ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ, СИСТЕМНИЙ АНАЛІЗ ТА КЕРУВАННЯ

UDC 519.766.4

P.I. Bidyuk, M.M. Konovalyuk, N.V. Kuznetsova, I.V. Pudlo

ADAPTIVE SHORT-TERM FORECASTING OF SELECTED FINANCIAL PROCESSES

A computer based system is proposed for adaptive modeling and forecasting of financial and economic processes, that is constructed with application of system analysis principles. A hierarchical structure of decision making process during forecasts estimation was taken into consideration and the methods were used for describing uncertainties of structural, parametric and statistical nature. To estimate model structure and parameters several mutually supporting estimation techniques were used as well as optimal state estimation procedure for dynamic systems that allowed take into consideration some types of structural and statistical uncertainties. Probabilistic modeling methods make it possible to consider uncertainties of probabilistic type. The problem of short term forecasting for gold price is considered as an example using a set of constructed regression models and Kalman filter for generating optimal estimates of states. The best forecasting results were achieved with optimal filter and autoregression models with trends. Also the models were constructed for conditional variance that provided acceptable quality forecasts for variance (volatility) that could be used for constructing decision making rules in trading operations.

Keywords: adaptive forecasting, system approach, nonlinear nonstationary processes, model structure and parameters estimation, complex criterion.

Introduction

To increase quality of risk management and managerial decisions, quality of automatic control for engineering systems and production technology it is necessary to develop and apply to practical control problems solving new forecasting techniques directed towards further improvement of short- and medium term forecasting. High quality forecasts help managers to produce reliable objective operational and strategic plans for future developments. Existing today forecasting methods that are based on various analytical procedures, logical rules and rational expert reasoning cannot provide in many cases desirable quality of forecasting results what requires from researchers new efforts to enhance the quality of forecasts estimates [1—3]. Quality of the forecasts estimates depends highly on quality of data itself, application of correct preliminary data processing techniques directed to improvement of their statistical characteristics, correct application of structure and parameters estimation procedures as well as techniques for generating the forecasts themselves. One of the main points during model construction and forecasts estimation is continuous control of all relevant computational procedures. This task can be solved with development and practical application of adaptive forecasting systems based on ideas of system analysis, such as hierarchic system structure, hiring of optimization procedures where possible, mathematical description and taking into consideration of uncertainties, generation of alternatives for making decisions etc [4].

Besides there exists a possibility for combining forecast estimates generated by ideologically different techniques, what results very often in better forecasts than quality of separate estimates generated by each specific approach. At any rate such approach to forecasting may lead to substantial decreasing of variance for forecasting errors if applied correctly.

In this study we consider the problem of adaptive system construction for solving the problems of model building and generating short-term forecasts using statistical data for selected financial processes. The most difficult financial processes for forecasting are the processes of price evolving on gold, stock shares and other securities.

Problem statement

The purposes of the study are as follows: 1) to develop adaptive model development and forecasting system based on the modern system analysis approaches; 2) to construct linear autoregression models with moving average (ARMA) for the process of coal and crude oil production and to compute short-term forecasts on their basis; 3) to apply Kalman filter (KF) for optimal estimation of the coal and crude oil production state and for short-term forecasting; 4) to perform comparative analysis of forecasts estimates computed with the models constructed and Kalman filter based on the models.

Some modern forecasting techniques

Adaptive Regression Analysis Approach. Correct application of modern modeling and adaptive estimation techniques, probabilistic and statistical data analysis provide a possibility for organizing computing process in such a way so that to get higher quality of forecast estimates in conditions of structural, parametric and statistical uncertainties. Such uncertainties arise due to availability of nonstationary and nonlinear processes under study, incomplete data records, noisy measurements, extreme values and short samples. One of the possibilities for adaptation provide Kalman filtering techniques that generate optimal state estimates together with short term forecasts in conditions of influence of external stochastic disturbances and measurement noise. However, such techniques require estimates of statistical parameters for stochastic disturbances and measurement noise in real time what creates extra burden and errors for forecasting procedures.

We propose a concept of a model adaptation for dynamic processes forecasting based on modern system analysis ideas that supposes hierarchical approach to modeling and forecasting procedures, taking into consideration of possible structural, parametric and statistical uncertainties, adaptation of mathematical models to possible changes in the processes under study and the use of alternative parameter estimation techniques aiming to model and forecast estimates improvement. The functional layout of adaptive forecasting system is given in Fig. 1. Here each step of data processing is controlled by appropriate set of statistical parameters each of which characterizes specific features of data, model as a whole, model parameters and finally quality of the forecast estimates generated.

