Multiscale structure of time series revealed by the monotony spectrum

Călin Vamoş*
"T. Popoviciu" Institute of Numerical Analysis, Romanian Academy, P.O. Box 68, 400110 Cluj-Napoca, Romania

Observation of complex systems produces time series with specific dynamics at different time scales. The majority of the existing numerical methods for multiscale analysis first decompose the time series into several simpler components and the multiscale structure is given by the properties of their components. We present a numerical method which describes the multiscale structure of arbitrary time series without decomposing them. It is based on the monotony spectrum defined as the variation of the mean amplitude of the monotonic segments with respect to the mean local time scale during successive averagings of the time series, the local time scales being the durations of the monotonic segments. The maxima of the monotony spectrum indicate the time scales which dominate the variations of the time series. We show that the monotony spectrum can correctly analyze a diversity of artificial time series and can discriminate the existence of deterministic variations at large time scales from the random fluctuations. As an application we analyze the multifractal structure of some hydrological time series.

I. INTRODUCTION

By observing natural complex systems with a hierarchical structure, such as those in geophysics, biology, finance, ecology, internet traffic, etc., one obtains time series which reflect their dynamics at different time scales [1]. Therefore a correct multiscale analysis of such time series is essential to design a proper model of complex systems [2,3].

The majority of the multiscale analysis methods first decompose the time series in simpler components which are associated with the time scales. For instance, the classical spectral Fourier method decomposes time series into a sum of trigonometric components while the multiresolution analysis uses wavelets [3]. Other methods, such as the empirical mode decomposition (EMD), build intrinsic mode functions (IMFs) which are not predefined and are applicable to arbitrary nonlinear and nonstationary time series [4–8]. An IMF is a simple oscillatory mode satisfying two conditions: (i) between two successive zeros there is a single local extremum; (ii) the envelopes of the local maxima and minima are symmetric. Hence it is a zero-mean oscillatory wave form modulated in both amplitude and frequency.

Obviously, the results of such multiscale analyses depend on the particular decomposition on which it is based. It is also possible to obtain information about the multiscale structure of a time series without decomposing it. For instance, the visibility graph algorithms map time series into graphs in order to apply graph theory to the time series analysis [9,10]. It uses a very general approach based on the relative positions of the local extrema.

The multiscale analysis method presented in this paper relies on the variation of monotony properties of time series during a succession of smoothings. The *local time scale* (LTS) is the length of a monotonic segment (it is also known as "characteristic time scale" [4]). The global spectral characterization of a time series is given by the mean LTS (MLTS). The successive averagings eliminate progressively the small LTSs until a monotonic averaged time series is obtained. The multiscale structure is described by the *monotony spectrum* (MS) which contains the variation of the mean amplitude of the monotonic segments as a function of the MLTSs for the averagings which reduces the number of monotonic segments. The maxima of the MS indicate the MLTSs which dominate the variations of the time series. This method does not depend on any preliminary hypothesis or any parameter to which we have to give a subjective, semiempirical value.

The MS is a condensed description of the monotony properties of a time series. A more detailed description is given by the *monotony graph* defined as the plot displaying the amplitudes of all the monotonic segments against their length. We used the detailed information contained in the monotony graphs to decompose a complex financial time series into intrinsic components with minimal spectral superposition [8]. First we separated the component with the maximum time scale (the monotonic component if it exists) and then we determined the other components such that their monotony graphs are disjoint one from another and from that of the first component. This algorithm was developed into an automatic one which can determine the component with the maximum monotonic variation [11]. We used it to decompose time series with self-similar properties such as the fractional Brownian motion into a self-similar set of components.

II. MONOTONY SPECTRUM

Let us consider a time series with equidistant sampling $\{x(n), n = 1,2,\ldots,N\}$. We denote by I_j, $j = 1,2,\ldots,J$ its monotony intervals. If n_j, $j = 1,2,\ldots,J+1$ are the positions of its local extrema, then $I_j = [n_j, n_{j+1})$ with the exception of the last interval which is closed at both ends, $I_J = [n_J, n_{J+1}]$. For a given time step $n \in I_j$, we define the LTS as the length of that interval $\theta_j = n_{j+1} - n_j$. The *amplitude of the local variation* is $a_j = |x(n_{j+1}) - x(n_j)|$. The *monotony graph* of the time series is the graph of all the couples (θ_j, a_j). We obtain a global description by averaging over the entire time

*cvamos@ictp.acad.ro

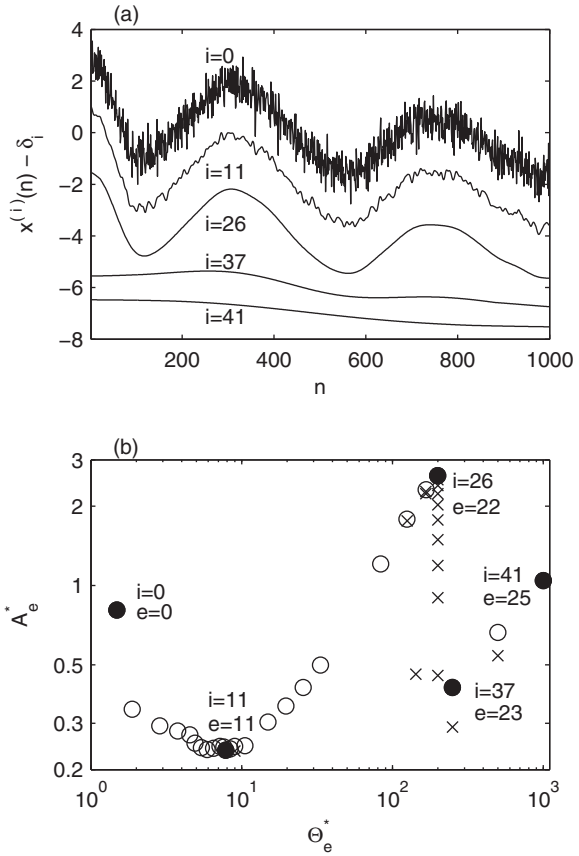

FIG. 1. (a) A noisy time series and some of its averaged forms. For the readability of the graph the plots are displaced vertically by the quantities $\delta_i \in \{2,4,6,7\}$. (b) MS of the time series in panel (a) ($\circ$ markers) and the noneffective MMVs ($\times$ markers). The local extrema ($\bullet$ markers) correspond to the smoothed time series plotted in panel (a).

series the local quantities defined above. The *mean local time scale* (MLTS) is equal with $\Theta = (N-1)/J$. The mean of a_j is denoted A. The *mean monotonic variation* (MMV) is the couple (Θ, A).

If the time series is a superposition of components with different time scales, then the MMV characterizes only the component with the smallest time scales. As an example we analyze the time series in Fig. 1(a) composed by a Gaussian white noise $\{z(n)\}$ superposed on a deterministic trend $\{f(n)\}$, $x(n) = f(n) + z(n)$. The trend has $J = 5$, $103 \leqslant \theta_j \leqslant 273$, and $(\Theta^{(f)}, A^{(f)}) = (199.8, 2.78)$ and for the noise $J = 674$, $1 \leqslant \theta_j \leqslant 5$, and $(\Theta^{(z)}, A^{(z)}) = (1.4822, 0.8056)$. (For an infinite Gaussian white noise, the probability that at a given time there exists a local time extremum is $2/3$ and then the mean of J is $2N/3$.) This is a two-level time series with two dominant time scales related to the trend and the noise. For $\{x(n)\}$ we have $J = 672$, $1 \leqslant \theta_j \leqslant 5$, and $(\Theta^{(x)}, A^{(x)}) = (1.4866, 0.8083)$, values almost identical with those for the noise taken separately.

