TREND, DECOMPOSITION AND COMBINED APPROACHES OF TIME SERIES FORECASTING BASED ON THE “CATERPILLAR”-SSA METHOD

Vitalii Shchelkalin

Abstract: *In presented paper mathematical models and methods based on joint applying ideas of the “Caterpillar”-SSA and Box-Jenkins methods are produced. This combination of models lead to a synergy and mutually compensate the opposite by nature shortcomings of each models separately and increases the accuracy and stability of the model. The further development of technique for models constructing, technique of Box-Jenkins, and improvement of themselves autoregressive integrated moving average (ARIMA) models, designed about forty years ago and remaining in present time as one of the most efficient models for modeling, forecasting and control exceeding their own rivals on whole row of criterions such as: economy on parameters quantity, labour content of models building algorithm and resource-density of their realization, on formalization and automation of models construction is produced. A novel autoregressive spectrally integrated moving average (ARSIMA) model which describes a wider class of processes in contrast to the Box-Jenkins models is developed. Decomposition and combined forecasting methods based on “Caterpillar”-SSA method for modeling and forecasting of time series is developed. The essence of the proposed decomposition forecasting method and combined forecasting method consist in decomposing of time series (exogenous and predicted) by the “Caterpillar”-SSA method on the components, which in turn can be decomposed into components with a more simpler structure for identification, in selection from any of these components of constructive and dropping the destructive components, and in identification of those constructive components that are proactive on the propagated time series, or vice versa if its delay interval is less than the required preemption interval of forecasting, mathematical models with the most appropriate structure (in the combined approach) or its ARIMAX models (in decomposition approach) models and calculation of their predictions to the required lead time, to use the obtained models, or as a comb filter (in the case of signals modeling), or as an ensemble of models, setting the inputs of MISO model or used as a component of the combined mathematical model whose parameters are adjusted to further cooptimization method. In such models as inputs can be also include the instantaneous amplitudes, obtained after applying the Hilbert transform to the components of the expansions. The advantages of the proposed methods for models of the processes constructing is its rigorous formalization and, therefore, the possibility of complete automation of all stages of construction and usage of the models.*

Keywords: *modeling, filtering, forecasting, control, "Caterpillar"-SSA method, ARIMA model, "Caterpillar"-SSA – ARIMA – SIGARCH method, ARSIMA model, ARSIMA – SIGARCH model, heteroskedasticity, Levenberg-Marquardt method, Davidon–Fletcher–Powell method, decomposition forecasting method, combined forecasting methods, Hilbert transform.*

Introduction

To this time a large amount of mathematical models and methods of time series analysis and forecasting is created and it is trend to combining mathematical models in order to obtain the best characteristics of the final models in the sense of formation and the translation of some of the theoretical developments of some methods to other methods where it is can be possible and appropriate. In this paper the synthesis of mathematical models based on the joint usage of ideas of two methods, to each of which individually skeptically pertain many specialists from area of forecasting, but sufficiently well grounded in theory is produced. The first method - a deterministic method for different types of analysis and forecasting has not been entered yet into the standard mathematical packages – a "Caterpillar"-SSA method. The second – a statistical method – a Box Jenkins method. The proposed models realize trended approach, which is concluded in modeling of the process as deflections of actual values relatively to the trend component and lead to synergies, mutually compensating of opposite in nature shortcomings of models which their forming.

The development of the methods of mathematical modeling, forecast and control is determined by the degree of mathematical description of processes, phenomena and objects taking place in various branches of science and technology, taking into account both the mathematical and technical advantages and achievements and shortcomings and limitations, as well as the quality and volume of the sample data and resource constraints, including the temporal of formation an adequate mathematical model of the process.

A general class of multivariate polynomial multiply stochastic models

Polynomial models can adequately describe an extremely wide class of multidimensional multiply stochastic systems and processes and to solve on their basis the tasks of modeling, forecasting and control. A general class of multivariate polynomial multiply stochastic models is presented below.

A general class of multivariate polynomial multiply stochastic models for the system having r exogenous variables, m outputs and m noises can be represented as:

$$F(q)\cdot Y(k)=\sum_{i=1}^{N}\frac{B_i(q)}{A_i(q)}\cdot\Theta_{X_i}\cdot X_i(k)+\frac{C'(q)}{D'(q)}\cdot W(k) \tag{1}$$

or

$$Y(k)=\sum_{i=1}^{N}P_i(q)\cdot X_i(k)+R(q)\cdot W(k),$$

where q - or backshift operator such that $q^i X(k)=X(k-i)$, here the time in parentheses; $Y(k)$, $X(k)$, $k=\overline{0,n-1}$ - centered multivariate time series; $P_i(q)=\frac{B_i(q)}{A_i(q)}\cdot\Theta_{X_i}$ – the models of transfer functions; $R(q)=\frac{C'(q)}{F(q)\cdot D'(q)}$, $R(q)\cdot W(q)$ – additive noise model in the form of moving average; $C'(q)=C^1_{q_1}\left(q^{s_1}\right)\cdot C^2_{q_2}\left(q^{s_2}\right)\cdot\ldots\cdot C^1_{q_n}\left(q^{s_n}\right)=\prod_{i=1}^{n}C^i_{q_i}\left(q^{s_i}\right)$, $C'(q)$ – a generalized moving average operator; $D'(q)=D^1_{p_1}\left(q^{s_1}\right)\cdot D^2_{p_2}\left(q^{s_2}\right)\cdot\ldots\cdot D^1_{p_n}\left(q^{s_n}\right)=\prod_{i=1}^{n}D^i_{p_i}\left(q^{s_i}\right)$, $D'(q)$ – generalized autoregressive operator; $C^k_{n^k_c}\left(q^{s_k}\right)=I+\sum_{j=1}^{n^k_c}C^k_j q^{j\cdot s_k}$ – matrix polynomial from q^{s_k} of order n^k_c, defining moving average component of the periodic component with period s_k; $D^k_{n^k_d}\left(q^{s_k}\right)=I+\sum_{j=1}^{n^k_d} D^k_j q^{j\cdot s_k}$ – the matrix polynomial from q^{s_k} of order n^k_d, defining autoregressive component of the periodic component with period s_k; $A_i(q)$, $B_i(q)$, $C(q)$, $D(q)$, $F(q)$, $G(q)$, $H(q)$ – matrix polynomials $A_i(q)=I+\sum_{j=1}^{n_a} A^i_j q^j$, $B_i(q)=\sum_{j=0}^{n_b} B^i_j q^j$, $F(q)=I+\sum_{j=1}^{n_f} F_j q^j$ of orders n_a, n_b, n_h, n_g and n_f respectively; Θ_{X_i} – delay matrix of exogenous variables with respect to the outputs of the system, which has the form $\Theta_{X_i}=\begin{bmatrix} q^{c^{11}_{X_i}} & q^{c^{12}_{X_i}} & \vdots & q^{c^{1r}_{X_i}} \\ q^{c^{21}_{X_i}} & q^{c^{22}_{X_i}} & \vdots & q^{c^{2r}_{X_i}} \\ \cdots & \cdots & \ddots & \cdots \\ q^{c^{m1}_{X_i}} & q^{c^{m2}_{X_i}} & \vdots & q^{c^{mr}_{X_i}} \end{bmatrix}$; $c^{kj}_{X_i}$ – time delay of j th exogenous variable relative to output; $X(k)=\left(x_1(k) \;\; x_2(k) \;\; \dots \;\; x_r(k)\right)^T$, $Y(k)=\left(y_1(k) \;\; y_2(k) \;\; \dots \;\; y_m(k)\right)^T$, $W(k)=\left(w_1(k) \;\; w_2(k) \;\; \dots \;\; w_m(k)\right)^T$ – vectors of inputs, outputs and noise, respectively, for the k th time moment. And as for the model with a scalar output value, a rough preliminary structural and parametric identification of the matrix polynomial $F(q)$ is supposed to find by the "Caterpillar"-SSA method.

If in the model (1) $A_i(q)=I$, $i=\overline{1,N}$ and $D'(q)=I$ then we obtain seasonal VARMAX model:

$$F(q)\cdot Y(k)=\sum_{\mathrm{i}=1}^{N} B_i(q)\cdot \Theta_X \cdot X(k)+C'(q)\cdot W(k) \tag{2}$$

and if $F(q)=I$ - the Box-Jenkins model:

$$Y(k)=\sum_{i=1}^{N}\frac{B_i(q)}{A_i(q)}\cdot \Theta_{X_i}\cdot X_i(k)+\frac{C'(q)}{D'(q)}\cdot W(k) \tag{3}$$

Let's write models (2), (3) and (1) in difference form.

