ISSN 2587-814X (print), ISSN 2587-8158 (online)
Russian version: ISSN 1998–0663 (print), ISSN 2587–8166 (online)

BUSINESS INFORMATICS

HSE SCIENTIFIC JOURNAL

Vol. 13 No 3 – 2019

CONTENTS

Publisher:
National Research University
Higher School of Economics

Subscription index in the Rospechat catalog – 80870

The journal is published quarterly

The journal is included into the list of peer reviewed scientific editions established by the Supreme Certification Commission of the Russian Federation

Address:
28/11, build. 4, Shablovka Street
Moscow 119049, Russia

Tel./fax: +7 (495) 772-9590 *26311
http://bijournal.hse.ru
E-mail: bijournal@hse.ru

Circulation:
English version – 300 copies,
Russian version – 300 copies,
online versions in English and Russian – open access

Printed in HSE Printing House
3, Kochnovsky Proezd, Moscow,
Russian Federation

ABOUT THE JOURNAL

Business Informatics is a peer reviewed interdisciplinary academic journal published since 2007 by National Research University Higher School of Economics (HSE), Moscow, Russian Federation. The journal is administered by School of Business Informatics. The journal is published quarterly.

The mission of the journal is to develop business informatics as a new field within both information technologies and management. It provides dissemination of latest technical and methodological developments, promotes new competences and provides a framework for discussion in the field of application of modern IT solutions in business, management and economics.

The journal publishes papers in the areas of, but not limited to:

- data analysis and intelligence systems
- information systems and technologies in business
- mathematical methods and algorithms of business informatics
- software engineering
- internet technologies
- business processes modeling and analysis
- standardization, certification, quality, innovations
- legal aspects of business informatics
- decision making and business intelligence
- modeling of social and economic systems
- information security.

The journal is included into the list of peer reviewed scientific editions established by the Supreme Certification Commission of the Russian Federation.

The journal is included into Web of Science Emerging Sources Citation Index (WoS ESCI) and Russian Science Citation Index on the Web of Science platform (RSCI).

International Standard Serial Number (ISSN): 2587-814X (in English), 1998-0663 (in Russian).

Editor-in-Chief: Dr. Alexey Golosov – President of FORS Development Center, Moscow, Russian Federation.

EDITORIAL BOARD

ABOUT THE HIGHER SCHOOL OF ECONOMICS

Consistently ranked as one of Russia's top universities, the Higher School of Economics (HSE) is a leader in Russian education and one of the preeminent economics and social sciences universities in Eastern Europe and Eurasia.
Having rapidly grown into a well-renowned research university over two decades, HSE sets itself apart with its international presence and cooperation.

Our faculty, researchers, and students represent over 50 countries, and are dedicated to maintaining the highest academic standards. Our newly adopted structural reforms support both HSE's drive to internationalize and the groundbreaking research of our faculty, researchers, and students.

Now a dynamic university with four campuses, HSE is a leader in combining Russian educational traditions with the best international teaching and research practices. HSE offers outstanding educational programs from secondary school to doctoral studies, with top departments and research centers in a number of international fields.

Since 2013, HSE has been a member of the 5-100 Russian Academic Excellence Project, a highly selective government program aimed at boosting the international competitiveness of Russian universities.

ABOUT THE SCHOOL OF BUSINESS INFORMATICS

The School of Business Informatics is one of the leading divisions of HSE's Faculty of Business and Management. The School offers students diverse courses taught by full-time HSE instructors and invited business practitioners. Students are also given the opportunity to carry out fundamental and applied projects at various academic centers and laboratories.

Within the undergraduate program, students participate each year in different case-competitions (PWC, E&Y, Deloitte, Cisco, Google, CIMA, Microsoft Imagine CUP, IBM Smarter Planet, GMC etc.) and some of them are usually as being best students by IBM, Microsoft, SAP, etc. Students also have an opportunity to participate in exchange programs with the University of Passau, the University of Munster, the University of Business and Economics in Vienna, the Seoul National University of Science and Technology, the Radbound University Nijmegen and various summer schools (Hong Kong, Israel etc.). Graduates successfully continue their studies in Russia and abroad, start their own businesses and are employed in high-skilled positions in IT companies.

There are four graduate programs provided by the School:

- ✦ Business Informatics
- ✦ E-Business;
- ✦ Information Security Management;
- ✦ Big Data Systems.

The School's activities are aimed at achieving greater integration into the global education and research community. A member of the European Research Center for Information Systems (ERCIS), the School cooperates with leading universities and research institutions around the world through academic exchange programs and participation in international educational and research projects.

DOI: 10.17323/1998-0663.2019.3.7.19

Fuzzy cognitive models for socio-economic systems as applied to a management model for integrated development of rural areas

Svetlana V. Podgorskaya [a]
E-mail: svetlana.podgorskaya@gmail.com

Aleksandr G. Podvesovskii [b]
E-mail: apodv@tu-bryansk.ru

Ruslan A. Isaev [b]
E-mail: Ruslan-Isaev-32@yandex.ru

Nadezhda I. Antonova [a]
E-mail: antonova_nadezhda@bk.ru

[a] Federal Rostov Agricultural Research Centre
Address: 52, Sokolov Avenue, Rostov-on-Don 344006, Russia

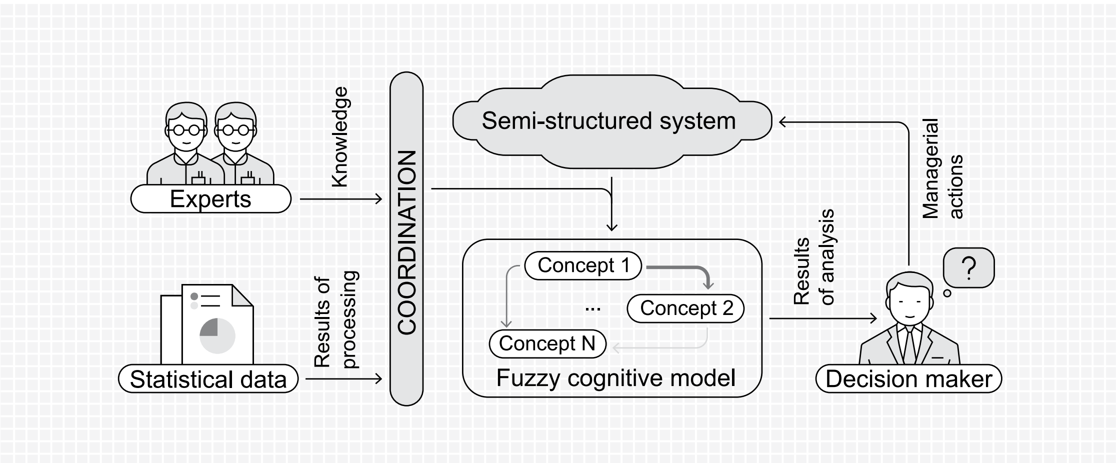
[b] Bryansk State Technical University
Address: 7, 50 Let Oktyabrya Avenue, Bryansk 241035, Russia

Abstract

The paper is devoted to fuzzy cognitive modeling, which is an effective tool for studying semi-structured socio-economic systems. The emphasis is on the process of developing (identification) fuzzy cognitive models, which are the most complex and critical stage of cognitive modeling. Existing identification methods are classified as either expert or statistical, depending on the source of information used. Typically, when constructing fuzzy cognitive models of semi-structured systems, the system under consideration possesses both quantitative (measurable) factors and factors of a relative, qualitative nature. While statistical data on the quantitative factors may be available, the only available source of information on the qualitative factors is expert knowledge.

However, each of the existing identification approaches focuses on just one source type, either expert or statistical. Thus, it is crucial to develop a more general approach to the development of fuzzy cognitive models for semi-structured systems to ensure reliable and consistent results by coordinated processing of information of both expert and statistical origins. We developed such an approach based on several identification methods with the subsequent coordination of intermediate results. To demonstrate the proposed approach, we applied it to a management problem of integrated development of rural areas. The fuzzy cognitive model we obtained can be used to predict the state of rural areas depending on initial trends and managerial actions, as well as to search and analyze effective managerial strategies for their development.

Graphical abstract



Key words: cognitive modeling; fuzzy cognitive model; identification of a cognitive model; pairwise comparison method; regression analysis; socio-economic system; integrated development of rural areas.

Citation: Podgorskaya S.V., Podvesovskii A.G., Isaev R.A., Antonova N.I. (2019) Fuzzy cognitive models for socio-economic systems as applied to a management model for integrated development of rural areas. *Business Informatics*, vol. 13, no 3, pp. 7–19. DOI: 10.17323/1998-0663.2019.3.7.19

Introduction

For socio-economic and other humanitarian systems, prediction and development of managerial strategies are complicated due to the following considerations:

✦ the multidimensional nature of the processes in the systems;

✦ the complexity of the connections between the processes, as well as their temporal variability;

✦ lack of quantitative information on dynamics of the processes.

This means that we can classify social, economic, and other similar systems as semi-structured. It can be complicated or even impossible to study and manage such systems based on analytical models that describe correlations between the input or output parameters. However, we can use models based on expert information, experience, judgement and intuition.

Semi-structured systems are often modeled using the cognitive approach. According to this approach, managerial solutions should be based on formal models and methods which use human cognitive abilities (perception, imagination, knowledge, understanding, and explanation) [1]. The term "cognitive modeling" refers to cognitive approach-based methods of structure and target as well as imitation modeling of systems. In other words, cognitive modeling allows us to study a system's structure and operation by analyzing its cognitive model.

Fuzzy logic is often used in cognitive modeling as a mathematical tool. There are a number of cognitive models which are based on fuzzy cognitive maps (FCM). A detailed review of these models is presented in [2]. In particular, Silov's FCMs, which are the extensions of signed cognitive maps [4], were introduced in [3] and have proved successful in applied problems of modeling and analysis of semi-structured systems.

In this paper, we present an approach to the development of cognitive models of semi-structured systems based on expert and statistical information. We also present an application of the method to a management model for integrated development of rural areas.

1. Key concepts in cognitive modeling

A cognitive model is based on a formal representation of the cause and effect links between the factors that describe a system. The system is represented as a cause and effect network (cognitive map), which has the following form:

$$G = < E, W >, \qquad (1)$$

where $E = \{e_1, e_2, ..., e_n\}$ – a set of factors (concepts);

W – a binary relation on the set E, which determines connections between its elements.

Concepts can reflect both absolute and measurable system characteristics (such as population or income), as well as relative qualitative parameters (popularity, competitiveness, etc.)

If changes in the e_i concept cause changes in the e_j concept, we say that ei influences e_j and denote this as $(e_i, e_j) \in W$ or $e_i W e_j$ (hereinafter $i, j = 1, ..., n$, where n is the number of concepts). An influence is called positive when an increase in the e_i state leads to an increase in the e_j state, and negative otherwise.

When we develop a fuzzy cognitive model, we assume that the intensity of mutual influence of the concepts can vary. In this case, W is determined as a fuzzy relation, and the corresponding cognitive map is called a fuzzy cognitive map (FCM).

Silov's FCM is an example of such a model, where the W relation is given as a set of numbers w_{ij} that reflect the direction and intensity level (weight) of the influence between the concepts e_i and e_j. The following assumptions are made:

a) $-1 \le w_{ij} \le 1$;

b) $w_{ij} = 0$ if e_i does not influence e_j;

c) $w_{ij} = 1$ for the maximum positive influence of e_i onto e_j;

d) $w_{ij} = -1$ for the maximum negative influence of e_i onto e_j;

e) w_{ij} takes other permissible values for other values of the intensity.

Such an FCM can be represented as an oriented weighted graph, where vertices and edges correspond to the concepts and cause and effect relations, respectively. Weights of the edges are determined by the corresponding w_{ij} values. The W relation can be represented as a $n \times n$ matrix (where n is the number of concepts in the FCM), which is called a cognitive matrix and is an adjacency matrix for the cognitive graph.

The first stage in fuzzy cognitive modeling is to create an FCM for the system under consideration. This can be done by using data from experts, or by analyzing the available statistical information. The next stage is the modeling, and the problems here can fall into the following two types:

✦ static (structure and target) analysis, where we look for the concepts that influence the modeling the most, reveal contradictions between the goals, analyze feedback loops, etc.;

✦ dynamic (scenario) analysis, aimed at the prognosis of the system's state under various managerial actions, as well as at generation and selection of optimal actions that will bring the system into the desirable state (various models of impulse processes can be used to describe the system's dynamics [5]).

The results of a modeling are usually represented by tables and plots. To be understood by an expert, the results must be interpreted using natural language and understandable terminology [6].

2. Identification of the parameters for a fuzzy cognitive model: existing approaches

The stage where we weight the links between the concepts is called a parametric identification; this is one of the most complex and important stages in cognitive modeling. The more reliable the results of this stage are, the more reliable is the final cognitive model.

Classification of the weighting methods is presented in *Figure 1.*

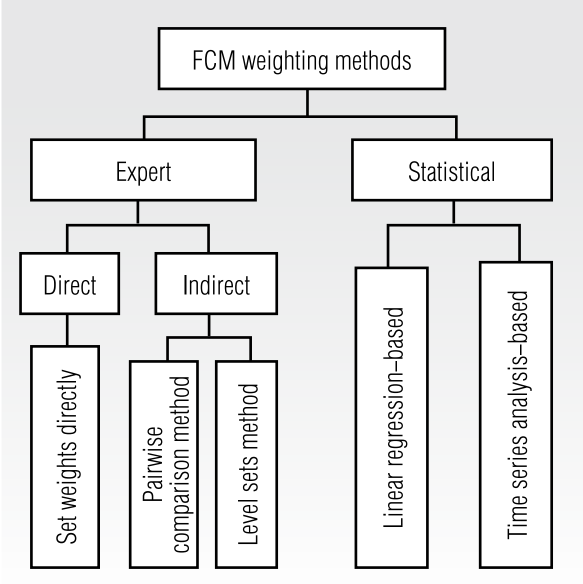


Fig. 1. Classification of the methods of FCM parametric identification

Weights for an FCM are usually determined by an expert method, either direct or indirect.

In the direct methods, the weights are directly specified by an expert [7]. This is the simplest way, but the results can be unreliable and unjustified due to the human factor. Indirect methods are less subjective, and the weighting problem can be represented as a series of less complicated subtasks. Some examples are the Saaty's pairwise comparison method [8], Yager's level sets method [9], as well as authors' modifications of these methods that improve the FCM's efficacy [10, 11].

As we have mentioned above, some concepts may reflect qualitative characteristics of the system being researched. If there are statistical data on the values of these characteristics, we can use them to weight the links between the concepts. In this case, expert estimates may or may not be used. Therefore, to identify the FCM's parameters, we can additionally use statistical methods [12, 13].

Let's consider the two methods used during the experimental stage of the research.

3. Pairwise comparison method for weighting the links in a fuzzy cognitive model

This chapter is based on the authors' research [10]. The approach proposed there is used in the present paper as part of a more general approach for FCM development.

When experts apply the pairwise comparison method for parametric FCM identification, they consider a certain concept *A* pairwise with all the concepts linked to it. Concepts that influence the concept *A* are considered separately from the concepts that are influenced by it. Concepts whose influences have different signs are also considered separately. For each of the pairs, the concept that influences the most gets selected, and the link with this concept is assigned a bigger weight. As a result, we obtain a pairwise comparison matrix *D*, where every element d_{ij} reflects the ratio of the links between the concepts e_i and e_j. In this matrix $d_{ii} = 1$ and $d_{ij} = 1/d_{ji}$.

To formalize d_{ij} estimations, the scales presented in *Table 1* may be used (the alternative scale is introduced and justified by the authors in [10]).

Table 1.

Values for two appropriate scales

Verbal description	classical scale	alternative scale
No advantage	1	9/9
Almost no advantage	2	9/8
Slight advantage	3	9/7
Almost considerable	4	9/6
Considerable advantage	5	9/5
Almost clear advantage	6	9/4
Clear advantage	7	9/3
Almost absolute advantage	8	9/2
Absolute advantage	9	9/1

Next, the matrix *D* must be checked for consistency. This is done by calculating the consistency index CI, and consistency ratio CR:

$$CI = \frac{\lambda_{\max} - n}{n - 1}, \quad (2)$$

$$CR = \frac{CI}{CIS}, \quad (3)$$

where CIS — an experimentally obtained estimation for the expected value of the consistency;

$\lambda_{\max}$ — the largest eigenvalue of the matrix *D*;

n — the dimension of the matrix *D*.

If CR values are greater than 0.1, the matrix *D* is inconsistent. In this case the model needs to be re-evaluated.

The weights vector W is calculated using the eigenvector of the matrix *D* at $\lambda_{\max}$:

$$W = \lambda_{\max} W. \quad (4)$$

Lastly, the vector W must be normalized by dividing all its elements by the maximum element and multiplying by the strength coefficient $k \in (0, 1]$. This coefficient is usually set by an expert and reflects the weight of the link.

4. Weights of the links based on regression analysis and elasticity coefficients

In [12, 14], the authors presented a method that weighs FCM links based on statistical data. The method includes the following steps:

1. development of a linear regression model (pair or multiple) using the available statistical data;
2. estimation of the model and importance of its parameters;
3. check for multicollinearity (in case of multiple regression) and its elimination;
4. using elasticity coefficients as the base for the links' weights:

$$E_i = b_i \frac{\overline{x_i}}{\overline{y}}, \quad (5)$$

where b_i is a regression coefficient in the linear regression model, $\bar{x}_i$ and $\bar{y}$ are the sample means;

5. normalization of the elasticity coefficients (to obtain the links' weights in the range [–1, 1]) using a sigmoid normalization function:

$$S(E_i) = \frac{1 - \exp(-bE_i)}{1 + \exp(-bE_i)}, \qquad (6)$$

where the coefficient b is the slope of the function.

One of the ways to obtain the value of b is to set the elasticity coefficient $E_0 > 0$ in such a way that the link's weight equals a certain value α. Then the coefficient b can be obtained from the following formula:

$$b = -\frac{1}{E_0}\ln\frac{1-\alpha}{1+\alpha}, \qquad (7)$$

where E_0 – the elasticity coefficient;

α – the link's weight (both are set by an expert).

5. Development of fuzzy cognitive models using coherent processing of expert and statistical information

Dealing with semi-structured systems (especially socio-economic ones), we typically have both quantitative (measurable) and relative, qualitative factors. We may have statistical data on quantitative factors, but we can only rely on expert knowledge when qualitative factors are concerned. However, every identification method available now uses a single type of information – either expert or statistical information.

Therefore, it is very important to develop a more general approach to fuzzy cognitive models of semi-structured systems which could use coherent processing of both types of information and ensure reliable and consistent results.

Such an approach may be based on combined use of available identification methods (several of which are developed by the authors) with subsequent coordination of the results. This means that to identify the links between the concepts we can use statistical methods when we have statistical data, and expert methods otherwise. For the results to be reliable and consistent, we need to coordinate the outcomes of both approaches.

Let us briefly describe some possible ways to coordinate the results of processing the expert and statistical information.

1. Statistical estimations may be used to improve and supplement the expert results. For example, for expert weighting by the pairwise comparison method, we need to set a strength coefficient for the links, and we can do this by using the available statistical estimations. This approach is preferable when statistical data are more reliable than the expert estimations.

2. As opposed to the above, expert estimations may be used to normalize statistical data, in particular, to identify a normalization function. This approach should be applied when expert estimations are highly reliable.

3. In some cases, statistical estimations may be used to coordinate separate sets of expert estimations.

4. Expert estimations may be used to coordinate statistical estimations for different parts of an FCM, as well as estimations obtained on the basis of statistical data of different types.

This approach is presented below as applied to the development of a fuzzy cognitive model for managing integrated development of rural areas.

6. Fuzzy cognitive model for managing integrated development of rural areas

Rural areas as research and management subjects are highly dynamical, and the processes in them are multidimensional. Many elements and links in this system are not fully researched and can be described only qualitatively. We also lack information on the system's behavior. The processes in the system vary over time, and they are typically nonlinear. Therefore, rural areas can be considered as semi-structured systems.

Hence, we propose to use a fuzzy cognitive model for studying the development of rural areas and finding effective managerial solutions for their sustainability.

We developed such a model using the IGLA decision support system. This system was developed by research at the "Information technology and software" department, Bryansk State Technical University [15].

The only way to obtain complete and reliable information on the structure and development of a rural socio-economic system is to use proper data collection technologies. So, we asked ten experts on rural socio-economic development to fill out a specially designed questionnaire. In addition, we researched Russian monographs on sustainable management of rural areas [16–23].

We concluded that in a cognitive model for managing integrated development of rural areas it is feasible to use eleven most important concepts. These concepts can be divided into four blocks (*Table 2*).

The next stage in the development of a cognitive model is finding links between the concepts. The experts used a cognitive map to reveal cause and effect links between the concepts and their influence type (either positive or negative).

The selected concepts are qualitative and quantitative elements of the system. Some of the quantitative elements can be described by statistical parameters, and we can estimate the strength of mutual influence by using statistical methods. This can essentially increase the model's objectivity and validity.

Table 2.

Concepts in a cognitive model for managing integrated development of rural areas

Institutional block
1. Development of market infrastructure (tax, credit, budget, and innovational politics)
2. Development of rural self–government
Socio-demographic block
3. Average population per year
4. Unemployment level
5. Development of the social sphere
Economic block
6. Per capita income
7. Agricultural production
8. Development of small and medium businesses
9. Investments in fixed capital
10. Economic diversification level
Ecological block
11. Negative influence on the environment

However, the concepts from the institutional and ecological blocks cannot be quantified, hence it is impossible to use statistical methods to estimate their influence on other concepts. Therefore, the influence was determined by experts. To obtain and process the expert knowledge, we used an authors' modification of the pairwise comparison method, where an alternative scale for the preference estimation is applied [10].

The cognitive model also includes the following concepts from the socio-demographic and economic blocks: population, unemployment level, agricultural production, per capita income and investments in fixed capital. These concepts reflect the parameters that are described statistically. In this case, to reduce the influence of experts' subjectivity and to increase reliability of the results, we processed the Russian Statistics Committee's data for the years 2000–2017 using the authors' method [12]. This method is based on the pair and multiple regressions [24].

Since the cognitive model includes both cost and nonmonetized parameters, we have to exclude the influence of inflation. The following indices for the years 2000–2017 were used as the deflators:

✦ consumer price index for 'Per capita income';

✦ agricultural produce price index for the 'Agricultural production';

✦ annual inflation index for the 'Investments in fixed capital'.

Based on the available statistical data, we set four regression equations that allow us to estimate five links in the cognitive model:

1). Influence of income on the population;

2). Influence of the agricultural production on the unemployment rate;

3). Influence of the unemployment rate on the incomes of the population;

4). Influence of the population and investments in fixed capital on the agricultural production (multiple regression).

The regression models we developed allow us to make the following conclusions:

✦ for the first case, the correlation is very high. However, the regression coefficient and the corresponding links have opposite signs (the sign of the link was set by experts). In our opinion, this means that there are other factors that negatively influence the demographics in rural areas, such as population makeup by sex and age, the number of reproductive age women in rural areas, and negative natural and migration growth of the rural population. These factors were not used as concepts in the model;

✦ the regression coefficient between the population and production is negligible, based on the *t*-statistic, and the corresponding influence is insignificant;

✦ for the remaining three links, the regression coefficients have correct signs (which coincide with the signs set by the experts), and they are significant, based on the *t*-statistic (influence of the production on unemployment, influence of the unemployment on income, and influence of the investments on production).

The values obtained for the regression coefficients and elasticity coefficients are presented in *Table 3.*

For the results to be consistent, we now have to coordinate the results obtained by processing statistical data and by the expert approach. To do this, let us consider how the concept "Investments in fixed capital" influences the concept "Agricultural production". Using the aforementioned method of identification of the normalization function, we calculate the value for the parameter b so the elasticity coefficient after the normalization will coincide with the experts' estimation of 0.8:

$$b = -\frac{1}{0.809}\ln\frac{1-0.8}{1+0.8} = 2.71. \tag{8}$$

Normalization of the elasticity coefficients allows us to calculate the weights of the links between the concepts. They are also presented in *Table 3.*

The final results of the parametric FCM identification are presented in *Table 4.*

The weights calculated from statistical data both supplement the experts' results and increase the objectivity and justification of the parameters in a fuzzy cognitive model.

A graphic FCM representation obtained by the visualization subsystem of the IGLA decision support system is presented in *Figure 2.*

Table 3.

Weights of some FCM links calculated based on statistical data

Influence	Regression coefficient	Elasticity coefficient	Weight of the link
‘Agricultural production’ on the ‘Unemployment rate’	–0.00236	–0.306	–0.39
‘Unemployment rate ’on the ‘Per capita income’	–1272.176	–3.378	–0.99
‘Investments in fixed capital’ on the ‘Agricultural production’	6.975	0.809	0.8

Table 4.

Fuzzy cognitive matrix

Influencing concepts	Influenced concepts										
	1	2	3	4	5	6	7	8	9	10	11
1	0	0.434	0	0	0	0	0	0.654	0.512	0	0
2	0	0	0	0	0.53	0	0	0.471	0	0.417	0
3	0	0	0	0	0	0	0.178	0	0	0	0
4	0	0	0	0	–0.55	–0.99	0	0	0	0	0
5	0	0	0.353	0	0	0	0	0	0	0	0
6	0	0	0.118	0	0.353	0	0	0	0	0	0
7	0	0	0	–0.39	0	0	0	0	0	0	0.564
8	0	0	0	–0.69	0	0.527	0.8	0	0	0	0.446
9	0	0	0	0	0	0	0.78	0.691	0	0	0.418
10	0	0	0	–0.52	0.531	0	0	0.393	0	0	0
11	0	0	–0.6	0	0	0	0	0	0	0	0

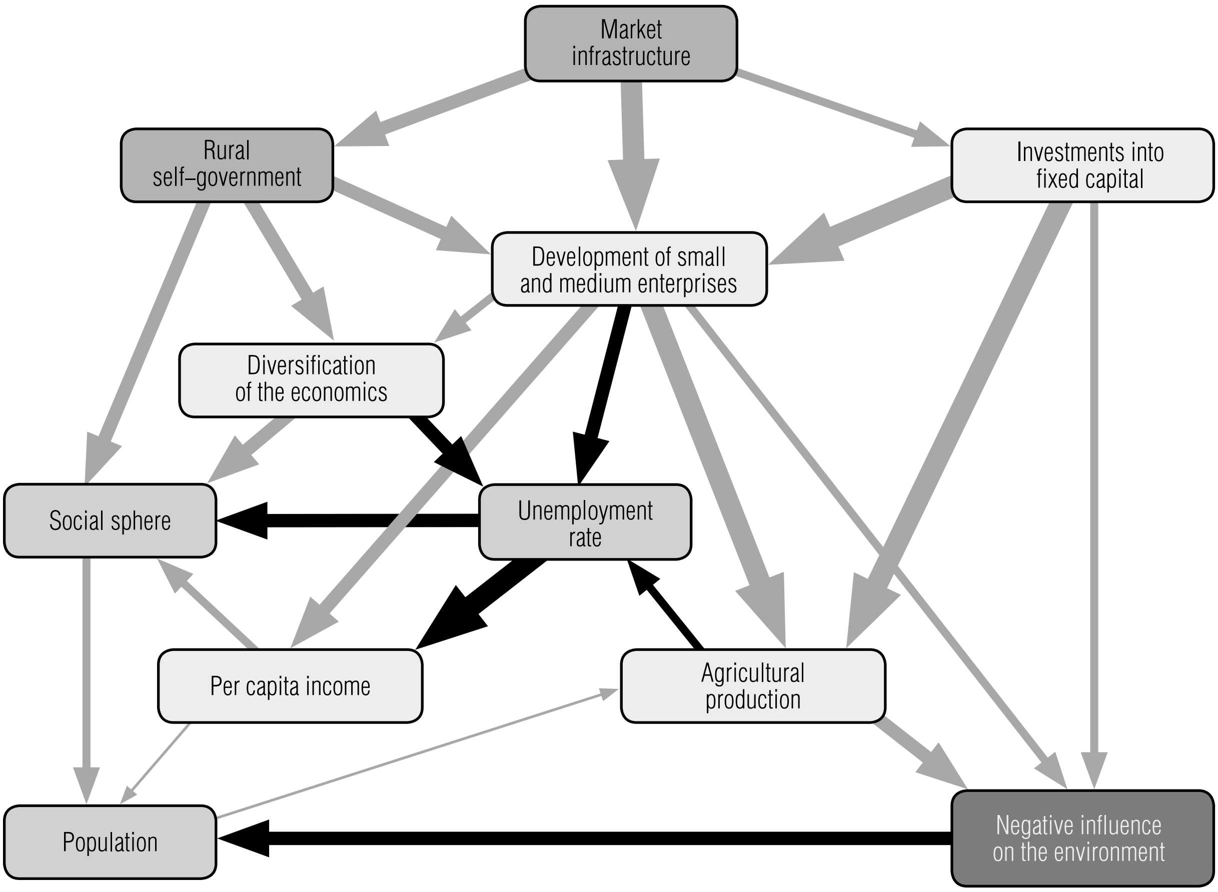


Fig. 2. Fuzzy cognitive model for rural integrated development management

Different colors correspond to different blocks (*Table 2*). Several other ways to visualize an FCM are described in [25].

The model we obtained may be used for the following purposes:

✦ to reveal the factors that influence the rural areas' development the most which must be addressed when taking managerial decisions;

✦ to predict the state of rural areas given initial tendencies and possible managerial decisions (or without them);

✦ to find effective strategies for rural areas' management that could help reach certain goals.

All these problems can be solved by structure and target and scenario analysis of the cognitive model so developed. These stages of cognitive modeling lie outside the scope of this paper and will be addressed in future research.

Conclusion

We present a method to develop fuzzy cognitive models for socio-economic systems containing both quantitative and relative, qualitative factors. It is assumed that statistics on the quantitative factors may be available. In this case, these factors may be processes using the authors' methods for FCM identification.

The method also utilizes an expert's knowledge (or that of a group of experts) obtained and processed using expert methods for FCM parametric identification, such as the Saaty's pairwise comparison method or Yager's level sets method.

The final important step of the method is to coordinate the intermediate results obtained by both expert and statistical methods. This step ensures that the final results are highly reliable and consistent. This, in turn, ensures a high quality of cognitive models as well as high efficiency of the subsequent managerial decisions.

The proposed approach is applied to the problem of managing integrated development of rural areas. The FCM obtained may be used to predict development of rural areas given different initial tendencies and managerial decisions as well as to find and analyze effective managerial strategies.

It is important that further research be devoted to structure and target, as well as scenario analysis of the model developed.

Another promising research direction is to update the approach presented to support group expertise during an FCM development. Methods that would permit us to evaluate the consistency of a group's opinions and to reach the necessary level of consistency need to be developed. ■

References

1. Avdeeva Z.K., Kovriga S.V., Makarenko D.I. (2007) Cognitive modeling for solving problems of managing semi-structured systems (situations). *Large-Scale Systems Control*, no 16, pp. 26–39 (in Russian).
2. Borisov V.V., Kruglov V.V., Fedulov A.S. (2012) *Fuzzy models and networks*. Moscow: Gorjachaja Linija – Telekom (in Russian).
3. Silov V.B. (1995) *Strategic decision making in fuzzy conditions*. Moscow: INPRO-RES (in Russian).
4. Roberts F.S. (1976) *Discrete mathematical models, with applications to social, biological, and environmental problems*. N.J.: Prentice-Hall.
5. Isaev R.A., Podvesovskii A.G. (2017) Generalized model of pulse process for dynamic analysis of Sylov's fuzzy cognitive maps. CEUR Workshop Proceedings of the *International Conference on Information Technology and Nanotechnology, Samara, Russia, April 25–27, 2017,* no 1904, pp. 57–63.
6. Podvesovskij A.G., Lagerev D.G., Korostelev D.A. (2009) Application of fuzzy cognitive models for alternatives set generation in decision problems. *Bulletin of Bryansk State Technical*, vol. 4(24), pp. 77-84 (in Russian).
7. Litvak B.G. (2004) *Expert technologies in management*. M.: Delo (in Russian).
8. Saaty T.L. (1978) Exploring the interface between hierarchies, multiple objectives and fuzzy sets. *Fuzzy Sets and Systems*, vol. 1, no 1, pp. 57–68.
9. Yager R.R. (1982) Level sets for membership evaluation of fuzzy subset. *Fuzzy Sets and Possibility Theory: Recent Developments* (ed. R.R. Yager). NewYork: Pergamon, pp. 90–97.
10. Isaev R.A. (2016) Modified pairwise comparison method for expert estimation of a fuzzy cognitive model parameters. *Modern Information Technology and IT Education*, vol. 12, no 2, pp. 35–42 (in Russian).
11. Isaev R.A., Podvesovskij A.G. (2017) Evaluation of expert judgements consistency when constructing a membership function of fuzzy set using the method of level sets. *Modern Information Technology and IT Education*, vol. 13, no 3, pp. 9–15 (in Russian).
12. Podvesovskij A.G., Isaev R.A. (2016) Application of multiple regression analysis for parametric identification of fuzzy cognitive models. Proceedings of the *4th International Conference on Information Technologies for Intelligent Decision Making Support, ITIDS'2016, Ufa, Russia, May 17–19, 2016*, vol. 2, pp. 28–33 (in Russian).

13. Isaev R.A., Podvesovskii A.G. (2018) Application of time series analysis for structural and parametric identification of fuzzy cognitive models. CEUR Workshop Proceedings of the *International Conference on Information Technology and Nanotechnology, Samara, Russia, April 24–27, 2018*, no 2212, pp. 119–125.

14. Averchenkov V.I., Kozhuhar V.M., Podvesovskij A.G., Sazonova A.S. (2010) *Monitoring and forecasting of regional needs for specialists of higher scientific qualification*. Bryansk: BSTU (in Russian).

15. Korostelev D.A., Lagerev D.G., Podvesovskij A.G. (2008) Decision support system based on fuzzy cognitive models "IGLA". Proceedings of the *11th Russian Conference on Artificial Intelligence (RCAI–2008), Dubna, Russia, September 29 – October 3, 2008*, vol. 3, pp. 329–336 (in Russian).

16. Soldatova I.Yu. (2015) State innovative measures for sustainable rural development. Proceedings of the *Conference on Problems of Improving the Efficiency of Local Self-Government in the Context of Modern Reforms and Political Processes in Russia, Rostov-on-Don, Russia, April 7–8, 2015*, pp. 323–327 (in Russian).

