Science of Law, 2025, No. 3, pp. 325-336
DOI: 10.55284/q9b67188

System Analysis of a Machine-Building Enterprise Development Indicators

Elena Katysheva[1]*, Natalia Bogdanova[2], Yulia Lyubek[3], Olga Batova[4], Alexander Nikiforov[5]

[1,2,3,4,5]Empress Catherine II Saint Petersburg Mining University, Russia; helene_la_belle@mail.ru

Keywords:
Bus production,
Correlation and regression analysis,
Expert assessment,
Forecasting,
Mechanical engineering,
Transport.

Abstract. The article is devoted to identifying the most important factors influencing the volume of bus production by Russian enterprises. The place of mechanical engineering in the country's economy is characterized, as well as the importance of bus manufacturing plants for the transport industry development. The object of the study is PJSC NEFAZ - one of the largest bus manufacturers in Russia. This enterprise occupies a leading position in the production of urban public transport in the country. The article presents the results of a system analysis of the indicators of the automobile manufacturing company PJSC NEFAZ. To solve the problem, a forecast of the volume of production for two reporting periods was built using trend models. Also, using the method of expert assessment and multivariate correlation and regression analysis, factors influencing the main indicators of the company were identified. The results obtained allow making informed management decisions, reducing time and money costs, and competently assessing risks based on forecasts.

1. INTRODUCTION

The mechanical engineering industry has historically occupied a very important place in the Russian economy. Mechanical engineering enterprises provide the means of production to all industries without exception, reflecting the level of development of scientific and technological progress [1]. It should be noted that recent years have been characterized by intensive growth in mechanical engineering production (in 2023 - more than 15%), which is explained by the growth of investments, an increase in demand for investment equipment and military equipment, import substitution processes in the context of economic sanctions, and the implementation of state subsidy programs. The structure of the mechanical engineering industry includes more than 70 sub-sectors, including heavy, energy, oil and gas engineering. Acceleration of growth rates of these sub-sectors contributes to the development of high-tech industries in Russia, including the petrochemical sector [2].

In addition to the production of means of production, one of the first places in the mechanical engineering industry belongs to enterprises specializing in the production of vehicles [3]. In particular, one of the most important elements of urban infrastructure is bus transport. It has a number of advantages: it is economical and accessible to the population, helps to reduce the number of traffic jams on the roads and the level of environmental pollution [4 – 7] due to the intensive use of personal vehicles within the city. In addition, bus transport is indispensable in the mineral resource complex, where daily delivery of personnel to remote field sites is required. The intensification of import substitution processes has led to the fact that today the main suppliers of buses to the regions are domestic automobile factories. Their development directly affects the state of public transport in Russia. Also, automobile manufacturing companies provide jobs, are often city-forming and directly affect the economic development of the region and the country as a whole.

The high importance of the efficient operation of domestic bus manufacturing plants for the development of the Russian transport industry [8] requires an increase in their production indicators, improved production organization and improved quality of manufactured products. The specifics of the activities of a machine-building company are reflected in the system of indicators characterizing the efficiency of the enterprise [9]. Analysis of the company's work allows determining the directions of its development, methods of rational use of labor and material resources [10]. Systems analysis is a key approach in strategic planning, contributes to the development of economically sound management decisions [11 – 13].

Despite the above-mentioned intensification of growth in the mechanical engineering industry and, in particular, the automobile industry, the Russian transport system is currently characterized by a number of problems [14, 15], among which the following can be highlighted:

- The level of transport infrastructure does not correspond to the needs of the population;
- The technological lag of the Russian transport complex in comparison with foreign competitors;
- Labor shortage both in transport companies and in automobile manufacturing enterprises;
- A reduction in the possibilities of acquiring new equipment on lease due to the tightening of the monetary policy of the Central Bank;
- A high degree of wear and tear of the transport complex.

To solve the problems, it is necessary to develop domestic plants, optimize their work, and introduce new technologies. The competitive advantages of a machine-building enterprise are manifested in an increase in the level of product quality; provision of additional after-sales services; compliance of products with consumer expectations; rational logistics policy; reasonable pricing; reliable business reputation and the degree of social significance of the company.

To improve the work of the enterprise, it is important to apply balanced management decisions, correctly allocate resources, highlight the necessary elements of the system and the connections between them. To determine the paths of development of the company, managers often use methods of system analysis. The main principle of system analysis is the representation of complex objects in the form of systems. System analysis includes the analysis of the structure and functions of the system, identifying the connections between the elements, and developing solutions for the functioning of the system [16].

*Corresponding author.

2. MATERIALS AND METHODS

The PJSC NEFAZ is a system for the buses production, the elements of which are the plant employees, technology and means of production, consumed resources. The system has integrity, in the absence of one of the elements it is difficult to implement the production cycle and obtain a finished product, which is the purpose of the system. For effective management of the company, it is necessary to understand how the elements are related to each other, and what factors affect the main indicators of the system. For this purpose, methods of system analysis have been developed. The article applies statistical analysis methods - multiple correlation and regression analysis, analytical forecasting method, expert assessment method.

The implementation plan for the above methods is described in the article [17]. The research scheme is shown in Figure 1. This article considers the stages of statistical analysis of the company's indicators.

Forecasting is the process of predicting the future state of a system based on the analysis of current indicators [18]. Forecasting in production is necessary for making management decisions, optimizing resource implementation and risk management. Forecasting models are used in various sectors of the economy and industry, healthcare systems, and education [19, 20]. The use of trend models is the main method of technical analysis of investors [21].

Expert forecasting methods are applied to systems for which it is difficult to develop an adequate mathematical model [22, 23]. In this article, this method is used as an auxiliary one before multivariate correlation and regression analysis. The objective of the method is to identify and quantitatively evaluate the relationship between the main indicator and the factors influencing it. The study [24] noted the effectiveness of using multifactor mathematical modeling of processes whose operation is characterized by many complex features.

