Comparing the performance of FA, DFA and DMA using different synthetic long-range correlated time series

Ying-Hui Shao,[1, 2] Gao-Feng Gu,[1, 2] Zhi-Qiang Jiang,[1, 2] Wei-Xing Zhou,[1, 2, 3, *] and Didier Sornette[4, 5, †]

[1] *School of Business, East China University of Science and Technology, Shanghai 200237, China*
[2] *Research Center for Econophysics, East China University of Science and Technology, Shanghai 200237, China*
[3] *School of Science, East China University of Science and Technology, Shanghai 200237, China*
[4] *Department of Management, Technology and Economics, ETH Zurich, Zurich, Switzerland*
[5] *Swiss Finance Institute, c/o University of Geneva, Geneva, Switzerland*
(Dated: August 22, 2012)

Notwithstanding the significant efforts to develop estimators of long-range correlations (LRC) and to compare their performance, no clear consensus exists on what is the best method and under which conditions. In addition, synthetic tests suggest that the performance of LRC estimators varies when using different generators of LRC time series. Here, we compare the performances of four estimators [Fluctuation Analysis (FA), Detrended Fluctuation Analysis (DFA), Backward Detrending Moving Average (BDMA), and centred Detrending Moving Average (CDMA)]. We use three different generators [Fractional Gaussian Noises, and two ways of generating Fractional Brownian Motions]. We find that CDMA has the best performance and DFA is only slightly worse in some situations, while FA performs the worst. In addition, CDMA and DFA are less sensitive to the scaling range than FA. Hence, CDMA and DFA remain "The Methods of Choice" in determining the Hurst index of time series.

A complex system, be it ecological, biological, technological, social, economic or financial, is usually embedded in a complex network, which is composed of a large number of interacting heterogeneous constituents linked via interwoven nonlinear heterogenous ties [1]. The observed signals of the physical quantities characterizing a complex system often exhibit long-range correlations [2]. It is of crucial importance and significance to quantify such long-range correlations to have a deep understanding of the dynamics of the underlying complex systems. More than ten techniques have been invented to detect long-range correlations in time series [3, 4], such as the rescaled range (R/S) analysis [5], the wavelet transform module maxima (WTMM) approach [6–10], the fluctuation analysis (FA) [11], the detrended fluctuation analysis (DFA) [12], the detrending moving average analysis (DMA) [13], and so on.

Our work focuses on three methods (FA, DFA and DMA) that are very popular especially in the econophysics community. Consider a time series $\{x(t) : t = 1, 2, \cdots, N\}$ with zero mean and its profile $y(t)$ constructed as the cumulative sum of $x(t)$. The three methods proceed to obtain fluctuation functions $F(s)$ specific to a timescale s. For long-range correlated time series, we have

$$F(s) \sim s^{\alpha}, \tag{1}$$

where α is a scaling exponent. In FA, the fluctuation function is computed as follows [11]

$$F(s) = \sqrt{\langle [y(t+s) - y(t)]^2 \rangle}, \tag{2}$$

which is actually a special case of the structure function in turbulence [14]. In contrast, both DFA and DMA adopt detrending techniques. The time series $y(t)$ is covered by N_s disjoint boxes of size s. When the whole time series $y(t)$ cannot be completely covered by N_s boxes, we can utilize $2N_s$ boxes to cover the time series by starting from both ends of the time series. In each box, a trend function $g(t)$ of the sub-series is determined. The residuals are calculated by

$$\epsilon(t) = y(t) - g(t), \tag{3}$$

where the trend $g(t)$ is a polynomial function in the DFA algorithm [12] and a moving average function over s data points in the DMA method [13]. The fluctuation function $F(s)$ is then obtained as the r.m.s. of the residual time series:

$$F(s) = \sqrt{\frac{1}{N} \sum_{t=1}^{N} [\epsilon(t)]^2}. \tag{4}$$

Note that all these methods have a multifractal version [15–19] and can be generalized to handle high-dimensional fractals and multifractals [19–21]. When $y(t)$ is a fractional Brownian motion (FBM), the scaling exponent α is identical to the Hurst index H [22–25].

