УДК 004.89:519.22](043.3)

O. Trofymchuk, P. Bidiuk, O. Terentiev, V. Klymenko

THE METHODOLOGY FOR ADAPTIVE MODELING AND FORECASTING NONLINEAR AND NONSTATIONARY PROCESSES

Oleksandr Trofymchuk

Institute of Telecommunications and Global Information Space of the National Academy of Sciences of Ukraine, Kyiv,
orcid: 0000-0003-3358-6274

Trofymchuk@nas.gov.ua

Petro Bidiuk

National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute»,
orcid: 0000-0002-7421-3565

pbidyuke_00@ukr.net

Oleksandr Terentiev

Institute of Telecommunications and Global Information Space of the National Academy of Sciences of Ukraine, Kyiv,
orcid: 0000-0002-4288-1753

o.terentiev@gmail.com

Victoriia Klymenko

Institute of Telecommunications and Global Information Space of the National Academy of Sciences of Ukraine, Kyiv,
orcid: 0000-0002-8911-5773

klimenko_vika@ukr.net

The study is directed towards development of systemic methodology for modeling and forecasting nonlinear and nonstationary processes in economy, finances and other areas of human activities. There exist multiple problems that are to be solved with the data of such type practically in all areas of human activities: such as constructing adequate models including estimation and forecasting state of a system under investigation; technical, economic and medical diagnostics; automatic control in technologies; moving objects control; financial and other risk estimation and management; risk factor interaction; control and monitoring of microclimate in greenhouses and industrial enterprises; decision making support in business; dynamic strategic planning in production; providing stability for critical economic infrastructure etc. The structure and parameter adaptation procedures for the regression and probabilistic models are proposed as well as respective information system architecture and functional layout are developed. The system development is based on the system analysis principles such as hierarchical architecture of the system, adaptive approach to model structure estimation, optimization of model parameter estimation procedures, functional completeness of the system providing for autonomous operation of the system, identification and taking into consideration of possible uncertainties available in the process of data processing and mathematical model development, application of appropriate sets of quality criteria that are guarantying high quality of intermediate and final results of data analysis. The possible uncertainties are inherent to data collecting, model constructing and forecasting procedures, and play the role of negative influence factors to the computational procedures of proposed information system. Reduction of their influence is favorable for enhancing the quality of intermediate and final results of computations. The illustrative examples of practical application of the methodology developed proving the system functionality are provided.

Keywords: nonlinear nonstationary processes, modeling, forecasting, Kalman filter, generalized linear models.

Introduction

Current research problems are with complicated processes in many areas of human activities such as ecology, climatology, technology, economy, finances, technical applications, space and ocean studies, construction etc. The most distinctive common features of the processes are availability of nonlinearity and nonstationary behavior that require development of new adequate models for solving the tasks of forecasting and making appropriate decisions [1, 2]. The nonstationary behavior is related to changing in time statistical parameters such as mathematical expectation and/or variance of a process under study [2–5]. Dynamics of mathematical expectation results in producing of a process trend that can be deterministic or stochastic. The first one is described with selected deterministic function, for example, polynomial, and the second type of trend is stochastic that can be described with combination of random processes sampled from appropriate distributions. The processes with trend are referred to as integrated ones due to the fact that linear trend reminds reaction of an integrator, electronic device used in analog computers. The processes with changing in time variance are considered as heteroscedastic [6]. Dynamics of variance can be described by selected mathematical models, for example, autoregression (ARCH — Autoregression with Conditional Heteroscedasticity) and autoregression with moving average [4, 6–8]. Here condition is imposed on availability of necessary information (measurements) for constructing the model of variance dynamics. Wide set of models is available in numerous special publications exhibiting formal description of variance dynamics, and some of them will be considered in the study.

As far as trend and variance are parameters of nonstationary processes both integrated and heteroscedastic processes require paying special attention to constructing adequate mathematical models providing the possibility for high quality forecasting. Besides, nonstationary behavior of the processes can exhibit nonlinearity with respect to variables or model (process) parameters. As an example, can be mentioned logistic regression that is widely used in analysis of financial risks. It can be stressed that nonstationary and nonlinear processes (NNP) create a wide class of processes available in many areas of human activities at least mentioned above. To reach high model adequacy and quality of forecasts based upon models of the processes mentioned we need to improve estimation of model structure and apply correctly the methods of the model parameter estimation. The methods can be available for immediate use like widely known in practice ordinary least squares (OLS), maximum likelihood (ML) or the methods can be modified for specific application. Among relatively new methods is Markov chain Monte Carlo (MCMC) technique that can be used in special simulation procedures and adjusted to specific distribution of measurement data [9].

There are multiple problems to be solved with the models of nonstationary and nonlinear processes including estimation and forecasting state of a system under investigation; technical, economic and medical diagnostics; automatic control in technologies, moving objects; financial and other risk estimation and management; risk factor interaction; control and monitoring of microclimate in greenhouses and industrial enterprises; vibration monitoring in construction; decision making support in business; dynamic strategic planning in production; providing stability for critical economic infrastructure etc. Practically all these problems require development and application of appropriate adaptation and optimization procedures to keep the models at necessary level of adequacy to the process under study. The known adaptation procedures include re-estimation of a model structure and its parameters using appropriate sets of criteria and computational procedures [10].

Today specialized methodologies implemented in the frames of decision making support systems (DMSS) are very popular as convenient instrumentation for problem solving in many areas where analysis of NNPs is performed [3, 11–14]. They belong to the class of information processing systems, and are constructed with application of system analysis principles providing the possibilities for reaching high quality of intermediate and final results of data and expert estimates analysis [15, 16]. The methodology is directed towards enhancement (thanks to application of appropriate preliminary processing) of statistical/experimental data quality, constructing model of higher adequacy, improvement of forecasts, and generating of alternative decisions. Design of system architecture in the form of flexible modular structure helps with fast modification and expansion of its functions so that to adjust the system to changing conditions of specific application. Analysis of NNP is usually simplified with the use of specialized intellectual methodology thanks to its capabilities to generate for user useful information including retrospective analysis of former results of analysis similar processes [3, 5].

The study will consider some models of NNPs relevant to analysis of financial processes, the methods of their parameters estimation and possibilities for constructing and application of specialized systemic methodology to get higher results of modeling, including structural and parametric adaptation, possible risk estimation and forecasting. To some extent the problem can be solved with application of appropriate simulation of selected financial processes directed towards generating the data exhibiting required statistical characteristics [3, 5].

Problem statement

The purpose of the study is in solving the following problems: to develop structure and parameter adaptation procedures for the regression and probabilistic models; to develop the information system architecture for modeling and forecasting nonlinear and nonstationary processes in economy, finances, ecology and other areas based on the system analysis principles; to consider possibilities for elimination of some uncertainties inherent in data collecting, model constructing and forecasting procedures; to develop the methodology for modeling and forecasting linear and nonlinear processes in the frames of the same system; providing illustrative examples of practical application of the methodology developed proving the system functionality.