17. Gataulina E.A. (2015) Development of small and medium businesses in agriculture – the key to sustainable rural development. *Scientific and Methodological Fundamentals of Sustainable Rural Development, Ser. "Proceedings of VIAPI"*, pp. 70–88 (in Russian).

18. Ovchinceva L.A., Kotomina M.A. (2013) The new role of the regions in sustainable rural development. *Nikonov Readings*, no 18, pp. 248–252 (in Russian).

19. Podgorskaja S.V. (2018) Increasing rural areas workforce capacity as the basis of their sustainable development. Proceedings of the *International Conference on Agriculture in Russia and Abroad: Modern Challenges of Economic Development, Rostov-on-Don, Russia, October 3–4, 2018*, pp. 336–341 (in Russian).

20. Podgorskaja S.V. (2008) *Conceptual model of investment support for village social infrastructure development (a case study of Rostov region)*. PhD Thesis. Rostov-on-Don (in Russian)

21. Podgorskaja S.V., Tarasov A.S. (2017) Assessment of marketing performance in an agricultural educational organization. *Polythematic Online Scientific Journal of Kuban State Agrarian University*, no 125, pp. 247–257 (in Russian).

22. Bahmatova G.A. (2014) Differentiation of Rostov region territories according to sustainable development level taking into account their geographical location. *Scientific Journal of Russian Research Institute of Land Improvement Problems*, no 1 (13), pp. 261–270 (in Russian).

23. Tarasov A.S., Antonova N.I., Markina E.D., Bahmatova G.A. (2017) Comprehensive forecast generation for sustainable development of engineering, market and environmental infrastructure of rural areas. *Problems of Regional Agricultural and Industrial Complex Development*, no 3 (31), pp. 147–152 (in Russian).

24. Kremer N.Sh., Putko B.A. (2010) *Econometrics*. Moscow: UNITI-DANA (in Russian).

25. Podvesovskij A.G., Isaev R.A. (2018) Visualization metaphors for fuzzy cognitive maps. *Scientific Visualization,* vol. 10, no 4, pp. 13–29.

About the authors

Svetlana V. Podgorskaya

Cand. Sci. (Econ.), Associate Professor;

Leading Researcher, Federal Rostov Agricultural Research Centre,
52, Sokolov Avenue, Rostov-on-Don 344006, Russia;

E-mail: svetlana.podgorskaya@gmail.com

ORCID: 0000-0001-8912-7865

Aleksandr G. Podvesovskii
Cand. Sci. (Tech.), Associate Professor;
Head of Informatics and Software Department, Bryansk State Technical University,
7, 50 Let Oktyabrya Avenue, Bryansk 241035, Russia;
E-mail: apodv@tu-bryansk.ru
ORCID: 0000-0002-1118-3266

Ruslan A. Isaev
Postgraduate Student, Informatics and Software Department, Bryansk State Technical University,
7, 50 Let Oktyabrya Avenue, Bryansk 241035, Russia;
E-mail: Ruslan-Isaev-32@yandex.ru
ORCID: 0000-0003-3263-4051

Nadezhda I. Antonova
Head of Department, Federal Rostov Agricultural Research Centre,
52, Sokolov Avenue, Rostov-on-Don 344006, Russia;
E-mail: antonova_nadezhda@bk.ru
ORCID: 0000-0003-3039-2837

DOI: 10.17323/1998-0663.2019.3.20.34

Approach to the organization of decision support in the formulation of innovative regional development strategies based on adaptive-simulation model

Lilia R. Chernyakhovskaya [a]
E-mail: lrchern@yandex.ru

Marsel M. Nizamutdinov [b]
E-mail: marsel_n@mail.ru

Vladimir V. Oreshnikov [b]
E-mail: VOresh@mail.ru

Alsouy R. Atnabaeva [b]
E-mail: Alsouy@mail.ru

[a] Ufa State Aviation Technical University
Address: 12, Karl Marx Street, Ufa 450008, Russia

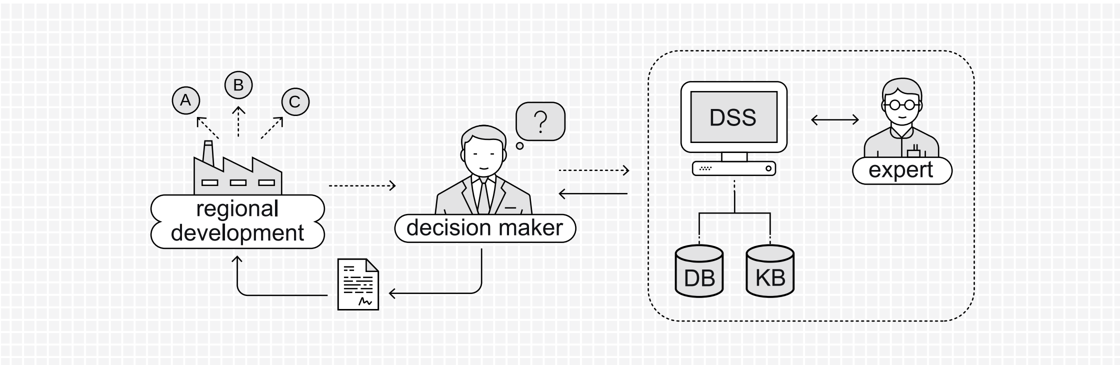
[b] Ufa Federal Research Center, Russian Academy of Sciences
Address: 71, Prospekt Oktyabrya, Ufa 450054, Russia

Abstract

This article deals with the formulation of a decision support system (DSS) in the field of regional development management. The review of existing approaches in this area presented here attests, on the one hand, to their diversity, and on the other hand allows us to draw conclusions about the need to address several methodological and practical issues of decision support in terms of innovative development of regions. Based on this, the goal of the research was to develop the concept of DSS to justify the parameters of an innovative development strategy for regional development based on adaptive mechanisms for coordinating the interests of economic agents.

The methodology of the study is based on the synthesis of various approaches in the framework of integration into the structure of adaptive simulation models of problem-oriented knowledge bases with the mechanism of logical inference, as well as intelligent technologies for processing semi-structured information used to find solutions in the process of shaping and adjusting the parameters for managing innovative development of a region. The result of the study is a theoretical justification for developing problem-oriented DSS, including a description of the interrelated stages that determine the main design features of this tool. In the framework of the study, a conceptual scheme for implementing DSS in the field of managing innovative development of regions is proposed, and the key functional blocks of the proposed tools are described. In addition, the place of existing tools in the structure of the regional development management system is determined, and we show the possibilities of their use in the formation of forecast-planned assessments of the development of the region, as well as in the evaluation of the effectiveness of alternative management actions. The proposed tools will expand the possibilities of applying the methods of management theory and decision support, intelligent information technology, economic and mathematical methods and modern computer simulation technologies for strategic planning of socio-economic systems of macro- and meso-level. In practice, the tools may be of interest to public authorities in solving problems in in the formulation of innovative regional development strategies for Russian regions, the formation of medium-term forecasts and the justification of the parameters of social, economic and budgetary policy.

Graphical abstract



Key words: decision support; innovative development; simulation modeling; adaptive modeling; regional management; forecasting and planning.

Citation: Chernyakhovskaya L.R., Nizamutdinov M.M., Oreshnikov V.V., Atnabaeva A.R. (2019) Approach to the organization of decision support in the formulation of innovative regional development strategies based on adaptive-simulation model. *Business Informatics*, vol. 13, no 3, pp. 20–34.
DOI: 10.17323/1998-0663.2019.3.20.34

Introduction

The regional socio-economic system by virtue of its nature is a complex, multi-aspect and multi-purpose system, including many elements and heterogeneous relationships in all spheres of society. Ongoing economic processes have a direct influence on the stability and sustainability of the development of the region as a whole. Moreover, in the context of global competition for various types of limited resources, the determining factor is the ability of regional authorities to facilitate the innovative potential of the population and enterprises. The formulation of appropriate strategies for innovative development needs in-depth comprehensive substantiation, determination of target reference points, a resource base and optimal ways to solve the assigned tasks. In this regard, an objective condition for the formation of an effective economic policy is the widespread use of scientifically sound management tools, which, in particular, include the tools for developing and substantiating the parameters of innovative development of the region. At the same time, accelerated economic growth should not contradict the more global task of achieving balanced development of both the real sector of the economy and the social sector of the region.

Research in the field of formulating economic and mathematical models and tools for managing the strategic development of complex social and economic systems and, in particular, planning innovative development to one degree or another, has been carried out in recent decades in most developed countries of the world. Currently many software products have been developed that allow us to solve the set of problems of forecasting regional development [1]. However, the vast majority of modern Russian technologies are still based on the prerequisites for the functioning of a planned economy and do not reflect current market conditions for conducting business in the Russian Federation. Moreover, existing technologies mainly reflect the inertial mechanisms of territorial development and are not aimed at solving urgent problems of ensuring accelerated economic growth based on the transition to an innovation-oriented development model.

Thus, there are objective prerequisites for improving the currently used simulation tools due to the wider use of simulation modeling, conducting scenario experiments, taking into account the objective possibilities and restrictions of the regions. In aggregate, this allows us to formulate specific mechanisms and set conditions for achieving socially significant development priorities.

1. Review of existing approaches

Analysis of the current state of research in the field of organizing decision support for the development of strategies promoting innovative development of the regions implies a comprehensive review of issues of strategic management, the theory of innovative economics and development of decision support systems (DSS). Separately, we should mention the issues of developing economic and mathematical models in relation to the problem under consideration.

These areas have a well-developed scientific base. In particular, the basic tenets of the theory of innovative economics were formed already by J. Schumpeter. A number of aspects are considered in studies of Kondratiev's long waves, diffusions of C. Freeman's innovations, and also the technological paradigm of S.Yu. Glazyev, C. Peres, D.S. Lvov, Yu.V. Yakovets et al. However, the majority of the available studies are not fully focused on quantitative assessment and forecast of the influence of factors of innovative development on the evolution of territorial social and economic systems.

In order to solve this problem, it seems necessary to use specialized tools, in particular, decision support systems (DSS) [2]. However, unfortunately, within the framework of the problems being studied, most of the existing technologies are focused on the tasks of managing the development of individual enterprises or activities. However, at the regional level, where issues of developing strategies for innovative development are being addressed, these systems do not find practical application. Certain works in this area (information and analytical DSS on managing innovative activity in regions [3], DSS on strategy for innovative development of regions [4], etc.) only partially reflect such an essential property of the regional system as adaptability of behavior of key economic entities [5], or have exclusively theoretical or industry-specific significance [6].

Various directions have also developed in the field of economic and mathematical modeling of social and economic processes (general economic equilibrium models, simulation modeling, probabilistic-statistical models, etc.). At the same time, the models used in practice are mainly mixed and combine the features of several different approaches [7]. The most famous foreign models in this area are the Wharton annual model of the US economy, the LIFT model, the Brookings model of the US economy and the model of US fiscal policy. Among Russian technologies, a special place is occupied by the modelling software systems “SIRENA” and “SIRENA-2” (Institute of economy and industrial engineering of the RAS Siberian branch), the modelling complex “POLIGON-2” (Novosibirsk State University), the “RIM” model (Institute for National Economic Forecasts of RAS), and the “Intersectoral Interactions Model” (Institute of the economy and scientific and technical progress forecasts, USSR Academy of Sciences), the CGE-model “RUSEC,” the “Econometric model of the Russian economy,” the “CGE-model of the social and economic system of Russia with integrated neural networks” (Central Economic and Mathematical Institute of RAS), the “Model of the Region of the Russian North” (Syktyvkar State University), the model “Governor,” the “Agent-based model of Moscow” (Central Economic and Mathematical Institute of RAS), etc. [8]. Many complex information and analytical systems, including the information and analytical complex “Prognoz” (CJSC “Prognoz”), are based on the use of simulation models [9].

Despite the advantages of the models presented, the parameters of innovative development of the regions are presented extremely incompletely in them [10, 11]. At the same time, many models, in particular, those based on econometric equations, do not reflect the adaptive properties of agents of the regional system, all of which causes a decrease in the accuracy of the forecasts obtained.

In this regard, the aim of this study is to develop the concept of a decision support system for the process of formulating strategies for innovative development of regions. To achieve this goal, it is necessary to solve a number of problems, including determining the conceptual framework for the development of DSS and considering the practical aspects of organizing the decision support process.

2. Conceptual framework and logic of DSS development

The set of methodological problems in the field of strategic planning of regional development determines the need to develop special management tools that will reflect the key principles of interaction of territorial units in the formation of an innovative economy and will quickly determine the possible consequences of decisions. A qualitative change is required in the methodological and instrumental base of the existing system for managing innovative development of regions through the integration of strategic management methods, economic and statistical analysis, mathematical modeling, decision support theory and simulation modeling.

In order to achieve this goal, first of all, it is necessary to consider the place of this tool in the regional management system. At the same time, the management of territorial development, in general, includes two components – strategic and operational. The level of strategic management sets priorities for the development of the region; its main parameters define goals for the operational level. The objective of operational management, on the other hand, is to solve current problems aimed at achieving certain particular parameters that characterize the strategic goals of the development of the region. A change in the parameter values of the regional social and economic system (RSES) is directly influenced by regulators (control parameters) and is the result of the activities of the subjects of operational management, as well as processes occurring within the system itself and in the external environment (*Figure 1*).

As a decision-maker in relation to the task of developing a strategy for innovative development of the region (indicated in *Figure 1* as the "Subject of Strategic Management"), the relevant government bodies should be considered: the composition and structure of the bodies differ slightly in different territorial entities of the Russian Federation. First of all, they include the Ministry of Economic Development and the Ministry of Industry and Innovation Policy.

In the course of solving the tasks of strategic management, the formulation and substantiation of long-term large-scale projects are carried out. Direct generation of a strategy for innovative development of a regional system entails the implementation of a number of stages, including the following:

✦ **Defining the goals of the strategy.** It should be noted that goals should be quantitatively measurable, have clear time characteristics, and be interconnected with more global parameters for the development of the regional economy. Therefore, as indicators characterizing the goals of implementing the innovative strategy for the

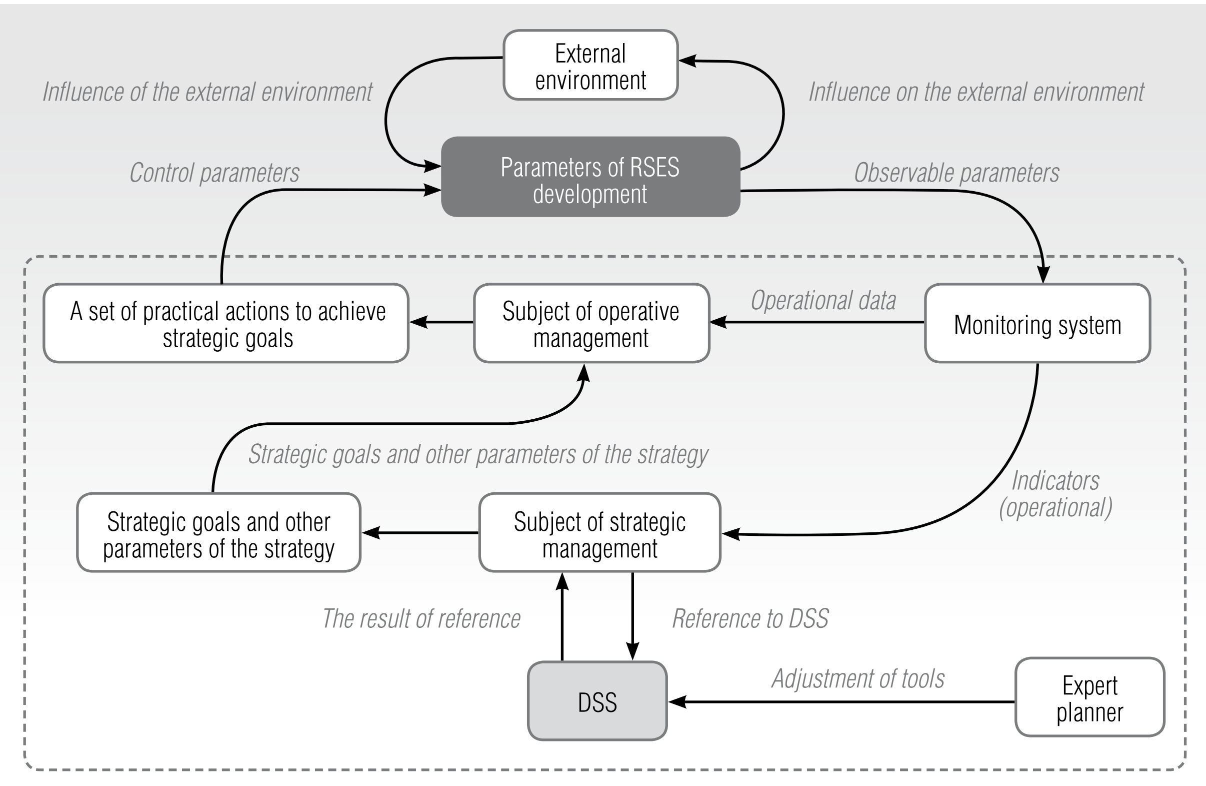


Fig. 1. The place of DSS in the management system for regional development

development of the region, we can have, for example, the share of innovative products, the region's GRP growth rate and the share of people employed in innovative types of economic activity;

✦ **Formulation of a list of indicators under consideration.** The list of these indicators is determined based on the afore-mentioned goals, the need to consider key factors that influence the innovative activity of economic entities, and the principles of the formation of a regional system model;

✦ **Analysis and identification of problems.** At this stage, the features of the innovative development of a region, its strengths and weaknesses, and limitations are determined;

✦ **Analysis of the influence of scenario parameters.** At this stage, the influence of scenario parameters on the level of innovative activity of economic entities is considered. Within the framework of the proposed approach, modeling the influence of external conditions differs only in the principle of determining the values of indicators, while for the economic agents themselves there is no fundamental difference;

✦ **Identification of controlled parameters and mechanisms for controlling them.** It should be noted that in the field of innovative development, management is carried out mainly by indirect methods, in particular, by creating favorable conditions for the formation and implementation of innovations;

✦ **Definition of development goals, a set of targets and their quantitative characteristics.** At this stage, using the methods of economic and mathematical modeling, the planned values of the previously considered indicators are substantiated;

✦ **Formulation of a set of measures aimed at achieving the goal.** The use of DSS at this stage provides the opportunity to take into account previously obtained results of state policy to enhance the innovative potential of a region. For these purposes, an appropriate precedent base may be applied;

✦ **Substantiation of the efficiency** of measures and assessment of the amount of necessary resources.

From the list presented, it follows that the task of substantiating the parameters of innovative development of the region is multi-faceted and must reflect the behavior of agents of the regional system. This statement of the problem determines the need for the synthesis of various approaches to research. A similar requirement is implemented when using DSS based on adaptive-simulation models (ASM) [12].

Within the framework of this study, the region's ASM is understood as a model based on the idea of consistent adaptation to change in the economic situation of not only the goals and behavior of individual agents (and, accordingly, their strategies), but also the management subsystem. It should be noted that the complex social and economic systems modeled in this case are also adaptive, which leads to the use of the ASM of the region to develop a strategy for its development.

In general, the theoretical substantiation for the development of such a problem-oriented decision support system should include the implementation of several successively interconnected stages that determine the main design features.

Within the framework of the conceptual stage, one must clearly define the basic principles and requirements for the development of decision support tools applicable to developing strategies for innovative development of regions.

The key principle of this study is a systematic approach, consisting in the need to analyze all the basic elements and interconnections of the regional system that form the basis for innovative development, as well as a description of their most significant properties and the determining factors for innovative development. At the same time, such principles as adaptability and consistency of decision-making procedures by agents and the regional management system in the framework of the implementation of measures of state social and economic policy are also significant. The principle of integrity requires consideration of various aspects that determine the degree to which the innovative potential of a region is implemented, including economic, social, institutional, organizational, technical and other factors. The principle of conformity with the objectives of the study predetermines limitations on the structure of the parameters under consideration and helps to avoid significant deviations associated with the diversity of life in society, including by focusing on the factors of innovative development of the region. The principles of scientific validity and practical applicability, complementing each other, provide orientation, on the one hand, on a scientific basis (including the previously mentioned technologies in the field of formulating, implementing and disseminating innovations), and on the other hand, on the tasks of economic development facing public authorities of the territorial entities of the Russian Federation. Another important principle is the comparability of data, due, inter alia, to the need to use heterogeneous and different-sized characteristics obtained from various sources and relating to different time periods.

It is also necessary to formulate a common research methodology, substantiate approaches, methods and technologies, as well as determine the concept of reflecting adaptive management schemes in the simulation model of the region's economic system (for example, using evolutionary procedures as part of decision support [13]). It should be noted that the research methodology reflects a systematic view of the region with the determination of the place, role and relationships of economic, social and innovative processes in the reproduction process, formalized on the basis of the adaptive-simulation model (ASM) of the region.

In the future, during the design and mathematical formalization of the simulation model, we see the integration of structural elements and functional relationships which are necessary to substantiate the parameters of the region's innovative development strategies. In this regard, on the basis of the methodology presented, one must develop logical, informational and mathematical models for the entire complex of economic agents under consideration. At the same time, it is necessary to develop a management subsystem designed to solve the problem of innovative development.

This subsystem, in our opinion, includes an indicative planning model and a set of control parameters that can be used to substantiate the parameters of economic development strategies by utilizing the innovative potential of the region. In addition, the development of an algorithm for classifying situations based on fuzzy logic methods [14, 15] and adjusting the indicative plan will make it possible to determine the necessary regulatory effects. This procedure is carried out to achieve the planned guidelines for the development of a region as part of adapting the functioning strategy of both the economic subsystem and the management subsystem.

Among the factors of innovative development, researchers distinguish the following:

✦ strategic (the quality of strategic planning, the level of formation of the national innovation system);

✦ economic (the availability of financial resources of organizations, the investment situation, the variety of forms of management);

✦ organizational and management (the level of training in the field of innovation management, the availability of scientific and research and technology potential, the level of development of innovative infrastructure);

✦ technical and technological (deterioration of fixed assets, complexity and features of production, development and implementation of innovations in economically strong businesses);

✦ social (the prestige of various types of labor activity, the lack of highly qualified specialists);

✦ institutional (the state of the regulatory framework, the level of state support, financing of scientific and technical programs);

✦ market (demand for an innovative product, the level of development of market infrastructure).

Turning to the direct development of DSS software, it is necessary to systematically integrate the above models and algorithms into a single tool. In particular, we are talking about databases on the state of economic agents and the external environment, the user interface, the control module, the system for outputting the results, etc. This approach is generally consistent with research by Russian and foreign authors [16]. It also requires the development of methodological support for the use of DSS when substantiating the parameters of the strategy for innovative development of a region. These methods should contain a description of the process of developing an indicative plan, determining the parameters of the regulatory impact, as well as conducting computational experiments.

Within the framework of the practical testing of the proposed tools, there are plans to implement a set of scenario experiments using the ASM of a region to substantiate the most appropriate strategy for innovative development in the mid-term.

The decision support tools thus developed differ from the information and analytical systems used today in that they use ASM, which is focused on substantiating the parameters of the strategy for innovative development of the economy. This class of models makes it possible to holistically reflect adaptive properties in the behavior of agents and the reaction of the control subsystem as part of the implementation of strategies. The proposed model is distinguished by the presence of integrated circuits of interaction of subjects in various situations, which provides the opportunity to adapt their resource strategies in the process of determining agreed development goals.

3. Practical aspects of the organization of decision support using the adaptive-simulation model

From a practical standpoint, the set of tasks to be solved with the use of the decision support tools we developed is determined based on the sequence of stages in the development of a strategy for innovative development of a region and the need to take into account the advantages and disadvantages of individual approaches [17]. In this regard the following are among the basic tasks:

✦ collection and storage of factual information, including structured (in the form of decision-making rules) and weekly structured information (in the form of ontologies and precedent bases);

✦ assessment of observed and predicted situations using a problem-oriented knowledge base with a logical inference mechanism;

✦ making a forecast of changes of the situation in certain conditions on the basis of an adaptive-simulation model;

✦ substantiation of a set of recommendations for the development and adjustment of key parameters to manage the region's innovative development.

Taking into account the features of the functioning and development of the control object, as well as taking into account the goals and objectives of the development of tools, it is advisable to use a simulation model reflecting the adaptive properties of the behavior of economic agents as its core. At the same time, the characteristics of the agent in question and decision-making rules based on both their own experience and the analysis of the behavior of other agents are important. The solution to these problems is ensured by using the capabilities of the ASM of the region, which integrates three key levels within the framework of a single design – economic agents, government, and the macro-environment.

Of the entire set of parameters describing the external environment of the agent (including the characteristics of contractors), only a limited part of them is perceived by the economic agent. In addition, information on the state of the agent is also available to him only partially. Based on this, decisions are made by the agent on the basis of limited knowledge (*Figure* 2). Based on this information, as well as taking into account the existing rule bases and knowledge base, the agent classifies situations and makes decisions. Simultaneously with the decision itself, the agent forms an image of the desired result of the implementation of the strategy. Moreover, the decision within the framework of this study also means the possible refusal of the agent to take any actions. The implementation of the decision and, therefore, the consequences of this decision are influenced by the parameters of the external environment, including the active reaction from the contractors. Thus, the consequences of the decision to one degree or another differ from the expected results. In the course of comparison by the agent of the expected and obtained results, either the rules and knowledge of the agent are confirmed (if the results coincide), or a correction of its rules and knowledge bases takes place.

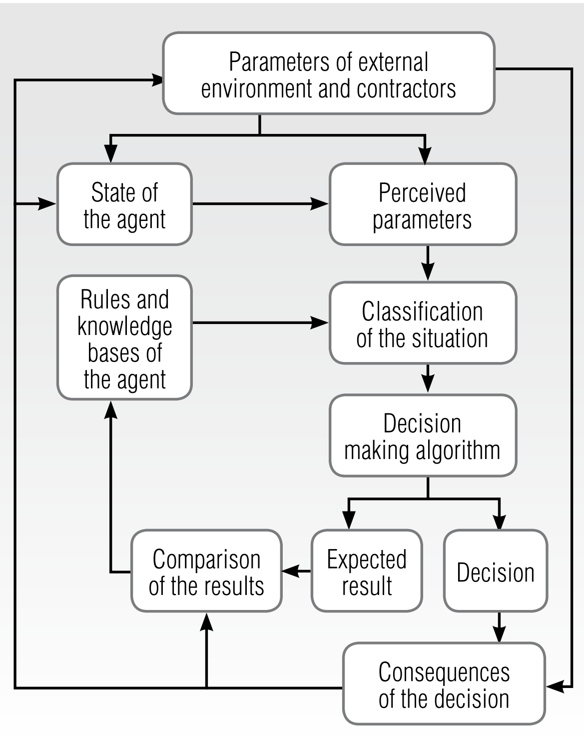


Fig. 2. Aggregate scheme of agent's performance

Moreover, in subsequent cycles of the situation analysis and decision-making, the agent in question is guided by a new set of rules and knowledge. In accordance with this algorithm, the agent is being trained and adapted. That is, having a "memory," he is able not only to adjust the quantitative parameters of the decision, but also to supplement the existing rule base in order to identify qualitatively new situations and develop appropriate decisions.

Another way of adjusting the rule base and agent knowledge base is the "exchange" of rules and knowledge with contractors. Such a mechanism does not imply a direct comparison of the expected and observed results obtained during decision-making. Through this process, the most important properties of agents in the social and economic system are implemented, mainly: adaptability and learning ability. The reflection of these properties allows us to model regional processes not from the position of a mechanistic approach, but reflecting the behavioral characteristics of various agents.

Within the framework of the proposed model, it seems appropriate to consider agents of three types – "Population", "Enterprises" and "State and municipal government bodies." In an aggregated form, the logical model of the regional system is presented in *Figure 3*. At the same time, the formation of an agent-based model implies the availability of many parameters characterizing each agent of the model belonging to one or another type. So, for the agent "Population," not only indicators such as "Sex" and "Age" can be considered, but also "Level of Education," "Tendency to saving," etc. Focusing on the innovative component of the economic development of the region, it should be noted that each of the agents of the "Enterprise" type has such characteristics as "Identifier of economic activity," "Volume of shipped products," "Profitability of production," etc., each of which anyhow affects the share of innovative products. Similar studies confirm the numerosity of factors that influence the level of innovative development of the region In particular, [18] demographic, economic, financial, labor, social, investment and some other factors were singled out in the research, the total number of which is 119 units.

The management of innovative development within the framework of the proposed model is an adjustment of the parameters that influence the behavior of agents of the regional system. The most important parameters in this case are the following:

- ✦ GRP of the region;
- ✦ the amount of budget expenses in the field of "National Economy;"
- ✦ level of average monthly accrued wages in the region;
- ✦ rates of growth of the amount of investments in fixed assets;
- ✦ deterioration level of fixed assets;
- ✦ the number of post-graduate students and doctoral students;

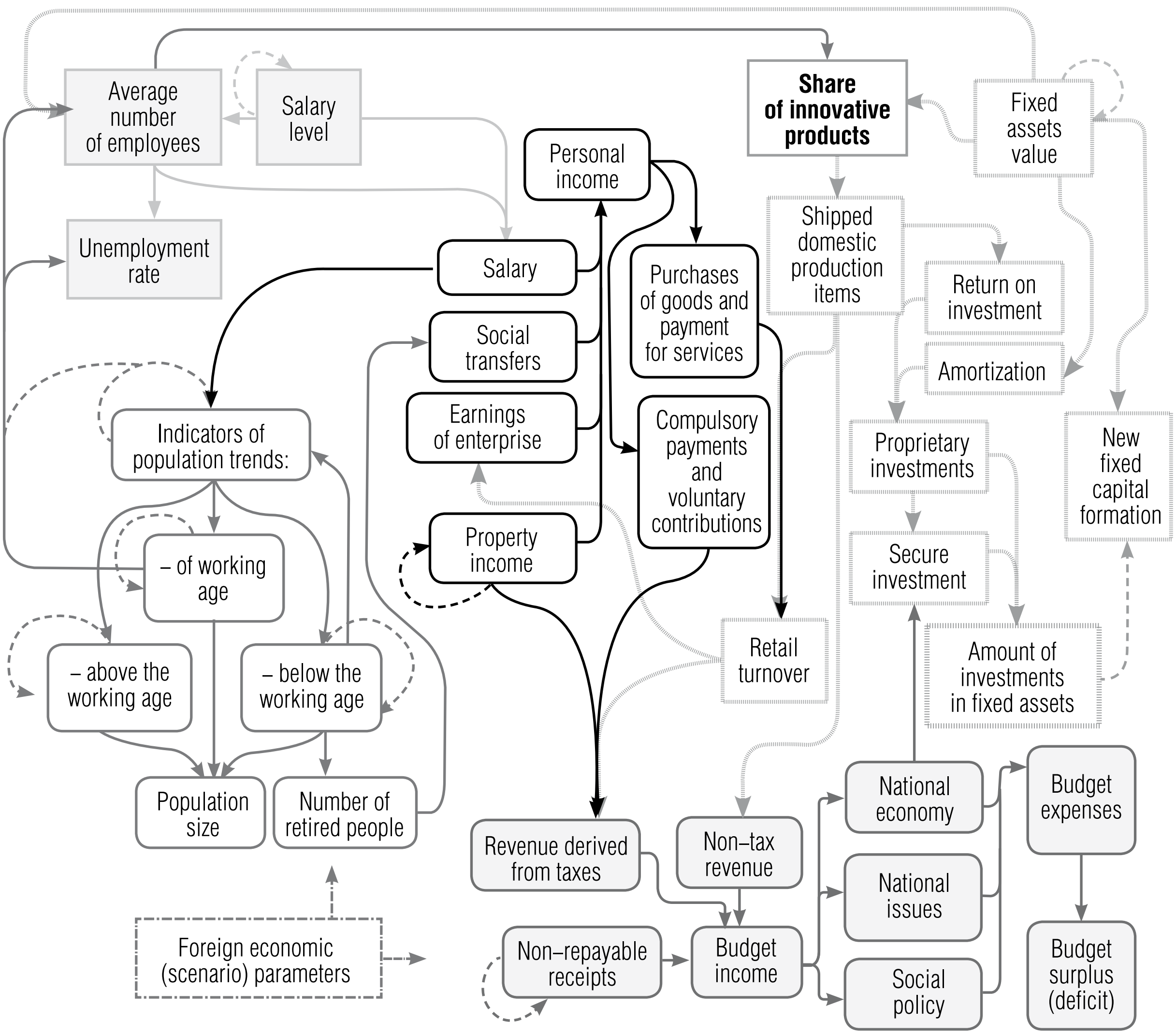


Fig. 3. Aggregated logical structure of a region model

✦ rates of taxes and fees, the value of which can be adjusted at the level of territorial units of the Russian Federation.

In addition to quantitative indicators, it is also necessary to take into account qualitative parameters, such as the quality of strategic planning and the level of formation of the national innovation system. It is suggested that these parameters be evaluated by an expert method and further used in calculations using fuzzy logic methods. The state of these parameters in the aggregate determines the class of the situation, which determines the degree of involvement of the innovative potential of the region. Another feature of the management of innovative development of a region taken into account within the framework of the model is the kind of spread of innovation. Thus, the interaction of an agent with innovatively active contractors increases its own level of innovative activity. The factor of diffusion of innovations is thereby taken into account. The question of taking into account such a factor as the demand for innovation(s) remains open. These parameters determine the specifics of the model for the considered task of managing the innovative development of a region.

Within the framework of the proposed model, the functioning of the agent is defined by a set of conditions and relevant rules of behavior. Adjustment of conditions can lead to both quantitative and qualitative changes in the characteristics of the agent. Such a structure of ASM formulation makes it possible to obtain balanced forecast estimates of regional development indicators. At the same time, the economic and mathematical model alone does not allow us to solve all the tasks assigned to the decision support system, and it needs to be supplemented with a complex of functional units. The composition and interrelation of the units is determined on the basis of the tasks defined for this tool (*Figure 4*).

Making decisions requires an appropriate reference database. For its formation within the framework of the proposed DSS, a data entry and adjustment unit is provided. Moreover, the features of the existing approach to assessing the innovative development of the region are such that subsequent adjustments to the applied methodological base are very likely. Therefore, one must provide for the possibility of changing the number and structure of indicators, as well as the adjustment of previously entered data.

