CHAPTER 1. ENGINEERING SCIENCES

FINANCIAL MARKETS RESEARCH. METHODS AND PERSPECTIVES OF THE ANALYSIS

Sytnyk Volodymyr[1]
Georgalina Olena[2]

DOI: dx.doi.org/10.30525/978-9934-571-30-5_1

Abstract. In the context of globalization, the very high dependence of the real sector of the economy and the welfare of individuals on the processes taking place in financial markets can be observed. Theoretical description of mass expectations of participants in the financial market and forecasting of quotations on this basis require the use of more effective formal methods. In this article the stock and currency markets are examined, since these markets are the closest to the state of information efficiency. This state is a goal for sovereign regulators of financial markets, that is reflected, in particular, in the documents of the International Organization of Securities Commissions (IOSCO), which makes developed markets the most interesting, "clean" subject of research. The subject of the study are time series of financial indicators – the value of securities and currency pairs. The aim of the work is to develop a combined methodology for analyzing the descriptive characteristics of financial markets. Since there is currently no method that fully allows analyzing and forecasting financial indicators of market relations, it is necessary to develop a stack of models and the corresponding forecasting methodology. The new methodology and the corresponding method should have a speed of calculation of predicted values, that is comparable with the analogues, and more accurate values of forecasting financial indicators in comparison with analogues.

[1] Candidate of Physical and Mathematical Sciences, Associate Professor,
Institute of Computer Systems, Odessa National Polytechnic University, Ukraine
[2] Candidate of Technical Sciences, Associate Professor,
Military Academy of Odesa, Ukraine

To achieve the goal, the following tasks are set and solved:

1. To carry out the best possible review of models and methods for forecasting financial indicators, to identify the advantages and disadvantages of each class of models. To identify the most commonly used classes of forecasting models and their main shortcomings, to identify promising approaches that help to address the model deficiencies.

2. To develop a new forecasting technique that addresses the identified shortcomings.

3. To perform for the developed methodology (stack of models) software implementation of algorithms.

4. To evaluate the effectiveness of the proposed methodology for solving the problem of forecasting the securities market and currency pairs.

The subject of the study are financial indicators of the securities market and currency pairs. The subject matter are methods and models for forecasting financial indicators. For the tasks at hand the methods of Data Mining, mathematical modeling, adaptive forecasting, exponential smoothing of models, methods of object-oriented programming were used in the work.

The following main results are obtained.

1. A combined methodology for assessing the financial performance of the securities market and currency pairs, related to the class of regression models with a base of partial description in the class of polynomial ones, was developed.

2. A set of algorithms for extrapolating time series, for identification of the model and obtaining predicted values was developed.

3. The forecasted values of financial indicators of the securities market and currency pairs, that confirming the effectiveness of the developed model are obtained.

The developed model and prediction technique can be used to predict time series of different subject areas. The developed algorithms for extrapolating time series with and without external factors are evident for software implementation. The prospect of further research is the optimization of the parameters of the obtained class of models through inductive modeling, the use of parameters of the obtained class of models for effective application of neural network technologies.

1. Introduction

Information about the dynamics of exchange rates of securities or currencies creates the impression of a chaotic movement, the direction of which is constantly changing under the influence of irregular and often unknown forces. The investigated object is fully exposed to the elements of the world market, and there is no precise information on the future of the course. You need to make a prediction. At the same time, it is quite obvious that it is very difficult to forecast even the sign of the rate change. Usually it is a matter of experts who analyze the current situation, as well as try to identify factors that are regularly associated with the movement of the course (this is a fundamental analysis). In constructing formal models, we are also trying to identify a range of significant factors and on their basis to construct any indicator or predictor.

However, neither expert practitioners nor formal methods have yet produced sustainable good results. This is due primarily to the fact that if there is indeed a row of factors that affect the course in a stable way, then their influence is reliably hidden both by the imposed random component and by the control actions of central banks and large financial players. As a result, these factors and their influence are difficult to identify.

It is possible to consider short-term forecasting of a course as a matter of fact a problem of forecasting of consecutive movement of the isolated time series which reason is mainly the mass behavior of small businessmen and large financial players which do the basic volume of financial operations in the foreign exchange market. This approach can be attributed to the so-called technical analysis.

This makes it possible to the time series only the model is based on without using additional information, and all the arguments about the mass behavior of market participants should be used only for qualitative interpretation. In this case, has been given such the task.

• First, find out the possibility of applying for short-term courses forecasting of any statistical methods which purpose is to describe recurring events or situations that are characterized by relatively stable links.

• Secondly, if statistical methods can be applied to solve the task, then to establish their most promising class, to indicate the characteristic features of these methods, special attention should be given to the simple ones.

• Thirdly, to show the practical results of at least some attempts.

2. Comparison of models

Regression models and methods. The advantages of these models include simplicity, flexibility, as well as the uniformity of their analysis and design. Using linear regression models, the prediction result can be obtained faster than using other models. In addition, the advantage is the transparency of modeling, that is, the availability for analysis of all intermediate calculations. The main disadvantage of nonlinear regression models is the difficulty in determining the type of functional dependence [1], as well as the laboriousness of determining the parameters of the model. The disadvantage of linear regression models is low adaptability.

Autoregressive models and methods. Important advantages of this class of models ($ARIMAX$) are their simplicity and transparency of modeling. Another advantage is the uniformity of analysis and design. This class of models is one of the most popular now, and therefore it is easy to find examples of application of autoregressive models for solving problems of forecasting time series in different subject areas. The disadvantages of this class of models are: a large number of model parameters, the identification of which is ambiguous, and the resource intensity; low adaptability and linearity of models and, as a consequence, the impossibility of modeling non-linear processes, often encountered in practice.

Models and methods of exponential smoothing. The advantages of this class of models is the simplicity and uniformity of their analysis and design. This class of models is most often used for long-term forecasting. The disadvantage of this class of models is the lack of flexibility.

Artificial neural network models and methods. The main advantage of artificial neural network models (ANN) is their non-linearity, other important advantages are: adaptability, scalability (parallel structure of ANN speeds up calculations) and uniformity of their analysis and design [3; 4]. The disadvantages of ANN are the lack of transparency in modeling, the complexity of the choice of architecture, the high requirements for the consistency of the training sample, the complexity of the choice of the learning algorithm, and the resource intensity of the learning process [1, p. 122-159].

