2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Multifractal and Multiscale Detrended Fluctuation Analysis of Cardiovascular Signals: how the Estimation Bias Affects Short-Term Coefficients and a Way to mitigate this Error

Paolo Castiglioni, Gianfranco Parati, and Andrea Faini

***Abstract*—The Detrended Fluctuation Analysis (DFA) is a popular method for quantifying the self-similarity of the heart rate that may reveal complexity aspects in cardiovascular regulation. However, the self-similarity coefficients provided by DFA may be affected by an overestimation error associated with the shortest scales. Recently, the DFA has been extended to calculate the multifractal-multiscale self-similarity and some evidence suggests that overestimation errors may affect different multifractal orders. If this is the case, the error might alter substantially the multifractal-multiscale representation of the cardiovascular self-similarity. The aim of this work is 1) to describe how this error depends on the multifractal orders and scales and 2) to propose a way to mitigate this error applicable to real cardiovascular series.**

***Clinical Relevance*— The proposed correction method may extend the multifractal analysis at the shortest scales, thus allowing to better assess complexity alterations in the cardiac autonomic regulation and to increase the clinical value of DFA.**

I. Introduction

The interest in the fractal analysis of the heart rate has been progressively growing aimed at describing the complex cardiovascular regulation, its adaptations to external stimuli and behavioral conditions, and its interactions with the autonomic control and the respiratory system. In this context, the Detrended Fluctuation Analysis (DFA) is a powerful tool for quantifying the fractal structure of cardiovascular series [1] based on the estimation of a coefficient, α, related to the Hurst's exponent. The DFA calculates the 2nd order moment of the fluctuations of the integrated series after polynomial detrending on consecutive blocks of n beats, $F(n)$. For fractal series, α is estimated as the slope of log $F(n)$ vs. log n [2].

Since the heart rate shows a multiscale behavior, it has been proposed to estimate α as a function of n, $\alpha(n)$, by calculating "local" slopes around n [3], [4]. However, the DFA may overestimate the local slopes at the shorter scales, likely because of the overfitting of the detrending polynomial (the overestimation increases with the polynomial order) [5]. A proposed correction method assumes the same α at all the scales [5] and thus cannot be applied on real cardiovascular series because of their multiscale nature. Alternatively, it was suggested to use high-order polynomials at the larger scales, where the overfitting is negligible, and the 1st order polynomial, which causes the lowest overfitting, at the shortest scales [6]. However, also the 1st order detrending polynomial may cause overestimations that limit the shortest analyzable n.

Recently, the multiscale DFA was modified to evaluate the multifractality [7] calculating the q^{th} moment order of the residual variances, $F_q(n)$, and the multifractal multiscale self-similarity coefficients, $\alpha(q,n)$, as the local slopes of log $F_q(n)$ vs. log n. It is possible that the estimation bias affecting the monofractal α also influences $\alpha(q,n)$ at $q\neq2$, but whether the error depends on both q and n and if it alters the quantification of multifractality has never been described systematically. Therefore, this work aims to describe the estimation bias in the multifractal multiscale DFA and to propose a method applicable to real cardiovascular series for mitigating this error.

II. DFA Estimation Bias

A. The Multifractal Multiscale DFA

The DFA evaluates the multifractality of a time series S_i of N samples (i=1,...,N), with mean μ and standard deviation σ, by calculating the summation y_i

$$y_i = \sum_{j=1}^{i} \frac{S_j - \mu}{\sigma} \quad (1),$$

by splitting the y_i series into M blocks of n samples, and by evaluating the variability function $F_q(n)$

$$\begin{cases} F_q(n) = \left(\frac{1}{M}\sum_{k=1}^{M}\left(\sigma^2{}_n(k)\right)^{q/2}\right)^{1/q} & for\ q \neq 0 \\ F_q(n) = e^{\frac{1}{2M}\sum_{k=1}^{M}\ln\left(\sigma^2{}_n(k)\right)} & for\ q = 0 \end{cases} \quad (2)$$

with $\sigma^2{}_n(k)$ the variance of the residuals of y_i in each block k after polynomial detrending [8]. In our work, n increased linearly on a log scale from 6 up to N/4 samples. Moreover, we considered integer q between -5 and +5; maximally overlapped blocks, which means that M=N–n+1; and linear detrending. The local slopes $\alpha(q,n)$ were calculated as the derivative of log $F_q(n)$ vs. log n [6].

