An Acad Bras Cienc (2021) 93(4): e20200859 DOI 10.1590/0001-3765202120200859
Anais da Academia Brasileira de Ciências | *Annals of the Brazilian Academy of Sciences*
Printed ISSN 0001-3765 I Online ISSN 1678-2690
www.scielo.br/aabc | www.fb.com/aabcjournal

PHYSICAL SCIENCES

Extended Methodology for DFA and DCCA: Application of Automatic Search Procedure and Correlation Map to the Weierstrass-Mandelbrot Functions

EULER B.S. MARINHO, AMIN BASSREI & ROBERTO F.S. ANDRADE

Abstract: Detrended fluctuation analysis and detrended cross-correlation analysis are used in this study to identify and characterize correlated data. The objective of these two techniques is to separate different fluctuations from the contributions due to external trends by evaluating the autocorrelation and cross-correlation exponents, in order to determine if scale properties persist with the size of the series. Two new methodologies were extended from cross-correlation coefficients for local analysis, which we call the *automatic search procedure* and *correlation map*. The simulations using the Weierstrass-Mandelbrot functions enabled us to identify several properties inherent to the magnitude, signal and original series. Also, the methodologies were applied to verify the existence of scale property in these data.

Key words: Detrended fluctuation analysis, detrended cross correlation analysis, local analysis, Weierstrass-Mandelbrot functions.

INTRODUCTION

In many situations, an observable variable is measured in successive intervals of time or space, which forms a data series. To determine if the series exhibits a complex behavior, such as self-affinity, we can express the autocorrelation by a power law. One possible analysis is the study of scale properties of the fluctuations in the series, which allows the verification of correlations present in a signal (autocorrelation) or across two distinct signals (cross-correlation).

Several mathematical tools that investigate and interpret the scale property fluctuations in space/time series have been developed. Hurst (1951), was the first to identify a power law in a time series. His study involved the problem related to water storage of the Nile River through rescaled (R/S) analysis. He found that many of these records could be described by a relation of the type $R/S \sim \nu^H$, where ν is the time scale used and H is the Hurst exponent, which determines the level of persistence of the physical parameter of interest.

Peng et al. (1994) described the DFA - detrended fluctuation analysis, to quantify correlation properties in non-stationary physiological time series. The application of this technique provided evidence of a crossover phenomenon associated with a change in short and long range scale exponents.

Podobnik & Stanley (2008) developed an extension of DFA, the DCCA - detrended cross correlation analysis, which is capable of investigating cross-correlations between different non-stationary temporal series that are recorded simultaneously.

Zebende (2011) proposed a new method that is capable of quantifying the level of cross-correlation between non-stationary time series, the CCC (cross-correlation coefficient), which relates the coefficients according to the scale used. This function was defined in terms of the DFA and DCCA methods and was initially applied to climate time series.

Two new methodologies are presented in this work for local analysis, the *automatic search procedure* and the *correlation map*, which indicate the dependence of the coefficients with the scales used and the number of records. Through the Weierstrass-Mandelbrot Functions we present several properties of the original, signal and magnitude series, and show how each can be applied in different situations. The original and signal series have linear properties and are related to fluctuations, which are useful for global analysis, whereas the magnitude series have nonlinear properties and is related to signal intensification, which is useful for local analysis.

We also used spectral analysis, which allowed us to observe that the equation $b = 2\alpha + 1$, Voss (1988), where b is the spectral exponent and α is the *DFA* exponent that is applied to the original series, can also be applied to the signal series.

DFA - DETRENDED FLUCTUATION ANALYSIS

Introduced by Peng et al. (1994), *DFA* is a method of scale analysis used to estimate exponents that characterize long-range correlations. Consider the input integrated series $y_k = \sum_{i=1}^{k} x_i$, where $x_i = y_{k+1} - y_k$ and $i, k \in \{1, 2, ..., N\}$. The integrated series y_k is divided into M_ν windows of sizes equal to ν for the calculation of the average or the linear trend in each window. Figure 1 illustrates this procedure. The fluctuations around the average are showed in Figure 1(a) and the fluctuations around the trend in Figure 1(b). The red lines represent, for each window m of size ν, the local averages in Figure 1(a) and the local trends in Figure 1(b).

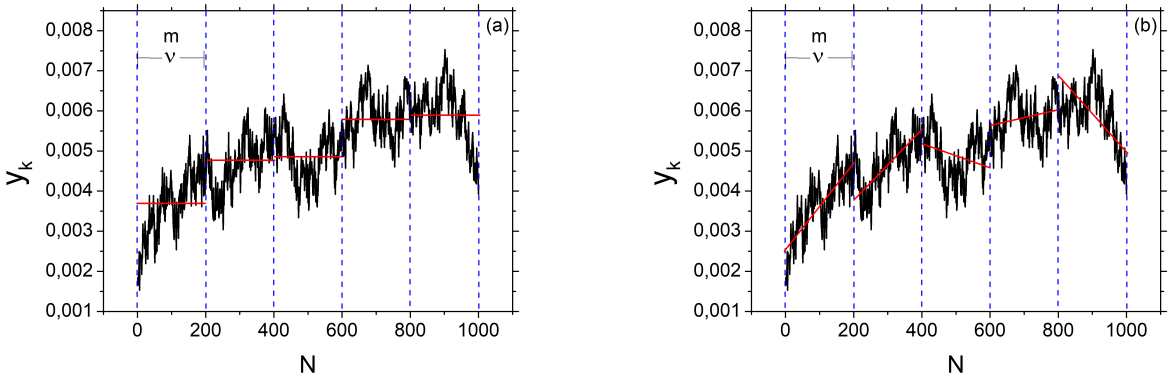

Figure 1. Example for autocorrelation analysis procedure: (a) fluctuation around the average and (b) fluctuations around the trend. y_k is the original input series (equivalent to the integrated series) and red lines represent, for each window m of size ν, the local averages in (a) and the local trends in (b).

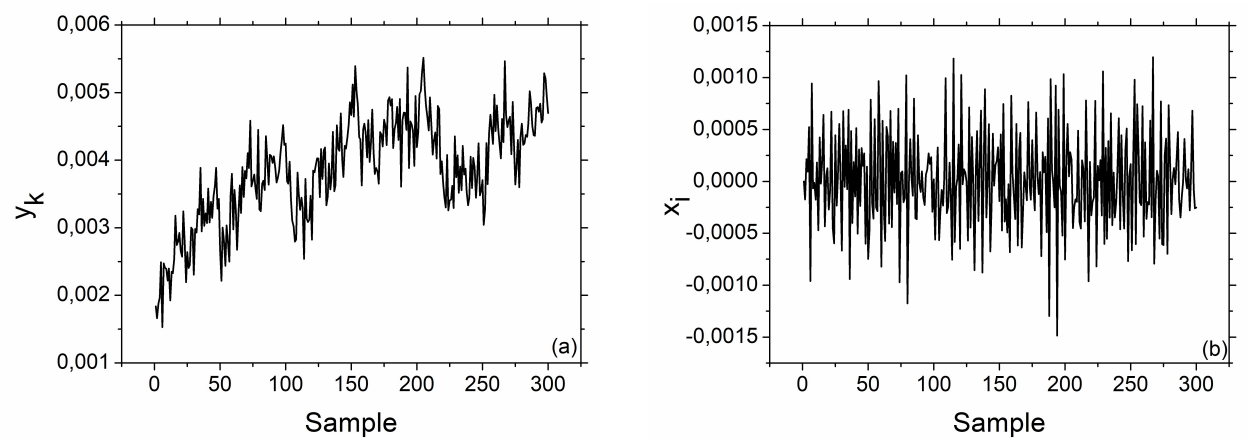

Figure 2. Example of transforming the series of successive differences into the integrated series: (a) the input series y_k; (b) the series of successive differences of y_k, that is, $x_i = y_{k+1} - y_k$. The integrated series $y_k = \sum_{i=1}^{k} x_i$, where $i, k \in \{1, 2, ..., N\}$, is obviously equal to (a).