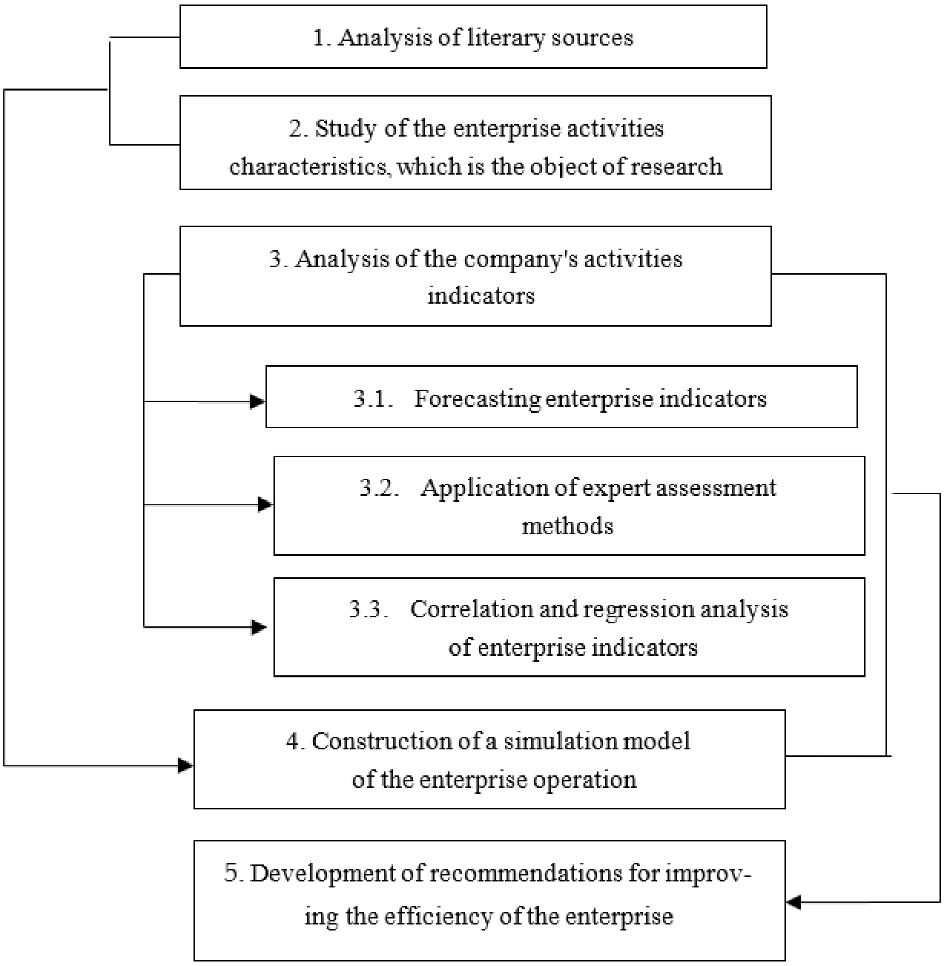

Figure 1. The main stages of the researh of the work of NEFAZ PTC.

This article is devoted to identifying the development trend of a bus manufacturing enterprise and determining the dependencies between the main performance indicators of the company and the factors influencing them. The need for the study is due to the importance of management decisions made based on these indicators.

The research object in this paper is the Public Joint Stock Company Neftekamsk Automobile Plant (PJSC NEFAZ), a Russian automobile manufacturing company that is part of the PJSC KAMAZ group of enterprises and produces passenger and crew buses and trolleybuses, electric buses, tanker trucks, flatbed trailers, and special superstructures on vehicle chassis. PJSC NEFAZ is a leader in the supply of buses for urban public transport in Russia. The company is involved in solving emerging transport problems. The company aims to produce the best comfortable buses in Russia by complying with international trends in the industry development.

PJSC NEFAZ shows intensively growing dynamics of development. Thus, the enterprise's revenue in 2023 exceeded the same indicator in 2022 by 44% [25]. In particular, revenue from the sale of passenger buses increased by 11%. The main share of revenue (73%) comes from passenger vehicles. The enterprise produces more than 40% of shift buses in Russia. The growth in production and sales of passenger transport is due to the implementation of import substitution programs, as well as the functioning of the state program for the renewal of public transport in the regions of Russia. PJSC NEFAZ is one of the main manufacturers of buses in Russia. The enterprise supplies buses, trolleybuses and electric buses to the largest cities of Russia - Moscow, St. Petersburg, Yekaterinburg, Novosibirsk. The enterprise's buses are also used by industrial companies, including PJSC Gazprom. The plant annually expands its range of manufactured equipment, introduces the latest technologies, and updates machines and equipment. In 2024, new models were included in serial production [25, 26]:

- All-wheel drive off-road bus KAMAZ-6250, recognized as the best commercial vehicle of the year in Russia, designed for use on intercity lines, as well as at the facilities of PJSC Gazprom;
- Extra large class bus KAMAZ-6299;
- Medium semi-low-floor bus KAMAZ-4290.

The company adheres to environmental standards and produces vehicles with engines running on natural gas. It was revealed that the company is developing and the organization's revenue is growing: an increase of more than 200% from 2018 to 2023.

The trend forecasting method (trend models using) consists of determining the type of function describing the development of system indicators in the past and present to predict the value of the indicator in the future. The main indicator of any production is the volume of production. The key goal of the enterprise is to improve the quality and quantity of manufactured goods. For successful operation, including increasing revenue and profit, it is necessary to know the planned production, which will be the focus of the enterprise's purchases and investment policy. Forecasting in production is necessary for making management decisions, optimizing the implementation of resources and risk management.

The forecasting process usually includes the following stages:

- Data collection: compiling a time series of indicator values [27];
- Identifying the presence of a trend component;
- Determining the forecast lead period - the time interval for which the forecast is based;
- Choosing a model and developing a forecast [28, 29];
- Checking the accuracy and adequacy of the forecast [30, 31].

The volume of bus production at the NEFAZ PJSC enterprise was selected as the studied indicator. The purpose of forecasting is to make point and interval forecasts of the volume of bus production for 2024-2025.

The forecast lead period is the time interval for which the forecast must be built. For the studied object, a lead period of two years was taken. In order for the trend line on the graph not to diverge significantly, it is necessary to correctly select the retrospection section (the number of parameter values on which the forecast is based). For this, the data sufficiency condition for forecasting was checked.

The mean square error of the equation k is found using the formula:

$$k = \sqrt{\frac{1}{n} + \frac{3\cdot(n+2\cdot z-1)^2}{n\cdot(n^2-1)}}, \quad ((1)$$

where z is the number of time units for which the trend is built; n is the number of measurements sufficient to anticipate the forecast.

It is necessary to make a forecast for z = 2 years (for 2024-2025). The mean square error k should not exceed 1. Substituting these data into formula (1), we obtain:

$$\sqrt{\frac{1}{n} + \frac{3\cdot(n+2\cdot 2-1)^2}{n\cdot(n^2-1)}} \leq 1. \quad ((2)$$

3. RESULTS AND DISCUSSION

As a result of solving inequality (2), the required number of measurements n = 9 (k = 0.93) was determined. Therefore, n = 10 selected years will be enough to forecast the volume of bus production for 2024-2025.

The time series of retrospective data is composed of indicators for 2014-2023 [25] and is presented in Table 1.

Table 1. Initial time series for forecasting.

Period No	1	2	3	4	5	6	7	8	9	10
Year (x)	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Production (y)	586	612	823	720	785	970	1602	1787	1425	1449

Source: compiled by the authors based on data from [25].

Forecasting was performed using trend models of the MS Excel environment. Before starting forecasting, the dynamic series was checked for the presence of a trend component. It was performed using the difference in mean levels method. For this, the original series is divided into two partial samples $n_1 = n_2 = 5$. For the difference in mean levels method, the procedure "Two-sample t-test with equal variances" was called at the significance level $\alpha =$ 0.05. The results of the procedure are presented in Table 2.

