Proceedings of the World Congress on Engineering 2017 Vol II
WCE 2017, July 5-7, 2017, London, U.K.

Long Range Correlation in Time Series of News Sentiments

Sergei Sidorov, *Member, IAENG,* Alexey Faizliev and Vladimir Balash

***Abstract*—Using artificial intelligence algorithms, providers of news analytics calculate the sentiment score of almost every economic and financial news in real time. The sentiment score of negative, neutral, positive news are assigned to be -1, 0, 1, respectively. We constructed time series of news sentiments as follows: a nine-month period of 2015 was divided into non-overlapping consecutive intervals of equal length, and then we calculated the sum of sentiment scores of all news within each time interval. In this paper we examine long-range dependance and self-similarity of time series of sentiments of economic and financial news using the Detrended Fluctuation Analysis of order 1 (DFA), Rescaled Range Analysis (R/S), Average Wavelet Coefficient Method (AWC) and Fourier Transform Method (FTM). Empirical results obtained by this methods show that time series of news sentiments exhibit self-similarity (as well as a long memory property). The Hurst exponent (as well as the long-range correlation exponent) is greater than 0.55 over four orders of magnitude in time ranging from several minutes to dozen of days. DFA and AWC methods allowed us to reveal a strong scaling behavior as well as to detect a distinct crossover effect. On the other hand, it turns out that for the classic R/S analysis and Fourier transform techniques, the scaling regimes and/or positions of crossovers are hard to define.**

***Index Terms*—long-range correlation, time series, detrended fluctuation analysis, Hurst exponent.**

I. Introduction

Long range dependance and self-similarity of financial time series has been the focus of attention for many researchers in the last several years [1]–[16]. However, there are very few works devoted to the analysis of the statistical properties of the news flow. Following papers [17] and [18], in this work we analyze time series of news sentiments.

News analytics is an unconventional approach to the analysis of news flow based on artificial intelligence techniques. Introduction to the news analytics tools and techniques can be found in [19] and [20]. For each news, news analytic providers find it sentiment score using artificial intelligence algorithms [19]. The sentiment score of a neutral news is assumed to be 0. Every positive news has sentiment score 1. If news is negative then its sentiment score is assigned to be -1. Then we constructed time series of news sentiments as follows. We use the final data sample and divide the whole period into non-overlapping consecutive intervals of equal length $\delta = 1$ and 5 minutes and calculated the sum of sentiments of all news within each time interval.

In contrast to the work [17], in this paper we will use Detrended Fluctuation Analysis of orders 1 (DFA), Rescaled Range Analysis (R/S), Average Wavelet Coefficient Method (AWC) and Fourier Transform Method (FTM) to examine long-range auto-correlation and self-similarity of time series of news sentiments. We suppose that the usage of the three estimators may prevent us from obtaining one-sided results. By using a bigger data period, we could study the dependence of the Hurst exponent not only on intra day intervals, but at intervals of a few days duration. Moreover, our analysis is based on a much bigger data period (189 instead of 43 trading days) from relatively recent news analytic data (from January 1, 2015 to September 22, 2015).

The results of this paper show that the long-range power-law correlation take place in time series of news sentiments. All four methods showed the presence of a long-range correlation in the time series of news sentiments (the Hurst exponent is greater than 0.5).

II. Methods

A. Auto-correlation types

Let $X = (x_t)_{t=1}^n$ be a time series with large n and let $s \in \mathbb{N}$, $s \ll n$. Let $C(s)$ denote the (auto) correlation between $X_1 = (x_t)_{t=1}^{n-s}$ and $X_2 = (x_{t+s})_{t=1}^{n-s}$. The correlation can be of several types:

1) x_t are *uncorrelated*; it is obvious that if X_1 and X_2 are uncorrelated then $C(s)$ must be zero, $C(s) = 0$;
2) *short-range correlations* of $(x_t)_{t=1}^n$ results in exponentially declining of $C(s)$, i.e. $C(s) \sim e^{-s/s_0}$, where s_0 is a parameter of decay;
3) long-range correlation of the $(x_t)_{t=1}^n$, $C(s)$ should follow a power-law dependence: $C(s) \sim s^{-\gamma}$, $0 < \gamma < 1$.

We use the following methods for estimation of the correlation.

B. DFA method

DFA was proposed in the papers [21], [22]. This method is used for studying the indirect scaling of the long-range dependence. DFA method was effectively applied for solving a great amount of problems in science as well as in engineering field, including DNA analysis [23]–[25], biomedical signal processing [26]–[30], study of daily internet traffic dynamics [31], analysis of economical and finance time series [1]–[5], human gait behavior [32], [33]. While DFA has some drawbacks [34], the work [35] remarks that DMA and DFA stay the options of choice for evaluating the long-range correlation of time series. DFA algorithm includes five stages:

1) *Integration.* We calculate $y_k = \sum_{t=1}^k (x_t - \overline{x})$, $k = 1, 2, \dots, n$, where $\overline{x} = \sum_{t=1}^n x_t$.
2) *Cutting.* We separate the $(y_k)_{k=1}^n$ into $n_s = [n/s]$ non-crossing intervals, each of length s, starting with y_1.
3) *Fitting.* For each interval $l = 1, \dots, n_s$ we construct a fitting linear function P_l (trend) by means of least-square fit of the data $(y_k)_{k=(l-1)s+1}^{ls}$. Denote $y_k^* = P_l(k)$.