Simplicity and monotony of analysis and design are the advantages of models based on Markov chains. The disadvantage of these models is the lack of the possibility of modeling processes with a long memory [5].

Models based on classification-regression trees. The advantages of this class of models are: scalability, which allows fast processing of extremely large data sets, speed and unambiguousness of the learning process of the tree (in contrast to ANN), as well as the ability to use categorical external factors. Disadvantages of these models is the ambiguity of the algorithm for constructing the tree structure [6].

It should be noted additionally that for none of the groups of models (and methods) considered, the advantages of forecasting accuracy are indicated. This is done due to the fact that the accuracy of forecasting a particular process depends not only on the model, but also on the researcher's experience, on the availability of data, on the hardware capacity available to the researcher, and many other factors.

Combined models. One of the popular modern trends in forecasting is the creation of combined models and methods. This approach allows us to compensate for the shortcomings of some models with the help of others and is aimed at improving the accuracy of forecasting as one of the main criteria for the effectiveness of the model.

In the review of prediction models [7], the following types of combinations are considered:

– ANN + fuzzy logic;
– $ANN + ARIMA$;
– $ARIMA$ + regression;
– $ANN + GA$ (genetic algorithm) + fuzzy logic;
– regression + fuzzy logic.

In most combinations, based models ANN are used to solve the clustering problem, and further for each cluster, a separate forecasting model based on $ARIMA$, GA, fuzzy logic, and the like is constructed. The application of combined models is a direction that, with a correct approach, can improve the accuracy of forecasting. The main disadvantage of combined models is the complexity and resource intensity of their development. A number of researchers followed an alternative path and developed autoregressive models based on the assumption that the time series is a sequence of repeating clusters (patterns). However, the developers did not create combined models, but determined the clusters and performed the forecast based on one model. In [8], a model for predicting the direction of movement of market indices (index movement), which takes into account clusters of the time series, is proposed. In this model it is assumed that if at some point in time in the past the market behaved in a certain way, then in the future its behavior will be repeated in connection with the fact that the time series is a sequence of clusters.

Two more works [9; 10] proposed a prediction model based on the autoregressive model, which takes into account the fragments of the time series. Here, the predicted value of the time series is determined by the expression $Z(t)=\alpha_0+\alpha_1 Z(t-1)+...+\alpha_M Z(t-M)$, which is a linear autoregressive order M. The coefficients of autoregression $\alpha_0,\alpha_1,...,\alpha_M$ are determined as follows. It is assumed that there are K vectors of the length M of the time series for which expression

$$\begin{pmatrix} Z(i_1) \\ Z(i_2) \\ ... \\ Z(i_k) \end{pmatrix} = \alpha_0 \begin{pmatrix} 1 \\ 1 \\ ... \\ 1 \end{pmatrix} + \alpha_1 \begin{pmatrix} Z(i_1-1) \\ Z(i_2-1) \\ ... \\ Z(i_k-1) \end{pmatrix} + ... + \alpha_M \begin{pmatrix} Z(i_1-M) \\ Z(i_2-M) \\ ... \\ Z(i_k-M) \end{pmatrix}$$

To determine the closest vectors $Z(i_1-1)$, $Z(i_1-2),...,Z(i_1-M),...,$ $Z(i_K-1),Z(i_K-2),...,Z(i_K-M)$, the value of the Pearson linear correlation between all possible vectors and the last available vector $Z(t-1)$, $Z(t-2),...,Z(t-M)$ is used in [9], and in [10] instead of linear correlation the Euclidean distance between vectors is calculated.

The developers of the models discussed above assert that the proposed models are simple, transparent and effective for the time series studied. It is obvious that the main disadvantages of these models are:

– impossibility to take into account external factors;

– the ambiguity of the criterion for determining a similar sample;

– the complexity of determining the effective combination of two parameters: (the length of the vectors) and (the number of vectors accepted for calculation) in [9, 10].

3. Statistical models of short-term forecasting

The object of the study is a time series of consecutive data, which we will write as $x_1,...,x_n$, where n is the length of the series. The task is to identify the presence of the dependence of the observation with i-nomber on the previous ones, and on this basis to make a forecast for ($n+1$) – moment.

Before starting the study of the mechanism of the connection between successive values of the series, it should be clarified that the series is not random, (so-called white noise), in which there is no dependence between the values of the series relating to different moments and the direction of motion of which in the future are equally probable, that is unpredictable. To do this, the selected series should be checked using randomness criteria, and as a result, get the first idea of the source material. As criteria for randomness, the following criteria are used: the criterion of turning points, the criterion for the distribution of the phase length, the criterion based on the signs of the differences, and the criterion based on rank correlation.

We now pose such problems.

• Explore the possibility of using statistical methods for short-term forecasting of courses;

• Establish their most promising class, identify the characteristics;

• Obtain a numerical analysis of financial indicators.

It is believed that the characteristic feature of adaptive prediction methods is their ability to continuously take into account the evolution of the dynamic characteristics of the processes under study, to “adapt” to this evolution, providing, in particular, the greater weight and the higher the information value of the available observations, the closer they are to the current moment of forecasting . However, the separation of methods and models into “adaptive” and “nonadaptive” is relatively arbitrary. In this sense, any forecasting method is adaptive, because they all take into account new information, including observations made since the last forecast. The general meaning of the term is that “adaptive” prediction allows you to update forecasts with minimal delay and with the help of simple mathematical procedures. The formulation of the prediction problem using the simplest version of the method of exponential smoothing is formulated as follows.

Let the time series $x_\tau,\ \tau = \overline{1,t}$ be represented in the simplest form

$$x_\tau = a_0 + \varepsilon_\tau \tag{*}$$

where is a_0 an unknown parameter that does not depend on time, but is ε_τ a random remainder with an average value of zero and a finite variance. As is known, the exponentially weighted moving average $\bar{x}_t(\lambda)$ of a series x_τ at a point with t a smoothing parameter (the adaptation parameter) $\lambda\ (0 < \lambda < 1)$ is given by

$$\bar{x}_t(\lambda) = \frac{1-\lambda}{1-\lambda^t} \sum_{j=0}^{t-1} \lambda^j x_{t-j}\,, \tag{1}$$

which gives the solution of the problem: $\bar{x}_t(\lambda) = \arg\min\limits_a \sum\limits_{j=0}^{t-1} \lambda^j \left(x_{t-j} - a\right)^2$.

