Indonesian Journal of Electrical Engineering and Computer Science
Vol. 22, No. 3, June 2021, pp. 1590~1600
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v22.i3.pp1590-1600

Information system for monitoring of urban air pollution by heavy metals

Zhanar Oralbekova[1], Zarına Khassenova[2], Bakhyt Mynbayeva[3], Makpal Zhartybayeva[4], Kazizat Iskakov[5]
[1,4,5]L. N. Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan
[2]Serikbayev Kazakhstan State Technical University, Ust-Kamenogorsk, Kazakhstan
[3]Abai Kazakh National Pedagogical University, Almaty, Kazakhstan
[5]National Research Nuclear University, Moscow Engineering Physics Institute (MEPhI), Moscow, Russia

Article Info

Article history:

Received May 23, 2020
Revised May 28, 2021
Accepted Jun 13, 2021

Keywords:

Air pollution
Data assimilation
Heavy metals
Information systems
Monitoring

ABSTRACT

The authors of the article formulated the information monitoring system of Almaty city (The Republic of Kazakhstan). One of the urgent problems of the modern world is the ecological safety of the urban environment. The well-known unfavorable natural and anthropogenic factors of the Almaty city include the following: geographical and climatic features of the area, the lack of technologies for monitoring atmospheric air and traffic jams. All of the above factors increase the emissions of pollutants into the atmosphere, in particular, toxic heavy metals. A promising direction for solving this problem is the development of real-time monitoring using a mathematical model of data assimilation and the creation of software that allows you to assess the state of the system in real time. In this regard, this article describes the developed information system for the algorithm of data assimilation that optimizes the process of monitoring atmospheric air pollution with heavy metals.

Corresponding Author:

Makpal Zhartybayeva
Department of Computerand Software Engineering
L. N. Gumilyov Eurasian National University
Pushkin St. 11, Nur-Sultan 010008, Kazakhstan
Email: makkenskii@mail.ru

1. INTRODUCTION

The actual problem of urban environment safety needs to be solved in the modern city of Kazakhstan Almaty. To the unfavorable natural and anthropogenic factors of the city (geographical and climatic features, the lack of modern technologies for monitoring atmospheric air, and the presence of a large number of cars), an increase in emissions of pollutants (EP) into the atmosphere, including toxic heavy metals (HM), was added. A specific direction of solving this urgent problem is the development of the existing system of monitoring of atmospheric air in the city of Almaty. A simple determination of the content of all pollutants in the air as objects of research is not informative for forecasts. Therefore, the creation of a new information system for monitoring atmospheric air will allow you to quickly and effectively respond to air pollution. The prerequisites for the development of a new effective information monitoring system are the preliminary experimental data of scientific studies on the pollution of the environment with heavy metals [1]. However, these studies were mainly concerned with determining the degree of atmospheric air pollution in Almaty with HM.

In [2]-[4], mathematical models were used to solve the problems of forecasting and estimating anthropogenic pollution in industrial regions of the Russian Federation (RF). Analysis of other research papers on mathematical modeling [5]-[13] showed that they mainly concerned system analysis and statistical models of limiting factors in ecology, and were outdated. The newer works of Russian scientists related to mathematical modeling on environmental problems, introduction to mathematical ecology, and others were of no practical value, since they are generalizing textbooks and monographs without proper experimental material [4]. Applied research of Russian scientists was associated with the use of methods of correlation and regression analysis in agrometeorology. For our research, the developed computer system "Ecoterra" was more suitable for developing solutions taking into account environmental factors [9], but it did not have meteorological factors for the diffusion of EP. The prerequisites for our research to a lesser extent were the work in the field of computer use [10], [11] and old methods of data analysis and processing [12], [13], unfortunately, not in the field of urban ecology.

As a mathematical support for the information system, we used solutions to data assimilation problems: algorithms for numerical modeling of pollutant propagation, which assess the state of the system in real time set-data assimilation problems [4], [5], [14], [15]. We needed to add the task to predict the value of the model state function in accordance with the obtained observation data, that is, to estimate the "real" state of the system using a mathematical model, a priori information and measurement data.

In other scientific studies, we have found that there are 2 main approaches to data assimilation: variational and dynamic-stochastic, based on the Kalman filter algorithm. The analysis of works on the variational principle [16], [17], the application of the Kalman algorithm in the problems of assimilation of meteorological and oceanic observations is given in [18]-[22] showed that the use of the apparatus of conjugate equations in application to various fields was effective in the problems of variational assimilation of data. Great contributions were made in own research [23]-[26]. Therefore, in our own research of the variational type of data assimilation with the use of the apparatus with conjugates, the topic of variational data assimilation is quite actively developed.

In the Republic of Kazakhstan, research in this direction is actively conducted in the industrial city of Ust-Kamenogorsk by scientists of D. Serikbayev East Kazakhstan Technical university [27]-[30]. In the Commonwealth of Independent States according to the direction involved: Institute of computational mathematics and mathematical Geophysics Siberian Branch of Russian Academy of Sciences (SB RAS) (Novosibirsk, RF), Institute of numerical mathematics of Russian Academy of Sciences (Moscow, RF), Institute of computational technologies SB RAS, M. V. Lomonosov Moscow State University (RF).

The purpose of these studies is: to formulate an information system of ecological monitoring of atmospheric contaminants, which can be used to assess the condition of the air in the atmospheric surface layer of the Almaty city, as well as make mathematically reasonable local prediction of air pollution subject to HM. Objectives of the study: i) To create a scientific basis for modeling mesometeorological atmospheric processes for the city on the basis of environmental monitoring of atmospheric pollution to create a information system; ii) To create the subsystem of software that uses mathematical tools to build a pollution model and to assess the state of the system in real time; iii) To develop an information system and database that optimizes the process of monitoring atmospheric air pollution of HM; and iv) To predict the concentration of HM.

