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

Vol. 18 No. 4 – 2024

BUSINESS INFORMATICS

HSE Scientific Journal

CONTENTS

Publisher:
HSE University

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:
26-28, Shabolovka Street,
Moscow 119049, Russia

Tel./fax: +7 (495) 772-9590 *28509
http://bijournal.hse.ru
E-mail: bijournal@hse.ru

Circulation:
English version – 100 copies,
Russian version – 100 copies,
online versions in English and Russian – open access

Printed in HSE Printing House
44, build.2, Izmaylovskoye Shosse,
Moscow, Russia

ABOUT THE JOURNAL

Business Informatics is a peer reviewed interdisciplinary academic journal published since 2007 by HSE University, Moscow, Russian Federation. The journal is administered by HSE Graduate School of Business. The journal is issued quarterly, in English and Russian.

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 following areas: modeling of social and economic systems, digital transformation of business, innovation management, information systems and technologies in business, data analysis and business intelligence systems, mathematical methods and algorithms of business informatics, business processes modeling and analysis, decision support in management.

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 Scopus, Web of Science Emerging Sources Citation Index (WoS ESCI), Russian Science Citation Index on the Web of Science platform (RSCI), EBSCO.

The journal is distributed both in printed and electronic forms.

EDITORIAL BOARD

ABOUT HSE UNIVERSITY

Consistently ranked as one of Russia's top universities, HSE University 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 HSE GRADUATE SCHOOL OF BUSINESS

HSE Graduate School of Business was created on September 1, 2020. The School will become a priority partner for leading Russian companies in the development of their personnel and management technologies.

The world-leading model of a ‘university business school’ has been chosen for the Graduate School of Business. This foresees an integrated portfolio of programmes, ranging from Bachelor’s to EMBA programmes, communities of experts and a vast network of research centres and laboratories for advanced management studies. Furthermore, HSE University’s integrative approach will allow the Graduate School of Business to develop as an interdisciplinary institution. The advancement of the Graduate School of Business through synergies with other faculties and institutes will serve as a key source of its competitive advantage. Moreover, the evolution and development of the Business School’s faculty involves the active engagement of three professional tracks at our University: research, practice-oriented and methodological.

What sets the Graduate School of Business apart is its focus on educating and developing globally competitive and socially responsible business leaders for Russia’s emerging digital economy.

The School’s educational model will focus on a project approach and other dynamic methods for skills training, integration of online and other digital technologies, as well as systematic internationalization of educational processes.

At its start, the Graduate School of Business will offer 22 Bachelor programmes (three of which will be fully taught in English) and over 200 retraining and continuing professional development programmes, serving over 9,000 students. In future, the integrated portfolio of academic and professional programmes will continue to expand with a particular emphasis on graduate programmes, which is in line with the principles guiding top business schools around the world. In addition, the School’s top quality and all-encompassing Bachelor degrees will continue to make valuable contributions to the achievement of the Business School’s goals and the development t of its business model.

The School’s plans include the establishment of a National Resource Center, which will offer case studies based on the experience of Russian companies. In addition, the Business School will assist in the provision of up-to-date management training at other Russian universities. Furthermore, the Graduate School of Business will become one of the leaders in promoting Russian education.

The Graduate School of Business’s unique ecosystem will be created through partnerships with leading global business schools, as well as in-depth cooperation with firms and companies during the entire life cycle of the school’s programmes. The success criteria for the Business School include professional recognition thanks to the stellar careers of its graduates, its international programmes and institutional accreditations, as well as its presence on global business school rankings.

DOI: 10.17323/2587-814X.2024.4.7.24

Neural network technologies in supply chain management: Consumer selection technique*

Valeriya A. Nazarkina
E-mail: valeria71@bk.ru

Vladislav Yu. Shchekoldin
E-mail: schekoldin@corp.nstu.ru

Novosibirsk State Technical University (NSTU), Novosibirsk, Russia

Abstract

The supply chain management's effectiveness depends, among other things, on the selection and coordinated interaction with product consumers. This article is devoted to the development of a method for selecting a consumer in the regional wholesale and retail fuel market. The methodological basis of the study is the theory of statistical analysis and neural networks. The main tool for developing the methodology was neural network technologies, with the help of which it is most likely possible to correctly estimate the boundaries for indicators' values that characterize consumers and reflect their history of purchasing behavior, to select potential clients and the possibility of further cooperation with existing ones. The information base for the work is the data on consumers of a given company's products, data from the 2GIS electronic directory, as well as the results of the primary statistical analysis and forecasts made based on neural networks of various topologies. The author presents his methodology for selecting a consumer. It has the potential for development and implementation for solving a number of other management problems. As part of the testing, the best configuration (topology) of the neural network was determined, and standard values of entry barriers when consumer choice accomplished were assessed. The methodology we developed

* The article is published with the support of the HSE University Partnership Programme

was tested using the example of a company operating in the wholesale and retail fuel market in Novosibirsk and the Novosibirsk region. When verifying the neural network model, the quality of client classification was compared based on logistic regression, decision tree and random forest models and we found that the neural network approach provides the best results for assessing the degree of client suitability. As a result of testing the methodology, recommendations for improving neural network models were developed, including expanding the set of factors that determine the characteristics of consumers, as well as optimizing the internal structure of neural networks.

Keywords: neural networks, supply chain management, distribution logistics, consumer selection, neural network training

Citation: Nazarkina V.A., Shchekoldin V.Yu. (2024) Neural network technologies in supply chain management: Consumer selection technique. *Business Informatics*, vol. 18, no. 4, pp. 7–24. DOI: 10.17323/2587-814X.2024.4.7.24

Introduction

The unpredictability of the influence of external environmental factors poses challenges to the management of organizations associated with the risk of choosing an incorrect management decision in various areas of the company's activities. Despite the desire to attract as many customers as possible and gain their loyalty, the organization faces problems of interaction complexity, transaction inefficiency and default on contractual obligations by counterparties. A sufficient number of methods for selecting a reliable resource supplier from theoretical and practical points of view have been developed, but scientists and business representatives do not deal with the issues of choosing a consumer in detail. The decision on the choice of consumers and further cooperation with them is often determined by trial and error, which ended up affecting the effectiveness of management decisions in the field of supply chain management in general, and interactions with consumers in particular.

Nowadays one of the topical methods for solving a wide variety of economic and management problems is the implementation of neural network technologies. From the point of view of logistics, significant objects for the application of neural networks are processes of supply chain management, because these methods can be used to evaluate, model and forecast the development with predetermined accuracy.

The purpose of this study is to develop a method for selecting consumers based on the use of neural networks and to test it using the example of a company operating in the regional wholesale and retail fuel market.

1. Theoretical basis of the research

Neural networks are a certain type of artificial intelligence models based on the structure, dynamics and functions of the human brain. Neural networks in the most general form consist of many interconnected nodes (neurons) and are used to process, model and analyze complex heterogeneous processes. Neural networks have attracted wide attention and are being intensively used in various fields due to their ability not only to analyze data, but also to independently build logical conclusions, form value judgments and develop predictive solutions of a wide variety of types.

Historically, the first works related to the application of theories describing the functioning of the brain appeared in the 1940–1950's, when it became necessary to build concepts of artificial neurons to determine their potential in the development of complex intelligent systems. One of the first successful representations of the nervous system was the perceptron model proposed by Rosenblatt in 1957 [1], which turned out to be capable of solving some recognition problems. Despite the initial successes of the perceptron model, it had certain shortcomings that did not allow it to be used to solve certain types of problems. In the 1970–80's, the backpropagation algorithm was developed [2], which made it possible to train more complex networks. Among the extensive research devoted to the issues under consideration, it is worth noting the significant contribution to the development of the theory of neural networks and pattern recognition by the Soviet and Russian scientist Galushkin [3].

The 21st century has also contributed to the development and dissemination of neural network technologies, with a rapid growth in research related to so-called deep learning. Pioneering research in this area is associated with the name of Hinton, who made a significant contribution to the development of modern deep neural network training algorithms [4]. Deep learning involves training neural networks with several hidden layers, including those with a dynamically varying structure, the presence of external and internal topologies, etc.

It should be noted that in addition to neural network models, other approaches are also encountered in practice that allow solving various classification problems. Such methods include, for example, the construction of logistic regression models, decision trees, random forest models, etc. [5], and each of these approaches has its own advantages and disadvantages. Let us consider some of them.

The advantages of logistic regression models include [6]: high result interpretability, high efficiency of statistical procedures used to estimate the model parameters; and flexibility in solving binary and multifactor classification problems. Logistic regression models also have a number of disadvantages: dependence of the results obtained on the structure of input factors; high sensitivity to outliers in the initial data; and frequent manifestation of the multicollinearity.

The application of decision tree models ensures the construction of an easy and convenient interpretation of the solutions obtained; universality of the computational scheme for any type of data; robustness of the classification results, i.e., independence on outliers in the initial data. The disadvantages of this approach are as follows: the resulting decision trees often have a complex and confusing structure (so-called tendency to overfitting); classification instability to changes in the initial data; and the appearance of locally optimal solutions, which are determined by the heterogeneity of the initial data or the processes being studied, and so on [7].

When using random forest models, the researcher obtains the following advantages: the ability to scale and conveniently parallelize the basic algorithm depending on the properties of the problem being solved; the ability to rank independent factors included in the model; and high efficiency of classifications for large-scale problems. However, such models have these disadvantages: the complex and ambiguous structure of the model due to a wide variety of averaging options; instability to fluctuations in the initial data; a large number of empirically determined algorithm parameters; and high requirements for the characteristics of computing equipment (memory, speed, etc.) [8, 9].

The development of neural network technologies has had a great impact on the field of logistics. Many formulations of logistics problems and methods for solving them have undergone significant changes in light of neural network theory. We will mention some of them.

In the research [10], the authors apply the idea of constructing a multilayer neural network to analyze order data, including the number of visits to the company's website, the time of visit taking into account working, weekdays, and holidays. The article [11] discusses the implementation of the Hopfield neural network to solve the problem of constructing a dynamically optimal route in a telecommunications network and proposes a heuristic rule for stopping the neural network training process, all of which allows for effective limitation of the training time.

Neural networks can be used to optimize warehouse operations. In the article [12] the authors prove that networks with special neural activation functions such as *Traingdx* are most effective in warehouse management when using three-layer neural networks with a 6-8-1 topology.

Another area of application of neural networks is the analysis and forecasting of risks in logistics. The article [13] considers the problem of safe flight around obstacles by manned and unmanned aerial vehicles, where the author proposes to use a multi-layer network of sequential error propagation with three layers.

It should be noted that among the large number of works devoted to solving logistics problems in which neural network technologies would be used, there is practically no research in which consumers of goods and services were studied. At the same time, studying the history of consumer behavior and constructing consumer classifications is given a lot of attention when solving problems outside the field of logistics [14–17], since understanding the structure of the customer base and the dynamics of its changes gives the company additional tools for improving the efficiency of interactions with consumers.

In today's rapidly changing and often complicated to predict economic environment, companies have to deal with situations where it is important to determine priorities when cooperating with certain customers. Thus, at the stage of concluding contracts, it is sometimes necessary to identify promptly whether a particular client is financially independent and solvent, whether there will be difficulties with fulfilling orders due to geographical remoteness, whether it is possible to use jointly warehouse capacities, etc. Undoubtedly, by analyzing the history of consumer behavior, it is possible to determine which of the already concluded and fulfilled contracts appeared to be profitable for the company, which did not lead to the achievement of the set goals, and which turned out to be a complete failure in terms of revenue, resource costs, and employee time, causing damage to the company's image, etc. This information can be used to determine in advance the degree of profitability of cooperation with the next potential client based on its contemporary characteristics. The study presented here is aimed at solving this problem.

2. Research design

The necessity in solving optimization problems of the logistics processes stipulated the possibility of using neural network technology in terms of determining the boundaries of certain factors' values for selecting potential consumers, and its significance for the company in terms of profitability and fulfillment of contractual obligations. This study was structured based on empirical data collected from secondary and primary sources. To solve the research problems, we used the method of primary statistical analysis, correlation and regression analysis, the method of back propagation of error and data mining methods [2, 6, 18, 19].

The content and expected results of the application of the consumer selection methodology are presented in *Table 1*.

At *the first stage* of the methodology, it is necessary to describe the essence of the problems in terms of interaction with the company's consumers. We analyze the stages of the logistics cycle typical for working with consumers. Problems arising in the process of product distribution may concern issues of information interaction between the company and consumers, quality control and transportation, documentation, etc. As a result of identifying the problem, the organization's management makes a decision to find the best methods for selecting consumers.

At *the second stage*, it is required to select the indicators that the organization is guided by when making a decision on interaction with the consumer. The factors typical for choosing a consumer include: distance to the delivery location; number of types of products subject to simultaneous sale; equipment capacity; rating; sales volume for a certain period, etc.

Table 1.

The main stages of the consumer selection methodology using a neural network

	Stage	Content	Results
1	Problem definition	Analysis of the logistics cycle stages, problem's identification	Determining the need to find relevant methods for obtaining information about consumers
2	Identifying factors influencing the consumer selection process	Composition of a set of factors subject to quantitative assessment by company specialists, experts, and research consumers	List of factors characterizing consumer properties that are used to build a neural network
3	Creating a database of current consumers of the organization	Determining the values of selected factors for consumer evaluation	A database of up-to-date data on consumers and their transactions for training the neural network
4	Primary statistical analysis	Calculation and interpretation of statistical characteristics of factors	Preliminary conclusions about the properties of observed objects (consumers)
5	Defining the roles of factors	Building a simple neural network	Ranked list of factors for making a consumer choice decision
6	Building a neural network of complex structure	Building neural networks of different structures and comparing them with results obtained on the basis of other classification models	Choosing the best configuration (topology) of a neural network
7	Using the "best" neural network to identify prospective clients	Evaluating the range of factor values that determine the consumer's status	Determining standard values of entry barriers for consumer selection

At *the third stage*, it is necessary to develop and fill a database for training the neural network. For all potential consumers of the company, the values of the factors selected at the second stage are determined. Some factors are assessed by direct measurements of the relevant indicators, while others require obtaining and using secondary information. In addition, each of the potential consumers must be assessed by a logistics specialist for the priority of choosing him as a real client.

The fourth stage involves calculating the main statistical characteristics determined by the values of the factors of the consumer database constructed in the third stage. To ensure the correctness of the statistical analysis, in particular, to determine homogeneous groups within which one can speak of a certain identity of the analyzed objects (consumers), the data must be distributed (grouped) by homogeneity classes, the number of which k is determined, for example, by the Sturges formula [20]:

$$k = [1 + 3{,}32\lg(N)], \tag{1}$$

where

N – total number of data (the sample size, or the database volume);

lg(.) – decimal logarithm;

[.] – operation of calculating the integer part of a number.

As a result, a statistical analysis of the values of the measures of central tendency, variation and shape is carried out, and preliminary conclusions are made about the properties of consumers and the history of their purchasing behavior.

At *the fifth stage*, the factors are ranked according to their degree of influence on the decision to interact with potential clients. For this purpose, a simple neural network is built, consisting of one neuron and input variables corresponding to the factors selected at the third stage. The degree and nature of the influence of the factors will be determined by the values of the network weights. As a rule, the efficiency of such a network is quite low, which does not allow it to be used to predict the priority of consumers. However, it allows us to rank the input factors, which makes it possible to build a high-quality interpretation of the decision-making process for choosing customers.

At *the sixth stage*, neural networks of various topologies are developed and trained to ensure the best degree of predictability of the customer's "utility." In this case, it is necessary to consider several network options that differ from each other in the number of hidden layers, their interrelations and the number of neurons in each layer [21]. The result of this stage will be the selection of a neural network configuration that ensures the lowest level of error in predicting the reliability of consumers.

To ensure the adequacy of the obtained results, it is necessary to compare the quality of the customers classification obtained based on neural networks and classifications constructed by other methods of data mining. The simplest way to compare different classifications is to use the contingency table method [22]. We will assume that the best classification will be the one that provides the least number of forecasting errors.

At *the seventh stage*, the best of the neural networks built at the sixth stage is used to determine the ranges of factor values at which the status of the current client is maintained. This will also allow us to determine the standard (base) values of the input factors to simplify the procedure for selecting a new customer.

3. Practical aspects of the research: testing the methodology

The methodology was tested based on information provided by a company specializing in the sale of liquefied gas to organizations and individuals in Novosibirsk and the Novosibirsk Region. The company also sells related products and services, carries out design work, provides maintenance services and organizes technical maintenance of natural gas pipelines.

At the *first stage*, the operations of the logistics cycle were considered. The consumers were defined as LPG filling stations (hereinafter referred to as LPGFS) located in the city of Novosibirsk and the Novosibirsk region. The study of the logistics process in the company revealed the need for regular verification of information on the characteristics of consumers, since there are no clear boundaries of the values of factors that make it possible to identify consumers with properties that allow them to be considered suitable for cooperation. In this regard, errors arise, the causes of which are, for example, the influence of external factors, limited resources, errors of specialists, etc., which can lead to incorrect identification of customers and, as a consequence, to apply incorrect decisions on working with them.

The organization's management decided to look for ways to optimize interaction with consumers in terms of determining standard values of factors for both current consumers and for selecting and interacting with potential consumers.

At the *second stage*, a pool of factors was compiled that determine the characteristics of consumers taken into account when making decisions. This pool included the following indicators:

- distance from the company to the consumer (km);
- range of products (number of types of liquefied gas purchased by the consumer);
- production capacity (number of dispensers at LPGFS);
- gas station rating;
- total volume (capacity) of the main and additional tanks at LPGFS;
- average volume of fuel sales per day (thousand liters).

At the *third stage*, a database structure was developed that included information on the company's transactions. The database contained the company's operating results for the period September–November 2023.

A company specialist who held the position of head of the logistics department assessed the profitability of completed transactions, as a result of which all database records were marked with the values of a binary variable that determined the "reliability" of clients: the designation "1" was used for "suitable" clients with whom it was profitable for the company to continue to cooperate, and "0" for "unsuitable" ones.

At the *fourth stage*, a primary statistical analysis was conducted for each of the factors that characterize the activity of LPGFS. The values of the measures of central tendency (mean, median), variation measures (standard deviation, lower and upper limits), measures of shape (skewness and kurtosis), as well as extreme values (minimum and maximum) were calculated. The results are summarized in *Table 2.*

Table 2.

Values of statistical characteristics of factors

Statistical characteristics	Factors and their designations					
	Distance, km	Assortment width, units	Production capacity, units	Rating, conv. units	Tank capacity, thousand liters	Sales volume, thousand liters
	x_1	x_2	x_3	x_4	x_5	x_6
Mean	118.469	2.563	4.547	2.469	19.983	5.610
Median	51.500	2.000	4.000	2.300	17.000	4.375
Standard deviation	136.189	1.344	2.949	0.920	14.351	3.749
Lower limit	0.000	1.219	1.598	1.549	5.632	1.861
Upper limit	254.658	3.907	7.496	3.389	34.334	9.359
Skewness	1.687	0.531	2.09	0.633	1.986	1.428
Kurtosis	2.535	1.035	8.054	0.141	5.649	2.335
Minimum	11	1	1	1	3.3	0.8
Maximum	628	5	18	5	85	18

Note that the lower and upper boundaries in *Table 2* correspond to the so-called "one sigma" interval and are defined as the difference and sum of the mean value and standard deviation, respectively. The interval between these boundaries contains the most probable values of the random variable being analyzed, which in the case of a normal distribution include about 70% of the sample observations [7, 23].

To ensure the correctness of the analysis, the data were distributed into homogeneity classes, the number of which was determined by the Sturges formula (1) and was found to be seven. The histogram, which is a graphical interpretation of the frequency distribution of LPGFS by distance to customers, is shown in *Fig. 1*.

Based on *Fig. 1*, we can assume about the exponential distribution of the distance to customers, which is explained by a significant increase in their number when approaching the city. 41 gas stations out of 64 studied (64%) are located within 100 km from the company. The average distance to customers is 118 km. Half of all values fit into the interval from 11 km to 51.5 km, which indicates a strong predominance of low values of this factor, which is also confirmed by the positive value of the sample's skewness. The most probable values of this indicator are in the interval (0–254), 79% of all consumers fall into it. The kurtosis coefficient is greater than zero, which indicates a good predictability of the indicator relative to the normal distribution, since a large number of gas stations are grouped in one class.

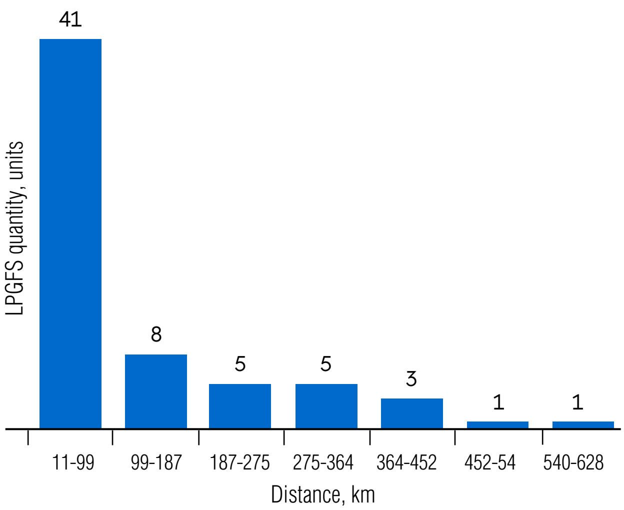

Fig. 1. Distribution of consumers by the factor "distance from the company to the client".

The frequency distributions of the number of LPGFS for the remaining factors were interpreted in a similar manner.

When analyzing the number of types of fuel sold, it turned out that the most common LPGFS are those with two or four types of liquefied gas (40%). For the "production capacity" indicator, it was noted that the total number of gas stations has no more than six dispensers. Analyzing the data on consumer ratings, one can notice some unpleasant statistics: most often, gas stations have a low rating (no higher than 3 points), while the average rating is around 2.5. The distribution of tank volume values is similar to exponential. The overwhelming majority of consumers (more than 80%) have installed tanks with a total volume of no more than 34 thousand liters. When analyzing the volume of daily sales, it turned out that the largest number of LPGFS is characterized by low sales relative to the rest – less than 6 thousand liters of gas per day. Only six filling stations out of 64 have average daily sales of more than 10 thousand liters (less than 10% of their total number).

After analyzing the initial data for the selected indicators, it is important to understand how significantly these factors affect the final result (the client's "suitability" for cooperation). For this purpose, we will use the idea of constructing the Rosenblatt perceptron [1, 2].

At the fifth stage, a neural network consisting of one neuron was constructed. As the activation function of the neuron (as for all other variants of neural networks considered in the work), the logistic function was taken in the form of

$$\sigma(x) = \frac{1}{1+e^{-x}}. \tag{2}$$

The choice of the logistic function is due to its continuity, which ensures smoothness in the transition region. The *ESS* value, the residual sum of squares [7] between the specialist's assessment (Y) and the assessment issued by the neural network ($\hat{Y}$), was used as a functional determining the correctness of the neural network operation:

$$ESS = \sum_{k=1}^{N}(Y_k - \hat{Y}_k)^2 \to \min. \quad (3)$$

In (3) the summation is carried out over all database records; N is the database size (number of records). For input factors, the notations x_1–x_6 are used in the order of listing (*Table 2*).

The calculation of the estimates determined by the neural network was carried out based on the value of the activation function (2) from the linear combinations of the values of the input variables of the model x_1–x_6, as well as any variables of the internal layers of the neural network, determined by the topology of the neural network. For example, for a neural network consisting of one neuron and having six input variables, the estimated value of the probability that the client will be recognized as "suitable" for cooperation will be determined as

$$\hat{Y} = \sigma(w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_5x_5 + w_6x_6) = \sigma(\Sigma), \quad (4)$$

where w_i, $i = 1, ..., 6$ are the weight coefficients of each of the input factors, determined by solving problem (3); Σ is the value of the linear combination of input factors, called the adder.

Minimization of the functional (3) is carried out by changing the unknown weight coefficients w_i in the adder. Since the problem (3) has no solution in analytical form, the weight coefficient estimates were found numerically [24]. *Figure 2* shows the diagram of the implemented model from one neuron.

The model of the simplest network includes the values of the input factors (circles on the left in *Fig. 2*), the adder Σ from formula (4), and the value of the activation function (2). The output of the neural network is determined by the proportion of correct forecasts – the ratio of the number of correct responses of the neural network to the total volume of the database N (in this case, N = 64). The degree of darkening of the arrows reflects the strength of the influence of the correspond-

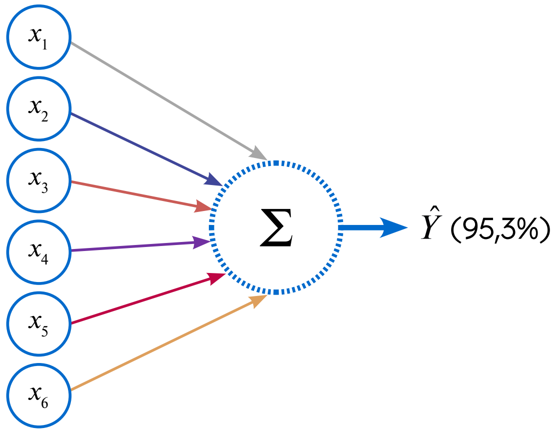

Fig. 2. Representation of a single neuron model.

ing input factor on the result of assessing the "suitability" of the consumer: the darker the corresponding arrow, the stronger the influence of the factor.

By minimizing the sum of squared deviations of the probabilities of making a decision on cooperation with a consumer, estimated by the company's expert, from its values predicted by the neural network, estimates of the values of weight coefficients were obtained. *Table 3* presents the values of the coefficient estimates and the results of ranking the factors analyzed by the absolute values of the weight coefficients.

The obtained values of weight coefficients can be interpreted. For example, the sign of the coefficient indicates the direction of the dependence. From *Table 3* it is evident that only the coefficient w_1 turned out to be less than zero, which means the presence of a negative dependence, i.e., with an increase in the distance to the consumer, the probability of making a decision on cooperation decreases on average. The remaining factors have a positive dependence on the probability of making a decision on cooperation.

The absolute values of the coefficients mean the "power of influence" of the corresponding factors on the decision to cooperate with a particular client. Let us analyze the results obtained.

Table 3.

Neuron weighting coefficients' estimators and factor ranking

Weighting coefficient	w_1	w_2	w_3	w_4	w_5	w_6
Coefficient estimator	–50.000	11.101	0.671	11.639	8.600	3.744
Degree of importance, %	58.305	12.944	0.783	13.573	10.029	4.366
Factor rank	1	3	6	2	4	5

1. The distance to LPGFS (factor x_1) has the strongest influence; its importance relative to all other factors is almost 60%. At the same time, the assessment of the corresponding weighting coefficient (w_1) rests on the limit of acceptable values (±50), which means its complete dominance in decision-making. It is quite logical that it makes no sense for the company to organize deliveries over long distances.

2. The gas station rating (factor x_4) is in second place, accounting for just over 13% of importance. Of course, the higher the reputation and consumer ratings of the company, the more reliable it is.

3. The variety of fuel types at LPGFS (factor x_2) is in third place, slightly inferior to the rating (significance less than 13%). The significance of this factor is due to the higher financial stability of the enterprise in case of need for a wide range of fuel for consumers.

4. Tank capacity (factor x_5) is in fourth place. The influence of this factor is confirmed by the "convenience" of the selected gas station for cooperation. To some extent, this indicator can be considered as the volume of a warehouse for a trading company: it must be sufficient (as well as the stocks in it) so that the company can carry out its activities without hourly delivery of products. The situation is similar for LPGFS. By optimization, it is possible to minimize the required capacity of tanks, but delivery of products using just-in-time systems [25] can be quite complicated due to the specifics of the goods being transported, so it will be important to have spare gas storage tanks.

5. Daily gas sales volume (factor x_6) is in the penultimate place (less than 5% importance), which is somewhat surprising. The influence of this factor should be obvious, because the higher the company's turnover, the higher its stability. However, it is important to understand that all companies (as well as LPGFS) are in different situations. Thus, for powerful stations with six or more dispensers located in the city center, a high, at first glance, turnover may actually be quite low compared to other consumers located nearby.

6. The production capacity of LPGFS (factor x_3) is the last (importance less than 1%). Like the previous indicator, it does not have clear limits, so it is not a strong factor. In general, a larger number of gas dispensers is an opportunity for the company to develop, which can play a role in the long term.

It should be noted that the estimates obtained using the neural network correctly reflect the thought process of a specialist who does this manually. Analyzing the results obtained, it should be noted that out of 64 records in the company's client database, the single neuron network made a mistake in only three, which is 4.7% of errors. To understand the causes of these errors, it is necessary to examine the results and identify the features of these gas stations. In all three cases, the probability assessment of the client's "reliability" issued by the neural network was more than 0.9, i.e. it was more than confident in their "suitability" for cooperation.

The first LPGFS has low values for almost all indicators, but it has a high rating, which is most likely the reason for such an assessment. Obviously, the ratings provided by special services are the most unreliable indicator, since in many cases they are either unrepresentative due to the small number of averaged assessments, or incorrect due to the use of certain schemes for "winding up" the necessary values.

The second gas station has good factor values, but the company's specialist noted that they do not cooperate with this station, since the other branch is closer to the client, and interaction is carried out through them.

The third gas station was rated positively, but the specialist rated the experience of working with this company negatively because the station is in the process of launching; the data on it is contradictory, and more time is needed to sign contracts. This situation is not an error, but in light of the increased recognition capabilities of the neural network, it is recommended to add a check for the operation time of LPGFS on the market.

At the *sixth stage*, to eliminate erroneous triggering of the neural network, it was decided to consider more complex options of network structure by adding internal layers with different numbers of neurons; the most suitable option turned out to be a topology with seven neurons 4-2-1 (*Fig. 3*).

The selected topology allowed us to successfully describe the interaction process of the factors considered, while no discrepancies were found between the expert's assessments and the results of the neural network. The darker arrows of the neuron connections in the network correspond to the third neuron of the first layer and the sixth neuron of the second (*Fig. 3*). The difference in the influence of the values of the neurons of the second layer on the network output is only 18.4%. It should also be noted that the values of the fifth neuron have a negative effect on the result, and the sixth – a positive one. Unfortunately, this fact cannot be used to interpret the results (as it was for the single neuron network), since the values of each of the neurons of the internal layers are added up under the influence of the previous layers in a nonlinear manner due to the selected activation function in the form of a logistic function (2).

In order to ensure the suitability of the results obtained, in addition to the neural network model, consumer classifications were constructed using logistic regression, decision trees and random forest. The freely distributed Orange Data Mining software package was used to develop and identify the corresponding models [19].

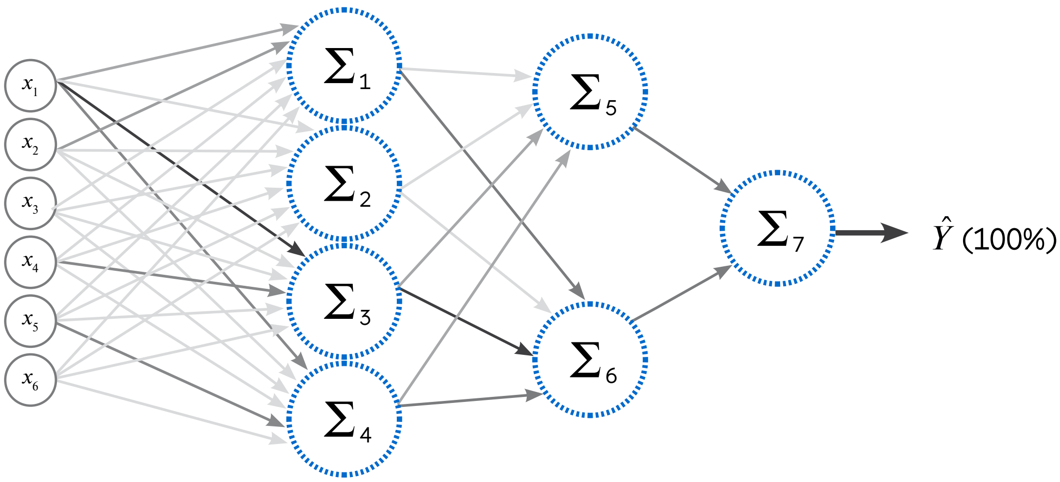

Fig. 3. Representation of the neural network model according to the 4-2-1 topology.

To construct the best logistic regression model, the maximum likelihood method was chosen as the most frequently used in such situations. Within the framework of this study, a standard version of regression model linear in both parameters and input factors was implemented.

As for the construction of the model using the decision tree and random forest methods, various tree options were considered. The quality of the classifications built on their basis was determined using a standard indicator – the F_1-measure [4, 19, 22]. In this case, to ensure the construction of effective classification trees, trees with different parameters were considered. The results are presented in *Table 4* and *Table 5*. In addition, *Table 6* shows the results of the comparison of classifications based on the best results of the models obtained.

Table 4 presents the results of assessing the quality of models constructed using the decision tree method. The following notations are used:

a is the minimum number of elements in one leaf of the tree;

b is the number of elements in a leaf at which no further splitting is performed;

F_1 is the value of the quality measure of classifications.

In *Tables 4* and *5*, the classification variants that are the best in terms of the maximum value of the quality measure are highlighted in bold. *Table 5* presents the results of assessing the quality of models constructed using the random forest method. The number of trees of the corresponding random forest is indicated in brackets after the F_1-measure.

Let us notice that for some combinations of random forest' parameters, the resulting classifications are the same, which is reflected in the F_1-measure values. In addition, note that further increase in the number of trees in the forest does not lead to changes in the values of the classification quality measure.

The classification results by the methods mentioned are presented in *Table 6*, where the designation "1" was used for "suitable" clients with whom it is profitable for the company to cooperate, and "0" for "unsuitable". In this case, the "observed values" correspond to the specialist's assessments, and the "predicted" ones correspond to the classification results by the methods under consideration.

Table 6 shows that the use of logistic regression resulted in six errors (9.4% of errors), three of which occurred for "suitable" clients, and the other three for "unsuitable" clients. When using the decision tree method, the number of errors for "suitable" clients was two, and for "unsuitable" clients – five (a total of 10.9% of errors). The random forest method resulted in four errors for "unsuitable" clients (a total of 7.8% of errors). In general, it can be noted that methods alternative to the neural network did not provide a complete match between the expert's assessments and the classification results. This allows us to conclude that the neural network is the most suitable for solving the problem under consideration for assessing the quality of clients.

It is important to understand that to check the correctness of the neural network, it is not enough to use only the data set that is currently available. It is important to ensure that its predictive properties are preserved for new portions of data, which from a statistical point of view means ensuring the adequacy of the model [6].

Table 4.

Parameters of the decision tree model and the quality of the resulting classifications

a	2			4		6	8	10
b	2, 4, 6	8	10	4, 6, 8	10	6, 8, 10	8, 10	10
F_1	0.876	**0.891**	0.857	0.889	0.857	0.842	0.842	0.844

Table 5.

Parameters of random forest models and the quality of the resulting classifications

a	2	2	2	2	2	4	4	4	4	6	6	6	8	8	10
b	2	4	6	8	10	4	6	8	10	6	8	10	8	10	10
$F_1(10)$	0.840	0.840	0.855	0.855	0.870	0.874	**0.889**	0.859	0.874	**0.889**	0.859	0.874	0.859	0.874	0.874
$F_1(50)$	**0.904**	**0.904**	**0.904**	0.889	0.889	0.874	**0.904**	0.874	0.874	0.889	0.874	0.874	0.874	0.874	0.874
$F_1(100)$	0.887	0.887	0.887	0.887	0.887	0.874	**0.889**	**0.889**	0.874	0.874	**0.889**	0.874	**0.889**	0.874	0.871

Therefore, after completing the training of the neural network, ten new LPGFS were taken to check its performance. Contracts with that consumers were concluded during the study and, naturally, were not included in the original database. The results of the neural network (in the form of estimates of the probability of the client's "suitability" and the final assessment) were compared with the specialist's assessments (*Table 7*).

The neural network correctly assessed nine positions out of ten. The error was made with the seventh station. The error in the probability assessment is less than one tenth (0.411 against 0.500, which would be enough to recognize the LPGFS as suitable for cooperation). This station has a high rating (4), sells five types of fuel and has an average capacity (four dispensers). However, it is 149 km from the company and has installed tanks of relatively low capacity (about 16 thousand liters). The volume of gas sales is less than 6 thousand liters, i.e. this LPGFS is quite average, but ambiguous in characteristics.

The neural network has assessed that cooperation with such a consumer will not be profitable, but the company's specialist has decided that it is suitable. This contradiction may be a sign that this LPGFS may need to conduct a more thorough study of its capabilities to ensure successful cooperation with the company. When a certain number of such "controversial" situations accumulate, a decision may be made to retrain the neural network on a larger database, or, if the results continue to be inconsistent with reality, a more radical action may be required – changing the topology of the neural network and its complete retraining.

Table 6.

Classification methods' quality estimation

		Predicted values by logistic regression		Predicted values by decision tree		Predicted values by random forest method		
		0	1	0	1	0	1	Σ
Observed values	0	20	3	18	5	19	4	**23**
	1	3	38	2	39	0	41	**41**
	Σ	**23**	**41**	**20**	**44**	**19**	**45**	**64**

Table 7.

Comparison of neural network and expert assessments for new customers

No. of LPGFS	1	2	3	4	5	6	7	8	9	10
Expert assessments	1	0	1	1	1	1	1	1	1	1
Probability assessment	0.954	0.121	0.851	0.884	0.732	0.999	0.411	0.970	0.804	0.925
Network assessment	1	0	1	1	1	1	0	1	1	1

At the *seventh stage*, for independent analysis of the consumer for compliance with the company's requirements, one can use a list of threshold values for each factor, which will help you quickly assess how clients with factor values close to average are suitable for further cooperation.

To build such a list, the neural network we developed was applied. The ranges of factor values were determined one by one, at which the client's "suitability" is maintained, while the values of all other characteristics were fixed at average levels (*Table 2*). Based on the weighting coefficients, the neural network independently calculates the values at which the probability assessment (assessment of the "client's reliability") will be in the range from 0.5 to 1, which corresponds to the decision that the client is suitable for cooperation. Thus, we will compile a table of threshold values (*Table 8*), which can be used as a hint for a specialist when assessing the degree of customer suitability.

Comparing the data in *Table 2* and *Table 8*, we can conclude that the modeling results obtained during the study allow us to speak about the correctness of using neural networks for the tasks of selecting consumers in the wholesale and retail trade of oil and gas products.

4. Discussion

In the process of testing the methodology we developed, it became necessary to construct an additional interpretation of the results, not only in terms of obtaining analytical and statistical conclusions, but also for developing visual and specific recommendations for interaction with the company's customers. For example, it is of interest to construct a geographic interpretation of the results obtained by depicting the approximate coverage area of the company on a map and analyzing it. The radius of this area according to *Table 8* is 156 km.

The distance threshold can be called the most important characteristic, since it is based on the factor that has the greatest weight in the neural network. It determines how far away potential consumers are and, particularly, how convenient it is for them to access LPGFS. This value plays a key role for the company, since it directly affects the cost of transporting gas to consumers, which, in turn, can significantly affect their own costs.

Table 8.

List of threshold values for customer characteristics

Factor	Admissible values	
	min	max
Distance, km	–	156
Assortment width, units	2	–
Production capacity, units	1	–
LPGFS rating, conv. units	1.1	–
Tank capacity, thousand liters	20	–
Sales volume, thousand liters	5.6	–

To clarify the factor of “the distance to the client”, one can add a new characteristic that takes into account the area of the client’s location relative to the company’s location. However, it is important to understand that the weight coefficient of the new factor will “take away” some of the strength of the old one since they based on similar customer characteristics.

Another solution to this problem is to replace the existing factor with four new ones, each of which is a distance in a certain direction (for instance, north-south-west-east). Then the distance to each LPGFS will be taken into account by one or two separate factors (for example, city gas stations will be predominantly located in the west, north or south direction due to the specific geography of Novosibirsk). This approach will allow us to determine more accurate boundary values of distances and will help in a more accurate assessment of the coverage area, which will presumably be stretched in different directions. For example, the maximum possible distance to the east, towards Novosibirsk, will be noticeably greater than to the west, since most of the city, and, accordingly, the company’s potential consumers are located on the eastern bank of the Ob River.

It is important to understand that determining the transition distance value is a complex task and may depend on many factors (transport infrastructure, population density, terrain topography, etc.). Therefore, when deciding on cooperation with a gas station, it is necessary to take these factors into account and conduct a detailed analysis of the market and the relevant infrastructure.