Table 2. Results of the procedure "Two-sample t-test with equal variances"

Indicators	Sample 1	Sample 2
Mean	705,2	1446,6
Dispersion	10839,7	91910,3
Observations	5	5
Pooled Dispersion	51375	
t-statistic	-5,17	
P(T<=t) one-tailed	0,0004	
t critical one-tailed	1,86	
P(T<=t) two-tailed	0,0009	
t critical two-tailed	2,31	

The obtained result is compared with the critical value: the critical value of the Student's criterion is 2.31, and the calculated value is 3.32. Since tSt.> tcr., the time series contains a trend component. The difference between the trend models is in the form of the presented dependence equation. The article applies linear, parabolic, power and logarithmic trends. To determine the trend functions, the Trend Line function in MS Excel is used. The graphs of the dependence of the volume of bus production on time are found, they are superimposed on a linear (Figure 2), power (Figure 3), logarithmic (Figure 4), exponential (Figure 5) trend with a forecast for 2 periods. Also, the determination coefficient R2 is displayed on each graph.

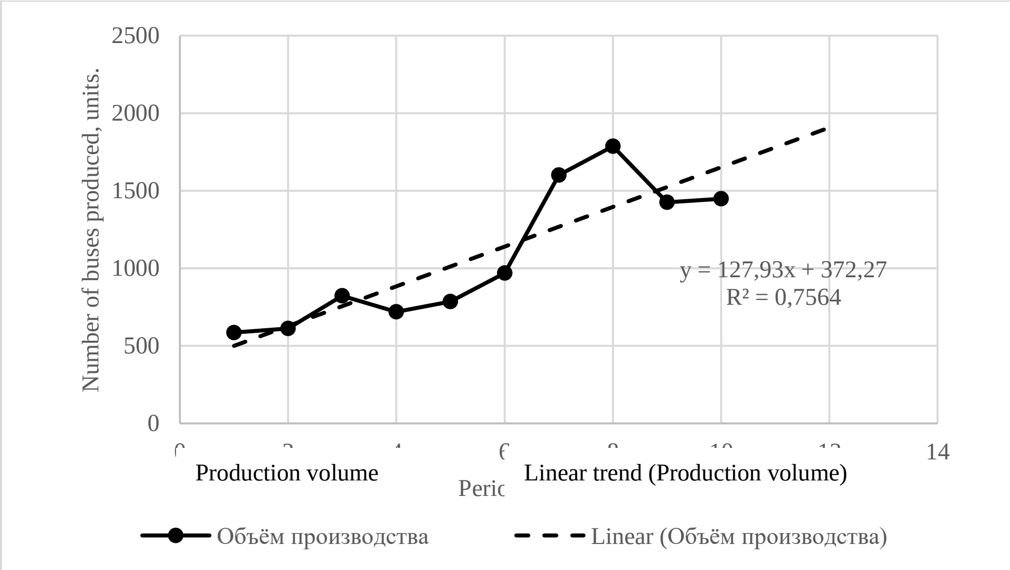

Figure 2. Dependence of the number of produced buses on the period number (linear trend).

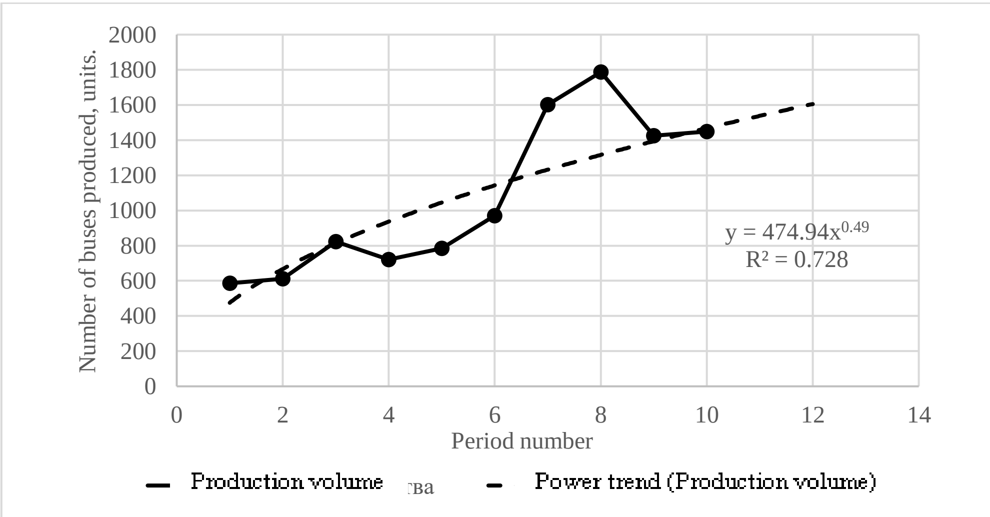

Figure 3. Dependence of the number of produced buses on the period number (power trend).

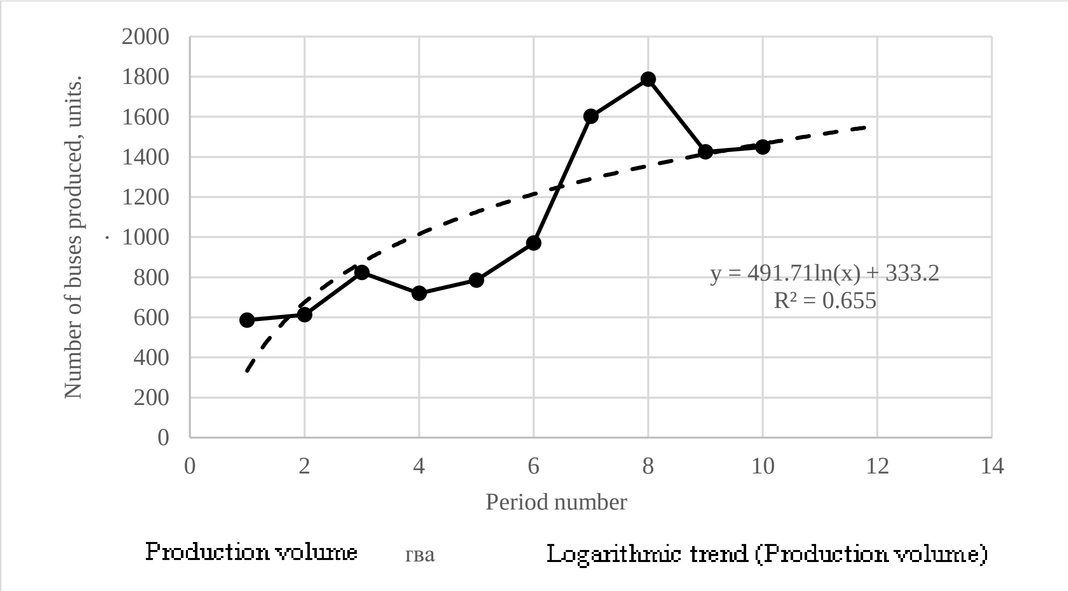

Figure 4. Dependence of the number of produced buses on the period number (logarithmic trend).

