ISSN 1068-3739, Russian Meteorology and Hydrology, 2015, Vol. 40, No. 10, pp. 658–666. Allerton Press, Inc., 2015.
Original Russian Text . P.F. Demchenko, A.S. Ginzburg, G.G. Aleksandrov, A.I. Vereskov, G.I. Gorchakov, N.N. Zavalishin, P.V. Zakharova, E.A. Lezina, N.I. Yudin, 2015, published in Meteorologiya i Gidrologiya, 2015, No. 10, pp. 31–43.

Statistical Modeling of Average Daily Concentration of Pollutants in the Atmosphere over Moscow Megalopolis by the Multiple Regression Method

P. F. Demchenko[a], A. S. Ginzburg[a, b, c], G. G. Aleksandrov[a], A. I. Vereskov[d], G. I. Gorchakov[a], N. N. Zavalishin[a], P. V. Zakharova[e], E. A. Lezina[e], and N. I. Yudin[†]

[a]*Obukhov Institute of Atmospheric Physics, Russian Academy of Sciences, Pyzhevskii per. 3, Moscow, 119017 Russia, e-mail: pasha@ifaran.ru*

[b]*Moscow Technological Institute, Leninskii pr. 38A, Moscow, 119334 Russia*

[c]*Aerokosmos Space Research Institute, Gorokhovskii per. 4, Moscow, 105064 Russia*

[d]*All-Russian Research Institute of Metrological Service, ul. Ozernaya 46, Moscow, 119361 Russia*

[e]*Mosekomonitoring State Nature Protection Organization, ul. Novyi Arbat 11b, str. 1, Moscow, 119019 Russia*

Received February 18, 2015

Abstract—Presented are the results of the construction of statistical models of the time series of air pollutants (particulate matter with the size of less than 10 m (PM_{10}), CO, and NO_2) for the network of automatic stations of air pollution control over Moscow megalopolis. The multiple nonlinear regression of pollutant concentration to external factors (meteorological and other) and values of concentration on previous days are taken as the statistical model of the time series of the average daily concentration of a certain pollutant. The nonlinear nature of the models of the time series can be caused with the dependence of pollutant concentration on wind speed and with other factors. Nonlinear regression based on the relatively short learning samples was used for simulating the series of average daily concentration of pollutants. The computations demonstrated that this gives much higher correlation between the computed and observed values of concentration and smaller standard deviation as compared with the model of inertial forecasting.

DOI: 10.3103/S1068373915100039

Keywords: Concentration of air pollutants, statistical model, Moscow megalopolis

INTRODUCTION

The modeling of the temporal evolution of pollutants in the urban environment is important for solving a number of problems, first of all, for producing the forecasts of their concentration with various lead times [14, 15]. Besides, to solve the problems of air quality control, the need arises in revealing relations between the rate and localization of emission sources (industry and transport), on the one hand, and the results of measurement of pollutant concentration obtained at stationary or mobile monitoring stations, on the other hand [16]. As a rule, the existing models for computing the pollutant concentration are divided into statistical, simplified univariate quasistationary [3], or integral box models [13, 27], and Eulerian or Lagrangian three-dimensional models (3D models) [11, 14]. In Roshydromet practice, the COSMO-ART three-dimensional photochemical model is used, for example, for the Moscow region [11].

Unlike 3D models, statistical models do not require information on the rate of emission sources. It is difficult to obtain such data. The computation costs of obtaining the prognostic values in statistical models are much lower than in 3D models. At present, the accuracy of the forecasts of pollutant concentration with the lead time of 24 hours based on statistical and 3D models almost does not differ [19, 21].

In statistical models, the initial data are the concentration of pollutants (internal predictors or endogenous variables) which were observed before and on the day of forecast production; the predicted and/or observed values of external predictors (exogenous variables). Exogenous predictors are meteoro-

logical parameters (temperature, wind direction, wind speed, etc.) and other factors (the calendar day and month, motor road workload, etc.). The model either is multiple (or linear) regression (MLR) with the values of coefficients computed from the past measurements [32] or contains implicit functional dependence between the predicted parameter and predictors as for the methods of artificial neural networks (ANNs) [12, 17, 20, 22–25, 28–33]. As a rule, to predict air pollution, hybrid models with the elements ANN or MLR are used; they are supplemented with the methods of fuzzy logics [16], clusterwise regression [26], and the nearest neighbor method [25] which is similar to the pattern recognition technique [5]. Hybrid models are the models with the joint use of statistical and 3D models of PM_{10} forecasting on the territory of Europe [21].

