FETDE 2020
International Conference on Finance, Entrepreneurship and Technologies in Digital Economy

FORECASTING AND STOCK MARKET CRITICAL POINTS ANALYSIS USING MODIFIED LOCAL HOLDER EXPONENTS

D. V. Kurganov (a)*, Yu. A. Kuperin (b), L. A. Dmitrieva (c), G. A. Chernykh (d), A. Heut (e)
*Corresponding author

(a) Limited liability company "Digitalist", Saint Petersburg, Russia, kyrganov@ya.ru
(b) Saint Petersburg State University, Saint Petersburg, Russia
(c) Saint Petersburg State University, Saint Petersburg, Russia
(d) Saint Petersburg State University, Saint Petersburg, Russia
(e) Joint-stock company "Gazprombank", Moscow, Russia, okolekk@gmail.com

Abstract

The article is devoted to a new indicator for forecasting critical points of financial time series based on modified Hölder indicators. The indicator was developed to predict large movements of financial instruments in the stock market. The analysis of the indicator's performance was conducted on the US and Russian stock markets using time series with a minute sampling frequency. It is shown that this indicator is able to predict large movements of financial time series with good enough statistics. The corresponding calculations, tables and statistics are presented. The paper shows that the developed predictor on average in 80% of cases in the US market and on average in 60% of cases on the Russian market correctly predicts large movements in the market. These results were obtained by statistical processing of all predictions of critical points in the markets of the USA and Russia. Also, a significant difference was found between the parameters of the developed indicator for the US and Russian markets.

Keywords: Forecasting, time series, Hölder indicators, local Hölder exponents

1. Introduction

The idea of predicting the dynamics of the prices of financial instruments in the stock and currency markets with the help of a mathematical approach arose long ago. To solve this problem researchers apply the whole variety of mathematical methods, beginning with simple linear models and ending with models of recurrent neural networks and hidden Markov processes (Khan & Gour, 2013). In particular, one of the possible methods for analyzing and predicting the dynamics of the stock market is based on the multifractal approach (Suárez-García & Gómez-Ullate, 2014; Kapecka, 2013; Kuperin & Schastlivtsev, 2008). In this paper we use the notion of the local Hölder exponents (LHE), which are closely related to multifractal analysis, which can be found, for example, in (Kuperin & Schastlivtsev, 2008). More specifically, LHE are defined below. Definition. Let a function $f(t)$ be defined on a domain $t \subset R$ and satisfy the relation

$$\left| f\left(t + \Delta t - f\left(t \right) \right) \right| : \ C_t \cdot \Delta t^{\alpha(t)}$$

at $\Delta t \to 0$, $0 < \alpha(t) \le 1$ and $C_t > 0$. Then the number $\alpha(t)$ is called the local Hölder exponent of the function $f(t)$ at the point t. Local Hölder exponents show how smooth the function is at a given point, in other words, the higher the value $\alpha(t)$, the higher the smoothness of the function $f(t)$. The method described in this article is based on the assumption that before a strong price movement in the market, the time series of prices is smoothed (Kuperin & Schastlivtsev, 2008). Thus, LHE allow us to estimate the smoothness of the series and make a forecast about the subsequent strong price change. There are two basic approaches to the calculation of local Holder exponents, which have a number of advantages and disadvantages (Abry et al., 2009). We will describe them briefly below. Method of oscillations. It is one of the simplest methods for computing local Hölder exponents (Legrand et al., 2006; Trujillo et al., 2010). The drawback of this approach lies in the slow convergence of the values of the computed local Hölder exponents to their theoretical values. In other words, in practice, it is necessary to use large neighborhoods of a time series point   to obtain good-accuracy LHE values. Wavelet transform method. This method of calculating Hölder exponents is one of the most frequently used methods (Los & Yalamova, 2004; Struzik, 2001a,b). This method can be based on both a continuous and discrete wavelet transform (Bacry et al., 2002). The drawback of the method is the presence of the so-called edge effect, which appears when the wavelet transform is implemented. The edge effect for finite time series is manifested in the fact that changing at least one value at the beginning or at the end of the time series leads to a strong change in the wavelet transform. In this paper, LHE are calculated by the modified local Hölder exponents (MLHE) method, proposed and described in detail in (Kuperin & Schastlivtsev, 2008). Advantages of this method in comparison with the methods described above consist in the absence of an edge effect and in the rapid convergence of the computed values to the corresponding theoretical values. It should be noted that in this paper the method of calculating MLHE in comparison with the computational procedure proposed in (Kuperin & Schastlivtsev, 2008) has been changed. Namely, we do not use the notion of a signal line. proposed in (Kuperin & Schastlivtsev, 2008), and we perform MLHE smoothing not with the help of the moving average, but with the help of the Hodrick-Prescott filter (Cogley & Nason, 1995).