Several groups have attempted to assess the performance and relative merits of these techniques. Xu et al. [26] compare the performances of DFA and DMA on long-range power-law correlated time series synthesized using the modified Fourier filtering method [27], and find that DFA is superior to different DMA variants. Bashan et al. [28] observe that the centred DMA performs as well as DFA for long time series with weak trends and slightly outperforms DFA for short data with weak trends. They conclude that DFA "remains the method of choice" when
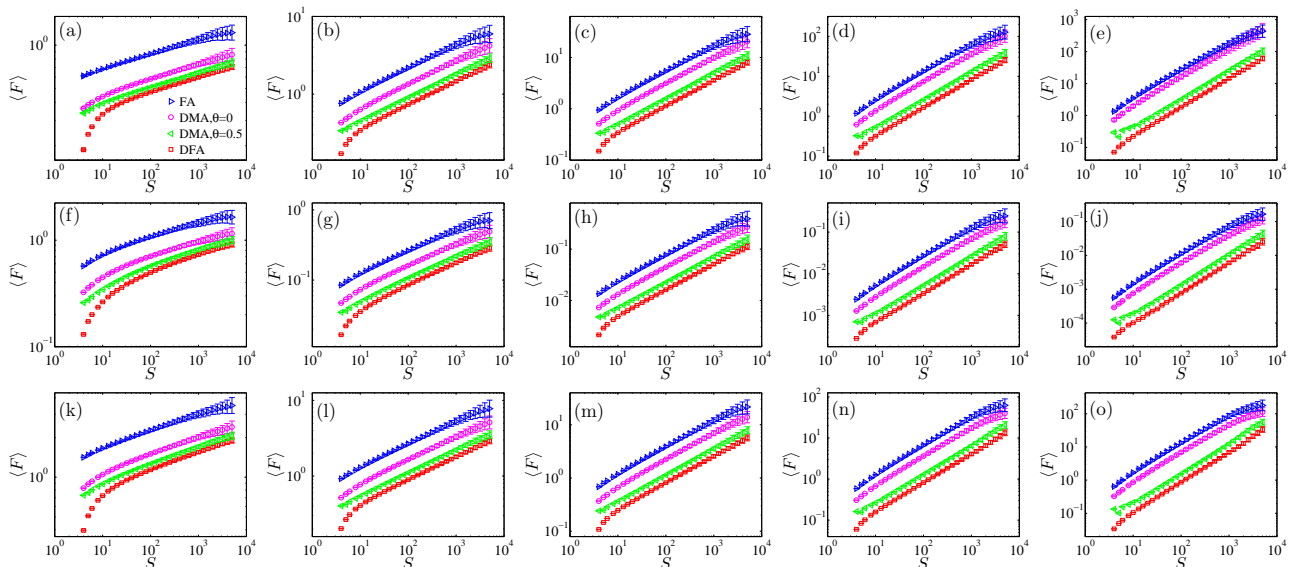

FIG. 1. **Scaling plots of $\langle F \rangle$ against s.** Each plot contains four curves obtained from four different analysis methods (FA, BDMA, CDMA and DFA) and each curve represents a fluctuation function averaged over 100 repeated simulated time series with the error bars showing the standard deviations. The three rows correspond to three generators (FGN-DH, FBM-RMD and WFBM from top to bottom). Each column corresponds to a fixed Hurst index ($H_{in} = 0.1, 0.3, 0.5, 0.7,$ and 0.9 from left to right). The curves have been shifted vertically for better visibility.

the trend is not *a priori* known. Serinaldi [29] uses the Davies-Harte algorithm to generate fractional Gaussian noises (FGNs) and FBMs by summing the FGNs [30], and find that DFA and DMA have comparable performances. Jiang and Zhou [31] report that DFA and the centred DMA perform similarly and both of them outperform the backward and forward DMA methods, when the FBMs are generated using the Fourier-based Wood-Chan algorithm [32]. Huang et al. [33] find comparative performances of FA and DFA for FBMs with $H = 1/3$, which are generated with the Wood-Chan algorithm [32]. In contrast, Bryce and Sprague [34] argue that FA outperforms DFA, for FGNs with $H = 0.3$ that are generated using the Davies-Harte algorithm [30].

We notice that these studies concentrate on DFA versus DMA or DFA versus FA and report what appears to be contradictory results when considered together. A careful reading unveils that these studies cannot be directly compared because they have adopted different synthesis algorithms (or generators) for the long-range correlated time series to be tested. Indeed, comparing the performances of long-range correlation detection methods is not an easy task for the following reasons. Firstly, there are many algorithms to generate FGNs and FBMs [35], and one should be careful not to draw too rapid conclusions on the relative performance of long-range correlation detection methods that may be sensitive to the micro-structure of the generated time series that depend on the specific synthesis algorithm. Secondly, real time series may contain *a priori* unknown nontrivial trends [36–39], which complicates significantly the detection of long-range correlations, because trends and long-range

correlations often lead to similar signals. Thirdly, there is no consensus on an objective determination approach of the scaling range, which plays a crucial role in the estimation of the scaling exponents. Often, studies use quite short scaling ranges (a decade or less), which is an hindrance for determining the genuine presence of long-range correlations [40–42].

In this work, we focus on comparing FA, DFA and two versions of DMA, where a linear detrending is adopted in DFA and the backward and centred versions of DMA (denoted BDMA and CDMA respectively) are investigated since the forward DMA performs the worst according to the literature. The comparison between FA, DFA and two versions of DMA is conducted on time series generated using three different algorithms, thus generating a 3×4 matrix of comparisons: (1) FGNs using the Davies-Harte algorithm (FGN-DH) [30] so that we can compare with the analysis by Bryce and Sprague [34], (2) FBMs using a wavelet-based generator (WFBM) [43], which input Hurst indexes are very close to the estimated DFA exponents even when $H < 0.5$ [44], and (3) FBMs using the random midpoint displacement algorithm (FBM-RMD) [45], because the numerical results of the generated time series are in excellent agreement with the analytical results for DMA [25]. Besides, we do not consider trends or other hidden nonlinear structures.