From the series of successive differences or steps, one can obtain two additional series, one of the magnitudes and one of the signals, respectively, defined by $M = |x_i|$ and $S = sign\,(x_i)$ Ashkenazy et al. (2001). Then, the corresponding integrated series y_k^M and y_k^S are calculated using $y_k = \sum_{i=1}^{k} x_i$. Figure 2 shows an example of transforming the series of successive differences into an integrated series. A given input series y_k is shown in Figure 2(a), the series of successive differences of y_k, that is, $x_i = y_{k+1} - y_k$, is shown in Figure 2(b). The integrated series $y_k = \sum_{i=1}^{k} x_i$, where $i, k \in \{1, 2, ..., N\}$, will be obviously the same of Figure 2(a). Then, one can compute the series of successive difference magnitudes, $|x_i|$, shown in Figure 3(a) and the series of successive difference signals, $sign\,(x_i)$, where $sign\,(x_i) = 1$, if $x_i > 0$; $sign\,(x_i) = 0$, if $x_i = 0$ and $sign\,(x_i) = -1$, if $x_i < 0$, shown in Figure 3(b). Then, these last two series are integrated: Figure 3(c) shows the integrated series of magnitudes, $y_k^M = \sum_{i=1}^{k} |x_i|$ and Figure 3(d) shows the integrated series of signals, $y_k^S = \sum_{i=1}^{k} sign\,(x_i)$. The decomposition of the original series y_k into magnitude and signal series represents a useful strategy that can lead to a better characterization of the data under study. Each of these elements may exhibit different behaviors that are mixed in the fluctuations of the original series.

The integrated series y_k is then subtracted from the average or local trend for each window of size ν, with the purpose to obtain the autocorrelation f for each window described by the pair of parameters (m, ν), with $1 \leq m \leq M_\nu$. There are two types of autocorrelation measures:

$$f^2_{DFA_0}\,(m, \nu) = \frac{1}{\nu} \sum_{k=I_{min(m,\nu)}}^{I_{max(m,\nu)}} [y_k - \bar{y}_k\,(m, \nu)]^2, \tag{1}$$

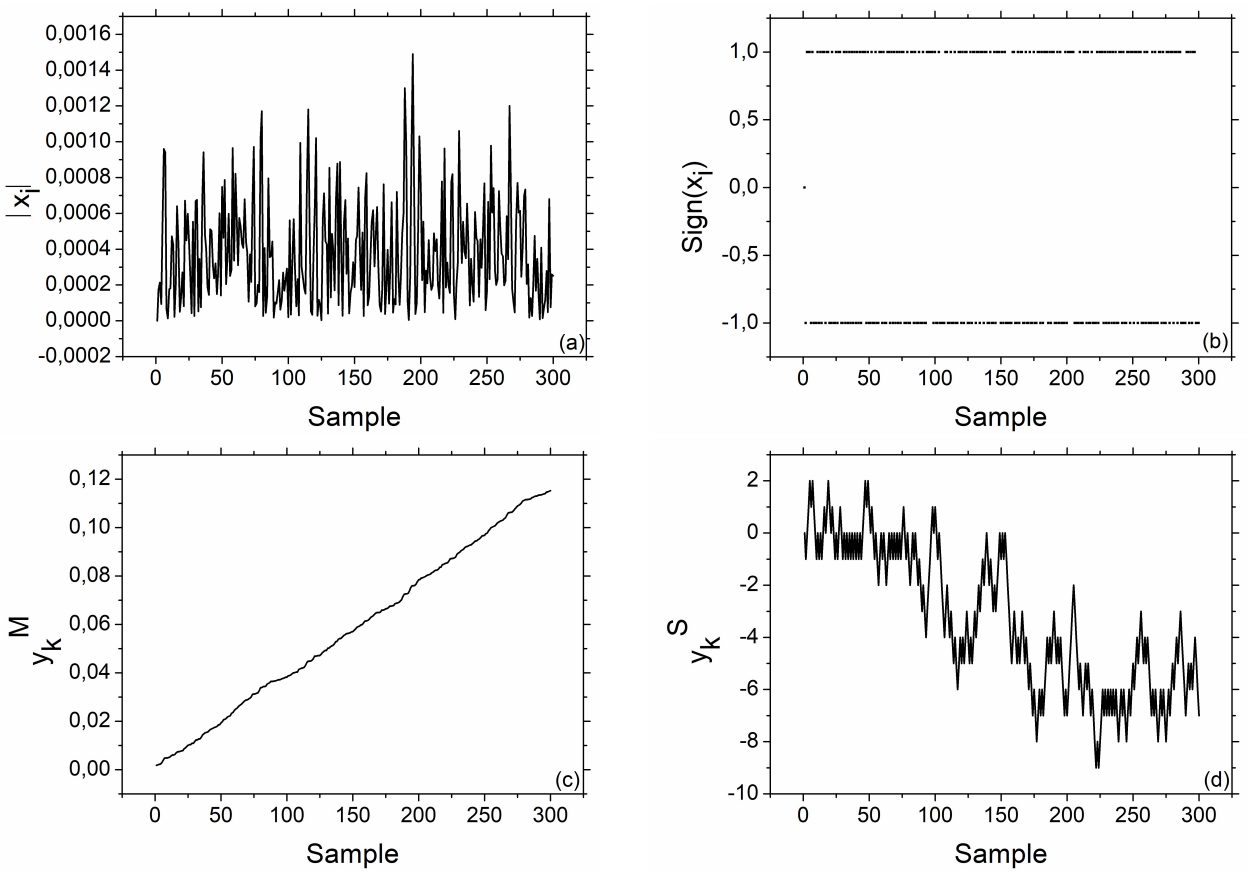

Figure 3. (a) the series of successive difference magnitudes, $|x_i|$**; (b) the series of successive difference signals,** $sign\,(x_i)$**, where** $sign\,(x_i) = 1$**, if** $x_i > 0$**;** $sign\,(x_i) = 0$**, if** $x_i = 0$ **and** $sign\,(x_i) = -1$**, if** $x_i < 0$**; (c) the integrated series of magnitudes,** $y_k^M = \sum_{i=1}^{k} |x_i|$**; (d) the integrated series of signals,** $y_k^S = \sum_{i=1}^{k} sign\,(x_i)$**.**

where the subscript o stands for the average, and

$$f^2_{DFA_1}(m,\nu) = \frac{1}{\nu} \sum_{k=I_{min(m,\nu)}}^{I_{max(m,\nu)}} [y_k - p_k(m,\nu)]^2, \tag{2}$$

for the trend, denoted by the subscript 1. In equation (1), the parameter $\bar{y}_k(m,\nu)$ is the average value of y_k in the window (m,ν), bounded by $I_{min(m,\nu)}$ and $I_{max(m,\nu)}$. In equation (2) $p_k(m,\nu) = a(m,\nu)\, z_k + b(m,\nu)$ is the first-order polynomial evaluated by the least squares method that expresses the linear local trend of the series in the window (m,ν).

Finally, we calculate the average of the fluctuations, expressed by F_X, for all $m - th$ windows of size ν:

$$F^2_X(\nu) = \frac{1}{M_\nu} \sum_{m=1}^{M(\nu)} f^2_X(m,\nu), \tag{3}$$

where the subscript X can be DFA_0 or DFA_1. From the integrated series the previous steps are repeated for other window sizes ν. The calculation for several window sizes is to obtain a relation between F_X

and ν. If the series has a scale property, we expect it to follow a power law given by $F_X \approx \nu^{\alpha}$. For a self-similar fractal process, F_X increases with ν. This function can be linearized through a log-log graph, where the value of α represents the angular coefficient of the best-fit straight line.

The increments are anti-correlated or anti-persistent if $0 < \alpha < 0.5$, which indicates that the values or phenomena tend to revert in the future. The increments are uncorrelated if $\alpha = 0.5$, that is, the series equals to white noise, which is a random sign that has a constant power spectral density. The increments are correlated or persistent if $0.5 < \alpha < 1.0$, which indicates that the values or phenomena tend to maintain their courses in the future.