One can make apparent the larger scale variations of such a noisy time series by successive smoothings with increasing strength [12]. We use a succession of averagings $i = 0,1,2,\ldots,I$ and we denote by $\{\overline{x}^{(i)}(n)\}$ the time series averaged i times with the convention $\overline{x}^{(0)}(n) = x(n)$. The ith moving average is defined as

$$\overline{x}^{(i)}(n) = \frac{1}{2K_i + 1} \sum_{k=-K_i}^{K_i} \overline{x}^{(i-1)}(n+k), \tag{1}$$

where $K_i > 0$ and $K_i < n \leqslant N - K_i$. If $n \leqslant K_i$ ($n > N - K_i$), then the average is taken over the first $n + K_i$ (the last $N - n + K_i + 1$) values of $\overline{x}^{(i-1)}(n)$. The properties of this moving average are analyzed in [13]. The averagings are stopped when there remains a single monotonic segment $J_I = 1$ and $\Theta_I = N$. If we know that the time series does not contain a monotonic component, then we can reduce the computing time by stopping the averagings when A_i becomes smaller than 0.1 of the global amplitude of the averaged time series.

There are averagings which only damp the variations of $\{\overline{x}^{(i)}(n)\}$ determining the decrease of A_i while J_i and Θ_i remain constant. The useful averages for the multiscale analysis are those which also modify the shape of the averaged time series by reducing the number J_i of the monotonic parts. We call them *effective averagings*. After an effective smoothing, several intervals I_j with small amplitudes are eliminated and their neighboring monotonic segments are joined. If in the range of such monotonic intervals there is a monotonic variation at a larger scale, then the resulting new monotonic segment has a greater amplitude than those of the preceding ones and it contributes to the increase of the mean amplitude A_i.

We denote by $E \leqslant I$ the number of the effective averagings so that their indices are i_e, $e = 0,1,2,\ldots,E$. We introduce the notations $J_e^* = J_{i_e}$, $\Theta_e^* = \Theta_{i_e}$, and $A_e^* = A_{i_e}$. Then Θ_e^* is the *effective MLTS* (effMLTS) and (Θ_e^*, A_e^*), is the *effective MMV* (effMMV) for the eth effective moving average. For the initial unsmoothed time series ($i = 0$) we do not have a previous value J_{i-1} and we make the convention $i_0 = 0$ so that $(\Theta_0^*, A_0^*) = (\Theta_0, A_0)$. (A list of these notations is given in Appendix B.) When the averagings are stopped $\Theta_E^* = \Theta_I = N - 1$.

In order to obtain an optimal computing time, we have to choose the succession of averagings such that the number of inefficient averagings is reduced as much as possible. As a general rule, we increase the averaging window by ΔK_i at the ith averaging, $K_i = K_{i-1} + \Delta K_i$, because after the elimination of the small scale fluctuations by the first smoothings, the remaining variations have larger scales and they need stronger smoothings to be damped. Because initially we do not know if the time series contain small scale fluctuations, during the first averaging $\Delta K_i = 0$ and $K_i = 1$ is kept at its minimum value. When for the first time the averaging is not effective, we make $\Delta K_i = 1$. Numerical tests showed that we can increase K_i faster than keeping $\Delta K_i = 1$: when the averaging is not effective we add a unity to ΔK_i, $\Delta K_{i+1} = \Delta K_i + 1$. In the case of the time series in Fig. 1(a), for the first 15 averagings $K_i = 1$ and it is increased gradually to $K_I = 172$ after $I = 41$ averagings.

We define the *monotony spectrum* (MS) of a time series as the graph of the couples (Θ_e^*, A_e^*), $e = 0,1,2,\ldots,E$. In previous papers [8,11] we used the designation of MS for the graph of (θ_j, a_j) which is called above monotony graph. We change the terminology because a spectrum usually describes

the distribution of the properties of a time series over several time scales and such information is obtained only by successive smoothings.

The MS of the time series in Fig. 1(a) is plotted in Fig. 1(b). The initial averagings determine the decrease of A_e^* as the noise is damped. The time series averaged $i = 11$ times corresponding to the minimum of A_e^* contains obvious fluctuations but with reduced amplitudes and longer LTSs compared with the initial time series [Fig. 1(a)]. All these averagings are effective such that $e = i = 11$. The following averagings eliminate these residual fluctuations and, due to the existence of the deterministic trend, A_e^* increases. When A_e^* reaches its maximum ($e = 22$ and $i = 26$), all the small scale variations are eliminated and the averaged time series has only $J_{22}^* = 5$ monotonic parts approximating very well the trend [Fig. 1(a)]. The next ten smoothings are not effective and the next efficient averaging eliminates one of the trend oscillations [$e = 23$ and $i = 37$ in Fig. 1(a)]. After $i = 41$ averagings, from which $E = 25$ are effective, a single monotonic segment remains.

This simple example illustrates the ability of the MS to reflect multiscale structures. The MS in Fig. 1(b) has three maxima corresponding to the MLTSs dominating the time series variations. The first maximum at $\Theta_0^* = 1.49$ is obtained for the unsmoothed time series and it corresponds to the noise fluctuations. The global maximum at $\Theta_{22}^* = 199.8$ corresponds to the oscillations of the deterministic trend with $J_{22}^* = 5$ monotonic segments. The two maxima are separated by almost two orders of magnitude of MLTS indicating that the analyzed time series is a superposition of two independent components. The last maximum at $\Theta_{25}^* = N - 1 = 999$ is related to the monotonic component of the trend.

As a conclusion, we may say that the maxima of the MS are the main characteristics of the *multiscale* or *multilevel structure* of a time series. The MLTSs corresponding to these maxima are called *dominant time scales*. The time series in Fig. 1(a) has a three-level structure with the greatest maximum corresponding to the deterministic trend. This characterization is in accordance with the intuitive perception that the deterministic trend is clearly visible and the noise has a smaller importance in the structure of this time series.

III. MULTISCALE STRUCTURE OF ARTIFICIAL TIME SERIES

In this section we present multiscale analyses using the MS for several types of time series: superposition of periodic signals, Gaussian white noise, red noise, deterministic trend with white noise, discrete fractional Brownian motion (fBm), and multifractal time series. For noisy time series we also make Monte Carlo simulations. We qualitatively compare our results to those of the EMD method, wavelet multiresolution, and Fourier analysis.

A. Deterministic time series

As an example of a deterministic time series we use the trend contained in the noisy time series analyzed in the previous section. Its MS has only two maxima very close to those of the MS of the noisy time series located at large time scales (Fig. 2). The main difference between the two MSs is the lack in the MS of the trend of the small time scales associated with the noise. The mean local amplitudes of the MS of the trend are slightly larger because the number of the averagings is smaller than when the noise has to be smoothed.

