

# Modeling and forecasting of cargo turnover at the railways of the Russian Federation according to statistics data

Petr Gerasimenko<sup>1,1</sup>, Valentin Khodakovsky<sup>1</sup>, Sergey Verteshev<sup>2</sup>, Sergey Lyokhin<sup>2</sup> and Alexander Khvattcev<sup>2</sup>

<sup>1</sup>Saint Petersburg State Transport University,

<sup>2</sup>Pskov State University, Pskov, Russia

**Abstract.** The approach of modeling and forecasting of cargo turnover of transport is described. Based on statistical information on the transport activities of the Russian railway, a study of the dependence of the autocorrelation coefficients was conducted and a trend in the change in cargo turnover was established. The quality of the trend was evaluated and the point and interval forecasting of cargo turnover was carried out.

The role of railway transport in the Russian economy cannot be overestimated. A number of advantages make it the main mode of transport in the country. Losing to air transport in terms of speed of transportation, railway transport surpasses it in economic indicators in mass transportation. The advantage of road transport over rail transport with respect to transportation of bulk goods is mainly in the case of transport over short distances. River transport is inferior to rail transport because of the low speed of cargo delivery and the irregularity of traffic during the year.

Therefore, for Russia, the development of railway transport is of particular importance. It is the basis of the Russian transport system, and in some cases, not only the main, but also the only mode of transport that connects different regions and regions of our country, contributing to the development of the state economy. The work of railway transport can be judged by the development of cargo turnover, as an important indicator for the entire Russian national economy.

The study of the main indicators of cargo transportation in the Russian Federation for the period from 1920 to 2018 is carried out on the basis of statistical data. Table 1 presents statistical data on cargo turnover in billion tons of km [1], which means that the turnover only in 2018 began to exceed the level of 1990.

It should be noted that after the collapse of the Soviet Union, the turnover in Russia fell sharply, but later before the global economic crisis began to recover. With the onset of the crisis, there was a new drop in cargo turnover (Table 1).

Since cargo turnover determines the level of development of economic relations, improvement of trade and industrial activity, the degree of economic activity, due to the economic crisis, the drop in cargo turnover with the lag in one year is quite natural.

---

<sup>1</sup> Corresponding author: [pv39@mail.ru](mailto:pv39@mail.ru)

**Table 1.** Statistics

No.	1	2	3	4	5	6	7	8	9	10
Year	1928	1940	1950	1960	1990	2000	2005	2006	2007	2008
cargo turnover	59	287	415	1047	2523	1373	1898	1951	2090	2116
No.	11	12	13	14	15	16	17	18	19	20
Year	2009	2010	2011	2013	2014	2015	2016	2017	2018	2018
cargo turnover	1885	2011	2128	2196	2301	2306	2344	2493	2596.4	2506.17

Assessing further changes in cargo turnover in the Russian Federation, its point and interval estimates is interesting. As you know, forecast values can be obtained by constructing regression and trend dependencies. For correct forecasting of cargo transportation with the help of multivariate regressions, it is necessary to know well the need of the national economy in transportation and the existing cargo flows, as well as the main patterns of growth in transportation. The volume of cargo transport, measured by the number of tons of cargo transported and the number of ton-kilometers net, depends on many factors. Therefore, evaluation of the factors determining the functional dependence is quite a complex task and requires a large amount of information.

As a rule, not all factors are usually known and their values in most cases have a high error, especially in a crisis period. Typically, no factor regression model can include all the factors that affect cargo. In addition, the impossibility of their inclusion in the regression equation is explained by the limitations of the mathematical apparatus due to multicollinearity and heteroskedasticity, especially when the number of factors is close to the sample size. Therefore, more often carry out a forecast on the trend. For large systems such as Railways, the trend forecast for the previous years is quite effective.

It should be noted that the trend equation has an advantage, as it allows to take into account, although in an implicit form, the influence of all factors on changes in the levels of the projected indicator of cargo transportation. Moreover, the trend equation simulates a dynamic process, which is the logical basis for forecasting. With respect to multivariate regression, it should be noted that it determines variations in the level of the indicator in the static population, i.e. it does not explain the change in the time of the volume of cargo transportation, but its compliance with a certain combination of factors. Thus, the logical base of the forecast for multifactor regression is not adequate to the forecasting problem.

A significant disadvantage of trend forecasting is the inability to control the process of choosing the optimal mode of transportation, since it is not possible to determine different forecast options with different combinations of factor values, which usually allows you to perform a regression model.

