Evolutionary methods in modelling behaviour of complex system

V V Mokshin¹ and L M Sharnin¹

¹Kazan National Research Technical University named after A. N. Tupolev – KAI, Karl Marx str.,10, Kazan, Russia, 420111

e-mail: vladimir.mokshin@mail.ru, lmsharnin@kai.ru

Abstract. In this paper, a recursive-regression approach is formulated in the formation of a functioning model of a complex system. The complex system consist of the object of study represented by a set of related factors. Some of these factors can be specified from the outside, whereas other factors contain information generated by system. We demonstrate that time-dependence of these factors can be reproduced by the nonlinear regression model. The proposed sub-course combines the method of group accounting of arguments with standard regression analysis adapted to fast dynamic processes. As an example, recursive-regression modelling of the functioning of an industrial enterprise was performed. This allows us to predict a possible behaviour of the system and to identify the so-called significant factors that have a significant impact on the behaviour of the system. To demonstrate validity of the method, we apply it to analyse the data characterizing a manufacturing company. Also paper includes meteorological data analysis. An important advantage of the proposed approach is that models for effective factors are formed automatically at each time step. The given results make possible to determine the optimal parameters of introduced algorithm and number of time observations for analysis.

1. Introduction

There are various approaches aimed at the formation of models of analysis and control of complex systems whose behavior is caused by various external factors. As complex systems can be considered as biological and economic systems, as well as entire manufacturing enterprises, as well as sociological objects (city, district, region, etc.) [1, 2]. Management of complex systems involves the analysis of the corresponding dynamic processes and the selection of the most optimal solutions aimed at achieving a specific end result. A detailed analysis of continuously available information about the “object under study” as a dynamic system [3, 4] creates the basis of operations for the development of new management decisions [5]. One of the main tasks is to find a methodology for the most correct prediction of the particular characteristics (parameters) of the system. For the analysis of complex dynamic systems, methods of linear and nonlinear programming, the method of neural networks, fractal and multifractal analysis, methods of nonlinear dynamics and chaos theory, spectral and correlation analyzes, concepts of self-organizing systems, balance methods, etc. are widely used today [6 - 11]. These methods are successfully used in various tasks related to optimization, prediction and recognition in physics, astrophysics, biology and physiology, geology, meteorology, economics, etc.
The most qualitative and reliable results are obtained with the combined use of several methods, when the possibilities of various methods are adjusted and mutually reinforced. In the present work, a recursive approach based on a combination of a multifactorial nonlinear regression model and the method of group accounting of arguments (MGAA) is proposed for analyzing and controlling the object of study [12, 13, 14]. In this case, the method of group accounting of arguments allows one to find the original solution of the problem associated with the choice of a certain set of model functions necessary for constructing a regression model. So, in accordance with the MGAA, on the basis of the criterion of reliability calculated from additional samples or using the sliding control method [15], it is possible to select this finite set of functions.

The recursive-regression approach proposed in this paper is a combination of standard regression analysis with the method of group accounting of documents adapted to dynamic, rapidly changing processes. The present approach is an iterative procedure in which, based on some “basic” model, as information is updated and statistics accumulate, it is automatically corrected: independent variables are added and / or removed. As a result, the regression model is dynamically updated [16] and at the same time its recursive self-organization occurs. The combination of the aforementioned methods makes it possible to adapt the well-known regression analysis to the study of complex dynamic systems and significantly improve the quality of forecasting. This significantly reduces the time spent on building the regression model [8].

2. Generalized analysis model of the object of study

2.1. Generalized model

As a system of research, let us consider some object and the final combination of attributes (variables) associated with it (see Fig. 1). It is assumed that this set of features can be divided into those that directly affect the object – independent factors (we denote them by \( x_i, i = 1, M \)), and those determined by the object itself – dependent factors (we denote them by \( y_j, i = 1, K \)). A set of independent factors \( x_i \) is also called input factors, and dependent factors \( y_j \) are called responses, effective factors, or indicators of the functioning of the system (see Fig. 1).

\[
\begin{align*}
 & x_1, x_2, \ldots, x_M \\
 \rightarrow & y_1, y_2, \ldots, y_K \\
 & \text{Object of study}
\end{align*}
\]

**Figure 1.** The object of study, which is influenced by factors \( x_i \), and which itself directly determines the factors \( y_j \).

