THE PUBLISHING HOUSE
OF THE ROMANIAN ACADEMY

PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,
Volume 23, Number 4/2022, pp. 399–406

PROCESSING, NEURAL NETWORK-BASED MODELING OF BIOMONITORING STUDIES DATA AND VALIDATION ON REPUBLIC OF MOLDOVA DATA

Radu-Emil PRECUP[1,2], Gheorghe DUCA[3], Sergey TRAVIN[4], Inga ZINICOVSCAIA[3,5,6]

[1] Politehnica University of Timisoara, Department of Automation and Applied Informatics,
Bd. V. Parvan 2, 300223 Timisoara, Romania
[2] Romanian Academy – Timisoara Branch, Center for Fundamental and Advanced Technical Research,
Bd. Mihai Viteazu 24, 300223 Timisoara, Romania
[3] Institute of Chemistry, Research Center of Physical and Inorganic Chemistry,
Str. Academiei 3, 2028 Chisinau, Republic of Moldova
[4] Russian Academy of Sciences, Semenov Federal Research Center for Chemical Physics,
Kosygina Str. 4, Building 1, 119991 Moscow, Russian Federation
[5] Joint Institute for Nuclear Research, Joliot-Curie Str. 6, 141980 Dubna, Russian Federation
[6] Horia Hulubei National Institute for R&D in Physics and Nuclear Engineering,
Str. Reactorului 30, MG-6, Bucharest-Magurele, Romania
Corresponding author: Radu-Emil PRECUP, E-mail: radu.precup@aut.upt.ro

Abstract. This paper suggests an approach to process and model the data obtained in biomonitoring studies. The approach is validated on data obtained from biomonitoring studies performed in the Republic of Moldova in 2015. Using the preliminary data, the decomposition on the basis of the pollution spectrum for the most polluted and cleanest sites is first carried out. The deviations of model predictions from the actual measurements are considered. A correlation analysis is next performed to evidence the correlation of two geographical coordinates with chemical elements. Factor analysis and regression analysis are applied to highlight the nonlinear mechanisms specific to the obtained data. A multilayer neural network-based model is derived to describe the relationship of the pollution rank to the geographic coordinates. The predictive capabilities of the model are represented graphically.

Key words: correlation analysis, factor analysis, neural networks, moss biomonitoring, regression analysis.

1. INTRODUCTION

Monitoring of air pollution using bioindicators is emerging as a potentially effective and more economical alternative performing by direct ambient air measurements [1]. It has been shown in [2] that mosses and lichens, in spite of all disadvantages, are good tools for air pollution monitoring. Moss biomonitoring is nowadays a part of the pollution monitoring programmes in most European countries as it gives evidence of the anthropogenic impact in urban and rural areas due to industrial activity, mining, transport and fossil fuel combustion, as pointed out in the recent papers [3–8], which discuss the atmospheric heavy metal deposition, the spatial patterns and the temporal trends in Europe [3], the air pollution in the Republic of Moldova [4, 5], the neutron activation analysis in assessing the pollution by trace elements in certain Polish national parks [6], the analysis of the spatial data measured in a heavily polluted region of the Czech-Polish border [7, 8], and the atmospheric deposition in Georgia [9].

Factor Analysis (FA) and Principal Component Analysis (PCA) are usually applied to evaluate the biomonitoring data [7, 9, 10]. Both FCA and PCA have some limitations because of their statistical features. For example, the weaknesses of FA consist in the ambiguity of the estimation of factor parameters and the need to specify the number of common factors before performing the analysis [7].

Building upon authors' recent papers on biomonitoring discussed above [4–10] and nonlinear modeling of various systems focusing on tensor product-based model transformation applied to tower crane systems [11], unified clustering applied to various data sets [12, 13], evolving fuzzy models applied to shape memory

alloy wire actuators [14], and linear and nonlinear observers for strip winding systems [15], this paper suggests an approach to describe the biomonitoring data obtained for moss samples collected in the Republic of Moldova in the 2015/2016 moss survey. The correlation analysis is first applied to reduce the number of determined chemical elements to two factors. The data smoothing is next performed using a discrete cosine transform. The paper also suggests an algorithm to calculate the number of linearly independent (basis) vectors in which the matrix itself can be decomposed. A multilayer neural network-based model is derived to describe the relationship of the pollution rank to the geographic coordinates.

The approach suggested in this paper is important with respect to the state-of-the-art briefly discussed above because of the advantages of the three techniques applied. These techniques, namely FA, regression analysis and neural networks, highlight the nonlinear mechanisms specific to the obtained data.