Results

Fluctuation functions. Figure 1 compares the fluctuation functions calculated with four different scaling analysis methods (FA, BDMA, CDMA, DFA) on time
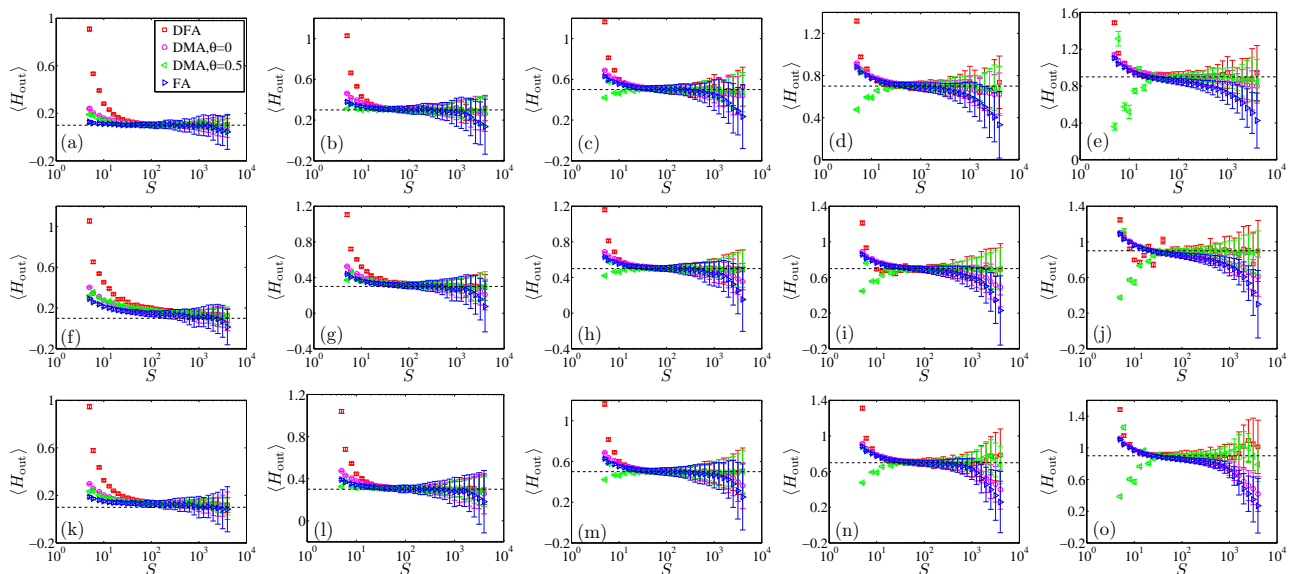
FIG. 2. **Local slopes of the fluctuation functions.** Each plot contains four curves obtained from four different scaling analysis methods (FA, BDMA, CDMA and DFA) and each curve represents a slope function averaged over 100 repeated simulated time series with the error bars showing the standard deviations. The three rows correspond to three generators (FGN-DH, FBM-RMD and WFBM from top to bottom). Each column corresponds to a fixed Hurst index ($H_{in} = 0.1, 0.3, 0.5, 0.7$, and 0.9 from left to right). The horizontal dashed lines indicates the exact value of the Hurst index used to generate the synthetic time series.

series generated using three different generators (FGN-DH, FBM-RMD and WFBM). We notice that panel (b) confirms the results in Ref. [34], which compares the performances of FA and DFA on FGNs with H_{in}. One can also notice that the error bar increases with s for each curve.

When the scale s is small and the Hurst index H_{in} is small, the curvature of the fluctuation function for DFA is remarkable, while the FA curve looks quite straight. In addition, the DMA curves also exhibit some mild curvature. With the increase of the Hurst index H_{in} of the analysed time series, the curvature of the DFA and DMA curves decreases. We thus confirm that FA performs best in most cases and DFA performs worst at small scales.

However, the conclusions are very different at large scales. The DFA curves have the smallest error bars, the centred DMA curves show the second smallest error bars, and the FA curves exhibit the largest error bars. More significantly, the DFA and CDMA curves are very straight, while the FA and BDMA curves exhibit some clear curvature with the magnitude of the curvature becomes larger with the increase of the Hurst index H_{in}.

These observations are qualitatively the same for different time series generators.

Local slopes. Figure 2 compares the local slopes, which are the estimates of the Hurst exponent, calculated with four different scaling analysis methods on the time series generated using three different generators. Comparing the three plots of each column, it is found that the relative performances are qualitatively the same for the three time series generators. For each scaling analysis method,

the error bars become larger with the increase of the scale for each fixed Hurst index H_{in} or with the increase of the Hurst index H_{in} at fixed scale. Again, the error bars of the DFA curve are the largest in each plot.

At large scales, we find that FA is the worst in the sense that the FA curves have the largest error bars and deviate the most from the theoretical line $\langle H_{out} \rangle = H_{in}$. In contrast, DFA and CDMA have comparable performances and perform best.

At small scales, the order of performance, as measured by the proximity of the estimates of the scaling exponents to the true Hurst values and by the size of the error bars, is FA $\succ$ CDMA $\succ$ BDMA $\succ$ DFA for $H_{in} = 0.1$ in the first column, CDMA $\succ$ FA $\succ$ BDMA $\succ$ DFA for $H_{in} = 0.3$ in the second column, CDMA $\simeq$ FA $\succ$ BDMA $\succ$ DFA for $H_{in} = 0.5$ in the third column, FA $\succ$ BDMA $\succ$ CDMA $\succ$ DFA for $H_{in} = 0.7$ in the fourth column, and FA $\succ$ BDMA $\succ$ CDMA $\simeq$ DFA for $H_{in} = 0.9$ in the fifth column, where A $\succ$ B means that A is superior to B.