The behavior of such a system, characterized by a combination of features, can be described by a model based on a multifactorial nonlinear regression analysis [17]. In accordance with the regression analysis, after determining all dependent and independent factors, it is necessary to form a model of system functioning. It is assumed that each of the dependent factors \( y_j \) can be represented as a function of the input factors \( x_i \):

\[
y_j = f_j(x_1, x_2, \ldots, x_M), \quad j = 1, K,
\]

where \( K \) is the number of dependent (effective) factors of the functioning of the system, \( M \) is the number of factors influencing the functioning of the system, \( (M + K) \) is the total number of factors. After determining the functioning model (1), it is possible to formulate the optimization problems of the research system functioning.

To construct the model (1), one can use the method of group accounting of arguments — the method of generating and choosing regression models of optimal complexity [13, 18]. For this purpose, an inductive approach is used, according to which models of increasing complexity are
sequentially generated until the minimum of a certain quality criterion of the model is found \[1,3\]. This means that the subsequent complication of the model does not lead to an improvement in its quality. As a result, achieving the minimum quality criterion will match the desired model.

One of the authors of this method, A.G. Ivakhnenko, writes \[5,13,16\], that at the same time “there is a purposeful search of many models-applicants of varying complexity according to a number of criteria. The result is a model of optimal structure in the form of a single equation or system of equations. The minimum of the criterion determines the mode of optimal structure”.

2.2. The method of group accounting of arguments

The method of group accounting of arguments consists in finding the functional relationship between the input factors \(x_i\), and the effective indicators of the functioning of the system \(y_j\), according to their temporal behavior. In the most general form, such a dependence can be represented as a generalized Kolmogorov-Gabor polynomial \[13,16\]:

\[
y = a_0 + \sum_{i=1}^{M} a_{ix_i} + \sum_{j=1}^{M} \sum_{i<j} a_{ijx_i}x_j + \sum_{i=1}^{M} \sum_{j=1}^{M} \sum_{k<j} a_{ijkx_ix_jx_k} + \ldots
\]

where the weights \(a\) are found using regression analysis and the least squares method for the training set \[7\]. Polynomial (3) actually contains all possible combinations of input factors \(x_i, i = 1, M\). Note that to improve the quality of models (1), it is also possible to use not only the power dependences of the input factors, but also the elementary functions of these factors.

As an example, consider the following situation. Suppose that there are two independent input factors \(x_1\) and \(x_2\) and one effective factor \(y\). In accordance with the method of group accounting of arguments, various possible combinations of limited complexity \(S\) arise, i.e. \(y^{(1)}, y^{(2)}, y^{(3)}, \ldots\). The complexity \(S\) here is the number of members of the polynomial (2) being formed \[16\]. So,

\[
y^{(1)} = a_{10} + a_{11}x_1,
y^{(2)} = a_{20} + a_{21}x_2,
y^{(3)} = a_{30} + a_{31}x_1 + a_{32}x_2,
y^{(4)} = a_{40} + a_{41}x_1 + a_{42}x_2 + a_{43}x_1x_2,
y^{(5)} = a_{50} + a_{51}x_1 + a_{52}x_2 + a_{53}x_1x_2 + a_{54}x_1^2x_2,
y^{(6)} = a_{60} + a_{61}x_1 + a_{62}x_2 + a_{63}x_1x_2 + a_{64}x_1^2x_2^2,
y^{(7)} = a_{70} + a_{71}x_1 + a_{72}x_2 + a_{73}x_1x_2 + a_{74}x_1^2x_2^2.
\]

The best model is selected based on the minimum value of the quality criterion \[15\]. Or, the threshold value of the quality criterion is determined and several models are selected whose quality criterion values do not exceed this threshold.

