(j_2^D)]$ and $\mathrm{mad}(\underline{j}^D) = [\mathrm{mad}(j_1^D), \mathrm{mad}(j_2^D)]$, computed over Monte Carlo realizations.
- $\mathrm{MSE}_{j_1}(j_2) = \min_{j_1} \mathrm{MSE}(j_1, j_2)$ and $\mathrm{MSE}_{j_2}(j_1) = \min_{j_2} \mathrm{MSE}(j_1, j_2)$.

4.1.2. Synthetic multifractal processes

In this section, the multifractal processes used in the numerical simulations are described. All the processes that are considered show perfect scaling across all available scales. In Sections 4.3–4.5 below, these processes with perfect scaling are further corrupted either with noise at fine scales or by filtering at coarse scales.

Canonical Mandelbrot Cascades: Multiplicative cascades were introduced in hydrodynamics to model turbulence and were later gathered into a unified framework by Mandelbrot [22]. They are here thus referred to as the Canonical Mandelbrot Cascades (CMC). Their construction relies on an iterative split-and-multiply random procedure on the interval. These cascades were shown to have a rich multifractal behavior (cf. e.g., [10]), controlled by the nature of the splitting procedure and the distribution of the random variables that are involved. In the present contribution, use is made of the binary CMC with log-Poisson variables $W_k = 2^{\gamma} \exp(\ln(\beta)\pi_\lambda)$, where π_λ is a Poisson random variable with parameter $\lambda = -\gamma\ln(2)/(\beta-1)$. In the following, this process will be denoted CMC(β,γ).

Lévy-stable process: The so-called α-stable Lévy process $L_\alpha(t)$ is built from a symmetric α-stable measure $M(ds)$ as $L_\alpha(t) = \int_{\mathbb{R}} f(t,s)M(ds)$, where $f(t,s) = \mathbb{1}(t-s>0) - \mathbb{1}(-s>0)$ [12]. Its multifractal properties are fully controlled by the parameter α, as shown in [23]. In the following, this process will be denoted Lévy(α).

Multifractal random walk: Multifractal random walk (MRW) was defined as a multifractal variation of fractional Brownian motion (fBm), cf. e.g., [24]: $X(k) = \sum_k G_H(k) \exp(\omega_\lambda(k))$, where $G_H(k)$ are the increments of fBm with parameter H, and ω_λ is a Gaussian process, independent of G_H, with a particular covariance structure [24] chosen to mimic that of CMC. MRW is not Gaussian and with stationary increments. Its multifractal properties are controlled by the parameters H and λ. In the following, this process will be denoted as MRW (H,λ).

Fractional Brownian motion in multifractal time: Fractional Brownian motion in multifractal time (MF-fBm) acts as a well-known and reference process for multifractal behavior [25] (as Fractional Brownian motion $B_H(t)$ does for self-similarity [1]). Use is made here of the compound Poisson cascade version of MF-fBm, referred to as CPC-fBm, $V_{H,A}(t)$, introduced in [26] and abundantly studied (cf. e.g., [27,28]): It is constructed by subordinating $B_H(t)$ to a compound Poisson cascade $A(t)$: $V_{H,A}(t) = B_H(A(t))$. The multifractal properties of CPC-fBm depend both on the Hurst parameter of the fBm and on the distribution of the random variables associated with the compound Poisson motion. For the purposes of the present contribution, these random variables are restricted to being log-normal, thus depending only on their mean μ_M and variance σ_M^2. Further, we will restrict ourselves to the case $\mu_M = 0$. In the following, this process will be denoted as CPC-fBm (H,σ_M^2).

4.1.3. Simulation setup

For the numerical simulations, various sample sizes are considered: $N \in \{2^{15}, 2^{18}\}$. The synthesis parameters of all processes are set to:

- CMC(β,γ): $\beta = 2/3$, $\gamma \in \{1/3, 2/3\}$.
- Lévy(α): $\alpha \in \{0.6, 1.6\}$.
- MRW(H,λ): $H=0.8$, $\lambda \in \{\sqrt{0.03}, \sqrt{0.08}\}$.
- CPC-fBm(H,σ^2): $H=0.5$, and $\sigma_M^2 \in \{0.01, 0.1\}$.

For the analysis, Orthonormal Daubechies wavelets with $N_\psi = 3$ vanishing moments are used. Bootstrap parameters are set to $B=500$ bootstrap resamples. For estimation, nonweighted linear regression (Eqs. (2) and (3)) is used (cf. Section 3.1).

It has been checked that results equivalent to those reported below are obtained when either the analysis or the synthesis parameters are varied.

4.1.4. Bootstrap parameter selection

A crucial issue in nonparametric block bootstrap procedure consists in selecting the block size (cf. [21]). The intrinsic time-scale nature of the wavelet Leaders leads to apply bootstrap to time-scale strips in the wavelet coefficient domain. An alternative, computationally less expensive choice, consists of using time-block bootstrap independently at each scale 2^j. An extensive and huge set of Monte Carlo simulations, not reported here, leads to conclude that, while it is crucial to make use of time-blocks, the use of time-scale strips does not bring significant improvement in performance. These simulations further showed that the time-block length W does not significantly impact the performance of the procedure on condition that it is chosen of the order of magnitude of the wavelet time-support. These results are in clear agreement with those reported in [8]. Therefore, in the sequel, all reported results are obtained with a bootstrap procedure applied independently, scale by scale, to time blocks of $L_X(j,k)$ of size $W = 2 \cdot N_\psi$.

4.2. Multifractal processes

4.2.1. Automated selection versus optimal-MSE

Performance is first examined for the nominal situation of perfectly scaling processes. Tables 2 and 3 show satisfactory matches between the automatically selected scaling range $\langle \underline{j}^D \rangle$ and the optimal-MSE range $\underline{j}^M$. Discrepancies between $\underline{j}^D$ and $\underline{j}^M$, that can be observed in Tables 2 and 3, are actually low or of minor practical impacts. Indeed, Figs. 1 and 2 (black solid lines) clearly illustrate, for several processes, that MSE is relatively flat around its minimum value, so that a selected scaling range that differs by 1, or even 2 octave(s), from the optimal-MSE value does not cause a significant increase in MSE and is thus an acceptable practical choice. Moreover, these figures also show the median of the bootstrap estimation of the MSE (blue dashed lines), on which the proposed procedure relies, that closely follows the MSE Monte Carlo estimation. Further, both figures display the histograms of the selected j_1^D and j_2^D. It can be seen that most values of j^D coincide with minimum MSE and, moreover, that MSE remains very flat in the interval that includes most of the selected values.

Tables 2 and 3 and Figs. 1 and 2 thus suggest that, for a large variety of different multifractal processes, the proposed procedure satisfactorily selects scaling ranges that yield close to optimal MSE for multifractal parameter estimation. These results further call for a number of comments detailed in the following subsections.

