China-USA Business Review, ISSN 1537-1514
July 2013, Vol. 12, No. 7, 702-717

Prediction of Non-stationary Time Series With Replacement Variables

Perminov Gennady Ivanovich
National Research University Higher School of Economics, Moscow, Russia

The experimental comparison of methods of forecasting non-stationary time series by replacing the variables. One of these variables is the phase of the process. Currently, in the direction of the change of variables there were two main approaches to the isolation of stable characteristics of the process, and the band-pass filtering on the empirical mode decomposition—The band-pass filtering. Using a band-pass filter that transmits only the frequencies in a small neighborhood of a selected frequency will provide components with clearly highlighted the basic rhythm. One option for the analytical component of the original series is a discrete wavelet transform (DWT) with getting more approximate and detail the factors on levels of decomposition—On empirical mode decomposition. Empirical method of decomposition (EMD) is based on the simple assumption that any data consist of a variety of simple domestic species fluctuations (MOD). Each "fashion" must be a signal with zero mean, Maxima and minima of positive-negative, i.e., between each maximum and minimum signal $x(t)$ necessarily have his schedule with the direct crossing $x = 0$. Continue to make the variable phase each fashion is Hilbert decomposition. In both cases, instead of the original series we receive several quasi-stationary series that are projected separately. Projection results are reverse-conversion with predictive component series, which gives the forecast. Difficulty at this stage is the selection of forecasting method for sampling small dimension, because according to our research, the length of the memory of the financial series is extremely small. For the same source data give calculations for both approaches using neural network forecasting and sliding fractal Caterpillar SSA. The results show satisfactory as a prediction.

Keywords: filter strip, decomposition to empirical fashion, Hilbert-Huang transform, discrete Wavelet transform, Forecast method of SSA.

Introduction

In practice, the ideal theorems of fixed processes are only approximate, and accurately they cannot be implemented in principle. On the other hand, it is well known that the methods of mathematical statistics, correctly apply only to stationary time series. If the time series varied, the theorem on the effectiveness, consistency, and asymptotic normality of sample estimates and their variances in general, are not met. When you define a non-stationary time series forecast error to consider two factors: limb difference and sampling distributions for different samples due to the non-stationary process.

Perminov Gennady Ivanovich, Associate Professor, Candidate of Technical Sciences, Department of Business Analytics, National Research University Higher School of Economics.
Correspondence concerning this article should be addressed to Perminov Gennady Ivanovich, 141220, Pushkinsky district, Moscow region, Cheluskinsky, Water stalled st., 9A, Russia. E-mail: gperminov@hse.ru.

In some stationary projection magnification window width (sample size) typically increases the accuracy of the forecast. For non-stationary series remove in time from the last point of the sample data used for projection usually leads to the increased error prediction. In other words, when you build the forecast you can use the model of stationary fit of available data, and to estimate the dynamic of this model quite quickly become unusable (see Figure 1).

Application of simulation to calculate the ongoing assessments of the statistical characteristics of non-stationary process $x(t)$ generates a fatal error: arising from the extremities of time averaging, e.g., due to a lack of representativity of the sample, as well as errors arising due to changes in statistics at the interval of averaging. According to the results of the observations cannot be statistically separated in a finite sample of non-stationary factor (denoted by Σ_2 error) of factor to materialise (denoted Σ_1 through error). Window width optimization occurs because error Σ_1 and $\Sigma_2(t)$, as a function of the sample size has different behavior. To reduce error $\Sigma_1(t)$ due to incomplete statistical representativeness should be increased so as to reduce errors of $\Sigma_2(t)$ non-stationary impact on the statistical characteristics of time series sample size should be reduced.

To confirm the above, consider a simple AR model of first-order to find errors and RTS Index weekly S&P, as well as one-day oil prices. Increase the width of the window will move to one observer, and initial size equal to 4, the final 40 observations. The typical dynamics model in the range error changes the width of the window from 6 to 40 (see Figure 1).

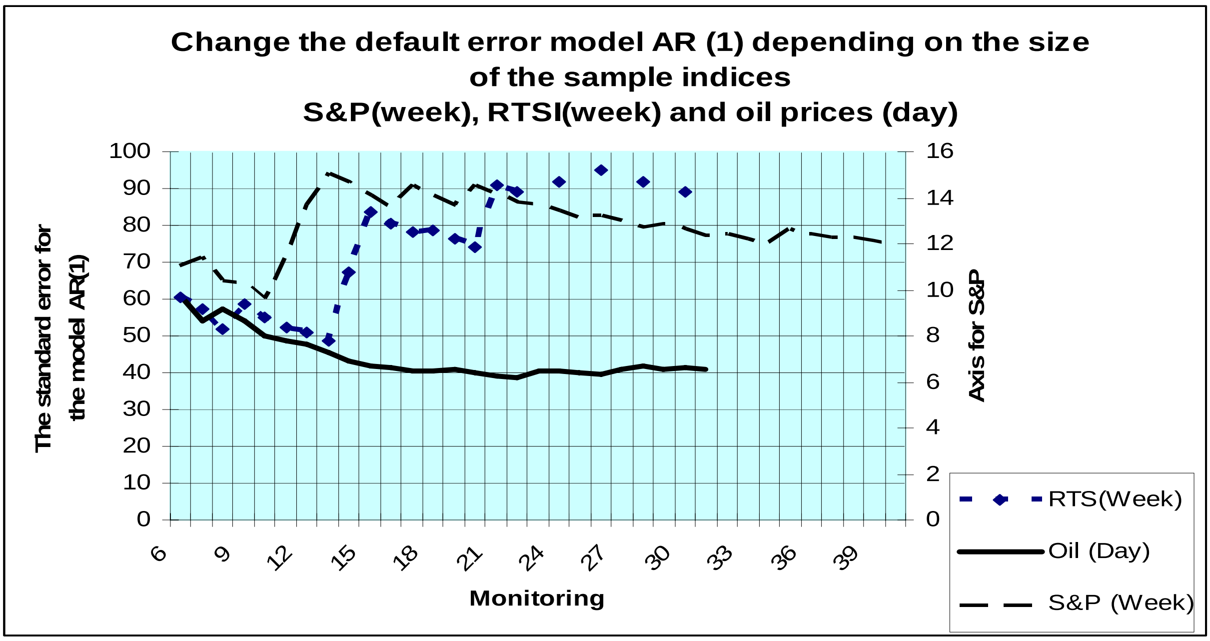


Figure 1. The sample values impact error model AR (1).

There are some non-stationary time series trend after which can be approximately fixed. Such data sets, for example, with polynomial trend, are not considered. The aim of our research is a series not given standard algorithms to stationary.

Non-stationary time series can be categorized by type not of stationarity into three fairly general classes:

- Non-stationary time series information is kept in a small act of distribution or its basic settings, and changing the variance or mean number;

- Large non-stationary time series when changing distribution;
- Significantly non-stationary time series when not only changing the law of distribution of a random variable, but also there is no analytical view of the trend of a time series. The third class of a series of economic time series is characterized by frequent changes in complex conditions defining the random process. The market is expected to meet with fractal hypothesis periodic interference by stakeholder groups, opposing the growth or drop in prices, the transition from one administration to another schema partial. Discontinuous component is a special case of significant non-stationary time series, which are not always visually seen. Non-stationary time series is essentially split into segments, which change little component properties then produce a consistent selection component of the time series for each segment separately.

Special Methods of Analysis and Forecasting Non-stationary Time Series

Traditional data analysis methods are based completely on the line and fixed assumptions. The analysis of non-stationary time series methods are inapplicable classic.

Only in recent years have new methods for analyzing non-stationary and nonlinear data (Bezruchko & Smirnov, 2005; Orlov & Osminin, 2011). These include wavelet-analysis (Kolodyazhny, 2006; Dobeši, 1999) methods for random allocation (Flandrin, 1999; Gröchenig, 2001) and a few other methods (Louis, 2008).

In real systems, the series can be both a non-linear, and not stationary. In addition, according to Figure 1, data with observations, suitable for prediction, usually too short.

For non-linear and non-stationary time series stretching or forecast based on historical retrospective data is risky, because the latter may belong to a segment with different parameters not stationary.