2. THE PROPOSED ALGORITHM

Our proposed algorithm for developing a new information system for optimizing monitoring of air pollution with HM on the example of Almaty city is as follows: i) Areas have been selected for sampling air (in this case, for Almaty city); ii) Determined the content of HM in them is by a well-known method using a special device “Shimadzu” and the degree of air pollution; iii) Created a database on air pollution by HM and meteorological data of the city is being; iv) An application is created that formulates netCDF (network common data form) files. The netCDF files created at this stage are used as input parameters for the mathematical support of the developed information system for environmental monitoring of atmospheric pollution; and v) The mathematical support of the information system for environmental monitoring of atmospheric pollution is used, taking into account the ways of interfacing the developed models and methods of atmospheric chemistry with real meteorological data of the created database on HM. It should be noted that the created software tools in the development of an information system for environmental monitoring and analysis of the results of monitoring changes in the chemical composition of the atmosphere allow you to manage calculations on a server platform.

3. RESEARCH METHODS

In the course of implementing the main goal of the article, an information system and database were formed to optimize monitoring of atmospheric air pollution on the server of the National Scientific Laboratory for Collective Use of Information and Space Technologies in Almaty city (The Republic of

Kazakhstan). Figure 1 shows a diagram of the main blocks of the developed system. The developed information system provides: i) Collection, processing of initial data on environmental pollution by heavy metals (concentration of impurities) in the industrial city of Almaty; ii) Collection and processing of meteorological parameters (wind speed, wind direction, temperature, and atmospheric pressure); iii) Formation of netCDF files; iv) Mathematical support of the information system for environmental monitoring of atmospheric air with heavy metals: variational algorithms for sequential assimilation of data in real time; v) Determination of the degree of atmospheric air pollution, assessment of the air condition of the surface layer of the city's atmosphere; vi) Visualization of the results obtained; and viii) Output of data in the form of reports on the degree of air pollution of the city based on the results obtained.

To collect and process the initial data on environmental pollution of HM, as well as meteorological parameters, the following tasks were implemented: selected sites for air sampling and determined the content of HM in the samples of Almaty city; created a structure and replenishment of the database on air pollution of Almaty city HM; included in the database results on air pollution of Almaty city, which are presented in the works [31], [32]. For the purpose of collecting and processing data to optimize monitoring of air pollution, the eco.mdf database was created, which is included in the state register of rights to objects protected by copyright of the Republic of Kazakhstan [33]. The relational model of the created database is shown in Figure 2. The developed environmental monitoring information system is available at http://5.63.119.49:8080 (Figure 3).

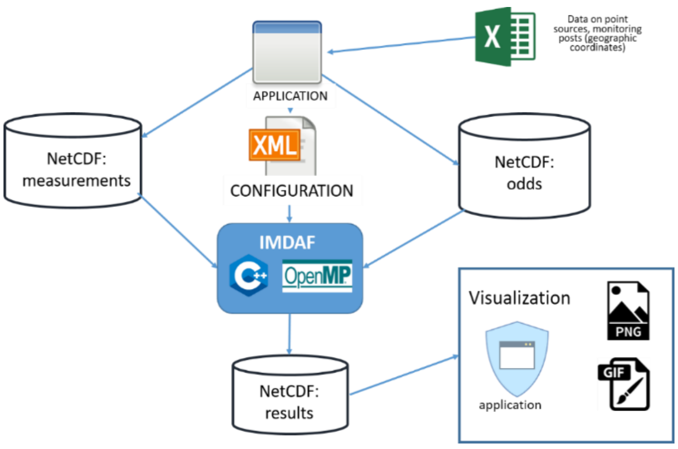


Figure 1. Diagram of the main blocks of the developed system

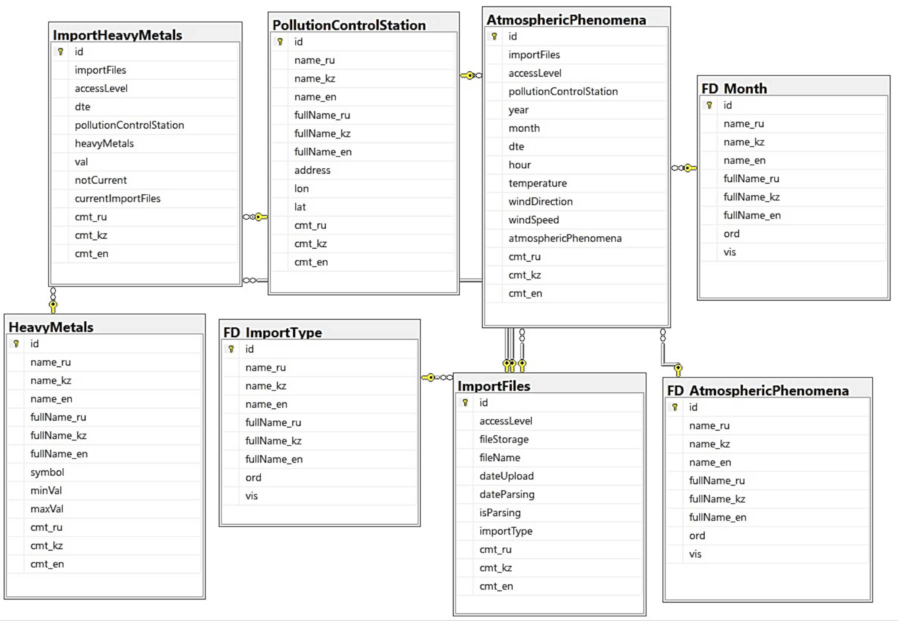
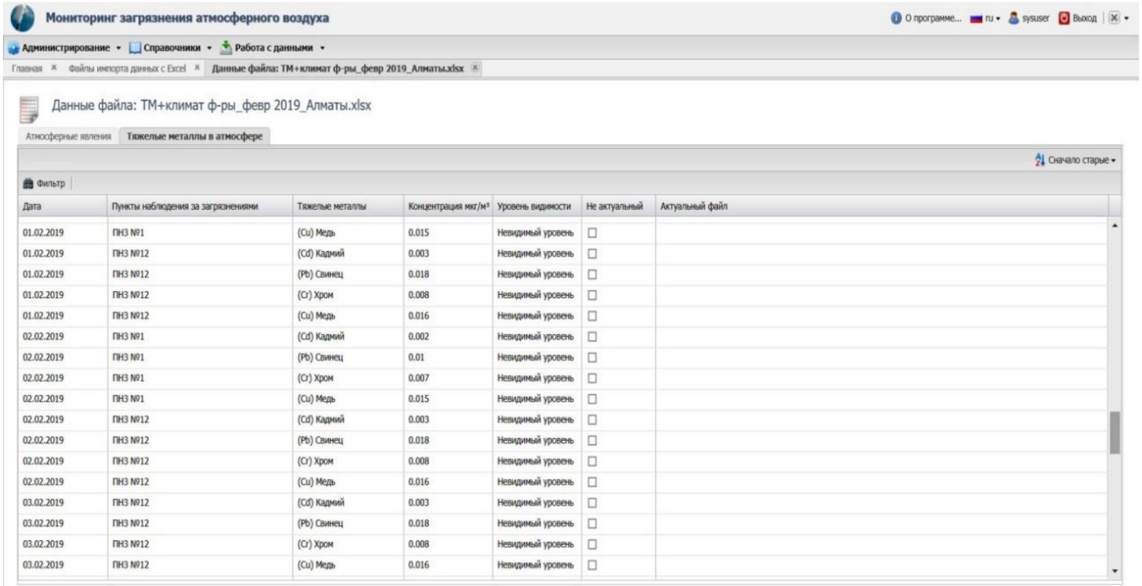