Table 2
Exact scaling: selected scaling range and minimum MSE: c_1 and c_2.

		c_1		c_2		c_1 and c_2	
		j_1	j_2	j_1	j_2	j_1	j_2
CMC(2/3,1/3)	j^M	2	4	2	5	2	4
	$\langle \underline{j}^D \rangle$	2	8	2	8	2	9
	mad(j^D)	1	1	0	2	0	1
CMC(2/3,2/3)	j^M	2	10	2	5	2	5
	$\langle \underline{j}^D \rangle$	2	6	2	7	2	8
	mad(j^D)	0	1	0	1	0	1
Lévy(0.6)	j^M	4	6	1	4	2	5
	$\langle \underline{j}^D \rangle$	4	8	2	6	4	9
	mad(j^D)	1	1	1	1	1	1
Lévy(1.6)	j^M	3	5	2	4	2	5
	$\langle \underline{j}^D \rangle$	3	8	2	7	3	8
	mad(j^D)	1	1	0	1	1	1
MRW(0.8, $\sqrt{0.03}$)	j^M	1	5	2	8	2	8
	$\langle \underline{j}^D \rangle$	3	9	3	8	3	9
	mad(j^D)	1	1	1	2	1	1
MRW(0.8, $\sqrt{0.08}$)	j^M	2	5	1	9	2	5
	$\langle \underline{j}^D \rangle$	2	8	2	8	2	8
	mad(j^D)	1	1	0	1	0	1
CPC-fbm(0.5, 0.01)	j^M	1	6	3	10	4	10
	$\langle \underline{j}^D \rangle$	4	9	4	9	4	9
	mad(j^D)	2	2	1	2	1	1
CPC-fbm(0.5, 0.1)	j^M	1	11	3	8	4	9
	$\langle \underline{j}^D \rangle$	2	8	3	8	3	8
	mad(j^D)	1	2	1	1	1	1

Table 3
Exact scaling: selected scaling range and minimum MSE: $\zeta(q)$.

	q	−2		−1		1		2		All q's	
		j_1	j_2	j_1	j_2	j_1	j_2	j_1	j_2	j_1	j_2
CMC(2/3,1/3)	j^M	2	11	5	11	1	3	1	3	2	11
	$\langle \underline{j}^D \rangle$	3	9	2	8	1	8	1	8	2	10
	mad(j^D)	1	2	0	2	0	1	0	2	0	2
CMC(2/3,2/3)	j^M	2	10	5	13	1	3	1	12	2	11
	$\langle \underline{j}^D \rangle$	4	10	2	8	2	7	1	8	3	10
	mad(j^D)	2	2	0	2	0	1	0	2	1	2
Lévy(0.6)	j^M	3	8	5	9	3	12	2	8	4	8
	$\langle \underline{j}^D \rangle$	4	9	4	9	2	6	1	5	4	10
	mad(j^D)	2	2	1	1	1	1	0	1	1	1
Lévy(1.6)	j^M	2	6	4	6	8	12	2	8	4	10
	$\langle \underline{j}^D \rangle$	3	9	3	9	3	7	2	6	3	9
	mad(j^D)	1	2	1	2	1	1	1	1	1	1
MRW(0.8, $\sqrt{0.03}$)	j^M	2	8	2	4	2	4	2	6	2	4
	$\langle \underline{j}^D \rangle$	3	9	3	9	3	8	3	8	3	9
	mad(j^D)	1	2	1	2	2	1	1	2	1	1
MRW(0.8, $\sqrt{0.08}$)	j^M	2	7	2	4	9	11	1	4	2	4
	$\langle \underline{j}^D \rangle$	2	9	2	8	2	8	1	8	2	9
	mad(j^D)	0	2	0	2	1	1	0	1	0	1
CPC-fbm(0.5, 0.01)	j^M	1	11	1	11	1	4	4	9	4	10
	$\langle \underline{j}^D \rangle$	4	9	4	9	4	9	4	9	4	9
	mad(j^D)	1	1	1	1	2	2	1	2	1	1
CPC-fbm(0.5, 0.1)	j^M	2	4	1	11	1	8	1	3	4	9
	$\langle \underline{j}^D \rangle$	3	8	3	8	3	8	3	8	3	9
	mad(j^D)	1	1	1	1	1	2	1	2	1	1

4.2.2. Automated selection versus common practice

These results shed an interesting light on actual practice in multifractal analysis conducted on real world data: Commonly, practitioners would select the *largest* range of

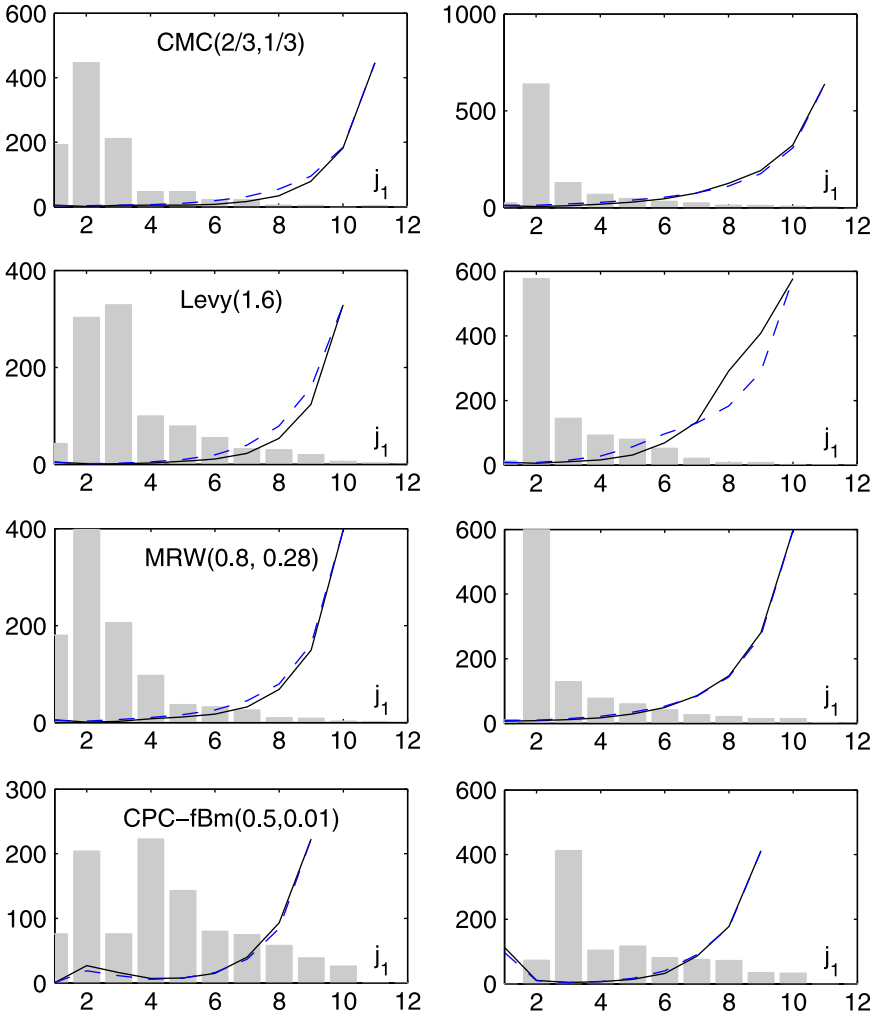

Fig. 1. *Selected* j_1^D: Histograms of selected lower cutoff octaves j_1^D for CMC(2/3,1/3) (top row), Lévy(1.6) (2nd row), MRW(0.8, $\sqrt{(}0.08)$) (3rd row) and CPC-fBm(0.5,0.01) (bottom row), and for estimates $\hat{c}_1$ (left column) and $\hat{c}_2$ (right column). Black solid lines: $\mathrm{MSE}_{j_2}(j_1)$ computed from Monte Carlo realizations. Blue dashed lines: median of $\mathrm{MSE}_{j_2}(j_1)$ computed from bootstrap replications for each realization. $\mathrm{MSE}_{j_2}(j_1)$ is shown in arbitrary units. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

scales where a scaling behavior is visually acceptable, with the underlying a priori intuition that the larger the scaling range the better the estimation performance.[1] On the example illustrated in Fig. 3, such practice would likely result in the selection of the range $\underline{j} = [2, 11]$ (dashed red). However, Tables 2 and 3 suggest that the optimal range is actually much narrower: $\underline{j} = [4, 9]$ (green mixed vertical lines) corresponds to the optimal-MSE range and $\underline{j} = [3, 9]$ (solid blue vertical lines) is consistently obtained from the proposed selection procedure. Therefore, the use of the proposed automated scaling range selection procedure leads to better estimations, even in the *simple* case of perfectly scaling processes, by handling the bias-variance trade-off in a better manner than intuition or visual inspection would do.

