

CONTROL OF THE REGIONAL SOCIAL RESOURCE DYNAMIC BASED ON THE REGRESSION-DIFFERENTIAL SIMULATION

A.V. Zatonzkiy, *z xenon@narod.ru*,
T.V. Safyanova, *yagyumunenori@mail.ru*

Perm National Research Politechnic University, Berezniki branch, Berezniki,
Perm region, Russian Federation

In this paper, we adopted a linear combination of normalized partial criteria from the publicly available statistical data as the assessment of the social resource development. We have shown the stability of their ranking to the small possible errors of the expert determination of ranks. We demonstrated impossibility of applying traditional econometric models to solve this task. We developed a model based on the second order regression differential equation allowing predict consequences of solutions on the development of the social resource in a tactical perspective. The peculiarity of this model is in use of the publicly available statistical factors, for which we determined their relative importance, coefficients of mutual influence, best approximation between values of the annual statistical series and other adaptation model parameters. We determined positive and negative forecast areas and gave recommendations for improvement of the social resource dynamics.

Keywords: social resource, regression-differential model, control, prognostication, software.

1. Introduction

Social resource (SR) is a totality of real or potential population resources determining the nature of social interactions, relationships and associations of people in socio-economic processes. Its existence and degree of development allow to carry out social and economic opportunities of any society, territory, or economic region. Forecast of SR dynamics is the basis for decision-making on development of social and economic opportunities of any society, including any region of the Russian Federation.

For example in conditions of the Perm region (Russian Federation) population of which has been slowly decreasing since 1990s, the question of provision of the region enterprises, social and public spheres with demographic, labor, educational, intellectual and cultural potential looks the live issue today. At present it acquires special urgency in several territorial-industrial complexes (TIC) of the region, where urbanization level is high, but at the same time the need in the SR upgrade exists. In particular, in Verkhnekamye (formed by Berezniki, Usolye and Solikamsk) there are several large enterprises and branch infrastructure that are currently beign developed and already face tasks with SR provision. To manage such TIC it is highly important to have the possibility of simulating and forecasting the SR dynamics. The same task is topical for enterprises and organizations, deprived of access to the protected data of the Ministry of Economic Development. They are able to use only open statistical data with its well-known disadvantages. Therefore, the objective of SR modeling, in the Perm Territory including, appears practicably significant and relevant.

Among the preliminary assessments of the solution quality on the development of socio-economic systems is mathematical simulation of their consequences with application of analytical models or special software. Having obvious disadvantages, these approaches however allow to estimate numerous management solutions for development of socio-economic system, and select the optimal or best one.

2. Review of the task and used econometric models

The most common regression economic and mathematical models, also used for the forecast purposes, are dynamics models of the following type:

$$y(\vec{x}(t), \vec{z}(t), t) = a_0 + \sum_i a_i x_i(t) + \sum_j b_j z_j(t),$$

where $\vec{x}(t) = \{x_1(t), x_2(t), \dots\}$ – the factor vector, $\vec{z}(t) = \{z_1(t), z_2(t), \dots\}$ – the disturbance vector, $y(\cdot)$ – response of the studied object. As an option, the following models are applied:

$$y(\vec{x}(t), \vec{z}(t), t) = a_0 + \prod_i a_i x_i(t) + \prod_j b_j z_j(t)$$

or, for the functions of one variable:

$$y(x(t), z(t), t) = a_0 + \sum_i a_i x(t)^i + \sum_j b_j z(t)^j$$

or the time series model in the form of $y(t) = \sum_{i=0}^L a_i t^i$. Such models are also used as a basis for PhD theses: [5, 11, etc.]. These models may be simply “read”, for example, as follows: if you invest in the company (industry) according to chart $x_1(t)$, at the output you will obtain the net present value (or other indicator of the economic efficiency) $y(x_1(t), z_1(t))$, taking into account demand for the products (disturbance effect) $z_1(t)$. Further we usually mind identification of a_i and b_j , the responses expressed by function $F(y)$, if more precisely:

$$y(\vec{x}(t), \vec{z}(t)) = a_0 + \sum_i a_i x_i(t) + \sum_k b_k z_k(t) - F(y(\vec{x}(t), \vec{z}(t)))$$

etc. At this, there is a tacitly accepted assumption that a direct link between the factor and the *value* of the response exists, and the only dynamic element in the model is the delay time (for example, in the following type models $y(\vec{x}(t)) = a_0 + \sum_i a_i x_i(t - \Delta t)$).

However, such assumption is not always close to reality. For example, if you manure your field according to the certain rules, you can get a crop growth (and future economic or social bonuses). That is, on the basis of several observations it is well known that the amount of fertilizers $x_1(t)$ applied in the right time t *accelerates* the crop growth in some period of fertilizer application:

$$\frac{\partial y(x, z, t)}{\partial t} \approx a_0 + a_1 x(t) \quad \forall t: 0 < x(t) \leq x_{\max}, \quad a_1 > 0,$$

and the decrease in the precipitation amount in certain circumstances, *reduces* the growth rate:

$$\frac{\partial y(x, z(t), t)}{\partial t} \approx a_0 + b_1 z(t) \quad \forall t: z_{\min} \leq z(t) \leq z_{\max}, \quad b_1 > 0, \quad \forall t: z(t) < 0.$$

For complex systems, especially those which take into account the natural processes, the regression identification of correlation between $y(t)$ and $x_i(t)$ without strong evidence of their mutual independence leads to generation of “parrot models” [9]. These models adequately interpolate the past, but are not able to forecast the future, that is, in fact, required for building the models of decision-making support.

They also include attempts to extrapolate $y(t)$ according to the time series (trends), especially when mind errors or unobservable external disturbances.

The question arises: what should be identified when building a dynamic economic and mathematical model: correlation between the factor and the *response* or correlation between the factor and the *dynamics of changes* in the response under the influence of this factor? At the level of general scientific reasoning, we can conclude that in some econometric models it is better to use differential equations as the platform, at least on the basis of an ordinary differential equation (ODE) of the 1st order with corresponding Cauchy task

$$\begin{cases} \frac{dy(t)}{dt} = a_0 + \sum_{i=1}^N a_i \cdot x_i(t) + a_{N+1} \cdot y(t) = f(y(t), x(t), t), \\ y(0) = y_0 \end{cases} \quad (1)$$

that has obvious advantages compared to algebraic equations of the system dynamics:

- 1) natural description of positive and negative changes in the dynamics as a result of the impact;
- 2) possibility of getting asymptotic solutions.

However, the econometric models based on differential equations are often complicated in their fulfillment, and what is more important, they do not allow to apply publicly available statistical data to forecast the object development. Many models artificially impose restrictions on the function form in the right side as a result of the deep analysis of the subject area. For example, work [3] proposes the model of the following type:

$$\frac{dy}{dt} = f_i^+ - f_i^- ,$$

where f_i^+ is the right side of the equation and includes all factors leading to the growth of variable y , and f_i^- includes all factors leading to decrease of f_i of the variable. The summands of the right formula side are presented as a product of functions depending only on the pre-identified factors. The authors explain this limitation by the fact that the simulating objective is simplified. However, it is necessary to priorly determine what factors lead to the reaction growth, what factors lead to the reaction decrease; how these factors can be presented in the form of products, etc. The authors of this work, as in [10], were interested in the global trends, emergence of bifurcations in the socio-economic system and their use in the control purposes, but not in “operational” tactical management of the territory, though based on its long-term prosperity. Therefore, this model is convenient for further analyzes but has a number of constraints, and cannot be considered as the best model for a number of complex economic systems, including simulation of the SR.