We propose the new adaptive scheme that is distinguished with several possibilities for adaptation using a complex quality criterion. The data collected should be correctly prepared for model structure and parameter estimation. The model structure estimation is a key element for reaching necessary quality of forecasts. It is proposed to define a model structure as follows:

$$S = \{r, p, m, n, d, z, l\},$$

where r is model dimensionality (number of equations); p is model order (maximum order of differential or difference equation in a model); m is a number of independent variables in the right hand side; n is a nonlinearity and its type; d is a lag or output reaction delay time; z is external disturbance and its type; l are possible restrictions for variables. For automatic search of the "best" model it is proposed to use the following combined criteria:

$$V_N(\theta, D_N) = e^{|1-R^2|} + \ln\left(1 + \frac{SSE}{N}\right) + e^{|2-DW|} + \\ + \ln(1 + MSE) + \ln(MAPE) + e^U, \quad (1)$$

where θ is a vector of model parameters; N is a power of time series used; R^2 is a determination coefficient; DW is Durbin–Watson statistic; MSE is mean squared error; $MAPE$ is mean absolute percentage error; U is Theil coefficient. The power of the criterion was tested experimentally and proved with a wide set of models and statistical data.

There are several possibilities for adaptive model structure estimation (Fig. 1): 1) automatic analysis of partial autocorrelation for determining order of autoregression; 2) automatic search for the exogenous variables lag estimates (detection of leading indicators); 3) automatic analysis of residual properties; 4) analysis of data distribution type and its use for selecting correct model estimation method; 5) adaptive model parameter estimation with hiring extra data; 6) optimal selection of weighting coefficients for exponential smoothing, nearest neighbor and some other techniques; 7) the use of adaptive approach to model type selection. The use of a specific adaptation scheme depends on volume and quality of data, specific problem statement, requirements to forecast estimates, etc. In some cases we used successfully logistic regression together with linear regression to describe the data. These models as well as classification trees and Bayesian networks have been used successfully to forecast direction of stock price movement and some macroeconomic processes.

Application of the concept described provides the following advantages: 1) automatic search for the "best" model reduces the search time; 2) it is possible to analyze much wider set of candidate models than manually; 3) the search is optimized thanks to the use of complex quality criterion; 4) in the frames of computer system developed it is possible to integrate ideologically different methods of modeling and forecasting and compute combined forecasts estimates that are distinguished with better quality. Testing of the system with stock price and macroeconomic data showed that it is possible easy to reach a value of absolute percentage error of about 3–7 % for short term forecasting.

Kalman filtering. Modern Kalman filtering algorithms could be easily hired for solving short term forecasting problems in the frames of adaptation procedure given above. The models construc-