Manuscript received February 24, 2017; revised March 14, 2017. This work is supported by RFBR (grant 16-01-00507).
V. A. Balash, S. P. Sidorov, A. R. Faizliev are with Saratov State University, Saratov, 410012 RUSSIA e-mail: sidorovsp@info.sgu.ru.

4) *Detrending.* The detranded time series are obtained by $\epsilon_k(s) = y_k - y_k^*$.
5) We calculate variance of residuals ϵ_i:

$$F_l^2(s) = \frac{1}{s}\sum_{i=1}^{s} \epsilon^2_{(l-1)s+i}(s),\ l = 1, \dots, n_s.$$

6) *DFA fluctuation function* is defined by

$$F(s) = F^{[m]}_{\text{DFA}}(s) = \left(\frac{1}{n_s}\sum_{l=1}^{n_s} F_l^2(s)\right)^{1/2},$$

where m denotes the degree of approximating polynomials use on stage 3.

Derivation of DFA can be found in the work [36].

C. The auto-correlation parameter

The fluctuation functions $F(s)$ deduced in DFA let us examine the s-dependance of F. In the case of long-range power-law correlation of x_t, the F must follow a power-law

$$F(s) \sim s^{\alpha}$$

for large enough s. The fluctuation parameter α is connected with the value of correlation exponent γ in the following way (see [37]) $\alpha = 1 - \gamma/2,\ 0 < \gamma < 1$. The correlation parameter α reflects the auto-correlation properties of time series as follows:

1) $\alpha = 1/2$, the time series uncorrelated (white noise) or short-range correlated;
2) $\alpha < 1/2$, it is anti-correlated;
3) $\alpha > 1/2$, it is long-range power-law correlated;
4) $\alpha = 1$, pink noise ($1/f$ noise).

D. Rescaled Range Analysis (R/S)

The R/S algorithm is one of the most widely used methods for scale exponent estimation. R/S algorithm estimates the value of the Hurst exponent based on an empirical data set for the long-range dependent process that generated the data set. R/S-analysis uses a heuristic approach developed in [38], [39], [40], [41], [42].

R/S algorithm consists of the following steps:

1) *Cutting.* We divide the $(x_k)_{k=1}^n$ into $n_s = [n/s]$ non-overlapping intervals $X^{(i)}(s) := \{x_{(i-1)s+1}, \dots, x_{is}\}$, $i = 1, \dots, n_s$, each of length s, starting with x_1.
2) *Accumulation.* We calculate the accumulated series for each window $X^{(i)}(s)$, $i = 1, \dots, n_s$, as follows

$$y_j^{(i)}(s) = \sum_{t=(i-1)s+1}^{j} (x_t - \overline{x}(s)), \qquad (1)$$

where $\overline{x}(s) = 1/s\sum_{t=(i-1)s+1}^{is} x_t$ is the mean of ith window of size s, and $j = (i-1)s+1, \dots, is$.
3) *Range calculation.* We compute the range of deviation within each window for the accumulated series as follows $R^{(i)}(s) = \max_{(i-1)s+1\le j\le is}\{y_j^{(i)}(s)\} - \min_{(i-1)s+1\le j\le is}\{y_j^{(i)}(s)\}$.
4) *Standard deviation calculation.* For each window $i = 1, \dots, n_s$ we find the standard deviation:

$$S^{(i)}(s) = \sqrt{\frac{1}{s}\sum_{t=(i-1)s+1}^{is} (x_t - \overline{x}(s))^2}.$$

5) *Computation of R/S statistics:*

$$\langle R(s)/S(s)\rangle := \frac{1}{n_s}\sum_{i=1}^{n_s} \frac{R^{(i)}(s)}{S^{(i)}(s)},\ s = s_1, \dots, s_L.$$

6) *Estimation of Hurst exponent.* We solve the following least mean square problem

$$\sum_{j=1}^{L} \left(\log\left(\langle R(s_j)/S(s_j)\rangle\right) - H\log s_j - b\right)^2 \to \min_{H,b}$$

R/S algorithm computes the R/S-statistic for different s and plots the resulting estimates versus the s on loglog scale. Then the Hurst parameter can be estimated via the slope of the resulting log-log plot. As it pointed out in [43], classical R/S-analysis is not suitable for small samples, but can be highly effective for quite large samples and it often provides a rather accurate picture of the presence or absence of long-range dependence in a given empirical data sets. Moreover, R/S-analysis demonstrates relative robustness under heavy tails with infinite variance in the marginal distribution of the data [38]–[41].