The threshold number of fuel types is important for LPGFS because it affects customer service and the efficiency of fuel management. This indicator may depend on various factors, such as the location of the gas station and the needs of local residents. Typically, remote stations offer no more than two or three types of fuel for cars, which corresponds to the threshold value found at two types of fuel. If the station carries out gas refueling, it is also sold in cylinders (pressure tanks), which can be convenient for car owners who use gas both as a fuel and for household needs.

The lower threshold value of the capacity of the LPGFS was found to be equal to one, which is most likely due to its low degree of importance (0.783%, *Table 3*). However, if the station has too few dispensers, it may be insufficient to meet the needs of all consumers. On the other hand, too high capacity may lead to excessive costs for the construction and maintenance of the gas station. Therefore, before selecting the capacity of the LPGFS, it is recommended to analyze the needs of consumers, the availability of gas pipelines and other factors in order to select the optimal capacity that will meet the requirements of all stakeholders and ensure maximum efficiency and cost-effectiveness of the LPGFS.

The minimum possible rating of the gas station was 1.1 points on a five-point scale, which is very low and is a consequence of the low level of customer satisfaction expressed in reviews of the LPGFS. To obtain a more objective picture of the quality of service and the level of customer satisfaction, it is necessary to study reviews from 2GIS and other sources, using the average weighted assessment, taking into account the reputation and reliability of the sources, as well as their quantity.

The threshold value of the volume of tanks installed at the LPGFS is 20 thousand liters, which is quite consistent with the volumes of average gas filling stations (*Table 2*, the average is 19.983 thousand liters). However, if the costs significantly exceed expectations, then the installed tanks may be sufficient for no more than a day, which will entail frequent refueling several times a day, and, therefore, may require the installation of additional tanks to increase the total capacity.

Minimum daily sales at LPGFS is an important factor that determines the minimum volume of fuel that must be sold (and, respectively, be available at the time of sale) for the station to remain profitable. For an average LPGFS the minimum daily sales are 5.600 liters. However, this value is quite low for city stations that service a large number of cars. At the same time, for LPGFS on the outskirts of the city or in sparsely populated areas, where the volume of fuel sales is lower, this value will be more than sufficient.

In further modifications of the neural network, it could be reasonable to replace this factor with a more objective one, such as the ratio of sales to the population of the area where the gas station is located, to take into account the potential demand for fuel in a particular area. In addition, it is possible to calculate separate factors for the city and region, based on the characteristics of the regional fuel market.

Of additional interest is the study using a neural network of how the threshold values can change when not only one, but also two (three, etc.) other factors change simultaneously. An example of such a calculation, carried out using regression analysis methods to construct the corresponding dependencies [6], is shown in *Fig. 4*.

Analyzing *Fig. 4*, we can say that the nature of the change in the maximum possible distance varies significantly: for consumers selling four types of fuel ($x_2 = 4$), with an increase in rating, the distance decreases, while with three or four types of fuel, it first increases and then decreases. Moreover, we can determine the maximum permissible distance, which for consumers with $x_2 = 2$ is 193 km, and for customers with $x_2 = 3$ is 138 km.

The threshold values so obtained can be applied for fast assessment of consumers when concluding a contract. During further development of the neural network, the factors can change both quantitatively and qualitatively; therefore timely updating of the pool of input factors and the contents of the database will allow us to obtain more correct assessments of consumers.

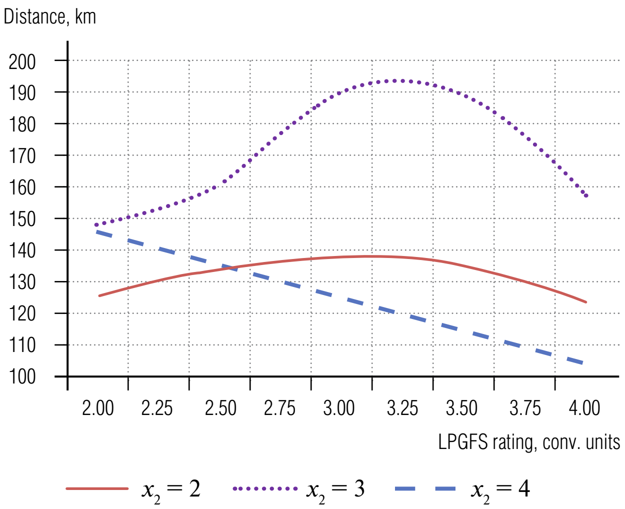

Fig. 4. Comparison of upper distance thresholds depending on rating and number of fuel types.

Conclusion

The result of using the author's methodology is the construction of a decision-making model for choosing consumers with whom the company plans further mutually beneficial cooperation. For this purpose, a pool of factors characterizing consumers and their purchasing behavior was formed. Analyzing the database of the company's clients and expert assessments determining their reliability, neural networks were developed that allow us to assess the prospects of cooperation with clients. With their help, the problem of ranking factors characterizing consumers was solved in relation to the degree and nature of their influence on the decision-making process on the reliability of a particular consumer.

Based on the results of neural network training, a network with the "best" topology was selected, which ensured correct forecasting for all database records. Comparison of the results of this neural network's outputs with classifications built on the basis of other data mining methods allowed us to conclude that the neural network model is the best fit for solving the problem under consideration.

The neural network we constructed was used to determine the threshold values of the input factors of the model. This allowed us to develop recommendations for the company's employees on selecting proposals for cooperation with consumers.

Further improvement of the proposed methodology may consist in expanding the pool of input factors by involving new consumer characteristics, including those proposed in this paper, as well as by splitting the existing factors into several components, each of which characterizes a certain specificity of consumers. In addition, for neural networks with a large number of input variables, it makes sense to consider more complex topologies that can include not only additional internal layers, but also feedback.

Due to its universality, the methodology we developed can be recommended for solving various classification problems not only in the field of logistics, but also for a wide range of economic and managerial problems. ■

References

1. Rosenblatt F. (1965) *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms*. Moscow: Mir (in Russian).

2. Rumelhart D.E., Hinton G.E., Williams R.J. (1986) Learning internal representations by error propagation. *Parallel Distributed Processing, Cambridge, MA, MIT Press*, vol. 1, pp. 318–362. https://doi.org/10.1016/B978-1-4832-1446-7.50035-2

3. Galushkin A.I. (2010) *Neural networks: basic theory.* Moscow: Hotline –Telecom (in Russian).

4. Hinton G., Le Cun Y., Bengio Y. (2015) Deep learning. *Nature,* vol. 521, pp. 436–444. https://doi.org/10.1038/nature14539

5. Sevastyanov L.A., Shchetinin E.Yu. (2020) On methods for increasing the accuracy of multi-class classification on unbalanced data. *Informatics and its applications*, vol. 14, no. 1, pp. 63–70 (in Russian). https://doi.org/10.14357/19922264200109

6. Timofeev V.S., Faddeenkov A.V., Shchekoldin V.Yu. (2016) *Econometrics.* Moscow: YURAYT (in Russian).

7. Tsoi M.E., Shchekoldin V.Yu. (2021) *Marketing research: methods for analyzing marketing information*. Novosibirsk: Publishing house of NSTU (in Russian).

8. Breiman L. (2001) Random Forests. *Machine Learning,* vol. 45, no. 1, pp. 5–32. https://doi.org/10.1023/A:1010933404324

9. Chistyakov S.P. (2013) Random forests: a review. *Proceedings of the Karelian Scientific Center of the Russian Academy of Sciences*, no. 1, pp. 117–136 (in Russian).

10. Abbate R., Manco P., Caterino M., et al. (2022) Demand forecasting for delivery platforms by using neural network. *IFAC-Papers OnLine*, vol. 55, no. 10, pp. 607–612. https://doi.org/10.1016/j.ifacol.2022.09.465

11. Danilchenko M.N., Muravnik A.B. (2021) Neural network approach to route construction in a special-purpose automated control system. *High-tech technologies in space exploration of the Earth,* vol. 13, no. 1, pp. 58–66 (in Russian). https://doi.org/10.36724/2409-5419-2021-13-1-58-66

12. Sustrova T. (2016) A suitable artificial intelligence model for inventory level optimization. *Trends Economics and Management,* vol. 10(25), pp. 48–55. https://doi.org/10.13164/trends.2016.25.48

13. Mikhailin D.A. (2017) Neural network algorithm for safe flight around air obstacles and prohibited ground zones. *Scientific Bulletin of MSTU GA*, vol. 20, no. 4, pp. 18–24 (in Russian). https://doi.org/10.26467/2079-0619-2017-20-4-18-24

14. Hughes A. (1996) *Boosting Response with RFM.* New York: Marketing Tools.

15. Griffin J. (2002) *Customer loyalty: how to earn it, how to keep it.* San Francisco, CA: Jossey-Bass.

16. Guo Li. (2011) A research on influencing factors of consumer purchasing behaviors in cyberspace. *International Journal of Marketing Studies*, vol. 3, no. 3, pp. 182–188. https://doi.org/10.5539/ijms.v3n3p182

17. Tsoi M.E., Shchekoldin V.Yu., Lezhnina M.N. (2017) Construction of segmentation based on modified RFM analysis to increase consumer loyalty. *Russian Entrepreneurship*, vol. 18, no. 21, pp. 3113–3134 (in Russian). https://doi.org/10.18334/rp.18.21.38506

18. Saunders M., Lewis F., Thornhill E. (2006) *Methods of conducting economic research.* Moscow: EKSMO (in Russian).

19. Demsar J., Curk T., Erjavec A., et al. (2013) Orange: Data mining toolbox in Python. *Journal of Machine Learning Research*, vol. 14, pp. 2349–2353.

20. Sturges H. (1926) The choice of a class-interval. *Journal of the American Statistical Association,* vol. 21, pp. 65–66. https://doi.org/10.1080/01621459.1926.10502161

21. Prieto A., Prieto B., Ortigosa E.M., et al. (2016) Neural networks: An overview of early research, current frameworks and new challenges. *Neurocomputing,* vol. 214, pp. 242–268. https://doi.org/10.1016/j.neucom.2016.06.014

22. Upton G. (1982) *Analysis of contingency tables.* Moscow: Finance and Statistics (in Russian).

23. Cochran W. (1976) *Sampling methods.* Moscow: Statistics (in Russian).

24. Bakhvalov N.S., Zhidkov N.P., Kobelkov G.M. (2020) *Numerical methods.* Moscow: Knowledge Laboratory (in Russian).

25. Liker J. (2005) *Dao Toyota: 14 principles of management of the world's leading company.* Moscow: Alpina Business Books (in Russian).

About the authors

Valeriya A. Nazarkina

Cand. Sci. (Econ.), Assoc. Prof.;

Associate Professor, Marketing and Service Department, Novosibirsk State Technical University (NSTU), 20, Karl Marx Ave., Novosibirsk 630092, Russia;

E-mail: valeria71@bk.ru

ORCID: 0000-0002-2207-5228

Vladislav Yu. Shchekoldin

Cand. Sci. (Tech.);

Associate Professor, Marketing and Service Department, Novosibirsk State Technical University (NSTU), 20, Karl Marx Ave., Novosibirsk 630092, Russia;

E-mail: schekoldin@corp.nstu.ru

ORCID: 0000-0001-8016-1282

DOI: 10.17323/2587-814X.2024.4.25.45

Development and testing of the toolkit of strategic planning of territorial development on the basis of an intelligent adaptive simulation model*

Marsel M. Nizamutdinov
E-mail: marsel_n@mail.ru

Vladimir V. Oreshnikov
E-mail: voresh@mail.ru

Zulfiya A. Davletova
E-mail: davletova11@mail.ru

Institute of Socio-Economic Research, Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, Russia

* The article is published with the support of the HSE University Partnership Programme

Abstract

Management of strategic development of Russian regions is a complex task, the solution of which is associated with a set of difficulties of methodological nature. In particular, there is a low quality of formed forecast assessments on the main parameters under consideration. Despite the availability of research in this area and regulatory legal framework, strategic planning documents in Russia are often not linked to each other, are repeatedly revised in the course of implementation and, ultimately, are not fully implemented. This is largely due to the fact that the available scientific potential is not fully utilized, including the development of relevant information systems. The aim of this research is to develop a decision support toolkit for strategic planning of regional development. Agent-based modeling, adaptive management, data mining and scenario modeling are used as the main research methods. In the course of the research, the concept of toolkit formation is proposed based on the construction of an intelligent adaptive simulation model (IASM), taking into account the theory of strategic planning and the ability to process heterogeneous data. The proposed structure of IASM includes four interrelated hierarchical levels – intelligent agents, macro-processes, management system and external environment. Special attention is paid to the development of a model of adaptive behavior of an intelligent agent. The proposed approach to implementation will make it possible to cover the whole range of tasks – from the analysis of input data to the development of management decisions. The software implementation of the model thus developed is carried out using the AnyLogic toolkit.

Keywords: strategic planning, territorial socio-economic systems, simulation model, adaptive management, intellectual analysis, decision support system, agent-based approach

Citation: Nizamutdinov M.M., Oreshnikov V.V., Davletova Z.A. (2024) Development and testing of the toolkit of strategic planning of territorial development on the basis of an intelligent adaptive simulation model. *Business Informatics*, vol. 18, no. 4, pp. 25–45. DOI: 10.17323/2587-814X.2024.4.25.45

Introduction

The basis for making any decision in the field of managing the development of territorial socio-economic systems is a clear scientifically substantiated idea of how its implementation will affect the change in the values of all key indicators that characterize this system. Thus, the functions of forecasting and planning as constituent elements of the management system are actualized. At the same time, the formation of long-term development forecasts requires from the decision maker a much greater amount of knowledge, background information and experience compared to the solution of operational management issues [1]. It is strategic management and the development of regional development strategies that require appropriate information support tools [2]. It should be noted that there are significant developments in this area in relation to manufacturing enterprises [3].

Under these conditions, the problems of developing approaches, methods and tools for forecasting and strategic planning in the field of development of the constituent entities of the Russian Federation that take into account these challenges are becoming more relevant. At the same time, today in this sphere there are a number of difficulties [4–7] associated with the inconsistency and contradiction of the goals outlined in the documents of strategic development, insufficient use of management information and the level of staff training. Also, the actual priority of tactical level issues against the background of low elaboration of strategic management of the region has a significant impact. As a result, the resulting forecast assessments and decisions often require revision and adjustment.

This is largely due to the existence of methodological problems in strategic planning and management which are based on a complex set of interagency approvals. Today, this problem has not been solved even at the highest level. The presentation by the Ministry of Economic Development of the Russian Federation of the updated Strategy of Socio-Economic Development until 2023, scheduled for May 14, 2021, never took place. In fact, it was announced that a transition to the development of a certain broad list of "new initiatives" was underway. One of the reasons for this situation is the harsh criticism of the draft Strategy by academic economists, who concluded that this document "continues the tradition of regularly violated plans." Similarly, the annual report of the Accounts Chamber of the Russian Federation states that "the system of strategic planning is currently unbalanced and inefficient, insufficiently methodologically ensured, with a low level of control and executive discipline; a system of strategic planning is absent; the goals of the authorities are not linked with each other and with the requirements of the President of the Russian Federation and they do not dock with the project activities of the Government of the Russian Federation; moreover, they are defined formally, and their achievement is not controlled" [8]. Back in 2017, the representative of the Ministry of Economic Development of the Russian Federation noted the discrepancies in the strategic planning documents, in particular, differences in the names of the same indicators, inconsistency in the values of target indicators, lack of balance and consistency within the documents, as well as lack of skills to work with such data. Among the main tasks was the creation of a digital platform by 2020 which would enable the automation of the strategic planning process from the moment a decision is made to develop a document to its completion and performance evaluation. To achieve this task, it was assumed to use big data, methods of simulation modeling, artificial intelligence and cloud technologies [9].

Many other experts also note that the current strategic planning documents of different levels in Russia (about 54 thousand units), although developed on the basis of the Federal Law No. 172 "On Strategic Planning in the Russian Federation", are not linked to each other [10–12]. In fact, scientists and specialists have already openly stated that in the current conditions the strategic level is subordinated to operational and tactical tasks, which is fundamentally contrary to the logic of formation of any effective management system.

High risks and costs associated with making wrong and inefficient decisions require careful consideration and evaluation of possible alternatives [13]. The availability and reliability of information plays an important role.

The aim of the research is to develop decision support tools for strategic planning of regional development. We must integrate management models, knowledge processing technologies and modeling tools in a single information field, as well as to propose methods for their incorporation into the existing planning system. At this stage, we propose to create a methodology and concept of information support for the processes of strategic planning of socio-economic development of macro- and meso-level systems on the basis of system integration of adaptive management models, intellectual knowledge processing technologies and simulation modeling.

AnyLogic system was chosen as a tool environment for simulation modeling. This software solution has a number of undeniable advantages: it is a professional tool for agent-based modeling, it is integrated with GIS maps, has extensive animation and visualization capabilities and is able to process big data as input for the model. A distinctive feature of AnyLogic is also the ability to combine different modeling paradigms; for example, system dynamics methods can be applied in agent-based models. Experiments based on simulation models built in AnyLogic allow us to perform scenario calculations and optimize resource planning.

1. Concept of research and development of an intelligent adaptive simulation model

The decision support (DS) toolkit in the creation of the development strategy of the constituent entities of Russia is primarily based on an understanding of the tasks that will be solved using it. For this purpose, the range of solutions that require justification should be defined. At the same time, the development of a regional development strategy is not a goal, but a mechanism for achieving the goals of regional development. That is, more general goals determine private goals. Thus, having analyzed the objectives of regional development strategy development as a document and its structure, we concluded that the ultimate goal of the DS toolkit should be a scientifically justified set of values of the managed parameters. At the same time, when using the developed toolkit, in our opinion, other related tasks can be solved, including:

1. Assessing the observed and retrospective state of the regional system, identifying its strengths and weaknesses, analyzing and monitoring indicators.
2. Development of a scenario development forecast, including a baseline scenario under current conditions, and analysis of possible results of the implementation of certain strategic management measures.
3. Determining the value of controlled parameters required to achieve a given value of a target parameter under given constraints, and solving planning problems.
4. Identify a set of recommended activities within the decision support task.

Based on this, the requirements for decision support tools are determined [14]. In the management of regional development, we can rather conditionally distinguish two levels – operational and strategic management. In the first case, the goal is to eliminate current problems, in particular, related to the need to achieve individual elements of strategic objectives. In turn, strategic management establishes priorities for the development of the regional socio-economic system (RSES), key parameters, as well as goals for the operational level. Thus, the parameters of the region's development are influenced by the control parameters – the results of the management subjects' activity. In addition to this, the dynamics of RSES parameters are affected by the processes occurring both in the system itself and in the external environment.

It should be clarified that in the framework of the ongoing research under intelligence in an intelligent adaptive simulation model (IASM) we understood the ability to work with knowledge, poorly structured information due to the presence of appropriate components (knowledge base, rule base, application of fuzzy logic methods, etc.) in addition to the capabilities of simulation models built on the basis of generally accepted approaches. At the same time, adaptability is implemented through behavioral algorithms that minimize planning errors through the step-by-step adjustment of economic agents' strategies [15].

The core of the toolkit is the economic and mathematical model of the RSES (*Fig. 1*). In this case, as it was shown above, the dynamics of regional socio-economic processes are determined by the set of decisions made by individual agents with the properties of adaptability capable of perceiving information, processing it and forming logically justified decisions on its basis, i.e., intellectual functions. In this regard, we propose to use the agent-based approach as the basis of the model developed.

Within the framework of building economic and mathematical models of the development of the RSE based on the principles of agent-based modeling, all agents should be described through a set of quantitative development indicators and mechanisms used to achieve them. The development of these mechanisms is one of the most difficult tasks, as they describe the agent's decision-making behavior. These actions are based on work with information related to receiving, storing, processing and using. Interaction with the external environment is also an exchange of information, i.e., values of some parameters.

In our opinion, it is the availability of the ability to form logically valid conclusions and adapt the agent through its training that allows us to obtain the most suitable model of the real RSES. In this regard, we propose the development of IASM.

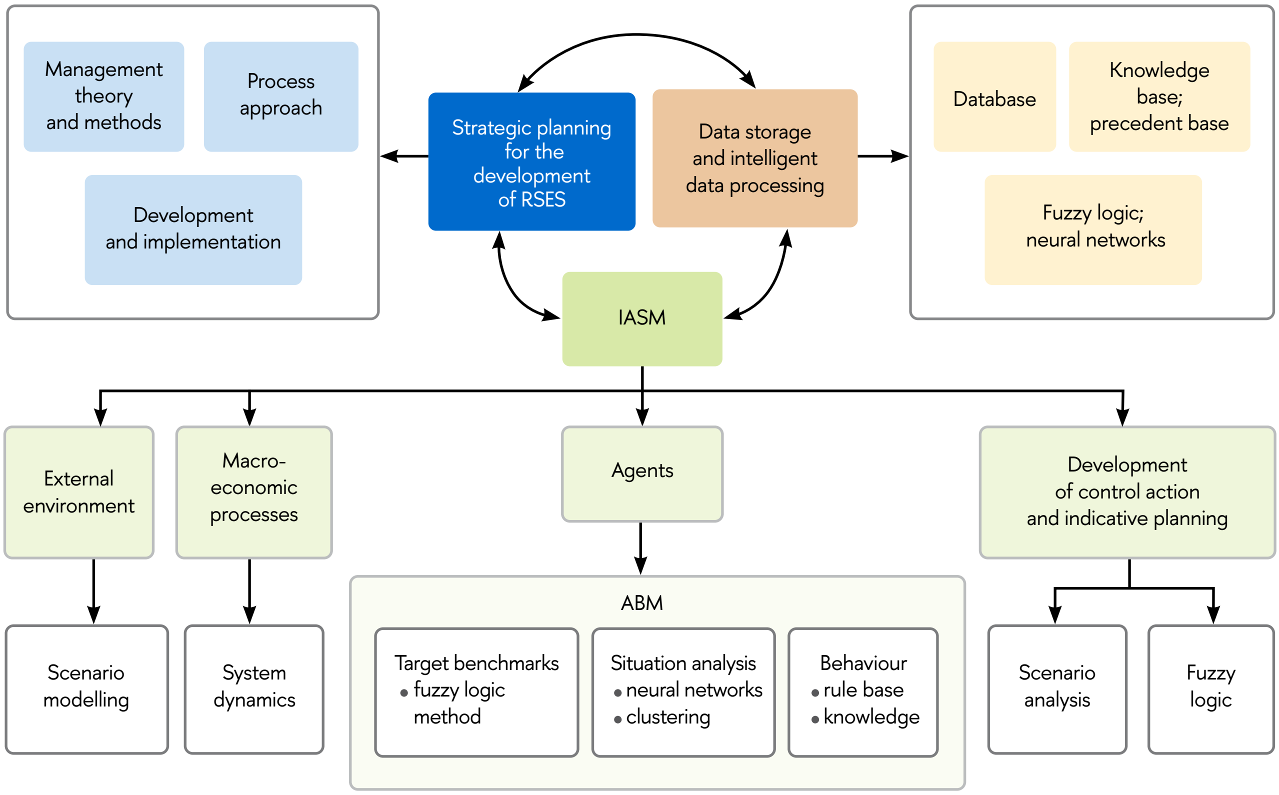

Fig. 1. Concept study and toolkit development.

At the same time, it is obvious that for the description of many macroeconomic processes it is more logical to use methods that do not require us to describe the behavior of individual agents, i.e., with a higher level of aggregation [16]. In this regard, it makes sense to reflect some macroeconomic processes using dynamic modeling and other methods. Such a combination of different approaches optimizes the costs of development and operation of the model of a socio-economic system.

The novelty of our statement of the problem and the proposed research concept, in our opinion, lies in the assessment of the possibility of applying the class of adaptive management models [17] and methods of intelligent knowledge processing widely used in the theory of decision-making in complex human-machine systems to formalize the bounded-rational logic of the behavior of economic agents at the regional level when implementing their development strategy in a competitive market environment.

In its theoretical aspect, the distinctive feature of the proposed approach is the consideration of the region not from the position of macroeconomics, within the framework of which aggregated processes do not allow us to see the contribution of their individual components, but from the position of the balance of macro- and micro-level representation, which is reflected both in the methodology of managing the development of the object and in the simultaneous use of two paradigms of simulation modeling – system dynamics and the agent-based approach to solve the relevant problems. This manifests itself in the focus of the proposed research on the integration of knowledge about the research object requiring specific methods and approaches to study its individual elements.

In its instrumental aspect, the originality of the proposed model of information support organization, in our opinion, consists in the logical-hierarchical composition of a set of behavioral, agent-oriented and management models in a single information environment. Unlike previously developed approaches, this allows us to form strategies for the development of regional systems coordinated by hierarchy levels and developed taking into account the goals and interests of agents and strategic development plans of the region as a whole.

2. Structure of an intelligent adaptive simulation model

The approach presented here is based on the application of formalized methods of analysis, forecasting and planning. The development of the corresponding toolkit implies, first of all, the formation of its core – an intelligent adaptive simulation model. To describe it, in our opinion, it is necessary to define the structure of the model, as well as the methodological and informational basis for its formation. The general structure of the IASM is presented in *Fig. 2*.

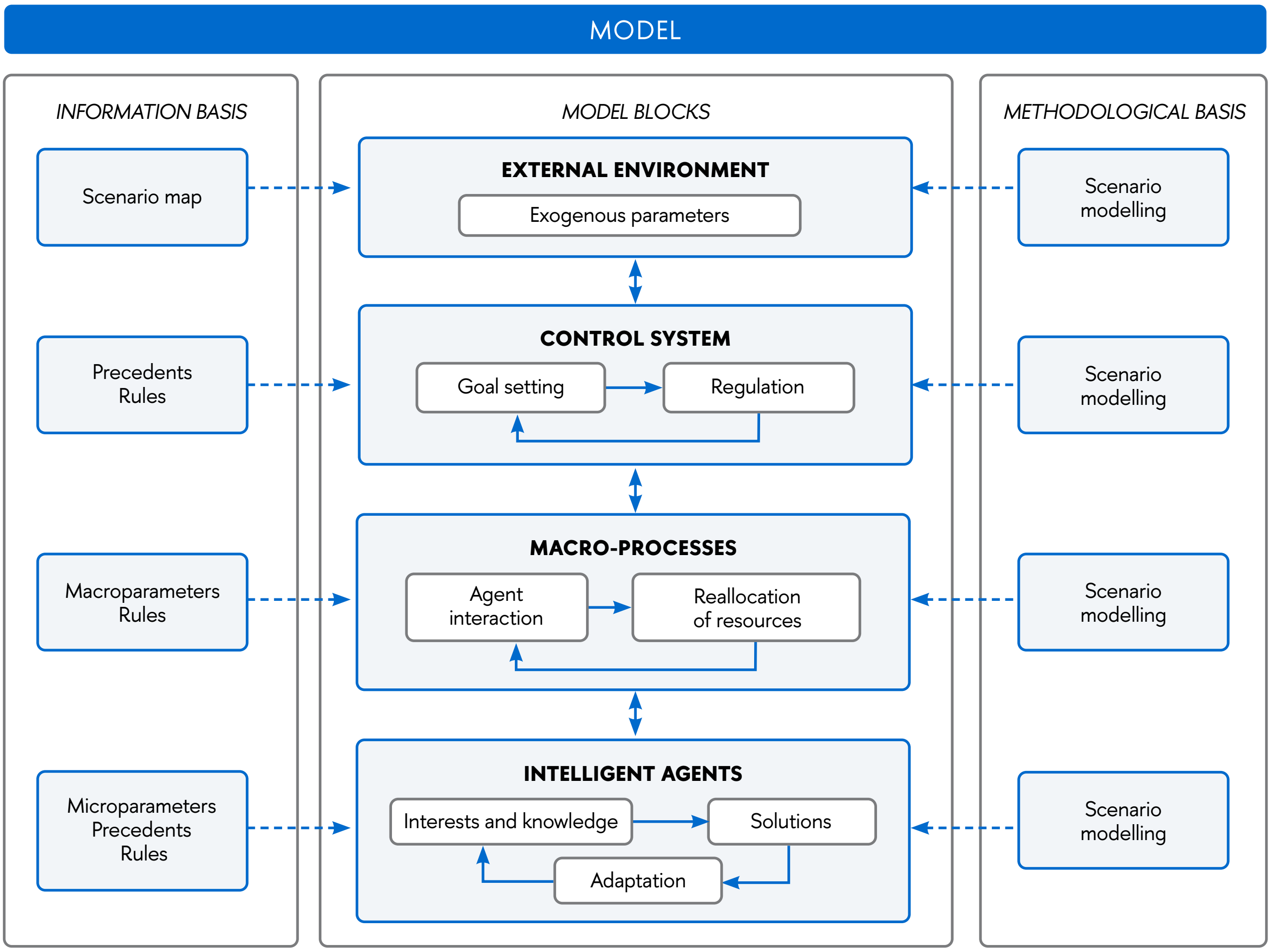

Fig. 2. Structure of the intelligent adaptive simulation model.

The RSES model includes several levels.

1. The level of intellectual agents reflecting the functioning of individual agents and small social groups. Each of them has its own interests, which are largely conditioned by the situation in which it finds itself. Within the framework of the model, an agent compares the parameters of a given situation with the expected characteristics and makes decisions based on this. To model the procedure of analyzing and classifying the situation, we propose to apply neural network technologies and methods that allow us not only to make a choice between already known options, but also to identify those belonging to previously unknown situations [18]. At the same time, decision making can be based both on the use of rules (in the case of clearly defined requirements for the agent's actions) and on the basis of precedents (in the case when there is no rule and there are many known options). If the situation is fundamentally new, a new solution must be developed, which is also a manifestation of the agent's intelligence. In this case, the description of its behavior is based on the theory of bounded rationality [19] within the framework of which the decisions taken will not necessarily be optimal. The implementation of each decision and the evaluation of its consequences lead to the expansion and adjustment of its knowledge base and precedent base, i.e. adaptive learning of the agent takes place. Due to this, with the appearance of new precedents, the next decisions can be made according to new rules. In general, the functioning and development of an agent is described within the framework of the agent-based modeling (ABM) methodology, the use of which in relation to complex socio-economic systems is considered in the works of a number of authors [20, 21].

2. The interaction of many individual agents generates macroeconomic processes. At this level of aggregation, it becomes too costly to consider each agent separately. In this regard, a set of macroeconomic agents is considered. The ratio of stocks of various resources of these agents forms the structural model of the economy. At the same time, resources can be transferred from one agent to another and transformed. All this is proposed to describe using the methods of system dynamics. In interaction each agent has its own interests. Interaction in market conditions is possible only if the interests are coordinated.

3. The control system is a setting above the model of the region and serves for setting target indicators and regulators. If the actual values of the indicators do not correspond to the previously set values, the regulators are adjusted first. However, if this does not have the expected effect under the existing constraints, the values of target indicators are adjusted. In this case, the choice of decision (to change the values of regulators or target indicators) is based on the analysis of the situation and adaptive management.

4. The development of the regional system is significantly influenced by the external environment, which is described by a set of exogenous parameters. In the framework of this study, they are independent of the processes occurring within the RSES. The values of exogenous parameters are set using scenario maps.

Thus, the structure of the intellectual adaptive simulation model describes a set of socio- economic relations implemented at different levels in the most generalized form. In this case, its basis is the functioning of intelligent agents.

3. Model of adaptive behavior of an intelligent agent

The behavior of an intelligent agent is described within the framework of the developed model through its decision-making aimed at satisfying its own interests under the existing conditions of functioning, taking into account the available knowledge (rules, precedents, known situations and solutions), as well as the possibility of agent adaptation. The adopted and realized decision determines the change in the agent's parameters and influences the characteristics of the external environment. As stated above, the agent functions under conditions of incomplete and contradictory information. Its actions can be described by the theory of bounded rationality. This concept in the classical sense characterizes the situation in which a person considers a small number of options that differ significantly and chooses the option closest to his aspirations, which does not guarantee the maximization of utility [22, 23]. In the framework of the proposed model, this situation can be represented by a set of possible decisions of the agent, realized with some probability. Based on previous experience and available knowledge, the agent chooses a solution that satisfies him, which does not indicate the optimality of the choice made, i.e., the consideration of the entire possible set of solutions.

Based on these assumptions, the research has developed a model of agent decision making (*Fig. 3*), based on the application of a set of methods of agent-based modeling, situation classification, theory of bounded rationality, application of rule bases and precedents, data analysis, and others.

The totality of information defines the situation in which the agent finds itself. Identification of a situation consists in attributing it to some known class or, if necessary, allocating a new class. The set of situations known to the agent forms a part of the agent's knowledge base – the "Situation Base".

Another part of it is the "Decision Base", which contains information about what actions the agent can perform (a list of controllable parameters and their characteristics). The functioning of an intelligent agent is a set of decisions it makes $Rh = \{Rh_t^1, Rh_t^2, ..., Rh_t^n\}$ based on the prevailing class of situation $K = \{K_t^1, K_t^2, ..., K_t^n\}$ at time t.

The choice of a decision option is based on the rules and precedents available to the agent, which are contained in the "Rules Database" and "Precedents Database," respectively. They describe what actions and with what probability the agent chooses in this or that situation. Referring to these databases allows us not only to correlate the observed conditions and decisions, but also to determine the expected consequences of their realization (expectations). At the same time, actions that cannot be performed (there are no conditions for their performance) have zero probability of choice, i.e., they are actually excluded from the list of possible decisions of the agent.

The rule base includes conditions $U = \{U_t^1, U_t^2, ..., U_t^n\}$, decisions and expected outcomes $Rz\_r = \{Rz\_r_t^1, Rz\_r_t^2, ..., Rz\_r_t^n\}$, and it is subscribed to by a tuple of elements $Ru = \langle U, Rh, Rz\_r \rangle$. In turn, the precedent database contains decisions $Rh\_b = \{Rh\_b_t^1, Rh\_b_t^2, ..., Rh\_b_t^n\}$, made in a previously observed situations $K\_b = \{K\_b_t^1, K\_b_t^2, ..., K\_b_t^n\}$, and information about previously observed outcomes $Rz\_b = \{Rz\_b_t^1, Rz\_b_t^2, ..., Rz\_b_t^n\}$. The precedent database is described by a tuple of elements $Pr = \langle K\_b, Rh\_b, Rz\_r \rangle$.

A precedent is an incident or event that took place in the past and serves as an example or a basis for subsequent actions in the present [24]. On this basis, precedents are a reflection of the agent's experience and knowledge about the relationship between events and phenomena. The result is a list of indicators that have changed their value after the realization of an action and are associated by the agent with the consequences of this action, as well as the significance of changes in these indicators.

In this case, priority is given to the use of the rule base, as it contains more stringent requirements for the definition of the solution laid down in normative form or as a result of repeated precedents. At the same time, the rule base is not unchangeable. It can be replenished and adjusted.

The use of the precedent database implies accessing it and searching for the situation most similar to the current one. For this purpose, the parameters of the current situation are compared with the existing conditions of known precedents. It should be noted that the search for a precedent is carried out not by an individual indicator, but by the whole set of indicators that form the situation. In this case, we are talking about reasoning on the basis of precedents, which in a broad sense is a method of solving new problems on the basis of already known solutions. Reasoning on the basis of known situations is a special case of reasoning by analogy.

Speaking about situation identification, it should be understood that the agent needs not only to determine its position among many other variants of situation development in this area, but also to determine how familiar it is to him. Thus, if the agent has previously encountered it, this situation can be called known and known solutions can be considered for it. Otherwise,

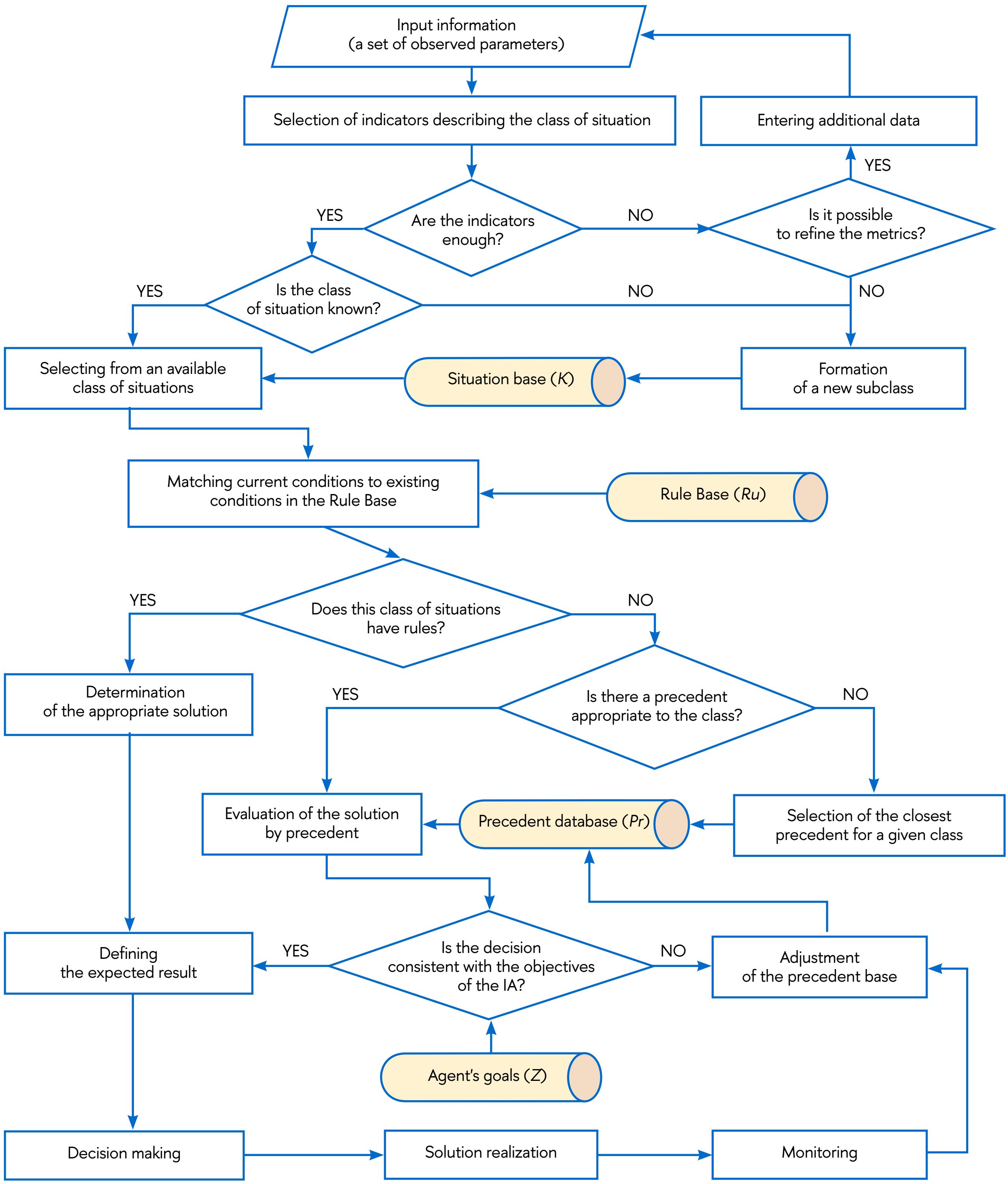

Fig. 3. Model of decision making by an intelligent agent.

the agent has no experience of behavior in this situation, and the decision can be made on the basis of the closest known one. Thus, the intelligent agent, analyzing the current situation, compares the actual conditions *U_f* with the conditions contained in the rule base *U_b*, and if a match is found, applies the corresponding rule *Rh*. If there is no rule, the available precedents are considered and the most coincidental to the actual and previously observed situation is identified among them.

One of the main properties of an agent is the presence of its target attitudes. In the framework of this study, the agent's target attitudes are understood as a set of the agent's ideas about which parameters are significant indicators and which of the directions of changes in the values of indicators of its development is more preferable, as well as what is the ratio of the significance of target indicators among themselves. Thus, in terms of the formation of the agent's target orientations, we can distinguish three fundamental issues:

1. What are the target benchmarks for this agent, i.e., the composition of parameters that are attributed to the target indicators?
2. What is the desired value of the agent's target benchmarks, i.e., some quantitative measurement of these target indicators;
3. How do the values of certain goals of the agent relate to each other?

The most important benchmarks for the development of society are defined in the Decree on the National Development Goals of the Russian Federation for the period until 2030 and in the perspective until 2036. Based on them, it is possible to identify a number of indicators for quantitative assessment of the functioning of economic agents, including the following:

- increase in the total fertility rate;
- increasing life expectancy;
- poverty reduction;
- provision of housing for citizens;
- ensuring the growth rate of the country's gross domestic product above the global average;
- increase in the volume of investments in fixed capital;
- ensuring sustainable growth of the population's income, etc.

