



Munich Personal RePEc Archive

**Forecasting the main economic indicators
for industry in the analytical system
"Horizon"**

Kitova, Olga and Dyakonova, Ludmila and Savinova,
Victoria and Fomin, Kiril

Plekhanov Russian University of Economics

16 October 2023

Online at <https://mpra.ub.uni-muenchen.de/118887/>
MPRA Paper No. 118887, posted 26 Oct 2023 04:52 UTC

Forecasting the main economic indicators for industry in the analytical system "Horizon"

Kitova, Olga

Associate professor, doctor of economics, head of the department of Informatics
Higher School of Cybertechnologies, Mathematics and Statistics
Plekhanov Russian University of Economics, Moscow, Russian Federation
Plekhanov Russian University of Economics, 36 Stremyanny Lane, office 441 (building 3),
Moscow, 117997, Russian Federation tel. 7 (499) 237-85-20
[0000-0002-1820-0954]
Kitova.OV@rea.ru

Dyakonova, Ludmila

Associate Professor, Cand. Sc. (Physics&Mathematics), Associate Professor of the department
of Informatics
Higher School of Cybertechnologies, Mathematics and Statistics
Plekhanov Russian University of Economics, Moscow, Russian Federation
Plekhanov Russian University of Economics, 36 Stremyanny Lane, office 420 (building 3),
Moscow, 117997, Russian Federation. tel. 7 (499) 236-73-73
[0000-0001-5229-8070]
Dyakonova.lp@rea.ru

Savinova, Victoria

Senior Lecturer of the Department of Informatics
Higher School of Cybertechnologies, Mathematics and Statistics
Plekhanov Russian University of Economics, Moscow, Russian Federation
Plekhanov Russian University of Economics, 36 Stremyanny Lane, office 441 (building 3),
Moscow, 117997, Russian Federation tel. 7 (495) 958-24-10
[0000-0002-0036-3675]
Savinova.VM@rea.ru

Fomin, Kiril

graduate student
Higher School of Cybertechnologies, Mathematics and Statistics
Plekhanov Russian University of Economics, Moscow, Russian Federation
Plekhanov Russian University of Economics, 36 Stremyanny Lane, office 420 (building 3),
Moscow, 117997, Russian Federation. tel. 7 (499) 236-73-73

Abstract

Industrial development is of great strategic importance for ensuring sustainable economic growth in Russia and solving social problems. Therefore, the development of approaches and methods for comprehensive analysis and forecasting of industrial indicators at the national and regional level is of particular importance, which will facilitate the adoption of scientifically based decisions in the field of industrial planning and management. What is needed is a system of indicator models that will allow for a comprehensive analysis of industrial development, identification of the main influencing factors, as well as the development of forecasting models and methods and their application to the indicators under study.

The hybrid forecasting system "Horizon", developed by the authors of the study, implements regression and intelligent models for most groups of indicators of the Russian economy. At the

same time, most researchers rely in their studies on autoregressive time series models based on ARIMA. The authors have developed a new module of the Horizon ARIMA system, which can be used when forecasting individual indicators. These forecasts can be considered as baseline when conducting comparative analysis with hybrid models.

This study is devoted to forecasting a group of main economic indicators of Russian industry using ARIMA time series models.

Key words: socio-economic indicators of the Russian Federation, industry indicators, forecasting, time series, hybrid information and analytical system.

Introduction

Russian industry is one of the key sectors of the country's economy. The development of this industry is of great strategic importance for ensuring sustainable economic growth and solving social problems. Sustainable development of the industrial complex at the present stage requires the implementation of a scientifically based industrial policy that ensures the attraction of investments, increased competitiveness in the foreign and domestic markets, the creation of a modern technological base and the implementation of structural reforms.

Therefore, the development of approaches and methods for comprehensive analysis and forecasting of industrial indicators at the national and regional level is of particular importance, which will facilitate the adoption of scientifically based decisions in the field of industrial planning and management. It is necessary to develop a system of indicator models that will provide the ability to conduct a comprehensive analysis of industry indicators, as well as identify the main influencing factors. Currently, such a system of models does not exist.