The rest of the paper is organized as follows in the order of the steps of the approach pointed out above: the factor analysis is described in the next section. The regression analysis and the neural network application are presented in Section 3. The predictive capabilities of the resulting model are discussed in terms of the modeling results given in Section 4. The conclusions are highlighted in Section 5.

2. FACTOR ANALYSIS

Preprocessing the data shows that the "pollution spectra"ranked in descending order of the total pollution load line up in a neat sequence illustrated in Figs. 1 and 2 given in [16]. This ranking leads to a list of data collection sites, and a part the list is exemplified in the right-hand parts of Figs. 1 and 2 [16].

Each "fingerprint" or spectrum of representation of each of the 32 elements with minor local extremes has a clear decreasing trend from elements representing the earth's crust (Ca, Al, Mg, Na) to trace elements present in small amounts (U, Th). At the same time, as the decrease of the total pollutant load from Singerei to Donduseni, the spectra firstly decrease, and secondly become flatter.

The generally accepted pollutant load index (PLI) is next replaced by a rank distribution, which is determined unequivocally as follows: the rank is an integer serial number of a sampling point in the above given ordered list. Assuming that the rank of matrix of the input data has a dimension of 2, according to the number of selected eigenvalues, it leads to a pair of eigenvectors, which form the basis according to Fig. 3 and Table 1 given in [16].

The vector $\mathbf{X}(R)$ is decomposed as follows as a linear combination of the two eigenvectors $\mathbf{b}_1$ and $\mathbf{b}_2$:

$$\mathbf{X}(R) = F_1\mathbf{b}_1 + F_2\mathbf{b}_2, \tag{1}$$

where F_1 and F_2 are the coefficients of the linear combinations. Fig. 4 [16] gives examples of the decomposition of experimental data in terms of the basis for most clean and most polluted data collection sites.

According to Fig. 4 (b) [16], the only significant and visible discrepancy between the experimental measurement and the model expansion are noticed only for Mg. All intermediate data collection sites with pollution ranks from 2 to 33 are laid out according to the basis in a similar way with no less quality.

Without exception, all spectrum deconvolution methods include at least two types of non-uniqueness, namely the scale one and the rotation one. Scale non-uniqueness implies that in any blind recognition method, even the spectrum with perfect resolution without reference to a specific concentration is determined at best with an accuracy of an unknown proportionality factor. The rotational non-uniqueness means that basis rotation through Euler angles can lead to both mixing and separation of individual absorption bands. It is theoretically impossible to separate them according to certain basic spectra using only algebraic techniques, without involving additional information, e.g. chemical information.

The brief description of the algorithm that implements the Moore-Penrose transform will be given as follows. This transform allows movement from some existing initial basis to a meaningful basis that will satisfy the formulated physicochemical constraints of the model. For example, this may be the requirement of the law of conservation of matter or the fulfillment of certain kinetic equations.

In this case, the idea is to reduce the two-factor task to just one-factor, so that the first factor changes from 1 to 0 as the pollution rank increases, and the second one is its complement to 1.

According to [17–19], a pseudoinverse matrix $\mathbf{C}^+$ of the matrix $\mathbf{C}$ satisfies the following relations:

$$\mathbf{C}\,\mathbf{C}^{+}\mathbf{C}=\mathbf{C}, \tag{2}$$

$$\mathbf{C}^{+}\mathbf{C}\,\mathbf{C}^{+}=\mathbf{C}^{+}, \tag{3}$$

$$(\mathbf{C}\,\mathbf{C}^{+})^{T}=\mathbf{C}\,\mathbf{C}^{+}, \tag{4}$$

$$(\mathbf{C}^{+}\mathbf{C})^{T}=\mathbf{C}^{+}\mathbf{C}, \tag{5}$$

where T indicates matrix transposition. For any matrix, its pseudoinverse exists and is unique. For a square nondegenerate matrix **C**, the pseudoinverse matrix corresponds to the inverse one $\mathbf{C}^{-1}$:

$$\mathbf{C}^{+}=\mathbf{C}^{-1}. \tag{6}$$

If **A** is a matrix of full column rank, i.e. all columns are linearly independent, then

$$\mathbf{A}^{+}=(\mathbf{A}^{T}\mathbf{A})^{-1}\mathbf{A}^{T}. \tag{7}$$

If **A** is a full string rank matrix, then

$$\mathbf{A}^{+}=(\mathbf{A}\,\mathbf{A}^{T})^{-1}\mathbf{A}. \tag{8}$$

It is convenient to assume that the wavelength sweep of the spectra is oriented horizontally, the set of basis vectors of the reference spectra is also horizontal, and the mixing matrix **C** is stretched vertically. Therefore:

