Methodological Approach to the Identification of Predictive Models of Socio-economic Processes for Investment and Innovative Development of Enterprises

T.L. Bezrukova, S.S. Morkovina, B.A. Bezrukov and E.G. Popkova

Doctor of Economics, Voronezh State Forest Engineering Academy, Russia, Voronezh,
Candidate of Economic Sciences, Voronezh State Forest Engineering Academy, Russia, Voronezh,
Doctor of Economics, Volgograd State Technical University, Russia, Volgograd,

Submitted: Sep 6, 2013; Accepted: Nov 5, 2013; Published: Nov 10, 2013

Abstract: In article the algorithm of management developed by authors is considered by development of investment and innovative activity of the enterprises which unlike the available is presented in the form of step by step realized actions, allowing to make administrative decisions on ensuring efficiency of creation of innovations at all stages of innovative process. The model of investment and innovative development of the enterprises of furniture branch in which basis the method of identification of the structural and dynamic systems, being more perfect instrument of identification and forecasting of a condition of investment and innovative processes is put is also scientifically proved. The model allows to define necessary level of expenses for development of innovative activity and to optimize structure of sources of its investment.

Key words: Social and economic processes • Investment and innovative development • Predictive model • Industrial enterprise

INTRODUCTION

Effective development of modern enterprises is determined by the ability of the entity to the creation of new products, transformation of better ideas in the production, etc., which in turn is possible only in conditions of having innovation and investment resources and therefore, makes the company attractive for investors.

Investment and innovation development (IID) is an organization of innovation activity as a continuous process based on individual circuit of invested funds, resulted in creation of a new improved quality, state of the object, its composition or structure.

RESULTS

The authors have developed a model for investment and innovation development of companies with limited financial resources.

The objectives of development of investment and innovation activities of industrial enterprises, as an object of mathematical modeling, should be considered, based on the following class of structures for the identification of models of socio-economic processes:

\[
\frac{\partial Y}{\partial t} = a_1 \frac{\partial x_1}{\partial t} + a_2 \frac{\partial x_2}{\partial t} + ... + a_m \frac{\partial x_m}{\partial t} \\
+ b_1 \frac{\partial v_1}{\partial t} + b_2 \frac{\partial v_2}{\partial t} + ... + b_k \frac{\partial v_k}{\partial t} \\
+ z_1 \frac{\partial w_1}{\partial t} + z_2 \frac{\partial w_2}{\partial t} + ... + z_m \frac{\partial w_m}{\partial t} + a_0
\]

(1)

where Y - identifiable function (for commercial companies Y is usually revenue), x1, x2, ... - external parameters, v1, v2,... - internal parameters and w1, w2,... - control parameters of socio-economic system, t - time.

For identification of the proposed models (1) should be used the methods of regression, factor and correlation analysis. The analysis of advantages and disadvantages of these methods showed that all the methods in varying degrees include a subjective principle. Therefore, as a method for identification of prognostic models of socio-economic processes of industrial enterprises was selected method of group data handling (GMDH - Group Method of Data Handling) theory of self-organization [1, 2].

Corresponding Author: T.L. Bezrukova, Doctor of Economics, Professor, Russia, Voronezh, Timiriyzeva St., 8
Basics of the theory of self-organization of mathematical models include:

- **Principle of “external addition”** [1]. To identify the optimal model external criteria for comparison of models is used, i.e., these criteria, the definition of which is based on new data samples in comparison with the sample used to build the model.

- **The principle of “freedom of choice-making”** [1]. Full description of the model (basic function) should allow make choice of several best solutions at each stage of identifying the optimal model, that is, the experiment should allow to identify as much input parameters as possible.