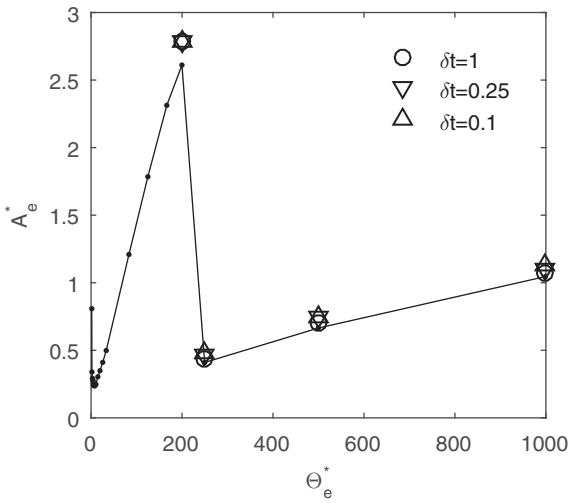

FIG. 2. MSs of the trend of the time series in Fig. 1(a) for three different samplings characterized by their time steps. The continuous line with point markers is the MS in Fig. 1(b).

Figure 2 also shows that the MS is not affected by changing the sampling rate of the deterministic trend with an order of magnitude. The main difference in the computation of the MS when the time series is resampled occurs due to the fact that the averaging window K_i is expressed as the number of time steps which change with the sampling. However, the effMMVs of the three samplings are almost identical.

An important problem related to deterministic time series is the ability of the numerical algorithm to separate two spectral components with comparable frequencies. Let us consider the continuous-time signal,

$$x(t) = \cos 2\pi t + a\cos(2\pi f t + \phi), \tag{2}$$

where $t \in [0,10]$ and $f < 1$. We discretize $x(t)$ with a time step $\delta t = 0.01$, so that $N = 1000$. It was proved theoretically that the EMD method can separate the two components only if $fa < 1$ [14,15]. But numerically the separation is achieved only if the additional condition $f < 0.55$ holds.

We consider time series given by Eq. (2) with $a = 1$, $f = 30/34$, and different values of the phase ϕ. This signal was given as an example for which the EMD method cannot separate the components [4]. The MSs of all these time series have a two-level structure indicating the presence of the two components [Fig. 3(a)]. The frequencies of the two components are estimated as $0.5/\Theta_e^*$ because for a sinusoid Θ_e^* is equal to its semiperiod. When $fa < 1$, the MS has two maxima for $f < 0.95$, a threshold much larger than that of the EMD method ($f < 0.55$).

For nonsinusoidal periodic components, similarly to the periodic case, their separation depends on the number and positions of the local extrema, but now the theoretical approach cannot give sufficient information [14]. A numerical analysis

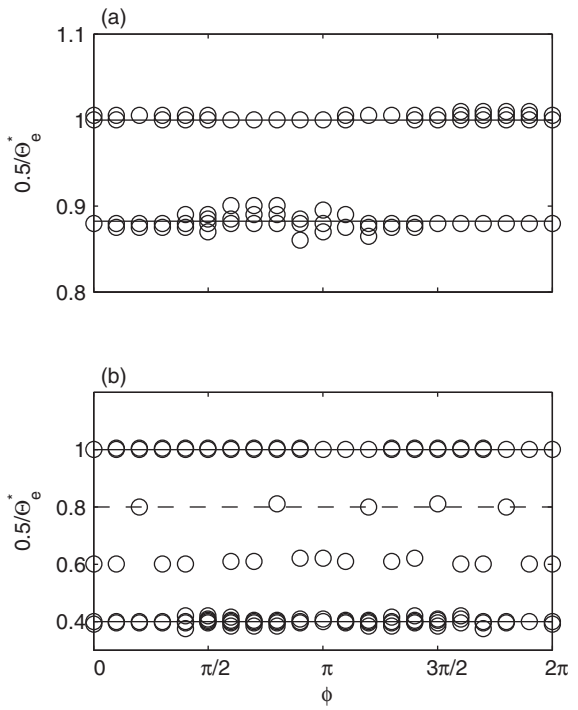

FIG. 3. Frequencies estimated by the MS for the sinusoidal (a) and nonsinusoidal (b) components, as a function of ϕ. The continuous lines mark the principal frequencies and the dashed line one of the secondary frequencies of the nonsinusoidal component.

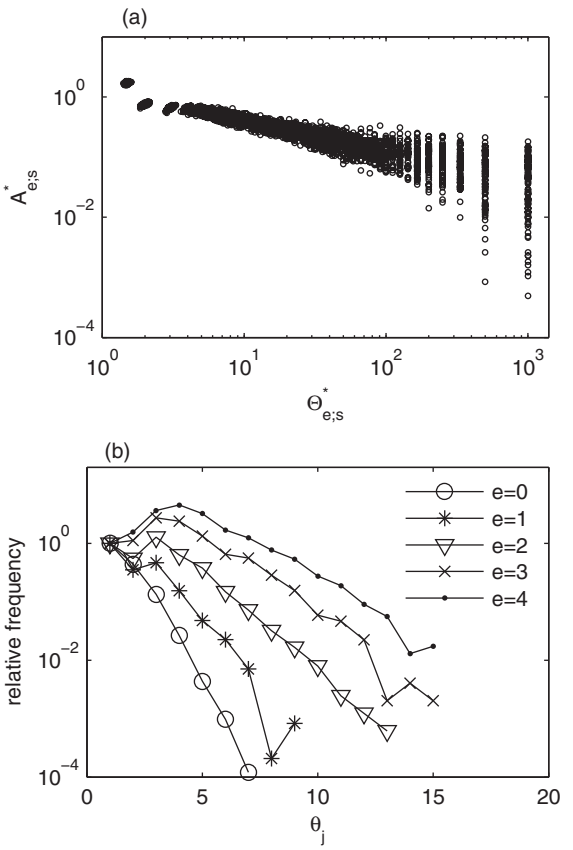

FIG. 4. (a) Global MS of 100 Gaussian white noises. (b) Relative frequencies of the number of the monotonic segments of a Gaussian white noise with $N = 20\,000$ with respect to their lengths for the first four effective averagings.

was performed for the signals,

$$x_1(t) = \cos(2\pi t) + 0.15\cos(6\pi t),$$
$$x_2(t) = \cos(2\pi t) + 0.15\cos(4\pi t) - 0.15, \qquad (3)$$

which are periodic, but not sinusoidal. When the condition $fa < 1$ is satisfied, the EMD method can decompose the signal $x_1(t) + ax_2(ft + \phi)$ only if $f < 0.32$ [14]. The MS indicates the existence of the two components for much larger frequencies, for example, for $f = 0.4$ [Fig. 3(b)]. It also shows that there exists a secondary component between the two main ones, related to the frequency $2f$ contained in the component $ax_2(ft + \phi)$. The MS cannot identify the frequency 3 contained in $x_1(t)$.

B. Gaussian white noise

We generated $S = 100$ Gaussian white noises with zero mean and unit variance with $N = 1000$ values and for each of them we computed the MS $(\Theta^*_{e;s}, A^*_{e;s})$, $e = 0,1,2,\ldots,E_s$, and $s = 1,2,\ldots,S$. All the effMMVs are plotted together in Fig. 4(a), so that we obtain a *global MS* characterizing the entire statistical ensemble. The global MS in Fig. 4(a) has a monoscale structure with the dominant time scale equal to the smallest MLTS. From the minimum MLTS to the maximum one the average amplitude monotonically decreases. Hence each new averaging only diminishes the existing monotonic variations because there are no trend or correlations which could be revealed by the successive averagings. The average amplitude drops abruptly after the first averaging and then for $e > 2$ the global MS has a power law dependence which in the loglog plot of Fig. 4(a) becomes linear.