2. Basic concepts

The method of MLHE allows to determine changes in the dynamics of financial time series, characterized by strong amplitude motions of the series. Let the amplitude of the MLHE be determined by the difference between the maximum and minimum.

Definition 1. We call the MLHE signal the amplitude of the MLHE exceeding a certain threshold value s, the value of which is determined by numerical experiment. We will assume that the signal appears at the maximum t_s of the amplitude of the MLHE. The point in time, in which a large motion of the time series begins, will be called the critical point of the time series.

An important step in calculating the statistics of the MLHE method is the definition of the term "movement of the financial time series". Obviously, depending on what definition to adopt, different results will be obtained. For objectivity and greater coverage of possible ways of applying the method of MLHE in this paper, it is proposed to use several variants of such definitions at once and to count statistics according to each of the definitions separately. Also in options trading often rely on the value of volatility at each point in time, as its change leads to a change in the value of the option. Therefore, it is reasonable to make calculations regarding volatility. Below are the definitions used in calculating statistics. We introduce the main definitions:

Definition 2. We define the motion $\Delta H(\varepsilon)$ of a time series over a certain time interval ε as the difference between the values of the final and the first points of this interval.

Definition 2.1. We define a large (significant) motion $\Delta H(\varepsilon)$ in a certain time interval ε as a motion (see definition 2) whose value exceeds the mean value of motion on the same time interval for a given time series.

Definition 2.2. We define a large (significant) motion $\Delta H(\varepsilon)$ as a motion (see definition 2) whose magnitude exceeds the median value of the motion for a given time series.

Definition 2.3. We define a large (significant) motion $\Delta H(\varepsilon)$ of a time series in a certain time interval ε as the maximum difference of values (in modulus) between points inside the interval ε and the first point of the interval.

Definition 3.1. We define a large (significant) motion of a time series as a motion $\Delta H(\varepsilon)$, the value of which exceeds the mean value of motion for a given time series.

Definition 3.2. We define a large (significant) motion as a motion $\Delta H(\varepsilon)$ whose value exceeds the median value of the motion for a given time series.

Definition 4. We determine the volatility V of the time series over a certain time interval ε as the standard deviation of the logarithmic increment of the instrument X under study on the interval ε (1).

Definition 4.1. We define large (significant) volatility as volatility, the value of which exceeds the average value of volatility for a given time series:

$$V = \sqrt{\frac{1}{1-\varepsilon}\sum_{i=1}^{\varepsilon}\left(\ln\frac{X_i}{X_{i-1}} - \left\langle\ln\frac{X_i}{X_{i-1}}\right\rangle\right)^2}, \tag{1}$$

where $\left\langle\ln\dfrac{X_i}{X_{i-1}}\right\rangle$ this averaged over all values of the series $\ln\dfrac{X_i}{X_{i-1}}$.