Figure 2. Relational model of the created database



Дата	Пункты наблюдения за загрязнениями	Тяжелые металлы	Концентрация мкг/м³	Уровень видимости	Не актуальный	Актуальный файл
01.02.2019	ПНЗ №1	(Cu) Медь	0.015	Невидимый уровень	☐	
01.02.2019	ПНЗ №12	(Cd) Кадмий	0.003	Невидимый уровень	☐	
01.02.2019	ПНЗ №12	(Pb) Свинец	0.018	Невидимый уровень	☐	
01.02.2019	ПНЗ №12	(Cr) Хром	0.008	Невидимый уровень	☐	
01.02.2019	ПНЗ №12	(Cu) Медь	0.016	Невидимый уровень	☐	
02.02.2019	ПНЗ №1	(Cd) Кадмий	0.002	Невидимый уровень	☐	
02.02.2019	ПНЗ №1	(Pb) Свинец	0.01	Невидимый уровень	☐	
02.02.2019	ПНЗ №1	(Cr) Хром	0.007	Невидимый уровень	☐	
02.02.2019	ПНЗ №1	(Cu) Медь	0.015	Невидимый уровень	☐	
02.02.2019	ПНЗ №12	(Cd) Кадмий	0.003	Невидимый уровень	☐	
02.02.2019	ПНЗ №12	(Pb) Свинец	0.018	Невидимый уровень	☐	
02.02.2019	ПНЗ №12	(Cr) Хром	0.008	Невидимый уровень	☐	
02.02.2019	ПНЗ №12	(Cu) Медь	0.016	Невидимый уровень	☐	
03.02.2019	ПНЗ №12	(Cd) Кадмий	0.003	Невидимый уровень	☐	
03.02.2019	ПНЗ №12	(Pb) Свинец	0.018	Невидимый уровень	☐	
03.02.2019	ПНЗ №12	(Cr) Хром	0.008	Невидимый уровень	☐	
03.02.2019	ПНЗ №12	(Cu) Медь	0.016	Невидимый уровень	☐	

Figure 3. Ecological monitoring system, data on heavy metals in atmospheric air for 2019

In this system of ecological monitoring for heavy metals, data processing is carried out, accordingly, at the stage of mathematical modeling. Since modeling is a powerful tool for studying the processes of atmospheric air pollution. Therefore, mathematical support of the information system for environmental monitoring of atmospheric pollution is based on the model of transport of impurities in the atmosphere [34]:

$$\frac{\partial \phi(\vec{x},t)}{\partial t} + \mathrm{div}(\phi(\vec{x},t)\vec{u}(\vec{x},t) - \mu(\vec{x},t)\mathrm{grad}\phi(\vec{x},t)) = f(\vec{x},t) + r(\vec{x},t), (\vec{x},t) \in D \times (0,T), \quad (1)$$

$$\mu(\vec{x},t)\frac{\partial \phi(\vec{x},t)}{\partial \vec{n}} = 0, (\vec{x},t) \in \Gamma_{out}, \phi(\vec{x},t) = \phi_b(\vec{x},t), (\vec{x},t) \in \Gamma_{in}, \quad (2)$$

$$\phi(\vec{x},t) = \phi_0(\vec{x}), \vec{x} \in D, t = 0 \quad (3)$$

Here, ϕ– impurity concentration [kg/m^3], $\vec{u}(u,v,w)$– wind velocity vector [m/s], μ –diffusion coefficient [m^2/s],ϕ_b– background concentration values [kg/m3], ϕ_0– initial concentration distributions [kg/m^3]. The function f characterizes the power of the impurity source. The variable r corresponds to the uncertainty function added to the model for assimilating the data. $D = (0,X) \times (0,Y) \times (0,Z)$, $X,Y,Z > 0$is the area in space,Γ_{out}– part of the border of $\partial D \times (0,T)$ sector, where the wind velocity vector directed outward from the sector,Γ_{in} – a part of the boundary where the vector speed of the wind directed into the sector. The obtained experiments and results using the model of the transport impurities in the atmosphere (1)-(3) are reflected in [34]. As part of further research, the ARIMA model, which is popular and widely used for forecasting time series, was applied to determine the level of concentration of heavy metals in the atmospheric air.

4. RESULTS AND DISCUSSION

For the formation of netCDF format files, the following tasks were implemented: was developed the application format for the container of the basic elements of the information modeling system; the organization of the database and access to it from problem-oriented applications, which are presented in the work [34]. The resulting netCDF files are used as source data for the implementation of the impurity transport model in the atmosphere. There are two ways to replenish the created database: manually and by importing it from external sources, such as MS Excel.