As can be seen, despite the fact that the indicators themselves are quantitatively measurable, the target attitudes have qualitative characteristics. This fact allows them to reflect more adequately the behavior of real people whose targets can be described as "increase in monetary income," "increase in the level of environmental safety of places of residence," etc., and the ratio of importance between them is described by the concepts of "significantly more important," "somewhat more important," etc. That is, we are talking about the application of indicators described using fuzzy logic methods. An example of combining this approach and agent-based modeling is available in [25].

The realization of the agent's set goals is carried out through the choice of decisions aimed at changing the controlled parameters under certain conditions that form the situation faced by the agent. Analyzing the existing conditions, the agent determines to which of the known classes this or that situation belongs. To do this, it needs to have an appropriate knowledge base, including a description of situations, and a tool for their identification and classification. It is proposed to use clustering methods to determine the class of a situation.

The behavior of an economic agent is largely determined by the information it possesses. The following sources of information can be distinguished: 1) information from the external environment, the parameters of which do not depend on the agent himself; 2) information that was formed during the interaction of the agent with the external environment; 3) information characterizing the agent himself, the receipt of which does not require interaction with other agents or the external environment.

Speaking about the set of parameters characterizing the agent itself, it should be noted that this group of indicators also includes the resources possessed by the agent. In this sense, there is a relationship between resources and controlled parameters of the agent. In fact, the agent's behavior is a change in the values of these parameters. As part of the formation of the economic-mathematical model, information about the list of these indicators and their characteristics are contained in the agent's decision base. It should be emphasized that, in a broad sense, the list of these indicators is not only not the same even for agents of the same type, but also is not constant over time. If we consider the agent "Human", the list of controllable parameters and possibilities to change their values depend on age, social status, place of residence, etc. However, from the point of view of solving practical modeling issues, it seems reasonable to establish a common list of controllable parameters for all agents of the same type, but defining for each of them different possibilities of adjusting its values. More difficult in terms of formalization, in our opinion, is the problem of the emergence of fundamentally new controllable parameters as a result of changes in external conditions, scientific and technological progress and human creativity. It should be noted that the list of controllable parameters depends to a significant extent on the level of detail of the model and its general focus on the solution of certain problems.

Based on the generalized scheme we have presented, the behavior of each of the agents under consideration was formalized.

The functioning of the economic agent "Human" is multidimensional. Decisions made by him affect issues in all spheres of life and, depending on the chosen level of detail, may include both strategic issues (e.g., place of residence, change of occupation, choice of training direction in higher education, etc.) and everyday issues, the solution of which is referred to the operational level and is often "mechanical" in nature.

For this economic agent, the research conducted proposes to consider decision making in three aspects:

- change of place of residence (migration);
- revenue generation;
- cost formation.

Let us consider the second question in more detail, since in many respects its solution determines the parameters of the situation for others (*Fig. 4*).

Adaptation of an economic agent to a changing situation consists in changing the rules of its behavior depending on certain conditions.

The key goal of the agent in this direction is to obtain income not lower than expected. It should be emphasized that in this case, as mentioned earlier, it is not about maximizing the values of the parameter, but about satisfying some given level.

From the set of characteristics of the agent "Human" it is required to select those that are significant from the point of view of solving the given problem. In this case, the characteristics can be selected from different groups – Demographics (gender, age), Education (level of education, direction of training), Work (length of service, experience of entrepreneurial activity, current status of professional activity), Finances of the population (current and retrospective level of income). On the other hand, the agent's decision is influenced by the state of the environment in which he/she is located, including the characteristics of the economic activities under consideration from the point of view of a potential employee (level of remuneration, availability of vacancies) and from the point of view of an entrepreneur (level of competition and availability of demand for FEA products). The combination of these characteristics determines certain situations which are supposed to be identified by cluster analysis. Given the multitude of possible values of these characteristics, it is not possible to determine in advance the exact number and parameters of each situation, but it is obvious that the decision made by a young specialist with no experience of entrepreneurial activity in the conditions of economic recession and excess of personnel on the labor market will be significantly different from the decision of an entrepreneur with experience in the market, where there is an increase in demand. At the same time, the differences in decisions are caused not only by the difference in the values of factors, but also by the rules applied by agents. Since the establishment of deterministic relationships between conditions and decisions in this case is not possible (including due to the impossibility of full formalization of conditions and their quantitative assessment, as well as unambiguous attribution of the observed combination of conditions to a particular class of situations), it seems reasonable to choose a solution using intelligent algorithms, including those based on fuzzy logic methods.

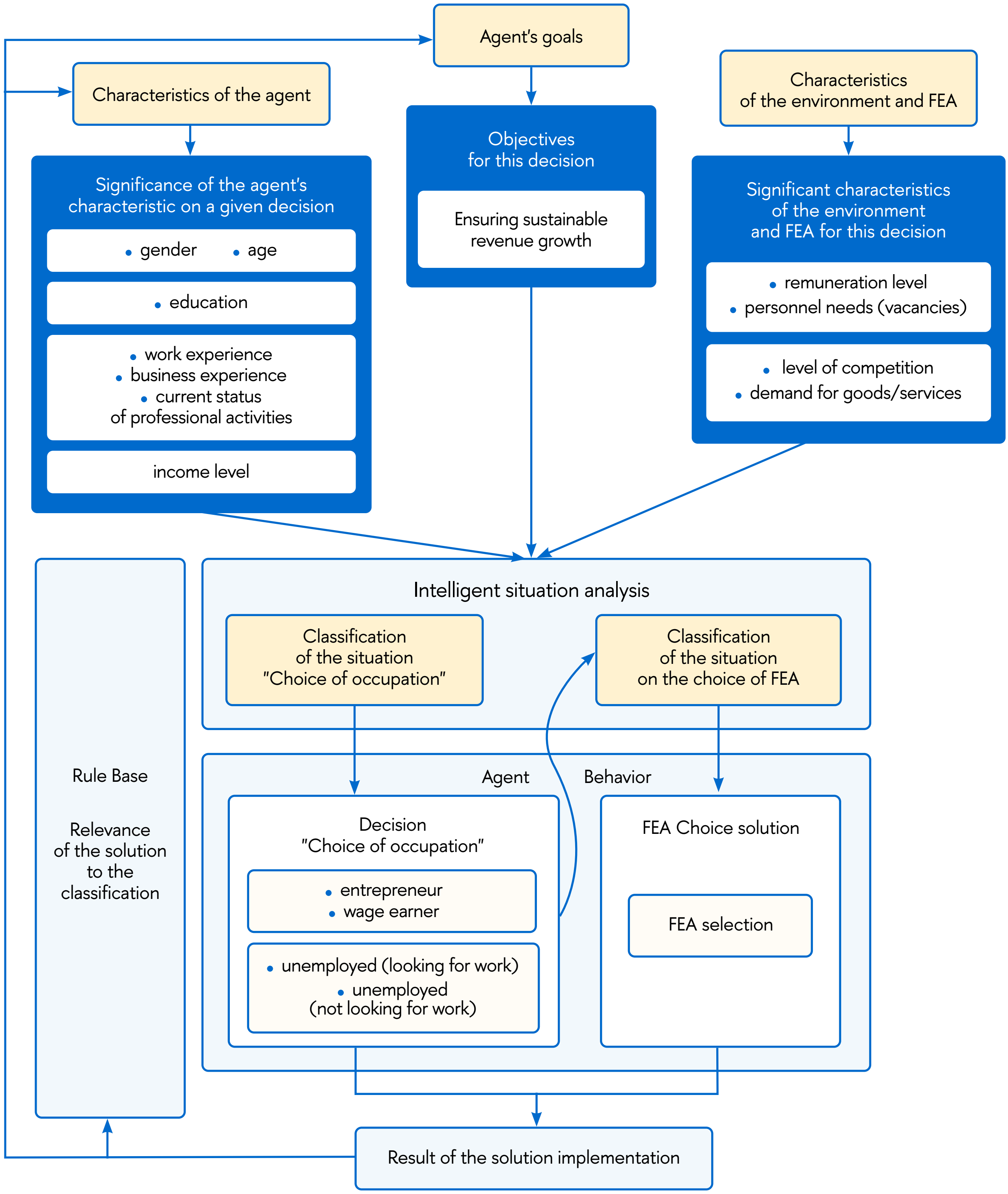

Fig. 4. Formation of income of the agent "Human".

These rules can be transformed as a result of experience gained during the implementation of previously made decisions. Moreover, this experience may lead to changes in the agent's target attitudes. This procedure implements not only the possibility of adaptation of the economic agent to changing conditions, but also the adaptation of the rules of behavior of the agent itself.

With respect to other economic agents and decisions, it seems reasonable to formalize the decision selection procedure in a similar way.

At the same time, the calculation of other parameters of functioning of the economic agent "Human" is realized within the framework of the set of equations presented below.

One of the key subsystems of the model is the demographic block, including the description of population movement. These processes lead to changes in the number of groups. The method of age movement is used in the modeling toolkit. Taking into account the gender aspect increases the adequacy of the reflection of the processes. N gender and age groups are included in the study: $\forall n, n = \overline{1;N}$.

$$P_t^n = P_{(t-1)}^{(n-1)} + B_t^n - D_t^n + M_t^n,$$

where P_t^n – population;

B_t^n – number of birth (for the first age group);

D_t^n – number of deaths;

M_t^n – balance of migration flows.

The number of births in the t-th year is calculated based on the birth rate kB_t using the formula:

$$B_t^n = kB_t \cdot P_t \cdot k_{m/w},$$

where $k_{m/w}$ is the gender distribution coefficient.

It should be taken into account that the fertility rate depends on a set of heterogeneous factors, including the number of women of fertile age, marriage rate, income level of the population and housing availability of the population.

The values of these factors are determined based on the logic of the model formation. For example, the number of wages, as an element of the level of income of the population, is determined by the interaction of the economic agent "Business" within the framework of the policy to attract labor force and the agent "Human" within the framework of making decisions on the formation of their own income. In turn, the volume of the total area of residential premises is determined by the results of the implementation of the consumer function by the agent "Human", i.e., his decision-making with regard to the use of own and borrowed financial resources.

The economic essence of the agent "Human" is realized in the formation and use of financial resources. The key sources of income InH_t^{Σ} are entrepreneurial activity InH_t^{en}, labor remuneration InH_t^{w}, social transfers InH_t^{tr}, property income InH_t^{prop}, and other income InH_t^{of}. In addition, credit resources can be considered InH_t^{cr}.

$$InH_t^{\Sigma} = InH_t^{en} + InH_t^{w} + InH_t^{tr} + \\ + InH_t^{prop} + InH_t^{of} + InH_t^{cr} + StH_{t-1}.$$

The most important source of monetary income today is labor remuneration. In this case, the level of labor remuneration and the number of employees of enterprises and organizations are determined within the framework of the proposed model on the basis of the market mechanism of interaction of economic agents.

To formally describe the income strategy of the economic agent "Human", a procedure of adaptive behavior on the labor market is provided in which a sequential choice of the type of activity an FEA takes place. The agent evaluates the attractiveness of each of the four options – employment, entrepreneur activity, status of unemployed, job seeker and unemployed who is not looking for a job at the moment. Based on a set of characteristics of the agent and the environment, a choice is made in favor of one of the options. From a macroeconomic point of view, the number of workers L_t^w depends on the value of wages Wg_t, the number of the working-age population P_t^{em} and the amount $InH_t^{\Sigma n-p}$. That is, the balance of supply and demand in the labor market is considered. However, in contrast to purely balance methods of calculation, the proposed approach is based on the adaptive behavior of an intelligent agent capable of assessing the economic situation and preferability of choice. Due to this, at each step of the iterative algorithm the adjustment of the number of employed in the economy by value λ^L is realized.

Similarly, the behavior of the agent with respect to the formation of its other parameters is described, and the description of the rest of the agents in the system is carried out.

Thus, one of the distinctive features of an intelligent agent is its ability to adapt to changes in the conditions of functioning. In this case, we are talking not only about adjusting quantitative parameters, but also about changing the qualitative component of the agent's behavior which manifests itself in two aspects:

1. Choosing different courses of action under different conditions.
2. An agent's ability to learn.

The first component is associated with the determination of changes in the values of the controlled parameters. However, it is the agent's learnability that provides the possibility of its representation as a full-fledged and independent subject of economic relations and thus forms the essence of the agent-based approach to modeling and the basis of an intelligent adaptive model.

In the framework of the current research, learnability is considered as the possibility of adjusting the rules of behavior. If the agent's rules of behavior are variants of actions in a given situation with a certain probability of choosing one or another action, then in this aspect learning is based on adjusting the probability of choosing each of them. The agent does not just make a choice of this or that action, but also forms a certain expectation. If the observed parameters confirmed his expectations, then the decision made (the chosen action) is perceived by the agent as correct and the probability of its use in a similar situation is preserved. If a different result was obtained than expected, the probability is corrected. At the same time, the agent's learning should not be limited only to changing the parameters of choosing one or another already known action. As noted earlier, the agent can supplement its knowledge base and, in some cases, adjust the goals of functioning and its own interests. It should be noted that adjusting the values of an agent's target indicators when they cannot be achieved is consistent not only with the notion of adaptation of agent behavior, but also with critical research on agent behavior [26].

As part of the testing of the proposed IASM toolkit, agent-based models were developed to describe one of the most important tasks of strategic planning – the regulation of migration and demographic processes at the regional level. *Figure 5* shows a fragment of the agent-based model interface, which allows us to determine the dynamics of migration flows between federal districts of the Russian Federation.

The regulated parameters in the model are the average wages in the region, the volume of investment in fixed capital, as well as the integral index of the population's quality of life. Such models allow us, in particular, to form balanced schemes of territorial population distribution, to justify the conditions and mechanisms for reducing disproportions in the socio-economic development of individual territories, caused, among other things, by low population density and lack of qualified personnel for the regional economy. The blue dots on the map show the agents representing the permanent population of the region, and the red ones – migrants. The drop-down menu makes it possible to compare territories by the parameters of population size and density, average wages, and volume of investment in fixed assets. The model outputs are diagrams illustrating the projected ratio of migrants to the permanent population, as well as the dynamics and direction of migration flows. A separate sheet calculates the statistics of arrivals/departures for each federal district in relation to the selected "base" territory (*Fig. 5* shows the example of the Republic of Bashkortostan).

4. Application of an intelligent adaptive simulation model in strategic planning

The use of IASM as a fundamental component of decision support systems (DSS) in the field of strategic planning implies the use of the toolkit's capabilities at all key stages of the process under consideration. At the same time, the most important function is the development of a forecast of socio-economic development of the territory, taking into account the peculiarities of the behavior of all parties involved, including

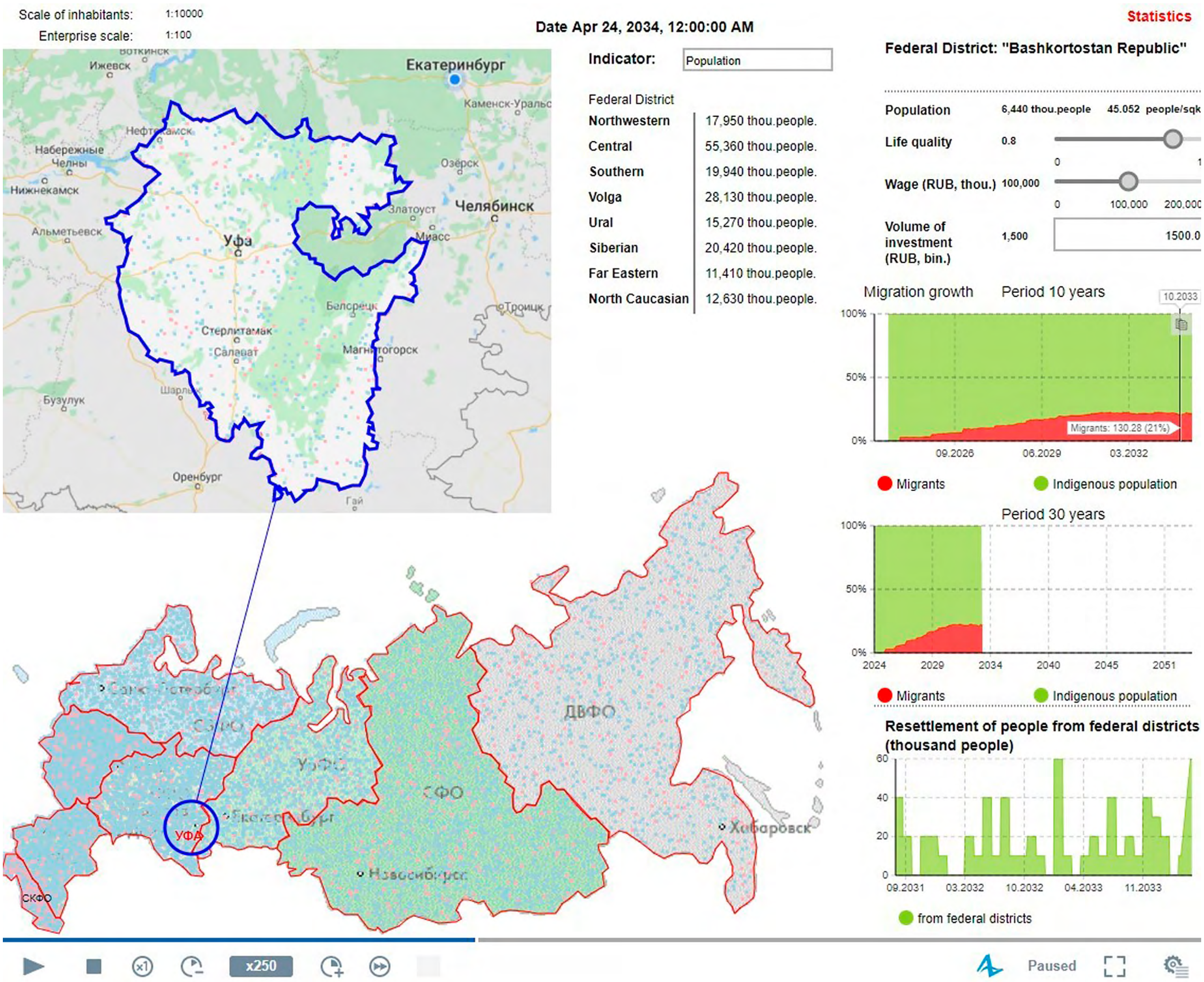

Fig. 5. Interface of the agent-based model for regulating migration flows between federal districts of the Russian Federation.

consideration of the adaptability of their behavior to changes in the operating environment. The integration of IASM into the existing procedures and algorithms of strategic decision-making should be based not on the replacement of the DSS with a model toolkit, but on complementing the capabilities and already available methods. Considering this aspect of the problem under study, it is necessary to briefly describe the stages of strategic planning and the expected place of IASM in these stages (*Fig. 6*). It should be noted that for full-fledged functioning of the IASM within the framework of the DSS, it should be supplemented with a number of functional blocks (data input and output, report generation, calculation of dynamics and structure indicators, etc.), but these issues are not considered in this study, as they are more engineering, rather than scientific in nature.

I. The preliminary stage includes problem definition, organizational support, data collection and pre-

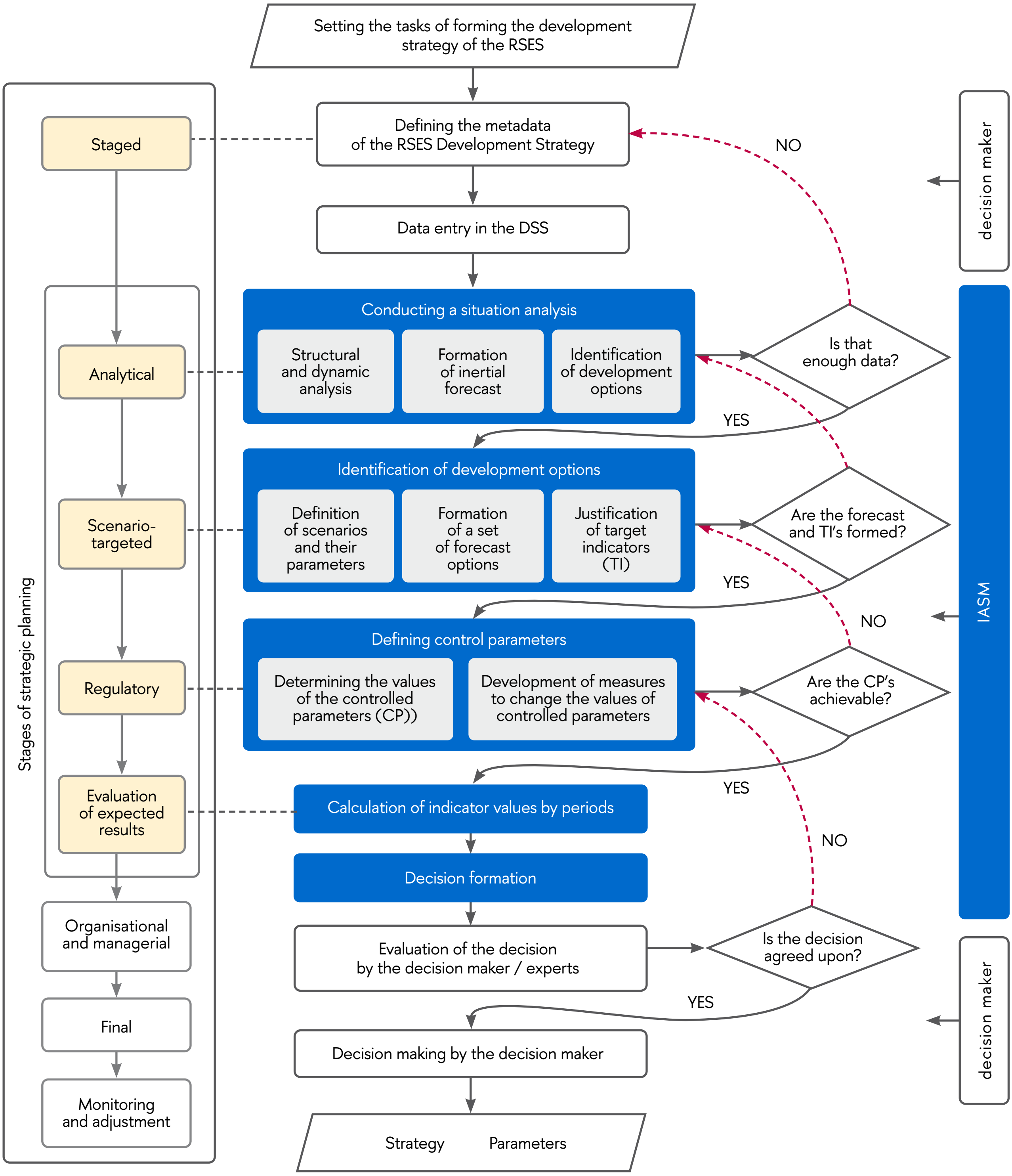

Fig. 6. Algorithm for determining the parameters
of the development strategy of the RSES using the IASM-based DSS.

liminary analysis. From the point of view of IASM work, at this stage the data are entered into the system and checked. At the same time, we are talking not only about statistical information reflecting the current situation and development of the territorial system, but also parameters related to the establishment of the planning horizon, exogenously imposed restrictions.

II. The analytical stage includes determining the state and direction of development of the territorial system. Based on the task at hand, it is necessary to assess economic, demographic, social, cultural, natural-geographical and other factors. This stage may involve the implementation of dynamic, structural, SWOT-analysis, the study of individual spheres of society, analysis of advantages, identification of problems and threats to development. The result is the definition of territorial development problems or identification of unused potential.

Integration of the IASM into strategic planning procedures is based on the formation of a conservative version of the forecast of key macroeconomic parameters, implying the preservation of existing conditions and existing values of managed parameters.

III. Scenario-target stage includes the definition of target benchmarks for the development of the territorial system. Within the framework of the IASM, the creation of development scenarios is carried out through the formation of an indicative plan. Thus, the tasks of forecasting and planning the development of the territorial socio-economic system are linked. The application of simulation modeling allows us to take into account the limitations and various conditions.

IV. Regulatory stage. Determination of parameters of the vector of control actions based on the value of target indicators is based on reduction of deviation of current values from those obtained as a result of modeling.

A set of measures is formed to achieve the required values of regulators in the areas. At this stage, the number of required resources is determined.

V. The stage of assessment of expected results. At this stage, the parameters of the territorial system development are determined as a result of multiple revisions of the controlled parameters of all economic agents considered in the framework of modeling. It is this stage that makes the greatest use of the IASM capabilities, because it makes it possible to obtain forecasting and planning estimates first at the level of individual agents and then at the level of aggregated indicators. Parameter estimation is based on a number of modeling experiments.

VI. Organizational and management stage, where the system of strategic decisions implementation is defined. This involves the description of management processes, delineation of authorities and areas of responsibility of executors, methods and mechanisms of interaction, etc. Application of the IASM is not assumed.

VII. Final stage. Before proceeding to the practical implementation of the results obtained from strategic planning, they must be approved in the form of a normative document and receive the appropriate legal status. At this stage, the use of the IASM is also not expected.

VIII. Stage of monitoring and adjustment. During the implementation of strategic decisions, the compliance of the values of the obtained results with the parameters of the indicative plan is monitored. If a deviation is detected, the initial parameters are revised and the modeling complex is adjusted. The frequency of monitoring and appropriate adjustment of the IASM is determined by the objectives of the study and the specifics of the management area.

It follows from the above that the proposed toolkit built on the basis of IASM can be applied at most stages of strategic planning and is an auxiliary tool, not replacing specialists, but improving their efficiency.

On the basis of the developed DS toolkit with the use of agent-based modeling technologies, we obtained a forecast assessment of the dynamics of migration flows between the federal districts of the Russian Federation in case of changes in the value of certain controllable parameters: the integral indicator of the population's quality of life, the volume of investment in fixed capital and average wages provided by the current strate-

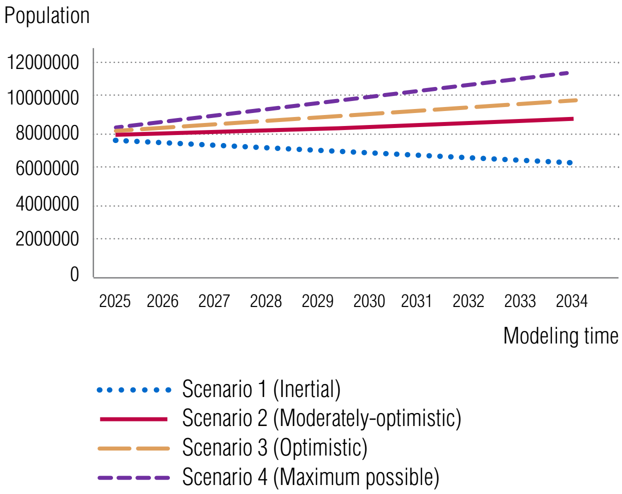

Fig. 7. Model dynamics of population change in the Far Eastern Federal District under different scenario options.

gic plans for the integrated development of territories (*Fig. 7*). Scenario experiments were conducted (using the Far Eastern Federal District as an example of the territory with the lowest population density), which made it possible to forecast for the period up to 2035 the dynamics of the district's population under four basic scenario options, including various combinations of controllable parameters. Preliminary results showed the variation of population change in the FEFD (*Fig. 7*) ranging from –20% (inertial), to +12% (moderately optimistic), +25% (optimistic) and +45% (maximum possible).

In general, the proposed toolkit forms the basis for the development of a balanced demographic and migration policy within the framework of the development and implementation of strategic development plans for the Russian regions in the medium and long term.

Conclusion

The approach to strategic planning of territorial development on the basis of an intellectual model developed in the framework of our research is based on the system integration of various instrumental approaches: methods of agent-based modeling in terms of formalizing the behavior of intelligent agents, which are characterized by adaptive behavior and incomplete rationality; methods of dynamic modeling in terms of formalizing flow data at the macro level.

The proposed IASM structure includes four interrelated hierarchical levels – intelligent agents, macroprocesses, management system and external environment. These levels form the blocks of the model, each of which requires a specific information and methodological basis. In the framework of the conducted research, a model of adaptive behavior of an intelligent agent based on its decision-making has been developed. There is no doubt that each economic agent has distinctive features conditioned by goals, position in the economic system, resources used, significant parameters, etc. The example presented of income generation of the agent "Human" includes only a minor, but important component of the functioning of this agent.

From a practical point of view, the implementation of the IASM-based DSS will increase the level of consistency and efficiency of the decisions taken. In our opinion, such a toolkit may be of interest, first of all, to public administration bodies engaged in the development of strategic decisions at the regional level. An algorithm for determining the parameters of the RSES development strategy with the use of DSS based on IASM is proposed, and the sequence of actions interconnected with the stages of strategic planning is determined. It is shown that most of the key tasks imply the possibility of increasing efficiency through the use of modern information systems. At the same time, the IASM is only a tool, not a substitute for the decision maker. ■

Acknowledgments

The article was prepared with the financial support of the Russian Science Foundation, project No. 23-28-00871.

References

1. Suvorov N.V., Treshchina S.V., Beletsky Yu.V. (2020) Problems of developing methods for long-term forecasting of the dynamics of the domestic economy (methodology and model tools). *Problems of Forecasting,* no. 6(183), pp. 66–80 (in Russian).

2. Lipushkina I.Yu. (2014) Formation of a unified system of strategic and territorial planning of regional development. *Strategy for sustainable development of Russian regions*, no. 24, pp. 37–47 (in Russian).

3. Makarov V.L., Bakhtizin A.R., Beklaryan G.L., Akopov A.S. (2021) Digital plant: methods of discrete-event modelling and optimization of production characteristics. *Business Informatics*, vol. 15, no. 2, pp. 7–20. https://doi.org/10.17323/2587-814X.2021.2.7.20

4. Belyanova A.M., Biryukov V.A., Cherkovets V.N. (2016) Strategic planning in conditions of modern Russia's economy (materials of research seminar on strategic planning). *Moscow University Economics Bulletin*, no. 3, pp. 141–158 (in Russian).

5. Leonov S.N. (2018) Problems of forming a system of regional planning in Russia at the present stage. *Trudy Bratskogo gosudarstvennogo universiteta. Seriya: Ekonomika i upravleniye,* vol. 1, pp. 3–8 (in Russian).

6. Shvetsov A.N. (2016) Destructive stereotypes of cardinal planning. *VIII international research-to-practice conference "From recession to stabilization and economic growth", Plekhanov Russian University of Economics*, pp. 8–16 (in Russian).

7. Reznikova A.V. (2014) Strategic planning of regional development as the main element of territorial development. *Issues of Regional Economy*, no. 4(21), pp. 148–154.

8. Accounts Chamber of the Russian Federation (2020) *Bulletin of the accounts chamber of the Russian Federation. State administration*, no. 1(226). Available at: https://ach.gov.ru/statements/byulleten-schetnoy-palaty-1-266-2020-g (accessed 15 April 2024) (in Russian).

9. Center for Development Strategy (2017) *Strategic planning and new technologies*. Available at: https://www.csr.ru/ru/news/2313/ (accessed 15 April 2024) (in Russian).

10. Karpov V.K., Lylov A.S. (2020) On the problems of main planning in Russia. *Theory and practice of the world science*, no. 3, pp. 2–9 (in Russian).

11. Pankevich N.V., Rudenko V.V. (2021) Selected problems of identity in planning management documents of the Russian Federation. *Surgut State University Journal*, no. 4(34), pp. 128–134 (in Russian).

12. Uskova T.V. (2020) Problems related to development and strategic planning in the research of VolSC RAS. *Economic and Social Changes: Facts, Trends, Forecast,* vol. 13, no. 6, pp. 9–23 (in Russian).

13. Nijkamp P. (1983) Information systems for regional development planning: A contemporary review. *Reprinted from Environment and Planning B: Planning and Design*, vol. 10. Available at: https://core.ac.uk/reader/33893640 (accessed 15 April 2024).

14. Borzykh N.Yu. (2022) Analysis of decision support systems, their classifications and decision-making methods. *Tendentsii razvitiya nauki i obrazovaniya*, no. 91-7, pp. 87–90 (in Russian).

15. Nizamutdinov M.M., Oreshnikov V.V. (2016) Determination of parameters for managing regional development based on fuzzy logic algorithms. *Economics and Mathematical Methods*, vol. 52, no. 2, pp. 30–39 (in Russian).

16. Galin D.M., Sumarokova I.V. (2021) Application of a dynamic macromodel of Russia to forecast the development of its economy. *Trudy Instituta sistemnogo analiza Rossiyskoy akademii nauk (ISA RAN) (Proceedings of the Institute for Systems Analysis Russian Academy of Sciences (ISA RAS))*, vol. 71, no. 3, pp. 11–23 (in Russian).

17. Nizamutdinov M.M., Oreshnikov V.V. (2018) Information support for the formation of strategies for innovative development of regions based on an adaptive artificial model. *Systems and Means of Informatics*, vol. 28, no. 2, pp. 154–169 (in Russian).

18 Borodin S.N. (2024) Forecasting the economic development of the region based on regression and neural network models. *Russian Economic Bulletin*, vol. 7, no. 1, pp. 290–300 (in Russian).

19. Poznyak K.V. (2020) The concept of bounded rationality in behavioral economics. *Universum: psikhologiya i obrazovaniye*, no. 10(76), pp. 14–15 (in Russian).

20. Makarov V.L., Bakhtizin A.R., Beklaryan G.L., Akopov A.S. (2019) Development of a software platform for large-scale agent-based research of complex social systems. *Software Engineering*, vol. 10, no. 4, pp. 167–177 (in Russian).

21. Novikova T.S., Tsyplakov A.A. (2020) Social policy in a multi-sectoral agent-based mode. *Economic and Social Changes: Facts, Trends, Forecast,* vol. 13, no. 3, pp. 129–142 (in Russian).

22. Novikov D.A. (2022) Bounded intelligence and control. *Matematicheskaya Teoriya Igr i Ee Prilozheniya*, vol. 14, no. 1, pp. 49–84 (in Russian).

23. Linsky D.V. (2021) Quantitative assessment of expectations and opinions of economic agents regarding socio-economic development. *Theoretical economics*, no. 4(76), pp. 66–75 (in Russian).

24. Maslovsky A.E., Kosinova A.V. (2023) Case-based reasoning in soft decision support. *Naukosfera*, no. 9–2, pp. 126–130 (in Russian).

25. Miigranova L.I., Atnabaeva A.R. (2020) An approach to modelling educational migration using fuzzy logic and agent-based research. *Economics, entrepreneurship and law*, vol. 10, no. 12, pp. 3389–3402 (in Russian).

26. Simon H.A. (1978) Rationality as a process and as a product of thought. *American Economic Review*, vol. 68, no. 2, pp. 1–16.

About the authors

Marsel M. Nizamutdinov

Cand. Sci. (Tech.), Assoc. Prof.;

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

E-mail: VOresh@mail.ru

ORCID: 0000-0001-5779-4946

Zulfiya A. Davletova

Cand. Sci. (Tech.);

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

E-mail: davletova11@mail.ru

ORCID: 0009-0008-4389-2113

DOI: 10.17323/2587-814X.2024.4.46.60

Building a system of dynamic norms for evaluating the functioning of complex systems on the example of the regions of the Central Federal District

Roman A. Zhukov [a]
E-mail: pluszh@mail.ru

Svetlana V. Prokopchina [b]
E-mail: svprokopchina@mail.ru

Maria A. Plinskaya [a]
E-mail: maria.plinskaya@gmail.com

Maria A. Zhelunitsina [a]
E-mail: maria202001@yandex.ru

[a] Financial University under the Government of the Russian Federation, Tula Branch, Tula, Russia
[b] Financial University under the Government of the Russian Federation, Moscow, Russia

Abstract

In this paper, we present a method of forming norms for evaluating the results of the functioning of complex systems applicable to socio-ecological and economic systems, taking into account the priorities of the development of the regions of the Russian Federation. The methodology involves the selection of normative values from a set of norms based on two methods: the first is based on the construction of econometric models using statistical data for a set of subjects (the first type) and for one selected subject (the second type). The second method uses the methodology of Bayesian intelligent measurements based on the regularizing Bayesian approach (the third and fourth types). Depending on the result of the calculations, a norm is selected that gives a higher (in the case of high priority), average (in the case of medium priority) and lower (in the case of low priority) normative value of the evaluated effective features characterizing the development of the subject. The implementation of the method is demonstrated by the example of the regions of the Central Federal District, including the Tula Region, for which econometric and fuzzy models of the relationship between the volume of gross regional product with the value of fixed assets and the number of employees for sections A (Agriculture, forestry, hunting, fishing and fish farming) and C (Mining) according to OKVED1 are constructed, forming the raw materials sector according to data for 2007–2022. The EFRA and Infoanalyst 2.0 software platforms are used as tools. The results obtained can be used by regional authorities in the formation of norms to assess the results of the functioning of the regions in the short and medium term.

Keywords: econometric model, fuzzy model, Bayesian intellectual measurements, norm, software platform, socio-ecological and economic system

Citation: Zhukov R.A., Prokopchina S.V., Plinskaya M.A., Zhelunitsina M.A. (2024) Building a system of dynamic norms for evaluating the functioning of complex systems on the example of the regions of the Central Federal District. *Business Informatics*, vol. 18, no. 4, pp. 46–60. DOI: 10.17323/2587-814X.2024.4.46.60

Introduction

When assessing the results of the functioning of complex systems, which are the subjects of the Russian Federation considered as socio-ecological and economic systems (SEES) [1], qualitative problems arise: how can one judge how satisfactory such results are, and with which norms to compare them. This article is a continuation of previous research, including research submitted for a dissertation defense [2]. By norm, we will understand the expected (planned) value of the effective feature, characterizing a given mode of operation of the system, its subsystem or element. If the developed standards act as norms for technical or software systems [3–5], then there are no such norms for SEES. This is so even if we consider the ecological subsystems of the SEES [6] which must operate within the framework of environmental protection standards [7]. For example, for enterprises – institutional units – residents of the subjects of the Russian

Federation, operating in accordance with the All–Russian classifier of economic activities (NACE), and the totality of which is an element of one of the subsystems within the framework of the accepted classification [8, 9], pollution is allowed above the established limits and maximum permissible concentrations of harmful substances (MPC). At the same time, this trend may persist for many years due to the specifics of production. This means at the regional level, some questions may arise from the acceptability (acceptability, expectation and, consequently, normativity) of such a result, allowed and regulated by "compensation" – payments to the budget for environmental pollution, which from the point of view of creating a favorable living environment. In addition, subjects operate in specific, unique and time-varying conditions, which should impose a number of restrictions on the results of their functioning, including the established norms. Therefore, to assess the activities and development of the SEES based on a system of indicators and norms, it is necessary to use methodologies, methods and techniques that take into account emerging constraints. An example is private and integral indicators of performance and efficiency, such as: average indicators over the evaluation period [10, 11]; technical efficiency indicators [12, 13]; indicators based on production functions [14–16] or within the framework of fuzzy modeling [17–19]. In conditions of incompleteness and fuzziness of data, the Bayesian intelligent measurement methodology (BIM) works well [20, 21], and therefore its application seems appropriate [22–24]. At the same time, the assessment indicators and norms used, calculated on the basis of different methodologies, can give different results, which determines the need to form a method for selecting norms among existing ones and improving methods for calculating indicators. In this study, we aim to develop a method that allows us to make such a choice, as well as its implementation taking the example of subjects of the Central Federal District with an emphasis on the Tula Region, using the EFRA software package [25] and a new version of the Infoanalyst 2.0 software platform developed in Python [26] that enables us to calculate soft norms and use them in the formation of dynamic norms.

1. Building standards based on econometric modeling

Let there be a k_i-th element from K ($k = 1..K_k \in N$) subsystem of type S_q, which is a set of economic units that contribute to the volume of gross regional product (GDP by region) according to the i-th of the I sections NACE 1 ($i = 1..I \in N$). Moreover, $k_i \subset K \cap I$. We will consider GDP by region to be a performance sign (result) of the functioning of the element k_i and denote it as $y_{k,i}$, and the corresponding value at t observation time ($t = 1..T \in N$) is $y_{k,i}(t)$. The element functions under certain conditions j ($j = 1..J \in N$), characterized by factorial features (factors) $x_{k,i,j}$ with values $x_{k,i,j}(t)$.

Then the relationship between the values of the results and the factors can be represented in the form of an econometric equation [27]:

$$y_{k,i}(t) = f_i(C_{i,j}, x_{k,i,j}(t)) + \varepsilon_{k,i}(t). \tag{1}$$

where $C_{i,j}$ are parameters of function $f_i(\cdot) = \hat{y}_i$;

$\varepsilon_{k,i}$ are the values of the stochastic random component ε_i, which is assumed to be normal in the first approximation: $\varepsilon_i \sim N(0; \sigma^2_{\varepsilon_i})$.