Definition 5. Let us determine the average value of the motion $\Delta H(\varepsilon)$ by the formula:

$$\langle\Delta H\rangle = \frac{1}{n(\varepsilon)}\sum_{i=1}^{n(\varepsilon)}\Delta H(\varepsilon_i), \tag{2}$$

where ε_i are disjoint time intervals of equal length ε, and $n(\varepsilon) = \left[\dfrac{N}{\varepsilon}\right]$ this is the number of intervals

divided by the time series, N this is the length of the time series. Here $\left[\dfrac{N}{\varepsilon}\right]$ is the integer part of the

number $\dfrac{N}{\varepsilon}$.

We define the median value of the motion $Me\left[\Delta H(\varepsilon)\right]$ as the mean value of an ordered series

of motions $\Delta H(\varepsilon_i)$ of the time series, where ε_i are disjoint time intervals of equal length ε,

$n(\varepsilon) = \left[\dfrac{N}{\varepsilon}\right]$ the number of intervals divided by the time series, N is the length of the time series.

The algorithm for calculating the predictor of large movements for financial instruments is implemented in several steps:

1. Calculation of the MLHE of the time series under study;

2. Smoothing of MLGP by Hodrick-Prescott filter (Cogley & Nason, 1995) filter, where the trend curve is used as a smoothing curve.

3. For each MLHE signal, the value of the change in the time series over the interval $\left[t_s, t_s + \varepsilon\right]$ introduced in Definition 2 and Definition 3.

4. If the value of movement is large (significant), then the signal is considered correct, otherwise the signal is considered incorrect.

The motion is considered large according to the definition with respect to which the calculations are made. These calculations were carried out for all definitions of large motions (listed above) in time series.

As a mechanism to control the objectivity of the results for each time series, the number of correct and incorrect signals obtained as a result of random numerical generation.

By random generation we mean the following. Imagine that we have received a proportion of the correct signals of the MLHE equal to $P.$ Now it is necessary to show that this proportion is not accidental. That is, if we choose random time points (random signals) instead of the MLHE signals on the same time series, then the proportion of the correct signals for random selection should be lower (statistically) than for the signals of the MLHE. Random numerical generation is the receipt of random points in the time

domain of the time series, which we apply to the algorithm "as if" they are signals of large movements. If the proportion of the correct signals with random signals would be the same as for signals of the MLHE, then the signals of the MLHE would have no value, i.e. themselves would be accidental and not capable of predicting large movements. The random generation of signals (RSG) allows us to show what proportion of the "correct signals" can be obtained on the time series under study if we calculate the motions of this series at random points. To calculate the optimal parameters of the MLHE, a simple optimizer is used that sorts out all possible combinations of parameters in a given range with a given step for each parameter. This algorithm was used to calculate the 13 most liquid shares in the US market, as well as 5 liquid shares of the MICEX stock exchange for the period from 01.01.15 to 17.08.15 for the minute sampling rate of time series. In this paper, only data with a high sampling rate (minute values) is used. This is due to the fact that for statistical analysis of data it is necessary to obtain a large number of signals, that is, series containing about 50,000 values and higher are needed. Such a number of values in the public domain is available only for the minute sampling frequency, that is, series containing about 50,000 values and higher are needed. Optimization is carried out in three parameters: the MLHE window w (the number of points in the series necessary to calculate one value of the MLHE), the interval ε and the threshold s above which the amplitude of the MLHE is considered to be a signal amplitude. Intervals and optimization steps are listed below:

1. window MLHE w : interval [200,350], step is 30 points (such a short interval and a large step are chosen because of the large time for calculating the MLHE about 500 minutes into one time series);

2. time interval ε : interval [20,220], step is 10 points;

3. threshold value for the amplitude of the MLHE: interval [0.1, 4], step 0.1 (measured with respect to the mean amplitude, i.e. 0.2 means 0.2 of the mean).

Below is a brief description of the optimization algorithm and the objective function.