According to these assumptions of the theory of self-organization the task of structural-parameter identification of model of the process solved by GMDH can be written in the general form [1, 2]:

$$ f^* = \arg\min_{f \in F} K_p(f) $$

(2)

where F - the set of the models, $K_p$ - external criterion of model quality $f$ in this set. Since the F-discrete set, (2) is a discrete task (integer) programming specifics of which is that for each $f \in F$ additional parameter extreme task of estimation of parameters $f$ by the minimum of internal criteria, does not coincide with $K_p$ is solved.

Obviously, (2) does not include a comprehensive statement of the task, so in addition it is necessary to: define the type and amount of the initial information, specify the class of basic functions (operators) that make up the set F; determine how to generate the models together with a method of estimating the parameters, select the external criterion compare models, specify the method of minimizing the $K_p$. This means that in general the process of solving the task of the form (2) includes the following steps:

- Monitoring of researched processes and obtain experimental data (sample cases) and other characteristics of the process.
- Class choice of basic functions of models based on analysis of data and the corresponding conversion.
- Generation of various structures of models of in the selected class.
- Estimation of the parameters generated structures and the formation of variety F.
- Minimizing of external criteria $K_p$ and choose the best model $f^*$.
- Verification of the adequacy of the resulting optimal model.

GMDH can be used to identify models reflecting consistent pattern of connections “input – output” (“physical” model), models of the best filtering (smoothing) of noisy output variables (“filter” model) and obtain of predictive models solving the problem of extrapolation of output variables (“predictive” model) [1]. Implementation of experiments to identify the three types of models is achieved by using external criteria GMDH.

Analysis of the stages of solving of the task of structural-parametric identification of model of social-economic process shows that to obtain a positive result we should correctly select the class of basic functions, external criteria compare models and consistency of their implementation [2].

Class of structures of models (F) (1) was chosen because of the condition of obtaining prediction models, which require the model to reflect some of the dynamics of the process under study and allow extrapolate this trend over time. The correct choice of the class of structures is confirmed by receiving the predictive model of significance. As a rule, for the models used for situational management of social and economic process relative prediction error should not exceed 2-3%. Otherwise the refinement of predictive models of class structures (F) should take place.

In this scientific paper the authors propos to use unbiasedness criteria (n), the convergence criteria (i) [1], the criteria of epignose prediction (P) [2], combined and scenario criteria [1] as external criteria of comparing models.

Let's define that:

- X - (n x m) matrix of values of the input variable (selection);
- Vector of the output variable $Y = (y_{o1}^1 + \eta_1, \ldots, y_{on}^n + \eta_n)$.

where, $y_{o1}^1, \ldots, y_{on}^n$ - the true values of the output value at n points,

$$ \eta = (\eta_1, \ldots, \eta_n) $$

- an error of observations and the values $\eta_i$, ($i=1,\ldots,n$) are assumed to be random, independent and identically distributed with zero expected value and finite variance.
• \( \Omega \)-output sample can be written as \( \Omega = D + C \), where \( D = A + B \).

In this notation, the unbiasedness criterion (criterion of a minimum displacement):

\[
n^2_{\text{gw}} = \| f_A - f_B \|^2, \quad \| \cdot \| \text{ in the space } \mathbb{R}^4,
\]

where \( f_A \) and \( f_B \) - models obtained on sample A and B, require the maximum coincidence of the output values of the two models obtained in two different parts of the original data table. This is especially actually for models of social and economic processes that are history confirms, exposed of various crises.

Convergence criterion has the form:

\[
i^2 = \| u_0 - f_D \|^2,
\]

where \( U_0 \) - modelling value of the output variable, obtained on a sample D, \( f_D \) - actual value of the output variable in the sample D. Convergence criterion is essentially a prediction error. Using of the convergence criterion allow check the model on the adequacy to the researched process in the range of interpolation of the function. Error criterion of epignose forecast:

\[
P^2 \| u_{DC} - f_C \|^2,
\]

where \( U_{DC} \) - model value of the output variable. The model is defined on a sample of D and the values calculated for the sample C, \( f_C \) - actual value of the output variable in the sample C. The use of the criterion of error epignose prediction model can check the adequacy of the process studied in the interval of extrapolation of the function.