This model is similar to the model of world dynamics by Forrester [2], successfully describing asymptotic approximations based on the differential analyzer, which is actually the differential equation (1). But for adjustment of the latter, significantly voluminous data are needed (the author used the data of 1900–1970 period), and their reliability is under serious doubt. Such high demand to the data may result from the fact that the system consisting of 5 differential equations describing the global processes is applied, and these equations are mutually correlated. The author concludes that his long-term forecast until 2100 is not actual, since it is impossible to take into account all changes taking place as a result of technical and technological progress, so the conclusions stemming from his model bear rather qualitative character. They allowed us distinguish global cycles in the economy, general trends, etc. But at the level of regional economics management such data is generally not available, and forecast of events for the further 130 years is not required.

Many authors proposes differential econometric models, the selection of factors in which is also resulted from the deep analysis of the production processes and correlation between aggregated variables (such as production volume, cost of the main business assets and their growth rates, gross and net profit, amount of tax deductions, etc.). Since when building the mathematical model of the regional SR development it is impossible (or very difficult) to investigate preliminary the underlying processes, to determine the correlation between the factors which data is available, this approach is not acceptable. Speaking properly, this is the set task: to create a model identifying which it will become obvious what factors influence positively, what factors influence negatively, and influence of what factors can be neglected at all.

We believe reasonable to consider the traditional approach of the regression analysis when only the model configuration is predetermined (in this case, the form and the order of the differential equation), but any assumptions on the importance, correlation of the factors etc. are not made. We would like to receive very **simple** instrument such as regression multifactor model, which has no serious tasks of identification, but at the same time it is free of the above mentioned drawbacks.

The differential models in ecology and economy have been applied for a long time. For example, work [4] formulates the model of ecological balance

$$\frac{\partial y(x, y, t)}{\partial t} = y - y^2 - x(t) \cdot y \quad / \quad 0 \leq x(t) < 1 , \quad (2)$$

corresponding in its form to (1) that is close both to the earning capacity (defined here by the catch quota $x(t)$) and the disasters in population development, which will be considered later in the book. Work [6]

clarifies forecast attempt according to the noise model (2) with the help of trends and models such as (1). It is concluded that in this particular case, the model based on regression of differential equations is more adequate. The failed attempts described in this work *do not prove* the impossibility of successful approximations and extrapolation of the time trends in ecological-economic simulating with application of traditional and widely used methods, but illustrate the advantages of applying the ODE models as a basis. However, the model is a priori based on the first order differential equation. It's not surprising then that predictive models based on the polynomials bring are worse than the ODE based models. However this conclusion can not be considered as definitively proving the superiority of the ODE based models over the ordinary regression models.

To achieve objectives set in the introduction it is necessary to solve two tasks: to get a generalized criterion for assessing the system quality that would be stable both for errors of experts and initial data errors (for example, as a linear combination of partial criteria), and to build an econometric regression system model based on the regression differential equation.

3. Building the quality criterion

The following statistical indicators of the SR assessment were selected among presented on the web-site of the Perm branch of the Federal State Statistics Service. In our opinion, they are relevant or may become relevant to this subject area:

- 1) y_1 population size (thousand people);
- 2) y_2 age composition of population (%);
- 3) y_3 crude birth rate (per thousand people);
- 4) y_4 crude death rate (per thousand people);
- 5) y_5 number of theatergoers (per 1,000 persons);
- 6) y_6 number of economically active population (thousand people);
- 7) y_7 number of unemployed people (thousand people);
- 8) y_8 number of graduates having secondary vocational education (persons);
- 9) y_9 number of graduates having higher professional education (persons).

For elimination of the influence of indicator dimensionality their values were normalized according to the standard formula

$$\tilde{y}_i(t_j) = \frac{y_i - \min_j y_i(t_j)}{\max_j y_i(t_j) - \min_j y_i(t_j)}, \quad (3)$$

where $\tilde{y}_i(t_j) \in [0,1]$ is normalized value, $y_i(t_j)$ is the original value of the i -th partial criteria in the j -th year, $j = \overline{1,12}$ $j = \overline{1,12}$, where in 2000, $j = 1$.

The traditional approaches to the assessment of such systems are the following:

1. SR decomposition into units and levels \rightarrow the integrated indicator of the quality based on assessment of these units.
2. Indirect SR assessment according to the well-known (well-established, long-standing) indicators such as the Human Development Index (HDI). As a rule, such approaches are not proved by anything except for the views of their authors and elementary checks.
3. SR assessment depending on directly measured (and partially publicly available) statistical data. In this case, as well as in the previous position, the SR assessment is a linear combination of partial assessments with empirically (expert) determined coefficients.

For all these approaches, the final assessment is a linear convolution of partial criteria y_i

$$y = \sum_{i=1}^N \alpha_i y_i, \quad (4)$$

where $\alpha_i \geq 0$ is the weighting value of the indicator.

There also exists the task of determining the weighting coefficients of the partial criteria in (4). The generally accepted approach is when the expert opinions, or additional information determine

the ranks of private criteria significance R_i , while the lower rank values correspond to the higher criteria significance. For example, a simple formula meets these requirements

$$a_i = 1/R_i, \quad (5)$$

Since selection of the particular criteria among the available was rather optional, we used the standard methods to assess its admissibility. One of the methods to check independence of samples $y_k = \{y_{k1}, y_{k2}, \dots, y_{k12}\}$ and y_m is calculation of their pair correlation $R(y_k, y_m)$ that can be obtained using function CORREL in MS Excel. Having made this calculation, we received $C_{12}^2 = \frac{9!}{2!(9-2)!} = 36$

combinations. At first glance, the results of calculation show that some particular criteria under No. 1, 2, 3, 4, 8, 9 should be excluded from consideration. In particular, exclusion of criterion y_2 supposedly gives the opportunity not to consider $R(y_1, y_2)$ and $R(y_2, y_3)$; the same can be said about criteria No. 3 that has two correlations with greater value. Further selecting the optimal criteria ranking we will check if assumption about the influence of pair correlation on SR assessment is confirmed or not.

Selection of weighting coefficients by expert examination method has significant disadvantages. Despite the use of different methods of the expert examination theory in order to exclude subjectivity of the experts (calculation of correlation, concordance of opinions, etc.) it is always impossible to eliminate completely the possibility of incorrect expertise. Actually, it is typical for the nature of the examination, when the expertise organizer voluntarily or not selects those experts, among all, who can provide ranking satisfactory for him. Therefore, we would like to have some sort of evaluation criteria for ranking that does not depend on the expert selection, on method of processing individual rankings, etc. In a random case there may not be such criterion, but the following can be proposed in relation to SR. The social system with a large number of people under the influence of any factors (excluding global disturbances such as war, resettlement or pestilence) can not be changed *quickly*. During the last 20 years population of the Perm territory has been changing slowly and not significantly, there was no mass inflow or outflow of non-recurrent specialists, other disturbances mentioned above did not take place either, so it makes sense to hypothesize that evaluation of the territorial SR shall also change *smoothly*. That means, the best ranking should be the one which gives the assessment minimum

$$\alpha_i : S = \sum_j (y_{j+1} - y_j)^2 \rightarrow \min. \quad (6)$$

For selection of rankings ensuring the smoothest change in SR assessment, we created MS Excel spreadsheet with criterial data and developed a special VBA-program providing us with complete enumeration of all ranks permutations for all partial criteria by Fisher-Yates method [1], as there are relatively few criteria ($N = 9$) and permutations $N! = 9! \approx 3,6 \cdot 10^5$. With more private criteria we will have to use more effective enumerative methods, for example, cutting off repeated enumerations in case of $\alpha_i = \alpha_j$ with $i \neq j$. For each enumeration we have calculated S according to (6) for selecting the best ranking.