The smoothing factor λ can also be interpreted as a discount factor, which characterizes the degree of depreciation of the observation per unit time. For series with an “infinite past”, formula (1) reduces to the form

$$\bar{x}_t(\lambda) = (1-\lambda)\sum_{j=0}^{\infty} \lambda^j x_{t-j} \,. \tag{2}$$

According to the simplest version of the method of exponential smoothing, the forecast $\hat{x}_t^1$ for an unknown value x_{t+1} from the trajectory of the series known by the time is constructed from

$$\hat{x}_t^1 = \bar{x}_t(\lambda) \tag{3}$$

where the value $\bar{x}_t(\lambda)$ is defined by formula (1) or (2), respectively for a short or long time series.

Formula (3) is convenient, in particular, in that when the next $(t+1)$ observation x_{t+1} occurs, the recalculation of the predictive function $\hat{x}_{t+1}^1 = \bar{x}_{t+1}(\lambda)$ of the next value is made using a simple relationship: $\bar{x}_{t+1}(\lambda) = \lambda\bar{x}_t(\lambda) + (1-\lambda)x_{t+1}$. The method of exponential smoothing can be generalized to the case of a polynomial nonrandom component of the time series, that is, to a situation where instead of (*) it is postulated

$$x_{t+\tau} = a_0 + a_1\tau + a_2\tau^2 + \ldots + a_k\tau^k + \varepsilon_\tau \tag{**}$$

where $k \geq 1$. In the ratio (**), the starting point of time is shifted at the current time t, which facilitates further calculations. Accordingly, in the scheme of the simplest version of the method, the forecast $\hat{x}_t^1$ of the value x_{t+1} will be determined by the relations

$$\hat{x}_t^1 = \bar{x}_{t+1}(\lambda) = \hat{a}_0(t,\lambda) + \hat{a}_1(t,\lambda) + \ldots + \hat{a}_k(t,\lambda),$$

where the estimates $\hat{a}_j(t,\lambda)$, $j = \overline{0,k}$ are determined from the solution of the optimization problem

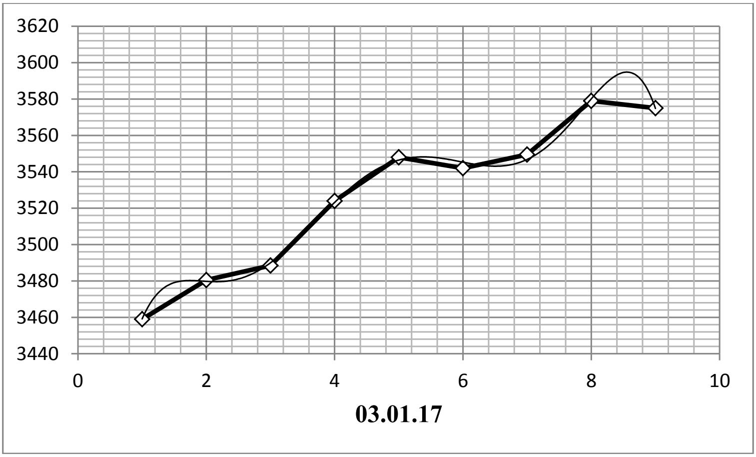
$$\sum_{j=0}^{\infty} \lambda^j \left(x_{t-j} - a_0 - a_1 j - \ldots - a_k j^k\right)^2 \to \min_{a_0,\ldots,a_k} \,. \tag{4}$$

The solution of problem (4) reduces to solving a system of $k+1$ linear equations and does not present any fundamental difficulties.Using the correlation-regression analysis, we determine the length t and order k of the best approximation sample. In this case, the problem

$$\sum_{j=0}^{t-1} \left(x_{t-j} - a_0 - a_1 j - \ldots - a_k j^k\right)^2 \to \min_{a_0,\ldots,a_k} \tag{5}$$

is solved.

Here, the parameters t and k are chosen according to the maximum of the determination coefficient. For a number of stocks of LUKOIL, it is established that the best values are $t = 9$, $k = 6$ (Fig. 1).



y = -0,0864x^6 + 2,5246x^5 – 28,543x^4 + 157,17x^3 – 435,91x^2 + 583,53x + 3180,4

R^2 = 0,9983

Fig. 1.

The solution of problem (5) reduces to solving a system of linear algebraic equations

$$\begin{cases} ta_0 + \sum\limits_{j=0}^{t-1} ja_1 + \ldots + \sum\limits_{j=0}^{t-1} j^k a_k = \sum\limits_{j=0}^{t-1} x_{t-j}, \\ \sum ja_0 + \sum j^2 a_1 + \ldots + \sum j^{k+1} a_k = \sum jx_{t-j}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \sum j^k a_0 + \sum j^{k+1} a_1 + \ldots + \sum j^{2k} a_k = \sum j^k x_{t-j}. \end{cases}$$

As a result, we obtain the first approximation of the parameters $\tilde{a}_j(t)$, $j = \overline{0,k}$ of the approximation model and the time-series model

$$x_{t+\tau} = \tilde{a}_0(t) + \tilde{a}_1(t)\tau + \ldots + \tilde{a}_k(t)\tau^k .$$

Next, we adapt this model to account for the next x_t point x_{t+1}. For this we solve problem (4), which leads us to the solution of a system of linear algebraic equations

$$\begin{cases} \sum_{j=0}^{t-1} \lambda^j a_0 + \sum_{j=0}^{t-1} j\lambda^j a_1 + \ldots + \sum_{j=0}^{t-1} j^k \lambda^j a_k = \sum_{j=0}^{t-1} \lambda^j x_{t-j}, \\ \sum j\lambda^j a_0 + \sum j^2 \lambda^j a_1 + \ldots + \sum j^{k+1} \lambda^j a_k = \sum j\lambda^j x_{t-j}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \sum j^k \lambda^j a_0 + \sum j^{k+1} \lambda^j a_1 + \ldots + \sum j^{2k} \lambda^j a_k = \sum j^k \lambda^j x_{t-j}. \end{cases}$$

The solution of this system gives us $\hat{a}_j(t,\lambda)$, $j = \overline{0,k}$ a forecast $\hat{x}_t^1 = \bar{x}_{t+1}(\lambda) = \hat{a}_0(t,\lambda) + \hat{a}_1(t,\lambda) + \ldots + \hat{a}_k(t,\lambda)$ of the value x_{t+1} (Fig. 2).