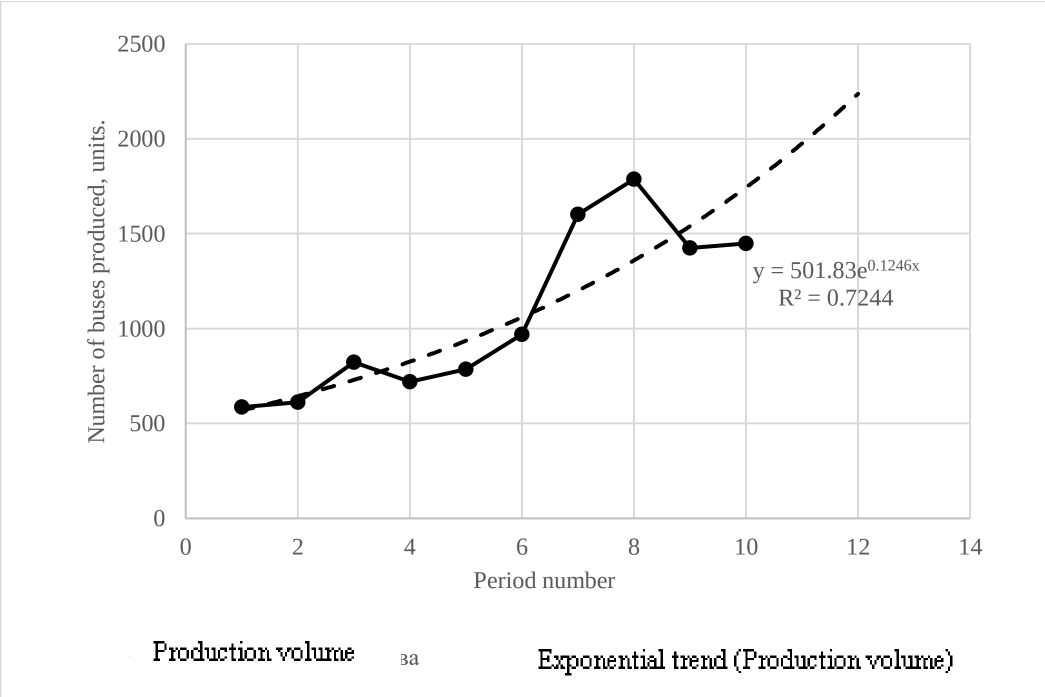

Figure 5. Dependence of the number of produced buses on the period number (exponential trend)

An upward trend in the company's development has been noted: the number of buses produced is gradually increasing. In 2022, the production volume decreased due to the foreign economic and political situation in the country: sanctions restrictions, increased costs for the acquisition of new fixed assets, raw material and logistics costs affected production. However, in 2023, PJSC NEFAZ began to return to production according to planned indicators: new suppliers were found, alternative technologies were obtained. To identify the best trend model, the determination coefficient is used, showing the share of explained variation of the dependent indicator [32]. The higher the determination coefficient, the more accurate the trend model. The equations of trend lines and the obtained determination coefficients are given in Table 3.

Table 3. Results of the procedure "Two-sample t-test with equal variances" Modeling results.

Trend line type	**Trend equation**	**Coefficient of determination R^2**
Linear	$y = 127{,}93t + 372{,}27$	0,7564
Power	$y = 474{,}94t^{0{,}49}$	0,728
Logarithmic	$y = 491{,}71\ln(t) + 333{,}2$	0,655
Exponential	$y = 501{,}83e^{0{,}1246t}$	0,7244

Comparison of the obtained coefficients of determination values allowed us to establish that the linear curve is the best model, since this curve has the highest coefficient of determination $R^2 = 0{,}7564$.

After obtaining the regression equation, the significance of its coefficients was assessed using Fisher's F-criterion: the hypothesis H0 is put forward about the statistical insignificance of the regression equation and the indicator of the tightness of the relationship.

Fisher's F-criterion is determined by the formula:

$$F = \frac{R^2}{1-R^2} \cdot (n - m - 1), \quad (3)$$

where n is the number of levels of the dynamic series, m is the number of coefficients for the parameter of the regression equation. The value of the Fisher F-criterion for the obtained series is found using formula (3):

$$F_{fact} = \frac{0{,}4031}{1 - 0{,}4031} \cdot (10 - 1 - 1) = 5{,}4$$

The tabular value of the F-criterion at a significance level of α = 0,05: Ftable = 4,96. Since $F_{fact} > F_{table}$, the hypothesis H_0 is rejected, the equation is considered statistically significant.

According to the method of presenting the predicted indicator, a distinction is made between:

- Point forecast - this is a prediction of the future that produces an unambiguous value of the predicted indicator;
- Interval forecast - this is a prediction of the future that assumes a certain range of values of the predicted indicator.

The point forecast is calculated using the linear trend model equation:

$$y_p = 127{,}93t + 372{,}27 \quad (4)$$

The point forecasts for 2024 and 2025 (periods 11 and 12) were obtained using formula (4):

$$y_{2024} = 127{,}93 \cdot 11 + 372{,}27 = 1779 \text{ units},$$
$$y_{2025} = 127{,}93 \cdot 12 + 372{,}27 = 1907 \text{ units}.$$

To find the confidence interval of the interval forecast, the standard forecast error s_{y_p} was used, which is found using the formula:

$$s_{y_p} = \sqrt{\frac{\sum_{i=1}^{n}(y - \hat{y})^2}{n - m - 1}} \cdot \sqrt{1 + \frac{1}{n} + \frac{\left(x_p - \bar{x}\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}} \quad (5)$$

Intermediate calculations are presented in Table 4.

Table 4. Intermediate calculations.

Year	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Average value
x	1	2	3	4	5	6	7	8	9	10	5,5
y	586	612	823	720	785	970	1602	1787	1425	1449	
$x - \bar{x}$	-4,5	-3,5	-2,5	-1,5	-0,5	0,5	1,5	2,5	3,5	4,5	Sum
$(x - \bar{x})^2$	20,25	12,25	6,25	2,25	0,25	0,25	2,25	6,25	12,25	20,25	82,5
$\hat{y}$	500	628	756	884	1012	1140	1268	1396	1524	1652	Sum
$(y - \hat{y})^2$	7362	260	4481	26893	51493	28849	111703	153108	9730	41035	434913

The calculated variables are substituted into formula (2), and the following is obtained:

$$s_{y_p} = \sqrt{\frac{434913}{10 - 1 - 1}} \cdot \sqrt{1 + \frac{1}{10} + \frac{(11 - 5{,}5)^2}{82{,}5}} = 282{,}4$$

We find the forecast confidence intervals based on the relationship:

$$y_p - t_{tabl.} \cdot s_{y_p} \leq y_p \leq y_p + t_{tabl.} \cdot s_{y_p} \quad (6)$$

where t_{tabl} is the tabular value of the Student's t-test at the significance level α with the number of degrees of freedom n-2 (ttable = 2,306).

Then, according to relation (6), the interval forecasts for 2024-2025 are:

$$1128 \leq y_{2024} \leq 2430$$
$$1256 \leq y_{2025} \leq 2558$$

Thus, the volume of bus production at the PJSC NEFAZ plant will be 1779 units in 2024 and 1907 units in 2025 according to the point forecast. However, values in the intervals [1128; 2430] and [1256; 2558] are also possible according to the interval forecast. The lower limit of the interval forecast corresponds to a deterioration in the economic condition of the enterprise or uncontrolled intervention of external negative factors. The upper limit of the interval forecast is fulfilled under favorable market conditions and successful implementation of investment programs.