We divide the original time series into two consecutive samples: primary (30% of the whole series) and secondary (70% of the whole series). On both samples, for all sets (w, s, ε) of parameters, we calculate the proportions $P(w, s, \varepsilon)$ of the correct signals. Let $\overline{P}(w, s, \varepsilon)$ is the proportion of correct MLHE signals received on the primary sample with the values of the MLHE window w, the signal threshold s and the time interval ε. Likewise let be $\hat{P}(w, s, \varepsilon)$ is the proportion of the correct MLHE signals received on the secondary sample. We define an objective function as follows:

$$F(w, s, \varepsilon) = \Delta P(w, s, \varepsilon) \cdot \hat{P}(w, s, \varepsilon), \quad (3)$$

where $\Delta P = \left| \hat{P}(w, s, \varepsilon) - \overline{P}(w, s, \varepsilon). \right|$ Then the optimal values of for w, s, ε can be determined from condition

$$F(w^*, s^*, \varepsilon^*) = \max_{\hat{P}} \min_{\Delta P} \left[\Delta P(w, s, \varepsilon) \cdot \hat{P}(w, s, \varepsilon) \right], \quad (4)$$

where w^*, s^*, ε^* are the optimal parameters. In other words, the optimal parameters are those parameters for which the values of $\overline{P}(w, s, \varepsilon)$ and $\hat{P}(w, s, \varepsilon)$ differ as little as possible, and $\hat{P}(w, s, \varepsilon)$ has value as large as possible. Such an optimization algorithm is necessary to obtain a stable parity of

parameters (w, s, ε), in which the proportion of correct signals as little as possible would change with time. Otherwise, it may turn out that the obtained values of the fractions of the correct signals are just the result of excessive optimization at a given time interval and do not bear any meaningful meaning.

3. Results and analysis

The result of the algorithm for the minute sampling frequency on the interval from 01.01.15 to 17.08.15 can be seen in Table 1. The result of the algorithm performance for the minute sampling frequency on the interval from 01.01.15 to 17.08.15 can be seen in Table 2. The abbreviations used are: RSG is random signal generator, RSG is a synonym for random numerical generation, that is, they are just random points on a time series. The generator uses a uniform distribution throughout the optimized interval, i.e. probability of signal appearance for all points of the interval on which the optimization is carried out are equal. We recall that (2.1), (2.2), (3.1), (3.2) are the sub clauses of Definitions 2 and 3 indicated at the beginning of the paper. The slash sign "/" separates the values of the parameters and the results of the algorithm work, obtained for different definitions, for example, the record (2.1) / (2.2) means that the values obtained for determining (2.1) go to the sign of the line, and after the sign the values obtained for (2.2). Values of normalized values of large movements of time series are added. The time series is normalized according to

$$\hat{X} = \frac{X}{\langle X \rangle},$$

the formula where X this is the original time series, $\langle X \rangle$ this is the average of the time series. The result of the algorithm for the minute sampling frequency on the interval from 01.01.15 to 17.08.15 can be seen in Table 1. The result of the algorithm performance for the minute sampling frequency on the interval from 01.01.15 to 17.08.15 can be seen in Table 2.