This method gave us the following rankings:

- full rankings (123 ... 9) where all original private criteria are saved;
- partial rankings (123 ...) where we assumed that 1, 2 or 3 private criteria can be discarded, and solved the task of finding the discarded private criteria;
- complete rankings like 111222333 and 112233444 where we have suggested that some private criteria may have the same importance (and hence the same weighting coefficients α_i in formula 4).

The smallest value of *non-smoothness* S was obtained at ranking 120406035, the chart of the SR assessment dependence on the year number is shown in Fig. 1. The same figure shows the trends of changes in the SR assessment with other similar criteria rankings. Obviously, the ranking 120406035 (the lower solid line) is really the best in terms of smoothness assessment.

The smoothness of the criteria ranking is caused by the nature of the system development that “can not” (excluding the fluctuating shifts) be changed rapidly. The optimal ranking does not contradict to the common sense (expert estimations) as the highest ranks in it belong to the population size (the SR

base), the age composition of population (i.e. its ability to act as SR) and the number of graduates with the secondary vocational education (the indicator common for supply of workers for the Perm territory enterprises). For example, at the beginning of 2000s, the territorial economy grew actively, and significant increase in revenue brought by the enterprises took place, which resulted in the level of GRP (largely owing to the favorable economic situation). The consequences resulted in growth of tax revenues, as well as increase in gross regional product per capita (revenue growth). At the same time, 2008–2010s were characterized by “failures” in the economy, we could observe the impact of the global economic crisis – the price drop for the exported resources, and as a result, decline in investments into the SR development, as well as the reduction of the real income of the population. However, such reduction was short-lived. This is largely due to the rapid restoration of the previous level of prices for the resource component, as well as the policies pursued by the state government in terms of the investment support of the real sector.

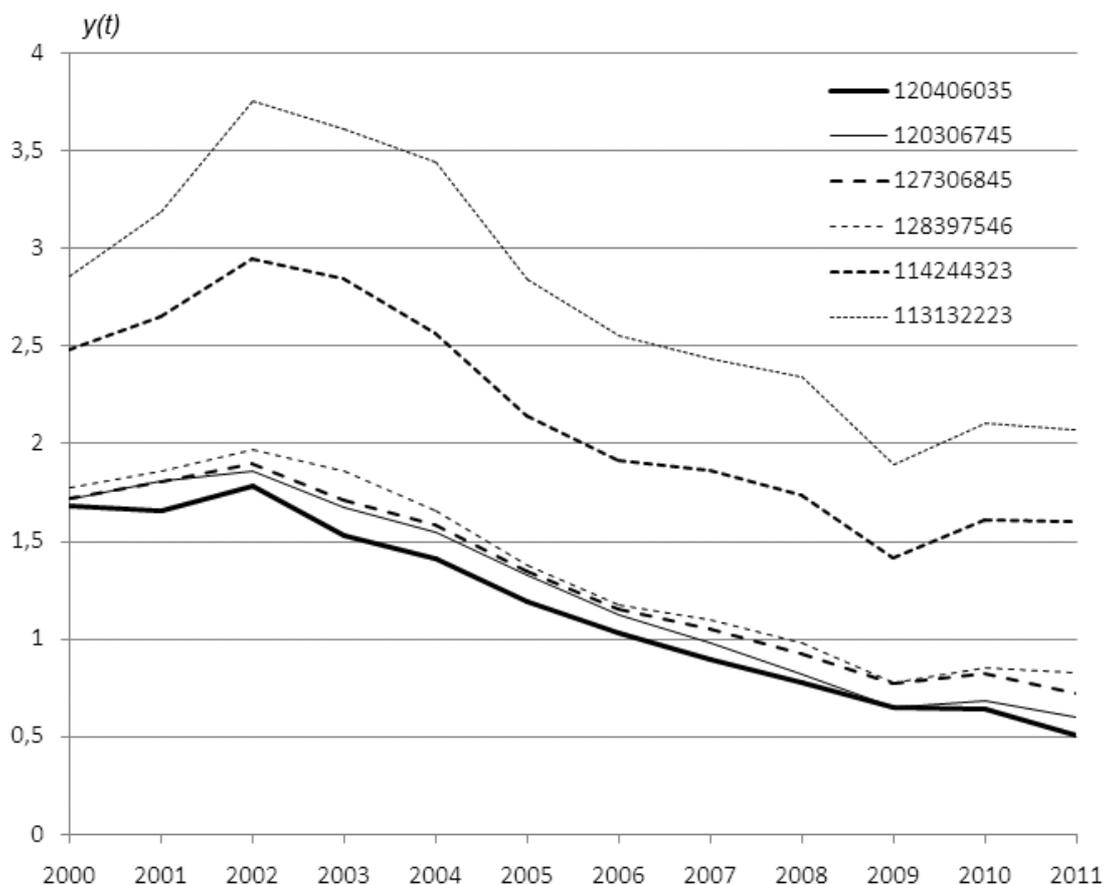


Fig. 1. Changes in the criteria of the SR assessment on the annual basis with various rankings of private criteria

Trend in Fig. 1 probably shows the Russian nationwide trend of the social resources quality drop by reducing the population size (for the Perm Territory – from 2,869 thousand people in 2000 to 2,651 thousand people in 2011) and for the reason of changes in its qualitative composition. Indeed, this assumption is more justified than the periodic assessment growth in 2002–2003 and 2010–2011 that would be given by ranking 120406035 and others. If the SR assessment growth in 2002–2003 can be explained by some aftereffects of the Soviet Union, the new significant (by 11 %) increase in the assessment in 2010 is unlikely to have any rational explanation.

4. Development of the differential regression model of the regional natural resources potential

The factors affecting the SR development which are characterized by the publicly available data, include the following:

- 1) x_1 rate of the positive migration balance;
- 2) x_2 GRP per capita;
- 3) x_3 residential area per inhabitant;
- 4) x_4 commissioning of the dwelling houses per 1,000 people;
- 5) x_5 share of the dwelling houses constructed by the population at their own expense,%;
- 6) x_6 per capita income of the population;
- 7) x_7 average monthly nominal salary;
- 8) x_8 deposits in Sberbank, at the beginning of the year;
- 9) x_9 average amount of pensions;
- 10) x_{10} amount of hospital beds;
- 11) x_{11} number of reported crimes per 100 thousand people;
- 12) x_{12} emissions of pollutants into the atmosphere;
- 13) x_{13} amount of the fixed assets at the end of the year;
- 14) x_{14} retail trade turnover per capita.

