

## What can wavelets reveal about SOFIX?

B. Lomev\*, I. Ivanov, and B. Bogdanova

Dep. of Statistics and Econometrics, University of Sofia, 125 Tsarigradsko chausse blvd., Sofia, Bulgaria.

### Abstract

The current paper investigates the major index of the Bulgarian Stock Exchange with respect to the presence of long-range dependence and principal predictability of the index. The wavelet transform is utilized in order to carry out the investigation since it is a well-suited tool for the analysis of fractal processes and sheds additional light to the term structure of the index.

*Keywords:* Wavelet transform, long-range dependence, prediction, market inefficiency.

### 1. Introduction

In recent papers Lomev and co-writers [1-3] confirm the presence of long-range dependence (LRD) of SOFIX on other East-European stock exchange indexes. The presence of LRD is closely related to persistency, that is, positive values are likely to be followed by positive values and negative values are likely to be followed by negative values. The latter implies that the chance to predict correctly the future direction of the process is greater than 50%. Actually, the problem of predictability of SOFIX has been already investigated by Lomev and Ivanov [4] and it turns out that the index is predictable. From this point of view the current paper utilizes the wavelet transform in two directions. Firstly, LRD presence for SOFIX is detected. Secondly, wavelet-based forecasts are developed for the index, in order to test it for principal predictability and to investigate its term structure.

A good reasoning behind the application of the wavelet transform is that it is a proper tool for the analysis of financial data. On one hand financial data incorporates information about decisions and actions taken by market participants, operating over different time horizons and on the other hand the wavelet transform of some series decomposes the original sequence over a range of frequency scales. The wavelet transform may be continuous or discrete; however for empirical purposes the discrete wavelets transform (DWT) is preferred. Different algorithms are developed for the implementation of DWT, each one having its advantages and drawbacks. A detailed discussion may be found for example in [5], [6]. The current investigation utilizes the triangle algorithm of Mallat [7] for the detection of LRD and the  $\acute{a}$  trous wavelet transform [6] for prediction purposes.

### 2. Long-range dependence

Long-range dependence is one of the important forms of

scale invariance. A long-range dependent process is characterized by a very slowly decaying autocorrelation function. Such processes are properly modeled as FARIMA(p,d,q) processes, introduced by Granger [8] and Hosking [9] which are defined as follows:

$$\Phi(B)\nabla^d X_t = \Theta(B)Z_t, \quad (1)$$

where

$$\nabla^d = (1-B)^d = \sum_{k=0}^d \binom{d}{k} (-1)^k B^k \quad (2)$$

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p, \quad (3)$$

$$\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q,$$

$B$  is the backshift operator and  $Z_t$  is a discrete white noise. LRD is present when the fractional differencing parameter  $d \in (0, 1/2)$ .

For the estimation of the fractional differencing parameter are developed a lot of methods based on frequency analysis. The method used in the current paper is based on wavelet analysis since the wavelet transform is a natural tool for studying scale invariant processes. The triangle algorithm of Mallat is applied. It filters the original data series  $\mathbf{X} = (X_0, X_1, \dots, X_{N-1})$  using a pair of high-pass and low-pass filters denoted respectively as  $\mathbf{h} = (h_0, h_1, \dots, h_{L-1})$  and  $\mathbf{g} = (g_0, g_1, \dots, g_{L-1})$ , each of length  $L$ ,  $L < N$ . The wavelet and scaling coefficients corresponding to the first level of decomposition ( $j=1$ ) are obtained via

$$w_{1,t} = \sum_{l=0}^{L-1} h_l X_{2t+1-l \bmod N} \quad (4)$$

$$c_{1,t} = \sum_{l=0}^{L-1} g_l X_{2t+1-l \bmod N} \quad (5)$$

\* E-mail address: b.museva@law.uni-sofia.bg

where  $t=0,1,\dots,N/2-1$ . It should be noted that a downsampling operation is involved. This can be seen from the subscript of  $\mathbf{X}$  in Eq.4 and Eq.5. More generally, the formula for the wavelet and scaling coefficients corresponding to the  $j^{\text{th}}$  level of decomposition of  $\mathbf{X}$  at scale  $2^{-j}$  is

$$w_{j,t} = \sum_{l=0}^{L-1} h_l c_{j-1,2t+1-l \bmod N} \tag{6}$$

$$c_{j,t} = \sum_{l=0}^{L-1} g_l c_{j-1,2t+1-l \bmod N} \tag{7}$$

for  $j=2,\dots,J$  and  $t=0,1,\dots,N/2^j-1$ .