The cumulative functions of the $\alpha(q,n)$ squared increments were calculated separately for positive q

$$\alpha^{+}_{CF}(n) = \sum_{q=1}^{Q}[\alpha(q,n) - \alpha(q-1,n)]^2 \quad (3)$$

and negative q

$$\alpha^{-}_{CF}(n) = \sum_{q=-Q+1}^{0}[\alpha(q,n) - \alpha(q-1,n)]^2 \quad (4).$$

The $\alpha^{+}_{CF}(n)$ and $\alpha^{-}_{CF}(n)$ functions represent scale-by-scale measures of the degree of multifractality [9].

P. Castiglioni is with IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy (email: pcastiglioni@dongnocchi.it). G. Parati and A. Faini are with Istituto Auxologico Italiano, IRCCS, Dept. of Cardiovascular, Neural and Metabolic Sciences Milan, Italy (email: a.faini@auxologico.it).

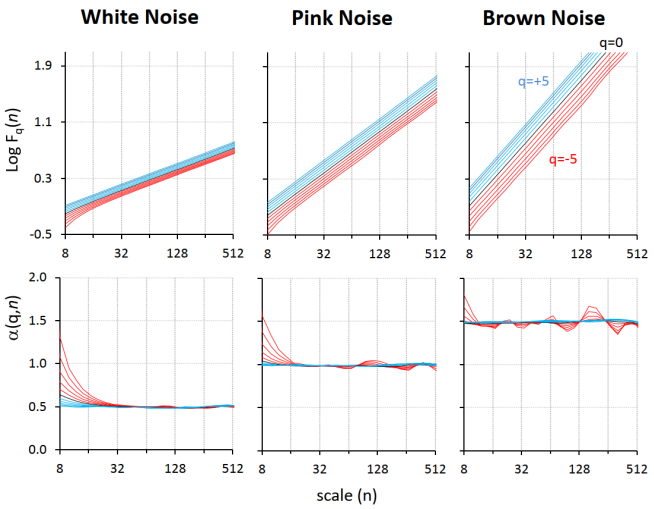

Figure 1. Upper panels: DFA variability functions $F_q(n)$ for white, pink, and brown noise; lower panels: corresponding multifractal multiscale coefficients $\alpha(q,n)$. Average over 10 realizations for $q>0$ in blue, $q<0$ in red, $q=0$ in black.

B. *Synthetic Signals*

To quantify the estimation errors, we synthesized time series from 3 monofractal-monoscale processes with a known self-similarity coefficient: white Gaussian noise (α=0.5), pink noise (α=1.0), and Brown noise (α=1.5). For each process, we synthesized 10 series of 2^{14} samples in Matlab R2020a. The white noise was generated by the *randn* function; the Brown noise by the *cumsum* function applied to the white noise; the pink noise by the *pinknoise* function [10].

Figure 1 shows the average of the 10 $F_q(n)$ functions and 10 derived $\alpha(q,n)$ coefficients for each noise. Underestimations of $F_q(n)$ at the smaller n, particularly for $q<0$, produce $\alpha(q,n)$ overestimations at the shorter scales. The overestimation is more pronounced for $q<0$ and thus it wrongly suggests multifractality at the shortest scales. Actually, the α_{CF}^{-} index (figure 2, upper panels) indicates a pronounced multifractal component at the shorter scales for the negative moment orders.

