ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

UDC 502.3:613.15:911:004(477)

doi: 10.32620/reks.2023.3.18

Olga BUTENKO, Anna TOPCHIY

National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine

ANALYSIS OF POLLUTANTS IN AIR WITHIN THE TERRITORY OF UKRAINE USING GEOSTATISTICAL METHODS

*Air quality has recently been of great concern, as it directly affects people's lives. Continuous monitoring of atmospheric air quality and forecasting the dynamics of its changes are essential steps in assessing its current state and determining the concentration of pollutants. Therefore, the development of an effective system for assessing and forecasting the quality of atmospheric air has become one of the most important tasks. The **subject matter** of this article is geostatistical methods for air quality analysis. The **goal** is to analyze pollutants in the air over Ukraine's territory from 1990 to 2021. The dataset on air pollutants was provided by the State Statistics Service of Ukraine in the form of aggregated tables, which were initially processed for subsequent modelling. Cartographic modelling of pollutants was performed using geostatistical **methods**. As a **result**, this study presents 13 cartographic models showing the spatial distribution of air pollutants for different regions of Ukraine. However, because of the lack of official information on the presence of military actions, the results of geostatistical methods cannot be interpreted in the context of the military situation in the eastern part of the country. Information about military actions can be gathered from various sources, but this would require a considerable time and effort to structure and systematize the dataset. **Conclusions**. The method considered in this study cannot simultaneously consider multiple parameters, such as the value of pollutant indicators and the presence of military actions. Additional methods, such as graph theory and regression analysis, are employed to obtain quantitative assessments of the modelling results considering all factors influencing the environmental condition. The chosen method is a straightforward tool for solving environmental problems. Thanks to available GIS systems like ArcGIS Pro, visualization of the applied geostatistical and mathematical methods is possible. The cartographic models presented in this study cover the entire territory of Ukraine and have administrative boundaries depending on the location of the pollutant collection station.*

Keywords: *air quality; data; modelling; GIS; Air Quality Index; kriging geostatistical models.*

Introduction

As is well-known, air is one of the primary resources essential for supporting life [1, 2]. Air pollution is a serious environmental problem worldwide [1, 3], and it has significant impacts on the quality of the atmosphere [1, 4] and human health [1, 5]. Rapid industrialization and urbanization have led to increased air pollution levels, with industrial emissions being the main cause of severe air pollution in industrial areas [1, 6].

Air pollution and its associated issues attract increasing public attention daily because the quality of the atmospheric environment affects every member of society. Consequently, there is a growing demand for the development of strategies to manage the environmental condition and combat pollution. The Donbas region is the most technogenically burdened area in Ukraine, and the ongoing military actions have significantly exacerbated the ecological situation in the eastern part of the country. Due to active military operations, environmental monitoring is not conducted, potentially leading to a catastrophic situation in the region. Today, there are many environmental monitoring systems [7, 8], however, to solve the problem, it is relevant to build a comprehensive system for monitoring the environmental situation [9, 10], considering the features [9], specifics and the use of multi-purpose strategies in the study areas. Today there are standardized provisions [10], scientific schools [11] that explore the issues of complex indicators of the quality of the functioning of information monitoring systems [12]. To effectively assess and forecast air quality, data from Earth remote sensing [13], ecological statistics, and mathematical statistics and modelling are appropriate. The widespread use of Earth remote sensing data is primarily due to the efficiency and visibility of processing data obtained from large areas. Modern satellites of remote sensing data (Terra, Aqua, Landsat 8, Sentinel-2) make it possible to obtain space images of the studied territories with a frequency of 1 to 8 days and with a spatial resolution of 250 to 10 meters [14]. Also, modern methods of data processing make it possible to create visual cartographic materials, which helps to effectively solve the problems of environmental monitoring [15, 16].

In this study, the research area is the territory of Ukraine, specifically focusing on the analysis of air quality as one of the parameters in the environmental monitoring system due to the military actions in the eastern part of the country. Many air pollutants, such as carbon monoxide (CO), sulfur dioxide (SO2), nitrogen dioxide (NO2), suspended particulate matter (SPM), and ozone (O3), have adverse effects on the atmosphere's quality and human health [1, 11]. High concentrations of these atmospheric pollutants can be hazardous to health, causing various health issues such as respiratory difficulties, headaches, dizziness [1, 12], asthma, lung function impairment, and cardiorespiratory diseases [1, 13]. Air pollution can arise from numerous anthropogenic sources, including household smoke and heating devices [1, 10], industrial emissions, vehicle and aircraft emissions, waste burning, and outdoor fires such as forest fires [1, 17]. These activities release significant amounts of gaseous emissions (SO2, NO2, CO, H2S, volatile organic compounds, and hydrocarbons) and particulate matter (smoke, soot, metallic particles, dust, vapors, and aerosols), all of which influence air quality. The combustion of automotive fuels (fossil fuels) contributes to the formation of most of these pollutants [1, 17].

Regarding air pollution measurement, it is essential to measure the overall air pollution, considering all major pollutants, to assess the general air quality and trends in atmospheric air quality across different regions. In this regard, the Air Quality Index (AQI) serves as a comprehensive tool that indicates air pollution levels based on multiple air pollutants. Compared to individual pollutant indicators, AQI facilitates the public's understanding of air quality levels and raises awareness about air pollution control. AQI is crucial in drawing attention to daily and monthly changes in pollutant concentrations in the environment and helps in educating and raising awareness about air quality issues [1, 18]. It also aids in forecasting and provides extended information about changes in pollutant concentrations in the air [1, 19]. The predictive assessment and extended information are based on standard reference values of pollutants in the atmosphere, corresponding to "good," "moderate," and "heavy pollution" levels.

Objectives and novelty

Thus, **the purpose of this work** is to perform cartographic modeling of pollutants based on statistical methods in the air on the territory of Ukraine to identify potentially hazardous areas in the regions for further mathematical modeling and building decision-making rules based on it.

The novelty of this work is the mapping of a large sample of pollutants throughout the study area using statistical methods for further evaluation.

There are a huge number of mathematical statistics methods for solving environmental monitoring problems, the results of which are presented in the form of cartographic models. Based on the obtained results, a decision is made to apply additional methods, or the considered method can be an independent tool for solving the problem.

1. Related studies

Most previous studies on air pollution have primarily focused on individual pollutants such as particulate matter and SO2 [1, 12]. However, in reality, people are rarely exposed to just one pollutant [1, 13]. Research on AQI (Air Quality Index) forecasting models is gradually expanding, and several statistical and machine learning models are being developed for air quality prediction. In recent years, a range of machine learning-based air pollution forecasting models have been used to assess various pollution levels in different locations.