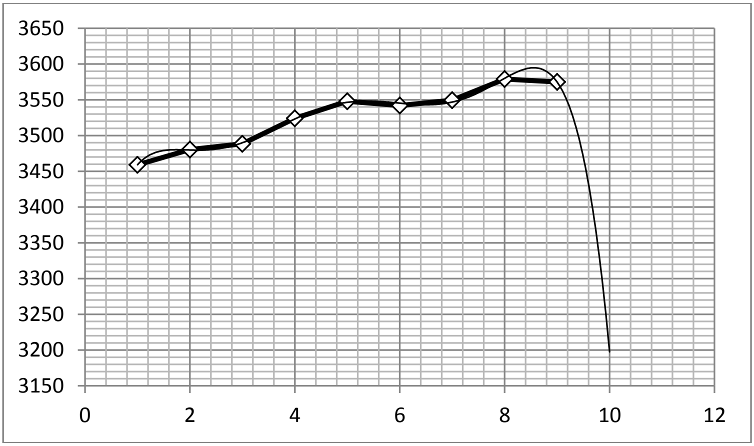
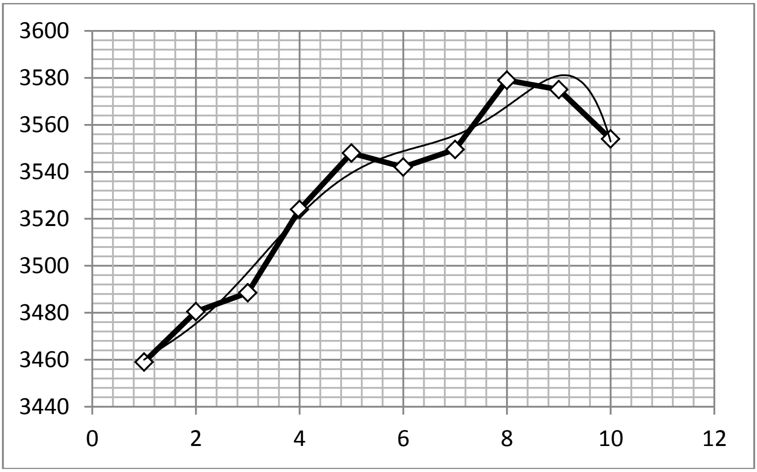


Fig. 2.

After obtaining the true value x_{t+1}, we do the recalculation of the predictive function $\hat{x}_{t+1}^1$ using the formula $\bar{x}_{t+1}(\lambda) = \lambda \bar{x}_t(\lambda) + (1-\lambda) x_{t+1}$, that is, we adapt the model to a new value (Fig. 3).

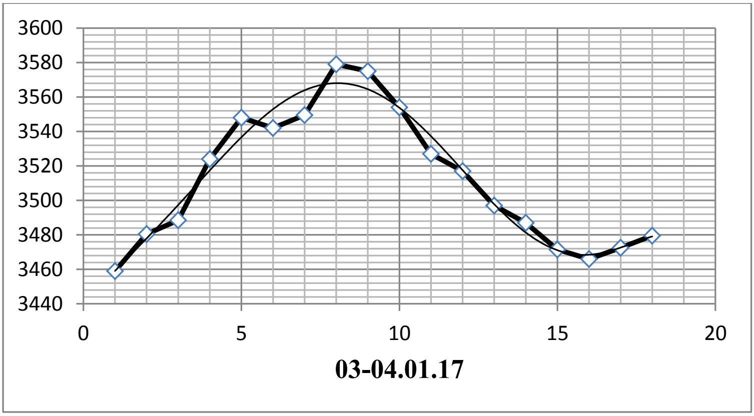
We repeat the adjustment of the model, while the coefficient of determination of the model remains smaller $0,9$. As a result, we get a fairly smoothed model, which approximates the time series fragment within the limits $\tau = \overline{1, t+M}$. For a number of securities of LUKOIL, for example, $M = 18$ (Figure 4.5).

In the future it is advisable to apply the sliding window method for forecasting, that is, without changing the length of the best likelihood sample, add this value with simultaneous removal of the first one (Fig. 6). Another



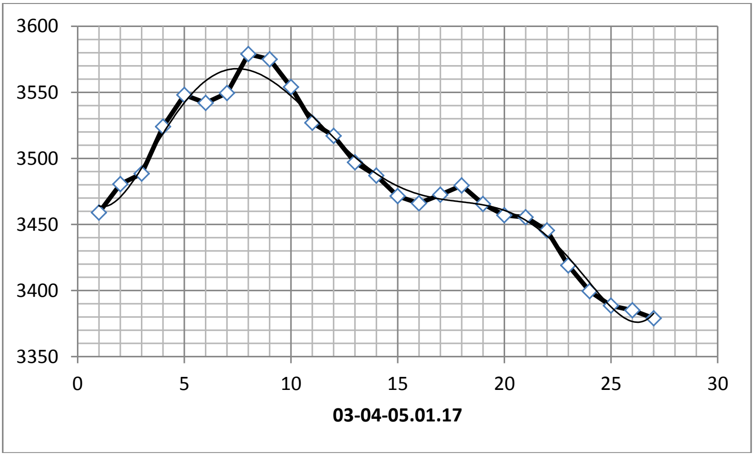
$$y = -0{,}0122x^6 + 0{,}3486x^5 - 3{,}7266x^4 + 18{,}237x^3 - 40{,}801x^2 + 56{,}036x + 3429{,}9$$
$$R^2 = 0{,}9719$$

Fig. 3.