Given the need for periodic updating of data, an objective condition for the effective operation of the tools is the arrangement of monitoring of ongoing processes according to strictly established parameters. This allows you to form a statistical base for solving managerial problems and track the consequences of decisions, that is, to improve the rule base and precedent base used. Today working with large amounts of information is impossible without using modern information technologies for data mining and processing. It seems appropriate to organize data storage based on the concepts of OLAP and Data Mining. Specialists in this field indicate that by using these technologies it is possible to identify hidden patterns in large amounts of information [19, 20].

In addition, the data entry and adjustment unit is used to set the task by the decision maker (DM). In *Figure 4*, the indicated problem is solved within the framework of the task formulation unit. The feasibility of including this unit is determined by the variety of practical tasks and areas of application of the tools developed, as well as the need to enter data characterizing the choice of a regional development strategy.

The strategy formulation unit ensures determination of a unified strategy for the development of a region and allows for the pre-adjustment of a number of model parameters in order to determine the priority direction of development. In particular, investment-oriented, socially-oriented, innovative [21], integrated and other strategies can be considered. Based on the foregoing, the data entry unit acts as a user interface for entering statistical information, scenario and control parameters.

The task of determining the parameters of innovative development of the regional system

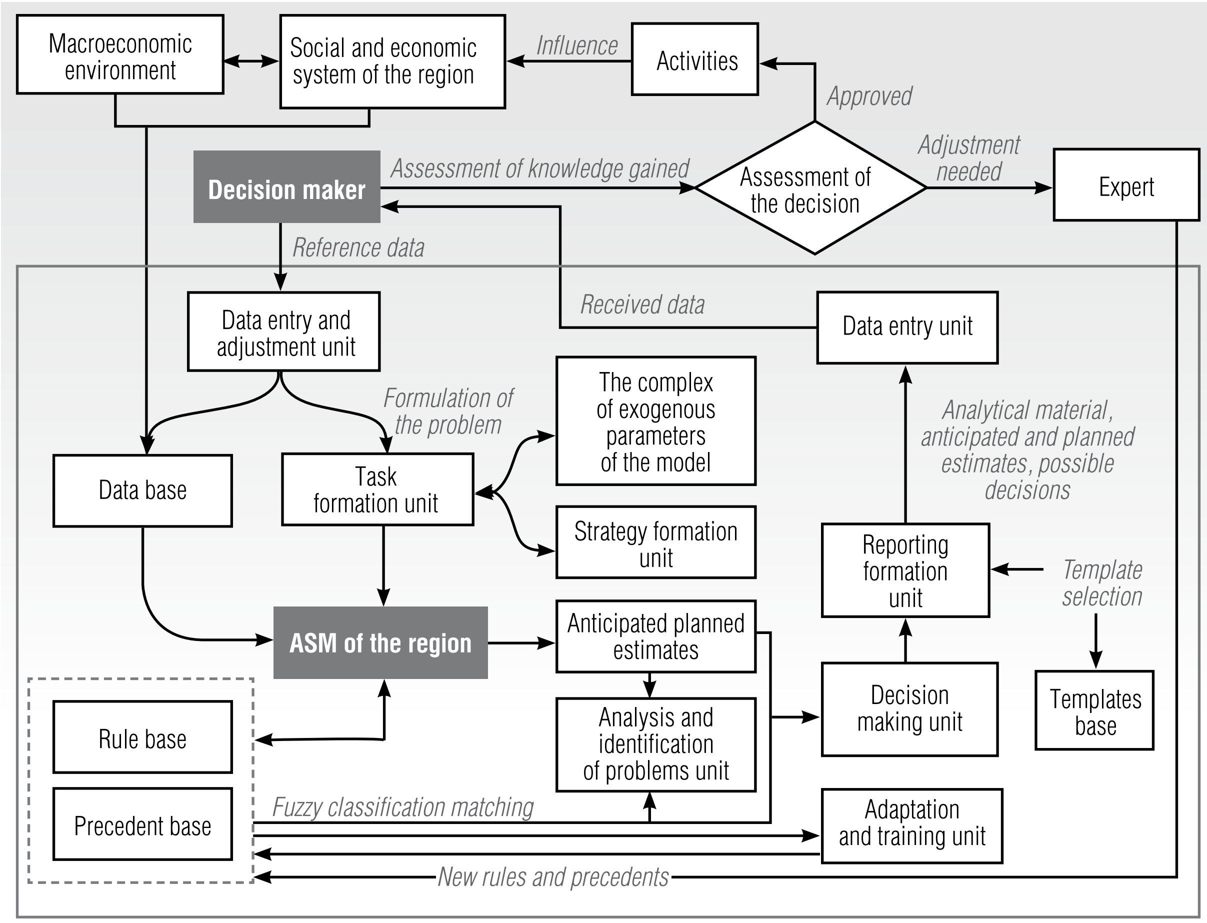


Fig. 4. DSS organization scheme based on the ASM of the region

(including the issues of information support for the procedures for developing a strategy for innovative development of the region [22]) imposes some limitations on model experiments. Thus, it seems necessary to determine the level of innovative development of the territory [23]. Specialists in this area note that within the framework of solving this problem, the following restrictions can be additionally established [24]:

✦ aggregation of a set of criteria of different sizes and multidirectional in their dynamics;

✦ consideration of the significance of the criteria within the framework of formulating the integrated development indicators;

✦ formalization of fuzzy characteristics for effective analysis of qualitative information, as well as clear quantitative data;

✦ linking the integral indicator with the target priorities for the strategic development of a region.

The simulation results so obtained serve the analysis and identification of problems unit. At this stage, a structural analysis, analysis of dynamics and other types of analysis are carried out aimed at the direct identification of the problem and the factors of its occurrence. To solve this problem, the precedent base [25] is used; it stores information about the possibility of using the accumulated experience to solve new problems. The enlarged precedent structure includes two elements – the identification and the teaching parts. On the basis of anticipated and planned estimates that were determined during the analysis of development problems, and a set of precedents, DSS forms a set of decisions for further assessment according to certain criteria.

One of the key components of the decision support system is the decision-making unit. Within the framework of this unit, it is suggested to implement a procedure for classifying situations and comparing them with known precedents to determine further actions. To solve this problem, methods from the theory of fuzzy logic can be used. This approach allows us to track gradual changes in the properties of agents of the regional system, as well as to consider not only quantitative, but also qualitative characteristics. The most significant step is the formulation of fuzzy set adjectives that describe the semantics of the basic values of fuzzy and linguistic variables. The heterogeneity of information sources should be taken into account. Moreover, the choice of decisions should reflect the specifics of innovative development.

The final stage is the formation of a report on the results of simulation experiments and the proposed directions for solving the problems identified. At the same time, it is advisable to use standard reports that allow decision makers to more quickly familiarize themselves with the received data. It should be taken into account that the decision support system is only a tool that allows you to prepare a decision, and it should not replace the decision-maker, who has full responsibility [26]. Based on this, an assessment of the knowledge gained is carried out. When coordinating the proposed decision on the basis of its key parameters, a set of measures is being developed, the implementation of which leads to an adjustment of the parameters of the regional system. If the proposed decision has not been agreed upon, then the rule bases and precedents bases are adjusted with the involvement of relevant experts, and the DSS is restarted.

In addition to these functional units and ASM, to use the decision support system requires appropriate methodological support associated with the practical use of tools. It should include the following methodologies:

- a methodology for collecting and processing information about the state of the region, economic agents and the environment;
- the methodology to set an experiment with ASM of the regional system;
- methodology for the development of managerial decisions based on DSS.

The proposed decision support system can become a universal tool that provides decision makers with important information for decision making on the retrospective, current and future conditions of the regional system, allowing one to analyze and identify possible decisions to existing problems.

Conclusion

In course of the study, a conceptual scheme for a decision support system in the field of managing a region's innovative development was developed. The place of the adaptive simulation model was justified. We identified the possibilities of its application within the framework of developing forecasts and plans for the development of the region and in assessing the effectiveness of the implementation of public policy measures. A procedure for decision-making and a set of functional units providing the final user's work with the ASM of the region are proposed.

The theoretical significance of the study is to expand the possibilities of complex use for solving the problems of strategic development of regional social and economic systems, both methods of management theory and decision support and intellectual information technologies, as well as economic and mathematical methods and modern technologies of simulation modeling. We should also note that the software tool is practice-oriented. The DSS so developed can be used by government bodies in determining the parameters of mid-term development of the regions of Russia, as well as other territorial systems. ■

Acknowledgments

This study was supported by the Russian Foundation for Basic Research (RFBR), project No 18-00-00345.

References

1. Legris P., Ingham J., Collerette P. (2003) Why do people use information technology? A critical review of the technology acceptance model. *Information and Management*, vol. 40, no 3, pp. 191–204.
2. Simankov V.S., Vladimirov S.N., Denisenko A.O., Cherkasov A.N. (2008) Methodological aspects of the expert decision-making support systems' construction. *Vestnik of Don State Technical University*, vol. 8, no 3, pp. 258–267 (in Russian).
3. Palyukh B.V., Kakatunova T.V. (2012) A fuzzy cognitive map as a tool to model innovations at the regional level. *Software & Systems*, no 4, pp. 128–131 (in Russian).
4. Zakharova A.A. (2007) Some aspects of the development of information system to support strategic decision-making for innovative development of the region. *Modern High Technologies*, no 9, pp. 27–29 (in Russian).
5. Akopov A.S. (2010) On one model of adaptive management of complex organizational structures. *Audit and Financial Analysis*, no 3, pp. 310–317 (in Russian).
6. Antamoshkin A.N., Antamoshkina O.I., Hodos D.V. (2010) Modeling of basic steps of innovative development program formation. *Siberian Journal of Science and Technology*, no 4, pp. 310–317 (in Russian).
7. Markus M.L., Majchrzak A., Gasser L. (2002) A design theory for systems that support emergent knowledge processes. *MIS Quarterly: Management Information Systems*, vol. 26, no 3, pp. 179–212.
8. Makarov V.L., Afanasiev A.A., Losev A.A. (2011) Computable simulation model for money circulation in the Russian economy. *Economics and the Mathematical Methods*, vol. 47, no 1, pp. 3–27 (in Russian).
9. Lychkina N.N. (2007) Simulation models in procedures and systems of strategic decision support at enterprises. *Business Informatics*, no. 1, pp. 29–35 (in Russian).
10. Burkov V., Burkova I., Irikov V. (2010) Managing regional innovative development: Modern approach. *Theoretical and Practical Aspects of Management*, no 11, pp. 8–12 (in Russian).
11. Tatarkin A.I. (2016) Innovative vector of the Russian economics: Behavior readiness of the population. *Business, Management and Law*, no 1, pp. 20–28 (in Russian).
12. Nizamutdinov M.M., Oreshnikov V.V. (2015) A toolkit for forecasting changes in the parameters of regional development on the basis of an adaptive simulation approach. Proceedings of the *3rd International Conference on the Information Technologies for Intelligent Decision Support (ITIDS' 2015), Ufa, 18–21 May 2015*, pp. 211–215 (in Russian).
13. Lomazov V.A., Lomazova V.I., Petrosov D.A. (2014) Evolutionary procedure of decision support in modeling of interacting processes. *Problems of Contemporary Science and Practice. Vernadsky University*, no 2, pp. 82–89 (in Russian).
14. Glushan V.M., Karelin V.P., Kuzmenko O.L. (2009) Fuzzy models and methods of multi-criterion choice in intelligent decision support systems]. *Izvestiya SFedU. Engineering Sciences*, no 4, pp. 106–113 (in Russian).
15. Nizamutdinov M.M., Oreshnikov V.V. (2016) Formation of the target indicators of regional development strategy using fuzzy logic algorithms. *Economics and the Mathematical Methods*, vol. 52, no 2, pp. 30–39 (in Russian).
16. Manenti L., Sartori F. (2011) Metadata support to retrieve and revise solutions in case-based reasoning. *International Journal of Metadata, Semantics and Ontologies*, vol. 6, no 3–4, pp. 185–194.
17. Kravchenko T.K., Seredenko N.N. (2012) The establishment of systems for decision support: integration of the advantages of individual approaches. *Artificial Intelligence and Decision Making*, no 1, pp. 39–47 (in Russian).
18. Kisurkin A.A. (2012) Factors affecting the innovative development of the region and their classification by management levels. *Modern Problems of Science and Education*, no 2. Available at: http://science-education.ru/ru/article/view?id=5762 (accessed 20 April 2019) (in Russian).
19. Barabanov I.N., Korgin N.A., Novikov D.A., Chkhartishvili A.G. (2010) Dynamic models of informational control in social networks. *Automation and Remote Control*, no 11, pp. 172–182 (in Russian).

20. Nemati H.R., Steiger D.M., Iyer L.S., Herschel R.T. (2002) Knowledge warehouse: An architectural integration of knowledge management, decision support, artificial intelligence and data warehousing. *Decision Support Systems*, vol. 33, no 2, pp. 143–161.
21. Gazhva V.O., Gubanova E.V. (2016) Innovative development of the Kaluga Region. Proceedings of the *International Conference on Directions of Social and Economic Development of Regional Economy*, *Kaluga, 15 March 2016,* pp. 122–125 (in Russian).
22. Tyushnyakov V.N., Zhertovskaya E.V., Yakimenko M.V. (2015) Information and analytical support of situational centers as the basis for developing strategies of a region's innovative development. *Fundamental Research*, no 11–6, pp. 1253–1257 (in Russian).
23. Gurieva L.K. (2013) New approaches to innovative regions development. *Humanities and Social Sciences*, no 5, pp. 35–42 (in Russian).
24. Zakharova A.A. (2014) Integrated assessment of innovative development of the region on the basis of fuzzy sets. *Scientific Review. Technical Sciences*, no 1, pp. 161–168 (in Russian).
25. Eremeev A.P., Varshavskii P.R. (2009) [Modeling of case-based reasoning in intelligent decision support systems. *Artificial Intelligence and Decision Making*, no 2, pp. 45–47 (in Russian).
26. Kravchenko T.K., Isaev D.V. (2016) Making strategic decisions under risk and uncertainties. *Bulletin of the Financial University*, vol. 20, no 4, pp. 22–31 (in Russian).

About the authors

Lilia R. Chernyakhovskaya

Dr. Sci. (Tech.);

Professor, Department of Technical Cybernetics, Ufa State Aviation Technical University,
12, Karl Marx Street, Ufa 450008, Russia

E-mail: lrchern@yandex.ru

ORCID: 0000-0002-2447-8864

Marsel M. Nizamutdinov

Cand. Sci. (Tech.), Associate Professor;

Head of the Sector of Economic and Mathematical Modeling,
Institute of Social and Economic Researches,
Ufa Federal Research Center, Russian Academy of Sciences,
71, Prospekt Oktyabrya, Ufa 450054, Russia

E-mail: marsel_n@mail.ru

ORCID: 0000-0001-5643-1393

Vladimir V. Oreshnikov

Cand. Sci. (Econ.);

Senior Researcher, Institute of Social and Economic Researches,
Ufa Federal Research Center, Russian Academy of Sciences,
71, Prospekt Oktyabrya, Ufa 450054, Russia

E-mail: VOresh@mail.ru

ORCID: 0000-0001-5779-4946

Alsouy R. Atnabaeva

Researcher, Institute of Social and Economic Researches,
Ufa Federal Research Center, Russian Academy of Sciences,
71, Prospekt Oktyabrya, Ufa 450054, Russia

E-mail: Alsouy@mail.ru

ORCID: 0000-0002-7042-1180

DOI: 10.17323/1998-0663.2019.3.35.51

Credit risk stress testing in a cluster of Russian commercial banks

Davit S. Bidzhoyan
E-mail: bidzhoyan_david@mail.ru

Tatiana K. Bogdanova
E-mail: tanbog@hse.ru

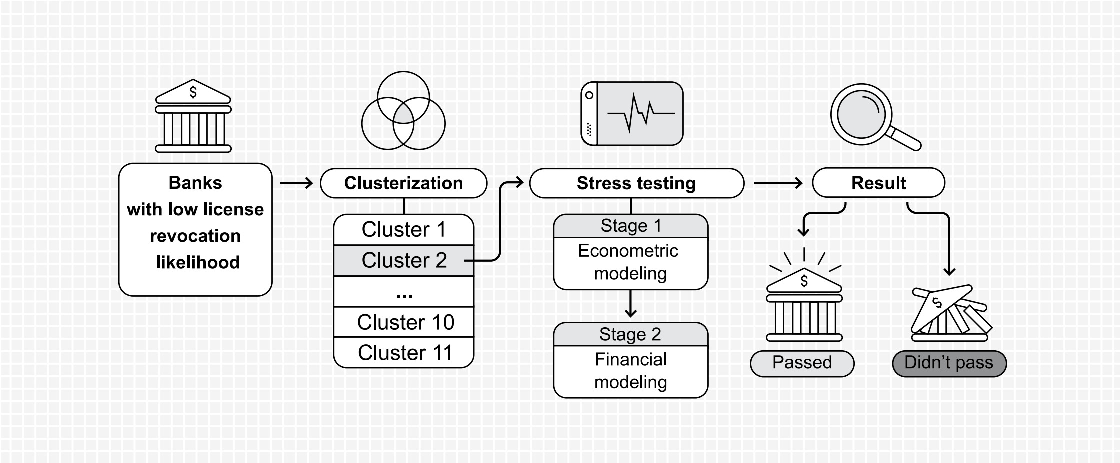
Dmitry Yu. Neklyudov
E-mail: nekludovmid@gmail.com

National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow 101000, Russia

Abstract

Stress testing as an instrument of risk evaluation is actively used in many international organizations, as well as by central banks in many countries. Some organizations (including the Bank of Russia) when conducting stress testing do not publish results of the tests, though they are interesting for the business community. They are reticent so to avoid causing panic on markets which could lead to a massive outflow of deposits from the banking sector as a whole or from some individual banks in particular. As a rule, stress testing is conducted relying on huge number of unpublished reporting forms, but the business community has no access to them. Only four reporting forms are presented on the Bank of Russia's website. In this paper we propose a simplified algorithm of credit risk stress testing of a banking cluster based on the four officially published reporting forms. The algorithm provides modelling of median values of banking variables depending on macroeconomic indicators, and subsequent retranslation of the received values for assessing the financial position of each bank included in the cluster. It is assumed that growth rates of banking indicators obtained from the econometrics models relying on median values are the same for each bank in the cluster. As of 1 January 2018, credit risk stress testing was conducted for 26 banks, nine of which are system-significant credit institutions. Within the stress testing, eight econometric time series models were developed. As a result, it was discovered that 11 out of 26 banks in the cluster will face certain difficulties regarding statutory requirements related to capital ratios or buffers.

Graphical abstract



Key words: stress testing; credit risk; macroeconomic indicator; econometric model; system-significant credit institutions; median.

Citation: Bidzhoyan D.S., Bogdanova T.K., Neklyudov D.Yu. (2019) Credit risk stress testing in a cluster of Russian commercial banks. *Business Informatics*, vol. 13, no 3, pp. 35–51. DOI: 10.17323/1998-0663.2019.3.35.51

Introduction

The financial statement of banks largely depends on the macroeconomic environment in which banks have to operate. This dependence is manifested both directly and indirectly. Direct impact is through the revaluation of assets and liabilities depending on the base parameter. For example, the open foreign-currency position of banks directly depends on the value of the exchange rate to the Russian ruble; the fair value of the bonds depends on the yield curve to maturity, which rises during crisis periods; and the value of shares falls as the market index falls. The indirect effect is through the financial statements of the bank's borrowers, which lead to a change in the assets quality. As a rule, during crisis periods the financial statements of borrowers deteriorate, setting the need to charge additional loan loss provisions (LLP), all of which leads to a decrease in the bank's capital.

Since banks play the central role in a market economy, being the "circulatory system" of the economy, all market participants are interested in assessing a bank's financial statement, including the Central Bank of the Russian Federation. It has all possible information on the activities of banks, due to the large number of reporting forms provided on a regular basis. In addition to the Bank of Russia, the financial statement of banks is interesting to their creditors, depositors and investors, and mostly to legal entities whose deposits and funds on settlement accounts are not insured by the state, unlike deposits of individuals. Moreover, creditors of Russian commercial banks are deprived of the possibility of deep analysis of the bank's financial statements due to limited information: there are only four reporting forms on the Bank of Russia website that allow one to assess the financial statements of banks.

Due to the current geopolitical situation and the continuing policy of the Bank of Russia to improve the banking sector, many banks are at risk of default. A rough estimate of the impact of a deteriorating macroeconomic environment on the financial condition of banks can be obtained using stress testing methods. The results of stress testing conducted by the Bank of Russia are not published, but nevertheless, they may be of interest to the bank's lenders and depositors. The paper [1] presents the stress testing algorithm for a bank's credit risk on an individual basis based on official, publicly available financial reports. However, conducting stress testing of the credit risk of a large number of banks is laborious. Therefore, it is important to develop a credit risk stress testing algorithm for a cluster of banks.

This paper has the following structure. The first paragraph provides a brief overview of papers by Russian and foreign scientists on the subject of the study; the second section presents the stress testing algorithm for a cluster of banks; and the third presents the stress testing results of certain bank clusters. In the conclusion, the main results of the study are formulated.

1. Methods and models for assessing the banks' financial position

A large number of scientific papers by both foreign and Russian authors are devoted to the analysis of a bank's financial statements. All mathematical models can be divided into several categories:

- estimating probability of default [2–6];
- evaluating the financial statements by ratings [7, 8];
- models of technical efficiency [9, 10];
- interest rate models [11–13].

The afore-mentioned methods model the current financial statement of a bank, while stress testing methods allow us to simulate the effect of various macroeconomic shocks on the banks' financial position.

Stress testing as a tool for bank risk assessment appeared in the 1990s. However, it became popular only after the financial crisis of 2007–2009. Today the central banks of almost all countries with developed and developing economies use stress testing to assess future risks during crisis periods. In addition, stress testing as a risk assessment tool is used by international organizations such as the Basel Committee on Banking Supervision and the International Monetary Fund. A classification of stress testing tools is given in [14].

Credit risk, as the main source of losses, can be measured in various ways. Organizations that possess the necessary data use non-performing loans (NPL) [15], loans of quality categories 4–5 (according to the Bank of Russia Regulation No 590-P dated 28 June 2017 "On the Procedure for Making Loss Provisions by Credit Institutions for Loans and Similar Debts"), as well as their share in the portfolio. However, paper [1] presents an algorithm for credit risk stress testing based on publicly available financial reports; it proposes to use the volume of LLP, which can be calculated relying on changings accounts balances, as a measure of credit risk. In this paper, we also use LLP to measure credit risk.

Stress testing based on scenario analysis should be conducted by regressing the volume of NPL on macroeconomic variables. As a rule, simple regression models are used, as shown in [16]. A more complex model based on panel data is presented in [15].

One of the key points in stress testing is the stress scenario development. One of the main criteria should be its severity and plausibility. Central banks use one or more stressful scenarios. This can lead to false positive results (the "false illusion" effect), since in one stressful scenario the bank may come to a state of default, and in another – will not. Calculating for all scenarios leads to computational burden. Therefore, in [17] the principle of "the most severe plausible scenario" was proposed. Plausibility can be calculated based on the Mahalanobis distance [18], the Kulbek–Leibler distance, the Bregman distance, and the *f*-divergence method [19].

The large number of macroeconomic factors that directly or indirectly affect the financial statements of banks should be taken into account when conducting stress testing. For example, the Bank of England when stress testing the seven largest banks uses 58 macroeconomic indicators[1] and work with them is extremely difficult. Many studies have proposed methods for reducing the dimensionality: in particular, the principal component method [20] and other methods that reduce the dimensionality [21] are used.

The bank can sell a part of its assets with a high risk ratio at a large discount in order to meet the capital adequacy ratios and premiums. The mathematical formalization of such behavioral models is described in detail in [22–24].

2. An algorithm of stress testing for a cluster of banks

In [25], an information-logical model for identifying a group of reliable banks is presented. This model has four stages:

Stage 1. Collection and aggregation of official public reporting data and the macroeconomic variables;

Stage 2. Building the logistic regression model for estimating the probability of license revocation of Russian commercial banks and determining the optimal cut-off value;

Stage 3. Grouping banks with a low probability of license revocation based on cluster analysis or on the basis of specified criteria;

Stage 4. Stress testing the credit risk of a selected group of banks in order to identify whether the capital is sufficient to cover losses associated with credit risk as the main source of losses. Banks whose capital adequacy ratios are above the minimum values at the end of a stressful period are considered to be reliable.

This paper proposes an algorithm for conducting credit risk stress testing for a group of banks which is an extension of the algorithm presented in [1]. The simulation includes eight banking indicators, depending on macroeconomic variables:

- corporate LLP;
- retail LLP;
- corporate loans;
- retail loans;
- individual deposits;
- corporate deposits;
- funds on individuals current accounts;
- funds on corporate settlement accounts.

The following macroeconomic variables are used as risk factors in the credit risk stress testing framework:

- mean, standard deviation and variance of the exchange rate of the US dollar, in rubles;
- mean, standard deviation and variance of the cost of a barrel of Brent oil, in US dollars;
- mean, standard deviation and variance of the interbank lending rate MIACR[2], in %;
- mean, standard deviation and variance of the MICEX[3] index, in points;
- mean value, standard deviation and variance of the RTS index, in points;
- real GDP growth rate (compared to the same quarter of previous year), in %;
- the consumer price index growth rate (compared to the same quarter of previous year), in %;
- the population real income growth rate (compared to the same quarter of previous year), in %;
- the growth rate of expenditures of the population (compared to the same quarter of previous year), in %;
- the growth rate of imports (compared to the same quarter of previous year), in %.

The whole process of stress testing is divided into two main blocks: econometric modeling and financial modeling.

[1] https://www.bankofengland.co.uk/stress-testing
[2] MIACR – Moscow Inter-Bank Actual Credit Rate
[3] MICEX – Moscow Interbank Currency Exchange

Within the first block, econometric modeling of bank indicators is carried out. At each time point (selected modeling step) for the analyzed cluster of banks, for each of the eight bank indicators, their median values are calculated. They are further used in calculating the growth rate of banking indicators depending on macroeconomic variables based on econometric models, which are estimated using the least squares method (OLS). In the case of the presence of autocorrelation and / or heteroscedasticity of residuals, the heteroskedasticity and autocorrelation consistent (HAC) Newey–West robust standard errors are used.

Parameters of the simulation procedure for credit risk stress testing are shown in *Table 1.*

Table 1.

Parameters of credit risk stress testing simulation

Interval	quarter
Simulation period	Q1 2008 – Q4 2018
Forecasting horizon	1 year (4 quarters)
Stress scenario	2015 year[4]

Table 2 shows the necessary data for conducting a credit risk stress test.

In the second block (financial modeling), based on the predicted values of banking indicators in the stress period obtained at the first stage, the adequacies of total, Tier 1 and Common equity Tier 1 capitals are calculated for each bank in order to assess whether the bank will be able to meet the capital adequacy requirements in stress.

Within this framework, the capital adequacy ratios are calculated in three steps.

Step 1. Calculation of capital adequacy ratios after additional charges of LLP on existing loan portfolios according to the formula:

Table 2.

Data for credit risk stress testing

Variable	Reporting form	Aggregation code
$N_{1.0}$	0409135[5]	H1.0
$N_{1.1}$	0409135	H1.1
$N_{1.2}$	0409135	H1.2
High liquid assets	0409135	LAM
Liquid assets	0409135	LAT
Total capital	0409123[6]	000
Tier 1 capital	0409123	106
Common equity Tier 1 capital	0409123	102

$$H^k_{1.i,t} = \frac{K^k_{i,t} - \sum_{j=1}^{2} \Delta\hat{R}^k_{j,t}}{RWA^k_i - \sum_{j=1}^{2} cr_j \cdot \Delta R^k_{j,t}}, \qquad (1)$$

where k – bank in the cluster;

$H^k_{1.i,t}$ – the value of i-th capital adequacy ratio of bank k at time t[7];

$K^k_{i,t}$ – i-th capital of bank k at time t;

$\Delta\hat{R}^k_{j,t}$ – the forecasted charge of *LLP* in existing portfolio j of bank k (j = 1 – corporate loans portfolio; j = 2 – retail loans portfolio);

cr_j – risk coefficient of portfolio j;

RWA^k_i – risk weighted assets of capital i of bank k.

Step 2. Calculate the growth of loan portfolios in three stages, as shown in *Figure 1.*

The following designations in the formulas in the figure are used:

$Z^k_{j,t}$ – the actual value of j-th portfolio of bank k at time t + 1 (j = 1 – corporate loans portfolio; j = 2 – retail loans portfolio);

[4] The historical scenario of 2015 means using the values of the parameters that were in 2015 as the basis of the stress scenario in 2018

[5] Reporting Form 0409135 "Information on Credit Institutions' Required Ratios and Other Performance Indicators"

[6] Reporting Form 0409123 "Capital Calculation (Basel III)"

[7] Hereafter under "at time *t*" we consider the end of quarter

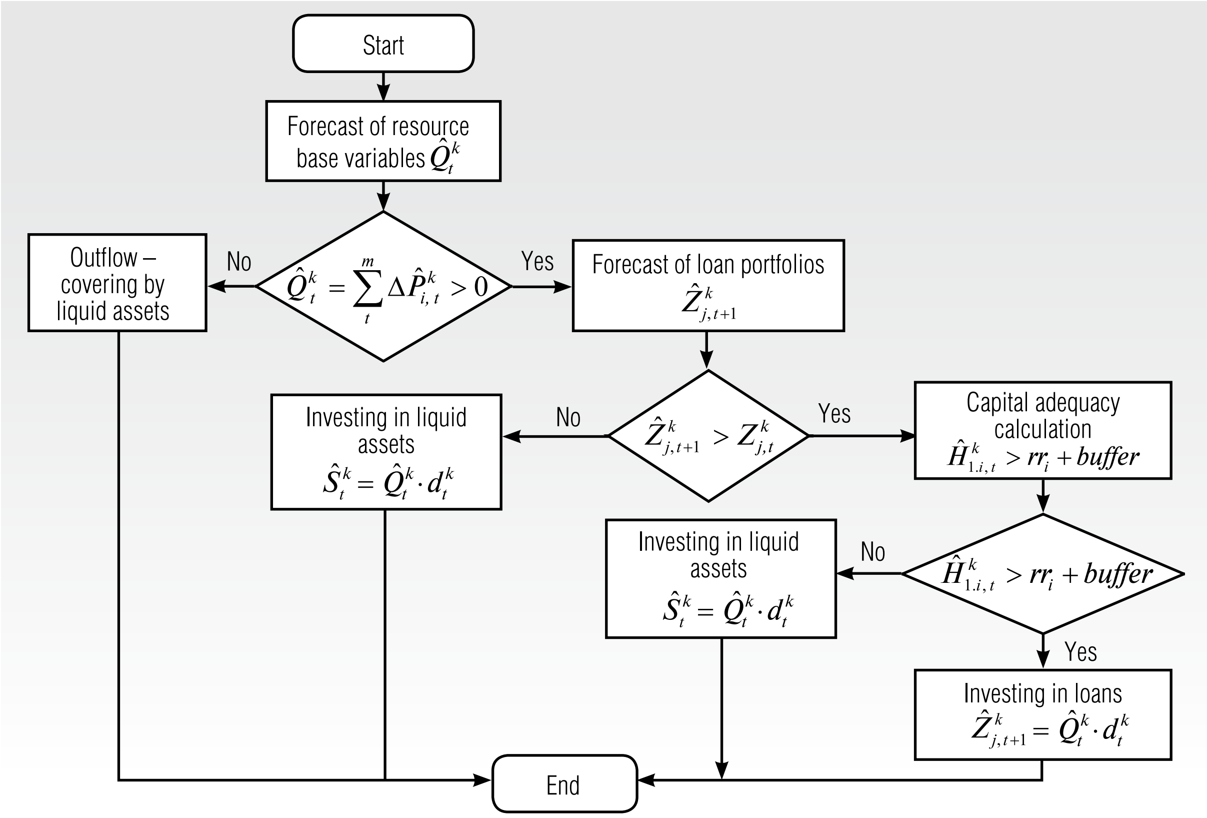


Fig. 1. The scheme for calculating the volume of portfolio increase

$\hat{Z}_{j,t+1}^k$ – the forecasted value of j-th portfolio of bank k at time $t + 1$ ($j = 1$ – corporate loans portfolio; $j = 2$ – retail loans portfolio);

$\hat{Q}_{t+1}^k$ – the surplus of resource base of bank k at time t;

$\hat{S}_t^k$ – the growth of resource base of bank k at time t;

d_t^k – the share of j-th portfolio in total assets of bank k;

$\hat{P}_{i,t}^k$ – the i-th resource base of bank k at time t;

$H_{1.i}^k$ – the estimated i-th capital adequacy ratio;

rr_i – the i-th capital minimum required ratio;

buffer – buffer to capital adequacy ratios.

At the first stage, one checks whether there is an inflow of the resource base. If there is an outflow of the resource base, the deficit is covered by liquid assets. If there is an inflow, then transition to stage 2, within the framework of which the demand for loans from banks is checked. If demand falls, then the surplus inflow is invested in liquid assets, otherwise transition to stage 3. The availability of the necessary capital to meet the capital requirements is checked. If all requirements are met, the funds are invested in the loan portfolio, otherwise – in liquid assets.

Step 3. Calculation the capital adequacy ratios taking into account the additional charge of LLP of newly issued loans using the formula:

$$\hat{H}_{1.i,t\_new}^k = \\ = \frac{\hat{K}_{i,t\_new}^k - cp \cdot \sum_{j=1}^{2} \Delta \hat{Z}_{j,t}^k}{\widehat{RWA}_{i,t\_new}^k + \sum_{j=1}^{2} cr_j \cdot \left(\Delta \hat{Z}_{j,t}^k - cp \cdot \Delta \hat{Z}_{j,t}^k \right)}, \quad (2)$$

where k – bank in the cluster;

$\hat{H}_{1.i,t\_new}^k$ – the estimated value of i-th capital adequacy ratio of bank k at time t taking into account newly issued loans;

$\hat{K}^k_{i,t\_new}$ — the estimated i-th capital of bank k at time t taking into account taking into account charged LLP of newly issued loans;

$\Delta\hat{Z}^k_{j,t}$ — the forecasted value of j-th portfolio growth of bank k at time t (j = 1 — corporate loans portfolio; j = 2 — retail loans portfolio);

cr_j — risk coefficient of portfolio j;

cp — LLP on newly issued loans;

$\widehat{RWA}^k_{i,t\ new}$ — the estimated risk weighted assets of capital i of bank k taking into account charged LLP of newly issued loans.

