Alternative Ways to Compare the Detendred Fluctuation Analysis and its Variants. Application to Visual Tunneling Detection

Bastien Berthelot, Eric Grivel, Pierrick Legrand, Jean-Marc André, Patrick Mazoyer

Alternative Ways to Compare the Detendred Fluctuation Analysis and its Variants. Application to Visual Tunneling Detection

Bastien Berthelot[b,c], Eric Grivel[a], Pierrick Legrand[b], Jean-Marc André[a], Patrick Mazoyer[c]

[a] *Bordeaux University - INP Bordeaux - IMS - UMR CNRS 5218, FRANCE*
[b] *Bordeaux University - IMB UMR CNRS 5251 - CQFD team, INRIA, FRANCE*
[c] *THALES AVS France, Campus THALES Merignac*

Abstract

The detrended fluctuation analysis (DFA) and its variants such as the detrended moving average (DMA) are widely used to estimate the Hurst exponent. These methods are very popular as they do not require advanced skills in the field of signal processing and statistics while providing accurate results. As a consequence, a great deal of interest has been paid to compare them and to better understand their behaviors from a mathematical point of view. In this paper, our contribution is threefold. Firstly, we propose another variant avoiding the discontinuities between consecutive local trends of the DFA by *a priori* constraining them to be continuous. Secondly, we show that, in all these approaches, the square of the fluctuation function can be presented in a similar matrix form. When the process is wide-sense stationary (w.s.s.), the latter can be seen as the power of the output of a linear filtering whose frequency response depends on the given method. In the general case, an interpretation of the square of the fluctuation function is also given by expressing it as the convolution between the 2D-Fourier transform of two matrices, one whose elements correspond to the instantaneous correlation function of the signal and the other which depends on the detrending method. To end up, an illustration is provided in the field of avionics for the detection of the visual tunneling, a deleterious cognitive state.

Keywords: DFA, DMA, Hurst coefficient, frequency analysis, tunneling.

1. Introduction

One of the main goals of signal processing algorithms is the characterization and classification of random processes, usually collected by using sensors. To this end, it is common to estimate some features such as the signal power in certain frequency bands, the zero-crossing rate, the entropy or the multiscale entropy [1], or the parameters of an *a priori* model that fits the data. In the latter case, an autoregressive moving average (ARMA) model can be considered to describe a process whose power spectral density exhibits both frequency resonances and rejections. In addition, as the values of the normalized autocovariance function of this model tend to decay to zero exponentially when the lag increases, this model is well-suited for short-memory data. Hence, this can be a relevant model in speech processing for instance. However, in other cases such as in the field of economics [2], the normalized autocovariance function decays slower than exponentially so that the sum of its terms goes to infinity. This corresponds to processes such as autoregressive fractionally integrated moving average (ARFIMA) processes [3, 4], which have long-range dependance (LRD). LRD is therefore another feature of interest when studying time series, and it can be characterized in many ways. One of them consists in analyzing the properties of the autocorrelation function or of the power spectral density (PSD), but it can also be done by estimating the Hurst exponent, denoted H [5]. Thus, a process is said to have LRD if $0.5 < H < 1$. When $0 < H < 0.5$, the process is short range dependant or anti-persistent. There are also some particular cases: H is respectively equal to 0.5, 0 and -0.5 for a Brownian noise, a pink noise and a white noise. The reader can refer to [6] for more information.

The different approaches that aim at estimating the Hurst coefficient can be sorted in two main families: the time-domain estimators and the frequency-domain estimators. Among the time-domain estimators, the rescaled range (R/S) analysis [7] was first proposed by Hurst in 1951. Then, other methods, such as the aggregated variance method, the absolute-value method and the variance-of-residuals method were introduced. See [8] and [9] for more details.

To estimate the Hurst exponent of a pure mono-fractal time series, the fluctuation analysis (FA) [10] was proposed in the early 90's as another time domain approach. After integrating the signal leading to the new sequence y_{int}, the fluctuation function, denoted as $F_{\mathbf{0}}(N_{\mathbf{0}})$ and equal to $\sqrt{< (y_{int}(i+N_{\mathbf{0}}) - y_{int}(i))^2 >}$, where $< . >$ denotes the temporal mean, is computed for different values of the lag $N_{\mathbf{0}}$. As $F_{\mathbf{0}}(N_{\mathbf{0}})$ is proportional to $N_{\mathbf{0}}^{H+1}$ [10], $\log(F_{\mathbf{0}}(N_{\mathbf{0}}))$ is represented as an affine function of $\log(N_{\mathbf{0}})$ to estimate H in the least-square (LS) sense[1]. Since the FA is sensitive to non-stationarities, the detrended fluctuation analysis (DFA) [11] was developed. This method operates with the following steps: after integration, the sequence y_{int} is split into small parts of size $N_{\mathbf{1}}$. Using a LS criterion, their trends are deduced. Most of the time, block estimation schemes are considered, but on-line methods based on Kalman filtering [12] could be used. The resulting piece-wise linear trend is then subtracted to the whole integrated signal y_{int}. The power of the subtraction residual, denoted as $F_{\mathbf{1}}^2$ and corresponding to the square of the fluctuation function, is computed for different values of $N_{\mathbf{1}}$. Due to the integration, the regularity of the integrated signal is equal to $H+1$. As $F_{\mathbf{1}}$ is proportional to $N_{\mathbf{1}}^{H+1}$, the log-log representation corresponds to a straight line with a slope denoted as $\alpha = H+1$ and called the scaling exponent.

Several variants of the DFA now exist based on different ways to obtain the global trend of the integrated process [13].

1. Instead of using a linear trend for each segment, polynomials with a degree larger than 1 can be considered. This leads to the quadratic DFA and the cubic DFA. These methods are called higher-order DFA.
2. When dealing with a signal whose regularity evolves over time, the scaling exponent can be estimated on a sliding window. For each windowed signal, either the standard DFA described above is used or the square of the fluctuation function is computed and α is updated by using a Kalman

[1] In this paper, log denotes the logarithm to the base 10.

filter. See [14].

3. As discontinuities between local trends can be debatable, various authors aimed at addressing this issue. Thus, the authors in [15] suggest searching the global trends of the integrated signal and of its shifted version. Then, both are combined to obtain a continuous global trend. This method is called AFA for adaptive fractal analysis.
4. The detrended moving average (DMA) is based on a low-pass filtering of the whole integrated signal in order to obtain the trend. In its standard version, the filter has a finite impulse response (FIR) of length $N_{\mathbf{2}}$ [16]. In economics, this way to deduce the trend is known as "moving average filtering" [17]. Using either a finite-impulse-response filter or an infinite-impulse-response filter leads to variants known as the simple moving average, including the backward moving average and the centered moving average (CDMA), the weighted moving average of order l, labeled as WDMA-l and the weighted centered detrended moving average (WCDMA) [18].

The reader can refer to [19] or [20] for a comparison between some of these approaches.

In addition to these time-domain estimators, many of the more recent methods aiming at estimating the Hurst exponent are based on the frequency analysis. They consist in studying the power spectral density (PSD) of the time series [21] and include the local Whittle method, the periodogram-based method, the empirical-mode-decomposition-based approach [22], the fractional Fourier transform [21], the wavelet-based method [23] and the semi-parametric method [24, 25]. Different comparative studies have also been led [26]. These methods have been compared with the DFA in specific usecases, both on monofractal and multifractal processes [27–30]. Even though they can outperform the DFA on a given synthetic data-set, the latter and its variants have been proven to be among the best estimators. Indeed, in [8], the authors compared most of the above-mentioned estimators, except the one based on wavelets, on different types of signals. The DFA appeared to be the most robust non-parametric es-

timator as it provided the smallest mean squared error in the estimation of H. Moreover, the DFA and its variants do not *a priori* require advanced skills in statistical signal processing as they are based on regressions and linear filtering. This is one of the main reasons of their popularity. Their applications to real processes have been extensively studied especially in the field of biomedical [31–38], econometrics [39, 40], meteorology [41, 42], or geophysics [43]. They can be considered as a trade-off between performance, computational cost and simplicity of implementation and use.

During the last years, the main contributions dealing with these approaches have been done on four aspects: providing extensions or generalizations of the algorithms [44], addressing the cases of multifractal time series [45], developing fast versions [46, 47], proposing mathematical analysis to better understand their behaviors [20, 48–51]. In [52], a relation between the fluctuation function and an estimation of the normalized autocorrelation function of the signal was given, by assuming that the process was wide sense stationary (w.s.s.) and ergodic and by making some approximations. The single-frequency responses of the DFA and the higher-order DFA [48] as well as the centered DMA [51] were analyzed. The authors concluded that for stochastic processes whose PSD is a function of the frequency of the form $f^{-\beta}$, using the higher-order DFA is convenient to estimate α as long as $\alpha = \frac{\beta+1}{2}$. It should be noted that the link between $F_{FA}(N_{FA})$ and the normalized covariance function was studied in [53].

In this paper, the contribution is threefold:

Firstly, we suggest studying another variant of the DFA. More particularly, we propose to model the global trend of the integrated signal by assuming that the consecutive local trends are continuous. The estimations of the parameters of the local trends are based on a constrained LS criterion. This method is called CDFA in the remainder of this paper.

Secondly, a theoretical comparison between the ways to deduce the square of the fluctuation function with the DFA and its variants is proposed. Two cases are addressed:

For one thing, the process is assumed to be w.s.s. In this case, the statistical

mean of the square of the fluctuation function is expressed from the autocorrelation function of the process and consequently from the PSD. Therefore, the matrix formulation we propose makes it possible to understand how the fluctuation function is modified when using one of the variants of the DFA. The differences between the various methods are hence highlighted. Although $N_{\mathbf{1}}$ corresponds to the local-trend length for the DFA and the CDFA and $N_{\mathbf{2}}$ corresponds to the filter order for the DMA, our purpose is to study the influence of these parameters. In the following, $N_{\mathbf{1}} = N_{\mathbf{2}} = N$. In addition, we compare our work with the approaches proposed by [48] and [52], in which approximations were done to express the fluctuation function.

Then, the case of non-stationary processes is addressed. We will see that our analysis generalizes the interpretation we provide for w.s.s. processes. More particularly, we show that the square of the fluctuation function can be deduced by expressing it as the convolution between the 2D-Fourier transform of two matrices : one whose elements corresponds to the instantaneous correlation function of the signal and another called weighting matrix, which depends on the detrending method. Comparing the methods amounts to comparing the 2D-Fourier transform of the weighting matrices.

To end up, examples are given to illustrate our analysis. These deal with ocular processes obtained on subjects under the state of visual tunneling. The Hurst exponent of the gaze position, estimated with both the DFA and CDFA, is shown to be a promising marker of this deleterious cognitive state.

The remainder of the paper is organized as follows: In section 2, the main steps of the approaches are recalled before expressing the square of the fluctuation function in a uniform matrix way. Section 3 provides a comparative analysis both in the stationnary and non-stationnary case. In section 4, simulations are presented.

2. Presentation and comparison of methods based on trend extraction

After giving the general steps of the DFA and its variants and providing some notations, this section deals with a uniform way to express the trend vector with the different variants of the DFA.

2.1. General steps of these approaches

Let us consider M consecutive samples $\{y(m)\}_{m=1,\ldots,M}$ of the signal. The DFA and its variants are defined by the following four steps [11, 16, 18]:

- **Step 1**. The so-called profile

$$y_{int}(m) = \sum_{i=1}^{m} (y(i) - \mu_y)$$

  is first computed, where $\mu_y = \frac{1}{M}\sum_{m=1}^{M} y(m)$ is the mean of y.

- **Step 2**. The trend of the profile is estimated. This step will be detailed for each method in the next subsections below.

- **Step 3**. The resulting trend is subtracted to the profile. This leads to a residual. Then the square root of the residual power, $F_{\boldsymbol{i}}(N)$, is computed, where the subscript $_{\boldsymbol{i}}$ refers to the method that is used *i.e.* DFA ($\boldsymbol{i} = \mathbf{1}$), DMA ($\boldsymbol{i} = \mathbf{2}$) or CDFA ($\boldsymbol{i} = \mathbf{3}$).

- **Step 4**. Steps 2 and 3 are repeated for different values of N. At this stage, as $F_{\boldsymbol{i}}(N) \propto N^{\alpha}$ [10], $\log(F_{\boldsymbol{i}}(N))$ is plotted as a linear function of $\log(N)$.

- **Step 5.** The final step is to search a straight line fitting the log-log representation. Its slope, denoted as α, is estimated in the LS sense.

The approaches differ in the way of deducing the trend, *i.e.* in step 2. In the following, let us present each of them and express the trend vector in a matrix form.