In order to explain why the power law is not satisfied for $e \leqslant 2$, we consider a single Gaussian white noise with $N = 20\,000$ values. In this way we obtain a significant statistics of monotonic segments. In Fig. 4(b) we plot the relative histograms of θ_j for the first four efficient averagings. For an easier comparison of the shapes of the histograms, they are rescaled by the sample frequency for $\theta_j = 1$. For the initial time series ($e = 0$) the distribution is dominated by the monotonic segments with a single time step ($\theta_j = 1$) and it decreases monotonically with respect to θ_j. For the first two averagings ($e = 1,2$) there are two maxima, one at $\theta_j = 1$ and the other at $\theta_j = 3$. After that ($e \geqslant 3$) the distributions have only one maximum for $\theta_j > 1$ and their shape is preserved. The existence of this maximum indicates that by successive averagings we have introduced a correlation between the values of the averaged time series.

In fact, any smoothing induces a spurious correlation in the smoothed time series. Because the first smoothings are always made with $K_i = 1$ constant, we consider the averaging (1) in this simple case. We denote by $w_k^{(i)}$, $-i \leqslant k \leqslant i$, the weighting coefficients corresponding to the ith moving average

$$\overline{x}^{(i)}(n) = \sum_{k=-i}^{i} w_k^{(i)} x(n+k)$$

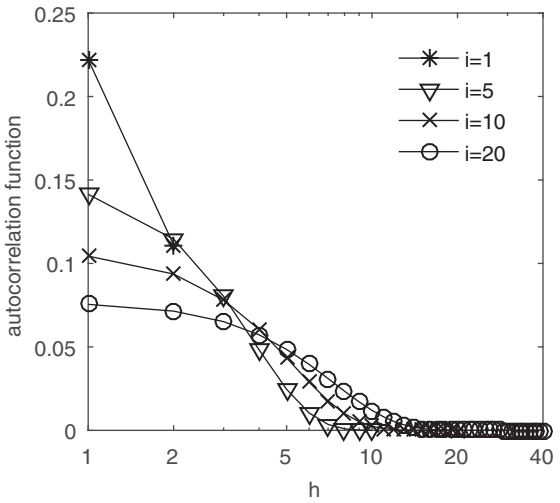

FIG. 5. Autocorrelation function induced on an i.i.d. time series by the averaging (1) with $K_i = 1$ for several values of i.

for $i < n < N - i$. By a simple calculation [13, Appendix B], the autocorrelation induced by this averaging on an i.i.d. (independent and identically distributed) time series with unit variance is given by

$$\gamma^{(i)}(h) = \sum_{k=-i}^{i-h} w_k^{(i)} w_{k+h}^{(i)}$$

for $h \leqslant 2i$ and zero for $h > 2i$. Figure 5 shows the autocorrelation function for several values of i. One notices that as the averagings accumulate, the correlation extends to larger time lags and the correlation becomes weaker. Hence the influence of the autocorrelation induced by averagings on the MS is limited at the smallest time scales. In the case of the Gaussian white noise this influence is noticeable for the first three averagings [Fig. 4(a)]. When the time series is not i.i.d. and contains an intrinsic autocorrelation (deterministic trend, autoregressive law, etc.), the influence of the autocorrelation induced by averagings becomes even weaker.

Because the points of the global MS are dispersed, in order to determine the power law associated to it we introduce the *average MS* $(\langle\Theta^*\rangle_k, \langle A^*\rangle_k)$ obtained by averaging the effMMVs which have $\Theta^*_{e;s}$ in K bins (see Appendix A). The resulting average MS is plotted with cross markers in Fig. 6(a). The straight line determined by linear regression for $\langle\Theta^*\rangle_k > 4$ has a slope equal to $b \approx -0.45$. We have chosen the values of S and K such that each bin contains enough values to make the relative standard deviation of $\langle A^*\rangle_k$ smaller than 0.03 [Fig. 6(b)]. Therefore, in this section and in the following ones, we plot the average MSs without error bars for standard deviations.

The shape of the average MS of the Gaussian white noise is similar to that of the mean power spectrum of the EMD [16]. In both cases, the amplitude (MS) and the energy (EMD) decrease by a power law with respect to the time scale. Such a behavior is related to the fact that a Gaussian white noise is predominantly composed by fluctuations lasting only a few time steps. On the other hand, a classical result states that the power spectrum of a Gaussian white noise has a constant value [17,18]. The conclusion we can draw from this result is that the harmonic components and the wavelets are inappropriate to describe the local structure of the Gaussian white noise because they have a smooth, deterministic variation and they approximate poorly the random local fluctuations.

C. Autoregressive red noise

In this section we analyze the average MS of a particular case of red noise. We numerically generated realizations $\{z_n\}$ of a stationary stochastic process $\{Z_n\}$ of type AR(1) (autoregressive of order 1). It is defined by

$$Z_n = \phi Z_{n-1} + G_n, \tag{4}$$

where $|\phi| < 1$ and $\{G_n\}$ is a Gaussian white noise with zero mean and unit variance. The properties of the AR(1) process are well known [17]. When the value of the parameter ϕ increases, the successive values of the noise become more correlated. In our numerical tests $\phi \in [0,0.99]$. For $\phi = 1$ Eq. (4) defines the usual Brownian motion.

Figure 6(a) shows the average MSs for five values of ϕ. The average MS of the Gaussian white noise ($\phi = 0$) follows the power law presented in the previous section. For $\phi < 0.7$ only the difference between the first two amplitudes $\langle A^*\rangle_1 - \langle A^*\rangle_2$ becomes smaller so that the rest of the average MS remains parallel to that for $\phi = 0$ [these average MSs are not plotted in Fig. 6(a)]. For $\phi \geqslant 0.7$ the average MS acquires a maximum which is located at larger MLTSs when ϕ is larger. If the time series are longer ($N > 1000$), then the maximum remains at the same position and the linear part of the average MS lengthens keeping the same slope. In the case of the Brownian motion ($\phi = 1$) the average MS has a constant positive slope.

In conclusion, for $\phi < 0.7$ the average MS of the AR(1) noise preserves the monoscale structure of the Gaussian white noise. For $\phi > 0.7$ a second maximum appears at intermediate time scales and the average MS acquires a two-level structure similar to that of a deterministic trend with a noise superposed (Sec. II). When $\phi > 0.9$, the average MS has a monoscale structure with a large dominant time scale. This behavior is similar to the EMD power spectrum [19]. As for the Gaussian white noise the local wavelet power spectrum and Fourier power spectrum have a different energy distribution over time scales [14,18].