For those banks that do not meet capital requirements, the additional capitalization volumes are calculated by the formula:

$$\hat{D}^k = \max\{\left(required\ H_{1,i} - \hat{H}^k_{1,i}\right) \cdot \widehat{RWA}^k_i\}, \quad (3)$$

where $\hat{D}^k$ — the estimated deficit of capital of bank k;

$required\ H_{1\,i}$ — minimum required capital adequacy ratio (i = 0 — total capital, i = 1 — Common Equity Tier 1 Capital, i = 2 — Tier 1 Capital);

$\hat{H}^k_{1\,i}$ — the final estimated value of i-th capital adequacy ratios of bank k;

$\widehat{RWA}^k_i$ — the final estimated risk weighted assets of i-th capital of bank k.

Banks that comply with capital adequacy requirements on them are considered reliable for investments.

3. Stress testing the largest Russian bank cluster

In [14], a cluster analysis was conducted on banks (334 banks) with a low probability of license revocation as estimated on the basis of a logistic regression model. According to cluster analysis results, 11 clusters were obtained. *Table 3* shows the distribution of banks by clusters.

In this paper, we carried out stress testing of the credit risk of the second cluster of banks, which includes the largest Russian banks. The list of banks forming the second cluster is given below in *Table 13.*

Table 3.

Distribution of banks by clusters

No of cluster	Number of banks
1	63
2	26
3	15
4	10
5	63
6	28
7	45
8	23
9	29
10	14
11	9

3.1. Econometric models for predicting banking variables

Modeling the volume of individuals' deposits

The banks' deposits from individuals in the second cluster positively depend on the real GDP growth rate, with a lag of one quarter. The results of the econometric model are presented in *Table 4.*

The following designations are used in the table:

y_t — individuals' deposits;

$x_{1,t-1}$ — real GDP growth at time $t-1$ (in %, compared to the same quarter of the previous year).

The model is fully appropriate, since tests for autocorrelation and heteroscedasticity indicate their absence, and the high value of the adjusted R^2 and a low value MAPE mean the model is good for prediction. The graph of forecast values of the volume of individual deposits of banks of the second cluster in the stress scenario is shown in *Figure 2.*

Table 4.

Results of modeling individuals' deposits

Variable	Coefficient	Test	Statistics	*p*-value
constant	1.34	Darbin–Watson	1.66	
pcy(y_{t-1})	0.86***	Breusch–Godfrey	0.53	0.59
$x_{1,t-1}$	0.37*	Breusch–Pagan–Godfrey	1.39	0.26
R^2 adj	0.68	White test	2.07	0.09
MAPE	4.33%			

Significance codes: «***» – 0.001; «**» – 0.01; «*» – 0.5

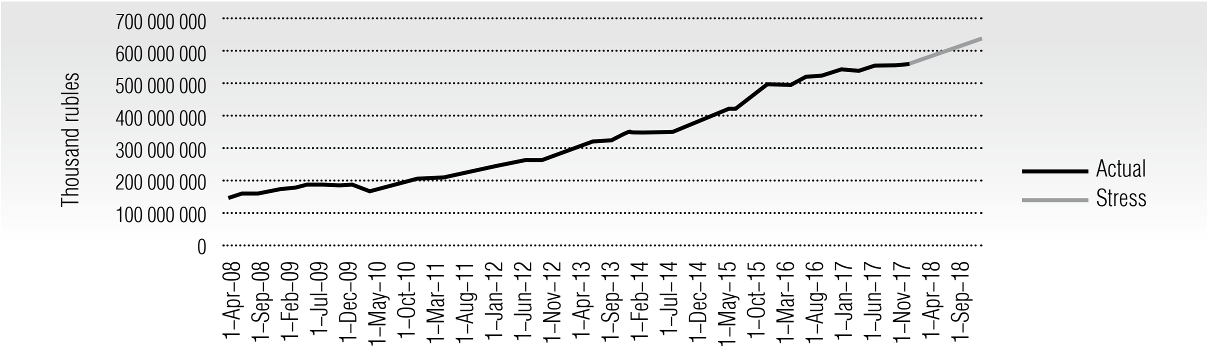


Fig. 2. The actual and forecasted values of individuals' deposits

Modeling the volume of corporate deposits

The volume of corporate deposits positively depends on the real growth rate of GDP with a lag of two quarters. The model also has an autocorrelation component. The simulation results are presented in *Table 5.*

The following designations are used in the table:

y_t – corporate deposits;

$x_{1,t-2}$ – real GDP growth at time $t-2$ (in %, compared to the same quarter of the previous year).

The model has autocorrelation, to eliminate the negative effects of which the Newey–West robust standard errors were applied. However, the model residuals are homoscedastic.

Table 5.

Results of modeling corporate deposits

Variable	Coefficient	Test	Statistics	*p*-value
constant	5.01	Darbin–Watson	2.24	
pcy(y_{t-1})	0.64***	Breusch–Godfrey	4.00	0.027
$x_{1,t-2}$	1.33*	Breusch–Pagan–Godfrey	0.92	0.40
R^2 adj	0.62	White test	1.35	0.26
MAPE	4.53%			

Significance codes: «***» – 0.001; «**» – 0.01; «*» – 0.5

The actual and forecasted values of corporate deposits of banks of the second cluster in the stress scenario are shown in *Figure 3.*

Modeling the volume of the retail loans portfolio

The real GDP growth rate with a one quarter lag has a positive effect on the volume of the retail portfolio of banks in the second cluster, while the unemployment rate is negative. The simulation results are presented in *Table 6.*

The following designations are used in the table:

y_t – retail loans portfolio;

$x_{1,t-1}$ – real GDP growth at time $t-1$ (in %, compared to the same quarter of the previous year);

$x_{2,t}$ – the unemployment rate at time t (in %).

The model residuals are autocorrelated, but homoscedastic. To eliminate the negative effects of autocorrelation, the Newey–West robust standard errors were applied. The high value of adjusted R^2 and the low MAPE indicate the suitability of the model for prediction. The actual and forecasted values of the retail loans portfolio of banks of the second cluster in the stress period are shown in *Figure 4.*

Modeling the volume of the corporate loans portfolio

Interest rates of corporate loans over a year (negatively) and prices for a barrel of Brent crude oil (positively) determine the volume of the corporate portfolio of banks in the second cluster. The simulation results are presented in *Table 7.*

The following designations are used in the table:

y_t – the volume of corporate loans portfolio;

$x_{1,t}$ – corporate loan's interest rate at time t (in %);

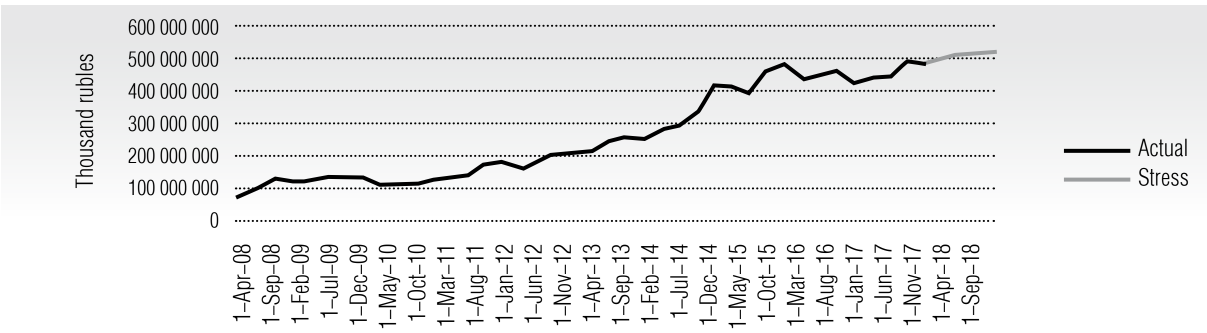


Fig. 3. The actual and forecasted values of corporate deposits

Table 6.

Results of modeling the retail loans portfolio

Variable	Coefficient	Test	Statistics	*p*-value
constant	–0.72***	Darbin–Watson	1.11	
pcy(y_{t-1})	0.73***	Breusch–Godfrey	3.98	0.028
$x_{1,t-1}$	1.57***	Breusch–Pagan–Godfrey	2.28	0.096
$x_{2,t}$	–0.34***	White test	1.44	0.21
R^2 adj	0.86			
MAPE	6.71%			

Significance codes: «***» – 0.001; «**» – 0.01; «*» – 0.5

$x_{2,t-1}$ – the mean price of Brent oil at time $t-1$ (in US dollars).

The model is adequate and suitable for prediction. The model residuals are not autocorrelated and homoscedastic, adjusted R^2 high, MAPE. The actual and forecasted values of the corporate loans portfolio of banks in the second cluster in the stress scenario are shown in *Figure 5.*

Modeling the volume of retail LLP

The retail LLP are negatively dependent on the real GDP growth rate. Also in the model there is an autocorrelation component. The simulation results are presented in *Table 8.*

The following designations are used in the table:

y_t – retail LLP;

$x_{1,t}$ – real GDP growth rate at time t (in %, compared to the same quarter of previous year).

The results of the Breush–Godfrey, Breush–Pagan–Godfrey, and White tests indicate that there is no autocorrelation and heteroscedasticity in the model residuals. The model is adequate and suitable for prediction. *Figure 6* shows a graph of actual and forecasted values of retail banks LLP in a stress scenario.

Modeling the volume of corporate LLP

The value of corporate LLP of banks of the second cluster is negatively dependent on the real GDP growth rate and the price of Brent crude oil. The results of the model are presented in *Table 9.*

The following designations are used in the table:

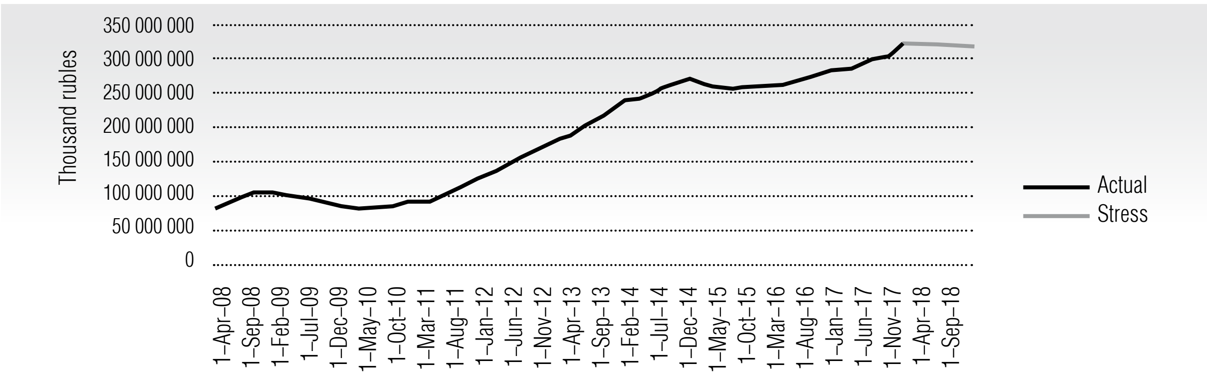


Fig. 4. The actual and forecasted values of retail loans portfolio

Table 7.

Results of modeling the corporate loans portfolio

Variable	Coefficient	Test	Statistics	*p*-value
constant	27.13	Darbin–Watson	1.55	
pcy(y_{t-1})	0.87***	Breusch–Godfrey	1.65	0.2
$x_{1,t}$	–2.34*	Breusch–Pagan–Godfrey	0.86	0.46
$x_{2,t-1}$	0.03*	White test	0.94	0.5
R^2 adj	0.80			
MAPE	7.61%			

Significance codes: «***» – 0.001; «**» – 0.01; «*» – 0.5

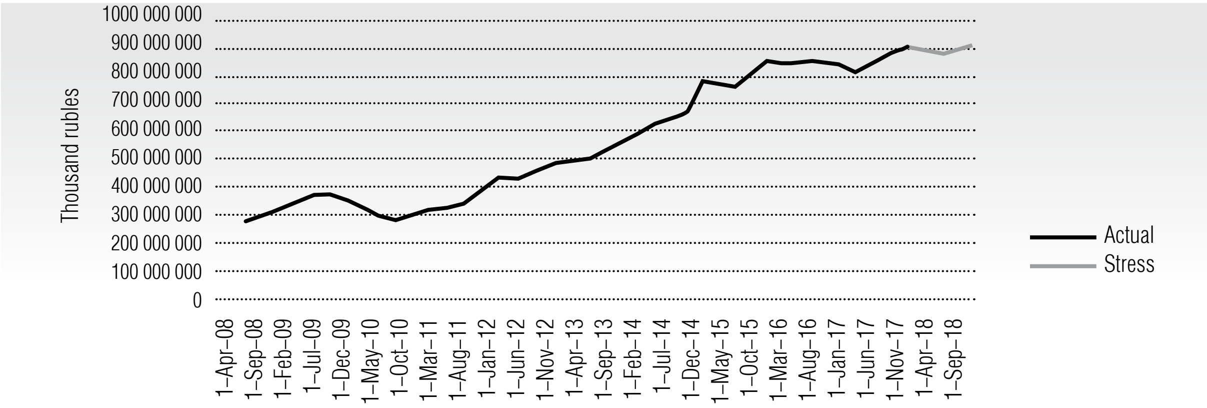


Fig. 5. The actual and forecasted values of corporate loans portfolio

y_t – corporate LLP;

$x_{1,t}$ – real GDP growth rate at time t (in %, compared to the same quarter of previous year);

$x_{2,t}$ – the mean price of Brent oil at time t (in US dollars).

The model is adequate and suitable for prediction. The adjusted R^2 value is high, and the average modeling error is low. There is no autocorrelation and heteroscedasticity. The actual and forecasted values of corporate loans loss provisions of banks of the second cluster in the stress period are shown in *Figure 7*.

Modeling the volume of funds on individuals' current accounts

Interest rates on deposits of individuals negatively affect the amount of funds on the current accounts of individuals of banks of the second cluster. This is because with the growth of interest rates on deposits, people prefer to transfer money from current accounts to deposits. The results of the model are presented in *Table 10*.

The following designations are used in the table:

y_t – funds of current individuals' accounts;

$x_{1,t}$ – real GDP growth rate at time t (in %, compared to the same quarter of previous year).

The results of the Breusch–Godfrey, Bruesh–Pagan–Godfrey and White tests show the absence of autocorrelation and heteroscedasticity. The model is adequate and suitable for prediction. The actual and forecasted values of funds on individual current accounts in banks included in the second cluster during the stress period are shown in *Figure 8*.

Table 8.

Retail LLP modelling resutls

Variable	Coefficient	Test	Statistics	*p*-value
constant	0.81*	Darbin–Watson	1.19	
pcy(y_{t-1})	0.74***	Breusch–Godfrey	3.36	0.004
$x_{1,t}$	–0.15*	Breusch–Pagan–Godfrey	0.92	0.40
R^2 adj	0.88	White test	1.3	0.28
MAPE	8.56%			

Significance codes: «***» – 0.001; «**» – 0.01; «*» – 0.5

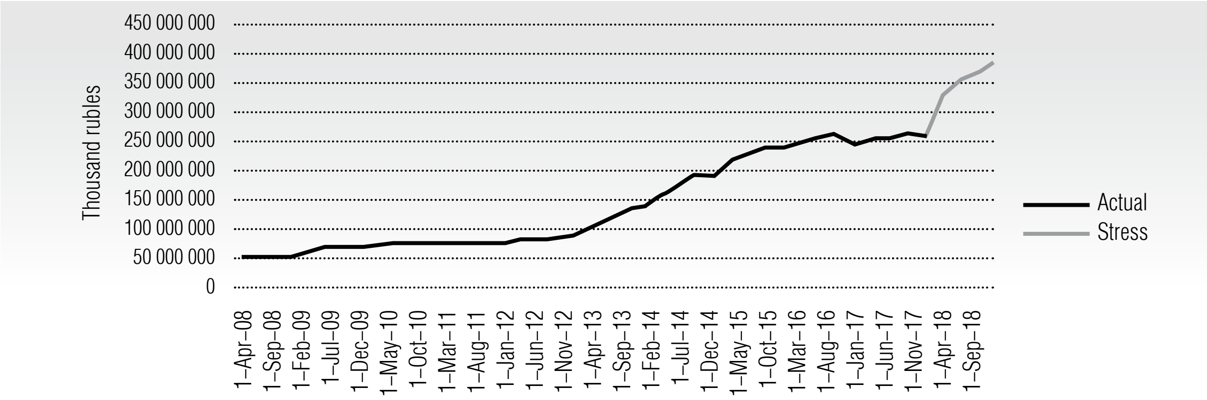


Fig. 6. The actual and forecasted values of the retail loans loss provision

Table 9.

Resutls of corporate LLP modeling

Variable	Coefficient	Test	Statistics	*p*-value
constant	–6.25*	Darbin–Watson	2.10	
pcy(y_{t-1})	0.6***	Breusch–Godfrey	0.12	0.88
$x_{1,t}$	–6.4**	Breusch–Pagan–Godfrey	0.91	0.44
$x_{2,t}$	–0.24**	White test	0.63	0.57
R^2 adj	0.92			
MAPE	6.43%			

Significance codes: «***» – 0.001; «**» – 0.01; «*» – 0.5

Modeling the volume of funds on corporate settlement accounts

The annual growth rate of investments in fixed assets is a factor determining the amount of funds in the corporate settlement accounts of banks of the second cluster. The results of the model are presented in *Table 11.*

The following designations are used in the table:

y_t – the volume of funds on corporate settlement accounts;

$x_{1,t}$ – the growth rate of investment in fixed assets at time t (in %, compared to the same quarter of previous year).

The adequacy of the model is confirmed by tests for the presence of autocorrelation and heteroscedasticity, indicating their absence. A small MAPE and a high value of corrected R^2 indicate the suitability of the model for prediction. The actual and forecasted volumes of funds in the settlement accounts of legal entities of banks of the second cluster in the stress scenario are shown in *Figure 9.*

The forecast values of quarterly growth rates of the analyzed bank indicators of the cluster are given in *Table 12* and in *Figure 10.*

It should be noted that the volume of loan portfolios have a certain tendency to decrease, while growth of corporate and individuals deposits in the stressful period is predicted. This growth can be explained by the fact that banks, mainly state-owned and systemically significant credit organizations, fall into the second cluster. This means that during crisis periods the probability of default of these banks is extremely low, which is a determining factor for depositors when choosing a bank.

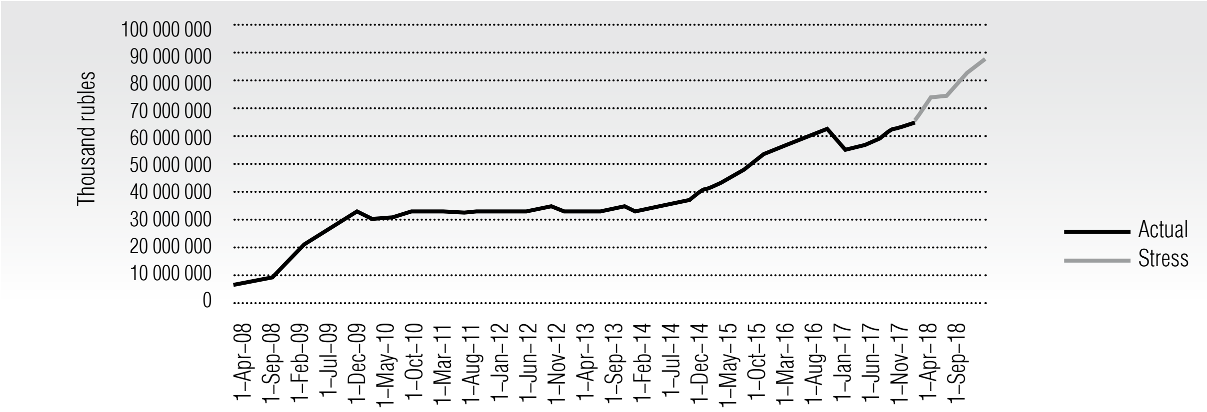


Fig. 7. The actual and forecasted values corporate LLP

Table 10.

Results of modeling funds on individuals' current accounts

Variable	Coefficient	Test	Statistics	*p*-value
constant	34.07**	Darbin–Watson	2.13	
pcy(y_{t-1})	0.54***	Breusch–Godfrey	0.36	0.68
$x_{1,t}$	–2.93**	Breusch–Pagan–Godfrey	0.37	0.69
R^2 adj	0.68	White test	0.45	0.8
MAPE	9.24%			

Significance codes: «***» – 0.001; «**» – 0.01; «*» – 0.5

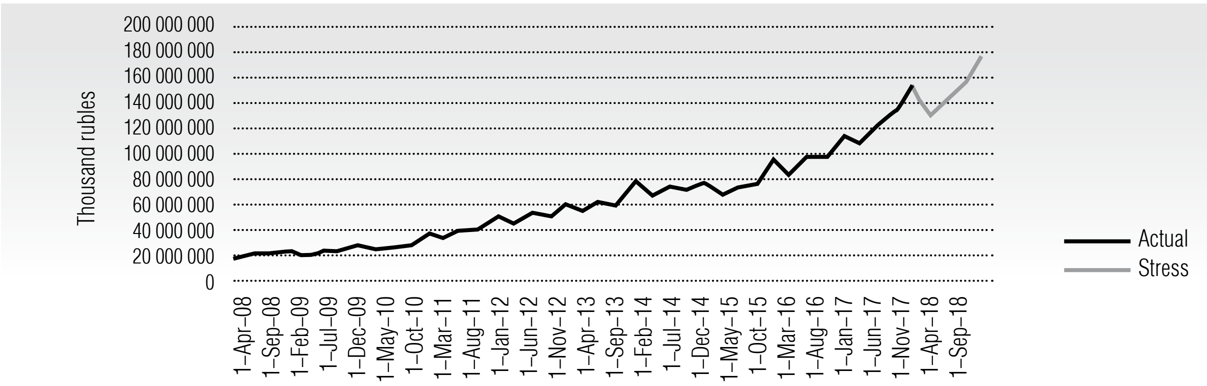


Fig. 8. The actual and forecasted values of funds on individuals' current accounts

3.2. Stress test results

Of the 26 banks of the second cluster, nine are systemically important credit organizations. According to the results of the stress testing of the banks of the second cluster, 11 banks (of which three are systemically significant), were observed to violate capital adequacy requirements. *Table 13* shows the values of capital adequacy prior to the start of stress testing (as of 01 January 2018 – "before stress") and according to the results of modeling a stressful situation (as of 01 January 2019 – "after stress").

Figure 10 shows graphs of average values of the capital adequacy $H_{1.0}$ (%), basic capital $H_{1.1}$, (%) and fixed capital $H_{1.2}$ (%) before the start of stress testing (as of 01 January 2018 – "before

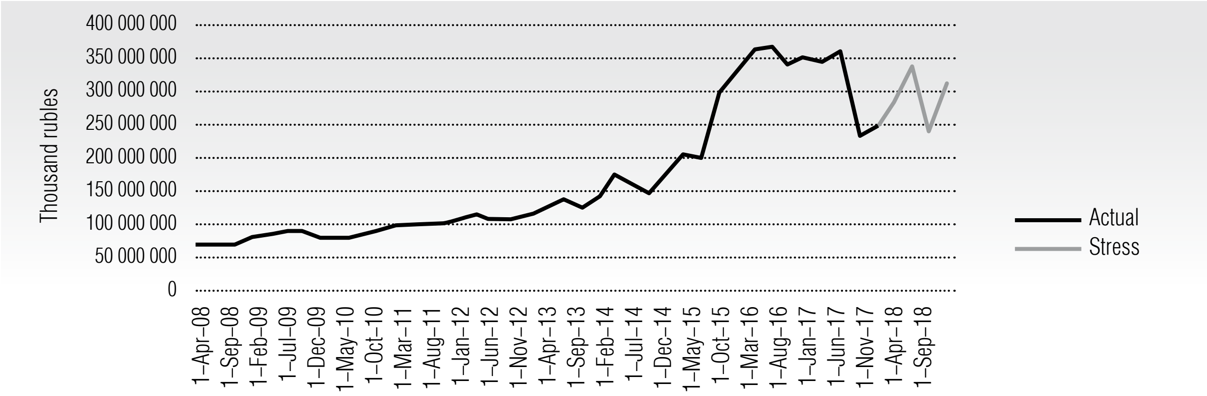


Fig. 9. The actual and forecasted values of funds on corporate settlement accounts

Table 11.

Results of modeling funds on corporate settlement accounts

Variable	Coefficient	Test	Statistics	*p*-value
constant	4.93*	Darbin–Watson	1.77	
pcy(y_{t-1})	0.71***	Breusch–Godfrey	0.71	0.49
$x_{1,t}$	–0.71*	Breusch–Pagan–Godfrey	1.45	0.24
R^2 adj	0.58	White test	0.57	0.71
MAPE	13.44%			

Significance codes: «***» – 0.001; «**» – 0.01; «*» – 0.5

Table 12.

Predicted values of quarterly growth rates of banking indicators of the second cluster in a stress scenario

Quarter of 2018	Retail loans	Corporate loans	Individuals' deposits	Corporate deposits	Retail LLP	Corporate LLP	Individuals' funds	Corporate funds
1	–0.16	–0.89	3.26	3.80	24.05	13.73	–16.12	15.27
2	–0.26	–1.17	2.85	2.43	8.48	1.22	8.16	19.33
3	0.71	0.71	2.52	1.09	4.21	10.13	9.60	–29.40
4	–1.03	0.39	2.68	0.03	4.32	6.77	15.63	30.98

stress") and based on the results of stress modeling (as of 01 January 2019 – "after stress"), since the forecasting horizon is 4 quarters), for banks forming the second cluster.

As can be seen from *Figure 10*, the adequacy of all types of capital for banks of the second cluster, on the basis of stress testing of credit risk, decreased on average by about 2%. Thus, the adequacy of total capital decreased from 13.37% to 11.32%, the adequacy of basic capital from 9.93% to 7.78%, and the capital adequacy ratio from 10.19% to 8.05%.

Table 13.

The results of stress testing of credit risk of banks of the second cluster

License Number	Bank	H1.0 (%)		H1.1 (%)		H1.2 (%)		Requirements violation	Capital deficit (thousands of rubles)
		Before stress	After stress	Before stress	After stress	Before stress	After stress		
1000	**VTB Group**	**11.28**	**10.59**	**8.87**	**8.19**	**9.09**	**8.39**	**1**	**3 927 623**
1326	Alfa Bank	12.04	9.90	7.88	5.62	9.14	6.91	0	0
1343	Bank Levoberezhny	14.97	12.67	9.66	7.15	9.66	7.15	0	0
1439	Vozrozhdenie Bank	13.53	10.41	9.42	6.10	9.42	6.10	0	0
1481	Sberbank	14.97	13.72	10.72	9.40	10.72	9.40	0	0
1637	**SDM-Bank**	**14.81**	**14.12**	**11.89**	**11.17**	**11.89**	**11.17**	**1**	**5 711 336**
1792	Rusfinance Bank	**13.41**	**10.71**	**13.31**	**10.60**	**13.31**	**10.60**	**1**	**4 953 08**1
1942	Globexbank	**14.11**	**10.99**	**13.27**	**10.11**	**13.27**	**10.11**	**1**	**12 979 623**
2210	TKB Bank	**11.08**	**8.39**	**7.70**	**4.88**	**7.70**	**4.88**	**1**	**3 130 759**
2272	Rosbank	13.10	11.93	9.22	7.96	9.22	7.96	0	0
2289	**Russian Standard Bank**	**12.50**	**7.78**	**9.54**	**4.53**	**9.54**	**4.53**	**1**	**98 297**
2306	Absolute Bank	12.36	10.70	8.35	6.60	8.35	6.60	0	0
2312	Bank DOM.RF	10.55	8.30	7.68	5.33	7.68	5.33	0	0
2440	Metallinvestbank	12.75	11.59	8.98	7.75	8.98	7.75	0	0
2590	**Ak Bars Bank**	**14.72**	**13.40**	**10.53**	**9.14**	**10.53**	**9.14**	**1**	**1 768 125**
2707	Loko Bank	13.89	12.63	12.01	10.71	12.01	10.71	0	0
316	Home Credit Bank	13.92	12.54	9.92	8.46	9.92	8.46	0	0
328	AB Rossia	12.01	11.23	7.77	6.95	8.43	7.61	0	0
3292	**Raiffeisen Bank**	**13.27**	**12.14**	**9.94**	**8.75**	**10.63**	**9.45**	**1**	**42 362 841**
3311	**Credit Europe Bank**	**14.41**	**10.45**	**11.80**	**7.65**	**11.80**	**7.65**	**1**	**1 586 504**
3349	Russian Agricultural Bank	15.55	13.99	10.41	8.74	10.94	9.28	0	0
354	**Gazprombank**	**12.75**	**11.48**	**8.70**	**7.39**	**9.07**	**7.75**	**1**	1 **330 903**
436	Saint–Petersburg Bank	14.27	11.91	10.05	7.55	10.05	7.55	0	0
588	**SNGB Bank**	**15.58**	**12.09**	**11.91**	**8.19**	**11.91**	**8.19**	**1**	**3 927 623**
705	SKB Bank	12.79	9.61	8.72	5.21	10.71	7.32	0	0
963	Sovkombank	12.94	11.15	10.03	8.16	11.02	9.18	0	0

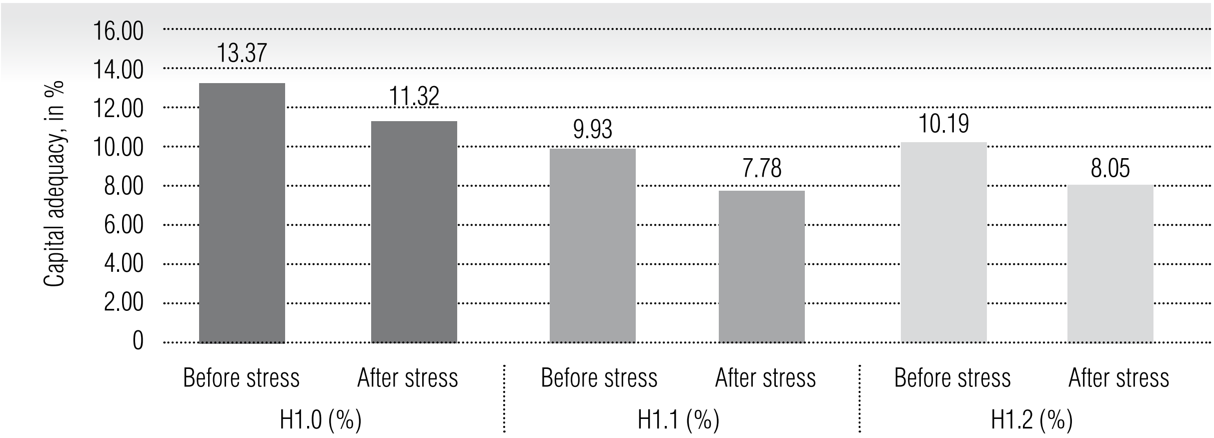


Fig. 10. The average values of the capital adequacy of all types of capital before the start of stress testing (as of 01 January 2018 – "before stress") and based on the results of stress modeling (as of 01 January 2019 – "after stress")

Conclusion

The proposed algorithm allows all counterparties of banks that have access only to public financial reports to conduct credit risk stress testing of the banks' cluster of interest, which can be used to evaluate which banks are most reliable in crisis situations. Using the example of credit risk stress testing conducted for a cluster of 26 banks, eight econometric models of banking indicators were developed, representing the median values of banking indicators included in this cluster of banks. According to the results of stress testing, 11 out of 26 banks forecast a deficit of all types of capital. ■

References

1. Bidzhoyan D., Bogdanova T. (2018) Russian banks credit risk stress-testing based on the publicly available data. *Advances in Intelligent Systems and Computing*, vol. 850, pp. 262–271.
2. Arena M. (2008) Bank failures and bank fundamentals: A comparative analysis of Latin America and East Asia during the nineties using bank-level data. *Journal of Banking and Finance*, no 32, pp. 299–310.
3. Makinen M., Solanko L. (2018) Determinants of banks closure: Do levels or changes of CAMEL variables matter? *Russian Journal of Money and Finance*, vol. 77, no 2, pp. 3–21.
4. Mitchell T. (2015) Bank default prediction: A comparative model using principal component analysis. *Journal of Stock and Forex Trading*, vol. 4, no 2, pp. 677–689.
5. Karminsky A.M., Kostrov A.V. (2014) Modelling the likelihood of Russian bank's default: extended opportunities. *Journal of the New Economic Association*, vol. 17, no 1, pp. 64–86 (in Russian).
6. Peresetsky A.A. (2010) *Models of Russian bank's license revocation reasons.* Working paper WP/2010/085. Moscow: RES (in Russian).
7. Karmisky A., Khromova E. (2016) Modelling banks' credit rating of international agencies. *Eurasian Economic Review*, vol. 6, no 3, pp. 341–363.
8. Goldewski C.J. (2007) Are ratings consistent with default probabilities? *Empirical evidence on banks in emerging market economies. Emerging Markets Finance and Trade*, vol. 43, no 4, pp. 5–23.
9. Avrikan N., McCrystal A. (2012) Sensitivity analysis of network DEA: NSBM versus NRAM. *Applied Mathematics and Computation*, vol. 218, no 22, pp. 11226–11239.
10. Karas A., Schoors K., Weill L. (2010) Are private banks more efficient than public banks? Evidence from Russia. *Economics of Transition*, vol. 18, no 1, pp. 209–244.
11. Gonzales F. (2005) Bank regulation and risk-taking incentives: An international comparison of bank risk. *Journal of Banking and Finance*, vol. 29, no 5, pp. 1153–1184.