In the task, the main problem is finding such a sustained, “stationary” characteristics of the process (variables) that can translate a number to become stationary, non-linear, for whom there are proven methods of analysis (e.g., neural networks, adaptive, or genetic algorithms). Analysis of the data in these systems can be reduced to finding a new variable, after conversion, which becomes fixed. In some works as such variables proposed: phase of the process, which has properties of stationarity (Huang et al., 1998) and the discrete wavelet transform (SMD) (Mallat & Hwang, 1992).

In both directions with the replacement of the variables has a first phase—time series in Kvazi-stationary. This is done by sustainable characteristics of flapper valve strip Seel-if filter and SMD fashion to empirical method of decomposition (EMD) phase.

Use of filter strips, vapour permissive only frequencies in a small neighborhood of a selected frequency will provide component with clearly identified the main rhythm.

In addition to band-pass filtering other approaches are possible. Increasingly popular in recent years proposed (Huang et al., 1998) signal decomposition method on so-called “empirical fashion” (this version of adaptive non-linear filter). For each of the “mod” phase is well defined by Hilbert transform as a spin trajectory around clearly selected centre on the complex plane (NNT) (V. A. Davydov & A. V. Davydov, 2010). The NNT is comprised of two parts: an EMD and the Hilbert’s spectral analysis (HSA). But in economic and financial series domestic development method is not available. It is therefore NNT appropriate to assess its applicability to this series.

Brief Conceptual Information on NNT

The first phase (EMD) method is based on a simple assumption NNT that any data consist of a variety of simple domestic species fluctuations (MOU). Each "fashion" must be a signal with zero mean, Maxima and minima of positive-negative, i.e., between each maximum and minimum time series $x(t)$ necessarily have his schedule with the direct crossing $x = 0$. At any time, there can be many different coexisting species fluctuations (MOU), superimposed one on another. The result is finite complex data. Each of these vibrations is an essential function of the regime (intrinsic mode function-IMF) with the following definition:

- Number of extremes in the data set and the number of zero-crossings must be equal or differ at most by one;
- At any point in the data the average envelope, certain local maximums, and local minima are zero.

EMD has the advantage that the prediction is not for the full range of data, but only for the IMF, which has a much narrower bandwidth, and all shall be equal to the number of IMFs extremes and zero-crossings.

The decomposition of a number of on empirical modes (EM) increases the degree of determinism. It allows you to present a number as a sum of a slow trend and set low-, medium-, and high-frequency adjustments or internal mod. These functions in-house mod (IMF) are calculated directly from the data and, at the same time, are empirically orthogonal basis to decay so that their full amount allows you to recover the signal. Decomposition has good local properties and adaptability, and is based on the following view:

Time series = fast oscillation + slow oscillation

The slow oscillation component is, so-called, which approximates the element (responsible for transmission of low frequencies), unlike quickly applying, which is drills (is responsible for transferring the treble), plays it local peculiarities.

$$f(k) = m_1[f](k) + d_1[f](k) \tag{1}$$

where $f(k)$ is an arbitrary function, $m_1[f](k)$ is approximating the integral (the local trend), $d_1[f](k)$ is detailing the component (local details).

Identification of informative component based on Hilbert-Huang spectrum.

After applying the Hilbert transform to each of EM (the procedure is not involved, the resulting balance), formed the analytical components $z_i(k)$:

$$z_i(k) = c_i(k) + jH[c_i(k)] = a_i(k)e^{j\varphi_i(k)} \tag{2}$$

The original number $s(k)$ is itself in the following way:

$$s(k) = \sum_{j=1}^{M-1} a_j(k) e^{\left(i\int \omega_j(k)\Delta k\right)} \tag{3}$$

where $a_j(k)$ is the gain envelope j is EM and $w_j(k)$ instant frequency.

The expression (3) allows to represent the amplitude and instantaneous frequency as a function of time in a three-dimensional (3D) space in which the amplitude (or corresponding energy) is depicted on the frequency-time plane. This time-frequency distribution is called a spectrum of Hilbert-Huang $H(w, t)$. It is generated by calculating the instantaneous frequencies and amplitudes of all component decomposition and applying these values on the 3D chart.

Therefore, after you convert the Hilbert each component of the time series, you must predict each of them. Then, for projected (extrapolated) values are inverse Hilbert transformation results are summed and is forecast for the original time series.

Thus, the entire algorithms for the forecast are as follows:

(1) Source time series $s(t)$ undergoes decomposition for finite number of EM $\{c_i(k)\}_{i=1}^{M-1}$ and the resulting balance $r_m(k)$;

(2) For all or some component of a Hilbert transform;

(3) For all or some of the converted using components or neural networks or genetic algorithms are forecasting (extrapolation) by the specified number of values;

(4) Results of prediction are the reverse transformation of Hilbert's;

(5) All used components are added together, produce a forecast for the original time series.

HHT is a frequency-time analysis of data and does not require a priori functional basis transformation.

Under Hilbert's transformation understand HSA. To select an arbitrary number of amplitude and phase $s(t)$ appears as the real part of the complex signal $s_a(t)$ (it is called analytical component):

$$s(t) = \mathrm{Re}(\dot{s}_a(t)) \tag{4}$$

The real part of the analytic signal, of course, must match the original signal $s(t)$. Imaginary part of $s \pm (t)$ is called a paired signal or Kvadraturnym supplement:

$$\dot{s}_a(t) = s(t) + js_{\perp}(t) \tag{5}$$

Includes signal is derived from the source by converting a Hilbert, calculated as follows:

$$s_{\perp}(t) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{s(t')}{t-t'}dt' \tag{6}$$

This integral is a rollup time series $s(t)$ and function $1/\pi t$. This means that the Hilbert's transformation can be made linear system with constant parameters. This in turn implies that you can determine the frequency characteristic of the Hilbert transform:

$$\dot{K}_{\perp}(\omega) = \int_{-\infty}^{\infty}\frac{1}{\pi t}e^{-j\omega t}dt = \begin{cases} j, \omega < 0, \\ 0, \omega = 0, \\ -j, \omega > 0. \end{cases} \tag{7}$$

So, amplitude-frequency (AF) of Hilbert transform is one everywhere except at zero frequency, i.e., a Hilbert transform does not change the amplitude ratios in the spectrum of the signal, only removing it from the continuous component. Phases of all spectral components in the positive frequencies are reduced to 90°, in the area of negative frequency is increased by 90°.

Thus, the Hilbert transform is an ideal phase shifter, which modifies the phase shift at all frequencies equal to 90°.

It is obvious that the inverse Hilbert should make the same phase shift, but of opposite sign, again while preserving amplitude ratios in the spectrum of the signal.

A Practical Prediction Method for NNT

As the original series was a one-day "S&P Index". Decomposing the EM made package Spectra-Analyzer (Lûbušin, 2006). This tool allows you to select (see Figure 2): (1) kernel-based trend with the specified width, or polynomial specified power, or a combination (see Figure 3); (2) remove high-frequency noises using criterion Donoho-Johnston (see Figure 4); and (3) select "fast oscillation" (see Figure 5).