Determination of factors affecting production volume. To optimize the constructed forecast, it is necessary to identify factors that directly or indirectly affect the studied indicator [33, 34]. The method of a priori ranking (expert assessments) will allow to highlight the most significant factors among others based on the opinion of specialists (experts) in the studied area. Employees of various departments of the enterprise acted as experts. In this article, the method is used as an auxiliary one to exclude insignificant factors before multiple correlation and regression analysis. The object of the study is the number of buses produced at the PJSC NEFAZ plant. The volume of production is affected by external and internal factors. External factors include political, natural and general economic factors. Internal factors include production factors (technical and technological equipment of production, organizational, economic and social factors).

Key factors influencing bus production volume:

X_1 – government subsidies to the industry, determined by federal budget expenditures on transport, million rubles;

X_2 – government financing of public transport in the regions, determined by regional budget expenditures on transport, million rubles;

X_3 – share of imported units and parts for production, %;

X_4 – share of imported buses from China and Belarus, %;

X_5 – average salary of plant employees, rubles;

X_6 – share of worn-out vehicles in the regions (service life over 15 years), %;

X_7 – number of units of production in the nomenclature, units;

X_8 – share of families without a personal car, %.

The main consumers of large-class buses are municipal public transport enterprises. The number of buses produced depends mainly on orders from regional authorities. Therefore, the factors influencing state financing of the industry (X_1, X_2) are taken into account. In addition, production factors (X_3, X_5, X_7) play an important role. Factors indirectly influencing demand for public transport services (X_6, X_8) are also taken into account. Expert assessment. Five experts were asked to rank 8 factors influencing the volume of buses produced at the NEFAZ plant. The lower the rank assigned by the expert, the greater the influence the factor has on the resulting indicator. The resulting rank matrix is presented in Table 5.

Table 5. Rank matrix.

Experts	Factor ranks							
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	1	3	2	4	5	7	6	8
2	1,5	1,5	5	4	3	7	8	6
3	1	2	3	7,5	5	4	6	7,5
4	2	1	4	7	6	3	5	8
5	1,5	3	1,5	5	6	4	7	8
Ranks sum	7	10,5	15,5	27,5	25	25	32	37,5
Deviation of the ranks sum from the average ranks sum	-15,5	-12	-7	5	2,5	2,5	9,5	15
Standard deviation	240,25	144	49	25	6,25	6,25	90,25	225

Table 5 also calculates the ranks sum for each factor and determines the deviation Δ of the ranks sum from the average ranks sum for each factor. The significance of a factor is determined by the ranks sum. According to the analysis, the volume of bus production is significantly affected by factors X_1 (state subsidization of the industry), X_2 (state financing of public transport in the regions) and X_3 (the share of imported units and parts for production).

To check the quality of the results obtained, the consistency of the opinions of the surveyed experts was assessed using the concordance coefficient. The coefficient expresses the ratio of the dispersion estimate of the obtained ranks to the maximum value of this estimate. This reveals how similarly the experts assess the situation. The closer the value of the concordance coefficient to one, the more objective the results.

The concordance coefficient *W* is calculated using the formula

$$W = \frac{S}{\frac{1}{12}m^2(k^3-k) - m\sum_j T_j} \qquad (7)$$

where $T_j = \frac{1}{12}\sum_u(t_u^{\,3} - t_u)$, where *u* is the number of groups formed by factors of the same rank in the j-th ranking; t_u is the number of identical ranks in the u-th group of the j-th ranking.

There are no duplicate ranks in the first and fourth rankings, so $T_1 = T_4 = 0$. The second, third, and fifth experts each assigned one related rank that occurs twice. Therefore, $T_2 = T_3 = T_5 = \frac{1}{12}(2^3 - 2) = 0{,}5$.

Having calculated T_j, we find the coefficient of concordance *W*=0,75. The significance of the coefficient of concordance *W* is calculated using the formula

$$\chi^2_p = m(k-1)W = 5 \cdot (8-1) \cdot 0{,}75 = 26{,}25.$$

At a 5% significance level and the number of degrees of freedom $f = k - 1 = 8 - 1 = 7$, the tabular value of the χ^2 - criterion is 14,1. Since $\chi^2_p = 26{,}25 > \chi^2_T = 14{,}1$, the opinions of the surveyed specialists are consistent. A rank diagram was constructed to identify significant factors (Figure 6).

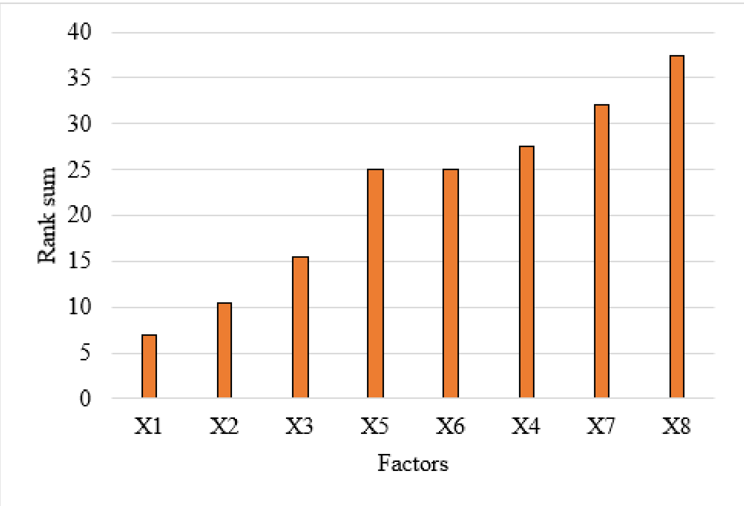

Figure 6. Rank sum diagram.

Factor X_8 (the share of families without a personal car) with the highest rank is excluded from further research, since it has the least impact on the volume of bus production.

Correlation analysis. After identifying significant factors, it is necessary to determine how the identified factors are related to each other [35], calculate the quantitative measure of their influence on each other and the resulting indicator. The correlation between phenomena occurs stochastically and shows how a change in one feature affects a change in another.

A correlation can occur [36]:

- when one feature depends on another;
- when there is a phenomenon that affects both features;
- when both features are the cause or consequence of each other.

Correlation analysis performing begins with determining a sample of parameters and compiling a correlation table [37]. The static data sample is made up of indicators for 2010 - 2022. Annual indicators are taken from Federal Statistics Service of the Russian Federation data, studies of All-Russian Center for the Study of Public Opinion, annual reports of PJSC NEFAZ [25, 26, 38]. The initial data are presented in Table 6.

Since many of the data in Table 6 are presented in different units of measurement, it is necessary to bring them into a comparable form. To do this, each data cell was divided by the sum of the entire corresponding column. In this way, the data were expressed as relative values. The new data are presented in Table 7.

Table 6. Data for correlation analysis.