12. Ungan E., Caner S., Ozyildirim S. (2008) Depositors's assessment of bank riskiness in the Russian Federation. *Journal of Financial Services Research*, vol. 33, no 2, pp. 77–110.
13. Peresetsky A.A. (2012) *Econometric methods in off-site analysis of Russian banks.* Moscow: HSE (in Russian).
14. Blaschke W., Jones T., Magnoni G., Peria S.-M. (2001) *Stress-testing of financial systems: An overview of issues, methodologies, and FSAP experience.* IMF Working Paper WP/01/88.
15. Melecky M. Podperia A. (2010) *Macro prudential stress-testing practices of central banks in central and eastern Europe: an overview and challenges ahead.* World Bank Policy Research Working Paper WP/2010/09/01.
16. Foglia A. (2009) Stress testing credit risk: a survey of authorities approaches. *International Journal of Central Banking*, vol. 5, no 3, pp. 9–45.
17. Breuer T., Summer M. (2017) Solvency stress testing of banks: *Current practice and novel options.* Report for the Sveriges Riksbank and Finansinspektionen.
18. Breuer T. Csiszar I. (2013) Systematic stress test with entropic plausibility constraints. *Journal of Banking and Finance*, no 37, pp. 1552–1559.
19. Breuer T., Csiszar I. (2016) Measuring distribution model risk. *Mathematical Finance*, vol. 26, no 2, pp. 395–411.
20. Kapinos P., Mitnik O. (2016) A top-down approach to stress-testing of banks. *Journal of Financial Services Research*, vol. 49, no 2, pp. 229–264.
21. Pritsker M. (2017) *Choosing stress scenarios for systemic risk through dimension reduction.* Federal Reserve Bank of Boston Working Paper WP17-04.
22. Shleifer A. Vishny R. (1992) Liquidation values and debt capacity: a market equilibrium approach. *Journal of Finance*, vol. 47, no 4, pp. 1343–1366.
23. Cont R., Wagalath L. (2016) Fire sales forensics: Measuring endogenous risk. *Mathematical Finance*, vol. 26, no 4, pp. 835–866.
24. Pedersen L. (2009) When everyone runs for exit. *International Journal of Central Banking*, vol. 5, no 4, pp. 177–199.
25. Bidzhoyan D.S., Bogdanova T.K., Neklyudov D.Yu. (2018) Evaluating reliability of a bank as an investment object. *St. Petersburg State Polytechnical University Journal. Economics*, vol. 11, no 4, pp. 70–84 (in Russian).

About the authors

Davit S. Bidzhoyan

Lecturer, Department of Business Analytics, National Research University Higher School of Economics, 20, Myasnitskaya Street, Moscow 101000, Russia;

E-mail: bidzhoyan_david@mail.ru

ORCID: 0000-0002-3668-1691

Tatiana K. Bogdanova

Cand. Sci. (Econ.);

Associate Professor, Department of Business Analytics, National Research University Higher School of Economics, 20, Myasnitskaya Street, Moscow 101000, Russia;

E-mail: tanbog@hse.ru

ORCID: 0000-0002-0018-2946

Dmitry Yu. Neklyudov

Senior Lecturer, Department of Business Analytics, National Research University Higher School of Economics, 20, Myasnitskaya Street, Moscow 101000, Russia;

E-mail: nekludovmid@gmail.com

ORCID: 0000-0002-9165-280X

DOI: 10.17323/1998-0663.2019.3.52.66

Comparative analysis of methods for forecasting bankruptcies of Russian construction companies

Alexander M. Karminsky
E-mail: karminsky@mail.ru

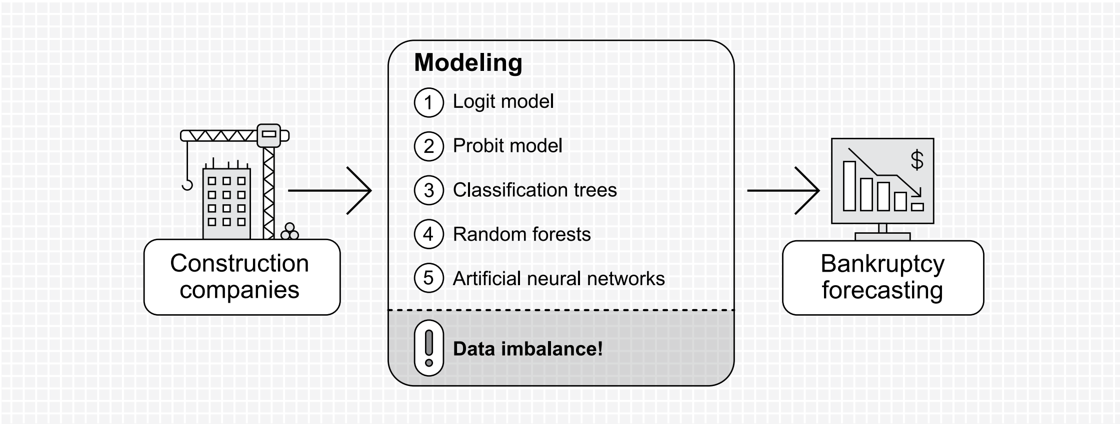
Roman N. Burekhin
E-mail: romanvia93@yandex.ru

National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow 101000, Russia

Abstract

This paper is devoted to comparison of the capabilities of various methods to predict the bankruptcy of construction industry companies on a one-year horizon. The authors considered the following algorithms: logit and probit models, classification trees, random forests, artificial neural networks. Special attention was paid to the peculiarities of the training machine learning models, the impact of data imbalance on the predictive ability of models, analysis of ways to deal with these imbalances and analysis of the influence of non-financial factors on the predictive ability of models. In their study, the authors used non-financial and financial indicators calculated on the basis of public financial statements of the construction companies for the period from 2011 to 2017. The authors concluded that the models considered show acceptable quality for use in forecasting bankruptcy problems. The Gini or AUC coefficient (area under the ROC curve) was used as the quality markers of the model. It was revealed that neural networks outperform other methods in predictive power, while logistic regression models in combination with discretization follow them closely. It was found that the effective way to deal with the imbalance data depends on the type of model used. However, no significant impact on the imbalance in the training set predictive ability of the model was identified. The significant impact of non-financial indicators on the likelihood of bankruptcy was not confirmed.

Graphical abstract



Key words: bankruptcy; the construction sector; imbalance data; machine learning models; parametric models of prediction of bankruptcy.

Citation: Karminsky A.M., Burekhin R.N. (2019) Comparative analysis of methods for forecasting bankruptcies of Russian construction companies. *Business Informatics*, vol. 13, no 3, pp. 52–66. DOI: 10.17323/1998-0663.2019.3.52.66

Introduction

In a market economy, forecasting financial insolvency is an important task for any company. To achieve this goal, different methods of assessing credit risks are used. Their purpose is to proactively and effectively forecast the onset of an adverse situation in the company. Typically, these methods are parametric models characterized by a relatively simple mathematical apparatus and a simple qualitative interpretation. These methods are static, do not take into account subtle economic or behavioral factors, and the predictive ability of the models decreases with the nonlinear nature of the relationships between the indicators.

Market models (structural models and shortened models) are often too complex or market dependent. To apply them, you need access to a large amount of data (market value of share capital, debt obligations, spreads of bond yields, etc.) Despite the widespread use of market models by Western companies, their use in the Russian market is difficult due to the small number of listed securities. To conduct an effective credit policy, new methods must be flexible and adaptable to the changing realities of a market economy. Therefore, there is currently an interest in models based on machine learning algorithms, including classification trees, random forests, gradient boosting, artificial neural networks, etc.

There are a number of common problems associated with predicting bankruptcy of companies. Firstly, the economic indicators describing the state of the company differ in various studies, and their integration into the most effective model causes additional difficulties. Secondly, there is a problem of data imbalance, since there are more solvent companies than bankrupt ones. As a result, the trained model tends to classify companies as reliable, although they may have signs of financial failure. Thirdly, the very concept of “bankruptcy” can be interpreted in different ways, so different companies can fall into this category. In this work, the category of bankrupts includes companies in respect of which legal bankruptcy proceedings have begun, as well as companies that have liquidated voluntarily.

Despite the importance of the task of forecasting bankruptcies using more advanced methods, there are not so many domestic works in this area, and works on forecasting reviews of bank licenses are more likely to be the exception [1, 2]. A feature of this work is the comparison of regression models and models based on machine learning methods in the tasks of predicting bankruptcies of companies based on one industry. Considerable attention is paid to the specifics of building machine learning models, the impact of data imbalances, as well as non-financial indicators on the predictive ability of models.

The construction industry is a link between other industries, which determines its importance in the national economy. Today in Russia there are more than two hundred and seventy thousand companies performing certain construction work (design, engineering calculations, construction, etc.). Their number, as well as the high level of defaults in this sector makes it difficult to choose a suitable partner. This industry is one of the most affected by the crisis. In particular, the volume of work in comparable prices has not ceased to fall since 2014, and by the end of 2017, construction turned out to be an industry with one of the highest share of bad debts. Lending to the construction sector represents a significant part of the Russian banking business. Therefore, an increase in the number of insolvent construction companies can cause instability in the banking sector. Moreover, national and international regulatory requirements (recommendations of the Basel Committee) force the use of an advanced approach based on internal ratings to quantify risks in order to reduce the burden on capital. Therefore, the problem of forecasting the future state of construction companies is relevant, and new tools for forecasting bankruptcies are in demand.

This paper answers the question whether models based on machine learning methods can be a worthy alternative to regression models when applied to the field of bankruptcy forecasting of companies in the non-financial sector, using the construction industry as an example. It is concluded that all the considered models are capable of predicting bankruptcy in the next 12 months, while neural networks are superior to other methods in identifying insolvent companies, and logistic regression models combined with discretization closely follow them. A negative effect of the imbalance of the training set on the predictive ability of the model was not found[1].

1. Models for predicting financial insolvency

Regression models (logit and probit models) are common in the problems of identifying solvent and insolvent borrowers [3]. Their advantage lies in the absence of severe restrictions on functioning, ease of interpretation and simplicity of calculations. An important drawback of these models is a decrease in prognostic ability with the non-linear nature of the relationships between the indicators, while machine learning algorithms are less sensitive to these problems. There are many works proving the possibility of using advanced methods for predicting company insolvency [4–7].

The authors [8] were among the first to use classification trees to predict company bankruptcies. They found that their classification trees outperform discriminant analysis. It was

[1] Preliminary results of the study were presented in the graduate work by Roman N. Burekhin, performed at the HSE Faculty of Economic Sciences in 2018

also noted that with the complication of the model (inclusion of a larger number of factors), its accuracy deteriorated due to overfitting. However, this success did not cause the widespread use of decision trees in this area. In the future, in most works, there is a comparison of the effectiveness of decision trees with other algorithms. The random forest algorithm was presented in [9] and applied in many areas: from marketing (predicting customer loyalty to a brand) and the criminal sphere (predicting homicide or relapse among parole), to credit scoring. Based on the financial reporting data, the authors [4] successfully use random forest models for forecasting defaults of companies from seven European countries (Finland, France, Germany, Italy, Portugal, Spain and the UK). In 1990, the authors [10] were among the first to use the neural network in predicting bankruptcies. A neural network was built with several hidden layers and using financial coefficients used in the Altman model as input. At the same time, the share of correctly classified companies was about 80%.

These algorithms often show higher efficiency, despite the fact that they are characterized by significant time and physical costs. Moreover, at present, there is a tendency in which algorithms based on one method are losing popularity, while ensemble or hybrid models are becoming more popular and demonstrate higher efficiency [11].

Since the 1970s, financial ratios derived from financial statements have been an important source for constructing default forecasting models. However, models based on accounting information are criticized because of the historical nature of the information used as input and not taking into account the volatility of the value of the company during the period analyzed. However, proponents of this approach argue that the inefficiency of capital markets can lead to more significant errors in predicting credit risks. In article [12], credit risk assessment models based on accounting and market information are compared. The authors conclude that the approaches considered do not have significant differences in the predictive ability, while these types of data are complementary, and the complex model shows the best result.

It can be concluded that market information can be a significant factor in predicting company insolvency. However, due to the fact that most of the companies examined do not have access to the stock market, financial statements become the only available source of information, and the use of market models becomes impossible.

2. Data description

The main data source in the work was the SPARK system (Interfax agency). Information about the default of companies was used in the "Unified Federal Register of Bankruptcy Information" database. In the study, the following companies were classified as construction companies (classification in SPARK):

✦ building;

✦ construction of engineering structures;

✦ specialized construction work (development and demolition of buildings, preparation of the construction site, finishing construction work).

An important issue is the definition of an insolvent company. In accordance with the Federal Law of October 26, 2002 No. 127-ФЗ (dated December 29, 2017) "On Insolvency (Bankruptcy)" [13], one sign of bankruptcy is considered to be a situation where the demands of creditors on monetary obligations are not fulfilled within three months from the date they were to be executed. The following definitions of bankruptcy are widespread in research: a company is not able to pay interest on a debt or part of its principal debt, the organization is monitored (a procedure that analyzes the financial situation and solvency of a debtor, as well as its ability to pay off debt), the company is not active for a long period of time, the company is in a state of liquidation. In our work, the category of bankrupts included companies in respect of which the legal bankruptcy procedure was launched, as well as companies that liquidated voluntarily. A similar classification is given in [4, 14]. It is noted that these companies are characterized by a critical financial situation and are often unable to fulfill their obligations.

Based on the financial statements, the values of fourteen coefficients reflecting the economic activity of the enterprise were calculated. At the same time, the following classification of financial indicators was proposed: profitability, liquidity, business activity, financial stability. A similar classification is given in [3]. Also included in the model are non-financial factors that reflect the size and age of the company. The variables are described in *Table 1.*

Table 1.

Variable description

Group	Variables	Variable description
Dependent variable	Bankruptcy	1 – if a default occurred in the next reporting period; 0 – otherwise
Profitability	Return on assets (ROA)	Net profit to assets ratio
	Return on equity (ROE)	Net profit to equity ratio
	Return on sales (ROS)	Net profit to revenue ratio
	Operating margin	Operating profit to revenue ratio
Liquidity	Current ratio	Current assets to current liabilities ratio
	Quick ratio	Receivables, financial investments and cash to current liabilities ratio
	Equity maneuverability ratio, net working capital (NWC) ratio	The difference between equity and non–current assets to equity ratio
Business activity	Accounts receivable (AR) turnover ratio	Revenue to receivables ratio
	Accounts payable (AP) turnover ratio	Cost of sales to accounts payable ratio
	Assets turnover ratio	Revenue to assets ratio
	Share of non–current assets	Non–current assets to total assets ratio
Financial stability	Autonomy ratio	Equity to assets ratio
	Share of retained earnings in revenue	Retained earnings to revenue ratio
	Interest coverage ratio (ICR)	Profit before tax and interest payable to interest payable ratio
Company size	Logarithm of company's assets	Assets logarithm
Age	Age	

For the analysis of default events, the time range of 2011–2017 was chosen. The time horizon was divided into two blocks: a training sample (period 2011–2015) and a test sample (period 2016–2017). At the next stage, the following selection procedure was carried out:

1) removal of observations with missing data (for example, for which there is no information on the value of assets and revenue) or filling in gaps in the data, where possible;

2) removal of observations with obvious errors (for example, where the size of assets or the size of receivables is negative);

3) identification and removal of outlier observations, since their presence leads to biased results. The main algorithm used for this procedure is the three sigma rule.

As a result, 3981 organizations fell into the final sample, 390 of which defaulted. The training sample (period from 2011 to 2015) included 3300 construction companies, of which 325 defaulted. The test sample (the period from 2016 to 2017) included 681 companies, of which 65 defaulted. *Figure 1* shows the total number of companies and the number of companies that went bankrupt in the next reporting year, by year.

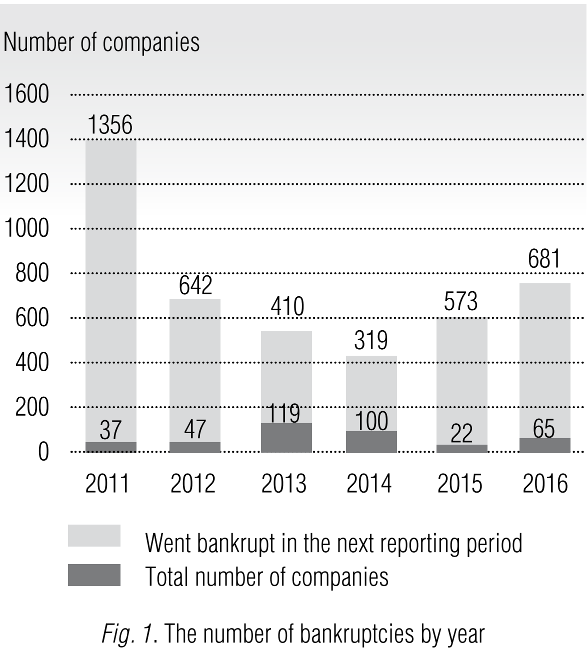


Fig. 1. The number of bankruptcies by year

The research data set is unbalanced (only 9.8% of companies defaulted). Therefore, when constructing the models, two techniques for working with unbalanced data were used: undersampling and oversampling.

Undersampling involves the use of input data containing all insolvent companies and a random selection of solvent companies. As a result, the proportion of insolvent to solvent companies increases. Also, when constructing such dependencies, it is recommended that this experiment be performed several times to obtain consistent results (in this study, the assumption is made that a consistent result is obtained after one experiment). Oversampling involves the use of input data containing all solvent companies, and "cloning" of insolvent companies until their number approaches the number of solvent companies. The search for the optimal share of the minority class in the training set is also the subject of research in this paper.

Cross-validations were used to find the optimal value of the share of insolvent companies in undersampling. When using oversampling, this approach is not recommended, since in this case, we see cloning of information which is used both in training and in testing the model (which leads to overfitting). To implement oversampling, the training set was divided into two subsets. The first (company default information for 2014) was used to test models, the second (remaining periods from 2011 to 2015) – to build models without cross-validation. The share of insolvent companies, in which the model is most important for the learning set, was used to compare models on the test set (2016–2017).

3. Description of models

Two parametric algorithms for constructing binary choice models were used in the work: logit and probit models with sampling corrections (using WOE) and without. These models are compared with algorithms based on machine learning methods (classification trees, random forests, artificial neural networks) which are described in the following sections.

Traditionally, the following metrics of model quality are distinguished: accuracy, sensitivity, specificity, area under the ROC curve, Gini coefficient, F-metric. The use of these metrics depends on the purpose of the analysis.

In this study, each of the models considered at the output has a range of values from 0 to 1; therefore, it is necessary to determine the cutoff threshold. The assignment of the cutoff threshold depends on the analyst's preferences regarding errors of the first and second kind, which leads to difficulties in comparing different models. Therefore, in this paper, the predictive power is estimated using ROC analysis, AUC (the area under the ROC curve), or Gini coefficient. The advantage of these metrics is that there is no need to determine the cutoff threshold and the ability to compare the quality of models regardless of the analyst's goals. The calculation of the Gini coefficient was carried out as follows:

$$Gini = (AUC - 0.5) \cdot 2 \cdot 100\%, \qquad (1)$$

where AUC is the area under the ROC curve.

Visual analysis of the effectiveness was carried out using the ROC curve. The greater the bend of the ROC curve, the higher the quality of the model, while the diagonal line corresponds to the complete indistinguishability of the two classes. Accordingly, the higher the value of the area under the ROC curve, the better the separation power of the model. Analysis of the ROC curve allows the user to select the ratio between sensitivity and specificity necessary for analysis. An example of constructing an ROC curve for one variable is presented in *Figure 2.*

In the work, as the calculation tool for econometric analysis and reflection of statistical conclusions, we used the programming language R, which is a free open source software environment.

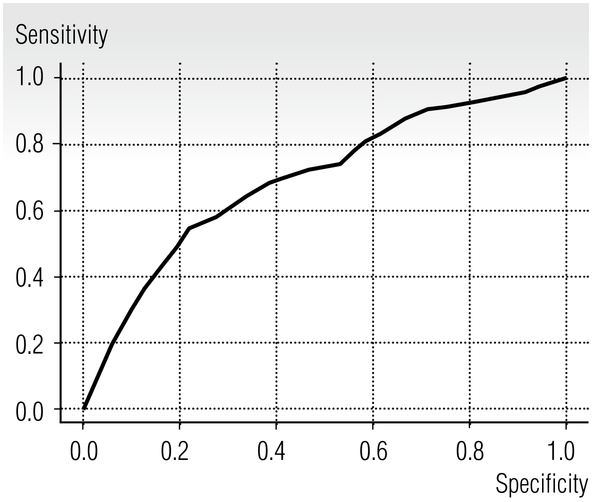


Fig. 2. ROC curve of the ROA factor

3.1. Binary selection models

Two algorithms for constructing binary selection models (logit and probit) were used in the work: without corrections for discretization and with corrections. Here is a description of the general algorithm (with discretization amendments) for constructing binary choice models.

Step 1. Reduction factors to the discrete form. In the process of solving a research question, most authors are faced with the problem of outliers. The given problem is no exception for this work. The traditional approach to solving it is the exclusion of such observations. However, the subjectivity of outlier determination and sample reduction are significant disadvantages of this approach. The paper uses the transition from discrete to continuous form, which leads to increased comparability of factors among themselves and the unity of approaches to assess the significance of factors. We carried out the quantile discretization procedure – replacing the initial values of factors with discrete values based on grouping by quantiles. The essence of this approach is as follows:

a) the values of the variables are ordered in ascending order;

b) the values of each indicator are divided into 10 groups (deciles are used);

c) the values in each group are replaced by points from 1 to 10 (the group with the lowest values gets 1 point, and the group with the highest values gets 10 points).

Stage 2. Transformation of factors. To convert factors, the WOE (weight of evidence) approach is used. The WOE indicator characterizes the degree of deviation of the level of defaults in this group from the average value in the sample. For each factor and for each group within the factors, it is necessary to calculate the number of companies in default and the number of companies not in default. The WOE for group i of a particular variable is calculated as follows:

$$WOE_i = \ln\left(\frac{d_i^{(1)}}{d_i^{(2)}}\right), \qquad (2)$$

where $d_i^{(1)}$ – the share of non-default companies belonging to group i in the total number of non-default companies; $i = 1, 2, ..., k$; k – number of variable categories;

$d_i^{(2)}$ – the share of companies in default owned by group i in the total number of companies in default; $i = 1, 2, ..., k$; k – number of variable categories.

In order to increase the linearity of variables and improve the accuracy of the model, all explanatory variables are replaced by WOE, which is a common technique in credit scoring [15].

Stage 3. Assessment of the predictive power of factors. After all values are converted to WOE, it is necessary to evaluate the importance of each factor. Two algorithms for assessing the significance of factors were used in the work: information value (information value, *IV*) and ROC analysis. The calculation of the value of information value (*IV*) is performed according to the following formula:

$$IV = \sum_{i=1}^{k}\left(d_i^{(1)} - d_i^{(2)}\right)\cdot WOE_i, \qquad (3)$$

where k – the number of categories of an independent variable (each factor has ten), the remaining notation – from formula (2).

Formula (3), which reflects calculation *IV*, is based on the summation of WOE_j, adjusted for the difference $\left(d_i^{(1)} - d_i^{(2)}\right)$. The main purpose of these calculations is to identify some indicator that reflects the ability of a variable to cluster some attribute. If this indicator is above 0.02, then the factor should be used in modeling [15].

The study applied the following criteria for selecting factors in the final model:

✦ acceptable quality of the model in accordance with the criterion of "information value" ($IV > 0.02$);

✦ the Gini coefficient in the one-factor model must be greater than 5%;

✦ economic assessment factor.

Stage 4. The analysis of correlations. When constructing a multi-factor model, factors with high correlation coefficients must be excluded. Correlation analysis avoids multicollinearity. Multicollinearity leads to model instability and increases standard deviations of factor estimates. The presence of multicollinearity is indicated by high values of pair correlation coefficients between the factors of variables. The criterion for determining high correlation may vary; for economic data, the threshold is usually set at 0.30–0.50. The criterion for high correlation in this model is a correlation coefficient greater than 0.5.

Step 5: Multivariate analysis. The modeling of the probability of the borrower's non-creditworthiness was carried out as follows:

$$\mathrm{P}\left(Y_i = 1|x_1, \ldots, x_n\right) = F\left(a_0 + a_1 x_1 + \ldots + a_n x_n\right) = = F\left(a_0 + \mathbf{x}'a\right). \qquad (4)$$

In the case of the logit model, F (*) represents the logistic distribution function:

$$F\left(a_0 + \mathbf{x}'a\right) = \Lambda\left(a_0 + \mathbf{x}'a\right) = \frac{e^{a_0 + \mathbf{x}'a}}{1 + e^{a_0 + \mathbf{x}'a}}. \qquad (5)$$

In the case of the probit model, F (*) is a normal distribution function:

$$F(a_0 + \mathbf{x}'a) = \Phi(a_0 + \mathbf{x}'a) = \int_{-\infty}^{a_0+\mathbf{x}'a} \varphi(\upsilon)d\upsilon, \quad (6)$$

where $\varphi(\upsilon) = \frac{1}{\sqrt{2\pi}} e^{\left(-\upsilon^2/2\right)}$.

The calculation of the coefficients is carried out by the maximum likelihood method, which maximizes the probability of the joint implementation of events (solvency and insolvency). The standard error of the coefficients was estimated with a Newey–West correction for heteroskedasticity and first-order autocorrelation.

Step 6. Model specification. To select the optimal combination of factors, the Backward Selection method was used – sequential exclusion of factors (i.e. insignificant variables are sequentially excluded from the model, which includes all factors selected in the one-factor analysis). At the same time, the level of statistical significance is tested using p-value calculated according to the results of logistic regression. As a result, factors with p-value less than 10% were selected.

Stage 7. Validation of the model. Choosing the best model. The problem of overfitting requires a model validation procedure. This problem is manifested in the fact that the "trained" model has good results on the training sample, but does not give accurate forecasts for the test sample. To solve this problem, two approaches were used.

The first approach is the "mixing algorithm", the idea of which is as follows:

1. 80% of the companies from the training set are randomly selected;

2. The coefficients of the model are estimated;

3. It is evaluated whether signs are preserved at coefficients, and whether the factors considered are significant;

4. Steps 1–3 are repeated 1000 times; the stability of the signs is checked.

Based on the results obtained, it can be concluded whether the signs of the coefficients for all variables are stable, and how the sign of the coefficient depends on the initial sample.

The second approach is ROC analysis. Analysis of the values of AUC and Gini on the test set helps to make a conclusion about the quality of the models obtained.

3.2. Machine learning models

Logistic analysis (as well as probit analysis) are traditional popular tools for predicting bankruptcies, but they have a number of disadvantages associated with low predictive power, the presence of restrictions on use. Therefore, at the moment, machine learning algorithms have become widespread.

Classification trees. Today classification trees are the foundation for building more complex machine learning algorithms, such as random forests and boosting algorithms (GBM, XGBoost). In this paper, the CART algorithm (classification and regression trees) is considered. A distinctive feature of this algorithm is that it provides only two possible options for the development of the event, which is suitable for realizing the purpose of this study. The main idea of CART is to split the primary set into two subsets so that the bankrupt companies are in one set, and solvent organizations are in the other. The difficulty in using this method is to determine the moment of stopping the "splitting of sets," since the problem of overfitting arises. The following stopping rules are distinguished:

✦ the measure of "purity" is less than a certain value;

✦ restriction on the number of nodes or layers of a tree;

✦ size of the parent node;

✦ size of the descendant node.

The rules themselves are set using cross validation. Despite the fact that there are a number of examples of the successful use of this method in forecasting defaults [5], this method has several disadvantages: high sensitivity to input data, susceptibility to overfitting and the difficulty of determining the optimal tree architects.

Random forests. Random forests appeared as a modification of decision trees and, accordingly, often provide more accurate predictive results. Random forests consist of a user-defined number of classification trees that are generated using a modified CART algorithm. The scheme of this algorithm is presented in *Figure 3.* Two approaches were used in the algorithm: each tree is trained on its own subsample of initial data (bootstrapped data); different subsets of factors are used in constructing classification trees. These actions lead to the construction, and then to the "voting of trees" regarding the belonging of an object to a certain class.

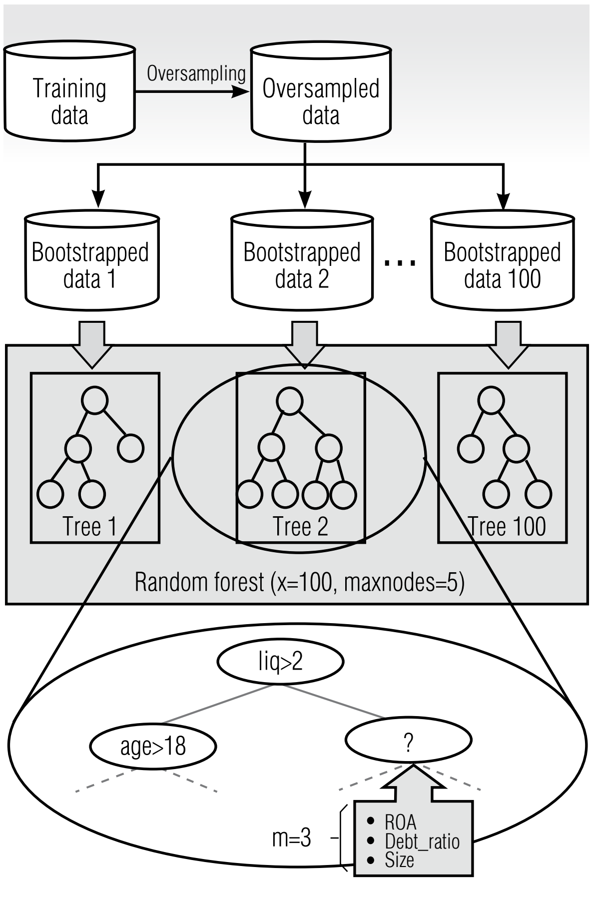


Fig. 3. Random forest algorithm [4]

Unlike regression models, which are quite sensitive to outliers, random forest (RF) is more robust to this problem. The advantage of random forest is higher efficiency in case of imbalance of data (which is relevant for our task), as well as less exposure to overfitting. The disadvantage of the algorithm is less transparency (in contrast to classification trees) and, accordingly, lower interpretation. There is relative difficulty in the process of determining the parameters of a random forest. The determination of the parameters (number of trees, number of factors used in building one tree of factors, maximum number of nodes in one tree) was carried out using cross-validation.

Random forests are often used to determine the significance of a variable. The idea of assessing the importance of a factor is based on the fact that a permutation of the values of an important variable should lead to a significant increase in the error rate on the test set.

Artificial neural networks. Currently, neural network modeling is gaining popularity, especially when predicting phenomena with uniform attributes. Using a set of input parameters, the network architecture is selected. When in the simplest version it is represented by three layers. The first layer contains nodes (neurons) for input variables (each neuron has only one input from the external environment). The second layer contains an arbitrary number of "hidden" neurons and is therefore called a hidden layer. The third layer contains neurons that are responsible for the result. Moreover, in the tasks of forecasting bankruptcies, the last layer contains only one neuron. Between the input and hidden neurons, a connection with certain weights is set. For example, for the *j*-th neuron in the intermediate layer and the input data, the following linear dependence will be determined:

$$a_j = \sum_{i=1}^{N} \omega_{ij} \cdot x_i\,, \qquad (7)$$

where a_j – the value of the j-th neuron;

ω_{ij} – weight of j-th neuron with variable x_i.

Each value a_j is converted using some activation function to obtain the actual resulting value z_j of neuron j. Since our study predicts two classes, it is convenient to use the logistic function as an activation function:

$$z_j = f(a_j) = \frac{1}{1+e^{-a_j}}, \qquad (8)$$

where z_j – normalized value of the j-th neuron;

a_j – the value of the j-th neuron.

A similar procedure is carried out for subsequent layers. The z_j values are again weighted, and then converted using the activation function to obtain the result in the final layer. Many minimization methods are distinguished. Their idea is that, starting from the initial value of weight ω^0, a sequence of vectors of weight coefficients ω^1, ω^2, ..., ω^k is generated, such that with each iteration of the algorithm the value of the function of the quality criterion decreases:

$$E\left(\omega^{k+1}\right) < E\left(\omega^{k}\right), \qquad (9)$$

where ω^k – weight value after the k-th step of training;

ω^{k+1} – weight value after the k+1-th step of training.

Figure 4 shows an example of one of the resulting neural network models.

One of the most common methods used to train neural network models is the steepest descent method. In this algorithm, the adjustment of the weights is performed in the direction of the maximum reduction of the quality criterion, i.e. in the opposite direction to the gradient vector. Despite the fact that the steepest descent method converges to the optimum value of ω^* rather slowly, it is a common method of finding the minimum in many statistical libraries.

One of the difficulties of training a neural network is that the quality criterion function can have many local minima. As a result, after the initialization of the model, one can come to a local minimum, which will negatively affect the results obtained on the test set. To overcome this problem, weights are randomly sorted, and the learning algorithm itself is repeated several times. The optimal parameters (the number of neurons in the inner layer, the number of inner layers), as well as for classification trees and random forests, were determined using cross-validation.

To increase the efficiency of the neural network, as well as speed up the learning process, preliminary data processing is necessary. A simple and efficient preprocessing step involves scaling and centering data.

4. Comparative analysis of the models

In accordance with the classification used, the models considered showed good quality. *Table 2* shows the sorting of the best models in the group in descending order of model quality. The best quality was shown by an artificial neural network with one hidden layer and four neurons using the oversampling algorithm. It is reflected that the use of a logistic model with sampling and transition to WOE leads to a significant increase in the accuracy of models (the Gini coefficient increases on average by 15%). It is noteworthy that the quality of the models corresponds to the AUC level in such works [4, 5, 11].

The results of one-way analysis using regression models indicate that all the factors considered can be used to build binary choice models, since for each of them the AUC value for univariate analysis is higher than 0.5. The final multivariate model included eight factors out of sixteen factors reflecting different aspects of the risks of construction companies: liquidity (current ratio, equity ratio), profitability (return on equity), solvency (interest coverage ratio), turnover (asset turnover), business activity (share of non-current assets), non-financial predic-

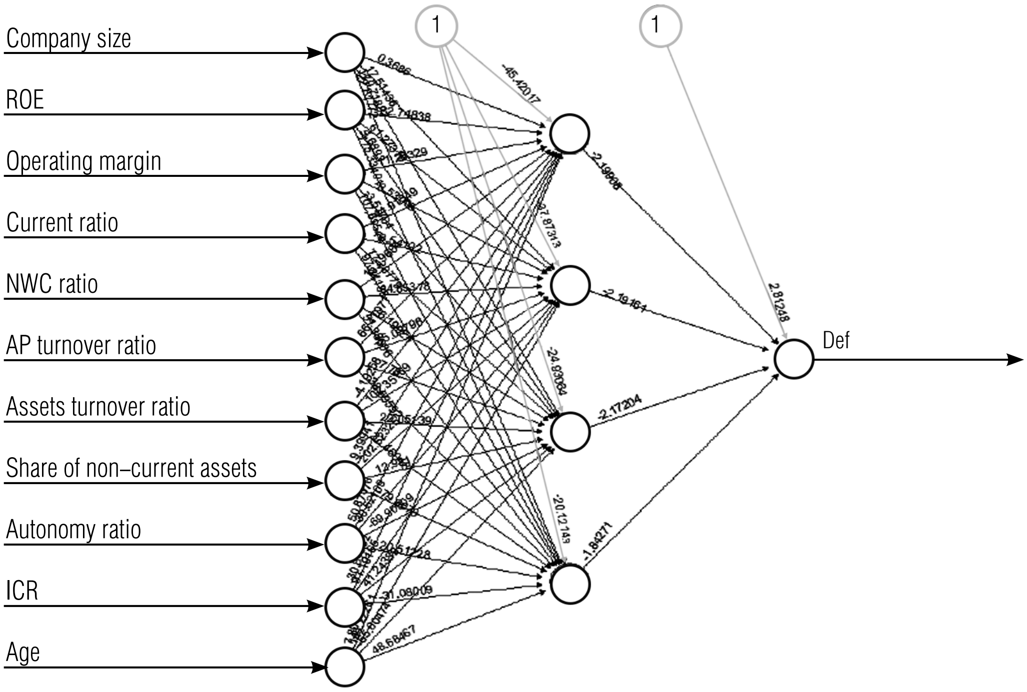


Fig. 4. Example of neural network architecture

Table 2.