The most noticeable air pollutants responsible for air pollution, according to the author's research, are PM2.5 [11]. PM2.5 concentrations can be estimated [11, 20] using logistic regression [11, 21] and autoregression [22]. Different authors have removed daily pollution level forecasts [11] by predicting hourly data [11, 23] using various algorithms [11, 24]. The quality of air in urban conditions is first evaluated by actively sampling air particles in each city area. Most countries currently use air quality assessment methods [11, 25] with stationary air pollution monitoring stations [11, 26]. These reference stations can provide very accurate readings from a small number of well-selected areas that reflect various environments [11, 12].

S. Zhu et al. proposed a two-stage forecasting model based on Complementary Ensemble Empirical Mode Decomposition (CEEMD), Grey Wolf Optimizer (GWO), and Support Vector Regression (SVR) to forecast two major sources of acid rain (NO2 and SO2) [1, 22].

Z. Qingping et al. developed a hybrid general regression neural network with Ensemble Empirical Mode Decomposition (EEMDGRNN), which relies on data preprocessing and analysis for one-day PM2.5 concentration forecasts. Although this approach yields accurate results, it has significant temporal complexity [26].

Accurate AQI forecasts are of significant value to governments, businesses, and the public in making informed decisions. Although several models have been proposed for air quality forecasting, the issue of their forecasting accuracy has not been fully resolved among the existing models, casting doubt on their effectiveness [27].

Statistical models, such as multiple linear regression and autoregressive [1, 28] integrated moving average, can be simple and effective tools for air quality forecasting [29, 30], but their forecasting accuracy remains low [31, 32]. Similarly, standalone artificial intelligence models, such as support vector machines, k-nearest neighbors (KNN), and artificial neural networks, cannot

ISSN 1814-4225 (print)
Radioelectronic and Computer Systems, 2023, no. 3(107)
ISSN 2663-2012 (online)

provide satisfactory and accurate forecasts because of the non-stationary nature of pollutant concentrations [33, 34].

In their research, P. Bhalgat et al. presented an integrated model for forecasting air pollution levels using Artificial Neural Networks (ANN) and kriging. This model uses a linear regression protocol and a multilayer perceptron (ANN) for forecasting the next day. The AR and ARIMA models have successfully predicted SO2 values, but further research is needed for PM2.5 forecasting and AQI calculations [33].

The suggested hybrid models attempt to combine different models for air quality forecasting; however, there is still a lack of research on the effectiveness and accuracy of hybrid models; therefore, their rationality requires further examination [1, 35].

2. Materials and methods of the research

2.1. Standardization of AQI

In recent years, the problem of air pollution has been exacerbated by several factors, such as population growth, urbanization, rapid economic development, industrialization, increased traffic, and energy consumption, as well as increased military actions. To reduce the level of air pollution, air quality indicators have been calculated, and national air quality standards have been established to protect public health. The Air Quality Index (AQI) is calculated using a linear function (1) based on the concentration of pollutants. The numerical values of AQI with corresponding pollution levels are presented in Table 1 [13].

$$I = I_{High} - \frac{C}{C_{High}} - C_{Low} * (C - C_{Low}) + I_{Low}, \quad (1)$$

where I – the (Air Quality) index, C – the pollutant concentration, C_{Low} – concentration breakpoint that is <= C, C_{High} – concentration breakpoint that is >= C, I_{Low} – the index breakpoint corresponding to C_{Low}, I_{High} – the index breakpoint corresponding to C_{High}.

2.2. Data sources

This study uses data on emissions of pollutants into the atmosphere from stationary sources of pollution provided by the State Statistics Service of Ukraine. The dataset contains information on sulfur dioxide (SO2), nitrogen oxide (NO2), ammonia (NH3), carbon monoxide (CO), PM10, PM2.5, lead (Pb), copper (Cu), zinc (Zn), chromium (Cr), cadmium (Cd), mercury (Hg), nickel (Ni), and arsenic (As) for the period from 1990 to 2021. For some pollutants, data are unavailable for the period from 1990 to 2004. PM10 and PM2.5 are represented as substances in the form of suspended solid particles with sizes greater than 2.5 micrometers and less than 10 micrometers. Due to a lack of prior information, the data on pollutants are generalized for the entire territory of Ukraine. The State Statistics Service does not provide region-specific information for the entire set of pollutants. An exception is made for carbon dioxide (CO2), for which regional information is available from 2004 to 2021, excluding temporarily occupied territories such as the Autonomous Republic of Crimea and Sevastopol. For Donetsk and Luhansk regions, the data are generalized and averaged because of difficulties or impossibility in collecting data throughout the entire territory.

Table 1

Air Quality Index Categories

Air Quality Index levels of health concern	Numerical Value	Meaning
Good	0 – 50	Air quality is considered satisfactory, and air pollution poses little or no risk.
Moderate	51 – 100	Air quality is acceptable; however, for some pollutants there may be a moderate health concern for a very small number of people who are unusually sensitive to air pollution.
Unhealthy for sensitive groups	101 – 150	Members of sensitive groups may experience health effects. The general public is not likely to be affected.
Unhealthy	151 – 200	Everyone may begin to experience health effects; members of sensitive groups may experience more serious health effects.
Very unhealthy	201 – 300	Health alert: everyone may experience more serious health effects.
Hazardous	>300	Health warnings of emergency conditions. The entire population is more likely to be affected [2]

Using built-in analysis tools in ArcGIS Pro, a preliminary analysis of air pollutants on the territory of Ukraine from 1990 to 2021 was conducted (Fig. 1). Such visualization tools make it easier to understand and analyze changes in air pollutants. The analysis of tabular data in ArcGIS Pro is performed in the Python programming language using the Matplotlib and Pandas libraries. The capabilities of the ready-made GIS environment allow to embed additional modules in the Python language, so that analysis and visualization tools can be modified to solve a specific problem. The standard set of features allows you to visualize data in any conventional way – a bar chart (histogram, line chart), line chart, pie chart, scatter plot, geographic, "sun rays". It is such a diagram as a geographical one that can be supplemented with an author's script to solve a specific task if this is not considered by the developers. All other types of charts have ready-made built-in tools for quick use and analysis of the original dataset. To analyze the distribution of pollutants in the air on the territory of Ukraine from 1990 to 2021, a grouped bar chart was built. The height of each column corresponds to the value of each pollutant in the air in the country. This type of primary analysis of statistical data has the following advantages: maximum clarity, simple construction of a bar chart even without the use of a special GIS and programming language, various visualization subtypes for solving specific problems, and the ability to combine with other types. The disadvantage is the lack of information content with many categories and the impossibility of providing continuous variables.