Then, to calculate the normative values i-th performance feature of the first type for k_i in period t, we will use $\hat{y}_i$, which is a production function (PF), the parameters of which can be estimated using standard methods (OLS, MLE) from a combined k and t sample. Here is an example of such a sample. We will consider data for the regions of the Central Federal District (CFD) for 2007–2022. The NACE section agriculture, forestry, hunting, fishing and fish farming (section A) has the following effective feature – the volume of GDP by region (in millions of rubles). The factor features are the number of employees (thousand people) and the cost of fixed assets at full accounting value at the end of the year (million rubles).

Substituting in (1) the actual values of factor features $x_{k,i,j}(t)$ for a specific k region in the time period t, it is possible to obtain normative (expected, planned) values of the performance feature $\hat{y}_{k,i}(t)$, which we will call dynamic standards of the first type and denote $\bar{y}_{k,i,\mathrm{I}}(t)$.

If we construct a PF based on data only for k_i (for example, for Tula Region), then we can obtain the values of dynamic norms of the second type $\tilde{y}_{k,i}(t)$ in the form of a model $\tilde{y}_{k,i}$, that was previously used in another capacity as a model of the functioning of k_i-th element which is a component of the optimization model within the framework of a multi–level optimization approach [28]. Let's denote the values of dynamic norms of the second type as $\bar{y}_{k,i,\mathrm{II}}(t)$.

2. Building norms based on the Bayesian intelligent measurement methodology

We will use the same notation as in paragraph 1. The values of the norms for $y_{k,i}(t)$ will be denoted $\bar{y}_{k,i}(t)$, and for $x_{k,i,j}(t)$ will be denoted $\bar{x}_{k,i,j}(t)$. Then, in order to build them within the framework of the BIM, the following steps must be carried out.

1. Setting a priori scales. This stage involves the formation of basic a priori numerical and linguistic scales with dynamic constraints (SDC) of the form $[(\cdot)_{\min} - \sigma_{(\cdot)}; (\cdot)_{\max} + \sigma_{(\cdot)}]$; $(\cdot)$ are $y_{k,i}$ or $x_{k,i,j}$ (in terms of BIM, these are influencing factors); $(\cdot)_{\min}$, $(\cdot)_{\max}$ are minimum and maximum values among $y_{k,i}(t)$ or $x_{k,i,j}(t)$ values; $\sigma_{(\cdot)}$ is correction of the interval determined by the formula:

$$\sigma_{(\cdot)} = \overline{(\cdot)}^{\,a} / 3. \tag{2}$$

Here a is the designation of the a priori scale, and $\overline{(\cdot)}$ is defined as:

$$\overline{(\cdot)} = ((\cdot)_{\max} + (\cdot)_{\min}) / 2. \tag{3}$$

We will call $\overline{(\cdot)}^{\,a}$, as well as the norms that are already known (set on the basis of standards or experts), norms of the third type, which we will designate as $\bar{y}_{k,i,\mathrm{III}}(t)$.

For a numerical scale, the user sets the number of reference points L_r, characterized by a pair of numbers $(h_{(\cdot),l}(t);\ p_{(\cdot),l}(t))$, where $h_{(\cdot),l}(t)$ is value of the factor $(\cdot)(t)$, corresponding to the position of the reference point on the scale with a range of dynamic constraints ($l = 1..L_r \in N$); $p_{(\cdot)}(t)$ is the probability that the value $(\cdot)(t) = h_{(\cdot),l}(t)$. The linguistic scale has nine classes L_c, ranging from extremely below the norm to extremely above the norm. For the linguistic scale, the procedure is similar.

The scale is divided into nine intervals, taking into account the previously established norm of the third type through the transition from numerical to linguistic representations [29].

In this case, the condition must be met:

$$\sum_{l=1}^{L} p_{(\cdot),l}(t) = 1. \tag{4}$$

2. Representation of influencing factors in the form of fuzzy numbers. The representation of a value $x_{k,i,j}(t)$ in the form of a fuzzy number will be determined by a set of pairs of numbers $(h_{(\cdot),l}(t);\ p^{ap}_{(\cdot),l}(t))$ using a formula similar to the Bayes formula [29]:

$$p^{ap}_{(\cdot),l}(t) = P^{ap}(H_{(\cdot),l} \mid S) = \\ = P^{a}(H_{(\cdot),l}) P(S \mid H_{(\cdot),l}) \Big/ \sum_{j=1}^{J} P^{a}(H_{(\cdot),l}) P(S \mid H_{(\cdot),l}), \tag{5}$$

where ap is the designation of a posteriori probability;

$H_{(\cdot),l}$ is set of hypotheses or alternative solutions;

S is event consisting in the joint appearance of estimates $\tilde{h}_{(\cdot),l}(t)$ for $(\cdot)$.

3. The formation of a soft norm for an effective feature. As in the previous step, the value for $y_{k,i}(t)$ will be represented by a set of pairs of numbers $(h_{y_{k,i},l}(t); \bar{p}^{ap}_{y_{k,i},l}(t))$, where the probability can be determined using a formula similar to (5):

$$\bar{p}^{ap}_{y_{k,i},l}(t) = \frac{P^{a}_{y_{k,i},l}(h_{y_{k,i},l}(t-1)) P(\tilde{h}_{y_{k,i},l}(t))}{\sum_{j=1}^{J} P^{a}_{y_{k,i},j}(h_{y_{k,i},j}(t-1)) P(\tilde{h}_{y_{k,i},j}(t))}. \tag{6}$$

Here $(t-1)$ means that the a priori representations for were obtained in the previous period, and is calculated by recurrent application of Bayesian convolution:

$$P(\tilde{h}_{y_{k,i},l}(t)) = \frac{1}{J}\sum_{j=1}^{J} P^{ap}_{x_{k,i,j},l}(h_{x_{k,i,j},l}(t) | \bigcup_{s=1}^{J} h_{x_{k,i,s},l}(t)). \quad (7)$$

In the case of two influencing factors $P^{ap}_{x_{k,i,j},l}(h_{x_{k,i,j},l}(t) | \bigcup_{s=1}^{J} h_{x_{k,i,s},l}(t))$ it can be represented as:

$$P^{ap}_{x_{k,i,1},l}(h_{x_{k,i,1},l}(t) | h_{x_{k,i,2},l}(t)) = \\ = \frac{P^{ap}(H_{x_{k,i,1},l} | S) P^{ap}(H_{x_{k,i,2},q} | S)}{\sum_{l=1}^{L}\sum_{q=1}^{Q} P^{ap}(H_{x_{k,i,1},l} | S) P^{ap}(H_{x_{k,i,2},q} | S)}, \quad (8)$$

where L, Q are significant alternative values for the first and second factors, respectively.

4. Calculation of the norm for the effective feature. The value of the norm $\overline{y}_{k,i}(t)$ for time t as a probability weighted average $\overline{p}^{ap}_{y_{k,i},l}(t)$ is determined as:

$$\overline{y}_{k,i}(t) = \sum_{l=1}^{L} \tilde{h}_{y_{k,i},l}(t) \cdot \overline{p}^{ap}_{y_{k,i},l}(t). \quad (9)$$

At the same time, information can come from various sources, which imposes restrictions in the form of metrological requirements $\{MFlow_{k,i}(t)\}$. In addition, measurement conditions $Conditions_{k,i}(t)$ can also be unique and depend on a variety of a priori information A, constraints and assumptions O. Therefore, in the general case, the relations presented (for probabilistic factors) must be supplemented with the expression $|\{MFlow_{k,i}(t)\}| \; Conditions_{k,i}(t)$. We will call this norm the norm of the fourth type $\overline{y}_{k,i,\mathrm{IV}}(t)$.

3. The choice of norms using the priorities of the development of subjects

Let there be a set of values of norms of the first, second, third and fourth types ($\overline{y}_{k,i,\mathrm{I}}(t)$, $\overline{y}_{k,i,\mathrm{II}}(t)$, $\overline{y}_{k,i,\mathrm{III}}(t)$, $\overline{y}_{k,i,\mathrm{IV}}(t)$) for elements k_i. Each k_i can be assigned a priority in the form of rank ($rank = 1..Rank \in N$), which on the qualitative side corresponds to the degree of importance of the k subject's development in the direction i. The more significant i is, the higher the priority. Priority can be determined based on expert assessments, using, for example, the pairwise comparison method. If, as performance sign, we take the contribution to the GDP by region of an entity by type of activity (NACE), then as a ranking criterion we can choose the share of GDP volume in section i, in the total GDP by region volume of the subject (region). In this case, it is advisable to use an interval estimate to determine the rank. The length of the interval for each subject k is determined by the formula:

$$l = (d_{k,\max} - d_{k,\min}) / Rank \quad (10)$$

where $d_{k,\max}$, $d_{k,\min}$ are the maximum and minimum shares of GDP according to NACE for the subject k.

Then the priority for the element is $k_i \to rank$, if the condition is met:

$$d_{k,i} \in [d_{k,\min} + l \cdot (invrank - 1); d_{k,\min} + l \cdot invrank)], \quad (11)$$

where $d_{k,i}$ is the share of the element k_i in GDP by region according to NACE i;

$invrank$ – reverse ranking;

$d_{k,\max}$ is included in the interval for the first priority.

Since the number of types of norms is 4, the number of ranks will be 4, respectively ($Rank = 4$) with priority terms: essential, important, unimportant, essentially unimportant. If only one period is considered, the ranking of norms will be carried out from a higher to a lower value among the norms $\overline{y}_{k,i,m}(t)$ (m = I, ..., IV). In the case of setting norms for a medium-term period, for example, for the period T_p, after the last known value in the period $t = T$, which occurs when forming socio-economic development programs, it is recommended to rank according to an aggregated indicator $\overline{y}_{k,i,m}(t)$, which is determined as:

$$\overline{y}_{k,i,m} = \frac{1}{T_p}\sum_{t=T-T_p}^{T} \overline{y}_{k,i,m}(t), \quad (12)$$

where $y_{k,i}$ is the actual value of the indicator; $\bar{y}_{k,i,m}(t)$ is the value of the m type norm.

According to this criterion, it is possible to choose the norm for $y_{k,i}$.

4. EFRA and Infoanalyst 2.0 software platforms

The EFRA software platform is a decision support system based on econometric modeling and a multilevel optimization approach [25]. The software environment is implemented in the Delphi programming language and using the MS SQL Server database.

Working with the EFRA software package consists of the following stages:

1. Uploading data.

2. Correlation and factor analysis of dependencies.

3. Construction of partial and integral indicators.

4. Optimization.

At the first stage, the user downloads data from an *xlsx* file or an embedded database for subsequent processing.

The second stage provides for:

1. Conducting a correlation analysis. It includes the calculation of paired correlation coefficients, t-statistics with indication of significant coefficients for a user-defined significance level.

2. Factor analysis of dependencies. It includes the construction of econometric models of various functional forms (linear, exponential, power multiplicative, logarithmic based on OLS) and statistical evaluation of models and their parameters based on classical statistical tests (Fisher is used for the estimation R^2, Student is used for the estimation of the model parameters, Farrar–Glober is used for the estimation the multicollinearity, Spearman is used for the estimation of the heteroscedasticity). In addition, it is possible to make a forecast of effective signs based on the expected values of factors that can be entered manually or downloaded from a file (*xlsx* format).

At the third stage, partial (for elements) and integral (for subsystems and the system as a whole) performance and efficiency indicators are calculated, as well as harmony coefficients (indicators of the balanced functioning of the SEES subsystems), which make it possible to draw conclusions about the nature of the state of the SEES, its subsystems and elements in the time period under consideration (if the value of the performance feature is greater than one, then the functioning is considered satisfactory, less – unsatisfactory).

At the fourth stage, optimization problems are solved with the choice of the type of restrictions, the result of which is to obtain quantified changes in the values of factors that will lead to an improvement in the values of the effective signs, that is, ensure their compliance with the standards with a given degree of deviation, taking into account the imposed restrictions.

The first and second EFRA modules make it possible to build norms of the first (when loading data for a set of k_i-th elements) and second (for the specific k_i element) types.

The Infoanalyst 2.0 software platform is also a decision support system (Python development language with integration with MS SQL Server), which is based on the methodology of Bayesian intelligent measurements and the methodology of fuzzy inference based on permissive rules, which makes it possible to make recommendations to decision makers, depending on the state of the complex system being studied, visualized in the form of a hierarchical information model with the ability to enable audit mode with cognitive graphics (infograms).

Working with the Inofanalyst 2.0 software platform consists of the following stages:

1. Model structure formation and data loading.

2. Setting the scales and calculating the dynamics.

3. Modeling and forecasting.

4. Making recommendations.

At the first stage, the hierarchical structure of the model is set. This is displayed in the corresponding frame of the software module. In addition, data is loaded into a software environment that provides for manual input or download from an *xlsx* or *xml* file. The values entered are displayed in the "Data" tab. The "Factors" tab provides a description of the factors. Each of the introduced factors is visualized in the frame as an element of a hierarchical information model.

At the second stage, an a priori scale is set for each of the factors (formula (4) is implemented); an a posteriori scale is generated (formula (5)); dynamics is calculated (scales are automatically calculated for all specified time periods); an integral or resultant factor is calculated (in this case, it is used to calculate norms of type four) based on formulas (6)–(8). The results of calculating a posteriori scales and norms can be saved to a file (*xlsx*, *csv*, *xml* and *json* formats), including factor values, their corresponding probabilities, terms (for linguistic scales), the left and right boundaries of possible factor values. The priori and posteriori scales, dynamics of factors and their interpretation, and metrological characteristics are displayed in the corresponding frames of the Infoanalyst 2.0 interface.

At the third stage, modeling and forecasting are carried out. The simulation is an approximation of fuzzy data by user-defined order polynomials (OLS). The result of such modeling are three models: the most probable model, the models of the lower and upper levels, the latter of which are determined by the threshold of significance for the probabilities of the terms (reference points) of the factor. The forecast includes setting the prediction horizon, the order of the derivative and the number of points for calculating the average finite differences of a given order based on actual or fuzzy data. The result is presented in the form of a posteriori numerical and linguistic scales, which can also be saved to a file.

The fourth stage includes the formation of recommendations, including auto-recommendations (for which factors it is necessary to strengthen or weaken their influence on the effective or integral feature) based on the built-in fuzzy inference algorithm, which provides for a comparison of the likely occurrence of the factor value in one of the classes of the linguistic scale (9 classes, starting from extremely below the norm to extremely above the norm).

All calculations are recorded in the MS SQL Server database.

Each of the stages is implemented through the appropriate software modules of the environment.

In this study, Infoanalyst 2.0 is used to build norms of the third and fourth types, which implies the use of a basic module (main.py) and the scale construction module (Scale.py).

5. The results of the method's implementation on the example of the regions of the Central Federal District

The information base of this study is publicly available statistics on 17 subjects of the Central Federal District (excluding Moscow). The source is: "Regions of Russia. Socio-economic indicators" [30] (published on the website of the Federal State Statistics Service (Rosstat)); data from the Unified Interdepartmental Information Statistical System (UIISS) [31]. GDP by region data for the Central Federal District regions in 2022 were obtained from [32], and the gross structure from [33]. All cost indicators were adjusted for the inflation rate [34] and brought to the level of 2007 within the framework of the hypothesis of invariance of processes relative to models [35]. Econometric models of power multiplicative form were used to construct dynamic norms of the first and second types. The performance indicators (according to NACE 1) are the gross domestic product by region (volume) under the section "Agriculture" (section A) and GDP by region under the section "Mining" (section C), forming the raw material sector in accordance with the sectoral classification [8]. The following factors are selected: the cost of fixed assets at full accounting value at the end of the year (FA, million rubles) and the number of employees (NE, thousand people). This model structure is similar to the classical Cobb–Douglas model. The base evaluation period was the period from 2007 to

2022, and for each of the models, in order to improve their characteristics, different periods were selected, starting from the base period and ending with the period 2020–2022. The models were based on a combined k ($k = 1, ..., 17$) and t ($t = 2007, ..., 2022$) sample (based on data from 17 regions of the Central Federal District), as well as data for each of the 17 regions (34 models in total).

The general view of the models is represented by the formula:

$$y_i = a_0 \cdot x_{i,1}^{a_1} \cdot x_{i,2}^{a_2} + \varepsilon. \qquad (13)$$

where y_i is GDP by region volume;

$x_{i,1}$ is the cost of fixed assets;

$x_{i,2}$ is the number of employees;

a_i are the coefficients of the model;

ε are residuals for sections A and C corresponding to NACE 1;

$i = 1, 2$.

The models were evaluated both as a whole and according to their parameters by p-value. The test results are presented on an external resource and contain a number of statistical tests, including: the Student's criterion (for evaluating model parameters); Fisher's (estimation of the coefficient of determination); classical tests for a number of residues (5 tests), homoscedasticity testing [36].

Figure 1 demonstrates the actual data and results of calculating the standards of four types of GDP by region volume in sections A (Agriculture, forestry, hunting, fishing and fish farming) and C (Mining) according to NACE 1 for the Belgorod, Bryansk, Vladimir and Voronezh regions of the Central Federal District. The beginning of the lines corresponds to the initial period that was used in the construction of models (13).

From *Figs. 1* and *2* it can be seen that the dynamics of norms $\overline{y}_{k,i,\mathrm{I}}(t)$ (the model is based on data from a combined k and t sample) and $\overline{y}_{k,i,\mathrm{IV}}(t)$ are similar. For the Bryansk and Tambov regions, due to the lack of data, the volume of GDP by region (section C) has not been considered.

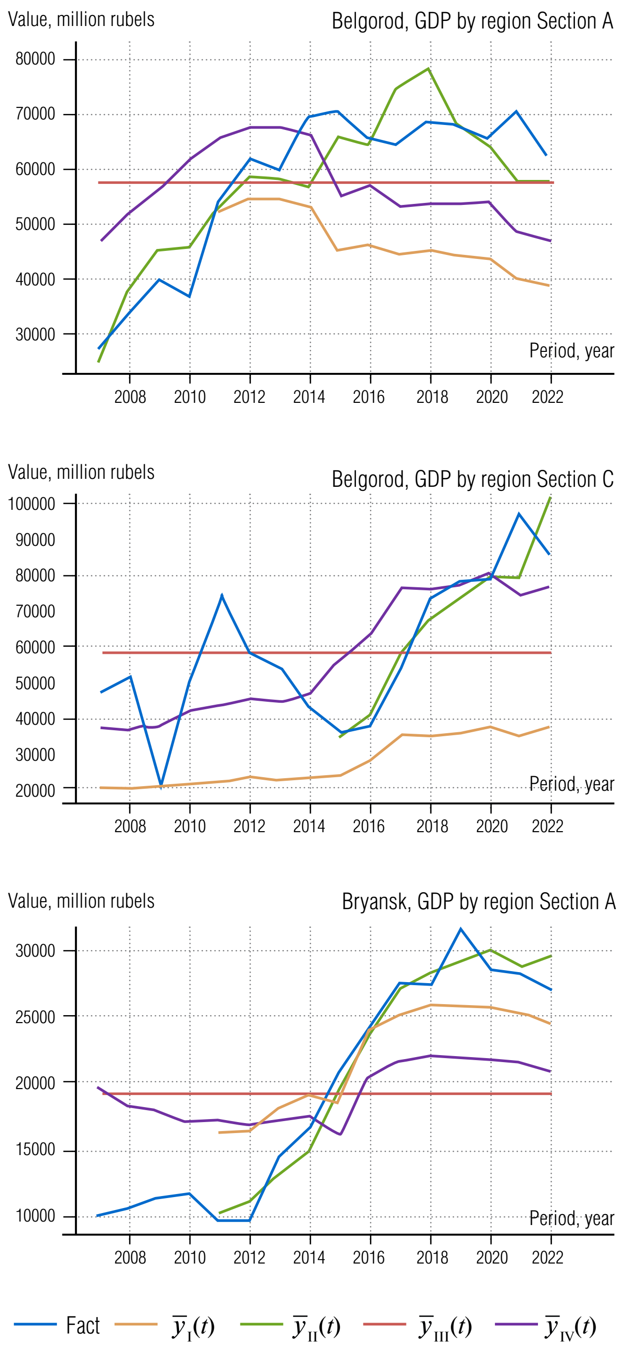

Fig. 1. Actual and normative values of GDP by region volume for sections A and C (Belgorod–Bryansk regions).
Fact is the actual values of the indicator;
$\overline{y}_{\mathrm{I}}(t), \overline{y}_{\mathrm{II}}(t), \overline{y}_{\mathrm{III}}(t), \overline{y}_{\mathrm{IV}}(t)$ –
are norms of types I, II, III and IV, respectively.

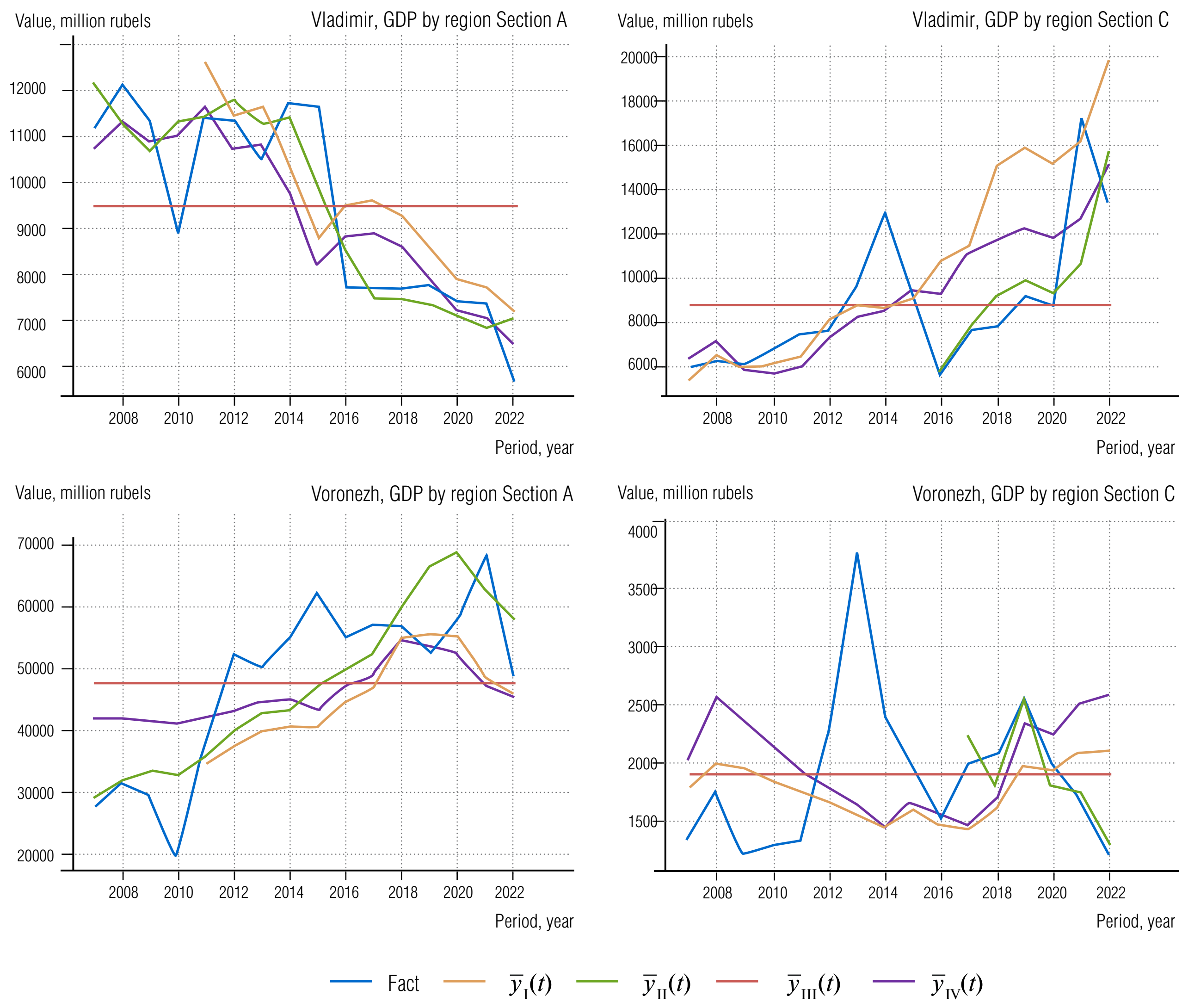

Fig. 2. Actual and normative values of GDP by region volume for sections A and C (Vladimir–Voronezh regions). Fact is the actual values of the indicator; $\overline{y}_{I}(t)$, $\overline{y}_{II}(t)$, $\overline{y}_{III}(t)$, $\overline{y}_{IV}(t)$ – are norms of types I, II, III and IV, respectively.

For all regions, the values of the norm of the second type are closer to the actual values of the others.

Figures 1 and *2* shows that sharp jumps in the actual GDP by region values for section C were observed in the Belgorod (2011), Vladimir (2014 and 2021) and Voronezh (2013) regions.

A similar situation was observed in the Kaluga (2015) and Kostroma regions (2020). Moreover, for the Kaluga region, the highest rates were observed in 2012. Three years before the jump in 2015 in the volume of GDP by region under section C. It can be assumed that the increase in the cost of the FA and the number of employees had a positive effect with a delay for this period.

In Moscow region, the value of the norm of the first type is significantly higher when compared with other regions of the Central Federal District. That is, a higher volume of GDP by region was expected, since the Moscow region has a large volume of fixed assets (in value terms), and the number of people employed in agriculture is commensurate with the number in the Voronezh and Belgorod regions. However, according to the actual data, it can be judged that the growth of the FA with a slight change in the number of employees from 2016 to 2021. It did not lead to a significant change in GDP by region, and therefore the increase in fixed assets turned out to be ineffective.

Throughout the entire assessment period (2007–2022), a similar pattern was observed in Tula Region.

In Yaroslavl region, from 2017 to 2022, there was a decrease in the rate of the first type, which is associated with a decrease in the cost of fixed assets (2017–2022) and the number of employees (2017–2021).

In addition, in 2022, in all regions, there was a decrease in the actual volume of GDP by region in sections A and C in relation to 2021, which is due, among other things, to foreign policy reasons. This was reflected in the values of norms of all types for most regions of the Central Federal District. That is, the developed norms are sensitive to the specific conditions of the functioning of the regions, which justifies the expediency of their use in assessing the performance of economic subjects.

The recommended types of norms for the Central Federal District regions for sections A (agriculture, forestry, hunting, fishing and fish farming) and C (mining) according to NACE 1 were determined in accordance with the proposed methodology for their selection using formulas (10) and (11) according to 2022 data. The results are presented in *Table 1.*

Thus, the method of forming norms for the results of the functioning of complex systems is presented, which makes it possible to take into account both specific conditions and priorities.

6. Discussion

The method proposed here can be used to evaluate, analyze and make decisions about the functioning of specific business entities: enterprises, organizations and their divisions operating under conditions of uncertainty and heterogeneity of information. Indeed, at present, issues related to the assessment and monitoring of indicators for which regulatory levels have not been defined have become relevant in the tasks of business and the economy as a whole.

These include, first of all, integral indicators that determine the state and dynamics of complex objects and systems; indicators that depend on many factors, some of which cannot be quantified and included in traditional models, as well as qualitative indicators. This makes it difficult to form standards. In addition, the information aspects of determining such norms are complicated due to significant uncertainty associated with inaccuracy, incompleteness and vagueness of the source data, their heterogeneity, spaciousness and time. Situational uncertainty inherent in modern business systems and economics also has a significant impact.

The following indicators can be given as examples.

The level of competitiveness of the enterprise. This complex non-quantitative indicator is determined, as a rule, by the rating of the enterprise in the relevant rating agencies. To determine it, agencies use quantitative information from the company's balance sheets, which is known to be incomplete and inaccurate in real practice. In addition, the standards used by the agencies are their own, calculated according to the methods developed by them. Therefore, competitiveness assessments are different and biased. In this regard, this paper proposes approaches, in particular the use of norms of the first type, which allow us to calculate norms that take into account the conditions in which both the assessed economic subjects and its environment (competitors) function.

Business efficiency. The classical approach to determining this indicator involves calculating key performance indicators (KPIs) expressed in quantita-

Table 1.

Ranks (priorities) and types of norms for the regions of the Central Federal District

N	Region	Section A			Section C		
		Priorities (rank)	Share in the GDP by region structure in the region	Type of norm	Priorities (rank)	Share in the GDP by region structure in the region	Type of norm
1	Belgorod	2	15.7%	$\overline{y}_{1,1,III}$	1	21.5%	$\overline{y}_{1,2,II}$
2	Bryansk	1	18.1%	$\overline{y}_{2,1,II}$	4	0.0%	–
3	Vladimir	4	3.0%	$\overline{y}_{3,1,IV}$	4	0.7%	$\overline{y}_{3,2,III}$
4	Voronezh	1	16.2%	$\overline{y}_{4,1,II}$	4	0.4%	$\overline{y}_{4,2,II}$
5	Ivanovo	4	3.0%	$\overline{y}_{5,1,IV}$	4	0.2%	$\overline{y}_{5,2,III}$
6	Kaluga	4	6.2%	$\overline{y}_{6,1,I}$	4	0.3%	$\overline{y}_{6,2,I}$
7	Kostroma	3	7.1%	$\overline{y}_{7,1,IV}$	4	0.2%	$\overline{y}_{7,2,III}$
8	Kursk	1	18.2%	$\overline{y}_{8,1,II}$	2	10.1%	$\overline{y}_{8,2,II}$
9	Lipetsk	3	10.0%	$\overline{y}_{9,1,III}$	4	0.6%	$\overline{y}_{9,2,III}$
10	Moscow	4	1.5%	$\overline{y}_{10,1,II}$	4	0.1%	$\overline{y}_{10,2,II}$
11	Oryol	1	23.5%	$\overline{y}_{11,1,II}$	4	0.1%	$\overline{y}_{11,2,I}$
12	Ryazan	3	9.8%	$\overline{y}_{12,1,IV}$	4	0.2%	$\overline{y}_{12,2,IV}$
13	Smolensk	4	4.0%	$\overline{y}_{13,1,IV}$	4	0.2%	$\overline{y}_{13,2,II}$
14	Tambov	1	30.3%	$\overline{y}_{14,1,II}$	4	0.0%	–
15	Tver	4	5.2%	$\overline{y}_{15,1,IV}$	4	0.2%	$\overline{y}_{15,2,IV}$
16	Tula	4	6.5%	$\overline{y}_{16,1,III}$	4	0.6%	$\overline{y}_{16,2,IV}$
17	Yaroslavl	4	4.0%	$\overline{y}_{17,1,IV}$	4	0.2%	$\overline{y}_{17,2,IV}$

tive form. However, not all indicators that determine business performance may have permanent norms. For example, an indicator of the sustainability of an enterprise both in terms of output and in terms of the number of employees. Such indicators are dynamic in nature, therefore, the standards for them should be determined depending on the current production situation. Thus, the number of employees in small and medium-sized enterprises may vary precisely to maintain a high level of business efficiency. The dynamic interval (soft) norms (norms of the fourth type) proposed in this work make it possible to determine this integral indicator more accurately and reliably.

The volume of sales. This indicator depends on numerous factors, such as seasonality, market situation, image characteristics of the properties of goods and enterprises, which are also not determined by a specific number, but can only be set within certain limits. In this case, interval dynamic norms, determined by the norm of the fourth type without weighing, can be very useful.

Customer satisfaction. This is a complex dynamic indicator that is strongly influenced by market conditions. Therefore, soft dynamic norms are also needed.

In the future, the proposed norms are supposed to be used in the construction of optimization models that make it possible to find the most profitable changes in factors from a business point of view, leading to an improvement in the goal indicators of the functioning of economic subjects This will enable managers at various levels to strengthen the rationale for their decisions.

Conclusion

This article presents a method for the formation of dynamic standards of complex systems, taking into account development priorities based on econometric modeling and the methodology of Bayesian intelligent measurements.

The application of this method for the regions of the Central Federal District according to data for 2007—2022 on the basis of the EFRA and Infoanalyst 2.0 software platforms confirmed the possibility of constructing norms (four types) taking into account factors affecting the results of the functioning of system elements, including in conditions of incompleteness and fuzziness of data.

The results of this study can be useful to regional governments in analyzing situations, forming socio-economic development programs in terms of setting targets for short and medium-term periods, as well as for synthesizing solutions, including using optimization models aimed at ensuring sustainable development of the state and its subjects. ■

Acknowledgments

The study was carried out at the expense of a grant from the Russian Science Foundation № 24-28-20020[1] and Tula Region.

References

1. Zhukov R.A. (2019) *Socio-ecological-economic systems: Theory and practice*. Moscow: INFRA-M (in Russian). https://doi.org/10.12737/monography_5b7516626665a8.43347695

2. Zhukov R.A. (2023) *Modeling the development of hierarchical socio-economic systems based on a multi-level optimization approach*. Moscow: Financial University under the Government of the Russian Federation (in Russian).

[1] https://rscf.ru/project/24-28-20020/

3. Andreev A.V. (2018) *Theoretical foundations of reliability of technical systems*. St. Petersburg: Publishing House of the Polytechnic University (in Russian).

4. Matushkina I.Y., Onishchenko L.A. (2018) *Technical regulation: technical regulations and standardization: a textbook*. Yekaterinburg: Ural University Publishing House (in Russian).

5. Belik A.G. (2018) *Quality and reliability of software systems: textbook*. Omsk: Publishing House of OmSTU (in Russian).

6. He D., Kang Hou K., Li X.X., Wu S.Q., Ma L.X. (2023) A reliable ecological vulnerability approach based on the construction of optimal evaluation systems and evolutionary tracking models. *Journal of Cleaner Production*, vol. 419, 138246. https://doi.org/10.1016/j.jclepro.2023.138246

7. Mezenina O.B. (2023) *Economics of environmental management: economic efficiency of environmental protection costs of an industrial enterprise: a textbook*. Yekaterinburg, UGLTU (in Russian).

8. Dmitriev V.V., Kaledin N.V. (2016) Russian Northwest: An integral assessment of the conditions of regional social, environmental and economic systems and quality of life. *Baltic Region*, vol. 8, no. 2, pp. 125–140 (in Russian). https://doi.org/10.5922/2074-9848-2016-2-7

9. Kleiner G.B., Rybachuk M.A. (2019) System balance of the Russian economy: Regional perspective. *Economy of Region*, vol. 15, no. 2, pp. 309–323 (in Russian). https://doi.org/10.17059/2019-2-1

10. Palash S.V. (2017) Structural balance of the economy: government programs for industrial development in the Russian Federation. *St. Petersburg State Polytechnic University Journal of Engineering Science and Technology*, vol. 10, no. 1, pp. 53–72 (in Russian). https://doi.org/10.18721/JE.10105

11. Sun X., Liu X., Li F., Tao Y., Song Y. (2017) Comprehensive evaluation of different scale cities sustainable development for economy, society, and ecological infrastructure in China. *Journal of Cleaner Production*, vol. 163, pp. 329–337. https://doi.org/10.1016/j.jclepro.2015.09.002

12. Makarov V.L., Aivazyan S.H., Afanasiev M.Yu., Bakhtizin A.R., Nanavyan A.M. (2014) The estimation ofthe regions' efficiency of the Russian Federation including the intellectual capital, the characteristics of readiness for innovation, level of well-being, and quality of life. *Economy of Region*, vol. 4, pp. 9–30 (in Russian). https://doi.org/10.17059/2014-4-1

13. Mikhalchuk A.A., Chistyakova N.O., Akerman E.A., Tatarnikova V.V. (2022) Evaluation methods of the technical efficiency of the Siberian Federal District regions using the production possibilities curve. *Vestnik NSUEM*, no. 4, pp. 76–91 (in Russian). https://doi.org/10.34020/2073-6495-2022-4-076-091

14. Mishra S.K. (2010) A brief history of production functions. *The UIP Journal of Managerial Economics, UIP Publications*, vol. 8, no. 4, pp. 6–34.

15. Zhukov R.A., Kozlova N.A., Manokhin E.V., Plinskaya M.A. (2022) Construction of an aggregated production function with implementation based on the example of the regions of the Central Federal District of the Russian Federation. *Business Informatics*, vol. 16, no. 3, pp. 7–23. https://doi.org/10.17323/2587-814X.2022.3.7.23

16. Zhukov R., Kuznetsov G., Manokhin E., Gorodnichev S., Plinskaya M. (2023) Balanced functioning of socio-economic systems: regional perspective. J*ournal of Law and Sustainable Development*, vol. 11, no. 12, e1922. https://doi.org/10.55908/sdgs.v11i12.1922

17. Pedroza C. (2006) A Bayesian forecasting model: predicting U.S. male mortality. *Biostatistics*, vol. 7, no. 4, pp. 530–550.

18. Long K.Q. (2021) Clustering lifestyle risk behaviors among Vietnamese adolescents and roles of school: A Bayesian multilevel analysis of global school-based student health survey 2019 / K.Q. Long, H.T. Ngoc-Anh, N.H. Phuong et al. *The Lancet Regional Health – Western Pacific*, vol. 15, 100225. https://doi.org/10.1016/j.lanwpc.2021.100225

19. Gallardo M. (2022). Measuring vulnerability to multidimensional poverty with Bayesian network classifiers. *Economic Analysis and Policy*, vol. 73, pp. 492–512. https://doi.org/10.1016/j.eap.2021.11.018

20. Prokopchina S.V. (2019) *Soft calculations and measurements. Methods, information technologies and means of intellectual information processing in digitalization tasks: monograph*. Moscow: Scientific Library (in Russian).

21. Prokopchina S.V., Shcherbakov G.A., Efimov Yu.V. (2019) *Modeling of socio-economic systems in conditions of uncertainty*. Moscow: Scientific Library (in Russian).

22. Berezin A.S., Zhukov R.A., Prokopchina S.V. (2022) Bayesian intellectual measurements of indices and indicators of regional availability of cultural objects. *Soft Measurement and Computing*, vol. 53, no. 4, pp. 5–15 (in Russian). https://doi.org/10.36871/2618-9976.2022.04.001

23. Zhukov R.A., Grigoriev E.V., Plinskaya M.A., Zhelunitsina M.A. (2022) On the dynamics of mortality in the Tula region. *Soft Measurement and Computing*, vol. 60, no. 11-2, pp. 30–38 (in Russian). https://doi.org/10.36871/2618-9976.2022.11-2.003

24. Zhukov R.A., Prokopchina S.V., Bormotov I.V., Manokhin E.V., Rudneva I.D. (2024) The model of the value system of Russian youth based on Bayesian intellectual measurements. *MIR (Modernization. Innovation. Research)*, vol. 15, no. 1, pp. 96–114 (in Russian). https://doi.org/10.18184/2079-4665.2024.15.1.96-114

25. Zhukov R.A. (2020) An approach to assessing the functioning of hierarchical socio-economic systems and decision-making based on the EFRA software package. *Business Informatics*, vol. 14, no. 3, pp. 82–95. https://doi.org/10.17323/2587-814X.2020.3.82.95

26. Zhukov R.A., Prokopchina S.V. *The Infoanalyst 2.0 software package. Certificate of state registration of the computer program No. 2024617544 dated 04.03.2024*. Available at: https://fips.ru/EGD/0b443bd3-982b-41db-a413-b1b035d9c596 (accessed 24 October 2024) (in Russian).

27. Zhukov R.A. (2021) Method for assessing the results of hierarchical socio-economic systems' functioning based on the aggregated production function. *Economics and Mathematical Methods*, vol. 57, no. 3, pp. 17–31 (in Russian). https://doi.org/10.31857/S042473880016428-9

28 Zhukov R.A., Kozlova N.O., Manokhin E.V., Myasnikova E.B., Melay E.A. (2023) Multi-criteria optimization as the methodology of ensuring sustainable development of regions: Tula region of the Russian Federation. *International Journal of Sustainable Development and Planning*, vol. 18, no. 4, pp. 1057–1068. https://doi.org/10.18280/ijsdp.180408

29. Prokopchina S.V. (2021) *Fundamentals of the theory of scaling in economics: textbook*. Moscow: Scientific Library (in Russian).

30. Federal State Statistics Service (2024) *Regions of Russia. Socio-economic indicators*. Available at: https://rosstat.gov.ru/folder/210/document/13204 (accessed 24 October 2024) (in Russian).

31. *Unified Interdepartmental Information and Statistical System (EMISS*). Available at: https://www.fedstat.ru (accessed 24 October 2024) (in Russian).

32. *Rating of socio-economic development of regions in 2022.* Available at: https://realnoevremya.ru/attachments/1760 (accessed 24 October 2024) (in Russian).