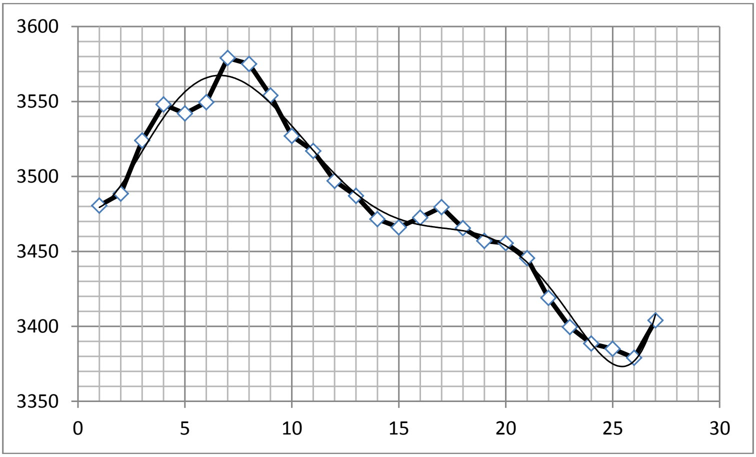
$$y = -0{,}0002x^6 + 0{,}0126x^5 - 0{,}2219x^4 + 1{,}5087x^3 - 4{,}2905x^2 + 24{,}116x + 3438$$
$$R^2 = 0{,}9629$$

Fig. 4.



$$y = 1E\text{-}04x^6 - 0{,}0085x^5 + 0{,}2719x^4 - 4{,}0097x^3 + 25{,}556x^2 - 45{,}71x + 3488$$
$$R^2 = 0{,}9796$$

Fig. 5.

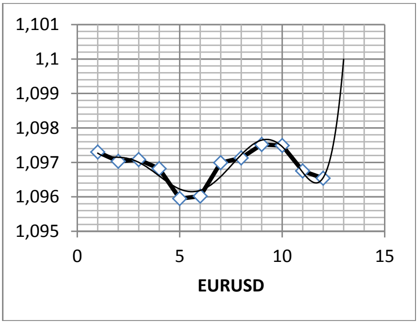


$$y = 0{,}0001x^6 - 0{,}0084x^5 + 0{,}2526x^4 - 3{,}4269x^3 + 19{,}065x^2 - 22{,}153x + 3485{,}7$$
$$R^2 = 0{,}9816$$

Fig. 6.

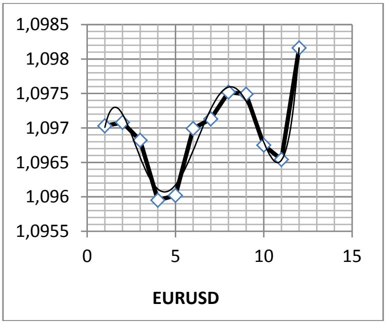
option is to reduce the adaptation parameter λ, but its reduction will also worsen the basic values of the model parameters.

It should be noted that the verification of the proposed adaptive model on currency pairs EURUSD, GBRUSD confirmed its effectiveness for $t = 12$, $k = 6$ (Fig. 7-10) and the variable sliding window method, which starts immediately with $M = 0$.



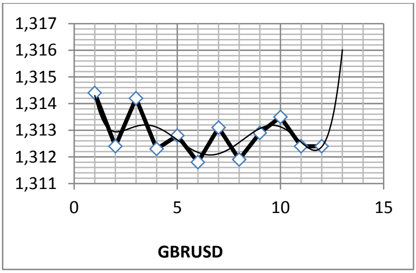
$y = 2E\text{-}07x^6 - 9E\text{-}06x^5 + 0{,}0001x^4 -$
$- 0{,}0007x^3 + 0{,}0022x^2 - 0{,}003x + 1{,}0987$
$R^2 = 0{,}8702$

Fig. 7.



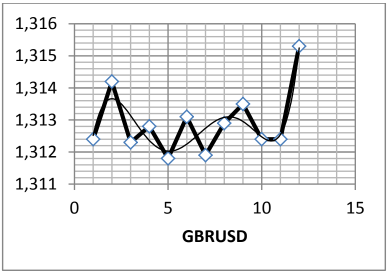
$y = 1E\text{-}08x^6 + 1E\text{-}06x^5 - 5E\text{-}05x^4 +$
$+ 0{,}0005x^3 - 0{,}0026x^2 + 0{,}0049x + 1{,}0942$
$R^2 = 0{,}9351$

Fig. 8.



$y = 4E\text{-}07x^6 - 2E\text{-}05x^5 + 0{,}0003x^4 -$
$- 0{,}002x^3 + 0{,}0074x^2 - 0{,}0132x + 1{,}3219$
$R^2 = 0{,}5014$

Fig. 9.



$y = 6E\text{-}08x^6 + 1E\text{-}07x^5 - 4E\text{-}05x^4 +$
$+ 0{,}0006x^3 - 0{,}0035x^2 + 0{,}008x + 1{,}3074$
$R^2 = 0{,}7283$

Fig. 10.

This is due to the less stable series of currency pairs. Here, it is possible to increase the order of the polynomial and, correspondingly, increase the number of parameters of model adaptation, which, however, is undesirable, since it leads to an increase in computational errors.

Another way to enhance the adaptation of the model to new values of the time series is to use the Holt method, that is, smoothing both the trend and the level of values of the time series.

4. Modeling by the method of selection of arguments

Mathematical models of various complex systems are also formed by means of selective choice from the best variants by the Group Method of Data Handling shell (GMDH) on the computer. The principle of model self-organization, underlying the GMDH, asserts that the optimal model corresponds to a minimum of an external criterion or a couple of criteria selected by the author of the model. Other grounds for the application of the method are the theorem of Gödel's incompleteness and the principle of preserving the freedom of choice of D. Gabor. The theory is completed by a new principle of multilevel modeling in several languages, differing in the level of detail [11].

There are observational data $Y_e(x)$. To describe them, we need to build a model $Y_m = F(a_0, a_1, ..., a_k, x)$. Here e – experiment, m – model, $a_0, a_1, ..., a_k$ – model parameters, F – model structure. Let the model be a polynomial: $Y_m = a_0 + a_1 x + ... + a_k x^k$. Consider the situation when the model is not specified (the order k of the polynomial and the parameters $a_0, a_1, ..., a_k$ of the model are not defined). It is necessary to find k, $a_0, a_1, ..., a_k$, $\min_k K(Y_e, Y_m)$, where K – a criterion reflecting the predictive properties of the model. This problem is the task of F structural optimization. To solve this problem, we apply the group method of data handling (GMDH). GMDH allows you to select the optimal complexity model from a given class of models to describe the available set of experimental data. The model constructed on these data does not contradict them, and therefore can be called plausible. GMDH has advantages when there is no or almost no a priori information about the structure of the model and the distribution of its parameters. GMDH also has advantages when there is very little observational data, up to the point that the model parameters are larger than the number of observations.