The number of air quality monitoring stations in Moscow megalopolis has continuously increased in recent 10 years. In 2014 the number of stationary automatic air pollution control stations (AAPCSs) reached 47 including four stations installed at different heights at the Ostankino Tower (Ostankino high-altitude meteorological complex), six stations on the territories which were later joined to Moscow, and one station in the Moscow oblast near Zvenigorod. Mobile AAPCSs were additionally used in the seven districts of Moscow [9, 10].

The acquired data of air pollution monitoring jointly with the results of meteorological observations give a unique opportunity for constructing and adjusting different statistical and hybrid models of pollutant evolution in order to predict the pollutant concentration in the air over the urban areas. Paper [18] deals with the first results of the construction of the model of multiple nonlinear regression for the determination of the variations of nitrogen dioxide NO_2 concentration at Biryulevo station. The given paper presents the results of the testing of the model that works with the concentration of particulate matter with the size of less than 10 m, NO_2, and CO based on the data from a number of the monitoring network stations during the cold and warm periods in different years.

INITIAL DATA

The authors used the data of Mosekomonitoring State Nature Protection Organization obtained from the AAPCS network in Moscow and adjoining areas. The measurements were carried out in real time mode round the clock with the quantization of output data with the interval of 20 minutes. The average daily values were computed by summing these data. The areas of station locations were divided into inhabited, natural, mixed, and those adjoining to motorways. The scheme of the location of stations in Moscow, Zelenograd, Zvenigorod, and New Moscow is presented in Fig. 1. The concentration of 22 pollutants typical of emissions from anthropogenic sources in Moscow was measured. The concentration of particulate matter with the size of <10 m (PM_{10}) and <2.5 m ($PM_{2.5}$), carbon oxide (CO), nitrogen dioxide (NO_2), nitrogen oxide (NO), total hydrocarbons (CH_x), sulfur dioxide (SO_2), and ozone (O_3) are measured at the majority of stations. The concentration of specific compounds (H_2S and NH_3) is measured close to the emission sources. The concentration of 16 pollutants (including formaldehyde, benzene, etc.) is additionally measured near the Third Ring Road [9].

The concentration of PM_{10} was measured by the TEOM instrument, the concentration of NO_2, by the ME-9841V, TD, and Thermo-17i instruments, and the concentration of CO, by the L-100 instrument. The measurement error did not exceed 20% in all cases.

The meteorological data on the surface air layer conditions (temperature, pressure, relative humidity, wind direction, and wind speed) were collected from all stations located in the megalopolis. Similar data (except pressure) obtained at the height of 503 m at the Ostankino Tower were also used.

Papers [7, 8] provide the data on the use of the statistical model of the ozone concentration forecast with the lead time of 24 and 48 hours for the area under study (the town of Dolgoprudny) using the multiple linear regression.

The given paper considers the series of CO, NO_2, and PM_{10}. To develop and test the statistical model, the data obtained in 2009 and 2011 were used (the data for 2010 were removed due to abnormal summer fires). To test the prognostic potential, the data of 2013 were used.

THE MODEL AND PREDICTORS

To carry out the statistical modeling of air pollution, let us consider the values of average daily concentration of any pollutant $Y_i = Y(t_i)$ averaged for the ith day for the separate monitoring station. Let us assume the multiple regression model (not necessarily the linear one) for the time series Y_i:



Fig. 1. The scheme of the location of Mosekomonitoring automatic control stations of air pollution. The location of stations in Zelenograd, Zvenigorod, and New Moscow are presented separately. (*1*) Stations close to motorways; (*2*) mixed territory; (*3*) inhabited territory; (*4*) natural territory; (*5*) multilevel monitoring station; (*6*) the station in TINAR; (*7*) the territory monitored by the mobile station.

$$Y_i \quad \hat{Y}_i \quad {}_i; \ \hat{Y}_i \quad \sum_{l\ 1}^{L} b_l Y_{i\ l} \quad \sum_{m\ 1}^{M} F^m(X_i^{(m)}, \mathbf{c}_m) \tag{1}$$

where $\hat{Y}_i$ is the value of concentration predicted for day t_i; ${}_i$ is the forecast error.