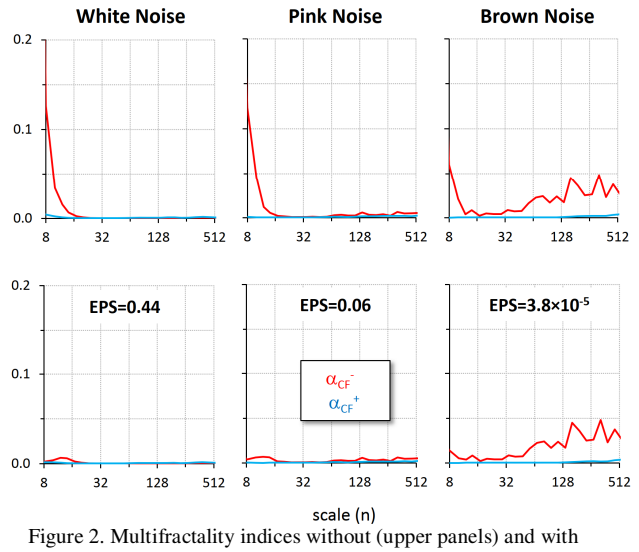

Figure 2. Multifractality indices without (upper panels) and with (lower panels) removal of overfitted blocks by the EPS threshold: average over 10 realizations of white, pink and brown noise.

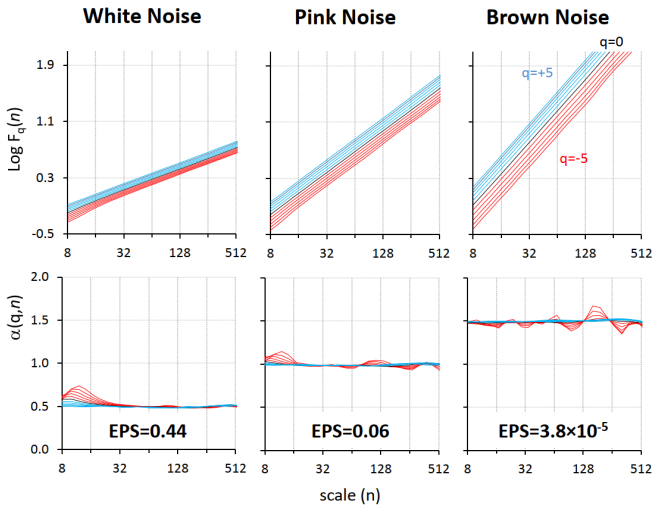

Figure 3. Variability functions $F_q(n)$ and the derived $\alpha(q,n)$ coefficients for the noises of figure 1 after the proposed correction; EPS is the identified threshold for removing overfitted blocks.

III. Bias Reduction by Removal of Overfitted Blocks

An explanation for the false multifractality at the shortest scales highlighted by figures 1 and 2 is the presence of blocks overfitted by the detrending polynomial. It is possible that the samples y_i in one or some of the M blocks of eq.2 align almost perfectly over a straight line just for chance, causing σ^2_n values close to 0, an event more likely for small n. The effect of these overfitted blocks is to underestimate $F_q(n)$ and since $q<0$ amplifies the smaller components, the effect is greater for negative than positive q. Therefore, the presence of overfitted blocks would explain the $F_q(n)$ underestimation at the smaller n with deviations from the straight line more pronounced for $q<0$. For this reason, our approach is to improve the $F_q(n)$ estimates by removing too low σ^2_n values.

To remove the overfitted blocks we were inspired by a procedure previously proposed to deal with the quantization errors of the recording device. For this scope, it was suggested to ignore residual variances lower than a given threshold EPS defined by the instrument precision [11]. Similarly, we suggest running the sums in eq. (2) only on the blocks with σ^2_n greater than a threshold EPS. The critical point in our application is to find the correct EPS value. Since the errors we want to minimize affect more the shortest scales and the negative q, we applied an iterative procedure based on the Golden Section search method to find the EPS value that minimizes $\alpha_{CF}^{-}(n)$ at the shortest scale, i.e. n=8. We also impose that $\alpha(q,8)\geq\alpha(0,8)$ for $q<0$, because we expect that the overestimation error is greater for $q<0$ than $q=0$.

Figure 3 shows the effects of the removal of the outliers on the same series of figure 1. The method automatically adapts the EPS threshold according to the self-similar characteristics of the random process and appears able to remove almost completely the $\alpha(q,n)$ overestimation at the shortest scales without affecting the estimates at larger scales. The lower panels of Figure 2 indicate clearly the positive effect of the removal showing that the artefactual multifractality affecting the shorter scales is largely reduced.