The main approach to building models is econometric approach. The hybrid forecasting system "Horizon", developed by the authors of the study, implements intelligent models along with regression ones. The system is open both from the point of view of including new indicator models and new modules that carry out forecasting based on data mining.

Our research is devoted to forecasting a group of main economic indicators of industry in Russia using ARIMA time series models.

Literature Review

Most often, statistical and econometric methods are used to predict socio-economic processes. Statistical and econometric methods involve building and testing many models for each time series, comparing them based on statistical quality indicators, and selecting the best one for forecasting. Among these methods, forecast extrapolation by trend, adaptive methods, regression methods, and time series analysis methods are distinguished (Dougherty, 2011).

The development of empirical models intended for applied macroeconomic forecasting is described in (Pestova and Mamonov, 2016), which compares modern structural and non-structural methods.

Forecasts of macroeconomic indicators, including industrial indicators, based on various methods of time series analysis have become widespread. Such analysis includes identifying trends and seasonal components, building adaptive models, models of stationary and non-stationary time series. Although a number of works use simple trend forecasts (Krasnyuk, 2013, Kalinovskaya and Kerryev, 2019), in most cases, researchers use the autoregressive model with integration and moving average ARIMA and its modification highlighting seasonality SARIMA.

Researchers from the Institute of Economic Policy named after E.T. Gaidar, over the last decade, forecast calculations have been carried out using ARIMA for the main macroeconomic indicators, including the index of industrial production of the Russian Federation (IPI) (Astafieva and Turuntseva, 2014, 2021), and the results of comparing the quality of the forecasts obtained with alternative forecasts (naive, naive seasonal and constructed using moving average). An

aggregated index of production by type of economic activity “Mining”, “Manufacturing”, “Providing electricity, gas and steam” is used as a projected monthly index of industrial production; Air conditioning”, “Water supply; drainage, organization of waste collection and disposal, pollution elimination activities”, presented on the website of the Federal State Statistics Service. The used methodology for analyzing the comparative quality of forecasts is described in detail in (Turuntseva and Kiblitckaya, 2010). In (Astafieva and Turuntseva, 2021), which includes a review of IPI forecasts for 2009-2021, it is noted that forecasts of the IPI indicator in general and by industry based on time series models demonstrate a lower level of errors in comparison with all the simplest forecasting methods. Work (Turuntseva et al., 2014) presents calculations of forecast values for a number of economic indicators of the Russian Federation in May-October 2022, performed on the basis of formal models of ARIMA time series, taking into account the existing trend and, in some cases, its significant changes. The constructed forecasts were compared with the results obtained on the basis of market surveys.

In (Khairutdinov, 2022) forecasting using the ARIMA model of the main indicators for the petrochemical industry: the volume of investments, the average annual number of personnel, the volume of shipped petrochemical products is discussed.

In (Mirolyubova et al., 2021) models were developed for forecasting demand volume for six product groups of a large industrial enterprise based on ARIMA models using Statistica package.

Many researchers use regression methods to predict socio-economic indicators.

(Petraikova and Zinovkina, 2019) conducted a correlation-regression analysis of the dependence of the volume of oil produced on such factors as the average annual number of industrial production personnel, the volume of production drilling for oil and the average daily amount of oil produced at a well, the degree of depreciation of fixed assets; commissioning of production facilities (oil wells); fixed assets renewal ratio; level of profitability.

(Sokolova and Sidorov, 2022) describe a regression model for predicting the volume of industrial production in a region (Vologda region) based on indicators of investment in fixed capital of industrial production and the number of people employed in industry.

In many studies, the authors use vector autoregression models and intelligent methods and models based on artificial neural networks (ANN) and regression decision trees to predict industrial indicators.

In (Treshchevsky et al., 2020) the results of forecasting production volumes in the manufacturing industry of the region (Voronezh region) based on neural networks. Data for 12 years (from 2005 to 2016) were chosen as the basis for forecasting are presented. As the authors correctly note, it is impossible to use a longer period, since until 2005 a different method of classifying types of economic activity (OKONKH) was used. OKVED began to be used precisely in 2005. For forecasting, 6 indicators of the region's manufacturing industry were taken, as well as 8 indicators from the development strategy of the Russian Federation. However, the authors were unable to build a satisfactory ANN-based model for the factors under study.