$$\mathbf{D}\approx\mathbf{C}_{old}\mathbf{S}_{old}^{T}. \tag{9}$$

The approximate equal sign in (9) means that the error matrix is neglected. The mixing matrix $\mathbf{C}_{old}$ is obtained in a rather arbitrary way, in the course of solving a system of linear equations with an arbitrarily chosen basis of eigenvectors $\mathbf{S}_{old}^{T}$ according to Frobenius. In addition to permutation arbitrariness, this basis also has a large-scale uncertainty and, most importantly, a pivotal ambiguity. Thus, this leads to reaching the minimum mutual presence of spectral lines in different standards, but this does not mean that the pseudo-standards considered in this paper are in fact the spectra of individual substances. They can also be their linear combinations, and cannot be identified at "blind recognition".

However, if some information about the kinetics of the processes occurring in a mixed system is available, this allows for the calculation of the expected calculated time sweep $\mathbf{C}_{new}$ for each of the individual substances. This sweep needs to correspond to the new basis to maintain the observed pattern of changes in the total spectra over time, and (9) is transformed into

$$\mathbf{D}\approx\mathbf{C}_{old}\mathbf{S}_{old}^{T}=\mathbf{C}_{new}\mathbf{S}_{new}^{T}. \tag{10}$$

The objective is to find a linear transformation matrix **W** that can transform the existing old basis $\mathbf{C}_{old}$ into the desired one $\mathbf{C}_{new}$ making use of the still unknown new $\mathbf{S}_{new}^{T}$ basis, which expansion will give the "correct" dependences of concentrations on time. For this purpose, the left- and right-hand terms of (10) are multiplied by the pseudoinverse matrix $\mathbf{C}_{new}^{+}$ leading to:

$$\mathbf{C}_{new}^{+}\mathbf{C}_{old}\mathbf{S}_{old}^{T}=(\mathbf{C}_{new}^{+}\mathbf{C}_{new})\mathbf{S}_{new}^{T}=\mathbf{S}_{new}^{T}. \tag{11}$$

Substituting $\mathbf{C}_{new}^{+}$ in the left-hand term of (11) in terms of the expansion formula (9), it is obtained that

$$\mathbf{S}_{new}^{T}=(\mathbf{C}_{new}\mathbf{C}_{new}^{T})^{-1}\mathbf{C}_{new}\mathbf{C}_{old}\mathbf{S}_{old}^{T}. \tag{12}$$

The linear transformation matrix **W** (i.e. related to the transition from the basis), which gives a certain expansion to the current basis, in which will give a meaningful kinetic picture, is carried out by an elementary algebraic transformation according to Moore-Penrose. This transformation is advantageous as it does not require the expenditure of serious computing resources and the use of capricious iterative procedures of minimization and fitting. The formula to calculate **W** is

$$\mathbf{W}=(\mathbf{C}_{new}\mathbf{C}_{new}^{T})^{-1}\mathbf{C}_{new}\mathbf{C}_{old}. \tag{13}$$

Thus, (12) and (13) lead to the expression of the transition:

$$\mathbf{S}_{new}^{T} = \mathbf{W}\,\mathbf{S}_{old}^{T}. \tag{14}$$

This one step (i.e. without any iterations) transformation proves to be useful. The application of the Moore-Penrose transformation led to obtaining such a variant of the basis, which corresponds to the simplest and most visual form of the mixing matrix **C**. In this paper, the mixing matrix consists of only two columns, which are also linearly dependent, and the general expression of this matrix is

$$\mathbf{C} = [c_{RS}]_{R=1\ldots R_{\max},\, S\in\{1,2\}}. \tag{15}$$

The value of an element of the second column is always just the addition of the first element to 1. As the pollution factor decreases from 1 to 0, the value of the element of the second column in parallel increases from 0 to 1:

$$c_{R1} + c_{R2} = 1, \tag{16}$$

$$c_{R1} \equiv F_1(R) \approx 1 - R / R_{\max}, \tag{17}$$

$$c_{R2} \equiv F_2(R) \approx R / R_{\max}. \tag{18}$$

As shown in (17) in (18) in a rough approximation, it can be assumed that the factors $F_1(R)$ and $F_2(R)$ linearly increase and decrease, respectively, as the serial number (rank) $i = 1\ldots m$ changes in the list of the most polluted (in the opposite direction – the cleanest) cities of the Republic of Moldova. Therefore, the model curve of the rank dependence of the pollution rank factor $F_1(R)$ and the additional purity rank factor $F_2(R)$ can be successfully approximated using the following formulae that do not have a special physical meaning, but describe sufficiently well the experiment:

$$F_1(R) = \alpha_0 + \alpha_1 e^{-\gamma(R+\delta)^2} - \beta R, \tag{19}$$

$$F_2(R) = 1 - F_1(R), \tag{20}$$

where $\alpha_0 = 0.53$, $\alpha_1 = 0.47$, $\gamma = 0.008$, $\delta = -1.2$ and $\beta = 0.016$.