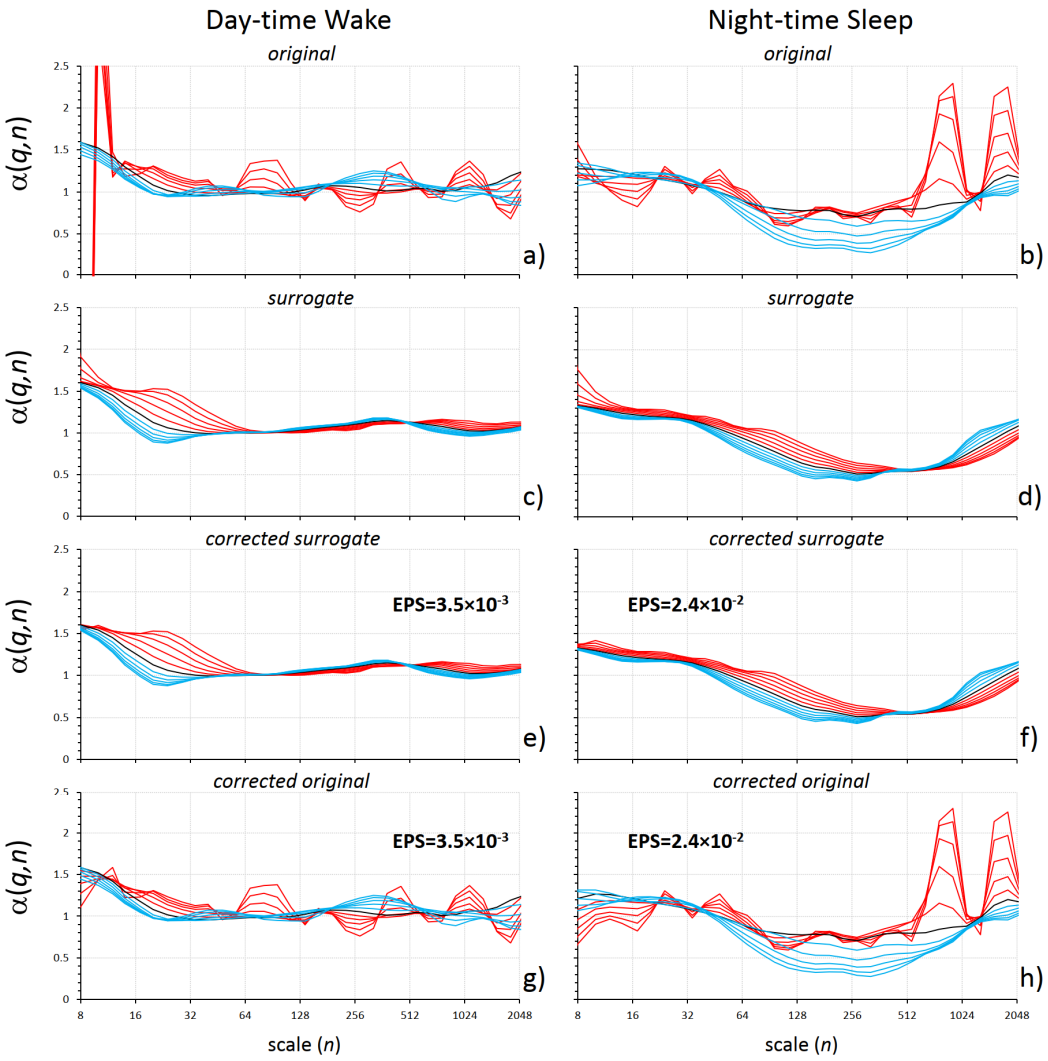

Figure 4. Multiscale multifractal DFA coefficients, $\alpha(q,n)$, of R-R intervals recorded in a volunteer during 4 hours of daytime wake (left) and nighttime sleep (right) vs. the scale n, in number of heartbeats, separately for moment orders q between -5 and +5. From top to bottom: $\alpha(q,n)$ of original (panels a and b) and phase-shuffled surrogate series (panels c and d); $\alpha(q,n)$ of the surrogate series after identification of the EPS threshold and overfitted blocks removal (panels e and f); $\alpha(q,n)$ of the original series after removal of the overfitted blocks by the EPS thresholds identified from the surrogate data (panels g and h). For each segment of the original data, we generated 100 phase-shuffled surrogate series and show the average $\alpha(q,n)$ of the 100 estimates, with $q>0$ in blue, $q=0$ in black, $q<0$ in red color.