Table 1. The interval from 01.05.15 to 17.08.15 definition 2

Assets	MLHE window	Interval, number of points (2.1)/(2.2)	Signal threshold in the number of mean values (2.1)/(2.2)	The normalized value of large motion (2.1)/(2.2)	Faithful signals of MLHE (2.1)/RSG	Faithful signals of MLHE (2.2)/RSG	Total number of signals (2.1)/(2.2)
AT&T	320	70/40	1.8/2.2	0.0043/0.0023	0.81/0.35	0.90/0.5	70/57
Apple	320	50/50	1.9/1.9	0.0047/0.0035	0.86/0.33	0.95/0.49	87/87
H&P	320	70/70	2.9/2.9	0.0066/0.0047	0.82/0.33	0.89/0.51	60/60
Coca-cola	320	40/40	2.8/2.8	0.0030/0.0024	0.85/0.36	0.89/0.49	56/56
IBM	320	90/40	4.0/3.4	0.0062/0.0029	0.75/0.35	0.80/0.49	32/35
Intel	320	50/50	2.9/2.9	0.0052/0.0037	0.93/0.35	0.95/0.49	61/61
Microsoft	320	70/70	1.2/1.2	0.0057/0.0042	0.82/0.34	0.89/0.50	89/89
Google	320	90/130	0.8/4	0.0091/0.0093	0.52/0.34	0.82/0.50	44/17
McDonalds	320	60/40	1.7/2.8	0.0041/0.0024	0.78/0.36	0.89/0.51	68/54
Exxon	320	50/50	4/4	0.0040/0.0030	0.92/0.35	1.0/0.50	52/52
Kraft Foods	320	40/40	3.9/3.9	0.0050/0.0038	0.80/0.35	0.92/0.50	26/26
Pfizer	320	40/50	3.7/2.8	0.0035/0.0032	0.89/0.38	0.93/0.49	45/53
Bank Of America	320	40/40	4/1.3	0.0043/0.0033	0.83/0.33	0.90/0.47	40/72
D&J-ind	320	70/80	2.2/3.1	0.0028/0.0023	0.83/0.33	0.95/0.51	48/39
Gazprom	320	210/200	2.4/2.4	0.0129/0.0113	0.51/0.39	0.6/0.5	77/77
Lukoil	320	200/220	3.1/2.7	0.0116/0.0108	0.55/0.39	0.7/0.5	38/43
Sberbank	320	220/190	2.9/1.1	0.0193/0.0149	0.53/0.38	0.6/0.5	65/110
RTSI index	320	190/220	2.6/2.6	0.0116/0.0110	0.69/0.37	0.78/0.5	45/45
MTS	320	220/220	1.5/3.9	0.0161/0.0131	0.56/0.37	0.66/0.47	83/36

The work of the MLHE method is based on the hypothesis of multifractality of financial time series, so one can test the method (and at the same time the hypothesis itself) based on the reverse assumption:

using the scrambling algorithm. "Let us destroy" the multifractal structure of the time series and apply the method to the new scrambled time series. For example, take the stock of AT & T, which is the first in the tables 1,2,3. As an algorithm for scrambling, we take an algorithm based on the iAAFT technique (iteratively Amplitude Adjusted Fourier Transform), proposed and described in (Schreibe & Schmitz, 2000). This scrambling algorithm is chosen because it allows to preserve the probabilistic distribution of the time series and its frequency spectrum (within the limits of calculation error). The error that determines the differences in the original spectrum and in the spectrum of the scrambled series will be determined by the formula:

$$e_{spec} = \frac{1}{N} \sum_{i=1}^{N} \left\| \left| \hat{F}(\omega_i) \right| - \left| \hat{S}(\omega_i) \right| \right\|, \qquad (5)$$

where $\left| \hat{F}(\omega_i) \right|, \left| \hat{S}(\omega_i) \right|$ are the Fourier transform amplitudes of the original and scrambled series, this is the total number of Fourier-expansion frequencies (for a discrete Fourier transform is equal to the number of points of the original series). Thus, the error is the average deviation of the amplitude modules of the Fourier spectrum.