On the other hand, R/S-analysis is quite sensitive relative to the presence of explicit short-range dependence structures and its bias. Therefore, these shortcomings of R/S analysis lead to the fact that many researchers do not consider this algorithm as a rigorous statistical method.

III. Wavelet-based methods

A. The Hurst exponent

The self-similarity parameter $0 < H < 1$ of self-affine processes is also called the Hurst (or roughness) exponent [44]. The Hurst exponent is commonly used for measuring the duration of long-range dependence of a stochastic process.

There are three possibilities:

- If $H = 0.5$ then $C(s) = 0$ which means that past and future increments are uncorrelated (Brownian motion);
- In case $H > 0.5$ we have $C(s) > 0$ and the increments are positively correlated (the process $\{X(t)\}_t$ is called persistent).
- If $H < 0.5$ then $C(s) < 0$ and increments are negatively correlated (the process $\{X(t)\}_t$ is called anti-persistent or anti-correlated).

B. The Average Wavelet Coefficient Method

The Average Wavelet Coefficient Method (AWC) was proposed in papers [45] and [46].

The method is used for measuring the temporal self-affine correlations of a time series by estimating its Hurst exponent. The AWC method is based on the wavelet transform (good review of the wavelet transform can be found in books [47] and [48]).

The strategy for the data-analysis by the AWC method consists of three main steps:

1) Wavelet transformation of the data $X(t)$ into the wavelet domain, $\mathcal{W}[X](a, b)$, where a, b are scale and location parameters, respectively [47], [48].
2) Then for a given scale a we can find a representative wavelet amplitude for that particular scale, and to study

ISBN: 978-988-14048-3-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

its scaling. To do so we calculate the averaged wavelet coefficient $W[X](a)$ according to the equation

$$W[X](a) = \langle |\mathcal{W}[X](a,b)| \rangle_b,$$

where $\langle \cdot \rangle_b$ denotes the standard arithmetic mean value operator with respect to the b.

3) For a self-affine process $X(t)$ with exponent H, the spectrum $W[X](a)$ should scale as $a^{H+0.5}$ [46]. Therefore, to find $H+0.5$ we should plot $W[h](a)$ against scale a in a log-log plot. As it is pointed out in [46], a scaling regime consisting of a straight line in this plot implies a self-affine behavior of the data.

C. Fourier transform method

Let us recall that the Fourier transform of the function f and inverse transform are defined by

$$\widehat{f}(\zeta) = \sum_t e^{-i\zeta t} f(t), \quad (2)$$

$$f(t) = \sum_\zeta e^{i\zeta t} \widehat{f}(\zeta). \quad (3)$$

Fourier inversion theorem states that f can be reconstructed from $\widehat{f}$.

If $\rho_x(s)$ is autocorrelation function of $x(t)$, $\rho_x(s) = \sum_t x(t)x(t+s)$, then the Fourier transform of $\rho_x(s)$ is

$$\widehat{\rho}(\zeta) = |\widehat{x}(\zeta)|^2. \quad (4)$$

Therefore, to find the value of $\rho_x(s)$ one should make the following steps

1) Find $\widehat{x}(\zeta)$, the Fourier transform of $x(t)$, using (2).
2) Find $\widehat{\rho}(\zeta)$, the Fourier transform of $\rho_x(s)$, using (4)
3) Compute $\rho_x(s)$ as the inverse transform of $\widehat{\rho}_x(\zeta)$ using (3).

Fourier transform is a classical way of correlation exponent estimation γ, but it is often not appropriate due to noisy nature, non-stationarity and imperfect measurement of data x_t.

$\widehat{\rho}(\zeta)$ is a power spectrum $P(\omega)$ of the Fourier transform. $\omega = 0 : Fs/n : Fs/2$ is a frequency increment (angular frequency), where Fs is sampling frequency, and n - number of samples.

IV. Empirical Results

A. *time series of news sentiments*

Providers of news analytics obtain and aggregate data from different sources (including news agencies and business reports) and social media. Our data covers the period from January 1, 2015 to September 22, 2015 (i.e. 189 trading days). We consider all the news released during this period. Initially we performed data selection and cleaning process as described in [17] or [18].

For each news, news analytic providers find it sentiment score using artificial intelligence algorithms. The sentiment score of a neutral news is assumed to be 0. Every positive news has sentiment score 1. If news is negative then its sentiment score is assigned to be -1.

Then we constructed time series of news sentiments as follows. We divided 189-day period Δ into n non-crossing consecutive segments $\Delta_1, \ldots, \Delta_n$ of the same longevity δ minutes, $\Delta = \Delta_1 \cup \ldots \cup \Delta_n$. We found x_t, the sum of all sentiment scores of all economical and finance news reported in the world during each interval Δ_t, $t = 1, 2, \ldots, n$. The sequence $x_1, x_2, \ldots, x_n$ is the time series of news sentiments with the δ minutes window. The overall sentiment of news for the 189-day period is 2011463. Table I shows the summary statistics.