The output of the  $j^{\text{th}}$  level of decomposition is a vector of wavelet coefficients  $\mathbf{w}_j = (w_{j,0}, w_{j,1}, \dots, w_{j,N/2^j-1})$  and a vector of scaling coefficients  $\mathbf{c}_j = (c_{j,0}, c_{j,1}, \dots, c_{j,N/2^j-1})$ , both of which are of length  $N_j = N/2^j$ ,  $j=1,\dots,J$ . In this application the finest scale (high frequency) corresponds to  $j=1$  and we will further refer to  $j$  as octave.

An important feature of the Mallat algorithm is that it is a decimated, orthogonal transform but it is not shift-invariant. Such a decomposition decorrelates the complicated structure of long-range dependent processes, which makes it suitable for estimation of the fractional differencing parameter. The estimation method that is to be presented follows closely the article of Arby and co-writers [10]. They investigate the relationship between the variance of each wavelet sequence  $\mathbf{w}_j$  to the octave  $j$  for scale invariant processes and define the exact log-scale diagram as the plot of

$$s_j = \log_2(\text{var}(\mathbf{w}_j)) \tag{8}$$

against the octave  $j$  in the interval  $[j_1, j_2]$  for which scaling behavior is present. The index  $j_1$  is an integer and is called the high frequency cutoff and the index  $j_2$  is also an integer, called low frequency cutoff. In this interval the plot will be a straight line with a slope coefficient

$$\gamma = 2d \tag{9}$$

It is obvious that from Eq.9 the estimate of the fractional differencing parameter can be derived from the estimate of the slope coefficient  $\gamma$ . Since  $s_j$  is not known an unbiased estimate of it is given by  $y_j$ , defined as:

$$y_j = \log_2(\mu_j) - g(j), \tag{10}$$

where

$$\mu_j = \frac{1}{N_j} \sum_{t=0}^{N_j-1} |w_{j,t}|^2 \tag{11}$$

and

$$g(j) \sim \frac{-1}{N_j \ln 2}. \tag{12}$$

Generally to the plot of  $y_j$  against the octave is referred to as the logscale diagram. An unbiased estimator of the slope coefficient  $\gamma$  is obtained through linear regression over  $y_j$  in the range  $[j_1, j_2]$ . The variance of  $y_j$  is denoted by  $\sigma_j^2$  and can be approximated as:

$$\sigma_j^2 \sim \frac{2}{N_j \ln^2 2} \tag{13}$$

It is reasonable to adopt weighted linear regression in order to diminish the variance of the estimate and thus the unbiased estimate of the slope coefficient is given by:

$$\hat{\gamma} = \frac{\sum_{j=j_1}^{j_2} y_j (j S - S_1) / \sigma_j^2}{SS_2 - S_1^2}, \tag{14}$$

where

$$S = \sum_{j=j_1}^{j_2} 1 / \sigma_j^2, S_1 = \sum_{j=j_1}^{j_2} j / \sigma_j^2, S_2 = \sum_{j=j_1}^{j_2} j^2 / \sigma_j^2.$$

From Eq.9 and Eq.14 the estimate of the fractional differencing parameter is:

$$\hat{d} = \frac{\hat{\gamma}}{2} \tag{15}$$

The cutoff  $j_1$  is empirically selected. One simple choice is to identify the point from which the line gets straight. Another option is to perform Monte Carlo experiments and to choose that value of  $j_1$  for which the mean squared error is minimal. We performed such experiments over 4500 simulated series of fractional Gaussian noise with  $d \in (0,0.5)$ , each of length 2048. The results show that if the true parameter value is in the range  $(0,0.20]$ , then the optimal choice is  $j_1 = 1$ . When the true value is in the range of  $(0.20,0.50)$ , then  $j_1 = 2$ . Theoretically,  $j_2$  is chosen to be as large as possible. Practically, it is prescribed that

$$j_2 = \lceil \log_2 N - Const \rceil, \tag{16}$$

where  $Const$  is a constant and when Daubechies wavelets are used, its value is  $\log_2(2n+1)$ , where  $n$  is the number of vanishing moments.