Table 2. The interval from 01.05.15 to 17.08.15 definition 3

Asset	MLHE window	Interval, number of points (3.1)/(3.2)	Signal threshold in the number of mean values (3.1)/(3.2)	The normalized value of large motion (3.1)/(3.2)	Faithful signals of MLHE (3.1)/RSG	Faithful signals of MLHE (3.2)/RSG	Total number of signals (3.1)/(3.2)
AT&T	320	40/50	3.5/2.8	0.0021/0.0016	0.72/0.34	0.79/0.5	43/54
Apple	320	40/40	2/2	0.0028/0.0019	0.72/0.33	0.77/0.5	83/83
H&P	320	60/60	3.5/3.5	0.0041/0.0027	0.67/0.35	0.82/0.5	49/49
Coca-cola	320	50/50	2.7/2.7	0.0023/0.0017	0.73/0.38	0.80/0.49	56/56
IBM	320	30/30	2.6/2.6	0.0023/0.0015	0.60/0.38	0.76/0.49	45/45
Intel	320	50/50	2.5/2.5	0.0035/0.0022	0.77/0.38	0.80/0.5	67/67
Microsoft	320	40/40	3.9/3.9	0.0028/0.0018	0.78/0.36	0.88/0.53	42/42
Google	320	40/140	3.2/3.9	0.0040/0.0066	0.57/0.36	0.76/0.53	23/19
McDonalds	320	60/70	3.2/3.2	0.0028/0.0021	0.73/0.36	0.89/0.51	48/47
Exxon	320	40/40	3.6/3.6	0.0024/0.0016	0.83/0.36	0.90/0.51	58/58
Kraft Foods	320	30/30	3.1/3.1	0.0030/0.0020	0.70/0.36	0.87/0.51	23/23
Pfizer	320	50/50	3.5/3.7	0.0026/0.0021	0.75/0.37	0.81/0.52	48/43
Bank Of America	320	30/30	4/4	0.0025/0.0018	0.7/0.36	0.8/0.47	40/40
D&J-ind	320	70/70	3.8/2.8	0.0019/0.0012	0.75/0.32	0.90/0.49	32/32
Gazprom	320	210/60	2.4/2.6	0.0085/0.0031	0.47/0.38	0.58/0.51	77/72
Lukoil	320	170/170	1.9/0.9	0.0070/0.0055	0.66/0.38	0.63/0.50	59/86
Sberbank	320	210/20	1.7/3.8	0.0118/0.0024	0.54/0.39	0.68/0.51	93/41
RTSI index	320	190/200	2.8/2.7	0.0077/0.0066	0.62/0.38	0.72/0.50	42/43
MTS	320	180/160	3.9/1.5	0.0093/0.0063	0.55/0.36	0.65/0.49	36/83

We now apply the MLHE algorithm to the scrambled series. The results are shown in Table 4, «Results of the application of the MLGR algorithm for a scrambled time series based on the AT & T time series».

Table 3. The interval from 01.05.15 to 17.08.15 definition 4

Asset	MLHE window	Interval, number of points	Signal threshold, in the number of mean values	The normalized value of a large motion, (4.1)	Faithful signals of MLHE (4.1)/RSG	Total number of signals
AT&T	320	50	1.6	5.38	0.85/0.25	72
Apple	320	60	3.7	7.84	0.80/0.19	50
H&P	320	50	2.3	8.35	0.76/0.18	67
Coca-cola	320	50	2.4	5.24	0.84/0.24	60
IBM	320	40	4	6.57	0.53/0.20	32
Intel	320	70	2.5	8.19	0.85/0.19	68
Microsoft	320	40	3.8	8.13	0.81/0.14	43
Google	320	30	4	9.41	0.30/0.11	17
McDonalds	320	60	1.7	5.52	0.73/0.19	68
Exxon	320	60	0.8	6.06	0.88/0.25	101
Kraft Foods	320	30	2	8.33	0.70/0.25	30
Pfizer	320	40	3.6	5.89	0.9/0.23	48
Bank Of America	320	40	4	7.52	0.83/0.20	40
D&J-ind	320	70	2	3.14	0.80/0.25	51
Gazprom	320	200	1.4	8.51	0.52/0.33	94
Lukoil	320	220	2.1	8.18	0.66/0.33	53
Sberbank	320	190	2.7	12	0.52/0.33	67
RTSI index	320	200	2.6	6.70	0.73/0.24	45
MTS	320	210	3.1	12	0.46/0.30	48