2.3. Kriging

Spatial interpolation techniques can be divided into two main categories: deterministic and geostatistical approaches. To put it simply, deterministic methods do not attempt to capture the spatial structure in the data. They only use predefined mathematical equations to predict values at unsampled locations (by weighing the attribute values of samples with known location). In contrast, geostatistical approaches intend to fit a spatial model to the data. This enables us to generate a prediction value at unsampled locations (like deterministic methods) and to provide users with an estimate of the accuracy of this prediction. Deterministic methods gather the triangulated irregular network (TIN), inverse distance weighting (IDW) and Trend surface analysis techniques. Geostatistical approaches include kriging and its variants – ordinary kriging, simple kriging, regression kriging or kriging with an external drift, co-kriging, and point kriging / block-kriging [10]. In this article, we will use ordinary kriging.

In general, kriging is the most commonly used geostatistical approach for spatial interpolation. Kriging techniques rely on a spatial model between observations (defined by a variogram) to predict attribute values at unsampled locations. One of the specificity of kriging methods is that they do not only consider the distance between observations but also capture the spatial structure in the data by comparing observations separated by specific spatial distances two at a time [10].

Ordinary kriging is a mathematical interpolation method that solves the problem of interpolating results (outputs, responses) obtained at a limited number of locations for air quality [10]. Kriging weights come from a semi-variogram that was developed by examining the spatial structure of the data. To create a continuous surface or map of the phenomenon, predictions are made for locations in the study area based on the Semi Variogram and the spatial arrangement of measured values that are nearby. Ordinary kriging approaches are available in various forms, but all are based on the concept of a basic linear regression algorithm.

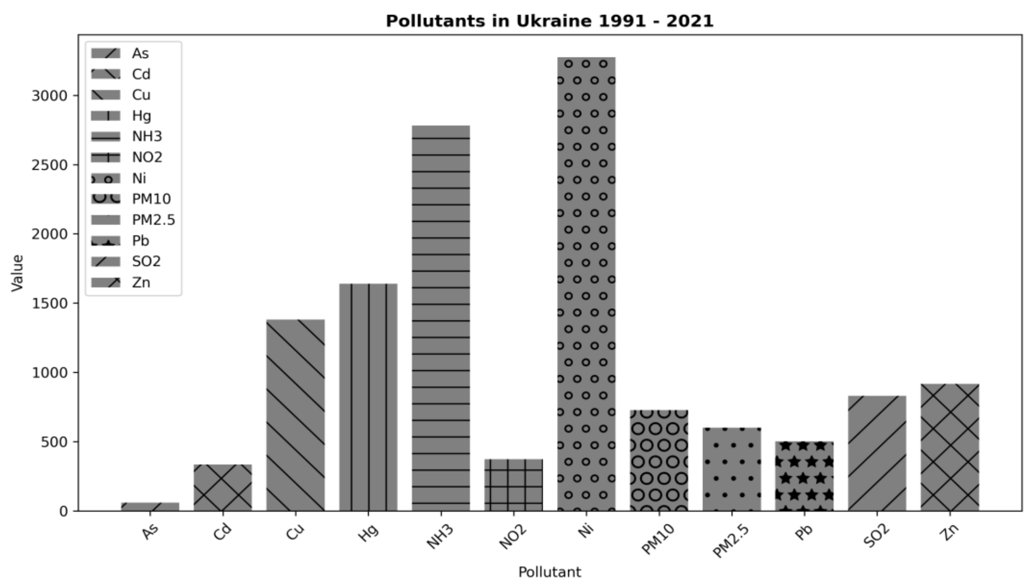

Fig. 1. Levels of air pollutants on the territory of Ukraine from 1990 to 2021

Radioelectronic and Computer Systems, 2023, no. 3(107)
ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

The estimate at location u can be expressed as (2).

$$Z(u) = \sum_{\alpha=1}^{n} \lambda(u) Z(u_{\alpha}), \tag{2}$$

where Z (u_α) is the random variable model at location u_α. The u_α's are the n data locations, the λ_α (u)'s are the ordinary kriging weights and Z(u) is the estimated value. Here the weights λ_α (u) are subject to the system.

The Semi Variogram approach is a geostatistical function that describes in numerical terms (or represents graphically) the predictability relationship between points of data at some distance from each other. This predictability constrains the confidence within which the estimation of a value is made; therefore, an estimate of the risk is quantified.

Simple kriging is the simplest kriging variant. In this case, the deterministic trend, m, is known and considered constant over the entire field under study.

Regression kriging is similar to ordinary kriging in that it considers that the deterministic trend is not constant over the whole field but depends on the spatial location of the observation.

When one disposes of a high spatial resolution auxiliary variable V_0 and wants to capture the spatial variability or correlation of a second variable V_1, co-kriging is of particular interest. In fact, the objective is to evaluate the spatial structure of V_1 regarding V_0 with the samples available and then interpolate this spatial structure at unsampled locations. As previously stated, V_1 is generally time-consuming and or expensive to obtain and it is much easier to use auxiliary data to improve the prediction of V_1 at unsampled locations. Co-kriging is more difficult to implement than the other kriging techniques, but it might result in better predictions if it is performed correctly. All the kriging techniques aim at predicting the value of a variable at specific unsampled locations. These locations can be considered spatial points (or more precisely as pixels in the grid of interpolation). Consequently, these kriging approaches are also referred to as point kriging methods. When the uncertainty is relatively large, one might want to smooth the interpolated results by performing kriging on a larger area than single pixels. This type of kriging interpolation is known as block kriging. This has the advantage of lowering prediction errors over the map. Obviously, it comes with the risk of losing some useful information but when the uncertainty is too important, it might be worth it [10].

Therefore, spatial interpolation techniques are summarized and structured in Table 2.

This work was carried out to achieve the following objectives:

– to quantify atmospheric emissions from major industrial sources located in the city and its vicinity.

– to evaluate the influence of these industrial sources on the air quality of the city [10].

– to study the influence of traffic on roads on air pollution [10, 27].

Table 2

Spatial interpolation techniques Methods		
Deterministic methods	Geostatistics methods	
1.Triangulated Irregular Network (TIN) – uses predefined mathematical equations to predict values at unsampled locations.	Kriging	1.Ordinary Kriging – solves the interpolation problem for air quality by considering a spatial model and variogram.
2. Inverse Distance Weighting (IDW) – Weighs attribute values of known samples based on distance to predict values at unsampled locations.		2. Simple Kriging – assumes a constant deterministic trend over the entire field.
3. Trend Surface Analysis – predicts values at unsampled locations using predefined mathematical trends.		3. Regression Kriging – considers spatially varying deterministic trends.
		4. Co-Kriging – evaluates spatial structure regarding an auxiliary variable (V0) and interpolates the spatial structure of the variable of interest (V1) at unsampled locations using samples of both V0 and V1.
		5. Point Kriging – predicts the value of a variable at specific unsampled locations.
		6. Block Kriging – smoothest interpolated results over a larger area to reduce prediction errors.

