UDC 519.23 https://doi.org/10.23947/2541-9129-2020-1-43-47

Mathematical analysis and forecasting of the dynamics of air pollution by stationary sources in the Russian Federation

E. S. Gorbacheva, I. M. Peshkhoev

Don State Technical University (Rostov-on-Don, Russian Federation)

Introduction. The article considers the dynamics of air pollution by stationary sources in the Russian Federation from 1998 to 2016. Harmful emissions into the atmosphere cause great harm to all living organisms. As a result, the life expectancy of the population is significantly reduced. Therefore, the assessment of pollution volumes and subsequent measures for the protection of atmospheric air are priority tasks of our time.
Problem statement. The objectives of the study are to analyze the dynamics of pollution, build a mathematical model of this process, and implement a forecast for a five-year period.
Theoretical part. Data for the work is taken from the official statistical book. Microsoft Excel and StatSoftStatistica computer technologies are used for calculations.
Conclusion. Based on the analysis, an adequate mathematical model is constructed, which may be of interest for predicting the anthropogenic impact on the environment.

Keywords: atmospheric pollution, mathematical model, least squares method, forecasting.

For citation: Gorbacheva E.S., Peshkhoev I.M. Mathematical analysis and forecasting of the dynamics of air pollution by stationary sources in the Russian Federation: Safety of Technogenic and Natural Systems. 2020;1: 43‒47. https://doi.org/10.23947/2541-9129-2020-1-43-47

Introduction. Air pollution causes great harm not only to living organisms, but also to the hydrosphere, soil and vegetation cover, buildings, structures and other objects. People who live in places with polluted air often have diseases such as allergies, cancer, and various lung diseases. As a result, life expectancy is significantly reduced. Therefore, the protection of atmospheric air is one of the priorities of our time.

Main sources of atmospheric pollution in Russia are: industry, transport, utilities and agriculture. The level of air pollution depends, as a rule, on the degree of urbanization and industrial development of the territory.

The paper considers the amount of harmful substances (hereinafter — pollution) released by stationary sources into the atmosphere in 1998-2016. The task of the authors is to analyze the dynamics of changes in pollution, to build a mathematical model of this process and to forecast it for a five-year period. The source data is taken from the book "Russia in numbers" [1].

Methodology. The values of atmospheric air pollution are presented in Table 1, and Fig. 1 shows their graphical representation.

Table 1

Air pollution values in the Russian Federation

Year	Pollution, million tons	Year	Pollution, milliontons
1998	18,7	2008	20,1
1999	18,5	2009	19
2000	18,8	2010	19,1
2001	19,1	2011	19,2
2002	19,5	2012	19,6
2003	19,8	2013	18,4
2004	20,5	2014	17,5
2005	20,4	2015	17,3
2006	20,6	2016	17,3
2007	20,6	-	-

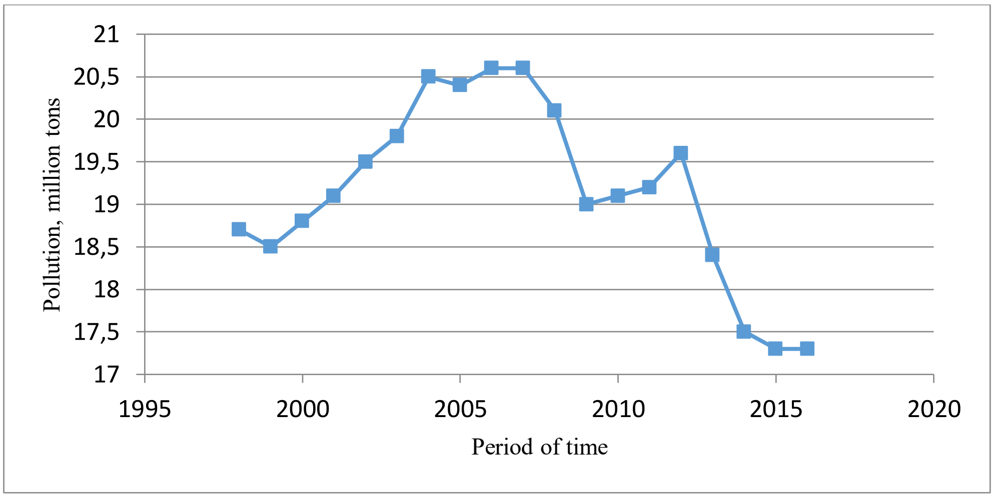

Fig. 1. Dynamics of changes in air pollution in the Russian Federation

The initial data show fluctuations in pollution against the downward trend, so the model dependence of pollution on time is a linear combination of monotonically decreasing and oscillatory functions. The monotonically decreasing function is an exponent, and the oscillating function is a sinusoid [2-4]:

$$\bar{X} = A + Be^{-C(t-1998)} + D \cdot \sin\left[\frac{2\pi}{E}(t+F)\right], \quad (1)$$

where $\bar{X}$ — the mathematical expectation of the pollution value; t — the year; A, B, C, D, E, F — the parameters of the exponential-harmonic function.