The first members in (1) in the forecast function $\hat{Y}_i$ contain information about the past values of the concentration under study (endogenous variables) and make up the autoregression part of the statistical model. The coefficients b_l are computed from the observational data using the accepted algorithms of construction of statistical models.

The second part of the expression for the function $\hat{Y}_i$ describes dependence on the external (relative to the time series under study) exogenous predictors $X^{(m)}$. As a rule, meteorological parameters characterizing air transport conditions and the internal processes of pollutant transformation are selected as external predictors. The functions $F^{(m)}(X^{(m)}, \mathbf{c}_m)$ are considered as prescribed in the model, and the vector of parameters $\mathbf{c}_m$ for

each of them is determined empirically. Both the predicted values $X^{(m)}(t_i)$ and the values obtained before as a result of observations $X^{(m)}(t_{i-1})$ can be prescribed for external predictors [34].

In the case of the linear approximation $F^{(m)} = c_m X^{(m)}$, prognostic equation (1) turns into the multiple linear regression model. In the case of the approximation $F^{(m)} = c_m \quad {}^{(m)}(X^{(m)})$, equation (1) corresponds to the multiple linear regression model where nonlinear relations have been preliminary removed [12]. We will use equation (1) in the general form; however, autoregression and external variables will be considered uniformly, and unknown coefficients will be determined based on the degree of similarity of the data of forecasts and observations using the selected algorithm of the realization of the least-squares method [1].

In the proposed model variant the main group of meteorological predictors includes wind speed *WA* and wind direction *WD*, air temperature *T*, relative humidity *q*, and pressure *p*. These parameters are measured at the majority of monitoring stations of the network under consideration. The values WA_{503}, WD_{503}, T_{503}, and q_{503} measured at the height of 503 m at the Ostankino Tower are of special importance. The measurements located at the Ostankino Tower minimize the impact of urban surface inhomogeneities. The horizontal variability of surface air pressure at the scale of Moscow megalopolis is low, and the values of p_s are assumed to be equal to those at Lomonosov Moscow State University (MSU) weather station. Along with the temperature at the height of 503 m, air temperature T_s at the ground-based monitoring station in Ostankino was also considered as a predictor in the test model variant.

Cyclic variables associated with the calendar month and day make up the special group of predictors in statistical forecasting. Wind direction is also referred to cyclic variables [34]. As a rule, the expansion using the functions $\cos[2 \ (x \quad x_{\min})/(x_{\max} \quad x_{\min})]$ and $\sin[2 \ (x \quad x_{\min})/(x_{\max} \quad x_{\min})]$ is used for cyclic variables *x* [34]. The calendar month can be present both as an additional predictor [23, 34] and as the dependence of linear regression coefficients on the season [7, 8]. The dependence of the concentration of air pollutants on the day of the week in big cities was called the "day-off effect" and is taken into account in the majority of statistical models. This effect is also observed in Moscow [6].

The proposed model takes into account the day-off effect in the form of an additional predictor (the summand in equation (1)) that is equal to zero on weekends and to one on working days. As a result, both holidays and long vacations (for example, on Christmas) can be taken into account. On those days the "sign of working day" predictor is equal to zero whereas it is equal to one on the rest of the days.

To determine the seasonal variations, the approximation by harmonic members is not used in the model. Instead of this, models with different coefficients for the warm and cold season and, perhaps, for the transition seasons are considered separately. Such approximation is used in the statistical forecasts of PM_{10} [24, 27] and $PM_{2.5}$ [25]. In the summand associated with the dependence of wind speed on wind direction, the nonlinear function $F^{(WD)} = c^{(WD)} \cos^{(WD)} \quad {}^{(WD)}$ is used in model (1). In the case of such approximation the wind direction giving the maximum contribution to the concentration variability can be singled out for each station.

The coefficient of correlation between the variability of air pollutant concentration and wind speed is negative under urban conditions, as a rule [23]. This was also corroborated by the results of computations carried out by the authors using the Mosekomonitoring network data. The negative sign of correlation is caused by the improvement in the ventilation of lower air layers at the high wind speed and by the formation of stagnant phenomena at the low wind speed.

At the same time, it was revealed [16] that at the wind speed of <4 m/s the micrometeorological processes defining the contribution of lower air layer to the ventilation should be taken into account. To take account of the possible stagnant phenomena, the dependence $F^{(WA)} = c^{(WA)}/(WA \quad 3.5)$ can be used in model (1). As the test computations based on the Biryulevo monitoring station data demonstrated, the empirical value of $F^{(WA)} = 3.5$ m/s described the variations of NO_2 in the best way.