33. *The gross structure of the regional product in 2022*. Available at: https://www.fedstat.ru/indicator/59450 (accessed 24 October 2024) (in Russian).

34. *Tables of inflation*. Available at: https://уровень-инфляции.рф/таблицы-инфляции (accessed 24 October 2024) (in Russian).

35. Zhukov R.A., Plinskaya M.A., Manokhin E.V. (2022) Assessment of the regions functioning based on production functions with the above cost factors. *Journal of Applied Economic Research*, vol. 22, no. 3, pp. 657–682. https://doi.org/10.15826/vestnik.2023.22.3.027

36. *The results of model testing*. Available at: https://disk.yandex.ru/i/CJFtrRDuWYsdsg (accessed 24 October 2024) (in Russian).

About the authors

Roman A. Zhukov

Dr. Sci. (Econ.), Cand. Sci. (Phys.-Math.), Assoc. Prof.;

Researcher, Professor, Department of Mathematics and Informatics, Financial University under the Government of the Russian Federation, Tula Branch, 1a, Oruzheynaya Street, Tula 300012, Russia;

E-mail: pluszh@mail.ru

ORCID: 0000-0002-2280-307X

Svetlana V. Prokopchina

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

Professor, Department of Modeling and System Analysis, Faculty of Information Technology and Big Data Analysis, Financial University under the Government of the Russian Federation, 49, Leningradsky Prospekt, Moscow 125167, Russia;

E-mail: svprokopchina@mail.ru

ORCID: 0000-0001-5500-2781

Maria A. Plinskaya

Master Student, Financial University under the Government of the Russian Federation, Tula Branch, 1a, Oruzheynaya Street, Tula 300012, Russia;

E-mail: maria.plinskaya@gmail.com

ORCID: 0000-0002-1307-0935

Maria A. Zhelunitsina

Master Student, Financial University under the Government of the Russian Federation, Tula Branch, 1a, Oruzheynaya Street, Tula 300012, Russia;

E-mail: maria202001@yandex.ru

ORCID: 0009-0006-3129-2749

DOI: 10.17323/2587-814X.2024.4.61.80

Development of a high-level design of an analytical software complex for an enterprise that provides end-to-end planning

Natalya N. Seredenko
E-mail: Seredenko.NN@rea.ru

Plekhanov Russian University of Economics, Moscow, Russia

Abstract

Currently, many companies engaged in the production of medium-term turnover goods, are faced with the need to create a high-level design of a software complex that allows them to support a full cycle of sales planning, production, logistics and marketing campaigns. This is due to the economic development of the enterprise and the integration of independent software systems/modules that allow for the implementation of limited business processes without connection with related functions and business processes of the enterprise. Thus, enterprises find themselves in a situation where various departments have implemented a disparate set of information systems and software modules within which local accounting and analytics of various operations are carried out, while the software and analytical complex as a whole does not provide a complete, connected and cyclical planning process. This paper presents a model of requirements for functions, information objects and data flows, providing end-to-end planning, as well as an approach to identify missing objects of the existing information complex of the enterprise. An analytical network consisting of missing elements has been developed, taking into account the dependencies and interrelationships of information objects and software modules, which makes it possible to form a priority vector of the relative importance of software components. This vector represents a set of priorities for improvements to the enterprise software package and allows you to more effectively allocate the resources of the development team for the software implementation of missing functions, information objects and integration data flows between software modules.

Keywords: high-level design of an enterprise analytical software complex, end-to-end planning at the enterprise, data flow modeling, design of the enterprise information architecture, integration of enterprise software modules, functional composition of enterprise systems, Analytic Network Process

Citation: Seredenko N.N. (2024) Development of a high-level design of an analytical software complex for an enterprise that provides end-to-end planning. *Business Informatics*, vol. 18, no. 4, pp. 61–80. DOI: 10.17323/2587-814X.2024.4.61.80

Introduction

Companies engaged in the production of medium-term turnover goods are constantly looking for effective solutions to improve the accuracy of sales planning. The quality of planning depends not only on the analysis of the company's historical sales and market conditions, but also on the analysis of the activities of internal processes: production, marketing, logistics. In order for business plans and operational sales plans to correspond to the future reality as much as possible, it is necessary to take into account the processes of related functional divisions of the enterprise when planning [1–3]. Thus, the task of building a system of end-to-end planning at the enterprise is relevant, which represents the interrelation of processes reflected in the system of plans and reports and ensuring end-to-end planning of all production, economic, financial and other processes of the enterprise's activities [4, 5].

An information complex supporting end-to-end interconnected cyclic planning requires the implementation of a large number of complex functions, as well as a high level of integration solutions for the operational exchange of information objects between modules [6, 7]. At the same time, the life cycle of the development of the information complex of an enterprise often represents a disparate, time-spaced implementation of separate modules to solve local problems of a separate division. There are corporate information systems on the software market that contain modules for automating the activities of production, sales, marketing and logistics departments, as well as having a complete and consistent data model at the heart of their architecture which allows you to organize complex interconnected processes within a single platform. These are ERP systems, monolithic integrated solutions that are designed to connect different tasks and business processes. However, often, by the time a company realizes the need to build end-to-end interconnected processes and has the financial opportunity to implement large platforms, it already has a large information complex of various, loosely interconnected solutions. A one-time abandonment of all available software solutions in order to implement a single platform can lead to a halt in the organization's activities and is extremely difficult to implement. In addition, monolithic solutions are difficult to scale, they have weak fault tolerance and the risk of dependence on a single platform and technologies.

If a company has a specific strategy for the architectural development of its software package, the solution to the problem of building an information complex that provides end-to-end planning may have its own characteristics. As part of the current study, the question of the applicability of the most common software architectures to the effective solution of the problem of end-to-end planning is considered (*Appendix 1*). Based on this analysis it can be concluded that enterprises faced with the need to solve end-to-end planning tasks should develop an information complex in accordance with the principles of microservice or hybrid architecture [8, 9], or layered architecture. The use of cloud services is also possible, but the task of securing the internal contour of the enterprise and access to it from cloud services should be addressed as a matter of priority. If an enterprise has a disparate information complex by the time it realizes the need to solve the problem of end-to-end planning, it makes sense to simultaneously build a strategy for the development of the enterprise's information complex by choosing an architectural approach.

Theoretical research and practical developments in the field of end-to-end planning consider various approaches to the organization of the enterprise's software and analytical complex [10]. Most of them offer unified solutions, such as 1C:Integrated Automation [11], 1C:ERP Enterprise Management [4], Oracle Hyperion Planning [12, 13], IBM Planning Analytics [14], SAP Analytics [15]. However, scientific research on this topic does not examine the state of an organization when, at a certain point in development, various systems have already been implemented in key functional units in accordance with some selected architectural principles, and interaction between them has been set up.

This paper examines enterprises that are faced with the need to refine an existing software package in order to increase the efficiency of the end-to-end planning process, and using microservice, hybrid, layered architecture, or a disparate software package. The process of finalizing the information complex of an enterprise is the formation and implementation of a portfolio of IT projects, in which each project is the configuration of the missing interaction (integration) between the systems / modules of the enterprise software complex.

According to the classifier developed in this study [16], the modernization of this kind of state of the organization is classified as an optimization task of finding competitive solutions to improve the infrastructure of the enterprise. This implies highlighting a set of unrealized tasks, as well as determining the priorities according to which it is advisable to implement them. This task is relevant for the company, since it leads to the achievement of strategic goals.

1. Statement of the problem

The development of an IT strategy in an organization is a key process and is of particular interest due to the huge role and high level of digitalization of most modern companies. The object of the study is companies whose information complex consists of separate modules, the interconnection of which is realized, among other things, through the transfer of information objects. Each microservice (module) includes its own stack of functions, technologies, ways of organizing data and software interfaces, depending on the software architecture of the enterprise [17]. The goal-setting of these organizations within the framework of the development of the IT landscape in the context of supporting end-to-end planning implies the development of a plan for the implementation of coordinated improvements to each microservice (module).

It is worth noting that the conceptual scheme of end-to-end planning may vary depending on the economic, production, financial and other processes of the enterprise [4]. Based on statistical observation of the main directions [18 example of the goals set by the company's management, 19], we formulate a possible in organizing end-to-end planning, and the resulting business requirements for the corresponding software package (*Fig. 1*).

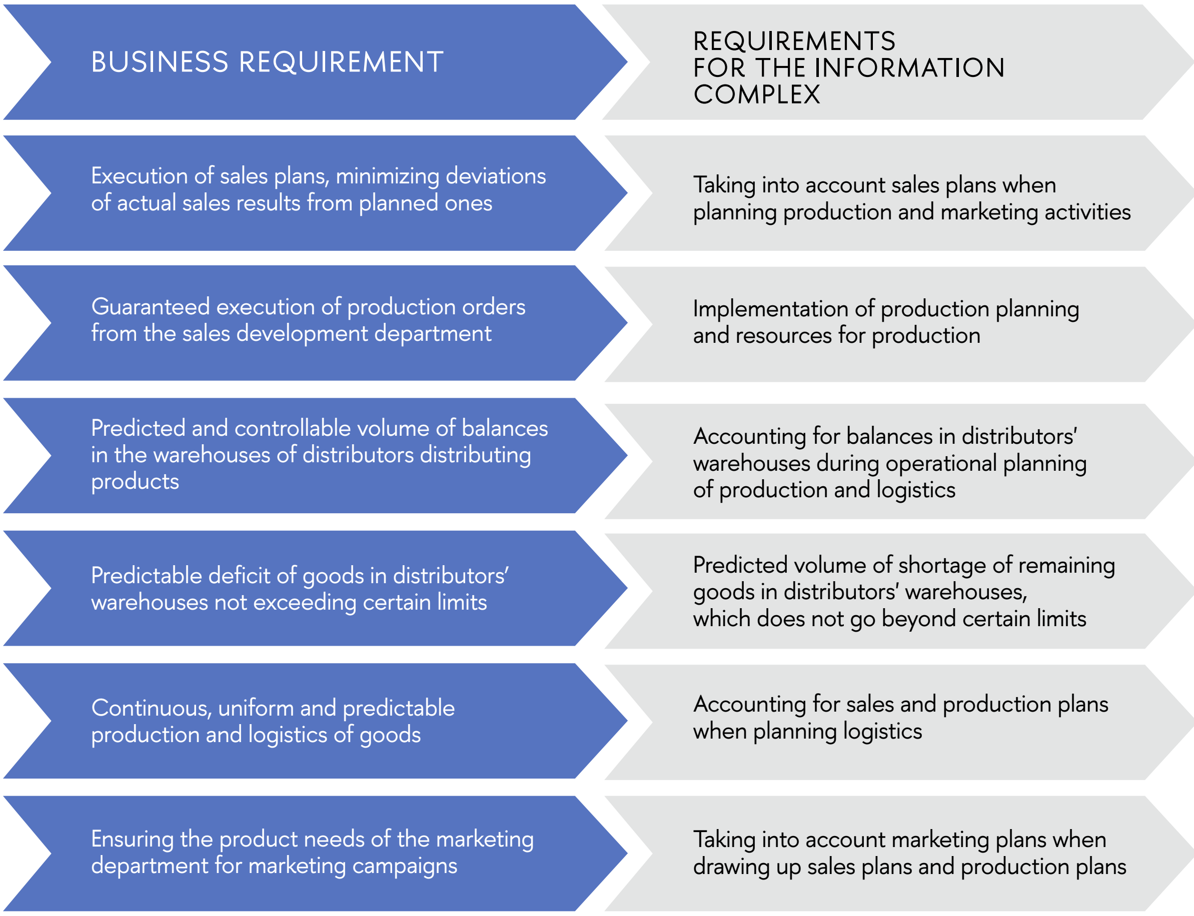

Fig. 1. Connection of the enterprise's business requirements in the field of end-to-end planning with the requirements for the information complex.

Achieving these goals is possible only if there is a software package that implements a full set of relevant functions, information objects and integrations.

In order to help organizations implement information technology development strategies that enable end-to-end planning, the current study proposes a methodology based on designing an effective high-level system design template and bringing the enterprise information complex to the developed template. According to the component methodology of business process reengineering proposed in the works by Telnov [20, 21], an enterprise system is a tuple of components $S = <G, E, E_n, T, F, R, Z>$. In relation to the area under consideration, let us clarify the definition of components:

G – a set of business goals related to the requirements for the information complex presented in *Fig. 1.* Within the scope of work, market and financial goals are relevant.

E – a set of information elements, namely, modules and systems that are part of the information complex that provides end-to-end planning.

E_n – a set of elements of the market and social environment. In the context of the problem being solved, they imply economic circumstances that force organizations to increase the efficiency of end-to-end planning in the enterprise.

T – a set of time periods characterizing the cyclical nature of end-to-end planning processes.

F – a set of functions included in the end-to-end planning process, implemented within the framework of many information modules.

R – a set of relationships that we will interpret as a set of interfaces (integration interactions) between information modules *E*. Each interface is a customized transfer of a separate information object from module E_n to module E_m.

Z – a set of patterns (strategies, methods) of the functioning of the information complex. As methods in the work, it is proposed to use the method of analytical networks by Saaty [22] for priority ranking of missing objects.

Thus, the statement of the problem within the proposed terms will sound as follows. We propose to identify a typical set of business processes (functions) F_{full}, divided into software modules E_{full}, and a complete typical set of transferred used data (information objects) R_{full} from the E_n module to the E_m, which the enterprise needs to organize end-to-end planning. Next, we propose to compare the actual set of implemented functions and transmitted information objects of the organization with the standard set of the model and select the set *F* and *R*, which represents the difference between the actual and standard set. As an example of a set *F*, one of the possible states of an enterprise's information infrastructure can be cited: lack of accounting and logistics planning functions; unrealized marketing planning function; formation of a production plan without taking into account the fact of sales. An example of a set *R* is the following set of missing information objects: lack of production plans in the logistics module, sales data in the marketing module, sales facts in the production module. Next, the task comes down to the implementation of the missing information objects in the corresponding modules, as well as the missing functions that use these objects. To increase the efficiency of implementation of this task, it is necessary to calculate the priorities of the portfolio of missing objects.

The stages of the described methodology are shown in *Fig. 2*.

2. Life cycle of the end-to-end planning process

In order to formulate requirements for data flows for a high-level design of an enterprise software and analytical complex that meets the stated goals, it is necessary to determine the requirements for the relationships between modules:

- Actual sales data must be taken into account when planning sales, marketing campaigns and production [23, 24].
- Plans for sales and marketing campaigns should be formed interdependently [25–27].
- Remains in the warehouses of departments performing the sales function (this can be their own chain of stores, sales through distributors or through any other sales channels [23]), must be taken into account when planning sales, planning production, planning logistics.
- Requests for production generated by the sales department must be included in plans for production orders, and the list of goods available for ordering for production must be available when generating requests for production [28–30].
- The balance of raw materials for production must be taken into account when planning production [28].
- The production plan should be taken as input in logistics planning [31].
- The fact of production and requests for delivery of goods must be taken into account when planning logistics.
- Additional applications for product release, generated by the marketing department based on the planning of marketing campaigns, must be taken into account when planning production [26, 32].

Thus, based on the listed requirements for data exchange between software modules, it is possible to create a data flow model that provides a sufficient

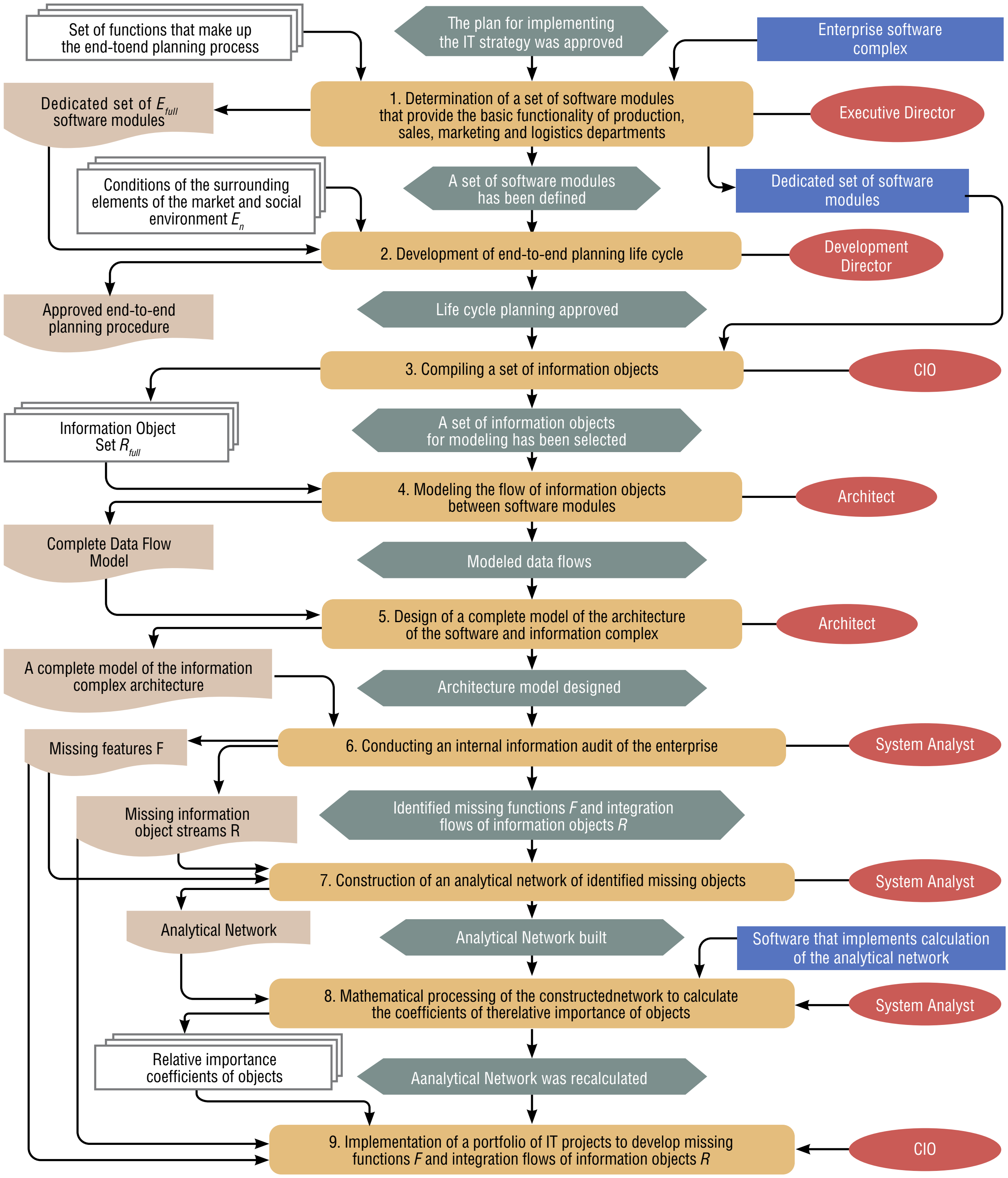

Fig. 2. Methodology for introducing architectural changes that ensure end-to-end planning into the enterprise information complex.

number of integration interactions to link all planning processes. The model of data exchange requirements is presented in *Fig. 3*.

In addition to the requirements for data exchange between modules, the model takes into account the requirements for the cyclicality and interconnectedness of business processes for sales planning, production, marketing campaigns and logistics.

3. Information architecture design

The previous sections described the prerequisites that lead large enterprises to carry out end-to-end planning of sales, production, marketing and logistics. A life cycle model of end-to-end planning is presented, and requirements for the software package are formulated. The current section contains the proposed model of a software and information complex that allows for a continuous planning process that links the activities of various divisions of the company, as well as providing a high level of analytical substantiation of plans.

In order to move from a data exchange model to a high-level system design model, a notation is needed that matches a set of requirements. Based on the analysis and generalization of existing types of diagrams and the experience of IT architects of various enterprises, we will highlight the requirements for notation based on the needs for solving the tasks:

1. The diagram must explicitly demonstrate the microservices of the enterprise, since they are the main structural elements of the software package.
2. The diagram should reveal the functional content of the modules.
3. Illustration of the complete set of information object flows between modules.
4. Demonstration that information objects belong to microservice databases. At the same time, it is important to understand in which microservices information objects are created, and in which they are used and not changed.
5. The diagram should allow you to demonstrate the nesting of modules.

A scheme that meets these requirements will make it possible to take into account key factors when making decisions on the development of a high-level design of an enterprise information complex and not to miss important entities. Such a representation will also make it possible to carry out an examination of the prioritization coefficients, which will be calculated as a result of applying the proposed methodology. This scheme can be a convenient auxiliary tool for managing and monitoring the implementation of a portfolio of projects to refine the information enterprise, since it corresponds to the language of both technical specialists and IT managers.

As a part of the study, existing notations were studied and popular solutions for visual modeling were considered. It was revealed that none of the structural diagrams, such as UML, STR, ERD, FDD [33, 34] satisfy all of the listed requirements. UML behavior diagrams [33] emphasize scenarios and business processes. Object-oriented UML and RUP diagrams do not demonstrate the structure of modules (microservices). IDEF and DFD [34, 35] emphasize system functions and the flow of information and physical objects, but provide little insight into the high-level design of the information complex. ARIS [36] focuses on event and function flows, but does not emphasize high-level systems design.

In order to satisfy the listed requirements for clarity and completeness, let's take the UML deployment component diagram as a basis: it implies a look at the enterprise architecture from the point of view of deploying modules (microservices). We propose the following structural additions to the diagram notation: we will include a list of implemented functions inside each module; we will add a list of specific transmitted information objects and the directions of their transmission; let's include an explicit indication of entities in the database and characteristics of master systems and recipient systems. We will call the resulting scheme "Information Architecture".

Let us formalize the proposed definition. Information architecture is a diagram that shows the architecture of a high-level design system, including nodes such as modules (microservices) of the information

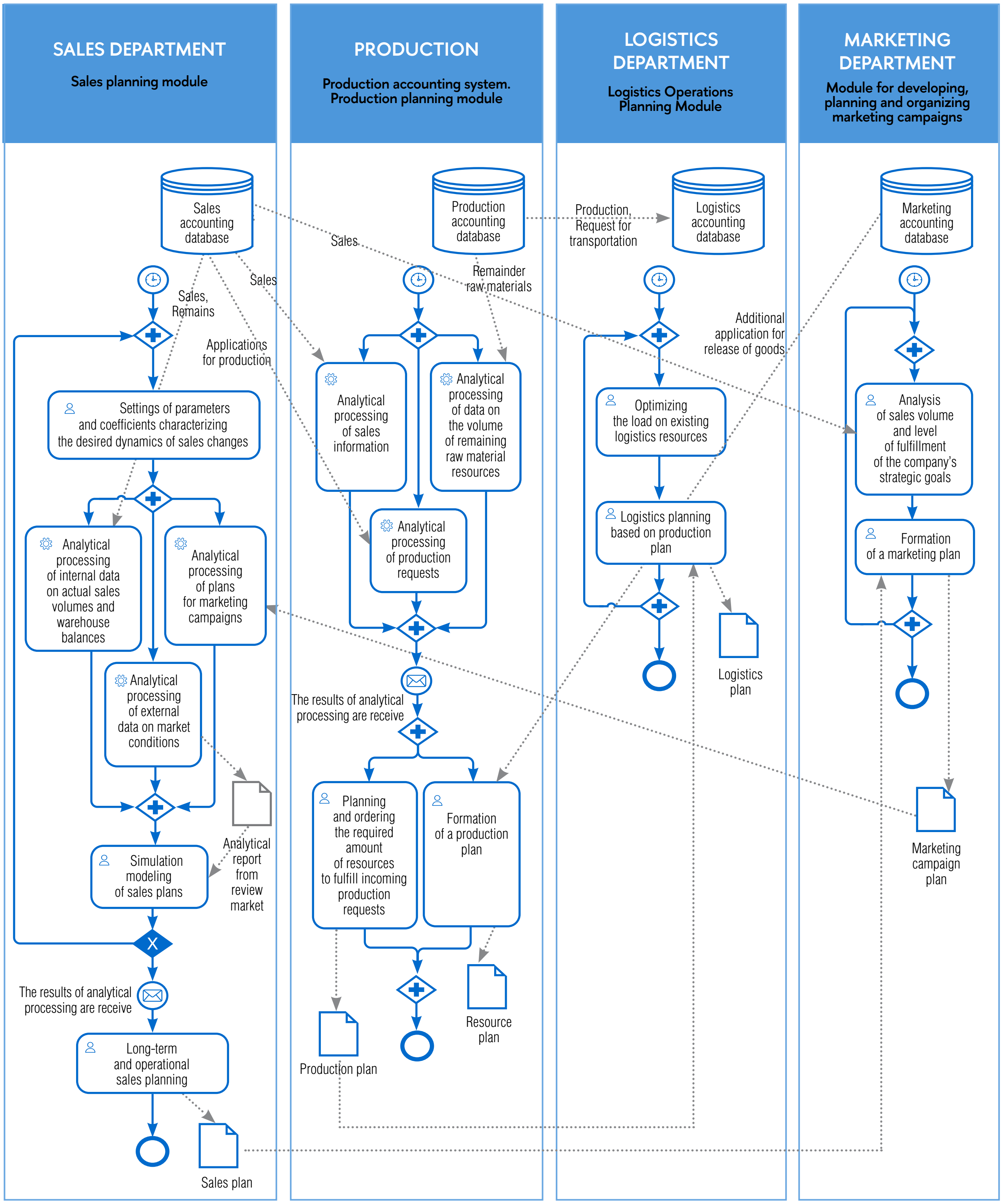

Fig. 3. Data flow model for high-level design of an enterprise software and analytical complex providing end-to-end planning.

complex; functions implemented inside the module; stored, processed and produced between measuring object modules.

Figure 4 suggests an information architecture model developed by the author which is an effective prototype for designing a software and information complex that ensures effective end-to-end planning in an enterprise.

Accompanying information on the revised notation:

1. Large rectangles represent systems/modules installed in the enterprise. It must be borne in mind that in real practice there is a different number of modules and distribution of functions between them. The proposed scheme is a unified average architectural solution.
2. Blocks with functions implemented within systems/modules. These functions are derived in section 3 based on the requirements for the information complex by enterprises producing goods with medium-term turnover.
3. Arrows emanating from some systems/modules included in other systems/modules are information flows transmitted within the framework of the described business process.
4. Databases in the form of "cylinders" and an indication of the information objects contained in these databases. It is necessary to keep in mind the following important factor: requirements for the integrity and a single version of master data at enterprises oblige strict monitoring of data source systems and data recipient systems. This aspect is important when developing integrations between systems, as well as when implementing an analytical warehouse and data model, on which analytical reporting is usually based. Therefore, this aspect is reflected in the information architecture, namely, information objects for which the system/module is the master system are depicted in a white rectangle inside the database, and objects for which the system/module is the recipient and within which these objects are not changed, are depicted in a gray rectangle inside the database.

Thus, we propose to include the following systems and modules in the information complex that ensures effective end-to-end planning:

1. **CRM system** (Customer Relationship Management). To support the production planning process, this system must contain two modules: *the Distributor Sales Accounting Module*, which collects data on distributor sales, and *the Production Order Formation Module*, which allows you to place product orders.

For the successful implementation of these functions in the modules of the CRM system, it is necessary to ensure the formation and transfer of the following information objects to consumer modules:

- Sales. Both sales to distributors and sales to end customers (through distributors or directly) can be accounted for and planned. To account for sales to end customers, separate CRM systems are being implemented that transmit data through distributors.
- Leftovers in distributors' warehouses.
- List of products available for production order.
- Application for production.

2. **Sales planning module**. This system is usually an analytical module that is part of the operating system, or a separate independent system, integration with which is configured through an ETL ("Extract, Transform, Load") solution. Taking into account the often-encountered complexity of the planning process in modern enterprises, we will assume that it is more correct to take this module into account as a separate independent module in the architecture requirements.

For the successful implementation of the listed functions in the *Sales Planning Module*, it is necessary to ensure the formation and transfer of the following information objects:

- Sales plans.
- Analytical summary with market analysis.

3. **Production accounting system**. This system is used to account for production processes, planning the volume of production of goods and raw materials. Unlike sales planning, which involves a large num-

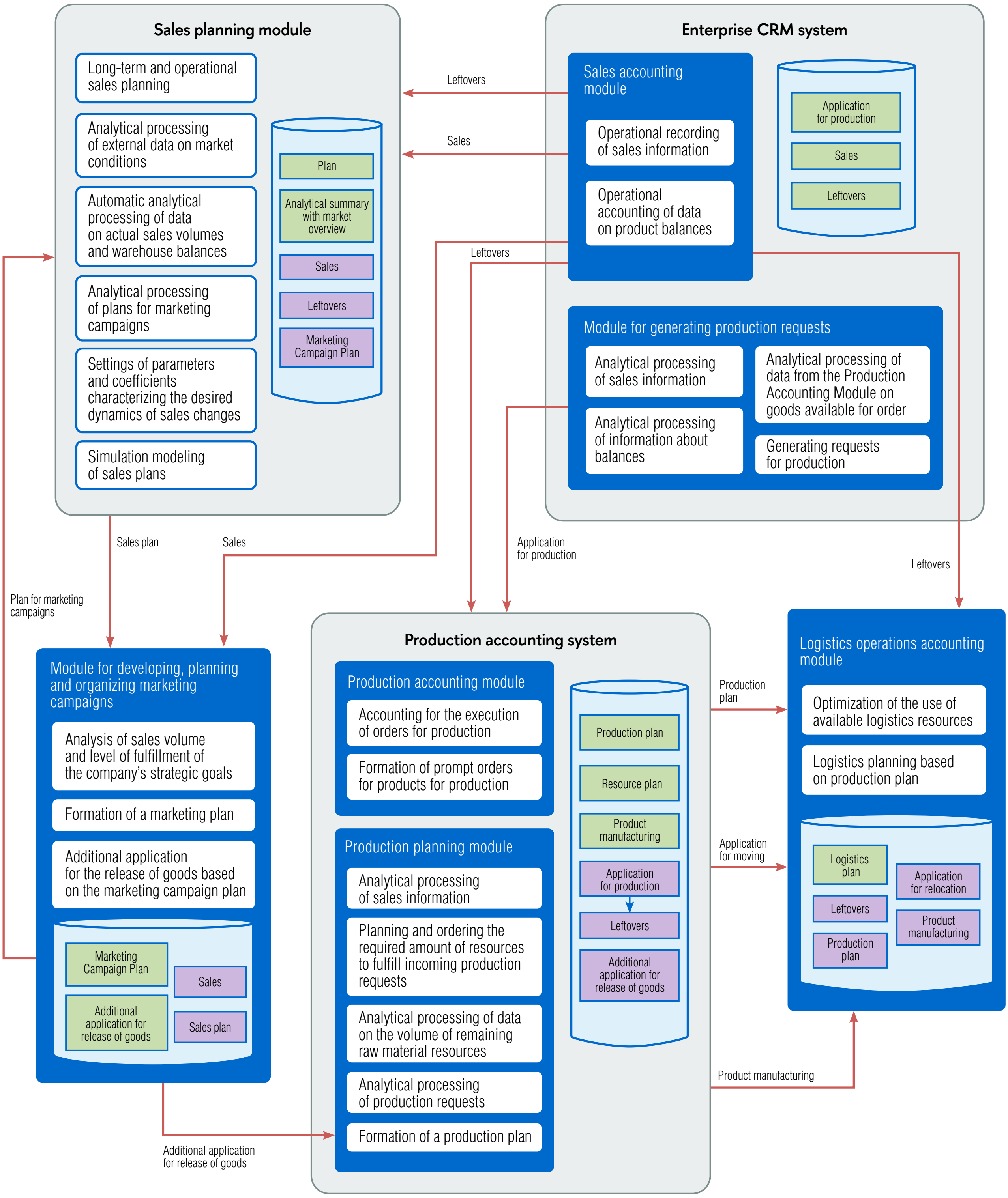

Fig. 4. Model of the architecture of a software and information complex that ensures effective end-to-end planning.

ber of factors, such as market dynamics, external analytical reports, simulation modeling and marketing campaign planning, production planning is primarily based on approved sales plans, and therefore planning is most often a module built into the Production Accounting System. Thus, it makes sense to present the architecture of the production accounting system with two built-in modules: *Production Accounting Module* and *Production Planning Module*.

For the successful implementation of these functions in the Production Accounting System, it is necessary to ensure the formation and transfer of the following information objects:

- Production plan.
- Ordering products for production.
- Ordering of raw materials for production.
- Production of products.
- The rest of the products in production.
- The rest of the raw materials for production.

4. **Logistics operations accounting module**. This module registers logistics movements in accordance with logistics plans, as well as optimizes the utilization of available logistics resources.

For the successful implementation of the listed functions in the *Logistics Operations Accounting Module*, it is necessary to ensure the formation and transfer of the following information objects:

- Logistics plan.
- Application for relocation.
- Moving.
- Balance in logistics.

5. **Module for the development, planning and organization of marketing campaigns**. The operations performed in the marketing department are directly related to the production and sales process, as they directly affect the structure and volume of goods sold. Sales planning cannot be carried out with a high degree of accuracy without taking into account the activities of the marketing department.

For the successful implementation of the listed functions in the *Module for the development, planning and organization of marketing campaigns*, it is necessary to ensure the formation and transfer of the following information objects:

- Plan for marketing campaigns.
- Additional application for the release of goods.

Thus, it can be concluded that the listed modules dealing with operational accounting and planning of the activities of these departments should be integrated with each other to transfer and receive information objects, as well as the successful implementation of functions using these objects.

4. Sequence of implementation of the IT project portfolio

The formulated requirements for the high-level design of a software-analytical complex that provides end-to-end planning in an enterprise clearly demonstrate the logical architecture. This architecture will provide a continuous and interconnected planning process for various divisions of the organization, taking into account factors and information objects arising from adjacent divisions, which ensures a high level of analytical justification and high accuracy of planning.

It is assumed that the enterprise has the opportunity to compare its current information architecture and business processes with the template diagram presented and, if necessary, launch projects to implement missing integrations between its current software modules. This will ensure the availability of the necessary information objects in the appropriate modules, as well as the integration into current business processes of new missing functions that use these information objects.

It must be borne in mind that information systems integration projects are labor-intensive and complex tasks, each of which requires the allocation of a large amount of resources and a high level of qualifications of team members [37, 38]. Therefore, after checking with the proposed basic layout and identifying weak points in the current information architecture and business processes, it is necessary to create a list of missing integrations and implementations of functions in the systems and prioritize them. Future projects to implement dedicated integrations and features should be carried out in accordance with certain priorities, as this will ensure a more optimal allocation of resources and higher efficiency.

Thus, the next management task that needs to be solved is ranking projects for the implementation of missing integrations/functions. The main requirement for ranking is to take into account the dependencies and interdependencies that exist between software modules and business process elements implemented in different modules. The Analytic Network Process (ANP), which is a continuation of the Analytic Hierarchy Process (AHP), developed by Saaty, has the ability to take these dependencies into account [22, 39]. In accordance with this method, it is necessary to build a network of information objects, software modules and their relationships that arise when developing a complex high-level design. Let's build a complete view of the network structure of information objects and program modules (*Fig. 5*).

Objects included in the analytical network will be interpreted as follows:

- An Alternatives Cluster is a complete set of Information Objects for which the priority of relative importance needs to be calculated. In accordance with these priorities, the management of the enterprise will be able to draw up a portfolio of IT projects for implementation.
- Arrows from software modules to information objects mean that the objects influence the modules. This influence is expressed in the implemen-

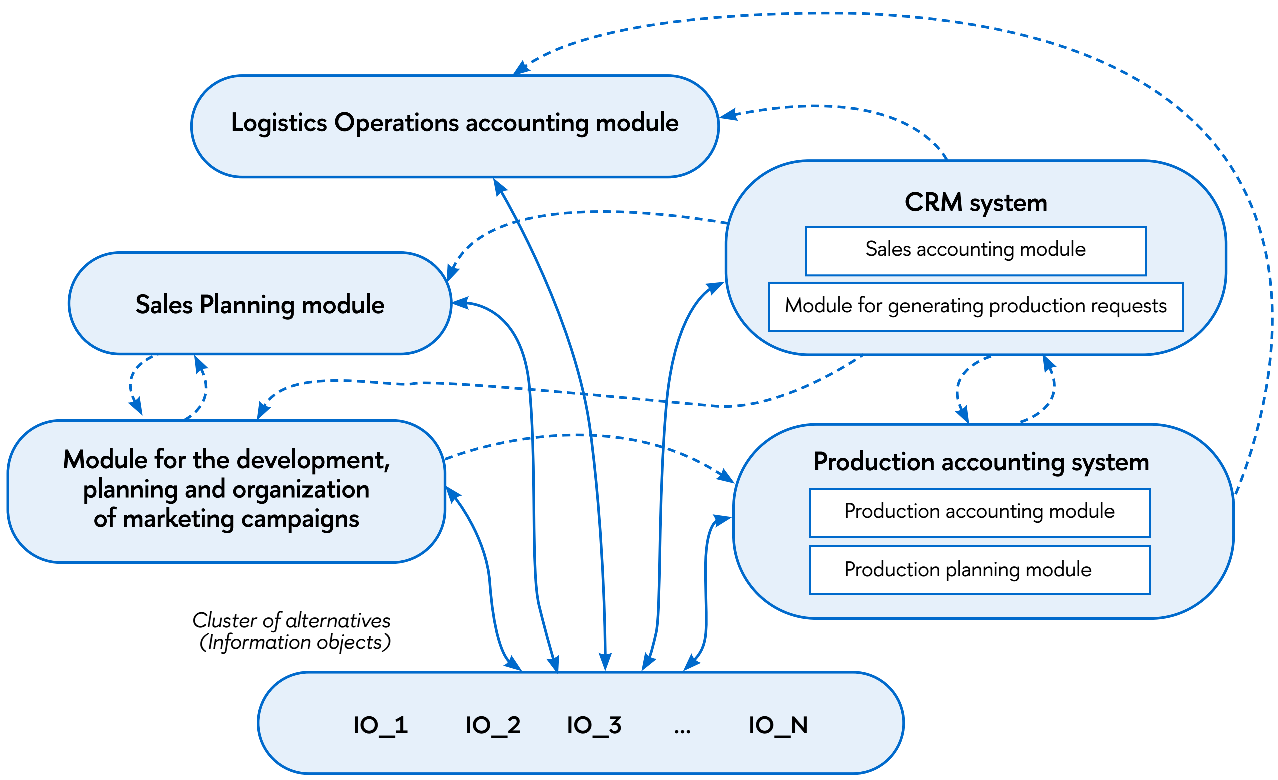

Fig. 5. Analytic network reflecting dependencies and relationships of program modules and information objects.

tation of new functions that can be implemented in these modules when new information objects appear in them.

- Arrows from information objects to software modules mean that the modules affect the objects. This impact is due to the complexity of implementing missing integrations and information object storage systems.
- Arrows between software modules indicate the flow of influence that software modules have on each other due to the presence of information flows between them.

Each dependence (each generated arrow) must be specified by a set of matrices of pairwise comparisons of the relative importance of the corresponding elements. When making pairwise comparisons, the following question is asked: "for a given network element and a pair of elements being compared, how much stronger is the influence of a given object from the pair on the element being evaluated compared to the other element?" Ratings are given on a nine-point scale, where a value of 1 point means that the compared elements are equivalent, and a score of 9 points means complete superiority.

Let $R = (R_1, R_2, ..., R_n, R_N)$ $n = (1, ..., N)$ – a set of alternatives (information objects), $E = (E_1, E_2, ... E_k, ..., E_K)$ $k = (1, ..., K)$ – a set of software modules within which it is necessary to make improvements to implement new business processes and integrations.

Relying on the structure of the constructed analytical network, we have the following types of dependencies:

- Dependencies of software modules relative to each other (type *I* dependencies, shown in *Fig. 5* with dotted arrows). When making pairwise comparisons of software modules relative to each other, the following question is asked: "which module software development has a higher priority for business at the present time?" Thus, we obtain a set of matrices of pairwise comparisons of relative importance $P^{I}_{(x,y)}$, where $x, y = (E_1, ..., E_K)$.
- Dependencies between alternatives and software modules:
  - Dependencies between software modules, taking into account the influence of information objects on them (Type *IIa* dependencies, shown in *Fig. 5* with solid arrows coming from a cluster of alternatives). When making pairwise comparisons of modules regarding the impact of alternatives on them, the following question is asked: "for this information object (alternative) in which module is it more important to implement integration for the subsequent implementation of missing functions in this module?". Thus, we obtain a set of matrices of pairwise comparisons of the relative importance of modules, the elements of which are denoted as $P^{IIa}_{n(p,t)}$ – is the result of comparing the relative importance of modules $p, t = (E_1, ..., E_K)$ relative to the importance of implementing the information object $n = (R_1, ..., R_N)$.
  - Dependencies between information objects (alternatives), taking into account the influence of software modules on them (dependencies of type *IIb*, shown in *Fig. 5* with solid arrows belonging to the cluster of alternatives). When making pairwise comparisons of alternatives regarding the influence of software modules on them, the following question is asked: "for this module, which information object (alternative) is more important to implement?". Thus, we obtain a set of matrices of pairwise comparisons of the relative importance of alternatives, the elements of which we denote as $P^{IIb}_{k(m,l)}$ – is the result of comparing the relative importance of alternatives $m, l = (R_1, ..., R_N)$ relative to the importance of implementing them in the software module .