Since there are many adjustable parameters in function (1): three parameters in the exponent (constant, amplitude, and time) and three parameters in the sine wave (amplitude, frequency, and shift), and the initial data is small, the process of model identification was divided into two stages.

At the first stage, ignoring the fluctuations, the exponential dependence was identified:

$$\bar{Y} = A + Be^{-C(t-1998)}.$$

The coefficients A, B and C are found by solving the optimization problem:

$$\sum_i (Y_i - Y(t_i))^2 \to min, \quad (2)$$

where t_i — the value of the argument (time); Y_i — the actual value of pollution:

$$Y(t_i) = A + Be^{-C(t_i-1998)}.$$

At the second stage the oscillatory part is considered:

$$\bar{Z} = D \cdot sin\left[\frac{2\pi}{E}(t+F)\right].$$

The coefficients D, E and F are found by solving the optimization problem:

$$\sum_i (Z(t_i) - (Y_i - Y(t_i)))^2 \to min, \quad (3)$$

where t_i — the value of the argument (time):

$$Z(t_i) = D \cdot sin\left[\frac{2\pi}{E}(t_i+F)\right];$$

Y_i – the actual value of pollution

$$Y(t_i) = A + Be^{-C(t_i-1998)}.$$

Results. To find unknown coefficients, we used the functions of the *Microsoft Excel* solution search package [5-7]. As a result of solving the optimization problem (2) by the least squares method [8], the coefficients were found: A = 15,4; B = 4,378; C = 0,017. As a result of solving the optimization problem (3), the coefficients were found: D = 1,253; E = 19,287; F = -2001,84.

The final type of the dependency (1):

$$\bar{X} = 15{,}4 + 4{,}378 \cdot e^{-0{,}017(t-1998)} + 1{,}253 \cdot sin\left[\frac{2\pi}{19{,}287}(t - 2001{,}84)\right]. \quad (4)$$

Table 2 shows [9] the characteristics obtained from the model (4).

Table 2

Characteristicsof themodel

Year	Actual pollution, million tons	Estimated pollution, milliontons	Absolutedeviation, milliontons	Relativedeviation, %	Square of the relative deviations, %
1998	18,7	18,587	0,113	0,603	0,363
1999	18,5	18,700	0,200	1,083	1,173
2000	18,8	18,920	0,120	0,639	0,408
2001	19,1	19,216	0,116	0,605	0,366
2002	19,5	19,548	0,048	0,247	0,061
2003	19,8	19,875	0,075	0,379	0,144
2004	20,5	20,155	0,345	1,684	2,836
2005	20,4	20,350	0,050	0,243	0,059
2006	20,6	20,434	0,166	0,805	0,648
2007	20,6	20,390	0,210	1,017	1,035
2008	20,1	20,217	0,117	0,581	0,337
2009	19	19,925	0,925	4,866	23,678
2010	19,1	19,538	0,438	2,293	5,258
2011	19,2	19,091	0,109	0,567	0,322
2012	19,6	18,624	0,976	4,977	24,774
2013	18,4	18,181	0,219	1,191	1,419
2014	17,5	17,801	0,301	1,718	2,951
2015	17,3	17,518	0,218	1,259	1,585
2016	17,3	17,356	0,056	0,323	0,105

The following indicators are calculated:

- the correlation coefficient is 0,936;
- the average value of the absolute deviation is 0,253 million tons;
- the average value of the relative deviation is 1,320 %;
- the maximum relative deviation is 4,977 %;
- the standard deviation is 1,885 %;
- the model's adequacy is 1,889 %.

Fig. 2 shows the atmospheric pollution values obtained from the model (4).

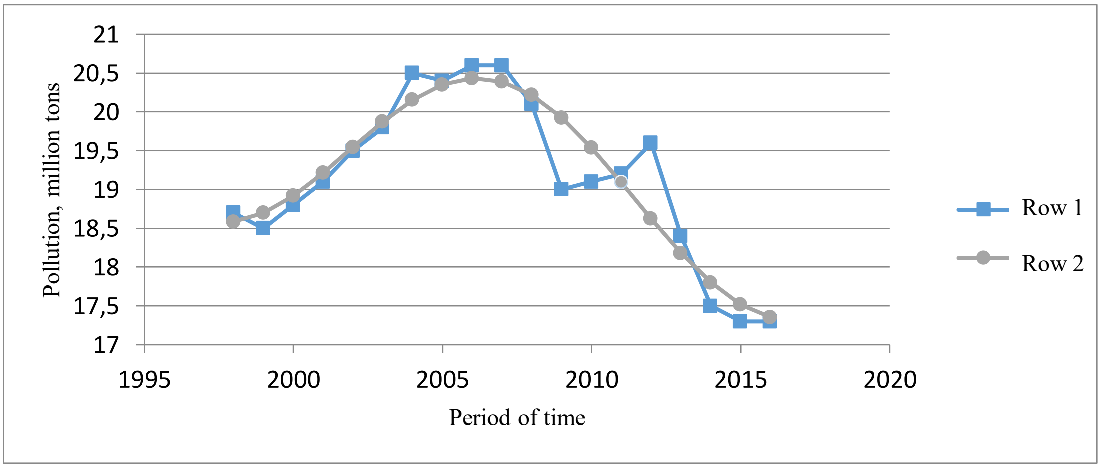

Fig. 2. Dynamics of changes in the atmospheric pollution in the Russian Federation:
row 1 — actual data; row 2 — calculated data