In (Boldyrevsky et al., 2019) a study was conducted of the time series of IPI of the manufacturing industry for the constituent entities of the Russian Federation based on the data of the Federal State Statistics Service (Rosstat) from 1999 to 2016. Clustering by constituent entities of the Russian Federation revealed two clusters: regions resistant to external changes and regions highly dependent on supply and sales markets. Next, predictive IPI models based on neural networks were built for each cluster.

In (Shinkevich et al., 2022) methods of descriptive statistics, correlation and regression analysis, polynomial trend line, and neural network training were used to predict indicators of trends in scientific research development of industry. Neural network models achieved the highest accuracy of predictions.

Of great interest is the study (Fokin, 2020), in which a large vector autoregression with regularization is constructed on monthly data of Russian macroeconomic indicators. Data on industrial production indices, producer prices, investments, exports, imports, interest rates,

indicators of the consolidated and federal budgets, etc. were used. In the work pseudo-out-of-sample forecasts for industrial production indices based on VAR-LASSO model, and the quality of the obtained forecasts is compared with the quality of forecasts using the classical ARIMA model. The author concludes that the constructed model outperforms ARIMA predictions. It should be noted that in the study a scenario approach is implemented: the real oil price was used as an exogenous (scenario) indicator.

The authors of this study are developing a system of hybrid models “Horizon”, which includes a regression model module, a neural network module and a forecasting module based on decision trees (Kitova et al., 2016, 2019, 2020), which allows you to build models for 500 or more socio-economic indicators of the Russian Federation. The constructed large country model of variant forecasting of key indicators of the Russian economy currently contains 8 interconnected blocks of indicators characterizing the relevant sectors of the economy. To implement variant forecasting, the model contains a block of expert indicators of scenario forecast conditions.

In our work (Kitova et al., 2021), we built and verified an ensemble of forecasting models for a block of industrial indicators of the Russian Federation based on a system of linear regression equations and artificial neural networks using the Horizon information and analytical system (2001-2021). This article examines 121 industrial indicators of the Russian Federation and constructs a system of interconnected multiple linear regression equations.

Regression models showed satisfactory results for more than 64% of indicators. For those indicators for which it was impossible to build a regression forecasting model with satisfactory accuracy and quality, the results were improved using artificial neural networks (multilayer perceptron). Scenario forecasting of the entire group of indicators for the 1st-4th quarters of 2021 showed good results overall. The theoretical significance of the results obtained lies in the development of a flexible predictive ensemble of machine learning models for scenario forecasting of industrial indicators of the Russian Federation.

However, as discussed above, many researchers use the ARIMA autoregressive model as a base model, which gives fairly good results. We have developed and included an ARIMA module in the Horizon system, which will allow us to further compare calculations using complex hybrid models with this model when forecasting socio-economic indicators.

In this study, we use the ARIMA module to generate forecasts for key industrial indicators.

Materials and methods

Structure of Russian industry indicators

The list of all indicators that are observed by Rosstat is presented in the regulatory and administrative document - the Federal Plan of Statistical Work (hereinafter - FPSW). At the moment, the FPSW includes more than 4 thousand indicators and surveys, including about 1.5 thousand indicators related to industry.

The structure of industrial indicators of the Russian Federation and regions is a systematic description of key indicators characterizing the activities of industrial enterprises in Russia and its regions.

Within the framework of the indicator structure, the following groups of industry indicators are analyzed:

- Main economic indicators of industry
- Enterprises and industrial organizations
- Material and technical base and investments in industry
- Labor in industry
- Energy consumption
- Industry prices
- Environmental protection

Within the framework of this study, the group “Main economic indicators of industry” is studied. Table 1 shows the industry indicators presented in the FPSW for this group.