It can be seen that the values of the proportions of the correct signals for the given series were practically equal to the values obtained for the RSG and much lower than the values obtained earlier for the original AT & T series. All values are slightly higher than for the RSG due to the optimization algorithm. It can be shown that each of the values does not exceed the limits of the confidence interval. For example, for the definition (2.2), a confidence interval with a confidence level of 95% is 0.9, so a value of 0.56 falls into this interval, and therefore cannot be considered statistically significant. For other values, the confidence interval is even greater, since these values are calculated on a smaller number of signals. From the results obtained it can be concluded that the financial time series have multifractal structure and method MLHE stops working for scrambled time series.

Table 4. Results of the application of the MLGR algorithm for a scambled time series based on the AT&T time series

Definition	Interval, number of points	Signal threshold in the number of average values	The proportion of the correct signals of the MLHE for the scrambled time series	The proportion of correct RSG signals	The proportion of correct MLHE signals for AT & T	The total number of signals for the scrambled series
2.1	180	1.1	0.44	0.35	0.81	69
2.2	220	0.2	0.56	0.5	0.90	127
3.1	210	0.2	0.43	0.34	0.72	127
3.2	90	3.7	0.55	0.5	0.79	18
4.1	210	3.0	0.45	0.25	0.85	28

The most important and almost obvious result observed in all the tables for all selected time series (especially for the US market) is the explicit statistical difference between the frequency of the correct signals of the MLHE and the frequency of the correct signals obtained by the random signal generator (RSG).

The optimal values of the MLHE window are the same for all analyzed time series. In general, this is due to the rough optimization of this parameter (as already described above due to the speed of the program), but somehow these values are within a small interval [290,350] points relative to the entire range for the window w

The optimal interval values $\mathcal{E}$ for different series of the US market and the Russian market (individually) lie in a small numerical interval relative to the entire range of values $\mathcal{E}$, which, on the one hand, speaks of the non-random work of the MLHE method, and on the other, the overall internal structure of the time series of the US market and the Russian market (separately).

The optimal values of the signal threshold depend on the initial definitions and within the same definition are also in a narrow numerical interval relative to the entire optimization interval. If we compare the definitions (2.1) / (2.2) and (3.1) / (3.2), then in each definition group, the values obtained from the definitions (x.1) (based on mean values) are always lower than (x.2) (based on median values). This is clearly seen from the values of the proportions of the correct signals for the RSG, which for (x.1) are on the average 0.36, and for (x.2) they are 0.5. This is explained by the fact that the average values are shifted to a larger side relative to the median, which in turn is explained by the greater sensitivity of the average to "out layers", i.e. to very rare, large values, to which the median is not sensitive. It is worth noting the results obtained on the basis of the definition (4.1) based on volatility, if we compare the probabilities of the correct signals of the MLHE with the RSG, it becomes clear that with respect to this definition the method gives the best results. Another important and interesting result is observed when comparing the shares of American and Russian markets. It is easy to see that the optimal parameters for these groups vary widely and are centered around values that are far from each other. It is also clear that the probability values of the correct signals for the Russian market are on average much lower than for the American market. The reason for this difference can be several factors at once. First, the presence in the US market of more modern trading technologies, for example, HFT. HFT is an abbreviation for High-frequency trading, i.e. trading at a very high frequency, several transactions per second. And the development of communications, which allows the American market to react more quickly to any information, in particular, this can be explained a much shorter time interval ε. Secondly, the presence of a larger number of market participants, which on average provides more traders on each investment horizon. If in detail, it is obvious that every player on the market has one or more investment horizons, on which trader conducts his trade. Since in real trading the number of traders is limited and the number of investment horizons is practically unlimited, then some of the investment horizons will be empty or almost empty, and the larger the number of traders, the denser the possible investment horizons will be filled, which brings the real distribution of traders horizontally closer to the multifractal one. On the other hand, an increase in the number of players leads to more transactions per unit of time, and, consequently, leads to an increase in volatility. Thus, the multifractal approach on such series should definitely work better, which is confirmed by the high probability values for the US market. For the Russian market, the most traded and liquid instrument is the futures on the RTSI index, which in these tables among Russian instruments shows the best results. Third, the US market is more independent in the context of direct state intervention on price regulation. For example, at the very beginning of the Russian crisis of 2014, the Central Bank of Russia restrained the growth of the dollar artificially, creating demand for the ruble. It is