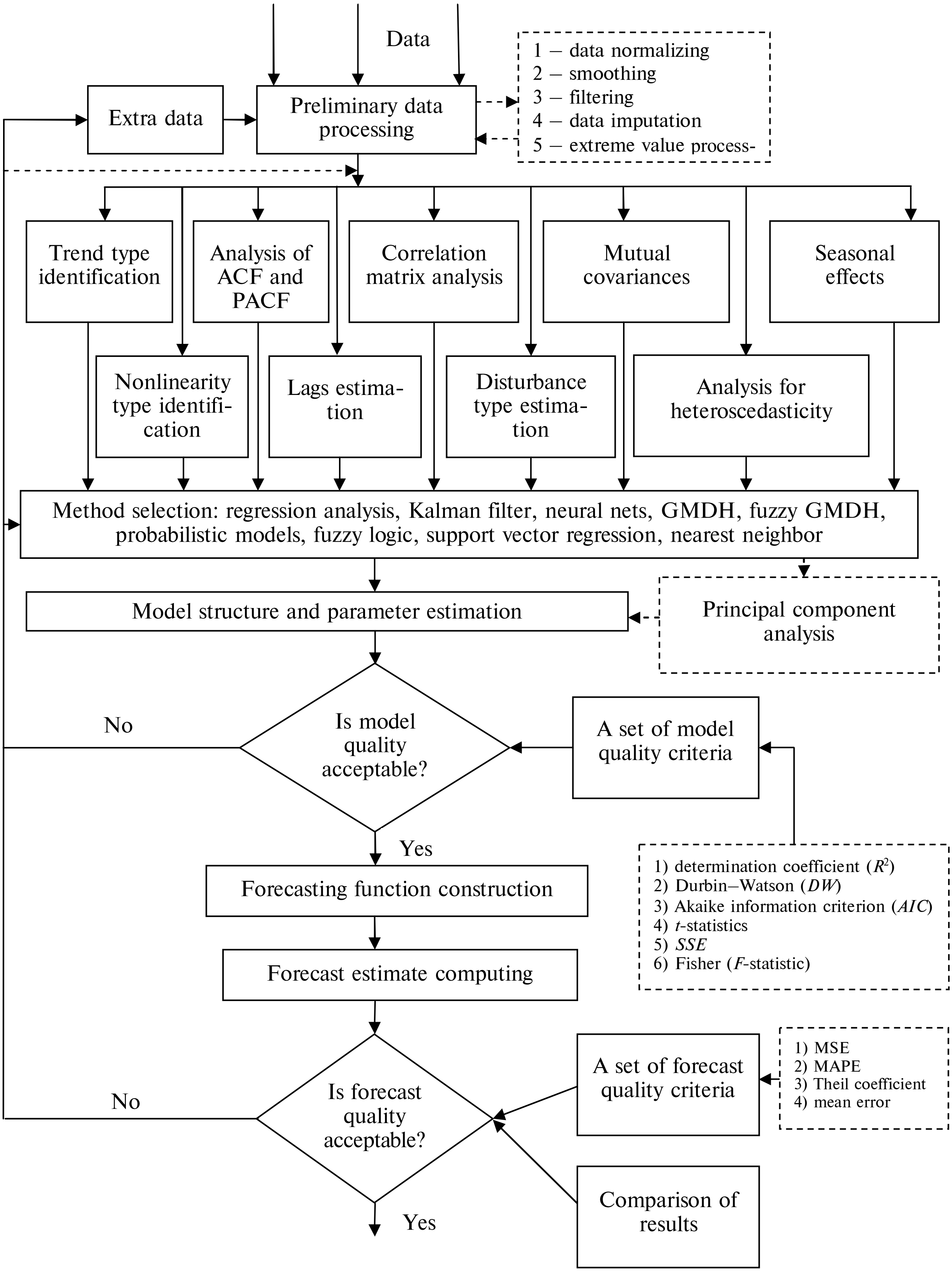

Fig. 1. Adaptive model estimation and process forecasting schema

ted according to the adaptation scheme considered should be transformed into the state space representation form that makes it possible further application of the Kalman type optimal filtering algorithms. An advantage of the approach is in the possibility of model adjusting to random external disturbances (state noise which is always available) and taking into consideration possible measurement errors (measurement noise). In most cases of practical applications such approach provides for high quality of short term forecasts thanks to availability of optimal state estimates computed by the filtering algorithm.

Bayesian networks. Bayesian networks (BN) or Bayesian belief networks are probabilistic models in the form of a directed acyclic graph the vertices of which represent selected variables and arcs reflect existing cause and effect relations between the variables [5]. Today BN find quickly expanding applications in various areas of human activities such as computer based medical and engineering diagnostic systems, process forecasting, classification problems, risk management, and many others. BN provides a possibility for discovering existing dependences between variables and for determining new conditional probabilities for states and situations after receiving new information by any node of a graph. Success of application of the approach depends on correctness of a problem statement, appropriate variables selection, availability of necessary data and/or expert estimates for the structure and parameter learning.

General problem statement touching upon application of Bayesian networks includes the following steps: 1) thorough studying of a process being modeled; 2) collecting of statistical data and expert estimates; 3) selection of known or development of a new method for model structure estimation (learning); 4) BN parameter learning (construction of conditional probability tables); 5) development of a new or selection of known inference method; 6) testing the BN constructed using actual and generated data; 7) application of the model to practical problem solving, i.e. state forecasting, classification, diagnostics etc.

In spite of the fact that general theory of BN has been developed quite well, usually many questions arise when a particular practical problem is solved. This is especially true regarding the problems of forecasting because the quality requirements to forecast estimates are continuously increasing what results in further refinement of relevant computing methods and algorithms.