3. Results

After the primary assessment of the statistical data, the evaluation of the air pollutants' samples is conducted using a GIS approach. The data on the concentrations of sulfur dioxide (SO2), nitrogen oxide (NO2), ammonia (NH3), carbon monoxide (CO), PM10, PM2.5, lead (Pb), copper (Cu), zinc (Zn), chromium (Cr), cadmium (Cd), mercury (Hg), nickel (Ni), and arsenic (As) from stationary stations on the territory of Ukraine for the period 1990–2021 were imported into ArcGIS Pro and interpreted using Kriging according to formula (1). The spatial distribution mapping of air pollutants on the territory of Ukraine presented in this work is shown in Fig. 2 to 5. US Geological Survey (USGS) data are used to determine the geographical boundaries of each air pollutant on Ukraine's territory [13].

In the context of the cartograms illustrating the spatial distribution of pollutants in Ukraine for the period from 1990 to 2021 (presented in Fig. 2 to 5), it's important to note that the data for the years 2022 and 2023 is currently unavailable in the public domain. The color gradations on the cartograms represent the quantitative differences between pollutants for the specified years up to 2021. Any changes or developments in air quality related to pollutants in the years 2022 and 2023 have not been incorporated into the visualization because of the unavailability of public data for that period. Regarding the years 2022 and 2023, similar cartographic representations and analyzes will be conducted once the relevant data becomes publicly accessible. The forthcoming cartograms will provide a comprehensive view of pollutant distribution, allowing for a comparative analysis similar to that conducted for the years 1990 to 2021. The intent is to maintain consistency and enable a holistic understanding of the evolution of air quality in Ukraine over the specified time range.

The data are classified into 15 ranges in which light shades correspond to low values and dark shades correspond to high values. A range of 15 classes is optimal for representing differences in pollutant amplitude, and the fewer quantitative characteristics for each pollutant, the fewer classes may be needed to make the mapping model informative. As seen from the obtained cartographic material, the concentration of each air pollutant ranges from 32 to 523 Kg/Mg.

Stations for collecting quantitative characteristics of a pollutant are also presented on cartographic models of spatial distribution and are designated in the legends as "Sample location".

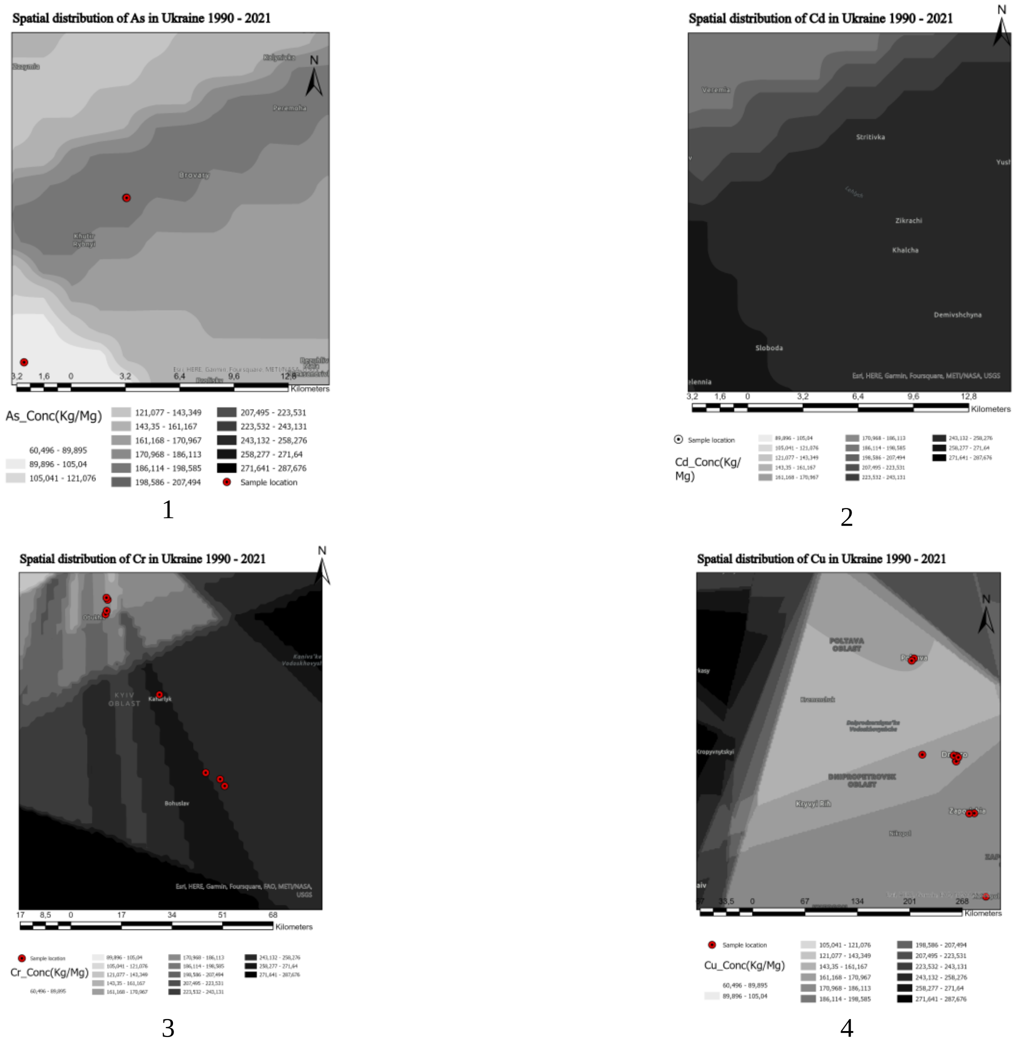

Fig. 2. Spatial distribution of pollutants on Ukraine's territory from 1990 to 2021:
1 – As; 2 – Cd ; 3 – Cr, 4 - Cu