[1] At least, this has been our regular practice in multifractal analysis conducted on the numerous datasets we have been analyzing.

4.2.3. Estimation versus analysis

Theoretically, multifractal analysis requires that the scaling range is the same for all orders m. The proposed automated range selection procedure can either be applied to all orders jointly (cf. Eq. (7)) or independently for each order (cf. Eq. (8)). This permits to address a very important practical issue: Practitioners inspect visually and independently $C_1(j)$ and $C_2(j)$, when scaling ranges observed on both plots match, this constitutes a satisfactory evidence in favor of the relevance of the multifractal paradigm to describe the data at hand. Conversely, significant mismatches between the scaling range automatically extracted independently from $C_1(j)$ and $C_2(j)$ constitute indications that multifractal models are strongly questioned by data. This is further discussed in Section 5. Table 2 shows excellent matches between the selected scaling ranges obtained from $C_1(j)$ or $C_2(j)$ independently and jointly, as expected for a purely multifractal process.

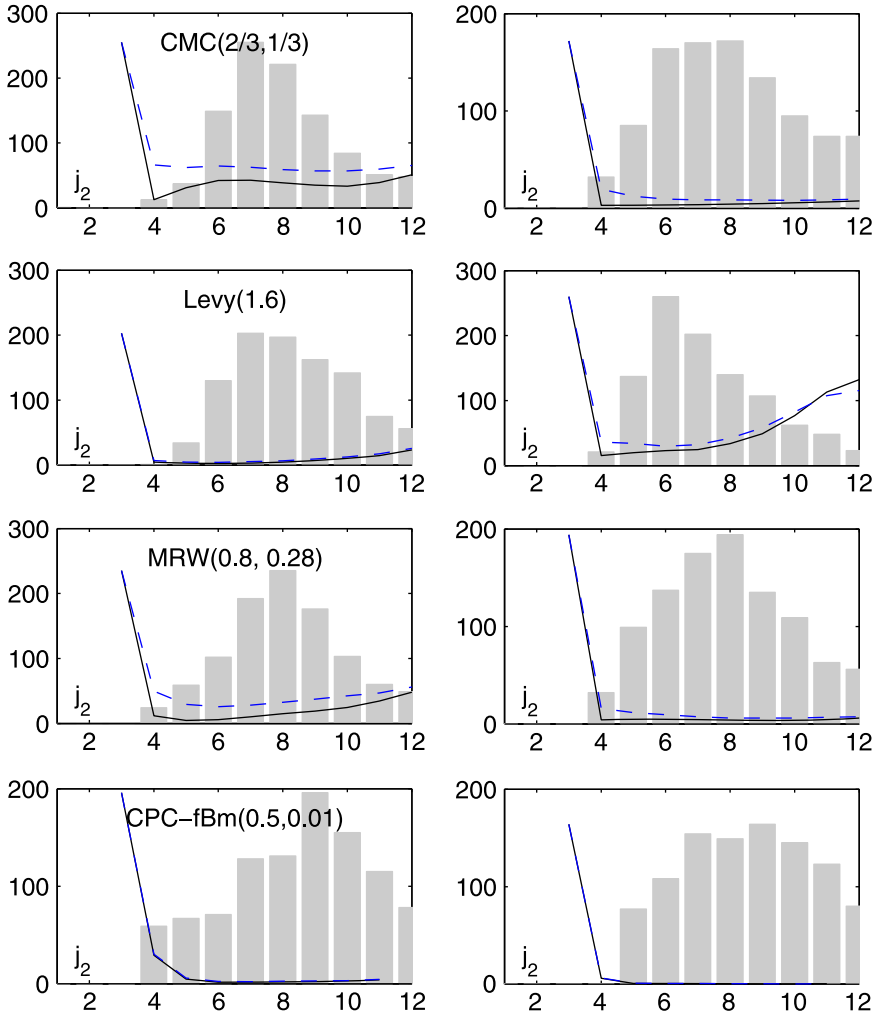

Fig. 2. *Selected* j_2^D: Histograms of selected upper cutoff octaves j_2^D for CMC(2/3,1/3) (top row), Lévy(1.6) (2nd row), MRW(0.8, $\sqrt{(}0.08)$) (3rd row) and CPC-fBm(0.5,0.01) (bottom row), and for estimates $\hat{c}_1$ (left column) and $\hat{c}_2$ (right column). Black solid lines: $\mathrm{MSE}_{j_1}(j_2)$ computed from Monte Carlo realizations. Blue dashed lines: median of $\mathrm{MSE}_{j_1}(j_2)$ computed from bootstrap replications for each realization. $\mathrm{MSE}_{j_1}(j_2)$ is shown in arbitrary units. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

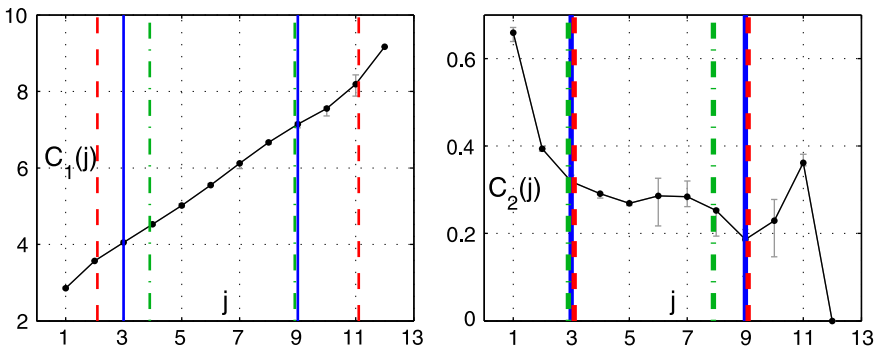

Fig. 3. *Exact scaling*: Logscale diagram for a perfectly scaling process (black line), as well as the detected scaling range (vertical blue line), MSE-optimal range (vertical green-mixed line) and visually acceptable range (vertical red-dashed line), for $C_1(j)$ (left) and $C_2(j)$ (right). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

4.2.4. *Importance of the lower and upper cutoffs*

Figs. 1 and 2 (and later Figs. 5 and 8) further show that the widths of the histograms for j_2^D are much larger than those for j_1^D. This is a direct consequence of the fact that structure functions themselves have much larger variance at coarse scales than they do at fine scales (mostly because

the number of available wavelet leaders decreases when scales increase). Therefore, accurately picking j_1 impacts estimation performance far more than carefully selecting j_2 does, another crucial fact for practitioners to be aware of, which constitutes a clear outcome of the analysis conducted here.

