International Journal of Phytoremediation

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/bijp20

An investigation on environmental pollution due to essential heavy metals: a prediction model through multilayer perceptrons

Murat Sari, Ibrahim Ertugrul Yalcin, Mahmut Taner, Tahir Cosgun & Ibrahim Ilker Ozyigit

To cite this article: Murat Sari, Ibrahim Ertugrul Yalcin, Mahmut Taner, Tahir Cosgun & Ibrahim Ilker Ozyigit (2023) An investigation on environmental pollution due to essential heavy metals: a prediction model through multilayer perceptrons, International Journal of Phytoremediation, 25:1, 89-97, DOI: 10.1080/15226514.2022.2059056

To link to this article: https://doi.org/10.1080/15226514.2022.2059056

Published online: 09 Apr 2022.

INTERNATIONAL JOURNAL OF PHYTOREMEDIATION
2023, VOL. 25, NO. 1, 89–97
https://doi.org/10.1080/15226514.2022.2059056

Taylor & Francis
Taylor & Francis Group

Check for updates

An investigation on environmental pollution due to essential heavy metals: a prediction model through multilayer perceptrons

Murat Sari[a], Ibrahim Ertugrul Yalcin[b], Mahmut Taner[c], Tahir Cosgun[a,d], and Ibrahim Ilker Ozyigit[e,f]

[a]Department of Mathematics, Faculty of Arts and Sciences, Yildiz Technical University, Istanbul, Turkey; [b]Department of Civil Engineering, Faculty of Engineering and Natural Sciences, Bahcesehir University, Istanbul, Turkey; [c]Department of Mathematics, Faculty of Engineering and Natural Sciences, Bahcesehir University, Istanbul, Turkey; [d]Department of Mathematics, Faculty of Arts and Sciences, Amasya University, Amasya, Turkey; [e]Department of Biology, Faculty of Arts and Sciences, Marmara University, Istanbul, Turkey; [f]Department of Biology, Faculty of Sciences, Kyrgyz-Turkish Manas University, Bishkek, Kyrgyzstan

ABSTRACT

This research is to predict heavy metal levels in plants, particularly in *Robinia pseudoacacia* L., and soils using an effective artificial intelligence approach with some ecological parameters, thereby significantly eliminating common defects such as high cost and seriously tedious and time-consuming laboratory procedures. In this respect, the artificial neural network (ANN) is employed to estimate the concentrations of essential heavy metals such as Fe, Mn and Ni, depending on the Cu and Zn concentrations of plant and soil samples collected from five different locations. The derived relative errors for the constructed ANN model have been computed within the ranges 0.041–0.051, 0.017–0.025, and 0.026–0.029 for the training, testing and holdout data regarding Fe, Mn, and Ni, respectively. In addition, it has been realized that the relative errors could be diminished up to 0.007 for Fe, 0.014 for Mn and 0.022 for Ni by considering the Cu, Zn, location and plant parts as independent variables during the analysis. The results produced seem instructive and pioneering for environmentalists and scientists to design optimal study programs to leave a livable ecosystem.

NOVELTY STATEMENT

The levels of essential heavy metals, Fe, Mn, Ni, based on Zn and Cu in plant and soil samples have been predicted through an AI-based prediction model, a class of feedforward artificial neural networks (ANNs) with a multilayer perceptron (MLP). Thereby common drawbacks such as high cost and severely time-consuming laboratory procedures have been significantly eradicated. In the evaluation of different pollution levels at locations, it has been shown that the ANN method can overcome several disadvantages of analytical element analyzers to monitor the amounts of heavy metals such as Fe, Mn, and Ni in soil and plants.