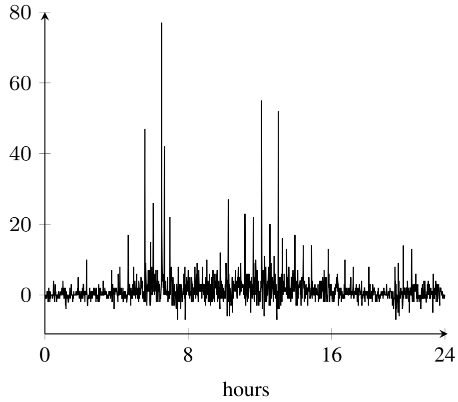

Fig. 1. Dynamics of sentiments (February 2, 2015)

TABLE I
SUMMARY STATISTICS OF TIME SERIES, $\delta = 1$ AND $\delta = 5$ MINUTES

δ, minutes	1	5
n	$2.7 \cdot 10^5$	$5.4 \cdot 10^4$
Sum sentiment	279750	126063
Mean	1.03	2.32
Minimum	-45	-45
Maximum	237	237
St. deviation	4.03	7.60
Median	0	1
Skewness	11.54	7.75
Kurtosis	282.18	104.38

Fig. 1 plots an example of time series of news sentiments with 1-min window. It is evident that time series is quite volatile and demonstrates a non-stationary behavior. The amount of positive news is much greater than the amount of negative news during the period.

B. Self-Similarity Analysis

In this subsection we demostrate the auto-correlation and self-similarity analysis with the use of DFA, R/S analysis, AWC and Fourier transform method.

First we use DFA (of order 1) method to evaluate the correlation and Hurst exponent. Then, we obtained values of $F(s_i)$ for different interval lengths $s_i \in [10^1, 10^{4.5}]$ using DFA. The task of selecting a needed length for scaling range has been examined in the papers [49], [50], [51]. The influence of scaling range on the efficiency of some detranding techniques has been investigated in the same works. We used two ranges $[10^1, 10^{4.5}]$ (for 1-min data) and $[10^1, 10^4]$ (for 5-min data) in our analysis. This values belong to the scaling range adviseded by these papers.

Figures 2, 3 plot the results of DFA application to 1-min or 5-minute time-series of sentiments respectively. Outcomes

ISBN: 978-988-14048-3-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

show that for 1-min and 5-min data sets the estimates of the Hurst exponent are equal to 0.7 for 1-min and to 0.62 for 5-minute time-series, which indicates the presence of long-range correlation. Linear regressions are highly significant, since the determination coefficients R^2 are close to 1. It must be pointed out that Detranded Fluctuation Analysis let us discover the crossover effect shown in the study [37].

AWC method gives similar results with DFA method in identical estimates of the Hurst exponent on the whole scale of s both for 1 minute or 5 minute time series (Table II). The differences between estimates of the Hurst exponent obtained by means of two different methods are less than 0.03. It should be noted that regression errors are very small and the value of R^2 is close to 1 for both methods. We remark that AWC method also allowed to detect the so called crossover effect (Figures 6, 7). it assessed the Hurst exponent in small and large scales separately. It should be noted that the crossover effect matches with 1 day period (or 24 hours). Estimates of the Hurst exponent, for small (about one day) as well as large scales (more than a day) using DFA and AWC methods yield similar results, the difference being less than 0.04, on the whole scale of for both 1 minute or 5 minute time series.

Figures 4, 5 demonstrate the log-log plot of dependance R/S statistics of s for 1-minute and 5-minute time series obtained by R/S analysis. In this case, though the values of the Hurst exponent are less than $H = 0.63$ and 0.56 respectively, they still demonstrate that there are positive long-range correlations in time series. Linear regressions are also highly significant. In contrast to the DFA, R/S analysis failed to reveal the presence of the crossover effect.

The Fourier transform analysis of the 1-min and 5-min time series of news sentiments presented in Figures 8 and 9. The figures show the raw power spectrum, $P(w)$ vs. angular frequency w for the news flow intensity data. We can see that the raw power spectrum is too noisy, which does not allow us to accurately estimate the Hurst exponent and certainly does not give us to detect the effect of the crossover. We can apply the log-binning technique [52] to reduce its noise. Results of the log-binning smoothing are shown in Figures 10 and 11. The dashed lines of figures corresponds to the slope of $(-2H - 1)$ with the values of $H = 0.61$ and $H = 0.56$ for 1-minute and 5-minute time series respectively. It should be noted that while regressions are significant, they are considerably worse than regressions obtained by DFA, AWC and R/S analysis. Moreover, it is not possible to identify the effect of the crossover even after the smoothing of power spectrum.

Table II presents estimates of the Hurst exponent obtained by DFA, R/S, AWC and Fourier transform methods for 1-minute and 5-minute time series.