Box-Jenkins model:

$$\prod_{i=1}^{N} A_i(q) D'(q) Y(k)=D'(q)\sum_{i=1}^{N} B_i(q)\cdot \Theta_{X_i}\cdot X_i(k)+\prod_{i=1}^{N} A_i(q) C'(q) W(k)$$

or

$$A'(q)Y(k)=\sum_{i=1}^{N} B'_i(q) X_i(k)+C''(q)W(k),$$

where $A'(q)=\prod_{i=1}^{N} A_i(q) D'(q)$, $B'_i(q)=D'(q)B_i(q)\cdot \Theta_{X_i}$, $i=\overline{1,N}$, $C''(q)=\prod_{i=1}^{N} A_i(q) C'(q)$.

The general polynomial model (1):

$$\prod_{i=1}^{N} A_i(q) D'(q) F(q) Y(k)=D'(q)\sum_{\mathrm{i}=1}^{N} B_i(q)\cdot \Theta_{X_i}\cdot X_i(k)+\prod_{i=1}^{N} A_i(q) C'(q) W(k)$$

or

$$A''(q)Y(k)=\sum_{i=1}^{N}B_i''(q)X_i(k)+C''(q)W(k),$$

where $A''(q)=\prod_{\mathrm{i=1}}^{N}A_i(q)D'(q)F(q)$, $B_i''(q)=D'(q)\sum_{i=1}^{N}B_i(q)\cdot\Theta_{X_i}$, $C''(q)=\prod_{\mathrm{i=1}}^{N}A_i(q)C'(q)$.

ARIMA model

A mathematical model of the processes that depend from several exogenous factors in the operator form can be presented as a model of the seasonal autoregressive integrated moving average (SARIMA) [Бокс, Дженкинс, 1974], [Евдокимов, Тевяшев, 1980]:

$$y_t=\sum_{i=1}^{N}\frac{\omega_{c_i}^i(B)}{\delta_{r_i}^i(B)}x_{t-b_i}^i+\frac{\theta_{q^*}^*(B)}{\Phi_{p^+}^+(B)}e_t \tag{4}$$

where

y_t , $t=\overline{1,n}$ – initial or transformed (normalized or logarithm) centered time series of the predicted values;

n – volume of time series;

x_t^i, $t=\overline{1,n}$, $i=\overline{1,N}$ – initial or transformed (normalized or logarithm) centered time series of the external (exogenous) factors;

N – the number of exogenous factors;

B – the lag operator or backshift operator which operates on an element of a time series x_t to produce the previous element such that $B^i x_t=x_{t-i}$;

b_i – the time delay of i th exogenous time series x_t^i relative to predicted time series y_t;

$\delta_{r_i}^i(B)$ and $\omega_{c_i}^i(B)$ – polynomials of the transfer functions form B of r_i and c_i degrees respectively;

$\delta_{r_i}^i(B)=1-\delta_1^iB-\delta_2^iB^2-\ldots-\delta_{r_i}^iB^{r_i}$;

$\omega_{c_i}^i(B)=\omega_0^i-\omega_1^iB-\omega_2^iB^2-\ldots-\omega_{c_i}^iB^{c_i}$;

$\Phi_{p^+}^+(B)$ – the generalized autoregressive operator of the order $p^+=p^*+\sum_{i=1}^{n_s}D_iS_i$, $p^*=\sum_{i=1}^{n_s}p_iS_i$;

$\Phi_{p^+}^+(B)=\Phi_{p^*}^*(B)\nabla_{S_1}^{D_1}\nabla_{S_2}^{D_2}\ldots\nabla_{S_{n_s}}^{D_{n_s}}$;

D_i , $i=\overline{1,n_s}$ – the order of difference capture S_i;

S_i , $i=\overline{1,n_s}$ – the period the i th periodic component and S_1 = 1;

n_s – the amount of periodic components;

∇_{S_i} and B^{S_i} – simplify operators such that $\nabla_{S_i}x_t=\left(1-B^{S_i}\right)\cdot x_t=x_t-x_{t-S_i}$.

$\Phi^*_{p^*}(B)$ – the generalized autoregressive operator of the order p^* of form $\Phi^*_{p^*}(B) = \prod_{i=1}^{n_s} \Phi^i_{p_i}\left(B^{S_i}\right)$;

$\Phi^i_{p_i}\left(B^{S_i}\right)$, $i = \overline{1, n_s}$ – polynomials from B^{S_i} of orders p_i respectively that define the components of autoregression of the periodic components with periods s_i respectively;

$\theta^*_{q^*}(B)$ – the generalized moving average operator of order $q^* = \sum_{i=1}^{n_s} q_i S_i$ of form $\theta^*_{q^*}(B) = \prod_{i=0}^{n_s} \theta^i_{q_i}\left(B^{S_i}\right)$;

$\theta^i_{q_i}\left(B^{S_i}\right)$, $i = \overline{1, n_s}$ – polynomials from B^{S_i} of orders q_i respectively that define the moving average components of the periodic components with periods S_i respectively;

e_t – error terms of the model that are generally assumed to be independent, identically distributed variables sampled from a normal distribution with zero mean.

The solution of task of structural identification of the parameters c_i, r_i, b_i $i = \overline{1, N}$ is performed automatically by the analysis of the response function to a unit impulse. The automated identification of parameters p_i, $i = \overline{1, n_s}$ - by analyzing of the partial autocorrelation function (PACF), and parameters q_i, $i = \overline{1, n_s}$ – by analyzing the autocorrelation function (ACF). The parameters n_s, D_i, S_i are also identified by the analysis of ACF and PACF of the process. Parametric identification is performed by the method of Levenberg-Marquardt or Davidon-Fletcher-Powell.

In [Седов, 2010], [Бэнн, Фармер, 1987] presents other additive, multiplicative and hybrid forecasting models. There also offered a variety of models for simulating and forecasting the trend, seasonal, weekly, residual components of the processes.

Trend approach of time series forecasting

The main requirements to the mathematical models construction in 70th – 90th years of the last century, is an economical number of parameters, the velocity of the model determination and its resource-intensity for use on available then computers with low productivity. However, modern computer technology and mathematical modeling methods provide a great possibilities for analysis, modeling and forecasting time series of the different nature. Therefore, at present these requirements are not crucial and modern computing tools and systems allow to stand on the first plan the requirement of modeling accuracy, quality of the analysis and forecasting.

One of the most widely used models that corresponding to above requirements are the ARIMA models. Box-Jenkins method works only with pre-reduced to the stationary form time series. Nonstationary series are usually characterized by the presence of high power at low frequencies. However, in many practical applications of interest information may be concentrated at high frequencies. In such cases, all that was done - it is filtered out non-stationary low-frequency components and was used the remainder of the series for further analysis. At the same time as a filter to eliminate low-frequency component in the ARIMA model used a filter of the first differences or maximum second. Watching the gain of the filter can be seen that low frequency considerably weakened and, therefore, be less visible at the filter output. So the method of seasonal ARIMA model was satisfactorily predicted only with a relatively simple structure time series.

In the 80th years of last century Granger and Joyeux [Granger, Joyeux, 1980] proposed a new class of ARFIMA models which is convenient to describe the financial and economic time series with the effects of long and short memory.

Known publications using the "Caterpillar"-SSA method in various branches of science and technology as a method of fairly good description of non-stationary time series with linear, parabolic or exponential trend with not always stable oscillatory component, however studies have identified a number of significant shortcomings of the method, greatly limiting its applicability. Method for modeling uses suboptimal in terms of accuracy of some time series of orthogonal basis vectors of the trajectory matrix. Therefore, the main idea of the origin of the proposed method was first concluded in joint use the "Caterpillar"-SSA method and models of autoregressive moving average, trained on a competitive base, with account generalized criterion of the accuracy and adequacy [Щелкалин, 2010]. Using such combination was dictated by the fact that individually, these approaches have several disadvantages, but their joint usage brings synergy, increasing their efficiency, robustness and adequacy. However, the trend separation by "Caterpillar"-SSA method, as well as any other method, the residual component of the series in most cases is non-stationary, and therefore hereinafter "Caterpillar"-SSA method has been used in combination with the model of autoregressive integrated moving average (ARIMA).