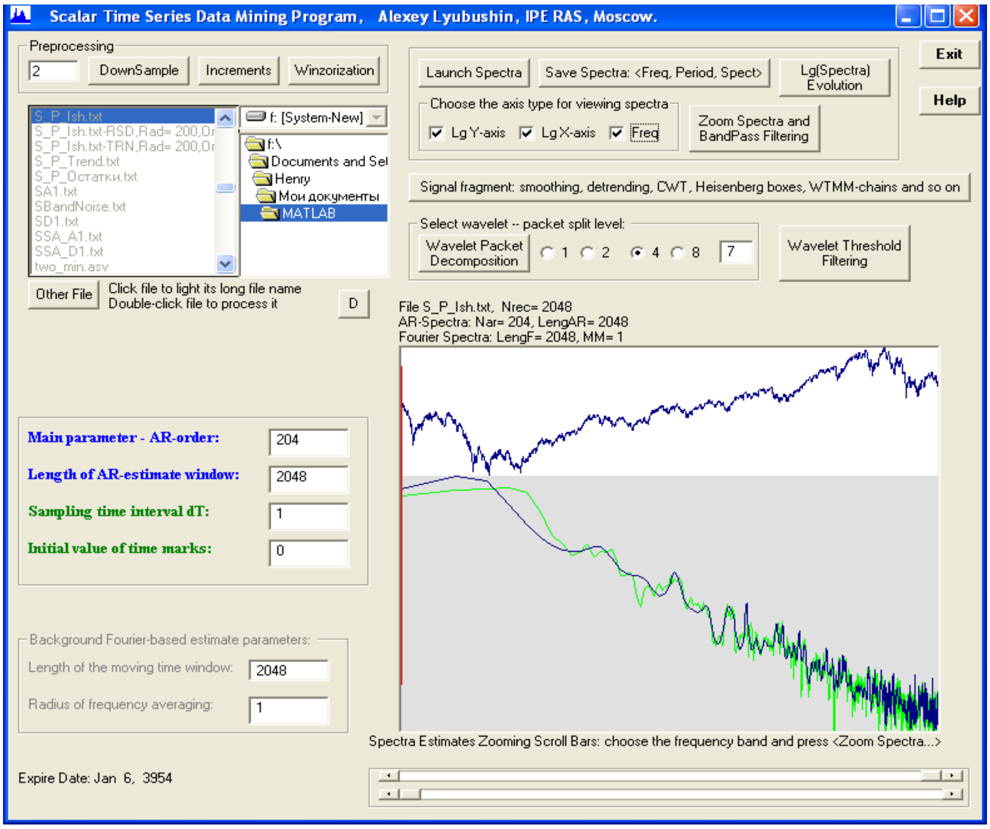


Figure 2. Source spectrum: time series.

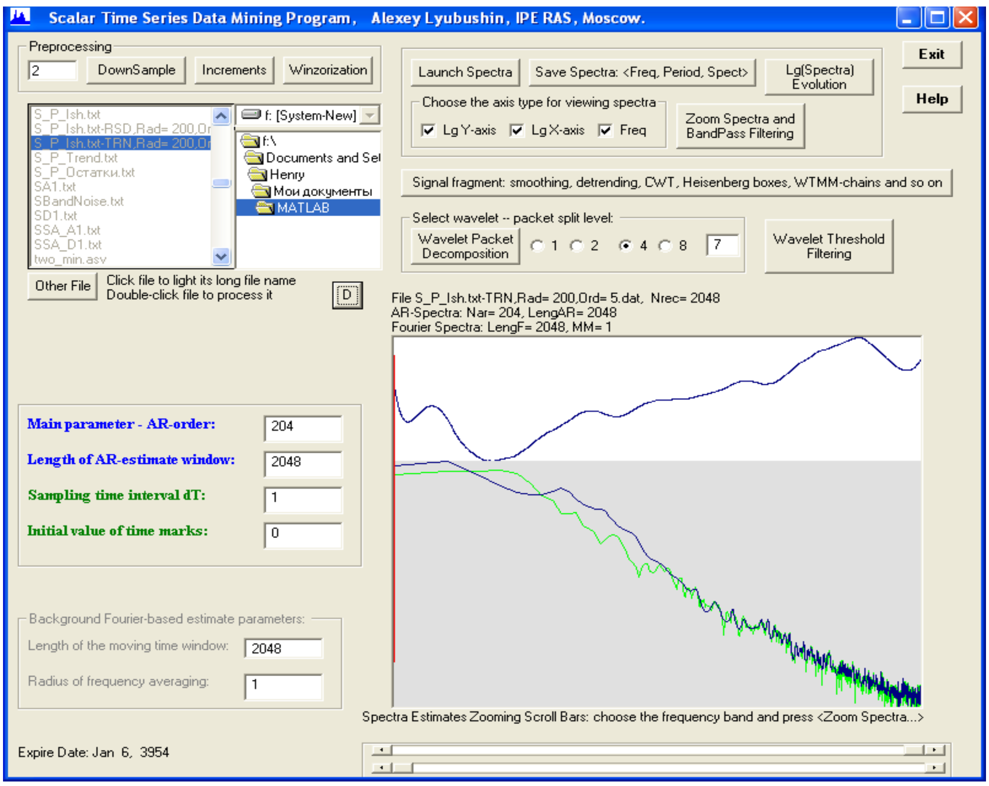


Figure 3. The selected trend of “S&P index”.

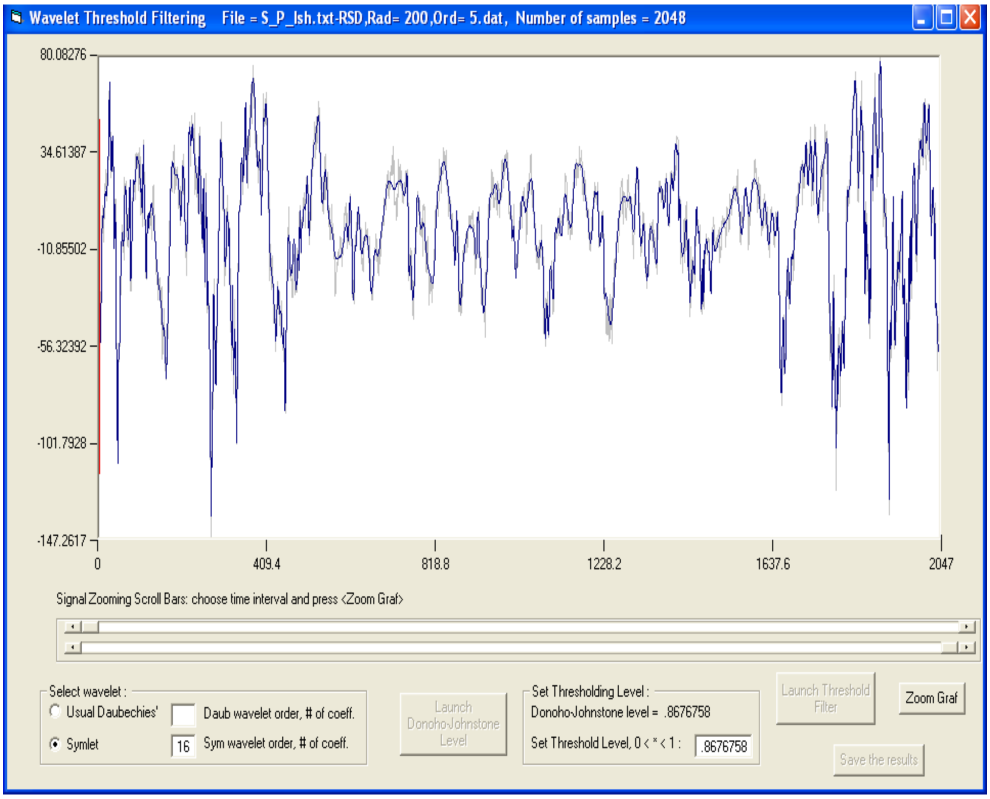


Figure 4. Frequency noise on allocation Donoho-Johnston.

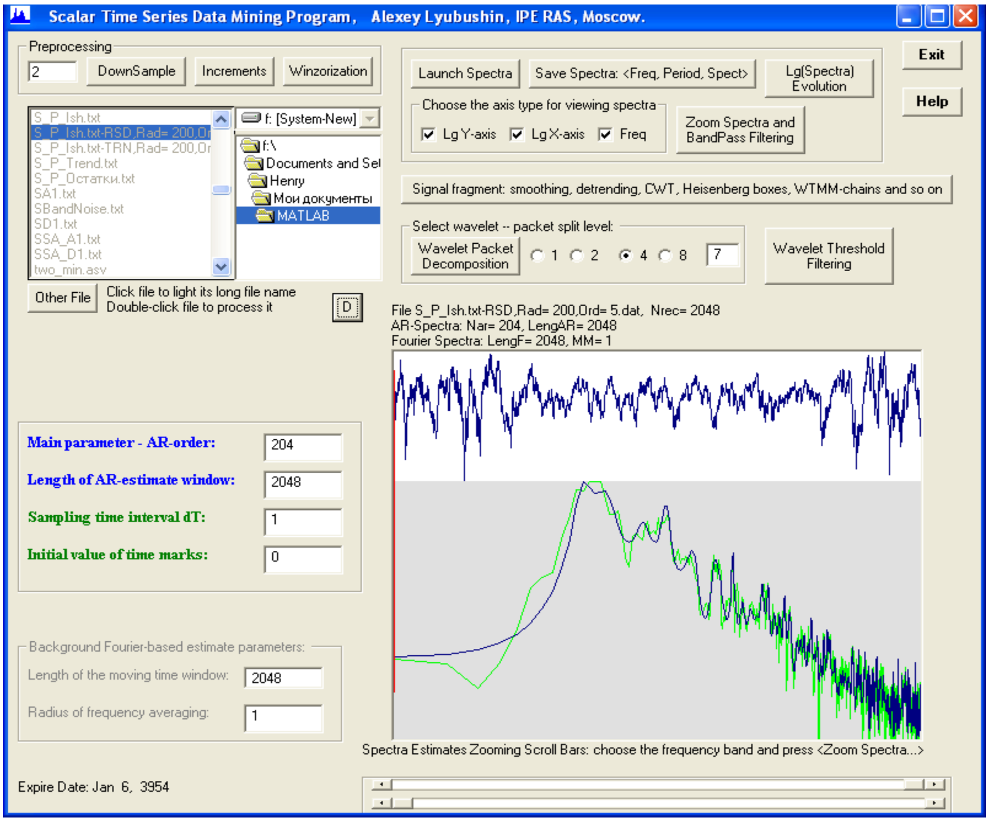


Figure 5. Dedicated “fast oscillation”.