The signs are often corrupted causing trends, and such trends should be well distinguished and eliminated from the fluctuations, allowing us to investigate the scaling behavior of these fluctuations. DFA is capable of determining the scale behavior of the data within possible trends without knowing its origin and form.

DCCA - DETRENDED CROSS CORRELATION ANALYSIS

DCCA was introduced by Podobnik & Stanley (2008), as a generalization of DFA, and it can be used in the analysis of two non-stationary series to verify cross-correlations and long-range memory. Consider two sets of data series, y_k and $y_k^{'}$, of the same size N, with increments x_i and $x_i^{'}$ respectively. The integration of these two series is given by,

$$y_k = \sum_{i=1}^{k} x_i \qquad and \qquad y_k^{'} = \sum_{i=1}^{k} x_i^{'}, \tag{4}$$

where $k = 1, 2, ..., N$.

Each integrated series y_k is now divided into M_ν windows of size ν. In each window, described by the pair of parameters (m, ν), with $1 \leq m \leq M_\nu$, we calculate the fluctuation measure with three types of cross-correlation estimators:

$$f^2_{DCCA_0}(m,\nu) = \frac{1}{\nu} \sum_{k=I_{min(m,\nu)}}^{I_{max(m,\nu)}} [y_k - \bar{y}_k(m,\nu)][y_k^{'} - \bar{y}_k^{'}(m,\nu)], \tag{5}$$

$$f^2_{DCCA_1}(m,\nu) = \frac{1}{\nu} \sum_{k=I_{min(m,\nu)}}^{I_{max(m,\nu)}} [y_k - p_k(m,\nu)][y_k^{'} - p_k^{'}(m,\nu)], \tag{6}$$

$$f^2_{|DCCA_1|}(m,\nu) = \frac{1}{\nu} \sum_{k=I_{min(m,\nu)}}^{I_{max(m,\nu)}} |[y_k - p_k(m,\nu)][y_k^{'} - p_k^{'}(m,\nu)]|. \tag{7}$$

In equations (5), (6) and (7), $\bar{y}_k$ and $\bar{y}_k^{'}$ are the average values of, respectively, y_k and $y_k^{'}$ in the box (m, ν) that is bounded by $I_{min(m,\nu)}$ and $I_{max(m,\nu)}$. Additionally, $p_k(m,\nu) = a(m,\nu) z_k + b(m,\nu)$ and $p_k^{'}(m,\nu) = a^{'}(m,\nu) z_k^{'} + b^{'}(m,\nu)$ are the first-degree polynomials. The subscript $|DCCA_1|$ in equation (7) uses the absolute value of the local fluctuations of each series.

In the sequence, the fluctuation function is calculated for each width ν by the equation (3), where the subscript X indicates the method to be used, i.e., $X = DFA_0, DFA_1, DCCA_0, DCCA_1 \quad or \quad |DCCA_1|$. If

the series presents a scale property related to the cross-correlations, we expect it to follow a power law $F_X \approx \nu^{\lambda}$, where the exponent λ represents the measure of the cross-correlation between the two analyzed series. In the case of autocorrelation analysis, the exponent $\lambda = \alpha$ also becomes equivalent to the Hurst (1951) exponent or roughness, which is generally indicated by H. Its main advantage is to allow the identification of the self-affinity and the long-range correlations in series with polynomial tendencies that can hide correlations. The value of λ indicates the type of correlation between the series, that is, as the exponent H, λ is related to the concept of persistence and antipersistence, as was seen for α.

The DCCA cross-correlation coefficient σ_{DCCA} was proposed to quantify the cross-correlation between two nonstationary series, Podobnik et al. (2011) & Zebende (2011). This coefficient is defined for the analysis of each scale ν (width of the windows) through the ratio

$$\sigma_{DCCA_1}(\nu) = \frac{F^2_{DCCA_1}(\nu)}{F_{DFA_1}(\nu)\, F'_{DFA_1}(\nu)}. \tag{8}$$

where σ_{DCCA_1} is a dimensionless quantity that varies in the range $-1 \leq \sigma_{DCCA_1} \leq 1$. Similar to the standard correlation coefficient, $\sigma_{DCCA_1} = 1$ indicates a perfect cross-correlation, while $\sigma_{DCCA_1} = -1$ indicates a perfect anticorrelation. If $\sigma_{DCCA_1} = 0$, there is no cross-correlation between the series. Since the DCCA coefficient is a function that depends on the scale ν, it is possible to observe how the cross-correlation between the series behaves for different scales and to determine, for example, if a high cross-correlation is valid for all scales or if there is a change of intensity in cross-correlation depending on the scale.

SPECTRAL ANALYSIS

The discrete Fourier transform of the sequence f_n of N terms is given by,

$$F_m = \frac{1}{N}\sum_{n=0}^{N} f_n e^{-i2\pi\frac{m}{N}n} \tag{9}$$

where $0 \leq m \leq N-1$ and $0 \leq n \leq N-1$. The power spectrum is expressed as,

$$P_m = [Re\,(F_m)^2 + Im\,(F_m)^2]^{\frac{1}{2}}. \tag{10}$$

The power spectrum of a completely random series is constant, that is, it does not exhibit any spectral structure. Therefore, when analyzing the spectrum of any series we are interested in detecting and interpreting the peaks that may exist and the frequencies to which they are associated. Many physical variables have a certain behavior that can be predicted. In this direction, we can obtain a certain proportionality of the power spectrum:

$$P(f) = \frac{1}{f^b}, \tag{11}$$

or $y = -bx$, where $y = log[P(f)]$ and $x = log(f)$, where b is the angular coefficient which is related to the spectral exponent.

In this study the power spectrum was generated as a function of frequency through a FFT (Fast Fourier Transform) algorithm. To reduce data dispersion, we smoothed the spectrum by calculating

half of the spectral value at a point by summing the residuals of the previous and posterior values, that is, $P^*(f) = \frac{1}{4}P(f-1) + \frac{1}{2}P(f) + \frac{1}{4}P(f+1)$, this smoothing is known as the Tukey-Hanning window. Voss (1988) related the exponents b and α through the relation $b = 2\alpha + 1$, where b is the spectral exponent and α the DFA exponent.

WEIERSTRASS-MANDELBROT FUNCTION

The Weierstrass-Mandelbrot function, used in this study, is defined as,

$$f(x) = \sum_{n=-\infty}^{\infty} b^{-n\alpha}[1 - \cos(b^n x)], \tag{12}$$

where $b > 1$ and $0 < \alpha < 1$. The function is continuous and has no derivative at any point. Ten curves (with α varying from 0.1 to 1.0) were constructed with 5000 samples each. Figure 4 shows a subset of these series, the first 300 samples. The original series are showed in Figures 4(a) and 4(b), where one can see that the fluctuations attenuate with the increase of the exponent α, that is, the greater the value of the exponent α, the smoother the resulting curve will be. For the sake of notation, we call the function with $\alpha = 0.1$ as W1, $\alpha = 0.2$ as W2, and so on.

To achieve a better understanding of the global analysis through the determination of exponents by DFA and DCCA, as well as to verify the reliability of the algorithm used, tests were performed to obtain the same α exponents presented in the curves in Figures 4(a) and 4(b). For the fluctuations analysis, a minimum window value $\nu = 5$ was chosen, and this value was gradually increased.

The behavior of the fluctuations in the signal series (Figures 4(c) and 4(d)) is very similar to the original series. Figures 4(e) and 4(f) show the magnitude series, where we can see that they exhibit few fluctuations and have a different behavior compared to the curves for the signal and original series. However, the magnitude series curves are similar to the original and signal series when the exponent α is greater than 0.8.

Figure 5 illustrates the DFA_0 and DFA_1 curves that are related to the autocorrelation for function W2 ($\alpha = 0.2$). The original series is displayed in Figure 5(a), the signal series in Figure 5(b) and magnitude series in Figure 5(c). The curves in Figures 5(a) and 5(b) tend to be parallel, which indicates that the exponents α for both techniques are similar. However, the curves of the magnitude series shown in Figure 5(c) have a different behavior; the DFA_0 curve is linear, while the DFA_1 curve shows ripples. The analysis was performed for all of the series presented in Figure 4, and similar results were obtained.