Methodologies.

Some models of nonstationary and nonlinear processes

The models of quasi linear processes (nonlinear in variables) are widely used in practice, as an example of the model which can be a polynomial describing deterministic trend. Consider in short each stage of regression model constructing using statistical/experimental data. Some elements of the sequence of actions can be used for constructing nonlinear models. For example, the following procedures can be used for data analysis:

— to estimate availability of dependency between variables, for this purpose linear and nonlinear correlation functions can be hired;

— possible autocorrelation lags for dependent variable can be estimated with autocorrelation function;

— appropriate correlation analysis of linear model errors provides the possibility for establishing nonlinearity and heteroscedasticity in data;

— some linear models are easily transformed into nonlinear (quasi linear) polynomial regression model; and linear regression can be transformed into nonlinear expanded autoregression with moving average (NARMA) as follows [3]:

$$y(k) = a_0 + \sum_{i=1}^{p} a_i \, y(k-i) + \sum_{j=1}^{q} b_j \, v(k-i) + c\, x^m(k) + \varepsilon(k); \quad (1)$$

or in more complex bilinear structure (2):

$$y(k) = a_0 + \sum_{i=1}^{p} a_i \, y(k-i)\, y(k-j) + \sum_{j=1}^{q} b_j \, v(k-i) + c\, x^m(k) + \varepsilon(k), \; j = 1, 2, \ldots, \quad (2)$$

where, m is integer, that determines degree of nonlinearity for independent variable; in the same way is constructed a model of nonlinear trend using polynomial or some other nonlinear function;

— parameter estimation for quasi linear models (nonlinear with respect to variables) can be performed with OLS or other methods of linear estimation [3, 17];

— to construct quasi linear models the same adaptive methodologies can be applied to estimate structure and parameters that are used for estimation of linear models, and with the same sets of statistical quality criteria and forecast estimates;

— some statistical criteria that are used for adequacy analysis of linear models can be used for analysis of nonlinear models, say the following: the coefficient of determination, sum of squared errors, Akaike information criterion, Fisher F-statistic, Bayesian information criterion and others;

— in general, to construct nonlinear models it is possible to apply the same principles of system analysis: the hierarchical principle for development specialized decision support systems on the basis of statistical data, the principles of adaptation, optimization, functional completeness, taking into consideration uncertainties etc.;

— if the process under study can be decomposed into separate simpler processes this is a good possibility for further modeling; developing their mathematical description will be simplified: some processes can be linear, and some — nonlinear;

— according to the levels of hierarchy the system of data collecting, analysis, modeling and management is developing and functioning;

— analysis of special literature can help with the search and further use of available models as well as their theoretical and experimental studies; this approach can decrease substantially the time necessary for modeling and other expenses.

As a result of preliminary analysis of the process under study all the information collected is used for estimation of model structure (or several candidates) with further parameter estimation using statistical data. The information coming from various sources usually is supporting each other and should be used correctly.

Some common features of the processes in economy, finances and ecology

A wide diversity of various processes exists in economy, finance, ecology, demography, technology and other areas of human activity. However, there are some general common features of the processes like linearity/nonlinearity, and stationarity/nonstationarity that allow divide them into practically understandable classes and select/develop appropriate modeling and forecasting techniques [3, 5, 18, 19]. Fig. 1 (a simplified classification of dynamic processes in economy and finances) shows simplified classification of the processes from which it is possible to make a conclusion about a wide variety of mathematical model structures that could be applied for formal description of the processes dynamics and solving the problem of forecasting their evolution in time.

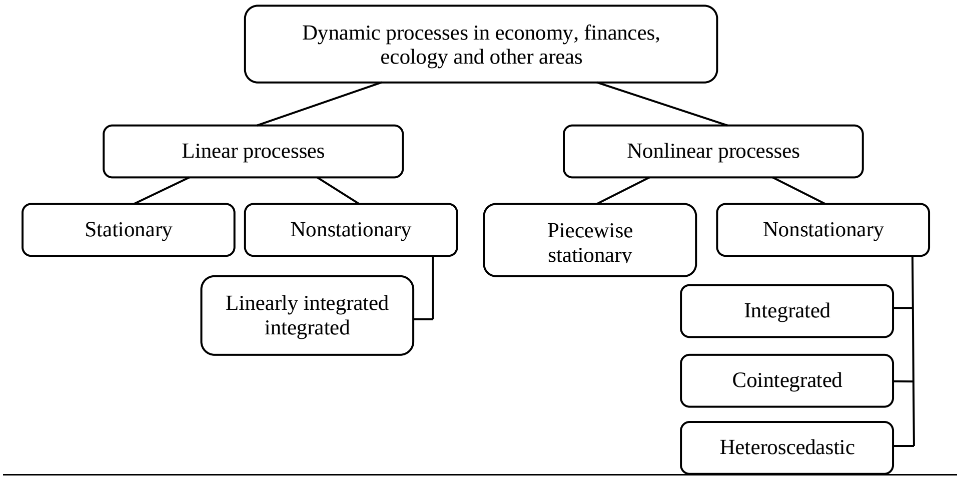

Fig. 1

Linear processes can be stationary without trend and nonstationary when they contain linear (first order) trend, I(1), where I(1) means integrated of the first order. If variance (covariance) of stochastic linear process is time dependent then it is classified as heteroscedastic and requires application of nonlinear models for describing the process variance and possibly the process itself [3, 20].

There also exists a wide variety of nonlinear processes though we selected only some of them that are more frequent in economy and finances. Generally the processes can be split into nonlinear regarding parameters and nonlinear regarding variables. The first type is more sophisticated with respect to modeling and parameter estimation and usually requires more efforts and time for their model development. Widely used logistic regression could be mentioned as an example of such a model.

Some nonlinear processes can exhibit linear behavior in their stable (nominal) mode of operation. This feature allows for linear description of the process in the vicinity of operating point. Generally NNPs are very often met in the areas of study mentioned above. The set of the processes includes integrated processes (IP) that contain a trend of order two or higher as well as cointegrated processes containing trends of the same order, and the processes with time dependent variance, i.e. heteroscedastic processes. Most of financial processes illustrating price evolution of stock instruments belong to this class. In engineering applications such processes are studied in diagnostic systems where appropriate decision is made regarding current system state. The models of heteroscedastic processes are necessary for forecasting conditional variance that is used in many applications: financial risk analysis and management; technical, economic and medical diagnostics; automatic control etc [6, 21].

As a result of availability in the processes under study of linear and nonlinear components most of nonlinear models are extended with linear autoregression terms or other linear elements. However, analysis of such mixed models meets substantial practical difficulties. Due to the difficulties values of lags in the models are usually selected not large. Generally, the problem of lags estimation for nonlinear models is rather complicated and should be considered separately.