KEYWORDS
Artificial neural network; essential heavy metal; network algorithm; plant location; plant part; prediction model

Introduction

Analysis and evaluation of heavy metal content in plants and soil is vital for ecological life, as the diversity of these elements affects the abundance and richness of vegetation. In this context, very close to nature mathematical models describing environmental processes have been becoming one of the most important targets carefully followed by researchers and their stakeholders. Mathematical models are frequently employed to understand, analyze and overcome problems especially regarding the biological systems (Gill *et al.* 2012). Mathematical modeling of a problem that represents a special part of nature allows scientists to make accurate predictions without conducting biochemical experiments. Advances in computer systems and software engineering lead to the arise of new research areas such as "Artificial Neural Network Modeling". This approach can be applied to better understand many complex problems that arise in various fields of science (Sari and Cetiner 2009; Sari *et al.* 2021).

Since the type and degree of pollution directly affect living organisms, the construction of a mathematical model to make scientific predictions concerning environmental pollution should be considered to be necessarily a useful tool for observation, comprehension, and control in furtherance of authorities' decision-making processes. Predictions concerning pollution have a crucial role in environmental sciences since they enable the academic community make significant estimations and hereby the decision-making authorities give environmentally friendly arrangements (Crouse *et al.* 2009; Peng *et al.* 2016; Cabaneros *et al.* 2017). The relationship between the metabolism of living systems and environmental pollution type and also its level might be well established

CONTACT Murat Sari sarim@yildiz.edu.tr Department of Mathematics, Faculty of Arts and Sciences, Yildiz Technical University, Istanbul, Turkey.

using mathematical models (Jaskulak *et al.* 2020; Sen *et al.* 2020).

As being one of the most important factors affecting living organisms, heavy metal pollution has a special importance in environmental pollution (Turan *et al.* 2020). However, while some heavy metals (Pb, Cr, Cd, and Hg) have a direct toxic effect, some other heavy metals such as copper (Cu), iron (Fe), manganese (Mn), nickel (Ni), and zinc (Zn) are considered as essential elements (Jothimani *et al.* 2017). Because of their beneficial effects on plant growth and development, the amounts of these essential heavy metals in the plant body are of vital importance. For this reason, examinations to monitor the amount of these elements in plants and soil are carried out in field studies (Ozyigit *et al.* 2018; Yalcin *et al.* 2020; Fang *et al.* 2021; Jeddi *et al.* 2021). Reduced amounts of these essential heavy metals can lead to deficiency symptoms, while excessive levels can cause toxicity. Therefore, a strict balance in plants is necessary to avoid health consequences (Ozyigit *et al.* 2018).

Although there are remarkable studies investigating the mathematical modeling of environmental pollution dynamics (Serbula *et al.* 2012; Cristaldi *et al.* 2020; Diaconu *et al.* 2020; Likus-Cieslik *et al.* 2020; Pulscher *et al.* 2020), to our knowledge, no study on predicting the degree and dynamics of environmental pollution with artificial intelligence approaches has been found yet as in this study. Albeit there are valuable studies conducted on determination of Cr pollution, it has not been demonstrated clearly via artificial intelligence approaches.

As mentioned in the literature, environmental pollution in various ecosystems is a great concern and has been studied with different approaches so far. However, as a traditional method, experimental analysis of pollution usually requires high costs and investment of labor. In the literature, optical emission spectroscopy (Yilmaz *et al.* 2015; Ozyigit *et al.* 2022), biosensors (Guo *et al.* 2018), sensor networks (Luo and Yang 2019), mass spectrometry (Matsui *et al.* 2020), nanostructures (Cui *et al.* 2020) etc. are used to evaluate pollution. Besides their benefits, many of these techniques are disadvantageous because of being expensive or time-consuming or not practical in use (Sari *et al.* 2021; Arshad *et al.* 2022; Sohrabi *et al.* 2022; Zhao *et al.* 2022). Thus, the mathematical modeling can be proposed as an alternative method for the scientific prediction of heavy metal pollution. The ANNs, a quite flexible tool and free of strict assumptions, has a growing interest in the academic society to obtain more accurate responses of environmental problems like heavy metal pollution. This study pursuing the examination regarding the excess amount of essential metals like Fe, Mn, or Ni (based on the locations of the plants, the parts of the plants, and the other heavy metals such as Zn, and Cu) is to appear for the first time in the literature.