In standard "Caterpillar"-SSA method the signal or the model are defined only from condition of the dispersions reproduction of the time series, but the nature of the errors of the reproduction in model is not included. The "Caterpillar"- SSA method, founded on LS or MV estimation, received for decision of this problem more applicable in the theory of filtration and does not include any information whether the noise is white or not. Therefore, at present there are two approaches to solve this problem. In base of the first approach lies the idea of the increasing the order of the linear recurrence formula (LRF) of the recurrence forecast of the "Caterpillar"-SSA method by HTLS method - a modification of the ESPRIT method [Голяндина, Шлемов, 2012]. The second - by adding to the SSA-forecasts the Box-Jenkins forecasts, constructed on the remaining part of the time series after removing from it the signal recovered by "Caterpillar"-SSA method [Щелкалин, 2012]. In fact - by adding to the LRF the ARIMA model. Beneficial effects appearing already as a result of arising from the using such additive model (5) is described in [Щелкалин, 2012]. Identified by the "Caterpillar"-SSA signal is well approximated by ARIMA models (see model (8)). Subsequently, the model has taken the form (7), and most importantly, all of its parameters were already adjusted by optimization method to target function of integrated criterion of accuracy and adequacy. The author called model (7) as autoregressive model spectrally integrated moving average (ARSIMA) [Shchelkalin, 2011]. For the case of several seasonal components and several exogenous variables in the model are summarized in a similar manner as, for example, model (8) applies to model (10).

$$\mathrm{y}_t = g(B) w_t + \frac{\theta(B)}{\nabla \varphi(B)} e_t ; \tag{5}$$

$$\mathrm{y}_t = g(B) y_t + \frac{\theta'(B)}{\nabla \varphi'(B)} e_t \tag{6}$$

$$g'(B) y_t = \frac{\theta''(B)}{\omega(B) \varphi''(B)} e_t \tag{7}$$

$$\begin{cases} \hat{w}_t = \dfrac{\theta'(B)}{\nabla\varphi'(B)}\varepsilon_t; \\ y_t = \hat{w}_t + \dfrac{\theta(B)}{\nabla\varphi(B)}e_t; \end{cases} \Rightarrow y_t = \frac{\theta'(B)}{\nabla\varphi'(B)}\varepsilon_t + \frac{\theta(B)}{\nabla\varphi(B)}e_t; \quad (8)$$

$$y_t = g(B)\frac{\theta'(B)}{\nabla\varphi'(B)}\varepsilon_t + \frac{\theta'''(B)}{\nabla\varphi'''(B)}e_t \quad (9)$$

$$y_t = \left(\sum_{i=1}^{N}\frac{\omega_i'(B)}{\delta_i'(B)} + \sum_{i=1}^{N}\frac{\omega_i(B)}{\delta_i(B)}\right)B^{b_i'+b_i}x_t^i + \frac{\prod_{i=1}^{S_n'}\theta'^{S_i'}(B)}{\prod_{i=1}^{S_n'}\nabla_{S_i'}^{D_i'}\prod_{i=1}^{S_n'}\varphi'^{S_i'}(B)}\varepsilon_t + \frac{\prod_{i=1}^{S_n}\theta^{S_i}(B)}{\prod_{i=1}^{S_n}\nabla_{S_i}^{D_i}\prod_{i=1}^{S_n}\varphi^{S_i}(B)}e_t, \quad (10)$$

where w_t - identified by the "Caterpillar"-SSA method deterministic component of the process; L - a window length of method SSA; B - an operator of delay; $g(B)= \sum_{i=1}^{L-1} g_{L-i}B^i$ - polynomial of the delay operator B the initial coefficients of which are the coefficients of minnormLRF of the method "Caterpillar»-SSA g_i, $i=\overline{1,L-1}$ [Голяндина, Шлемов, 2012]; $g'(B)=1-g(B)$ = $\sum_{i=1}^{R-1} g''_{R-i}B^i$, where R - the order of a minimal length LRF; The operator $g(B)$ can be applied directly to the process y_t, then (5) can be written as (6), and in (7) $\omega(B)$ = ∇, and in the most cases $\omega(B)$ can be taken equal to $(1-B)$, because after applying of the method "Caterpillar»-SSA residual time series or immediately becomes stationary, or becomes after a single applying of the operator $(1-B)$, what increases the stability of the model, because repeated applying leads to the appearance of multiple roots of characteristic polynomial of the model that lying on the boundary of stability [Shchelkalin, 2011].

Structural identification of the models is performed, in addition to the analysis of the ACF, PACF and the response function to a unit impulse, by selecting of the moving window at formation of embedding vectors of time series, by analysis of the eigenvalues and eigenvectors of the SVD-decomposition of the trajectory matrix of the processes.

By analyzing of changes in time of the first principal components of the SVD decomposition of the trajectory matrix of the time series we can detect the change-points in time series structure.

In the scientific literature have long been known combined probability and deterministic model presented in [Бэнн, Фармер, 1987]. In [Щелкалин, 2011] is offered next modification of the ARIMA and the GARCH models and at first proposed "Caterpillar"-SSA – ARIMA – SIGARCH method and combined probabilistic and deterministic model of autoregression spectrally integrated moving average with spectrally integrated generalized autoregressive heteroskedasticity (ARSPSS – SIGARCH) and the method of its construction, which allows greater flexibility in time series analyzing, modeling and forecasting in comparison with the ARIMA – GARCH models.

The method "Caterpillar"-SSA is applied by author:

for structural identification of the determined component of processes;

for division into long and short memory of processes;

for cointegration interconnected processes;

for time series decomposition on independence components;

for time series decomposition on trend, weekly and seasonal components;

for separation for a finite and separately for deadbeat regulators in the case of use the proposed models in control theory;

for preliminary rough structural identification of the decomposing artificial neural network.

And the Box-Jenkins method is used for structural identification of a remaining stationary (or kvazi-stationary) additional part of processes.

In the time series analysis and forecasting that depending from several other time series it is essential the balancing of the dynamic properties of the variables that appearing in the left and right sides of the model equation. In this case, the idea of the "Caterpillar"-SSA method stand in preliminary generalized cointegration of time series and the model can be represented as follows:

$$\begin{aligned}
\widetilde{w}_t^{y} &= \frac{\omega_{y\;c_y^{w^y}}^{w^y}(B)}{\delta_{y\;r_y^{w^y}}^{w^y}(B)} \cdot y_{t-b_y^{w^y}} + \sum_{i=1}^{N} \frac{\omega_{x^i\;c_{x^i}^{w^y}}^{w^y}(B)}{\delta_{x^i\;r_{x^i}^{w^y}}^{w^y}(B)} \cdot x^i_{t-b_{x^i}^{w^y}} + \frac{\theta_{q_{w^y}^*}^{w^y\;*}(B)}{\varphi_{p_{w^y}^+}^{w^y\;+}(B)} \cdot e_t^{w^y}; \\
y_t &= g^y(B) \cdot \widetilde{w}_t^y + \sum_{i=1}^{N} \frac{\omega_{x^i\;c_{x^i}^y}^{y}(B)}{\delta_{x^i\;r_{x^i}^y}^{y}(B)} \cdot x^i_{t-m_{x^i}^y} + \frac{\theta_{q_y^*}^{y\;*}(B)}{\varphi_{p_y^+}^{y\;+}(B)} \cdot e_t
\end{aligned} \tag{11}$$

or

$$g'^y(B) y_t = \sum_{i=1}^{N} \frac{\omega'^{y}_{x^i\;c_{x^i}^y}(B)}{\delta'^{y}_{x^i\;r_{x^i}^y}(B)} \cdot F^i_{t-m^Q_{F^i}} + \frac{\theta'^{y\,*}_{q_y^*}(B)}{\varphi'^{y\,+}_{p_y^+}(B)} \cdot e_t \;.$$

In such statement of the problem this models is closely connected with so named subspace based methods, founded on estimation of signal subspace and then use characteristics of this subspace both for analysis of the signal, and for its continuations.

Decomposition approach of time series forecasting

Recently, in various branches of science and technology the models and methods of digital signal processing for modeling and forecasting of processes are beginning to be used.