The forecast for the period 2017-2021 was also obtained using formula (4) [10]. Table 3 presents the forecast results.

Table 3

Comparison of actual and predicted data on pollution

Year	Real value, milliontons	Predictedvalue, milliontons
2017	17,5	17,326
2018	17,1	17,426
2019	-	17,639
2020	-	17,937
2021	-	18,284

The forecast error is 0,994 % for 2017, and for 2018 it is 1,906 %.

Conclusion:

- the dynamics of air pollution from stationary sources in the Russian Federation in 1998-2016 is analyzed;
- a relationship was found that reflects the change in the amount of harmful emissions over the specified period;
- an adequate mathematical model has been constructed that may be of interest for predicting anthropogenic impacts on the environment and can be used as one of the tools for developing alternative models for such forecasting;
- the forecast of the dynamics of changes in atmospheric pollution in the Russian Federation for a five-year period has been implemented.

References

1. Malkov P.V. (ed.) Russia in numbers. Summary of statistics [Rossiya v tsifrakh. Kratkiy statisticheskiy sbornik] Moscow, Federal State Statistics Service, 2019.549 p. (In Russ.)
2. Gmurman V.E. Probability theory and mathematical statistics: textbook for universities [Teoriya veroyatnostey i matematicheskaya statistika: uchebnik dlya vuzov] Moscow, Yurayt, 2018. 479 p. (In Russ.)
3. Minashkin V.G. Statistics: textbook for universities [Statistika: uchebnik dlya vuzov] Moscow, Yurayt, 2018. 448 p. (In Russ.)
4. Mkhitaryan V.S. Statistics: textbook and workshop for universities [Statistika: uchebnik i praktikum dlya vuzov] Moscow, Yurait, 2018. 464 p. (In Russ.)
5. Sobol B.V., Peshkhoev I.M., Borisova L.V., Ivanochkina T.A. Workshop on statistics in Excel: textbook for universities [Praktikum po statistike v Excel: uchebnoe posobie dlya vuzov] Rostov-on-Don, Feniks, 2010. 381 p. (In Russ.)
6. Vadzinskiy R.N. Statistical calculations in Excel [Statisticheskie vychisleniya v srede Excel] Saint-Petersburg, Piter, 2008.608 p. (In Russ.)
7. Yakovlev V.B. [Statistika. Raschety v Microsoft Excel: uchebnoe posobie dlya vuzov] Moscow, Yuray, 2018.353 p. (In Russ.)
8. Zenkov A.V. Numerical methods: textbook for universities [Chislennye metody: uchebnoe posobie dlya vuzov] Ekaterinburg, publishing house of the Ural University, 2016. 124 p. (In Russ.)
9. TyurinYu.N., Makarov A.A. Data analysis on a computer [Analiz dannykh na kompyutere] Moscow, MTSNMO, 2016. 367 p. (In Russ.)
10. Sadovnikova N.A., Shmoylova R.A. Time series analysis and forecasting. Issue 2: textbook for universities [Analiz vremennykh ryadov i prognozirovanie. Vyp 2: uchebnoe posobie dlya vuzov] Moscow State University of Economics, Statistics, and Informatics, Moscow, 2004.199 p. (In Russ.)

Submitted 14.12.2019
Scheduled in the issue 10.01.2020

Information about the authors:

Elena S. Gorbacheva, Master's degree student, Don State Technical University (344000, Russian Federation, Rostov-on-Don, Gagarin square, 1), ORCID: https://orcid.org/0000-0002-7459-2980, lukashewa.alena@yandex.ru

Isa M. Peshkhoev, Associate professor, Department of Information Technologies, Don State Technical University (344000, Russian Federation, Rostov-on-Don, Gagarin square, 1), Candidate of physical and mathematical sciences, Associate professor, ORCID: https://orcid.org/0000-0003-0352-9586, peshkhoev@rambler.ru

Contribution of the authors

E. S. Gorbacheva — statement of the problem, collection and analysis of literature, research methodology, derivation of the mathematical dependence, identification model and calculation of its coefficients, checking the adequacy of mathematical models, receiving a forecast and verification of data errors, text editing, literary analysis, preparation of the manuscript; I.M. Peshkhoev — scientific supervision, formulation of the basic concept of the study and structure of the article.