Figure 6 shows the fluctuation analysis of the cross-correlation between the functions W3 ($\alpha = 0.3$) and W7 ($\alpha = 0.7$), where a similarity is observed with the curves in Figure 5. Like in Figure 5, the original series is presented in Figure 6(a), the signal series in Figure 6(b) and the magnitude series in Figure 6(c). The cross-correlation analyses were conducted for all pairs, for which similar results were obtained. We chose the functions W3 and W7 to illustrate because they are not very close, like for example the functions W2 and W3.

We plotted the best-fit straight lines for the points on the curves shown in Figures 5 and 6 to obtain the autocorrelation and cross-correlation exponents presented in Table I. The exponents α_O that were obtained for the autocorrelation in W2 and W3 were practically identical to those shown in Figure 4, validating the algorithm used. For the function W7, the obtained α_O (0.644) through DFA_1

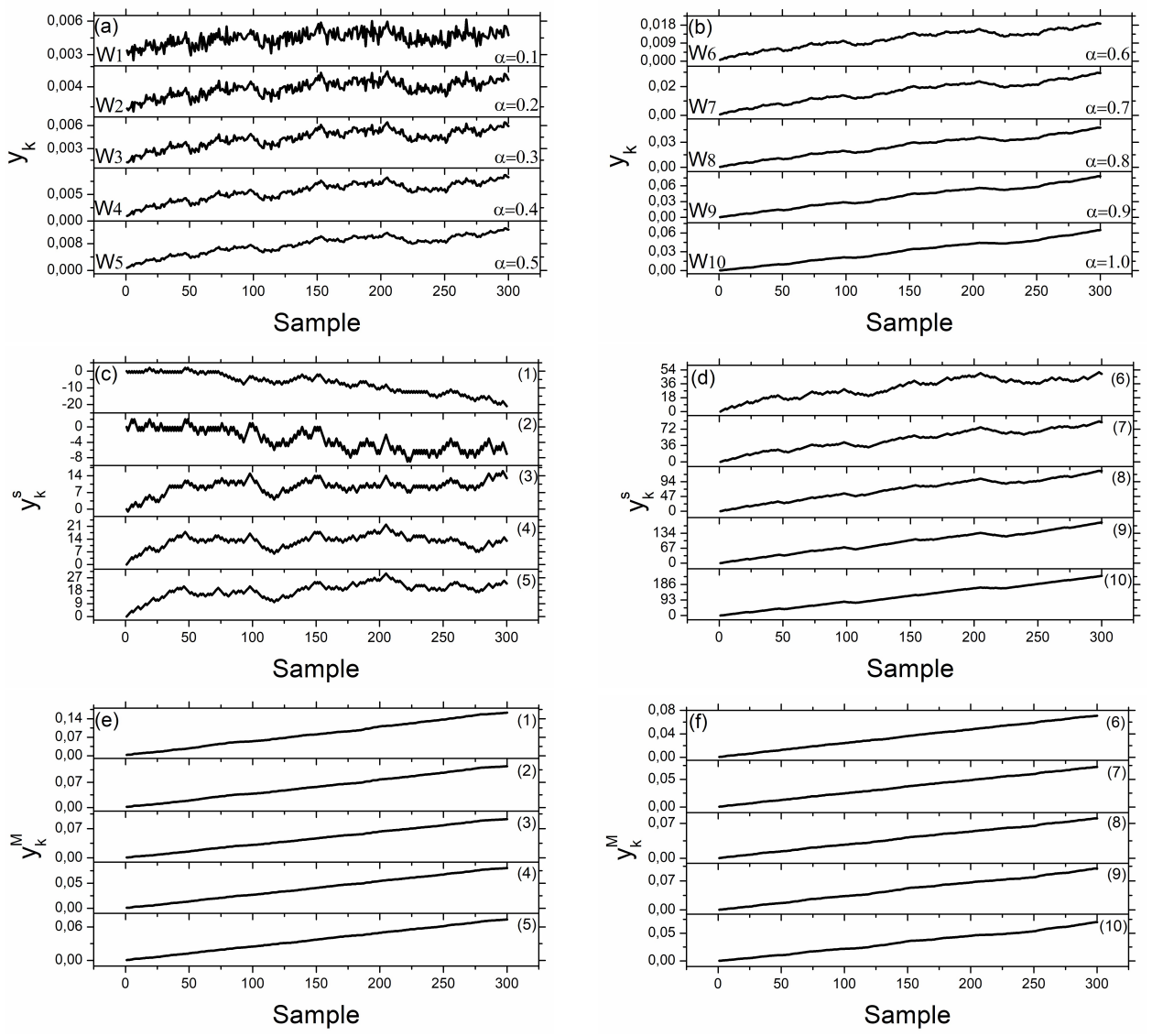

Figure 4. Records of 300 samples of the Weierstrass-Mandelbrot function: original series in (a) and (b) with various α exponent, signal series in (c) and (d) and magnitude series in (e) and (f). The numbers from (1) to (10) indicate the order in which each signal and magnitude series was obtained from the original series. For example, series (3) was obtained from the original series with $\alpha = 0.3$.

was a little far from the theoretical (0.7). The DFA_0 and DFA_1 curves provided similar exponents α_S for the functions W2 and W3, but not for the function W7. As expected, both the DFA_0 and DFA_1 estimates confirmed that W2 and W3 are anti-persistent and W7 is persistent. Also, the relation $\lambda \approx \left(\alpha + \alpha^{'}\right)/2$ of Podobnik & Stanley (2008) was confirmed, where α and $\alpha^{'}$ indicate the autocorrelation exponents of the analyzed series. To be specific, for the original series, using the DFA_0 estimate: $(0.304 + 0.725)/2 = 0.514 \approx 0.517 = \lambda_O$ and using the DFA_1 estimate: $(0.289 + 0.664)/2 = 0.477 \approx 0.478 = \lambda_O$. And, for the original series, using the DFA_0 estimate: $(0.435 + 0.720)/2 = 0.578 \approx 0.590 = \lambda_S$ and using the DFA_1 estimate: $(0.384 + 0.644)/2 = 0.514 \approx 0.504 = \lambda_S$.

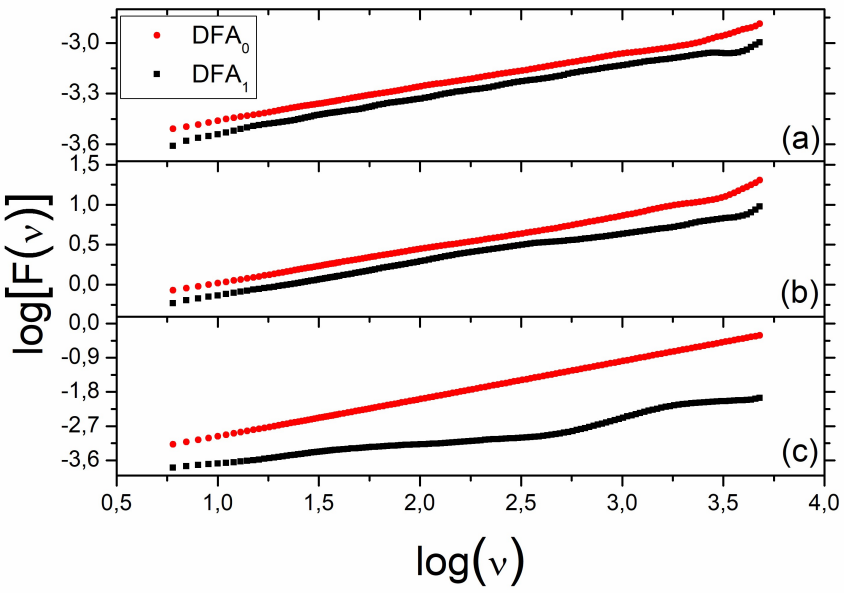

Figure 5. DFA_0 **and** DFA_1 **curves for the autocorrelation of the function W2. The original series is showed in (a), the signal series in (b) and the magnitude series in (c).**