Tasks solved with the help of GMDH.

• Approximation of table-defined functions.
• Description and prediction of time series.
• Choice of a computational scheme for solving differential equations.
• Structural identification of objects (input-output models).
• Classification / pattern recognition (search for structures with the teacher).
• Clustering / taxonomy (searching for structures without a teacher).
• Setting the structure of neural networks (self-organization of the structure).

Subject domains of GMDH.
• Geophysics (description of spatial distribution of geo-parameters).
• Computer linguistics (verification of similarity of words, construction of dictionaries).
• Business applications (forecast of time series, grouping of similar economic objects).

Classes of models. Let us consider experimental data, which must be described by a formula about which there is no a priori information. In this case, inductive modeling (IM) is used. IM deals with a pre-fixed class of models. The class of models depends on the problem under consideration. It can be:
– polynomials and trigonometric polynomials of one variable;
– linear and nonlinear functions of many variables;
– clusters of objects, etc.

The complexity of the model also depends on the task. It can be:
– number of parameters to be evaluated;
– number of clusters, etc.

An example of a class of models is a polynomial of one variable (t):

$$Y_0 = a_0 ,$$

$$Y_1 = a_0 + a_1 t \quad Y_1 = a_0 + a_2 t^2 \ \ldots\ldots\ldots \ Y_1 = a_0 + a_{100} t^{100} ,$$

$$Y_2 = a_0 + a_1 t + a_2 t^2 \quad Y_2 = a_0 + a_1 t + a_3 t^3 \ \ldots\ldots\ldots \ Y_2 = a_0 + a_{99} t^{99} + a_{100} t^{100} ,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Calculation of model parameters. When the model is specified, its parameters are selected according to the data. For this, a system of equations is constructed and solved. Let us look for the function of the ith variable in the form of a polynomial. Let there be observational data at the points (0.1, 0.2, 0.3, …). Suppose that the maximal degree of a polynomial is bounded $k \le 10$. Let at the given stage of a search of models only models of the second order are considered. Then we have:

$$Y_2 = a_0 + a_1 t + a_2 t^2 \;\ldots\ldots\ldots\; Y_2 = a_0 + a_9 t^9 + a_{10} t^{10} \quad \Rightarrow$$

$$\Rightarrow \quad Y_2(0,1) = a_0 + a_1 \cdot 0,1 + a_2 (0,1)^2 \;\ldots.\; Y_2(0,1) = a_0 + a_9 (0,1)^9 + a_{10} (0,1)^{10} \ldots$$

The parameters $a_0, a_1, ..., a_k$ of the model are determined from the internal quality criterion of the model: $\min\limits_{a_i} \|Y_e - Y_m\|$, where $\|Y_e - Y_m\|$ is the norm of distances.

As the norm of distances are used:

a) the method of least squares: $\sum (Y_e - Y_m)^2 \to \min$ if the parameters enter linearly: $Y_m = a_0 + a_1 t + a_2 t^2 + ...$

b) nonlinear programming, if the parameters enter nonlinearly: $Y_m = a_0 + a_1 \cos(b_1 t) + a_2 \cos(b_2 t) + ...$

Requirements of GMDH. Requirement 1: resistance to new data. We must ensure a good predictive property of the model, that is, resistance to new data.

Measure the stability of the model. Since the sensitivity of the model should be checked on the data novices, it is natural to split the entire set of values into two parts.

– The first data set will be used to build the model. This data set is called Training data.

– The second set of data will play the role of new data and it will be used to verify the quality of the model built. This data set is called Checking data.

To implement this idea, one must have a criterion that would assess the quality of the model on the new data. This criterion is called the criterion of regularity. It is an external criterion for the quality of the model.

Requirement 2: independence of the description of the model from the data. We must provide a good descriptive property of the model, that is, the independence of the description of the model from the data. Since the independence of the data must be checked by constructing the model on different data, it is natural to split the entire set of values into two parts.

– The first data set will be used to build one model. The second set will be used to build another model. Both sets are equal.

– As a first set, you can take the Training data.

– As a second set, you can take the Checking data.

To implement this idea one must have a criterion that would allow two models to be compared. This criterion is called the unbiasedness criterion. It is an external criterion for the quality of the model.

Steps of GMDH.

1. Determine the class of models of increasing complexity.
2. Divide the experimental data into two parts: Training and Checking.
3. For a given level of complexity, estimates of the model parameters on the first data set are found. An internal criterion is used.
4. This model is tested on the second set of data. Somean external criterion (here the regularity) is used.
5. If the external criterion reaches a minimum, then STOP, otherwise we increase the complexity of the model, and go to step 3.

The regularity criterion (K_{reg}) can be taken in the form

$$K_{reg} = \sum_{C} \left(Y_e - Y_m(t) \right)^2 ,$$

where $\sum_{C}$ is the summation over the points of the Checking set.

Comment. One can normalize K_{reg} by dividing it into an expression $\sum_{C} Y_e^2$ .The model trained on the first data set and the model trained on the second set should be close (T – first set, C – second set).

Forms of the criterion for unbiasedness.

1) Criterion focused on decision analysis.

$$K_{unbias} = \sum \left(Y_m(T) - Y_m(C) \right)^2$$

Here: $Y_m(T)$ – the values of the model trained on the Training set, $Y_m(C)$ – the values of the model trained on the Checking set, $\sum$ – the summation over all points.

Comment. One can normalize K_{unbias} by dividing it into an expression $\sum Y_e^2$.

2) Criterion focused on the analysis of parameters.

$$K_{unbias} = \frac{1}{|a(T)| \cdot |a(C)|} \sum \left(a_i(T) \cdot a_i(C) \right) .$$

Here: $a_i(T)$ – Parameters of the model built on Training;
$a_i(C)$ – parameters of the model built on Checking;
$\sum$ – summation over all points.

Two criteria; convolution rule.