2.3. Quality assessment and model selection

To assess the quality of models (for example, for models of the form (3)) one can use the variance \[10,11,13\]:

\[
\bar{\varepsilon}^2 = \frac{1}{n} \sum_{i=1}^{n} \left( y_i - f(x_i) \right)^2,
\]

where \(n\) is the number of considered time intervals of input factors \(x_i, i = 1, M\) and effective indicators of the functioning of the system \(y_j, i = 1, K\).

Then only those values of the weighting factors \(a\) of the model polynomial (3) will be acceptable, for which the value \(\bar{\varepsilon}^2\) takes the minimum values, i.e.

\[
\bar{\varepsilon}^2 \rightarrow \min
\]
In this case, the principle of multiplicity of models can be realized \cite{15, 16}, which consists in the following: there are many models on this sample with zero error. To do this, it is enough to increase the degree of the polynomial model. In other words, you can build a whole family of models, each of which, when passing through experimental points \( y_j \), where \( j = \text{const} \), will give a zero error, and, therefore,

\[
\overline{e^2} = 0
\]  

(6)

Obviously, the variance \( \overline{e^2} \) depends on the complexity of \( S \). Moreover, with increasing complexity, the variance \( \overline{e^2} \) will first fall, and then grow. Then you need to choose the optimal degree with a minimum value. In addition, the object of study can be affected by various interferences, which in the present work will mean unaccounted input factors affecting the system, or insufficient complexity of the desired polynomial. Considering the effect of interference, the following situations can be distinguished:

1. At different levels \( \theta \) of interference, the \( \overline{e^2} \) dependence on complexity \( S \) will change, while maintaining the overall directivity (meaning that with increasing complexity, \( \overline{e^2} \) will first decrease and then increase).

2. As the level of interference \( \theta \) increases, the value \( \min_{S} \overline{e^2} \) will increase.

As the level of interference increases, \( S^* = \arg \min_{S} \overline{e^2} \) will decrease, i.e. the optimal difficulty value will shift to the left (see fig. 2). Moreover, \( \overline{e^2}(S^*) > 0 \) if the noise level \( \theta \) is nonzero.

Figure 2. The dependence \( \overline{e^2} \) of the variance on the complexity of the model \( S \) at different levels of interference \( \theta \). The minima on the curves correspond to models with optimal complexity \( S^* \).

3. Evolutionary approach of formation models for research object functioning

The above-stated basic ideas of the method of the basic accounting of arguments allow determining the interrelations between input and effective factors, which are expressed with the help of functional dependencies of the form (3). Based on this, you can perform (i) analysis of the current state of the object, (ii) forecasting the values of the result factors, (iii) influence the resultant factors through the input factors, thereby controlling the object. However, all this is possible only in the case when the entire system remains unchanged over time. This means the preservation of the number of input factors, their independence from each other and the absence of new external factors. However, this situation is, in fact, idealized. Therefore, there is the task of creating such a model of the functioning of the system, which could adapt to changing external factors. In our case, this means that for each subsequent time step, a new model is created based on the previous model (see Fig. 3).

The area indicated by the dotted line determines the formation of the functioning model of the object of study at a certain point in time. The whole scheme reflects the possibility of changing the already created model in time, i.e. it becomes possible to add new input factors and re-form the model (1). In fact, this is the formulation of the basic ideas of recursive approaches \cite{17, 18, 19}. Such
considerations will be used by us in the dynamic formation of the model of the functioning of the object.

Figure 3. Schematic representation of the formation of the structure of a multidimensional additive series, taking into account changes in the system over time.

4. Application of the recursive and regression approach to evaluating the efficiency of managing a manufacturing enterprise

Consider the results of the above approach, in relation to the problem of improving the management efficiency of some industrial enterprises, cooperating with the enterprises of the automotive and petroleum industries. In order to form a model for the functioning of a production enterprise, we will select the values given in Table 1 as input external factors and effective factors $y_j$. In the study, quarterly values of selected factors were used, and a particular task was to increase the revenue from the sale of goods, products, works, and services. As a result, a functioning model was obtained, including a system of ten regression equations. This number of equations corresponds to the number of selected effective factors $y_j$.