4.2.5. Influence of sample size

To study the influence of the size of the realization on the detection performance, MRW with parameters $H=0.8$ and $\lambda=\sqrt{0.08}$ and of length $N \in \{2^{10}, 2^{12}, 2^{14}, 2^{16}, 2^{18}\}$ was studied. Table 4 reports the statistics of the automatically selected scaling ranges. It shows excellent matches between the automatically selected scaling range $\langle \underline{j}^D \rangle$ and the optimal-MSE range $\underline{j}^M$, for all sample sizes. Table 4 shows increased discrepancies between $\langle \underline{j}^D \rangle$ and $\underline{j}^M$ for large sample sizes. This is a direct consequence of the fact that MSE is more flat around its minimum, and therefore the algorithm has more "good" scaling ranges to choose from. To the contrary, and interestingly enough, Table 4 indicates that the proposed procedure performs even better for short sample sizes, since in this case MSE is sharper around its minimum.

Table 4
Exact scaling: selected scaling range and minimum MSE: Influence of data length on selected ranges on MRW($0.8, \sqrt{0.08}$) for c_1 and c_2.

$\log_2(N)$		c_1		c_2		c_1 and c_2	
		j_1	j_2	j_1	j_2	j_1	j_2
10	j^M	2	5	2	6	2	5
	$\langle \underline{j}^D \rangle$	2	5	2	5	2	5
	mad(j^D)	1	1	1	1	1	1
12	j^M	2	6	2	7	2	6
	$\langle \underline{j}^D \rangle$	2	7	2	6	2	6
	mad(j^D)	1	1	0	1	0	1
14	j^M	2	7	2	7	2	7
	$\langle \underline{j}^D \rangle$	2	7	3	7	2	7
	mad(j^D)	1	1	1	1	0	1
16	j^M	2	8	2	8	2	8
	$\langle \underline{j}^D \rangle$	3	9	3	8	3	9
	mad(j^D)	1	1	1	1	1	1
18	j^M	3	6	2	10	2	9
	$\langle \underline{j}^D \rangle$	4	10	4	10	3	10
	mad(j^D)	1	2	2	2	1	2

4.3. Lower cutoff for scale invariance and robustness to noise

To assess the ability of the proposed procedure to bring robustness against departures from perfect scaling across all scales, the situation where an additive noise corrupts the finest scales is first studied. Departure from exact scaling behavior at fine scales can easily be achieved by additively superimposing a white Gaussian noise, with controlled and prescribed signal to noise ratio (SNR), as illustrated in Fig. 4. The range of finer scales contaminated by noise depends on the SNR and can thus be expressed as $[1, j_N(\mathrm{SNR})]$, with j_N a monotonously decreasing function of SNR. Results are reported here for CPC-fBm only (consistent results were obtained for all processes).

Fig. 4 displays $C_1(j)$ and $C_2(j)$, averaged across Monte Carlo realizations, comparing the perfect scaling ('+'-blue) to the noise-corrupted ('*'-black) cases. Fine scale corruption is obviously visible. The MSE-optimal range $\underline{j}^M$ and the median selected range $\langle \underline{j}^D \rangle$, both obtained from the noisy data, are reported with dashed-black and mixed-red lines, respectively, again showing an excellent match of the latter to the former. This is further comforted by the results reported in Table 5 for various SNRs, $j_N(\mathrm{SNR}) \in \{4, 5, 6, 7, 8\}$. A large sample size is used so

Table 5
Fine scale scaling departure: Selected lower cutoff j_1^D for $\hat{c}_1$ and $\hat{c}_2$.

		j_N				
		4	5	6	7	8
c_1	j_1^M	5	6	7	8	9
	$\langle j_1^D \rangle$	6	6	7	7	9
	mad(j_1^D)	1	1	1	1	1
c_2	j_1^M	5	6	6	7	9
	$\langle j_1^D \rangle$	6	6	7	7	8
	mad(j_1^D)	1	1	1	1	1
c_1&c_2	j_1^M	5	6	7	8	9
	$\langle j_1^D \rangle$	5	6	6	7	9
	mad(j_1^D)	1	1	1	1	1

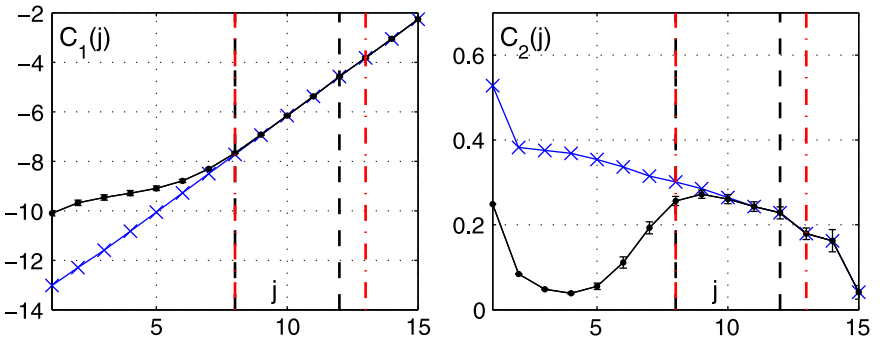

Fig. 4. *Fine scale scaling departure*: Average logscale diagrams (solid black) of noise-corrupted CPC-fBm with $j_N=6$, optimal MSE range (vertical mixed red line) and median selected range (vertical dashed black line), for $C_1(j)$ (left) and $C_2(j)$ (right), $\sigma_M^2=0.1$ and $N=2^{18}$. The blue solid line with crosses shows the noise-free case. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

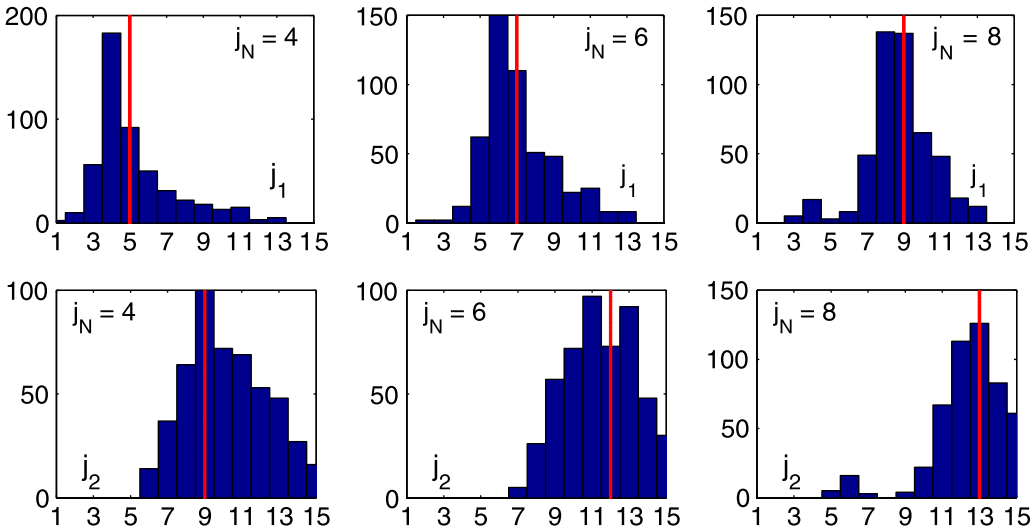

Fig. 5. *Fine scale scaling departure*: Histograms of selected j_1^D (upper row) and j_2^D (lower row) for $C_1(j)$, for CPC-fBm corrupted with white noise, $j_N=4$ (left), $j_N=6$ (middle) and $j_N=8$ (right). Red vertical lines indicate j_1^M. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

that enough scales remain available for estimation despite noise.