Table 1 – Main economic indicators of the industrial sector

N	Group of indicators
1.	Indicators of the Russian Federation economy
2.	Main indicators of the organization's performance in foreign trade activities of industrial production
3.	Selected performance indicators of foreign trade organizations by type of ownership
4.	Structure of the volume of shipped goods of own production, performed works and services by own forces in foreign trade activities
5.	Turnover of organizations in foreign trade activities
6.	Gross value added by foreign trade activities
7.	The volume of shipped goods of own production, performed works and services of own. by foreign trade forces
8.	Gross value added by foreign trade activities
9.	Production index by foreign trade activities
10.	Average annual number of employees of organizations engaged in foreign trade activities
11.	Concentration of production by foreign trade activities
12.	The level of use of the average annual production capacity of organizations producing individual types of products
13.	Organizational Entrepreneurial Confidence Index
14.	Assessment of the general economic situation at foreign trade enterprises
15.	Assessment of the main factors limiting production growth at foreign economic activity enterprises

The industry indicators described above are collected and stored at several time periods: month, quarter and year.

Studying a set of groups of industrial indicators allows us to obtain a characteristic of the technological and economic level of industry in the Russian Federation and its regions, as well as draw conclusions about the effectiveness of investments and the development strategy of individual industries.

Forecasting methods

ARIMA

Along with the regression model, Horizon implements the widely used ARIMA time series forecasting method.

Time series prediction methods using ARIMA were used for time series forecasting. We also used time series methods to analyze seasonality and trends [8].

ARIMA (Autoregressive Integrated Moving Average) is a statistical model that is used for time series analysis and forecasting. The ARIMA model is a combination of the autoregressive (AR) and moving average (MA) methods, as well as the integration method (I), designed to work with non-stationary time series.

The ARIMA model has three main parameters: p, d, q.

Parameter p (AR, autoregressive) specifies the order of autoregression and describes how many past values of the time series will be used to predict its future values. The larger the p-value, the more difficult it is to determine the model, but the accuracy of the predictions increases.

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + \varepsilon_t,$$

where:

Y_t – level of the time series at time t (dependent variable);

$Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ – levels of the time series at points in time $t-1, t-2, \dots, t-p$, respectively (independent variables);

$\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_p$ – estimated coefficients;

ε_t – random disturbance describing the influence of variables not taken into account in model.

The coefficient φ_0 determines the constant level of the series and is related to the mathematical expectation μ by the relation $\varphi_0 = \mu(1 - \varphi_1 - \varphi_2 - \dots - \varphi_p)$

Parameter d (I, integration) determines the order of integration and shows how many times the difference between the values of the time series must be applied to make it stationary. Stationarity means that the mean and variance of a time series do not depend on time. If the time series is not stationary, this can lead to forecasting errors. The higher the value of d , the more times it must be applied to the time series to make it stationary.

Parameter q (MA, moving average) determines the order of the moving average and determines the number of past errors that will be used to predict future values of the time series. The higher the q value, the more errors are averaged to predict future values.

$$Y_t = \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \dots - \omega_q \varepsilon_{t-q},$$

where,

Y_t – series level at time t ;

ε_{t-i} – values of residuals of i time periods ago (independent variables);

$\omega_1, \omega_2, \dots, \omega_q$ – estimated coefficients.

Moving average MA models predict Y_t values based on a linear combination of a limited number q of residuals, while autoregressive AR models predict Y_t values based on a linear function approximation of a limited p number of past Y_t values.

Using the concept of a moving average in this case means that the deviation of the dependent variable from its average, i.e. the value $Y_{t-\mu}$, is a linear combination of current and past values of the random disturbance vector.

The ARMA (p, q) model has the general form:

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + \dots + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \dots - \omega_q \varepsilon_{t-q}$$

These three parameters define the ARIMA model class (p, d, q), which in turn describes the characteristics of the time series and the number of parameters that are needed to accurately predict future values. Different ARIMA classes can be used for different types of time series, depending on their characteristics.

ARIMA models can be used to predict future values of a time series and to estimate the likelihood of certain values occurring in the series. They can also be used to analyze and study the structure of time series, highlighting the trend, cyclicity and seasonality in the series.

To create an ARIMA model, it is necessary to collect a sufficient amount of data and conduct preliminary analysis to determine the structure and stationarity of the time series. The ARIMA model can also be extended by adding exogenous variables that can improve the quality of the forecast.