One of the important problems of non-stationary time series is to define the width of the sliding window. The majority of researchers and financial time series are using a priori width window depending on the used methods of predicting—from 64 up to 360 days for a one-day series. Therefore, in a separate study, we examined the methods of non-stationary time series window width (Perminov, 2011).

The length of the last segment of the series "S&P Index" was 21. This value set the segment for further prediction using neural network method, NNT, genetic algorithm and method of moving fractal Caterpillar SSA.

Figure 6 illustrates the trend in the last segment and its value after converting to Hilbert and Figure 7 shows the trend of the last segment of residues and their conversion to Hilbert.

After conversion to Hilbert series of one-dimensional "Trend" and "Balances" the multidimensional, taking into account the width of the window, we received the following neural network model (see Figures 8 and 9).

As you can see from Figures 8 and 9 that the component model is quite acceptable quality. After receiving the predictive values of trend and residues should be transformed to apply the inverse Hilbert transform to forecast values. Their combination yielded values of non-stationary time series S&P on the last segment (see Figure 10).

Similar settlements with the use of Hilbert transform and genetic prediction for the last one-day segment is illustrated index S&P. Figures 11 and 12 show the results of the prediction method of sliding fractal Caterpillar SSA (Golândina, 2006). In the latter case, the results of prediction are reciprocally incorporated with valid data to render the error prediction.

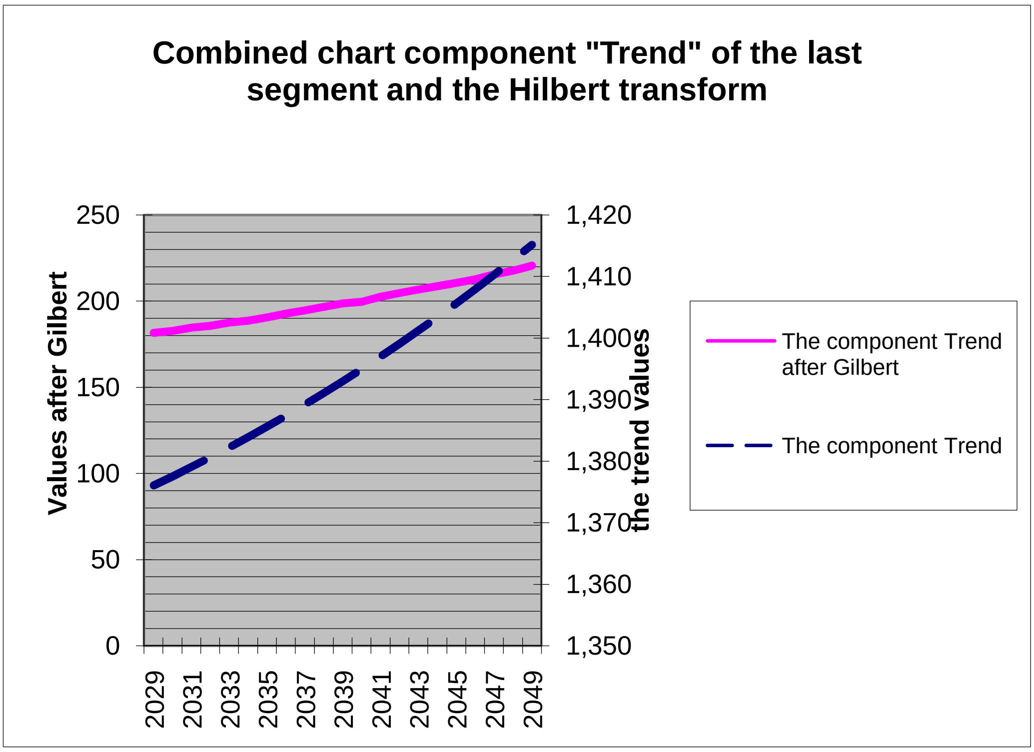


Figure 6. Trend the last segment and its conversion to Hilbert.

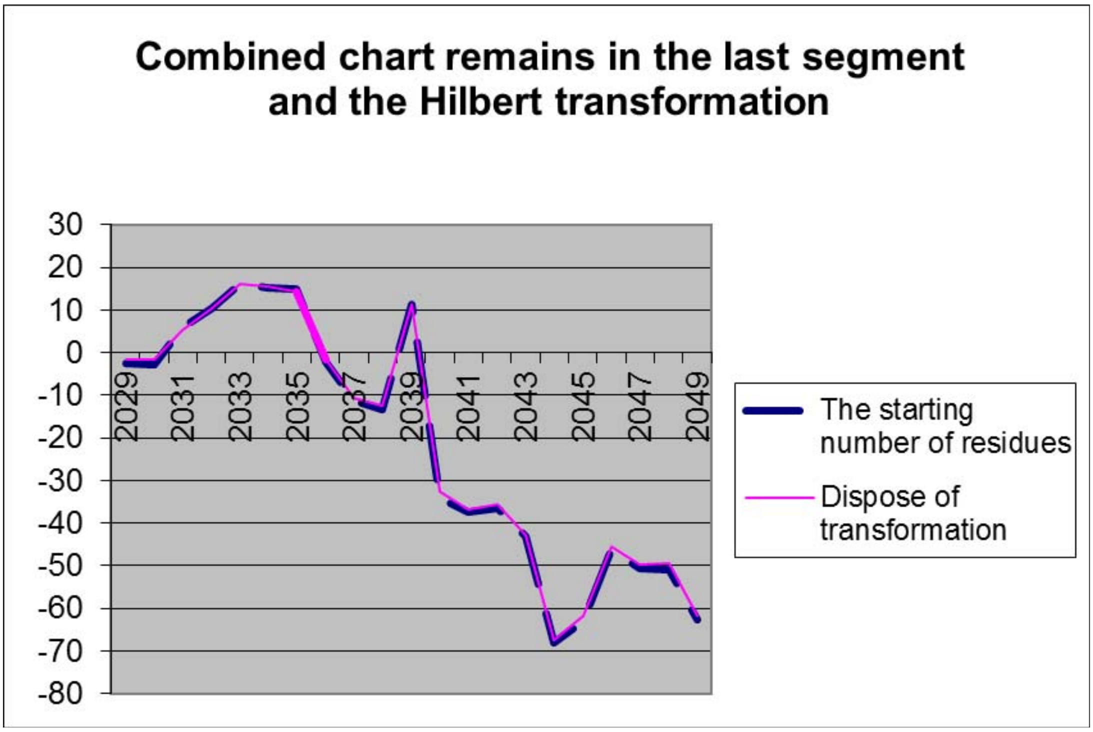


Figure 7. Remains of the last segment and their conversion to Hilbert.