Quality assessment of models on a test set

No	Model	Gini coefficient, %	The share of insolvent companies on the test set, %
1	Artificial neural networks (oversample)	59.6	50
2	Artificial neural networks	58.9	9.8
4	Logit model (oversample)	57.9	25
3	Probit model (oversample)	57.6	20
5	Logit model	57.6	20
6	Logit model (undersample)	57.3	20
7	Artificial neural networks (undersample)	56.0	50
8	Random forests (undersample)	52.4	15
9	Random forests	50.6	9.8
10	Random forests (oversample)	48.7	10
11	Classification trees (oversample)	45.0	15
12	Log model without discretization	42.2	9.8
13	Classification trees (with penalties for incorrect classification of a minority class)	40.0	9.8
14	Classification trees	38.0	9.8
15	Classification trees (undersample)	38.0	50

tors (age and size of the company). The inclusion of these factors leads to an increase in the efficiency of traditional models (the Gini coefficient increases from 0.38 to 0.58). Moreover, these models showed resistance to overfitting. In a multivariate analysis of the hypothesis regarding the sign of the dependence of the probability of insolvency on the alleged regressors were confirmed. Significant differences in accuracy indicators between the logit and probit models were not found.

This conclusion is consistent with many works, since the logistic distribution function and the distribution function of the standard normal random variable behave approximately the same, and the differences are associated with more "heavy tails" of the logistic distribution function.

Due to the stability of nonparametric algorithms to multicollinearity, all the factors considered earlier were used to build models based on machine learning methods. Analysis of classification trees and random forests showed that among the most influential factors were the coefficient of maneuverability of equity and the coefficient of autonomy (the largest drop in the Gini index in the random forest algorithm, the first partition in classification trees). This means that if a company has a significant amount of debt burden and it shows a negative financial result (in the balance sheet its equity is negative), this is an important indicator of the company's insolvency in the next reporting period. At the same time, non-financial factors (age, company size) turned out to be practically insignificant, which is reflected in *Figure 5*. Thus, the large size and long life of the company in the market cannot guarantee stability in the Russian market.

The dynamics of the average value of the Gini coefficient depending on the share of insolvent companies with optimal parameters on the training set using undersampling and oversampling (*Figure 6*) shows that in the forecasting problem with this data structure, the influence of the share of insolvent companies on the training set does not significantly affect the forecast potential one or another method. This conclusion is consistent with the work of Demeshev and Tikhonova [14]. The negative dynamics of the quality metric with an increase in the share of insolvent companies indicates a bias in the construction of the algorithm (for example, in the "random forest using oversampling" algorithm).

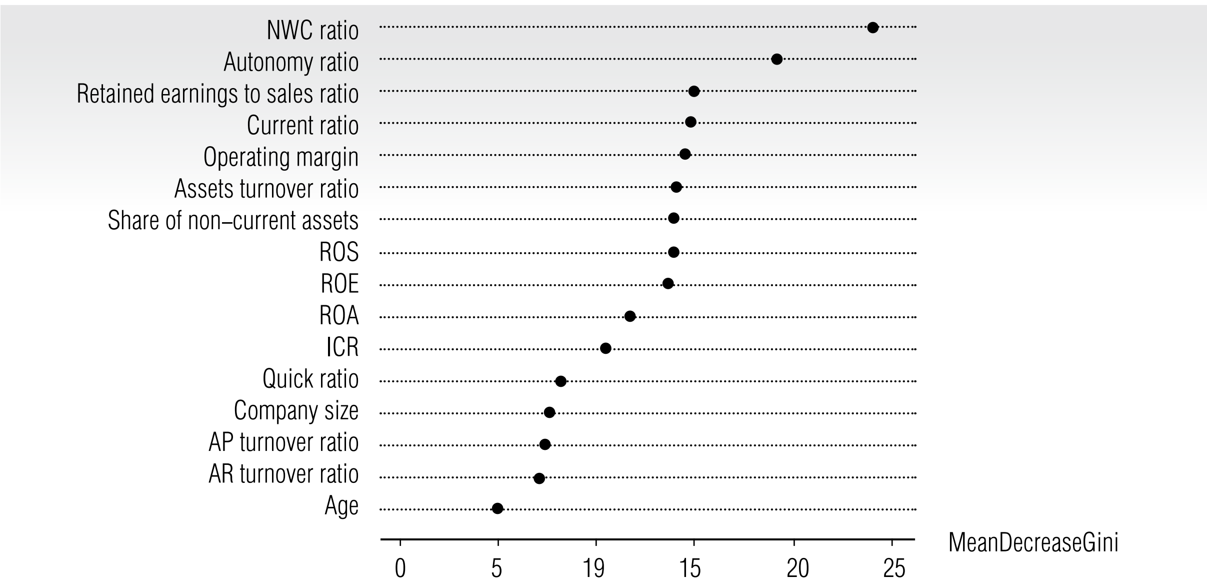


Fig. 5. Determination of the most significant parameters. Random forest algorithm

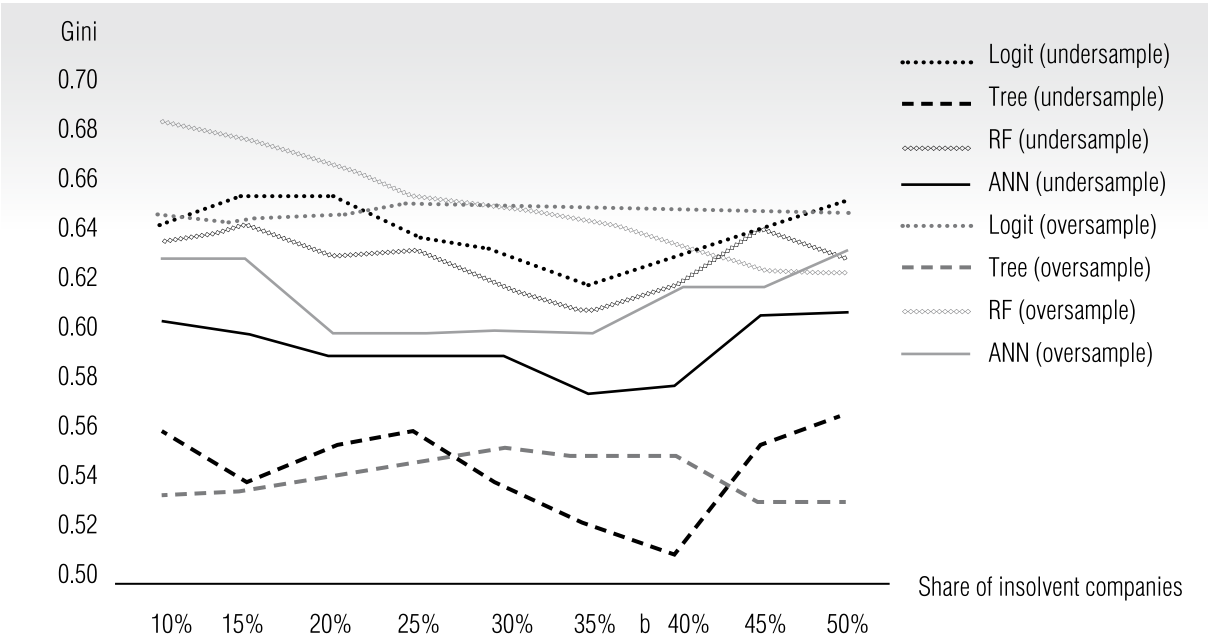


Fig. 6. The average value of the Gini coefficient on the training set

The use of the method of combating imbalance depends on the type of model used. For logistic regression, artificial neural networks and classification trees, oversampling has shown higher quality. However, the use of oversampling in the random forest method leads to overfitting. Therefore, for random forests, undersampling is more effective.

Conclusion

The use of a particular model depends on the goal of the analyst. In forecasting problems, nonlinear algorithms, as a rule, show a higher result. Therefore, the use of neural networks and random forests is more acceptable for this type of task. However, these models lose to the binary choice models in costs (time, computational) for calculations, as well as in interpretation.

The algorithms we examined showed acceptable quality for use in the tasks of forecasting bankruptcies of construction companies. As expected, the best model was an artificial neural network. Traditional sampling models have shown good results, while their results can be easily interpreted, and the calculation time is minimal. Despite the advantages of classification trees (ease of interpretation, the absence of restrictions on the type of variables, the absence of the need to specify the relationship form in an explicit form), this algorithm showed instability and low accuracy of predictions.

In the future, it seems promising to include other nonlinear algorithms in comparison, for example, models based on boosting (GBM, XGBoost), support vector models, etc. Moreover, in this work, the category of bankrupts includes companies in respect of which the legal bankruptcy procedure has begun, as well as companies that have liquidated voluntarily. In the future, it seems possible to distinguish between these categories using a single federal register of bankruptcy information and identify companies in respect of which the legal bankruptcy procedure has begun. It also seems possible to conduct an intersectoral comparison of the methods considered, determine the maximum forecasting horizon at which signs of bankruptcy appear, diversify within individual industries and use macroeconomic variables in modeling. ■

References

1. Karminsky A.M., Kostrov A.V., Murzenkov T.N. (2012) *Approaches to evaluating the default probabilities of Russian banks with econometric methods.* Working paper WP7/2012/04 (Series "Mathematical methods of decision analysis in economics, business and policies"). Moscow: HSE (in Russian).
2. Kostrov A.V. (2016) Comparison of statistical classification methods to predict Russian banks failures. *Management of Financial Risks*, vol. 47, no 3, pp. 162–180 (in Russian).
3. Tserng P., Chen P.-C., Huang W.-H., Lei M.C., Tran Q.H. (2014) Prediction of default probability for construction firms using the logit model. *Journal of Civil Engineering and Management*, vol. 20, no 2, pp. 247–255. DOI: 10.3846/13923730.2013.801886.
4. Behr A., Weinblat J. (2017) Default patterns in seven EU countries: A random forest approach. *International Journal of the Economics of Business*, vol. 24. No 2. pp. 181–222. DOI: 10.1080/13571516.2016.1252532.
5. Gepp A., Kumar K. (2015) Predicting financial distress: A comparison of survival analysis and decision tree techniques. *Procedia Computer Science*, vol. 54, pp. 396–404.
6. Tam K.Y., Kiang M.Y. (1992) Managerial applications of neural networks: the case of bank failure predictions. *Management Science*, vol. 38, no 7, pp. 926–947. DOI: 10.1287/mnsc.38.7.926.
7. Bogdanova T.K., Shevgunov T.Ya., Uvarova O.M. (2013) Using neural networks for predicting solvency of Russian companies on manufacturing industries. *Business Informatics,* no 2, pp. 40–48 (in Russian).
8. Breiman L., Friedman J., Olshen R., Stone C. (1984) *Classification and regression trees.* Wadsworth, New York: Chapman and Hall.
9. Breiman L. (2001) Random forests. *Machine Learning,* vol. 45, no 1, pp. 5–32.
10. Odom M.D., Sharda R. (1990) A neural network model for bankruptcy prediction. Proceedings of *International Joint Conference on Neural Networks. San Diego, USA, 17–21 June 1990*, vol. 2, pp. 163–168. DOI: 10.1109/IJCNN.1990.137710.
11. Kumar P.R., Ravi V. (2007) Bankruptcy prediction in banks and firms via statistical and intelligent techniques: A review. *European Journal of Operational Research*, vol. 180, no 1, pp. 1–28. DOI: 10.1016/j.ejor.2006.08.043.
12. Trujillo-Ponce A., Samaniego-Medina R., Cardone-Riportella C. (2014) Examining what best explains corporate credit risk: accounting-based versus market-based models. *Journal of Business Economics and Management*, vol. 15, no 2, pp. 253–276. DOI: 10.3846/16111699.2012.720598.
13. The Federal Law of 26 October 2002, No 127-FZ (as amended on 29 December 2017) *"On insolvency (bankruptcy)"* (in Russian).
14. Demeshev B.B., Tikhonova A.S. (2014) *Default prediction for Russian companies: intersectoral comparison.* Working paper WP2/2013/05 (Series WP2 "Quantitative Analysis of Russian Economy"). Moscow: HSE (in Russian).
15. Siddiqi N. (2012) *Credit risk scorecards: developing and implementing intelligent credit scoring.* John Wiley & Sons.

About the authors

Alexander M. Karminsky

Dr. Sci. (Econ.), Dr. Sci. (Tech.);

Professor, Department of Finance, National Research University Higher School of Economics, 20, Myasnitskaya Street, Moscow 101000, Russia;

E-mail: karminsky@mail.ru

ORCID: 0000-0001-8943-4611

Roman N. Burekhin

Doctoral Student, Doctoral School on Economics, National Research University Higher School of Economics, 20, Myasnitskaya Street, Moscow 101000, Russia;

E-mail: romanvia93@yandex.ru

ORCID: 0000-0003-1130-0175

DOI: 10.17323/1998-0663.2019.3.67.77

How can an IT organization earn its customers' trust: A practical approach

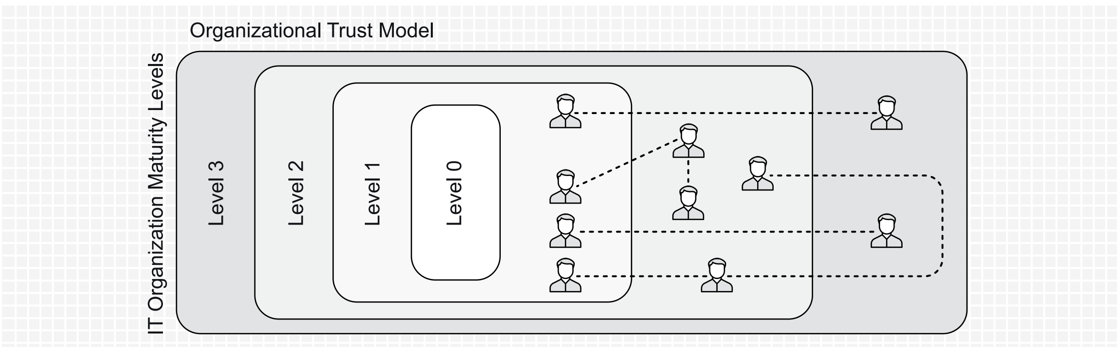
Alexander N. Biryukov
E-mail: anbiryukov@fa.ru

Financial University under the Government of the Russian Federation
Address: 49, Leningradskiy Prospect, Moscow 125993, Russia

Abstract

The paper describes a practical approach which can be used by internal IT organizations to gain their business customers' trust. The variety of customers of the internal IT service provider is limited to internal customers only. The distinguishing feature of the proposed approach is that it is completely practice-oriented, i.e. primarily aimed at building trust among IT service providers and their customers in a particular organization. The approach is based on the idea that there are measurable prerequisites for the emergence of a customer's trust which allow you to partially formalize the IT organization's intention to earn its customers' trust. A model of intra-organizational trust is proposed; it is progressively improved as the IT organization develops its trust-building capabilities. The model comprises all IT service customers in an organization along with their communications and accounts for internal organizational IT service market specifics. A high-level blueprint of the trust model is described which can serve as a starting point when developing a full-scale trust model in a particular IT organization. We present an approach to the trust model improvement which builds upon principles adopted from widely recognized CMMI model. With this approach, an internal IT service provider can benefit from maturity assessment methods to improve its trust building capabilities.

Graphical abstract



Key words: IT service provider; internal IT service market; trust; IT organization; trust criteria; intra-organizational trust model; trust management; earning trust; maturity; capability level.

Citation: Biryukov A.N. (2019) How can an IT organization earn its customers' trust: A practical approach. *Business Informatics*, vol. 13, no 3, pp. 67–77. DOI: 10.17323/1998-0663.2019.3.67.77

Introduction

Over the past few decades, the perception and understanding of trust and its role in organizations has significantly progressed. Nowadays, trust is considered to be a highly complicated multy-dimensional phenomenon occuring in many forms, such as interpersonal trust [1, 2], inter-organizational trust [3], intra-organizational trust [2, 4, 5]. Trust is studied from different viewpoints, among which are psycholigical, economic, philosophical and managerial viewpoints [2, 4–6]. Many researchers study trust in different areas including both technological, such as e-business, and non-technological, e.g. political and cultural areas. In spite of lots of exciting ideas and deep insights concerning trust proposed by different researchers and much effort expended, no holistic interdisciplinary approach to trust study has been developed thus far. There is no definition of trust, which most (or at least significant number of) researchers would accept as the common basis in their studies. Instead, almost every researcher proposes his/her own definition of trust, thus making the whole volume of knowledge about trust even more fragmentary and less cohesive [7].

We believe that this is not only due to the complex nature of trust itself, but also to the fact that most researchers adopt narrative and qualitative approaches to the study of trust. Traditionally, research builds upon a statistical analysis of answers to multi-item surveys which are designed to reveal the opinions of respondents on different factors affecting their trust. A review of modern trust study methods, which demonstrates ontological, teleological and epistemological aspects of the research process, which incudes questioning and further qualitative and quantitative analysis, can be found in [5]. Some of the up-to-date methods of trust research, mainly based on different interviewing technics are also presented in [8].

In spite of the fact that the definition derived from such an approach is inevitably biased, it seems to be a widely accepted idea that one cannot manage trust in practice without first formulating a universal trust definition. Lack of progress in constructing the definition therefore results in the lack of studies concerning practical aspects of trust management. This paper adopts an alternative approach. We believe that one can successfully undertake trust building activities without thorough trust definition.

Further on we concentrate on intra-organizational trust only. This kind of trust is relatively well studied [9–14]. A lot of different trust measures have been defined in the literature [7]; some are new concepts, such as the organizational citizenship behavior proposed, and behavioral patterns affecting intra-organizational trust have been studied. A universal definition of intra-organizational trust, however, does not exist, since the coherence of different definitions is still too low [7].

There are only a few studies of trust in the IT area (e.g. [15]); these mainly focus on inter-organizational trust in the context of IT outsourcing. As we show below, the results obtained are irrelevant for internal IT organizations and their customers due to principal differences between internal and open markets, where IT organizations and IT companies operate.

Finally, we find it important to note that after being mainly a topic for academic research trust appears to be becoming a point of interest for commercial companies; they have started offering services related to building, maintaining and measuring trust in organizations [16, 17]. This is yet further evidence of the growing importance of inter-organizational trust building activities.

1. The aim of the study and statement of the problem

The aim of this paper is twofold. First, we prop se a practical approach to trust management in the specific context of collaboration between an internal IT service provider and its customers. Second, we describe a quantitative non-empirical approach to trust definition in this special case. Our method is based upon the assumption that each organization will develop and use the local trust definition of its own, which most fits its requirements.

We call this definition the trust model. While the universal trust definition may only stem from the experience accumulated by a number of organizations, the trust model is just a local tool, a data structure assisting the members of an IT organization to build trust with business customers and quantitatively estimate the results of this activity.

It is important to note that inter-organizational trust cannot result from the execution of formal business processes, since both the trust itself and trust building activities heavily depend on the individualities of organizational members. Our approach is based upon the following principal assumptions:

1. Trust building is studied in the context of collaboration of organizational roles adopting general organizational policies, business rules and corporate culture, which heavily affects their behavior;

2. The collaborating parties need trust if they cannot predict each other's course of action;

3. Each role has its own unique view on how to achieve trust with its counterparties different from that of other roles.

There are two reasons that we believe that it is important to focus on the trust building between an IT organization and its customers. First, provided there is no reliable business case to decide if IT investments are necessary, a consensus between managers is needed to make an investment decision. If the average level of trust between business managers and an IT organization is high enough, the consensus is more likely to be achieved.

Second, due to the intrinsic complexity of an IT organization's work, most organizational members cannot be considered as IT experts who are able to correctly estimate the quality of IT-related products and services. Moreover, we believe this is true even if an IT organization provides its customers with detailed information on its processes, policies, guiding principles, etc. Trust can help both sides to make a Service Level Agreement, which forms the basis of their collaboration.

One can argue this logic because business units different from the IT organization can be considered to be service providers for their customers as well, even though this approach to organizational analysis is not commonly accepted. There are important differences, however, that make the IT organization unique in its position within the organization. For example, accounting services or HR services are clear even for non-professionals, while the corresponding processes are mostly aligned with external legislative requirements. That's why the organizational members have no reasons to distrust the corresponding business functions. On the other hand, there are business units, which provide complicated professional services, for example, the Technical Directorate at a factory. However, its customers, such as manufacturing managers, are normally technical experts themselves. Their collaboration is not based on trust, because both parties can completely validate each other's ideas.

The specificity of IT services is that they, on the one hand, are provided to practically every member of an organization and, on the other hand, may require deep professional skills to be correctly estimated. This makes it hard for non-IT professionals to correctly estimate how diligent an IT organization is.

Th strategic importance of trust for an IT organization has already been considered in [8]. The present paper focuses on operational activities, which facilitate permanent maintenance of trust and its support at an acceptable level from the IT organization's side. We further assume that the IT organization adheres to ITSM principles, especially as for Service Level Agreements, which form the basis of the collaboration between an IT organization and its customers.

The following terms will be used below:

✦ **IT organization** – an internal IT function of an organization, no matter what form it takes (e.g., a single IT department, group of IT departments or even an independent legal entity);

✦ **internal IT service provider** – same as the IT organization;

✦ **trust building** – an activity in the IT organization aimed at earning trust from its customers.

2. The IT organization and the internal IT service market

Following [19], we distinguish between different IT service customers and IT service users. At least four groups of customers and users can be identified in an organization:

✧ **participants of business processes who use IT in their everyday work.** They are mostly IT service users. The IT service customer in this case is normally one of the business managers responsible for collaboration with IT organization on behalf of a business unit or organization as a whole;

✧ **owners and managers of business processes who use IT to improve their respective business processes.** A typical IT service for those customers is business process automation;

✧ **senior management and top managers.** IT services for these customers include, for example, implementation of corporate-wide IT solutions such as ERP-systems or Master Data Management systems, IT budget optimization, increasing IT organization efficiency or ensuring proper level of information security;

✧ **shareholders and investors.** IT services for this category of customers may be intended to increase the return on IT investments. This implies, for example, the use of best practices and proven IT solutions whenever possible, contracting high-class suppliers and so on.

Some business process owners may hold positions as functional managers in the organizational hierarchy. For instance, the owner of the purchasing business process is normally the Head of the Acquisitions business unit, while the owner of the budgeting process is the Head of the Budget or Finance unit etc.

The internal organizational IT service market has some unique features that make it principally different from the open market. The differences are as follows:

✦ the IT organization is, by its nature, a monopoly on the market. It provides IT services either on its own or by contracting with external companies on behalf of the organization, thus accumulating all business requirements corporate wide. As a result, the IT organization cannot abandon a service request from any business customer;

✦ customers on the open market are able to contract with those IT service providers who render the highest quality services in a specific area. Instead, the customers on the internal market are forced to rely on the single IT service provider which may not be skilled in all areas, to say nothing of the emerging IT-services;

✦ customers on the open market use legal commercial agreements to manage relationships with service providers. The Internal Service Level Agreement cannot be used as a legal tool to manage relationships with the IT organization.

The above specifics of the internal IT service market result in specific relationships between the IT organization and its customers. Both sides may significantly benefit from being trusted partners. For instance, they can:

✦ minimize mutual control costs;

✦ reduce external risks by solving problems locally without addressing higher positioned managers;

✦ adopt long-term relationship models based on agile technologies and task managers instead of traditional project activities.

Since the customers and users of the internal IT service provider are members of the organization and there is a CIO who represents the IT organization for them, we can treat the trust between the IT organization and its customers as a kind of interpersonal trust.

3. Trust measurement and trust building

The fundamental review [7] summarizes the most important principles of intra-oraganizational trust measurement. According to [7], most researchers consider intra-organizational trust to be a multi-dimensional notion and trust measurement is treated as the process of a multi-item survey that includes a number of questions or metrics intended to capture different dimensions of trust. The metric specifies one aspect of trust or trust dimension. Metrics are grouped according to the different dimensions. For example, the metric may look like "My colleagues who collaborate with my counterparty consider him/her as trustworth." Since the respondent's perception of a metric is always biased, a kind of averaging is necessary to correctly interpret a respondents' view on the proposed trust dimension. Basically, the trust measurement in accordance with [7] is an approach aimed at constructing a comprehensive trust definition.

Authors of [7] have not discussed a more traditional view of trust measurement such as a quantitative estimation of intra-organizational trust although there are examples of such an approach for the case of inter-organizational trust [16]. Moreover, the authors of [7] revealed a large variety of dimensions and metrics proposed by different authors along with a very limited degree of replication of them by the authors other than their originators.

There are five trust categories that have been recommended in [7] as the most recognized and accurately defined. They therefore can de used as good starting points for future work. These categories correspond to five different statements of the trust research problem. Two categories describe inter-personal trust [20, 21], one the trust among business units and other work groups, the other [22] the trust between a person and his/her subordinates. The last category considers trust as a psychological state, which implies one's intention to accept vulnerability based upon positive expectations of another's intentions. We are interested in the first two categories, which can both be used for our purposes although different in nature.

The first category [20] defines trust as having two dimensions – cognition-based trust and affect based trust. The respective groups of metrics describe the trustor's perception of the ability of the trustee to achieve the expected result and the trustor's confidence in trustee's emotional support and expressions of care and concern for the trustor's welfare.

The second category [21] comprises four dimensions, namely free information exchange, striving for informal agreements, surveillance, and task coordination.

For our purposes, we combine the two categories into a single one that has five of the Now we are ready to proceed to our main goal, i.e. designing a practical method of trust building between the IT organization and its customers. Since the IT organization cannot emotionally affect its customers' behavior, we only can outline its business activities aimed at gaining customers' trust. That's why we further focus on information exchange, striving for informal agreements, surveillance, and task coordination from the IT organization's side only. Moreover, the IT organization should manage its customers' cognition-based trust. We also postulate a uniform approach from the IT-organization's side to all its customers.

Table 1 shows the quantitative metrics we have defined for the case of trust building between the IT organization and its customers. They are based on the recommendations of [7] and should not be considered as a complete set of metrics, rather as just an example illustrating the potential ability to develop such metrics.

We assume that the customer's trust is always the combination of two factors: 1) his/her personal trust based on the current and previous experience and/or personal trustfulness, and 2) influence of the customer's environment. Metrics that describe the influence of the environment are italicized in *Table 1.* We do not analyze an individual customer's behavior, only assume that the metrics in *Table 1* help the customer to improve his/her perception of the IT organization, thus resulting in increased trust. The environmental influence is important because a non-IT specialist often cannot identify the true reason of a low value of a metric (for instance, IT service quality or the value of IT risk). Our assumption is that in this case he/she will consult with his/her colleagues to account for their opinions on the IT organization and its work.

The way the customer processes the information received depends on many factors, such as colleagues' trustworthiness, the level of the customer's IT skills, the availability of resources and the desire to control the IT service provider etc. In any case, the IT organization has to account for the influence of the environment and strive to manage it. In other words, the IT organization should be fairly aware of the organizational communication structure.

The foregoing suggests that intra-organizational trust is highly individual for every particular organization by its very nature and depends on the unique structure of inter-personal intimate trust among different members of the organization, including not only customers of the IT organization but also their trusted parties. Moreover, the structure of inter-personal trust is highly volatile over time, which is equally true for the behavior of individual customers. This is the principal challenge to the aforementioned psychometric approach to trust measurement, which uses static surveys to define the trust.

Figure 1 shows an example of an inter-personal trust structure.

The four rectangles in *Figure 1* represent four members of an environment (A–D); the arrows indicate information flows. There are two types of information in the flow from A to C, e.g. opinions concerning IT service quality and striving for informal agreements. The levels of trustworthiness of the sources and their respected opinions are shown within the rectangles. This structure does not describe any particular picture of information exchange at a specific moment, but all potential ways the information can flow across the organization, i.e. the historical blueprint of information exchange among customers.

At any particular moment, the degree some customer trusts the IT organization depends on the following reasons:

✦ personal perception of current relationships with the IT organization;

✦ an individual way of assessment of the information received from the environment;

✦ the current state of the inter-personal structure in the organization.

Assume, for example, that customer C requests opinions of trusted customers A and B concerning the quality of IT services provided by the IT organization. Generally speaking, the information he/she receives depends

Table 1.

Quantitative metrics for building trust between the IT organization and its customers

Trust dimension	Metrics	Measurement method
Cognition–based trust	Quality of IT services	Number of SLA violations
	Satisfaction with previous collaboration	Percentage of prolonged SLAs
	Percentage of customers satisfied by IT services	*An estimate of the share of satisfied customers*
Free information exchange	Readiness to share information	Number of violations of agreements concerning information access
Striving for informal agreements	Customer's comprehension of the level of accuracy of system requirements in SLA	Percentage of identified risks approved by the customer
	Customer's comprehension of the requirements concerning schedule of work and resources needed	Percentage of other customers who approve the IT organization's plans
	Percentage of customers striving to establish informal relationships with the IT organization	*An estimate of the share of such customers*
Surveillance	Customer's cost of monitoring the schedule and work status	Analyze customer's information
	Customer's cost of quality control	A profile of divergencies of costs in the quality plan of the IT organization
	Percentage of customers striving to cut costs of monitoring and control	*An estimate of the share of such customers*
Task coordination	Percentage of changes of work schedule and/or resources needed initiated by the customer and waived by the IT organization	Percentage of changes initiated by the customer and abandoned by other customers presented in IT organization's work plan
	Number of changes of work schedule and/or resources needed for a customer initiated by the IT organization and not related to other customers	Analysis of IT organization's internal documentation
	Percentage of customers satisfied with task coordination with IT organization	*An estimate of the share of such customers*

on whether B asked A for a similar opinion. If yes, C will receive from B a combination of opinions of B and C, otherwise B's opinion will be highly individual. The structure in *Figure 1* does not help to sort out the situation. Customer C then changes his/her perception of the IT organization depending on the information received. For instance, he/she may increase his/her level of trust if the weighted sum of the opinions received is close to his/her own perception. Or he/she may ignore the external opinions if the weghted sum is small. The logic observed by customer C at any particular moment is unknown.

In spite of the fact that the structure in *Figure 1* cannot predict the result of communication between the customer and his/her environment, it may help the IT organization to identify those customers who potentially most affect the others, how large are the risks of low quality IT services provided to a particular customer, what groups of customers trust each other most, etc.

This information may prove useful for the internal IT service provider even if the structure is so complicated as not to allow it to analyze each customer's behavior. On the other hand, the structure itself may result from everyday monitoring of stable customers' behavior patterns.

We call this structure the intra-organizational trust model. It is not by any means universal being is relevant for building trust between the IT organization and its customers.

4. Trust building management and trust model improvement

Of course, the above arguments should be somehow validated, i.e., compared with the practical experience accumulated in IT organizations which attempt to become trusted counterparties for their respective customers. It should be noted that the trust building activity cannot be considered to be a common business process for at least two reasons. First, it heavily depends on the particular individuals involved and their emotional backgrounds. Second, any formal output of the activity may be of no help to trust building since there is no guarantee that this output really affects the customer's trust.

At the same time, it is less likely that any best practices of trust building can exist. This is indirectly supported by the fact that no universal trust definition has yet been found. All of this does not imply, however, that the IT organization is not able to effectively gain the confidence of its customers. To do this, it can follow the principles the well-known process improvement CMMI model [23–25] is based upon. The key idea is to replace the set of process areas in CMMI for Services [24] with the aforementioned intra-organizational trust model. This model becomes more and more validated and standardized in the same way as process areas in CMMI do, while the IT organization proceeds through

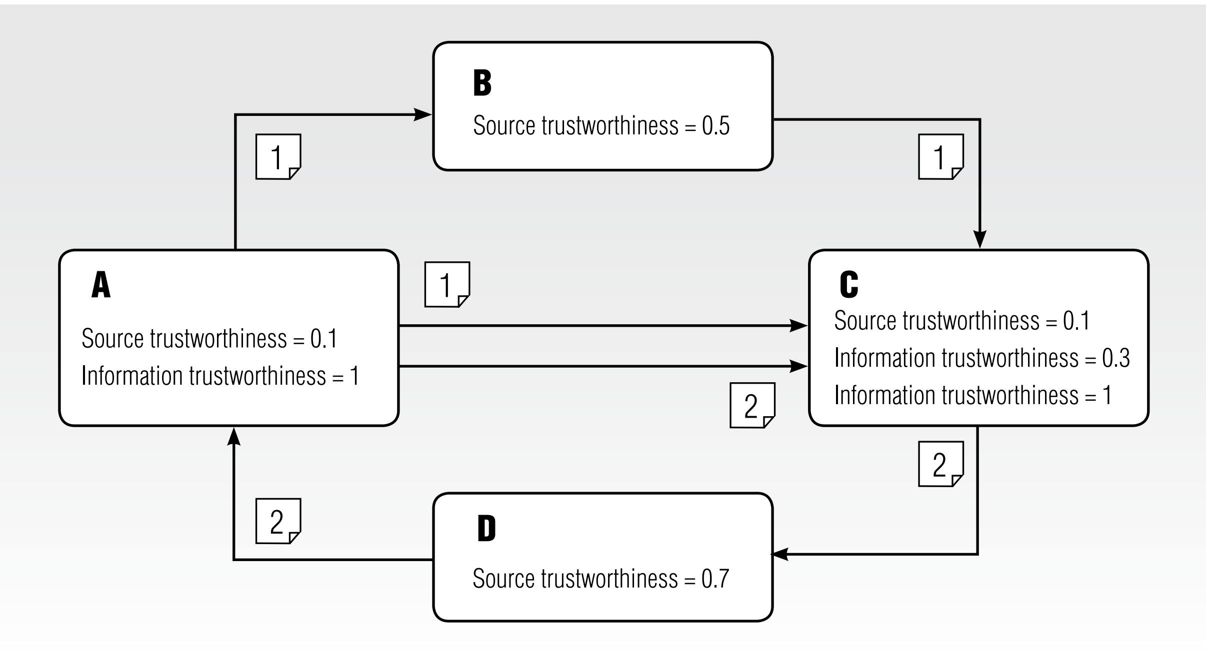


Fig. 1. An example of the inter–personal trust structure

trust building maturity levels. The trust model at any moment describes the current perception of customers' trust similarly to process areas which describe current institutionalization of business processes. In other words, the maturity level of the IT organization is defined it terms of the accuracy and completeness of the trust model itself and the accumulated experience of its use by IT specialists and IT managers. Note that the IT organization is interested in improving the trust model no matter what happens to the relations with any particular customer and the IT services provided. This is not the case with CMMI process areas that should not be improved unless there are customers interested in such an improvement.