Methodology of modeling nonlinear nonstationary processes

The methodology proposed for modeling NNPs illustrates Fig. 2. Estimation of a model structure using statistical and probabilistic (mutual) information analysis that provides a possibility for estimation of the following elements of a model structure: dimension of a model — number of equations creating the model; model order (the highest order of difference or differential equation of the model); nonlinearity and its type;

estimate of input delay time, and type of probabilistic distribution for the model variables. It is always appropriate to perform structure estimation for several candidate-models so that to have a possibility for selecting the best one of the candidates estimated.

Preliminary data processing is necessary for suppressing some data uncertainties that are linked to missing measurements; influence of stochastic disturbances and measurement errors; availability of extreme values; analysis of distribution types; structuring of available sets of data. To fulfill this task a wide set of instruments is available: digital and optimal filters [10, 22, 23], methods of missing measurements imputation, probabilistic procedures for data processing etc. Trend type identification (deterministic or stochastic) is performed as a result of appropriate analysis of available data. The correlation analysis is performed with making use of respective computational procedures available in specialized libraries performing necessary data analysis. Generally, the computational procedures given in Fig. 2 (functional layout of the modeling and forecasting system proposed) usually are collected in the frames of specialized information technology that could be easily expanded with new functions to satisfy changing requirements.

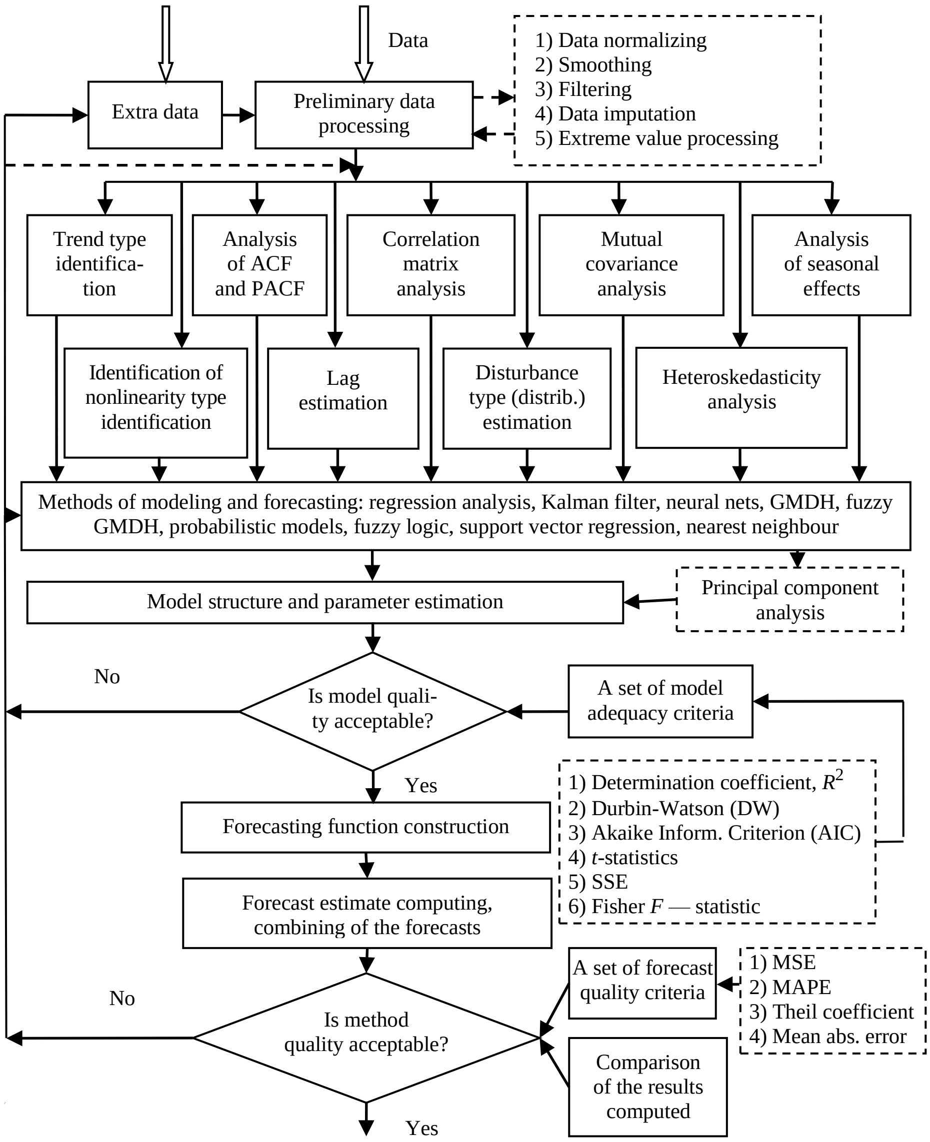

Fig. 2

Some types of possible linear and nonlinear models for nonstationary and nonlinear processes generalized on the basis of [3] are given in Table 1 (some nonlinear models for describing process dynamics).

Table 1

N	Model description	Formal model structure
1	AR + polynomial of time	$y(k) = a_0 + \sum_{i=1}^{p} a_i\, y(k-i) + b_1 k + ... + b_m k^m + \varepsilon(k),\ \ k = 0, 1, 2, ...$ is discrete time; $t = kT_s$; T_s is sampling time.
2	Generalized bilinear model	$y(k) = a_0 + \sum_{i=1}^{p} a_i\, y(k-i) + \sum_{j=1}^{q} b_j\, v(k-i) +$ $= \sum_{i=1}^{m} \sum_{j=1}^{s} c_{ij}\, y(k-i) v(k-j) + \varepsilon(k),$
3	Logistic regression	$\varphi(x(k, z)) = \dfrac{1}{1 + \exp(-x(k, z))},$ $x(k) = \alpha_0 + \alpha_1 z_1(k) + ... + \alpha_m z_m(k) + \varepsilon(k),$
4	Nonlinear extended econometric autoregression	$y_1(k) = a_0 + a_1 y_1(k-1) + b_{12} \exp(y_2(k)) + a_2 x_1 x_2 + \varepsilon_1(k),$ $y_2(k) = c_0 + c_1 y_2(k-1) + b_{21} \exp(y_1(k)) + c_2 x_1 x_2 + \varepsilon_2(k),$
5	Generalized autoregression with conditional heteroscedasticity (GARCH)	$h(k) = \alpha_0 + \sum\limits_{i=1}^{q} \alpha_i \varepsilon^2(k-i) + \sum\limits_{i=1}^{p} \beta_i h(k-i)\cdot$
6	Exponential generalized autoregression with conditional heteroscedasticity (EGARCH)	$\log[h(k)] = \alpha_0 + \sum\limits_{i=1}^{p} \alpha_i \dfrac{\lvert\varepsilon(k-i)\rvert}{\sqrt{h(k-i)}} + \sum\limits_{i=1}^{p} \beta_i \dfrac{\varepsilon(k-i)}{\sqrt{h(k-i)}} +$ $+ \sum\limits_{i=1}^{q} \gamma_i \log[h(k-i)] + v(k)$
7	Nonparametric model with functional coefficients	$y(k) = \sum\limits_{i=1}^{p} \{\alpha_i + (\beta_i + \gamma_i\, y(k-d)) \cdot \exp(-\theta_i y^m(k-d))\} + \varepsilon(k)$
8	Radial basis function	$f_\theta(x(k)) = \sum\limits_{i=1}^{M} \lambda_i \exp\left(-\dfrac{(x(k) - \mu_i)^2}{2\sigma_i^2}\right) + \varepsilon(k),$ $\theta = [\mu_i, \sigma_i, \lambda_i]^{\mathrm{T}};\ M = 2, 3, ...$
9	State-space representation	$\mathbf{x}(k) = \mathbf{F}[\mathbf{a}(k), \mathbf{x}(k-1)] + \mathbf{B}[\mathbf{b}(k), \mathbf{u}(k-d)] + \mathbf{w}(k)$
10	Convolutional neural networks	Selected (constructed) network structure
11	Fuzzy sets and neuro-fuzzy models	Combination of fuzzy variables and neural network model
12	Dynamic Bayesian networks	Probabilistic Bayesian network structure constructed with data and/or expert estimates
13	Multivariate distributions	Example: copula application for describing multivariate distribution
14	Immune systems	Immune algorithms and combined models