The ANNs mimicking a human brain offers higher reliability concerning learning from data day by day. When performing this, the ANNs are learning from a given real dataset called the training set, and thus, they reveal the hidden relationships between the input and output data. The most prominent objective of this research is the use of locations and plant parts in an artificial intelligence-based prediction model to estimate the heavy metal levels such as Fe, Mn, Ni depending upon Zn and Cu levels. The performance of the created ANNs is measured by using a different real dataset called the holdout data. The holdout data has been utilized only for comparison purposes during the investigation. In this respect, the ANNs have been employed to forecast Fe, Mn, and Ni levels of different plant parts based on the location of the plants and the pre-measured Zn and Cu levels in the current study. Since the actual values of Fe, Mn, or Ni will not be required, once the MLP has been trained, the needs for time and labor-power are going to decrease considerably in case of the implementation of the mentioned biochemical procedures once more. Due to the required preliminary preparations for spectroscopy methods, an unneglectable decrease is expected in the need for chemicals, although it is not at the same level with time and labor-power. An additional reduction in the required technical maintenance of the devices used can also be foreseen. Unmistakably, the importance of such predictions would be appreciated better when the concepts of sustainable, smart, and precision agriculture are taken into consideration.

Materials and methods

Data structure

Robinia pseudoacacia L. (black locust), is a medium-sized deciduous tree in the Fabaceae family. It is of high economic and ecological value but is also has a highly invasive nature. Therefore, it is distributed worldwide, from Australia to Asia, and Turkey (Martin 2019). It is generally planted for esthetic purposes in urban areas such as parks, gardens and roadsides in many countries of the world as well as in Istanbul and Kocaeli. It has also been planting in industrial areas that are responsible from heavy metal spread to the environment, as it is a good biomonitor (Tzvetkova and Petkova 2015).

In this study; soil and plant samples are gathered from four different stations in Istanbul and one station in Kocaeli province. All of these five stations are categorized by pollution intensity. The first group of samples are collected from the Prince Islands where there is no pollution due to the traffic or industry. The 2nd, 3rd and 4th stations in various regions of Istanbul are determined according to the traffic density. Dilovasi, a heavy industrial zone and one of the districts of Kocaeli province, is determined as the 5th station.

Washed leaf, unwashed leaf, stem, and bark samples of plant, along with soil samples were taken from related locations. The mineralization of plant and soil samples was performed according to EPA 3051 A Analytical Method for ICP-OES using Berghof-MWS2 microwave. Elemental concentrations in plant and soil samples were determined using calibration standards prepared by diluting 1,000 mg L^{-1} ICP multi-element standard solution (Merck) (Ozyigit *et al.* 2022). The calibration curves with high accuracy ($R^2 > 0.999$) were created using the data obtained from the analysis of the calibration standards prepared at 8 different concentrations for each element. The Fe, Mn, and Ni concentrations were figured out utilizing inductively coupled plasma optical emission spectroscopy, PerkinElmer-Optima 7000DV. Eight hundred data from Istanbul and Kocaeli provinces of Turkey were included in this study. In the collected data of 800 samples; the location, plant part (including soil), Zn, and Cu values (mg kg^{-1}) are accepted as inputs, while Fe, Mn, and Ni heavy metal values (mg kg^{-1}) are considered as output.

Artificial neural network

Throughout the study, the Multilayer Perceptron (MLP) Module of IBM SPSS Statistics 26 is exploited to build the required network model and verify its accuracy. The constructed ANN is trained via a back-propagation learning algorithm, and the gradient descent method is employed for minimization of the root mean square error (RMSE) by updating the weights.