D. Noise superposed on a deterministic trend

The shape of the MS of the Gaussian white noise is modified by the existence of a deterministic trend. We exemplify this situation by the artificial time series numerically generated by the method described in [13,20]. We construct a nonmonotonic trend from random monotonic semiperiods of sinusoid joined together such that the trend is continuous. The trend is characterized by three parameters: the length of the time series N, the number of monotonic segments P, and the minimum number of points in a monotonic segment $\Delta N_{\min}$. We superpose on this trend a white Gaussian noise with the amplitude q times the amplitude of the trend variations. Figure 1(a) shows such an artificial time series with $N = 1000$, $P = 5$, $\Delta N_{\min} = 50$, and $q = 1.5$. When $q = 0$ the time

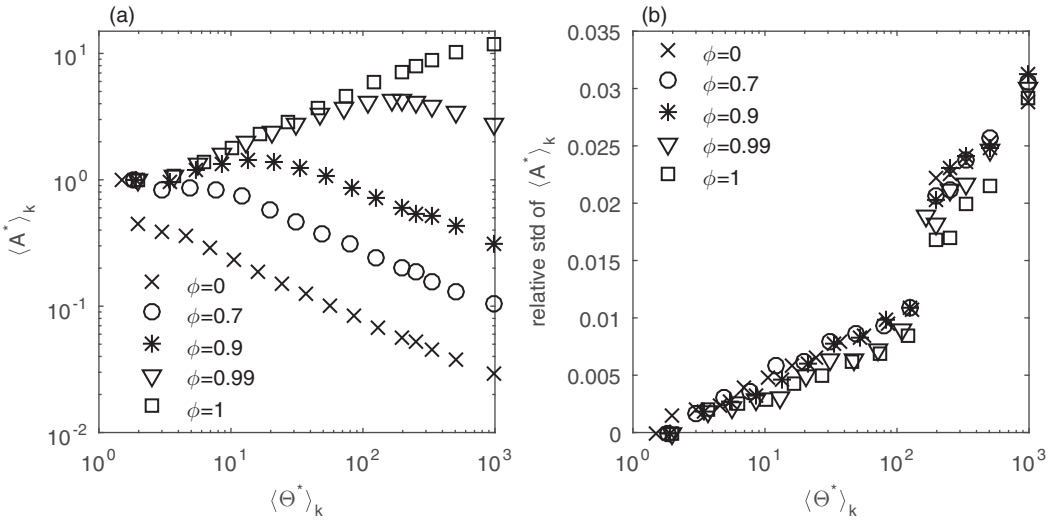

FIG. 6. (a) The average MSs of statistical ensembles with 1000 time series of AR(1) type with different values of $\phi < 1$ and 1000 Brownian motions ($\phi = 1$). (b) The relative standard deviations of $\langle A^* \rangle_k$.

series contains only noise and when $q \to \infty$ we have only a deterministic trend.

The results obtained for statistical ensembles of 1000 time series with $N = 1000$, $P = 10$, $\Delta N_{\min} = 50$ and different values for q are plotted in Fig. 7. For all the values of $q > 0$ the average MS has a two-level structure, not a three-level structure as that of the time series in Fig. 1(a). Because the graphs in Fig. 7 are averages over statistical ensembles of individual MSs, their shapes at large MLTS are strongly smoothed and they do not resemble anymore that in Fig. 1(b). The individual extrema are averaged over the statistical ensembles and the result is a monotonic increase. The length of this region becomes larger when q increases, i.e., when the influence of the trend is stronger.

At small MLTS the average MSs for $q > 0$ coincide with the MS of Gaussian white noises (x markers in Fig. 7) proving that at small MLTS it is dominated by noise. When the amplitude of the noise is equal to that of the trend variations ($q = 1$) this coincidence holds for $\langle \Theta^* \rangle_k < 10$. If the time series is dominated by noise ($q = 0.1$), then the interval of coincidence with the Gaussian white noise is longer, while if the time series is dominated by trend ($q = 10$), the coincidence interval is shorter. Hence, in average, at small time scales this type of time series is dominated by noise, while at large time scales it is dominated by the trend.

The MS method cannot extract the trend from a noisy time series. It only indicates the existence of a deterministic component and estimates its statistical significance. Then one may use any of the many numerical methods to estimate the deterministic trend from a noisy time series [3,21], including those based on the EMD method [22].

In Fig. 8 we analyze the average MS of the deterministic signal composed by the components given by Eq. (3) with

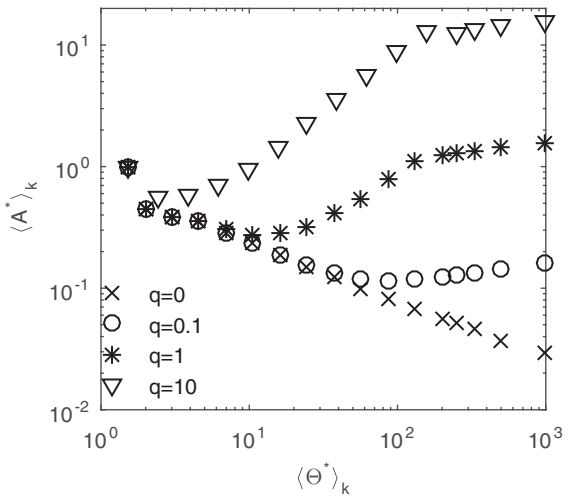

FIG. 7. The average MSs of statistical ensembles with 1000 time series obtained by superposing a Gaussian white noise on a deterministic trend with different amplitudes.

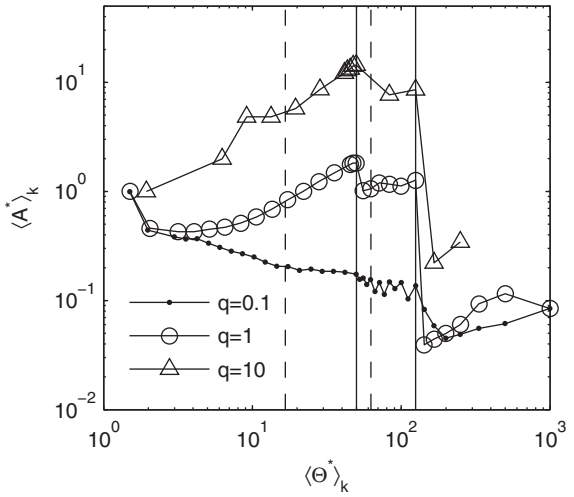

FIG. 8. The average MSs of the signal in Fig. 3(b) for $\phi = 0$ on which Gaussian white noises with different amplitudes are superposed. The continuous lines mark the semiperiods of the largest terms in Eq. (3), while the dashed lines mark those of the small terms.