Year	№	Number of buses produced, units. Y	Federal budget expenditure on transport, million rubles. X_1	Regional budget expenditures on transport, million rubles. X_2	Share of imported units and parts, % X_3	Share of imported buses on the market, % X_4	Average salary of employees, rub. X_5	Share of worn-out buses in the regions, % X_6	Number of units of products in the nomenclature, units. X_7
2010	1	855	301,3	161,5	12,4	35,9	12939	46,5	25
2011	2	545	298,2	153,9	13,3	47	17375	46,9	25
2012	3	740	229,4	249,6	14,1	56	18137	47,2	28
2013	4	664	182,2	243,9	14,7	36	20502	48,1	28
2014	5	586	221,3	288,6	10,5	33,1	21821	48,9	36
2015	6	612	305	400,5	11,5	22,8	25557	49,8	20
2016	7	823	296,9	436,9	7,2	32,5	29098	49,1	25
2017	8	720	283,4	560,3	11,7	36	31733	48,6	25
2018	9	785	254	560	17,5	34	33160	47,5	25
2019	10	970	275,7	686,6	19,6	38	38439	45,9	22
2020	11	1602	323,3	758,6	21,6	25	39677	44	27
2021	12	1787	518,1	891,9	20,2	23	43353	41,3	24
2022	13	1425	647	1056	18,3	27	50376	41	16

Source: compiled by the authors based on data from [25; 26; 38].

Table 7. Data presented in general form.

№	Y	X_1	X_2	X_5	X_6	X_7	X_8	X_9
1	0,0706	0,0729	0,0250	0,0644	0,0804	0,0339	0,0769	0,0767
2	0,0450	0,0721	0,0239	0,0691	0,1053	0,0455	0,0775	0,0767
3	0,0611	0,0555	0,0387	0,0732	0,1255	0,0475	0,0780	0,0859
4	0,0548	0,0441	0,0378	0,0763	0,0807	0,0536	0,0795	0,0859
5	0,0484	0,0535	0,0448	0,0545	0,0742	0,0571	0,0809	0,1104
6	0,0505	0,0737	0,0621	0,0597	0,0511	0,0669	0,0823	0,0613
7	0,0679	0,0718	0,0678	0,0374	0,0728	0,0761	0,0812	0,0767
8	0,0594	0,0685	0,0869	0,0607	0,0807	0,0830	0,0804	0,0767
9	0,0648	0,0614	0,0868	0,0909	0,0762	0,0868	0,0785	0,0767
10	0,0801	0,0667	0,1065	0,1018	0,0851	0,1006	0,0759	0,0675
11	0,1322	0,0782	0,1176	0,1121	0,0560	0,1038	0,0728	0,0828
12	0,1475	0,1253	0,1383	0,1049	0,0515	0,1134	0,0683	0,0736
13	0,1176	0,1564	0,1638	0,0950	0,0605	0,1318	0,0678	0,0491

The next stage of correlation analysis is finding the correlation coefficients for each pair of features. The correlation coefficient is a measure of the closeness of the relationship between two correlated values. The correlation coefficient varies within the range [-1; +1]. When $r_{xy} = \pm 1$, the dependence is linear functional. If there is no relationship between the features, the value of the coefficient will be close to zero. The sign of the correlation coefficient determines the type of relationship: the sign is positive for a direct relationship between the features, and negative for an inverse relationship. To perform correlation analysis, the Correlation add-in was selected in the Data Analysis package of the MS Excel program. After entering the range of initial data, the analysis results were obtained (Figure 7).

	Y	X1	X2	X3	X4	X5	X6	X7
Y	1							
X1	0,73438	1						
X2	0,82298	0,78144	1					
X3	0,74216	0,42182	0,66784	1				
X4	-0,54553	-0,49313	-0,59875	-0,24601	1			
X5	0,77451	0,73045	0,99069	0,64707	-0,59204	1		
X6	-0,89296	-0,83622	-0,7573	-0,77388	0,31594	-0,70084	1	
X7	-0,33715	-0,66542	-0,54777	-0,29497	0,30898	-0,52961	0,43324	1

Figure 7. Results of correlation analysis.

In Figure 7, column Y shows the degree of correlation between the factors and the resulting parameter, and columns X_1-X_7 show the degree of correlation between the factors. Pairs are considered correlated if $|r| > 0,75$.

The matrix allows us to understand that factors X_1 and X_2, X_1 and X_6, X_2 and X_5, X_2 and X_6, X_3 and X_6 are multicollinear, since the correlation coefficients exceed 0,75 in absolute value. Therefore, they duplicate each other.

Factors X_1, X_2 and X_3 are excluded from the model (as less significant relative to X_6) to avoid strong mutual correlation. For further analysis, 4 factors remain in the forecast model: X_4, X_5, X_6 and X_7.

Factor X_6 has the greatest influence on the resulting indicator (-0,89), and factor X_7 has the least (-0,35).

Correlation analysis showed a strong connection between state subsidies to the automotive industry and regional funding of public transport, as well as the dependence of the share of worn-out transport on these two factors.

Regression analysis. After determining the presence of a connection between the features, it is necessary to move on to regression analysis. Regression analysis is one of the main methods of system analysis. It is designed to find the form of connection between random variables [39, 40]. The essence of regression analysis is to build mathematical models of objects or phenomena based on data obtained as a result of observations or experiments [40]. Only those factors that are closely related to the resulting feature are included in the model [41]. The purpose of regression analysis is to numerically express the contribution of independent features to the variation of the dependent for further forecasting of the dependent indicator. Regression analysis begins with determining the dependent indicator Y from the independent factors $x_1, x_2, ..., x_m$, which determine the change in this indicator. For regression analysis Y is the volume of produced buses, X_4 is the share of imported buses from China and Belarus, X_5 is the average salary of employees, X_6 is the share of worn-out buses in the regions, X_7 is the number of units of production in the nomenclature. To conduct regression analysis, the Regression add-in in the Data Analysis package of the MS Excel program was selected. The data used for the analysis are presented in Table 8.

After entering the range of initial data, the analysis results were obtained (Figure 8).

Table 8. Data for regression analysis.

№	Y	X_4	X_5	X_6	X_7
1	0,0706	0,0804	0,0339	0,0769	0,0767
2	0,0450	0,1053	0,0455	0,0775	0,0767
3	0,0611	0,1255	0,0475	0,0780	0,0859
4	0,0548	0,0807	0,0536	0,0795	0,0859
5	0,0484	0,0742	0,0571	0,0809	0,1104
6	0,0505	0,0511	0,0669	0,0823	0,0613
7	0,0679	0,0728	0,0761	0,0812	0,0767
8	0,0594	0,0807	0,0830	0,0804	0,0767
9	0,0648	0,0762	0,0868	0,0785	0,0767
10	0,0801	0,0851	0,1006	0,0759	0,0675
11	0,1322	0,0560	0,1038	0,0728	0,0828
12	0,1475	0,0515	0,1134	0,0683	0,0736
13	0,1176	0,0605	0,1318	0,0678	0,0491

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,947853371
R Square	0,898426014
Adjusted R Square	0,847639021
Standard Error	0,013119422
Observation	13

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	0,012179212	0,00304	17,6901	0,000488982
Residual	8	0,001376954	0,00017		
Total	12	0,013556166			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0,479700213	0,10800635	4,44141	0,00216	0,230637122	0,72876	0,23064	0,72876
X4	-0,396186069	0,226691046	-1,74769	0,11865	-0,918936558	0,12656	-0,91894	0,12656
X5	0,209964629	0,224316026	0,93602	0,37665	-0,307309053	0,72724	-0,30731	0,72724
X6	-5,439558274	1,159121253	-4,69283	0,00156	-8,112496676	-2,76662	-8,1125	-2,76662
X7	0,389676944	0,315029575	1,23695	0,25118	-0,336782559	1,11614	-0,33678	1,11614

Figure 8. Results of regression analysis.