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

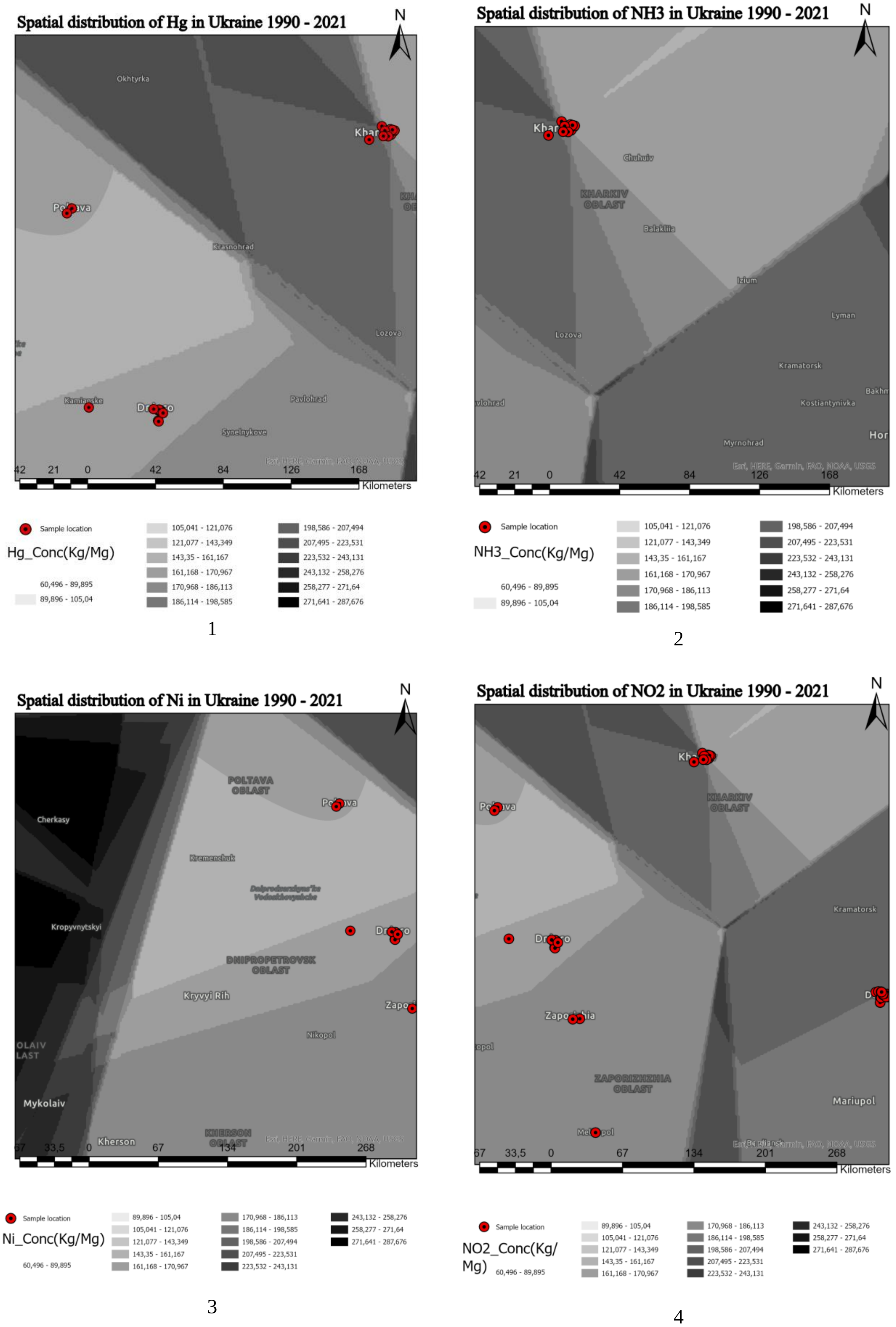

Fig. 3. Spatial distribution of pollutants on Ukraine's territory from from 1990 to 2021:
1 – Hg; 2 – NH3; 3 – Ni; 4 – NO2

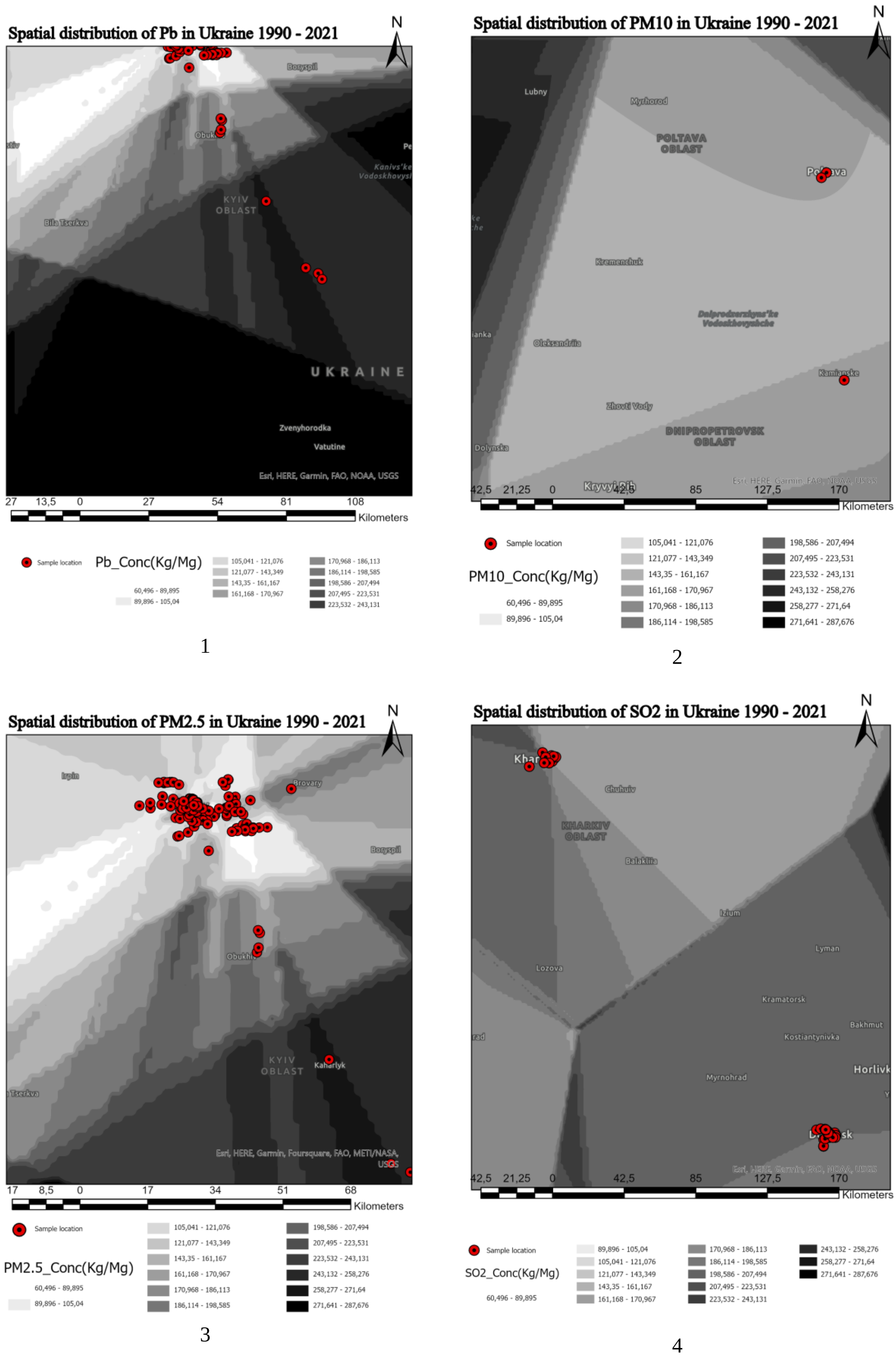

Fig. 4. Spatial distribution of pollutants Ukraine's territory from 1990 to 2021:
1 – Pb; 2 – PM10; 3 – PM2.5; 4 – SO2