The models 1–8 presented in Table 1 have known structure that can be modified in the process of adaptation to available data. Model 1 was successfully applied for modeling deterministic trend of various orders together with short-term deviations from conditional mean. Models 2, 4 can describe bilinear and exponential nonlinearities or nonlinearity with saturation (logistic model 3). Models 5, 6 are used for description of conditional variance dynamics while modeling heteroscedastic process. The last one turned out to be the best model for short term forecasting of conditional variance in most applications performed by the authors. Models 7–9 can describe arbitrary nonlinearities with respect to variables (usually of order 3–5). Fuzzy sets based approach to modeling supposes generating of a set of rules that could describe with acceptable quality functioning of selected processes and formulate appropriate logical inference. Neural networks and fuzzy neural networks are suitable for modeling sophisticated nonlinear functions in conditions of availability of some unobservable variables [24, 25]. Model of this type are related to the intellectual data analysis. Dynamic Bayesian networks and multivariate distributions are statistical/probabilistic models that could describe complex multivariate processes (systems) with generating final result of their application in the form of conditional probabilities (probabilistic inference). Bayesian data analysis is also related to artificial intelligence.

Various types of neural networks can be adjusted for modeling and forecasting time series data. For example, convolutional neural networks (CNN) and recursive neural networks (RNN) can be successfully applied to the problem of time series data forecasting [24, 25]. Here it is reasonable to replace 2D convolutions with 1D. In this case time axis can be treated as a spatial dimension, like the height or width of a 2D image. Such 1D network can be quite competitive with RNN on certain problems of sequence-processing, usually at a considerably cheaper computational cost. Small 1D CNNs can offer a fast alternative to RNNs for tasks such as time series forecasting. The convolution layers introduced previously were 2D convolutions, extracting 2D patches from image tensors and applying an identical transformation to every patch. In the same way, we can use 1D convolutions, extracting local 1D patches (subsequences) from sequences (see Fig. 3 illustration of 1D convolution: each output time-step is obtained from a temporal patch in the input sequence).

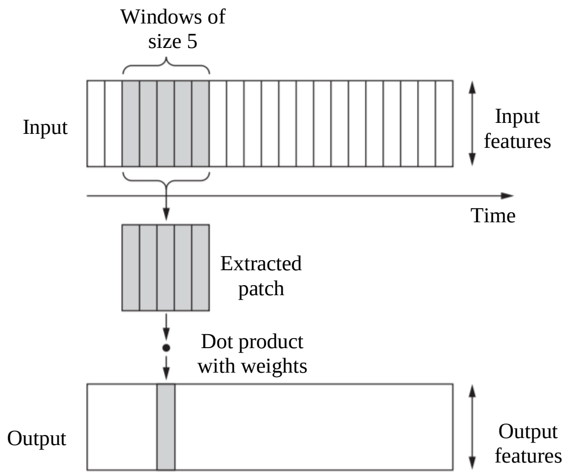

Fig. 3

Such 1D convolution layers can recognize local patterns in a sequence. Because the same input transformation is performed on every patch, a pattern learned at a certain position can later be recognized at a different position, making 1D CNNs translation invariant (for temporal translations). For instance, a 1D CNN processing time series using convolution windows of size 5 should be able to learn pattern fragments of length 5 or less, and it should be able to recognize them in any other context in an input series.

Formally, to detect nonlinearity in collected statistical/experimental data are available statistical tests and techniques that should be applied. Fig. 4 (some techniques for testing data for nonlinearity) shows some known techniques for testing the data for nonlinearity.

Along with application of known technics we proposed a simplified empirical criterion for detecting nonlinearity in data *R*. which is maximum deviation of the process under study from its linear approximation; σ is sample standard deviation of the process. The criterion does not require sophisticated computations but provides for additional information regarding availability of nonlinearity. Analysis of spectral function, the correlation procedures, and Fisher test are applied most often in practice.

The sequence of actions necessary for constructing nonlinear model illustrates Fig. 4 (the search for formal description of nonlinear process). The linear model is constructing using known methodology, and its residuals (containing nonlinear components) are used to construct nonlinear part of the model. Then complete (combined) model is built consisting of linear and nonlinear parts, and possibly other elements such as filters, estimators of non-measurable components.