Along with the nonlinear dependence of wind speed effects, the linear model $F^{(WA)} = c_L^{(WA)} WA$ can also be considered. The data of computations were compared with the linear and nonlinear parameterizations of the dependence on the wind speed. The comparison demonstrated that either these parameterizations give similar results or the nonlinear parameterization improves the agreement between the observed and simulated data. As the computation costs of the realization of these two versions do not differ, the nonlinear dependence is used hereinafter for correcting the computations under light wind conditions.

Below we provide the final form of the nonlinear regression function for modeling the average daily concentration of a certain pollutant in the atmosphere taking account of its concentration on the day of the forecast production

$$\hat{Y}_i \quad a_1 X_1 \quad a_9 \cos(X_2 \quad a_2) \quad \frac{a_3}{X_3 \quad 3.5} \quad a_4 X_4 \quad a_5 X_5 \quad a_6 X_6 \quad a_7 X_7 \quad a_8 . \qquad (2)$$

In (2), the single designations for the variables and coefficients with the consecutive numbering are introduced for the sake of convenience. The predictors referred to the day under prediction are as following: X_1 is the sign of the day (1 for the working day and 0 for the day off); X_2 and X_3 are the average daily values of wind direction (degree) and wind speed (m/s) at the height of 503 m at the Ostankino Tower; X_4 is the average daily relative humidity of surface air (the Ostankino Tower, %); X_5 is the average daily temperature at the height of 503 m (the Ostankino Tower, C); X_6 is the average daily surface pressure (MSU weather station, hPa).

There is one more predictor, X_7 is the average daily concentration of a pollutant on the day of the forecast production (mg/m^3). The additional computations were carried out when the concentration of pollutants a day before the forecast production was included into the list of predictors. They demonstrated that the inclusion of this predictor does not improve the results of the model computations.

To assess the quality of simulation of the observed variations of pollutant concentration by the model (1) using the series of N readings, the coefficient of correlation r between the predicted $\hat{Y}_i$ and observed Y_i values was used as well as the root-mean-square error of the forecast . The squares of these variables (determination coefficient and variance) are computed from the formulae

$$r^2 \quad \frac{(\ (Y_i \quad \overline{Y}_i)(\hat{Y}_i \quad \overline{\hat{Y}}))^2}{(\ (Y_i \quad \overline{Y}_i)^2)(\ (Y_i \quad \overline{Y}_i)^2)}, \quad ^2 \quad \frac{1}{N} \ (Y_i \quad \hat{Y}_i)^2. \qquad (3)$$

TESTING THE STATISTICAL MODEL OF THE SERIES OF POLLUTANT CONCENTRATION IN THE AIR OVER MOSCOW MEGALOPOLIS

The testing of the statistical model (2) was carried out using the data from several control stations of air pollution located in the areas with typical life conditions: in the districts of natural (MSU), inhabited (Ostankino), mixed (Biryulevo), and motorway (Kutuzovskii Prospect, MADI) types. The data were selected for cold (winter and spring) and warm (summer and autumn) three-month periods in 2009 and 2011 with the negative and positive temperature, respectively. The statistical model was constructed for the separate pollutant at every station. Meteorological parameters were taken from the data of observations on the day of the forecast. They were the analogs of the values predicted by meteorological models; however, the error related to the inaccuracy of these forecasts was removed at this stage.

First of all, it was interesting to consider the feasibility of construction of the statistical model using the relatively short learning samples. The selection of short learning samples was caused by the fact that the annual variations are not taken into account explicitly in the proposed variant of the statistical model. At the same time, the interannual variability of air pollution conditions is not taken into account either, though, as the test results demonstrated, it can be significant. For the same reason the data of abnormal periods, for example, of the summer 2010 characterized with significant fires in the adjoining areas, require special consideration going beyond the limits of the present paper.

To estimate the unknown parameters of the regression–autoregression mixed model (1) [1], the algorithm of the Gauss–Newton type was used for solving the nonlinear regression problems [2]. To avoid the computational singularity associated, for example, with the excess synchronicity in the evolution of separate predictors (multicollinearity), the specially developed iterative algorithm of the Gauss–Newton type was used with the detailed diagnosis of the results of computation of the parameters. The algorithm does not require the direct computation of the regression function derivatives and stipulates for the step selection and special procedure of regularization during iterations [4]. Besides, the algorithm was used with direct iterations based on the Gauss–Newton method for the special case (2) of the general model (1). This algorithm was realized in the R software package available for the users (http://www.r-project.org).