IV. Application on Real Heart Rate Series

While a previously proposed method for correcting the DFA overestimation of the monofractal-monoscale α at the shortest scales cannot be applied to real multiscale heart rate recordings [5], our procedure does not require monoscale fractal series and thus can be also applied on self-similarity coefficients that change with n, as those of the heart rate. However, our EPS threshold optimization is done under the hypothesis of monofractality, which may not hold for real cardiovascular signals. Thus, a further step is required to apply our correction method to real signals. The step consists of the EPS identification not on the original series but on the surrogate ones obtained after Fourier phase shuffling [12], which preserves the 2nd order statistics removing the multifractal components.

Figure 4 illustrates our method on real heart rate series. The example considers the same subperiods of 4-hour duration from a 24-hour recording of R-R intervals in a healthy volunteer used in a previous study: readers can find details on the data collection in [6]. The Ethical Review Board of Istituto Auxologico Italiano, IRCCS (Milan, Italy) approved these experimental procedures involving human subjects. The first subperiod was extracted during daytime activities (Wake), the second one during nighttime rest (Sleep). The panels a) and b) show the $\alpha(q,n)$ coefficients in these subperiods. Then we generated 100 phase-shuffled surrogate series with the code provided in [13] and calculated their multifractal-multiscale coefficients: panels c) and d) show the average $\alpha(q,n)$ over the 100 surrogates. The comparison between $\alpha(q,n)$ of the original and surrogate series points out the clear presence of multifractality at night, in particular at scales $n>256$ beats. By contrast, it is difficult to derive any conclusion at the shorter scales. During Wake, the original coefficients are extremely dispersed for $q<0$ but this seems due to the instability of the estimate rather than to multifractality. During Sleep, the original and surrogate coefficients show a similar dispersion, a result that might

suggest the absence of multifractality at the sorter scales; however, the $\alpha(q,n)$ coefficients of both the original and surrogate series show a pattern that is compatible with the reported overestimation error, more pronounced for $q<0$, resulting from the detrending overfitting, and that might hide a multifractal dynamics at the shorter scales present in the original series.

We removed the overfitted blocks identifying the EPS threshold on the surrogate series: the panels e) and f) show their corrected $\alpha(q,n)$ coefficients. Finally, we used these EPS thresholds to remove too low residual variances when estimating the $\alpha(q,n)$ coefficients of the original series: see the results in panels g) and h). In this way, the comparison between the original and surrogate series after the removal of the overfitted blocks provides a clearer picture of the multifractal components at the shorter scales. In Wake, the critical instability of the estimates disappears and the coefficients converge at around $\alpha=1.5$, with a similar dispersion among moment orders for the original and surrogate data. This suggests that during daytime multifractal components, if present, are negligible at the shorter scales. In Sleep, the removal of overfitted blocks eliminated the overestimation affecting the self-similarity coefficients with negative q. After the correction, the dispersion of the coefficients at scales between 8 and 16 beats seems substantially larger for the original than the surrogate series, which is a potential marker of multifractality during nighttime also at the shorter scales.

V. Conclusion

Analyzing synthesized series with known fractal properties we showed that the recently proposed multifractal-multiscale DFA can be affected by an estimation error and that this error might wrongly suggest the presence of multifractality at the shortest scales or hide a true multifractal component. Hypothesizing that the error is due to the overfitting of the detrending polynomials, we proposed a correction method based on the empirical identification of overfitted blocks for monofractal processes. The advantages of our method are that 1) it does not require that the recorded series is characterized by a monoscale self-similarity, and 2) the threshold for the outlier removal can be identified on the phase-shuffled surrogate series if the recorded series is multifractal. These properties make the proposed method versatile enough to be applied to real cardiovascular recordings. Removing overestimation errors from the surrogate series, the comparison between original and surrogate $\alpha(q,n)$ allows detecting true multifractal components also at the shorter scales.