For the mathematical software of the information system were implemented the following tasks: carried out the implementation and research of algorithms of data assimilation for monitoring the chemical composition of the atmosphere of Almaty city; studied variational algorithms for sequential data assimilation in real-time; developed ways of interface of the developed models and methods atmospheric chemistry with real meteorological data created a database on HM [34].

To visualize the obtained results, were implemented the following tasks: software tools for a new information system for environmental monitoring and analysis of the results of monitoring changes in the chemical composition of the atmosphere were developed; a human-machine interface for software tools for visualizing the results of monitoring was developed [34]. To output data in the form of reports on the degree of atmospheric air pollution in the city, based on the results obtained, the following tasks were implemented: analysis of selected air samples, determination of the degree of atmospheric air pollution in Almaty city; development of software tools for a new information system for environmental monitoring and analysis of the results of monitoring changes in the chemical composition of the atmosphere; human-machine interface for software tools, assessment of the ecological state of the air of the surface layer of the atmosphere for Almaty city. The application of mathematical methods for processing environmental information allowed us to obtain scientific foundations for modeling mesometeorological atmospheric processes .

During the implementation of the work, the system inverse modeling and data assimilation framework (IMDAF) was used [30]. The IMDAF system solves a hierarchy of increasingly complex mathematical modeling problems: direct problems, sensitivity assessment problems, inverse problems, and data assimilation problems. Through the use of splitting schemes and an ensemble approach, the algorithms that make up IMDAF are effectively mapped to parallel computing architectures. To enable efficient development, IMDAF is developed in an object-oriented paradigm. Specific examples of the system's operation are given for advection-diffusion-reaction models, which are used in a wide range of practical applications, including atmospheric chemistry and developmental biology. During the research, data forecasting was implemented using the ARIMA model in the R environment. About 27,000 data were processed, which were collected from observation points from 2006 to 2020 (1205 fees from one OPP, 1844 fees from the second OPP).

For further data forecasting, we first build a graph (date, heavy metal concentration). Data on atmospheric concentrations of the following heavy metals were collected from observation points: Cd, Pb, As, Cr, Cu (Figure 4). The analysis revealed that heavy metals Pb, Cr, and Cu have the least significant concentration in the air. We presented a method for processing information about the dynamics of stationary time series on the example of analyzing data on heavy metals in the atmospheric air of Almaty city (Figure 5) using an autoregression model and an integrated moving average. Methods for testing and verifying the adequacy of the model are given, and methods for predicting future values of time series levels based on the model are described.

Graphs 1, 2 show that the series are non-stationary. Such a graph is typical for a random walk, i.e. a process of approximately the form $y_t = y_{t-1} + u_t$. The integrated moving average autoregression model (ARIMA model) can fairly well describe the behavior of non-stationary time series, including those containing a seasonal component. ARIMA models are largely suitable for predicting the time series that characterize our data, since these series are stochastic due to the specifics of the influencing factors [33]. For a non-stationary series, it does not make sense to calculate ACF and PACF, since the correlation between random variables separated by k steps can vary over time (Figure 6). Next, we analyze the Pb data and graph of the approximation function for making forecasts based on statistical data (Figure 7).

	date	Cd	Pb	As	Cr	Cu	Ni
1196	2020-05-23	0.001	0.033	0	0.006	0.023	0.000
1197	2020-06-01	0.001	0.016	0	0.006	0.035	0.000
1198	2020-06-02	0.001	0.016	0	0.006	0.035	0.000
1199	2020-06-03	0.001	0.016	0	0.006	0.035	0.000
1200	2020-06-11	0.000	0.022	0	0.008	0.036	0.000
1201	2020-06-12	0.000	0.022	0	0.008	0.036	0.000
1202	2020-06-13	0.000	0.022	0	0.008	0.036	0.000
1203	2020-06-20	0.000	0.015	0	0.005	0.031	0.000
1204	2020-06-21	0.000	0.015	0	0.005	0.031	0.000
1205	2020-06-22	0.000	0.015	0	0.005	0.031	0.000

Figure 4. Data on atmospheric concentrations for heavy metals: Cd, Pb, As, Cr, Cu

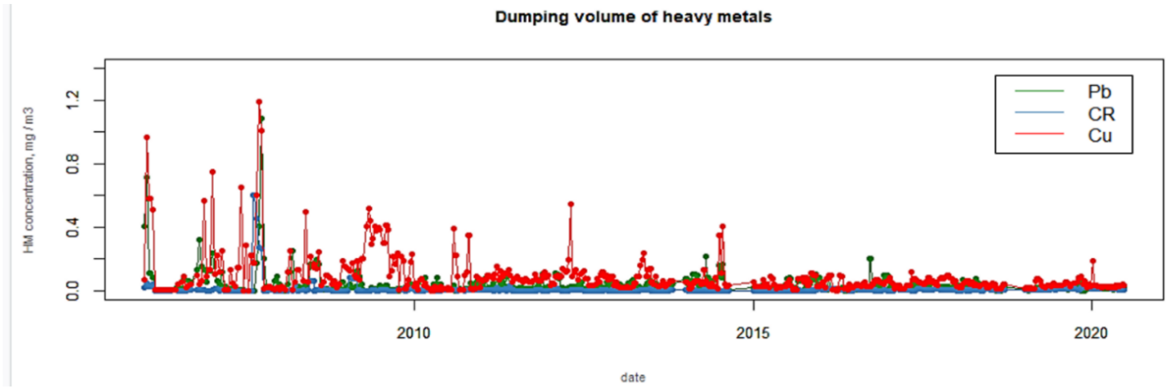
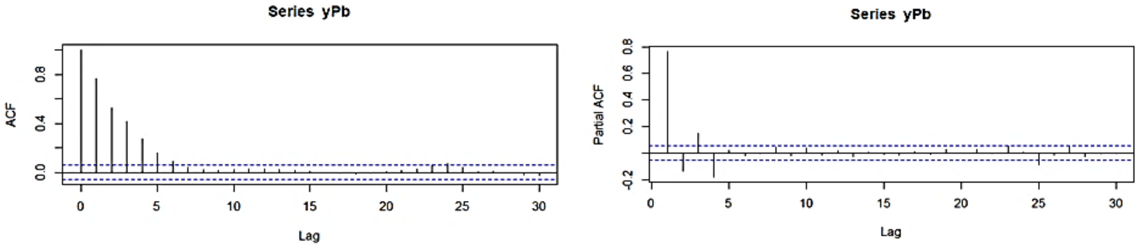