The important characteristic of the model is its ability to increase the determination coefficient r^2 as compared with the inertial forecast $\hat{Y}_i^{\text{per}} \quad Y_{i \ 1}$. For this purpose, the corresponding scattering diagrams are also compared separately in Figs. 2 and 3 along with the curves of the synchronous temporal variations of observed and simulated (using equation (2)) values of the concentration of some pollutants.

Figure 2 presents the results of comparison between the simulated and observed data for two samples. The absence of dramatic outliers is typical of the temporal variations of NO_2 concentration. The correlation

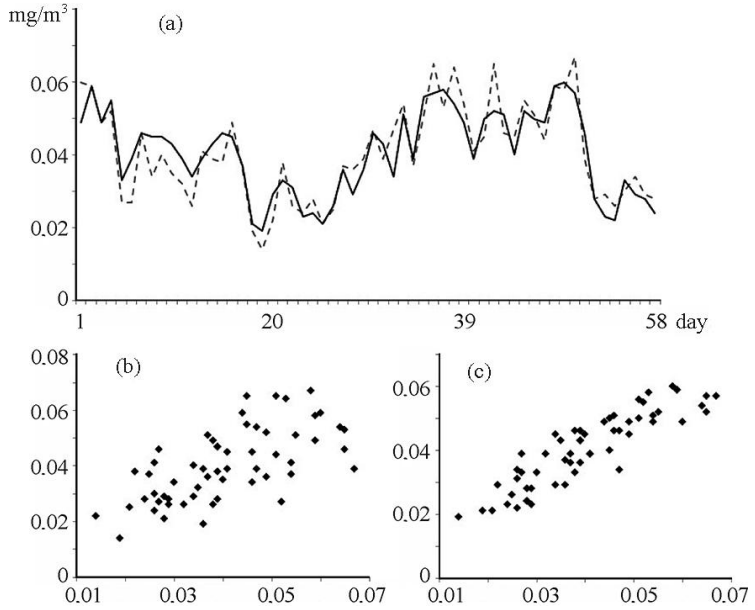


Fig. 2. The verification of the statistical model of the average daily values of NO_2 concentration. Kutuzovskii Prospect station, summer–autumn 2009. (a) Temporal variations of simulated by model (2) (the solid line) and observed (the dash line) concentration; (b, c) scattering diagrams for the inertial model ($r = 0.67$) and for model (2) ($r = 0.90$), respectively.

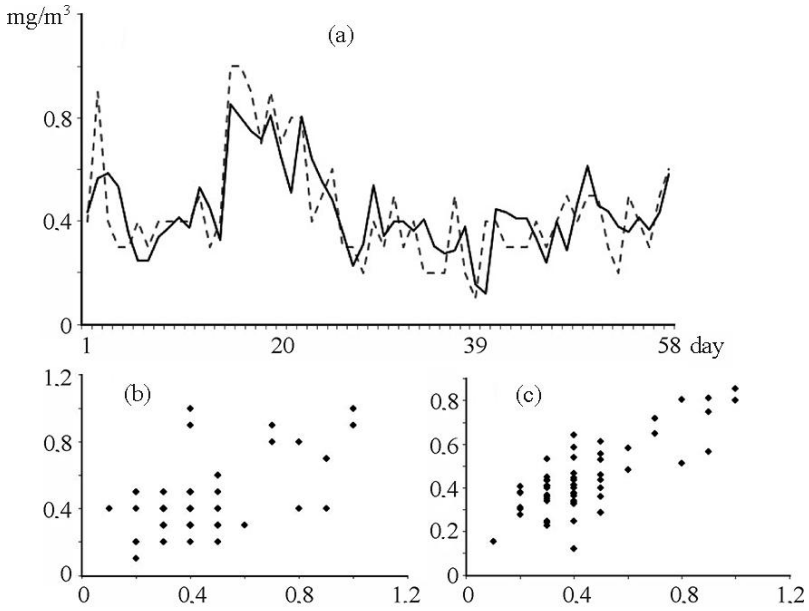


Fig. 3. The same as in Fig. 2 for the concentration of CO. The Ostankino Tower, winter–spring 2011. $r = 0.57$ for the inertial model and $r = 0.78$ for model (2).

coefficients can be high for the inertial forecast as well ($r = 0.67$). However, $r = 0.9$ for the model forecast that gives the increase in the determination coefficient r^2 by 80% as compared with the inertial forecast.