2.2. Notations

Some notations that will be used in the rest of the document are listed below:

- $A(i:j, k:l)$ is the part of the matrix A corresponding to the elements belonging to the rows i to j and to the columns k to l.
- $\mathbb{1}_{j\times k}$ and $\mathbf{0}_{j\times k}$ are matrices of size $j \times k$ filled with 1s and 0s respectively.
- $diag([.], l)$ is a matrix whose l^{th} diagonal is equal to $[.]$.
  Thus, $I_j = diag(\mathbb{1}_{1\times j}, 0)$ is the identity matrix of size j. $diag(\mathbb{1}_{1\times N-1}, 1)$ is the square matrix of size N whose 1^{st} sub-diagonal above the main one has its elements equal to 1.
- $J_j = I_j - \frac{1}{j}\mathbb{1}_{j\times j}$.
- T_l is a $N \times 1$ vector storing the values of the l^{th} local trend $t_l(n)$.
- Y and Y_{int} are two column vectors storing respectively the samples $\{y(n)\}_{n=1,...,M}$ and $\{y_{int}(n)\}_{n=1,...,M}$. This leads to :

$$Y_{int} = [y_{int}(1), y_{int}(2), ..., y_{int}(M)]^T = \mathrm{H}_M J_M Y \tag{1}$$

  with $H_M = \sum_{r=0}^{M-1} diag(\mathbb{1}_{1\times M-r}, -r)$ a lower triangular matrix filled with 1s.

- Depending on the approach used to estimate the trend of the profile, all the samples of the profile are not necessarily considered. In addition, some other transformations will be required to get the matrix form of the trend. Therefore, let us introduce the following matrix of size (j, M):

$$C_{j,k} = [\mathbf{0}_{j\times k} \quad I_j \quad \mathbf{0}_{j\times(M-(j+k))}] \tag{2}$$

  In this case, one can express the first LN elements of the vector Y_{int} as follows:

$$Y_{int}(1:LN) = [y_{int}(1), y_{int}(2), ..., y_{int}(LN)]^T = C_{LN,0}Y_{int} \underset{(1)}{=} C_{LN,0}\mathrm{H}_M J_M Y \tag{3}$$

- Finally, for the sake of simplicity, let us define $N' = \frac{N-1}{2}$.

2.3. Extraction of the trend vector

2.3.1. Matrix form of the trend vector with the DFA

When dealing with the DFA, the profile is split into L non-overlapping segments of length N, denoted as $\{y_{int,l}(n)\}_{l=1,\dots,L}$ with $n \in [\![1; N]\!]$. As M is not necessarily a multiple of N, the last $M - LN$ samples of the profile are not used. In this case, the l^{th} local trend, corresponding to the trend $t_l(n)$ of the l^{th} segment $y_{int,l}(n)$, is modeled by a straight line $\forall l \in [\![1; L]\!]$ and $\forall n \in [\![1; N]\!]$:

$$t_l(n) = a_{l,1}[(l-1)N + n] + a_{l,0} \tag{4}$$

Then, $\forall l \in [\![1; L]\!]$, the parameter vector $\theta_l = \begin{bmatrix} a_{l,0} \; a_{l,1} \end{bmatrix}^T$ is estimated in the LS sense from $\{y_{int,l}(n)\}_{n=1,\dots,N}$. The global trend $T_{\mathbf{1}}$ is then deduced by aggregating the local trends $\{T_l\}_{l=1,\dots,L}$. Using a vector form of (4), one has:

$$T_l = A_l \theta_l \; \forall l \in [\![1; L]\!] \tag{5}$$

where A_l is a $N \times 2$ matrix whose first column corresponds to a vector of 1s and whose second column is defined by the set of values $\{(l-1)N + n\}_{n=1,\dots,N}$. By introducing the parameter vector $\Theta_{\mathbf{1}} = \begin{bmatrix} \theta_1 \dots \theta_L \end{bmatrix}^T$ of size $2L \times 1$, and the $(LN \times 2L)$ matrix $A_{\mathbf{1}}$ which is block diagonal defined from the set of matrices $\{A_l\}_{l=1,\dots,L}$, the parameters of the local trends satisfy:

$$\arg\min_{\Theta_{\mathbf{1}}} \left|\left| C_{LN,0} Y_{int} - A_{\mathbf{1}} \Theta_{\mathbf{1}} \right|\right|^2 \tag{6}$$

This leads to:

$$\hat{\Theta}_{\mathbf{1}} = (A_{\mathbf{1}}^T A_{\mathbf{1}})^{-1} A_{\mathbf{1}}^T C_{LN,0} Y_{int} \tag{7}$$

Then, the trend vector $T_{\mathbf{1}}$ can be deduced as follows:

$$\begin{aligned} T_{\mathbf{1}} = A_{\mathbf{1}} \hat{\Theta}_{\mathbf{1}} &\underset{(7)}{=} A_{\mathbf{1}} (A_{\mathbf{1}}^T A_{\mathbf{1}})^{-1} A_{\mathbf{1}}^T C_{LN,0} Y_{int} \\ &\underset{(1)}{=} A_{\mathbf{1}} (A_{\mathbf{1}}^T A_{\mathbf{1}})^{-1} A_{\mathbf{1}}^T C_{LN,0} \mathrm{H}_M J_M Y \end{aligned} \tag{8}$$

The trend vector corresponds to the orthogonal projection of $C_{LN,0}\mathrm{H}_M J_M Y$ onto the space spanned by the columns of $A_{\mathbf{1}}$.

2.3.2. Matrix form of the trend vector with DMA

When dealing with the DMA, known as "simple moving average" [18], the profile is low-pass filtered. Indeed, the impulse response of the filter is given by:

$$h(n) = \frac{1}{N} \; for \; n = 0, ..., N-1. \tag{9}$$

Due to its symmetry, it leads to a linear-phase filter with a constant group delay equal to $\frac{N-1}{2}\frac{1}{f_s}$, where f_s denotes the sampling frequency. As the trend has to be subtracted to the integrated signal, N is chosen odd in the following. Moreover, the frequency response satisfies:

$$H(f) = \begin{cases} \frac{1}{N}\frac{sin(\frac{\pi N f}{f_s})}{sin(\frac{\pi f}{f_s})} e^{-j\frac{\pi(N-1)f}{f_s}} \text{ if } f \neq 0 \\ 1 \text{ otherwise} \end{cases}$$

Note that $|H(f)| = 0$ when $\frac{\pi N f}{f_s} = k\pi$, or equivalently when $f = \frac{k f_s}{N}$ for $k = 1, ..., N-1$. This amounts to saying that the FIR filter is defined by zeros of the transfer function which are equal to $e^{j\frac{2\pi k}{N}}$ with $k = 1, ..., N-1$. When N increases, the width $\frac{2f_s}{N}$ of the main lobe decreases. This corresponds to a low-pass filtering more and more selective when N increases.

The M samples of the profile are filtered. Instead of using a convolution at each time step, let us express the vector storing the filter output samples. This can be done by premultiplying Y_{int} by a filtering matrix M_{filt} defined as:

$$M_{filt} = \frac{1}{N}\sum_{r=0}^{N-1} diag(\mathbb{1}_{1\times M-r}, -r) \tag{10}$$

In addition, the group delay corresponding to N' samples and induced by the filter has to be compensated. This can be done by introducing another pre-multiplication by the following $M \times M$ matrix:

$$M_{comp} = diag\Big(\mathbb{1}_{1\times(M-N')}, N'\Big) \tag{11}$$

The resulting trend vector is equal to $M_{comp}M_{filt}Y_{int}$.

However, the last N' elements of this vector are equal to 0. In addition, due to the transient behavior of the filtering which corresponds to the first $N-1$

samples and the delay compensation introduced above, the first N' elements of the current trend vector should not be taken into account. For the above reasons, only a vector of size $M - N + 1$ should be considered. This amounts to adding another pre-multiplication by the matrix $C_{M-N+1,N'}$.
Therefore, the trend vector satisfies:

$$T_{\mathbf{2}} = C_{M-N+1,N'} M_{comp} M_{filt} Y_{int} \underset{(1)}{=} C_{M-N+1,N'} M_{comp} M_{filt} \mathrm{H}_M J_M Y \quad (12)$$

2.3.3. Matrix form of the trend vector with the CDFA

Instead of *a posteriori* correcting the discontinuities in the global trend of the data by using combinations of the consecutive local trends as this is done in the AFA [15], we propose to model the global trend of the profile by assuming that the consecutive local trends are continuous (see Fig. 1). The estimations of the trend parameters are then based on a constrained LS criterion. In the following, let us detail the proposed variant.

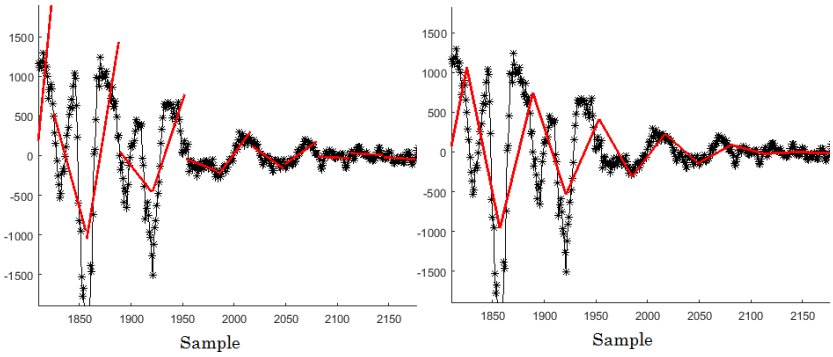


Figure 1: Global trends with DFA (a) and CDFA (b)

Minimization approach defining the CDFA. For the L segments under study, our purpose is to ensure continuity between the consecutive local trends $\forall l \in [1; L-1]$. Therefore, there are two possibilities that can be considered;

either, one has:

$$x_{l+1}(1) = x_l(N+1) \tag{13}$$

or

$$x_{l+1}(0) = x_l(N) \tag{14}$$

Given (4), defining the constraints (13) or (14) amounts to having $\forall l \in [1; L-1]$:

$$a_{l+1,0} = \beta(l)(a_{l,1} - a_{l+1,1}) + a_{l,0} \tag{15}$$

with $\beta(l) = lN + 1$ (resp. lN) if the first constraint (13) (resp. the second constraint (14)) is taken into account. Note that for both constraints, $\beta(l) - \beta(l-1) = N$. This remark will be useful for the parameter estimation step. Instead of using (15), one can consider $\forall l \in [1; L-1]$:

$$a_{l+1,0} = a_{1,0} + \sum_{j=1}^{l} \beta(j)(a_{j,1} - a_{j+1,1}) \tag{16}$$

The joint estimations of the $2L$ parameters $\{a_{l,1}\}_{l=1,...,L}$ and $\{a_{l,0}\}_{l=1,...,L}$ consists in minimizing the following criterion:

$$\begin{aligned} J(a_{1,1},..,a_{L,1},a_{1,0},..,a_{L,0}) &= \frac{1}{LN}\sum_{m=1}^{LN}(y_{int}(m) - t(m))^2 \\ &= \frac{1}{LN}\sum_{l=1}^{L}\sum_{n=1}^{N}(y_{int}((l-1)N+n) - t_l(n))^2 \end{aligned}$$

under $L-1$ constraints defined by (16). This amounts to minimizing this new criterion:

$$\begin{aligned} J(a_{1,1},..,a_{L,1},a_{1,0}) &= \sum_{n=1}^{N}(y_{int}(n) - a_{1,1}n - a_{1,0})^2 \\ &+ \sum_{l=2}^{L}\sum_{n=1}^{N}[y_{int}((l-1)N+n) - a_{l,1}[(l-1)N+n] - a_{1,0} - \sum_{j=1}^{l-1}\beta(j)(a_{j,1} - a_{j+1,1})]^2 \end{aligned} \tag{17}$$

Matrix form of the CDFA approach. Let us introduce the $(L+1) \times 1$ column parameter vector $\Theta_{\mathbf{3}} = [a_{1,1},..,a_{L,1},a_{1,0}]^T$ and the $LN \times (L+1)$ matrix $A_{\mathbf{3}}$

defined as follows:

$$A_{\mathbf{3}}(1:N,1:L+1)=\begin{bmatrix}1 & 0 & \dots & 0 & 1\\ 2 & \vdots & & \vdots & 1\\ \vdots & \vdots & & \vdots & \vdots\\ N & 0 & \dots & 0 & 1\end{bmatrix} \tag{18}$$

and $\forall l \in [2;L]$:

$$A_{\mathbf{3}}((l-1)\times N+1:lN,1:L+1)= \tag{19}$$
$$\begin{bmatrix}\beta(1) & N & \cdots & N & (l-1)N+1-\beta(l-1) & 0 & \cdots & 0 & 1\\ \beta(1) & N & \cdots & N & (l-1)N+2-\beta(l-1) & & \vdots & & \vdots\\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots\\ \beta(1) & \underbrace{N \quad \cdots \quad N}_{l-2} & & & lN-\beta(l-1) & \underbrace{0 \quad \cdots \quad 0}_{L-l} & & & 1\end{bmatrix}$$

The criterion introduced in (17) can be defined as follows:

$$\begin{aligned}J(a_{1,1},..,a_{L,1},a_{1,0}) &= \left|\left|C_{LN,0}Y_{int}-A_{\mathbf{3}}\Theta_{\mathbf{3}}\right|\right|^2 \\ &= [C_{LN,0}Y_{int}-A_{\mathbf{3}}\Theta_{\mathbf{3}}]^T[C_{LN,0}Y_{int}-A_{\mathbf{3}}\Theta_{\mathbf{3}}]\end{aligned} \tag{20}$$

Therefore, the estimate $\hat{\Theta}_{\mathbf{3}}$ satisfies:

$$\hat{\Theta}_{\mathbf{3}}=[A_{\mathbf{3}}^TA_{\mathbf{3}}]^{-1}A_{\mathbf{3}}^TC_{LN,0}Y_{int} \tag{21}$$

The trend vector $T_{\mathbf{3}}=A_{\mathbf{3}}\hat{\Theta}_{\mathbf{3}}$ can be expressed this way:

$$\begin{aligned}T_{\mathbf{3}} &\underset{(21)}{=} A_{\mathbf{3}}[A_{\mathbf{3}}^TA_{\mathbf{3}}]^{-1}A_{\mathbf{3}}^TC_{LN,0}Y_{int} \\ &\underset{(1)}{=} A_{\mathbf{3}}[A_{\mathbf{3}}^TA_{\mathbf{3}}]^{-1}A_{\mathbf{3}}^TC_{LN,0}\mathrm{H}_MJ_MY\end{aligned} \tag{22}$$

The trend vector can be seen as the orthogonal projection of $C_{LN,0}\mathrm{H}_MJ_MY$ onto the space spanned by the columns of $A_{\mathbf{3}}$.