The most significant obtained regression equations are given below:

$$
y_1 = -665 \cdot 10^5 \cdot 0.65 x_1 + 6.35 \cdot 10^{-2} x_2 x_3 + 3.59 \cdot 10^{-4} x_4 x_5 + 3.36 x_6^{0.5} x_7^{0.5} - 2.15 \cdot 10^{-7} x_8^{-1} + 5.78 x_9 - 4.56 \cdot 10^2 x_9^{0.5} + 1.65 \cdot \ln(x_10 x_12) + 84.04 x_{14} + 22.25 x_{15}^{0.5} x_{16}^{0.5} + 45.29 \cdot x_{14} + 4.99 x_{16} - 4.62 \cdot 10^{-6} e^{x_{17}},$$  

$$
y_2 = 5.95 \cdot 1.1 \cdot 10^{-4} e^{x_{1}} - 8.18 x_2 + 0.6 x_3 - 3.73 x_4 x_5 + 2.03 \cdot 10^6 e^{x_{6}} - 8.09 \cdot 10^{-4} x_6 x_7 - 4.59 x_8^{0.5} x_9^{0.5} - 1.86 \cdot 10^4 x_9 + 9.53 \cdot 10^4 \ln x_{10} + 2.53 e^{x_{11}} - 9.34 \cdot 10^{-4} x_{12} x_{14} + 7.41 \cdot 10^3 x_{13} x_{14} + 85.44 x_{16}^{0.5},$$  

$$
y_3 = 3.96 \cdot 10^{-2} - 0.45 x_1 + 0.16 x_2 + 6.1 \cdot 10^6 x_2^2 - 9.94 \cdot 10^6 x_2^2 + 3.14 \ln(x_3) - 5.56 \cdot 10^{-5} x_5^2 + 0.67 x_7 + 1.99 x_8 - 3.65 x_9 + 0.18 x_{10} + 23.2 x_{14} + 5.63 x_{15} - 50.93 x_{13} - 78.89 x_{14} - 2.77 \cdot 10^7 \ln x_{15} + 11.99 x_{16} - 94.33 x_{17}. $$

The whole set of obtained regression equations forms a recursive-regression model, in which the input factors $x_i, i = \bar{1}, M$, have a controlling influence on the resulting factors $y_j$.

The proposed model allows us to determine how much a change in an input factor $x_i, i = \bar{1}, M$ (or a combination of them) can affect the resulting factor $y_j, i = \bar{1}, K$. To do this, we estimated the degree of influence of factors by their weight and elasticity coefficients. The specific weights of input factors determine, as a percentage, their influence on the resulting factors (see Fig. 4-6). The coefficient of elasticity determines the degree of change of the $j$-th effective coefficient when the value of the $i$-th input coefficient changes by one percent. Positive values of elasticity coefficients indicate the positive influence of the corresponding external factor, and negative values – vice versa (see Fig. 7). Note that the elastic coefficients were calculated under the conditions of the last considered time interval. When
calculating them, all factors were assigned values that they took into account in the last time interval, and the coefficient that was taken into account changed over the entire time interval.

Table 1. The list of input and effective factors selected for the study.