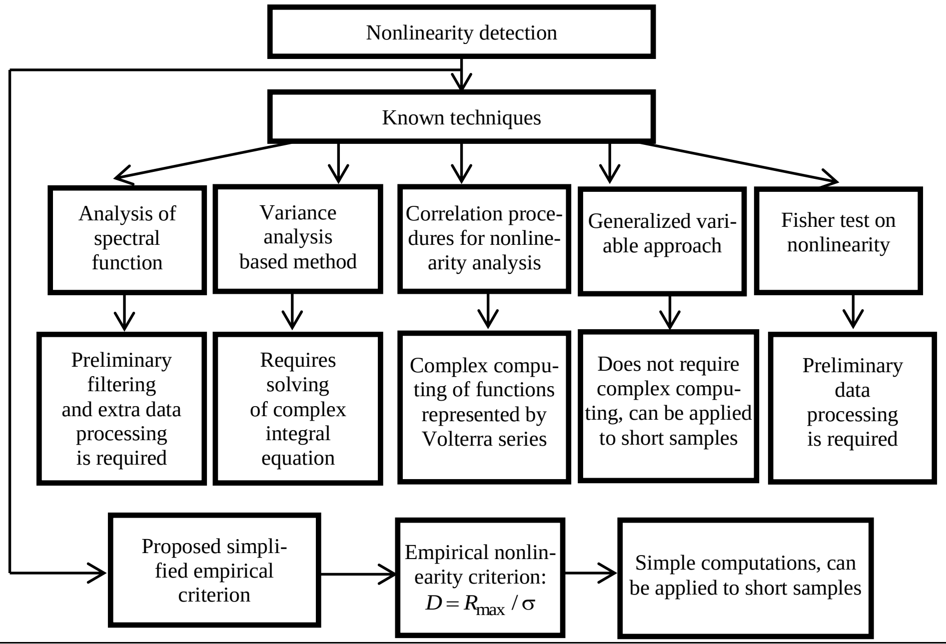

Fig. 4

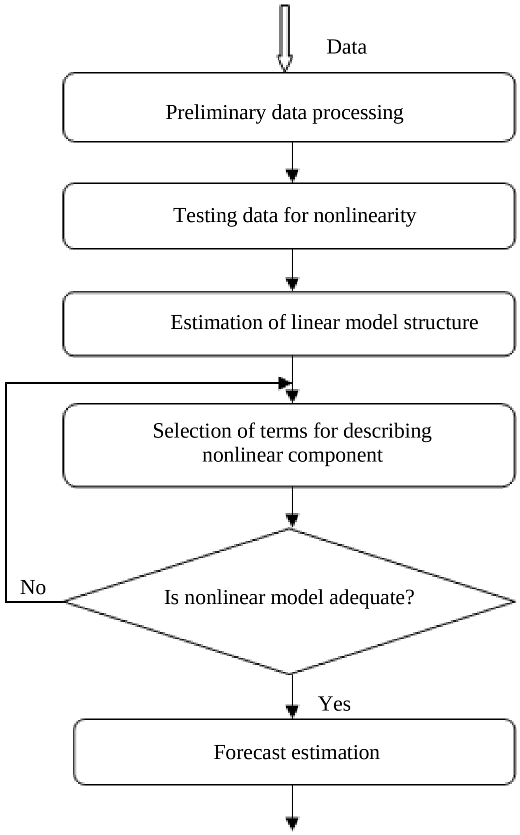

Fig. 5

Consider the possibility for formal description of the nonlinearities with respect to model variables. The nonlinearities could be identified as follows: linear part of the model can be estimated first using linear structures like autoregressive equations with moving average (ARMA(*p*, *q*)) possibly with linear trend or multiple regression [3, 20]. Then nonlinear part is added to the linear elements using the following possibilities: description of nonlinear deterministic trend, bilinear or higher order terms. Sometimes nonlinear terms describing cyclic changes of the main variable are added etc. The practice of model constructing indicates that required model adequacy can be reached using the combination of linear and nonlinear regression, linear regression and Bayesian network, linear regression and nonlinear functions in the form of nonparametric kernels or other nonlinear elements. According to this approach to model constructing several candidate models could be created with subsequent choice of the best one on the basis of appropriate set of statistical adequacy criteria as shown in Fig. 2. It is clear that formal possibilities for determining the type of nonlinearity in a unique way not always exist, for example, when the data samples are short. It is necessary to analyze constructed model adequacy with different nonlinear functions some of which may satisfy stated requirements. The bilinear functions are most often used, radial basis function (RBF), recursive neural networks (RNN) [24, 25], some types of generalized linear models (GLM), [6] selected polynomials.

Model parameter estimation

The next step is model parameter estimation by making use of alternative techniques; in linear case these methods are the following ones: ordinary least squares and its clones, maximum likelihood and some others. In a case of nonlinear model estimation the following methods can be useful: ML, Markov chain Monte Carlo (MCMC) procedures [8], nonlinear least squares (NLS) and other suitable methods that can generate unbiased parameter estimates under specific probabilistic distributions of model variables and model structures. Correct application of alternative parameter estimation techniques provides a possibility for further comparison of the candidate models and selection of the best one. It is also possible to trace the reasons for existing parametric uncertainties in the following form: parameter estimates computed with statistical data cannot be consistent, they may contain bias, and/or can be inefficient. All these effects finally result in poor adequacy of the model constructed.

As an example consider the problem of parameter estimation using Bayesian approach that requires selecting suitable prior density, $p(\theta)$, for unknown vector θ, and then estimate joint posterior density $p(x_{0:} : x_N, \theta | y_0 : y_N)$ in off-line case, or the sequence of posterior densities $\{p(x_{0:} : x_n, \theta | y_0 : y_n)\}$ in the on-line case.

Parameter estimation using the method of particle Markov chain Monte Carlo (PMCMC) [26]. The PMCMC techniques rely upon sequential Monte Carlo (SMC) methods so that to construct efficient proposal distributions. Consider particle marginal metropolis-hastings (PMMH) sampler that approximates the ideal MMH sampler using the following proposal density:

$$q((x'_{0:N}, \theta') | (x_{0:N}, \theta)) = q(\theta' | \theta)\ p_{\theta'}(x'_{0:N} | y_{0:N}), \tag{3}$$

where, $q(\theta' | \theta)$, is proposal density that can be used to generate a candidate for parameter estimate, θ'. Here the acceptance probability for sampler is defined as follows [27]:

$$1 \wedge \frac{p(x'_{0:N}, \theta' | y_{0:N})\, q((x_{0:N}, \theta) | (x'_{0::N}, \theta'))}{p\ (x_{0:N}, \theta | y_{0:N})\, q((x'_{0:N}, \theta') | (x_{0:N}, \theta))} =$$

$$=1\wedge\frac{p_{\theta'}(y_{0:N})\,p(\theta')\,q(\theta\,|\,\theta')}{p_{\theta}(y_{0:N})\,p(\theta)\,q(\theta'\,|\,\theta)}\,. \tag{4}$$

It should be stressed that this algorithm cannot be directly implemented because it is impossible to generate exactly from the distribution, $p_{\theta'}(x'_{0:N}\,|\,y_{0::N})$, and it is impossible to compute the terms, $p_{\theta}(y_{0:N})$, $p_{\theta'}(y_{0:N})$, related to the acceptance probability. The PMMH sampler provides the possibility for approximating the ideal MMH sampler in the way as shown below [27].

The PMMH sampling algorithm [28]: set initial condition at, $k=0$, $\theta(0)$ with random values. Compute, $p_{\theta(0)}(x_{0:N}\,|\,y_{0:N})$, generate $X_{0:N}(0)\sim\hat{p}_{\theta(0)}(dx_{0::N}\,|\,y_{0:N})$, and compute the value of marginal likelihood, $\hat{p}_{\theta(0)}(y_{0:N})$.

At iterations, $k\geq 1$, perform the following:

Generate the proposal parameter estimate: $\theta'\sim q(\theta\,|\,\theta(k-1))$.

Compute, $p_{\theta'}(x_{0:N}\,|\,y_{0:N})$, using sequential Monte Carlo algorithm, sample the distribution, $X'_{0:N}\sim\hat{p}_{\theta'}(dx_{0::N}\,|\,y_{0:N})$, and compute the estimate of marginal likelihood, $\hat{p}_{\theta'}(y_{0:N})$.