Remark. *We have done similar developments for the AFA approach [15] as well, but they are not given in this paper for the sake of space.*

3. Comparative analysis

This section is the core of this paper. Indeed, given the matrix form of the trend vector for the DFA, the DMA and the CDFA, we first propose to express the square of the fluctuation function. When the signal is w.s.s. it corresponds to the power of the signal which has been filtered. The frequency response of the filter depends on the approach. In addition, we give an interpretation of the fluctuation function by using the frequency domain when the signal is no longer w.s.s..

3.1. Towards a uniform expression of the residual power

Given (8), (12) and (22) in the above section, the trend vector $T_{\boldsymbol{i}}$, with $\boldsymbol{i} = \mathbf{1}, \ldots, \mathbf{3}$, has been expressed as a function of the signal vector Y. The next step is to deduce the expression of the residual vector $R_{\boldsymbol{i}}$, *i.e.* the expression of the difference between the integrated signal and the trend vector of the integrated signal. To take into account the fact that the trend vector is not necessarily of the same size, the matrix $\underline{C}_{\boldsymbol{i}}$ is introduced. Thus, one has for the subscript $\boldsymbol{i} = 1, ..., 3$:

$$R_{\boldsymbol{i}} = \underline{C}_{\boldsymbol{i}} \mathrm{Y} - T_{\boldsymbol{i}} = B_{\boldsymbol{i}} Y \tag{23}$$

where $\underline{C}_{\mathbf{1}} = \underline{C}_{\mathbf{3}} = C_{LN,0} \mathrm{H}_M J_M$ for the DFA and the CDFA and $\underline{C}_{\mathbf{2}} = C_{M-N+1,N'} \mathrm{H}_M J_M$.
In table 1, the expressions of the matrices $B_{\boldsymbol{i}}$ are summarized.

Table 1: Summary of the expressions and sizes of the matrix $B_{\boldsymbol{i}}$

Methods	Definition of the matrix $B_{\boldsymbol{i}}$	Size $S_{\boldsymbol{i}}$ of the trend vector
DFA	$\left(I_{LN} - A_{\mathbf{1}}(A_{\mathbf{1}}^T A_{\mathbf{1}})^{-1} A_{\mathbf{1}}^T\right) C_{LN,0} \mathrm{H}_M J_M$	LN
DMA	$C_{M-N+1,N'} \left(I_M - M_{comp} M_{filt}\right) \mathrm{H}_M J_M$	$M - N + 1$
CDFA	$\left(I_{LN} - A_{\mathbf{3}}(A_{\mathbf{3}}^T A_{\mathbf{3}})^{-1} A_{\mathbf{3}}^T\right) C_{LN,0} \mathrm{H}_M J_M$	LN

Then, given $S_{\boldsymbol{i}}$ the size of the trend vector and introducing $\Gamma_{\boldsymbol{i}} = \frac{1}{S_{\boldsymbol{i}}} B_{\boldsymbol{i}}^T B_{\boldsymbol{i}}$, the power of the residual $F_{\boldsymbol{i}}^2(N)$, also called the square of the fluctuation function,

can be deduced as follows:

$$F_{\boldsymbol{i}}^2(N) = Tr\left(\Gamma_{\boldsymbol{i}} Y Y^T\right) \tag{24}$$

By taking advantage of the symmetry of $\Gamma_{\boldsymbol{i}}$, (24) becomes:

$$F_{\boldsymbol{i}}^2(N) = \sum_{k=1}^{S_{\boldsymbol{i}}} \Gamma_{\boldsymbol{i}}(k,k) y^2(k) + \sum_{r=1}^{S_{\boldsymbol{i}}-1} \sum_{k=1}^{S_{\boldsymbol{i}}-r} [\Gamma_{\boldsymbol{i}}(k,k+r) + \Gamma_{\boldsymbol{i}}(k+r,k)] y(k) y(k+r) \tag{25}$$

In section 3.2, this formalism will be useful to express the power of the residual from the autocorrelation function of a process assumed to be w.s.s., and consequently from its power spectral density. Nevertheless, before addressing this issue, let us express the slope α to provide all the steps of the DFA and its variants in a matrix form.

3.2. Link between the power of the residual and the PSD of a w.s.s. process

By assuming that y is w.s.s. and taking the statistical mean of (25), one has:

$$E[F_{\boldsymbol{i}}^2(N)] = \sum_{r=-S_{\boldsymbol{i}}+1}^{S_{\boldsymbol{i}}-1} Tr(\Gamma_{\boldsymbol{i}}, r) R_{y,y}(r) \tag{26}$$

where $R_{y,y}(r)$ is the autocorrelation function of the process y and $Tr(\Gamma_{\boldsymbol{i}}, r)$ denotes the r^{th} diagonal of the matrix $\Gamma_{\boldsymbol{i}}$.

As the autocorrelation function for real signals is symmetric and by denoting $g_{\Gamma_{\boldsymbol{i}}}(r) = Tr(\Gamma_{\boldsymbol{i}}, r)$, the above equation can be expressed as the result of a convolution:

$$E[F_{\boldsymbol{i}}^2(N)] = g_{\Gamma_{\boldsymbol{i}}} * R_{y,y}(\tau)|_{\tau=0} \tag{27}$$

Given the Wiener-Khintchine theorem and using the inverse Fourier transform (TF^{-1}), $E[F_{\boldsymbol{i}}^2(N)]$ can be expressed from the PSD of y, denoted as $S_{yy}(f)$:

$$\begin{aligned} E[F_{\boldsymbol{i}}^2(N)] &= TF^{-1}\Big(\big(\sum_{r=-S_{\boldsymbol{i}}+1}^{S_{\boldsymbol{i}}-1} Tr(\Gamma_{\boldsymbol{i}}, r) e^{-j2\pi f_n r} \big) S_{yy}(f) \Big)\big|_{\tau=0} \\ &= TF^{-1}\Big(\Psi_{\boldsymbol{i}}(f) S_{yy}(f) \Big)\big|_{\tau=0} \end{aligned} \tag{28}$$

In (28), $\Psi_{\boldsymbol{i}}(f) = \sum_{r=-S_{\boldsymbol{i}}+1}^{S_{\boldsymbol{i}}-1} Tr(\Gamma_{\boldsymbol{i}}, r) e^{-j2\pi f_n r}$ corresponds to the Fourier transform of the sequence $\{Tr(\Gamma_{\boldsymbol{i}}, r)\}_{r=-S_{\boldsymbol{i}}+1,\ldots,S_{\boldsymbol{i}}-1}$. Let us look at the properties

of the latter: first of all, as it is real and even, $\Psi_{\boldsymbol{i}}(f)$ is necessarily real and even. Moreover, as $\Gamma_{\boldsymbol{i}}$ is a Gramian matrix since it is the product between $\frac{1}{\sqrt{S_{\boldsymbol{i}}}}B_{\boldsymbol{i}}$ and its transpose, the element $\Gamma_{\boldsymbol{i}}(i,j)$ located at the i^{th} row and the j^{th} column of $\Gamma_{\boldsymbol{i}}$ corresponds to the scalar product between the i^{th} and the j^{th} rows of $\frac{1}{\sqrt{S_{\boldsymbol{i}}}}B_{\boldsymbol{i}}$. Therefore, taking advantage of the properties of the scalar product, one has:

$$|\Gamma_{\boldsymbol{i}}(i,j)| \leq |\Gamma_{\boldsymbol{i}}(i,i)| \tag{29}$$

As a corollary, using the inequality (29), one has:

$$\begin{aligned} |Tr(\Gamma_{\boldsymbol{i}},r)| &\leq \sum_{k=1}^{S_{\boldsymbol{i}}-r} |\Gamma_{\boldsymbol{i}}(k,k+r)| \leq \sum_{k=1}^{S_{\boldsymbol{i}}-r} \Gamma_{\boldsymbol{i}}(k,k) \\ &\leq \sum_{k=1}^{S_{\boldsymbol{i}}-1} \Gamma_{\boldsymbol{i}}(k,k) = Tr(\Gamma_{\boldsymbol{i}},0) = Tr(\Gamma_{\boldsymbol{i}}) \end{aligned}$$

In the above, note that $Tr(\Gamma_{\boldsymbol{i}})$ corresponds to the square of the Froebenius norm of the matrix $\Gamma_{\boldsymbol{i}}$. It is necessarily positive and the maximum value of the sequence of the traces. As a consequence, the sequence can be seen as the convolution of a vector with its flipped version and its Fourier transform $\Psi_{\boldsymbol{i}}(f)$ is necessarily positive.

Therefore $\Psi_{\boldsymbol{i}}(f)S_{yy}(f)$ can be seen as the PSD of the signal y filtered by a filter whose transfer function $H_{filter,\boldsymbol{i}}(z)$ satisfies: $\Psi_{\boldsymbol{i}}(f) = |H_{filter,\boldsymbol{i}}(z)|^2_{z=exp(j\theta)}$, with $\theta = 2\pi f/f_s$ the normalized angular frequency. Consequently, we can conclude that $E[F^2_{\boldsymbol{i}}(N)]$ corresponds to the autocorrelation function of the filter output calculated for the lag equal to 0, *i.e.* the power of the filter output.

Remark: about the FA approach

Let us recall the definition of the criterion introduced by Peng [10] and denoted as $F_{\boldsymbol{0}}(N)$:

$$\begin{aligned} F^2_{\boldsymbol{0}}(N) &= \Big\langle (y_{int}(i+N) - y_{int}(i))^2 \Big\rangle = \left\langle \left(\sum_{j=i+1}^{i+N} y(j) \right)^2 \right\rangle \\ &= \left\langle \sum_{j=i+1}^{i+N} y^2(j) + 2 \sum_{j=i+1}^{i+N-1} \sum_{k=j+1}^{i+N} y(j)y(k) \right\rangle \end{aligned} \tag{30}$$

The expectation of $F^2_{\mathbf{0}}(N)$ can then be expressed from the autocorrelation function $R_{y,y}(\tau)$ of the signal as follows:

$$E[F^2_{\mathbf{0}}(N)] = \sum_{r=1-N}^{N-1} \Big(Tr(\Gamma_{\mathbf{0}}, r) \Big) R_{y,y}(r) \tag{31}$$

where $\Gamma_{\mathbf{0}}$ is a square matrix of size N whose every element is equal to 1. From the definition of $\Gamma_{\mathbf{0}}$, the sequence $g_{\Gamma_{\mathbf{0}}}(r) = Tr(\Gamma_{\mathbf{0}}, r)$ is a triangular function. The filtering induced by the FA approach is low-pass and reduces to the integration of the process under study.

In this subsection, the signal is assumed to be w.s.s., but, in several applications, this is not necessarily the case. For this reason, we propose to address this issue in the next subsection.

3.3. Link between the power of the residual and the Wigner-Ville transform of the process, when it is not w.s.s.

As the signal is no longer assumed to be w.s.s., we suggest introducing a time frequency analysis, namely the Wigner-Ville transform, which can be seen as the instantaneous correlation function of the signal [54].

From (25), it comes that $F^2_{\boldsymbol{i}}(N)$ can be expressed as a weighted sum of the instantaneous correlations terms $\{y(k)y(k+r)\}_{r=-M+1,\dots,M-1}$. Let us rewrite it by using the following two matrices of size $(2M-1 \times 2M)$:

On the one hand, Y_{corr} has its k^{th} column storing the instantaneous correlations of the signal y at time k for the lags varying from $M-1$ to $1-M$:

$$Y_{corr} = \begin{bmatrix} y(1)y(1-(M-1)) & \dots & y(M)y(M-(M-1)) \\ y(1)y(1-(M-2)) & \dots & y(M)y(M-(M-2)) \\ \vdots & \vdots & \vdots \\ y(1)y(1) & \dots & y(M)y(M) \\ \vdots & \vdots & \vdots \\ y(1)y(1+(M-2)) & \dots & y(M)y(M+(M-2)) \\ y(1)y(1+(M-1)) & \dots & y(M)y(M+(M-1)) \end{bmatrix} \tag{32}$$

On the other hand, the weighting matrix $W_{\boldsymbol{i}}$ is filled with zeros. Its i-th anti-diagonal[2], with $i = 1, ..., M$, is of length M and equal to $[\Gamma_{\boldsymbol{i}}(1,i), \Gamma_{\boldsymbol{i}}(2,i), \ldots, \Gamma_{\boldsymbol{i}}(M,i)]$. Thus, for $M = 3$, one has:

$$W_{\boldsymbol{i}} = \begin{bmatrix} 0 & 0 & \Gamma_{\boldsymbol{i}}(3,1) \\ 0 & \Gamma_{\boldsymbol{i}}(2,1) & \Gamma_{\boldsymbol{i}}(3,2) \\ \Gamma_{\boldsymbol{i}}(1,1) & \Gamma_{\boldsymbol{i}}(2,2) & \Gamma_{\boldsymbol{i}}(3,3) \\ \Gamma_{\boldsymbol{i}}(1,2) & \Gamma_{\boldsymbol{i}}(2,3) & 0 \\ \Gamma_{\boldsymbol{i}}(1,3) & 0 & 0 \end{bmatrix} \tag{33}$$

Let us now represent some examples of $W_{\boldsymbol{i}}$ in Fig. 2 with $N = 9$ and $N = 21$. The colormap was chosen to distinguish the null values from the negative ones and the positive ones.