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

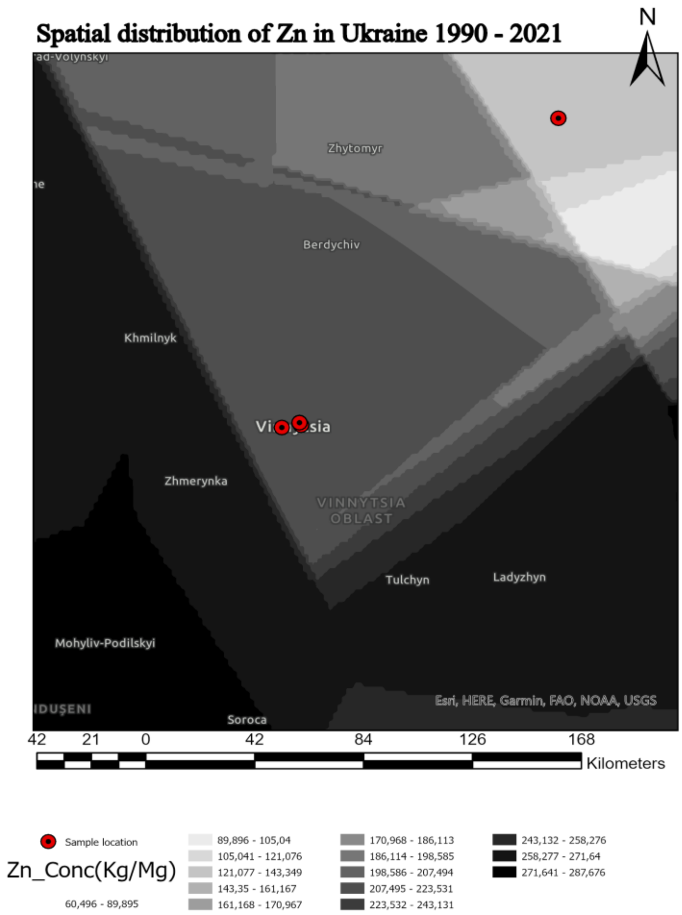

Fig. 5. Spatial distribution of Cr on Ukraine's territory from 1990 to 2021

The largest number of collection stations are presented in the central part of Ukraine; the greater the number of collection stations, the larger the set of a priori information, which means the results of geostatic modeling should be better. However, no quantitative assessment of the results of geostatistical modelling is performed in this work.

Because not all regions have stations for collecting information on air pollutants, and during periods of armed conflict, environmental monitoring in the eastern part of the country is difficult or not conducted, the constructed cartographic model of spatial distribution of air pollutants for the eastern part of Ukraine may not be 100% informative. Linear regression methods are applied to assess the effectiveness of the modelling in the interval from [0,1].

4. Discussion

The significance of environmental air quality monitoring necessitates continuous improvement in assessing and forecasting air pollutants. Kriging, the current approach, relies on routine monitoring data, resulting in a sizable dataset with uneven spatial distribution. To enhance accuracy, employ complementary mathematical methods is crucial. In this study, we propose the integration of non-linear regression methods tailored for non-uniform spatially distributed data. This addition offers distinct advantages, allowing the assessment and prediction of air quality levels with parameters exhibiting non-linear dependencies. Consequently, more accurate simulation results can be obtained over a vast geographical area.

Mathematical statistical methods, supplemented by mathematical modeling such as regression, serve as versatile and universal tools for assessing and predicting air quality. This blend provides a robust foundation for comprehensive analysis and forecasting, accommodating the complex dynamics inherent in air quality assessment.

The presented methodology for constructing cartographic models of air pollutants is universally applicable, effectively addressing challenges in environmental monitoring and aiding in decision-making processes. Its versatility enables its application across diverse territories, serving as an asset in various environmental contexts.

During times of active hostilities, traditional monitoring stations may be incapacitated or inaccessible to the public, particularly in conflict-ridden regions such as Ukraine. Consequently, access to real-time environmental data and monitoring is confined to specialized professionals. In such critical circumstances, real-time environmental monitoring is imperative to gather crucial information about external factors influencing the overall regional situation. This information is vital for informed decision-making, enabling effective planning and implementation of measures aimed at the restoration and stabilization of affected regions.

Conclusions

The main contribution of this research lies in the integration of geostatistical analysis and cartographic modeling to comprehensively characterize the spatial distribution of air pollutants across Ukraine from 1991 to 2021. Despite limited prior data availability, the study creatively combines geostatistical modeling with regression analysis techniques, allowing for a quantitative assessment of the mathematical modeling results. This innovative approach enhances our ability to visualize and interpret air pollutant patterns through the development of effective cartographic models.

Furthermore, this research justifies the suitability of employing Kriging, a well-established geostatistical method widely used in the field. The rationale for its adoption stems from its ease of implementation and ready availability as a built-in tool in Geographic Information Systems (GIS). Kriging’s strength lies in its applicability for deterministic modeling without imposing significant constraints. However, it's essential to acknowledge and address the limitations associated with Kriging, such as challenges in estimating internal variance for new study points, use of Gaussian or Matérn functions within existing GIS systems, and the complexity of applying low-order polynomial trends while commonly relying on universal constant mean values.

By navigating these methodological considerations and limitations, this study aims to significantly contribute to a more nuanced understanding of air pollution distribution in Ukraine, providing valuable insights for effective environmental management and informed policy-making in the region.

Contributions of authors: conceptualization, methodology – **Anna Topchiy**; formulation of tasks, analysis **– Anna Topchiy, Olga Butenko**; development of model **– Anna Topchiy**; analysis of results, visualization of models **– Anna Topchiy**; original draft preparation – **Anna Topchiy**, review and editing article **– Olga Butenko.**

All the authors have read and agreed to the published version of this manuscript.

References

1. Maltare, N., & Vahora, S. Air Quality Index prediction using machine learning for Ahmedabad city. *Digital Chemical Engineering*, 2023, vol. 7, iss. 1, article no. 100093. DOI: 10.1016/j.dche.2023.100093.

2. Akinfolarin, O. M., Boisa, N., & Obunwo, C. Assessment of particulate matter based air quality index in Port Harcourt Nigeria. *Journal of Environmental Analytical Chemistry*, 2017, vol. 4, iss. 4, article no. 1000224. DOI: 10.4172/2380-2391.1000224.

3. Kurt, A., & Oktay, A. B. Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks. *Expert Systems with Applications*, 2010, vol. 37, iss. 12, pp. 7986-7992. DOI: 10.1016/j.eswa.2010.05.093.

4. Kulkarni, R., Bharadwaj, R., & Sakhare, K. Machine Learning Approach to Analyze Sensor Data of Air Pollutants for Sustainable Smart Cities. *European Chemical Bulletin*, 2019, vol. 12, iss. 10, pp. 2537-2549. ISBN: 2063-5346.