Effect of scaling range. In order to perform the scaling analysis onto real systems using any of the above methods, it is of crucial importance to determine the scaling range. This is because the estimate of the scaling exponent may vary dramatically if one changes the scaling range. We now investigate the effect of the scaling range on the estimation accuracy of the Hurst index performed with the four scaling analysis methods applied to time series synthesized by the three different generators.

Let us first consider the FGNs. We find that the FA gives accurate estimates when $H_{in} < 0.5$, while the es-
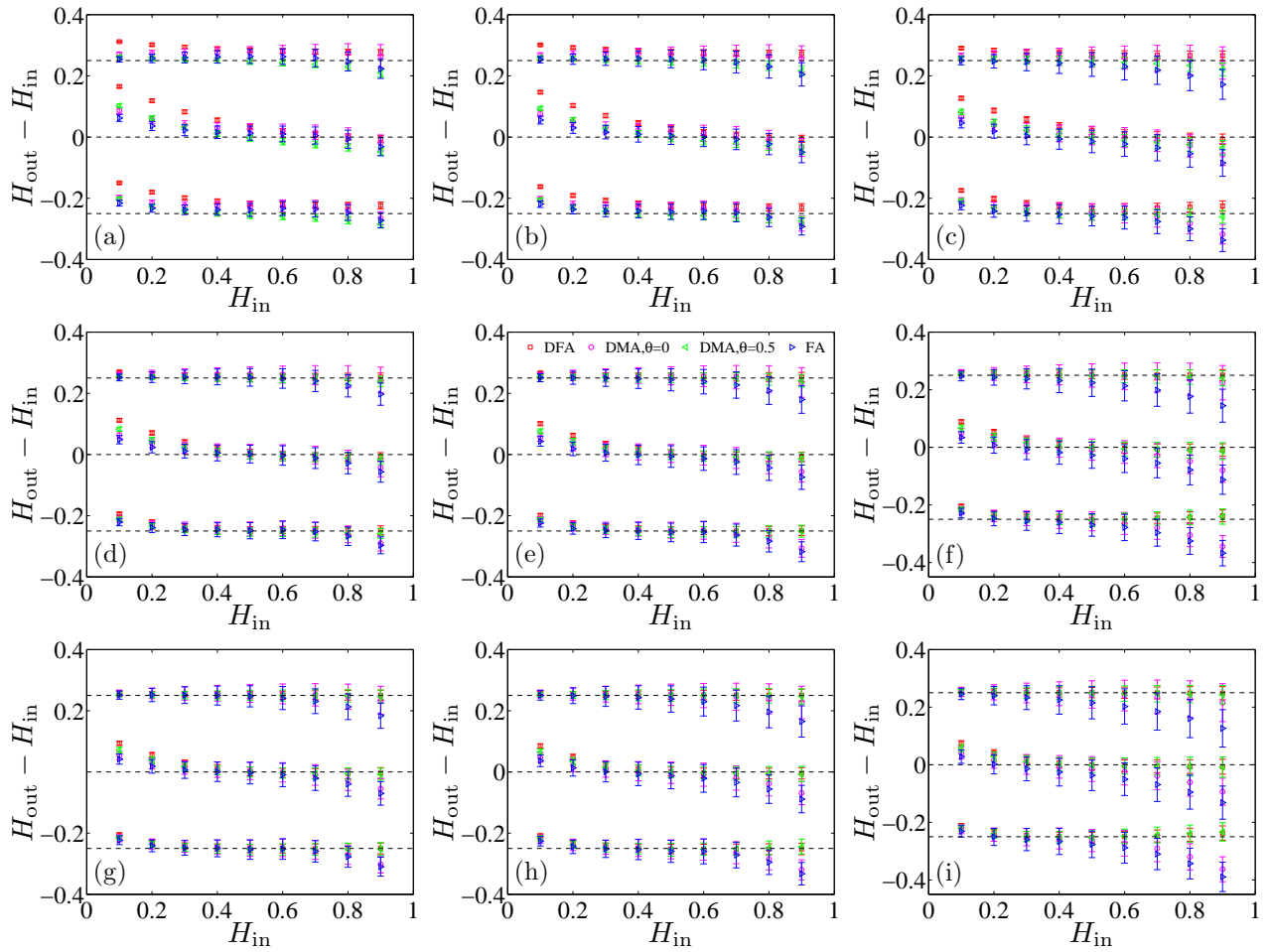
FIG. 3. **Impacts of the scaling range on the Hurst index estimates.** Each plot has a different scaling range $[s_{\text{left}}, s_{\text{right}}]$, where $s_{\text{left}} = 4, 10, 20$ from left column to right column and $s_{\text{right}} = 999, 1992, 5000$ from top row to bottom row. In each plot, there are three clusters of curves. Each cluster corresponds to the three generators (FGN-DH, FBM-RMD and WFBM from top to bottom). The top and bottom clusters have been shifted vertically by +0.25 and -0.25 respectively for better visibility. In each clusters, there are four sets of points with their error bars that are obtained from four different analysis methods (FA, BDMA, CDMA and DFA). Each point shows the average slope of the Hurst index estimates over 100 simulated time series. The error bars show the standard deviations.

timated indexes deviate more and more from the theoretical values when H_{in} increases in the persistent time series range, for all nine scaling ranges. The DFA estimates are not accurate only when $s_{\text{right}} = 999$ (first row) and $H_{\text{in}} < 0.5$ and DFA outperforms FA for all the other cases. More intriguingly, CDMA gives very accurate estimates of the Hurst indexes and performs the best almost in all situations. Overall, DFA outperforms BDMA and FA is the worst estimator.