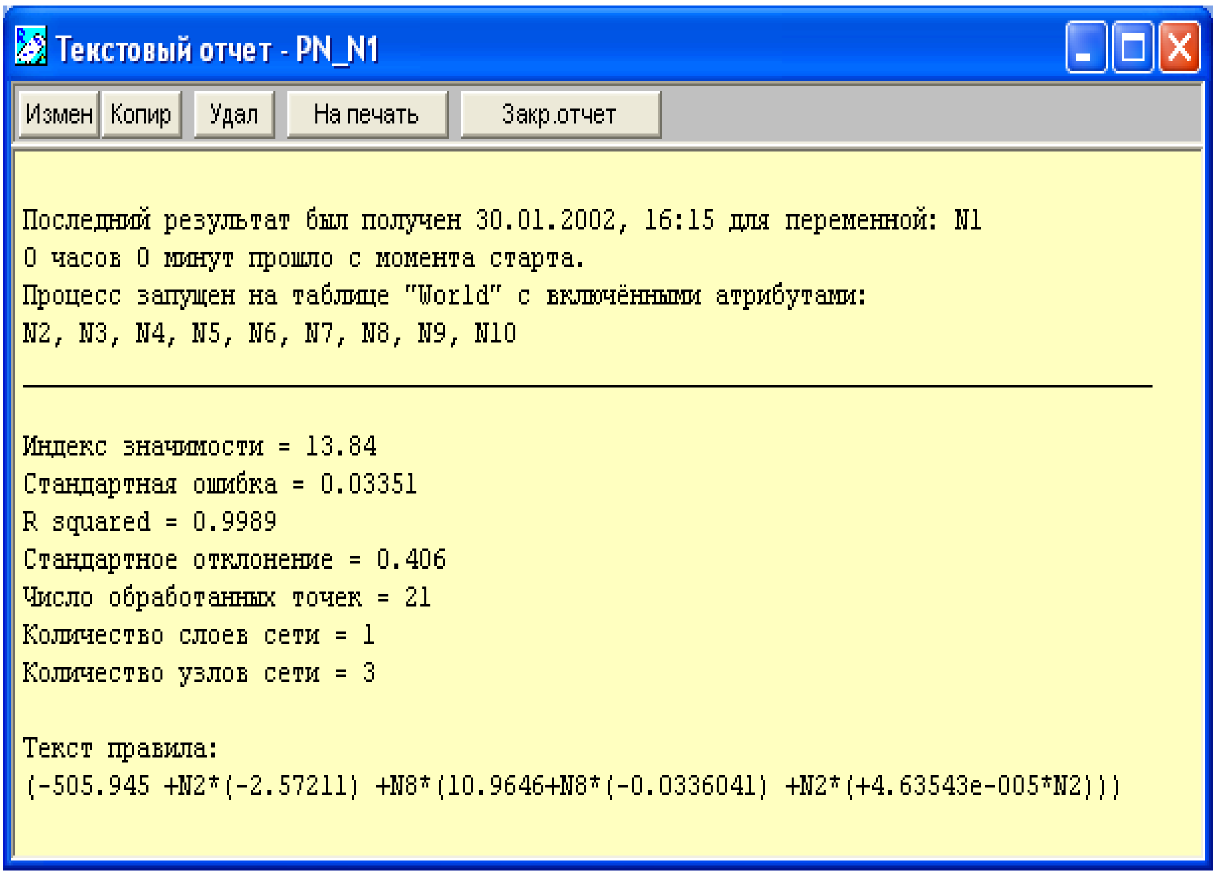
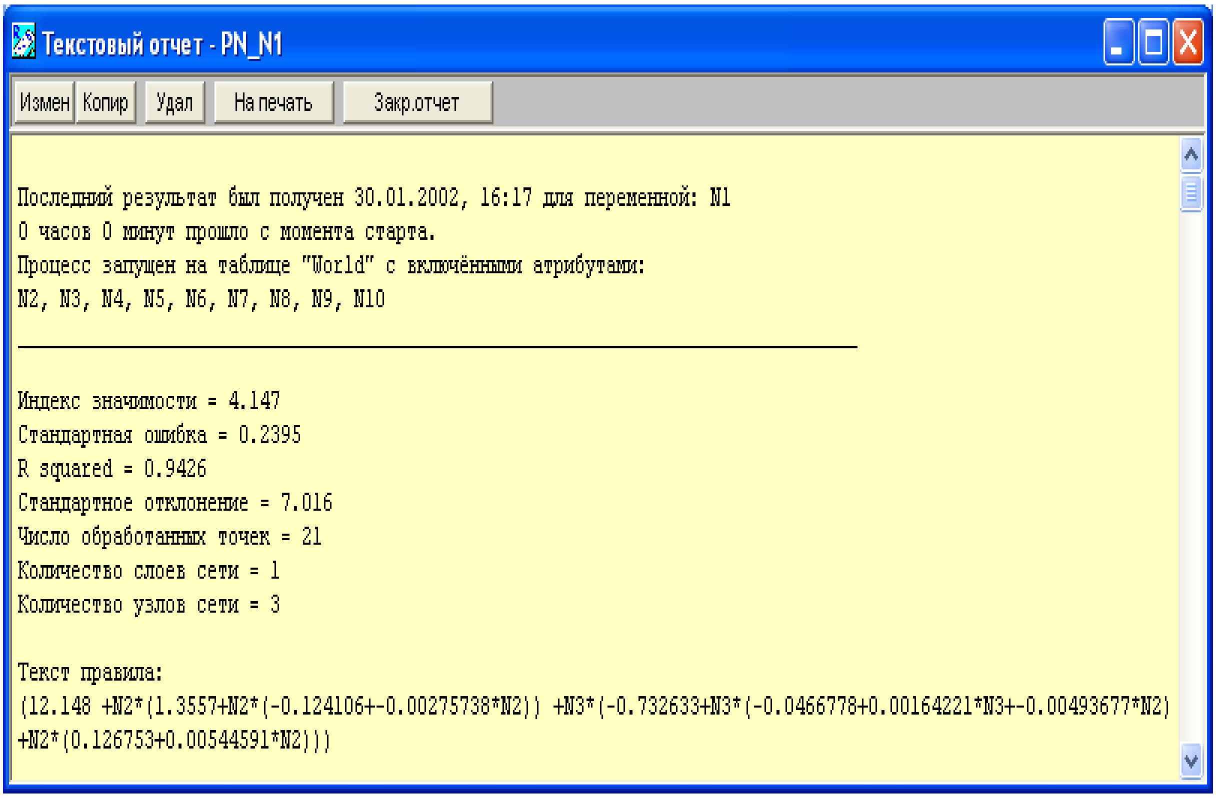


Figure 8. Neural network model trend at Hilbert transform.


Текстовый отчет - PN_N1

Измен Копир Удал На печать Закр.отчет

Последний результат был получен 30.01.2002, 16:17 для переменной: N1
0 часов 0 минут прошло с момента старта.
Процесс запущен на таблице "World" с включёнными атрибутами:
N2, N3, N4, N5, N6, N7, N8, N9, N10

Индекс значимости = 4.147
Стандартная ошибка = 0.2395
R squared = 0.9426
Стандартное отклонение = 7.016
Число обработанных точек = 21
Количество слоев сети = 1
Количество узлов сети = 3

Текст правила:
(12.148 +N2*(1.3557+N2*(-0.124106+-0.00275738*N2)) +N3*(-0.732633+N3*(-0.0466778+0.00164221*N3+-0.00493677*N2)
+N2*(0.126753+0.00544591*N2)))

Figure 9. Neural network model of the transformed balances Hilbert.

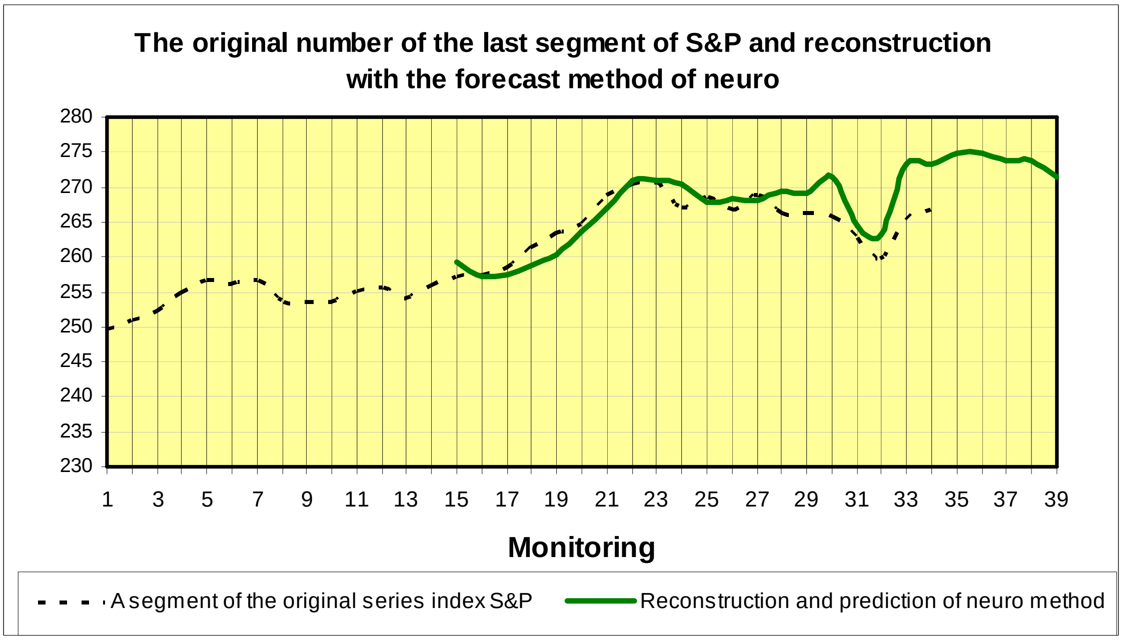


Figure 10. Results of Hilbert transformation and the neural network prediction for the last segment of a one-day index S&P.