Therefore, the proposed correction method may allow extending the multifractal-multiscale analysis at the shorter scales where strong interactions among the respiratory and the cardiovascular systems, as well as the influence of the sympathovagal autonomic regulation, are expected. This potentially improves the value of the DFA method in clinical settings and physiological studies.

References

[1] P. Castiglioni, M. Di Rienzo, and A. Faini, 'Self-Similarity and Detrended Fluctuation Analysis of Cardiovascular Signals', in *Complexity and Nonlinearity in Cardiovascular Signals*, R. Barbieri, E. P. Scilingo, and G. Valenza, Eds. Cham: Springer International Publishing, 2017, pp. 197–232. doi: 10.1007/978-3-319-58709-7_7.

[2] C. K. Peng, S. Havlin, J. M. Hausdorff, J. E. Mietus, H. E. Stanley, and A. L. Goldberger, 'Fractal mechanisms and heart rate dynamics. Long-range correlations and their breakdown with disease', *J Electrocardiol.*, vol. 28 Suppl, no. 0022-0736 (Print), pp. 59–65, 1995.

[3] J. C. Echeverria, M. S. Woolfson, J. A. Crowe, B. R. Hayes-Gill, G. D. Croaker, and H. Vyas, 'Interpretation of heart rate variability via detrended fluctuation analysis and alphabeta filter', *Chaos.*, vol. 13, no. 1054-1500 (Print), pp. 467–475, Jun. 2003.

[4] P. Castiglioni, G. Parati, A. Civijian, L. Quintin, and M. Di Rienzo, 'Local scale exponents of blood pressure and heart rate variability by detrended fluctuation analysis: effects of posture, exercise, and aging', *IEEE Trans Biomed Eng*, vol. 56, no. 1558-2531 (Electronic), pp. 675–684, Mar. 2009.

[5] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, and A. Bunde, 'Detecting long-range correlations with detrended fluctuation analysis', *Physica A: Statistical Mechanics and its Applications*, vol. 295, no. 3–4, pp. 441–454, Jun. 2001.

[6] P. Castiglioni and A. Faini, 'A Fast DFA Algorithm for Multifractal Multiscale Analysis of Physiological Time Series', *Frontiers in Physiology*, vol. 10, p. 115, Mar. 2019, doi: 10.3389/fphys.2019.00115.

[7] J. Gieraltowski, J. J. Zebrowski, and R. Baranowski, 'Multiscale multifractal analysis of heart rate variability recordings with a large number of occurrences of arrhythmia', *Phys.Rev.E Stat.Nonlin.Soft.Matter Phys.*, vol. 85, no. 1550-2376 (Electronic), p. 021915, Feb. 2012.

[8] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley, 'Multifractal detrended fluctuation analysis of nonstationary time series', *Physica A: Statistical Mechanics and its Applications*, vol. 316, no. 1–4, pp. 87–114, Dec. 2002.

[9] A. Faini, G. Parati, and P. Castiglioni, 'Multiscale Assessment of the Degree of Multifractality for Physiological Time Series', *Philosophical Transactions A*, p. in press, 2021.

[10] H. Zhivomirov, 'Pink, Red, Blue and Violet Noise Generation with Matlab', *MATLAB Central File Exchange*, vol. https://www.mathworks.com/matlabcentral/fileexchange/42919-pink-red-blue-and-violet-noise-generation-with-matlab, 2021.

[11] E. A. F. Ihlen, 'Introduction to multifractal wavelet and detrended fluctuation analyses', *Frontiers in physiology*, vol. 3, 2012, doi: 10.3389/fphys.2012.00141.

[12] T. Schreiber and A. Schmitz, 'Surrogate time series', *Physica D: Nonlinear Phenomena*, vol. 142, no. 3–4, pp. 346–382, Aug. 2000, doi: 10.1016/S0167-2789(00)00043-9.

[13] T. Gautama, *Surrogate Data*. MATLAB, 2005. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/4612-surrogate-data