Next, for a matrix of pairwise comparisons of the relative importance of software modules, the eigenvector $Z_{\text{own}.E} = (Z_{\text{own}.1}, ..., Z_{\text{own}.e}, ..., Z_{\text{own}.E})$, corresponding to the maximum eigenvalue of the matrix. A general view for calculating the eigenvector based on the definition of the eigenvector:

$$P^{I}_{(x,y)} \cdot Z^{I}_{(\text{own}.E)} = \lambda^{I}_{\max E} \cdot Z^{I}_{(\text{own}.E)}. \quad (1)$$

The elements of the resulting vector are transformed according to the following rule:

$$w_m^I = \frac{Z^I_{(\text{own.}\,m)}}{\sum_{m=1}^{M} Z^I_{(\text{own.}E)}}. \quad (2)$$

The vector $W^I = (w_1, \ ..., w_E)$ is a vector of coefficients of relative importance of software modules. Similarly, eigenvectors are calculated for matrices of pairwise comparisons of type *IIa* and type *IIb*. A general view for calculating eigenvectors:

$$P^{IIa}_{n(p,t)} \cdot Z^{IIa}_{\text{own.}\,n(p,t)} = \lambda^{IIa}_{\max n(p,t)} \cdot Z^{IIa}_{\text{own.}\,n(p,t)}; \quad (3)$$

$$P^{IIb}_{k(m,l)} \cdot Z^{IIb}_{\text{own.}\,k(m,l)} = \lambda^{IIb}_{\max k(m,l)} \cdot Z^{IIb}_{\text{own.}\,k(m,l)}. \quad (4)$$

The vector elements are subject to the following transformation:

$$w^{IIa}_{n(p,t)} = \frac{z^{IIa}_{\text{own.}\,n(p,t)}}{\sum_{m=1}^{M} z^{IIa}_{\text{own.}\,n(p,t)}}; \quad (5)$$

$$w^{IIb}_{k(m,l)} = \frac{z^{IIb}_{\text{own.}\,k(m,l)}}{\sum_{m=1}^{M} z^{IIb}_{\text{own.}\,k(m,l)}}. \quad (6)$$

The resulting vectors w_m^I, $w^{IIa}_{n(p,t)}$, $w^{IIb}_{k(m,l)}$ are further grouped into the supermatrix (the concept of a supermatrix was introduced in [22]). The coefficients of the relative importance of software modules will serve to weigh the blocks of the supermatrix. Next, the supermatrix is raised to the limiting power until the result stabilizes:

$$P^{\lim}_{SuperMatr} = \lim_{k\to\infty} \frac{1}{N} \sum_{k=1}^{N} P^{k}_{SuperMatr}. \quad (7)$$

The values of the desired coefficients of the relative importance of information objects will be calculated in the appropriate blocks and can be used when planning resources for the implementation of missing software improvements.

For the practical application of the proposed methodology in the enterprise, various software can be used that implements the method of analytical networks. In the work of Latypova, a comparative analysis of software tools implementing AHP and ANP was carried out [40]. As part of the current work, the free educational software SuperDecisions was used [41, 42].

Conclusion

The paper presents an approach to the development of a high-level design of a software and analytical complex that provides end-to-end enterprise planning. A step-by-step methodology for introducing architectural changes is proposed, which will help enterprises engaged in the production of goods with medium-term turnover to carry out an internal audit of the software package that provides key functions of the enterprise, and to launch a portfolio of improvements that will ensure the implementation of the identified missing functions.

It should be noted that the proposed high-level design is a generalization and supports the average process of end-to-end production planning at a logical level. In real practice, deviations from the presented architecture may occur. These deviations are due to different initial levels of informatization of enterprises, differences in software systems, financial resources available to the enterprise, the choice of specific software solutions for each module, and other factors that are individual for each specific company. However, the architecture developed within the framework of the article is proposed to be used as a basis for conducting an information audit in order to optimize end-to-end planning business processes.

The following areas can be identified as directions for further development of the methodology:

- expansion of the functional composition of program modules, as well as development of the attribute composition of the described information objects;
- use of the described approach to evaluate the software implementation of missing functions in the enterprise software package;
- development of an approach to assessing the effectiveness of proposed improvements to the information complex;
- development of the proposed analytical network in accordance with a methodology based on an assessment of benefits, opportunities, costs and risks;
- generalization of the proposed approach to other key functions of the enterprise.

References

1. Yankovskaya V.V. (2023) *Planning at an enterprise*. Moscow: INFRA-M (in Russian).
2. Ilyin A.I. (2014) *Planning at an enterprise*. 9th ed. Moscow: INFRA-M; Minsk: New Knowledge (in Russian).
3. Telnov Yu.F., Kazakov V.A., Danilov A.V. (2021) Technology for designing innovative processes for creating products and services of a network enterprise using a knowledge-based i4.0 system. *Business Informatics*, vol. 15, no. 4, pp. 76–92. https://doi.org/10.17323/2587-814X.2021.4.76.92
4. Danilochkina N.G. (2019) Concept of end-to-end planning at industrial enterprises. *Current Problems of Socio-economic Development of Russia*, no. 4. pp. 38–41 (in Russian).
5. Hahn D. (1999) *Planning and control. The concept of controlling*. Moscow, Finance and Statistics (in Russian).
6. Bragina A.V., Vertakova Yu.V., Evchenko A.V. (2020) Development of end-to-end technologies for planning the activities of an industrial enterprise in the context of digitalization of the economy. *Production Organizer*, vol. 28, no. 1, pp. 24–35 (in Russian).
7. Minaev V.A., Mazin A.V., Zdiruk K.B., Kulikov L.S. (2019) Digital twins of objects in solving control problems. *Radio Industry*, vol. 29, no. 3, pp. 68–78 (in Russian). https://doi.org/10.21778/2413-9599-2019-29-3-68-78.
8. Kopeliovich D.I., Kurguz M.A., Lebedev V.V. (2022) Microservice architecture as a type of service-oriented architecture. *Scienceosphere*, no. 4-2, pp. 230–235 (in Russian).
9. Ibatulin M.Yu., Terentyev V.A. (2018) Application of microservice architectures for building enterprise business architecture in the era of digital transformation. Proceedings of the *XXI Russian scientific conference on Enterprise Engineering and Knowledge Management (EEKM 2018). Moscow, April 26–28, 2018*. Vol. 1, pp. 376–381 (in Russian).
10. Ford N., Richards M. (2020) *Fundamentals of software architecture: An engineering approach*. O'Reilly Media.
11. Kurzova Yu.A. (2022) Overview of the 1C program: Integrated automation. *1C-Business Architect*. Available at: https://www.1ab.ru/blog/detail/1s-kompleksnaya-avtomatizatsiya-obzor/ (accessed 06 July 2024) (in Russian).
12. *Certificate of state registration of a computer program No. 2022665892 Russian Federation. SSC. System-2 of business planning and budgeting based on the Oracle Hyperion Planning 11.1.2.4*, software package No. 2022664704, application 08/03/2022, publ. 08/23/2022; applicant Joint Stock Company "Siberian Service Company (in Russian).
13. Minyailo A.S. (2020) Study of information risks of Russian industrial companies in budget process management systems. *Corporate Economics*, no. 4(24), pp. 4–21 (in Russian).
14. Pirogov M.V., Abdulganiev A.N., Martinovich D.A. (2022) Application of multidimensional data warehouse technology (OLAP) based on the system for managing financial performance and enterprise activity IBM Planning Analytics. *Regional Economy Issues*, no. 4(53), pp. 144–162 (in Russian).
15. Nazarov D.M., Nazarov A.D., Kovtun D.B. (2020) Building technology and predictive analytics models in the SAP analytic cloud digital service. Proceedings of the *2020 IEEE Conference on Business Informatics (CBI 2020). Antwerp, June 22–24 2020*. Vol. 2. pp. 106–110 (in Russian). https://doi.org/10.1109/CBI49978.2020.10067
16. Bragina A.V. (2021) *Technological modernization of an industrial enterprise using end-to-end planning technology*. Dissertation of a Cand. Sci. (Econ.). Kursk (in Russian).
17. Smirnov M. (2017) Microservice architecture in the corporate IT landscape. *Open Systems. DBMS*, no. 4, pp. 38–41 (in Russian).
18. Terentyev A.V., Polyakov Yu.N. (2020) Concept of design and construction of the basic algorithm of the business planning process in small and medium-sized businesses. *Management: Theory and Practice*, no. 1-3, pp. 99–105 (in Russian).
19. Shataeva O.V., Akimova E.N., Nikolaev M.V. (2021) *Economics of an organization (enterprise)*, 2nd ed. Moscow, Berlin: Direct-Media (in Russian).

20. Telnov Yu.F. Fedorov I.G. (2017) *Enterprise engineering and business process management. Methodology and technology*. Moscow: UNITY-DANA (in Russian).

21 Telnov Yu.F. (2005) *Business process reengineering: Component technology* (2nd ed.). Moscow: Finance and Statistics (in Russian).

22. Saaty T.L. (2015) *Decision making with dependence and feedback: The Analytic Network Process*. Moscow: LENAND (in Russian).

23. Gusakov I.V. (2014) *Analysis and planning of sales in FMCG market companies*. Moscow: Nobel Press (in Russian).

24. Lehmann D.R. (2017) *Product management*. Moscow: UNITY-DANA (in Russian).

25. Zhukova T.N. (2021) *Management and organization of marketing activities*. Moscow: INFRA-M (in Russian).

26. Kotler P., Saunders D., Armstrong G., Veronica V. (2007) *Fundamentals of marketing*. Moscow: Williams (in Russian).

27. Toymentseva I.A., Mikhailov A.M. (2015) The influence of marketing and management decisions on the process of optimizing the enterprise budget. *Bulletin of the Samara Economic University*, no. 12(134), pp. 16–19 (in Russian).

28. Smurnov E.S. (2010) *Automation of material accounting*. Moscow: Laboratory of Books (in Russian).

29. Semenova N.V., Baygulova A.A. (2014) *Fundamentals of production management: electronic training course*. Ulyanovsk: UlSU (in Russian).

30. Farakhutdinov Sh.F. (2021) *Modern trends and innovative methods in marketing research*. Moscow: INFRA-M (in Russian).

31. Gorsky M.A., Khalikov M.A. (2020) Models and methods for assessing the optimal size of an enterprise’s production segment. *Bulletin of the Altai Academy of Economics and Law*, no. 1-1, pp. 23–32 (in Russian).

32. Oyner O.K. (2022) *Marketing performance management*, 2nd ed. Moscow: Urait (in Russian).

33. Petrova I.R., Fakhrtdinov R.H., Suleymanova A.A., Razzhivin I.O., Fazulzyanov A.G. (2018) *Methodology of object-oriented modeling. The UML language*. Kazan: Kazan University (in Russian).

34. Freeman E., Robson E., Sierra K., Bates B. (2024) *Head First design patterns*. Saint Petersburg: Piter (in Russian).

35. State Standard of Russia (2000) *Methodology of functional modeling IDEF0. The guidance document*. Moscow.

36. Nazarova O.B., Novikova T.B., Maslennikova O.E. (2023) *ARIS: Theory and practice of business modeling*. Moscow: FLINT (in Russian).

37. Pervukhin D.V., Isaev E.A., Rytikov G.O., Filyugina E.K., Airapetyan D.A. (2020) Comparative analysis of theoretical models of cascade, iterative and hybrid approaches to managing the life cycle of an IT project. *Business Informatics*, vol. 14, no. 1, pp. 32–40. https://doi.org/10.17323/2587-814X.2020.1.32.40

38. Matveev A.A., Novikov D.A., Tsvetkov A.V. (2005) *Models and methods of project portfolio management*. Moscow: PMSOFT (in Russian).

39. Saaty T.L. (1993) *Decision making. The Analytic Hierarchy Process. Moscow: Radio and Communications* (in Russian).

40. Latypova V.A. (2018) A comparative analysis and a choice of tools implementing analytic hierarchy process. *Modeling, Optimization and Information Technologies*, vol. 6, no. 4(23), pp. 322–347 (in Russian). https://doi.org/10.26102/2310-6018/2018.23.4.024

41. Saaty R.V. (2016) *Decision making with dependencies and feedbacks. A tutorial for SuperDecisions software*. Pittsburgh: Foundation for Creative Solutions.

42. Creative Decisions Foundation (2024) *Super Decisions software*. Available at: https://www.superdecisions.com/ (accessed 06 July 2024).

Appendix 1.

Analysis of the applicability of software architecture for solving end-to-end planning problems

Architecture	Researches	Advantages	Disadvantages	Applicability for solving end-to-end planning problems
Monolithic architecture	Edsger Dijkstra formulated the principles of structural programming in the 1970s and 1980s. With the advent of personal computers and the development of the graphical interface, software has become more accessible and widespread. In the 1990s, monolithic architecture became the dominant approach for application development. The principles of monolithic architecture did not develop as a separate concept, but rather were applied and improved in the context of software development and evolution and were the prevailing approach. The very concept of «monolithic architecture» appeared later, when alternative approaches began to develop.	• Effective for small and simple systems • Easy to deploy • Stable operation	• Difficult to scale • Have weak fault tolerance • The risk of dependence on a single platform and technologies	Monolithic complexes are applicable to solving end-to-end planning problems only for small organizations in which the approach to planning is poorly developed and a small number of factors are taken into account when planning. Monolithic complexes can provide a certain process within their framework, but are extremely difficult to develop and integrate with external modules.
Event-driven architecture (EDA)	The development of event-driven architecture is related to object-oriented programming (EDA). In the 1970s, the concept was developed by Alan Kay and David Parnas. Further in the 1980s, Björn Stroustrup combined elements of procedural programming with an object-oriented approach. In the early 2000s, OOP was actively developing, and microservice architecture began to develop based on these principles.	• High overall adaptability • High performance • Scalability	• Difficult to implement due to asynchronous, distributed nature • It is necessary to solve problems of accessibility of remote processes, • The need to build logic for reconnecting the broker • Lack of atomic transactions for one business process	Event-driven architecture of information systems is used for basic enterprise tasks related primarily to operational accounting and building business processes. In order to design an information complex that allows solving end-to-end analytical planning problems, it is necessary to use approaches focused on the transfer and use of information objects, rather than on processing events/operations.

Architecture	Researches	Advantages	Disadvantages	Applicability for solving end-to-end planning problems
Microservice architecture	Microservice architecture is based on the concepts of Service Oriented Architecture (SOA). Microservices architecture is the result of the collective effort of many scientists, engineers and practitioners who contributed to its development. The main ideas were laid down in 2014 by Robert Martin (formulated the principles of modularity and independence in software development), James Lewis and Martin Fowler (defined the basic principles of this architecture), Fred George (proposed the use of the concept of "services" in the context of distributed systems), Dan Rosen (formulated the principles of "self-directed teams", which are an important element of microservice architecture). Modern Russian research on the topic of using microservice architecture belongs to Yu.F. Telnov, D.I. Kopeliovich, M.A. Kurguz, V.V. Lebedev, M.Yu. Ibatulin, V.A. Terentyev.	• Architecture provides flexibility and scalability • Independent deployment • Increased fault tolerance	• Challenges of development and deployment complexity • Increased control complexity • The need to implement interservice interaction	Microservice architecture is well suited to solve the end-to-end scheduling problem. It is within the framework of microservice architecture that it becomes possible to develop the complex in any direction, which means that the task of prioritizing the order of software development of new services and transferred information objects to ensure a comprehensive architecture is relevant.
Hybrid architecture	The concept of «hybrid architecture» arose as a result of a natural evolutionary process associated with the emergence of new technologies and an understanding of the advantages and disadvantages of both monolithic and microservice architectures. The beginning of the formation of hybrid approaches can be considered the late 1990s and early 2000s, when new technologies and architectural patterns began to appear: the spread of the Internet, the emergence of distributed systems, the development of containerization technologies. Important contributions to these methodologies were made by Ernest	• Allows for a gradual transition to microservices • Provides a flexible choice of optimal solutions for different parts of the system • Combines the advantages of monolithic and	• Difficult to control • Complexity of technical implementation • Compatibility issues between the components of the complex	Companies in a state of hybrid transition (either in a state of transition from one solution to another, or introducing new business models that are implemented based on microservice approaches and which integrate with monolithic applications) must solve the problem of prioritizing the development of modules for new services and order development of integrations for the exchange of information objects between systems.

Architecture	Researches	Advantages	Disadvantages	Applicability for solving end-to-end planning problems
	Clarke, who proposed the concept of "service orientation" in the 1990s, Martin Fowler, who identified SOA as an important design pattern in the early 2000s, and Jim Brandel developed the concept of "containers" in the 2010s-2020s.	microservice architectures		
Cloud computing	In the early 1980s, Ken Thompson created the first version of the multi-user UNIX operating system. Then virtual machines appeared, providing platform-independent programming environments. The term «cloud computing» appeared in the 1990s in the specifications of Compaq and Apple. In the first half of the 2000s, there was a tendency to transfer local software to cloud programs operating on the SaaS principle – Software as a Service.	• Flexibility • Automatic scalability • Availability of resources on demand • Reduced infrastructure costs	• Heavy dependence on cloud service provider • Limitations in the choice of technologies • Difficulty in debugging and data security aspects	Integration of enterprise information systems with modules deployed in the cloud is always a complex project, which, in addition to technical issues, involves solving security issues. To carry out integration, it is necessary to provide access to the enterprise's internal network, which imposes additional risks. If enterprises deploy solutions in the clouds, integration issues are resolved separately and have their own limitations, and therefore the set of integration tasks of cloud services is difficult to consider in a single portfolio with other integrations.
Layered architecture	Components within the layered architecture model are organized into horizontal layers, each of which performs a specific role in the application. Most layered architectures consist of four standard layers: presentation layer, business logic layer, data access layer, and database abstraction layer.	• Separation of tasks between components • Each layer of the layered architecture pattern has a specific role and responsibility in the application	• Failure to document the system and communicate which layers are open and which are closed leads to tightly coupled and fragile architectures that are difficult to deploy and maintain.	The ability of each module to interact with any other module without a clear structure and hierarchy can lead to a complex and confusing network of connections. In particular, cyclic dependencies may appear, which can lead to endless recursion and mutual locking. This architecture is characterized by high connectivity and complexity of interactions between components. This makes any changes problematic, as it is not completely cleaFor layered architecture, techniques are especially relevant to take into

Architecture	Researches	Advantages	Disadvantages	Applicability for solving end-to-end planning problems
	In the 1980s, design patterns emerged, for example, that implement layer structure in applications. In the 1990s, layered architecture became a popular approach for developing various software systems, including web applications and enterprise systems. In the 2010s, with the advent of new technologies such as web services, REST APIs, microservices, layered architecture became even more flexible and extensible.	• Changes made to one layer of the architecture generally do not affect components in other layers	• Layered architecture tends to create monolithic applications, even if the presentation layer and business logic layer are separated into separate deployable units • Low productivity	account the influence of components on each other and the dependencies between them, as well as to prioritize the software development of new dependencies and business logic taking into account these dependencies.

About the author

Natalya N. Seredenko

Cand. Sci. (Econ.);

Associate Professor, Department of Applied Informatics and Information Security, Plekhanov Russian University of Economics, 36, Stremyanny lane, Moscow 115054, Russia;

E-mail: Seredenko.NN@rea.ru

ORCID: 0009-0009-8113-0758

DOI: 10.17323/2587-814X.2024.4.81.97

Demystifying the digital transformation of the real estate brokerage industry in China: A case study of Lianjia (Beike)

Fengchen Wang [a, b]
E-mail: fwang@hse.ru

Olga V. Stoianova [a]
E-mail: ostoyanova@hse.ru

Angel Barajas [c]
E-mail: abarajas@uvigo.gal

[a] St. Petersburg School of Economics and Management, HSE University, Saint Petersburg, Russia
[b] Université Côte d'Azur, Nice, France
[c] University of Vigo, Ourense, Spain

Abstract

The rapid digital transformation of the real estate brokerage industry in China has revolutionized traditional business models, with Lianjia (Beike) at the forefront of this shift. This study explores Lianjia's journey from a conventional brokerage firm to a leading digital platform, analyzing the strategic digitalization of housing data through the creation of the Housing Dictionary, the development of the agent cooperation network (ACN), and the implementation of a management information systems (MIS) based Offline-to-Online (O2O) business model. Through a qualitative case study methodology, this research highlights how Lianjia's innovative use of technology has enhanced operational efficiency, customer satisfaction and competitive advantage. The findings provide valuable insights into the potential of digital platforms to drive continuous innovation and transformation in the real estate industry. This study also discusses the broader implications for the digital economy and offers recommendations for businesses aiming to undergo similar transformations.

Keywords: digital transformation, real estate brokerage, platform ecosystem, digital platform, Chinese housing market, case study, Lianjia (Beike)

Citation: Wang F., Stoianova O.V., Barajas A. (2024) Demystifying the digital transformation of the real estate brokerage industry in China: A case study of Lianjia (Beike). *Business Informatics*, vol. 18, no. 4, pp. 81–97.
DOI: 10.17323/2587-814X.2024.4.81.97

Introduction

The development of China's real estate industry, just like Chinese culture, is full of mystery. In addition to the impressive macroeconomic-growth indicators of real estate brought by China's economic development, academia and industry seem to possess only modest knowledge about China's real estate industry. Indeed, the real estate market has grown rapidly and played an important role in the economy over the past two decades. Investment in fixed assets (excluding rural households) nationwide totaled 50,303.6 billion yuan in 2023, marking a 3.0 percent increase over the previous year [1].

Real estate brokerage plays an important role in real estate promotion and sales by being the key channel of circulation and transaction in the real estate industry. The real estate brokerage industry in China has experienced four significant evolutions since the Chinese government restructured the real estate industry from planned to market-oriented in 1998, coinciding with the explosive growth of the real estate industry as a whole [2]. As of now, the real estate brokerage industry has become the most significant and effective element in the digital transformation of the entire real estate industry, which is remarkable for the industry as a whole because the real estate industry is believed to be difficult for carrying out technological innovation due to the existence of hard barriers in solidified processes and systems [3].

At the current stage, real estate brokerage in China is undergoing the digital transformation of the offline-to-online (O2O) commerce model under the Internet economy, in which the most prominent feature is the transformation from traditional offline brokerages to digital brokerage platforms. Moreover, the trend of evolution to digital platforms is spreading and penetrating at an accelerated pace due to the impact of the COVID-19 crisis [2, 4].

This study utilizes a qualitative case study methodology to explore Lianjia's digital transformation journey. Lianjia (Beike) stands out as a prime example of how traditional real estate firms can leverage digital technologies to transform their operations and service delivery. By exploring Lianjia's journey, this research seeks to understand the mechanisms and outcomes of digital transformation in the real estate brokerage industry. The primary objectives are to analyze the strategies employed by Lianjia, such as the creation of the Housing Dictionary, the development of the agent cooperation network (ACN), and the implementation of a management information systems (MIS) based Offline-to-Online (O2O) business model, as well as to evaluate the impact of these strategies on operational efficiency and customer satisfaction.

1. Literature review

1.1. Review on management information systems (MIS)

Management information systems (MIS) are integral to modern business operations, supporting decision-making, streamlining processes and enhancing organizational efficiency [5]. These systems encompass a range of technologies and processes used to collect, store, manage, and analyze data, providing valuable insights for management. Over the past few decades, the field of MIS has evolved significantly. Initially focused on basic data processing and record-keeping, MIS now incorporates advanced technologies such as big data, artificial intelligence (AI) and cloud computing [6]. This evolution has expanded the scope of MIS to include enterprise resource planning (ERP) systems, supply chain management (SCM), customer relationship management (CRM) systems and knowledge management systems (KMS) [7].

The successful implementation and operation of MIS are influenced by several key factors, including technological infrastructure, data quality and management, user training and engagement, organizational support and integration with business processes. A robust technological foundation, comprising reliable hardware, up-to-date software and a secure network, is essential for the effective functioning of MIS. This infrastructure supports the collection, processing and dissemination of information across the organization [8].

The accuracy, relevance, completeness and timeliness of data are crucial for the effectiveness of MIS. Proper data management practices ensure that high-quality data is available for analysis and decision-making [9]. Additionally, ensuring that users are adequately trained and engaged with the MIS is vital for its success. Users need to understand how to utilize the system effectively to support their tasks and decision-making processes [10].

Strong support from top management and alignment with organizational goals are essential for the successful implementation and operation of MIS [11, 12]. This includes providing the necessary resources and fostering a culture that values data-driven decision-making [12]. Furthermore, MIS must be integrated with existing business processes to provide relevant and actionable insights, ensuring that the system supports the organization's strategic and operational objectives effectively [11].

1.2. Offline-to-Online (O2O) business models in China

The Offline-to-Online (O2O) business model emerged to address the limitations of traditional offline operations by leveraging online capabilities to attract and retain customers. This approach seamlessly integrates online marketing and transactions with offline fulfillment and services, creating a cohesive customer experience. The rapid proliferation of smartphones and internet infrastructure in China has provided an ideal environment for O2O businesses to flourish across various sectors, for example, retailing [13, 14], manufacturing [15], food delivery [16, 17], wooden furniture [18] and real estate [4]. Key drivers of this transformation include supportive government policies, technological advancements and changing consumer behaviors.

First, the Chinese government has played a crucial role in the development of O2O businesses by implementing favorable policies and regulations that promote e-commerce and digital innovation while ensuring consumer protection and market stability [19]. This supportive regulatory environment also helped in shaping the service quality standards across industries [20].

Secondly, technological progress, particularly in big data, artificial intelligence (AI) and mobile internet, has been fundamental to the evolution of O2O business models [15]. These technologies enable synergistic interactions between online and offline platforms, enhancing operational efficiency and customer engagement.

Furthermore, consumer behavior has significantly influenced the adoption and success of O2O models. Chinese consumers increasingly favor the convenience of online shopping combined with the tangible benefits of offline experiences [21]. This preference reveals the importance of effectively integrating online and offline channels to provide personalized and seamless interactions, thereby enhancing customer satisfaction [18].

1.3. Data management in the application of MIS in O2O business models

The integration of management information systems (MIS) in Offline-to-Online (O2O) business models has revolutionized business operations by bridging the gap between traditional physical operations and digital platforms. MIS plays a crucial role in optimizing business processes by providing real-time data, improving inventory management, and streamlining workflows [22]. The integration of offline and online data ensures that businesses can efficiently manage their resources and respond promptly to market demands [23, 24].

MIS enables the integration of virtual online behavior with real-world offline activities, meeting clients' personalized requirements across multiple dimensions such as time, location, media, manner and cost. This integration optimizes resources and enhances the overall effectiveness of marketing campaigns [25]. By analyzing customer data, businesses can tailor their services to meet individual preferences, thereby enhancing customer satisfaction [21].

Moreover, MIS provides valuable insights that support strategic decision-making by integrating and analyzing data from both offline and online operations [26]. This comprehensive view enables businesses to make informed decisions that enhance their competitiveness and growth [27, 28]. However, implementing MIS in O2O business models comes with challenges, such as ensuring data accuracy, integrating different systems and maintaining customer privacy [13, 25]. Businesses must adopt advanced technologies and develop robust strategies to address these challenges effectively [29].

Data management is the cornerstone of digital transformation in business operations [30]. It involves the systematic collection, organization, storage and utilization of data to support decision-making processes and business operations. For example, the integration of big data and analytics allows managers to make informed decisions based on real-time data insights, improving market analysis and customer relationship management, which leads to more accurate and timely decision-making. Effective data management within MIS enhances operational efficiency and facilitates strategic renewal in O2O business models [2, 30, 31]. Through effective data management, MIS ensures that O2O business models can leverage accurate and timely data to drive continuous improvement and innovation, ultimately leading to sustained competitive advantage [32, 33].

2. Methodology

2.1. Selection of case study method

This study employs a qualitative case study methodology, which is ideal for investigating the research question within its specific context. As Yin [34] suggests, the qualitative case study method has several advantages, including the ability to provide rich, contextualized insights, capture complex phenomena and adapt to new findings. This approach allows for an in-depth exploration of specific cases, facilitating a comprehensive understanding of organizational processes.

2.2. Selection of case study company

The research design involved a single-case study focusing on Lianjia (Beike). This design was chosen due to Lianjia's exemplary digital transformation, which provides a rich context for exploring the mechanisms and outcomes of such transformations

Beijing Lianjia Real Estate Brokerage Co., known as Lianjia, was founded in 2001. From its establishment until 2017, Lianjia grew rapidly from a local entity to a nationwide real estate brokerage company with more than 8000 offline stores and 130000 employees across 28 cities in China, achieving a gross transaction volume of more than one trillion CNY [35].

In 2018, Lianjia launched a digital platform called Beike, also known as BEKE (KE Holdings Inc.). This platform is open to all real estate brokerage companies, including Lianjia, and provides housing transaction information, transaction support, housing financial support and other real estate services [36]. Beike demonstrated its strong market appeal with 279 resident brokerage brands, nearly 47000 community-centric stores, and nearly half a million real estate agents in over 100 cities in China by 2020, contributing to more than 3.6 million transactions [37]. In May 2020, Beike was listed on the New York Stock Exchange, becoming the largest real estate e-commerce platform in the Chinese market by market capitalization, at one point reaching five times the combined market capitalization of its competitors in China, such as E-House, Fangdd, Fang Holdings, 58.com and 515j.com.

2.3. Data source

This research relies on two primary sources of data: KE Holdings Inc.'s prospectus and annual reports [38, 39], which provide public information regarding Lianjia (Beike). The publicly available research data is carefully reviewed, adjusted and incorporated into this study ensuring a comprehensive case analysis of Lianjia's digital transformation process.

3. Case analysis of digital transformation of Lianjia (Beike)

Step I: Strategic digitalization of housing data – creation of the Housing dictionary.

Lianjia has been preparing since 2008 in order to complete its digital transformation, which has enabled it to seize the market opportunity in this industry. This was nine years before China officially adopted the AI industry as a national strategy, when in December 2017 the Chinese Ministry of Industry and Information Technology released the "Three-Year Action Plan for the Development of China's Artificial Intelligence Industry from 2018–2020" [40].

Since 2008, Lianjia has invested heavily in the electronic database of brokerage listings, later known as Housing Dictionary. With the help of a large number of human resources, digital technology and tools to verify and annotate every listing in its system, including house number, house type, orientation and location conditions, etc. From 2008 to the birth of Beike in 2018, this digital housing information has enabled Lianjia to accumulate the most authentic and large-scale data assets in China and also laid the foundation of big data to reconstruct the business model and standardized service process of the entire real estate brokerage industry through the transformation of digital platforms, and rapidly enhancing its technological capabilities in artificial intelligence (AI), virtual reality (VR), and Internet of Things (IoT), and further improving the accessibility and richness of data available to the platform participants. The housing dictionary of Lianjia contains comprehensive information on approximately 226 million homes by 2020, which is seen as the most comprehensive number of residential homes in China.

Step II: Preparation for Offline-to-Online (O2O) model – development of the agent cooperation network (ACN).

The well-established housing database makes Lianjia be able to pioneer the industry standard of "authentic property listings" in the real estate brokerage industry in China to counteract the malicious marketing tactic of fake listings, which is anathema to customers. In the absence of a unified regulatory framework for a long time, unscrupulous agents often post below-market or non-existent listings to increase customer traffic, leaving customers unable to distinguish between genuine and fake listings, and brokerage companies unable to truly control their core asset of listings. To standardize the transaction service process and promote cross-regional and cross-store cooperation to avoid vicious competition, Lianjia innovatively launched the agent cooperation network (ACN) internally in 2014. This network was the second key factor in the success of Lianjia's digital transformation, and later became the foundation of Beike's MIS-based O2O business model to connect online and offline infrastructure and manage various industry platform participants.

The agent cooperation network (ACN) defines the housing transaction process by real estate brokerage as ten steps. Each step has a clear role definition and corresponding responsibilities (*Fig. 1*). Under the premise of abiding by the rules of ACN Network, real estate agents have access to full sharing of property listings data and can participate in a transaction in different roles and interact with their upstream and downstream roles organically to bind interests. It allows multiple agents to process the same transaction to facilitate more frequent cross-store and cross-brand collaboration, with each agent earning a portion of commission based on their role and contribution, instead of the traditional real estate brokerage model where the only agent who ultimately facilitates the transaction gets paid the commission. It allows multiple agents to process the same transaction or one agent to participate in multiple transactions to facilitate the exchange of cross-store and cross-brand information and customer resources and improve transaction efficiency and customer satisfaction. Each agent receives a portion of the commission based on their role and contribution, as opposed to the traditional real estate brokerage model where only the broker who ultimately brokers the deal gets paid.

Step III: Implementation of the MIS-based O2O business model framework of Beike.

The MIS-based O2O business model framework makes Beike a platform that consists of three parts: platform infrastructure, platform participants and the services provided through the platform (*Fig. 2*). The platform infrastructure ensures that the platform participants can smoothly enjoy the O2O (offline and online) commerce model and platform usage experience, which in turn promotes the provision of higher quality services. On the other hand, through customer service experience and user feedback from industry platform participants, the iteration and upgrade of platform infrastructure are thus facilitated.

Platform infrastructure. The infrastructure core and operation system of the Beike platform is the agent cooperation network, which connects the online and offline infrastructure based on standardized and digitalized transaction processes and different roles, and each online and offline infrastructure contains four basic modules [37].

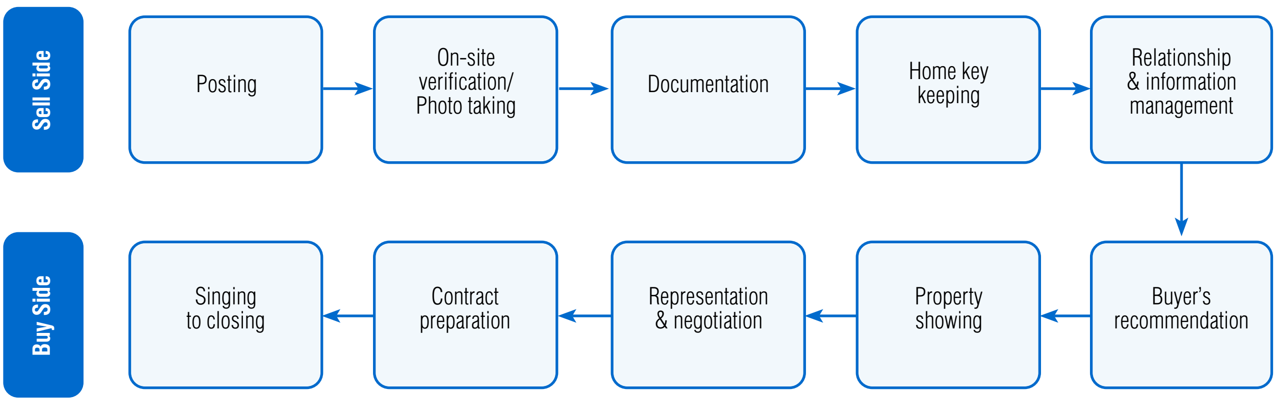

Fig. 1. Brokerage service flow under ACN [23].

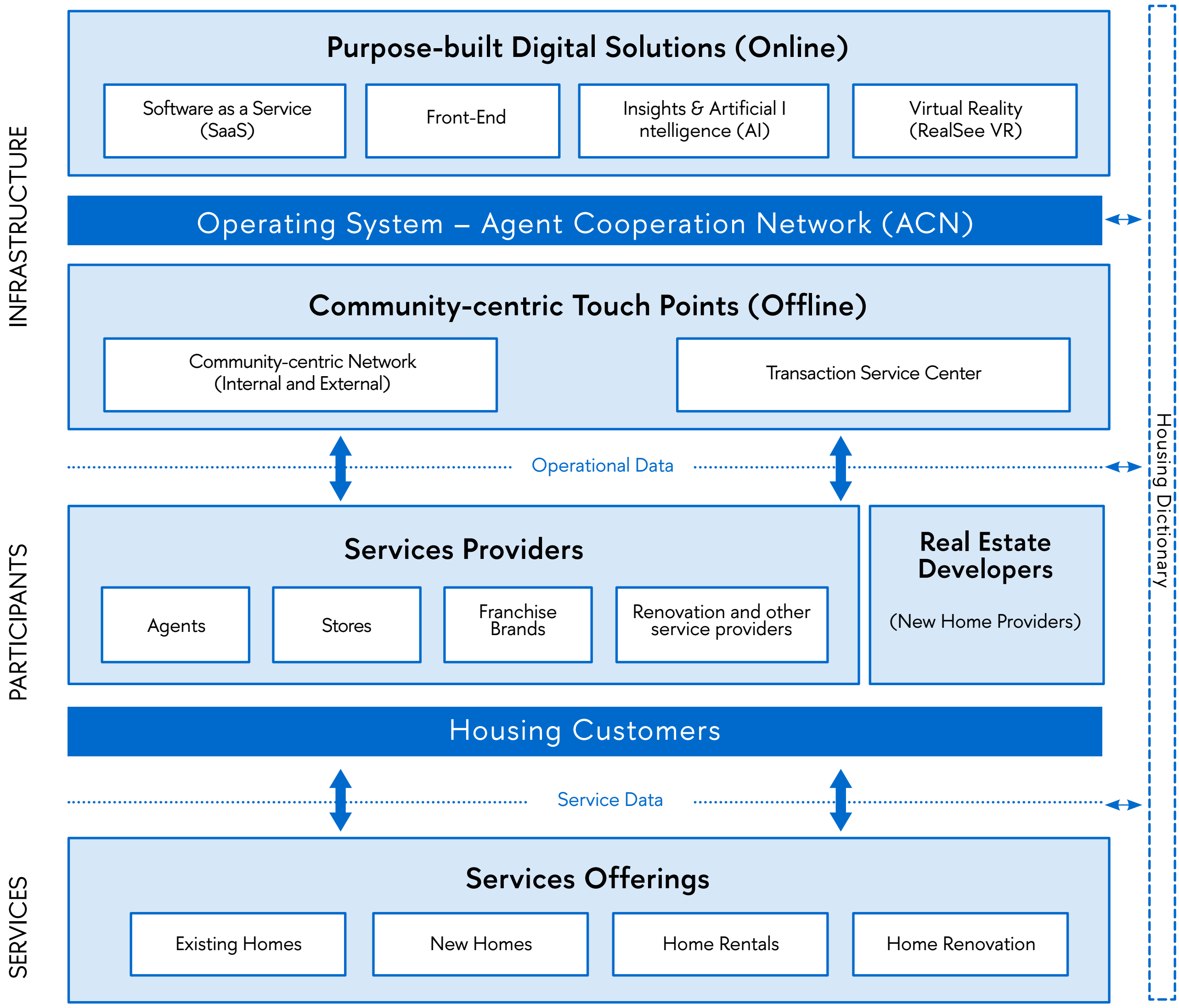

Fig. 2. MIS-based O2O business model framework of Beike [22].

The online infrastructure modules were built to support the digital transformation of its operational systems online, so-called purpose-built digital solutions, including Software as a Service (SaaS); customer front-end; data insights and artificial intelligence (AI); and virtual reality (VR).

- The SaaS system is for agents and store managers and is a carrier to execute each transaction section of the ACN, thereby ensuring that the whole transaction process is visualized and standardized.
- The front end of the Beike platform is to interact with its housing customers, including the 'ke.com' website, the Beike app, the customer service hotline in Mandarin Chinese, and other online interfaces.
- Data insights and AI technologies are applied to support the platform to carry out the foundation of the

data processing business. That helps the platform to process and analyze the historical property and transaction data and to provide analysis, including pricing, supply, and demand for agents and housing customers to better complete demand matching to facilitate housing transactions.

- The VR technology of the platform is provided by RealSee, a subsidiary of Beike, as one of the four online digital solutions. It brings offline homes to the online scene, which makes up for the deficiency caused by the limitations of offline physical space. Through VR live tours, housing customers could enjoy a 360-degree immersive display of property listings through the Beike front ends, with an agent online touring, and even reach a transaction intention. According to KE Holdings Inc. [39], as of 2022, a total of 1.508 billion times had been accumulated using VR tours by users of the Beike platform.