SARIMAX model is well applicable for predicting both deterministic and stochastic processes. So, returns to the models (8), (11) in decomposition forecasting method (DFM) involves identifying and building ARIMAX model is not only of time series $\hat{\widetilde{w}}_t^y$ and time series $\widetilde{w}^{x^i}{}_0$, $\widetilde{w}^{x^i}{}_1$, ..., $\widetilde{w}^{x^i}{}_{n-1}$, $i=\overline{1,N}$, $t=\overline{1,n}$ obtained at the fourth stage of diagonal averaging of "Caterpillar"-SSA method of matrixes Z^i, $i=\overline{1,N+1}$ that consisting of K columns from $(i-1)\cdot K$ th to $(i-1)\cdot K$ th of matrix Z, where $Z=\widetilde{Z}^i+\ldots+\widetilde{Z}^j$ - the amount of decompositions of the matrix $\widetilde{Z}^i = \left(U^i \cdot \left(U^i\right)^T \cdot X_1 \quad U^i \cdot \left(U^i\right)^T \cdot X_2 \quad \ldots \quad U^i \cdot \left(U^i\right)^T \cdot X_N \quad U^i \cdot \left(U^i\right)^T \cdot Y \right)$, that selected by standard analysis of the eigenvalues of the SVD-decomposition of trajectory matrix in the method of "Caterpillar"-SSA, but to identify and build ARIMAX models of each time series $w_t^{(j)}$ of decomposition (diagonal averaging of submatrixes $\widetilde{Z}^{j\,i}$, $i=\overline{1,L^y}$, which consisting of K columns from $(j-1)\cdot K$ th to $j\cdot K-1$ th of matrix $\widetilde{Z}^i$, $i=\overline{1,L^y}$), which can be distributed on time series with more simple structure (model (12), figure 1). Thus, it is can potentially to generate $L^y \times (N+1)$ time series of expansion (components) of the original predicted time series and exogenous time series, which are, as was mentioned above, subject to modeling, after selection from these components of the constructive and remove destructive. However, there is no need create all $L^{\mathrm{y}} \times (N+1)$ models, but are formed only meaningful, ie which modeling the structural components. To selection of the most meaningful components (most correlated with the projected time series) author proposed to use the algorithm of fast orthogonal search (FOS) [Ahmed, 1994] or some other method. In this case the columns of the algorithm matrix are the values obtained by the "Caterpillar"-SSA components and exogenous expansions of projected time series, the initial values and exogenous and projected time series, as well as their delay up to m order. The values of the constructed models of constructive components necessary for their further use or as an ensemble of models by applying the model to the input of Box-Jenkins model, or use a combination of models, adjusting their parameters together $y_t = \sum_{k\in M} \hat{w}_t^{(k)}$, where M - the set of indices of selected structural components. In such models as inputs can be also include the instantaneous amplitudes, obtained after applying the Hilbert transform to the components of the expansions (see figure 1 and model (12)). Thus, the model (11) becomes:

$$\hat{w}_t^{(k)} = \frac{\omega_y^{w^{(k)}}{}_{c_y^{w^{(k)}}}(B)}{\delta_y^{w^{(k)}}{}_{r_y^{w^{(k)}}}(B)} \cdot y_{t-b_y^{w^{(k)}}} + \sum_{i=1}^{N} \frac{\omega_{x^i}^{w^{(k)}}{}_{c_{x^i}^{w^{(k)}}}(B)}{\delta_{x^i}^{w^{(k)}}{}_{r_{x^i}^{w^{(k)}}}(B)} \cdot x^i_{t-b_{x^i}^{w^{(k)}}} +$$

$$+ \sum_{i=1, i\neq j}^{M} \frac{\omega_{w^{(i)}}^{w^{(k)}}{}_{c_{w^{(i)}}^{w^{(k)}}}(B)}{\delta_{w^{(i)}}^{w^{(k)}}{}_{r_{w^{(i)}}^{w^{(k)}}}(B)} \cdot \hat{w}^{(i)}_{t-b_{w^{(i)}}^{w^{(k)}}} + \frac{\theta^{w^{(k)}\,*}_{q^*_{w^{(j)}}}(B)}{\varphi^{w^{(k)}\,+}_{p^+_{w^{(k)}}}(B)} \cdot e_t^{w^{(k)}}, \ k\in M; \tag{12}$$

$$y_t = \sum_{k\in M} \frac{\omega^y_{w^{(k)} c^y_{w^{(k)}}}(B)}{\delta^y_{w^{(k)} r^y_{w^{(k)}}}(B)} \cdot \hat{w}^{(k)}_{t-b^y_{w^{(k)}}} + \sum_{i=1}^{N} \frac{\omega^y_{x^i c^y_{x^i}}(B)}{\delta^y_{x^i r^y_{x^i}}(B)} \cdot x^i_{t-b^y_{x^i}} + \sum_{k\in \mathrm{M}} \frac{\omega^y_{A^k c^y_{A^k}}(B)}{\delta^y_{A^k r^y_{A^k}}(B)} A^{\mathrm{k}}_{\mathrm{t}} + \frac{\theta^{y\,*}_{q^*_y}(B)}{\varphi^{y\,+}_{p^+_y}(B)} \cdot e_t$$

where M - is the set of indices of selected structural components; A_t^k, $k\in\mathrm{M}$ - instantaneous amplitudes obtained after applying of the Hilbert transform to the selected structural components.

Also in the next paper will be offered the decomposition artificial neural network. These models will be able to determine the relationship between the internal structural components defined by the decomposition of time series.

In contrast to SARIMAX model this decomposition model uses SARIMAX for modeling and forecasting of simplified time series of decomposition, while previously model SARIMAX modeled the entire process. However, most processes are heterogeneous nonstationary stochastic processes with polyharmonic, polynomial and stochastic trends, modulated in amplitude and frequency, have a complex correlation structure and for an adequate description of the SARIMAX model time series were reduced to stationary form by taking first differences or the maximum of the latter. Therefore, the SARIMAX method satisfactorily modeled time series only with a relatively simple structure, and the capture of higher-order differences, the stability of the model was lost. Thus, the use of (12) to simulate of signals increases the stability of the model and its accuracy and allows to simulate the signals which are modulated by amplitude and frequency.

Combined approach of time series forecasting

2000th years are characterized by the using for a wide range of models for time series analyzing and forecasting, as well as ensembles of models with different structures. With the advent of high-speed computer the transition from ensembles of predictive models to its combinations was occurred. The difference between the combined models and its ensembles lies in the simultaneous adjustment of model parameters.

For each specific area, selecting the most appropriate models and methods of processes forecasting taking into account the following characteristics:

the way of modeling the various components (trend, cyclical, seasonal, residual, etc.) [Бэнн, Фармер, 1987], [Седов, 2010] or less interpretable components of time series decomposition;

the way of accounting for the influence of external factors on the process;

the way of accounting the influence of expansion components, corresponding to different time series, each to other;

the method of modeling the relationships of latent time series;

the way of modeling the relationship between latent time series and forecasted time series;

the methods of modeling of the random components of time series;

the way of time series clustering on areas with a similar structure, etc.

The combined forecasting method (CFM) consists in decomposing by any method (in this case by the "Caterpillar"-SSA method) of exogenous and forecasted time series into components, which in turn can be decomposed into components with a simpler to identify structure (figure 1) or grouped into more interpretable components of the time series (figure 2), such as: trend, seasonal, weekly and residual, and in the selection by any methods the constructive and dropping the destructive of these components, and in the identification of those constructive components that have a pre-emptive nature on the forecasted time series, or vice versa which delay interval is less than the required interval of pre-emption of forecasting, the mathematical models with the most appropriate structures for each specific component of the time series and calculation of their predictions with the required lead time, to use the obtained models or as an ensemble of models, setting the inputs of MISO-models, or used as components of a model which parameters are adjusted collectively by optimization method.

The essence of the combined approach is to expand in any method of forecasted time series into components: y_t^T (trend), y_t^A (seasonal), y_t^D (weekly) и y_t^R (remain), and subsequently finding the forecasts of each of the components $y_t^T(l)$, $y_t^A(l)$, $y_t^D(l)$ и $y_t^R(l)$, and in finding the generalized prediction. Moreover, when calculating the forecasts of each of the components of time series are recommended certain models.

In order to predict the trend-seasonal component is proposed to use:

– a model of the moving average;
– a model based on decomposition on a finite Fourier series;
– a model based on the Kalman filter;
– a model based on polynomial interpolation;
– an exponential smoothing model;
– a weighted moving average model;
– the neural network model;
– Group Method of Data Handling (GMDH);
– structurally flexible polynomials and harmonic series [Крисилов, Побережник, 2003];
– bordering functions [Седов, 2010];
– wavelets;
– others.