Typically, the basis for regression differential simulation is an ordinary first order differential equation of the following type:

$$\frac{dy(t)}{dt} = a_0 + \sum_{i=1}^m a_i x_i(t) + b \cdot y(t), \quad (7)$$

where a_i and b are coefficients of factor influence and system response to the dynamics of changes. Further the minimization of the square deviation of the calculated value $y(t_k)$ from the criteria $y_{src}(t_k)$ values known at the nodes of the annual series

$$S = \sum_{k=1}^K (y(t_k) - y_{src}(t_k))^2 \quad (8)$$

determines the coefficients of equation (1) that is integrated within the annual series or forecasting time-frame by the modified Euler 1st order method. The task of minimizing (8) can be solved by any means of optimization, for example, by coordinate-wise or gradient descent. The peculiarity of the author's approach is that the publicly available annual series of statistical data is used as the basis of the model.

It is assumed as a prerequisite condition of the forecasting capability that the model should be able to make post-forecast at least for the last year, if coefficients of the model are obtained from the data of the previous years and the value of the response, a little different from the known one, is calculated on the basis of the well-known values of factors. Reduction of the number of "well-known" years allows to determine the allowable forecasting time-frame in the first approximation.

If you are not limited to the first order ODE, it is possible to use the higher orders of the regression differential simulation

$$\begin{aligned} \frac{d^n y(t)}{dt^n} + \sum_{i=1}^{n-1} g_i \frac{d^i y(t)}{dt^i} &= a + b \cdot y(t - \tau_0) + \sum_{i=1}^m c_i \cdot x_i(t - \tau_i) + \\ &+ \sum_{i=1}^m \sum_{j=1}^m d_{ij} \cdot x_i(t - \tau_i) \cdot x_j(t - \tau_j) + \sum_{i=1}^m f_i \cdot [x_i(t - \tau_i)]^2, \end{aligned} \quad (9a)$$

where g_i is the influence coefficients of the lower derivatives, a is a constant describing the influence of one n-th derivative response during the trend formation, b is coefficient of "feedback" describing the impact of the response value on its n-th derivative, c_i refers to coefficients of factor influence, $d_{ij} : i \neq j$ are coefficients of the factor mutual influence, $f_i \equiv d_{ii}$ – the coefficients of the factor square influence, τ_i – the influence delay of i-th factor, τ_0 – delay in feedback. The regression differential

simulation is supplemented by the initial conditions $\frac{dy(0)}{dt} = y_0', \frac{d^2 y(0)}{dt^2} = y_0'', \dots, \frac{d^{(n-1)} y(0)}{dt^{(n-1)}} = y_0^{(n-1)}$.

In particular, if not take into account the factor feedback delay we obtain the following equation for regression differential simulation on the basis of the 2nd order

$$y''(t) + g \cdot y'(t) = a + b \cdot y(t) + \sum_{i=1}^m c_i \cdot x_i(t) + \sum_{i=1}^m \sum_{j=1}^m d_{ij} \cdot x_i(t) \cdot x_j(t) + \sum_{i=1}^m f_i \cdot [x_i(t)]^2, \quad (9b)$$

used further in this work.

The equation (9b) is solved by the modified Euler method, in this case, providing the design scheme:

$$y(0) = y_0,$$

$$y'(0) = y'_0,$$

$$\tilde{y}'(t + \Delta t) = y'(t) + \Delta t \cdot \left(a + b \cdot y(t) + \sum_{i=1}^m c_i \cdot x_i(t) + \sum_{i=1}^m \sum_{j=1}^m d_{ij} \cdot x_i(t) \cdot x_j(t) + \sum_{i=1}^m f_i \cdot [x_i(t)]^2 - g \cdot y'(t) \right),$$

$$y'(t + \Delta t) = y'(t) + \Delta t \cdot \left(a + b \cdot y(t) + \sum_{i=1}^m c_i \cdot x_i(t) + \sum_{i=1}^m \sum_{j=1}^m d_{ij} \cdot x_i(t) \cdot x_j(t) + \sum_{i=1}^m f_i \cdot [x_i(t)]^2 - g \cdot \frac{y'(t) + \tilde{y}'(t + \Delta t)}{2} \right),$$

$$y(t + \Delta t) = y(t) + \Delta t \cdot \frac{y'(t + \Delta t) + y'(t)}{2}.$$

The standard deviation of this method is checked by solving several test equations and does not exceed the fractions of a percent. Optimization (4) was performed by the modified method of coordinate-wise descent and gradient method.

Using described above methods we developed special software that implements the model, numerical methods and provides us with convenient instruments for study of socio-economic systems, including SR. Applying it we checked the forecasting capacity and quality with use of the linear regression model

$$y(x(t), t) = a_0 + \sum_i a_i x_i(t) \quad (10)$$

the coefficients of which were found by solution of task (6).

We established the linear interpolation of the factors as common for linear model. As a result, while setting all 11 known years we get a fairly accurate approximation, but when we reduce the amount of conventionally known years by 1 year and get a “forecast” for the last year, we get a completely inadequate trend of the following type (Fig. 2).

The relative deviation of forecast is 498% with the forecast for 1 year and 754% with forecast for 2 years. Such significant deviation is not acceptable for the practical SR forecast.

The common means of simulating socio-economic systems also include autoregressive models of different orders

$$y_{AP}(t_k) = a_0 + \sum_{i=1}^l a_i \cdot y(t_{k-i}), \quad (11)$$

and a state space model. Use of these models (Fig. 3) leads to the similar effects: the models approximate the original data very well, but can not forecast SR dynamics. Application of the Kalman adaptive filter for the state-space models just slightly improves the post-forecast for one year.

The attempt to use the regression differential simulation based on the 1st order ODE leads to similar effects – forecast for 1 year is rather accurate, whereas increase in the forecasting period up to 2 years leads to 184% forecast error which excludes any use of this-type model.