For the uniqueness of the model selection combined criteria should be used. It is proposed to use the combined criterion of the form:

\[
K = n^2_{\text{gw}} + i^2.
\]

In the space of 4 external criteria of comparison of models 2-3 models are selected, which are then verified by the consistency of the physical picture of the process (“scenario criteria”) to determine the optimal model of the process. Long-term projections of economic processes at time \( 2T \), where \( T \) - time prediction and analysis of the physical picture of the process of the process complete the process of structural identification of models of economic processes. If the forecasted physical picture of the modeled process adequate to the real, we can say that during the time \( T \), we get a qualitative prediction [8].

As criteria for selecting the optimal structure in this case, we should pay attention to the change of the desired function, for example in the case of negative values of the unknown value, we should choose the structure which do not produce such results. Bust structures with the use of “scenario criterion” does not exceed 10% of the models selected by the combined criteria. Typically, the structures with minimum of errors of epignose projection or convergence criterion are optimal model of socio-economic processes [2].

Based on the proposed methodological approach to the identification of predictive models of socio-economic processes of investment and innovation for the development of enterprises has been developed algorithm for identification of prognostic models of socio-economic processes (Figure 1).

Implementation of a methodological approach to the identification of predictive models of socio-economic processes for investment and innovation development of enterprises is a modernized combinatorial GMDH algorithm, which is based on the deductive-inductive approach to modeling the processes of socio-economic systems [2].

According to the sixth stages of the method of identification of prognostic models of socio-economic processes in the first step of the algorithm of identification of prognostic models of socio-economic processes, it is necessary to enter the actual values of the independent and dependent variables, with the release of the control parameters.

An experiment of identification of the predictive models of socio-economic processes in accordance with a principle of self-organization theory should include as much as possible variables for the selection of the optimal implementation of the model. In this case, we have restricted the number of 100 independent variables, which reflects the real opportunity for researchers to organize continuous sampling of all the independent variables [7].

The second block of the algorithm determines on the basis of (1) the basic class of structures of models in the form of:
Fig. 1: The method of identification of prognostic models of socio-economic processes for investment and innovation development of enterprises
According to (7) dependent and independent variables are formed.

Considering the principle of “freedom of choice-making”, the method of modeling (using this term we mean the identification of predictive models) socio-economic processes of industrial enterprises admit to the experiment of structural identification as much parameters of different dimensions as possible. In this regard, centering and standardization of initial data is made in accordance with (8).

\[
X_{ij} = \frac{X_{j,i} - X_{c,i}}{S_{c,i}},
\]

(8)

where: \(X_{j,i}\) - the current value; \(i = 1, ..., n\), where \(n\) - the length of the sample, \(j = 1, ..., m\), where \(m\) - the number of variables, \(X_{c}\) - the mean value of the original variable; \(S_{c}\) - standard deviation of the original variable.

To determine the parameters of \(m\) input variables and their exclusion from the experiment in order to prevent ill-conditioning of the coefficient matrix we use pair-correlation coefficient (9) [3].

\[
R_{ij} = \frac{\sum_{y=M_1}^{M_2} X_{y,i} X_{y,j}}{\frac{M_2-M_1+1}{2}}, j \neq i; \quad 1 \leq j \neq i,
\]

(9)

where \(M_1\) - the beginning of the sample and \(M_2\) - the end of the sample. Bivariate correlation coefficient allow to exclude from the process of identification of the model with the dependent variable, i.e. with a coefficient \(R_{ij} > 0.7\).