Figure 5. Graph of heavy metal concentrations (PB-"forest green", CR-"steel blue", CU-"red")



(a) (b)

Figure 6. Graphical representation of Pb concentration calculations: (a) ACF and (b) PACF

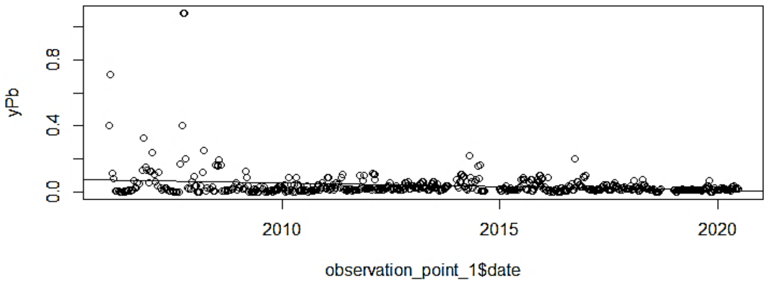


Figure 7. Graph of the approximating function (Pb)

Trend shows that Pb decreases year by year =>time series of Pb is non-stationary. The average value in each month is not much different = > the seasonal effect is weak. Schedule for Z it already looks like a stationary one, Z it turns around the middle in a band of approximately constant width. ARIMA applications for predicting heavy metal concentrations. ARIMA has the following parameters: i) AR: p; ii) I: d (the amount of " difference" to bring the initial non-stationary series to a stationary one); and iii) MA: q.

We evaluate the model arima(1, 1, 0), arima(3,0,3), arima(1,1,2) using the arima function in R. To evaluate the first model, you need to remove the main arguments using the print function. The obtained arguments showed that the estimated ARMA (1,1) or (ARIMA (1,1,0)) equation has the form:

$$\begin{cases} z_t = y_t \\ z_t = 0{,}0001 \cdot z_{t-1} \end{cases}$$

Next, we will do the same with the second and third models arima (3, 0, 3), arima (1, 1, 2). Then the equation of the second model arima (3, 0, 3) has the form:

$$\begin{cases} z_t = y_t - 0{,}0364 \\ z_t = -0{,}42 \cdot z_{t-1} + 0{,}49 \cdot z_{t-2} + 0{,}34 \cdot z_{t-3} + \varepsilon_t + 1{,}36\varepsilon_{t-1} + 0{,}44\varepsilon_{t-2} - 0{,}07\varepsilon_{t-3} \end{cases}$$

equation of the 3rd arima model(1,1,2):

$$\begin{cases} z_t = y_t \\ z_t = 0{,}6 \cdot z_{t-1} + \varepsilon_t - 0{,}64\varepsilon_{t-1} - 0{,}34\varepsilon_{t-2} \end{cases}$$

Let's choose the best model based on the punitive criterion of the AIC:

$$AIC = -2\, ln\, l + 2k,$$

where $\ln l$ - logarithm of the probability function, k - number of model parameters.

The more parameters k , the more complex the model, the higher the AIC. The lower the probability function, (the lower the probability of obtaining the available data for a given model) the higher the AIC. As a result, `AIC (model1) = -4000.906, AIC (model2) = -4207.222, AIC (model3) = -4173.849`. According to the AIC criterion, the best model is the second model. You can also use BIC $BIC = 2 \ln L + \ln n \cdot k$.

According to the BIC criterion, the second model was also the comparative best. For ARMA (p, q), you can see visually where the roots of the AR and MA parts lie as shown in Figure 8 (`autoplot`(model2)). Based on the selected best model, you can build a forecast graph with predicative intervals and a forecast (Figure 9). According to the described principle, we will do the same for Cr and Cu data (Figures 10 and 11). The predicted data on the concentration of heavy metals Pb, Cr, Cu in atmospheric air obtained using the model are shown in Table 1.

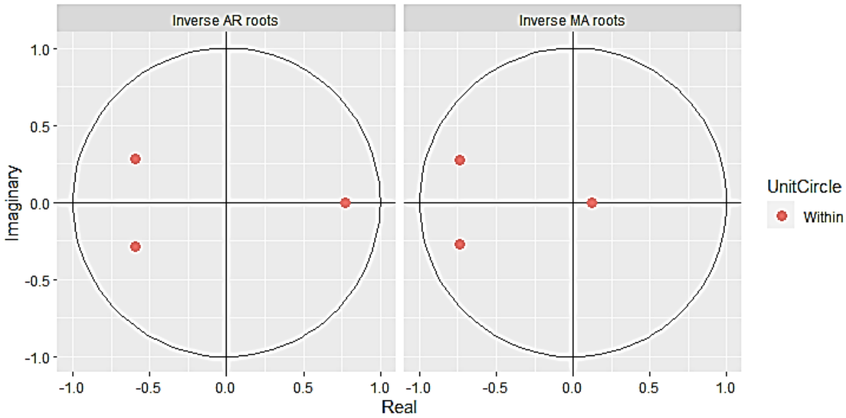


Figure 8. Visual representation of the location of AR, MA roots

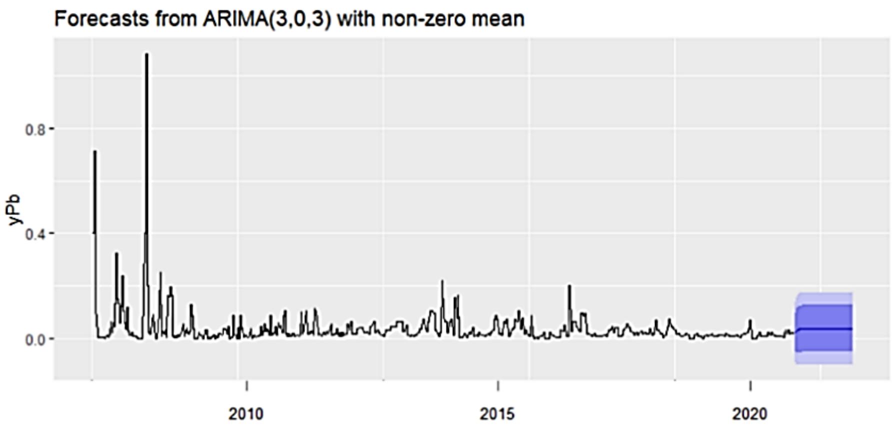


Figure 9. Graph of forecasts for the ARIMA model (3,0,3) with productive intervals according to Pb data