– Weights are assigned λ_1, λ_2 : $\lambda_1 + \lambda_2 = 1$ and a combined criterion is calculated $K = \lambda_1 K_{reg} + \lambda_2 K_{unbias}$.

– The model is chosen, the best one by the combined criterion (Fig. 11).

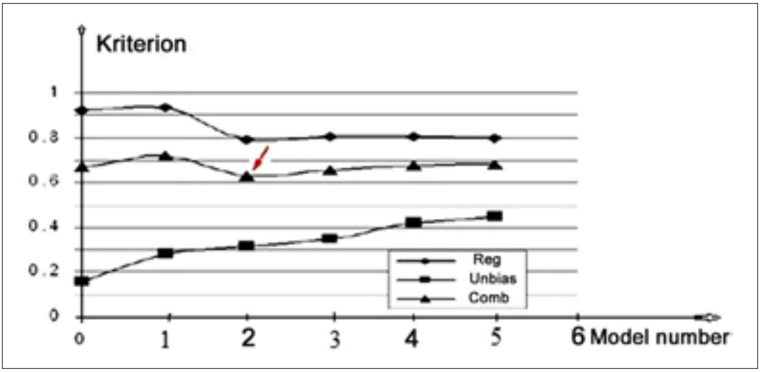


Fig. 11.

In this example, according to K_{reg} the best models are 2, 3, 4, 5, according to K_{unbias} – model 0. The combined criterion is selected by model 2.

Two criteria; consecutive selection.

Instead of selecting a model for a combined criterion $K = \lambda_1 K_{reg} + \lambda_2 K_{unbias}$, you can use a different strategy.

– Choose the best models on K_{reg} .

– The best ones are chosen from them on K_{unbias} .

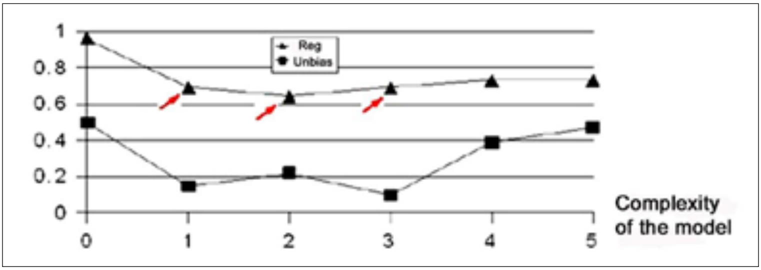


Fig. 12.

In this example, models 1, 2 and 3 are best on the criterion K_{reg} . The criterion K_{unbias} selects model 3 (Fig. 12).

Verification of the model. To implement external criteria, you need to have two sets of data.

– In the regularity criteria, they serve to build and verify the model, and are called Training Data and Checking Data.

– In the unbiasedness criterion, they serve to construct two variants of models that are tested on a full set.

However, the values of external criteria do not give an objective assessment of the quality of the model, they serve only for the selection of models. To assess the quality, you need to have one more independent data set – control sample, the so-called Verifying Data.

Specifically, on this sample, the quality / adequacy of the model is checked. Thus, in order to verify the model, the GMDH should use not two but three samples from the original data.

The partitioning does not have to be uniform. A typical example (V. S. Stepashko): Verifying 20%, Training: Checking = 2: 1.

There are 4 basic algorithms of the GMDH and many of their modifications.

o COMBI is a combinatorial algorithm.

o MULTI – combinatorial-selection algorithm.

o MIA is a multi-row iterative algorithm.

o RIA is a relaxation iterative algorithm.

All algorithms are multi-row.

Each series is associated with one level of complexity of the model. Obviously, in each row there can be several models. Algorithms differ in the conditions for the formation and selection of terms in the transition from one series to another. Algorithms have settings, among which the number of the best models of each series. Parameters are set by the user.

As an example, consider the RIA algorithm, from idea to illustration.

The main ideas of the RIA:

– reduce the number of models considered on each row, but not lose a successful combination of variables;

– reduce the number of rows, and thereby accelerate the output to the optimal level of complexity.

Therefore, on each series:

– a fixed number of best models is selected (each model is considered as a variable);

– each best model is combined with one variable of the initial set and this pair generates a new variable when going to the next level.

RIA – implementation. The algorithm has two parameters:

– the number of the best models in each series – p. For example, let $p=5$.

– the function of converting a pair of variables (Y,x) taken from one of the current series and from the original series into variables of the next series

$$Z(Y,x)=a_0+a_YY+a_xx$$

You can use other functions.

The number of the best models and the conversion function is set by the user.In the first series we have:

$$Y_1^{(1)}=a_0+a_1t \quad Y_1^{(2)}=a_0+a_2t^2 \ \ldots \ Y_1^{(20)}=a_0+a_{20}t^{20}.$$

Choose the top 5 models (because $p=5$), let it be, for example, $Y_1^{(3)},Y_1^{(5)},Y_1^{(7)},Y_1^{(9)},Y_1^{(11)}$. These models are treated as variables.

In the second row, functions from all combinations of selected variables and source variables, that is, models of the form

$Z\left(Y_1^{(j)},x_i\right)$, where $x_i=t^i$, $i=\overline{1,20}$, for example: $Z\left(Y_1^{(3)},x_5\right)=$ $=a_0+a_3Y_1^{(3)}+a_5x_5$.

The number of variables M in each series: $M=p\cdot k=100$ .RIA – illustration.

Fig. 13 shows the formation of new variables from the best variables of the current series and the initial set of variables

The application of the selection algorithm to the structure and parameters of the regression model has made it possible not only to clarify the values of parameters, but also to rate the quality of the results obtained. In particular, the quality of the forecast for LUKOIL securities was 67.65%, for the EURUSD series – 60.84%, and for the GBRUSD series – 60.53%. For comparison, in [12] for the LUKOIL series, the quality of adaptive forecasting is 60.2%.

In the future, it is possible to apply the GMDH algorithm to the structure of the neural network, which will eliminate problems of the uncertainty of the structure of internal neurons and allow to draw conclusions about the behavior of stock and financial markets more justified through more exact and effective instruments.