For the time series generated with FBM-RMD and WFBM, the relative performances of the four scaling analysis methods are qualitatively the same. When $H_{\text{in}} \ll 0.5$, FA $\succ$ BDMA $\succ$ CDMA $\succ$ DFA. For other situations, DFA and CDMA give very accurate estimates of the Hurst indexes and perform the best, while FA performs the worst.

Taking all these observations together, we conclude that CDMA has the best performance and DFA is slightly

worse. When the scaling range is properly determined, DFA and CDMA have similar performances. In contrast, FA has the worst performance, especially in the sense that it cannot provide accurate estimations of the Hurst index for persistent time series.

Discussion

We have investigated the performances of four estimators (FA, DFA, BDMA, and CDMA) for the characterization of long-range power-law correlated time series synthesized with three different generators (FGN-DH, FBM-RMD and WFBM). We have illustrated that, overall, CDMA and DFA are the best and exhibit comparable performances, while FA performs the worst. In particular, CDMA and DFA are less sensitive than FA to the choice of the scaling range. We depart significantly from

the conclusion of Ref. [34] that FA is superior to DFA, by showing that this statement holds only for very special cases (FGNs with $H_{in} = 0.3$) that cannot be extended to other situations.

An important issue is the effect of the length of time series on the results and conclusions, especially for short time series. We repeated the analysis by generating time series of length 2000, which corresponds to time windows of 8 years of trading at the daily scale, or less than a week of data sampled at the minute time scale. The analysis comparing the results for windows of 2000 time steps to those for windows of 20000 time steps is presented in Supplementary Information and confirms that the conclusions remain unchanged, because the corresponding plots for the two cases with different time series lengths are almost indistinguishable, except that the results for shorter time series have larger fluctuations.

Methods

Description and preprocessing of the data: For each generator (FGN-DH, FBM-RMD or WFBM), we synthesize 100 time series of length 20000 for a given Hurst index H_{in}. These time series are used in all the analyses. The discrete values of the fluctuation function $F(s)$ of each time series for each scaling analysis method are calculated at 32 s-values logarithmically sampled in the interval $[4, 5000]$.

Figure 1 details: Each point $(\langle F(s)\rangle, s)$ shows the average of 100 $F(s)$ values over the 100 time series for each H_{in} at scale s for a given generator and a given estimator.

Figure 2 details: For each time series, we calculate the local slope of $\ln F(s)$, which is the centred difference using two adjacent data points. Each point shows the average and the standard deviation estimated over the corresponding 100 local slopes.

Figure 3 details: For each time series, we calculate the slope of $\ln F(s)$ using the data points within the chosen scaling range. Each point shows the average and the standard deviation over the corresponding 100 slopes.