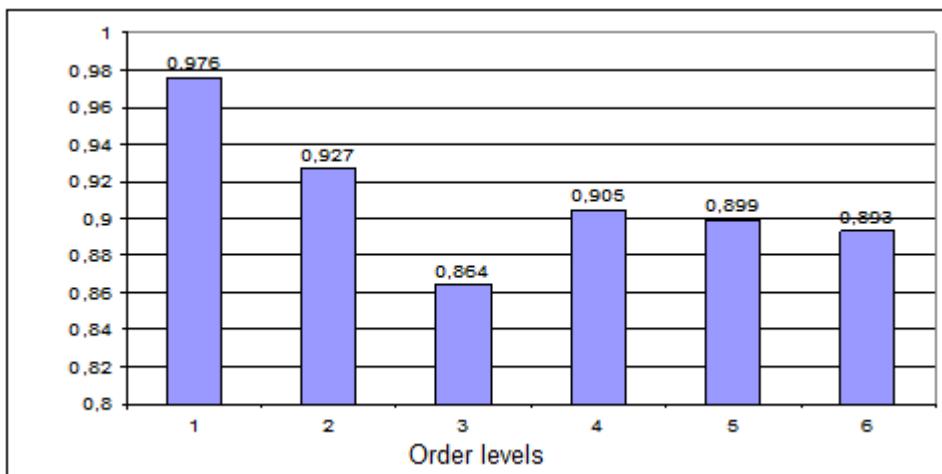
The peculiarity of trend forecasting is related to the reliability of its parameters. If the parameters of the trend is not robust enough, it will be unreliable and prognosis.

According to the statistical information (Table 1), since the beginning of the crisis, namely since 2009, there have been significant changes in a number of economic factors that have a significant impact on the turnover of goods.

The analysis of Table 1 shows that there is a change in the structure of the transportation process and the turnover of goods will essentially change abruptly. It is obvious that a single trend will not provide qualitative and reliable forecasting. In this case, it is necessary to investigate the nature of the change in the trend of the dynamic series, namely whether it is continuous or piecewise continuous.

In the simulation of the dynamic series trend, the initial statistical data are divided into two sets: before the time of the jump and after that moment. Further, continuous regression functions are constructed (separately for each set).

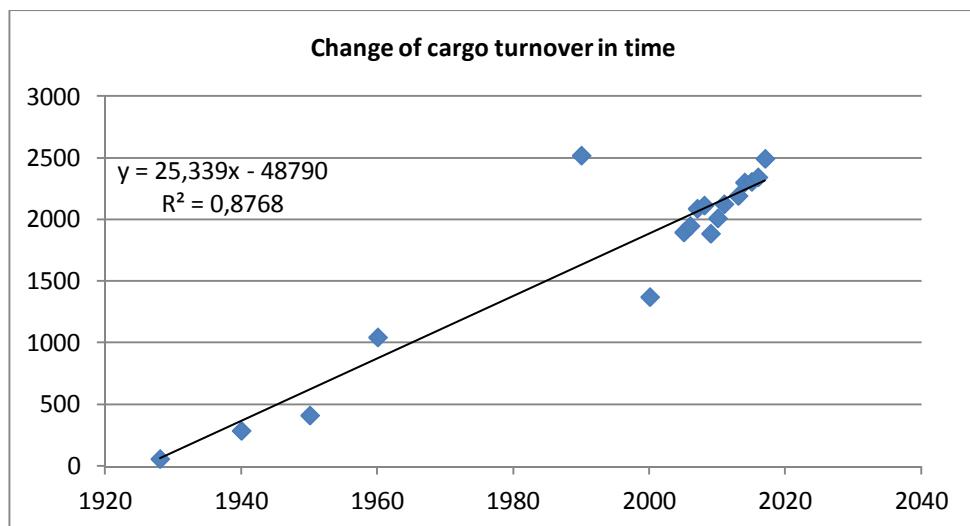
On the basis of statistical data considered in the work as the values of the time series, the autocorrelation coefficients presented on the correlogram are determined (Fig. 1).