[2]The i^{th} anti-diagonal of the matrix $W_{\boldsymbol{i}}$ corresponds to the set of elements located at the $(2M + 1 - j - i)^{th}$ row and the j^{th} column, with $j = 1, ..., M$.

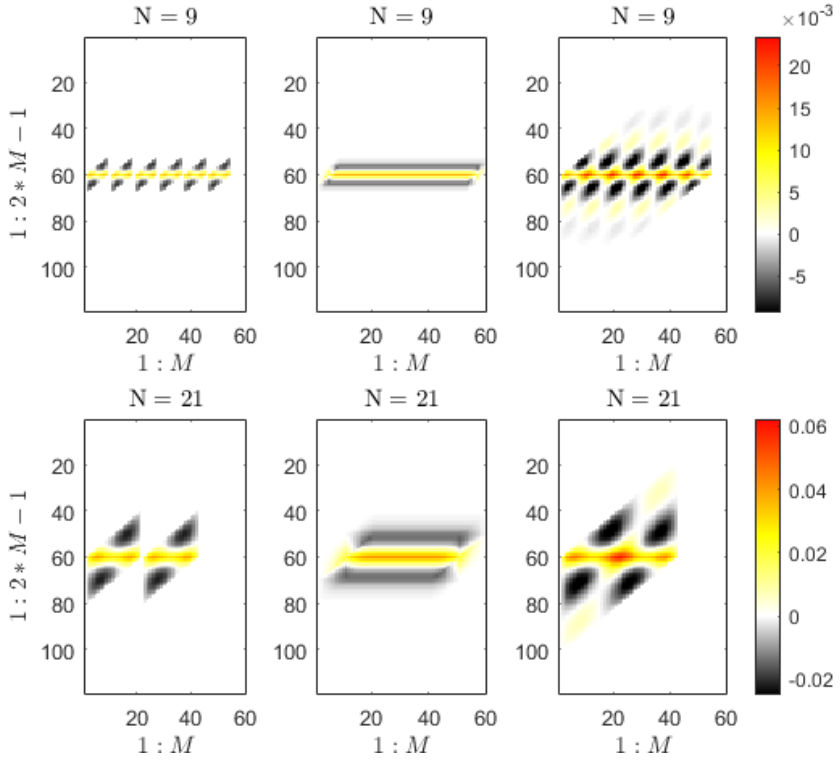


Figure 2: Weighting matrices $W_{\mathbf{1}}$ (left), $W_{\mathbf{2}}$ (center) and $W_{\mathbf{3}}$ (right) for $M = 60$.

Given (32) and (33), the square of the fluctuation function (25) can be seen as the two-dimensional Fourier transform (2D-FT) of the element-wise multiplication of Y_{corr} and $W_{\boldsymbol{i}}$ for the set of the spatial frequencies $(u, v) = (0, 0)$. Given the properties of the Fourier transform, $F^2_{\boldsymbol{i}}(N)$ can also be expressed as the convolution between the 2D-FTs of Y_{corr} and $W_{\boldsymbol{i}}$, for $(u, v) = (0, 0)$:

$$\begin{aligned} F^2_{\boldsymbol{i}}(N) &= \mathscr{F}(Y_{corr} W_{\boldsymbol{i}})|_{(u=0,v=0)} \\ &= \Big(\mathscr{F}(Y_{corr}) \circledast \mathscr{F}(W_{\boldsymbol{i}})\Big)|_{(u=0,v=0)} \end{aligned} \tag{34}$$

where u and v are the normalized spatial frequencies, $\mathscr{F}$ denotes the 2D-FT and $\circledast$ the convolution.

Therefore, one way to compare the DFA with its variants is to compare the properties of the 2D-FT of $W_{\boldsymbol{i}}$.

Remark: link with the w.s.s. case.

Let us analyze $E[F_{\boldsymbol{i}}^2(N)]$ when y is w.s.s. In this case, $E[Y_{corr}]$ is a matrix where each row contains the same element, which corresponds to the autocorrelation function of y for a specific lag. By neglecting the windowing influence, the modulus of the 2D-FT of $E[Y_{corr}]$ is null for every couple (u, v), except for $u = 0$, where it corresponds to the PSD of the process y at the frequency v. Thus, convolving $\mathscr{F}(W_{\boldsymbol{i}})$ with $E[Y_{corr}]$ amounts to filtering the process y with a band-pass filter whose central frequency decreases with N.

3.4. Summary of the section

This section aimed at providing an interpretation of the fluctuation function for both stationary and non-stationary processes. This interpretation is given for the DFA and the DMA, as well as the CDFA we proposed in 2.3.3. It is based on a uniform way of expressing the trend vector extracted by the approaches. Moreover, unlike in [48] and [52], no approximation is made. In the next section, we compare the DFA and its variants by using both our framework and simulations.

4. Simulation results

Let us compare the DFA, the DMA and the CDFA. The usual way would be to evaluate the performance of each approach by estimating the Hurst exponent of synthetic pure mono-fractal signals. Our work provides another type of comparison, based on both the filtering interpretation and the analysis of the 2D-FT of $W_{\boldsymbol{i}}$ we introduced in the above section.

4.1. When dealing with w.s.s. case

Given section 3, let us first look at the properties of $\Psi_{\boldsymbol{i}}(f)$ where the subscript $\boldsymbol{i}$ refers to the method that is used, *i.e* the DFA ($\boldsymbol{i} = \mathbf{1}$), the DMA ($\boldsymbol{i} = \mathbf{2}$) or the CDFA ($\boldsymbol{i} = \mathbf{3}$), before highlighting the differences between our work and the approaches presented in [48] and [52].

4.1.1. Comparative study based on the filtering interpretation

As $\Psi_{\boldsymbol{i}}(f)$ *a priori* depends on N, let us study the influence of N for a given signal of length M. Using (28) and the expressions of $\Psi_{\boldsymbol{i}}(f)$ summarized in 3.1 as well as Fig. 3, the following comments can be made:

1. When $N = 3$, the filters associated with the methods DFA and DMA are high-pass whereas they become band-pass for larger values of N. In the case of the CDFA, the filter is always band-pass. The null frequency is always rejected, which is consistent with the purpose of detrending. According to the simulations we carried out, the orders of magnitude of $\Psi_{\mathbf{1}}(0)$, $\Psi_{\mathbf{3}}(0)$ are equal to 10^{-16} whereas the one of $\Psi_{\mathbf{2}}(0)$ is equal to 10^{-17}.

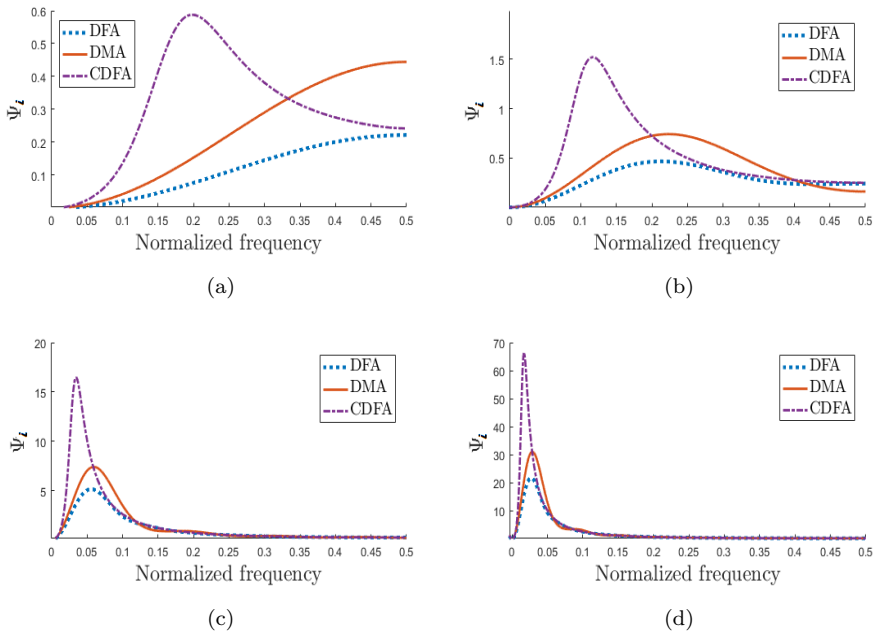
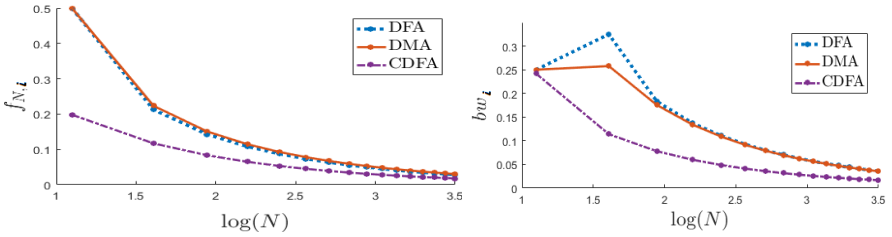


Figure 3: Comparison between frequency responses of $\Psi_{\boldsymbol{i}}(f)$. (a) $N = 3$, (b) $N = 5$, (c) $N = 17$ (d) $N = 35$.

2. In the following, let $bw_{\boldsymbol{i}}$ be the -3 dB bandwidth[3] of the filter associated to $\Psi_{\boldsymbol{i}}(f)$. Fig. 4a and 4b respectively show the evolution of the resonance frequency and the -3 dB bandwidth as a function of $\log(N)$. For every method, when looking at the right-hand side of Fig. 4b, the larger N, the smaller $bw_{\boldsymbol{i}}$ and the spikier the resonances of the frequency responses. The latter also move to low frequencies when N increases according to Fig. 4a.
3. For each value of N, the CDFA provides the spikiest and lowest resonance among the three approaches. See Fig. 3 and 4.



(a) (b)

Figure 4: Evolution as a function of $\log(N)$ of: (a) the resonance frequency of $\Psi_{\boldsymbol{i}}(f)$; (b) the frequency bandwidth (-3 dB) of $\Psi_{\boldsymbol{i}}(f)$.

4. As $\Psi_{\boldsymbol{i}}(f) = |H_{filter,\boldsymbol{i}}(z)|^2_{z=exp(j\theta)}$, measuring the difference between two frequency responses can be of interest. This can be done in many ways. In Fig. 5, the log spectral distances (LSD) [55] between the frequency responses of two among the three approaches have been computed for different values of N. In its standard definition, the LSD is expressed with an integral where the variable is the frequency. However, in practical cases, only the discrete Fourier transform (DFT) of the sequence $g_{\Gamma_{\boldsymbol{i}}}(r) = Tr(\Gamma_{\boldsymbol{i}}, r)$ can be computed. Therefore, the integral is approximated by a

[3]It corresponds to the frequencies for which $10 \log \frac{\Psi_{\boldsymbol{i}}(f)}{\Psi_{\boldsymbol{i}}(f_{N,\boldsymbol{i}})} > -3$

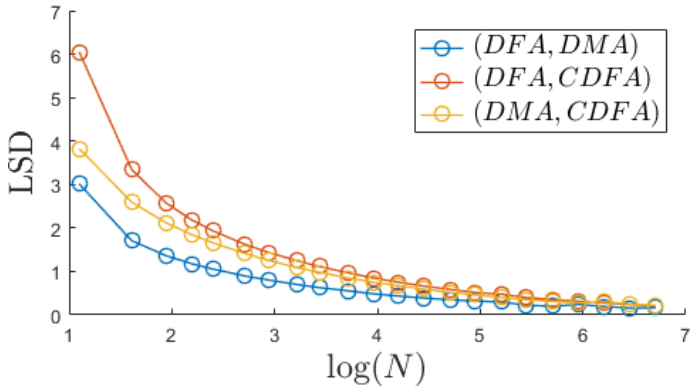


Figure 5: Evolution as a function of $\log(N)$ of the log spectral distance (LSD) [55] between $\Psi_{\boldsymbol{i}1}(f)$ and $\Psi_{\boldsymbol{i}2}(f)$.

discrete sum:

$$LSD(\mathscr{M}_1, \mathscr{M}_2) = \sqrt{\frac{1}{\mathcal{M}} \sum_{i=1}^{\mathcal{M}} \left(10 \log \left(\frac{\mathscr{M}_1(i)}{\mathscr{M}_2(i)} \right) \right)^2} \tag{35}$$

with $\mathscr{M}_1$ and $\mathscr{M}_2$ of size $\mathcal{M}$ the vectors storing the two discrete spectra. In order to minimize the approximation error, the two spectra can be computed by taking the DFT of the sequence $g_{\Gamma_{\boldsymbol{i}}}(r) = Tr(\Gamma_{\boldsymbol{i}}, r)$ with a zero-padding. In our simulations, $\mathcal{M}$ is taken equals to 2^{13}.