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[2] D. Sornette, *Critical Phenomena in Natural Sciences*, 2nd ed. (Springer, Berlin, 2004).
[3] M. Taqqu, V. Teverovsky, and W. Willinger, Fractals **3**, 785 (1995).
[4] J. W. Kantelhardt, in *Encyclopedia of Complexity and Systems Science*, Vol. LXXX, edited by R. A. Meyers (Springer, Berlin, 2009) pp. 3754–3778.
[5] H. E. Hurst, Trans. Amer. Soc. Civil Eng. **116**, 770 (1951).
[6] M. Holschneider, J. Stat. Phys. **50**, 963 (1988).
[7] J.-F. Muzy, E. Bacry, and A. Arnéodo, Phys. Rev. Lett. **67**, 3515 (1991).
[8] E. Bacry, J.-F. Muzy, and A. Arnéodo, J. Stat. Phys. **70**, 635 (1993).
[9] J.-F. Muzy, E. Bacry, and A. Arnéodo, Phys. Rev. E **47**, 875 (1993).
[10] J.-F. Muzy, E. Bacry, and A. Arnéodo, Int. J. Bifur. Chaos **4**, 245 (1994).
[11] C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons, and H. E. Stanley, Nature **356**, 168 (1992).
[12] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, Phys. Rev. E **49**, 1685 (1994).
[13] E. Alessio, A. Carbone, G. Castelli, and V. Frappietro, Eur. Phys. J. B **27**, 197 (2002).
[14] A. N. Kolmogorov, J. Fluid Mech. **13**, 82 (1962).
[15] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, and Y. Dodge, Nature **381**, 767 (1996).
[16] A. Castro e Silva and J. G. Moreira, Physica A **235**, 327 (1997).
[17] R. O. Weber and P. Talkner, J. Geophys. Res. **106**, 20131 (2001).
[18] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley, Physica A **316**, 87 (2002).
[19] G.-F. Gu and W.-X. Zhou, Phys. Rev. E **82**, 011136 (2010).
[20] G.-F. Gu and W.-X. Zhou, Phys. Rev. E **74**, 061104 (2006).
[21] A. Carbone, Phys. Rev. E **76**, 056703 (2007).
[22] P. Talkner and R. O. Weber, Phys. Rev. E **62**, 150 (2000).
[23] C. Heneghan and G. McDarby, Phys. Rev. E **62**, 6103 (2000).
[24] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, and A. Bunde, Physica A **295**, 441 (2001).
[25] S. Arianos and A. Carbone, Physica A **382**, 9 (2007).
[26] L. M. Xu, P. C. Ivanov, K. Hu, Z. Chen, A. Carbone, and H. E. Stanley, Phys. Rev. E **71**, 051101 (2005).
[27] H. Makse, S. Havlin, M. Schwartz, and H. E. Stanley, Phys. Rev. E **53**, 5445 (1996).
[28] A. Bashan, R. Bartsch, J. W. Kantelhardt, and S. Havlin, Physica A **387**, 5080 (2008).
[29] F. Serinaldi, Physica A **389**, 2770 (2010).
[30] R. B. Davis and D. S. Harte, Biometrika **74**, 95 (1987).
[31] Z.-Q. Jiang and W.-X. Zhou, Phys. Rev. E **84**, 016106 (2011).
[32] A. T. A. Wood and G. Chan, J. Comput. Graph. Stat. **3**, 409 (1994).
[33] Y.-X. Huang, F. G. Schmitt, J.-P. Hermand, Y. Gagne, Z.-M. Lu, and Y. L. Liu, Phys. Rev. E **84**, 016208 (2011).
[34] R. M. Bryce and K. B. Sprague, Sci. Rep. **2**, 315 (2012).
[35] W.-X. Zhou and D. Sornette, Int. J. Modern Phys. C **13**, 137 (2002).
[36] A. Montanari, M. S. Taqqu, and V. Teverovsky, Math. Comput. Modell. **29**, 217 (1999).
[37] K. Hu, P. C. Ivanov, Z. Chen, P. Carpena, and H. E. Stanley, Phys. Rev. E **64**, 011114 (2001).
[38] Z. Chen, P. C. Ivanov, K. Hu, and H. E. Stanley, Phys. Rev. E **65**, 041107 (2002).

[39] Z. Chen, K. Hu, P. Carpena, P. Bernaola-Galvan, H. E. Stanley, and P. C. Ivanov, Phys. Rev. E **71**, 011104 (2005).

[40] O. Malcai, D. A. Lidar, O. Biham, and D. Avnir, Phys. Rev. E **56**, 2817 (1997).

[41] B. B. Mandelbrot, Science **279**, 783 (1998).

[42] D. Avnir, O. Biham, D. Lidar, and O. Malcai, Science **279**, 39 (1998).

[43] P. Abry and F. Sellan, Appl. Comp. Harmonic Anal. **3**, 377 (1996).

[44] X.-H. Ni, Z.-Q. Jiang, and W.-X. Zhou, Phys. Lett. A **373**, 3822 (2009).

[45] B. B. Mandelbrot, *The Fractal Geometry of Nature* (W. H. Freeman, New York, 1983).