Set, $\theta(k)=\theta'$, $X_{0:N}(k)=X'_{0:N}$, $\hat{p}_{\theta(k)}(y_{0:N})=\hat{p}_{\theta'}(y_{0:N})$, with probability

$$1\wedge\frac{\hat{p}_{\theta'}(y_{0:N})\,p(\theta')\,q(\theta(k-1)\,|\,\theta')}{\hat{p}_{\theta(k-1)}(y_{0:N})\,p(\theta(k-1))\,q(\theta'\,|\,\theta(k-1))}\,. \tag{5}$$

Otherwise set, $\theta(k)=\theta(k-1)$, $X_{0:N}(k)=X_{0:N}(k-1)$, $\hat{p}_{\theta(k)}(y_{0:N})=\hat{p}_{\theta(k-1)}(y_{0:N})$.

The advantage of the algorithm presented above is that the invariant distribution of the Markov chain, $\{X_{0:N}(k),\theta(k)\}$, is $\hat{p}(x_{0:N}),\theta\,|\,y_{0:N})$, independent on the number of particles, M, used in the SMC approximation. In other words the SMC approximation used does not introduce any bias. On the other side, it is clear that the higher is M, the better will be mixing properties of the algorithm.

Analysis of model adequacy

At the next stage a set of statistical parameters characterizing model quality (adequacy) is computed and selecting the most suitable model out of the set of candidate models is performed. There is no need to leave only one model for computing forecasts (or solving control problem). Again, it can be a set of the best models constructed using different methods. The final choice is usually made after model application for solving the problem according to the initial problem statement [3].

After computing forecasts for the process (under study) using candidate models another set of forecast quality criteria is applied to select the best result of forecasting, say mean absolute percentage error (MAPE), Theil coefficient, mean absolute error (MAE), minimum and maximum errors of forecasting etc. The models constructed should also be tested on similar process, i.e. model calibration process performed [3, 5].

Dealing with data uncertainties

An important point to be considered in the procedures of model constructing, forecast estimation and decision making is the problem of dealing with data uncertainties. We consider the uncertainties as the factors of negative influence on the whole procedure of data processing and model constructing that may result in non-satisfactory intermediate and final results of computational experiments. The factors inevitably appear when the statistical data is used for model constructing.

The sources for appearing possible data uncertainties in the process of model constructing and forecast estimation are as follows [14, 15, 19]:

— some data is not available or lost due to various reasons what requires application of appropriate data imputation procedures;

— the data is generated with the system model under study which is influenced by the random external disturbances that distort actual values of system state;

— the observations are always measured (collected) with some errors the influence of which should be minimized before further use of the measurements;

— the parameters defining model structure are also assigned (estimated) random values what results in approximate model structure;

— very often there exists a difficulty of selecting a method for model parameter estimation, especially in cases of dealing with short samples, or samples with outliers, or when data probability distribution is poorly defined; as a result the parameter estimates can be biased or non-effective;

— the multistep forecasting requires the use of intermediate forecast estimates what may lead to substantial deterioration of the final forecast estimates.

Thus, the model constructing procedures that can be successfully implemented in the frames of appropriately designed decision support systems should contain the means for uncertainty identification and minimization of their negative influence. In Table 2 (possible types of data uncertainties in modeling and forecasting) types of possible data uncertainties and reflects some possible means for dealing with them are summarized [3].

Table 2

N	Uncertainty type	Reason for uncertainty	Methods of minimizing uncertainty influence
1	Uncertainty of a model structure	impossibility for establishing all possible causal relations between variables; approximate values for model structure parameters	expert approach; application of statistical tests; application of hypothesis testing theory
2	Statistical uncertainty	measurement errors; stochastic disturbances; outliers; missing data values	digital and optimal filters; refining the type of distribution; extremum value theory; imputation of missing values
3	Parametric uncertainty	incorrect choice of parameter estimation method; short samples	application of alternative parameter estimation techniques; expansion of data samples
4	Probabilistic uncertainty	complex mechanisms of causal relations between variables	Bayesian networks; Markov models; probabilistic filters; conditional distributions
5	Amplitude uncertainty	nonmeasurable variables; high measurement errors	Bayesian data processsing; fuzzy logic

Results & Discussion.
The application of usage of adaptive Kalman filter

The proposed methodology was tested to simulate the price of gold within the period between the 1 January–31 December 2023 (sample contains 365 daily values). The table 3 (statistical characteristics for models for case when training and testing data was not preprocessed using an adaptive Kalman filter) shows the obtained statistical characteristics of the constructed mathematical models, as well as their predictive characteristics. For this set of models, the training and testing data was not preprocessed using an adaptive Kalman filter [22, 23].

Table 3

Model type	**Model adequacy**			**Forecast quality**			
	R^2	$\sum e^2(k)$	DW	MSE	MAE	MAPE	Theil
AR(1)	0,99	25510,43	2,21	71,05	44,56	11,29	0,058
AR(1,4)	0,99	25477,20	2,18	68,97	42,94	9,95	0,056
AR(1) + 1st order trend	0,99	25420,17	2,14	48,56	35,39	6,74	0,047
AP(1,4) +1st order trend	0,99	25286,54	2,11	45,31	30,62	5,82	0,043
AR(1) + 4th order trend	0,99	25067,54	2,09	41,12	25,59	4,57	0,038

As can be seen from the comparison of the constructed mathematical models, the best model is AR(1) plus trend of 4th order. The resulting one-step forecasting model has an error mean absolute percentage error of about 4,57 %, and Theil coefficient is $U = 0{,}038$. The obtained Theil coefficient value indicates that the resulting model is best used specifically for short-term forecasting tasks

The results of constructing mathematical models using data preprocessing using an adaptive Kalman filter are given in the table 4 (statistical characteristics for models for case when training and testing data was preprocessed using an adaptive Kalman filter).

Table 4

Model type	**Model quality**			**Forecast quality**			
	R^2	$\sum e^2(k)$	DW	MSE	MAE	MAPE	Theil
AR(1)	0,99	24691,92	2,12	68,54	41,94	10,08	0,053
AR(1,4)	0,99	24579,71	2,11	62,15	39,27	8,27	0,051
AR(1) + 1st order trend	0,99	24485,37	2,09	42,69	32,38	5,94	0,041
AR(1) + 4th order trend	0,99	23943,38	2,07	38,86	21,18	3,96	0,032

As can be seen from the obtained statistical characteristics, the use of an optimal filter improves the statistical characteristics of mathematical models.

As in previous case the best model is AR(1) plus trend of 4th order. The resulting one-step forecasting model has an error mean absolute percentage error of about 3,96 %, and Theil coefficient is $U = 0{,}032$. The main conclusion that can be drawn from the obtained modeling results is that the use of an optimal filter increases the predictive properties of the mathematical model.