The rank dependences of the mixing matrix components of the most polluted site (Singerei) to the cleanest site (Donduseni) are illustrated in Fig. 1, which shows the almost perfect agreement between the results of decomposition of the experimental data and the predictions of the model.

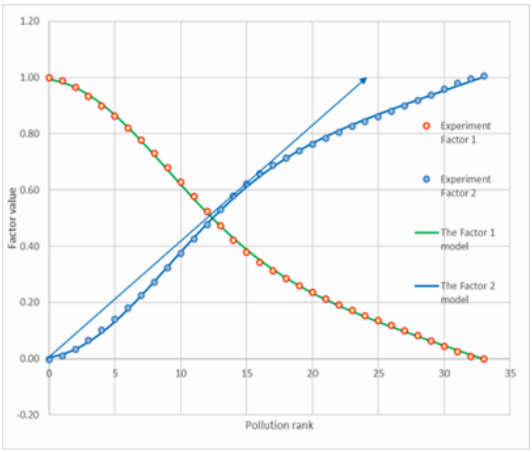

Fig. 1 – Rank dependences of the mixing matrix components from the most polluted site (Singerei) to the cleanest site (Donduseni).

All $1088 = 34 \times 32$ results of measurement are described with satisfactory accuracy using a one-factor dependence on the basis of two support vectors according to the results given in this section. The support vectors are summarized in Table 1 [16].

Table 2 given in [16] offers the deviations of model predictions from actual measurements. Each cell contains data for a specific locality (row) and a specific chemical element (column). For visual appearance, the cells are colored green for deviations not exceeding 15%, light green for deviations in the range from 15% to 30%, yellow for the range of 30% to 50%, and pink for the range of 50% to 100%. Finally, cells with a deviation of more than 100%, which were 12 and represents about one percent of the experimental data are colored bright red. Table 3 given in [16] summarizes the number of cells that fell within a certain deviation range.

3. REGRESSION ANALYSIS AND NEURAL NETWORK APPLICATION

In the next step, the data obtained for 34 points is spread to the entire territory, i.e. a geographic forecast is made. If the correlation diagram is expanded from a 32×32 matrix (by the number of elements) to 34×34 by adding two geographical coordinates, it will become obvious that neither latitude nor longitude correlates with any of the chemical elements, nor does any of the element concentrations show geographical preferences. This is suggestively shown in Table 4.

Similarly, the diagrams of the change in the rank factor F_1 (abscissa) from geographic latitude and longitude (ordinate) given in Fig. 2 demonstrate a multifold character. This sophisticates selection of a function for the regression dependence.

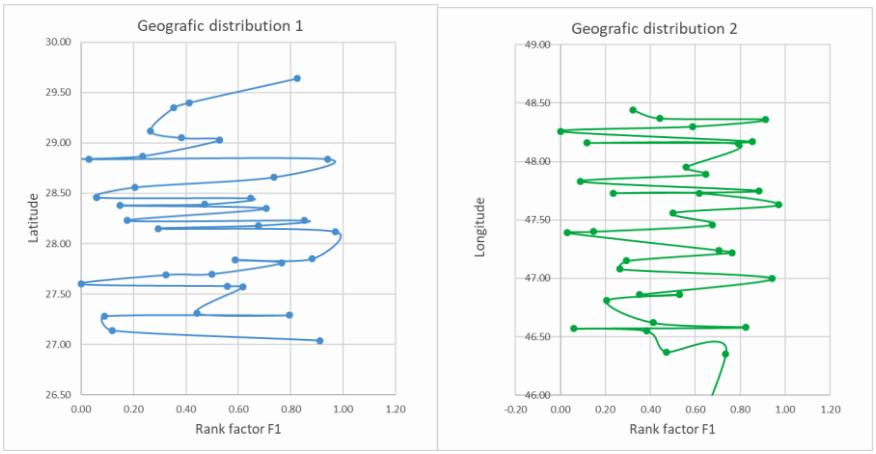

Fig. 2 – Diagrams showing the variability of geographic latitude and longitude as the pollution rank factor F_1 changes.