The collected data of 800 samples are randomly distributed into training (70%), testing (20%), and holdout (10%) subsets as in Table 1. These datasets are utilized to optimize the weights and construct the model, to compute the relative errors and RMSE besides to prevent the overfitting, to validate the model, respectively. All the covariates, location, plant parts, Zn, and Cu levels, are standardized before the training via adjusted normalization. The MLP setup could be summarized as follows:

The number of hidden layer is chosen as 1, and the number of units therein is chosen to be 3 as in Figure 1 and Table 2. The hyperbolic tangent function is employed as an activation function both for the hidden and output layers. The stopping rules are imposed as 10 consecutive steps with no decrease in the error, 15 minutes for maximum training time, 100 maximum number of epochs, 0.0001 minimum relative change in training error, and lastly 0.001 minimum relative change in training error ratio (Table 3).

The gradient descent algorithm can achieve the optimum solution with two different implementations: batch or online. The online algorithm updates the synaptic weights after each single training data record; in other words, it utilizes information from one record at one time. Online training continuously updates the weights until one of the stopping rules is met. Whenever all the records are utilized once and none of the stopping criteria is fulfilled, then the process is reiterated over the data records. Online training has superiorities over batch mode for relatively larger datasets with associated predictors; that is if there are a large number of records and too many inputs, and their values are not independent of each other, then online training could attain a plausible answer than batch training within a remarkably less amount of time. Since the number of collected samples is 800, the batch mode is utilized to obtain the results of this paper.

Results and discussion

This study aims to examine the Fe, Mn, or Ni levels based on the plant parts, plant location, Zn level, and Cu level. In the literature, the possible effects of heavy metals such as Cd, Co, Cr, Cu, Fe, Mn, Ni, Pb, Zn on environmental pollution have been on the agenda of researchers. According to the results obtained by the ICP-OES analysis, heavy metals

Table 1. Case processing summary.

	Fe		Mn		Ni	
	N	Percent	*N*	Percent (%)	*N*	Percent (%)
Sample						
Training	548	68.5%	561	70.1	540	67.5
Testing	174	21.8%	156	19.5	172	21.5
Holdout	78	9.8%	83	10.4	88	11.0
Valid	800	100.0%	800	100.0	800	100.0
Excluded	0		0		0	
Total	800		800		800	

Table 2. Network information.

Input layer	Covariates	1	Plant parts
		2	Location
		3	Zn
		4	Cu
	Number of units[a]		4
	Rescaling method for covariates		Standardized
Hidden layer(s)	Number of hidden layers		1
	Number of units in hidden layer[a]		3
	Activation function		Hyperbolic tangent
Output layer	Dependent variables	1	Fe, Mn, or Ni
	Number of units		1
	Rescaling method for scale dependents		Adjusted normalized
	Activation function		Hyperbolic tangent
	Error function		Sum of squares

[a]Excluding the bias unit.

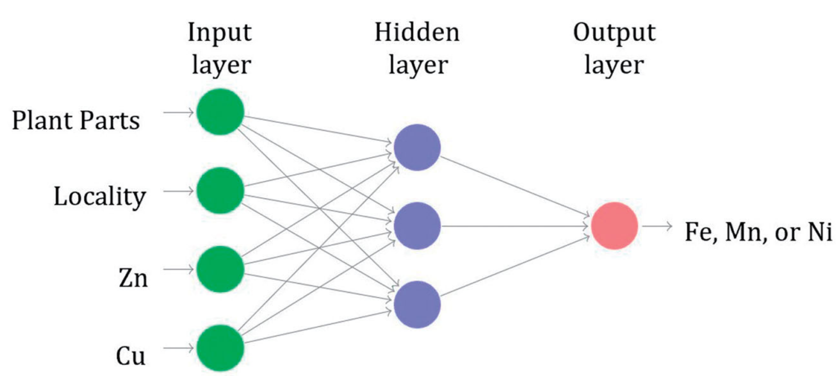

Figure 1. The standard architecture of the network diagram.

Table 3. Model summary.