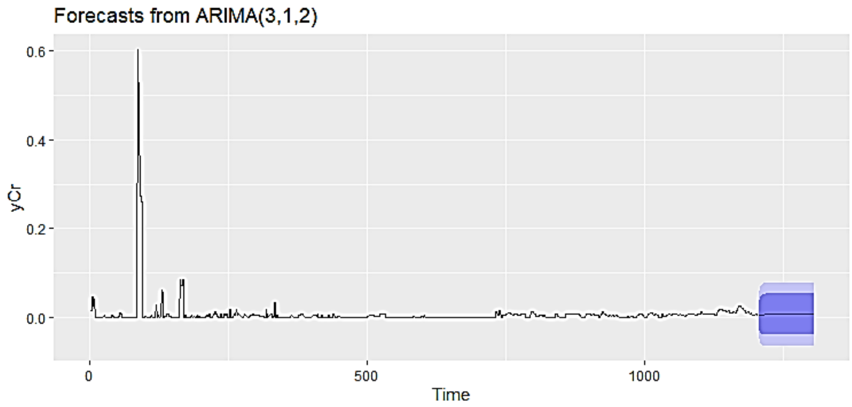


Figure 10. Prediction ARIMA (3,1,2) based on data (Cr)

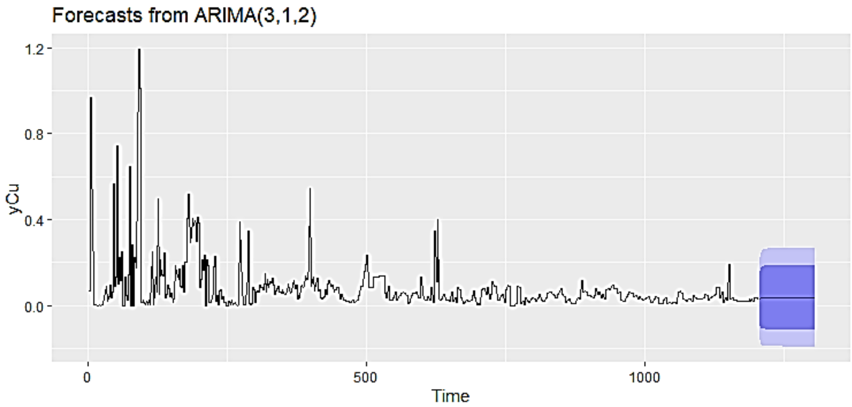


Figure 11. Prediction ARIMA (3,1,2) based on data (Cu)

Table 1. Predicted data on the concentration of heavy metals Pb, Cr, Cu in atmospheric air

predictdata	PbARIMA (3,0,3)	CrARIMA (3,1,2)	CuARIMA (3,1,2)
11.07.2020	0.03540145	0.004942237	0.03248529
12.07.2020	0.03561267	0.005537033	0.03396680
13.07.2020	0.03579848	0.005616783	0.03482298
20.07.2020	0.03592947	0.005909523	0.03543398
21.07.2020	0.03603708	0.006029903	0.03595130
22.07.2020	0.03611905	0.006194036	0.03635135
………	………	………	………
14.02.2021	0.03640180	0.006844717	0.03761635
21.02.2021	0.03640180	0.006844798	0.03761637
22.02.2021	0.03640181	0.006844865	0.03761637
23.02.2021	0.03640181	0.006844920	0.03761638
01.03.2021	0.03640181	0.006844966	0.03761639
02.03.2021	0.03640181	0.006845004	0.03761639
03.03.2021	0.03640181	0.006845035	0.03761639
04.03.2021	0.03640182	0.006845061	0.03761639
05.03.2021	0.03640182	0.006845082	0.03761640
12.03.2021	0.03640182	0.006845100	0.03761640
19.03.2021	0.03640182	0.006845114	0.03761640

Thus, the structure for modeling and making forecasts from time series data on heavy metals were discussed, where, first of all, the series were used itself.The purpose of the framework is to differentiate short-term and long-term dynamics in the series to improve the accuracy and reliability of forecasts. As a result of the program, a model was implemented that can predict the level of concentration of heavy metals in the atmospheric air with some error. Forecasts showed that the level of concentration of heavy metals in the atmospheric air continues to gradually increase.

5. CONCLUSIONS

As a result of the work, an information system was createdto optimize the monitoring of air pollution with heavy metals from the standpoint of speed, availability, efficiency and precision; the estimation of the state of the atmosphere of the city of modern mathematical methods and is composed of local prediction on the change in the atmospheric surface layer, developed by the application format of netCDF container of the basic elements of information-modeling system organized database and access it from the problem-oriented application, supplemented by a database on air pollution Almaty with HM; a human-machine interface for software tools has been developed. Our research made it possible to implement a variational algorithm for sequential assimilation of data in real-time monitoring of air pollution, taking into account the geographical and climatic characteristics of Almaty and its pollution by transport means.