To obtain the best predictive models we should determine the type and sequence of external criteria of selection. It is proposed to use the following criteria:

Unbiasedness criteria [4]:

\[
n_{cm} = \sqrt{\frac{\sum_{i=1}^{D} (Y_{A,i} - Y_{B,i})^2}{\sum_{i=1}^{D} Y_i^2}},
\]

(10)

where: \(Y_A\) and \(Y_B\) - models obtained on samples A and B, \(Y\) - the actual value of the output variable (researched microeconomic process, such as the cost of innovation);

Convergence criterion [4]

\[
I = \sqrt{\frac{\sum_{i=1}^{D} (Y_{im,i} - Y_{it,i})^2}{\sum_{i=1}^{D} Y_{it,i}^2}},
\]

(11)

where: \(Y_{im}\) - model value of the output variable, \(Y_{it}\) - table value of the output variable;

Criteria of error of epignose prediction [5]:

\[
P = \sqrt{\frac{\sum_{i=D+1}^{C} (Y_{im,i} - Y_{it,i})^2}{\sum_{i=D+1}^{C} Y_{it,i}^2}},
\]

(12)

where: \(Y_{im}\) - a model obtained from a sample D;
Fig. 2: Algorithm of identification of prognostic models of providing investment and innovation development of industrial enterprises

Combined criterion [2]:

\[ K = \frac{n}{n_{\text{max}}} + \frac{i}{i_{\text{max}}} \]  \hspace{1cm} (13)

where \( n_{\text{max}}, i_{\text{max}} \) - maximum displacement and minimum criteria for convergence.

Priority use of an external criterion is determined in an experiment of model of socio-economic processes by the quality of the predictive model and in the case of poor quality (accuracy epignose forecast more than five percent) the order of application of external criteria changes [9].

Summarizing all the stages of the technique of identification of prognostic models of socio-economic processes the algorithm of identification of predictive models of socio-economic processes can be represented as a chain of 8 blocks (Figure 2).

Let’s list them.

**I Block**

1. Data entry.

**II Block**

2. Forming of an array of external parameters according to the number of independent variables.
3. Forming of the base class of models based on the expression 7

III Block

4. Normalization and alignment of data for the above data to the same species (8).
5. Determination of pair correlations, removing the dependent variables (9).

IV Block

6. Assignment of a particular form of the model from the base class of models (7).
7. The definition of coefficients of the model by the method of least squares [4].
8. Determination of the value criterion of unbiasedness of the model (10).

V Block

9. Choice of N, best by unbiasedness criteria of the models.
10. Determining of the value of convergence criterion of the model (11).
11. Choice of N, best by the criteria of convergence of the models.

VI Block

12. Certain of criteria of error of epignose prediction (12).

VII Block

13. Determination of the combined criteria (13).

VIII Block

15. Determination of the optimal model of the researched process based on the “scenario criteria”.

The algorithm proposed above is implemented as a software system for the identification of predictive models of socio-economic processes EKONOM 2.1 language Object Pascal 7.1 environment in Delphi 7.0. Time of a single experiment for 4-year-old sample of 22 independent variables was about 2 hours on a PC with an Intel Core 2 Duo CPU E8400 3.0 GHz.

CONCLUSIONS

The authors carried out modeling of investment and innovation of industrial enterprises. Detailed analysis of the industry allowed to offer basic class of models (1), reflecting all of the typical characteristics of industrial enterprises and the engendering of predictive models of socio-economic processes, which is implemented on the basis of effective innovative business management.

Based on the study of modern information technology modeling socio-economic processes, the authors developed an original technique for the identification of predictive models under uncertainty influencing factors based on the theory of self-organization and implementation - method of group calculation of data for effective management of business processes of the industrial enterprise.

The authors create the effective algorithm of identification of socio-economic processes of production that implements efficient computational process on the basis of the systematic approach to the definition of the dependent variables. The proposed structural-parametric approach, the selection and sequence of application of external criteria of comparison models provide predictive models of socio-economic processes, able to implement contingency management of an industry.

Error of prediction models (the value criterion of epignose forecast) varies in the range in the normal range, which allows give the desired function of economic and social processes of production for the purposes of management, including the management decisions.

REFERENCES


