Improving the accuracy of the calculated environmental monitoring of atmospheric air quality in cities based on a corrective neural network

Y A Tunakova¹, S V Novikova¹, A R Shagidullin², V S Valiev² and K N Novikova¹

¹Kazan National Research Technical University named after A. N. Tupolev (KNRTU-KAI), Kazan, Russia
²Institute for Environmental Problems and Subsoil Use of the Academy of Sciences of Tatarstan, Kazan, Russia

juliaprof@mail.ru

Abstract. To obtain the values of the concentrations of atmospheric impurities with a spatial-temporal resolution throughout the city, which is not achieved during systematic experimental observations, it is recommended to use calculation methods. A significant advantage of the calculated concentrations, in contrast to the experimental ones, is the ability to determine the concentration at any point in the study area at the present time and in the future when the characteristics of stationary pollution sources change. To calculate the dispersion of emissions of pollutants in the air and determine the surface concentrations of impurities, the order of the Ministry of Natural Resources and Environment of the Russian Federation regulates the use of the universal program for calculating atmospheric pollution (UPCAP) “Ecolog”. To increase the convergence of the measured concentrations of impurities and calculated concentrations obtained using UPCAP “Ecolog-Gorod” we propose the use of neural network technologies.

1. Introduction
Artificial neural networks trained on an array of impurity concentrations allow considering all the hidden relationships between the levels of pollution of the surface layer of the atmosphere and the factors that form it [1]. We set the task of neural network adaptation correction of impurity concentrations calculated with the help of the universal program for calculating atmospheric pollution (UPCAP) “Ecolog-Gorod”, taking into account the main meteorological parameters [2]. With the help of the UPCAP “Ecolog-Gorod” software package, calculations were performed at the points of air sampling and measuring the concentrations of priority impurities at automated stationary posts of the Ministry of Ecology and Natural Resources of the Republic of Tatarstan (MENR RT). The maximum one-time concentrations of impurities coming from all stationary sources of enterprises were determined, taking into account the simultaneous operation of the sources. Comparison of calculated by UPCAP “Ecolog” and experimentally measured concentrations for almost all impurities shows a significant error, reaching about 30 % and even higher for individual impurities.

In this regard, we set the task of increasing the convergence of experimental and calculated concentrations of atmospheric impurities through the use of neural network technologies.

2. Methods
Significant meteorological predictor parameters were selected on the basis of an analysis of observational data and the real type of relationships between meteorological factors and the...
concentrations of impurities in the surface layer of atmospheric air. An analysis of the dependences between the measured concentrations of impurities and meteorological parameters in the territory of Nizhnekamsk showed the presence of close relationships with wind direction, wind speed, temperature, pressure, and air humidity. They formed clear correlation pleiades with a high level of statistical significance ($p = 0.01-0.005$). Therefore, these meteorological parameters were chosen as inputs to the neural network. A series of independent experiments were conducted to train and test possible structures, correcting neural networks for each impurity. As a result, the following general structure of the corrective neural network was determined:

1) Paradigm – multilayer perceptron (a universal approximator of regression dependence)
2) The number of hidden layers – 1 (determined according to the Kolmogorov theorem on the structure of the perceptron as a universal approximator)
3) Number of input neurons – 8
   - Neurons 1 and 2 – coordinates of the calculation point in absolute coordinates in metres;
   - Neuron 3 – air temperature;
   - Neuron 4 – atmosphere pressure;
   - Neuron 5 – wind speed;
   - Neuron 6 – direction of the wind;
   - Neuron 7 – air humidity;
   - Neuron 8 – calculated value of impurity concentration (software UPCAP “Ecolog-Gorod”).

All input parameters are normalized by the “hyperbolic tangent” function (based on the different scales of the initial data and the types of activation function). The proposed approach was tested separately for each of the main impurities. The network was trained using the Resilient backpropagation algorithm. From the original dataset, a subset of test data is randomly allocated to test the adequacy of the model and assess its generalizability. The volume of test suites was 10% of the original dataset [3]. The impurity concentrations measured at the automated stationary posts of the Ministry of Ecology and Natural Resources of the Republic of Tatarstan in Nizhnekamsk were used to train the neural network. To verify the adequacy of the calculation results, a comparison was made of the impurity concentrations obtained by calculation and experiment at the sampling points.

3. Results

3.1. Correction model for calculated values of nitric oxide (NO) concentrations

The average calculation error on the test set was 14.3 %. In comparison with the initial calculation error of 256.4 %, the accuracy is increased by 18 times. Graphically, the accuracy parameters of the constructed model are well demonstrated in Figure 1.

![Figure 1. Comparison of calculated (“Ecolog-Gorod”) – blue line, experimentally measured (red line), and corrected (green line) data for nitric oxide.](image-url)
3.2. Correction model for calculated values of nitrogen dioxide (NO$_2$) concentrations

Average margin of error on the test set was 14.68 %. Accuracy increased by 19 times. A graphical illustration of the model’s accuracy is shown in Figure 2.

![Figure 2](image_url)

**Figure 2.** Comparison of calculated (“Ecolog-Gorod”) - blue line, experimentally measured (red line), and corrected (green line) data for nitrogen dioxide.

3.3. Correction model for calculated values of sulphur dioxide (SO$_2$) concentrations

Average error on the test set was 20.81 %. Accuracy increased by 19 times. A graphical illustration of the model’s accuracy is shown in Figure 3.

![Figure 3](image_url)

**Figure 3.** Comparison of calculated (“Ecolog-Gorod”) – blue line, experimentally measured (red line), and corrected (green line) data for sulphur dioxide.

3.4. Correction model for calculated values of carbon monoxide (CO) concentrations

Average error on the test set was 20.88 %. Accuracy increased by 4 times. A graphic illustration of the model's accuracy is shown in Figure 4.
Figure 4. Comparison of calculated (“Ecolog-gorod”) – blue line, experimentally measured (red line), and corrected (green line) data for carbon monoxide.

4. Conclusion
Neural network correcting models are able to increase the accuracy of calculating the concentrations of impurities in the atmospheric air in comparison with the regulated calculation implemented in the “Ecolog-Gorod” software package in several times, in the experiments performed – from 2.5 to 22 times. Each admixture requires the construction of its own corrective neural network model, the topology of which is selected empirically.

To build corrective models, it is necessary to use an extended list of meteorological parameters, as well as to have a representative sample of experimentally measured values of concentrations of pollutants in the atmospheric air.

Acknowledgements
The research was carried out with the financial support of the Ministry of education and science of the Russian Federation in the framework of fulfilling obligations under Agreement number 075-03-2020-051/3 dated 09.06.2020 (topic number FZSU-2020-0021).

References
[1] Cabaneros S M, Calautit J K and Hughes B 2019 Environ. Model. Software 119 285
[2] Ramsey N, Klein P and Moore B 2014 Atmos. Environ. 86 58
[3] Krundyshev V 2019 Automat. Control Comput. Sci. 53 1012