The use of mathematical methods for processing environmental information allowed us to obtain scientific foundations for modeling mesometeorological atmospheric processes, and through the use of splitting schemes and an ensemble approach, the algorithms that make up IMDAF are effectively mapped to parallel computing architectures.The use of stochastic seasonal ARIMA models is a fairly effective method for short - and medium-term forecasting of time series with a periodicity that characterizes the change in the indicators of heavy metals in the air. Application of autoregression models of the integrated moving average with seasonality allows us to accurately approximate a wide class of random processes, the non-stationarity of which is due to the presence of a periodic component in the time series. The main downside of models of the ARIMA class is that when adding new information to the original series, it is necessary to correct the model. At the same time, the values of the AR and MA orders of processes can change significantly, which will lead to the construction of a completely different model. Forecasts showed that the level of concentration of heavy metals in the atmospheric air continues to gradually increase.

To solve one of the problems of environmental safety of the urban environment, in particular, air pollution with toxic heavy metals, we have proposed an information system for monitoring in real time using a mathematical model of data assimilation. Software has been created that uses mathematical tools to build a pollution model and to assess the state of the system in real time; the state of the city's atmosphere was assessed and a local forecast was made for changes in the surface layer of the atmosphere. The format of applications for the container of basic elements of the information modeling system is developed, the database and access to it from the problem-oriented application are organized, the database on air pollution of Almaty city is replenished.

ACKNOWLEDGEMENTS

The authors express their gratitude to Alexey Penenko for his attention to the work and for valuable advice in the software implementation of the data assimilation algorithm. This work was supported by a grant from Ministry of Education and Science of the Republic of Kazakhstan within the framework of the project AP05135992.

REFERENCES

[1] Report of the Abai KazNPU project, “Mathematical methods for assessing the quality of the environment in Almaty when it is contaminated with heavy metals,” agreement no. 4, April 1, 2010.
[2] P. N. Antokhin, A. Gochakov, A. B. Kolker, and A. Penenko, “Comparison of WRF-CHEM chemical transport model calculations with aircraft measurements in Norilsk,” *Atmospheric and Oceanic Optics*, vol. 31, no. 4, pp. 372-380, 2018, doi: 10.1134/S1024856018040024.
[3] V. Penenko, A. Penenko, E. A. Tsvetova, and A. Gochakov, “Methods for Studying the Sensitivity of Air Quality Models and Inverse Problems of Geophysical Hydrothermodynamics,” *Journal of Applied Mechanics and Technical Physics*, vol. 60, no. 2, pp. 392-399, 2019, doi: 10.1134/S0021894419020202.
[4] V. Penenko and A. E. Aloyan, “Models and methods for environmental protection,” Novosibirsk: Nauka, pp. 1-250, 1985.
[5] J. Jeffers, “Introduction to system analysis: application in ecology,” Moscow: Mir, pp. 1-256, 1981.
[6] G. I. Marchuk, “Mathematical Models in Environmental Problems*,”* Elsevier Science, pp. 1-220, 1986.
[7] L. A. Petrosyan and V. V., “Zakharov, Introduction to mathematical ecology,” Moscow, pp. 1-316, 1986.
[8] E. S. Ulanova and V. N., Zabelin, “Methods of correlation and regression analysis in agrometeorology,” *L. Gidrometeoizdat*, pp. 1-206, 1990.
[9] V. P. Gryaznov and Grishin N. N., “Development of the computer system “Ecoterra” for taking into account the ecological factor in the development of solutions fundamentals of urban optimization and recreational environment,” *Proc. Dokl. int. slave. the meeting. Togliatti,* 1992 May 30-June 2, 1992, pp. 33-36.
[10] Y. N. Tyurin and A. A. Makarov, “Data Analysis on a computer,” Moscow: INFRA-M, Finance and statistics, pp. 1-38, 1995.
[11] I. Gaidyshev, “Data analysis and processing,” *St. Petersburg: Piter*, pp. 1-752, 2001.

[12] J. Opsomer, P. Speckman, M. Kaiser and J. Jones, "Statistical models for limiting factors in ecology," *J. Amer. Statist. Assoc.*, v. 89, pp. 410-423, 1994, doi: 10.2307/2290841.

[13] A. A. Bykov and N. V. Murzin, "Problems of analysis of human security, society and nature," *St. Petersburg: Nauka,* pp. 1-247, 1997.

[14] A. V. Penenko, Z. S. Mukatova, V. V. Penenko, A. V. Gochakov, and P. N. Antokhin, "Numerical investigation of the direct variational algorithm of data assimilation in the urban scenario," *Atmospheric and Oceanic Optics*, vol. 31, no. 6, pp. 678-684., 2018, doi: 10.1117/12.2504639.

[15] V. V. Penenko, A. V. Penenko, and E. A. Penenko, "Variational approach to the study of processes of geophysical hydro-thermodynamics with assimilation of observation data," *Journal of Applied Mechanics and Technical Physics*, vol. 58, no. 5, pp. 771-778, 2017.

[16] I. Sasaki, "An objective analysis based on variational method," *J. Met. Soc. Japan,* vol. 36, no 3, pp. 29-30, 1958, doi: 10.2151/jmsj1923.36.3_77.

[17] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. AME. J. Basic Eng.*, vol. 82, no. 1, pp. 34-35, 1960, doi: 10.1115/1.3662552.

[18] E. G. Klimova, "An Efficient Algorithm for Stochastic Ensemble Smoothing," *Numerical Analysis and Applications,* vol. 13, no. 4, pp. 321-331, 2020.

[19] M. E. Gharamti, K. Raeder, J. Anderson, and X. Wang, "Comparing adaptive prior and posterior inflation for ensemble filters using an atmospheric general circulation model," *Monthly Weather Review*, vol. 147, no. 7, pp. 2535-2553, 2019, doi: 10.1175/MWR-D-18-0389.1.

[20] R. Ménard, J. F. Cossette, and M. Deshaies-Jacques, "On the complementary role of data assimilation and machine learning: An example derived from air quality analysis," Lecture *Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),*12142 LNCS, pp. 212–224, 2020, doi: 10.1007/978-3-030-50433-5_17.

[21] G. Scheffler, J. Ruiz, and Pulido M., "Inference of stochastic parametrizations for model error treatment using nested ensemble Kalman filters," *Quarterly Journal of the Royal Meteorological Society*, vol. 145, no. 722, pp. 2028-2045, 2020, doi: 10.1002/qj.3542.