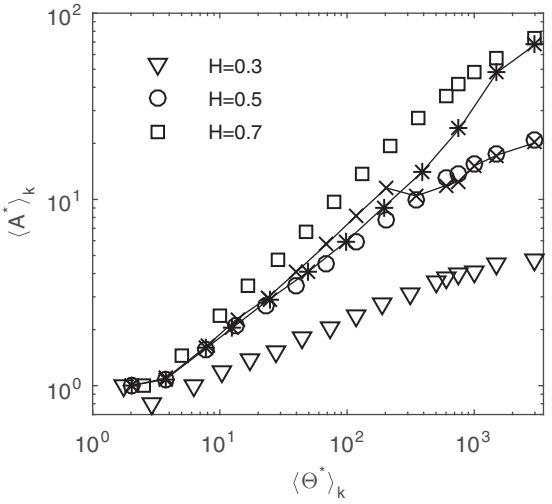

FIG. 9. The average MSs of statistical ensembles with 1000 discrete fBms with different values H and of two Brownian motions ($H = 0.5$) on which a linear trend (asterisk markers) and a sinusoid (x markers) are superposed.

a Gaussian white noise superposed on it. When the noise amplitude is very small ($q = 10$) the presence of the two components is identified by the two maxima located at the correct semiperiods. The two maxima are clearly visible if the noise has the same amplitude as the deterministic signal ($q = 1$), but a third maximum appears at the smallest time scale related to the noise presence. When the noise is much larger than the deterministic signal ($q = 0.1$), only the maximum at the smallest time scales remains. However, between the two principal semiperiods of the deterministic signal (the continuous vertical lines in Fig. 8) the average MS contains fluctuations as remnants of the influence of the two IMF components. Our numerical simulations showed that the MS can correctly identify the two components for $q > 0.3$.

E. Discrete fractional Brownian motion

The fBm is a continuous stochastic process [23] and we generated its discrete paths by means of the algorithm based on the wavelet decomposition [24]. We computed the average MSs for statistical ensembles of 1000 discrete fBms with $N = 3000$ and three different values of the Hurst exponent H (Fig. 9). They have a monoscale structure with the maximum located at the largest MLTS, each new averaging revealing variations with larger amplitudes due to the nonstationary nature of the fBm.

If we increase the length N of the time series, then the average MS maintains its linear shape in a loglog representation and the dominant time scale becomes equal to the new value of N. Hence, in this case, the dominant scale is not specific to the multiscale structure of the time series, its value being a consequence of the self-similarity of the time series. Except for the very large and very small $\langle \Theta^* \rangle_k$, the average MS has a power law variation with the exponent very close to the Hurst exponent: for $\langle \Theta^* \rangle_k \in [4,100]$ we obtain $b = 0.298$, $b = 0.496$, and $b = 0.686$ when $H = 0.3$, $H = 0.5$, and $H = 0.7$, respectively.

We extended this numerical experiment by superposing a linear trend on a Brownian motion ($H = 0.5$). In this case, the variation of $\langle A^* \rangle_k$ is almost identical to that of the Brownian motion for $\langle \Theta^* \rangle_k < 100$ (Fig. 9). When $\langle \Theta^* \rangle_k > 100$ they gradually depart from each other and the average MS for the modified Brownian motion becomes close to the average MS of the fBm with $H = 0.7$. Hence the linear trend with a single LTS equal with N influences the average MS not only at $\langle \Theta^* \rangle_k = N$, but also at intermediate scales.

We also superposed on the Brownian motion ($H = 0.5$) a sinusoid with the period equal to 500 time steps. The corresponding average MS of a statistical ensemble with 1000 members (Fig. 9) has a clear deviation from that of the Brownian motion at time scales smaller than 200 time steps, i.e., approximately half the period of the sinusoid.

These two examples show that the average MS can indicate the existence of deterministic components superposed on a Brownian motion and in general on a fBm. They can be recognized in the MS because they disturb the power law characteristic to a time series with long memory. This particularity could be used as a complement for algorithms such as the detrended fluctuation analysis (DFA) which estimate Hurst exponent even in the presence of a trend [25,26]. The DFA fits a straight line to the fluctuation function in a log-log representation, but it is important that the existence of a scaling relation characteristic to long memory is previously ascertained [27]. The MS could help to determine the correct model of the time series, because it does not make any preliminary assumption and it is not restricted by nonstationarity or nonlinearity. However, we remark that there are essential differences between MS and DFA. We have shown that even for simple time series the MS has an opposite behavior to that of the power spectrum (Secs. III B and III C) and theoretical and numerical results show that DFA is equivalent to spectral measures for long-memory time series [28].

F. Multifractal time series

As a multifractal structure we analyze the time series describing the transverse component of a passive transport process in a random velocity field. We consider a typical model of the hydraulic conductivity for saturated aquifers consisting of a two-dimensional statistically homogeneous space random function, with exponential correlation of correlation length z_c and variance σ^2 [29,30]. Details of the numerical algorithm used to generate the time series can be found in [11,31].

In Fig. 10 we plot the average monotonic spectra obtained for statistical ensembles of 1000 time series for three values of z_c and σ^2. These average MSs do not present self-similarity as the discrete fBm does (Fig. 9), however, those with $z_c = 1$ m have the loglog slopes only slightly varying, especially for the minimum and maximum time scales. For instance, when $\sigma^2 = 0.15$ the global slope is 0.3650, while for its most linear part ($\langle \Theta^* \rangle_k \in [10,1000]$) the slope is 0.3491. When $\sigma^2 = 12$ the global slope is 0.4728, while for its most linear part ($\langle \Theta^* \rangle_k \in [10,330]$) the slope is 0.4881. For $z_c = 30$ m the slope of the MS has large variations and over the approximately linear part ($\langle \Theta^* \rangle_k \in [10,200]$) the slope is 0.8973.

The multiscale structure of the hydrological time series determined using the MS is similar to that obtained with the

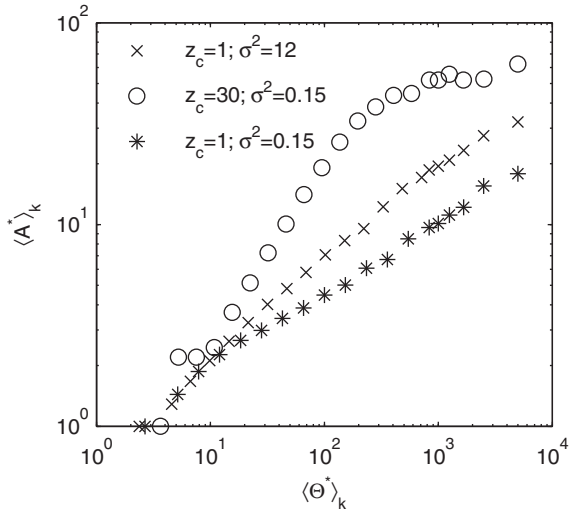

FIG. 10. Three average MSs of statistical ensembles of multifractal time series.

maximum monotonic variation [11]. The centers of mass of the monotony graphs of the intrinsic components were aligned on curves arranged in the same order as in Fig. 10. Their slopes in loglog representation were slightly larger, for instance, the slopes of their linear parts were 0.4946, 0.5701, and 0.9502.

Although the average MSs in Fig. 10 are not self-similar, their dominant time scale is always equal to the length of the time series. Hence each averaging reveals monotonic variations at larger time scales with larger amplitudes than those found previously. But the slopes of the average MSs are not constant, behavior specific for multifractal time series [32]. The same type of variable slope occurs when a deterministic trend is superposed on an fBm (Fig. 9). In the case of the hydrological time series a deterministic trend cannot exist, because they are generated by a random numerical algorithm without any deterministic element. In general, in order to determine the type of a time series, besides the results of some numerical analysis method, we need theoretical or experimental indications on the mathematical model behind observation data.

In practice it is usually necessary to analyze a single time series and then a statistical test has to be designed in order to select the appropriate mathematical model. For example, the existence of nonrandom components in a nonlinear geophysical time series was proved by comparing the observed IMFs with those of an ensemble of Gaussian white noises [16]. The MS could be used in a similar manner to test the existence in a financial time series of large scale nonrandom variations and of business cycles [33].