ARIMA (Auto Regressive Integrated Moving Model) and FA (Regression Factor Analysis) are popular methods for time series forecasting. They have the following advantages:

Advantages of ARIMA:

- ARIMA is used to model linear time series that have a trend, seasonality and a random component.
- ARIMA can be used to forecast and analyze data of any frequency (hourly, daily, weekly, monthly or annual).
- ARIMA allows you to model dependencies between different time periods, leading to more accurate forecasts.

Results

ARIMA

The time series forecasting module based on the ARIMA model is written in Python. In order for experts to carry out in-depth analysis, an expert report is generated in the system using VBA in the form of an Excel file: the original series, data excluding the seasonal component, trend and data visualization in the form of charts (see Figure 1).

The report contains the following information:
ARIMA model.

- Bayesian Information Criterion (BIC). The lower the BIC value, the better the model. BIC takes into account both the quality of the model and its complexity, which makes it more preferable than other statistical criteria.
- Setting interval – shows for what period of time the data for the indicator being studied was used.
- Percentage of emissions present, with date.
- Linear diagram of dynamics.

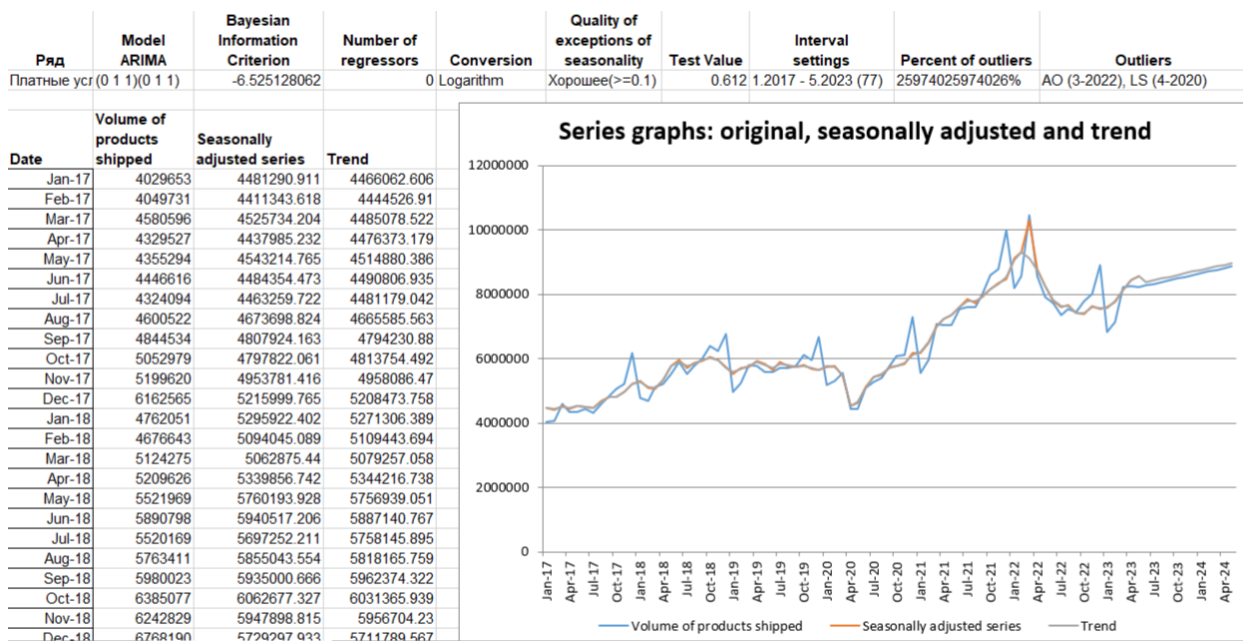


Figure 1 – Expert report for the indicator Volume of shipped products

The Horizon system allows you to simultaneously forecast several indicators that have a common trend and carry out factor analysis.

In Fig. Figure 2 shows the forecast of the indicator Volume of shipped products using the ARIMA method.