TABLE II
ESTIMATES OF THE HURST EXPONENT

Method	$\delta = 1$	$\delta = 5$
DFA method	0.690	0.622
R/S analysis	0.631	0.558
AWC method	0.683	0.609
Fourier transform	0.609	0.56

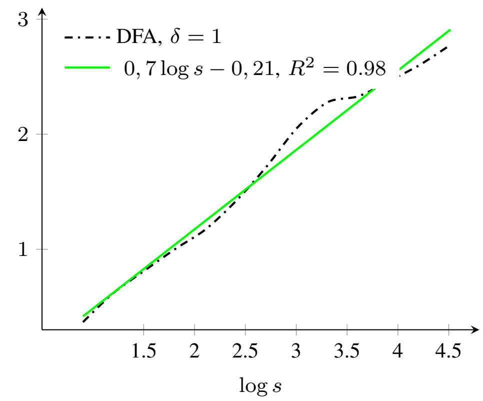

Fig. 2. $\log F(s)$ versus $\log s$ for the DFA estimation method, $\delta = 1$

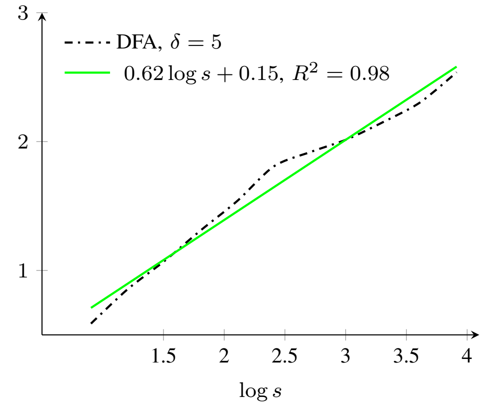

Fig. 3. $\log F(s)$ versus $\log s$ for the DFA estimation method, $\delta = 5$

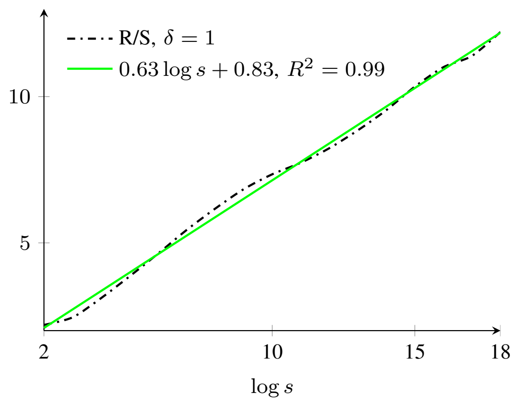

Fig. 4. $\log R/S$ versus $\log s$ for the R/S method, $\delta = 1$

V. CONCLUSION

In this paper we use the DFA, R/S analysis, AWC method and Fourier transform technique to examine the presence of long-range correlation of financial and economic time series of news sentiments. The results of this paper show that the long-range power-law correlation take place in time series of news sentiments. The paper shows that the behavior of long range dependence for time series of sentiment intensity most similar to the news flow intensity [17]. The results show that the self-similarity property is a stable characteristic of the sentiment of news information flow which serves the financial industry and stock markets. All four methods showed the presence of a long-range correlation in the time series of news sentiments (the Hurst exponent is greater than 0.5).