The raw data consists of the daily close values of SOFIX for the period between Jan, 1, 2002 and Feb, 28, 2011. The data is transformed into logarithmic returns. For the purposes of our investigation, the Daubechies low-pass and high-pass filters, as derived from the Daubechies wavelet function with 3 vanishing moments [11], are applied. From Eq.16 the level of decomposition was determined to be 8. Following Eq.10, we evaluated  $y_j$ ,  $j=1,\dots,8$ , and the log-scale diagram is presented in Fig.1. It can be seen that the

plotted values are scattered relatively close to a straight line; it is difficult to determine the value of  $j_1$ , though. Considering the results of the numerical experiments, the fractional differencing parameter is estimated from Eq.15 for  $j_1=1$  and  $j_1=2$  and the estimates are presented in Table 1. The results clearly show that there is a long-range dependence. Similar to those results are obtained by utilizing the Whittle method.

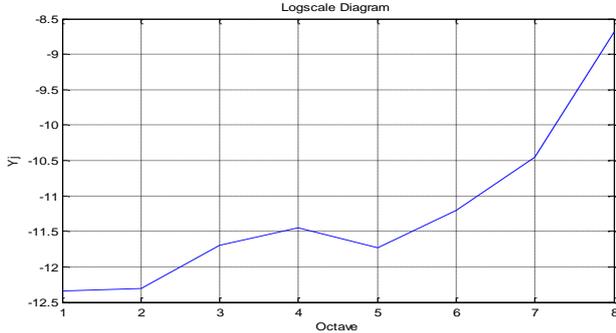


Fig. 1. Log-scale Diagram

Table 1. Estimates of the fractional differencing parameter for SOFIX

	$j_1 = 1$	$j_1 = 2$
$\hat{d}$	0.1447	0.1804

3. Predictability

For prediction purposes the so called “á trous” algorithm is used. Contrary to the Mallat algorithm, the “á trous” algorithm is a non-orthogonal, redundant, oversampled transform, but it is shift-invariant and hence, well-suited for forecasting purposes. In the current paper the Haar “á trous” wavelet transform, introduced by Zheng and co-writers [12]

is applied; utilizing the simple low-pass filter  $g = \left(\frac{1}{2}, \frac{1}{2}\right)$ . It decomposes the original sequence  $\mathbf{X} = (X_0, \dots, X_{N-1})$  as a superposition of the form:

$$X_t = c_{J,t} + \sum_{j=1}^J w_j, \tag{17}$$

$t = 0, 1, \dots, N-1$ . By  $\mathbf{c}_j = (c_{j,0}, c_{j,1}, \dots, c_{j,N-1})$  is denoted a smooth version of the original signal  $\mathbf{X}$  and  $\mathbf{w}_j = (w_{j,0}, w_{j,1}, \dots, w_{j,N-1})$  represents the detail of  $\mathbf{X}$  at scale  $2^{-j}$ ,  $j = 1, \dots, J$ . The creation of the first wavelet resolution level is derived from the convolution of the input data with  $g$ . For the  $j^{\text{th}}$  level of decomposition we have:

$$c_{j,t} = \frac{1}{2} (c_{j-1,t-2^{j-1}} + c_{j-1,t}) \tag{18}$$

and

$$w_{j,t} = c_{j-1,t} - c_{j,t}, \tag{19}$$

for  $t = 0, \dots, N-1$ . The sequence of detail and smooth coefficients are given by  $N$ -length vectors for every scale of decomposition.

The prediction model that is to be used is developed by Renaud and co-writers [13]. They present the multiscale autoregressive (MAR) model and show in a simulated study that the MAR technique outperforms the ordinary autoregressive models when the process of interest is long-range dependent. The one step ahead forecast for a MAR model is calculated as follows:

$$\hat{X}_N = \sum_{j=1}^J \sum_{k=1}^{A_j} \hat{a}_{j,k} w_{j,N-1-2^j(k-1)} + \sum_{k=1}^{A_{J+1}} \hat{a}_{J+1,k} c_{J,N-1-2^j(k-1)} \tag{20}$$

In this model there are  $Q = \sum_{j=1}^{J+1} A_j$  unknown parameters

which are estimated by least squares regression.