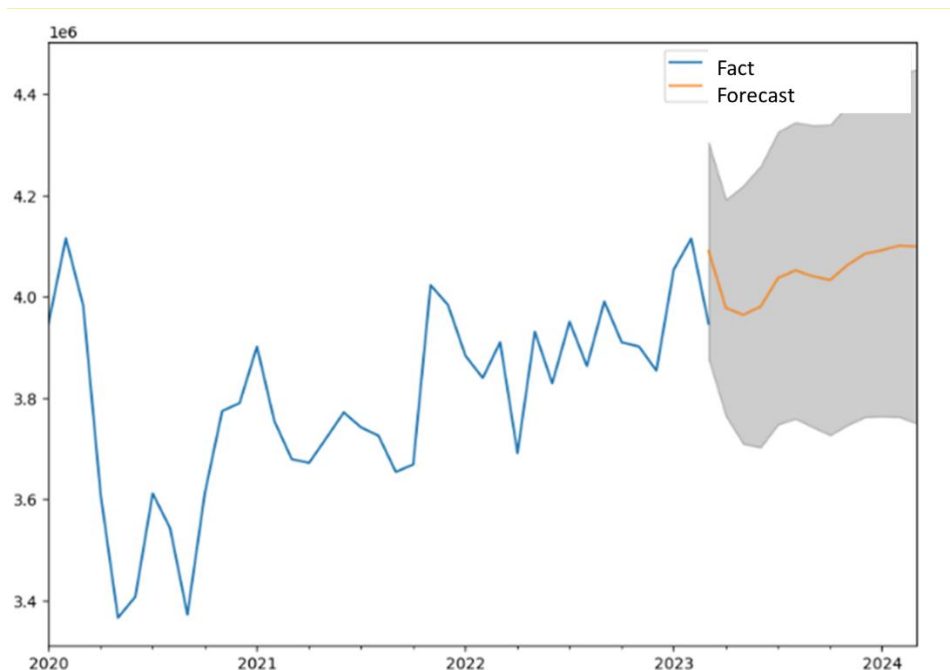


Figure 2. Forecast chart for the indicator Volume of shipped products

Discussion

The Horizon system supports regression and hybrid forecasting models based on a description of the dependencies between indicators both within one block and from different blocks, based on previously conducted research and economic sense, while calculations are based on a scenario approach using exogenous indicators such as price for oil, GDP, etc.

At the same time, as an analysis of sources has shown, many researchers use the autoregressive ARIMA model as a base model for forecasting socio-economic indicators, including industrial indicators, which in most cases gives fairly good results. In this regard, the ARIMA forecasting module was included in the Horizon system, and calculations were carried out for a number of industry indicators until 2024, which showed satisfactory quality and accuracy.

In the future, when forecasting socio-economic indicators, we plan to compare calculations using complex hybrid models with this model.

We are currently changing the structure of scenario indicators in accordance with the Central Bank document and plan to forecast a group of industry indicators using hybrid models.

Acknowledgments

This research was performed in the framework of the state task in the field of scientific activity of the Ministry of Science and Higher Education of the Russian Federation, project "Models, methods, and algorithms of artificial intelligence in the problems of economics for the analysis and style transfer of multidimensional datasets, time series forecasting, and recommendation systems design", grant no. FSSW-2023-0004.

References

Astafieva, E. and Turuntseva, M. Assessing the quality of short-term forecasts of industrial production indices of Rosstat.//Bulletin of the The Gaidar Institute for Economic Policy. 2014, vol. 12, pp. 34-37.

Astafieva, E. and Turuntseva, M. Assessing the quality of short-term forecasts of industrial production indices of the National Research University Higher School of Economics (April 2009 - August 2021). // Bulletin of the The Gaidar Institute for Economic Policy. 2021, vol. 9, pp. 32-35.

Boldyrevsky, P. B., Igoshev, A. K., Kistanova, L. A. Cluster analysis and neural network modeling for Movements of Industrial Production Index of the Russian Manufacturing Industry // Economic Analysis: Theory and Practice. 2019, vol. 18, No. 11(494), pp. 2158-2171.
Dougherty, Christopher. Introduction to econometrics. 2011. Oxford; New York : Oxford University Press.