Whatever the two compared methods, the LSD tends to decrease when N increases. There is a value N_{min} from which the LSD between two filter responses remain more or less unchanged. In our simulations, this value of LSD was set at 2. This induces that the difference between the fluctuation functions are negligible when $N \geq N_{min}$. Therefore, computing the slope of a set of values $\{F_{\boldsymbol{i}}(N)\}$, with values of N that are large enough ($N \geq N_{min}$), will provide similar results between the three approaches. According to our analysis, $N_{min} = 35$ (*i.e.* $\log(N_{min}) = 3.5$).

4.1.2. Comparison with existing works

Let us now show the differences between our work and the approaches presented in [52] and [48]. Thus, Höll [52] suggests approximating $F^2_{\boldsymbol{1}}(N)$ by a

weighted sum of the estimates of the autocorrelation function $\hat{R}_{y,y}(r)$ as follows:

$$F_{\mathbf{1}}^2(N) \approx \sum_{r=1}^{LN-1} L_r(N)\hat{R}_{y,y}(r) \tag{36}$$

where:

$$\begin{cases} L_r(N) = \frac{1}{3N^2}(-r^3 + 3r^2N + (-3N^2+1)r + N^3 - N) \text{ for } r \neq 0 \\ L_0(N) = \frac{N^2-1}{6N} \end{cases} \tag{37}$$

whereas in our work and according to (26), $R_{y,y}(r)$ is weighted by $Tr(\Gamma_{\mathbf{1}}, r)$. Therefore, by taking the Fourier transform of the set of weights $\{L_r(N)\}_{r=1,...,LN-1}$, the frequency response of the DFA, $\hat{\Psi}_{\mathbf{1},[52]}(f)$, which would be obtained from [52] can be defined and compared with $\Psi_{\mathbf{1}}(f)$.

In addition, we propose to define the frequency response $\hat{\Psi}_{\mathbf{1},[48]}(f)$ that can be deduced from Kiyono's work [48] by using our formalism. After some mathematical developments, we can show that:

$$\begin{aligned} \hat{\Psi}_{\mathbf{1},[48]}(\frac{k}{M}f_s) \approx \frac{2}{16(\frac{k}{M})^6\pi^6N^4}\big(&2\pi^4(\frac{k}{M})^4N^4 - 4\pi^2(\frac{k}{M})^2N^2 - 3 \\ &+(3 - 2\pi^2(\frac{k}{M})^2N^2)\cos(2\pi(\frac{k}{M})N) + 6\pi(\frac{k}{M})N\sin(2\pi(\frac{k}{M})N)\big) \end{aligned} \tag{38}$$

for the frequency $\frac{k}{M}f_s$, $\forall k = \{1, ..., \frac{M-1}{2}\}$ with f_s the sampling frequency.

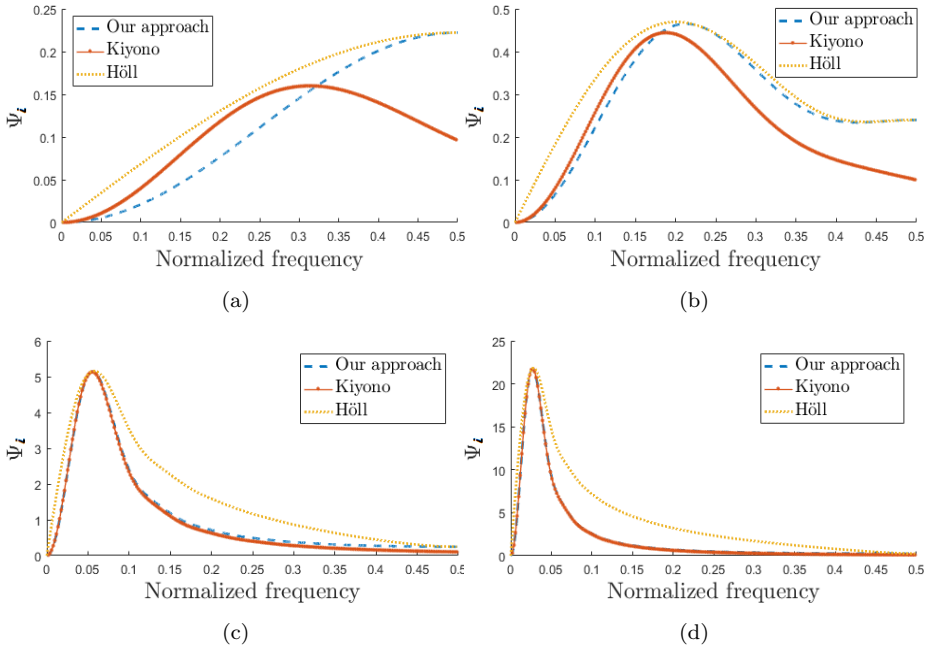


Figure 6: Comparison between frequency responses. (a) $N = 3$, (b) $N = 5$, (c) $N = 17$ (d) $N = 35$.

Given the comparison between the frequency responses in Fig. 6 for various values of N, we can conclude that Kiyono's approach [48] leads to frequency responses that are the closest to ours. However, for small values of N, the frequency responses deduced from [52] and [48] are far from the ones we obtain with our approach where no approximation is made. There is a sharp difference, especially in high frequency.

In the next subsection, another way of comparison between these approaches is provided in the general case.

4.2. When dealing with the general case

In this part, we extend our previous analysis to any kind of process. For a non-stationary process, the frequency content may vary over time. For this reason, a time frequency analysis must be considered in order to study the influence of each method on this type of process. Since comparing the methods

amounts to comparing the matrices $W_{\boldsymbol{i}}$, this section presents their analysis, in both the time and frequency domains. The 2D-FT of the weighting matrix $W_{\boldsymbol{i}}$ are represented in Fig. 7 for the DFA, the DMA and the CDFA.

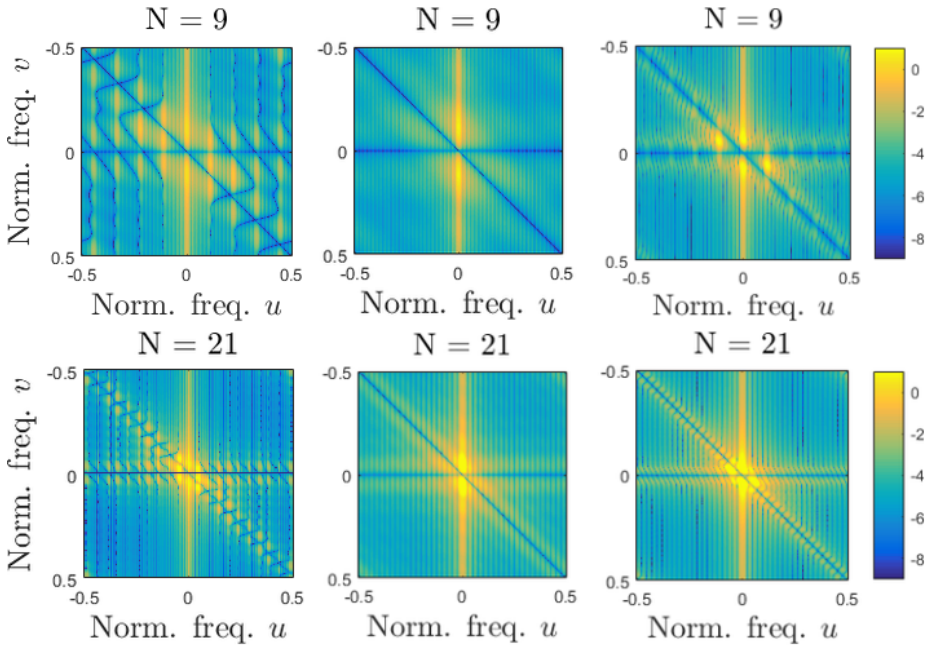


Figure 7: $\log(|\mathscr{F}(W_{\boldsymbol{i}})|)$ with $\boldsymbol{i} = \mathbf{1}$ for DFA (left), $\boldsymbol{i} = \mathbf{2}$ for the DMA (center) and $\boldsymbol{i} = \mathbf{3}$ the CDFA (right) for $M = 60$ and $N = \{9, 21\}$.

Let us first give some comments on $\log(|\mathscr{F}(W_{\mathbf{1}})|)$. By taking advantage of the properties of the 2D-FT[4], the large values of $\log(|\mathscr{F}(W_{\mathbf{1}})|)$ appearing when the normalized frequency u is a multiple of $1/N$ are representative of two features of the weighting matrices $W_{\mathbf{1}}$:

[4]Let us recall the interpretation of a 2D-FT analysis. Let two Dirac pulses be located at (u, v) and $(-u, -v)$ in the frequency domain. This leads to a vector of coordinates $(2u, 2v)$ which defines a position. In the spatial domain, this corresponds to a sinusoid of frequency $\sqrt{(u^2 + v^2)}$ along this direction, and to a constant when looking perpendicular to the direction. When dealing with images, this corresponds to equally-spaced bands along the direction.

1. The periodicity of $W_{\mathbf{1}}$ along the x-axis induces the large values of $\log(|\mathscr{F}(W_{\mathbf{1}})|)$ on both sides of the axis $v = 0$.
2. The non-null anti-diagonals of $W_{\mathbf{1}}$ induce the large values located on both sides of the axis $v = -u$.

For the DMA, $\log(|\mathscr{F}(W_{\mathbf{2}})|)$ is mainly characterized by normalized frequencies around $u = 0$ and $v = -u$. The same comments can can be done for the CDFA. Let us now study how $\log(|\mathscr{F}(W_{\boldsymbol{i}})|)$ evolves when N increases.

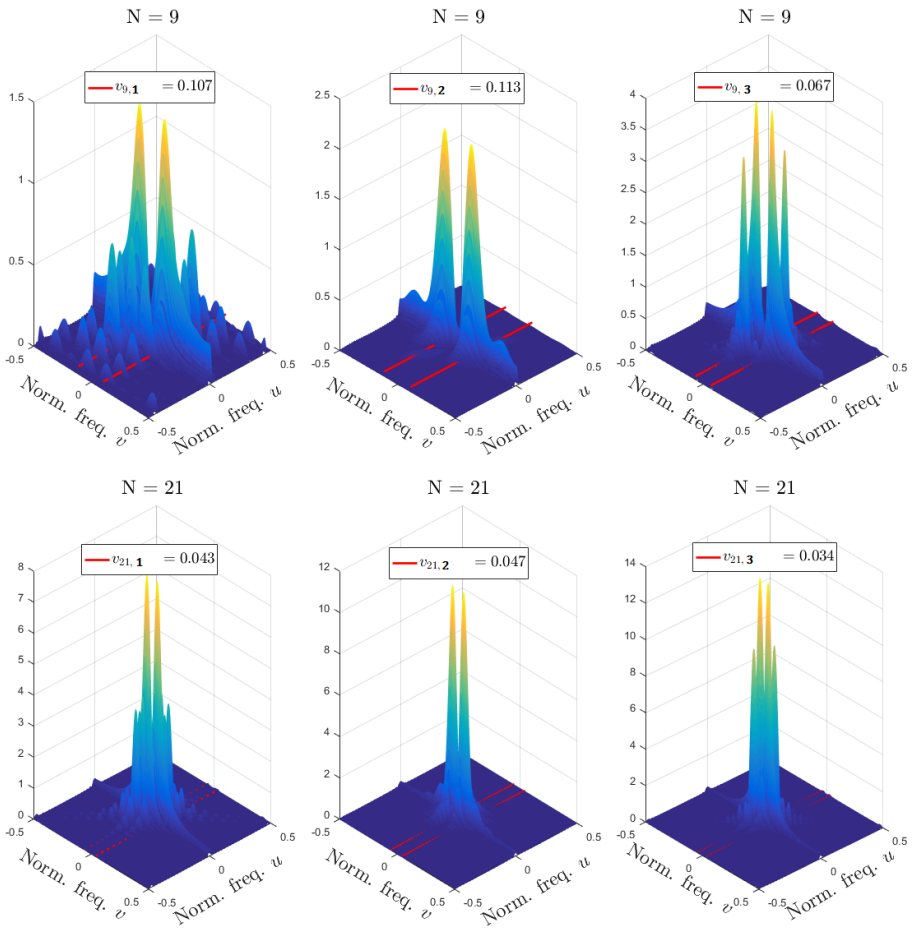


Figure 8: $|\mathscr{F}(W_{\boldsymbol{i}})|$ with $\boldsymbol{i} = \mathbf{1}$ for the DFA (left), $\boldsymbol{i} = \mathbf{2}$ for the DMA (center) and $\boldsymbol{i} = \mathbf{3}$ for the CDFA (right) vs. normalized frequencies u and v, for $M = 60$ and $N = \{9,\ 21\}$.