Fig. 5 shows the histograms of the selected j_1^D and j_2^D for three different SNRs. The modes of the distributions of j_1^D shift towards coarser scales as the SNR decreases. Also, the distributions are highly peaked, meaning that the algorithm performs satisfactorily in the detection of the corruption at fine scales. Interestingly, histograms for j_2^D also shift towards coarser scales, because the algorithm tries to compensate the loss of fine scales to achieve better estimation performance. Fig. 5 thus clearly illustrates the adaptivity of the procedure, in a nonsupervised manner, to the increasing level of noise at fine scales, by selecting both higher j_1^D and j_2^D.

Fig. 6 (top row) illustrates estimation performance as a function of SNR and compares it against estimation in the pure scaling case (SNR $=+\infty$, with fixed scaling range $\underline{j}=[3,15]$ (F) and with the automatically selected one (denoted by $j_N=0$)). To the contrary, Fig. 6 (bottom row) displays estimation performance using the fixed scaling range $\underline{j}=[3,15]$ for all j_N. Overall, Fig. 6 clearly shows that the automated selection of the scaling range permits to avoid the significant bias that stems from the use of the a priori fixed range of scales $\underline{j}=[3,15]$. An expected increase in variance is also observed as SNR decreases because (1) there are fewer octaves available for estimation, and (2) the procedure is forced to select coarser octaves, which themselves show larger variances.

The analyses reported in this section all yield similar conclusions: The scaling range is correctly, automatically and adaptively selected to ensure robustness against noise and (close to) optimal MSE.

4.4. Upper cutoff for scale invariance

Let us now turn to the situation where scale invariance is broken at coarse scales. In applications, this naturally arises as the physical, physiological, etc., mechanisms that produce scale invariance may cease to act at coarse enough scales: this is e.g., the case in hydrodynamic turbulence where the mechanisms permitting to inject energy in flows are bounded above by the so-called *integral* scale (cf. e.g., [5]).

To produce the coarse scale cutoff effect in a controlled manner, CPC-fBm(0.5,0.01) with $N=2^{15}$ is high-pass filtered. The frequency response of the filter is chosen such that its cutoff scale varies as 2^{j_C}, with $6\le j_C\le 10$ and so that it is maximally flat within the band-pass domain. Results are reported here for the sole CPC-fBm processes (consistent results were obtained for all processes). Fig. 7 displays $C_1(j)$ and $C_2(j)$, averaged across Monte Carlo realizations, comparing the perfect scaling (blue) to the high-pass-filtered (black) case. Coarse scales corruption is obviously visible. The MSE-optimal range $\underline{j}^M$ and the median selected range $\langle \underline{j}^D\rangle$, both obtained from the filtered data, are illustrated with dashed-black and mixed-red lines, respectively, again showing a satisfactory match of the latter with the former, selecting scaling ranges that differ in one or two scales from the optimal, but still inside the scaling region in both cases. This observation is further comforted by the results reported in Table 6 for various cutoff scales (corresponding to $j_C\in\{6,7,8,9,10\}$). Table 6 however shows that, for c_2, the match between $\underline{j}^M$ and the median selected range $\langle \underline{j}^D\rangle$ is less accurate. Fig. 7 provides a potential explanation: The practically achieved corruption of coarse scales, consisting of linear filtering, appears to be much less effective for c_2 (than it is for c_1).

Fig. 8 shows the histograms of the selected j_1^D and j_2^D. As expected, j_2^D is selected larger roughly following the increase of the filter cutoff scale 2^{j_c}. To the contrary, the procedure mostly selects the lowest available $j_1^D=2$ as scaling at fine scales is not altered and using the finest available scales always leads to lowest estimation variance. Again, this illustrates the ability of the proposed procedure to accommodate in a nonsupervised manner departures from exact scale invariance.

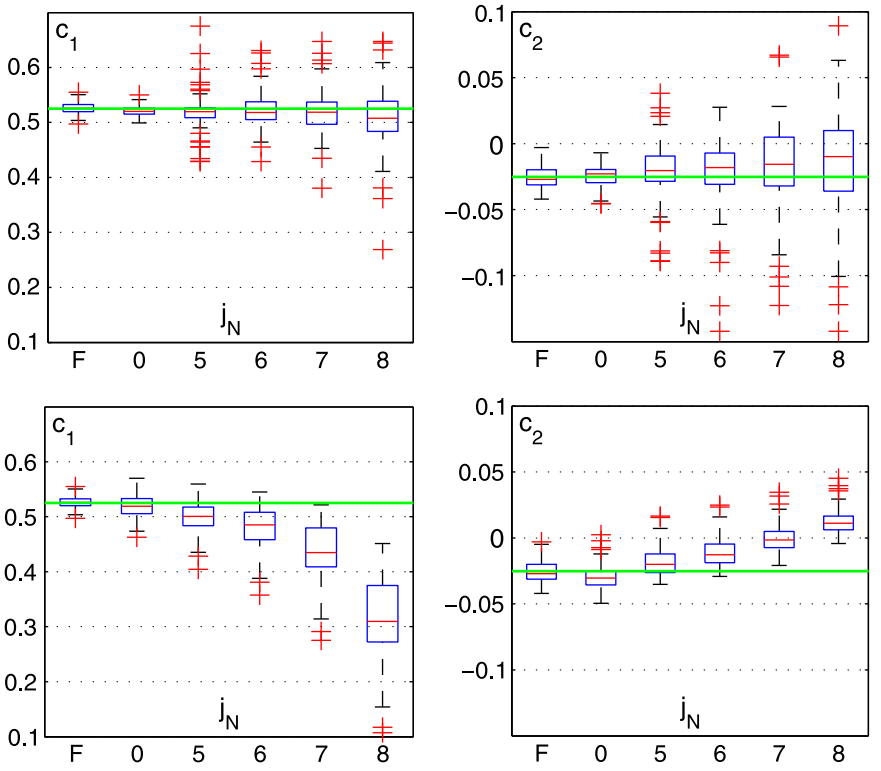

Fig. 6. *Fine scale scaling departure*: Boxplots for $\hat{c}_1$ (left) and $\hat{c}_2$ (right) for increasing SNR, using the selected scaling ranges (top row) and the fixed scaling range $\underline{j}=[3,15]$ (bottom row). The green horizontal line indicates the theoretical value. F and $j_N=0$ indicates the noise-free case, analyzed using a fixed scaling range and the selected scaling ranges, respectively. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

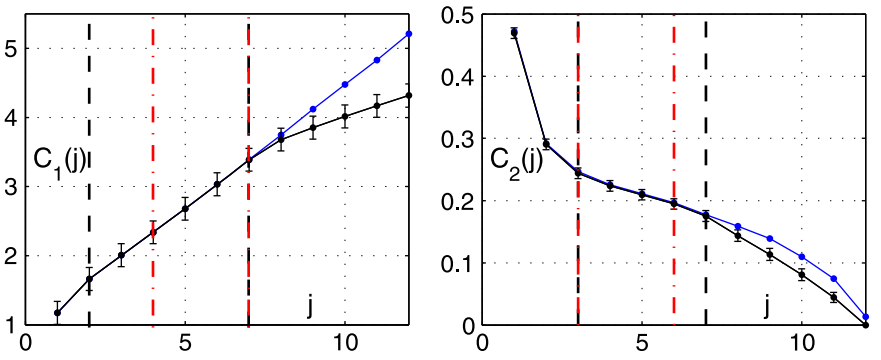

Fig. 7. *Coarse scale scaling departure*: Average logscale diagrams (solid black) of high-pass filtered CPC-fBm with cutoff octave $j_C=8$, optimal MSE range (mixed red vertical line) and median selected range (dashed black vertical line), for $C_1(j)$ (left) and $C_2(j)$ (right). The blue dotted-lines show $C_1(j)$ and $C_2(j)$ for the nonfiltered case. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