	Fe	Mn	Ni
Training			
Sum of squares error	3.518	1.014	1.304
RMSE	0.080	0.043	0.049
Relative error	0.044	0.017	0.028
Stopping rule used	10 Consecutive step(s) with no decrease in error	Maximum number of epochs (100) exceeded	10 Consecutive step(s) with no decrease in error
Training time	0:00:00.05	0:00:00.06	0:00:00.06
Testing			
Sum of squares error	1.283	0.300	0.352
RMSE	0.086	0.044	0.045
Relative error	0.051	0.022	0.029
Holdout			
Relative error	0.041	0.025	0.026

Dependent Variable: Fe, Mn, or Ni.

such as Cd, Co, Cr, Mn, Ni and Pb were found in vital concentrations in the leaf and bark samples of Turkish red pine (Yalcin *et al.* 2020). Again, it was recorded by Alani *et al.* (2020) that severe heavy metal pollution (Cd, Pb, Zn and Cr) was encountered in edible vegetables collected from agricultural areas. Observation of excessive heavy metals, especially Cr, in the soil due to the wastewater near the tanneries does not go unnoticed (Sinduja *et al.* 2022). Previously, Golui *et al.* (2021) discussed the same issue in terms of the fractionation of Zn, Cu, Ni, Pb and Cd compounds in various samples from polluted river water, industrial wastes, municipal solid wastes and sewage sludges. It has been remarked that especially leafy vegetables and fruits exceed the maximum acceptable heavy metal levels recommended by the FAO/WHO (Moyo *et al.* 2020). According to their results, it is seen that especially Cd and Ni levels pose a possible threat to human health. As a result of those previous studies, it can be concluded that excessive amounts of heavy metals pose *a priori* risk for both the environment and human health. In this context, this paper develops an extremely fast and cost-effective MLP-based prediction model for monitoring the concentration of essential heavy metals in soil and various parts of plants. The most outstanding aspect of the study is the determination of heavy metals included in various parts of a plant such as the washed leaf, unwashed leaf, branch, body, and soil particularly depending on the location of the plant. Especially, the data collection strategy deserves a seperate interest. The locations including the control group are categorized with respect to their pollution level, and the level of pollution is enumerated from 1 to 5 for each location. Then, the collected samples are analyzed to reach the necessary data by using the previously mentioned conventional methods. The received data is divided into three groups (for training, testing, and holdout) and is structured to train the previously constructed ANNs in Figure 1. Throughout this study, the holdout data have been used for comparison purposes only. Hence, the comparisons have been performed with the actual data which are obtained from the conventional biochemical analysis, namely laboratory-based elemental analyses. Thereby, due to the inherence of the proposed method, the results produced have been compared with the real data.

The qualitative results in Figures 2 and 3 reveal that the ANNs could predict the actual values of Fe, Mn, or Ni almost perfectly. The predicted values versus actual values, depicted in Figure 2, show that there is an accumulation near the perfect match line for each of the three heavy metals. Moreover, the computed residuals are seen to be in good agreement with this inference. Strikingly, the quantitative results indicate that the accuracy of the current method is between 94.9 and 98.3% when the relative errors in Table 3 are considered. The level of accuracy for the present method is seen to be in an encouraging and highly acceptable interval. The following implications could be drawn from the results in Tables 3–5.

Regarding the prediction of Fe level, the relative errors are 4.4, 5.1, and 4.1% respectively for training, testing and holdout data. The RMSE is obtained as 0.080 and 0.086, consecutively, for the training and testing data. The algorithm is seen to stop due to 10 consecutive steps with no decrease in the error. In addition, the sum of squares error is calculated as 3.518 that is a reasonable value when it is compared with the actual Fe values (between 62.282 and 3865.672). Independent variable importance analysis reveals that the location of the plant is as important as the levels of Zn or Cu (Table 4). On the other hand, plant parts, in which some of the samples are gathered, has relatively less importance concerning the Fe-level prediction.

Secondly, the relative errors in the prediction of Mn level are seen to be 1.7, 2.2, and 2.5% for training, testing, and holdout data, correspondingly. Although the algorithm stopped when the maximum number of epochs (100) exceeded, remarkably, the sum of squares error is found to be 1.014 for the training sample. Besides, the RMSE is obtained as 0.043 and 0.044, respectively, for the training and testing data. The parameter estimation has uncovered a strong relation between Mn level and Zn level, that is to say, the level of Zn is as important as the sum of the importances of plant parts, location, and Cu level in the prediction of Mn level.