Therefore, traditional methods of regression analysis are completely useless within the framework of the problem under study. Fundamentally new opportunities are provided by artificial intelligence techniques such as machine learning techniques, which include neural networks. It is no need to know all the details of process, in order to write out flawed regression equation, even if it is multifactorial, with high degree polynomials and numerous cross terms. Instead, using the multilayer neural network architecture with backpropagation of error algorithm involved in training, it will be proved in this section that it is possible to digitally represent all the hidden patterns as delicately as possible. Nevertheless, such a model would actually have predictive and statistical power greater than experimenter could take into account.

With a small amount of data, it is necessary to have a model whose parameters will be optimized. The model describes a set of assumptions and the main task is to choose the most appropriate one. The model should cover the possible options which can be met in practice. With a small amount of data, the model will produce a large variance for the output parameters, but as the amount of data increases, the variance will decrease and the forecast will be more unambiguous.

The model cannot take into account all factors. It is created by a person and puts limited possibilities into it. Such a model is more suitable as an element of more complex calculations. The general structure of a neural network-based model is presented in Fig. 3 focusing on the two inputs represented by the pollution rank factor F_1 and the purity rank factor F_2, and the output (or the response) represented by the pollution rank R actually applied here.

The question related to the number of layers and perceptrons (i.e. neurons) in a layer is not simple and it is usually solved heuristically. The neural network considered in this paper was configured with the parameters 2:9:9:1, i.e. 2 neurons in the input layer, two hidden layers each with 9 neurons and one neuron in the output layer, as suggestively illustrated in Fig. 3. The activation functions are of sigmoid type.

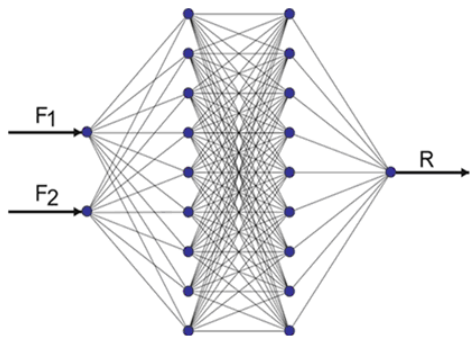

Fig. 3 – Structure of a multilayer neural network with two internal (hidden) layers, applied to "unfold" the folded relationship of pollution rank with geographic coordinates.

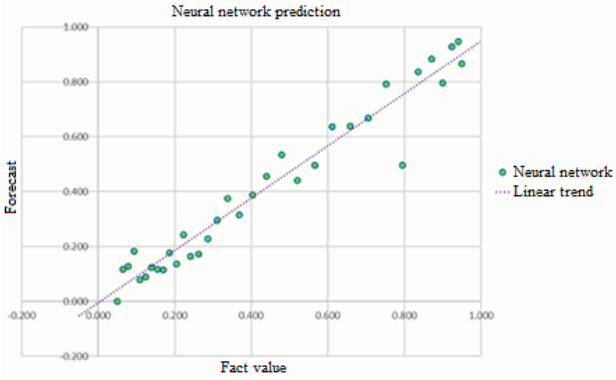

Fig. 4 – Pollution rank *R versus* pollution rank factor F_1 computed by the neural network on the testing data and the linear trend.

Figure 4 gives a sample of the neural network results on the testing data, namely the nonlinear input-output map $R(F_1)$. In other words, Fig. 4 gives the correspondence diagram of the actually determined values of the input-output-map $R(F_1)$ and the model predictions produced by the neural network.

The coefficient of determination is the proportion of the explained variance of the deviations of the dependent variable from its mean value. The dependent variable is explained (predicted) using a function of the explanatory variables, in our case, geographic coordinates. The coefficient of determination is analogous to the square of the correlation coefficient between the dependent variable and its predictive values using explanatory variables.

Figure 4 suggestively shows that high coefficient of determination values were attained by the neural network. Thus, with the help of a neural network as a representative artificial intelligence technique, artificial intelligence provides results that regression analysis could not cope with.

4. MODELING RESULTS

The predictive capabilities of the resulting model will be discussed in this section. The diagrams presented in Figs. 5 to 8 show the estimated data for the entire territory of the Republic of Moldova (and partly beyond). In fact, a square grid was used for coordinates from 27.0°N to 30.0°N and from 43.9°W to 48.5°W.

The results will be different if other neural network architectures will be trained, validated and tested on this application. Such architectures and recent applications include FPGA-implementation of deep neural networks and associated learning algorithms [20], support vector machines in functional verification [21], deep neural networks [22, 23], neuro-fuzzy models [24], and evolving fuzzy models [14, 25, 26]. However, the conclusions will be generally kept.