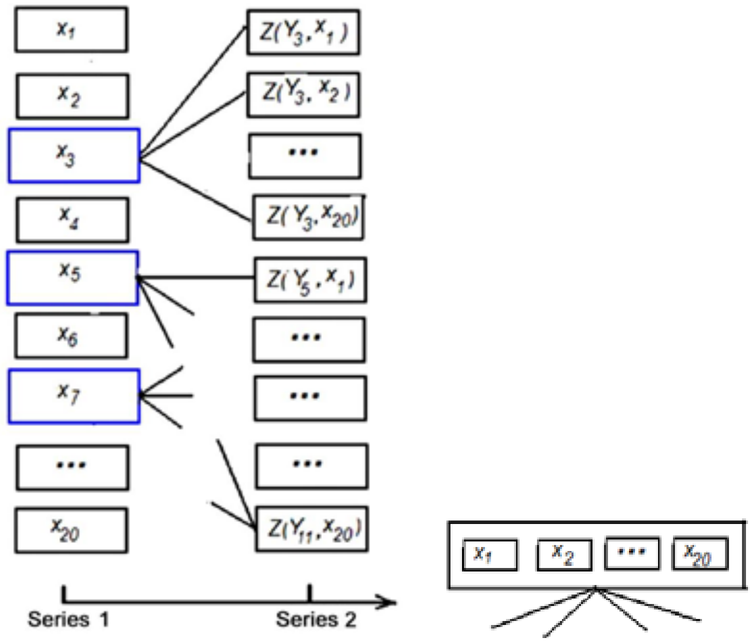


Fig. 13.

5. Conclusions

The most complete analysis of theoretical and calculated studies of market rates of financial indicators was carried out. The methodology for forecasting financial indicators is developed. The necessity of choosing a certain set of models for forecasting fluctuations in the rates of securities and exchange rates is substantiated. A complete analysis of a number of currency fluctuations for randomness has been performed. A preliminary analysis of the directions of the movement of securities and currency pairs was made and on the basis of this study economic efficiency indicators of the proposed models were formed. The selected models received the necessary adjustments with respect to the time series considered. In particular, the concept of the dynamic length of the sliding window of the statistical base is introduced and its necessity is justified. An information system for forecasting fluctuations in the exchange rate of securities and currency pairs was designed.

An analysis of the series of market fluctuations by regression methods was made, which led to the need to apply a nonlinear high-order model that had not previously been applied. At the same time, the reliability of the model was established at a fairly high level – at least 90%. In accordance with the level of reliability, the dynamic parameter of the statistical base was adjusted. This was particularly true for currency pairs, as the fluctuations in the rates of securities revealed sufficient resistance to changes in the statistical base. For securities in general it is established that for the forecasting of a course it is possible to create a fairly stable model with an easy correction of the coefficients of the model itself. As for currency pairs, their fluctuations are very unstable, therefore the size of the statistical base, that is constantly changing to maintain the quality of the model at a level of at least 90% of the reliability, is significant. As a result, for currency pairs, the number of forecast parameters increases by one in comparison with the forecast of the securities market.

Effective forecasts of the courses at the level of 60.53% for the adaptive model of non-linear forecasting of currency pair rates, 67.65% for the adaptive model of non-linear forecasting of the securities rate were obtained.

As a result, the high projected figures were obtained. These indicators are not lower than in [1, 12] and are without the drawbacks – the uncertainty in the choice of the degree of complexity of the model, as in neural networks, or the highly specialized application of the best likelihood sampling method. The proposed methodology works at a determination of not less than 90% and at the same time gives a forecast quality of at least 60%. The proposed methodology makes it possible in the future to easily proceed to minimize the number of model parameters, which is also an important prediction task.

References:

1. Tikhonov E.E. (2006) *Prognozirovanie v usloviyakh rynka* [Forecasting in market conditions]. Nevinnomyssk. (in Russian)
2. Mazengia D.H. Forecasting Spot Electricity Market Prices Using TimeSeries Models: Thesis for the degree of Master of Science in Electric PowerEngineering. Gothenburg. Chalmers University of Technology. 2008. 89 p.
3. Pradhan R.P., Kumar R. Forecasting Exchange Rate in India: An Application of Artificial Neural Network Model // Journal of Mathematics Research. 2010. Vol. 2. No. 4. P. Ill – 117.
4. Yildiz B., Yalama A., Coskim M. Forecasting the Istanbul Stock Exchange National 100 Index Using an Artificial Neural Network // An International Journal of Science. Engineering and Technology. 2008. Vol. 46. P. 36-39.

5. Zliu J., Hong J., Hughes J.G. Using Markov Chains for Link Prediction in Adaptive Web Sites // 1st International Conference 011 Computing in ajn Imperfect World. UK. London. 2002. P. 60-73.

6. Hannes Y.Y., Webb P. Classification and regression trees: A User Manual for IdentifyingIndicators of Vulnerability to Famine and Chronic FoodInsecurity // International Food Policy Research Institute. 1999. 59 p.

7. Alfares H.K., Nazeemddin M. Electric load forecasting: literature survey and classification of methods // International Journal of Systems Science. 2002. Vol. 33. P. 23-34.

8. Singh S. Pattern Modelling in Time-Series Forecasting // Cybernetics and Systems-AnInternational Journal. 2000. Vol. 31. No. 1. P. 49-65.

9. Scherer Perlin M. Nearest neighbor method // Revista Eletronica de Administracao. 2007. Vol. 13. No. 2. 15 p.

10. Fernandez-Rodriguez F., Sosvilla-Rivero S., Andrada-Felix J. Nearest-Neighbour Predictions in Foreign Exchange Markets // Fundacion de Estudios de Economia Aplicada. 2002. No. 5. 36 p.

11. Ivakhnenko A.G. (1975) *Dolgosrochnoe prognozirovanie i upravlenie slozhnymi sistemami* [Long-term forecasting and management of complex systems]. Kiev: Naukova dumka. (in Russian)

12. Chuchueva I.A. (2010) Prognozirovanie vremennykh ryadov pri pomoshchi modeli ekstrapolyatsii po vyborke maksimal'nogo podobiya [Forecasting time series using the model of extrapolation from a sample of maximum similarity]. Proceedings of the *Nauka i sovremennost: sbornik materialov Mezhdunarodnoy nauchno-prakticheskoy konferentsii (Russia, Novosibirsk, March 3, 2010)* (eds. S.S. Chernova), Novosibirsk: Sibprint, pp. 187-192.