Fokin, N.D. VAR-LASSO model for the Russian economy on a large data set. 2020. Available at SSRN: <https://ssrn.com/abstract=3695047> or <http://dx.doi.org/10.2139/ssrn.3695047>

Kalinovskaya, I. N., Kerryev, K. N. Forecasting the production of light industry goods in the Republic of Belarus. // Innovative technologies in textile and light industry: Proceedings of the International Scientific and Technical Conference, Vitebsk, 13–14 November 2019. Vitebsk: Vitebsk State Technological University. 2019, pp. 341-344.

Khairutdinov, I.S. Statistical analysis and long-term assessment of the development of the petrochemical industry in Russia. Dissertation for the degree of Candidate of Economic Sciences. St. Petersburg. 2022.

Kitova, O. V., Dyakonova, L. P. , Kitov, V. A., Savinova, V. M. Forecasting industrial indicators in Russia in the information and analytical system "Horizon". //Voprosy Istorii. 2021, vol. 10 (1), pp. 268 -287.

Kitova, O.V., Dyakonova, L.P., Kitov, V.A., Savinova, V.M. Application of neural networks for forecasting socio-economic time series.//Russian Economic Bulletin. 2020, vol. 3, No. 5, pp. 188-201.

Kitova, O.V., Kolmakov, I.B., Dyakonova, L.P., Grishina, O.A., Sekerin, V.D., Danko, T.P. Hybrid intelligent system of forecasting of the socio-economic development of the country.// International Journal of Applied Business and Economic Research. 2016, vol. 14, No.9, pp. 5755-5766.

Kitova O.V., Savinova V., Dyakonova, L., Kitov, V. Development of hybrid models and a system for forecasting the indicators of the Russian economy// Espacios. 2019, vol. 40, No 10, pp. 18-24.

Krasnyuk, L. V. Analysis of economic development and forecasting of the main indicators of the industry of the Russian Federation. // St. Petersburg State Polytechnical University Journal. Economic Sciences. 2013, vol. 4(175), pp. 38-44.

Miroljubova, A. A., Ermolaev, A. D., Prokofiev, M. B. ARIMA - forecasting the demand of the production plant. // Modern high technologies. Regional application. 2021, vol. 2(66), pp. 50-55.

Pestova, A. A. and Mamonov, M. E. Estimating the Influence of Different Shocks on Macroeconomic Indicators and Developing Conditional Forecasts on the Basis of BVAR Model for the Russian Economy. // Ekonomicheskaya Politika. 2016, vol. 11, no. 4, pp. 56-92.

Petrakova, A. M. and Zinovkina, S. A. Modeling and forecasting of key indicators of oil-producing industry of Russia // Krasnoyarsk Science. 2019, vol. 8, No. 4-2, pp. 62-66.

Shinkevich, A. I., Lubnina, A. A., Bronskaya, V. V. Analysis of trends in scientific research development of the industry on the basis of mathematical modeling methods // Izvestia of Samara scientific center of the Russian Academy of Sciences. 2022, vol. 24, No. 4(108), pp. 68-74.

Sokolova, O. A. and Sidorov, M. A. Forecasting the dynamics of industrial production indicators in the region (case study of the Vologda oblast). // Territorial Development Issues. 2022, vol. 10, No. 2.

Treshchevsky, Yu. I., Makarov, M. V., Klimov, N. A., Nikulnikov, P. D. Forecasting the development of the manufacturing industry in the region using standard formation programs neural networks.// Proceedings of the Southwest State University. Series: Economics. Sociology and Management. 2020, vol. 10, No. 1, pp. 118-129.

Turuntseva, M. Yu., Astafieva, E. V., Baeva, M. A. et al. Model calculations of short-term forecasts of socio-economic indicators of the Russian Federation.// Bulletin of the The Gaidar Institute for Economic Policy. 2014, vol. 5, pp. 2-31.

Turuntseva, M. Yu. and Kiblitckaya, T.R. Qualitative properties of various approaches to forecasting socio-economic indicators of the Russian Federation. 2010. Moscow: The Gaidar Institute for Economic Policy.