As shown in Fig. 8, for these three approaches, the contribution of $(u = 0, v = 0)$ remains close to 0. When N increases, $W_{\boldsymbol{i}}$ is more and more composed of frequencies $\sqrt{u^2 + v^2}$ that become smaller and smaller. For each approach, $|\log(F(W_{\boldsymbol{i}})|$ has two main lobes located at $(0, v_{N,\boldsymbol{i}})$ and $(0, -v_{N,\boldsymbol{i}})$ where $v_{N,\boldsymbol{i}}$ becomes smaller when N increases.
After these general comments, more specific ones are given in the next subsection.

4.2.1. Impact of the shape of the weighting matrices on the power of the residual

The difference between the weighting matrices causes a difference in the computation of the power of the residual. Let us thus analyze what mainly impacts on the computation of $F^2_{\boldsymbol{i}}(N)$.

1. **DMA case:** According to (34), the square of the fluctuation function corresponds to the convolution between the 2D-FT of Y_{corr} and $W_{\boldsymbol{2}}$ at the normalized frequencies $u = 0$ and $v = 0$. As the modulus of the 2D-FT of $W_{\boldsymbol{2}}$ exhibits two main peaks at $\pm v_{N,\boldsymbol{2}}$ (see Fig. 8), the frequency components of Y_{corr} at $u = 0$ and $v = \mp v_{N,\boldsymbol{2}}$ are the most amplified during the convolution step. In other words, the convolution in (34) amounts to emphasizing the frequency of Y_{corr} located at $v = \pm v_{N,\boldsymbol{2}}$.
2. **DFA case:** When the DFA is used, there is not only the above phenomenon but others that are due to the various lobes of $|\mathscr{F}(W_{\boldsymbol{1}})|$ appearing at each normalized spatial frequency u multiple of $\frac{1}{N}$.
3. **CDFA case:** From Fig. 8, when looking at the modulus of the 2D-FT of $W_{\boldsymbol{3}}$, there are two main lobes at normalized frequencies $(0, \pm v_{N,\boldsymbol{3}})$, and two secondary lobes with non-negligible amplitudes located at the same normalized frequencies than the two main secondary lobes of $|\mathscr{F}(W_{\boldsymbol{1}})|$, *i.e.* at $u = 1/N$ and $u = -1/N$. Therefore, the CDFA appears to be intermediate between the DFA and the DMA.

After this rough analysis, we suggest introducing objective criteria to compare the three methods.

4.2.2. Comparison based on the log spectral distance.

As we did for $\Psi_{\boldsymbol{i}}$, let us study the evolution of the similarities between $|\mathscr{F}(W_{\boldsymbol{i}})|$, as a function of $\log(N)$ for the DFA, DMA and CDFA. For this purpose, we define an extension of the LSD, applied to two square matrices $\mathscr{M}_1$ and $\mathscr{M}_2$ of size $\mathcal{M}$:

$$LSD_2(\mathscr{M}_1, \mathscr{M}_2) = \sqrt{\frac{1}{\mathcal{M}^2} \sum_{i=1}^{\mathcal{M}} \sum_{j=1}^{\mathcal{M}} \left(10 \log\left(\frac{\mathscr{M}_1(i,j)}{\mathscr{M}_2(i,j)}\right)\right)^2} \tag{39}$$

as well as another distance metric defined as follows :

$$D(\mathscr{M}_1, \mathscr{M}_2) = \sqrt{\frac{1}{\mathcal{M}^2} \sum_{i=1}^{\mathcal{M}} \sum_{j=1}^{\mathcal{M}} \frac{|\mathscr{M}_1(i,j) - \mathscr{M}_2(i,j)|}{|\mathscr{M}_1(i,j) + \mathscr{M}_2(i,j) + \epsilon|}} \tag{40}$$

with $\epsilon \ll 1$. The evolution of both criteria are respectively represented in Fig. 9a and Fig. 9b. As in 4.1.1, the computation of the spectra is done using zero-padding, with $\mathcal{M} = 2^{13}$.

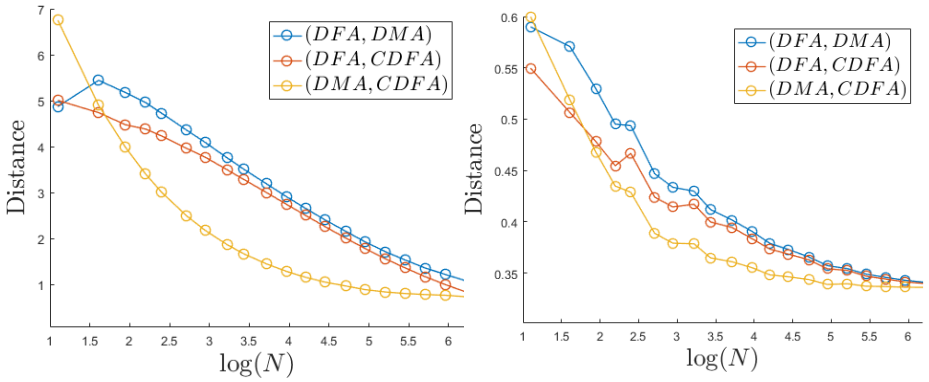


Figure 9: Evolution of the distance between $|\mathscr{F}(W_{\boldsymbol{i}1})|$ and $|\mathscr{F}(W_{\boldsymbol{i}2}|$: (a) with the metric LSD_2; (b) with the metric D.

No matter the metric used, the criterion decreases as $\log(N)$ increases. As in the stationary case, the larger N, the closer the estimation of $F_{\boldsymbol{i}}(N)$ by the three methods. As a consequence, computing the slope of a set of values $\{F_{\boldsymbol{i}}(N)\}$,

with values of N that are large enough, will provide similar results between the three approaches.

In the next section, a comparison is done between the behaviours of these approaches on stationary and non-stationary mono-fractal processes.

4.3. Comparative study based on the estimation of the Hurst exponent of mono-fractal signals

In this subsection, the comparative study is made between the DFA and CDFA, but the same type of analysis could also be done with the DMA . The synthetic mono-fractal signals studied in this section consist of two types of signals. The first are white Gaussian noises known to have a prescribed value of the Hurst exponent H equal to -0.5. The second are Weierstrass functions (WEI) with prescribed values[5] of $H = 0.9$.

4.3.1. Comparison between the DFA and the CDFA on 500 white noises

In Fig. 10, $\log(F_{\mathbf{i}}(N))$ is represented as a function of $\log(N)$ for the DFA and the CDFA for one realization of a white noise. Two slopes are computed. The first is based on the smallest values of N whereas the second is computed by using the largest. The slopes obtained with the DFA and the CDFA tend to be the same if large values of N ($N \geq N_{min}$) are used. It is coherent with the filtering analysis we did in the previous section, where we noticed that the LSD between $\Psi_{\mathbf{1}}(f)$ and $\Psi_{\mathbf{3}}(f)$ becomes smaller and smaller as N increases.

[5]Several simulations on processes characterized by different Hurst coefficients were conducted. However, for the sake of simplicity, only results are presented for $H = 0.9$ in the paper.

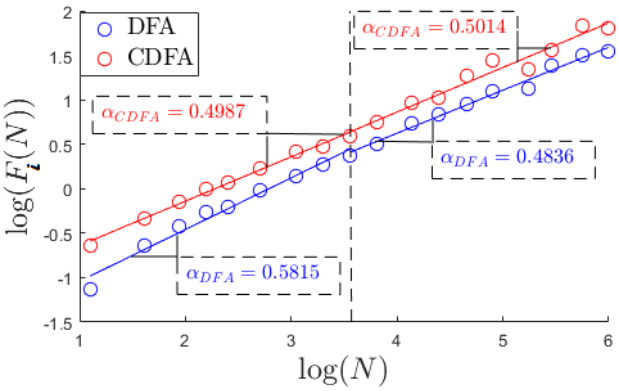


Figure 10: Evolution of $\log(F_{\boldsymbol{i}}(N))$ as a function of $\log(N)$ for both the DFA and the CDFA, in the case of one realization of a white noise. Theoretical expected value: $\alpha = 0.5$.

In Table 2, the mean and the variance obtained on 500 white noises are given. The CDFA provides more accurate estimations of α for small values of N ($N \leq N_{min}$) and when large values of N ($N \geq N_{min}$) are considered. In addition, the difference between the estimation based on small and large values of N is smaller with the CDFA than with the DFA. Therefore, the CDFA is more reliable than the DFA[6].

	Mean	Variance	% err.
DFA	0.592	3.29×10^{-4}	18.4
CDFA	0.507	5.16×10^{-4}	1.40

	Mean	Variance	% err.
DFA	0.487	3.02×10^{-3}	2.60
CDFA	0.491	5.43×10^{-3}	1.80

Table 2: Comparison of the mean and variance values of α for each approach, estimated on 500 white noises for different values of N: when $N \leq N_{min}$ (left) and $N \geq N_{min}$ (right). Theoretical expected value: $\alpha = 0.5$.

[6] As an alternative, one could use the approach proposed in section 3.1 of [49] correcting the values of the fluctuation function by multiplicating them with a corrective term for small values of N.

4.3.2. Comparative study on Weierstrass functions

Weierstrass functions (WEI) are continuous nowhere-differentiable functions [56]. Each WEI is basically a sum of damped sines with increasing frequencies. As its Holder exponent is the same at each time instant, the value of the Hurst exponent is equal to the Holder exponent. In our experiments, 500 stochastic WEI are generated[7] with a prescribed value $H = 0.9$.

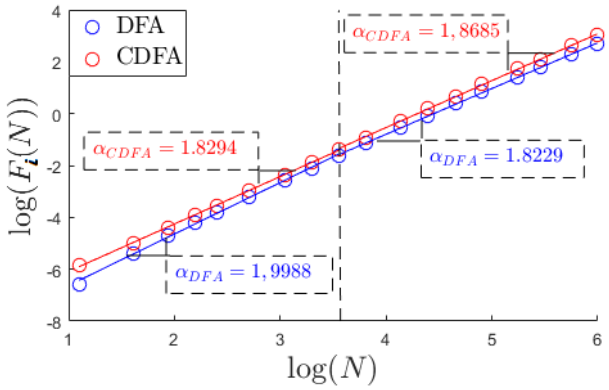


Figure 11: Evolution of $\log(F_{\boldsymbol{i}}(N))$ as a function of $\log(N)$ for both the DFA and the CDFA, in the case of one realization of a WEI process with $H = 0.9$. Theoretical expected value: $\alpha = 1.9$.

Fig. 11 shows the evolution of $\log(F_{\boldsymbol{i}}(N))$ as a function of $\log(N)$ for one realization whereas Table 3 provides the mean values and the variances of α for the DFA and the CDFA.

The results we obtain can be explained by the following reasons: unlike a white process, the power of a WEI with $H = 0.9$ is rather located in low frequencies. This means that the values of $F_{\boldsymbol{i}}(N)$ mainly depend on the properties of $\Psi_{\boldsymbol{i}}(f)$ in low frequencies. The filtering analysis we did in the previous section showed that the resonance of $\boldsymbol{\Psi_3}$ is located in lower frequencies that the one of $\boldsymbol{\Psi_1}$. In addition, the difference between them tends to be smaller when N increases.

[7] This can be done by using the free Matlab Toolbox FracLab available at the following url: `https://project.inria.fr/fraclab/`. See also [57].

	Mean	Variance	% err.		Mean	Variance	% err.
DFA	1.997	4.51×10^{-3}	5.11	DFA	1.816	1.29×10^{-2}	4.42
CDFA	1.832	1.18×10^{-2}	3.59	CDFA	1.868	2.53×10^{-2}	1.68

Table 3: Mean and variance values of α for each approach, estimated on 500 WEI processes with $H = 0.9$ for different values of N: when $N \leq N_{min}$ (left) and $N \geq N_{min}$ (right). Theoretical expected value: $\alpha = 1.9$.

This explains that $\log(F_{\mathbf{3}}(N))$ is always larger than $\log(F_{\mathbf{1}}(N))$. In addition, as the LSD between $\Psi_{\mathbf{1}}(f)$ and $\Psi_{\mathbf{3}}(f)$ becomes smaller and smaller as N increases, the difference between the estimations of H tends to become smaller and smaller.

4.4. Application on human biological processes

4.4.1. Presentation of the experiments

This subsection aims at presenting an application of our work on real processes. In particular, our purpose is to show how our framework could explain the results obtained with the DFA and its variants for the estimation of α in the following use-case: the analysis of the regularity of an airline pilot's gaze position for the detection of the visual tunneling state.

This cognitive state is defined as the inability for an individual to perceive an unexpected change in the visual scene [58]. It has been proven to be a major factor of accidents in specific areas [59]. Our analysis is based on the regularity of the gaze position, as some other ocular features are good indicators of the visual tunneling [60–62], and is collected from a non intrusive sensor, which is crucial to make this system acceptable to a pilot[8].

Our experiments were based on the multitasking simulator NASA MATB-II software [63], because of its recognized similarities with pilot activities in the cockpit. See Fig. 12 for an illustration of the software MATB-II interface.

[8]Other processes, such as ElectroEncephaloGraphy (EEG), ElectroCardioGraphy (ECG), ElectroDermal Activity (EDA), functional Magnetic Resonance Imaging (fMRI) are relevant as well in this application, but they are intrusive, and therefore not acceptable.