| Factor | Factor name |
|--------|-------------|
| $y_1$  | Proceeds from the sale of goods, products, works, services, thousand rubles |
| $y_2$  | Cost of goods sold, products, works, services, thousand rubles |
| $y_3$  | Profit from sales, thousand rubles |
| $y_4$  | Profit before tax, thousand rubles |
| $y_5$  | Current income tax, thousand rubles |
| $y_6$  | Interest receivable, RUB thousand |
| $y_7$  | Other operating income, thousand rubles |
| $y_8$  | Non-sale income, thousand rubles |
| $y_9$  | Income per employee, thousand rubles |
| $y_{10}$ | Profit per employee, thousand rubles |
| $x_1$  | Fixed assets, thousand rubles |
| $x_2$  | Inventories, thousand rubles |
| $x_3$  | Costs in work in progress (distribution costs), thousand rubles |
| $x_4$  | Finished products and goods for resale, thousand tons |
| $x_5$  | Accounts receivable (payments for which are expected within 12 months after the reporting date), thous. Rub. |
| $x_6$  | Bank account, thousand rubles |
| $x_7$  | Currency account, thousand rubles |
| $x_8$  | Retained earnings of the reporting year, thousand rubles |
| $x_9$  | Borrowed funds, thousand rubles |
| $x_{10}$ | Accounts payable, thousand rubles |
| $x_{11}$ | Selling expenses, thousand rubles |
| $x_{12}$ | Administrative expenses, thousand rubles |
| $x_{13}$ | Other operating expenses, RUB thousand |
| $x_{14}$ | Non-operating expenses, thousand rubles |
| $x_{15}$ | The number of employees with permanent employment contracts (contracts) in people. |
| $x_{16}$ | Payroll, thousand rubles |
| $x_{17}$ | Average salary per employee, thousand rubles |

Figure 4. The diagram of the relative influence of input factors on the revenue from the sale of goods, products, works and services, $y_1$.

Figure 5. The diagram of the relative influence of input factors on the cost of goods sold, products, works and services, $y_2$. 
From fig. 4 clearly shows that the effective factor \( y_1 \) (revenue from the sale of goods, products, works and services) is most influenced by the input factors \( x_2, x_7, x_{11}, x_{12} \) and \( x_{16} \), which for the system under study evaluate accounts receivable, currency account, selling expenses, operating expenses and payroll, respectively. From fig. 7 that for the growth of the values of \( y_i \), i.e. increase in revenue from the sale of goods, products, works and services; it is necessary to increase the fixed assets of the enterprise \( x_1 \), finished goods and goods for resale \( x_3 \), the current account \( x_4 \), retained earnings of the reporting year \( x_5 \), the average wage per employee \( x_{17} \), and reduce inventory \( x_8 \), costs in work in progress \( x_9 \), receivables \( x_{15} \), currency account \( x_{17} \), business expenses \( x_{11} \), operating expenses \( x_{13} \), non-operating expenses \( x_{14} \) and the number of employees with permanent labor contracts \( x_{15} \).

Thus, it is obvious that the resulting model can also be used to optimize all the effective factors \( y_j \) taken into account in the framework of our model. So, for example, if we consider the revenue from the sale of goods, products, works and services \( y_1 \) as an optimized resultant factor (see Fig.4 and 7), then, in accordance with our results, it is fair to say that increase by 40%.

5. Meteorological data

As a second example of application of the proposed approach we consider the time-dependent weather data for Szeged – one of the largest city of Hungary located on the Tisza River in the central part of the Carpathian Basin [17]. The available data set represents the hourly values of the nine weather parameters for the period from April 1, 2006 (midnight) to April, 28, 2016 (6 pm). We are focused on the temperature dynamics in this geographic place, therefore, it is reasonable to take the temperature as the output parameter \( y \). The other eight parameters (\( x_i \) with \( i = 1, 2, \ldots, 8 \)) such as humidity, wind speed, visibility etc. are considered as input factors. Additionally, by analogy with the previous example, we introduce again five noise factors, i.e. \( x_9, x_{10}, x_{11}, x_{12} \) and \( x_{13} \), whose time dependencies are generated from a set of random numbers (see Table 2). Thus, we are dealing with one output factor and 13 input factors, i.e. \( M = 1 \) and \( K = 13 \). We note that the factors \( x_2 \) and \( x_3 \) characterize the cloudiness and the daily summary precipitate, respectively, and take possible standard wordings.