Table I. Autocorrelation and cross-correlation exponents for W2, W3 and W7. α_O **and** α_S **are the autocorrelation exponents for the original and signal series, respectively, while** λ_O **and** λ_S **are the cross-correlation exponents for the original and signal series, respectively.**

Exponents	W2 α_O	W2 α_S	W3 α_O	W3 α_S	W7 α_O	W7 α_S	W3 & W7 λ_O	W3 & W7 λ_S
DFA_0 or $DCCA_0$	0.199	0.460	0.304	0.435	0.725	0.720	0.517	0.590
DFA_1 or $DCCA_1$	0.194	0.392	0.289	0.384	0.664	0.644	0.478	0.504

For the local analysis we used the coefficients DFA_1 and $DCCA_1$ with a sliding box. This procedure is illustrated for an arbitrary function in Figure 7(a). In this operation we move a box S, which is a series subset, of a certain size N' along with the series. We calculate σ_{DCCA_1} for this same box S of size N'. As can be seen in Figure 7(b), the box S is then divided into $M_\nu = 5$ windows, each one with size $\nu = 20$. From this, we calculate F to obtain, through equation (8), the autocorrelations and cross-correlation of y_k and y'_k in the interval N'.

After calculating this coefficient, the box S is moved one step forward, maintaining the same number M_ν of windows and calculating a new coefficient σ_{DCCA_1}. This analysis is repeated until the entire series of size N is surveyed with the same box S. Finally, the box of the same size N' is returned to the beginning of the series of size N to make the same analyzes using other window values ν. With this procedure we extended the methodology of analysis of the scale coefficients, where, by relating the window ν to time (or space), we can obtain the coefficients σ_{DCCA_1} as a function of the scale ν and the time t (or space z). This will result in the *correlation map*.

From the sliding box technique we also extended another methodology, which we call *automatic search procedure*, in which the algorithm searches the average $\bar{\sigma}$ of the coefficients σ, calculated for

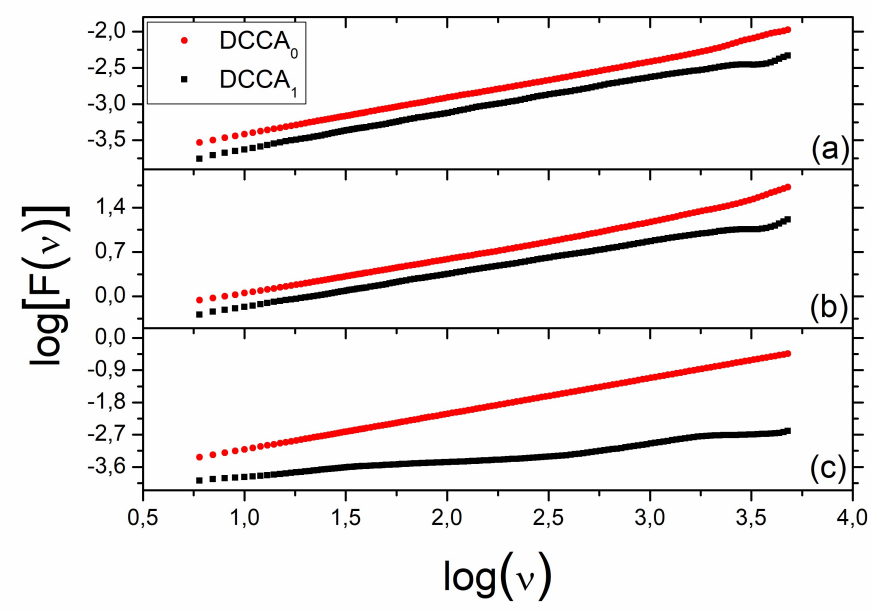

Figure 6. $DCCA_0$ **and** $DCCA_1$ **curves for cross-correlation between the functions W3 and W7. The original series is showed in (a), the signal series in (b) and the magnitude series in (c).**

all window sizes ν in a single box S of size N'. This is performed by limiting the interval of $\bar{\sigma}$. From the initial position of box S we obtain the first value of σ_ν, that is, σ_1. The second value of the coefficient, σ_2, is obtained by creating a new division of the box S for other values of windows ν. This procedure is repeated for various window sizes ν in the same box S in the same position. The same procedure as before is followed by moving the box S one step forward, until the entire series of size N is surveyed. The purpose here is to obtain the value of $\bar{\sigma}$. This methodology is unique, because it can be performed for all of the relative displacements between the series. The size of the box is an important parameter; it should be small enough to allow local analysis, and at the time, large enough to ensure the statistical significance of the fluctuation analysis.

In the automatic search procedure we evaluated the cross-correlation coefficients for the magnitude, signal and original series for the Weierstrass-Mandelbrot functions. For example, for the functions W2 and W3, with a sliding box of size $N' = 200$ points, Figures 8(a) and 8(b) show the procedure considering $\bar{\sigma} > 0.75$ and $\bar{\sigma} > 0.80$, respectively. According to Figure 8(a), it is possible to find values of $\bar{\sigma} > 0.75$ for the three series with a relative displacement $D_R = 0$. For different displacements between the functions W2 and W3 we noticed that the magnitude series identified a greater number of coefficients, compared to the signal and original series. However, Figure 8(b) shows that some points in the magnitude series have been lost, whereas there is no record for $D_R \neq 0$ in both the signal and original series. These results showed that the magnitude series is more favorable to determine correlations with high-valued coefficients. Other tests were performed with different types of series with similar results.

One can notice that Figure 8(a) has many points, for the magnitude series, because the function W2 and W3 are close. However, if the functions are not very close, the diagram would be much less populated. For instance, Figure 9 shows, only for the magnitude series, the automatic search procedure for the average coefficient $\bar{\sigma}$ between the functions W1 and W3. The condition $\bar{\sigma} > 0.75$ is shown in Figure 9(a) and the condition $\bar{\sigma} > 0.80$ in Figure 9(b).

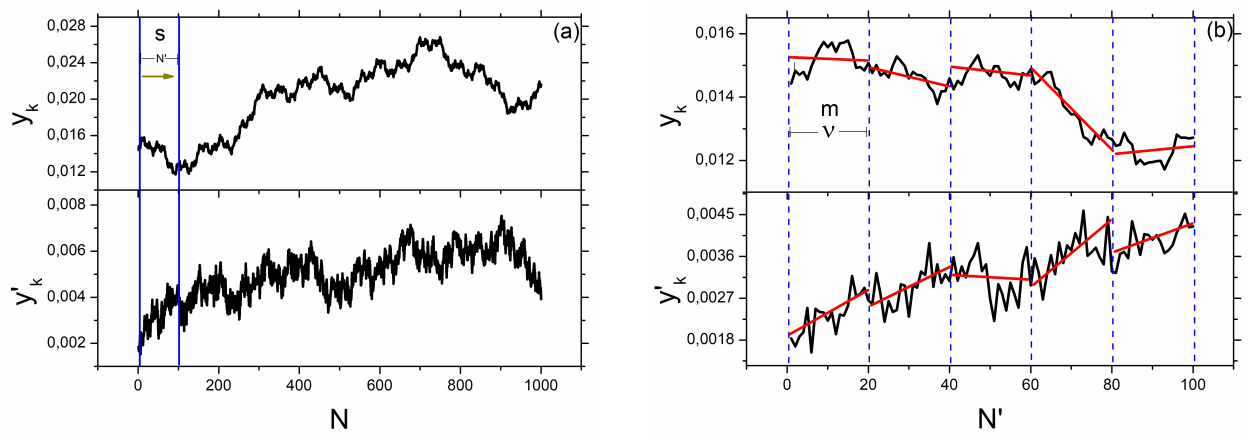

Figure 7. Procedure for the local analysis procedure: (a) Two series, y_k and $y_k^{'}$, with a moving box S with size $N^{'}$. (b) Detail of the same moving box S, now divided into $M_\nu = 5$ windows, that is, m varying from 1 to 5, with $\nu = 20$ in each sample.