obvious that with such interference, the natural processes occurring on the market are violated, which leads to a violation of the multifractal structure of the market.

4. Conclusions

In this study, several definitions of the concept of large motion were used, on the basis of which the corresponding calculations of the probability of the correct (according to the definition) forecast of the dynamics of the time series were made. This method showed good performance, several times exceeding the probabilities of the correct signals received by the generator of random signals, which are necessary for an objective evaluation of the effectiveness of the method. The best results are obtained with respect to the definition (3.1). In practice, the average value of volatility is used to form the price of options, i.e. MLGP can be considered a good tool for option trading. Also valuable results are differences in the values of optimal parameters and probabilities of correct signals for instruments of American and Russian financial markets. In addition, based on the results obtained, one more important conclusion can be drawn: the greater the liquidity of the instrument, the more the number of players trades in this instrument, the better the performance of the MLGR (which is in some contradiction with the theory of an effective market).

References

Abry, P., Gonçalves, P., & Véhel, J. L. (Eds.). (2009). *Scaling, fractals and wavelets*. ISTE.

Bacry, E., Muzy, J. -F., & Arneodo, A. (2002). Wavelet-based estimators of scaling behavior. *IEEE Transactinons on Information Theory, 48*, 2938-2954.

Cogley, T., & Nason, J. M. (1995). Effects of the Hodrick-Prescott filter on trend and difference stationary time series: Implications for business cycle research. *Journal of Economic Dynamics and Control, 19*, 253-278.

Kapecka, A. (2013). Fractal Analysis of Financial Time Series Using Fractal Dimension and Pointwise Hölder Exponents. *Dynamic Econometric Models, 13*(1), 107-126.

Khan, A. U., & Gour, B. H. U. P. E. S. H. (2013). Stock market trends prediction using neural network based hybrid model. *International Journal of Computer Science Engineering and Information Technology Research, 3*(1), 11-18.

Kuperin, Yu. A., & Schastlivtsev, R. R. (2008). Modified Holder Exponents Approach to Prediction of the USA Stock Market Critical Points and Crashes. *Statistical Finance, 15*. http://arxiv.org/abs/0802.4460

Legrand, P., Lutton, E., & Olague, G. (2006). Evolutionary denoising based on an estimation of Holder exponents with oscillations. *Lecture Notes in Computer Science, 3907*, 520-524.

Los, C., & Yalamova, R. (2004). Multifractal spectral analysis of the 1987 stock market crash. *Working Paper, Department of Finance, Graduate School of Management, Kent State University*, 1-39.

Schreibe, T., & Schmitz, A. (2000). Surrogate time series. *Physica D, 142*, 346-382.

Struzik, Z. R. (2001). Revealing local variability properties of human heartbeat intervals with the local effective Hölder exponent. *Fractals*, *9*(01), 77-93.

Struzik, Z. R. (2001). Wavelet methods in (financial) time-series processing. *Physica A: Statistical Mechanics and its Applications*, *296*(1-2), 307-319.

Suárez-García, P., & Gómez-Ullate, D. (2014). Multifractality and long memory of a financial index. *Physica A: Statistical Mechanics and its Applications, 394*, 226-234.

Trujillo, L., Legrand, L., & Lévy-véhel, J. (2010). The Estimation of Hölderian Regularity using Genetic Programming. *Conference: Genetic and Evolutionary Computation Conference, GECCO 2010*, 861-868.