Group Method for Data Handling. The group method for data handling (GMDH) is a powerful modern instrument for process modeling and forecasting developed at the Ukrainian National Academy of Sciences in the second half of the last century by O.G. Ivakhnenko [6]. It generates the forecasting model in the form of the Kolmogorov–Gabor polynomial that could be used for describing linear and nonlinear systems. The main positive feature of the method is that it selects automatically the best model structure in the class of preselected linear or nonlinear structures. The latest versions of the fuzzy GMDH techniques provide better possibilities for increasing the quality of forecasts estimates.

The problem statement for application of the technique should include the following elements: 1) selection of partial descriptions that create a basis for the possible final model; 2) selection and adaptation of the model parameters membership functions for a particular application; 3) development of a new or application of known model parameter estimation technique; 4) selection of a model quality criteria for the use at intermediate computation steps and for the final model selection. The models constructed with appropriately developed and tuned GMDH approach usually provide medium or high quality of short and medium term forecasts.

Generalized Linear Models. Generalized linear models (GLM) is a class of models that extend the idea of linear modeling and forecasting to the cases when pure linear approach to establishing relations between process variables cannot be applied [7]. The GLM approach also extends the possibilities for modeling in cases when statistical data exhibit distribution different from normal. GLM constructing can be considered from classical statistics or a Bayesian perspective. Usually the problem statement regarding such type of model construction is touching upon the following elements: selection of type of prior distribution for model parameters; a method for parameters estimation using appropriate simulations techniques; necessity for hierarchical modeling, posterior simulation etc. GLM could be successfully applied to solving the problems of classification and prediction of nonlinear process development. For example, they are used widely in scoring systems for predicting returning of loans by clients of a bank.

Combination of forecasts. The problem of forecasts combination arises in the cases when one selected technique is not enough for achieving desirable quality of forecasting. In such cases it is advisable to select two or more ideologically different forecasting techniques and to compute combined estimate using appropriately selected weights. In a simple case equal weights are applied to the individual forecasts. Other approaches to computing these weights are based on previously found prediction errors for each method or computed by optimization procedures. Especially good results of combination are achieved in cases when the error variances for individual forecasting techniques do not differ substantially from each other.

The models constructed

To describe coal and oil production in Ukraine several models had been constructed and tested. The first one was linear AR(p) model of the form:

$$y(k) = a_0 + \sum_{i=1}^{p} a_i y(k-i) + \varepsilon(k),$$

where p is autoregression order; $y(k)$ is a measurement at k-th moment of time; $\varepsilon(k)$ is a normal disturbance; a_i is i-th parameter of the model. The autoregressive moving average model (ARMA(p, q)) constructed is of the form:

$$y(k) = a_0 + \sum_{i=1}^{p} a_i y(k-i) + \sum_{j=0}^{q} b_j \varepsilon(k-j) + \varepsilon(k).$$

where q is moving average order; a_i, b_j are model parameters.

The models developed have been used to construct forecasting functions allowing to generate multistep forecasts. As an example below is given forecasting function (for three steps) constructed for the process ARMA(2, 1):

$$\hat{y}(k+3) = E_k[y(k+3)] =$$

$$= a_0 + a_1 E_k[y(k+2)] + a_2 E_k[y(k+1)] =$$

$$= a_0(1 + a_1 + a_1^2 + a_2) + (a_1^3 + 2a_1 a_2) y(k) +$$

$$+ (a_1^2 a_2 + a_2^2) y(k-1) + \beta_1 (a_1^2 + a_2) \varepsilon(k).$$

Recursive expression for arbitrary s steps forecasting can be written in this case in the form:

$$\hat{y}(k+s) = E_k[y(k+s)] =$$

$$= a_0 + a_1 E_k[y(k+s-1)] + a_2 E_k[y(k+s-2)].$$

Kalman filter application. The linear Kalman filter was applied for computing optimal estimates of states and short-term forecasting based on ARMA-type models. Optimal filtering of states and forecasting process includes the steps given below.