To validate the automatic search algorithm we propose to find the same points highlighted in Figure 8(b) for the magnitudes series using the correlation map, in order to analyze the coefficients σ_ν that are related in the calculation of each $\bar{\sigma}$. We generate a map where the vertical axis shows the variation of window ν, which ranges from 10 to 60 points. The horizontal axis shows the number of samples and the colors indicate the values of the coefficient σ_ν, in the range from the minimum value (black) to the maximum value (red). In Figure 10(a) we show the graph of the automatic search procedure and the map of correlations is presented in Figure 10(b). It is important to mention that in Figure 10(b), the values of σ_ν over the entire N extension of the series are associated in Figure 10(a) to the condition $D_R = 0$, again over the entire N extension. For any value of N, it is observed in Figure 10(b) that σ_{10}, σ_{11}, ..., σ_{60} are all greater than 0.8. The high correlations indicate that the subseries contained in a box S have similar properties.

Similarly, Figure 11(a) shows three points or regions concentrated at $D_R = 190$ with $\bar{\sigma} > 0.80$. This information can be seen in more detail in Figure 11(b) as three regions, for different values of σ_ν under the condition $\bar{\sigma} > 0.80$. The correspondence between Figures 11(a) and 11(b) is indicated by dashed black arrows. Negative correlations indicate that fluctuations in this interval tend to revert for values of $\nu < 15$. On the other hand, the banding pattern indicates a change in correlation σ_ν for a fixed window ν, for most of the range of 4000 analyzed points.

The third region in Figure 8(b), with $D_R = 690$ under the condition $\bar{\sigma} > 0.80$ is detailed in Figure 12. In Figure 12(a) the single point is associated in Figure 12(b) with the region with the whole range of values of σ_ν. The few high correlations seen in Figure 12(b) indicate the bands given by the box S where the subseries have similar properties. To be specific, the single point at $N = 2300$ and $\bar{\sigma} > 0.80$ in Figure 12(a) is associated with the region in Figure 12(b) with N between 2300 and 2500 with $\bar{\sigma} > 0.79$, or if we extend the range, with the region with N between 2200 and 2700 points with $\bar{\sigma} > 0.68$.

With the automatic search procedure and the correlation map we obtain new information showing that the magnitude series identify local correlations that the signal and original series are not able to

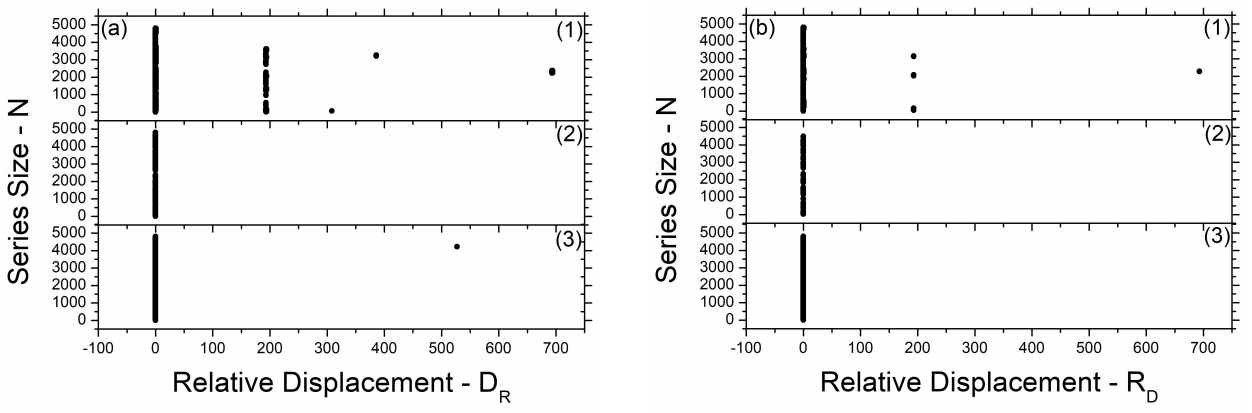

Figure 8. Automatic search procedure for the average coefficient $\bar{\sigma}$ between the functions W2 and W3, as a function of the series size N and the relative displacement D_R. (a) With the condition $\bar{\sigma} > 0.75$. (b) With the condition $\bar{\sigma} > 0.80$ in (b). The magnitude series is displayed in (1), the signal series in (2) and original series in (3).

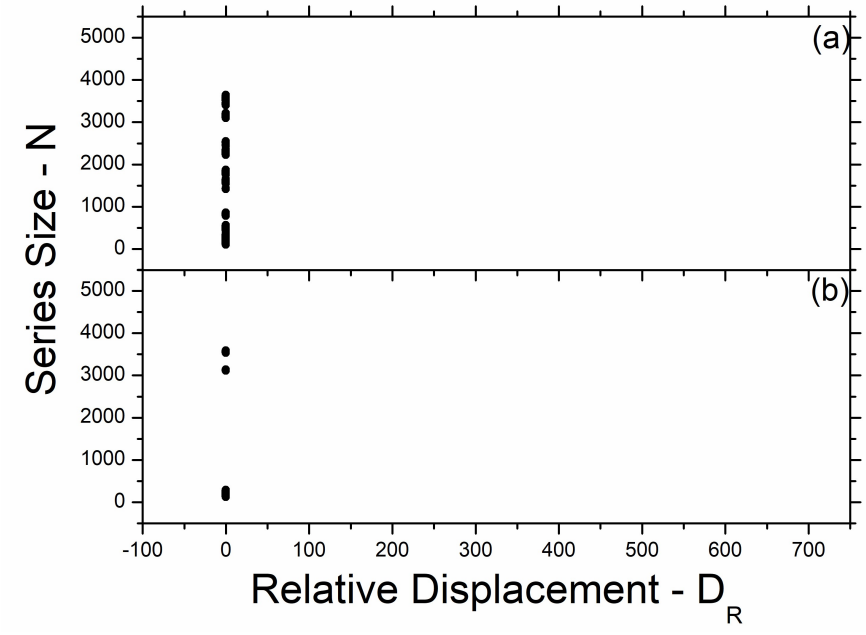

Figure 9. Automatic search procedure for the average coefficient $\bar{\sigma}$ between the functions W1 and W3 as a function of the series size N and the relative displacement D_R, only for the magnitude series. (a) With the condition $\bar{\sigma} > 0.75$. With condition $\bar{\sigma} > 0.80$. Notice a much less populated diagram in this figure when compared to Figure 8. This is due to the fact that now the two functions are less close.

identify. As previously mentioned, the magnitude series is related to the intensification of the original information, while the signal series is related to the direction of the information.

For the Weierstrass-Mandelbrot function W3, the non-smoothed power spectrum for the signal series is shown in Figure 13(a) and the smoothed one in Figure 13(b). For the original series, the non-smoothed power spectrum is displayed in Figure 14(a) and the smoothed one in Figure 14(b). We applied the Tukey-Hanning filtering in both spectra. There is a clear linear behavior even after smoothing, which is shown in the lower graphs. Therefore, it was possible to draw the best-fit straight line for the points to obtain the spectral exponent b. We could confirm the Voss relation, $b = 2\alpha+1$, by comparing the spectral exponents indicated in each graph in Figures 13 and 14 with the autocorrelation exponents listed in Table I. For the signal series, Table I provided $b = 2\,(0.435) + 1 = 1870$,