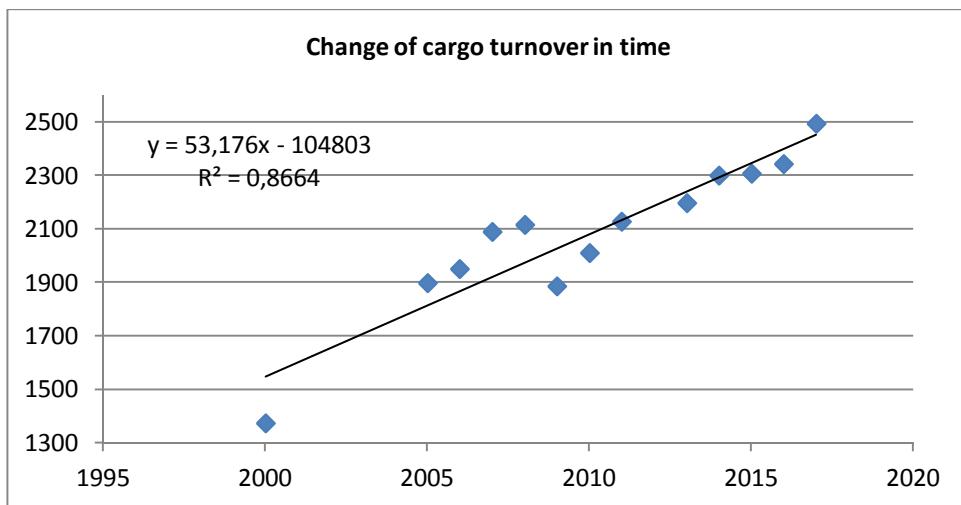
**Fig.1.** Correlation of autocorrelation coefficients.

From the correlogram follows the presence of an increasing trend and the absence of oscillatory changes in cargo turnover over time [2]. Based on this conclusion, a mathematical model of the dependence of cargo turnover on time (year).

Modeling in the work is performed using the least squares method [2], according to statistical data [1] (Fig. 2) and on the part of data (Fig. 3). The linear model is considered as mathematical models. The Regression tool of the Excel spreadsheet analysis Package add-in was used to model linear regression [3].