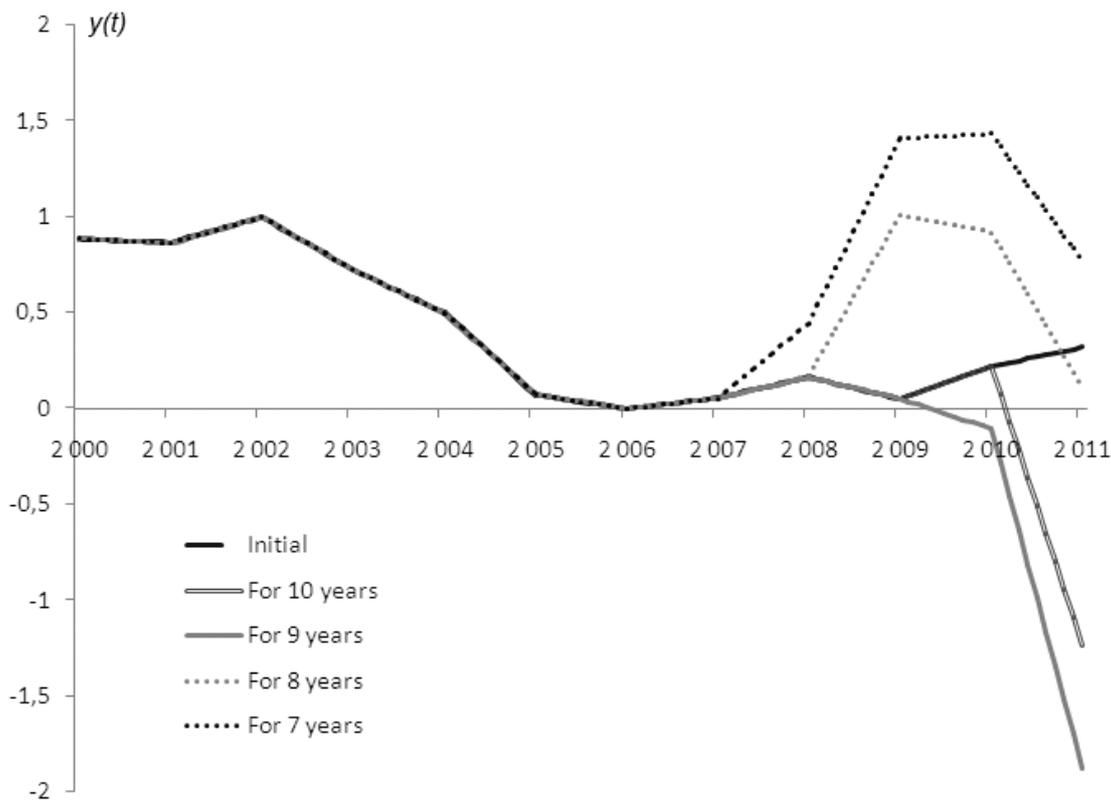


Fig. 2. Inadequate forecasts according to the data of 7-10 years with application of model (6)

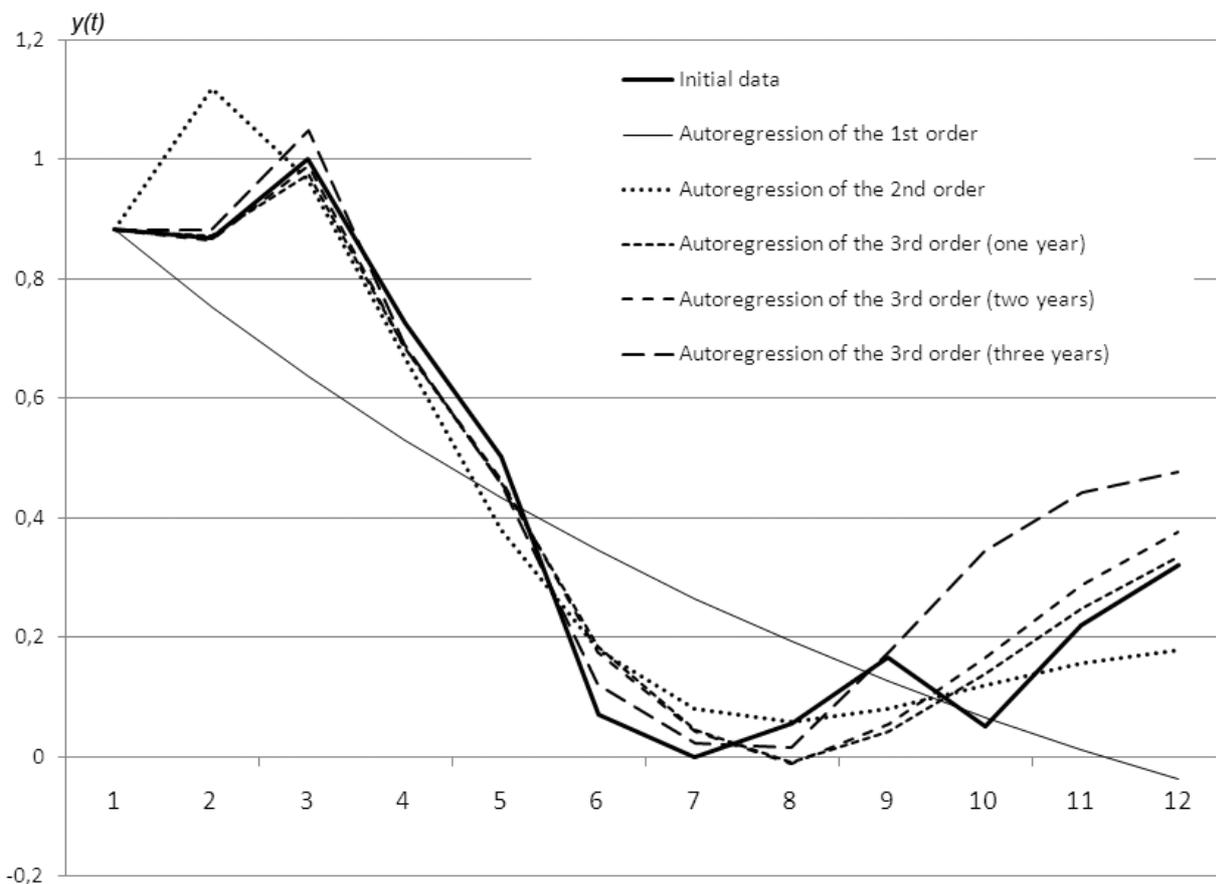


Fig. 3. The post-forecast results with application of models (7)

The model selection based on the ODEs of the 2nd and higher orders for describing the complex systems is caused by the nature of socio-economic system characterized by high response time. However, integration of the 1st order ODE at any chain approximation of the factors leads to the chain trend and bounces of the derivative value. This is unacceptable taking into account the SR system properties. The resulting curve should be smooth and have no breaks. This can be achieved by the following means:

1) interpolation of the SR assessment criterion $y_{\text{orig}}(t)$ in the intermediate points by a smooth cubic spline;

2) use of ODE of high (2nd and higher) order when the smooth curve is obtained by integrating it, regardless of the nature of the factor change.

The model based on the 2nd order ODE with the linear approximation of the factors without mutual interference coefficients makes good approximation of the initial data spline ($S = 0.058$). However, when we reduce the duration of the initial data series by one year, we get forecast for the 11th year with about 17% deviation.

As the number of “well-known” years is reduced not only to 10, but also to 6 years (Fig. 4) the forecast quality for the remaining years is rather satisfactory: $S \leq 0.001$, the forecast relative error