In order to predict weekly component:

– a model of the moving average;
– an exponential smoothing model.
– a model based on polynomial interpolation;
– ARIMA model;
– others.

To account for the exogenuouse factors:

– a model of the moving average;
– a model based on decomposition into finite Fourier series.

To predict the residual component:

– an autoregressive model;
– an exponential smoothing model;
– a model of spectral decomposition;
– ARIMAX model;

Thus, to obtain adequate models of complex non-stationary processes, particularly generated by multiple sources, and high-quality predictions it is necessary to produce their decomposition and combine the models with miscellaneous structure. On these considerations are based the combined model that based on deterministic and statistical models, which differ from those offered models by the simultaneous computation of the model coefficients by the optimization method such as Levenberg-Marquardt or Davidon–Fletcher–Powell algorithms.

As shown in [Смоленцев, 2008], [Щелкалин, 2011] for filtering and signal simulation it is well used principal component analysis (PCA), which lies at the basis of the "Caterpillar"-SSA with wavelets. A promising development of such methods is the combination of multiscale principal component analysis (MPCA), in which PCA is applied not only to the signal, but also to the wavelet coefficients and the wavelet components of a multidimensional signal. Thereby, such way it is possible to create the method "Caterpillar"-SSA based on wavelet analysis. In Kurbatskii, Sidorov, Spiryaev, Tomin, 2011] proposed a two-stage adaptive approach for time series forecasting, whose ideas are similar to the MPCA. In this approach, the first stage involves the decomposition of the initial time series into basis functions by the method of empirical mode decomposition and application to them of the Hilbert transform. In the second step the obtained functions and their instantaneous amplitudes are used as input variables of neural network forecasting.

Since in the proposed decomposition methods accurate convergence of the sum of all components of the process to the original process has not been proved mathematically rigorous, we introduce weights β_T, β_A, β_D and β_R for each of the components of the expansion, respectively, then the resulting prediction is $y_t(l) = \beta_T \cdot y_t^T(l) + \beta_A \cdot y_t^A(l) + \beta_D \cdot y_t^D(l) + \beta_R \cdot y_t^R(l)$. However, for the extraction of qualitative dependencies between predicted time series and components of the expansion of predicted and exogenous time series is necessary to build the next Box-Jenkins model:

$$y_t(l) = \frac{\omega^{y^T}_{c^{y^T}}(B)}{\delta^{y^T}_{r^{y^T}}(B)} \cdot y^T_{t-b_{y^T}}(l) + \frac{\omega^{y^A}_{c^{y^A}}(B)}{\delta^{y^A}_{r^{y^A}}(B)} \cdot y^A_{t-b_{y^A}}(l) + \frac{\omega^{y^D}_{c^{y^D}}(B)}{\delta^{y^D}_{r^{y^D}}(B)} \cdot y^D_{t-b_{y^D}}(l) +$$

$$+ \frac{\omega^{y^R}_{c^{y^R}}(B)}{\delta^{y^R}_{r^{y^R}}(B)} \cdot y^R_{t-b_{y^R}}(l) + \sum_{j=1}^{N} \left(\frac{\omega^{x^{jT}}_{c^{x^{jT}}}(B)}{\delta^{x^{jT}}_{r^{x^{jT}}}(B)} \cdot x^{jT}_{t-b_{x^{jT}}}(l) + \frac{\omega^{x^{jA}}_{c^{x^{jA}}}(B)}{\delta^{x^{jA}}_{r^{x^{jA}}}(B)} \cdot x^{jA}_{t-b_{x^{jA}}}(l) + \frac{\omega^{x^{jD}}_{c^{x^{jD}}}(B)}{\delta^{x^{jD}}_{r^{x^{jD}}}(B)} \cdot x^{jD}_{t-b_{x^{jD}}}(l) + \right.$$

$$\left. + \frac{\omega^{x^{jR}}_{c^{x^{jR}}}(B)}{\delta^{x^{jR}}_{r^{x^{jR}}}(B)} \cdot x^{jR}_{t-b_{x^{jR}}}(l) \right) + \frac{\theta^*_{q^*}(B)}{\varphi^+_{p^+}(B)} \cdot e_t \qquad (13)$$

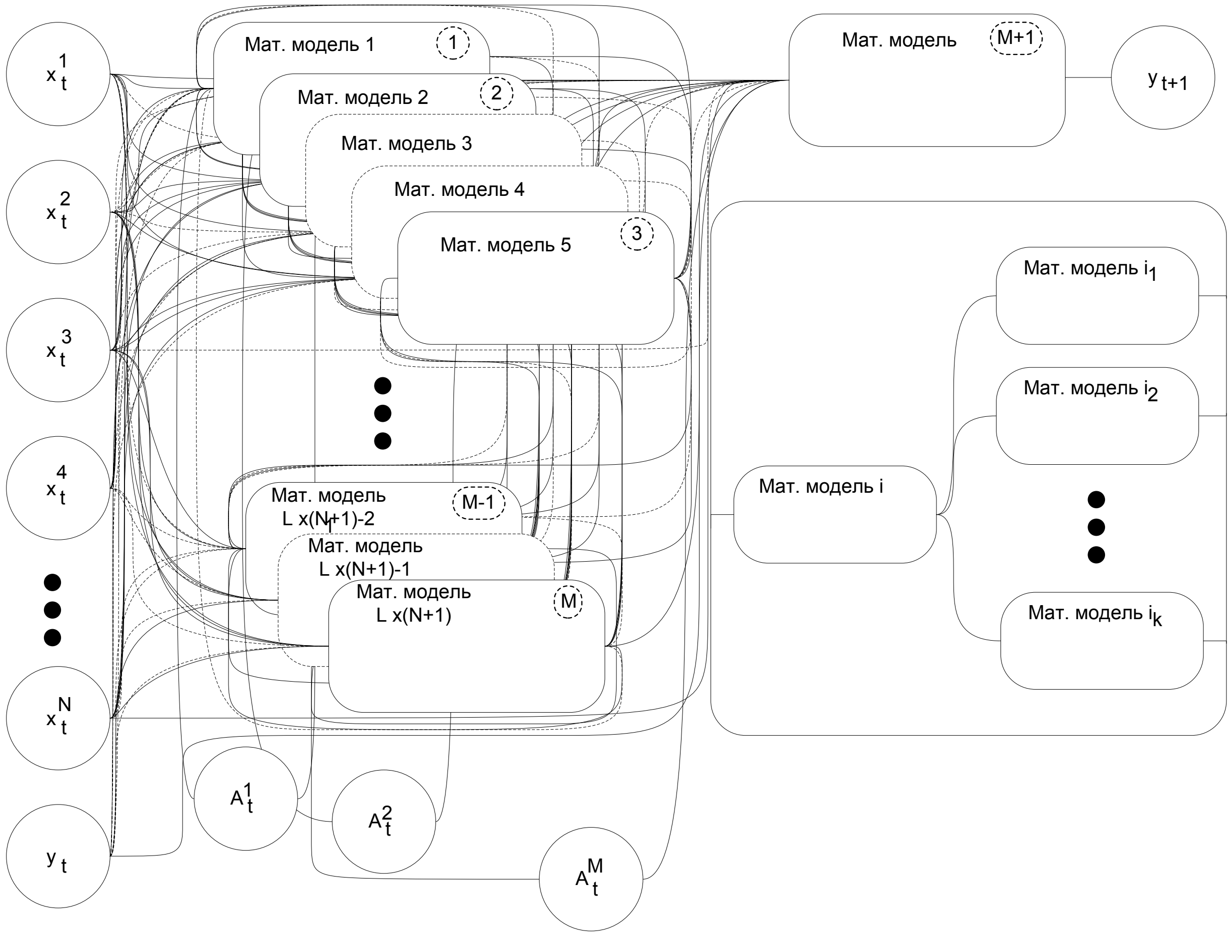

Figure 1. The first block diagram of the model of decomposition/combined forecasting method