Below is the high-level definition of the maturity levels of the IT organization. It is important to note that in contrast to the universal set of process areas in CMMI for Services, the trust model should be elaborated individually for each organization. The definition below is intended just to outline for IT practitioners an approach based on the trust building maturity assessment and aimed at trust building activity improvement. We identify the following maturity levels:

✧ **incomplete:** the very need to intentionally earn customers' trust has not been realized by the IT organization. Trust building activities are spontaneous and incomplete;

✧ **initial:** There is general understanding of the necessity of trust building. Different IT practitioners gain their customers' trust in different ways. Every customer's increase or decrease in trust is clearly recognized. There is no consensus on how trust should be built, i.e., no common trust model exists;

✧ **managed:** the IT organization has elaborated some proven practices of earning the customer's trust. Inter-personal trustful relationships are established between several customers and particular IT specialists. There are internal policies and business rules concerning trust building which are followed by all customer managers in the IT organization. Some elements of the trust model are used to earn some customers' trust;

✧ **defined:** the IT organization has adopted a common view on how trust has to be built. This view has the form of a single standard trust model. Business rules exist which allow the customer manager to adjust the model to his/her specific relationships with the respective customers.

As was stated above, the trust model assumes that each customer's trust at any moment is known. Obviously, this cannot be achieved through the use of questionnaires and other psychometric technics. There is some indirect evidence, however, which reports on increase/decrease in trust. The corresponding events can be watched and analyzed. Below are some examples of such events:

✦ SLA change which introduces higher/lower level of control from the customer's side;

✦ increase/decrease of the number of vertical escalations of problems between the customer and the IT organization;

✦ increase/decrease of costs of mutual monitoring and control from a customer's and IT organization's sides;

✦ increase/decrease of the number ✦of changes from the customer's side without SLA re-assessment;

✦ customer's refusal to deal with particular IT specialists.

Of course, the most reliable informal way to become aware of a particular customer's trust is to establish inter-personal trustful contact between the customer and the corresponding customer manager.

At the managed maturity level, the IT organization accounts for trustful relationships between some customers and correspondingly builds relationships with those customers. Trusted IT specialists play a major role in estimating customer's trust.

At the defined level, the IT organization exercises a common approach to trust building aimed at all current and potential customers, not at a customer of particular interest. The standard trust model is used which accounts for all interactions between customers.

It does not make sense to attempt to describe higher maturity levels at the current stage of understanding the role of the trust model. Any standard model improvement may only be based on real experience in its use, which should be accumulated first. Provided formal trust building business processes do not exist, the trust model looks like a practical alternative allowing one to capture, formalize and analyze the results of trust building activities, even though those activities are informal and vague by their very nature. The trust model may be implemented in a variety of ways from a simple set of files to a sophisticated data structure maintained with the use of formal business processes.

Conclusion

The approach presented in this paper is just a first step to practical trust management in IT organizations. The main result of the approach is the demonstration of an ability to practically build trust in the absence of a comprehensive definition of intra-organizational trust.

Further steps can only be made by IT practitioners responsible for customers' trust management in IT organizations. It is especially important therefore to get feedback from customer managers, presale managers and other IT specialists involved in customer relationship management. Much preparatory work should be done to elaborate practical tools that would allow them to validate the proposed approach.

The last, but not least challenge is related to the trust management business case. Even though trust benefits seem clear, ensuring a balance between them and the cost of ongoing trust building activities is not a trivial problem. ■

References

1. Paliszkiewicz J. (2011) Trust management: Literature review. *Management*, vol. 6, no 4, pp. 315–331.
2. Feltman C. (2008) *The thin book of trust: An essential primer for building trust at work*. Bend: Thin Book.
3. Paliszkiewicz J. (2011) Inter-organizational trust: Conceptualization and measurement. *International Journal of Performance Meaurement*, vol. 1, pp. 15–28.
4. Kramer R., Tyler T. (1996) *Trust in organizations: Frontiers of theory and research*. Sage Publications.
5. Muhl J.K. (2014) *Organizational trust: Measurement, impact and the role of management accountants*. Springer.
6. Shaw R.B. (1997) *Trust in the balance: Building successful organizations on results, integrity, and concern*. San Francisco: Jossey-Bass.
7. McEvily B., Tortoriello M. (2011) Measuring trust in organisational research: Review and recommendations. *Journal of Trust Research*, vol. 1, no 1, pp. 23–63.
8. Lyon F., Mollering G., Saunders M. (eds.) (2015) *Handbook of research methods on trust*. Edward Elgar.
9. Zaheer A., McEvily B., Perrone V. (1998) Does trust matter? Exploring the effects of interorganizational and interpersonal trust on performance. *Organization Science*, vol. 9, no 2, pp. 141–159.
10. Vangen S., Huxham C. (2003) Nurturing collaborative relations. *Journal of Applied Behavioral Science*, vol. 39, no 1, pp. 5–31.
11. Fulmer A., Dirks K. (2018) Multilevel trust: A theoretical and practical imperative. *Journal of Trust Research*, vol. 8, no 2, pp. 137–141.
12. Dar O.L. (2010) Trust in co-workers and employee behaviours at work. *International Review of Business Research Papers*, vol. 6, no 1, pp. 194–204.

13. Paille P., Bourdeau L., Galois I. (2010) Support, trust, satisfaction, intent to leave and citizenship at organizational level. A social exchange approach. *International Journal of Organizational Analysis*, vol. 18, no 1, pp. 41–58.

14. Fulmer A., Gelfand M.J. (2012) At what level (and in whom) we trust: Trust across multiple organizational levels. *Journal of Management*, vol. 38, no 4, pp. 1167–1230.

15. Ali Babar M., Verner J.M., Nguyen P.T. (2007) Establishing and maintaining trust in software outsourcing relationships: An empirical investigation. *Journal of Systems and Software*, no 80, pp. 1438–1449.

16. Paine K.D. (2013) *Guidelines for measuring trust in organizations*. Gainesville, FL: Institute for Public Relations, University of Florida.

17. The Ken Blanchard Companies (2017) *Building trust. The critical link to a high-involvement, high-energy workplace begins with a common language*. Available at: https://www.connexpartners.com/hubfs/assets/Building%20Trust.pdf (accessed 15 February 2019).

18. Biryukov A.N. (2017) Strategic management in the IT department. *Business Informatics,* no 2, pp. 17–24.

19. The Stationery Office (2011) *ITIL Service Strategy, 2011 edition*. London, UK: The Stationery Office.

20. McAlister D.J. (1995) Affect and cognition-based trust as foundation for interpersonal cooperation in organizations. *Academy of Management Journal*, vol. 38, no 1, pp. 24–59.

21. Curral S.C., Judge T.A. (1995) Measuring trust between organizational boundary role persons. *Organizational Behavior and Human Decision Processes*, vol. 64, no 2, pp. 151–170.

22. Cummings L.L., Bromiley P. (1996) The organizational trust inventory (OTI): Development and validation. *Trust in organizations: Frontiers of theory and research*. Thousand Oaks, CA: Sage Publications, pp. 302–330.

23. Carnegie Mellon University (2010) *CMMI for Development, Version 1.3. CMMI Product Team. Improving processes for developing better products and services. Technical Report. CMU/SEI-2010-TR-033. ESC-TR-2010-033.* Pittsburgh, PA: Carnegie Mellon University.

24. Carnegie Mellon University (2010) *CMMI for Services, Version 1.3. CMMI Product Team. Improving processes for developing better products and services. Technical Report. CMU/SEI-2010-TR-034. ESC-TR-2010-034.* Pittsburgh, PA: Carnegie Mellon University.

25. Carnegie Mellon University (2010) *CMMI for Acquisitions, Version 1.3. CMMI Product Team. Improving processes for developing better products and services. Technical Report. CMU/SEI-2010-TR-032. ESC-TR-2010-032.* Pittsburgh, PA: Carnegie Mellon University.

About the author

Alexander N. Biryukov

Cand. Sci. (Phys.-Math.);

Associate Professor, Department of Business Informatics,
Financial University under the Government of the Russian Federation,
49, Leningradskiy Prospect, Moscow 125993, Russia;

E-mail: anbiryukov@fa.ru

DOI: 10.17323/1998-0663.2019.3.78.96

Standardization in digital manufacturing: implications for Russia and the EAEU

Yulia V. Turovets
E-mail: yturovecz@hse.ru

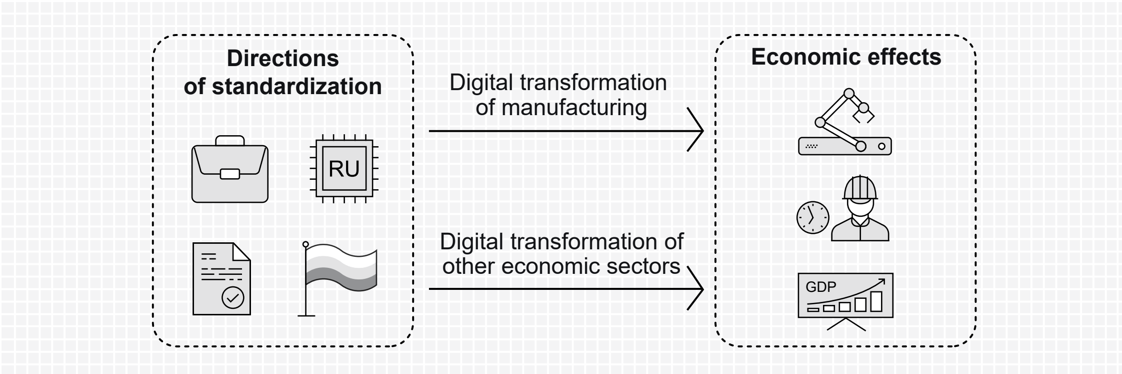
Konstantin O. Vishnevskiy
E-mail: kvishnevsky@hse.ru

National Research University Higher School of Economics
Address: 11, Myasnitskaya Street, Moscow 101000, Russia

Abstract

The shift to digital technologies in various industries is one of the key goals in the digital agenda. Due to the essential role of interoperability of products and elements in complex systems, standardization stays in the forefront of government policy and business. In manufacturing systems, standards are of a prime importance, since they serve as a channel for modernization and innovation speedup. This paper makes a contribution to the currently rare literature on digital manufacturing standardization as a policy tool to promote digital technologies in business. By comparing five national cases of China, Germany, Japan, the Republic of Korea and the USA, we introduce national models of standardization in smart manufacturing according to the extent of state participation in standardization. In doing so, we examined initiatives in industry, digitalization, the development of a national system of standards, the reference architecture of digital production, as well as the countries' cooperation in the field. Along with this, an overview of international initiatives in the field is presented, namely the ISO and the IEC. Taking into account the existing landscape, an assessment of the Russian case of digitalization in manufacturing and standardization is presented. Like China, Russia follows the third model of standardization. Given the results, we developed recommendations for Russia with the aim of intensifying efforts at standardization and the country's presence in the international agenda, as well as to develop a Russian framework for digital transformation in sectors and achieve related economic effects.

Graphical abstract



Key words: digital economy; digital manufacturing; standardization; technical committee; international standardization bodies.

Citation: Turovets Yu.V., Vishnevskiy K.O (2019) Standardization in digital manufacturing: implications for Russia and the EAEU. *Business Informatics*, vol. 13, no 3, pp. 78–96.
DOI: 10.17323/1998-0663.2019.3.78.96

Introduction

Currently digitalization is becoming a key issue in the industrial and innovation policy perspective and generates significant changes. A comprehensive intellectualization through information and communication technologies (ICT) leads to drastic rearrangement in production and business-processes facilitating cooperation in a global market [1, 2]. The absence of standards is often seen as a barrier to digitalization.

Digitalization and its implications for upgrading the technological sector has several dimensions. On the one hand, some digital technologies could bring new concepts of business process management without significant change in the set of technologies. Nevertheless, automated systems and management systems should be integrated with existing equipment [3]. On the other hand, a complex digital transformation entirely restructures a company's technology architecture [4, 5].

Digitalization as a scientific field is in its infancy. There are several notions of digital manufacturing in existing literature, which include smart manufacturing, cyber-physical systems, Industry 4.0, the smart/digital factory and some others. It is a complex system that integrates the pull of production and information technologies that help to optimize on the production floor and drive product development in a virtual environment. In a more broad sense, it is a concept of how to exploit an extensive set of data in a most effective way. Cyber-physical systems (CPS), cloud computing, and the Internet of Things (IoT), big data, digital modeling, additive manufacturing, virtual reality constitute its technological core [6, 7]. Despite the common set of technologies, their application and business models in sectors differ significantly and affect innovative activity of enterprises [8].

In this respect, we suggest that standards serve as a mechanism to boost innovation and adoption of digital technologies in manufacturing industries. Based on a comparative analysis of national initiatives in China, Germany, Japan, the Republic of Korea, the United States in the field, we distinguished three main models describing digital manufacturing standardization in countries. We relied on several parameters for the country analysis, including the national standardization system, industrial and digitalization related strategies, projects in digital manufacturing and standardization, elaboration of reference architecture and international cooperation in the field. This study also gives an overview of the international standardization landscape with respect to the Russian position in it. In order to present recommendations for Russia, we analyzed the Russian and foreign digital agenda and the role of standardization.

1. Digital manufacturing from the standardization perspective

Digitalization becomes indispensable for technological upgrading in different sectors [8]. There is a growing interest in smart manufacturing from government, business and academia. This is evidenced by national digitalization programs, corporate transformation strategies, as well as by a growing number of scientific publications. The available literature combines several approaches to the study of digital production, including technological trends [9–11], design principles [6, 12], the effects of its adoption [7, 13]. Other papers address mainly technological issues introduced by the integration of information technologies in production systems [14, 15]. However, the number of studies focused on innovation outcomes of digital manufacturing standardization is still limited.

As the innovation cycle accelerates, standards and related activities become a tool to solve global challenges, especially in the high-tech sectors of industrial production [16, 17]. The process of standards development is a consensus-based, open and transparent one that facilitates the agreement of stakeholders on technical specifications and implementation [18, 19]. Moreover, standards promote innovation spread by harmonization of technological solutions in complex systems, with replicability and conformity assessment (security, compatibility, etc.) [20].

Thus, standards facilitate knowledge transfer, its dissemination and promote further innovation [18, 21]. Massive use of information technologies requires compatibility of systems, products and services in the global market, the key tool for which are the standards [22, 23]. From the evaluation perspective, standards could serve as an indicator of project effectiveness, including projects supported by the state. Therefore, standards can stimulate development of new technological solutions and improvement of existing ones [20].

Standards in the field of digital technologies have particular characteristics in comparison with other domains. Firstly, digital technologies constitute complex systems comprising a range of parts and elements, both hardware and software, which can be developed by different suppliers in different ways, though interoperability must be ensured [22]. Secondly, modern information systems are designed with high switching costs. For the customer, its implementation means significant expenses for integration, learning, etc. Thirdly, intrinsic network effects are achievable only in case the number of users is growing [23].

Most studies referring to the ICT industry analyze standards elaborated predominantly within industrial consortia. By comparison, industrial and more precisely the machine-building sector receives not enough attention. Machine-building industries generate high added value, and companies of the sector are actively participating in official standardization organizations [24].

In the literature, two classes are distinguished — de facto and de jure standards. De jure or formal standards are developed by standard-setting organizations (SSO) or standard developing organizations (SDO) [24]. De facto standards arise as a result of market competition between several specifications and are introduced mostly by private companies [17]. Due to technology convergence, consortia and alliances become important modes of cooperation in standardization and submit their standards to formal organizations [19, 25].

The role of government in the field of standardization is mainly focused on coordination of standards development and mitigating risks for participants. Public entities contribute to creation of partnerships and alliances at the national level, often being an initiator [20]. Modernization of industries in turn leads to improvement of standardization itself; in particular, it fosters the shift of services standardization to the digital environment and the elaboration of digital standards.

In this paper, standardization is considered to be an efficient government mechanism of digitalization in industries and brings new evidence to the strand of literature dedicated to innovation policy in digital era. Based on national cases of smart manufacturing and comparison with Russia, a set of recommendations for Russia was developed.

2. International landscape in the field of digital manufacturing and standardization

2.1. National policy for digitalization

To select countries for the analysis, we used the WEF Readiness for the Future of Production index. It comprises several dimensions — structure and drivers of production; each of them includes a number of sub-indicators. For our purposes, we assessed countries by two parameters — scale (refers to the production structure parameter) and technology and innovation (refers to the drivers of production). The first one sheds light on the general volume of production and, thus, on the scope of standards application by manufacturing companies, since standards fulfill an economic function only in case of wide recognition. The second reflects the level of technological development in countries and, respectively, allows us to anticipate which countries are likely to be leaders in the field (*Table 1*) [26].

Table 1.

Scores of the selected countries according to the WEF Readiness for the Future of Production

Countries in the ranking	Structure of production – Scale	Drivers of Production –Technology & Innovation
USA	10	1
Germany	4	8
Japan	5	16
Republic of Korea	2	17
China	1	25

Source: [26].

Germany and the USA are on the intersection of these two facets and thus are leaders; China, the Republic of Korea and Japan occupy leading positions in terms of scale. Their strong positions in the field of digital production are supported by the existence of national industrial strategies and a standardization approach.

On the national level, the following parameters were analyzed: the national system of standards development, initiatives in industry, digitalization and standardization, reference architecture of digital production, cooperation among countries. Based on this, we distinguished three main models in the manufacturing standardization landscape: market-centered (the USA, the Republic of Korea), a balanced approach based on a private-public partnership (Germany, Japan) and a government-centered approach (China). The results of the analysis are given below.

National system of standardization

In the USA, standardization is driven mostly by business players, whereas state bodies, including various agencies and institutes, perform a coordinating role by providing the general regulatory framework, related research and expertise [27].

The second model, public-private partnerships, involves greater participation of government, industry associations and major research organizations. This model is attributed to Germany, the Republic of Korea and Japan [28]. It is important to mention that the German approach is characterized by a large focus on technological aspects [29]. With Japan's high involvement in global value chains, the country is under-represented in transnational consortia and alliances led by the U.S. firms. Nevertheless, there are several collaborations of Japanese firms with European, Asian and American partners [30].

The Chinese model is based on mostly state activities: the government raises funds and coordinates various projects within the framework of the "Chinese Standards 2035" strategy. Its main focus is to intensify research and development. Development of their own standards in China is driven by both external factors and internal purposes related to industrial upgrading. China is actively involved in international standards activities, especially in the field of 5G technologies [31].

Initiatives in the field of digital manufacturing and standardization

Standardization is considered to be an important policy tool to tackle economic challenges. The German case reveals that along with research activities and information infrastructure standards ensure competitive advantages for the country, where German business traditionally plays the main role [32]. Overall, digitalization in manufacturing helps to keep Germany high-value manufacturing export, ensures competitiveness in the global value chains for Korea [33]. GVC domination is also the main concern for the U.S., since manufacturing plays a key role for international competitiveness and is supported by the leading position in R&D. China's efforts focus mainly on structural change, since information technologies in manufacturing should favor innovation enhancement of national economy and its technological upgrading [34].

Digital agendas of most leading countries, in addition to the strategies themselves, encompass separate initiatives for industry standardization as well. Such initiatives have aimed at creation of standards and international promotion, i.e. its replication and building smart factories. Projects in standardization cover a wide range of activities, including testing of business models, elaboration of scenarios of technology adoption (e.g., the Labs Network

Industrie 4.0 in Germany), assistance with operational issues related to introduction of digital technologies in manufacturing (Plattform Industrie 4.0 in Germany). Within the framework of such public-private partnerships, government provides strategic guidance and integrates participants into a single ecosystem [35].

Currently there are two dominant approaches in the manufacturing digitalization – German and American. The former focused more on production aspects (hardware) of digitization and integration of cyber-physical systems, while the latter considers digitalization of industry more broadly as part of the system of the industrial internet in sectors.

A market-centered model represents another pattern, where a large part of standards is developed in consortia and alliances, the most significant of which both in the U.S. and international level is the Industrial Internet Consortium (IIC) established in 2014 by leading corporations. The IIC developed the reference architecture for the Internet of Things in industry, healthcare, energy, transport and public services. Participation in the consortium allows companies to get access to test-beds and to get acquainted with the regulatory framework in the field of new technologies. Other international industry consortia and alliances provide a similar support [36].

Along with business initiatives, there are also other U.S. state programs related inter alia to standardization. Under the "Manufacturing USA" program, the Digital Manufacturing and Design Innovation Institute (DMDII) funded by the U.S. Department of Defense and coordinated by the National Institute of Standards and Technology was established. It is a place for collaboration in R&D, commercialization and testing new solutions for different industries [36]. Institutes create effective and replicable solutions for enterprises by consolidating efforts of ministries, private companies, universities and research organizations [37].

Other countries are implementing similar projects. The Korea Manufacturing Innovation 3.0 is part of the comprehensive Korean Creative Economy (CEI) strategy. With respect to specialization of regions, 17 innovation centers that cover a range of digital technologies across the country were created. These centers are managed by large Korean corporations like Samsung, Hyundai, etc. ("Smart machinery" led by Doosan, "Shipbuilding/machinery", "Textile/electronics" by Hyundai Heavy Industries, "Smart Factory" by Samsung, automobile by Hyundai and Kia Motors) [37, 38].

Japan has both business-led (IoT Acceleration Consortium) and state-led (Industrial Value Chains Initiative – IVI, "Connected Industries", the Robot Revolution Initiative of Japan – RRI) initiatives in smart manufacturing, which are in general coordinated by government. The Industrial Value Chain Initiative (IVI) was introduced in 2015 as a cooperation platform of national manufacturing firms coordinated by the Ministry of the Economy, Trade and Industry (METI). It accumulates joint use cases and scenarios in smart manufacturing thus helping to implement flexible standards and models in different sectors. There is a strong cooperation across working groups on standardization of the initiatives [39]. The government makes a particular effort to promote standards for robotic systems across different sectors within the country and globally, since robotics is one of the five priority areas for the national economy in the long-run, according to the Robot Revolution Initiative of Japan (RRI) [40].

China follows the third approach, which is characterized by the development of its own model of digitalization based on the combination of the best world practices with a leading government role. The main task is not only modernization, but also radical structural transformation of industry, with an increase of value added and reorientation towards high-tech products [36]. Active state support is based on bilateral cooperation with leaders (Germany and the United States), while the "Standards of China 2035" strategy is aimed at creating Chinese standards by intensifying efforts in the field of research and development [41].

Despite the similarity of national goals for standardization, the countries' differences in economic and technological competences require more adaptive mechanisms in standardization. Public-private partnerships are the most appropriate way to realize digital projects.

Reference architecture of digital manufacturing

Reference architectures are more conventional to information and communication technologies. From smart manufacturing perspective, the term means "a model for a class of architectures," i.e. a pool of rules and principles to describe physical systems in a digital world [42].

Leading countries are striving to develop their own reference models. The Reference architecture model Industrie 4.0 (RAMI 4.0), introduced by the Plattform Industrie 4.0 and the Industrial Internet Reference Architecture (IIRA) by the IIC are main approaches. Japan has the harmonized with the RAMI principles the Industrial Value Chain Reference Architecture (IVRA) [43]. Korea is also undertaking efforts to elaborate own reference model with a particular attention to SME digitalization (Smart Factory Reference Model) [44].

China is also developing its own model – the China Intelligent Manufacturing System Architecture (IMSA). In the framework Sino-Germany bilateral cooperation the IMSA and the RAMI are harmonized, which is determines Chinese orientation on the German approach [45].

Cooperation in the field of digital manufacturing

In order to promote national approaches to digitalization of production, the leading countries are building networks with economic and trade partners in a bilateral way.

More precisely, Germany collaborates with China (a joint Sino-German Commission on Standardization cooperation); Japan in the framework of the Plattform I4.0 and Japanese Robot Revolution Initiative, with Korea within the Smart Factory Web, which is a joint Korean-Germany initiative supported by government and some others [45–48]. In addition at the national level, Germany (Plattform Industrie 4.0) with France (Alliance Industrie du Futur) and Italy (Piano Industria 4.0) have established a trilateral initiative and a working group based on the German reference architecture (The Paris Declaration for Smart Manufacturing) [49].

The IIC network of partners is also large and comprises cooperation between the Japanese IVI and the IIC (the Liaison Working Group) [50], the Korean Smart Factory Web (a testbed of the Industrial Internet Consortium) and some others [48]. Moreover, alignment between the RAMI and the IIRA shapes the digital agenda in industry and helps to secure global interoperability [51].

A summary analysis of the countries is shown in *Figure 1.*

Standardization receives greater attention from national governments and international organizations, occupying an important place in bilateral and multilateral relations. Currently, standards development in production industries is led by Germany and the USA, with reference architectures and cross-country cooperation. The major part relies on wide public-private partnerships.

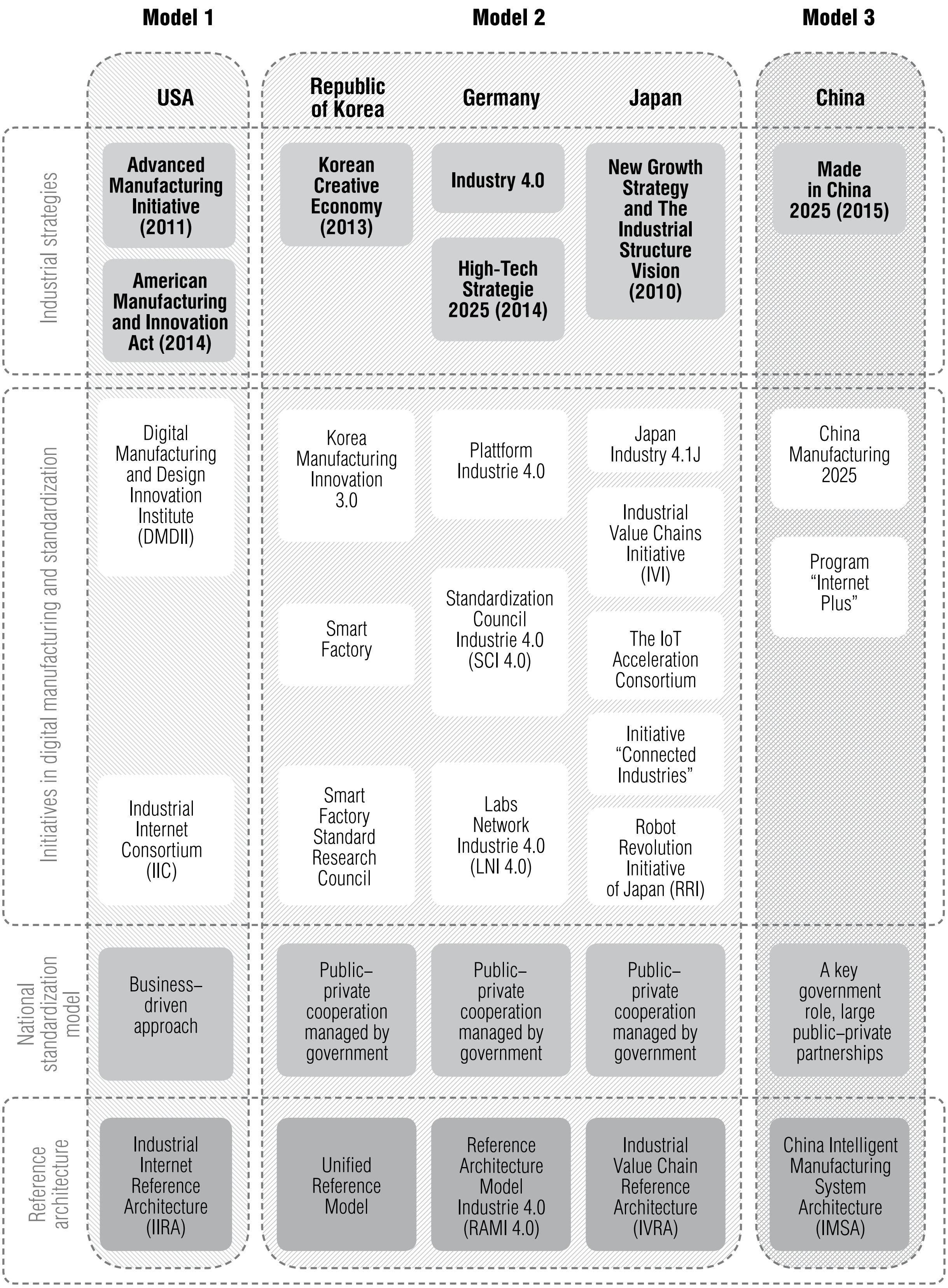


Fig. 1. Standardization digital manufacturing landscape in selected countries

2.2. International bodies involved with standardization in digital manufacturing

International organizations in standardization, primarily the ISO and IEC, pay significant attention to digital production and related issues. Some initiatives are performed via joint technical committees (JTC) (*Figure 2*).

Technical Committees (TC) are working on functional compatibility of systems and cybersecurity, as well as the extension and systematization of existing approaches by proposed participants. In this regard, within the ISO and IEC two strategic bodies have been established – ISO/SMC and IEC/SC, which bring together representatives of all technical committees related to production and information technologies [35]. In addition to joint initiatives of the ISO and IEC, there are several technical committees focused on smart production – the ISO TC 184 (product data, compatibility, integration, architecture for industrial automation, and TC 65 IEC (control of industrial processes and their automation and integration of product data and processes).

Along with international official organizations in the field of digital production, there are a number of industrial consortia and alliances that are becoming important participants in developing standards. Among them are the MTConnect, OPC Foundation and MESA. Standards developed by alliances may be submitted to the official international organizations for standardization. Some of them provide free access to their standards and technical information; for example, the OPC Foundation conducts its own certification and testing [29]. With a significant number of players in the international scope, standards and norms alignment remains a key point [51].

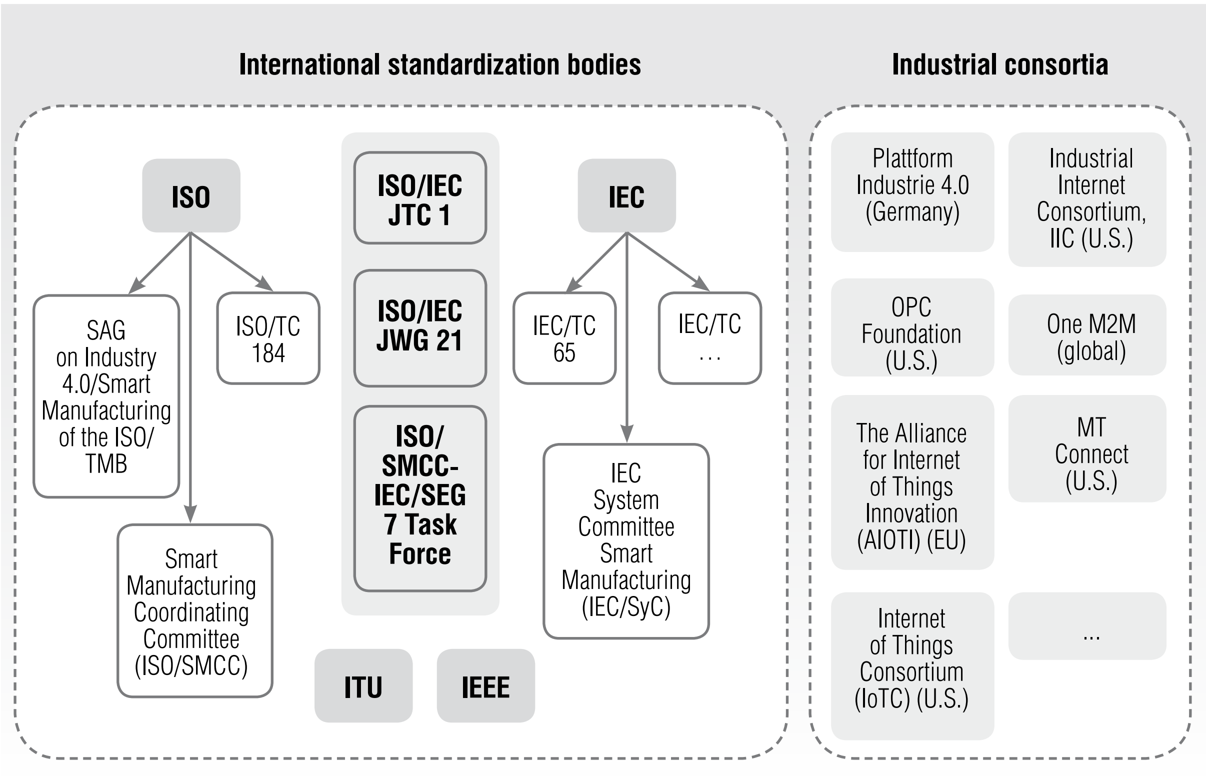


Fig. 2. International landscape of digital manufacturing

3. Russian perspectives in standardization of digital manufacturing

3.1. Factors and conditions of digitalization in Russia

The digital agenda in Russia is associated with significant socio-economic effects. By 2030 digitalization may become the key driver of economic growth: from 2017 to 2030 the contribution of sectoral digital transformation may account for as much as a 30% increase in GDP. The most significant effects related to digital technologies may be observed in the machine-building and chemical industries, where production efficiency will be about 5% due to the total factor productivity and the contribution of capital to value added [52].