5. CONCLUSIONS

This paper proposed an approach to describe the biomonitoring data obtained for moss samples collected in the Republic of Moldova. Several data processing techniques such as correlation analysis, factor analysis and regression analysis were applied prior to train a multilayer neural network-based model in the framework of the popular machine learning techniques.

The results presented in the paper are encouraging. They prove that regression analysis can be considered as useless tool in biomonitoring data description. Therefore, a multilayer neural network-based was developed and its capabilities in terms of the geographic distribution of the pollution rank were presented. According to performed data processing, only one factor describing the association of chemical elements was highlighted.

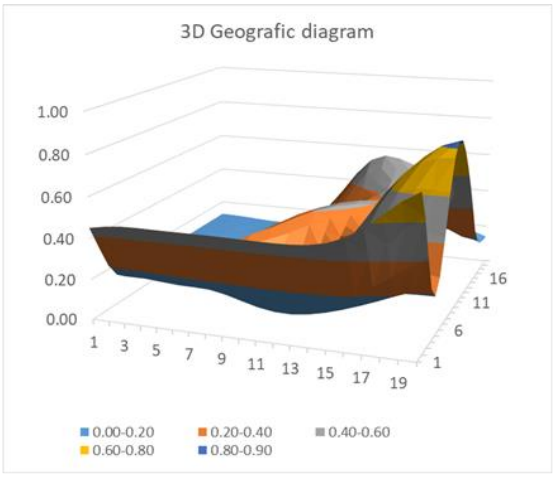

Fig. 5 – The values of the pollution rank *R* calculated using a neural network for coordinates on a 20×20 grid (relative step in latitude and longitude 0.05 of the range for each coordinate).

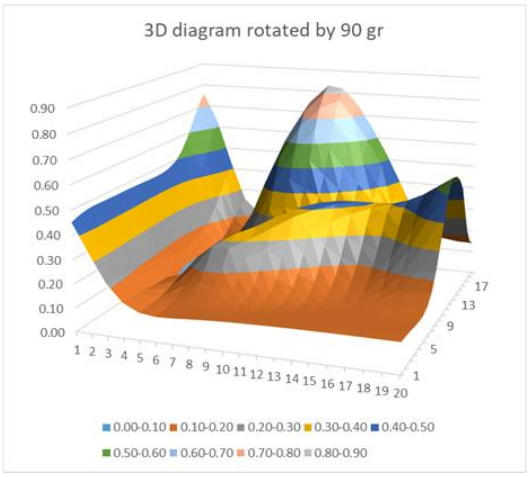

Fig. 6 – Diagram similar to Fig. 5, but rotated by 90°.

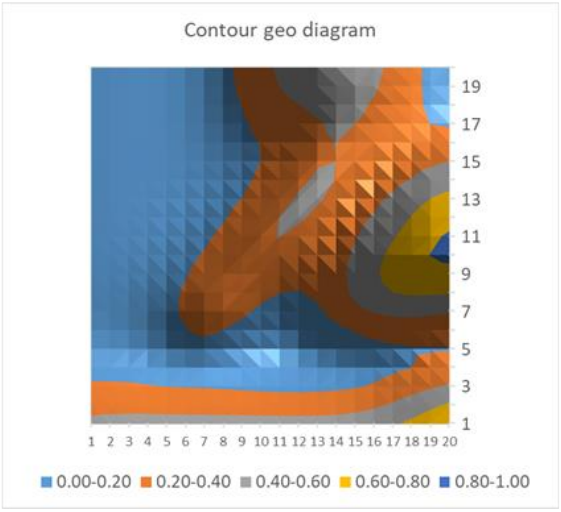

Fig. 7 – Colored contour diagram for the geographic distribution of the pollution rank.

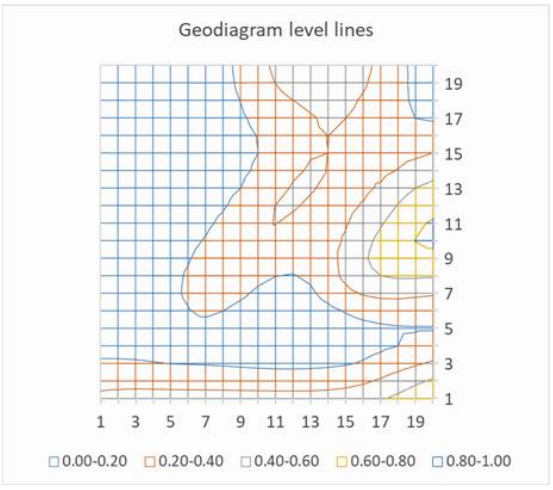

Fig. 8 – Contour diagram of level lines for the geographic distribution of the pollution rank.