IV. CONCLUSIONS

The MS quantifies the evolution of the mean monotonic variation of a time series during a succession of smoothings designed to eliminate gradually the fluctuations, beginning with those at the smallest time scales. The repartition of the maxima and minima of the MS describes the multiscale structure of the time series. We have proved the ability of the MS to reveal the correct multiscale structure of different types of artificial time series: superposition of periodic signals, Gaussian white noise with and without trend, AR(1) noise, fractional Brownian motion with and without trend, and multifractal time series.

These results have been obtained without any assumption regarding the stationarity or the linearity of the time series. Hence the MS method has the same generality as the EMD method, but it analyzes time series without decomposing them. Therefore it can be applied to identify the main characteristics of a time series such that the most appropriate algorithm can be chosen to decompose it.

The MS gives a global analysis of a time series, because it is an average representation of a much more detailed description. When we construct the MS we know all the values of LTSs $\theta_j^{(i)}$ and amplitudes $a_j^{(i)}$ for all the averagings $i = 1,2,\ldots,I$ and all the monotony intervals $I_j^{(i)}$, $j = 1,2,\ldots,J_i$. In addition, each interval $I_j^{(i)}$ has a given position in the time series. This information can be used to construct a monotony spectrogram organized as a three-dimensional graph. At each time step n on the abscissa we associate on the ordinate all the time scales $\theta_j^{(i)}$ for which $n \in I_j^{(i)}$ and the corresponding amplitudes $a_j^{(i)}$ on the vertical coordinate. In comparison with the Fourier spectrogram or the Hilbert spectrogram the frequency is replaced by the local time scale and the energy by the local amplitude.

The shortcoming associated with the generality of the LTS is its weak temporal resolution. One can increase the temporal resolution, but then the generality is restricted. For example, the instantaneous frequency can be computed by applying several methods: Hilbert transform, zero crossing, Teager energy operator, etc. But an extended comparative study of the different definitions of the instantaneous frequency [34] shows that the accuracy of the results of these numerical methods depends essentially on the conditions imposed by the Bedrosian [35] and Nuttall [36] theorems. In wavelet analysis, the time scale is given by the scaling coefficient which describes the dilation transformations [3,18,37]. But in this way the shape of the wavelets is neglected even if they present internal oscillations as, for example, Morlet wavelets. It is also difficult to define a frequency analogous to that in Fourier analysis [4,18,38,39].

APPENDIX A: AVERAGE MONOTONY SPECTRUM OF A STATISTICAL ENSEMBLE

The global MS of a statistical ensemble of time series has some characteristics that are not observable in the MS of a single time series as that in Fig. 1(b). At small and large MLTSs the effMMVs $(\Theta^*_{e;s}, A^*_{e;s})$ arrange themselves on specific formations [Fig. 4(a)]. This behavior occurs identically for all statistical ensembles studied in this paper.

At large MLTS the couples $(\Theta^*_{e;s}, A^*_{e;s})$ form vertical segments because the number of the monotonic segments $J_{e;s}$ is an integer, so that the effective MLTS takes only the discrete values $\Theta^*_{e;s} = N/J_{e;s}$. The maximum value $\Theta^*_{E_s;s} = N$

is obtained for the last averaging when the averaged time series is monotonic ($J_{E_s;s} = 1$). The previous value $\Theta^*_{E_s-1;s} = N/2$ occurs when $J_{E_s-1;s} = 2$ and the distance on the abscissa in log coordinates between the two vertical segments is $\log_{10}(N/2)$. When $J_{e;s}$ increases, these distances decrease. For the large values of $J_{e;s}$ the values of $\log_{10}\Theta^*_{e;s}$ are separated by so small intervals that the corresponding vertical segments cannot be perceived separately on the graph.

The other alignment of the effMMVs at small $\Theta^*_{e;s}$ is due to the discrete variations of the averaging window K_i in Eq. (1). For small values of K_i these jumps are perceptible and the effMMVs separate from each other. These discrete variations could be eliminated using a smooth averaging kernel [40], but the results are not essentially changed.

In order to determine the average MS we have to establish how the couples $(\Theta^*_{e;s}, A^*_{e;s})$ can be grouped such that the structure of the global MS is followed. The difficulty consists of the fact that at the two extremities the couples cluster themselves after different laws, while at intermediate MLTS they are homogenously distributed. Before proceeding to their grouping, we rescale the mean amplitude of local variations to unity $(\Theta^*_{e;s}, A^*_{e;s}/A^*_{0;s})$. In this way we compare only the shapes of the MSs of individual time series, eliminating the possible global amplitude variations.

At the minimum MLTS we group the couples $(\Theta^*_{e;s}, A^*_{e;s})$ which for fixed e satisfy the relation $\Theta^*_{e;s'} < \Theta^*_{e+1;s''}$ for all s' and s'' and we obtain groups containing disjoint MLTSs corresponding to two successive effective averagings. For each grouping we compute the center of mass $(\langle\Theta^*\rangle_{k'}, \langle A^*\rangle_{k'})$ where $k' = 1,2,\ldots,K_{\inf}$ satisfies the relation $k' = e+1$. For the global monotony spectrum in Fig. 4(a) we have $K_{\inf} = 3$.

At the maximum MLTS we group the couples $(\Theta^*_{e;s}, A^*_{e;s})$ with the same value $\Theta^*_{e,s}$ and satisfying the condition that their number is larger than half the number S of the time series in the statistical ensemble. In this way the amplitudes $A^*_{e;s}$ form a representative sample to compute the averages $\langle A^*\rangle_k$. For $e = E_s$ the condition $\Theta^*_{e;s} = N$ to stop the successive averagings holds for all s and then the number of effMMVs in this grouping is equal to S. Hence this group always contains more than $S/2$ effMMVs. For the next value $\Theta^*_{e;s} = N/2$, there are time series which do not contain two monotonic segments during the successive averagings and then the number of time series in this group is smaller than S. When $\Theta^*_{e;s}$ decreases, the number of the time series in the groups decreases and after several averagings it becomes smaller than $S/2$. For each grouping we compute the center of mass $(\langle\Theta^*\rangle_{k''}, \langle A^*\rangle_{k''})$ where $k'' = 1,2,\ldots,K_{\sup}$. For the global MS in Fig. 4(a) we have $K_{\sup} = 6$.

The values of the effMMVs which are not included in the groups at the extremities are grouped in equidistant bins with the length equal to the average length of the bins already existing. Finally the three sets of averages are joined and we obtain the average MS $(\langle\Theta^*\rangle_k, \langle A^*\rangle_k)$ with $k = 1,2,\ldots,K$. For the global MS in Fig. 4(a) we have $K = 19$ and its average MS is plotted with cross markers in Fig. 6(a). We see that even with the above precautions, between the large MLTSs and the intermediate region of the graph there is a small discontinuity.

APPENDIX B: HIERARCHY OF MS NOTATIONS

In the construction and analysis of the monotony structure of time series there is a hierarchy of procedures, each of them continuing the preceding one. We have introduced notations reflecting this dependence so that confusions between different indices are avoided.