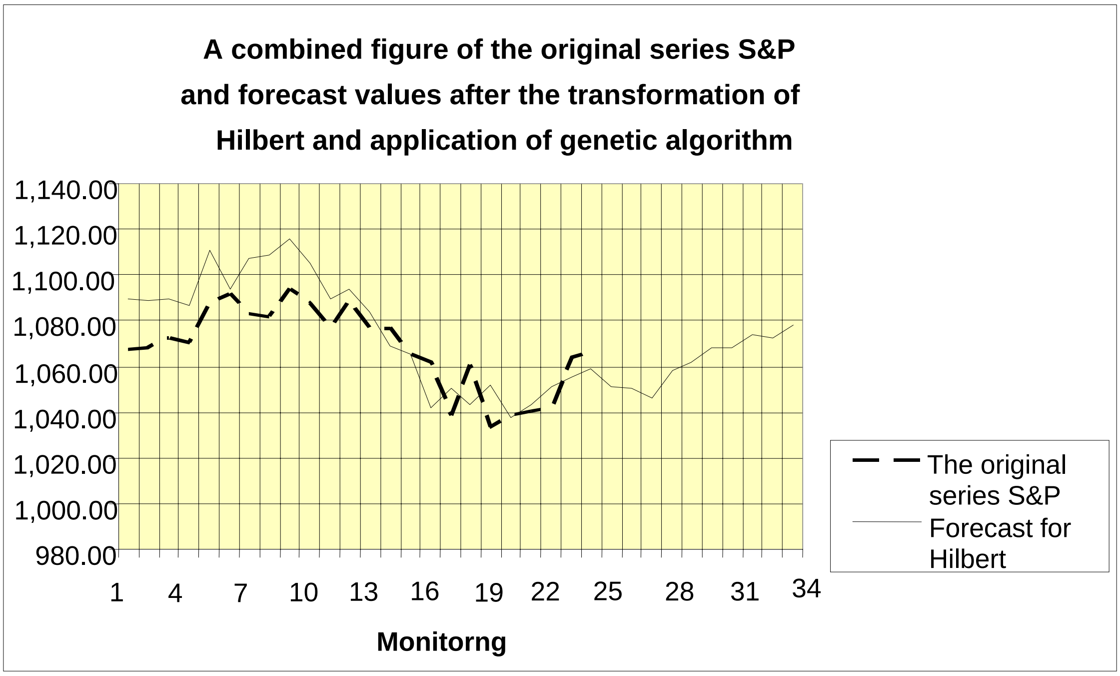


Figure 11. Results of Hilbert transformation and genetic prediction method for the last segment of a one-day index S&P.

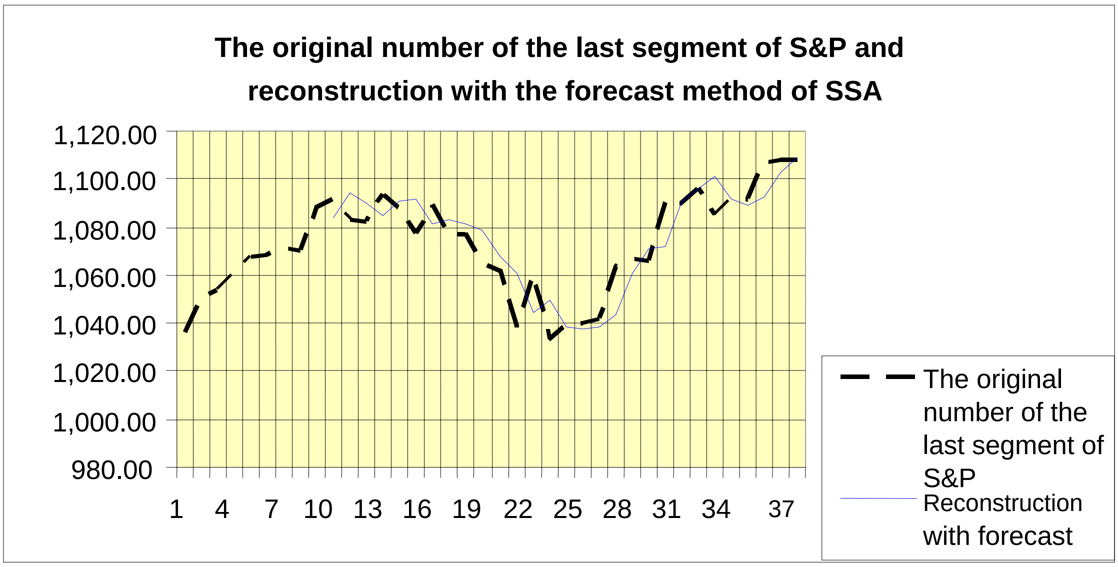


Figure 12. Results of Hilbert transformation and prediction method of rolling one-day segment for the last fractal index S&P.

Brief Conceptual Information on Discrete Wavelet Transform (DWT)

A significant advantage of DWT, is the property of locality of wavelets. The wavelet transformation of multiplication on the window as is contained in the core function that reduces or enlarges the window with growth parameter an increase authorized by frequency and time resolution, and reducing this parameter reduces the resolution by frequency and increases over time (see Figure 13). From here you will be able to adaptive signal in the options window. Moving the window time-frequency and low-frequency allocated equally well, and high-frequency characteristics of signals. This property of VP gives it a great advantage when analyzing the local properties of time series.

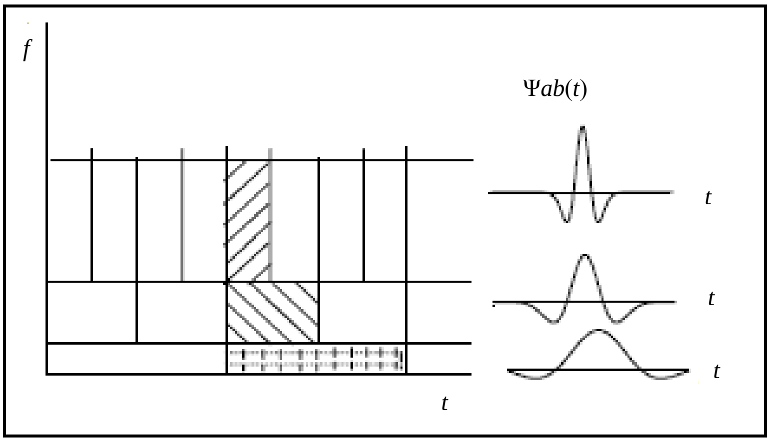


Figure 13. Adaptive wavelet transforms.

Algorithm of forecasting method using discrete wavelet for consists of the following steps:

- Step 1: Any method is determined by the length of the last segment. We repeat that in this work, the problem is not discussed. This step is necessary because, unlike the SSA forecasting method with NNT, good running short segments, fiberboard with each level reduces the sample. So the definition of discrete coefficients for three levels will reduce the sample in eight times and how they predict will remain in doubt;
- Step 2: A source of time series is a discrete wavelet decomposition and detailing the approximate conversion rates;
- Step 3: MDF allows you to analyze using wavelet coefficients separately each component of a time series that corresponds to a particular frequency band. Therefore, the last segment of the coefficients, is projected separately genetic, nejrosetevym, or sliding fractal (Caterpillar SSA) (Golândina, 2006). If necessary, truncating the high-frequency noise, some small-scale detailing rates, cD2, cD3 and cannot predict cD1;
- Step 4: For the coefficients of the last segment with the addition of projected values is the inverse DWT.