ISBN: 978-988-14048-3-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

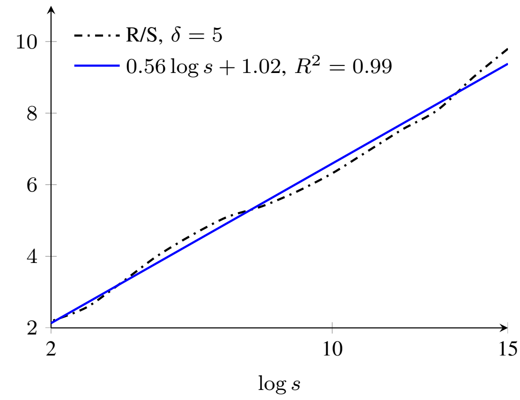

Fig. 5. $\log R/S$ versus $\log s$ for the R/S method, $\delta = 5$

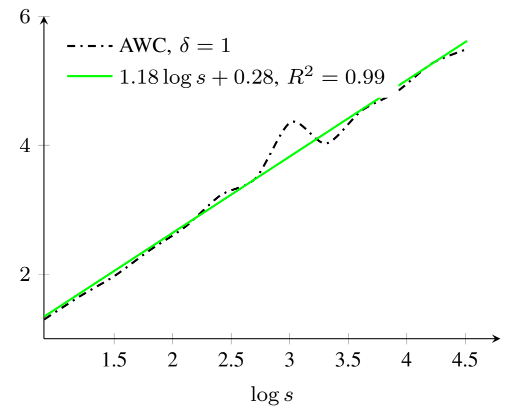

Fig. 6. $\log W[X](a)$ versus $\log a$ for AWC estimation method, $\delta = 1$

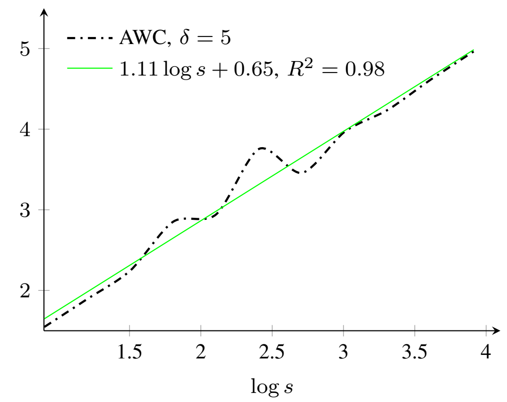

Fig. 7. $\log W[X](a)$ versus $\log a$ for AWC estimation method, $\delta = 5$

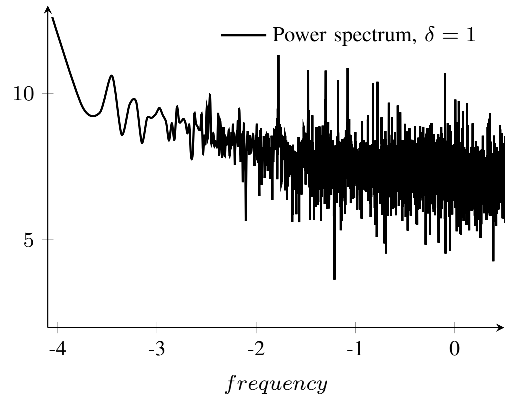

Fig. 8. Results for the Fourier transform method, $\delta = 1$

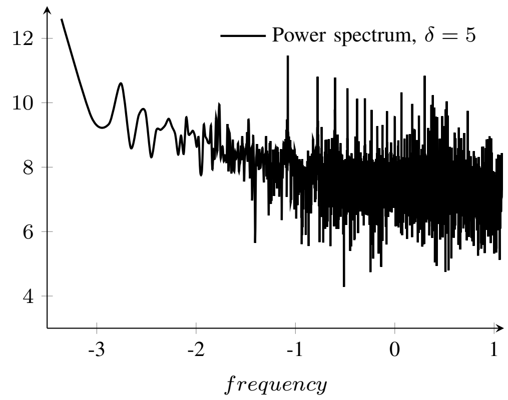

Fig. 9. Results for the Fourier transform method, $\delta = 5$

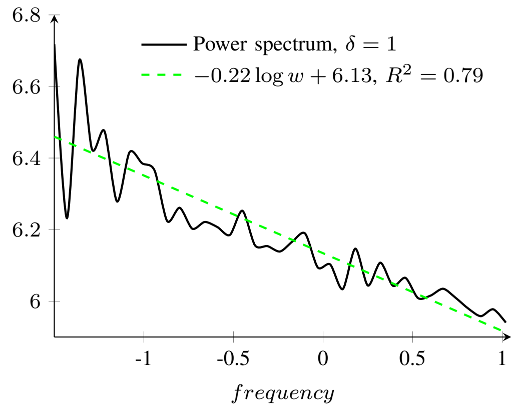

Fig. 10. Results for the Fourier transform method, $\delta = 1$

Results obtained by the DFA and AWC for sufficiently large scale have revealed the effect of crossover (corresponding to one day). Shortcomings of the classic R/S analysis and Fourier technique do not allow us to determine the effect of crossover.

References

[1] Y. Liu, P. Cizeau, M. Meyer, C.-K. Peng, and H. E. Stanley, "Correlations in economic time series," *Physica A: Statistical Mechanics and its Applications*, vol. 245, pp. 437–440, 1997.

[2] P. Cizeau, Y. Liu, M. Meyer, C.-K. Peng, and H. E. Stanley, "Volatility distribution in the S&P500 stock index," *Physica A: Statistical Mechanics and its Applications*, vol. 245, pp. 441–445, 1997.

[3] M. Ausloos, N. Vandewalle, P. Boveroux, A. Minguet, and K. Ivanova, "Applications of statistical physics to economic and financial topics," *Physica A: Statistical Mechanics and its Applications*, vol. 274, pp. 229–240, 1999.

[4] W. Willinger, M. S. Taqqu, and V. Teverovsky, "Stock market prices and long-range dependence," *Finance Stochast*, vol. 3, pp. 1–139, 1999.

[5] M. Ausloos and K. Ivanova, "Introducing false EUR and false EUR exchange rates," *Physica A: Statistical Mechanics and its Applications*, vol. 286, pp. 353–366, 2000.

[6] X. Lu, X. Sun, and J. Ge, "Dynamic relationship between japanese yen exchange rates and market anxiety: A new perspective based on mf-dcca," *Physica A: Statistical Mechanics and its Applications*, vol. 474, pp. 144–161, 2017.

[7] X. Zhang and W. Xiao, "Arbitrage with fractional gaussian processes," *Physica A: Statistical Mechanics and its Applications*, vol. 471, pp. 620–628, 2017.