Author contributions

ZQJ, WXZ and DS conceived the study, YHS, WXZ and DS designed the study, and YHS, GFG, ZQJ, WXZ and DS performed the study. WXZ and DS wrote the paper and reviewed the manuscript.
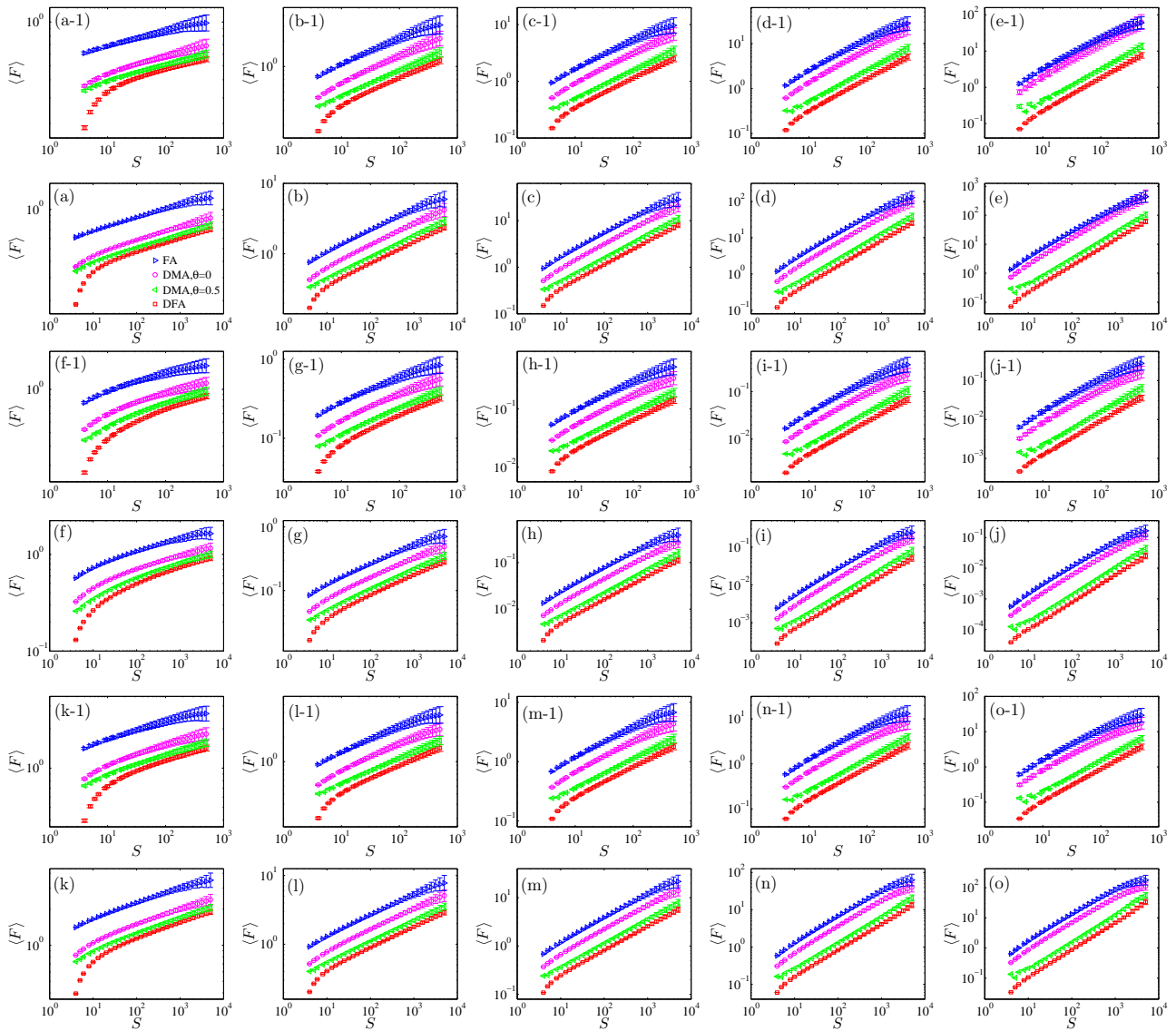
FIG. 4. Comparing plots of $\langle F \rangle$ against s. The plots labeled (a)-(o) are the same as in the paper where the length of time series is 20000, while the plots labeled with (a-1) to (o-1) are the results where the length of time series is 2000.
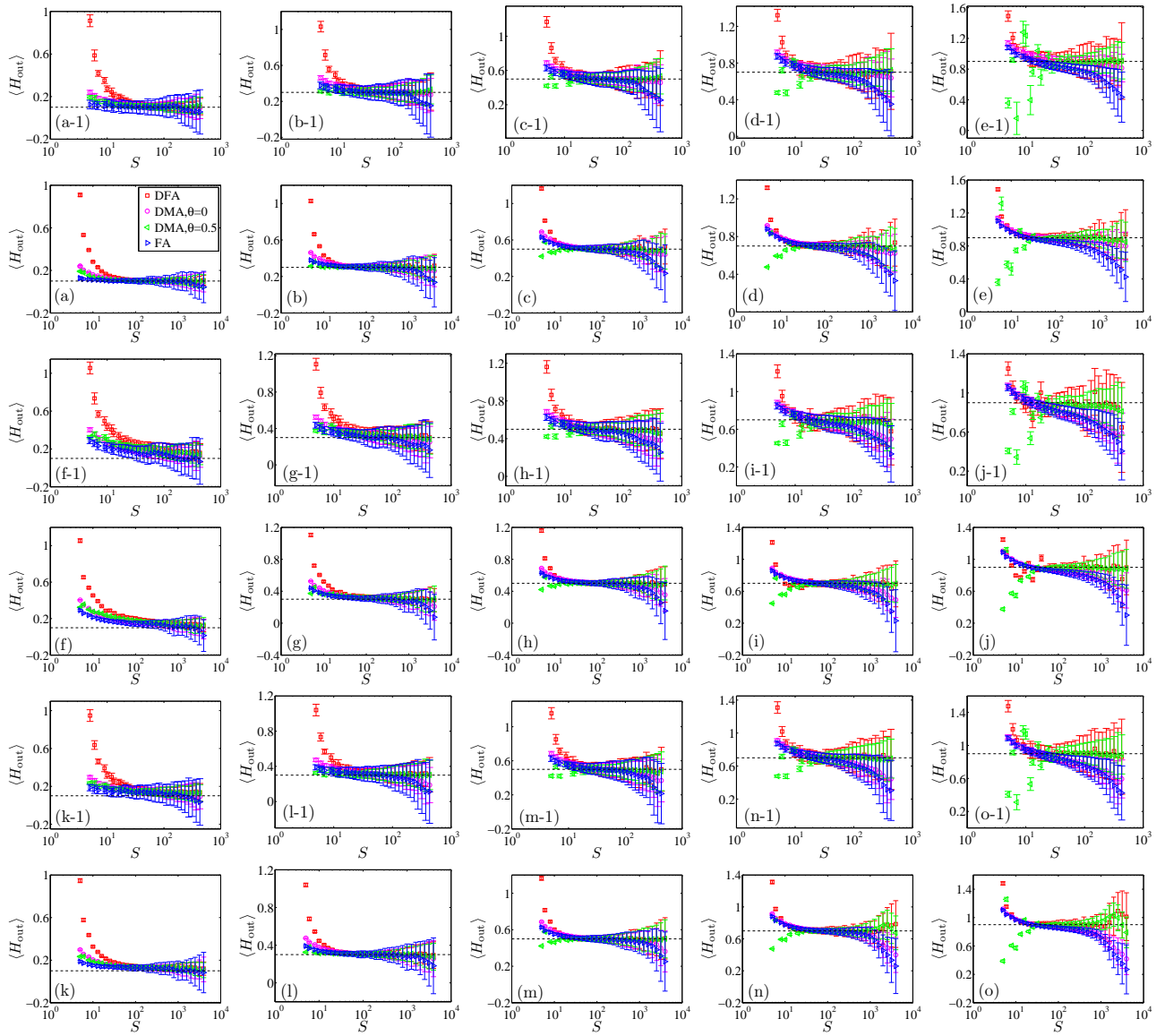