The typical example of the comparison between the measured and computed (using model (2)) values of concentration if there are dramatic variations for CO is given in Fig. 3. As clear from the figure, significant increase in the correlation coefficient is still observed for the model series as compared with the inertial forecast. Sometimes the situation can be improved by using certain special clustering methods [25].

The results of the testing of model (2) using the data of 29 AAPCS stations on NO_2 concentration in the air in May and June 2013

Station	Correlation coefficient		Root-mean-square error, mg/m^3	
	Model	Inertial forecast	Model	Inertial forecast
Polyarnaya	0.689	0.460	0.008	0.008
Ostankino				
348 m	0.626	0.572	0.004	0.012
248 m	0.525	0.398	0.005	0.010
130 m	0.677	0.569	0.010	0.010
MSU	0.822	0.776	0.013	0.011
Mar'inskii Park	0.647	0.528	0.007	0.010
MADI	0.822	0.758	0.011	0.015
Lyublino	0.806	0.692	0.011	0.013
Letnaya	0.644	0.520	0.016	0.009
Kosino	0.693	0.374	0.009	0.014
Kozhukhovo	0.687	0.497	0.009	0.016
Zelenograd				
microdistrict 16	0.682	0.516	0.007	0.024
microdistrict 11	0.660	0.551	0.011	0.014
microdistrict 6	0.694	0.591	0.006	0.008
Zvenigorod	0.776	0.671	0.006	0.012
Dolgoprudnaya	0.778	0.613	0.009	0.010
Gur'evskii Proezd	0.688	0.540	0.007	0.010
Biryulevo	0.729	0.506	0.007	0.011
Chayanova	0.707	0.568	0.011	0.013
Turistskaya	0.737	0.640	0.009	0.009
Tolbukhina	0.734	0.634	0.007	0.014
Spiridonovka	0.676	0.549	0.010	0.016
Ploshchad' Gagarina	0.669	0.502	0.010	0.014
Nizhnyaya Maslovka	0.781	0.604	0.011	0.024
Butlerova	0.891	0.879	0.021	0.014
Khamovniki	0.775	0.487	0.011	0.008

To reveal the potential of the modeling of the average daily values of air pollutant concentration at separate monitoring stations using model (2), the sample of observational data selected for the analysis was divided into two samples, namely, the learning and test samples. The model coefficients are computed from the learning sample. The test sample is used for the comparison of simulated and observed values of concentration using the independent observational data which were not used for computing the coefficients.

The results of the testing of model (2) as the prognostic model for the NO_2 concentration at 29 AAPCS network stations (see the scheme in Fig. 1) are presented in the table. The ratio of the size of learning and test samples equal to 2:1 is often used for testing statistical models [23]. The table presents the values of the correlation coefficient and root-mean-square error computed from (3). Similar computations were carried out for the series of concentration of CO and PM_{10}. For example, the values of the correlation coefficient and root-mean-square error are the following:

Forecast	Model (2)			Inertial		
Pollutant	CO	NO_2	PM_{10}	CO	NO_2	PM_{10}
r	0.62	0.67	0.74	0.42	0.53	0.64
, mg/m^3	0.11	0.009	0.011	0.18	0.012	0.015

Thus, the use of model (2) increases the correlation coefficient by 2, 1.5, and 1.3 times for CO, NO_2, and PM_{10}, respectively, and reduces the variance of the forecast error by 2.5, 1.9, and 1.8 times for these pollutants as compared with the inertial forecast.

The analysis of the skill scores of the statistical model in simulating the concentration of three pollutants at all stations demonstrated that the worse the inertial model operation is, the more significant the relative improvement is in the forecast based on formula (2) as compared with the inertial forecast.

CONCLUSIONS

The use of the statistical model for forecasting the concentration of various pollutants in the air at the locations of Moscow monitoring stations revealed that in this case the coefficient of correlation between the computed and observed values of concentration increases as compared with the inertial forecast. In the future it is desirable to include into the list of predictors the height of the mixed layer [28] or the height of the surface and raised inversion (in addition to the wind speed).