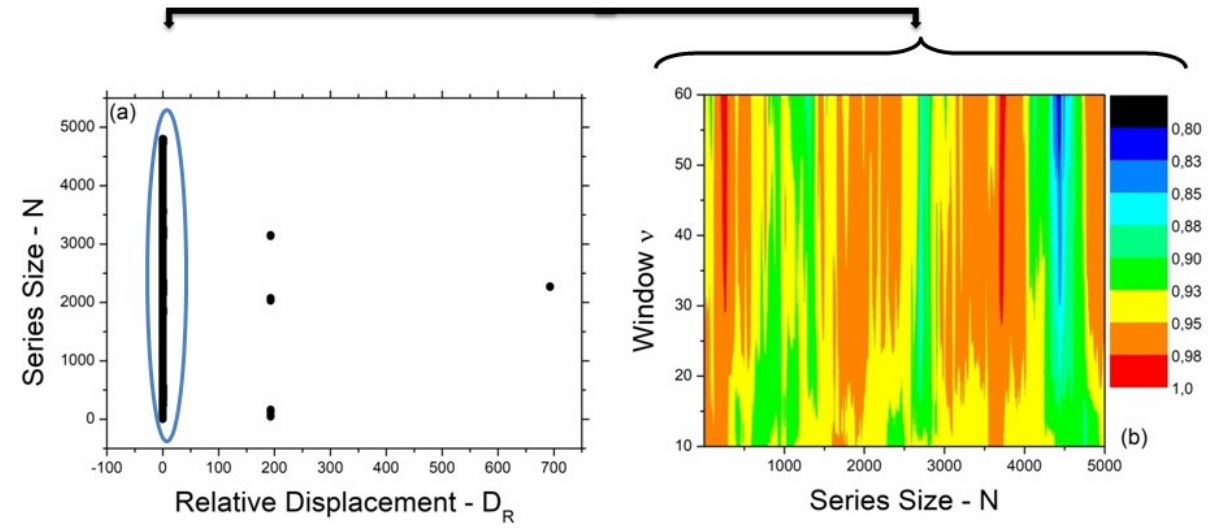

Figure 10. Analysis between the functions W2 and W3 for the relative displacement $D_R = 0$ point. In (a) the vertical axis represents the series size N and the horizontal axis represents the relative displacement D_R between the series for the condition $\bar{\sigma} > 0.80$. In (b) the vertical axis represents the window variation ν and the horizontal axis is the size of the series N . The colors represent the values of σ. The solid black arrow indicates that the map of correlations was generated from the automatic search procedure for $D_R = 0$.

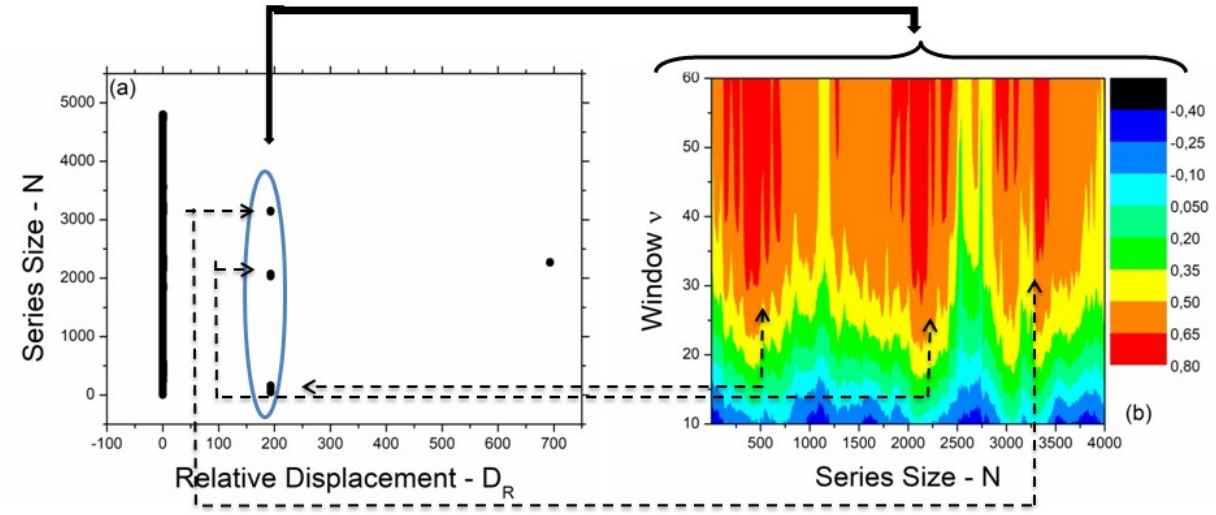

Figure 11. Analysis between the functions W2 and W3 for the relative displacement $D_R = 190$ points. In (a) the vertical axis represents the series size N and the horizontal axis represents the relative displacement D_R between the series for the condition $\bar{\sigma} > 0.80$. In (b) the vertical axis represents the window variation ν and the horizontal axis represents the size of the series N. The colors represent the values of σ. The solid black arrow indicates that the map of correlations was generated from the automatic search procedure for $D_R = 190$. Each black dashed arrow shows the relationship between the coefficients $\sigma_\nu > 0.80$ in (b) with the average $\bar{\sigma}$ of these coefficients in (a).

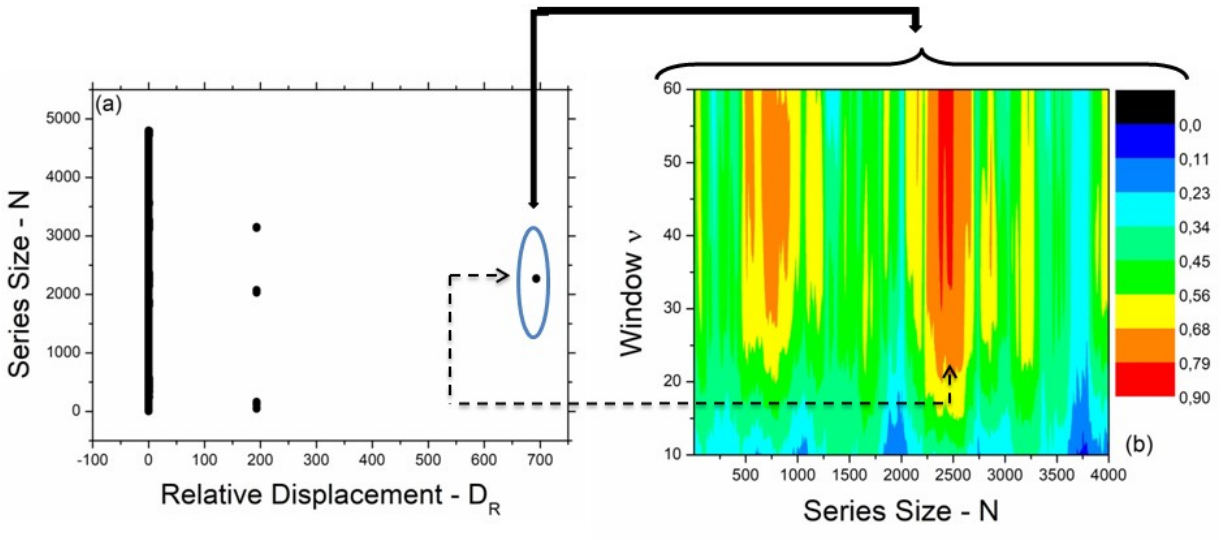

Figure 12. Analysis between the functions W2 and W3 for the relative displacement $D_R = 690$ points. In (a) the vertical axis represents the series size N and the horizontal axis represents the relative displacement D_R between the series for the condition $\bar{\sigma} > 0.80$. In (b) the vertical axis represents the window variation ν and the horizontal axis represents the size of the series, N . The colors represent the values of σ. The solid black arrow indicates that the map of correlations was generated from the automatic search procedure for $D_R = 690$. The black dashed arrow shows the relationship between the coefficients $\sigma_\nu > 0.80$ in (b) with the average $\bar{\sigma}$ of these coefficients in (a).