Size of the learning regime. – In fact, the whole time range for the available weather data contains 96 454 time points, where the elementary time interval between two time points is 1 hour. Such a large number of available time points allows one to determine the suitable size of the learning sampling. To do this we shall construct regression model (2) for the samplings of the different sizes and do estimates of correspondence between values of the output factor \( y_{res} \) resulted from this regression model and the target (experimental) values of \( y \). Here, the so called coefficient of determination \( R^2 \), the Fisher \( F \)-criterion and the coefficient of variation \( CV \) can be taken as measures of this correspondence [18]. We propose to reconsider samplings of the different sizes as follows. The size of the learning sampling is increased cyclically by 4 time points, i.e. by 4 hours, starting with a sampling of the size of 24 time points. The correspondence of the results is evaluated for the subsequent time
range with 72 time points that define the size of the test regime. In Fig. 4, we present values of the coefficient of determination $R^2$ and coefficient of variation $CV$ obtained with learning samplings of the different sizes.

Table 2. Input and output factors for analysis of the meteorological data

| Factor name | Factor’s notation | Factor’s description (in specific units) |
|-------------|-------------------|-----------------------------------------|
| Temperature | $y_1$             |                                         |
| Humidity    | $x_1$             |                                         |
| Summary-cloudiness | $x_2$       |                                         |
| Daily summary precip type | $x_3$       |                                         |
| Apparent temperature | $x_4$       |                                         |
| Wind speed  | $x_5$             |                                         |
| Wind bearing| $x_6$             |                                         |
| Visibility  | $x_7$             |                                         |
| Pressure    | $x_8$             |                                         |
| 1th noise factor | $x_9$       | 1th noise factor                        |
| 2th noise factor | $x_{10}$     | 2th noise factor                        |
| 3th noise factor | $x_{11}$     | 3th noise factor                        |
| 4th noise factor | $x_{12}$     | 4th noise factor                        |
| 5th noise factor | $x_{13}$     | 5th noise factor                        |

Recall that if two series with model results and ‘experimental’ data are characterized by identical values, i.e. demonstrate the perfect correspondence to each other then one obtains $R^2 = 1$ and $CV = 0$ [18]. As seen from the presented results, with increase of the size of the learning sampling up to the size with 64 time points (in Fig. 4, the corresponding boundary marked by vertical dotted line), accuracy of the regression model to reproduce the ‘experimental’ data becomes better. When the time window size contains 64 time points, these measures $R^2$ and $CV$ approach unit and zeroth values, respectively, and further increase of the time window size does not change practically the values of both the measures $R^2$ and $CV$. This is direct evidence that the size of the learning sampling for this data set has to contain not less than 64 time points.

Figure 8. Occurrence frequency $P(x_i)$ of the input factors that form polynomial (9) for the output factor $y_3$. (a) Occurrence frequency $P(x_i)$ vs. index $i$ of the input factor; (b) Distribution $P(x_i)$ vs. reordered indices of the input factors (see the text for details).

In our analysis we shall take the size of the learning sampling equal to 96 time points that corresponds to four days, and the test regime will contain 72 time points. Thus, the data set for seven
days (96 + 72 time points) will be taken to construct a regression model, whereas the data for the subsequent time points will be taken to check predictive abilities of the model.

**Figure 9.** (a) Coefficient of determination $R^2$ and (b) coefficient of variation $CV$ evaluated for a regression model, which is constructed for the learning samplings of the different sizes. The zeroth time point corresponds to the data for April 1, 2006 (midnight). Vertical dotted line corresponds to results obtained for the learning sampling with 64 time points, i.e. 64 hours.

In Fig. 10, we present dynamics of the humidity $x_1$, the pressure $x_8$ and the temperature $y$ for the considered time range. Here, the first 168 time points (hours) correspond to the learning and test regimes of the regression model. For the time range starting from 168 hours in Fig. 10 (c), we compare the measured values of the temperature and our predictions. The regression model reproduces all the features of the temperature dynamics for the first 48 time points for the time range from 168 to 216 hours, i.e. hourly temperature values for two days.