Note that to implement the Kalman filter, a standard framework from the SAS Company [29] was used. Any data scientist can download this framework from the link below and, if necessary, modify it for their research tasks.

The application of GARCH for modeling the gold price

For checking the time serious of gold price for the presence of heteroscedasticity was used Goldfeld–Quandt test, which confirmed the presence of time-varying conditional variance in the dataset [7, 8]. When developing trading strategies for trading gold, the value of price variance is also taken into account, for which it is necessary to fit appropriate mathematical models. For constructing a heteroscedasticity models were used the GARCH and EGARCH models [6–8, 21]. Despite the fact that fairly high orders of autoregressive components were used for constructing the GARF models, their quality turned out to be insufficient compared compared with the EGARCH model [8, 21]. The EGARCH model for one-step forecasting showed better results according to the statistics obtained, namely values of MAPE (adapt.) given in the 6th column for the mode of operation with adaptation. Descriptive and predictive statistical characteristics of the obtained models are given in the table 5 (descriptive and predictive statistical characteristics of models for forecasting conditional variance).

Table 5

Model type	**Model quality**			**Forecast quality**			
	R^2	$\sum e^2(k)$	DW	MSE	MAPE (adapt.)	MAPE	Theil
GARCH(1,8)	0,92	115673	0,248	792,3	348,1	361,3	0,158
GARCH (1,12)	0,94	76498	0,372	372,5	142,7	158,2	0,103
GARCH (1,16)	0,96	60285	0,419	329,1	78,4	85,7	0,079
EGARCH (1,8)	0,99	35734	0,631	58,9	8,62	9,37	0,041

The best results among of all models were presented by the model in the form of exponential GARCH (1,8). For this model were obtained fairly good statistical characteristics for predicting the conditional variance, namely MAPE = 9,37 % (and 8,62 % in the adaptation mode). The proposed adaptation scheme allow it to improve the predictive effectiveness of the model in the range between 0,75–2 %, which is an excellent demonstration of the advantages of the proposed modeling approach. The using of ensembles approach of models based on various forecasting methods allows further reducing the value of the average absolute percentage forecast error for about 0,4–1,2 %.

Conclusion

The systemic methodology was proposed for modeling and forecasting economic, financial and other processes exhibiting nonstationary and nonlinear behavior, based on the following system analysis principles: hierarchical system structure, identification and taking into consideration probabilistic and statistical data uncertainties, availability of model adaptation and optimization procedures, generating multiple decision alternatives, and tracking the computational processes at all the stages of data processing with appropriate sets of statistical quality criteria. The methodology developed has a possibility for easy extension of its functional possibilities with new (or modified) parameter estimation techniques, forecasting methods, financial risk estimation, and generation of decision alternatives. High quality of the final result is achieved thanks to appropriate tracking of the computational processes at all data processing stages: preliminary data processing, model structure and parameter estimation, computing short- and middle-term forecasts, generating of decision alternatives. The methodology is based on the ideologically different methods of dynamic processes modeling: regression analysis and probabilistic approach, what creates appropriate basis for hiring various methods to achieve the best results of modeling and forecasting. The illustrative examples of the methodology application show that it can be used successfully for solving practical problems of forecasting dynamic processes evolution and risk estimation. The results of computational experiments lead to the conclusion that today nonlinear regression, state-space models, and Bayesian networks are appropriate instruments for financial data analysis due to the fact that they provide a possibility for reaching high quality models and forecasts. It also should be stressed that the methodology proposed turned out to be very useful instrument for a decision maker that helps to perform quality processing of statistical data using ideologically different techniques, appropriate sets of statistical quality criteria, generate alternatives and select the best one. The methodology can be used for supporting decision making process in various areas of human activities including development of strategy for industrial and agricultural enterprises, financial enterprises regarding risk management, investment companies etc.

Further extension of the methodology functions is planned with new modeling, forecasting and decision making techniques based on probabilistic methodology, fuzzy sets, neural networks and other artificial intelligence methods of data analysis. Appropriate decision support system will be developed based upon the ideas proposed in the study and systemic approach to implementation.

О.М. Трофимчук, П.І. Бідюк, О.М. Терентьєв, В.І.Клименко

МЕТОДОЛОГІЯ АДАПТИВНОГО МОДЕЛЮВАННЯ І ПРОГНОЗУВАННЯ НЕЛІНІЙНИХ НЕСТАЦІОНАРНИХ ПРОЦЕСІВ

Трофимчук Олександр Миколайович

Інститут телекомунікацій і глобального інформаційного простору НАН України, м. Київ,

Trofymchuk@nas.gov.ua

Бідюк Петро Іванович

Національний технічний університет «Київський політехнічний інститут імені Ігоря Сікорського»,

pbidyuke_00@ukr.net

Терентьєв Олександр Миколайович

Інститут телекомунікацій і глобального інформаційного простору НАН України, м. Київ,

o.terentiev@gmail.com

Клименко Вікторія Ігорівна

Інститут телекомунікацій і глобального інформаційного простору НАН України, м. Київ,

klimenko_vika@ukr.net

Дослідження спрямоване на створення системної методології для моделювання і прогнозування нелінійних нестаціонарних процесів в економіці, фінансах та інших галузях людської діяльності. Існує безліч проблем, які необхідно вирішувати за допомогою даних такого типу практично в усіх галузях людської діяльності, такі як побудова адекватних моделей, включаючи оцінювання і прогнозування станів досліджуваних систем; технічна, економічна та медична діагностика, автоматичне керування технологічними процесами; керування рухомими об'єктами; оцінювання і менеджмент фінансовими та іншими типами ризиків; дослідження взаємодії факторів ризику; керування мікрокліматом в теплицях та промислових підприємствах; підтримка рішень у бізнесі; динамічне стратегічне планування на виробництві; забезпечення стійкого розвитку критичної економічної інфраструктури та ін. Запропоновано процедури для адаптивного оцінювання структури і параметрів регресійних і ймовірнісних моделей, а також архітектура і функціональна схема відповідної інформаційної системи. Розробка інформаційної системи ґрунтується на принципах системного аналізу, таких як ієрархічна архітектура системи, адаптивне оцінювання структури моделей, оптимізація процедур оцінювання параметрів моделей, функціональна повнота системи, яка забезпечує її автономне функціонування, ідентифікація та врахування можливих невизначеностей, які зустрічаються при обробці даних і побудові математичних моделей, застосування належних множин критеріїв аналізу якості, які гарантують досягнення високої якості проміжних та остаточних результатів аналізу даних. Невизначеності зустрічаються при зборі даних, оцінюванні структури і параметрів моделей, в процедурах прогнозування і відіграють роль факторів негативного впливу на обчислювальні процедури в запропонованій інформаційній системі. Зменшення їх впливу сприяє підвищенню якості проміжних та остаточних результатів обчислень. Розглянуто ілюстративні приклади практичного застосування розробленої методології, що підтверджують її функціональність.