The overall level of digital technologies use age varies widely across industries and enterprises. In recent years, large companies have introduced specialized software for functional fields, including management of finance operations, customer interaction, etc. Russian business considers digital technologies to be a source of strengthening market positions and gaining new opportunities in the long term [53]. However, the degree to which digital solutions are adopted in business remains low. From the sectoral perspective, manufacturing industry demonstrates a higher level of digitalization. Manufacturing firms rank first in electronic data interchange (72.3%), but they lag far behind the ICT sector in terms of cloud services use (23.2% versus 34.7%). In almost half of manufacturing enterprises (46%), the level of digitalization is considered to be low [54].

Large-scale digitalization in industry is constrained by the lack of domestic equipment and specialized software. Russian industrial companies are not actively investing in domestic digital solutions and purchase predominantly foreign products and services. For example, the share of foreign robots, as well as smart control systems, is almost 100%, CNC systems – 65% [55]. Imported industrial software, including PLM, CAD, CAM, CAE, accounted for 88% in 2014 [56]. As part of the sectoral plans for import substitution, the share of foreign products should be reduced by 2020 (industrial robots to 69%, CNC systems – 20%, engineering software – 60%) [55, 56].

Nevertheless, Russia has the opportunities and resources for a transition to a digital economy. This includes a highly dynamic information technology market, the existence of public research and engineering centers, development of private research organizations, availability of scientific schools, and a high level information and communication infrastructure. The greatest demand is expected in sectors with complex products and high modeling needs (aerospace, automotive, shipbuilding); with large production capacities (production of equipment and general purpose machinery, electrical equipment), as well as transport engineering [57].

3.2. Government policy in the digital economy on the national and supranational levels

In Russia, goals related to digitalization are assigned a strategic priority and are set down in the Decree of the President of the Russian Federation No. 204 of 07.05.2018 "On national goals and strategic objectives for the development of the Russian Federation for the period up to 2024."

The main initiative is the program entitled "The digital economy of the Russian Federation" adopted in 2017 and transformed into a national project in 2018. The program includes six Federal Projects in the areas of regulations governing the digital environment, information infrastructure, information security, personnel for the digital economy, digital technologies and digital public administration. The activities of the national program are aimed inter alia at promoting the creation and implementation of digital technologies in the economic and social sectors. At the same time, the main measures for standardization are set down in the Federal project "Normative regulation of the digital environment" [58].

Along with the federal programs, a number of ministerial projects are being developed, including on the digital industry managed by the Ministry of Industry and Trade. It also includes incentives for standardization and use of digital technologies [59].

A set of measures for industrial modernization is also included in the National Technological Initiative (NTI) in the area of "TechNet" and includes improvement of activities in the field of standardization and certification, the development of new approaches to standardization and the introduction of "factories of the future" standards [57].

The digital agenda including industry is one of the key measures to boost economic growth in the Eurasian Economic Union (EAEU). In accordance with the decision of the Supreme Eurasian Economic Council of October 11, 2017 No 12 "On the main directions of the digital policy of the EAEU until 2025," sectoral and cross-sectoral digital transformation are the main vectors for deepening economic cooperation in the EAEU. Technological modernization of production chains requires the creation of a single digital environment between national business and public authorities [60]. The Eurasian digital platform and its integration with the unified information system of the EAEU will be a mechanism to implement this. Deployment of digital platforms will facilitate efficient use of data throughout the value chain. Thus, the integrated application of digital technologies in sectors opens opportunities for development of new business models. The positive economic effect of the joint digital initiative can increase the total GDP of the participating countries by 11% by 2025. This figure is twice higher compared to implementation of digital initiatives separately [61]. In this regard, standardization is the precondition of digitalization projects.

At the supranational level, as well as at the national level, it is necessary to remove legal barriers to the deployment of digital technologies, establish a common digital environment and agree on the basic terminology and concepts related to the digital economy.

3.3. The Russian approach to standardization of digital manufacturing

National system of standardization

Standardization in Russia is now driven mostly by the government, which follows the third model described in this paper. Rosstandart and its technical committees play a key role in the standardization system. Along with Federal Projects implementation, the national standardization system will be modernized by computer models during the lifecycle, sectoral initiatives like smart manufacturing, smart cities, machine-building industry upgrading, information security, etc. Taking into account global trends in alliances and consortia, convergence of digital agendas and approaches to standardization within the EAEU is an important direction for strengthening Russia's position in the area. This requires consolidation of methodological foundations of digital transformation and business involvement [62, 63].

Russian initiatives in the field of digital manufacturing and standardization

Within the framework of the national program, standards are regarded as a mechanism to stimulate innovative activity of companies. In order to encourage business participation, development procedures of standardization documents will be simplified. This will reduce costs to participate in the standardization processes [58]. However, to enlarge business initiatives it also important to create conditions where standards will inevitably succeed both in external and in the internal markets. In other words, the role of standards within the country should be enlarged.

For wide use of digital technologies decision-making based on the experiments with the digital environment, use of digital design and operational documentation, digital product models should be approved. It is anticipated to develop standards in the field of the Internet of Things, cyber-physical systems, big data analysis, etc. [58]. At the beginning of 2019, Rosstandart approved the first national standard of the Internet of Things entitled "A Protocol of wireless data transmission based on narrowband modulation of radio signal (NB-Fi)." This was developed by the Technical Committee 194 "Cyber-physical systems" [64].

In addition to the EAEU level, digitalization issues are being discussed in the Intergovernmental Technical Committee 22 "Information technologies," which operates within the framework of the Interstate Council for Standardization, Metrology and Certification of the CIS [65]. Given the fact that the Committee is a permanent body within the framework of the ISO/IEC 1 joint committee, harmonization of approaches at the level of the CIS is one of the possibilities to promote the Russian approach in digital manufacturing. The Russian standardization system in digital production is presented in *Figure 3. Table 2* lists the Russian technical committees (TC) involved in the development of technologies standards.

In 2014–2019, TC focused on new digital technologies were created in artificial intelligence, cyber-physical systems, hardware and software distributed registry and blockchain technologies, robotics, additive manufacturing and digital modeling. An important point is that the technical committee 194 operates on the international level. Along with the harmonization of international standards, the TC develops national standards and promotes them in the ISO and IEC [66]. Russian system standardization also includes national professional consortia and associations, as well as development institutions (JSC RVC, etc.).

Table 2.

Russian technical committees (TC) involved in the development of digital technologies

No. of the TC	Name of the TC and year of establishment
164	Artificial intelligence (2019)
194	Cyber–physical systems (2017)
159	Hardware and software distributed registry and blockchain technologies (2017)
141	Robotics (2016)
182	Additive manufacturing (2015)
700	Mathematical modeling and high–performance computing technologies (2014)

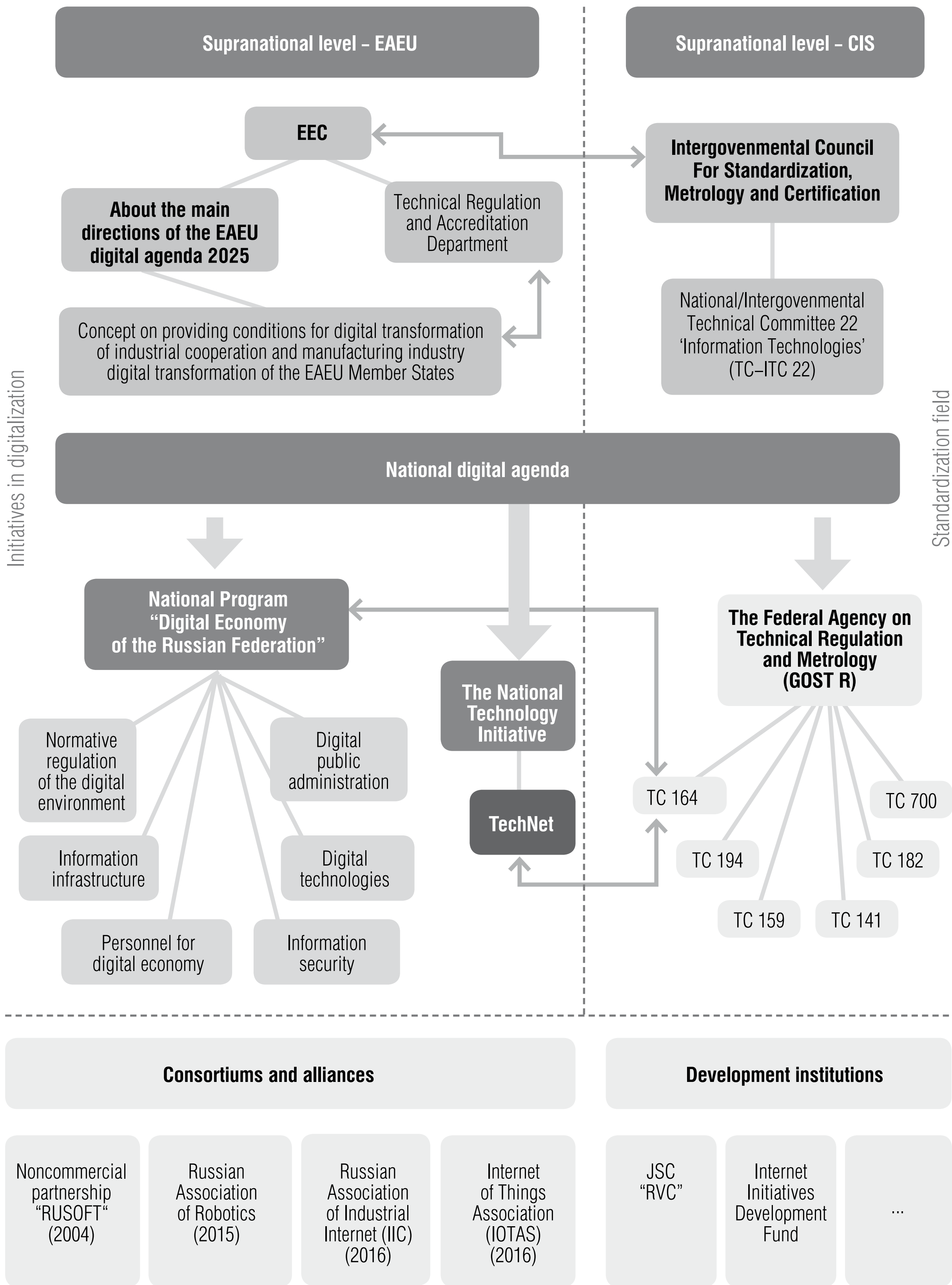


Fig. 3. The Russian standardization landscape in digital manufacturing

Russia's cooperation with other countries in the field of digital manufacturing standardization and participation in international organizations

Currently, Russian business has not established a national approach in digital manufacturing. To this end, it cooperates with German industrial companies and implements their experience in advanced manufacturing technologies. The most obvious examples of cooperation fall within the automotive industry, railway engineering, machine tool industry and some others. By the joint German–Russian GRID initiative, German business shares its experience and uses cases in digital transformation of manufacturing. German business also helps to engage in activities small and medium enterprises that are laggards in digitalization in comparison with large companies [67].

Russia also participates in the ISO and IEC activities focused on digital technologies. National representatives are members of the Smart Manufacturing Coordination Committee (SMCC), the joint ISO and IEC technical Committee and its subcommittees, ISO TC 184. In the IEC Committee TC 65 "Measurement, management and automation of production processes" Russia acts as an observer country.

4. Implications for Russia in the field of digital manufacturing standardization

Russian initiatives in transformation of production industries through digital technologies are in line with global trends. It is anticipated to develop standards in such areas as the Internet of Things, cyber-physical systems, etc. At the same time, Russia has not fully established a national model of smart manufacturing digitalization: there is no standardization framework for digital transformation in sectors which will ensure compatibility of systems and consistency with particular sectoral needs.

To date, the level of absorption of digital technologies in industry remains low, since enterprises use a traditional set of information and communication solutions. To a large extent, this is due to the lack of domestic technologies and related standards governing their implementation and use. In order to change the situation, business must be incentivized to create new solutions and further advocate them within the country and outside it. In turn, this requires an appropriate policy that will encourage companies to invest in digital solutions. A broad approach should be implemented, since integration of hardware and software (physical and digital components) becomes essential for complex production systems.

The leading countries in smart manufacturing are trying to ensure market dominance of their standards, references and models. Russia has a number of conditions for sectoral digital transformation, namely dynamically developing ICT: the value added growth of the ICT-sector (2.8%) in 2017 was almost double the GDP growth (1.6%) [68]. It also includes the availability of digital products and services that can replace comparable foreign ones, highly qualified personnel in the field of information and communication technologies, as well as talented STEM (Science, Technology, Engineering and Mathematics) graduates. In the framework of the national program "Digital economy of the Russian Federation," a set of new measures of financial support for developing digital technologies is provided. With regard to standards, the program provides an impetus to upgrading the national standardization system and could raise the role of standards in the overall innovation system.

Standardization in digital manufacturing contributes to modernization of the national standardization system and could significantly increase the relevance of standards to the innovation system. Use of standards as one of the tools to stimulate digital transformation of industry will require:

✦ development of a national model of digital manufacturing standardization;

✦ coordination of measures related to standards development within the strategic and program documents;

✦ launch and harmonization of sectoral digitalization projects;

✦ development of public-private mechanisms for implementing digital projects;

✦ closer cooperation of working groups of the national project "Digital economy of the Russian Federation," the National Technology Initiative, industrial associations and unions, government agencies and other stakeholders;

✦ support for cooperation between technical committees, government agencies, businesses and the scientific community to overcome the fragmentary digitalization in sectors;

✦ alignment of mechanisms between the national model of digitalization and standardization with that at the EAEU level;

✦ exploration of the areas where internationally recognized standards are not currently developed in order to enhance its development.

Conclusion

The effectiveness of adopting and using digital technologies depends to a large extent on the availability of standards. The main approaches to standardization in digital manufacturing revealed in this paper allowed us to identify mechanisms and directions to digitalization. Standards in the digital era ensure technology transfer and interoperability of systems, including hardware and software elements.

The efforts of the leading countries in the field of smart manufacturing are focused on standards and reference architectures in global market, expansion of bilateral and multilateral cooperation with partner countries concerning standardization issues of digital technologies, as well participating in relevant international organizations.

The Russian digital agenda largely coincides with the international one, however business involvement is still insufficient. With regard to state support, Russia is close to the China model, where standards are a mechanism of structural transformation in industry. In this regard, the development of a common framework for sectoral digitization is a key task of the policy. Its absence creates risks associated with dependence on foreign standards and solutions. Creating a national approach facilitates not only the achievement of the goals set down in the national project "Digital economy of the Russian Federation," but also provides an impetus to the innovation system and technological capabilities of Russian business in the EAEU and CIS. Russia is represented in all key organizations for standardization, but to promote Russian initiatives in digital manufacturing at the international level, especially in the ISO and IEC, it is necessary to provide additional measures and strongly enhance business participation.

Standards have always played a major role in industrial firms. With the rise of digital technologies, standardization patterns in manufacturing draw closer to the ICT. Therefore it could stimulate innovation performance by using diffusion mechanisms attributed to the ICT. This is particularly important for Russia, where low innovation activity hampers digitalization. By implementing new measures of the National Project, standards become an important channel for knowledge diffusion in industry and more broadly for upgrading technology.

The results of the paper can be useful for policymakers responsible for digitalization and digital transformation in industries, as well as for companies and other interested in participating in the national digital agenda stakeholders. ■

References

1. European Commission (2018) *Re-Finding Industry Report from the High-Level Strategy Group on Industrial Technologies*. Available at: https://ec.europa.eu/research/industrial_technologies/pdf/re_finding_industry_022018.pdf (accessed 10 June 2019).
2. OECD (2017) *The Next Production Revolution. Implications for Governments and Business*. Available at: https://espas.secure.europarl.europa.eu/orbis/sites/default/files/generated/document/en/9217031e.pdf (accessed 14 April 2019).
3. Ananyin V.I., Zimin K.V., Lugachev M.I, Gimranov R.D., Skripkin K.G. (2018) Digital organization: Transformation into the new reality. *Business Informatics*, no 2, pp. 45–54.
4. Park H., Kim H., Joo H., Song J. (2016) Recent advancements in the Internet-of-Things related standards: A oneM2M perspective. *ICT Express*, vol. 2, no 3, pp. 126–129.
5. Tokareva M.S., Vishnevskiy K.O., Chikhun L.P. (2018) The impact of the Internet of Things technologies on economy. *Business Informatics*, no 3, pp. 62–78.
6. Ghobakhloo M. (2018) The future of manufacturing industry: a strategic roadmap toward Industry 4.0. *Manufacturing Technology Management*, vol. 29, no 6, pp. 910–936.
7. Strange R., Zucchella A. (2017) Industry 4.0, global value chains and international business. *Multinational Business Review*, vol. 25, no 3, pp. 174–184.
8. Akatkin Y.M., Karpov O.E., Konyavskiy V.A., Yasinovskaya E.D. (2017) Digital economy: Conceptual architecture of a digital economic sector ecosystem. *Business Informatics*, no 4, pp. 17–28.
9. Esmaeilian B., Behdad S., Wang B. (2016) The evolution and future of manufacturing: A review. *Journal of Manufacturing Systems*, no 39, pp. 79–100.
10. Müller J., Voigt K. (2018) Sustainable industrial value creation in SMEs: A comparison between Industry 4 and Made in China 2025. *International Journal of Precision Engineering and Manufacturing-Green Technology*, vol. 5, no 5, pp. 659–670.
11. Kagermann H., Anderl R., Gausemeier J., Schuh G., Wahlster W. (2016) I*ndustrie 4.0 in a Global Context: Strategies for Cooperating with International Partners (Acatech STUDY)*. Munich: Herbert Utz Verlag.
12. Gilchrist A. (2016) Industry 4.0. Apress.
13. Szalavetz A. (2019) Industry 4.0 and capability development in manufacturing subsidiaries. *Technological Forecasting and Social Change*, vol. 145, pp. 384–395.
14. Weyer S., Schmitt M., Ohmer M., Goreck D. (2015) Towards Industry 4.0 – Standardization as the crucial challenge for highly modular, multi-vendor production systems. *IFAC-PapersOnLine*, vol. 48, no 3, pp. 579–584.
15. Choi S., Jung C., Kulvatunyou B., Morris K.C. (2016) An analysis of technologies and standards for designing smart manufacturing systems. *Journal of Research of the National Institute of Standards and Technology*, vol. 121, pp. 422–433.
16. Wakke P., Blind K., De Vries H. (2015) Driving factors for service providers to participate in standardization: Insights from the Netherlands. *Industry and Innovation*, vol. 4, no 22, pp. 299–320.
17. Blind K., Mangelsdorf A. (2016) Motives to standardize: Empirical evidence from Germany. *Technovation*, no 48–49, pp. 13–24.
18. Narayanana V.K., Chen T. (2012) Research on technology standards: Accomplishment and challenges. *Research Policy*, no 41, pp. 1375–1406.
19. Baron J., Ménière Y., Pohlmann T. (2014) Standards, consortia, and innovation. *International Journal of Industrial Organization*, no 36, pp. 22–35.
20. Zoo H., De Vries H., Lee H. (2017) Interplay of innovation and standardization: Exploring the relevance in developing countries. *Technological Forecasting & Social Change*, no 118, pp. 334–348.
21. Wang Z., Zhang M., Sun H., Zhu G. (2016) Effects of standardization and innovation on mass customization: An empirical investigation. *Technovation*, no 48–49, pp. 79–86.
22. Foster C., Heeks R. (2013) Innovation and scaling of ICT for the bottom-of-the-pyramid. *Journal of Information Technology*, no 28, pp. 296–315.
23. Kim D.-h., Leeb H., Kwak J. (2017) Standards as a driving force that influences emerging technological trajectories in the converging world of the Internet and things: An investigation of the M2M/IoT patent network. *Research Policy*, no 46, pp. 1234–1254.
24. Shin D.-H., Kim H., Hwang J. (2015) Standardization revisited: A critical literature review on standards and innovation. *Computer Standards & Interfaces*, no 38, pp. 52–57.

25. Delcamp H., Leiponen A. (2013) Innovating standards through informal consortia: The case of wireless telecommunications. *International Journal of Industrial Organization*, no 36, pp. 36–47.

26. World Economic Forum (2018) *Readiness for the Future of Production Report 2018*. Available at: http://www3.weforum.org/docs/FOP_Readiness_Report_2018.pdf (accessed 10 June 2019).

27. United States Government (2018) *Strategy for American leadership in advanced manufacturing*. Available at: https://www.whitehouse.gov/wp-content/uploads/2018/10/Advanced-Manufacturing-Strategic-Plan-2018.pdf (accessed 10 June 2019).

28. ITIF (2018) *Manufacturing digitalization: Extent of adoption and recommendations for increasing penetration in Korea and the U.S.* Available at: http://www2.itif.org/2018-korean-manufacturing-digitalization.pdf (accessed 10 June 2019).

29. ITIF (2018) *Why manufacturing digitalization matters and how countries are supporting it*. Available at: http://www2.itif.org/2018-manufacturing-digitalization.pdf (accessed 10 June 2019).

30. Ministry of Economy, Trade and Industry (2017) *New industrial structure vision*. Available at: http://www.meti.go.jp/english/publications/pdf/vision_171222.pdf (accessed 10 June 2019).

31. Second European Standardization Expert in China (2018) *Chinese Standards 2035, the standardization strategy research is kicked off*. Available at: http://www.sesec.eu/24-05-2018-chinese-standards-2035-the-standardization-strategy-research-is-kicked-off/ (accessed 10 June 2019).

32. Mattauch W. (2017) *Digitising European industries – Member states profile: Germany*. Available at: https://ec.europa.eu/futurium/en/system/files/ged/de_country_analysis.pdf (accessed 10 June 2019).

33. Federal Ministry for Economic Affairs and Energy (2016) *Digital Strategy 2025*. Available at: https://www.de.digital/DIGITAL/Redaktion/EN/Publikation/digital-strategy-2025.pdf?__blob=publicationFile&v=8 (accessed 10 June 2019).

34. Standardization Administration of the P.R.C. (2017) *Situation of Chinese intelligent manufacturing standardization & international cooperation suggestions*. Available at: https://www.bmwi-registrierung.de/gpg20sc/pdf/07.10.%20Presentation%20-%20Ms%20Hong%20Dai.pdf (accessed 10 June 2019).

35. DIN (2018) *The Reference Architectural Model RAMI 4.0 and the Standardization Council as an element of success for Industry 4.0*. Available at: https://www.din.de/blob/271306/340011c12b8592df728bee3815ef6ec2/06-smart-manufacturing-jens-gayko-data.pdf (accessed 10 June 2019).

36. Kagermann H., Anderl R., Gausemeier J., Schuh G., Wahlster W. (Eds.) (2016) *Acatech. Industrie 4.0 in a global context. Strategies for cooperating with international partners.* Munich: Herbert Utz Verlag.

37. Export.Gov (2018) *Korea – Manufacturing Technology – Smart Factory*. Available at: https://www.export.gov/article?id=Korea-Manufacturing-Technology-Smart-Factory (accessed 10 June 2019).

38. Ministry of Science and ICT (2018) *Smart Korea Creative Economy*. Available at: https://www.msit.go.kr/SYNAP/skin/doc.html?fn=8a5c9cc0262b0da349d6d641227368a7&rs=/SYNAP/sn3hcv/result/201908/ (accessed 10 June 2019).

39. IVI (2016) *The Industrial Value Chain Initiative. A Japanese contribution to smart manufacturing*. Available at: https://iv-i.org/en/docs/doc_160428_hannover.pdf (accessed 10 June 2019).

40. Ministry of Economy, Trade and Industry (2017) *"Connected Industries" Tokyo Initiative 2017*. Available at: http://www.meti.go.jp/english/press/2017/pdf/1002_004b.pdf (accessed 10 June 2019).

41. Second European Standardization Expert in China (2015) *China Internet Plus Strategy*. Available at: http://www.sesec.eu/app/uploads/2015/06/2015_05_SESECIII_Newsletter_April_2015_Annex02_China_Internet_Plus_Strat....pdf (accessed 10 June 2019).

42. DIN (2018) *German Standardization Roadmap Industrie 4.0. Version 3*. Available at: https://www.din.de/blob/65354/57218767bd6da1927b181b9f2a0d5b39/roadmap-i4-0-e-data.pdf (accessed 14 January 2019).

43. IVI (2016) *Industrial Value Chain Reference Architecture*. Available at: https://iv-i.org/en/docs/Industrial_Value_Chain_Reference_Architecture_170424.pdf (accessed 10 June 2019).

44. Park S. (2016) Development of innovative strategies for the Korean manufacturing industry by use of the Connected Smart Factory (CSF). *Procedia Computer Science, Asan, Korea, 16–18 August 2016*, no 91, pp. 744–750. Available at: https://core.ac.uk/download/pdf/82146226.pdf (accessed 10 June 2019).

45. Bundesministerium für Wirtschaft und Energie (2018) *Fortschrittsbericht 2018. Industrie 4.0 anwenden. Wegweisend. Praxisnah. Vernetzt*. Available at: https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/hm-2018-fortschrittsbericht.pdf?__blob=publicationFile&v=5 (accessed 10 June 2019).

46. Robot Revolution Initiative (2017) *What's RRI / WG1? – toward horizontal dynamic manufacturing*. Available at: https://www.jmfrri.gr.jp/content/files/Open/2017/20170315_RRI_forCeBIT/RRI_for_CeBIT_IoT_panel.pdf (accessed 10 June 2019).

47. Industrial Value Chain Initiative (2018) *The Allianz Industrie 4.0 Baden-Württemberg and the Industrial Value Chain Initiative Sign MoU.* Available at: https://iv-i.org/wp/wp-content/uploads/2018/03/Allianz-I4.0-IVI-MOU-press-release.pdf (accessed 10 June 2019).
48. Fraunhofer IOSB (2019) *Smart Factory Web*. Available at: https://www.iosb.fraunhofer.de/servlet/is/81769/ (accessed 10 June 2019).
49. Plattform Industrie 4.0 (2018) *Paris Declaration for Smart Manufacturing – by the Working Group "Standardization and Reference Architecture" of the Trilateral Cooperation, Digitizing European Industry*. Available at: https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/wg3-trilaterale-coop.pdf?__blob=publicationFile&v=6 (accessed 10 June 2019).
50. Industrial Internet Consortium (2017) *The Industrial Internet Consortium and The Industrial Value Chain Initiative Sign MoU.* Available at: https://www.iiconsortium.org/press-room/04-26-17.htm (accessed 10 June 2019).
51. Industrial Internet Consortium, Platform Industrie 4.0 (2017) *Architecture Alignment and Interoperability and Industrial Internet Consortium and Plattform Industrie 4.0 Joint Whitepaper*. Available at: https://www.iiconsortium.org/pdf/JTG2_Whitepaper_final_20171205.pdf (accessed 10 June 2019).
52. HSE (2018) *The contribution of digitalization to the growth of the Russian economy*. Express-information on digital economy, no. 91 (8). Available at: https://issek.hse.ru/data/2018/07/04/1152915836/NTI_N_91_04072018.pdf (accessed 10 June 2019) (in Russian).
53. McKinsey (2017) *Digital Russia: A new reality*. Available at: http://www.tadviser.ru/images/c/c2/Digital-Russia-report.pdf (accessed 10 June 2019) (in Russian).
54. HSE (2018) *Digitalization of business processes. Express-information on digital economy*, no 98 (9). Available at: https://issek.hse.ru/data/2018/08/22/1154862864/NTI_N_98_22082018.pdf (accessed 10 June 2019) (in Russian).
55. Ministry of Industry and Trade of the Russian Federation (2017) *Action plan for import substitution in the machine tool industry of the Russian Federation*. Order of the Ministry of Industry and Trade of the Russian Federation, no 4358, 12 December 2017. Available at: https://gisp.gov.ru/plan-import-change/8679776/#popdoc (accessed 10 June 2019) (in Russian).
56. Ministry of Digital Development, Communications and Mass Media of the Russian Federation (2015) *On approval of the import substitution plan]. Order of the Ministry of Digital Development, Communications and Mass Media of the Russian Federation*, no 96, 01 April 2015. Available at: http://minsvyaz.ru/uploaded/files/prikaz-ot-01-04-2015--96.pdf (accessed 10 June 2019) (in Russian).
57. NTI (2018) *Action plan ("roadmap") "Technet" (advanced manufacturing technology) of the National technology initiative*. Approved by the Presidium of the Council for economic modernization and innovative development of Russia, protocol no 1, 14 February 2017]. Available at: http://www.nti2035.ru/docs/%D0%94%D0%9A%20%D0%A2%D0%B5%D1%85%D0%BD%D0%B5%D1%82%20-%20%D0%BF%D1%80%D0%B8%D0%BB%D0%BE%D0%B6%D0%B5%D0%BD%D0%B8%D0%B5%20%D0%BA%20%D0%BF%D1%80%D0%BE%D1%82%D0%BE%D0%BA%D0%BE%D0%BB%D1%83%20%D0%B7%D0%B0%D1%81%D0%B5%D0%B4%D0%B0%D0%BD%D0%B8%D1%8F%20%D0%BF%D1%80%D0%B5%D0%B7%D0%B8%D0%B4%D0%B8%D1%83%D0%BC%D0%B0%20%D0%A1%D0%BE%D0%B2%D0%B5%D1%82%D0%B0.pdf (accessed 10 June 2019) (in Russian).
58. The Government of Russia (2018) *The National program "Digital economy of the Russian Federation"*. Available at: http://static.government.ru/media/files/urKHm0gTPPnzJlaKw3M5cNLo6gczMkPF.pdf (accessed 11.06.2019).
59. Comnews (2019) *The Ministry of Industry and Trade outlined the contour of "Digital industry"*. Available at: https://www.comnews.ru/content/119733/2019-05-23/minpromtorg-ochertil-kontur-cifrovoy-promyshlennosti (accessed 10 June 2019) (in Russian).
60. The World Bank (2017) *The Digital agenda of the Eurasian economic Union until 2025: Outlook and recommendations*. Available at: http://www.eurasiancommission.org/ru/act/dmi/SiteAssets/%D0%9E%D0%B1%D0%B7%D0%BE%D1%80%20%D0%92%D0%91.pdf (accessed 10 June 2019) (in Russian).
61. EAEU (2017) *About the main directions of implementation of the digital agenda of the Eurasian Economic Union till 2025.* Available at: http://www.eurasiancommission.org/ru/act/dmi/workgroup/Documents/%D0%9E%D1%81%D0%BD%D0%BE%D0%B2%D0%BD%D1%8B%D0%B5%20%D0%B4%D0%BE%D0%BA%D1%83%D0%BC%D0%B5%D0%BD%D1%82%D1%8B/%D0%A0%D0%B5%D1%88%D0%B5%D0%BD%D0%B8%D0%B5%20%D0%92%D0%95%D0%AD%D0%A1%20%E2%84%9612_%D0%9E%D1%81%D0%BD%D0%BE%D0%B2%D0%BD%D1%8B%D0%B5%20%D0%BD%D0%B0%D0%BF%D1%80%D0%B0%D0%B2%D0%BB%D0%B5%D0%BD%D0%B8%D1%8F%20%D1%80%D0%B5%D0%B0%D0%BB%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D0%B8%20%D1%86%D0%B8%D1%84%D1%80%D0%BE%D0%B2%D0%BE%D0%B9%20%D0%BF%D0%BE%D0%B2%D0%B5%D1%81%D1%82%D0%BA%D0%B8%20%D0%95%D0%90%D0%AD%D0%A1.pdf (accessed 10 June 2019) (in Russian).

62. The Government of Russia (2019) *On management system of implementation of the national program "Digital economy of the Russian Federation"*. Resolution no 234, 2 March 2019. Available at: http://government.ru/docs/35964/ (accessed 10 June 2019) (in Russian).
63. Aviation Union (2018) *Draft of "The development concept of the national standardization system up to 2027"*. Available at: http://www.aviationunion.ru/Files/Cogc_razv_19032018.pdf (accessed 10 June 2019) (in Russian).
64. The Government of Russia (2018) *Action plan ("road map") to improve legislation and eliminate administrative barriers in order to ensure the implementation of the National Technology Initiative in the field of "TechNet" (advanced production technologies)*. Available at: http://www.rvc.ru/upload/doc/NDK_technet.pdf (accessed 10.06.2019) (in Russian).
65. Euro-Asian Council for Standardization, Metrology and Certification (2018) *Euro-Asian Council for Standardization. Information about Interstate Technical Committee 22*. Available at: http://www.mgs.gost.ru/TKSUGGEST/%D0%9C%D0%A2%D0%9A2014.nsf/84eb0d5919ea20bac325653100289c4a/f26ec86e4be3c5614325801e00504d7e?OpenDocument (accessed 10 June 2019) (in Russian).
66. Rosstandart (2019) *Technical Committee 194*. Available at: http://webportalsrv.gost.ru/portal/TKSUGGEST/TK2006.nsf/84eb0d5919ea20bac325653100289c4a/112ceb451b548663432580f900394dfd?OpenDocument (accessed 10 June 2019) (in Russian_.
67. Germany Trade & Invest (2018) *Deutschland und Russland kooperieren bei Industrie 4.0 Nachholbedarf bei Automatisierung der Produktion bietet Geschäftschancen*. Available at: https://www.gtai.de/GTAI/Navigation/DE/Trade/Maerkte/suche,t=deutschland-und-russland kooperieren-bei-industrie-40,did=1905828.html (accessed 10 June 2019).
68. HSE (2018) *ICT sector in Russia. Express-information on digital economy*, no 110 (10). Available at: https://issek.hse.ru/data/2018/11/14/1141212573/NTI_N_110_14112018.pdf (accessed 10 June 2019) (in Russian).

About the authors

Yulia V. Turovets

Expert, Digital Economics Center; Research Assistant, Research Laboratory for Science and Technology Studies, Institute for Statistical Studies and Economics of Knowledge, National Research University Higher School of Economics, 20, Myasnitskaya Street, Moscow 101000, Russia;

E-mail: yturovecz@hse.ru

ORCID: 0000-0002-6336-1255

Konstantin O. Vishnevskiy

Cand. Sci. (Econ.), Associate Professor;

Director, Digital Economics Center; Senior Research Fellow, Research Laboratory for Science and Technology Studies, Institute for Statistical Studies and Economics of Knowledge, National Research University Higher School of Economics, 20, Myasnitskaya Street, Moscow 101000, Russia;

E-mail: kvishnevsky@hse.ru

ORCID: 0000-0003-3621-0504