**Figure 10.** (Color online) Panels (a) and (b): humidity and pressure as input factors $x_1$ and $x_8$, respectively, vs. time. Panel (c):
Temperature vs. time; here, solid line corresponds to ‘experimental’ data and red crosses are results obtained on the basis of regression model (10) for temperature taken as the input factor $y$. The zeroth time point is associated with the data for April 1, 2006 (midnight), and time range $t = 250$ hours corresponds to more than 10 days. The time ranges associated with learning regime (time $\in [0; 96]$ hours), test regime (time $2 (96; 168]$ hours) and predictive regime (time $> 168$ hours) are separated by vertical dotted lines.

6. Conclusions
This paper presents a recursive-regression approach to the formation of a complex system functioning model. As such, it can be the object of study (biological organism, enterprise, ecosystem, etc.), whose behavior is determined by a combination of external influencing factors. This approach allows you to differentiate the level of entry of each of the original external factors and, in general, allows you to control the entire system under study.

An important advantage of the proposed approach is that models for effective factors are formed automatically at each time step. A recursive self-organization of models occurs, and the need for their constant adjustment “by hand” disappears. Thus, the proposed approach makes it possible to automate
the process of forming the structure of an additive functional series of the form (3), while flexibly adjusting the necessary level of complexity of the series in terms of individual parameters. Moreover, here it also remains possible to go beyond the basis of the classical types of power and harmonic series by increasing their flexibility by introducing polynomial members (3) with fractional multiplicity. The integrated approach presented in the paper provides, in principle, a powerful and flexible tool for analyzing and controlling the object of research, the behavior of which is influenced by many different external factors. It should be noted that there is also the possibility of using the presented approach in conjunction with the methods and algorithms of neural networks. The proposed method of analyzing complex systems can also be used in image processing. Since the objects in the image can also be represented as a set of features [22-26]. And then carry out the analysis of the objects. A promising direction is also the use of the considered methodology for setting up neural networks and evaluating the educational process [27-29].

7. References

[1] Mantegna R N and Stanley H E 1999 An Introduction to Econophysics: Correlations and Complexity in Finance (Cambridge: Cambridge University Press)
[2] Chakraborti B K, Chakraborti A and Chatterjee A 2006 Econophysics and Sociophysics: Trends and Perspectives (Berlin: Wiley-VCH)
[3] Yulmetyev R M, Mokshin A V and Hanggi P 2005 Universal approach to overcoming nonstationarity, unsteadiness and non-Markovity of stochastic processes in complex systems Physica A 345 303-325
[4] Mokshin A V, Yulmetyev R M and Hanggi P 2005 Simple Measure of Memory for Dynamical Processes Described by a Generalized Langevin Equation Physical Review Letters 95 200601
[5] Ivakhnenko A G and Ivakhnenko G A 1996 Simplified Linear Programming Algorithm as Basic Tool for Open-Loop Control System Analysis Modeling Simulatio, SAMS 22 177-184
[6] Lemke F 1995 SelfOrganize! - software tool for modelling and prediction of complex systems SAMS 20(1-2) 17-28
[7] Vapnik V N 1984 Algorithms and programs for recovering dependencies (Science) p 816
[8] Ivakhnenko A G and Stepashko V S 1985 Simulation noise immunity (Naukova Dumka) p 216
[9] Draper N R and Smith H 1998 Applied regression analysis (New York: Wiley)
[10] Venables W N and Ripley B D 2002 Modern Applied Statistics with S (New York: Springer)
[11] Sen A and Srivastava M 2011 Regression Analysis Theory, Methods, and Applications (Berlin: Springer-Verlag)
[12] Mokshin V V, Saifudinov I R, Sharnin L M, Trusfus M V and Tutubalin P I 2018 Parallel genetic algorithm of feature selection for complex system analysis J. Phys.: Conf. Ser. 1096 012089 DOI: 10.1088/1742-6596/1096/1/012089
[13] Yakimov I, Kirpichnikov A, Mokshin V, Yakhina Z and Gainullin R 2017 The comparison of structured modeling and simulation modeling of queuing systems Communications in Computer and Information Science, CCIS 800 DOI: 10.1007/978-3-319-68069-9_21
[14] Yakimov I M, Trusfus M V, Mokshin V V and Kirpichnikov A P 2018 AnyLogic, extendsim and simulink overview comparison of structural and simulation modelling systems Proceedings of the 3rd Russian-Pacific Conference on Computer Technology and Applications, RPC DOI: 10.1109/RPC.2018.8482152
[15] Saifudinov I R, Mokshin V V, Tutubalin P I, Sharnin L M and Hohlov D G 2018 Visible structures highlighting model analysis aimed at object image detection problem CEUR Workshop Proceedings 2210 139-148
[16] Mokshin A V, Mokshin V V and Sharnin L M 2019 Adaptive genetic algorithms used to analyze behavior of complex system Communications in Nonlinear Science and Numerical Simulation 71 174-186 DOI: 10.1016/j.cnsns.2018.11.014
[17] Tutubalin P I and Mokshin V V 2017 The Evaluation of the cryptographic strength of asymmetric encryption algorithms Second Russia and Pacific Conference on Computer Technology and Applications, RPC 180-183 DOI: 10.1109/RPC.2017.8168094
[18] Lyasheva S A, Medvedev M V, Shleymovich M P and Mokshin V V 2018 The analysis of
image characteristics on the base of energy features of the wavelet transform CEUR Workshop Proceedings 2210 96-102