Ключові слова: нелінійні нестаціонарні процеси, моделювання, прогнозування, фільтр Калмана, узагальнені лінійні моделі.

REFERENCES

1. Efficient decision support systems — practice and challenges from current to future / in ed. Jao C.S. Rijeka : Intech, 2011. 492 p. DOI: https://doi.org/10.5772/682
2. McNeil A., Frey R., Embrechts P. Quantitative risk analysis. Princeton series in finance / in ed. Duffie D., Schaefer S. Princeton, NJ : Princeton. Princeton University Press, 2005. 554 p.
3. Trofymchuk O.M., Bidyuk P.I., Prosyankina-Zharova T.I., Terentiev O.M. Decision support systems for modeling, forecasting and risk estimation: a monograph. Riga : LAP Lambert Academic Publishing, 2019. 176 p.
4. Xekalaki E., Degiannakis S. ARCH models for financial applications. New York : John Wiley&Sons, Ltd., 2010. 535 p.
5. Бідюк П.І., Терентьєв О.М., Просянкіна-Жарова Т.І., Ефендієв В.В. Прогнозне моделювання нелінійних нестаціонарних процесів у рослинництві з використанням інструментів SAS Enterprise Miner. *Наукові вісті НТУУ «КПІ»*. 2017. № 1(111). С. 24–36. DOI: https://doi.org/10.20535/1810-0546.2017.1.87423
6. Bollerslev T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics.* 1986. Vol. 31. P. 307–327. URL: https://public.econ.duke.edu/~boller/Published_Papers/joe_86.pdf
7. Glosten L., Jagannathan R., Runkle D. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*. 1993. Vol. 48, N 5. P. 1779–1801. DOI: https://doi.org/10.1111/j.1540-6261.1993.tb05128.x
8. Lama A., Jha G.K., Paul R.K., Gurung B. Modelling and forecasting of price volatility: an Application of GARCH and EGARCH models. *Agricultural Economics Research Review*. 2015. Vol. 28, N 1. P. 73–82. DOI: https://doi.org/ 10.5958/0974-0279.2015.00005.1
9. Luengo D., Martino L., Bugallo M., Elvira V., Sarkka S. A survey of Monte Carlo methods for parameter estimation. *EURASIP Journal on Advances in Signal Processing.* 2020. N 25. P. 1–130. URL: https://arxiv.org/pdf/2107.11820.pdf
10. Harvey A.C. Forecasting, structural time series models, and the Kalman filter. *The Journal of the Operational Research Society*. 1991. Vol. 42, N 11. P. 1031–1035. DOI: https://doi.org/10.2307/2583225
11. Sugumaran V. Intelligent support system technology. London : IRM Press, 2002. 318 p.
12. Hollsapple C.W., Winston A.B. Decision support systems. Saint Paul : West Publishing Company, 1996. 860 p.
13. Бідюк П.І., Гожий О.П., Коршевнюк Л.О. Комп'ютерні системи підтримки прийняття рішень. Миколаїв : Вид-во ЧДУ ім. Петра Могили; К., 2012. 379 с.
14. Довгий С.О., Бідюк П.І., Трофимчук О.М. Системи підтримки прийняття рішень на основі статистично-ймовірнісних методів. Київ : Логос, 2014. 419 с.
15. Згуровский М.З., Панкратова Н.Д. Системный анализ: проблемы, методология, приложения. Київ : Наукова думка, 2005. 745 с.
16. Bidiuk P.I., Prosiankina-Zharova T.I., Terentieev O.M., Lakhno V.A., Zhmud O.V. Intellectual technologies and decision support systems for the control of the economic and financial processes. *Journal of Theoretical and Applied Information Technology.* 2019. Vol. 96, N 1. P. 71–87. URL: http://www.jatit.org/volumes/Vol97No1/7Vol97No1.pdf

17. Gilks W.R., Richardson S., Spiegelhalter D.J. Markov chain Monte Carlo in practice. New York : Chapman & Hall/CRC, 2000. 486 p.
18. Aizi S. Panel regression with nonstationary variables: Graduate Research Theses & Dissertations. URL: https://huskiecommons.lib.niu.edu/cgi/viewcontent.cgi?article=5250&context=allgraduate-thesesdissertations
19. Бідюк П.І., Романенко В.Д., Тимощук О.Л. Аналіз часових рядів. Київ : Політехніка, 2010. 317 с.
20. Dickey D.A. Stationarity issues in time series models. *In Proceedings of the Thirtieth Annual SAS Users Group International Conference*. Cary, NC : SAS Institute Inc., 2005. URL: https://support.sas.com/resources/papers/proceedings/proceedings/sugi30/192-30.pdf
21. Shamiri A., Mohd A.H. Shaari Nor. Modeling and forecasting volatility of the Malaysian and the Singaporean stock indices using asymmetric GARCH models and non-normal densities. *Malaysian Journal of Mathematical Sciences*. 2007. Vol. 1(1). P. 83–102.
22. De Jong P. The diffuse Kalman filter. *Annals of Statistics*. 1991. Vol. 19, N 2. P. 1073–1083. DOI: https://10.1214/aos/1176348139
23. Nelson C.R., Plosser C.I. Trends and random walks in macroeconomic time series: some evidence and implications. *Journal of Monetary Economics*. 1982. Vol. 10. P. 139–162. URL: https://hedibert.org/wp-content/uploads/2015/03/nelson-plosser-1982.pdf
24. Karpathy A. Convolutional neural networks for visual recognition (course website). URL: https://cs231n.github.io
25. Zeiler M., Fergus R. Visualizing and understanding convolutional networks. *Computer Vision and Pattern Recognition*. 2013. Vol. 1311.2901. DOI: https://doi.org/10.48550/arXiv.1311.2901
26. Zikmundová M., Helisová K., Benes V. On the use of particle Markov chain Monte Carlo in parameter estimation of space-time interacting discs. *Methodology and Computing in Applied Probability*. 2014. Vol. 16. P. 451–463. DOI: https://doi.org/10.1007/s11009-013-9367-2
27. Kantas N., Doucet A., Singh S.S., Maciejowski J.M. An overview of sequential monte carlo methods for parameter estimation in general state-space models. *IFAC Proceedings Volumes Symposium on System Identification*. 2009. Vol. 42, N 10. P. 774–785. DOI: https://doi.org/10.3182/20090706-3-FR-2004.00129
28. Tuyet Anh Thi Vu, Ba-Ngu Vo, Rob Evans. A particle marginal Metropolis-Hastings multi-target tracker. *IEEE Transactions on Signal Processing*. 2014. Vol. 62(15). P. 3953–3964. DOI: https://doi.org/10.1109/TSP.2014.2329270
29. SAS Documentation. URL: https://documentation.sas.com/doc/en/pgmsascdc/v_047/imlug/imlug_timeseriesexpls_sect003.htm

Submitted 29.01.2024