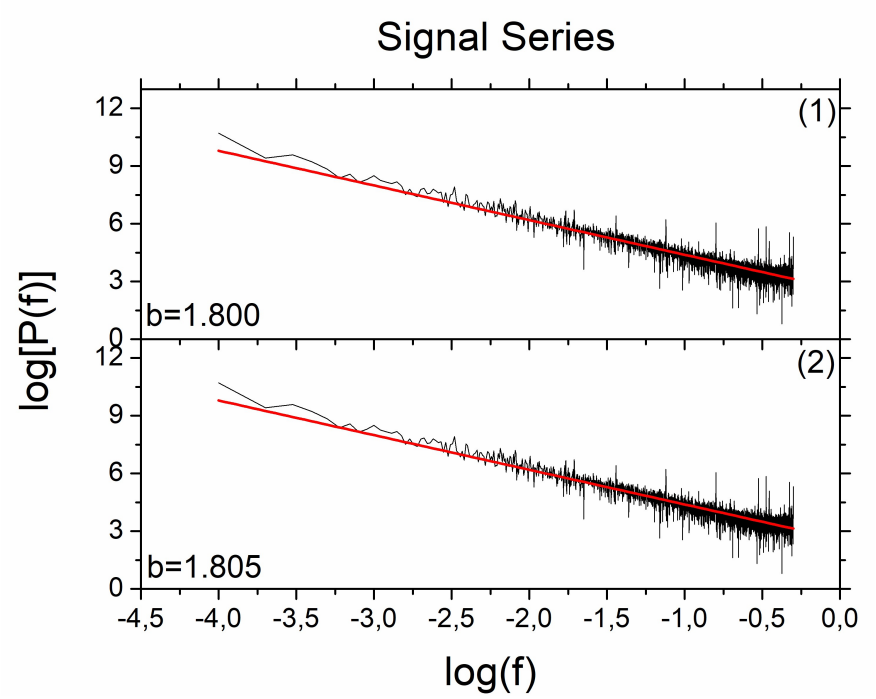

Figure 13. Power spectrum of the function W3 for the signal series, with the non-smoothed spectrum at the top and the smoothed one at the bottom. The values obtained for the spectral exponent b are indicated at lower left of each panel.

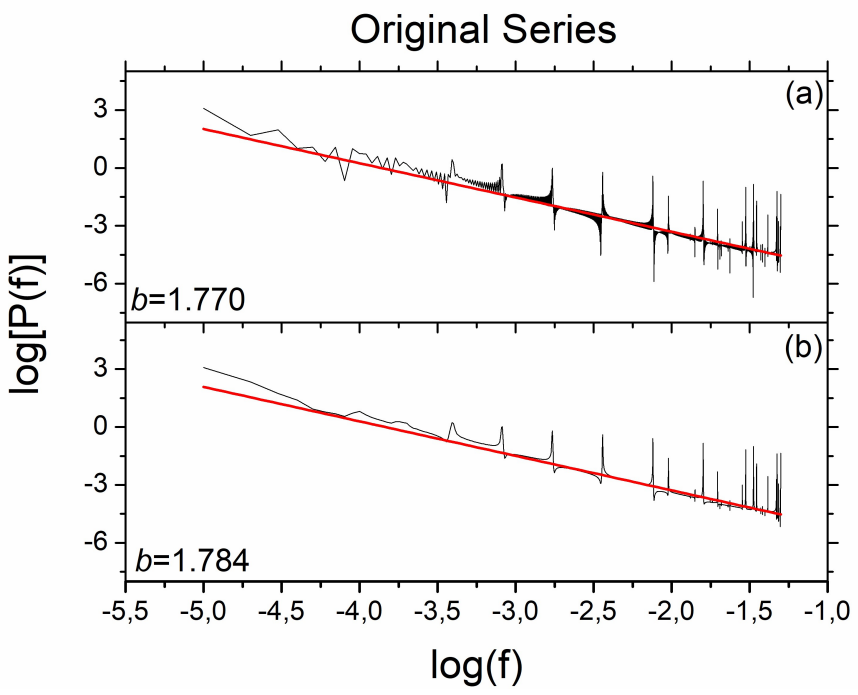

Figure 14. Power spectrum of the function W3 for the original series, with the non-smoothed spectrum at the top and the smoothed one at the bottom. The values obtained for the spectral exponent b are indicated at lower left of each panel.

whereas $b = 1800$ resulted from Figure 13(a). And, for the original series, we obtained, respectively, $b = 2\,(0.304) + 1 = 1608$ and $b = 1770$. Thus, the signal series yielded more accurate results than the original series.

CONCLUSIONS

We have extended two methodologies for the local analysis of the coefficients: the automatic search procedure and the correlation map, which are able to provide more detailed information about the dependence of the coefficients with the size of the series and the used window. The main characteristics inherent to the original, magnitude and signal series were presented, and the results indicated that they have their own scaling properties. We note that very small and very large windows ν are not suitable for fluctuation analysis. The autocorrelation exponents yielded $DFA_0 \approx DFA_1$, for both the signal and original series. On the other hand, the exponents for the magnitude series cannot be considered for the global analysis, but they are essential for the local analysis. Through simulations using the Weierstrass-Mandelbrot functions, the relation $\lambda \approx \left(\alpha + \alpha^{'}\right)/2$ was confirmed, where α and $\alpha^{'}$ indicate the autocorrelation exponents of the analyzed series and λ is the cross-correlation exponent. Also, the spectral analysis showed that the relation $b = 2\alpha + 1$ is valid for the signal and original series.

Acknowledgments

The authors thank the financial support of the following Brazilian agencies: *Financiadora de Inovação e Pesquisa* (FINEP), grant Rede 01 - Phase 5; *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES) - Finance Code 001; *Fundação de Amparo à Pesquisa do Estado da Bahia* (FAPESB), grant PIE00005/2016 (Infrastracture Call FAPESB 003/2015); *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq), grant INCT-GP.

REFERENCES

ASHKENAZY Y, IVANOV PC, HAVLIN S, PENG CK, GOLDBERGER AL & STANLEY HE. 2001. Magnitude and sign correlations in heartbeat fluctuations. Phys Rev Lett 86(9): 1900-1903. URL http://dx.doi.org/10.1.1/jpb001.

HURST HE. 1951. Long-term storage capacity of reservoirs. Trans ASCE 116(1): 770-808.

PENG CK, BULDYREV SV, HAVLIN S, SIMONS M, STANLEY HE & GOLDBERGER AL. 1994. Mosaic organization of DNA nucleotides. Phys Rev E 49(2): 1685-1689.

PODOBNIK B & STANLEY HE. 2008. Detrended cross-correlation analysis: a new method for analyzing two nonstationary time series. Phys Rev Lett 100(8): 084102.

PODOBNIK B, JIANG ZQ, ZHOU WX & STANLEY HE. 2011. Statistical tests for power-law cross-correlated processes. Phys Rev E 84(6): 066118.

VOSS RF. 1988. Fractals in nature: from characterization to simulation. In: The science of fractal images, p. 21-70. Springer.

ZEBENDE GF. 2011. DCCA cross-correlation coefficient: Quantifying level of cross-correlation. Physica A: Statistical Mechanics and its Applications 390(4): 614-618.

How to cite

MARINHO EBS, BASSREI A & ANDRADE RFS. 2021. Extended Methodology for DFA and DCCA: Application of Automatic Search Procedure and Correlation Map to the Weierstrass-Mandelbrot Functions. An Acad Bras Cienc 93: e20200859. DOI 10.1590/0001-3765202120200859.

Manuscript received on June 3, 2020;
accepted for publication on April 25, 2021

EULER B.S. MARINHO[1]
https://orcid.org/0000-0002-0945-8742

AMIN BASSREI[1,2]
https://orcid.org/0000-0002-4653-2016

ROBERTO F.S. ANDRADE[3]
https://orcid.org/0000-0002-9323-1400

[1]Universidade Federal da Bahia, CPGG/IGEO, Rua Barão de Jeremoabo, s/n, Ondina, 40170-115 Salvador, BA, Brazil
[2]Instituto Nacional de Ciência e Tecnologia de Geofísica do Petróleo, Rua Barão de Jeremoabo, s/n, Ondina, 40170-115 Salvador, BA, Brazil
[3]Universidade Federal da Bahia, Instituto de Física, Rua Barão de Jeremoabo, s/n, Ondina, 40210-340 Salvador, BA, Brazil

Correspondence to: **Amin Bassrei**
E-mail: bassrei@ufba.br

Author contributions

Euler Marinho developed this study, processed the data, generated the figures, interpreted the results and wrote the first version of the manuscript. Amin Bassrei supervised the research, contributed to the results discussion, reviewed and translated the manuscript. Roberto Andrade co-supervised the research and contributed to the results discussion.