One of the attractive features of this forecasting technique is that it is parsimonious and preserves just as many terms as necessary. The MAR type of models are very flexible tool for data analysis since when the original data shows high frequency variation  $A_j$  is selected to be greater for the small values of  $j$  and is set to zero or one for larger  $j$ . The converse is true in the case of  $1/f$  behavior.

We have already confirmed that SOFIX is long-range dependent so now it is interesting to see whether the latter statement is true. The data consists again of the logarithmic return series of SOFIX for the same period. To the data is applied the Haar “á trous” transform following Eq.18 and Eq.19 for  $J = 4$ . The first 1500 observations (approximately 2/3 of the data) are used to estimate the parameters of the model, the rest of the observations are retained for test purposes. As a starting point is considered the MAR(1) model which for short will be referred to as M1 model and its general form is given by Eq.21:

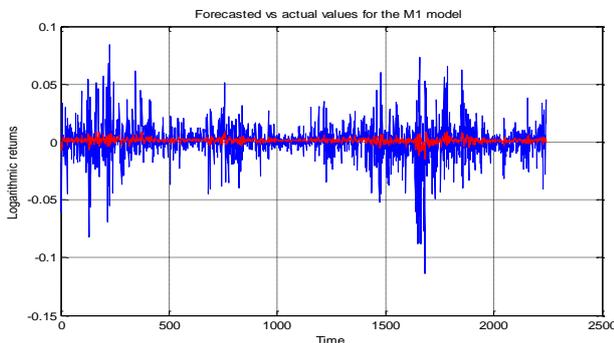
$$X_t = \sum_{j=1}^4 a_j w_{j,t-1} + a_5 c_{4,t}, \tag{21}$$

It turns out that the estimated coefficient corresponding to the finest detail series  $w_{1,t-1}$  is insignificant. This was anyway an expected result. More surprising is the fact that the coefficient corresponding to the coarsest detail series  $w_{4,t-1}$  is also insignificant. Thus another model is estimated, further referred to as M2 model, for which the finest and coarsest detail series are skipped, i.e. only the series  $w_{2,t-1}, w_{3,t-1}$ , and  $c_{4,t-1}$  are retained as regressors. The information criteria of Akaike and Schwarz for this model are improved, compared to M1 model. This implies that the moderate frequency scales influence the index future values. Besides these two models, many other alternative models were fitted, but none of them performed better, in terms of the information criteria. In addition, for each detail series  $\mathbf{w}_j, j = 1, \dots, 4$ , and for the smooth series  $\mathbf{c}_4$  individual ARMA models were fitted and thus one step ahead individual forecasts were calculated. The forecast of the logarithmic return series was derived as the sum of the individual predictions following Eq. 17 and it will be further referred to this model as M3 model.

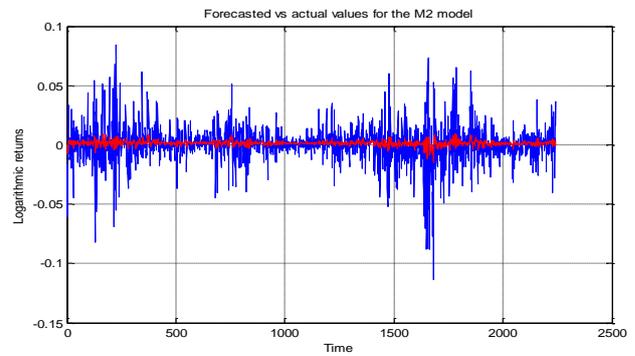
**Table 2.** Forecasting errors and directional symmetry

	MSE	DS
Zero forecast	3.3258e-004	-
M1 model	3.2868e-004	53.11%
M2 model	3.2844e-004	52.70%
M3 model	3.3225e-004	51.55%