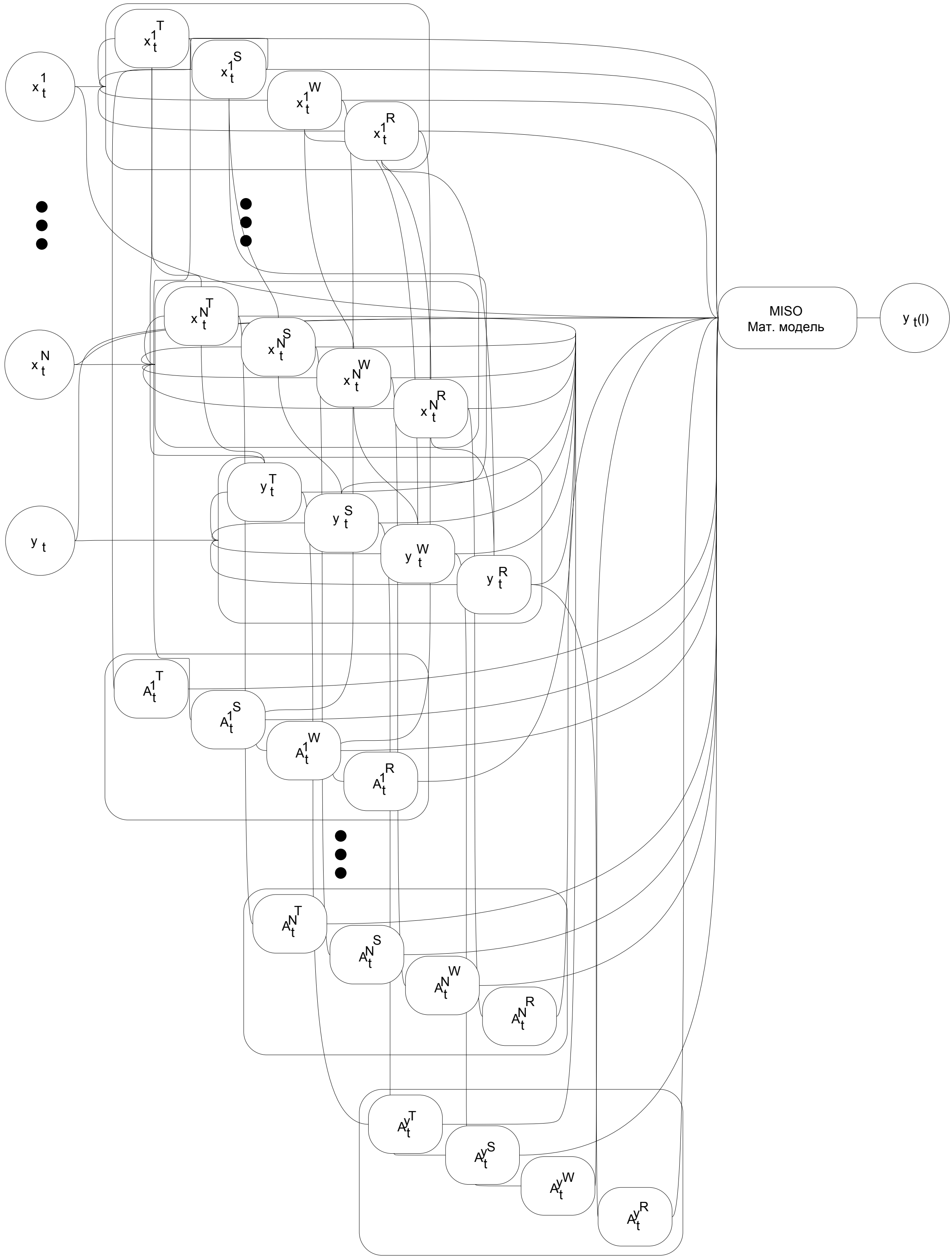

Figure 2. The second block diagram of the model of combined forecasting method

Results

In [Shchelkalin, 2011] the trend approach based on "Caterpillar"-SSA method for predicting natural gas consumption taking into account the changes in air temperature was used.

Testing of described models (12), (13) was conducted on real data of daily average electricity consumption and air temperature changes over a three year period of time.

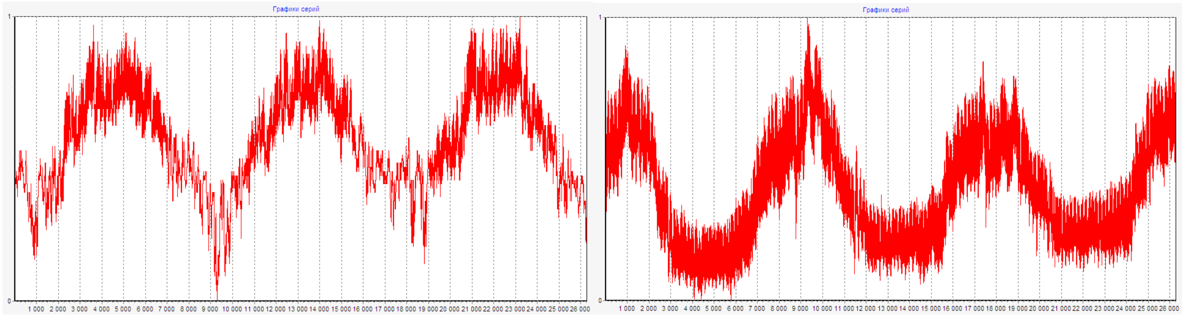

Figure 3. Chart of normalized hourly data of air temperature changes (left chart) and electricity consumption changes (right chart)

The "Caterpillar"-SSA method is made time series decomposition on trend-seasonal, weekly and remaining components.

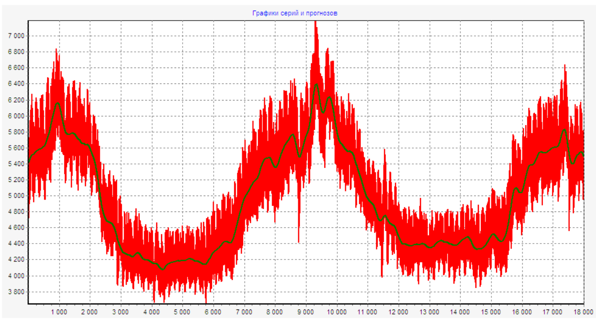

Figure 4. Trend-seasonal component time series of electricity consumption changes

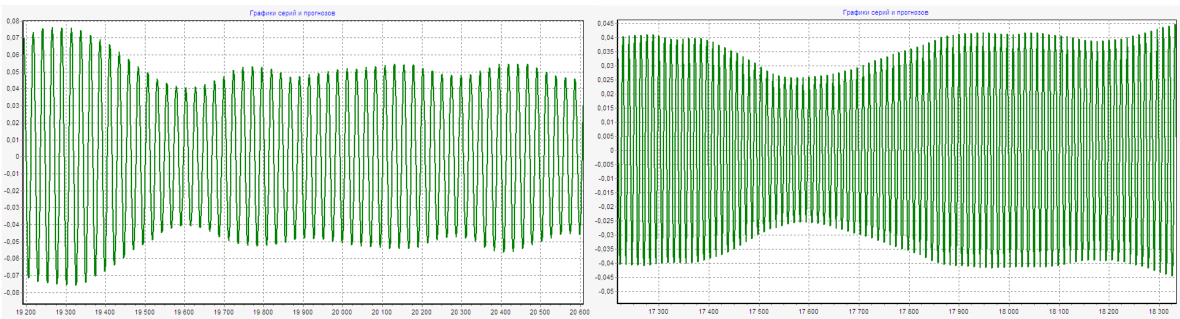

Figure 5. Monthly and weakly components of time series of electricity consumption changes

By analyzing of time series of the variances of residual errors of the model, we found that it is necessary to calculate the GARCH model of the analyzed residual time series to account a heteroscedasticity and to adequate computing of forecasts confidence intervals.

To calculate the qualitative forecasts of the electricity consumption processes it is necessary to apply cluster analysis to form training samples with similar daily images (figure 7).

Considering the charts in figure 8 we conclude that it is need to multiplication of time series weekend elements on the correction coefficients.