4.5. Random lower and upper departures from scaling

Let us now turn to a case study that very much resembles the real-world situation where practitioners analyze a large dataset, consisting of a large number of time series: Departures from exact scale invariance may occur at both fine and coarse scales simultaneously and may involve scales that vary from one time series to another. This is notably the case in biomedical applications, such as the one described in Section 5 below, where the level of noise varies from one subject to another, thus changing the lower cutoff, and where either the mechanisms producing scaling or seasonal trends may vary from subject to subject, thus modifying the upper cutoff. For such situations, practitioners must either perform a systematic check for each subject or adopt a fixed scaling range that globally matches the dataset. The former choice is time consuming, tedious and prone to error and subjectivity, while the latter option may lead to poor estimation performance with either dramatic biases for given subjects or large overall variance.

To reproduce that situation and explore the benefit of the proposed automated selection procedure, a large number of realizations of synthetic data are considered,

corrupted with white noise, where SNR (and thus j_N) is drawn at random (from the set $j_N = \{2,3,4\}$) and then high-pass filtered, where the high-pass cutoff octave j_C is also and independently drawn at random (from the set $j_C = \{7,8,9\}$). For each realization, estimation is performed using: (i) the automated scaling range selection procedure, (ii) a fixed conservative range $\underline{j} = (4,6)$ (FC), carefully chosen to be inside the scaling range for all realizations, and (iii) a fixed nonconservative range $\underline{j} = (3,9)$ (FNC), chosen as the largest scaling range observed across all time series. Results are reported here for the sole CPC-fBm processes (consistent results were obtained for all processes). Fig. 9 shows that estimations stemming from the proposed automated selection procedure benefit from performance far better than those of the nonconservative fixed range (which shows both large biases and variances and is thus useless in practice). Fig. 9 also shows that the automated scaling range selection yields more robust statistics with slightly better performance when compared to those obtained from the FC range, while avoiding the tedious and prone to error burden of scanning all time series independently. These results clearly illustrate the benefits of the proposed automated scaling range selection: Its adaptivity to data. When data allow it, a large range of scales is used (thus permitting low variance), and conversely, scaling range is narrowed when data suffer from severe departures from scaling (thus ensuring low bias).

Table 6
Coarse scale scaling departure: Selected j_2^D for $\hat{c}_1$ and $\hat{c}_2$.

		j_C				
		6	7	8	9	10
c_1	j_2^M	4	6	7	7	8
	$\langle j_2^D \rangle$	5	6	7	7	7
	mad(j_2^D)	0	1	1	1	1
c_2	j_2^M	5	7	6	8	8
	$\langle j_2^D \rangle$	10	9	9	9	9
	mad(j_2^D)	0	1	1	1	1
c_1&c_2	j_2^M	5	6	7	8	8
	$\langle j_2^D \rangle$	6	6	7	8	8
	mad(j_2^D)	1	0	1	1	1

4.6. Existence of several scaling regimes

If the data under analysis were to show several scaling regimes, the algorithm would choose the one with the best goodness-of-fit. This situation occurs, for example, in Section 4.3, where multifractal processes are corrupted with white noise that introduces a new scaling region at fine scales. However, all other regimes could be easily detected by running multiple instances of the algorithm with different restrictions on the search space $\mathcal{V}$. For instance, if data are suspected to show a scaling regime at fine scales and a different one at coarse scales, the algorithm could be run twice with search regions $\mathcal{V}_1$ and $\mathcal{V}_2$ such that $j_1 \le j_1^{max}$ in $\mathcal{V}_1$ and $j_2 \ge j_2^{min}$ in $\mathcal{V}_2$, for proper choices of j_1^{max} and j_2^{min}.

5. Heart rate variability

5.1. Scale invariance in heart rate variability

Heart rate variability (HRV) data is now commonly used to asses the cardiologic health status, with the leading theme that a *strong* variability indicates *good* health [14]. It is now well documented that the variability of heart rate can be interestingly quantified by scale invariance and multifractal properties and attributes, cf. e.g., [29,13] and references therein. Recently, it has been shown that the

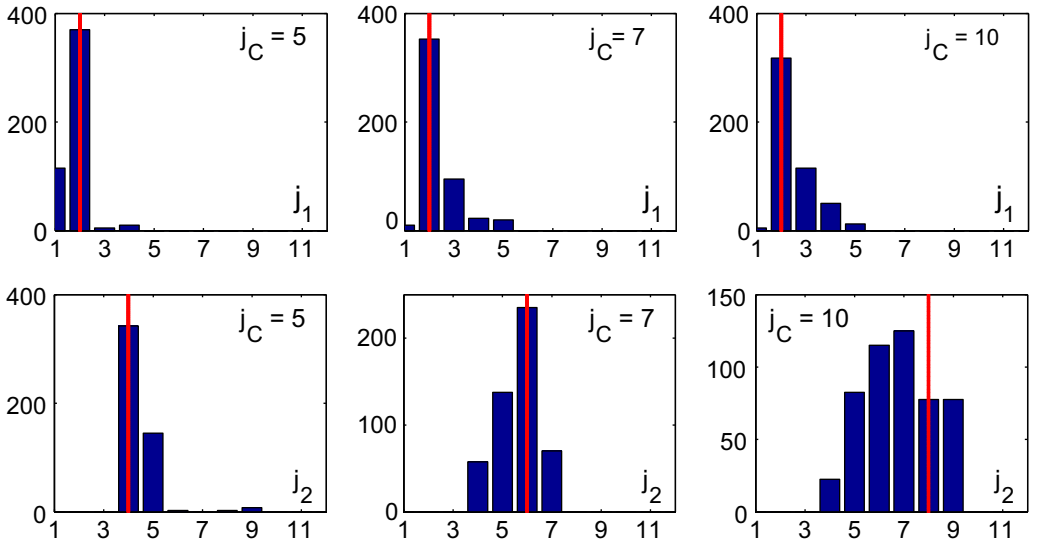

Fig. 8. *Coarse scale scaling departure*: Histograms of selected j_1^D (upper row) and j_2^D (lower row) for $C_1(j)$ of high-pass filtered CPC-fBm with j_C=6 (left), j_C=7 (middle) and j_C=10 (right). Red solid vertical lines indicate j_2^M. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

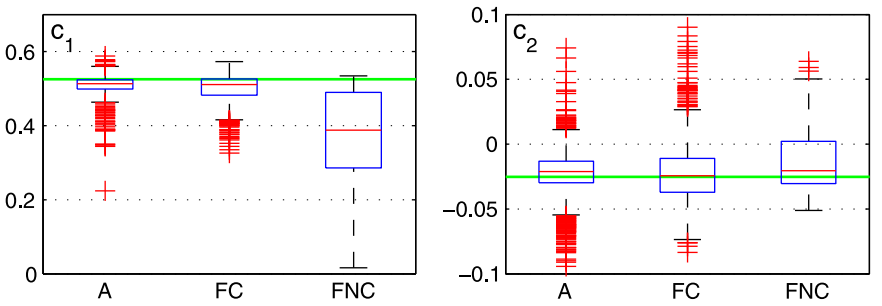

Fig. 9. *Fine and coarse scale scaling departures*: CPC-fBm corrupted by additive white noise with random SNR and high-pass filtered with random cutoff frequency, $\hat{c}_1$ (left) and $\hat{c}_2$ (right) from automated selection (A), and from fixed conservative $\underline{j}=(4,6)$ (FC) and nonconservative range $\underline{j}=(3,9)$ (FNC) scaling ranges. The horizontal solid green lines indicate the theoretical c_1 and c_2. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

wavelet leader multifractal formalism yields fruitful analysis of HRV (cf. [30,31]).