$$\delta y(11) = \left| \frac{0.3215 - 0.3121}{0.3215} \right| \approx 2.9\% .$$

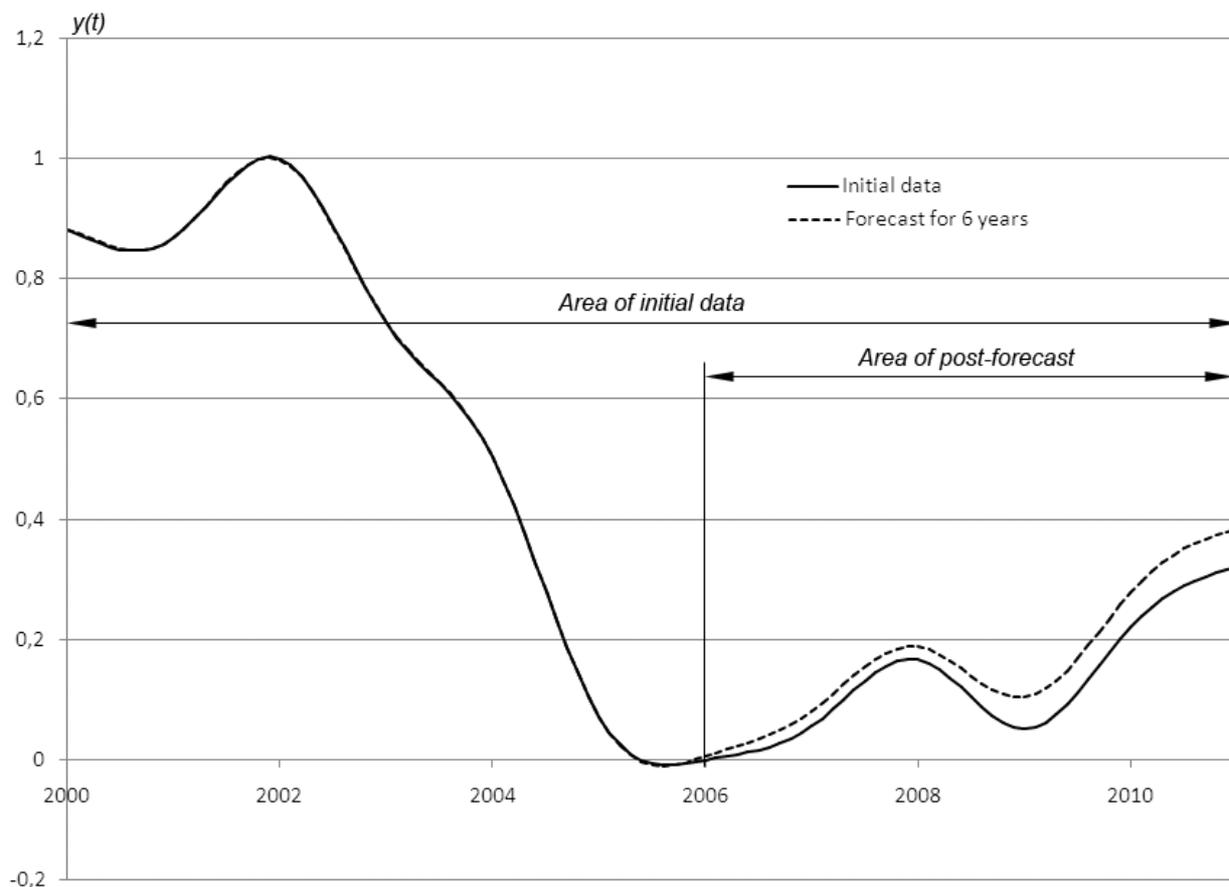


Fig. 4. Trends of initial and forecast data obtained from the regression differential simulation on the basis of the 2nd order ODE with coefficients obtained by 6 known values of the annual series

Therefore, we can assume that the model based on the 2nd order ODE is more applicable for simulating the development of the regional social resource than the linear model, the state space model and the regression differential simulation based on the 1st order ODE. Further complication of the model is not required, as the sufficient forecast quality within a reasonable forecast time-frame is already provided.

The study of the resulting model properties has given answers to the questions stated to those who manage the SR development from the state part.

Question 1. What factors are most important for the SR development, that is, to changes of what factors the model is most sensitive?

To answer the question we made a slight change in coefficients of the model (by $\pm 4\%$), while controlling the relative change in the system response.

It was found out that the greatest impact on the development dynamics of the SR system belong to the factor of “the number of hospital beds” (x_{10}). This indicator “leads” to many other factors: the number and area of the hospitals, the average incidence rate, the level of funding the healthcare system, etc. The regression differential simulation (in contrast to the linear multivariate model) “feels” the impact of these factor components. The influence of other factors on the response fluctuation is comparable, although factors with large coefficients c_i have a greater impact than others.

Question 2. What changes in the factors (or combinations thereof) allow to forecast improvement of the SR development within the operational (4-year) forecast time-frame?

We suppose that the factors will be developed approximately like “they have behaved” over the past few years. That is, for their forecast we must first:

a) Determine the regression equation of their behavior in the given retrospective view (usually greater in duration than the forecast time-frame specified above).

b) To calculate their values in future in accordance with equation obtained.

We distinguish three possible factor regression equations.

1. The factor varies quadratically, that is $x(t) \approx a + b \cdot t + c \cdot t^2$.

2. The factor varies linearly, that is $x(t) \approx a + b \cdot t$.

3. Both quadratic or linear laws are not acceptable for the factor. For example, the value of the factor is chaotically changing around some mean value $x(t) = \bar{x} \pm \Delta x$. In this case, we assume that this factor will remain at the level of one of the values within $\bar{x} \pm \Delta x$. In particular, you can use the last value of the annual series.

In order to determine the type of appropriate factor regression we have calculated the linear correlation coefficient R^2 that characterizes the proximity of the initial data to the data obtained by regression. It is assumed that if $R^2 \geq 0.95$, the selected type of the regression dependence satisfactorily describes the factor behavior in the retrospective. Moreover, since the linear regression is a special case of quadratic regression, algorithm for selecting the type of regression has the following form:

1. If for the linear regression $R^2 \geq 0.95$, the type of regression is linear.

2. If for the quadratic regression $R^2 \geq 0.95$, and for linear regression $R^2 < 0.95$, the type of regression is quadratic.

3. Otherwise, it is necessary to use the last value of the annual row.

Further the factors varied independently and then together in pairs by $\pm 5\%$. The results of the SR development forecasts with various independent factor changes are displayed in Fig. 5.

As in case with the study of sensitivity, the factors with small values of their coefficients in the model have little influence on changes in the forecast scenario for the SR system development. The greatest influence belong to the factors 2 (GRP per capita), 4 (the number of commissioned dwelling houses), 7 (the average monthly salary) and 9 (average pension). Thus, these factors have a greater impact on the future SR development than on its assessment at any particular point of time (as opposed to x_{10}).

We had the similar assessment of the mutual (synergistic) effect of the factors on the system behavior. The surfaces of the forecast changes are close to the flat changes, that indicates stability of the system as a whole in the sense that small mutual factor changes do not result in large changes in the system response.