FIG. 5. **Comparing local slopes of the fluctuation functions.** The plots labeled (a)-(o) are the same as in the paper where the length of time series is 20000, while the plots labeled with (a-1) to (o-1) are the results where the length of time series is 2000.
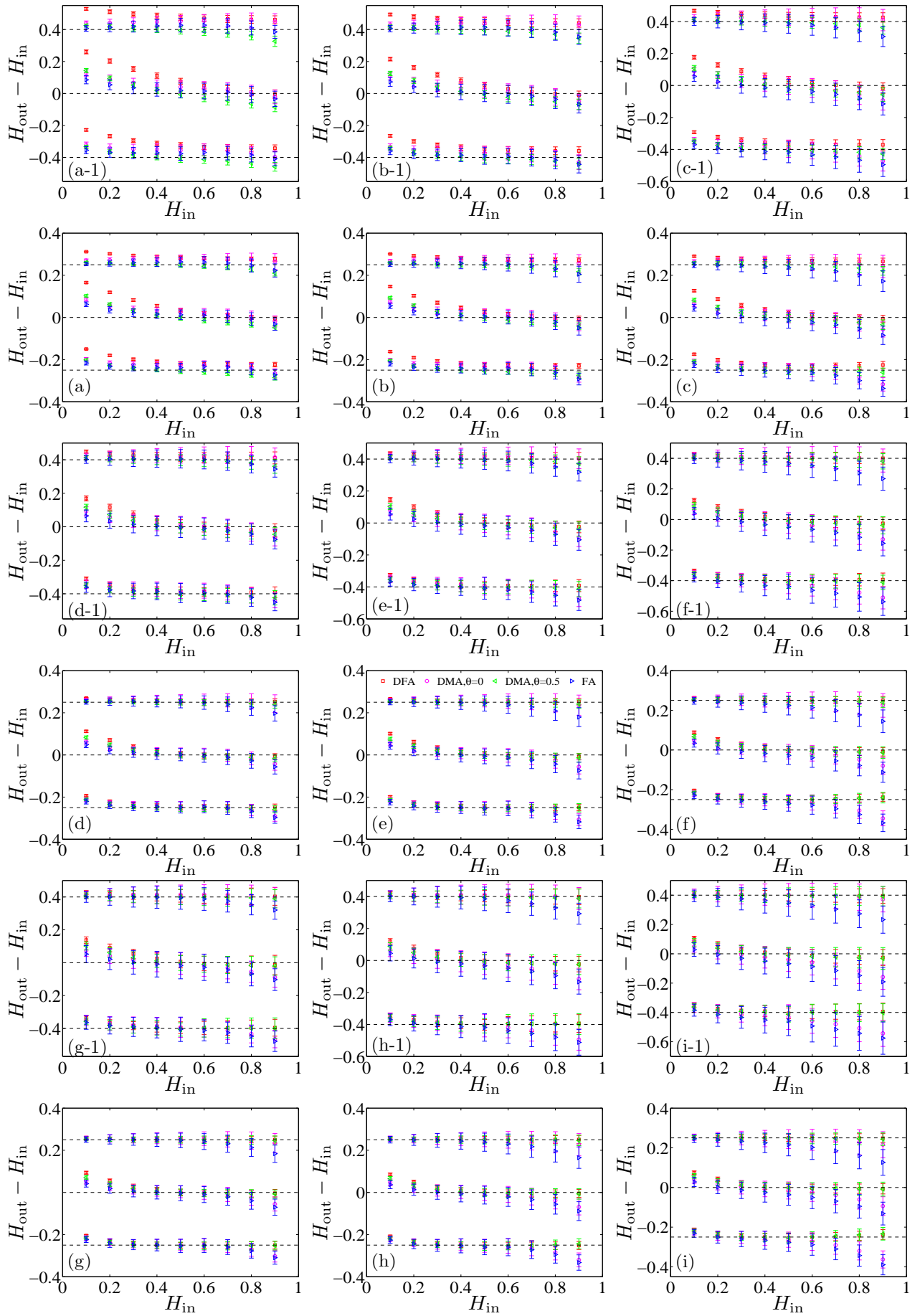
FIG. 6. **Comparing impacts of the scaling range on the Hurst index estimates.** The plots labeled (a)-(o) are the same as in the paper where the length of time series is 20000, while the plots labeled with (a-1) to (o-1) are the results where the length of time series is 2000.