In the present contribution, 54 (2-hour long) records are studied, taken from the *Normal Sinus Rhythm RR Interval Database*, made available by Physionet (http://www.physionet.org/physiobank/database/nsr2db/, cf. [32]). All records correspond to subjects with normal sinus rhythm, with a balanced male–female ratio (30/24), and with mean age ($\pm$ standard deviation) of 61.36 ($\pm$ 11.63) years. Heart beats are extracted by a standard and automated procedure, with correction and revision by experts, and are provided for analysis with annotation files. Heart beat lists are interpolated, with cubic splines, into a regularly sampled time series, at sampling frequency $f_s=4$ Hz. Healthy subjects only are studied here for simplicity.

5.2. Scaling range selection and scaling parameter estimation

The wavelet Leader multifractal formalism, as described in Section 2, is applied to each subject to estimate parameters c_1 and c_2, with 3 different settings for the scaling range selection: (i) automated selection (A), independently for each time series, and with automation performed on $C_1(j)$ and $C_2(j)$ independently, and $C_1(j)$ and $C_2(j)$ jointly; (ii) a fixed conservative range $\underline{j}=(8,10)$ (FC) carefully chosen to be well into the scaling region for all series; and (iii) a fixed nonconservative range $\underline{j}=(4,13)$ (FNC), corresponding to the largest scaling range observed across all subjects.

Fig. 10 compares estimation performance for $\hat{c}_1$ and $\hat{c}_2$ obtained with these 3 different choices of scaling ranges and Tables 7 and 8 report the median of selected j_1^D and j_2^D (with median absolute deviations and extreme values) and the statistics for $\hat{c}_1$ and $\hat{c}_2$, leading to the following comments and discussion.

The FNC range yields, for both c_1 and c_2, estimates with (slightly) lower variance (compared to the estimates stemming from the A and FC ranges), thus indicating a reduced variability in estimation resulting from the use of a large range of scales for estimation. Yet, the means (and medians) of the estimates significantly differ from those computed with the A and FC ranges, thus suggesting a strong bias, resulting from the poor accounting of the inter-individual variations of the actual scaling range bounds.

Table 7 shows that the average selected scaling range, from the automated selection procedure with a joint detection for c_1 and c_2, reads (6, 11), with little dispersion around the lower bound, thus showing a posteriori a relative homogeneity of the database and a relatively low inter-individual variability. The dispersion around the upper bound appears larger, yet, as already discussed in Section 4.2, optimal MSE as a function of j_2 is flat, thus the precise determination of the upper bound j_2 turns less critical in estimation performance as the variances of the quantities over which linear regression is performed increase with scales. Moreover, the automatically selected scaling range is found to yield an average scaling range close to the FC range, (8, 10), as well as estimation performance which is very comparable to those obtained with the FC range. The automated scaling range selection thus performs as well as the conservative scaling range choice, yet avoiding the issues raised by a systematic visual inspection of each subject in the database: Feasibility in terms of time devoted to the task and size of the database; inter-practitioner and even intra-practitioner subjectivity, etc.

Further, the automated scaling range selection procedure not only yields quasi-MSE-optimal estimates but also constitutes a data analysis tool, by outputting the range of scales within which scale invariance holds: Scaling in a range (6, 11) indicates scale invariance over time scales ranging from 20 s to 10.6 min (or equivalently for frequencies ranging from 0.0016 to 0.05 Hz), which is in very satisfactory agreement with findings reported on adults HRV in earlier contributions (cf. e.g., [13,29]). Note that visually, inspection of scaling plots for each subject would lead to conclude in most cases that scaling holds at finer scales as well (again an observation consistent with earlier findings), yet, as observed on numerical simulations in the pure scaling case (cf. Section 4.2), the automated scaling range selection procedure tends to narrow the scaling range involved into estimation to optimize estimation performance (MSE). In addition, the scaling ranges selected independently for $\hat{c}_1$ and $\hat{c}_2$ closely match in a remarkable way for these real-world data: This is a requirement from the theory underlying the multifractal paradigm that is not always observed on empirical data,

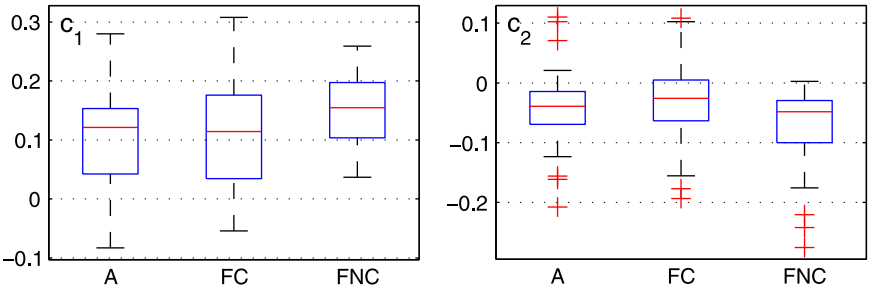

Fig. 10. *HRV data*: Boxplots for $\hat{c}_1$ (left) and $\hat{c}_2$ (right) from the automated selection (A), and from fixed conservative $\underline{j}=(8,10)$ (FC) and nonconservative range $\underline{j}=(4,13)$ (FNC) scaling ranges.

Table 7
HRV data: Statistics for j_1^D and j_2^D. The upper half shows the detected octaves. The lower half shows the detected scales in physical units (seconds).

		j_1^D				j_2^D			
		median	mad	min	max	median	mad	min	max
Octaves	c_1	7	1	5	8	10	1	8	14
	c_2	6	1	4	8	11	1	7	14
	c_1&c_2	6	1	5	8	11	1	8	14
Seconds	c_1	40	0.625	10	80	320	0.625	80	5120
	c_2	20	0.625	5	80	640	0.625	40	5120
	c_1&c_2	20	0.625	10	80	640	0.625	80	5120

Table 8
HRV data: Statistics for $\hat{c}_1$ and $\hat{c}_2$.

		mean	median	std	min	max
$\hat{c}_1$	A	0.10	0.12	0.09	−0.08	0.28
	FC	0.11	0.11	0.09	−0.05	0.30
	FNC	0.15	0.15	0.05	0.04	0.26
$\hat{c}_2$	A	−0.04	−0.04	0.06	−0.22	0.11
	FC	−0.03	−0.03	0.05	−0.18	0.15
	FNC	−0.07	−0.05	0.06	−0.28	0.00

and which thus constitutes a solid evidence supporting the use of the multifractal paradigm to model HRV at least on that dataset. To finish with, the automated scaling range selection procedure yields estimates c_2 that are quasi-systematically strictly negative, excluding 0 or positive estimates but for a few subjects: This constitutes another clear and solid evidence that scale invariance in HRV is better modeled by multifractal models rather than by the more standard self-similar processes, such as fractional Brownian motion (cf. e.g., [13,29] for detailed discussions).