For the offline infrastructure, Beike has set up two modules for its community-centric touchpoints which include transaction service centers, and both internal and external community-centric networks. This offline infrastructure is defined as a community of agents, stores, housing customers and other public institutions that support the majority of the operations that need to be done offline.

- The transaction service center is one of the most important offline infrastructure modules of the Beike platform and connects to its online transaction supporting system. It helps housing customers and agents with transaction services such as providing home purchase agreements; tax payments to tax authorities; title transfers; home ownership registration; bank loans; and guarantees through cooperation and staffing with government agencies, real estate valuation agencies, banks and guarantee companies. The module is designed to address trust issues such as the failure of the buyer to pay the seller as agreed; the failure of the seller to transfer the title to the buyer after receiving payment; the failure to complete the transfer due to ownership issues; and the failure of the buyer and seller to pay the agent's commission, etc.
- The community-centric network refers to the offline contact points formed by Lianjia (internal) and its affiliated offline stores (external). These offline stores are distributed in large and small residential communities, focusing on community outreach and engagement, and through the ACN and SaaS system of the Beike platform, form a network that interacts both online and offline.

Platform participants. The Beike's platform is composed of three types of participants. These are service providers, real estate developers and housing customers. The service providers are connected and interacted with by the agent cooperation network as the infrastructure, which includes agents, brokerage stores, brokerage brands, home renovation providers and other home-related service providers. By joining the Beike platform, they automatically join and become part of the agent cooperation network, which enables them to proactively allocate the role-based commissions and enjoy the online purpose-built digital solutions and offline community-centric touch points to facilitate their transactions. Unlike service providers, real estate developers are not directly involved in the agent cooperation network, but rather cooperate with Beike to put their new housing project listings on the platform to attract and match customers, while the platform divides and matches different roles for different stores, brands and agents based on the agent cooperation network. The platform housing customers are defined as home buyers and sellers, landlords and tenants, and other customers with housing-related needs.

Services offerings. As an e-commerce platform, Beike does not directly provide services and products related to real estate, but rather serves as a platform to connect its resident real estate brands, stores, home renovations and other service providers, as well as real estate developers with potential housing customers through information matching. The platform services corresponding to the service providers include existing home transactions (also called second-hand house transactions), home rentals, home renovation services, etc. For this part of the service, the main source of revenue for the Beike platform is to extract a portion of the commission from the housing transaction and the platform service fee for the resident brands and stores. The other part of the service is for real estate developers. The Beike platform acts as a sales channel for new homes (first-hand houses), matching developers with new home buyers and providing sales and marketing solutions including brokerage services, sales plans, online marketing, digital tools, etc. For new home services, Beike generates platform revenue through developer commissions.

4. Discussion

4.1. Key factors for Lianjia's successful digital transformations

A company will be more competitive and able to compete with the whole industry ecosystem when it transforms into a platform [2, 41]. By transforming into a digital Beike platform, Lianjia opens its infrastructure and operation system to the entire real estate industry, enabling it to gain competitiveness and industry voice through compatibility with both internal and external ecosystem participants. Lianjia's successful digital transformation to Beike can be attributed to a combination of key factors. These include the favorable macro-environment of the Chinese market, characterized by the government's emphasis on technological innovation and a regulatory framework that allowed for flexibility. This environment enabled Lianjia to challenge traditional industry barriers and establish itself as a disruptor.

What's more, data plays a key role in connecting the traditional and digitalized business processes. The transformation of real estate brokerage business processes from traditionally offline operations to digital online operations has led to significant improvements in efficiency, accuracy, security, and customer experience [29]. Essentially, as shown in *Fig. 3*, the management and application of business data, alongside the digital infrastructure that carry this data, are crucial to this transformation [42]. Leveraging digital technology, this shift streamlines property listings, viewings, negotiations, and transactions, ensuring efficiency and compliance with regulatory standards [37, 38].

According to KE Holdings Inc. [37, 38] and Zuo [4], there are four main processes structured as the major activities or services provided by real estate brokerage: market research and property listing (**A1**), property viewing and interaction (**A2**), negotiation and agreement (**A3**), and transaction and documentation (**A4**).

In the traditional real estate brokerage, the initial stage of market research and property listing (**A1**) relied heavily on manual data collection, printed materials and in-person visits. Real estate agents would physically visit properties to gather information, take photographs and create listings. These listings were then promoted through printed brochures and advertisements distributed to potential buyers. This method was time-consuming and prone to errors and delays. The digital transformation of this stage leverages platform-based automated listing systems and online promotion algorithms. Real estate agents can now use online platforms to instantly upload property details, photographs and virtual tours. These platforms use advanced algorithms to promote listings to targeted audiences, significantly increasing the reach and efficiency of marketing efforts.

The second stage, property viewing and interaction (**A2**), traditionally involved physical showings, phone calls for scheduling and manual coordination between agents and potential buyers. This approach was highly limited by geographical and logistical constraints, often making it difficult for buyers to view multiple properties in a short period. The digital revolution has drastically improved this stage, particularly in terms of time-saving and reducing management loops. Potential buyers are able to take virtual tours of properties, interacting with agents through video calls and chat features. This reduction in process loops cuts the time cost from weeks to minutes and hours, saving resources and enhancing overall operational efficiency and customer experience.

Negotiation and agreement (**A3**), the third stage, traditionally involved face-to-face negotiations and

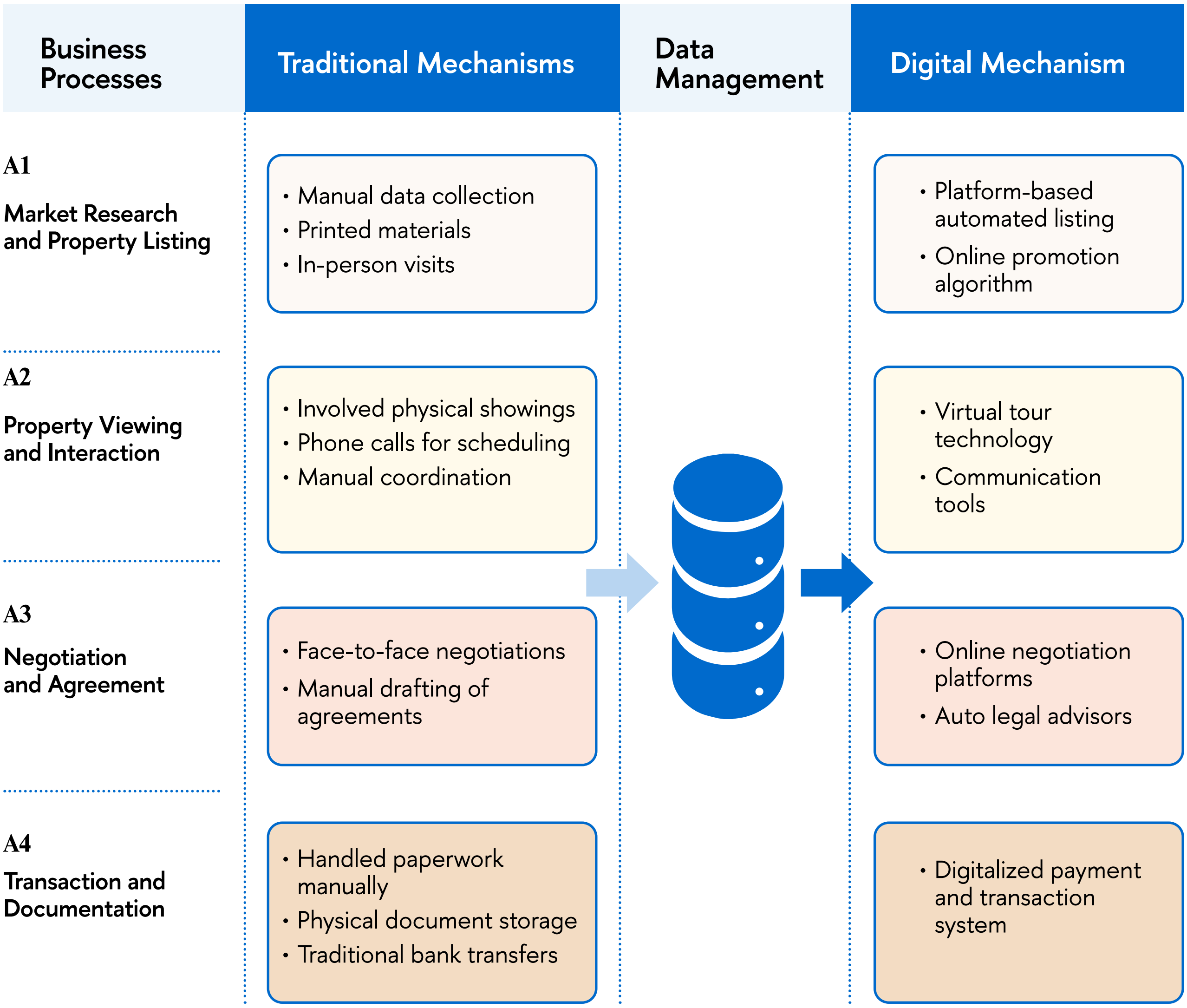

Fig. 3. Embedded mechanisms of transformative business processes.

manual drafting of agreements. Real estate agents and legal advisors would meet with buyers and sellers to negotiate terms, draft contracts and ensure compliance with legal requirements [10]. Digitalization has streamlined this stage with online negotiation platforms and automated legal advisors. These platforms facilitate real-time negotiations through chat and video features, while automated legal advisors use predefined AI databases to draft and review contracts, ensuring accuracy and compliance [44].

The final stage, transaction and documentation (**A4**) traditionally involved manually handling paperwork, physical document storage, and traditional bank transfers for payments. This approach was labor-intensive and vulnerable to errors, fraud and delays. Digital transformation has introduced digitalized payment and transaction systems, significantly improving the efficiency and security of this stage. Digital payment gateways and blockchain technology ensure secure and transparent transactions, while digital signature services facilitate the electronic signing of documents [45, 46].

4.2. Key differentiators in Lianjia's digital transformation

Several key differentiators have enabled Lianjia (Beike) to maximize the benefits of digital transformation, setting new standards in the real estate brokerage industry, and so to become an industry leader. These differentiators include an early start and accumulation of data, a wide stack of technologies in use and cross-cutting coverage of all business processes.

Firstly, Lianjia's digital transformation journey began with an early focus on accumulating and leveraging extensive data. This initiative was spearheaded by the creation of the Housing Dictionary in 2008. The Housing Dictionary is a comprehensive database that dynamically includes detailed information on millions of homes across China. It serves as the foundation for Lianjia's digital knowledge base, integrating historical transaction data, performance indicators of agents and other critical information.

- Comprehensive Data Repository: The Housing Dictionary provides a detailed and accurate repository of property data, enhancing the reliability of property listings and facilitating advanced analytics.
- Historical Transaction Data: By integrating historical transaction data, Lianjia can offer precise property valuations and market trend analysis, improving customer trust and decision-making.
- Performance Indicators: Tracking agents' performance metrics helps in identifying top performers, optimizing agent allocation and improving overall service quality.

Secondly, Lianjia's ability to integrate a broad range of advanced technologies has been pivotal in its digital transformation. The company employs a diverse technological stack to enhance operational efficiency, customer experience, and service delivery.

- Big Data Analytics: Utilizes extensive data to gain insights into market trends, customer preferences, and pricing strategies, enabling data-driven decision-making.
- Artificial Intelligence (AI) and Machine Learning (ML): Implements AI and ML for predictive analytics, personalized customer recommendations, and automation of routine tasks, increasing efficiency and customer satisfaction.
- Cloud Computing: Leverages cloud infrastructure to ensure scalable, secure and efficient data storage and processing, supporting the company's extensive digital operations.
- Virtual Reality (VR): Offers immersive property viewing experiences through VR technology, bridging the gap between online and offline interactions and enhancing customer engagement.
- Blockchain: Ensures secure and transparent transactions, reducing fraud and increasing trust in the digital transaction process.

Thirdly, Lianjia's comprehensive digital transformation approach extends across all its business processes, integrating both internal operations and interactions with various stakeholders.

- Agent cooperation network (ACN): Standardizes the transaction process and promotes collaboration among agents across different regions and stores. It facilitates data sharing, resource allocation and cooperative transactions, enhancing efficiency and service quality.
- Management information system (MIS)-based O2O Business Model: Supports internal workflows, enhances communication and optimizes resource management, ensuring seamless integration of digital processes within the organization. It bridges online and offline activities by integrating digital platforms with physical services. This model ensures a consistent and comprehensive customer experience both online and offline.

- Customer Front-end Platforms: Includes multiple channels such as the 'ke.com' website, Beike app and customer service hotlines, ensuring accessible and efficient customer interactions.

4.3. Criticisms of Lianjia's digital transformation

Traditional Chinese real estate brokerage companies, represented by Lianjia, have been able to gain tremendous platform advantages and create a competitive platform ecosystem through digital platform transformation, but the industry drawbacks brought about by the transformation or not changed through the transformation are also obvious. While the digital transformation of the real estate brokerage industry is not unique to different market characteristics, these drawbacks seem to have common features across different markets.

Increasing commissions and transaction costs. Barwick and Wong [47] criticized Lianjia, which has rapidly increased its market share in China after transforming into a digital platform and has increased its commissions while bundling its brokerage service products. This has undoubtedly increased eventual home transaction costs. Customers also tend to stick with only one platform and a single home, which strengthens the market dominance of digital platforms, particularly the dominating prices. Similarly, when home buyers in the United States realize that the final closing price will be adjusted, home buyers' decisions will be swayed and the final decision will no longer depend on who pays [48].

Negative effects of low entry barriers. Hsieh and Moretti [49] and Barwick and Pathak [50] pointed out that due to the low barriers to entry and the large number of participants in the real estate brokerage industry, especially in the residential brokerage industry, has led to extremely fierce competition in the industry. The real estate brokerage industry is considered to be a highly profitable industry due to the substantial brokerage commissions. This has attracted a lot of brokerage practitioners and companies to join the industry, but the number of properties available for transactions, both new and existing homes, is limited. This has led to the competition of different agents for the same housing listing, including many inexperienced agents, resulting in a long housing transaction cycle, which has led to lower labor productivity and indirectly affects the loss of social welfare [51].

Concern about service quality. Another concern caused by the lower entry barrier is the uneven quality of brokerage services. The low industry entry barrier makes the service quality of real estate agents vary. For example, the brokerage industry in China has developed for only 20 years since its formation, which makes the whole industry ignore the overall quality of practitioners, such as professional attitude, level of education, and professional training. In this industry, brokerage companies or agents could interface with potential housing customers as long as they have listing information, and digital brokerage platforms make it easier for them to assess this information and reach out to the customers. However, it is difficult for customers without home buying experience to distinguish which agents would provide better brokerage services, resulting in a huge gap in service experience and even the image of the entire industry [52].

Conclusion

This study has shown that Lianjia's comprehensive digital transformation, through the creation of the Housing Dictionary, the agent cooperation network (ACN) and the implementation of an MIS-based O2O business model, has significantly enhanced operational efficiency and customer satisfaction in the Chinese real estate brokerage industry. The innovative use of digital platforms has not only provided Lianjia with a competitive advantage but also offers a blueprint for other real estate firms aiming to undertake similar transformations.

However, the focus on a single case study presents certain limitations, as the findings may not fully represent the diverse experiences of digital transformation across the entire industry. Future research could address these limitations by conducting comparative studies across different markets or examining the long-term impact on profitability and market share.

In conclusion, this study underscores the critical role of digital transformation in modernizing the real estate brokerage industry and suggests that other firms could achieve similar success by adopting strategic digital innovations. The insights gained from Lianjia's experience contribute to the broader understanding of digital transformation and its potential to drive significant improvements in business performance and customer engagement. ■

References

1. National Bureau of Statistics of China (2024) *Investment in fixed assets in 2023*. Available at: https://www.stats.gov.cn/english/PressRelease/202402/t20240201_1947074.html (accessed 2 November 2024).
2. Zhang Y., Yao L. (2022) How the leading Chinese real estate brokerage transformed into a digital platform business. *Strategy and Leadership*, vol. 50, no. 1, pp. 19–24. https://doi.org/10.1108/SL-10-2021-0107
3. Kassner A.J., Cajias M., Zhu B. (2022) The PropTech investors' dilemma – What are the key success factors that secure survival? *Journal of Property Investment & Finance*, vol. 41, no. 1, pp. 76–91. https://doi.org/10.1108/JPIF-01-2022-0007
4. Zuo H. (2021) The rise of platform-based networks: Outlook for the real estate brokerage industry. *Understanding China's Real Estate Markets*, pp. 299–305.
5. Gupta H. (2011) *Management information system – An insight.* Hitesh Gupta.
6. Laudon K.C., Laudon J.P. (2004) *Management information systems: Managing the digital firm.* Pearson Education.
7. Costa A., Ferreira C., Bento E., Aparicio F. (2016) Enterprise resource planning adoption and satisfaction determinants. *Computers in Human Behavior*, vol. 63, pp. 659–671. https://doi.org/10.1016/j.chb.2016.05.090
8. Karim A.J. (2011) The significance of management information systems for enhancing strategic and tactical planning. *Journal of Information Systems and Technology Management*, vol. 8, pp. 459–470. https://doi.org/10.4301/S1807-17752011000200011
9. Janssen M., van der Voort H., Wahyudi A. (2017) Factors influencing big data decision-making quality. *Journal of Business Research*, vol. 70, pp. 338–345. https://doi.org/10.1016/j.jbusres.2016.08.007
10. Venkatesh V., Speier C., Morris M.G. (2002) User acceptance enablers in individual decision making about technology: Toward an integrated model, decision sciences. *Decision Sciences*, vol. 33, no. 2, pp. 297–316. https://doi.org/10.1111/j.1540-5915.2002.tb01646.x
11. Bush M., Lederer A.L., Li X., et al. (2009) The alignment of information systems with organizational objectives and strategies in health care. *International Journal of Medical Informatics*, vol. 78, no. 7, pp. 446–456. https://doi.org/10.1016/j.ijmedinf.2009.02.004
12. Marsh J.A., Farrell C.C. (2015) How leaders can support teachers with data-driven decision making: A framework for understanding capacity building. *Educational Management Administration & Leadership*, vol. 43, no. 2, pp. 269–289. https://doi.org/10.1177/1741143214537229

13. Kim Y., Wang Q., Roh T. (2021) Do information and service quality affect perceived privacy protection, satisfaction, and loyalty? Evidence from a Chinese O2O-based mobile shopping application. *Telematics and Informatics*, vol. 56, 101483. https://doi.org/10.1016/j.tele.2020.101483

14. Shen C., Chen M., Wang C. (2019) Analyzing the trend of O2O commerce by bilingual text mining on social media. *Computers in Human Behavior,* vol. 101, pp. 474–483. https://doi.org/10.1016/j.chb.2018.09.031

15. Sun S., Cegielski C.G., Li Z. (2015) Amassing and analyzing customer data in the age of big data: A case study of Haier's online-to-offline (O2O) business model. *Journal of Information Technology Case and Application Research*, vol. 17, no. 3–4, pp. 156–165. https://doi.org/10.1080/15228053.2015.1095017

16. Li M., Wang M., Wang T. (2021) Analysis on influencing factors of e-loyalty of a Chinese O2O E-commerce food delivery platform. Proceedings of the *2020 4th International Conference on E-Business and Internet, Association for Computing Machinery, New York, USA*, pp. 6–10. https://doi.org/10.1145/3436209.3436885

17. Xue G., Wang Z., Wang G. (2021) Optimization of rider scheduling for a food delivery service in O2O business. *Journal of Advanced Transportation*, vol. 2021, no. 1, 5515909. https://doi.org/10.1155/2021/5515909

18. Lin M., Wang Z., Zhang Z., Cao Y. (2019) Research on consumers' attitudes in China about using online-to-offline mode for purchasing wooden furniture. *Forest Products Journal*, vol. 69, no. 2, pp. 159–172. https://doi.org/10.13073/FPJ-D-18-00039

19. Junyong X., Linbo J. (2014) Electronic commerce in China: current status, development strategies, and new trends. *China Finance and Economic Review,* vol. 3, no. 3, pp. 71–94. https://doi.org/doi:10.1515/cfer-2014-030306

20. Li Y., Shang H. (2020) Service quality, perceived value, and citizens' continuous-use intention regarding e-government: Empirical evidence from China. *Information & Management,* vol. 57, no. 3, 103197. https://doi.org/10.1016/j.im.2019.103197

21. Yang Y., Gong Y., Land L.P.W., Chesney T. (2020) Understanding the effects of physical experience and information integration on consumer use of online to offline commerce. *International Journal of Information Management*, vol. 51, 102046. https://doi.org/10.1016/j.ijinfomgt.2019.102046

22. Guha S., Kumar S. (2018) Emergence of big data research in operations management, information systems, and healthcare: Past contributions and future roadmap, production and operations management. *SAGE Publications*, vol. 27, no. 9, pp. 1724–1735. https://doi.org/10.1111/poms.12833

23. Oh L.-B., Teo H.-H., Sambamurthy V. (2012) The effects of retail channel integration through the use of information technologies on firm performance. *Journal of Operations Management*, vol. 30, no. 5, pp. 368–381. https://doi.org/10.1016/j.jom.2012.03.001

24. Rialti R., Marzi G., Silic M., Ciappei C. (2018) Ambidextrous organization and agility in big data era. *Business Process Management Journal*, vol. 24, no. 5, pp. 1091–1109. https://doi.org/10.1108/BPMJ-07-2017-0210

25. Xu Y., Teng F., Liu Q. (2017) Management information systems for advertisement based on online-to-offline strategy. *IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*, pp. 197–202. https://doi.org/10.1109/ICIS.2017.7959993

26. Grover V., Chiang R.H.L., Liang T.-P., Zhang D. (2018) Creating strategic business value from big data analytics: A research framework. *Journal of Management Information Systems*, vol. 35, no. 2, pp. 388–423. https://doi.org/10.1080/07421222.2018.1451951

27. Akter S., Bandara R., Hani U., et al. (2019) Analytics-based decision-making for service systems: A qualitative study and agenda for future research. *International Journal of Information Management*, vol. 48, pp. 85–95. https://doi.org/10.1016/j.ijinfomgt.2019.01.020

28. Shet S.V., Poddar T., Wamba S.F., Dwivedi Y.K. (2021) Examining the determinants of successful adoption of data analytics in human resource management – A framework for implications. *Journal of Business Research*, vol. 131, pp. 311–326. https://doi.org/10.1016/j.jbusres.2021.03.054

29. Wang F. (2022) The necessity of strategic transformation of real estate digital operations from the perspective of information economic management. *Cuadernos de Economía*, vol. 45, no. 129, pp. 60–69. https://doi.org/10.32826/cude.v1i129.807

30. Chanias S., Myers M.D., Hess T. (2019) Digital transformation strategy making in pre-digital organizations: The case of a financial services provider. *The Journal of Strategic Information Systems*, vol. 28, no. 1, pp. 17–33. https://doi.org/10.1016/j.jsis.2018.11.003

31. Anejionu O.C.D., Thakuriah P., McHugh A., et al. (2019) Spatial urban data system: A cloud-enabled big data infrastructure for social and economic urban analytics. *Future Generation Computer Systems*, vol. 98, pp. 456–473. https://doi.org/10.1016/j.future.2019.03.052

32. Garcia-Teruel R. (2020) Legal challenges and opportunities of blockchain technology in the real estate sector. *Journal of Property Planning and Environmental Law*, vol. 12, no. 2, pp. 129–145. https://doi.org/10.1108/JPPEL-07-2019-0039

33. Marona B., Tomal M. (2020) The COVID-19 pandemic impact upon housing brokers' workflow and their clients' attitude: Real estate market in Krakow. *Entrepreneurial Business and Economics Review*, vol. 8, no. 4, pp. 221–232. https://doi.org/10.15678/EBER.2020.080412

34. Yin R.K. (2018) *Case study research and applications*. Sage.

35. Beike Real Estate (2024) *Introduction to Beijing Lianjia*. Available at: https://bj.lianjia.com/about/aboutlianjia (accessed 2 November 2024).

36. Beike Real Estate (2024) *Beijing Lianjia Chronicle*. Available at: https://bj.lianjia.com/about/lianjiaevents (accessed 2 November 2024).

37. KE Holdings Inc. (2020) *Initial Public Offering of American Depositary Shares. Prospectus of KE Holdings Inc*, no. 333-240068. Available at: https://www.sec.gov/Archives/edgar/data/1809587/000104746920004582/a2242260z424b4.htm (accessed 2 November 2024).

38. KE Holdings Inc. (2022) *Listing by way of introduction of the main board of the stock exchange of Hong Kong*. Available at: https://www1.hkexnews.hk/listedco/listconews/sehk/2022/0505/2022050500082_c.pdf (accessed 2 November 2024).

39. KE Holdings Inc. (2023) Beike carbon neutrality target and roadmap report. Available at: https://about.ke.com/news/article/1702633307019155 (accessed 2 November 2024).

40. Ciemniak R. (2021) Navigating the property data landscape. *Understanding China's Real Estate Markets. Management for Professionals. Springer, Cham,* pp. 21–34. https://doi.org/10.1007/978-3-030-49032-4_3

41. van Alstyne M.W., Parker G.G., Choud*ary S.P. (2016) Pipelines, platforms, and the new rules of strategy. Harvard Business Review*, vol. 94, no. 4, pp. 54–62.

42. Stoianova O., Lezina T., Ivanova V. (2020) The framework for assessing company's digital transformation readiness. *St. Petersburg University Journal of Economic Studies*, vol. 36, no. 2, pp. 243–265. https://doi.org/10.21638/spbu05.2020.204

43. Cubric M. (2020) Drivers, barriers and social considerations for AI adoption in business and management: A tertiary study. *Technology in Society*, vol. 62, 101257. https://doi.org/10.1016/j.techsoc.2020.101257

44. Li L., Wang C. (2006) Real estate agency in China in the information age. *Property Management*, vol. 24, no. 1, pp. 47–61. https://doi.org/10.1108/02637470610643119

45. Wang F. (2023) The present and future of the digital transformation of real estate: A systematic review of smart real estate. *Business Informatics,* vol. 17, no. 2, pp. 85–97. https://doi.org/10.17323/2587-814X.2023.2.85.97

46. Wang F. (2023) Block by block: a bibliometric analysis of blockchain in real estate. *Journal of Information Systems Engineering and Management,* vol. 8, no. 2, 21498. https://doi.org/10.55267/iadt.07.13880

47. Barwick P.J., Wong M. (2019) *Competition in the real estate brokerage industry: A critical review*. The Brookings Institution Publication.

48. Best M.C., Kleven H.J. (2018) Housing market responses to transaction taxes: Evidence from, notches and stimulus in the UK. *The Review of Economic Studies*, vol. 85, no. 1, pp. 157–193. https://doi.org/10.1093/restud/rdx032

49. Hsieh C.-T., Moretti E. (2003) Can free entry be inefficient? Fixed commissions and social waste in the real estate industry. *Journal of Political Economy*, vol. 111, no. 5, pp. 1076–1122. https://doi.org/http://doi.org/10.1086/376953

50. Barwick P.J., Pathak P.A. (2015) An empirical study of real estate agents in Greater Boston. *The RAND Journal of Economics*, vol. 46, no. 1, pp. 103–145. https://doi.org/10.1111/1756-2171.12082

51. Gilbukh S., Goldsmith-Pinkham P.S. (2019) Heterogeneous real estate agents and the housing cycle. *National Bureau of Economic Research*, working paper no. 31683. https://doi.org/10.3386/w31683

52. Wang F., Stoianova O.V. (2023) Modelling the effect of E-service quality and knowledge culture on employee performance: A SEM approach based on survey data from Chinese real estate industry. *Journal of Applied Informatics*, vol. 18, no. 5, pp. 5–19. https://doi.org/10.37791/2687-0649-2023-18-5-5-19

About the authors

Fengchen Wang

Doctoral Student, St. Petersburg School of Economics and Management, HSE University, 3A-1, Kantemirovskaya Street, St. Petersburg 194100, Russia;

MSc Student, Campus Valrose, Université Côte d'Azur, 28, Avenue Valrose, Nice 06000, France;

E-mail: fwang@hse.ru

ORCID: 0000-0003-3103-5049

Olga V. Stoianova

Dr. Sci. (Tech.);

Professor, Head of the Department of Business Informatics and Operations Management, St. Petersburg School of Economics and Management, HSE University, St. Petersburg 194100, Russia;

E-mail: ostoyanova@hse.ru

ORCID: 0000-0002-7281-2708

Angel Barajas

Ph.D. in Business Administration;

Associate Professor, Faculty of Business Administration and Tourism, University of Vigo, Ourense 32004, Spain;

E-mail: abarajas@uvigo.gal

ORCID: 0000-0002-0198-4457

DOI: 10.17323/2587-814X.2024.4.98.111

Management of pricing policy of a timber enterprise considering the problems of formation of raw material supply chains and determining production volumes*

Rodion S. Rogulin
E-mail: rafassiaofusa@mail.ru

Vladivostok State University, Vladivostok, Russia

Abstract

This paper considers a mathematical model that allows managers of a timber enterprise to develop supply chains and manage the pricing policy of the organization. This model is a modification of the model developed earlier and differs from it by taking into account the technology of raw material cutting. The model takes into account the consumption rates of raw materials, purchases on the commodity exchange, transportation of products and pricing policy of the enterprise taking into account the demand. The purpose of the model is to maximize the value of operating profit of the enterprise. When searching for a solution, an optimization strategy is applied which includes two stages: application of linear optimization at the first stage and genetic algorithm at the second stage. As a result of testing the model at one of the timber processing enterprises in the Primorsky Territory, data were obtained, based on which recommendations are formulated for managers of the company regarding cooperation with loggers. This work represents an

* The article is published with the support of the HSE University Partnership Programme

important step in the development of supply chain management methodology in the timber industry, taking into account the technology of raw material cutting. Further research may include modification of the model using stochastic factors, improving decision-making methods and development of more accurate product demand functions. The work has practical significance for enterprises of the timber processing industry, since it can contribute to the improvement of their production processes and increase profits.

Keywords: supply chain, production volume, timber enterprises, optimality of decisions, mathematical model, commodity and raw materials exchange, share of the useful volume of raw materials, transit time, rational raw material transactions, increasing efficiency

Citation: Rogulin R.S. (2024) Management of pricing policy of a timber enterprise considering the problems of formation of raw material supply chains and determining production volumes. *Business Informatics*, vol. 18, no. 4, pp. 98–111. DOI: 10.17323/2587-814X.2024.4.98.111

Introduction

Timber enterprises, which play a significant role in the economy, face ever-increasing challenges and difficulties in organizing efficient supply chains of raw materials and the organization of production using cutting technologies. At timber enterprises, the issues of price policy optimization taking into account the formation of supply chains and determining production volumes are topical and critically important [1–10].

The timber industry is of strategic importance, supplying raw materials and products necessary for a variety of industries – from construction to paper and furniture production. However, under the conditions of changing market requirements, environmental constraints and intense global competition, timber companies have to constantly adapt and optimize their production and logistics processes [2, 11–15].

The importance of this topic is emphasized by its relevance in the modern world. A sharp increase in competition and external factors, such as climate change and legislation, put pressure on timber enterprises, forcing them to strive to increase production efficiency, reduce costs and improve product quality [3].

Analysis of existing studies [1–22] confirms the complexity of the problem of optimizing raw material supply chains and production volumes at timber enterprises. Effective management of these processes requires taking into account many variables, constraints and uncertainties related with production and logistics operations.

Relying on the above, it can be argued that the development of mathematical models and optimization methods for the formation of effective supply chains and production volumes using cutting technologies is becoming an integral part of the development strategy of timber enterprises.

The research presented in this article is relevant in the context of an ever-changing forest industry environment and rapidly evolving technologies that can improve production efficiency and reduce negative environmental impacts. These solutions can help timber companies to reduce costs, optimize production processes and improve product quality, as well as to meet new sustainability and environmental standards.

The current study, on the one hand, differs from its counterparts [1–20] in that it allows us to take into account the technology of cutting round wood entering production in conjunction with other important factors of forestry production. Particularly, these factors include the formation of raw material supply chains from the commodity and raw material exchange (since the enterprise does not have its own plots) and determining production volumes and pricing policy of the enterprise over the entire planning horizon. On the other hand, this paper complements the work [21] because it presents an optimization algorithm that allows us to speed up the process of finding an optimal solution.

The structure of the work provides justification of the topicality of the study, review of the literature, setting the goals and objectives of the study, development of a mathematical model and optimization algorithm, description of the results of approbation of the model on the data of the commodity and raw materials exchange and one of the timber enterprises of the Primorsky Territory, as well as the formulation of proposals for further development and modification of the model.

1. Literature review

The topic of optimizing the formation of raw material supply chains and production volumes using cutting technology at timber enterprises is a multifaceted field of research which is becoming increasingly relevant in the modern world. Timber industries operating in a changing economic and ecological environment face ongoing challenges in managing supply chains and optimizing production processes.

The main objectives of research in this area are to increase production efficiency [1–3], reduce production and logistics costs [4, 5], reduce waste [6, 7], and improve product quality [8, 9], taking into account the relationship between demand and prices. Over the last three years, only one paper [13] was published that would combine the main factors of production of the timber industry complex: the relationship between demand and prices for finished products, the production process, the formation of raw material supply chains and the construction of routes for transportation of finished products to customers. However, the second derivative of the demand function in this model at point 0 is not continuous, which indicates that there is an instantaneous rate of change in demand with an extremely small change in price. In addition, the real production of the timber industry is more often engaged not in the production of OSB boards, but in the cutting of lamellas and other workpieces from the incoming round wood for further production.

In recent decades, researchers have been actively working on the development of mathematical models [10] and optimization methods [11] to solve complex problems related to supply chain management and production management in timber enterprises. These models and methods are able to account for the many variables, constraints and uncertainties that are inherent in this industry.

One of the key aspects of the research is the integration of nesting technology into supply chain and production optimization [12]. Cutting technology allows producers to maximize the efficient use of forest raw materials and minimize waste, which is important from the point of view of production sustainability.

Analysis of the literature allows us to identify several key areas of research. First, it is the development of mathematical models that enable optimization of supply chains and production volumes, taking into account various parameters and constraints [3, 8, 9]. Second, it is research in the field of inventory management and production processes optimization [1, 12, 13], including cutting technology [14–16]. Third, it is research into the impact of external factors such as climate change and environmental standards [17–19] on supply chain and production management in the timber industry [20, 21].

In addition, it is worth noting the significant advantages of applying information technology and modern data analysis methods in solving optimization problems. The implementation of digital solutions and analytical tools allows companies to more accurately forecast demand, optimize inventories and resources, and manage complex supply chains [22].

Despite significant advances, there remain unresolved challenges and problems in this area. Resistance to change on the part of employees and suppliers, as well as the changing global economic environment, present major obstacles. Moreover, there is a need to develop innovative approaches to supply chain and manufacturing management given the rapidly changing market conditions and requirements.

In view of the above, it can be stated that the problem that affects the determination of the price of finished products in conjunction with the tasks of determining production volumes, the formation of supply chains of raw materials and transportation of finished products to customers is important and topical.

The criterion for a more efficient model will be the number of iterations required for the algorithm to reach a solution that does not change significantly with increasing iterations. The closest work in the literature is the study [13]. The results of the model proposed by the author are mainly compared with the output of the model from [13].

2. Research goals and objectives

In [13], a model of the activities of enterprises in the timber industry is considered on which we rely in this study. The model presented in this paper takes into account three key production processes: supplies and volumes of raw material purchases from the domestic market of the region; production volumes, taking into account the demand for each type of product and available stocks of raw materials; as well as ways of delivery of finished products to customers. The production technology based on the cutting of incoming round wood in the warehouse is considered.

Usually, companies receive orders from customers in advance, which allows them to plan activities for long periods. This is important for optimizing production processes, material procurement and resource allocation. However, it should be noted that the demand for timber products is subject to seasonal fluctuations, such as increased demand for heating materials during the winter and demand for construction materials during the summer months. These seasonal changes create additional challenges for production planning and inventory management. However, with adaptive strategies and analysis of market trends, a company can effectively adapt to changing demand and successfully manage production processes [13].

The purpose of this study is to develop a mathematical model that allows for optimizing the production processes of the timber industry complex, including determining the volume of output using round wood cutting technology, purchasing raw materials on the domestic market, delivery of finished products to the buyer, as well as the formation of pricing policy of the enterprise at different planning periods. The main task is to assess the feasibility of interaction between the enterprise and the forest commodity exchange in order to optimize production processes and improve the efficiency of the company.

To achieve the goal of the work the following research objectives were put forward:

1) Design of a two-stage economic and mathematical model including:

a) determining a suboptimal vector of production volumes by days on a given planning horizon, raw material supply chains and transportation volumes of raw materials and finished products;

b) determining sale prices for finished products.

2) Formation of a consistent solution for the two stages of the mathematical model.

3) Development of software that allows us to solve the set tasks.

4) Analyzing the results of model testing.

3. Mathematical model

To ensure uninterrupted operation of the timber enterprise, timely supplies of raw materials are required. Each timber harvester (located in certain area of the region) notifies the company that it will prepare a given volume of raw materials by a predetermined date and put it up for sale. The buyer can buy a part of the lot and select the raw material required by characteristics from the stack.

After a sufficient volume of raw materials is delivered to the production warehouse, the enterprise has the task of determining the optimal set of production operations for cutting technology and setting prices for products, taking into account the current dynamics of market demand. It is important to note that the change in prices for products is limited by the value of μ per week, which requires careful planning and analysis of market trends. In addition, once the production plan has been developed, it is necessary to organize efficient logistics for delivering products to customers (taking into account their individual needs and preferences) through known transportation hubs. This process requires coordinated work between production, sales and logistics departments to ensure smooth order fulfillment and customer satisfaction [13].

Achieving this goal requires the development of a mathematical model that optimizes all the factors described above.

M – planning horizon under consideration (days);

k – type of product produced, $k = 1, ..., K$;

len_e – length of workpiece of type e (m);

width, *height* – width and thickness of the workpieces, respectively (m);

$\mathbb{w}(m)$ – week number $\mathbb{w}(m)$ depending on day number m, where

$$\mathbb{W}(m) = \frac{m}{7}; \quad (1)$$

$\mathbb{W}(M) = \frac{M}{7}$ – – number of weeks depending on planning horizon M;

c_{imrl} – purchase price of 1 m^3 from the i-th application for raw materials of the l-th type in the r-th region on the m-th day (rubles, shipping costs are included);

$c_{\jmath}$ – total costs of transportation of finished goods to the buyer at the point $\jmath$ (rubles);

b_{et} – number of times a workpiece of type e is encountered in a cutting of type t (in the literature, the set $\{b_{et}\}$ is usually referred to as a cutting map);

V_{imrl} – volume of raw materials of type l in application i from region r on day m (m^3);

$\bar{u}$ – maximum warehouse capacity (m^3);

$\bar{O}$ – maximum number of cuttings per day (units);

v'_{imr} – volumes of raw materials of type l purchased in the previous period, about which it is known that they will arrive at the warehouse on day m (m^3);

p_{kml} – the selling price of goods of type k made from raw materials of type l on day m (rubles);

I' – the number of orders that were purchased in the previous period (up to $m = 0$, and the date of their arrival at the warehouse is known in advance);

R – number of districts;

T_r – time consumption (in days) for delivery of any volume of raw materials from the region r by railroad;

$Q_{\jmath k\mathbb{w}l}$ – demand of $\jmath$-th retail company for product k from resource type l in week $\mathbb{w}$;

$\dot{Q}_{\jmath k\mathbb{w}l}$ – mathematical expectation of the volume of demand of the retailer $\jmath$ for products type k, produced from the resource of the type l, during the week $\mathbb{w}$;

$\mathcal{J}^*$ – number of retailers (final delivery points of manufactured products);

Bud_0 – initial budget;

$month(m)$ – month number depending on the day number;

$A_{ekl}^{month(m)}$ – amount of workpieces of type e, used to produce a unit of goods k from a resource of type l during the month $month(m)$ (units);

$Iter$ – number of independent iterations (units);

FC – fixed costs per day of operation (rubles);

$\mathbb{V}$ – slab volume (m^3);

$\tilde{L}$ – slab length (m);

x_{kml} – volume of production of goods of type k per day m made from raw materials of type l (pcs);

$\mathbb{z}_{\jmath k\mathbb{w}l}$ – volume of transportation of goods of type k made from raw materials of type l to point j during the week $\mathbb{w}(m)$ (units);

q_{tml} – number of cuttings of type t made from raw materials of type l per day m;

v_{imrl} — purchased volume of raw materials of type l from application i from region r on day m (m^3);

u_{ml} — volume of stock of raw materials of type l on day m in the warehouse (m^3);

$\tilde{u}_{eml}$ — volume of stock of workpieces of type e made from raw materials l by day m.