5. Zhu, S., Lian, X., Liu, H., Hu, J., Wang, Y., & Che, J. Daily air quality index forecasting with hybrid models: a case in China. *Environmental Pollution*, 2017, vol. 231, iss. 2, pp. 1232-1244. DOI: 10.1016/j.envpol.2017.08.069.

6. Sowlat, M. H., Gharibi, H., Yunesian, M., Mahmoudi, M. T., & Lotfi, S. Anovel, fuzzy-based air quality index (FAQI) for air quality assessment. *Atmospheric Environment*, 2011, vol. 45, iss. 12, pp. 2050-2059. DOI: 10.1016/j.atmosenv.2011.01.060.

7. Wu, Q., Lin, H. A novel optimal-hybrid model for daily air quality index prediction considering air pollutant factors. *Science of The Total Environment*, 2019, vol. 683, pp. 808–821. DOI: 10.1016/j.scitotenv.2019.05.288.

8. Suroshe, S., Dharpal, S. V., & Ingole, N. W. Prediction of Air Quality Index Using Regression Models. *GIS science journal*, 2022, vol. 9, iss. 8, pp. 576-591. Available at: https://drive.google.com/file/d/19tDn5c87f6q6BHq_n1vlnYQaaTQQhwHn/view. (accessed 3.02.2023).

9. Kleijnen, J. P. C. Kriging: Methods and Applications. *CentER Discussion Paper Series*, 2017, vol. 2017-047, pp. 1-16. DOI: 10.2139/ssrn.3075151.

10. Tyagi, A., & Singh, P. Applying kriging approach on pollution data using gis software. *International Journal of Environmental Engineering and Management*, 2013, vol. 4, iss. 3, pp. 185-190. ISSN: 2231-1319.

11. Sokolov, D., Merlak, V., Oryekhov, O., & Plakhteev, A. Ekolohichnyy monitorynh z vykorystannyam bezprovidnykh sensornykh merezh: rozroblennya ta eksperymenty [Environmental monitoring with wireless sensor networks application: development and experiments]. *Radioelektronni i komp'uterni sistemi – Radioelectronic and computer systems*, 2019, vol. 3, iss. 91, pp. 40-47. DOI: 10.32620/reks.2019.3.04. (In Ukrainian).

12. Ponochovnyy, Yu., & Kharchenko, V. Metodolohiya zabezpechennya harantozdatnosti informatsiyno-

keruyuchykh system z vykorystannyam bahatotsil'ovykh stratehiy obsluhovuvannya [Dependability assurance methodology of information and control systems using multipurpose service strategies]. *Radioelektronni i komp'uterni sistemi – Radioelectronic and computer systems*, 2020, vol. 3, iss. 95, pp. 43-58. DOI: 10.32620/reks.2020.3.05. (In Ukrainian).

13. Maryushko, M., & Pashchenko, R. Fraktal'nyy analiz kosmichnykh znimkiv Sentinel-2 dlya monitorynhu sil's'kohospodars'kykh kul'tur [Fractal analysis of Sentinel-2 satellite imagery for monitoring of agricultural crops]. *Radioelektronni i komp'uterni sistemi – Radioelectronic and computer systems*, 2020, vol. 4, iss. 96, pp. 34-47. DOI: 10.32620/reks.2020.4.03. (In Ukrainian).

14. Khaliq, A., Peroni, L., & Chiaberge, M. Land cover and crop classification using multitemporal Sentinel-2 images based on crops phenological cycle. *IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS)*, Salerno, Italy, 21-22 June 2018, IEEE, 2018, pp. 1-5. DOI: 10.1109/EESMS.2018.8405830.

15. Banerjee, T., & Srivastava, R. Assessment of the ambient air quality at the Integrated Industrial Estate-Pantnagar through the air quality index (AQI) and exceedence factor (EF). *Asia-Pacific Journal of Chemical Engineering*, 2011, vol. 6, iss. 1, pp. 64-70. DOI: 10.1002/apj.450.

16. Kumar, A., & Goyal, P. Forecasting of air quality in Delhi using principal component regression technique. *Atmospheric Pollution Research*, 2011, vol. 2, iss. 4, pp. 436-444. DOI: 10.5094/APR.2011.050.

17. Mishra, D., & Goyal, P. Analysis of ambient air quality using fuzzy air quality index: a case study of Delhi, India. *International Journal of Environmental and Pollution*, 2016, vol. 58, iss. 3, pp. 149-159. DOI: 10.1504/IJEP.2015.077173.

18. Soni, H. B., & Patel, J. Assessment of Ambient Air Quality and Air Quality Index in Golden Corridor of Gujarat, India: a case study of Dahej port. *International Journal of Environment*, 2017, vol. 6, iss. 4, pp. 28-41. DOI: 10.3126/ije.v6i4.18908.

19. Choi, G., Heo, S., & Lee, J-T. Assessment of environmental injustice in Korea using synthetic air quality index and multiple indicators of socioeconomic status: a cross-sectional study. *Journal of the Air & Waste Management Association*, 2015, vol. 66, iss. 1, pp. 28-37. DOI: 10.1080/10962247.2015.1107657.

20. Shelestov, A., Kolotii, A., Lavreniuk, M., Yailymov, B., Shumilo, L., & Korsunska, Y. tsyfrovizatsiya rozvytku mist: urban atlas na osnovi vidkrytykh danykh dlya mist Ukrayiny [Digitalization of city development: urban atlas on the basis of open data for cities of Ukraine]. *Radioelektronni i komp'uterni sistemi – Radioelectronic and computer systems*, 2021, vol. 3, iss. 99, pp. 19-28. DOI: 10.32620/reks.2021.3.02. (In Ukrainian).

21. Dominici, F., Peng, R. D., Barr, C. D., & Bell, M. L. Protecting human health from air pollution: shifting from a single-pollutant to a multi-pollutant approach. *Epidemiology*, 2010, vol. 21, iss. 2, pp. 187-194. DOI: 10.1097/EDE.0b013e3181cc86e8.

22. Zhu, S., Qiu, X., Yin, Y., Fang, M., Liu, X., Zhao, X., & Shi, Y. Two-step hybrid model based on data preprocessing and intelligent optimization algorithms (CS and GWO) for NO2 and SO2 forecasting. *Atmospheric Pollution Research*, 2019, vol. 10, iss. 4, pp. 1326-1335. DOI: 10.1016/j.apr.2019.03.004.

23. Mahendra, H. N., Mallikarjunaswamy, S., Mahesh Kumar, D., Kumari, S., & Kashyap, S. Assessment and Prediction of Air Quality Level Using ARIMA Model: A Case Study of Surat City, Gujarat State, India. *Nature Environment and Pollution Technology*, 2023, vol. 22, iss. 1, pp. 199-210. DOI: 10.46488/NEPT.2023.v22i01.018.