The limitation of authors' approach is the need to measure sufficient data to train the neural network-based model. This will be mitigated in the future research by the derivation of neural network-based models considering architectures that work with small datasets.

ACKNOWLEDGEMENTS

This work was supported by the grant of the Romanian Ministry of Education and Research, CNCS - UEFISCDI, project number PN-III-P4-ID-PCE-2020-0269, within PNCDI III, and the national project of Republic of Moldova number 20.80009.5007.27.

REFERENCES

1. S. CHAKRABORTTY, G.T. PARATKAR, *Biomonitoring of trace element air pollution using mosses*, Aerosol and Air Quality Research, **6**, *3*, pp. 247–258, 2006.
2. K. SZCZEPANIAK, M. BIZIUK, *Aspects of the biomonitoring studies using mosses and lichens as indicators of metal pollution*, Environmental Research, **9**, *3*, pp. 221–230, 2003.

3. H. HARMENS, D.A. NORRIS, E. STEINNES, E. KUBIN, J. PIISPANEN, R. ALBER, Y. ALEKSIAYENAK, O. BLUM, M. COŞKUN, M. DAM, L. DE TEMMERMAN, J.A. FERNÁNDEZ, M. FROLOVA, M. FRONTASYEVA, L. GONZÁLEZ-MIQUEO, K. GRODZIŃSKA, Z. JERAN, S. KORZEKWA, M. KRMAR, K. KVIETKUS, S. LEBLOND, S. LIIV, S.H. MAGNÚSSON, B. MAŇKOVSKÁ, R. PESCH, A. RÜHLING, J. M. SANTAMARIA, W. SCHRÖDER, Z. SPIRIC, I. SUCHARA, L. THÖNI, V. URUMOV, L. YURUKOVA, H.G. ZECHMEISTER, *Mosses as biomonitors of atmospheric heavy metal deposition: Spatial patterns and temporal trends in Europe*, Environmental Pollution, **158**, *10*, pp. 3144–3156, 2010.
4. I. ZINICOVSCAIA, C. HRAMCO, O.G. DULIU, K. VERGEL, O.A. CULICOV, M.V. FRONTASYEVA, G. DUCA, *Air pollution study in the Republic of Moldova using moss biomonitoring technique*, Bulletin of Environmental Contamination and Toxicology, **98**, *2*, pp. 262–269, 2017.
5. I. ZINICOVSCAIA, C. HRAMCO, O. CHALIGAVA, N. YUSHIN, D. GROZDOV, K. VERGEL, G. DUCA, *Accumulation of potentially toxic elements in mosses collected in the Republic of Moldova*, Plants, **10**, *3*, art. 471, 2021.
6. G. KOSIOR, M. FRONTASYEVA, Z. ZIEMBIK, I. ZINCOVSCAIA, A. DOŁHAŃCZUK-ŚRÓDKA, B. GODZIK, *The moss biomonitoring method and neutron activation analysis in assessing pollution by trace elements in selected Polish national parks*, Archives of Environmental Contamination and Toxicology, **79**, *3*, pp. 310–320, 2020.
7. A. S. KRAKOVSKÁ, V. SVOZILÍK, I. ZINICOVSCAIA, K. VERGEL, P. JANČÍK, *Analysis of spatial data from moss biomonitoring in Czech-Polish border*, Atmosphere, **11**, *11*, p. 1237, 2020.
8. O. MOTYKA, I. PAVLÍKOVÁ, J. BITTA, M. FRONTASYEVA, P. JANČÍK, *Moss biomonitoring and air pollution modelling on a regional scale: delayed reflection of industrial pollution in moss in a heavily polluted region?*, **27**, *26*, pp. 32569–32578, 2020.
9. O. CHALIGAVA, S. SHETEKAURI, W.M. BADAWY, M.V. FRONTASYEVA, I. ZINICOVSCAIA, T. SHETEKAURI, A. KVLIVIDZE, K. VERGEL, N. YUSHIN, *Characterization of trace elements in atmospheric deposition studied by moss biomonitoring in Georgia*, Archives of Environmental Contamination and Toxicology volume, **80**, *20*, pp. 350–367, 2021.
10. I. ZINICOVSCAIA, O. CHALIGAVA, N. YUSHIN, D. GROZDOV, K. VERGEL, C. HRAMCO, *Moss biomonitoring of atmospheric trace element pollution in the Republic of Moldova*, Archives of Environmental Contamination and Toxicology, **82**, *3*, pp. 355–366, 2022.
11. E.-L. HEDREA, R.-E. PRECUP, R.-C. ROMAN, E.M. PETRIU, *Tensor product-based model transformation approach to tower crane systems modeling*, Asian Journal of Control, **23**, *3*, pp. 1313–1323, 2021.