As an objective function, we will consider the value of the enterprise's operational profit on the planning horizon M (2):

$$\max_{p,x,v,\mathbb{z}} \sum_m \left(\sum_{k,l} p_{kml} x_{kml} - \sum_{i,l,r} c_{imrl} v_{imrl} - \sum_{j,k,l} c_j z_{jk\mathbb{w}(m)l} \right). \quad (2)$$

The optimization problem has the following constraints:

$$\tilde{u}_{eml} = \tilde{u}_{e(m-1)l} + \sum_t b_{et} q_{tml} - \sum_k A_{ekl}^{month(m)} x_{kml},$$
$$e = 1:E, m = 1:M, l = 1:L, \quad (3)$$

$$\sum_{t,l} q_{tml} \le \bar{O}, m = 1:M, \quad (4)$$

$$\sum_{m=7(\mathbb{w}-1)+1}^{7\mathbb{w}} x_{kml} \le \sum_j \mathbb{z}_{jk\mathbb{w}l},$$
$$m = 1:M, k = 1:K, l = 1:L, \quad (5)$$

$$\mathbb{z}_{jk\mathbb{w}l} \le Q_{jk\mathbb{w}l}, j = 1:\mathcal{J}^*, k = 1:K, \mathbb{w} = 1:\mathbb{W}, l = 1:L, \quad (6)$$

$$\sum_l \left(u_{ml} + \sum_e \tilde{u}_{eml} \cdot len_e \cdot height \cdot width \right) \le \bar{u}, m = 1:M, \quad (7)$$

$$x_{kml}, q_{tml}, \mathbb{z}_{jk\mathbb{w}l} \in Z^+, \quad (8)$$

$$\tilde{u}_{eml}, u_{ml}, v_{imrl} \ge 0, \quad (9)$$

$$v_{imrl} \le V_{imrl}, i = 1:I, m = 1:M, r = 1:R, l = 1:L, \quad (10)$$

$$Bud_0 + \sum_{m=1}^{m^*} \left(\sum_{k,l} p_{kml} x_{kml} - \sum_{i,r,l} c_{imrl} v_{imrl} - \right.$$
$$\left. - \sum_{j,k,l} c_i \mathbb{z}_{jk\mathbb{w}(m)l} - FC \right) \ge 0, \ m^* = 1:M, \quad (11)$$

$$u_{ml} = u_{(m-1)l} + \sum_{i,r} v_{i(m-T_r)rl} - \mathbb{V} \sum_t q_{tml}, \quad (12)$$

$$Q_{jk\mathbb{w}l} = \left(\dot{Q}_{jk\mathbb{w}l} \pm \varepsilon \right) \cdot e^{\left(\sum_{m=7(\mathbb{w}-1)+1}^{7\mathbb{w}} \frac{\left(p_{k(m+7)l} - p_{kml}\right)}{7 \cdot p_{kml}} \right)}, \quad (13)$$

$$p_{k(m+1)l} = p_{kml} \cdot \left(1 + \varepsilon^{(1)} \right),$$
$$m = 1:M-1, \varepsilon^{(1)} = \left[-\varepsilon_1^{(1)}; \varepsilon_2^{(1)} \right], \quad (14)$$

$$A_{ekl}^{month(m)} = \max\left(0; \min\left(A_{ekl}^{month(m)}, A_{ekl}^{month(m)} + \varepsilon^{(2)} \right)\right),$$
$$\varepsilon^{(2)} = \left[-\varepsilon_1^{(2)}; \varepsilon_2^{(2)} \right], \quad (15)$$

$$FC_\mu \in \left[\alpha^1, \beta^1 \right], \quad (16)$$

$$Z_{ij\mathbb{w}} \in \left[\alpha^2, \beta^2 \right], \quad (17)$$

$$c_{ij\mathbb{w}} \in \left[\alpha^3, \beta^3 \right], \quad (18)$$

where ε, $\varepsilon^{(1)}$, $\varepsilon^{(2)}$ — uniformly distributed random variables of continuous type;

$\mathbb{V}, Bud_0, \tilde{u}_{e0l}, u_{0l}, A_{ekl}^{month(0)}, p_{kml} = const;$

$\varepsilon^{(1)} \le \varepsilon \le \varepsilon^{(2)}; 0 < \varepsilon_1^1 < 1; 0 < \varepsilon_2^1 < 1; \varepsilon_1^2 > 0; \varepsilon_2^2 > 0;$

$\varepsilon^{(1)}, \varepsilon^{(2)}, \alpha^1, \alpha^2, \alpha^3, \beta^1, \beta^2, \beta^3 = const.$

Let us consider in more detail the constraints (3–18) of the optimization problem.

Constraint (3) gives an indication of what stock of workpieces should be kept in stock throughout the planning period to ensure continuous production.

Constraint (4) sets the maximum number of sheets available for cutting per day, which is important for optimizing material usage and production.

Constraint (5) controls both the amount of product production during each week and the amount of transportation, although with a target function of the form (2) it can be treated as an equality. This constraint has a direct impact on inventory management and logistics.

Constraint (6) ensures that the transportation volume to the end points does not exceed the demand volume at those points, which is important for efficient product delivery and customer satisfaction.

Constraint (7) describes the degree of warehouse occupancy.

Further, constraints (8–9) define the type of variables, and (10) limits the amount of raw materials purchased daily to the size of bids on the exchange. These constraints are the basis for procurement planning and inventory management.

Constraint (11) ensures that daily profits are non-negative, which is important for the financial stability of the firm.

Constraint (12) determines the availability of raw material stocks in the warehouse, which is necessary to ensure uninterrupted production.

Restriction (13) reflects the interdependence of demand and prices for raw materials (unlike the similar function used in [13], the second derivative of this function does not have a gap).

Constraint (14) defines the recurrent dependence of price on the day number, which helps to take into account the dynamics of price changes on the market.

Constraint (15) reflects the rates of consumption of billets for the production of each unit of output, which is a key factor for the efficient use of resources and optimization of production processes.

Finally, constraints (16–18) are necessary to "play out" the values of fixed costs, the maximum throughput of the transportation graph in delivering goods to buyers, and the values of costs of goods delivery.

The model (3–18) is a stochastic nonlinear mathematical programming problem. To solve this problem, it is planned to iteratively search for a suboptimal solution using two subproblems: optimization of the production plan, delivery of raw materials and finished products, and search for a suboptimal price vector for product sales. This approach will make it possible to effectively manage complex processes and minimize the cost of production of products and their delivery to customers:

1) Generate price vectors ($p^{(iter)} = \left\{ p_{kml}^{(iter)} \right\}_{kml}$, $iter = 1 : Iter$) and estimate (13) and p_{kml}. $Cnt = 0$, $iter_1 = 0$, generate values for (15–18). Move to step 2.
2) Solve task (3–13) with fixed parameters $p^{(iter)}$, $A_{ekl}^{month(m)}$, FC_μ, $Z_{i\jmath\mathbb{w}}$, $c_{i\jmath\mathbb{w}}$ with the goal function (19) applying Chvatal-Gomory[1] algorithm. Let $iter_1 = iter_1 + 1$.

$$\pi(iter_1) = \left\{ \max_{x,v,\mathbb{Z}} \sum_m \left(\sum_{k,l} p_{kml}^{(iter)} x_{kml} - \sum_{i,l,r} c_{imrl} v_{imrl} - \sum_{\jmath,k,l} c_\jmath z_{\jmath k \mathbb{w}(m) l} \right) \right\}_{iter}. \quad (19)$$

Estimate (20) and move to step 3.

$$\tilde{p}_{kml} = p_{kml}^{\arg\max\left(\pi\left(iter_1\right)\right)}. \quad (20)$$

3) If $\arg\max(\pi(iter)) = 1$, then $Cnt = Cnt + 1$, otherwise $Cnt = 0$. Move to step 4.
4) If $Cnt = \mathcal{g}$, leave algorithm (end), otherwise move to step 5.
5) Generate price vectors (($p^1 = \{\tilde{p}_{kml}\}_{kml}$, $p^{(iter)} = \left\{ p_{kml}^{(iter)} \right\}_{kml}$, $iter = 2 : Iter$) (21–22).

$$\varepsilon^1 \in \left[-\left(\varepsilon_1^{(1)}\right)^{Cnt+1}, \left(\varepsilon_2^{(1)}\right)^{Cnt+1} \right]. \quad (21)$$

$$p_{kml}^{(iter)} = \min \left[\max \begin{pmatrix} \tilde{p}_{kml} \cdot \left(1 - \varepsilon_1^{(1)}\right); \\ \tilde{p}_{kml} \cdot \left(1 + \varepsilon^{(1)}\right) \end{pmatrix}; \tilde{p}_{kml} \cdot \left(1 + \varepsilon_2^{(1)}\right) \right]. \quad (22)$$

6) Estimate (13–14). Move to step 2.

The model (2–19) at played out values (13–19) belongs to the class of mixed-integer linear programming problems. We choose Matlab as the programming environment. At the first stage we use the method of branches and bounds to solve the problem (2–19)

[1] MathWorks. Documentation. Mixed-Integer Linear Programming Algorithms. https://it.mathworks.com/help/optim/ug/mixed-integer-linear-programming-algorithms.html

with the values (13–19) played out, and at the second stage (price change) the calculations are performed using the genetic algorithm. Note that the degree of closeness of the found price vector to the optimal one is achieved due to expression (22).

In practice, the program implementation in Matlab environment can be translated into any other programming language (for example, Python of any version) using, for example, programming language converters or neural networks (for example, ChatGPT, etc.).

4. Calibration

We will test the model using data from DNS-Les LLC2. At the end of each trading day, data on completed transactions is recorded. Based on these statistics, we will assess the feasibility of interaction between one of the large and at the same time young enterprises of the Primorsky Territory.

Based on the statistics of the enterprise, logging enterprises from five districts participated as sellers of raw materials in ensuring the uninterrupted operation of its operation: $r = 1{:}5$. On the planning horizon from Feb. 1, 2018 to Nov. 31, 2018, an array of the following data was received from the enterprise: prices (c_{imrl}), dates of the appearance of raw materials, volumes on these days, application prices (p_{kml}), number of applications for each type raw materials. In addition, the demand $\dot{Q}_{jkwl}$ in each week for each type of product is known based on the company's sales statistics [12].

All data that affects the input initial and constant values, the sheet cutting map and the standard costs of raw materials for the production of each unit of goods are presented in [12].

Assume that the price cannot change by more than ten percent every day. *Table 1* shows the limits of the values of random variables.

Table 1.

Stochastic values intervals

Parameters	$\varepsilon_1^{(1)}$	$\varepsilon_1^{(2)}$	$\varepsilon_2^{(1)}$	$\varepsilon_2^{(2)}$	$\varepsilon^{(1)}$	$\varepsilon^{(2)}$
Values	0.03	0.03	1	10	2	2

To find a solution, we will use the Matlab programming language, namely the *intlinprog* function[3]. The problem is of significant size, which makes it impossible to guarantee finding an optimal solution in a short time due to the algorithmic complexity (the number of calculations grows non-polynomially) of the branch and bound method, which is the default method for finding a solution in *intlinprog*. Therefore, we decided to limit the number of solution options to be explored to 107. If the algorithm fails due to the constraint despite finding a feasible solution, we consider it suboptimal rather than optimal. The results obtained are presented in *Figs. 1–5*.

Separately, we note that due to corporate ethics, the work does not indicate the actual recorded demand $\dot{Q}_{nkm}$. However, data regarding demand $ave\left(\sum_n Q_{nkm}\right)$ is presented in *Figs. 3, 4*.

5. Test results

Let us to look at *Figs. 1–5*. They reflect the main results of testing the model. In *Fig. 1* you can see how the profit value changed at each iteration $iter_1$. We introduce the notation $\pi_m(iter_1)$, which reflects the profit value by day m at iteration $iter_1$. The main change in the value of the objective function was observed in the first seven iterations, which is quite fast for a problem of this size. However, after the 2nd iteration, the value of the objective function changes slightly, which indicates that the new method solves problems of this kind more effectively compared to the algorithm in [13].

---

[2] Official website of the LLC DNS-Les Enterprise. Russia, Primorsky Krai, Spassk-Dalniy. http://dns-les.ru/

[3] Intlinprog. Documentation. MathWorks. https://www.mathworks.com/help/optim/ug/intlinprog.html

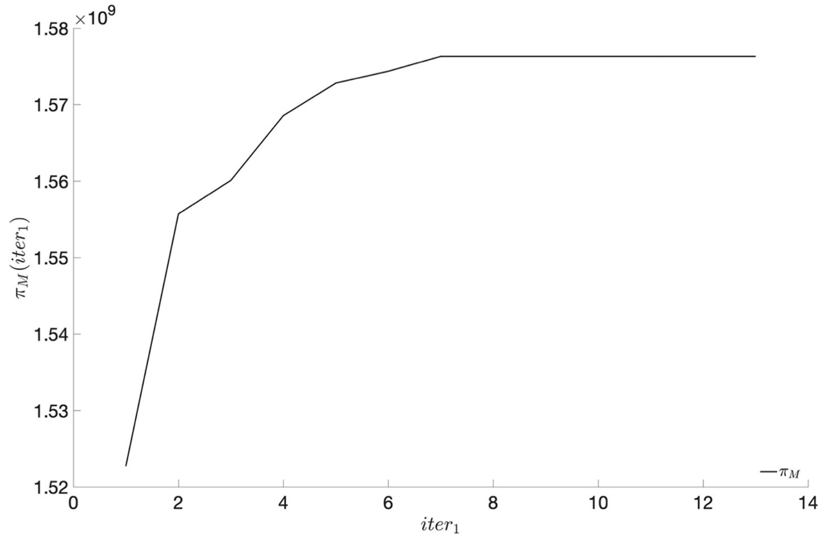

Fig. 1. Visualization of profit values depending on the iteration number.

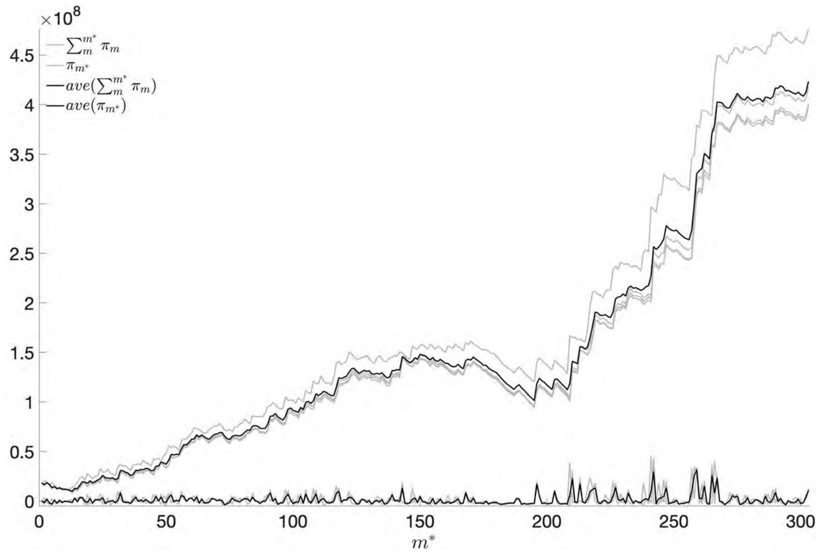

Fig. 2. Visualization of the testing process and the resulting solution.

Any solution requires a stability test. To do this, we will re-launch the developed model 4 times and find solutions. The results are presented in *Fig. 2.* Despite the rather large gap in the value of the objective function at the end of the planning horizon between solutions, the deviations in relative values are insignificant for all trajectories of values π_m for each decision.

Of particular interest is the observation that after the end of the summer period, closer to the winter period, the profit of the enterprise increases. This is due to the fact that production volumes increase in the winter and prices for raw materials fall.

An important factor is not only the prices of raw materials, but also the prices of final consumer goods

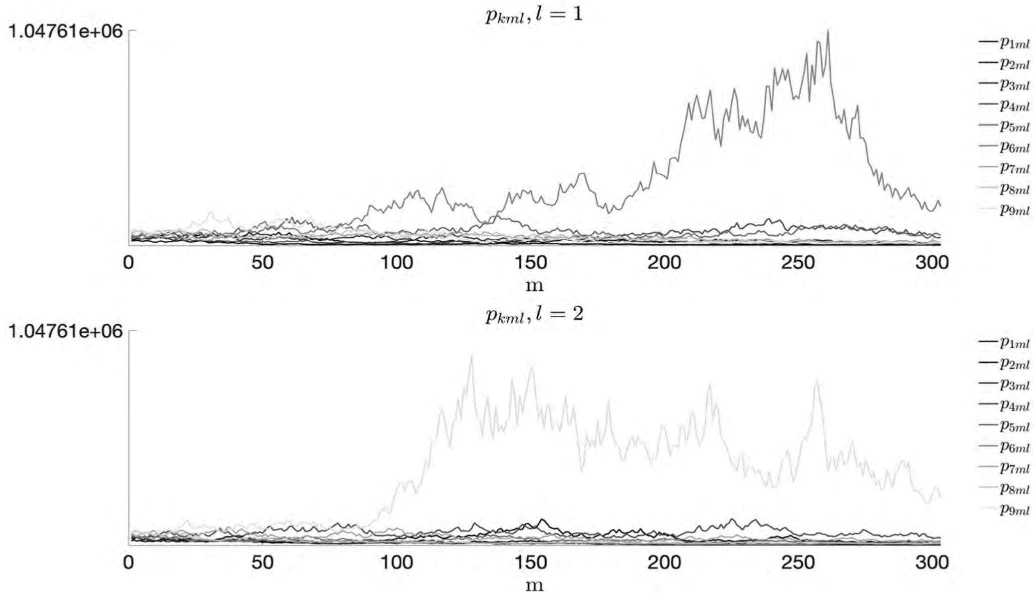

Fig. 3. Visualization of price behavior over the entire planning horizon.

that the enterprise produces. To do this, let's look at *Fig. 3*. It clearly shows that the "market" considers the most interesting product to be goods No. 6, made from raw materials of type $l = 1$, and No. 9, made from raw materials of type $l = 2$. This may be due to the cost of raw materials for the production of this product $A_{ekl}^{month(m)}$.

However, rising prices also imply a change in demand for goods. We introduce the notation Q^{0}_{jkwl}, which sets the trajectory of the changed volume of demand in accordance with the new calculated price values. *Figures 4, 5* show the behavior of demand for goods over the entire planning horizon on a weekly basis.

Of particular interest are the volumes of waste (cubic meters) that appear when cutting incoming raw materials. To do this, consider *Fig. 6*. Here, as with the behavior of the profit trajectory depending on the decision, you can observe a slight deviation from the average values for waste. Minor deviations in both cases with profit and waste means a high degree of predictability of the situation in the enterprise, which is an extremely important characteristic in the current turbulent times.

Let us consider the positive and negative aspects of the developed model and algorithm and propose options for their modification.

The following can be attributed to the advantages of the proposals developed:

1. Increased speed of solution search compared to the scheme proposed in [13].

2. Taking into account the technology of round wood cutting.

3. Compared to [13], the second derivative of the demand function is continuous, which guarantees that there is no instantaneous acceleration of the demand value when crossing the point zero.

The negative sides include the following:

1. Despite the fact that the function (13) has continuous first and second derivatives, it has certain drawbacks. *Figures 4*, *5*, which reflect the relationship between demand and prices, show that prices for the products that are of most interest to the market are rather overvalued and remain at this level for quite a long time. This is due to the fact that demand

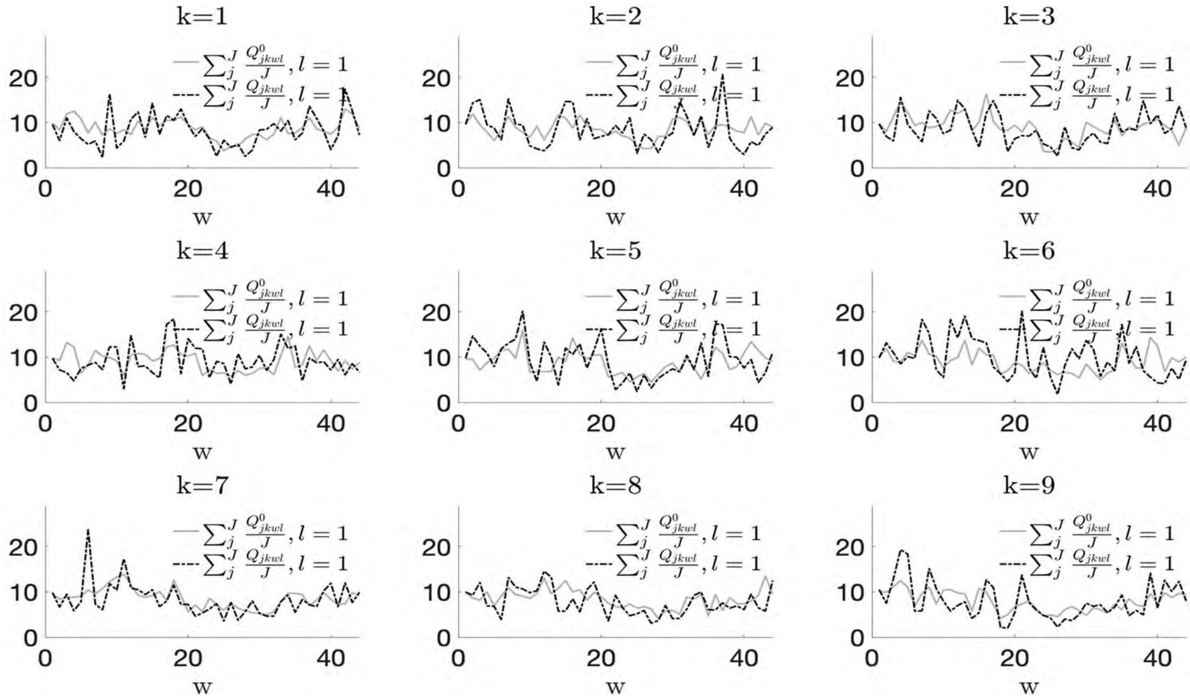

Fig. 4. Visualization of the volume of initial demand and the current volume of demand, which changed when searching for the price vector for $l = 1$.

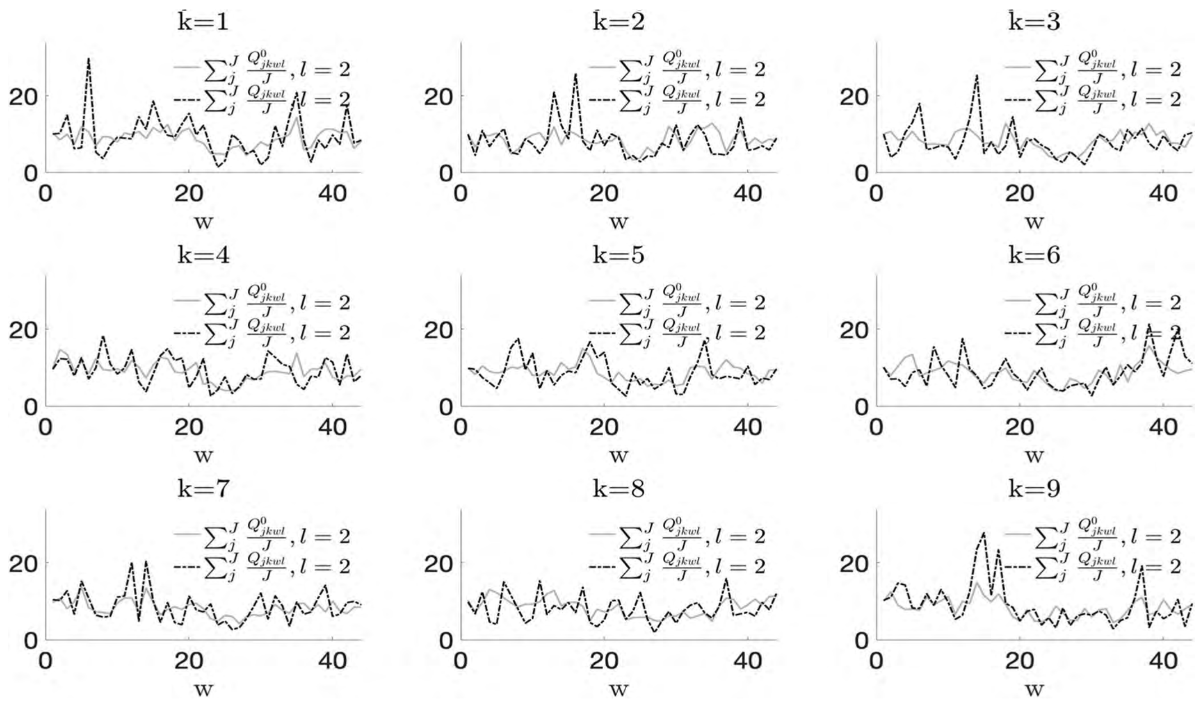

Fig. 5. Visualization of the volume of initial demand and the current volume of demand, which changed when searching for the price vector for $l = 2$.

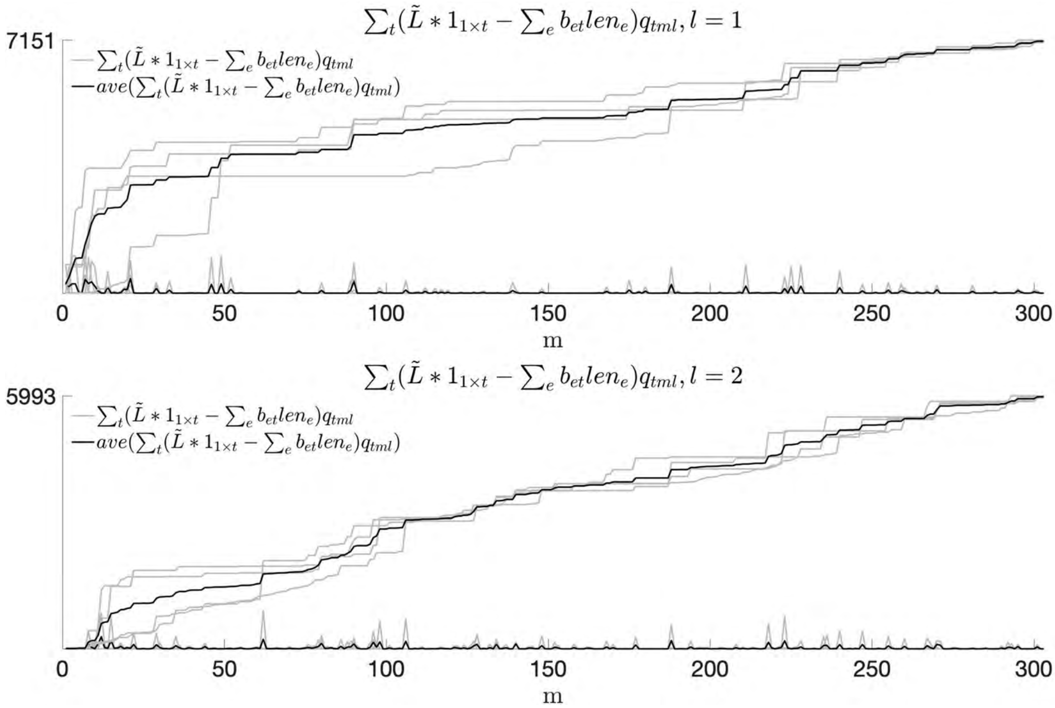

Fig. 6. Visualization of the volume of total waste of raw materials.

on the respective days decreased only during the period of price growth, after which its value kept in a small neighborhood of the initial value, since prices almost did not change. However, this is not possible in the real economy, as demand is also under pressure from high prices during subsequent planning periods. In order to take this aspect into account, it is necessary to modernize the demand-price relationship function.

2. The convergence of the algorithm has not been investigated.

3. Rapidly increasing dimensionality of the problem.

4. Price changes rarely occur every day.

5. To achieve diversification of raw material sources, the ability to purchase raw materials from the regional market should be added, as described in [13].

6. There is no consideration of the possibility of making borrowings.

Conclusion

The paper presents a mathematical model for solving the complex problem of suboptimal formation of the pricing policy of a timber industry enterprise, taking into account the process of developing stable supply chains from the commodity and raw material exchange and calculation of production volumes and transportation of finished products to consumers. This model is different in that it allows us to take into account the technology of cutting raw materials, which is important for this industry.

The central objective of the model is to maximize the operating profit of the enterprise, which is a complex mathematical problem. The model takes into account various aspects such as production rates of raw material consumption, procurement strategy at the commodity exchange, transportation volumes of finished products and pricing policy of the company. To solve such a complex problem, a two-stage optimization scheme including linear optimization and application of genetic algorithm has been developed.

The model was tested using real data from the timber processing complex of the Primorsky Territory. As the testing showed, the new optimization scheme allows us to find a solution faster than the previous version of the model based on the gradient descent method.

The results of the experiments made it possible to formulate valuable recommendations for the company's management staff in terms of interaction with loggers and improvement of production processes.

Further research may include modifying the model to account for probabilistic economic factors and improving solution search methods. It is also necessary to consider the possibilities of accelerating the search for a solution and to develop more accurate demand functions for the company's products. ■

References

1. Wieruszewski M., Turbański W., Mydlarz K., Sydor M. (2023) Economic efficiency of pine wood processing in furniture production. *Forests*, vol. 14, article 688. https://doi.org/10.3390/f14040688
2. Trigkas M., Papadopoulos I., Karagouni G. (2012) Economic efficiency of wood and furniture innovation system. *European Journal of Innovation Management*, vol. 15, no. 2, pp. 150–176. https://doi.org/10.1108/14601061211220959
3. Larsson M., Stendahl M., Roos A. (2016) Supply chain management in the Swedish wood products industry — a need analysis. *Scandinavian Journal of Forest Research*, vol. 31, pp. 777–787. https://doi.org/10.1080/02827581.2016.1170874
4. Chang H., Han H.-S., Anderson N., et al. (2023) The cost of forest thinning operations in the Western United States: A systematic literature review and new thinning cost model. *Journal of Forestry*, vol. 121, no. 2, pp. 193–206. https://doi.org/10.1093/jofore/fvac037
5. Buka-Vaivade K., Serdjuks D., Pakrastins L. (2022) Cost factor analysis for timber–concrete composite with a lightweight plywood rib floor panel. *Buildings*, vol. 12, no. 6, article 761. https://doi.org/10.3390/buildings12060761
6. Aryal U., Neupane P.R., Rijal B., Manthey M. (2022) Timber losses during harvesting in managed Shorea Robusta forests of Nepal. *Land*, vol. 11, article 67. https://doi.org/10.3390/land11010067
7. Numazawa C.T.D., Numazawa S., Pacca S., John V.M. (2017) Logging residues and CO_2 of Brazilian Amazon timber: Two case studies of forest harvesting. *Resources Conservation and Recycling*, vol. 122, pp. 280–285. https://doi.org/10.1016/j.resconrec.2017.02.016
8. Zubair M., Abbas Z., Hussain S. B. (2022) Factor affecting purchase of quality wood: Understanding perceptions of wood workers using logistic regression model. *Asian Journal of Research in Agriculture and Forestry*, vol. 8, pp. 243–248. https://doi.org/10.9734/ajraf/2022/v8i4184
9. Hosseini S.M., Peer A. (2022) Wood products manufacturing optimization: A survey. *IEEE Access*, vol. 10, pp. 121653–121683. https://doi.org/10.1109/ACCESS.2022.3223053
10. Chen J., Yu H., Jiang D., et al. (2021) A novel NIRS modelling method with OPLS-SPA and MIX-PLS for timber evaluation. *Journal of Forestry Research*, vol. 33, pp. 369– 376.
11. Stefańska A., Cygan M., Batte K., Pietrzak J. (2021) Application of timber and wood-based materials in architectural design using multi-objective optimization tools. *Construction Economics and Building*, vol. 21, pp. 1–5. https://doi.org/10.5130/AJCEB.v21i3.7642

12. Rogulin R.S. (2021) The influence of the cutting subtask when assessing the feasibility of purchasing forest raw materials from commodity exchanges. *Bulletin of Voronezh State University. Series: System analysis and information technologies*, no. 2, pp. 109–125 (in Russian).

13. Rogulin R.S. (2021) Mathematical model for the formation of pricing policy and plan for the production and transport system of a timber industry enterprise. *Business Informatics*, vol. 15, no. 3, pp. 60–77. https://doi.org/10.17323/2587-814X.2021.3.60.77

14. Lapshin V.P., Turkin I.A., Omelechko V.Yu. (2021) Development of a mathematical model for drilling wood taking into account the mutual influence of electric drives for feeding and cutting. *Electrical Technologies and Electrical Equipment in the Agro-industrial Complex*, no. 2, pp. 48–53 (in Russian).

15. Jiang S., Buck D., Tang Q., et al. (2022) Cutting force and surface roughness during straight-tooth milling of walnut wood. *Forests*, vol. 13, article 2126. https://doi.org/https://doi.org/10.3390/f13122126

16. Curti R., Marcon B., Denaud L., et al. (2021) Generalized cutting force model for peripheral milling of wood, based on the effect of density, uncut chip cross section, grain orientation and tool helix angle. *European Journal of Wood and Wood Products*, no. 79, pp. 667–678. https://doi.org/10.1007/s00107-021-01667-5

17. Merve E.K., Abhijeet Gh., Umit S.B. (2021) Modelling the impact of climate change risk on supply chain performance. *International Journal of Production Research*, vol. 59, pp. 7317–7335. https://doi.org/10.1080/00207543.2020.1849844

18. Abdullah A.H.A., Umar M.M., Ali Ak.Sh., Irfan A. (2021) Multi-objective optimization modelling of sustainable green supply chain in inventory and production management. *Alexandria Engineering Journal*, 2021, vol. 60, pp. 5129–5146. https://doi.org/10.1016/j.aej.2021.03.075

19 Astanti R.D., Daryanto Y., Dewa P.K. (2022) Low-carbon supply chain model under a vendor-managed inventory partnership and carbon cap-and-trade policy. *Journal of Open Innovation Technology Market and Complexity*, vol. 8, article 30. https://doi.org/10.3390/joitmc8010030

20. Pichler G., Sandak J., Picchi G., et al. (2022) Timber tracking in a mountain forest supply chain: A case study to analyze functionality, bottlenecks, risks, and costs. *Forests*, vol. 13, article 1373. https://doi.org/10.3390/f13091373

21. Abdollah B., Peiman Gh., Adel P.Ch., et al. (2022) A new wooden supply chain model for inventory management considering environmental pollution: A genetic algorithm. *Foundations of Computing and Decision Sciences*, vol. 47, pp. 383–408. https://doi.org/10.2478/fcds-2022-0021

22. Rogulin R.S. (2022) Review of applied fundamentals of using data analytics and machine learning in demand forecasting. *Economic and Social-Humanitarian Studies*, no. 3, pp. 115–126 (in Russian).

About the author

Rodion S. Rogulin

Cand. Sci. (Econ.);

Associate Professor, Department of Mathematics and Modeling, Vladivostok State University, 41, Gogolya St., Vladivostok 300012, Russia;

E-mail: rafassiaofusa@mail.ru

ORCID: 0000-0002-3235-6429

25th YASIN (APRIL) INTERNATIONAL ACADEMIC CONFERENCE ON ECONOMIC AND SOCIAL DEVELOPMENT

Dear colleagues,

HSE University is pleased to announce a call for applications to present academic reports at the **25th Yasin (April) International Academic Conference on Economic and Social Development** (hereinafter, the “25th Yasin Conference” or the “Conference”).

All events on the 25th Yasin Conference programme will take place in Moscow **from April 15 till 18, 2025**. All the information is available in the Conference webpage at: **https://conf.hse.ru/en/2025/**.

The Conference programme will focus on the following key academic themes:

- Economics;
- Human Capital and Society;
- Instrumental Methods in the Social Sciences;
- Foresight Research;
- International Relations.

Applications to present academic reports at the Conference should be submitted in accordance with the following thematic sections:

Under the academic theme of Economics:

- Macroeconomics and Economic Growth;
- Methodology of Economic Sciences;
- Theoretical Economics;
- Companies and Markets;
- Finance and Banking.

Under the academic theme of Human Capital and Society:

- Social Policy and Healthcare;
- Demography and Labour Markets;
- Political Processes;
- Sociology;
- Psychology.

Under the academic theme of Instrumental Methods in the Social Sciences:

- Instrumental Methods in Economic and Social Research.

Under the academic theme of Foresight Research:

- Russia’s Growth Scenarios in the Context of a Rapidly Changing External Environment;
- New Models and Methods for Technological and Socioeconomic Forecasting;
- International Symposium on “Foresight Research Amid Rapid Global Change”.

Under the academic theme of International Relations:

- Ideas and Civilizations in a Multipolar World;
- The World Economy;
- Asian Studies.

The deadline for filing applications to present academic reports at the Conference is **Monday, January 20, 2025.** In addition, applications will also be accepted to register for the Conference as an attendee: from Russian citizens – **by Thursday, April 10, 2025;** from foreign citizens – **by Monday, March 10, 2025**.

The events on the Conference programme will be held in Russian or English. Certain discussions will be held in a bilingual format with simultaneous interpretation provided.

The Conference will take place mainly in a live, face-to-face format. However, in exceptional cases, the Programme Committee retains the right to invite certain speakers and participants to join discussions online.

How can I apply to present an individual academic report at the Conference?

Reports delivered under specific thematic sections must describe the results of original research performed with the use of contemporary methodologies. Report presentations should last from 15 to 20 minutes.

The final deadline for online filing of applications to present reports via the HSE University Conference system is **Monday, January 20, 2025.**

All applications must be accompanied by **detailed abstracts**. Please see the requirements for formatting academic report abstracts below:

- The abstract must clearly state the issue at hand, as well as the level at which it has been analysed, the selected research methods, and the key findings, as well as their novelty and validity;
- The abstract format should be either Word or RTF;
- The character count should be between 3,000 and 7,000 characters;
- For Russian-language reports, the accompanying abstracts must be written in both Russian and English, while English-language reports require abstracts in English only.

Speakers may present **one individual report** and a maximum of two co-authored reports.

How can I submit a group application?

A team of authors filing individual applications, which are duly registered in the HSE University Conference system, may file a request with the Programme Committee, by Monday, **January 20, 2025**, to present their reports at the same session. For this purpose, they must submit a group application form via the HSE University Conference system.

Group applications should meet the following requirements:

- from at least two up to a maximum of five reports to be delivered during the given session;
- no more than two reports presented by representatives of the same organisation;
- all individual applications of group members must be registered in the Conference system.

Please note that an application to present a single report with co-authors is deemed an individual application and will not be considered as a group application.

Once applications have been submitted for the respective thematic sections of the Conference, individual sessions will be put together for each thematic section. Each session will last for 90 minutes. The Programme Committee may take into consideration proposals concerning the content of specific sessions during the expert review of applications stage and in the process of drawing up the final Conference programme.

Expert reviews of applications

Applications will be selected in three stages as follows:

1. First, all applications that do not meet the aforementioned formatting criteria will be excluded;
2. Second, expert assessments of report abstracts will be performed, with a specific focus on verifying the novelty and validity of findings (an application may be reassigned to another thematic section by decision of the head of the given expert group);
3. Third, the Programme Committee will make its final decisions on the inclusion of reports in the Conference programme based on the expert assessments.

Results of the expert review and confirmation of an applicant's participation

Applicants will be informed of the decisions to include their reports in the programme of the 25th Yasin Conference, or to decline their applications, as soon as the respective expert review results are available, but no later than by **Wednesday, March 19, 2025**.

Authors of reports selected for presentation at the Conference must confirm (RSVP) their participation by **Wednesday, March 26, 2025** (inclusive), via their personal account in the HSE University Conference system, and then, by **Friday, April 4, 2025** (inclusive), upload their presentation slides in English to their personal account. This process is a mandatory prerequisite for including reports in the final version of the programme.

Registration fees

The Conference registration fee for both presenters and attendees is RUB 3,000.

The following participants will be exempt from the registration fee:

- Students and PhD students from any university (upon presentation of their student ID and completing the registration process);
- HSE University staff members (upon presentation of their staff ID badge and completing the registration process);
- Participants who were individually invited by the Conference Programme Committee (e.g., honorary academic presenters, moderators, speakers at roundtables, discussion participants, and other Conference guests).

Detailed information on how to pay the registration fee is available in the Registration Fee Payments section of the Conference webpage.