The Coefficients column in Figure 8 shows the obtained coefficients of the multiple regression equation. According to the obtained results, the multiple regression equation will have the form

$$Y = 0{,}48 - 0{,}396x_4 + 0{,}21x_5 - 5{,}43x_6 + 0{,}39x_7$$

The regression coefficients show the change in the resultant feature with a change of 1 unit of the dependent factor.

The obtained results show that an increase in the share of imported buses on the market contributes to a decrease in the volume of buses produced by PJSC NEFAZ by 0,396 units. The volume of production is more significantly affected by the share of worn-out buses in the regions: if it grows, the indicator decreases by an average of 5,43 units. A positive effect on the volume of buses produced is exerted by an increase in employee salaries or an addition of units to the nomenclature.

At the final stage of correlation and regression analysis, the quality of the model is assessed and the obtained data are interpreted. It is necessary to find out how reliable the model is, whether it should be used, or whether it is a coincidence.

The quality of the model is determined using the coefficient of determination R^2. In this model, $R^2 = 0{,}898$. 89.8% of the number of buses produced (Y) is explained by the share of imported buses X_4, the amount of employee salary X_5, the share of worn-out buses in the regions X_6 and the number of units of production in the nomenclature X_7.

The static reliability of the model is assessed using Fisher's F-criterion. In Figure 7, $F_{fact} = 17{,}69$. According to the table of values of Fisher's F-criterion at the significance level of α = 0.05 and the degrees of freedom $k_1 = m = 4$, $k_2 = 13 - 4 - 1 = 8$, $F_{tabl} = 3{,}84$ was obtained. Since $F_{tabl} < F_{fact}$, the equation is statistically significant.

The statistical significance of the model is determined using the Student's t-test. The tabular value for 13 measurements and a significance level of 0,05 is 2,16. For some factors, the *t*-test value does not satisfy the condition, which means that not all coefficients are statistically significant.

The quality of the equation was assessed based on the calculation of the average approximation error, calculated using the formula

$$A = \frac{1}{n}\sum\left|\frac{y - y_x}{y}\right| \cdot 100\% \quad (8)$$

where *y* is the initial data by year, y_x is the predicted value of *y* according to the obtained equation, *n* is the number of measurements.

Intermediate calculations of y_x are presented in Table 9.

Table 9. Calculating the average approximation error.

y	**y_x**	**$\|(y-y_x)/y\|$**
0,0706	0,067679	0,041097
0,0450	0,056676	0,259772
0,0611	0,050005	0,181408
0,0548	0,06097	0,112333
0,0484	0,066656	0,377934
0,0505	0,050626	0,002106
0,0679	0,056232	0,172306
0,0594	0,059063	0,00626
0,0648	0,071498	0,103346
0,0801	0,081626	0,019396
0,1322	0,116881	0,11617
0,1475	0,141328	0,041946
0,1176	0,13476	0,145606
	A =	12,2%

Based on the obtained results (table 9), the average approximation error $A = 12{,}2\%$. Therefore, the quality of this model is satisfactory.

It was revealed that the enterprise's activity significantly depends on the amount of transport with high wear. The company's task is to replace such transport, thereby improving the infrastructure of cities.

4. CONCLUSIONS

The main goal of the company's development is to increase the volume of buses produced. The article develops a forecast for the volume of bus production for 2024-2025 based on data for the past 10 years, which amounted to 1779 and 1907 buses, respectively. An upward trend in the company's development is noted.

To determine the factors influencing this indicator, methods of system analysis were selected, allowing not only to determine the relationship, but also to describe it quantitatively. Methods of system analysis allow to identify shortcomings in the production process of the enterprise, to develop recommendations for optimizing its work, reducing risks and improving the quality of products. From the obtained results it is clear that the domestic automotive industry primarily depends on the situation in the regions of Russia: the degree of wear of vehicles and financing. The development of the largest bus manufacturing plants determines the state of the industry, and, as a consequence, the quality of public transport. PJSC NEFAZ is developing, the volume of buses produced is growing every year. With an increase in the volume and area of production, the introduction of automated technologies, other enterprises in the automotive industry will also be modernized.

Thus, the analysis will help optimize the company's work. It provides a more accurate assessment of the indicators for further strategic and practical planning, reduces equipment downtime.

REFERENCES

Milyakin, S. R., Skubachevskaya, N. D., & Valeev, S. V. (2024). The automotive market in the new conditions: Current state and outlook. *Studies on Russian Economic Development, 35*(2), 255–265. https://doi.org/10.1134/S1075700724020096

Ponomarenko, T. V., Gorbatyuk, I. G., & Cherepovitsyn, A. E. (2024). Industrial clusters as an organizational model for the development of Russia's petrochemical industry. *Journal of Mining Institute, 270*, 1024–1037.

RIA Rating. (2024, June 26). *Quarterly bulletin "Mechanical engineering."* https://riarating.ru/macroeconomics/20240626/630265272.html

Afanaseva, O. V., Bezyukov, O. K., & Ignatenko, A. A. (2024). Method for assessing the relationship between the characteristics of vibroactivity and the design parameters of a marine diesel. *Academic Journal of Manufacturing Engineering, 22*(1), 68–75. https://doi.org/10.24874/PES.SI.02.017

Martini, D., Bezzini, P., & Longo, M. (2024). Environmental impact of electrification on local public transport: Preliminary study. *Energies, 17*, 5886. https://doi.org/10.3390/en17235886

Liu, Z., & Xie, T. (2024). The impact of policy thematic differences on industrial development: An empirical study based on China's electric vehicle industry policies at the central and local levels. *Energies, 17*, 5805. https://doi.org/10.3390/en17225805

Vasilev, Y., Vasileva, P., & Tsvetkova, A. (2020, August). The study of spreading information on carbon capture, utilization and storage technologies in social media. In *Proceedings of the 20th International Multidisciplinary Scientific GeoConference SGEM 2020* (Vol. 5.1, pp. 833–839).