The monotony structure of a single time series:

J: the number of the monotonic segments;

θ_j: *local time scale* (LTS) of the jth monotonic segment;

a_j: amplitude of the variation of the jth monotonic segment;

(θ_j, a_j): monotonic variation of the jth monotonic segment;

(θ_j, a_j), $j = 1,2,\ldots,J$: *monotony graph* of the time series.

The mean monotony structure of a single time series:

$\Theta = (N-1)/J$: *mean local time scale* (MLTS);

A: mean amplitude of the local variation;

(Θ, A): *mean monotonic variation* (MMV) or the center of mass of the monotony graph.

The evolution of the monotony structure of a time series during I successive averagings:

J_i: J of the ith averaging;

$\Theta_i = (N-1)/J_i$: MLTS of the ith averaging;

A_i: mean amplitude of the local variation of the ith averaging;

(Θ_i, A_i): MMV of the ith averaging.

The evolution of the monotony structure of a time series during E successive effective averagings:

i_e: the indices of the averagings which are effective;

$J^*_e = J_{i_e}$: J of the eth effective averaging;

$\Theta^*_e = \Theta_{i_e}$: MLTS of the eth effective averaging (effMLTS);

$A^*_e = A_{i_e}$: mean amplitude of the local variation of the eth effective averaging;

(Θ^*_e, A^*_e): MMV of the eth effective averaging (effMMV);

(Θ^*_e, A^*_e), $e = 1,2,\ldots,E$: *monotony spectrum* (MS) of the time series.

The monotony structure of a statistical ensemble with S time series:

$J^*_{e;s}$: J^*_e of the sth member of the statistical ensemble;

$\Theta^*_{e;s}$: effMLTS of the sth member of the statistical ensemble;

$A^*_{e;s}$: mean amplitude of the local variation of the eth effective averaging of the sth member of the statistical ensemble;

$(\Theta^*_{e;s}, A^*_{e;s})$, $e = 1,2,\ldots,E_s$, $s = 1,2,\ldots,S$: *global MS* of the statistical ensemble.

The average monotony structure of a statistical ensemble obtained by averaging the global MS over K bins:

$\langle\Theta^*\rangle_k$: average effMLTS of the kth bin;

$\langle A^*\rangle_k$: average mean amplitude of the local variation of the eth effective averaging of the kth bin;

$(\langle\Theta^*\rangle_k, \langle A^*\rangle_k)$, $k = 1,2,\ldots,K$: *average MS* of the statistical ensemble.

[1] L. F. Costa, O. Oliveira, G. Travieso, F. A. Rodrigues, P. R. V. Boas, L. Antiqueira, M. P. Viana, and L. E. C. da Rocha, Adv. Phys. **60**, 329 (2011).

[2] J. Gao, Y. Cao, W. Tung, and J. Hu, *Multiscale Analysis of Complex Time Series* (Wiley, Hoboken, NJ, 2007).

[3] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis* (Cambridge University Press, Cambridge, 2000).

[4] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, Proc. R. Soc. London, Ser. A **454**, 903 (1998).

[5] L. Lin, Y. Wang, and H. Zhou, Adv. Adapt. Data Anal. **1**, 543 (2009).

[6] I. Daubechies, J. Lu, and H.-T. Wu, Appl. Comput. Harmon. Anal. **30**, 243 (2011).

[7] T. Y. Hou and Z. Shi, Adv. Adapt. Data Anal. **3**, 1 (2011).

[8] C. Vamoş and M. Crăciun, Eur. Phys. J. B **87**, 301 (2014).

[9] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, Proc. Natl. Acad. Sci. USA **105**, 4972 (2008).

[10] J. Iacovacci and L. Lacasa, Phys. Rev. E **93**, 042309 (2016).

[11] C. Vamoş, M. Crăciun, and N. Suciu, Eur. Phys. J. B **88**, 250 (2015).

[12] P. Chaudhuri and J. S. Marron, Ann. Stat. **28**, 408 (2000).

[13] C. Vamoş and M. Crăciun, *Automatic Trend Estimation* (Springer, Dordrecht, 2012).

[14] G. Rilling and P. Flandrin, IEEE T. Signal Proces. **56**, 85 (2008).

[15] Y. Yang, J. Deng, W. Tang, C. Wu, and D. Kang, Chinese J. Electron. **18**, 759 (2009).

[16] Z. Wu and N. E. Huang, Proc. R. Soc. London, Ser. A **460**, 1597 (2004).

[17] P. J. Brockwell and R. A. Davies, *Introduction to Time Series and Forecasting* (Springer-Verlag, New York, 2003).

[18] C. Torrence and G. P. Compo, Bull. Am. Meteorol. Soc. **79**, 61 (1998).

[19] C. Franzke, Nonlinear Proc. Geoph. **16**, 65 (2009).

[20] C. Vamoş and M. Crăciun, Phys. Rev. E **78**, 036707 (2008).

[21] J. D. Hamilton, *Time Series Analysis* (Princeton University Press, Princeton, 1994).

[22] Z. Wu, N. E. Huang, S. E. Long, and C.-K. Peng, Proc. Natl. Acad. Sci. USA **104**, 14889 (2007).

[23] B. B. Mandelbrot and J. W. V. Ness, SIAM Rev. **10**, 422 (1968).

[24] Y. Meyer, F. Sellan, and M. S. Taqqu, J. Fourier Anal. Appl. **5**, 465 (2000).

[25] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, Phys. Rev. E **49**, 1685 (1994).

[26] K. Hu, P. C. Ivanov, Z. Chen, P. Carpena, and H. E. Stanley, Phys. Rev. E **64**, 011114 (2001).

[27] D. Maraun, H. W. Rust, and J. Timmer, Nonlinear Proc. Geoph. **11**, 495 (2004).

[28] C. Heneghan and G. McDarby, Phys. Rev. E **62**, 6103 (2000).

[29] N. Suciu, C. Vamoş, J. Vanderborght, H. Hardelauf, and H. Vereecken, Water Resour. Res. **42**, W04409 (2006).

[30] N. Suciu, Adv. Water Resour. **69**, 114 (2014).

[31] N. Suciu, L. Schüler, S. Attinger, and P. Knabner, Adv. Water Resour. **90**, 83 (2016).

[32] L. Calvet, A. Fisher, and B. Mandelbrot, technical report, Cowles Foundation Discussion Paper No. 1165, 1997.

[33] S. J. Taylor, *Asset Price Dynamics, Volatility, and Prediction* (Princeton University Press, Princeton, 2007).

[34] N. E. Huang, Z. Wu, S. R. Long, K. C. Arnold, X. Chen, and K. Blank, Adv. Adapti. Data Anal. **1**, 177 (2009).

[35] E. Bedrosian, Proc. IEEE **51**, 868 (1963).

[36] A. H. Nuttall and E. Bedrosian, Proc. IEEE **54**, 1458 (1966).

[37] J. B. Ramsey and C. Lampart, Stud. Nonlinear Dyn. E. **3**, 23 (1998).

[38] S. D. Meyers, B. G. Kelly, and J. J. O'Brien, Mon. Weather Rev. **121**, 2858 (1993).

[39] M. B. Priestley, J. Time Ser. Anal. **17**, 85 (1996).

[40] C. Goodall, in *Modern Methods of Data Analysis*, edited by J. Fox and J. S. Long (Sage Publications, Newbury Park, CA, 1990), pp. 126–176.