Wavelet function is created on the basis of a reference (parent) function, which determines the type of wavelet scale. One of the fundamental ideas of the wavelet transform decomposition of signal waveforms is divided into two components—gross (approximation) and the qualification (drill through)—and their fragmentation in order to change the level of decomposition. It is possible, as in the interim and in frequency fields source of wavelets.

The application of continuous wavelet transforms directly (for example, the algorithm implements a fast Malla wavelet transform (BVP)) is not unambiguous and depends on temporary shift that is not valid for the task of forecasting non-stationary time series.

This limitation does not exist in the DWT. For its application should choose an algorithm of wavelet transformation.

The main results in the building of bases were Ingrid Dobeši. How to build and analyse the properties of multiple site collections an orthonormal bases “Dobeši” (Dobeši, 1999). These wavelets have well localized spectrum in the frequency domain.

The multiresolution analysis of decompositions on orthonormal basis of wavelets is considered as a split functions of successive layers of scale clarifying, each of which is more detailed than the previous one.

Therefore the amount is presented as a double-time shifts (*k*) magnitude and (*j*). Moves to lower levels of accuracy of the signal is reduced, but appears to signal filtering of wavelet removing noise from it and effective compression signals. Schematically in wavelet analysis of time series (*s*) shall be divided into "rough" (A1) and "the details" (D1). In turn, the approximate signal (A1) can also be divided into two levels—close (A2) and taking into account details (D2), and then to close part of the process is repeated (A3 and D3).

$$s = A1 + D1 = A2 + D2 + D1 = A3 + D3 + D2 + D1 \quad (8)$$

Approximate number of decomposition levels can be calculated by the formula:

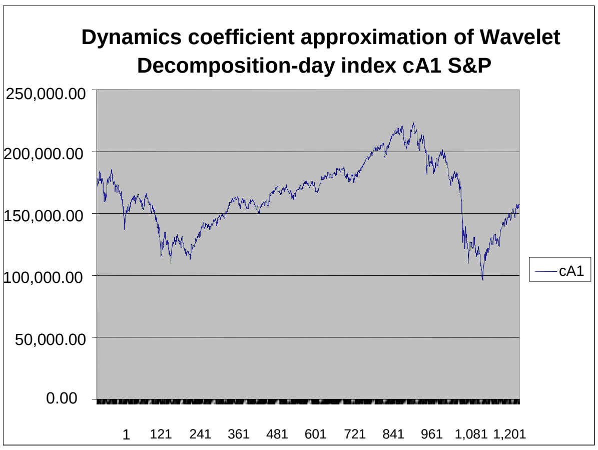
$$u = [\log 2\ (N)] \quad (9)$$

where N is the sample size $X(t)$, [ ] means taking the integer part.

Is believed that DWT allows you to get a good time resolution (low frequency) at high frequencies and a good resolution in frequency (bad timing) at low frequencies?

Practical Calculation of the Proposed Algorithm

To implement the DWT with a number arriving at Matlab2009 source three-level obtaining approximate cA1 (see Figure 14), cA2, and cA3, and detail coefficients sD1 (see Figure 14), cD2, and cD3.



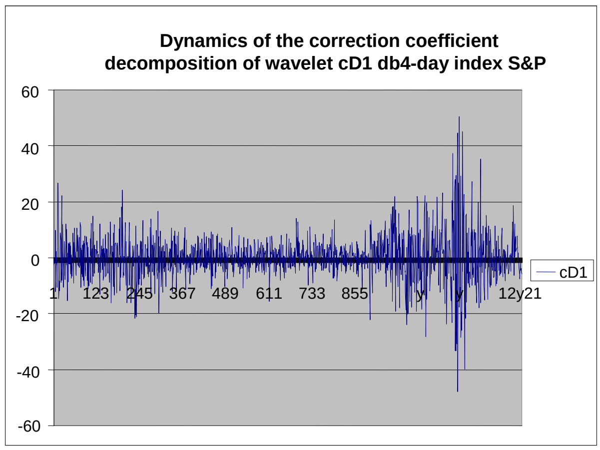


Figure 14. Dynamics coefficient approximation of cA1 and detailed implementing a series of one-day to cD1 factor "index of S&P".

Dynamics coefficient approximation cA2 and cA3, detail coefficients cD2 and cD3 in figures and tables are provided for rare, poor visibility, and save space.

Forecast Values of Coefficients of DWT

As the use of DWT series "S&P Index" with each level reduces the number of observations, the application of neural network and genetic methods to transform a one-dimensional array to multi-dimensional. Because here is the DWT with two levels and methods of AR (1) and Caterpillar (SSA).

Figure 15 shows the results of a prediction model for SSA and cA1 ratio approximation in Figure 16 detailed implementing rate cD1.

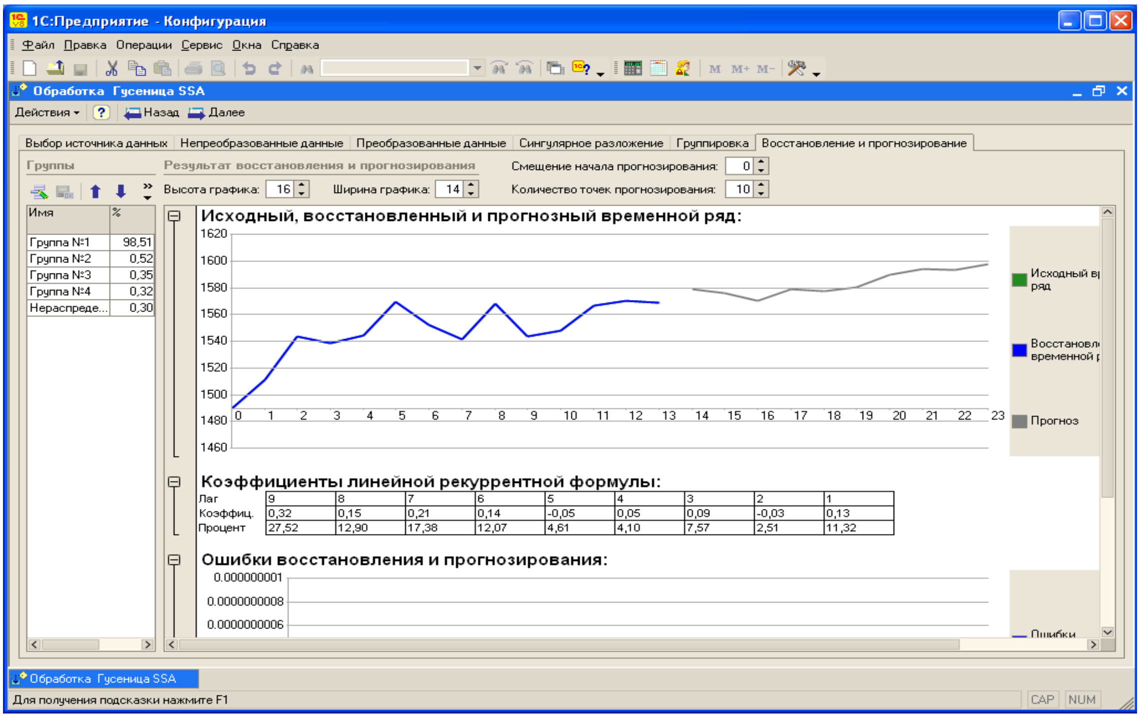


Figure 15. SSA model for approximation in cA1.