Yury, I., & Martirosyan, A. (2024). The development of the Soderberg electrolyzer electromagnetic field's state monitoring system. *Scientific Reports, 14*, 3501. https://doi.org/10.1038/s41598-024-52002-w

Danilova, N. (2016). Key success factors of machine-building enterprises activity. *Koncept, 1*. http://e-koncept.ru/2016/16014.htm

Martirosyan, A. V., Martirosyan, K. V., Grudyaeva, E. K., & Chernyshev, A. B. (2021). Calculation of the temperature maximum value access time at the observation point. In *Proceedings of the IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)* (pp. 1014–1018). IEEE. https://doi.org/10.1109/ElConRus51938.2021.9396287

Yakovleva, V. D., Mengelishev, D. S., Alfimov, D. N., & Konyshev, M. Y. (2024). Application of system analysis methods in organization management: Analysis and optimization of production processes. *Science Prospects, 6*(177), 63–66. https://elibrary.ru/download/elibrary_68586832_27980172.pdf

Kulik, S. (2023). Elements of system analysis (sequential analysis). *Modern Science: Current Problems of Theory and Practice. Series: Natural and Technical Sciences, 7*, 102–107. https://doi.org/10.37882/2223-2966.2023.07.23

Shakhbanov, R. B., & Azizova, L. R. (2019, December 24). Improving systems analysis through enterprise performance analysis. In *Current issues in the theory and practice of scientific research development: Proceedings of the international scientific and practical conference* (Part 1, pp. 12–16). OMEGA SCIENCES.

Averina, L. M., Matushkina, N. A., & Lavrikova, Y. G. (2010). Transition of the regional transport complex to an innovative path of development. *Regional Economy, 4*, 102–110.

Ushakova, M. A., & Sviridov, D. A. (2017). Problems of operation of obsolete vehicles in urban passenger transport. *Symbol of Science, 3*, 123–125.

Popova, T., & Popov, A. (2021). A formalised approach to optimal adoption of a complex of technical means. *World of Transport and Transportation, 19*, 45–52. https://doi.org/10.30932/1992-3252-2021-19-3-5

Bogdanova, N. A. (2024, April 23–25). Justification of development directions of the enterprise for the production of motor transport using predictive models (on the example of PJSC NEFAZ). In *Proceedings of the XXIV International Scientific and Practical Conference of Young Scientists, Students and Graduate Students: Analysis and forecasting of control systems in industry, transport and logistics* (St. Petersburg).

Chetyrkin, E. M. (1977). *Statistical methods of forecasting*. Statistics.

Prata, M., Masi, G., & Berti, L. (2024). Lob-based deep learning models for stock price trend prediction: A benchmark study. *Artificial Intelligence Review, 57*, 116. https://doi.org/10.1007/s10462-024-10715-4

Kurganov, V. M., Gryaznov, M. V., & Kolobanov, S. V. (2020). Assessment of operational reliability of quarry excavator-dump truck complexes. *Journal of Mining Institute, 241*, 10–21. https://doi.org/10.31897/PMI.2020.1.10

Bykowa, E. N., Khaykin, M. M., Shabaeva, Y. I., & Beloborodova, M. D. (2023). Development of methodology for economic evaluation of land plots for the extraction and processing of solid minerals. *Journal of Mining Institute, 259*, 52–67. https://doi.org/10.31897/PMI.2023.6

Matrokhina, K. V., Trofimets, V. Y., Mazakov, E. B., Makhovikov, A. B., & Khaykin, M. M. (2023). Development of methodology for scenario analysis of investment projects of enterprises of the mineral resource complex. *Journal of Mining Institute, 259*, 112–124. https://doi.org/10.31897/PMI.2023.3

Barsukov, D. P., & Afanasyeva, O. V. (2013). Using forecasting methods to solve problems of information and statistical analysis of enterprise activities under risk conditions. *Petersburg Economic Journal, 1*, 73–77.

Galkin, V. I., Martyushev, D. A., Ponomareva, I. N., & Chernykh, I. A. (2021). Developing features of the near-bottomhole zones in productive formations at fields with high gas saturation of formation oil. *Journal of Mining Institute, 249*, 386–392. https://doi.org/10.31897/PMI.2021.3.7

PJSC NEFAZ. (2023). *Annual report of PJSC NEFAZ 2023*. https://nefaz.ru/shareholders/annual_report/

PJSC KAMAZ. (2024). *Official website*. https://kamaz.ru

Marinina, O., Malikov, A., Lyubek, Y., Pasternak, S., Reshneva, E., & Stolbovskaya, N. (2024). Selection of enhanced oil recovery method on the basis of clustering wells. *Processes, 12*, 2082. https://doi.org/10.3390/pr12102082

Schipachev, A., Fetisov, V., Nazyrov, A., Donghee, L., & Khamrakulov, A. (2022). Study of the pipeline in emergency operation and assessing the magnitude of the gas leak. *Energies, 15*, 5294. https://doi.org/10.3390/en15145294

Fetisov, V. (2024). Analysis of numerical modeling of steady-state modes of methane–hydrogen mixture transportation through a compressor station to reduce CO2 emissions. *Scientific Reports, 14*, 10605. https://doi.org/10.1038/s41598-024-61361-3

Asadulagi, M.-A. M., Pershin, I. M., & Tsapleva, V. V. (2024). Research on hydrolithospheric processes using the results of groundwater inflow testing. *Water, 16*, 487. https://doi.org/10.3390/w16030487

Velikanov, V. S. (2020). Mining excavator working equipment load forecasting according to a fuzzy-logistic model. *Journal of Mining Institute, 241*, 29–36. https://doi.org/10.31897/PMI.2020.1.29

Pervukhin, D., Kotov, D., & Trushnikov, V. (2024). Development of a conceptual model for the information and control system of an autonomous underwater vehicle for solving problems in the mineral and raw materials complex. *Energies, 17*, 5916. https://doi.org/10.3390/en17235916

Larichkin, F. D. (2014). Methodical approaches to the factorial analysis of changes of parameters of mining production. *Journal of Mining Institute, 208*, 132.

Andreeva, E. S., Marinina, O. A., & Turovskaya, L. G. (2024). Nanofluid flooding as a method of enhancing oil recovery: Mechanism, advantages. *Bulletin of the Tomsk Polytechnic University, Geo Assets Engineering, 335*(6), 189–202. https://doi.org/10.18799/24131830/2024/6/4408

Ilyushin, Y. V., & Nosova, V. A. (2024). Methodology to increase the efficiency of the mineral water extraction process. *Water, 16*, 1329. https://doi.org/10.3390/w16101329

Popova, V. B., & Fetskovich, I. V. (2021). *Statistical analysis and forecasting using application packages: A tutorial*. Michurinsky SAU.

Semenova, T., & Martínez Santoyo, J. Y. (2025). Determining priority areas for the technological development of oil companies in Mexico. *Resources, 14*, 18. https://doi.org/10.3390/resources14010018

Autostat. (2024). *Analytical agency “Autostat” official website.* https://www.autostat.ru

Ilyushin, Y. V., & Novozhilov, I. M. (2017). Analyzing distributed control system with pulse control. In *Proceedings of the 20th IEEE International Conference on Soft Computing and Measurements* (pp. 296–298). IEEE. https://doi.org/10.1109/SCM.2017.7970565

Kozlovskiy, V., Blagoveschenskiy, D., Aydarov, D., Panyukov, D., & Farisov, R. (2022). The concept of comprehensive improvement program methodology. *Standards and Quality, 7*(1021), 36–42.

Nevskaya, M., Shabalova, A., Kosovtseva, T., & Nikolaychuk, L. (2024). Applications of simulation modeling in mining project risk management: Criteria, algorithm, evaluation. *Journal of Infrastructure, Policy and Development, 8*(8), 5375. https://doi.org/10.24294/jipd.v8ix.5375