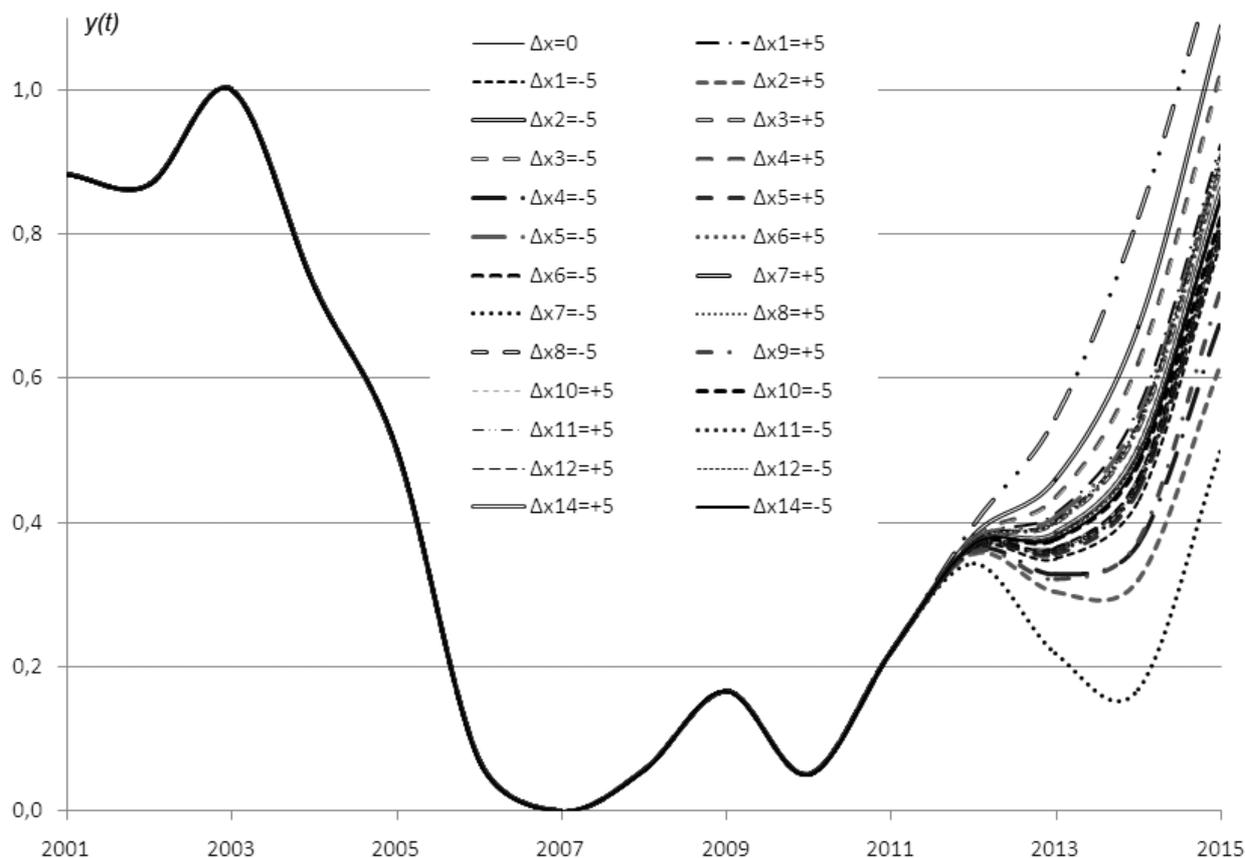


Fig. 5. Dynamics of the SR development forecast scenario with small-scale independent factor changes

Question 3. What controlling influences of the decision-makers can remove the system out of the negative forecast areas, if such areas are detected during the forecast?

To investigate the possible influence of the regional government on the negative forecast areas we have conducted the forecast with the changes in controllable factors x_2 , x_7 , x_9 and x_{10} by $\pm 5\%$ (Table), where $y(2015)$ is the estimated criterion value in the 15th year; $y_0(2015)$ is the initial criterion value in the 15th year.

The negative scenario trends when the amount of deposits (x_8) decreases and the crime rate increases (x_{11}) simultaneously can be overcome by increasing salaries by 5% that will lead to an increase in the SR by 29% (but it is necessary to increase the salaries only by 2% for the removal of the SR trend out of the negative forecast area). The reduction of pension by 5% leads to an increase in the SR by 3%.

Positive and negative forecast areas and possible ways to overcome the negative forecasts

Changes of the determinant trends $\Delta x_i, \%$	$\frac{y(2015)}{y_0(2015)}$ without compensation	Compensating action $\Delta x_i, \%$	$\frac{y(2015)}{y_0(2015)}$ with compensation
Amount of pensions ($\Delta x_9 - 5\%$)	+15%	—	—
Salary ($\Delta x_7 + 5\%$)	+41%	—	—
GDP per capita ($\Delta x_2 + 5\%$)	-27%	$\Delta x_7 + 5\%$	+14%
		$\Delta x_7 + 5\%$ $\Delta x_9 - 5\%$	+29%
Amount of deposits ($\Delta x_8 + 5\%$) and number of the reported crimes ($\Delta x_{11} - 5\%$)	-13%	$\Delta x_7 + 5\%$	+28%

The negative scenario trends when the level of nominal GRP (x_2) is increasing can be overcome by increasing the amount of nominal salaries by 5% that will lead to increase in the SR assessment by 14%. Simultaneous 5% decrease in the amount of pensions of the population and the same increase in the salaries will ensure the improvement in the SR dynamics by 28%.

5. The main results and conclusions

We proposed and justified method to support the decision-making on management of the regional social resource using the developed software that consists in a sequential

- 1) clarification of partial criteria system and determination of their rankings;
- 2) specification of the factor system;
- 3) joint determination of the order of regression differential simulation, approximation of the factors between the values of the annual series and the model coefficients;
- 4) determination of the forecast time-frame and error;
- 5) building the forecast scenarios depending on the impact of individual factors and their synergistic combinations;
- 6) selection of control actions that improve dynamics of the SR development in the case of adverse forecasts.

An important additional result of our work is development and testing of the multipurpose software that allows to simulate multi-factor socio-economic systems and forecast their development options under various impacts. We have demonstrated the advantage of the developed method over other prevalent approaches in terms of their forecast quality.

References

1. Donald E. Knuth. An analysis of optimum caching. *Journal of Algorithms*, June 1985, vol. 6, iss. 2, pp. 181–199.
2. Jay W. Forrester, World Dynamics, Wright-Allen Press, Inc., 238 Main Street, Cambridge, Massachusetts 02142. 142 p.
3. Akayev A.A., Sadovnichiy V.A. *Matematicheskoe modelirovanie global'noy, regional'noy i natsional'noy dinamiki s uchetom vozdeystviya tsiklicheskih kolebaniy* [Mathematical Modeling of Global, Regional and National Dynamics Given the Impact of Cyclical Fluctuations]. Available at: http://socmodel.com/sites/socmodel.com/files/articles/3A00_Akaev_Sadovnichiy.pdf.
4. Arnold V.I. *Teoriya katastrof* [Catastrophic Theory]. Moscow, Nauka., 1990. 128 p.
5. Dziuba S.A. *Modeli upravleniya podsistemami predpriyatiya v sfere srednego biznesa i ikh instrumental'noe obespechenie. Avtoref. diss. dokt. ekon. nauk* [Management Models for the Enterprise Subsystems in the Field of Medium-Sized Businesses and Their Instrument Maintenance. Abstract of Doct. (Econ.) Diss.]. Available at: <http://econom.nsc.ru/ieie/news/zashiti/avtoref/mart12/dzuba.pdf>.
6. Zatonkiy A.V. [The Advantages of Differential Models in Ecological and Economic Simulating]. *Bulletin of the Tomsk Polytechnic University*, 2012, no. 5, pp. 134–139. (in Russ.)
7. Zatonkiy A.V., Sirotina N.A. Forecasting the Economic Systems According to the Model Based on the Regression Model Differential Equation. *Economics and Mathematical Methods*, 2014, no. 50 (1), pp. 91–99. (in Russ.)
8. Zatonkiy A.V., Sirotina N.A., Yanchenko T.V. [On the Approximation of Factors of the Differential Model of Socio-Economic System]. *Current Researches on Social Tasks (E-Journal)*, 2012, no. 11 (19). Available at: <http://sisp.nkras.ru/e-ru/issues/2012/11/sirotina.pdf>. (in Russ.)
9. Losev K.S. *Mify i zabluzhdeniya v ekologii* [Myths and Misconceptions in Ecology]. Moscow, Nauchnyy mir Publ., 2010. 224 p.
10. Malinetskiy G.G. *Matematicheskie osnovy sinergetiki: Khaos, struktury, vychislitel'nyy eksperiment* [Mathematical Foundations of Synergetics: Chaos, Structures, Computational Experiment]. Moscow, Publishing House LKI, 2007. 312 p.
11. Mirolyubova A.A. *Metodologiya modelirovaniya investitsionnogo protsessa v real'nom sektore ekonomiki regiona. Avtoref. diss. dokt. ekon. nauk* [Simulation Methodology for the Investment Process in the Real Sector of the Regional Economy. Abstract of Doct. (Econ.) Diss.]. Available at: <http://vak.ed.gov.ru/common/img/uploaded/files/MirolubovaAA.docx>.