The adjustment of the model based on relatively short samples was a peculiarity of the model testing. This makes possible its use at new monitoring stations or in the cases when changes in life conditions take place in some areas. The interannual variability of computed coefficients should be considered as a shortcoming of such approach. The size of learning samples can be increased if needed.

The given paper is the first attempt to verify the potential of the statistical prediction of air pollutant concentration on the whole territory of Moscow megalopolis. So, a number of essential problems were left beyond its scope. First of all, the values of meteorological parameters observed on the day under prediction are used instead of those predicted on the day of the forecast issue. Thereby the contribution of the error of meteorological forecast is not taken into account. At the same time, the improvement of meteorological forecasts with the lead time exceeding 24 hours should expand the potential of the statistical forecasting of the air pollutant concentration for the period of several days in advance.

The paper does not contain data on the spatial distribution of predicted concentrations. The authors hope that the spatial interpolation carried out in one or another way, levels the random errors caused by the mesoscale turbulence formed as a result of the urban development. In the future it is important to analyze the spatial structure of model coefficients as well. For example, the prevailing wind direction making the maximum contribution to the variability of pollutant concentration in the air over the megalopolis, can be judged by the coefficient a_2 in formula (2).

Another promising research area can be the analysis of the residual between the data of observations and forecasts if taking into account the hydrometeorological conditions causes significant discrepancy between them.

ACKNOWLEDGMENTS

The research was supported by the Moscow Department for Environmental Management and Protecion (state contracts 0604-07/12 and 0604-07/13) and Ministry of Education and Science of Russian Federation (the unique identifier of the project RFMEFI58614X0004).

REFERENCES

1. T. W. Anderson, *The Statistical Analysis of Time Series* (Mir, Moscow, 1976) [Transl. from English].
2. Y. Bard, *Nonlinear Parameter Estimation* (Statistika, Moscow, 1979) [Transl. from English].
3. M. E. Berlyand, *The Forecast and Control of Air Pollution* (Gidrometeoizdat, Leningrad, 1985) [in Russian].
4. A. I. Vereskov, "Mathematical Support of Nonstandard Problems of Regression Analysis," in *Multidimensional Statistical Analysis and Probabilistic Modeling of Real Processes* (Nauka, Moscow, 1990) [in Russian].
5. E. L. Genikhovich, V. A. Gushchin, and L. R. Son'kin, "On the Possibility of Air Pollution Forecasting by the Pattern Recognition Method," Trudy GGO, No. 293 (1973) [in Russian].
6. G. I. Gorchakov, E. G. Semutnikova, A. V. Karpov, et al., "Air Pollution Week-long Cycle in Moscow: Refinement of Quantitative Parameters and Statistical Forecasting of Impurity Concentration," Optika Atmosfery i Okeana, No. 9, **23** (2010) [Atmos. Oceanic Optics, No. 9, **23** (2010)].
7. A. M. Zvyagintsev and G. M. Kruchenitskii, "An Empirical Model of Surface Ozone Concentration near Moscow (The Town of Dolgoprudnyi)," Izv. Akad. Nauk, Fiz. Atmos. Okeana, No. 1, **32** (1996) [Izv., Atmos. Oceanic Phys., No. 1, **32** (1996)].
8. A. M. Zvyagintsev and I. N. Kuznetsova, "Surface Ozone Variations in Moscow Environs: The Results of Continuous Ten-year Observations," Izv. Akad. Nauk, Fiz. Atmos. Okeana, No. 4, **38** (2002) [Izv., Atmos. Oceanic Phys., No. 4, **38** (2002)].
9. *The Report on Environmental Conditions in Moscow in 2012*, Ed. by A. O. Kul'bachevskii (Spetskniga, Moscow, 2012) [in Russian].