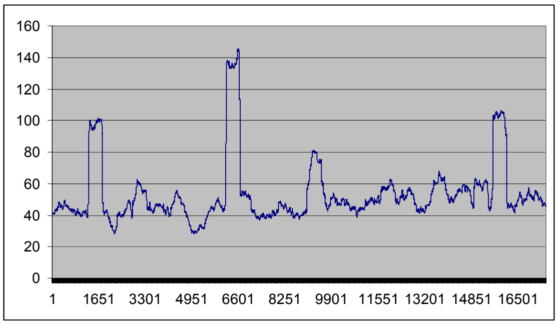

Figure 6. Time series of the variances of residual errors of the model

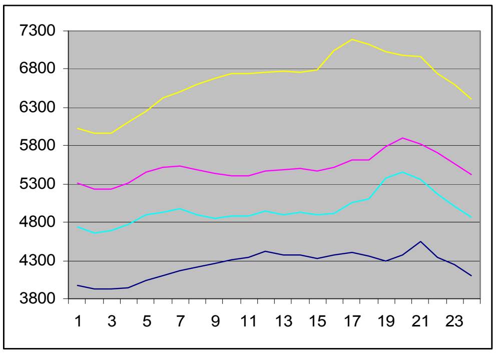

Figure 7. Charts of the daily electricity consumption at different times of the year

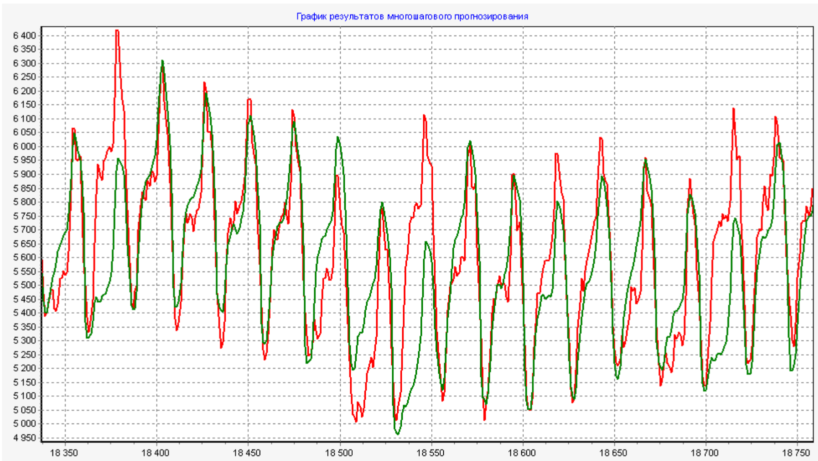

Figure 8. The original time series and time series obtained in the group stage of the "Caterpillar»-SSA method

The charts of different components forecasts of the energy consumption process are presents below.

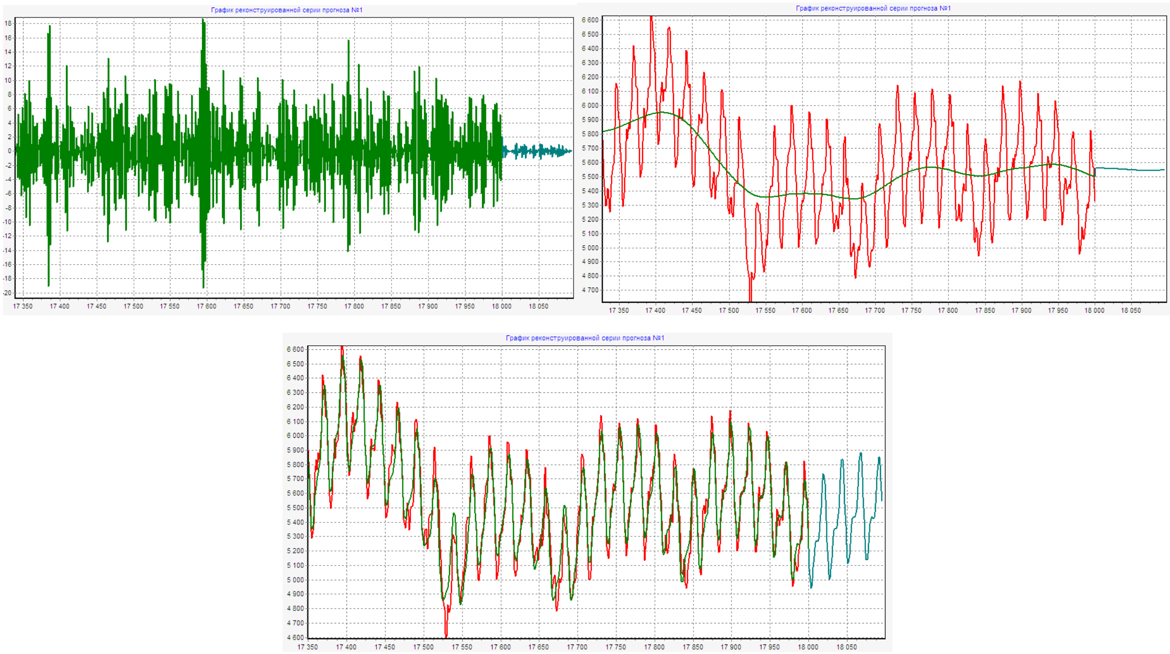

Figure 9. Charts of different components forecasts of the energy consumption process

On figure 9 the general forecast of the energy consumption process is presented.

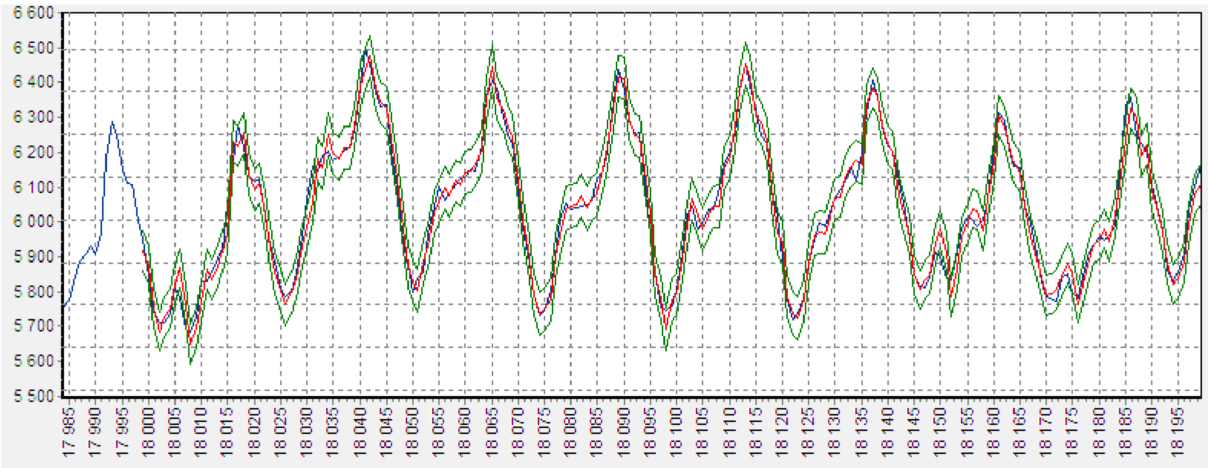

Figure 10. Charts of electricity consumption forecasts with confidence intervals.

Proposed decomposition forecasting method has reduced the mean absolute percentage error (MAPE) of presented time series of electricity consumption from 1.96% to 1.58% respectively in comparison to the prediction by SARIMAX models.

Conclusion

Thereby, to obtain adequate models of the complex processes, high-quality forecasts it is necessary to combine the models with miscellaneous structures, including nonlinear models, which are complementary in their competitive learning.

The main advantage of the proposed method of constructing an adequate model of the process under study is its rigorous formalization and, consequently, the ability to fully automate all phases of construction and use of the model.

The proposed methods in the construction of models are some intermediate approach between the classical and modern regression neural network, and a more formalized in structure and economical choice for time-consuming, and their models are quasi-optimal for detail.

Bibliography

[Бокс, Дженкинс, 1974] Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. Пер. с англ. – М. : Мир, 1974. Вып. II. – 197 с.

[Granger, Joyeux, 1980] Granger C.W.J., Joyeux R. An Introduction to Long-Memory Time Series Models and Fractional Differencing // Journal of Time Series Analysis. 1980. N 1(1). P. 15-29.

[Седов, 2010] Седов А.В. Моделирование объектов с дискретно-распределёнными параметрами: декомпозииoнный подход / А.В. Седов; Южный научный центр РАН. – М. : Наука, 2010. – 438 с.

[Бэнн, Фармер, 1987] Бэнн Д.В., Фармер Е.Д. Сравнительные модели прогнозирования электрической нагрузки: Пер. с англ. – М.: Энергоатомиздат, 1987. – 200 с.

[Shchelkalin, 2011] Vitalii Shchelkalin. ARSIMA Model // Business and Engineering Applications of Intelligent and Information Systems. Rzeszow – Sofia. 2011. Pp. 57 – 69.