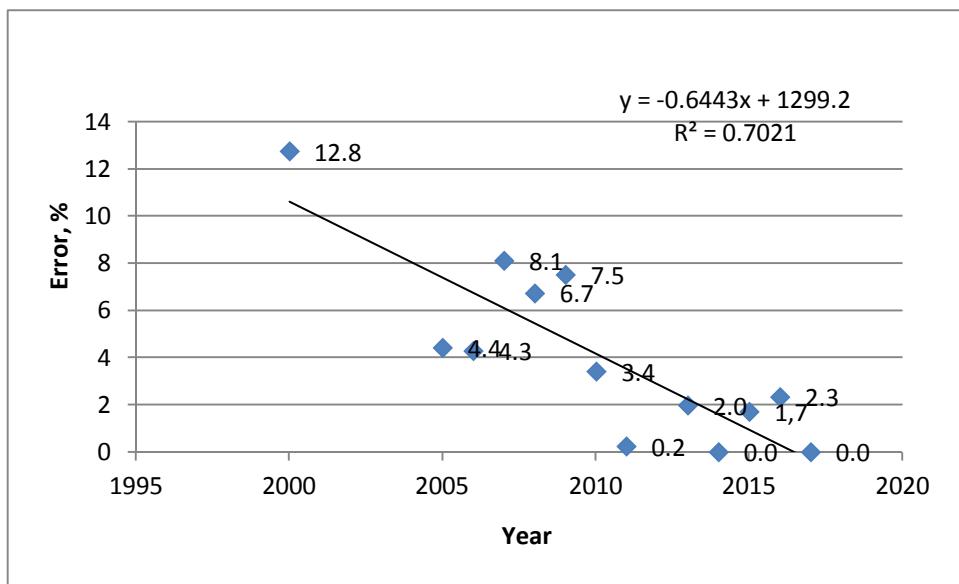


**Fig. 2.** “Trend line” based on data from 1920 to 2017.



**Fig. 3.** “Trend line” based on data from 1998 to 2017

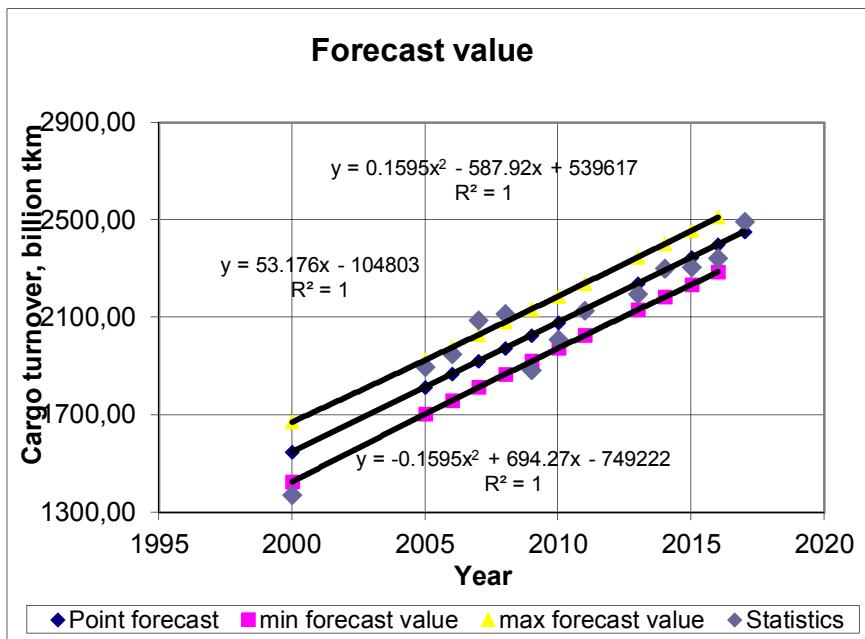
Trend line analysis (Fig. 2 and Fig. 3) and the values of their parameters allows us to conclude that the quality of the constructed regression functions is high. The coefficient of determination is 0.88 and 0.87. Further research was conducted on the dependence of rice. 3. The estimation of time series approximation errors showed that for the linear dependence the local maximum error is 12.8% and the integral average is 4.1%. Fig. 4 shows the graph and the error values in percent.



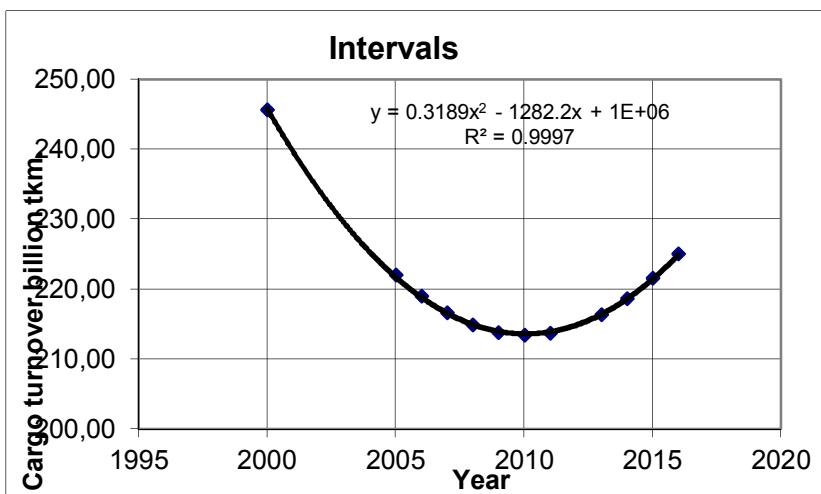
**Fig. 4.** Statistics and graph of the error of regression

To assess the adequacy of the regression equations obtained, the forecast for 2018 and its comparison with the real value were used. In Table 1 the penultimate value of cargo turnover at number 19 is real, and the last at number 20 is calculated according to the received model (Fig. 3). The error was 3.48%.

Assessment of the statistical reliability of the regression equation generally produced with the help of F-Fisher criterion. Evaluation of the models by Fisher's criterion allowed to reject the null hypothesis about the random nature of the equations and regression coefficients, and, consequently, for the estimated models, to accept an alternative hypothesis about the statistical significance of the regression equations [4]. On the basis of the linear model constructed and recommended for practical application, point and interval forecasting of cargo turnover was performed [5]. Fig. 5 and Fig. 6 show the graphs of the forecast.



**Fig.5.** Graph of point and interval forecast values of cargo turnover



**Fig. 6.** The values of the confidence interval

Since the parameters characterizing the quality of the models and its adequacy have shown high performance, it is advisable to apply the described methodology of cargo turnover modeling in order to predict it in practice.

## References

1. Russian statistical yearbook. 2018: Stat. SB.Rosstat, M., 694 (2018)
2. P. V.Gerasimenko, V. A. Khodakovsky. Sat. Tr. International scientific and methodological conference. St. Petersburg.: PGUPS, 84-88 (2014)
3. P. V. Gerasimenko, G. B. Staresina. In the book: Government and business. Modern problems of economy. Proceedings of the VII International scientific practical conference. North-Western Institute of Management Ranepa under the President of the Russian Federation, Faculty of Economics and Finance, 111 – 116 (2015)
4. P. V. Gerasimenko, G. B. Titov. Proceedings of St. Petersburg University of railway engineering. **3 (40)**, 162-167 (2014)
5. P. V. Gerasimenko. Proceedings of the International Academy of Sciences of higher school, **56.2**, 64-70 (2011)