12. Sirotina N.A., Zatonkiy A.V. [Multi-Factor Model of a Complex System Based on the First Order Differential Equation]. *Bulletin of the Kamsk Institute of Humanities and Engineering Technologies*, 2012, no. 7, pp. 9–18. (in Russ.)

13. Yanchenko T.V., Zatonkiy A.V. [Determination of the Optimal Ranking of Private Assessment Criteria of the Regional Social Resource]. *Economy and Management of Control Systems*, 2013, no. 4 (10), pp. 99–104. (in Russ.)

Received 15 March 2016

УДК 004.94:338.27

DOI: 10.14529/ctcr160211

УПРАВЛЕНИЕ ДИНАМИКОЙ РЕГИОНАЛЬНОГО СОЦИАЛЬНОГО РЕСУРСА НА ОСНОВЕ РЕГРЕССИОННО-ДИФФЕРЕНЦИАЛЬНОГО МОДЕЛИРОВАНИЯ

А.В. Затонский, Т.В. Сафьянова

*Пермский национальный исследовательский политехнический университет,
Березниковский филиал, г. Березники, Пермский край*

Обоснована важность моделирования и прогнозирования динамики регионального социального ресурса применительно к замкнутым территориально-промышленным комплексом на примере Верхнекамского региона Пермского края РФ. В качестве оценки развития социального ресурса в данном случае принимается линейная комбинация нормированных частных критериев из числа общедоступных статистических данных. Для уточнения ранжирования частных критериев использована гипотеза о гладкости результирующего критерия вследствие большой инерционности рассматриваемой системы. Показана устойчивость финальной ранжировки к малым возможным ошибкам экспертного определения рангов. Показана невозможность использования традиционных эконометрических моделей и моделей на основе регрессионного дифференциального уравнения первого порядка для решения этой задачи. Разработана модель на основе регрессионного дифференциального уравнения второго порядка, позволяющая в тактической перспективе прогнозировать последствия решений по развитию социального ресурса. Качество модели проверено не только по отклонению от исходных данных, но и постпрогнозом в пределах разумного горизонта. Особенностью модели является использование общедоступных статистических факторов, для которых определяются их сравнительная важность, коэффициенты взаимного влияния, наилучшая аппроксимация между значениями годового статистического ряда и другие параметры настройки модели. Определены положительные и отрицательные области прогноза, даны рекомендации по улучшению динамики социального ресурса. Важным побочным результатом работы является создание универсального программного инструмента моделирования динамики социально-экономических систем на основе разреженных общедоступных статистических данных, компенсирующего в некоторой степени недостатки подобных данных.

Ключевые слова: социальный ресурс, регрессионно-дифференциальная модель, управление, прогнозирование, программное обеспечение.

Литература

1. Knuth, D.E. *An analysis of optimum caching* / D.E. Knuth // *Journal of Algorithms*. – June 1985. – Vol. 6, iss. 2. – P. 181–199.

2. Forrester, J.W. *World Dynamics* / J.W. Forrester. – Wright-Allen Press, Inc., 238 Main Street, Cambridge, Massachusetts 02142. 142 p.

3. Акаев, А.А. *Математическое моделирование глобальной, региональной и национальной динамики с учетом воздействия циклических колебаний* / А.А. Акаев, В.А. Садовничий. – http://socmodel.com/sites/socmodel.com/files/articles/3A00_Akaev_Sadovnichy.pdf.

4. Арнольд, В. И. Теория катастроф / В.И. Арнольд. – М.: Наука, 1990. – 128 с.
5. Дзюба, С.А. Модели управления подсистемами предприятия в сфере среднего бизнеса и их инструментальное обеспечение: автореф. дис. ... д-ра экон. наук / С.А. Дзюба – <http://economy.nsc.ru/ieie/news/zashiti/avtoref/mart12/dzuba.pdf>.
6. Затонский, А.В. Преимущества дифференциальных моделей в эколого-экономическом моделировании / А.В. Затонский // Известия Томского политехнического университета. – 2012. – № 5. – С. 134–139.
7. Затонский, А.В. Прогнозирование экономических систем по модели на основе регрессионного дифференциального уравнения / А.В. Затонский, Н.А. Сиротина // Экономика и математические методы. – 2014. – № 50 (1). – С. 91–99.
8. Затонский, А.В. Об аппроксимации факторов дифференциальной модели социально-экономической системы / А.В. Затонский, Н.А. Сиротина, Т.В. Янченко // Современные исследования социальных проблем (электронный журнал). – 2012. – № 11 (19). – <http://sisp.nkras.ru/e-ru/issues/2012/11/sirotina.pdf>.
9. Лосев, К.С. Мифы и заблуждения в экологии / К.С. Лосев. – М.: Научный мир, 2010. – 224 с.
10. Малинецкий, Г.Г. Математические основы синергетики: Хаос, структуры, вычислительный эксперимент / Г.Г. Малинецкий. – М.: Изд-во ЛКИ, 2007. – 312 с.
11. Миролюбова, А.А. Методология моделирования инвестиционного процесса в реальном секторе экономики региона: автореф. дис. ... д-ра экон. наук / А.А. Миролюбова. – <http://vak.ed.gov.ru/common/img/uploaded/files/MirolubovaAA.docx>.
12. Сиротина, Н.А. Многофакторная модель сложной системы на основе дифференциального уравнения первого порядка / Н.А. Сиротина, А.В. Затонский // Вестник КИГИТ. – 2012. – № 7. – С. 9–18.
13. Янченко, Т.В. Определение оптимальной ранжировки частных критериев оценки краевого социального ресурса / Т.В. Янченко, А.В. Затонский // Экономика и менеджмент систем управления. – 2013. – № 4 (10). – С. 99–104.

Затонский Андрей Владимирович, д-р техн. наук, профессор, зав. кафедрой автоматизации технологических процессов, Пермский национальный исследовательский политехнический университет, Березниковский филиал, г. Березники, Пермский край; z Xenon@narod.ru.

Сафьянова Татьяна Васильевна, канд. техн. наук, доцент кафедры экономики, Пермский национальный исследовательский политехнический университет, Березниковский филиал, г. Березники, Пермский край; ayagumunenori@mail.ru.

Поступила в редакцию 15 марта 2016 г.

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Zatonskiy, A.V. Control of the Regional Social Resource Dynamic Based on the Regression-Differential Simulation / A.V. Zatonskiy, T.V. Safyanova // Вестник ЮУрГУ. Серия «Компьютерные технологии, управление, радиоэлектроника». – 2016. – Т. 16, № 2. – С. 101–115. DOI: 10.14529/ctcr160211

FOR CITATION

Zatonskiy A.V., Safyanova T.V. Control of the Regional Social Resource Dynamic Based on the Regression-Differential Simulation. *Bulletin of the South Ural State University. Ser. Computer Technologies, Automatic Control, Radio Electronics*, 2016, vol. 16, no. 2, pp. 101–115. DOI: 10.14529/ctcr160211