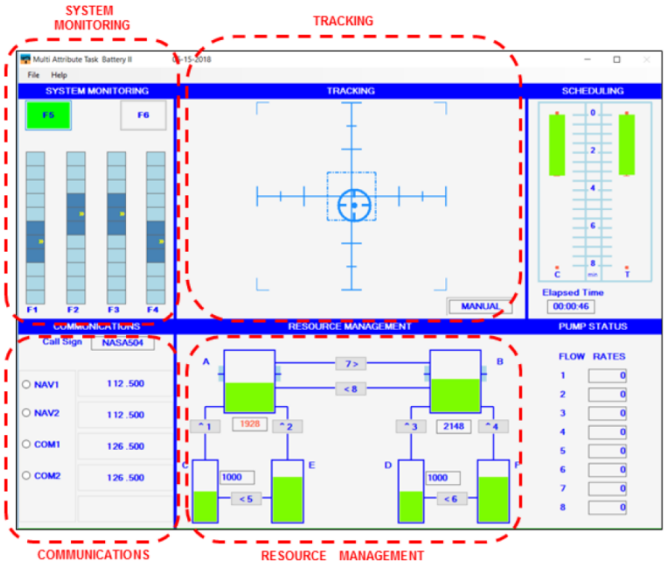


Figure 12: Interface of the MATB-II software

Given the literature on mental workload [64–66], we decided to expose subjects to short-duration scenarios leading to three levels of mental workload ("Low" , "Medium" and "High"). The latter differ by the frequency and difficulty of the tasks asked to the subjects.

Each experiment operates with the following steps:

1. The initial training phase:

   During this phase lasting thirty minutes, the experiment is first presented. Then, the subject uses the simulator, but only one of the tasks is active at a time. Then, a four-minute multitask scenario is launched.

2. The reference phase:

   This phase is decomposed into two parts: a first one lasting two minutes during which the subject is asked to relax and do nothing. Afterwards, a two-minute scenario "Low" is launched.

3. The experimental phase:

   This phase is divided into two periods, with a short break in-between. Following the procedure used in [38], each period is composed of a four-minute

scenario “Low”, a four-minute scenario “Medium”, and a six-minute scenario “High”. The order of the three in each block was set randomly for every subject, to avoid fatigue and order-effect bias in the results.

At the end of the experiment, subjects were asked to fill a NASA-TLX form [67] to subjectively evaluate their mental workload during the scenario. This form is used to ensure that the scenarios are well-designed in terms of ability to provoke different levels of mental workload.

Due to their complexity of implementation[9], the experiments were first carried out on thirteen subjects.

Each subject had to complete a specific scenario composed with periods of nominal activity (labelled *Nom*) and periods with high mental workload and engagement into the tasks, in order to undergo visual tunneling (labelled VT). A situation was labelled as visual tunneling if the following condition was met: no reaction to visual alarms, nor to secondary tasks, for a duration of more than twenty seconds. Thus, during the experiments, eight out of the thirteen subjects underwent visual tunneling. Two of them experienced this phenomenon twice. As a result, ten cases of visual tunneling were identified. During VT and *Nom* occurrences, the gaze position along both x and y-axis was collected from an eye-tracker. It was a VT2 model from the company EyeTech, with a sampling rate at 40Hz. Examples of gaze position processes are presented in Fig. 13.

4.4.2. Analysis of the gaze position processes and comments

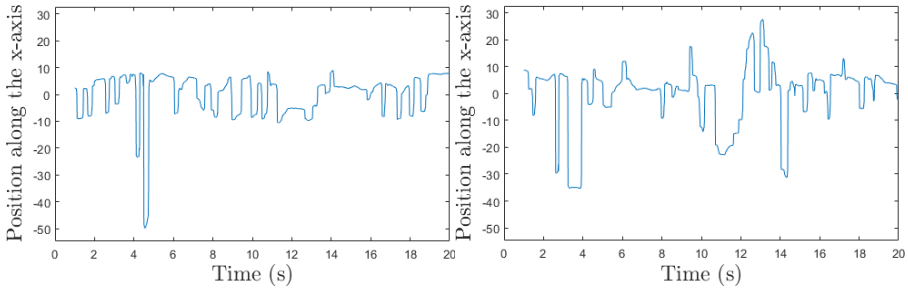
For each of the forty processes[10], the value of the slope α is computed. This value is estimated with the DFA, the DMA and the CDFA.

As presented in Fig. 13, the duration of an occurrence is set at twenty seconds. In the VT case, the gaze is mainly located towards the value 0 pixel, *i.e.* the center of the interface, whereas the subject's scanning is more spread in the *Nom*

[9] Piloting experience, neither glasses nor contact lenses allowed, etc.

[10] ten occurrences × two labels × two type of processes (gaze position along the x and y-axis)

case. In this particular example, the standard deviation is equal to 8.25 and 11.75 in the VT and Nom cases respectively. This feature, although relevant, is interface-dependant and thus cannot be used in another simulator. As the amplitude of the process does not contribute to the estimation of the slope α, the DFA or one of its variants can be considered.



(a) (b)

Figure 13: Time representation of the gaze position along the x-axis, (a) for the VT case, (b) for the Nom case.

Fig. 14 corresponds to the Wigner-Ville transform of the two above-represented processes. This illustrates the fact that the gaze position is a non-stationary process. In addition, most of the frequency content is located in the low frequencies.

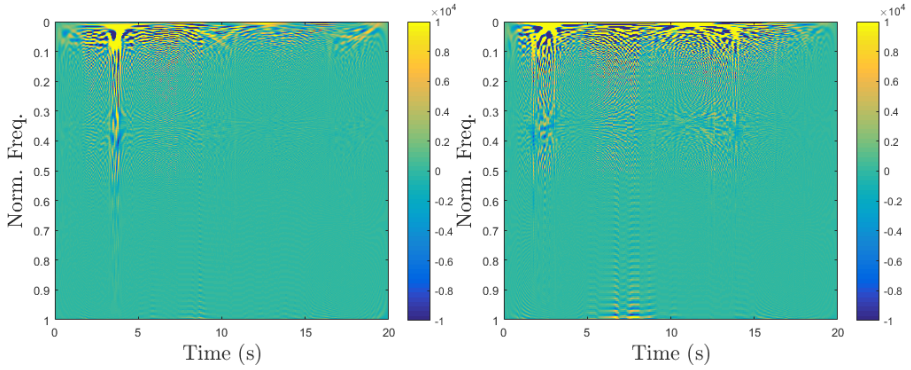


Figure 14: Wigner-Ville transform of the gaze position along the x-axis, (a) for the VT case, (b) for the *Nom* case.

As done in the previous section, $\log(F_{\boldsymbol{i}}(N))$ is plotted as a function of $\log(N)$ in Fig. 15. In this representation, the mean and standard deviation values over all processes are given. Only the DFA and the CDFA cases are pictured for the sake of clarity.

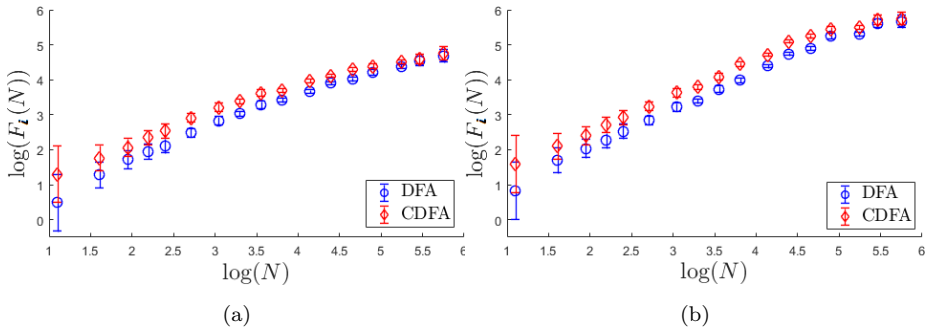


Figure 15: $\log(F_{\boldsymbol{i}}(N))$ vs. $\log(N)$, (a) for the VT case, (b) for the *Nom* case.

There is an increasing divergence between the mean values of $\log(F_{\mathbf{1}}(N))$ and $\log(F_{\mathbf{3}}(N))$ when $\log(N)$ becomes smaller. As a consequence, the values of α estimated on a set of small values of N is very dependant on the approach. This causes a deviation in the estimation of the slope that tends to reduce when

N increases. In the present application, due to reactivity constraints[11], the processes are of short length. Therefore, one could expect a difference in the estimation of the regularity of gaze positions.

When looking at the distributions obtained in the VT and Nom cases deduced from the set of 40 processes, with all approaches, in Fig. 16, it appears that:

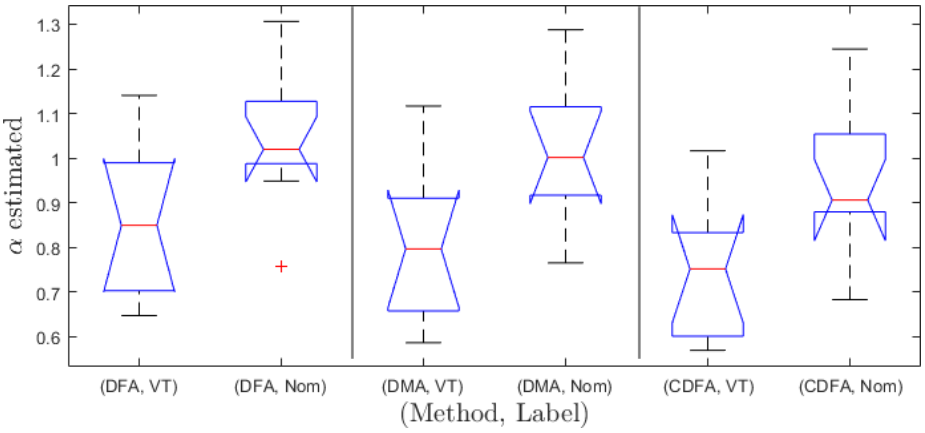


Figure 16: α of the gaze position estimated along the x-axis with both the DFA, the DMA and the CDFA.

1. The values estimated with the DFA are the largest for both the VT and Nom cases. This is due to the underestimation of $F_{\mathbf{1}}(N)$ that the DFA causes when using small values of N, as identified by Kantelhardt *et al.* in [49]. Indeed, the shorter the time scales, the larger the underestimation of the fluctuation function. As a consequence, when computing the slope α from a set of short time scales, the DFA overestimates α. However, the estimation of this quantity is not significantly different between the three approaches for both labels. In order to analyze this difference, an ANOVA test was performed[12] [68]. The p-values obtained were equal to $p = 0.359$

[11] In critical situations such as go-around, etc., the pilot has to make a decision in a very short time. Therefore, the system must be able to quickly detect any possible deleterious cognitive state from the pilot.

[12] An ANOVA test aims at determining whether data from several groups of a feature are

and $p = 0.439$ in the VT and the Nom case respectively.

2. No matter the estimation method used, there is a significant difference between the values obtained in the VT and the Nom cases. An ANOVA test conducted on these values provided the following p-values: $p = 0.0259$, $p = 0.0168$ and $p = 0.0125$ in the DFA, the DMA and the CDFA case respectively. As a consequence, there is a correlation between the visual tunneling of a subject and the regularity of his gaze position along the x-axis. The CDFA provides the most significant separation, as the p-value is the smallest.

5. Conclusions and perspectives

In this paper, the DFA and some of its variants, namely the DMA and the CDFA, which is the proposed method where the trend is constrained to be continuous, are compared. This comparison is based on a uniform way of expressing the square of the fluctuation function from the instantaneous correlation function of the process. Firstly, when the process under study is wide-sense stationary, the statistical mean of the square of the fluctuation function can be expressed as a weighted sum of the autocorrelation function of the signal under study, without any approximation. Secondly, if the process is non-stationary, the square of the fluctuation function can be expressed from the Wigner-Ville transform of the process. To summarize, thanks to our analysis, it is possible to separate the contribution of the process from the influence of the method in the computation of the fluctuation function. As a consequence, our framework makes it possible to better understand the different behaviours between the DFA and its variants, thanks to both the filter-based analysis and the 2D-FT-based analysis that we respectively propose in the stationary and non-stationary cases. Our way of expressing the methods can bring a more in-depth comparison, since it is associated with an intermediate step in the computation of α.

significantly different from each other.

Furthermore, we address the comparison between our analysis where no approximation is made and previous studies carried out on the DFA where an expression of the square of the fluctuation function was obtained using some approximations.

The different way the methods act on a given process leads to a difference in the evolution of the fluctuation function with regards to the value of N. To illustrate this phenomenon, simulations were carried out on both stationary and non-stationary processes. The stationary processes were built with a prescribed value of regularity. The non-stationary processes were ocular processes and their regularity was estimated to characterize a given human cognitive state: the visual tunneling of a airline pilot. The results showed that the values obtained with the three approaches were not significantly different, meaning that they are indeed an estimator of the same quantity. In addition, from our simulations, there seems to be a correlation between the regularity of the gaze position of an airline pilot and the visual tunneling. The CDFA appears to be the most relevant approach for this application, as the slope estimated with it is the most discriminatory factor. Among our perspectives, we plan to propose new variants of the DFA and analyze how to estimate the regularity by using particle-filtering based approaches.

6. Acknowledgements

We would like to thank Marc Donias and Thierry Ferreira for the fruitful discussions we had. These works were done within the scientific network GIS Albatros between Thales AVS France, Thales DMS France, University of Bordeaux, Bordeaux INP, University of Poitiers, University of Limoges, CNRS and INRIA.