24. Angelin, J., & Sasi Kumar, A. PM2.5 prediction using machine learning hybrid model for smart health. *International Journal of Engineering and Advanced Technology (IJEAT)*, 2019, vol. 9, iss. 1, pp. 6500-6504. DOI: 10.35940/ijeat.A1187.109119.

25. Zhu, D., Changjie, C., Tianbao, Y., & Xun, Z. A machine learning approach for air quality prediction: model regularization and optimization. *Big Data and Cognitive Computing*, 2018, vol. 2, iss. 1, article no. 5. DOI: 10.3390/bdcc2010005.

26. Qingping, Z., Haiyan, J., Jianzhou, W., & Jianling, Z. A hybrid model for PM2.5 forecasting based on ensemble empirical mode decomposition and a general regression neural network. *Science of The Total Environment*, 2014, vol. 496, pp. 264-274. DOI: 10.1016/j.scitotenv.2014.07.051.

27. Kumar, K., & Pande, B.P. Air pollution prediction with machine learning: A case study of Indian Cities. *International Journal of Environmental Science and Technology*, 2022, vol. 20, pp. 5333-5348. DOI: 10.1007/s13762-022-04241-5.

28. Cosma, A., & Simha, R. Machine learning method for real-time non-invasive prediction of individual thermal preference in transient conditions. *Building and Environment*, 2019, vol. 148, pp. 372-383. DOI: 10.1016/j.buildenv.2018.11.017.

29. Samal, K.K., Babu, K.S., Das, S.K., & Acharaya, A. Time Series-Based Air Pollution Forecasting using SARIMA and Prophet Model. *IEEE International Conference on Information Technology and Computer Communications*, Singapore, 2019, pp. 124-129. DOI: 10.1145/3355402.3355417.

30. Li, Y., & He, J. Design of an intelligent indoor air quality monitoring and purification device. *IEEE Information Technology and Mechatronics Engineering*

Conference (ITOEC), 3-5 October 2017, Chongqing, China, 2017, pp. 1147-1150. DOI: 10.1109/ITOEC. 2017.8122535.

31. Santos, J.A., Jiménez, M., & Espinosa, F. Effect of event-based sensing on IoT node power efficiency: Case study: Air quality monitoring in smart cities. *IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC)*, 2017. DOI: 10.1109/ACCESS.2019.2941371.

32. Amado, T. M., & Dela Cruz, J. C. Development of Machine Learning- based Predictive Models for Air Quality Monitoring and Characterization. *TENCON IEEE Region Conference*, 16-19 Nov 2018, Osaka, Japan, 2018, pp. 0668-0672. DOI: 10.1109/TENCON.2018.8650518.

33. Bhalgat, P., Bhoite, S., & Pitare, S. Air quality prediction using machine learning algorithms. *International Journal of Computer Applications Technology and Research*, 2019, vol. 8, iss. 9, pp. 367-370. DOI: 10.7753/IJCATR0809.1006.

34. Dastoorpoor, M., Idani, E., Goudarzi, G., & Khanjani, N. Acute effects of air pollution on spontaneous abortion, premature delivery, and stillbirth in Ahvaz, Iran: a time-series study. *Environmental Science and Pollution Research*, 2018, vol. 25, iss. 4, pp. 5447–5458. DOI: 10.1007/s11356-017-0692-9.

35. Taylan, O. Modeling and analysis of ozone concentration by artificial intelligent techniques for estimating air quality. *Atmospheric Environment*, 2017, vol. 150, pp. 356–365. DOI: 10.1016/j.atmosenv.2016.11.030.

Received 01.08.2023, Accepted 20.09.2023

АНАЛІЗ ЗАБРУДНЮЮЧИХ РЕЧОВИН У ПОВІТРІ НА ТЕРИТОРІЇ УКРАЇНИ ЗА ДОПОМОГОЮ ГЕОСТАТИСТИЧНИХ МЕТОДІВ

Ольга Бутенко, Анна Топчий

Якість повітря останнім часом викликає величезне занепокоєння, оскільки воно безпосередньо впливає на життя людей. Постійний моніторинг якості повітря та прогнозування динаміки його змін є необхідним етапом оцінювання його поточного стану та визначення концентрації забруднюючих речовин. Тому розробка ефективної системи оцінки та прогнозування якості повітря стала одним із найважливіших завдань. У цій роботі проводиться аналіз забруднюючих речовин на території України у період з 1990–2021 років. Набір даних про забруднюючі речовини у повітрі представлений державною службою статистики України у вигляді узагальнених таблиць, які були первинно опрацьовані для подальшого моделювання. Картографічне моделювання забруднюючих речовин проведено із застосуванням геостатистичних методів. Як результат, представлено 13 картографічних моделей просторового розподілу забруднюючих речовин повітря для різних регіонів України. У зв'язку з відсутністю офіційної інформації про наявність бойових дій, результати застосування геостатистичних методів не можуть бути інтерпретовані в контексті військової ситуації на сході країни. Інформація про бойові дії може бути зібрана з різних джерел, проте це вимагатиме великих витрат за часом для структурування та систематизації набору даних. Метод, що розглядається в роботі, не може бути застосований з урахуванням декількох параметрів одночасно, а саме значення показника забруднюючої речовини та наявність бойових дій. Для цього застосовуються додаткові методи теорії графів, регресійного аналізу для одержання кількісних оцінок результатів моделювання з урахуванням усіх факторів впливу на екологічну ситуацію. Вибраний метод є простим інструментом для вирішення екологічних завдань. Завдяки готовим ГІС системам, таким як ArcGISPro, є можливість візуалізації застосовуваних геостатистичних та математичних методів. Картографічні моделі, представлені в роботі, покривають усю територію України та мають адміністративні кордони в залежності від розташованої станції збору про забруднювальну речовину.

Ключові слова: якість повітря; дані; моделювання; ГІС; індекс якості повітря; геостатистичні моделі крігінгу.

Бутенко Ольга Станіславівна – д-р техн. наук, проф., проф. каф. геоінформаційних технологій та космічного моніторингу Землі, Національний аерокосмічний університет імені М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Топчий Анна Сергіївна – асп. каф. геоінформаційних технологій та космічного моніторингу Землі, Національний аерокосмічний університет імені М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Olga Butenko – Doctor of Technical Sciences, Professor, Professor at the Department of Geoinformation Technologies and Space Monitoring of the Earth, National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine,
e-mail: OS-B@ukr.net, ORCID: 0000-0001-9486-8633.

Anna Topchiy – PhD Student at the Department of Geoinformation Technologies and Space Monitoring of the Earth, National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine,
e-mail: annatop4y@gmail.com, ORCID: 0000-0003-0448-4543.