[19] Gabor D, Wilby W R and Woodcock R A 1961 A universal nonlinear filter, predictor and simulator which optimizes itself by a learning process Proc. Inst. Electr. Engrs. 108(40) 85-98

[20] URL: https://www.kaggle.com/будинчевит/szeged-weather

[21] Ross S M 2014 Introduction to probability and statistics for engineers and scientists (Academic Press)

[22] Goshin Ye V and Kotov A P 2017 Parallel implementation of multi-view image segmentation algorithm using the hough transform Computer Optics 41(4) 588-591 DOI: 10.18287/2412-6179-2017-41-4-588-591

[23] Fursov V A, Goshin Ye V and Kotov A P 2016 The hybrid CPU/GPU implementation of the computational procedure for digital terrain models generation from satellite images Computer Optics 40(5) 721-728 DOI: 10.18287/2412-6179-2016-40-5-721-728

[24] Gaidel A V 2016 Coordinated polynomial signs for the analysis of semi-meteoric biomedical images Computer Optics 40(2) 232-239 DOI: 10.18287/2412-6179-2016-40-2-232-239

[25] Agafonov A A and Myasnikov V V 2016 Method of determining a reliable short method in a dependent stochastic network and its application in geoinformational tasks of transport management Computer Optics 40(2) 275-283 DOI: 10.18287/2412-6179-2016-40-2-275-283

[26] Sergeev V V and Iuzkiv R R 2016 Parametric model of the autocorrelation function of cosmic hyperspectral images Computer Optics 40(3) 416-421 DOI: 10.18287/0134-2452-2016-40-3-416-421

[27] Novikova S V 2016 Structural Optimization of the Neural Network Model for the Gas Turbine Engine Monitoring Iz. VUZ. Ser. Russian Aeronautics 59(2) 263-270 DOI 10.3103/ S1068799816020185

[28] Tunakova Y, Novikova S, Shagidullin A and Shagidullin R 2017 Model Development for the Determination of PM2.5 and PM10 Fine Particles in Ground Layer of Atmospheric Air Ecology and Industry of Russia 21(12) 38-43 DOI: 10.18412/1816-0395-2017-12-38-43

[29] Snegurenko A P, Sosnovsky S A, Novikova S V, Yakchina R R, Valitova N L and Kremleva E Sh 2019 Using E-Learning Tools to Enhance Students-Mathematicians’ Competences in the Context of International Academic Mobility Programmes Integratsiya obrazovaniya = Integration of Education 23(1) 8-22 DOI: 10.15507/1991-9468.094.023.201901.008-022