[8] J. Kwapień and S. Drożdż, "Physical approach to complex systems," *Physics Reports*, vol. 515, no. 3–4, pp. 115–226, 2012.

[9] D. Stošić, D. Stošić, T. Stošić, and H. E. Stanley, "Multifractal analysis of managed and independent float exchange rates," *Physica A: Statistical Mechanics and its Applications*, vol. 428, pp. 13–18, 2015.

ISBN: 978-988-14048-3-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

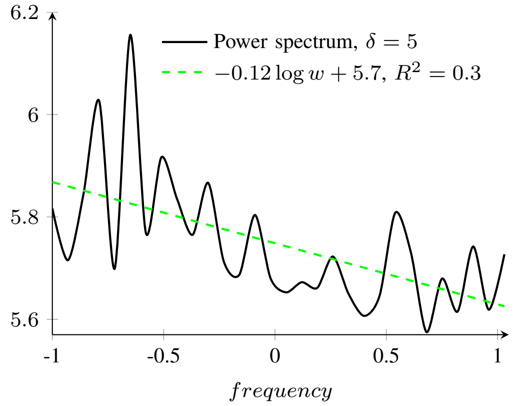

Fig. 11. Results for the Fourier transform method, $\delta = 5$

[10] P. Suárez-García and D. Gómez-Ullate, "Multifractality and long memory of a financial index," *Physica A: Statistical Mechanics and its Applications*, vol. 394, pp. 226–234, 2014.
[11] X. Lu, J. Tian, Y. Zhou, and Z. Li, "Multifractal detrended fluctuation analysis of the chinese stock index futures market," *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 6, pp. 1452–1458, 2013.
[12] W. Wang, K. Liu, and Z. Qin, *Multifractal Analysis on the Return Series of Stock Markets Using MF-DFA Method*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 107–115.
[13] ——, "Multifractal analysis on the return series of stock markets using mf-dfa method," in *International Conference on Informatics and Semiotics in Organisations*. Springer, 2014, pp. 107–115.
[14] C. Chen and Y. Wang, "Understanding the multifractality in portfolio excess returns," *Physica A: Statistical Mechanics and its Applications*, vol. 466, pp. 346–355, 2017.
[15] J. R. Thompson and J. R. Wilson, "Multifractal detrended fluctuation analysis: Practical applications to financial time series," *Mathematics and Computers in Simulation*, vol. 126, pp. 63–88, 2016.
[16] P. Mali and A. Mukhopadhyay, "Multifractal characterization of gold market: A multifractal detrended fluctuation analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 413, pp. 361–372, 2014.
[17] S. Sidorov, A. Faizliev, V. Balash, and E. Korobov, "Long-range correlation analysis of economic news flow intensity," *Physica A: Statistical Mechanics and its Applications*, vol. 444, pp. 205–212, 2016.
[18] S. Sidorov, A. Faizliev, and V. Balash, "Scale invariance of news flow intensity time series," *Nonlinear Phenomena in Complex Systems*, vol. 19, no. 4, pp. 368–377, 2016.
[19] G. Mitra and Y. Xiang, Eds., *The Handbook of Sentiment Analysis in Finance*. Albury Books, 2016.
[20] G. Mitra and L. Mitra, Eds., *The Handbook of News Analytics in Finance*. John Wiley & Sons, 2011.
[21] C. Peng, S. Buldyrev, S. Havlin, M. Simons, H. Stanley, and A. Goldberger, "Mosaic organization of DNA nucleotides," *Phys. Rev. E*, vol. 49, pp. 1685–1689, 1994.
[22] C. Peng, S. Buldyrev, A. Goldberger, S. Havlin, F. Sciortino, M. Simons, and H. Stanley, "Long-range correlations in nucleotide sequences," *Nature*, vol. 356, pp. 168–170, 1992.
[23] S. Buldyrev, A. Goldberger, S. Havlin, R. Mantegna, M. Matsa, C. Peng, M. Simons, and H. Stanley, "Long-range correlation properties of coding and noncoding DNA sequences: GenBank analysis," *Phys. Rev. E*, vol. 51, pp. 5084–5091, 1995.
[24] C. Peng, S. Buldyrev, A. Goldberger, R. Mantegna, M. Simons, and H. Stanley, "Statistical properties of DNA sequences," *Physica A*, vol. 221, pp. 180–192, 1995.
[25] S. Buldyrev, N. Dokholyan, A. Goldberger, S. Havlin, C. Peng, H. Stanley, and G. Viswanathan, "Analysis of DNA sequences using methods of statistical physics," *Physica A*, vol. 249, pp. 430–438, 1998.
[26] R.-G. Yeh, G.-Y. Chen, J.-S. Shieh, and C.-D. Kuo, "Parameter investigation of detrended fluctuation analysis for short-term human heart rate variability," *Journal of Medical and Biological Engineering*, vol. 30, pp. 277–282, 2010.
[27] R. Acharya, C. Lim, and P. Joseph, "Heart rate variability analysis using correlation dimension and detrended fluctuation analysis," *ITBM-RBM*, vol. 23, pp. 333–339, 2002.
[28] T. Penzel, J. W. Kantelhardt, L. Grote, J.-H. Peter, and A. Bunde, "Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea," *Trans. Biomed. Eng.*, vol. 50, pp. 1143–1151, 2003.