12. I.-D. BORLEA, R.-E. PRECUP, A.-B. BORLEA, D. IERCAN, *A unified form of fuzzy C-means and K-means algorithms and its partitional implementation*, Knowledge-Based Systems, **214**, art. 106731, 2021.
13. I.-D. BORLEA, R.-E. PRECUP, A.-B. BORLEA, *Improvement of K-means cluster quality by post processing resulted clusters*, Procedia Computer Science, **199**, pp. 63–70, 2022.
14. R.-E. PRECUP, C.-A. BOJAN-DRAGOS, E.-L. HEDREA, R.-C. ROMAN, E.M. PETRIU, *Evolving fuzzy models of shape memory alloy wire actuators*, Romanian Journal of Information Science and Technology, **24**, *4*, pp. 353–365, 2021.
15. A.-I. SZEDLAK-STINEAN, R.-E. PRECUP, E.M. PETRIU, R.-C. ROMAN, E.-L. HEDREA, C.-A. BOJAN-DRAGOS, *Extended Kalman filter and Takagi-Sugeno fuzzy observer for a strip winding system*, Expert Systems with Applications, **208**, art. 118215, 2022.
16. R.-E. PRECUP, G. DUCA, S. TRAVIN, I. ZINICOVSCAIA, *Supplementary material of the paper R.-E. Precup, G. Duca, S. Travin, I. Zinicovscaia, "Processing, neural network-based modeling of biomonitoring studies data and validation on Republic of Moldova data"*, Proceedings of the Romanian Academy, Series A: Mathematics, Physics, Technical Sciences, Information Science, 2022, http://www.aut.upt.ro/~rprecup/Supplementary_material_PRA.doc, accessed October 12, 2022.
17. E.H. MOORE, *On the reciprocal of the general algebraic matrix*, Bulletin of the American Mathematical Society, **26**, *9*, pp. 394–395, 1920.
18. R. PENROSE, *A generalized inverse for matrices*, Mathematical Proceedings of the Cambridge Philosophical Society, **51**, *3*, pp. 406–413, 1955.
19. S.L. CAMPBELL, G.D. MEYER, Jr., *Generalized Inverses of Linear Transformations*, Dover, New York, 1991.
20. A. DINU, S. GHEORGHE, G.M. DANCIU, P.L. OGRUTAN, *Debugging FPGA projects using artificial intelligence*, Romanian Journal of Information Science and Technology, **24**, *3*, pp. 299–320, 2021.
21. M.-C. CRISTESCU, *Machine learning techniques for improving the performance metrics of functional verification*, Romanian Journal of Information Science and Technology, **24**, *1*, pp. 99–116, 2021.
22. J. ZHENG, J.-F. WANG, J. LI, S.-P. CHEN, L. SHU, Y.-K. PENG, *Deep neural networks for detection of abnormal trend in electricity data*, Proceedings of the Romanian Academy, Series A: Mathematics, Physics, Technical Sciences, Information Science, **22**, *3*, pp. 291–298, 2021.
23. I. A. ZAMFIRACHE, R.-E. PRECUP, R.-C. ROMAN, E.M. PETRIU, *Reinforcement learning-based control using Q-learning and gravitational search algorithm with experimental validation on a nonlinear servo system*, Information Sciences, **583**, pp. 99–120, 2022.
24. S. MCLEAN CABANEROS, B. HUGHES, *Methods used for handling and quantifying model uncertainty of artificial neural network models for air pollution forecasting*, Environmental Modelling & Software, **158**, art. 105529, 2022.
25. S. BLAŽIČ, D. DOVŽAN, I. ŠKRJANC, *Cloud-based identification of an evolving system with supervisory mechanisms*, Proceedings of 2014 IEEE International Symposium on Intelligent Control, France, 2014, pp. 1906–1911.
26. I. ŠKRJANC, S. BLAŽIČ, P. ANGELOV, *Robust evolving cloud-based PID control adjusted by gradient learning method*, Proceedings of 2014 IEEE Conference on Evolving and Adaptive Intelligent Systems, Linz, Austria, 2014, pp. 1–6.

Received October 13, 2022