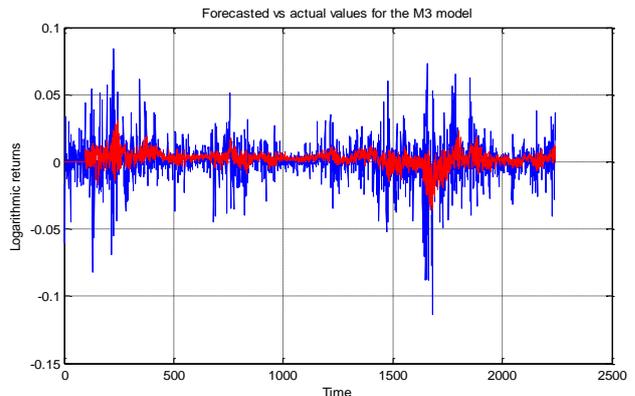
Table 2 presents the mean squared errors (MSE) of the one step ahead forecasts over the test set for the three models just discussed. Also the MSE for the zero forecast is given since this is the best forecast, if the index prices are following random walk process without drift. For all of the models the calculated MSE is smaller than the MSE of the zero forecast which confirms the principal predictability of SOFIX. The minimal MSE is for M2 model. The last column of Table 2 (denoted by DS) gives the directional symmetry, i.e. the percentage of correctly predicted directions in respect to the index. It is interesting that for all of the three models, the directional symmetry is greater than 50% and the highest value was achieved for the first model. Fig.2, Fig.3 and Fig.4 present the forecasted series by M1, M2, and M3 models, respectively, versus the logarithmic returns of SOFIX; the red lines being the forecasted series.



**Fig. 2.** One step ahead forecasts obtained by the M1 model versus actual values



**Fig. 3.** One step ahead forecasts obtained by the M2 model versus actual values



**Fig. 4.** One step ahead forecasts obtained by the M3 model versus actual values

#### 4. Conclusions

The current investigation confirms the presence of LRD for SOFIX and the index turns out to be principally predictable, which is indicative for market inefficiency. An interesting finding is that the future values of SOFIX are affected mostly by the detail series corresponding to moderate frequency scales.

#### References

1. Б.М.Ломев, И.Г.Иванов, Б.Ц.Богданова, Изследване на перфектността на възникващи фондови пазари от източноевропейския регион посредством дискретна трансформация с вълнички, Научни трудове на Русенския университет, 49, 5.1, 194 (2010).
2. Б.М.Ломев, И.Г.Иванов, Б.Ц.Богданова, Изследване на перфектността на фондовите борси в България и Русия посредством честотни методи, Шеста научна конференция „Мениджмънт и предприемачество“, Пловдив, Сборник доклади, 95 (2009).
3. Б.М.Ломев, И.Г.Иванов, Е.Туйова. Дългосрочна зависимост при индекса SOFIX на Българската Фондова Бора. Научни трудове на русенския университет, 47, 6.1 (2008).
4. B.M.Lomev and I.G.Ivanov., Self-similarity, Heavy Tails and Long-range Dependence as Measures for Financial Market Inefficiency-the Case of Bulgaria, the 9th Hellenic European Research on Computer Mathematics and its Applications Conference, Athens, (2009) (a conference talk).
5. R. Gençay, F.Selçuk, and B.Whitcher. An Introduction to Wavelets and Other Filtering Methods in Finance and Economics. Academic Press, New York, (2002).
6. D. B. Percival and A. T. Walden, Wavelet Methods for Time Series Analysis, Cambridge Press, Cambridge, (2000).
7. S. Mallat, A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11, 674 (1989).
8. C. W. Granger and R. Joyeux , An Introduction to Long-memory Time Series Models and Fractional Differencing, J.Time Series Analysis, 1, 15 (1980).
9. J. R. M. Hosking, Fractional differencing, Biometrika, 68 165 (1981).
10. P. Arby, P. Flandrin, M. S. Taqqu and D. Veitch, Self-similarity and long-range dependence through the wavelet lens, Theory and applications of long-range dependence, Birkhäuser, 527 (2003).
11. I. Daubechies, Ten Lectures on Wavelets, Vol. 61 of CBMS-NSF Regional Conference Series in Applied mathematics, Society for Industrial and Applied Mathematics, Philadelphia, (1992).
12. G. Zheng, J. L .Starck, J. G. Campbell and F. Murtagh, Multiscale transforms for filtering financial data streams, Journal of Computational Intelligence in Finance, 1999, 7, 18-35.
13. O. Renaud, J.-L. Starck, and Murtagh, F., Wavelet-based forecasting of short and long memory time series, Department d'Econometrie, University of Genève, Tech. Rep.2002.04, [http://www.unige.ch/ses/metri/cahiers/2002\\_04.pdf](http://www.unige.ch/ses/metri/cahiers/2002_04.pdf).