[29] P. Castiglioni, L. Quintin, A. Civijian, G. Parati, and M. D. Rienzo, "Local-scale analysis of cardiovascular signals by detrended fluctuations analysis: effects of posture and exercise," *Conf. Proc. IEEE Eng. Med. Biol. Soc. Ann.*, vol. 29, pp. 5035–5038, 2007.
[30] R. G. Yeh, G. Y. Chen, J. S. Shieh, and C. D. Kuo, "Detrended fluctuation analysis of short-term heart rate variability in late pregnant women," *Auton. Neurosci-Basic Clin.*, vol. 150, pp. 122–126, 2009.
[31] K. Fukuda, L. Amaral, and H. Stanley, "Dynamics of temporal correlation in daily internet traffic," *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, vol. 7, pp. 4069–4073, 1999.
[32] C.-K. Peng, J. Hausdorff, S. Havlin, J. Mietus, H. Stanley, and A. Goldberger, "Multiple-time scales analysis of physiological time series under neural control," *Physica A: Statistical Mechanics and its Applications*, vol. 249, pp. 491–500, 1998.
[33] J. Hausdorff, S. Mitchell, R. Firtion, C.-K. Peng, M. Cudkowicz, J. Wei, and A. Goldberger, "Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntingtons disease," *J. Appl. Physiology*, vol. 82, pp. 262–269, 1997.
[34] R. Bryce and K. Sprague, "Revisiting detrended fluctuation analysis," *Scientific Reports*, vol. 2:315, 2012.
[35] Y.-H. Shao, G.-F. Gu, Z.-Q. Jiang, W.-X. Zhou, and D. Sornette, "Comparing the performance of FA, DFA and DMA using different synthetic long-range correlated time series," *Scientific Reports*, vol. 2:835, 2012.
[36] M. S. Taqqu, V. Teverovsky, and W. Willinger, "Estimators for long-range dependence: An empirical study," *Fractals*, vol. 3, pp. 785–798, 1995.
[37] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. Rego, S. Havlin, and A. Bunde, "Detecting long-range correlations with detrended fluctuation analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 295, pp. 441–454, 2001.
[38] B. Mandelbrot and J. Wallis, "Computer experiments with fractional gaussian noises. parts 1,2,3," *Water Resour. Res.*, vol. 5, pp. 228–267, 1969.
[39] ——, "Robustness of the rescaled range R/S in the measurement of noncyclic long-run statistical dependence," *Water Resour. Res.*, vol. 5, pp. 967–988, 1969.
[40] ——, "Some long-run properties of geophysical records," *Water Resour. Res.*, vol. 5, pp. 321–340, 1969.
[41] B. Mandelbrot and M. Taqqu, "Robust R/S analysis of long-run serial correlation," in *Proc. of the 42nd Session of the Int. Statistical Institute*, vol. Bull. Int. Statist. Inst. 48 (Book 2), Manila, 1979, pp. 69–104.
[42] J. Bassingthwaighte and G. Raymond, "Evaluating rescaled range analysis for time series," *Ann. Biomed. Eng.*, vol. 22, pp. 432–444, 1994.
[43] "A critical look at lo's modified R/S statistic," *Journal of Statistical Planning and Inference*, vol. 80, no. 1–2, pp. 211–227, 1999.
[44] J. Feder, *Fractals*. New York: Plenum Press, 1988.
[45] A. R. Mehrabi, H. Rassamdana, and M. Sahimi, "Characterization of long-range correlations in complex distributions and profiles," *Phys. Rev. E*, vol. 56, pp. 712–722, 1997.
[46] I. Simonsen, A. Hansen, and O. M. Nes, "Determination of the hurst exponent by use of wavelet transforms," *Phys. Rev. E*, vol. 58, pp. 2779–2787, Sep 1998.
[47] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia: SIAM, 1992.
[48] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*. Cambridge: Cambridge University Press, 2000.
[49] D. Grech and Z. Mazur, "On the scaling ranges of detrended fluctuation analysis for long-term memory correlated short series of data," *Physica A*, vol. 392, pp. 2384–2397, 2013.
[50] ——, "Scaling range of power laws that originate from fluctuation analysis," *Phys. Rev. E*, vol. 87, p. 052809, May 2013.
[51] ——, "Impact of scaling range on the effectiveness of detrending methods," *Acta Physica Polonica A*, vol. 127, no. 3-A, pp. 59–65, 2015.
[52] I. Simonsen, "Measuring anti-correlations in the nordic electricity spot market by wavelets," *Physica A: Statistical Mechanics and its Applications*, vol. 322, pp. 597–606, 2003.

ISBN: 978-988-14048-3-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)