References

[1] J. Gao, J. Hu, F. Liu, Y. Cao, Multiscale entropy analysis of biological signals: A fundamental bi-scaling law, Frontiers in Computational Neuro-

science 9 (2015) 1–64.

[2] C. L. Chang, M. McAleer, Econometric analysis of financial derivatives: an overview, Journal of Econometrics 187, (2) (2015) 403–407.

[3] C. W. J. Granger, R. Joyeux, An introduction to long-memory time series models and frafgaoctional differencing, Journal of Times Series Analysis 1, (1) (1980) 15 – 29.

[4] J. R. M. Hosking, Fractional differencing, Biometrika 68, (1) (1981) 165–176.

[5] B. B. Mandelbrot, J. W. Van Ness, Fractional brownian motions, fractional noises and applications, SIAM Review 10(4) (1968) 422–437.

[6] V. Pipiras, M. S. Taqqu, Long-range dependence and self-similarity, Cambridge University Press, 2017.

[7] H. E. Hurst, Long-term storage capacity of reservoirs, Transactions of the American Society of Civil Engineers 116 (1951) 770–799.

[8] M. S. Taqqu, V. Teverovsky, W. Willinger, Estimators for long range dependence: an empirical study, Fractals 3, (4) (1995) 785–788.

[9] M. S. Taqqu, V. Teverovsky, On Estimating the Intensity of Long Range Dependence in Finite and Infinite Variance Time Series. A practical guide to heavy tails: statistical techniques and applications, R. Adler, R. Feldman, M. S. Taqqu, Editors. Birkäuser, 1998.

[10] C. K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons, H. E. Stanley, Long-range correlations in nucleotide sequences, Nature 356 (1992) 168–170.

[11] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, A. L. Goldberger, Mosaic organization of DNA nucleotides, Physical Review E 49, (2) (1994) 1685–1689.

[12] M. Najim, Modeling, estimation and optimal filtering in signal processing, Wiley, 2010.

[13] S.-J. Kim, K. Koh, S. Boyd, D. Gorinevsky, l_1 trend filtering, SIAM Review 51, (2) (2009) 339–360.

[14] L. Berthouze, S. Farmer, Adaptive time-varying detrended fluctuations analysis: a new method for characterizing time-varying scaling parameters in physiological time series, BMC Neuroscience 12 (2011) 1–2.

[15] M. A. Riley, S. Bonnette, N. Kuznetsov, S. Wallot, J. Gao, A tutorial introduction to adaptive fractal analysis, Frontiers in Physiology 3 (2012) 371.

[16] E. Alessio, A. Carbone, G. Castelli, V. Frappietro, Second-order moving average and scaling of stochastic time series, The European Physical Journal B 27, 2 (2002) 197–200.

[17] D. Osborn, Moving average detrending and the analysis of business cycles, Oxford Bull. Econom. Statist. 57 (1995) 547–558.

[18] L. Xu, P. C. Ivanov, K. Hu, Z. Chen, A. Carbone, H. E. Stanley1, Quantifying signals with power-law correlations: A comparative study of detrended fluctuation analysis and detrended moving average techniques, Physical Review E 71, 5 (2005) 051101.

[19] A. Bashan, R. Bartsch, J. W. Kantelhardt, S. Havlin, Comparison of detrending methods for fluctuation analysis, Physica A: Statistical Mechanics and its Applications 387 (2008) 5080–5090.

[20] Y.-H. Shao, G.-F. Gu, Z.-Q. Jiang, W.-X. Zhou, D. Sornette, Comparing the performance of FA, DFA and DMA using different synthetic long-range correlated time series, Scientific Reports 2 (2012) 835.

[21] R. Sun, Fractional order signal processing: techniques and applications, Thesis of Master of science in electrical Engineering, Utah State University (2007).

[22] G. Rilling, P. Flandrin, P. Gonçalves, Empirical mode decomposition, fractional gaussian noise and hurst exponent estimation, ICASSP (2005) 489–492.

[23] P. Abry, P. Flandrin, M. S. Taqqu, D. Veitch, Self-similarity and long-range dependence through the wavelet lens, Theory and Applications of Long-range Dependence (2003) 527–556.

[24] J.-M. Bardet, Testing for the presence of self-similarity of gaussian time series having stationary increments, Journal of Time Series Analysis 21 (2000) 497–515.

[25] E. Moulines, F. Roueff, M. S. Taqqu, Central limit theorem for the log-regression wavelet estimation of the memory parameter in the gaussian semi-parametric context, Fractals 15 (2007) 301–313.

[26] F. Esposti, M. G. Signorini, Evaluation of a blind method for the estimation of Hurst's exponent in time series, EUSIPCO (2006) 1–5.

[27] H. Salat, R. Murcio, E. Arcaute, Multifractal methodology, Physica A: Statistical Mechanics and its Applications 473 (2017) 467–487.

[28] E. Cirugedan–Roldan, A. Molina-Picó, D. Cuesta-Frau, S. Oltra–Crespo, P. Miró–Martınez, Comparative study between sample entropy and detrended fluctuation analysis performance on EEG records under data loss, EMBS (2012) 4233–4236.

[29] X. Navarro, A. Beuchée, F. Porée, G. Carrault, Performance analysis of Hurst's exponent estimators in higly immature breathing patterns of preterm infants, ICASSP (2011) 701–704.

[30] B. Audit, E. Bacry, J. . Muzy, A. Arneodo, Wavelet-based estimators of scaling behavior, IEEE Transactions on Information Theory 48 (2002) 2938–2954.

[31] S. Sanyal, A. Banerjee, R. Pratihar, A. K. Maity, S. Dey, V. Agrawal, R. Sengupta, D. Ghosh, Detrended fluctuation and power spectral analysis of alpha and delta EEG brain rhythms to study music elicited emotion, International Conference on Signal Processing, Computing and Control (ISPCC) (2015) 206–210.

[32] A. A. Pranata, G. W. Adhane, D. S. Kim, Detrended fluctuation analysis on ECG device for home environment, Consumer Communications and Networking Conference (CCNC) (2017) 4233–4236.

[33] A. G. Ravelo-Garcia, U. Casanova-Blancas, S. Martin-González, E. Hernández-Pérez, I. Guerra-Moreno, P. Quintana-Morales, N. Wessel, J. L. Navarro-Mesa, An approach to the enhancement of sleep apnea detection by means of detrended fluctuation analysis of RR intervals, Computing in Cardiology (2014) 905–908.

[34] R. Acharya, C. Lim, P. Joseph, Heart rate variability analysis using correlation dimension and detrended fluctuation analysis, ITBM-RBM 23 (2002) 333–339.

[35] R. Kumagai, M. Uchida, Detrended fluctuation analysis of repetitive handwriting, in: 2017 International Conference on Noise and Fluctuations (ICNF), 2017, pp. 1–4.

[36] W. Mumtaz, A. Malik, S. S. Ali, m. a. mohd yasin, H. U. Amin, Detrended fluctuation analysis for major depressive disorder, in: EMBC, 2015, pp. 4162–4165.

[37] A. Adda, H. Benoudnine, Detrended fluctuation analysis of EEG recordings for epileptic seizure detection, in: 2016 International Conference on Bioengineering for Smart Technologies (BioSMART), 2016, pp. 1–4.

[38] A. Kitlas Golińska, Detrended fluctuation analysis (DFA) in biomedical signal processing: Selected examples, Studies in Logic, Grammar and Rhetoric 29 (2012) 107–115.

[39] A. Tiwari, C. Albulescu, S.-M. Yoon, A multifractal detrended fluctuation analysis of financial market efficiency: Comparison using dow jones sector ETF indices, Physica A: Statistical Mechanics and its Applications 483 (2017) 182–192.

[40] A. Serletis, O. Y. Uritskaya, V. M. Uritsky, Detrended fluctuation analysis of the us stock market, International Journal of Bifurcation and Chaos 18 (2008) 599–603.

[41] P. Talkner, R. Weber, Power spectrum and detrended fluctuation analysis: Application to daily temperatures, Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics 62 (2000) 150–60.

[42] A. Király, I. Jánosi, Detrended fluctuation analysis of daily temperature records: Geographic dependence over australia, Meteorology and Atmospheric Physics 88 (2005) 119–128.

[43] W.-w. Tung, J. Gao, J. Hu, L. Yang, Detecting chaos in heavy-noise environments, Physical review. E 83 (2011) 046210.

[44] S. Arianos, A. Carbone, C. Türk, Self-similarity of higher-order moving averages, Physical Review E 84(4 Pt 2) (2011) 046113.

[45] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, A. Bunde, S. Havlin, H. E. Stanley, Multifractal detrended fluctuation analysis of nonstationary time series, Physica A: Statistical Mechanics and its Applications 316, (1-4) (2002) 87–114.

[46] Y. Tsujimoto, Y. Miki, S. Shimatani, K. Kiyono, Fast algorithm for scaling analysis with higher-order detrending moving average method, Physical Review E 93 (5) (2016) 053304.

[47] Y. Tsujimoto, Y. Miki, E. Watanabe, J. Hayano, Y. Yamamoto, T. Nomura, K. Kiyono, Fast algorithm of long-range cross-correlation analysis

using savitzky-golay detrending filter and its application to biosignal analysis, International Conference on Noise and Fluctuations (ICNF) (2017) 1–4.

[48] K. Kiyono, Establishing a direct connection between detrended fluctuation analysis and fourier analysis, Physical Review E 92 (2015) 042925.

[49] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, A. Bunde, Detecting long-range correlations with detrended fluctuation analysis, Physica A: Statistical Mechanics and its Applications 295, (3-4) (2001) 441–454.

[50] M. Höll, H. Kantz, Y. Zhou, Detrended fluctuation analysis and the difference between external drifts and intrinsic diffusionlike nonstationarity, Physical Review E 94 (2016) 042201.

[51] K. Kiyono, Theory and applications of detrending -operation -based fractal-scaling analysis, International Conference on Noise and Fluctuations (ICNF) (2017) 1–4.

[52] M. Höll, H. Kantz, The relationship between the detrendend fluctuation analysis and the autocorrelation function of a signal, The European Physical Journal B 88 (2015) 327.

[53] P. Carpena, M. Gomez-Extremera, C. Carretero-Campos, P. Bernaola-Galvan, A. V. Coronado, Spurious results of fluctuation analysis techniques in magnitude and sign correlations, Entropy 19 (2015) 2–61.

[54] E. P. Wigner, On the quantum correction for thermodynamic equilibrium, Physical Review 40 (1932) 749–759.

[55] L. R. Rabiner, B.-H. Juang, Fundamentals of speech recognition, PTR Prentice Hall, 1993.

[56] G. Hardy, On weierstrass' non-differentiable functions, Trans. of the American Mathematical Society 17, (3) (1916) 301–325.

[57] J. Levy-Vehel, P. Legrand, Signal and image processing with FRACLAB, FRACTAL04, Complexity and Fractals in Nature, International Multidisciplinary Conference (2004) 321–322.

[58] C. F. Simons, D. J.and Chabris, Gorillas in our midst: sustained inattentional blindness for dynamic events, Perception (1999) 1059–1074.

[59] S. Shappell, D. Wiegman, A human error analysis of general aviation controlled flight into terrain accidents occurring between 1990-1998, U.S. Federal Aviation Administration, Office of Aerospace Medicine (2003) 25.

[60] N. Regis, F. Dehais, E. Rachelson, C. Thooris, S. Pizziol, M. Causse, C. Tessier, Formal detection of attentional tunneling in human operator–automation interactions, IEEE Transactions on Human-Machine Systems 44 (2014) 326–336.

[61] G. Pourtois, M. De Pretto, C.-A. Hauert, P. Vuilleumier, Time course of brain activity during change blindness and change awareness: Performance is predicted by neural events before change onset, Journal of cognitive neuroscience 18 (2007) 2108–29.

[62] K. Mathewson, G. Gratton, M. Fabiani, D. Beck, T. Ro, To see or not to see: Prestimulus phase predicts visual awareness, The Journal of neuroscience 29 (2009) 2725–32.

[63] Y. Santiago-Espada, R. R. Myer, K. A. Latorella, J. R. Comstock, The multi-attribute task battery ii (matb-ii) software for human performance and workload research: A user's guide, Hampton, Virginia : National Aeronautics and Space Administration, Langley Research Center, 2011.

[64] G. Wilson, C. Russell, Real-time assessment of mental workload using psychophysiological measures and artificial neural networks, Human factors 45 (2003) 635–43.

[65] B. Cinaz, B. Arnrich, R. Marca, G. Tröster, Monitoring of mental workload levels during an everyday life office-work scenario, Personal and Ubiquitous Computing 17 (2013) 229–239.

[66] M. Fallahi, R. Heidarimoghadam, M. Motamedzade, M. Farhadian, Psycho physiological and subjective responses to mental workload levels during n-back task, Journal of Ergonomics 06 (2016).

[67] S. Hart, L. Staveland, Development of Nasa-Tlx (task load index): Results of empirical and theoretical research, Advances in Psychology 52 (1988) 139–183.

[68] R. V. Hogg, J. Ledolter, Engineering statistics, New York : Macmillan ; London : Collier Macmillan, 1987.