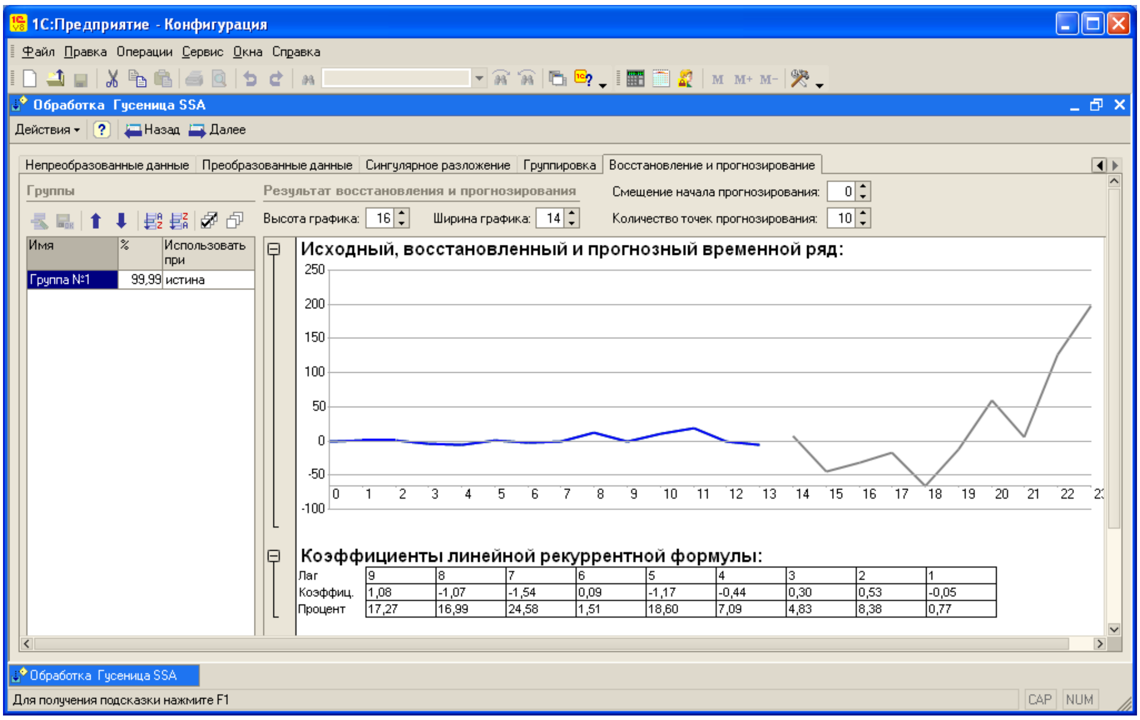


Figure 16. SSA model for detailed implementing rate sD1.

Other models (for cA2, cD2) to save space not illustrated.

Inverse Wavelet Transform of Predictive Coefficients DWT

Inverse wavelet transform produced using Matlab 2009 using the DWT. Results are presented in Figure 17 in conjunction with 13 observations from the last segment of the time series.

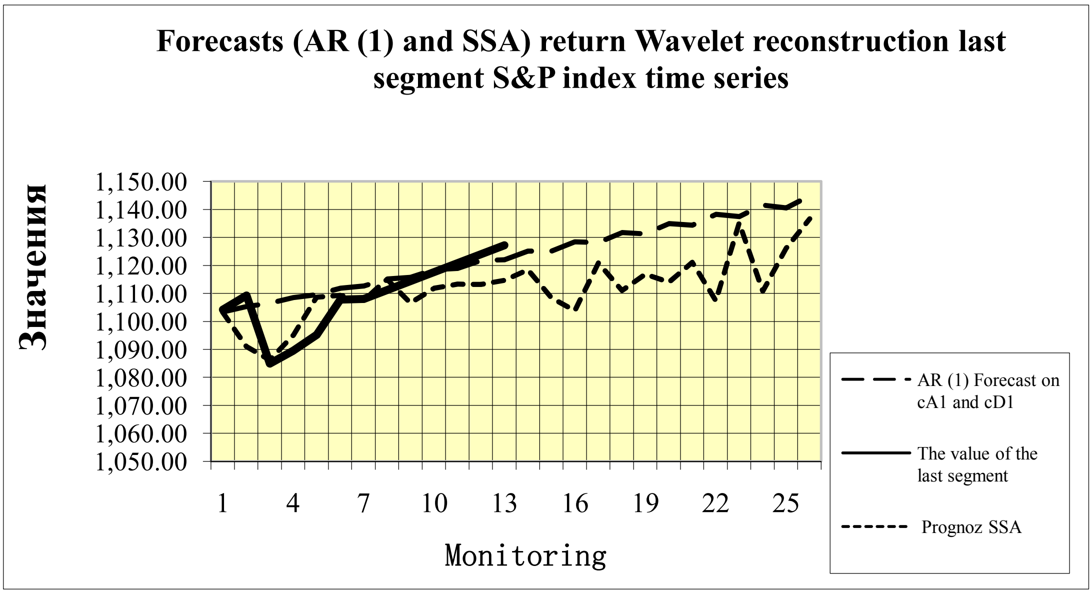


Figure 17. Joint chart values the last segment and forecast values from methods of AR (1) and the SSA and the inverse wavelet transform.

From a comparison of the results of the two methods with the use of DWT (see Figure 17), one can conclude the AR (1) linear prediction. Forecast method of SSA, by taking into account the large number of lags has a significant number of oscillatory components.

Conclusions

The main results of the work can be summarized as follows:

- Analysis of the contemporary approaches of non-stationary brings the stationarity of a variable substitution method: Hilbert-Huang transform and DWT;
- The results of the calculations of several dozen (about 20) financial series types of indices of prices for shares of large companies and prices of basic resources that showed they had very small length of segments with relatively stable settings of stationarity (about 10-20);
- A small length of segments homogeneous preference determines the alignment of the Hilbert-Huang transform stationarity, due to multiple cuts the length of the segment with increasing levels of DWT;
- After the installation of the stationarity of the short-term forecasting models is implemented using a variety of methods of non-linear forecasting: neural network, genetic, sliding fractal, and first-order autoregressive;
- Forecasting methods of the best were: neural network and genetic method when the length of the segment more than 20 fractal and sliding length of segment less than 20;
- Is a practical example of calculating the last method using prediction coefficient methods AR (1) and SSA;
- The methodology can be used to predict the wider class of non-stationary time series with pronounced trend-seasonal ingredients.

References

Bezruchko, B. P., & Smirnov, D. A. (2005). *Mathematical modeling and chaotic time series*. Soviet Union: Saratov State University Press.

Chandra, K., You, C., Olowoyeye, G., & Thompson, C. (1998). Non-linear time-series models of Ethernet traffic. *Technical report*, CACT.

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory, 36*(5), 961-1005.

Davydov, V. A., & Davydov, A. V. (2010). *Cleaning of geophysical data from the noise using the Hilbert-Huang transform*. Retrieved from http://www.actualresearch.ru

Dobeši, I. (1999). Wavelets and other methods of localization in phase space. Proceedings from *the International Congress of Mathematicians in Zurich* (pp. 84-108). Switzerland.

Flandrin, P. (1999). *Time-frequency/Time-scale analysis* (p. 386). San Diego: Academic Press.

Golyandina, N. E. (2004). *Method caterpillar (SSA): Time series analysis* (p. 76). St. Petersburg: St. Petersburg State University.

Gröchenig, K. (2001). *Foundations of time-frequency analysis: Applied and numerical harmonic analysis*. Boston: Birkhäuser.

Huang, H. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., … Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A., 454,* 903-995

Kolodyazhny, A. N. (2006). The use of wavelets to detect singular emissions trajectory measurements. *Computer Science, Management and Computer Technology, 1,* 95-100.

Louis, Y. L. (2008). Intrinsic mode decomposition of time series data with sawtooth transform. Patent Application No. 20090037147.

Lûbušin, A. A. (2006). Exploratory analysis of the properties of time series: On the basis of an interactive program spectra-analyzer (Uchebnoe guide for senior students of the Geophysical Department, Russian State Geological Prospecting University, Sergo Ordzhonikidze, Moscow).

Mallat, S., & Hwang, W. L. (1992). Singularity detection and processing with wavelets. *IEEE Trans. Info. Theory, 38*(2), 617-642.

Orlov, G. N., & Osminin, K. P. (2011). *Non-stationary time series*. Moscow: Book House Librokom.

Osminin, K. P. (2008). Algorithms for predicting non-stationary time series (Dissertation of the candidate of physical and mathematical sciences, Institute of Math. Modelling, Moscow).

Perminov, G. I. (2011). Methods for determining the width of the window in non-stationary time series. Proceedings from *the International Conference Problems of Development of the Social Sciences: Issues, Decisions and Perspectives* (pp. 155-169). Volgograd, Soviet Union.

Perminov, G. I. (2012). Experimental determination of length of memory in the financial series. *Financial Research, 1*(34), 110-121.