[22] E. Kalnay, "Atmospheric Modeling, Data Assimilation and Predictability," *Cambridge University Press: Cambridge,* UK, 2003.

[23] A. C. Lorenc, "Data Assimilation and Predictability: Data Assimilation," *Encyclopedia of Atmospheric Sciences: Second Edition,* pp. 237–240, 2020.

[24] A. C. Lorenc and M. Jardak, "A comparison of hybrid variational data assimilation methods for global NWP," *Quarterly Journal of the Royal Meteorological Society*, vol. 144, no. 717, pp. 2748-2760, 2020, doi: 10.1002/qj.3401.

[25] M. Jardak and O. Talagrand, "Ensemble variational assimilation as a probabilistic estimator - Part 1: The linear and weak non-linear case," *Nonlinear Processes in Geophysics*, vol. 25, no. 3, pp. 565-587, 2018, doi: 10.5194/npg-25-565-2018.

[26] M. Jardak and O. Talagrand, "Ensemble variational assimilation as a probabilistic estimator - Part 2: The fully non-linear case," *Nonlinear Processes in Geophysics*, vol. 25, no. 3, pp. 589-604, 2018, doi: 10.5194/npg-2018-6.

[27] A. L. Karchevsky, Y. M. Turganbayev, S. J. Rakhmetullina, and Zh. T. Beldeubayeva, "Numerical solution of an inverse problem of determining the parameters of a source of groundwater pollution," *Eurasian Journal of Mathematical and Computer Applications,* vol. 5, no. 1, pp. 53-73, 2017, doi: 10.32523/2306-3172-2017-5-1-53-73.

[28] W. Wojcik *et al*, "Probabilistic and statistical modelling of the harmful transport impurities in the atmosphere from motor vehicles," *Rocznik Ochrona Srodowiska,* vol. 19, pp. 795-808, 2017.

[29] Z. Beldeubayeva, S. Rakhmetullina, E. Turganbayev, and V. Krivykh, "Information system of groundwater monitoring," *International Multidisciplinary Scientific Geo Conference Surveying Geology and Mining Ecology Management, SGEM,* vol.1, no. 2, pp. 139-146, 2015.

[30] N. T. Danaev, A. N. Temirbekov, and E. A. Malgazhdarov, "Modeling of pollutants in the atmosphere based on photochemical reactions," *Eurasian Chemico-Technological Journal*, vol. 16, no. 1, pp. 61-71, 2014, doi: 10.18321/ectj170.

[31] B. N. Mynbayeva, Zh. O. Oralbekova, D. N. Isabayeva, K. T. Iskakov, and Khassenova Z. T., "An employment of mathematical toolkit for ecological information processing during heavy metals pollution monitoring in particular reference to Almaty city's atmosphere," *In: Conference proceeding: 18th International Multidisciplinary Scientific Geoconference SGEM, 30 June-9 July, 2018, Bulgaria, Varna, vol. 18: Ecology, Economics, Education and Legislation. Issue: 5.1. Ecology and Environmental Protection*, 2018, pp. 27-34, doi: 10.5593/sgem2017/52/S20.055.

[32] Z. O. Oralbekova, Z. T. Khassenova, and M. G. Zhartybayeva, "Collection and processing of data to optimize the monitoring of atmospheric air pollution," *Communications in Computer and Information Science,* 2019, pp. 161-170, doi: 10.1007/978-3-030-12203-4_16.

[33] Z. T. Khassenova, K. T. Iskakov, Zh. O. Oralbekova, I. I. Furaeva, and S. Zh. Rakhmetullina, "Copyright on the type of object "Databases", the name of the object "Monitoring of the chemical composition of the atmosphere," No.7458 of 9 January 2020.

[34] A. V. Penenko, Z. T. Khassenova, V. V. Penenko, and E. A. Pyanova, "Numerical study of a direct variational data assimilation algorithm in Almaty city conditions," *Eurasian J. Math. Comput. Appl.*, vol. 7, pp. 53–64, 2019, doi: 10.32523/2306-6172-2019-7-1-53-6.

BIOGRAPHIES OF AUTHORS

Zhanar Oralbekova PhD, assistant professor, L. N. Gumilyov Eurasian National University, Department of Computer and Software Engineering, Nur-Sultan city, Kazakhstan. Address: 11 Pushkin st., r. 210, Nur-Sultan city, Kazakhstan. Research interests: data assimilation algorithms, inverse problems, programming, georadar. Email: oralbekova@bk.ru

Zarina Khassenova, master of Technical Sciences, head of Project management office D. Serikbayev East Kazakhstan technical university, Ust-Kamenogorsk city, Kazakhstan. Address: 69 Protozanov st., r. 210, Ust-Kamenogorsk city, Kazakhstan. Research interests: data assimilation algorithms, programming. Email: zthasenova@mail.ru

Bakhyt Mynbayeva, Doctor of Biological Sciences, Professor. Abai Kazakh National Pedagogical University, Almaty city, Kazakhstan. Address: 66 Kabanbai batyr st., apt. 3, Almaty city, Kazakhstan. Research Interests: monitoring heavy metal pollution of environmental components: air, surface water and soil. Email: bmynbayeva@gmail.com

Makpal Zhartybayeva PhD, Senior Lecturer, L.N. Gumilyov Eurasian National University, Department of Computer and Software Engineering, Nur-Sultan city, Kazakhstan. Address: 11 Pushkin st., r. 210, Nur-Sultan city, Kazakhstan. Research interests: data assimilation algorithms, programming, robotics, georadar. Email: makkenskii@mail.ru

Kazizat Iskakov Doctor of Physical and Mathematical Sciences, Professor L. N. Gumilyov Eurasian National University, Department of Computer and Software Engineering, Nur-Sultan city, Kazakhstan; National Research Nuclear University, MEPhI (Moscow Engineering Physics Institute), Moscow city, Russia. Address: 11 Pushkin st., r. 210, Nur-Sultan city, Kazakhstan. Research interests: data assimilation algorithms, programming, inverse problems, georadar. Email: kazizat@mail.ru