6. Conclusions and perspectives

In the present contribution, we have defined a procedure for the automated selection of the scaling range where scaling parameter estimation is to be conducted. Elaborating on a previous contribution [8], the proposed procedure relies on a nonparametric bootstrap resampling scheme, performed on the wavelet leaders. The performance of the procedure has been assessed by means of Monte Carlo simulations, using perfectly scaling processes as well as processes with significant departures from exact scale invariance above or below a randomly set scaling range. For all situations, the proposed procedure was observed to yield estimates showing the lowest (or close to the lowest) achievable MSE, thus optimizing the bias-variance trade-off. Such satisfactory performance is obtained using wavelet leaders; however, simulations not reported here show that equivalent results are obtained when estimation is based on wavelet coefficients rather than leaders, as needed in specific applications. Also, the procedure can be extended straightforwardly to multivariate analysis or to higher dimensional field (image) analysis. At the moderate price of an increased computational cost (essentially due to the bootstrap operation), this procedure avoids the subjective and tedious (or not achievable for very large databases) task consisting of the systematic visual inspection of all subjects, while yielding comparable and even improved results. In addition to optimal estimation, the procedure also contributes to the analysis of scaling in data: By outputting the lower and upper bounds of the scaling range, it provides practitioners with indications of possible departures of scaling; also, the comparison of the scaling ranges selected from different multifractal attributes (c_1 and c_2 for instance) yields indications with respect to the validity of the multifractal paradigm underlying scale invariance analysis; further, the selected scaling range itself may constitute an information as important as the scaling exponent estimates themselves, discrimination between patients and healthy subjects may come from the change in the scaling range rather than changes in the scaling exponents. Finally, the proposed procedure was shown at work on a real biomedical database. Its methodological nature grants the procedure a general level of validity and it can thus be applied to real-world data produced by data of very different natures. A MATLAB procedure, designed by

ourselves, will be made publicly available from our webpages at the time of publication.

References

[1] B. Mandelbrot, J. van Ness, Fractional Brownian motion fractional noises and applications, SIAM Rev. 10 (1968) 422–437.
[2] B. Mandelbrot, Multifractals and 1/*f* noise, Selected Works of Benoit B. Mandelbrot, Wild Self-Affinity in Physics, Springer-Verlag, New York, 1998.
[3] P. Ciuciu, et al., Scale-free and multifractal dynamic properties of fMRI signals during rest and task, Front. Physiol. (2012) 3–186.
[4] P. Abry, et al., Multiscale nature of network traffic, IEEE Signal Process. Mag. 19 (3) (2002) 28–46.
[5] U. Frisch, Turbulence, the Legacy of A.N. Kolmogorov, Cambridge University Press, Cambridge, 1993.
[6] E. Foufoula-Georgiou, P. Kumar (Eds.), Wavelets in Geophysics, Academic Press, San Diego, 1994.
[7] L. Calvet, A. Fisher, Multifractal Volatility: Theory, Forecasting and Pricing, Academic Press, San Diego, CA, 2008.
[8] H. Wendt, P. Abry, S. Jaffard, Bootstrap for empirical multifractal analysis, IEEE Signal Process. Mag. 24 (4) (2007) 38–48.
[9] M. Doret, et al., Multifractal analysis of fetal heart rate variability in fetuses with and without severe acidosis during labor, Am. J. Perinatol. 28 (4) (2011) 259–266.
[10] R. Riedi, Multifractal processes, in: P. Doukhan, G. Oppenheim, M. Taqqu (Eds.), Theory and Applications of Long Range Dependence, Birkhäuser, 2003, pp. 625–717.
[11] S. Jaffard, Wavelet techniques in multifractal analysis, in: M. Lapidus, M. van Frankenhuijsen (Eds.), Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, Proceedings of Symposia in Pure Mathematics, vol. 72 (2), AMS, 2004, pp. 91–152.
[12] G. Samorodnitsky, M. Taqqu, Stable Non-Gaussian Random Processes, Chapman and Hall, New York, 1994.
[13] K. Kiyono, et al., Multiscale probability density function analysis: non-Gaussian and scale-invariant fluctuations of healthy human heart rate, IEEE Trans. Biomed. Eng. 53 (1) (2006) 95–102.
[14] S. Akselrod, et al., Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control, Science 213 (4504) (1981) 220–222.
[15] D. Veitch, P. Abry, M.S. Taqqu, On the automatic selection of the onset of scaling, Fractals 11 (2) (2003) 377–390.
[16] J. Park, C. Park, Robust estimation of the Hurst parameter and selection of an onset scaling, Stat. Sinica 19 (4) (2009) 1531–1555.
[17] X. Xia, G. Lazarou, T. Butler, Automatic scaling range selection for long-range dependent network traffic, Commun. Lett. 9 (10) (2005) 954–956.
[18] S. Mallat, A Wavelet Tour of Signal Processing, second ed. Academic Press, Cambridge, 2009.
[19] J. Delour, J. Muzy, A. Arneodo, Intermittency of 1d velocity spatial profiles in turbulence: a magnitude cumulant analysis, Eur. Phys. J. B 23 (2001) 243–248.
[20] A. Zoubir, D. Iskander, Bootstrap Techniques for Signal Processing, Cambridge University Press, Cambridge, 2004.
[21] S. Lahiri, Resampling Methods for Dependent Data, Springer New York, 2003.
[22] B. Mandelbrot, Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier, J. Fluid Mech. 62 (1974) 331–358.
[23] S. Jaffard, The multifractal nature of Lévy processes, Probab. Theory Relat. Fields 114 (1999) 207–227.
[24] E. Bacry, J. Delour, J. Muzy, Multifractal random walk, Phys. Rev. E 64 (2) (2001) 026103.
[25] B. Mandelbrot, A multifractal walk down Wall Street, Sci. Am. 280 (2) (1999) 70–73.
[26] J. Barral, B. Mandelbrot, Multifractal products of cylindrical pulses, Probab. Theory Relat. Fields 124 (3) (2002) 409–430.
[27] E. Bacry, J. Muzy, Log-infinitely divisible multifractal processes, Commun. Math. Phys. 236 (2003) 449–475.
[28] P. Chainais, R. Riedi, P. Abry, On non scale invariant infinitely divisible cascades, IEEE Trans. Inf. Theory 51 (3) (2005) 1063–1083.
[29] P.C. Ivanov, et al., Multifractality in human heartbeat dynamics, Nature 399 (6735) (1999) 461–465.
[30] R.F. Leonarduzzi, G. Schlotthauer, M.E. Torres, Wavelet leader based multifractal analysis of heart rate variability during myocardial ischaemia, in: Conference Proceedings on IEEE Engineering in Medicine and Biology Society, 2010, pp. 110–113.
[31] P. Abry, et al., Methodology for multifractal analysis of heart rate variability: from lf/hf ratio to wavelet leaders, in: Conference Proceedings on IEEE Engineering in Medicine and Biology Society, 2010, pp. 106–109.
[32] A.L. Goldberger, et al., PhysioBank, PhysioToolkit and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) e215–e220.