10. *The Report on Environmental Conditions in Moscow in 2014*, Ed. by A. O. Kul'bachevskii (Spetskniga, Moscow, 2014) [in Russian].
11. A. P. Revokatova, G. V. Surkova, et al., "Air Pollution Forecast for the Moscow Region Based on the COSMO-ART Model," Vestnik Moskovskogo Universiteta, Ser. 5, Geografiya, No. 4 (2012) [in Russian].
12. L. R. Son'kin, *Synoptic and Statistical Analysis and Short-range Forecasting of Air Pollution* (Gidrometeoizdat, Leningrad, 1991) [in Russian].
13. E. Agirre-Basurko, G. Ibarra-Berastedi, and I. Madariaga, "Regression and Multilayer Perception-based Models to Forecast O_3 and NO_2 Levels in Bilbao Area," Environ. Mod. Soft., **21** (2006).
14. S. P. Arya, *Air Pollution Meteorology and Dispersion* (Oxford Univ. Press, Oxford, 1990).
15. A. Baklanov, O. Hanninen, L. H. Stordal, et al., "Integrated Systems for Forecasting Urban Meteorology, Air Pollution, and Population Exposure," Atmos. Chem. Phys., **7** (2007).
16. R. D. Bornstein and D. S. Johnson, "Urban–Rural Wind Velocity Differences," Atmos. Environ., No. 7, **11** (1977).
17. C. Carnevale, G. Finzi, E. Pisoni, and M. Votta, "Neuro-fuzzy and Neural Network Systems for Air Quality Control," Atmos. Environ., No. 31, **43** (2009).
18. P. F. Demchenko, A. S. Ginzburg, A. I. Vereskov, et al., "Forecasting Model for Moscow Agglomeration: Statistical Algorithm and Preliminary Tests," in the *Proceedings of Pollution and Environment–Treatment of Air (PETRA-2013), June 4–6, 2013, Prague*.
19. L. A. Dias-Robles, J. C. Ortega, J. S. Fu, et al., "Hybrid ARIMA and Artificial Networks Model to Forecast Particulate Matter in Urban Areas: The Case of Temuco, Chile," Atmos. Environ., **42** (2008).
20. H. J. S. Fernando, M. C. Mammarella, G. Grandoni, et al., "Forecasting PM_{10} in Metropolitan Areas: Efficacy of Neural Neworks," Environ. Pollut., **163** (2012).
21. M. W. Gardner and S. R. Dorling, "Artificial Neural Networks (The Multilayer Perception)—A Review of Applications in the Atmospheric Sciences," Atmos. Environ., No. 14/15, **32** (1998).
22. I. B. Konovalov, M. Beekmann, F. Meleux, et al., "Combining Deterministic and Statistical Approaches for PM_{10} Forecasting in Europe," Atmos. Environ., **43** (2009).
23. D. K. Papanastasiou, D. Melas, and I. Kioutsoukis, "Development and Assessment of Neural Network and Multiple Regression Models in Order to Predict PM_{10} Levels in Medium-sized Mediterranean City," Water, Air, Soil Pollut., **182** (2007).
24. P. Perez, "Combined Model for PM_{10} Forecasting in a Large City," Atmos. Environ., **60** (2012).
25. P. Perez and J. Reyes, "An Integrated Neural Network Model for PM_{10} Forecasting," Atmos. Environ., **40** (2008).
26. P. Perez and G. Salini, "$PM_{2.5}$ Forecasting in a Large City: Comparison of Three Methods," Atmos. Environ., **42** (2008).
27. J.-M. Poggi and B. Portier, "PM_{10} Forecasting Using Clusterwise Regression," Atmos. Environ., **45** (2011).
28. M. Righby and R. Toumi, "London Air Pollution Climatology: Indirect Evidence for Urban Boundary Layer Height and Wind Speed Enhancement," Atmos. Environ., **42** (2008).
29. A. Russo, F. Raischel, and G. P. Lind, "Air Quality Prediction Using Optimal Neural Networks with Stochastic Variables," Atmos. Environ., **79** (2013).
30. A. Russo and A. O. Soares, "Hybrid Model for Urban Air Pollution Forecasting: A Stochastic Spatio-temporal Approach," Math. Geosci., No. 1, **46** (2014).
31. V. Singh, C. Carnevale, G. Finzi, et al., "A Stochastic Modeling System to Forecast Ozone and PM_{10} over Milan Metropolitan Area," in *ACCENT-GLOREAM Workshop, November 26–27, 2009*.
32. V. Singh, C. Carnevale, G. Finzi, et al., "A Cokriging Based Approach to Reconstruct Air Pollution Maps, Measurement Station Concentration and Deterministic Model Simulations," Environ. Mod. Soft., **26** (2011).
33. T. Slini, K. Karatzas, and A. Papadopulos, "Regression Analysis and Urban Air Quality Forecasting: An Application for the City of Athens," Global Nest, No. 2/3, **4** (2002).
34. D. Voukantsis, K. Karatzas, J. Kukkonen, et al., "Intercomparison of Daily Quality Data Using Artificial Neural Networks in Thessaloniki and Helsinki," Sci. Total Environ., **409** (2011).