Step 1. Formulation of mathematical model of the process under study:

$$x(k) = F(k)x(k-1) + w(k-1);$$

$$z(k) = H(k)x(k) + v(k),$$

where $x(k) = \begin{pmatrix} y(k) \\ y(k-1) \end{pmatrix}$ is 2-dimensional state vector; $F(k) = \begin{bmatrix} a_1 & a_2 \\ 1 & 0 \end{bmatrix}$ is state transition matrix for AR(2) model; $\{w(k)\} \sim N(0, Q(k))$ is random disturbance that is supposed to be normal; $H(k)$ is measurements matrix; $\{v(k)\} \sim N(0, R(k))$ is a sensor (measurement) noise; initial state and respective covariances are as in the standard problem statement:

$$E\{x_0\} = \hat{x}_0,\ E\{\hat{x}_0, \hat{x}_0^{\mathrm{T}}\} = P_0 = P_0';$$

$$E\{w(k), v^{\mathrm{T}}(k)\} = 0,\ E\{w(k), x_0^{\mathrm{T}}(0)\} = 0,$$

$$E\{v(k), x_0^{\mathrm{T}}\} = 0.$$

Step 2. State extrapolation (one-step ahead projection of a state):

$$\hat{x}(k) = F(k)\hat{x}(k-1).$$

Step 3. Extrapolation for the estimation errors covariance matrix:

$$P'(k) = F(k)P(k-1)F^{\mathrm{T}}(k) + Q(k-1).$$

Step 4. Compute matrix filter gain:

$$K(k) = P'(k)H^{\mathrm{T}}(k)\{H(k)P'(k)H^{\mathrm{T}}(k) + R(k)\}^{-1}.$$

Step 5. Optimal state estimation with taking into consideration the last measurement $z(k)$:

$$\hat{x}(k) = F(k)\hat{x}(k-1) + K(k)[z(k) - H(k)\hat{x}(k)].$$

Step 6. Compute covariance matrix for estimates errors for the next iteration:

$$P(k) = [I - K(k)H(k)]P'(k).$$

Step 7. Go to step 2.

Models construction and their applications

To select the best models constructed the following statistical criteria were used: determination coefficient (R^2); Durbin–Watson statistic (DW); Fisher F-statistic; and Akaike information criterion (AIC). The forecasts quality was estimated with making use of the following criteria: mean squared error (MSE); mean absolute percentage error ($MAPE$); and Theil inequality coefficient (U).

As an example of methodology application a time series was studied, the values of which were gold prices within the period between 2005–2006 (504 values all in all)). The statistical characteristics that show constructed models and forecasts quality (without Kalman filter application) are given in table 1.

Thus, the best model turned out to be AR(1) + + trend of 4th order. It provides a possibility for one step ahead forecasting with mean absolute percentage error of about 3,19 %, and Theil coefficient is: $U = 0{,}024$. The Theil coefficient shows that this model is generally good for short-term forecasting.

Statistical characteristics of the models and respective forecasts constructed with Kalman filter application are given in table 2.

Again the best model turned out to be AR(1) + + trend of 4th order. It provides a possibility for one step ahead forecasting with mean absolute percentage error of about 2,71 %, and Theil coefficient is: $U = 0{,}019$. Thus, in this case the results achieved are better than in previous case without filter application.

Statistical analysis of the time series selected with application of Goldfeld–Quandt test proved that this is heteroskedastic process with time varying conditional variance. As far as the variance is one of the key parameters that is used in the rules for performing trading operations it is necessary to construct appropriate forecasting models. Table 3 contains statistical characteristics of the models constructed as well as quality of variance forecasting. To solve the problem we used generalized autoregressive conditionally heteroskedastic (GARCH) models together with description of the processes trend the form of which is rather sophisticated (high order process). The models of this type turned out to demonstrate quite acceptable one-step ahead forecasting properties.

***Table* 1.** Models and forecasts quality without Kalman filter application

Model type	Model quality			Forecast quality			
	R^2	$\sum e^2(k)$	DW	MSE	MAE	$MAPE$	Theil
AR(1)	0,99	25644,67	2,15	49,82	41,356	8,37	0,046
AR(1,4)	0,99	25588,10	2,18	49,14	40,355	8,12	0,046
AR(1) + 1st order trend	0,99	25391,39	2,13	34,39	25,109	4,55	0,032
AP(1,4) +1st order trend	0,99	25332,93	2,18	34,51	25,623	4,67	0,032
AR(1) + 4th order trend	0,99	25173,74	2,12	25,92	17,686	3,19	0,024

***Table* 2.** Models and forecasts quality with application of Kalman filter

Model type	Model quality			Forecast quality			
	R^2	$\sum e^2(k)$	DW	MSE	MAE	$MAPE$	Theil
AR(1)	0,99	24376,32	2,11	45,21	39,73	7,58	0,037
AR(1,4)	0,99	24141,17	2,09	47,29	38,75	7,06	0,035
AR(1) + 1st order trend	0,99	23964,73	2,08	31,15	22,11	3,27	0,029
AR(1) + 4th order trend	0,99	22396,83	2,04	21,35	13,52	2,71	0,019