[Щелкалин, Тевяшев, 2011] Щелкалин В.Н., Тевяшев А.Д. Модель авторегрессии – спектрально проинтегрированного скользящего среднего со спектрально проинтегрированной обобщенной авторегрессионной условной гетероскедастичностью для моделирования, фильтрации, прогнозирования и управления процессами в современных системах автоматизации. // Труды Международной научно-практической конференции «Передовые информационные технологии, средства и системы автоматизации и их внедрение на российских предприятиях» AITA-2011. Москва, 4 – 8 апреля 2011 г. М.: Институт проблем управления им. В.А. Трапезникова РАН, 2011. с. 996 – 1022.

[Евдокимов, Тевяшев, 1980] Евдокимов А. Г., Тевяшев_А._Д. Оперативное управление потокораспределением в инженерных сетях. – Х. : Вища школа, 1980. – 144 с.

[Голяндина, 2004] Голяндина Н. Э. Метод «Гусеница»-SSA: прогноз временных рядов: Учеб. пособие. – СПб., 2004. – 52 с.

[Крисилов, Побережник, 2003] Крисилов В. А., Побережник С. М. Аппроксимация сложных зависимостей структурно-гибкими полиномиальными и гармоническими рядами // Управляющие Системы и Машины, №2, 2003. С. 1 – 11.

[Ahmed, 1994] Ahid Ahmed. Fast Orthogonal Search For Training Radial Basis Function Neural Networks // Abstract of a Thesis Presented in Partial FulFillment of the Requirements for the Degree of Master of Science (in Electrican Engineering). August, 1994. – 69 p.

[Перцовский, 2003] Перцовский О.Е. Моделирование валютных рынков на основе процессов с длинной памятью: Препринт WP2/2004/03 – М.: ГУ ВШЭ, 2003. – 52 с.

Bollerslev T. Generalized autoregressive conditional heteroscedasticity // Journal of econometrics. – 1986. – V. 31. – PP. 307 – 327.

[Щелкалин, 2011] Щелкалин В. Н. От идей методов «Гусеница»-SSA и Бокса-Дженкинса до декомпозиционного метода прогнозирования и декомпозиционной ИНС // Восточно–Европейский журнал передовых технологий. № 4/4 (52) – Харьков: Технологический центр, 2011. С. 59–69.

[Голяндина, Шлемов, 2012] Голяндина Н. Э., Шлемов А. Ю. Повышение точности SSA-прогноза зашумленного сигнала за счет увеличения порядка линейной рекуррентной формулы // Материалы IX Международной конференции «Идентификация систем и задачи управления», SICPRO ’12, Москва, 30 января - 2 февраля 2012 г. С 1035 – 1048.

[Щелкалин, 2012] Щелкалин В. Н. Математические модели и методы, основанные на совместном использовании идей методов «Гусеница»-SSA и Бокса-Дженкинса // Материалы IX Международной конференции «Идентификация систем и задачи управления», SICPRO ’12, Москва, 30 января - 2 февраля 2012 г. С 1035 – 1048. там же. С. 728 – 773.

[Смоленцев, 2008] Смоленцев Н. К. Основы теории вейвлетов. Вейвлеты в MATLAB. – М.: ДМК Пресс, 2008. 448 с.

[Щелкалин, 2011] Щелкалин В. Н. От идей методов «Гусеница»-SSA и Бокса-Дженкинса до декомпозиционного метода прогнозирования и декомпозиционной ИНС // Восточно-Европейский журнал передовых технологий. – 2011. - № 4/4 (52). – С. 59 – 69.

[Щелкалин, Тевяшев, 2011] Щелкалин В. Н., Тевяшев А. Д. Трендовый и декомпозиционный подходы прогнозирования процессов потребления электроэнергии // Восточно-Европейский журнал передовых технологий. – 2011. - № 5/4 (53). – С. 30 – 37.

[Щелкалин, Тевяшев, 2011] Щелкалин В. Н., Тевяшев А. Д. Цифровое моделирование радиосигналов комбинированными нелинейными моделями, основанными на моделях метода «Гусеница»-SSA и сезонной АРССЭ // Труды 13-й Международной конференции «Цифровая обработка сигналов и ее применение – DSPA-2011», 30 марта – 2 апреля 2011 г. Москва – С. 165 – 168.

[Щелкалин, Тевяшев, 2011] Щелкалин В.Н., Тевяшев А.Д. Модель VARSIMAX. Синтез финитного апериодического регулятора // Сборник статей одиннадцатой Международной научно-практической конференции “Фундаментальные и прикладные исследования, разработка и применение высоких технологий в промышленности”, 27 – 29 апреля 2011 г. Санкт-Петербург, Россия / под ред. А.П. Кудинова. – СПб. : Изд-во Политехн. ун-та, 2011. – С. 457 – 464.

[Щелкалин, 2011] Щелкалин В. Н. Метод «Гусеница»-SSA – АРПСС – СПОАРУГ и модель АРСПСС – СПОАРУГ для анализа и прогнозирования финансово-экономических временных рядов // Науковий вісник Чернівецького національного університету: Збірник наукових праць. Вип. 557 – 558. Економіка. – Чернівці: Чернівецький нац. ун-т., 2011 г. – С. 174 – 186.

[Щелкалин, 2011] Щелкалин В.Н. Метод «Гусеница»-SSA – АРПСС – СПОАРУГ и модель АРСПСС – СПОАРУГ для анализа и прогнозирования финансово-экономических временных рядов [Текст] : сборник трудов второй Международной научно-методической конференции “Математические методы, модели и информационные технологии в экономике”, 4 – 6 мая 2011 г. Черновцы. – С. 306 – 308.

[Тевяшев, Щелкалин, 2009] А. Д. Тевяшев, В. Н. Щелкалин. Сравнительный анализ методов прогнозирования процессов потребления природного газа // Восточно-европейский журнал передовых технологий. – 2009. №42. - С. 8 – 16.

[Щелкалин, Тевяшев, 2010] Щелкалин В. Н., Тевяшев А. Д. Системный подход к автоматическому построению комбинированных математических моделей, основанных на совместном использовании моделей сезонной АРССЭ и метода «Гусеница»–SSA с нелинейным усложнением передаточной функции, для моделирования и прогнозирования широкого класса процессов, описанных временными рядами.

Сезонная модель нелинейной авторегрессии – спектрально проинтегрированного скользящего среднего (сезонная нелинейная АРСПССЭ модель, seasonal NARSIMAX model) // Сборник трудов X Международной научно–практической конференции «Исследование, разработка и применение высоких технологий в промышленности». 9–11 декабря 2010 г. Т.IV. СПб., 2010. С. 361–369.

[Щелкалин, Тевяшев, 2010] Щелкалин В. Н., Тевяшев А. Д. Методика построения комбинированных математических моделей для описания и прогнозирования широкого класса физиологических и психофизиологических процессов // Сборник трудов I Международной научно–практической конференции «Высокие технологии, фундаментальные и прикладные исследования в физиологии и медицине». 23-26 ноября 2010 г. Т. III. СПб., 2010. С. 165-173.

[Щелкалин, Тевяшев, 2011] Щелкалин В.Н., Тевяшев А.Д. Модель VARSIMAX. Синтез финитного апериодического регулятора // сборник статей одиннадцатой Международной научно-практической конференции "Фундаментальные и прикладные исследования, разработка и применение высоких технологий в промышленности", 27 – 29 апреля 2011 г. Санкт-Петербург, Россия / под ред. А.П. Кудинова. – СПб. : Изд-во Политехн. ун-та, 2011. – С. 457 – 464.

[Kurbatskii, Sidorov, Spiryaev, Tomin, 2011] V. G. Kurbatskii, D. N. Sidorov, V. A. Spiryaev, and N. V. Tomin. On the Neural Network Approach and Forecasting of Nonstationary Time Series on the Basis of the Hilbert-Huang Transform // Automation and Remote Control, 2011, Vol. 72, No. 7, pp. 1405 – 1414. Original Russian Text published in Automatika i Telemekhanika, 2011, No. 7. pp. – 58 – 68.

Authors' Information

Vitalii Shchelkalin *– graduate student, Department of Applied Mathematics, Kharkiv national university of radioelectronics, Lenina str., 14, Kharkiv, Ukraine; e-mail: vitalii.shchelkalin@gmail.com*

Major Fields of Scientific Research: Mathematical modeling, Prediction, Control theory, Data analysis, Neural networks, Data mining