***Table* 3.** Results of modeling and forecasting conditional variance

Model type	Model quality			Forecast quality			
	R^2	$\sum e^2(k)$	DW	MSE	MAE	$MAPE$	Theil
GARCH(1,7)	0,99	153639	0,113	972,5	–	517,6	0,113
GARCH (1,15) + 2nd order trend	0,99	102139	0,174	458,7	–	211,3	0,081
GARCH (1,15) + 5th order trend	0,99	80419	0,337	418,3	–	121,6	0,058
EGARCH (1,7) + 9th order trend	0,99	45184	0,429	67,8	–	8,74	0,023

Thus, the best model constructed was exponential GARCH(1,7) + 9^{th} order trend. The achieved value of MAPE = 8,74% comprises very good result for forecasting conditional variance.

Further improvements of the forecasts were achieved with application of the adaption scheme given in Fig. 1. An average improvement of the forecasts was in the range between 0,5–1,5%, what justifies advantages of the approach proposed. Combination of forecasts generated with different forecasting techniques helped to further decrease mean absolute percentage forecasting error for about 0,3–0,6 % in this particular case.

Conclusions

The forecasting methodology based on application of adaptation scheme proposed, including structural and parametric adaption, proved to be useful for forecasting financial processes selected. The methodology has also been applied successfully to forecasting macroeconomic processes. The main features of the approach proposed are as follows: testing data quality with a set of statistical parameters; continuous analysis of data directed towards identification of forecasting model structure and its parameters; generation of candidate models and selection of the best one with another set of model quality parameters; and generation of a set of forecasts estimates on the basis of candidate models. The best forecast estimate is selected with a set of forecasts quality parameters. Thus, the whole computational process is controlled with three sets of quality parameters what guaranties acceptable quality of final result.

The linear Kalman filter was successfully applied for generating optimal estimates of states and short-term forecasts based on the models selected. The state noise covariances were estimated recursively with new data coming what corresponds to the general ideology of adaptation. The best results of gold price forecasting were received with Kalman filter application (one- and two-step predictions) for the models with AR(1) + trend of 4^{th} order. The results achieved in this case are better than in the case without optimal filter application. The best model constructed for forecasting conditional variance was exponential GARCH(1,7) + 9^{th} order trend. Achieved value of MAPE = 8,74% comprises very good result for forecasting conditional variance.

Further improvements of the forecast estimates were achieved with application of the adaption scheme given in Fig. 1 with complex statistical model quality criterion (1). An average improvement of the forecasts was in the range between 0,5–1,5 %, what justifies advantages of the approach proposed. Combination of forecasts generated with different forecasting techniques helped to further decrease mean absolute percentage forecasting error for about 0,3–0,6 % in this particular case.

The future research should be directed towards expanding of the adaptive forecasting scheme with new methods for adaptive model parameters estimation, and alternative forecasting techniques based on intellectual data processing schemes.

1. *R.H. Shumway and D.S. Stoffer,* Time Series Analysis and its Applications. New York: Springer Verlag, 2006, 588 p.
2. *P.I. Bidyuk et al.,* Time Series Analysis. Ukraine, Kyiv: Polytechnika, NTUU KPI, 2013, 607 p. (in Ukrainian).
3. *R. Harris and R. Sollis,* Applied Time Series Modelling and Forecasting. West Sussex: John Wiley & Sons Ltd., 2005, 313 p.
4. *M.Z. Zgurovsky and N.D. Pankratova,* The System Analysis: Problems, Methodology, Applications. Ukraine, Kyiv: Naukova Dumka, 2011, 726 p. (in Ukrainian).
5. *F.V. Jensen and Th. Nielsen,* Bayesian Networks and Decision Graphs. New York: Spinger-Verlag, 2009, 457 p.
6. *M.Z. Zgurovsky and Yu.P. Zaichenko,* An Introduction to Computing Intelligence. Ukraine, Kyiv: Naukova Dumka, 2013, 406 p. (in Ukrainian).
7. *A. Dobson,* An Introduction to Generalized Linear Models. New York: CRC Press Company, 2013, 407 p.

Рекомендована Радою
Навчально-наукового комплексу
Інститут прикладного системного
аналізу НТУУ "КПІ"

Надійшла до редакції
25 листопада 2013 року