Lastly, the relative errors are computed as 2.8, 2.9, and 2.6% subsequently for the training, testing, and holdout data during the prediction of Ni levels. The stopping criterion satisfied to halt the algorithm is 10 consecutive steps with no decrease in the error. Notably, the sum of squares error

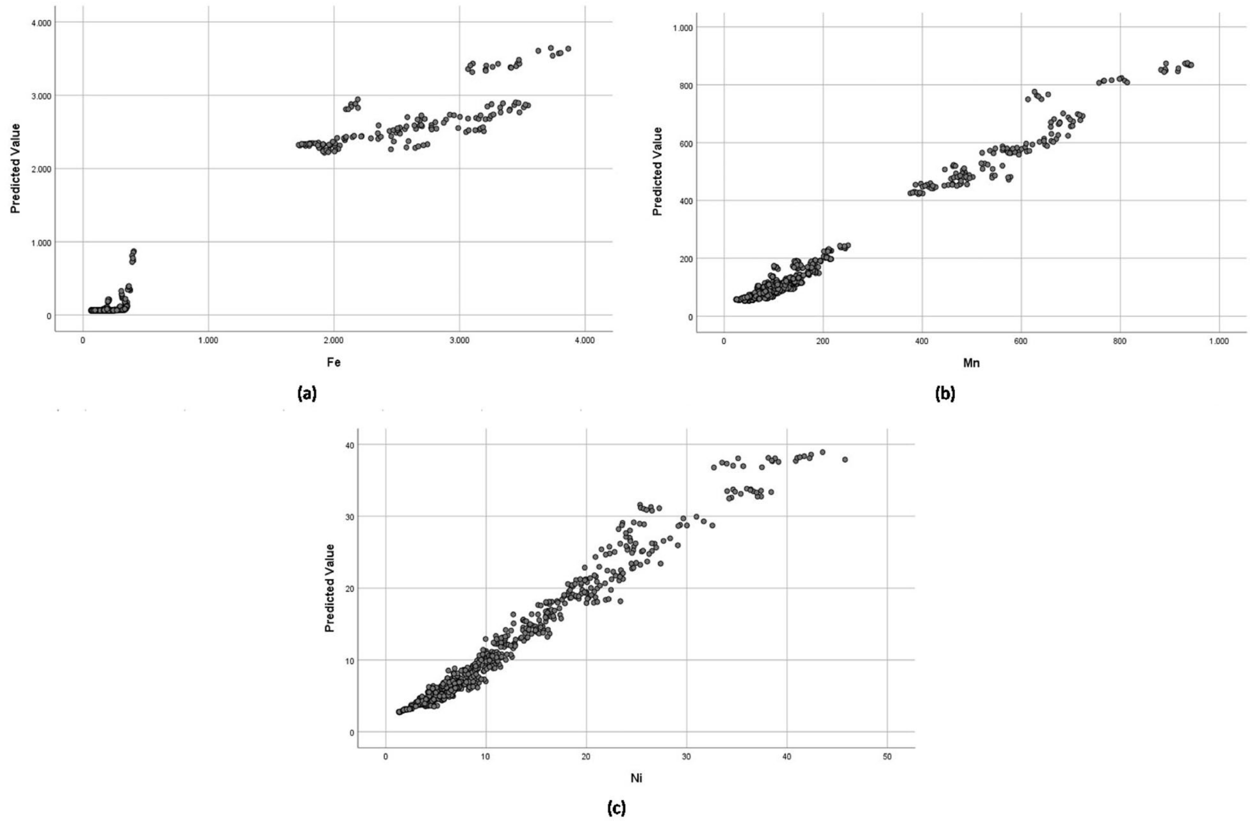

Figure 2. Predicted vs actual Fe (a), Mn (b), Ni (c) amounts in mg kg^{-1}.

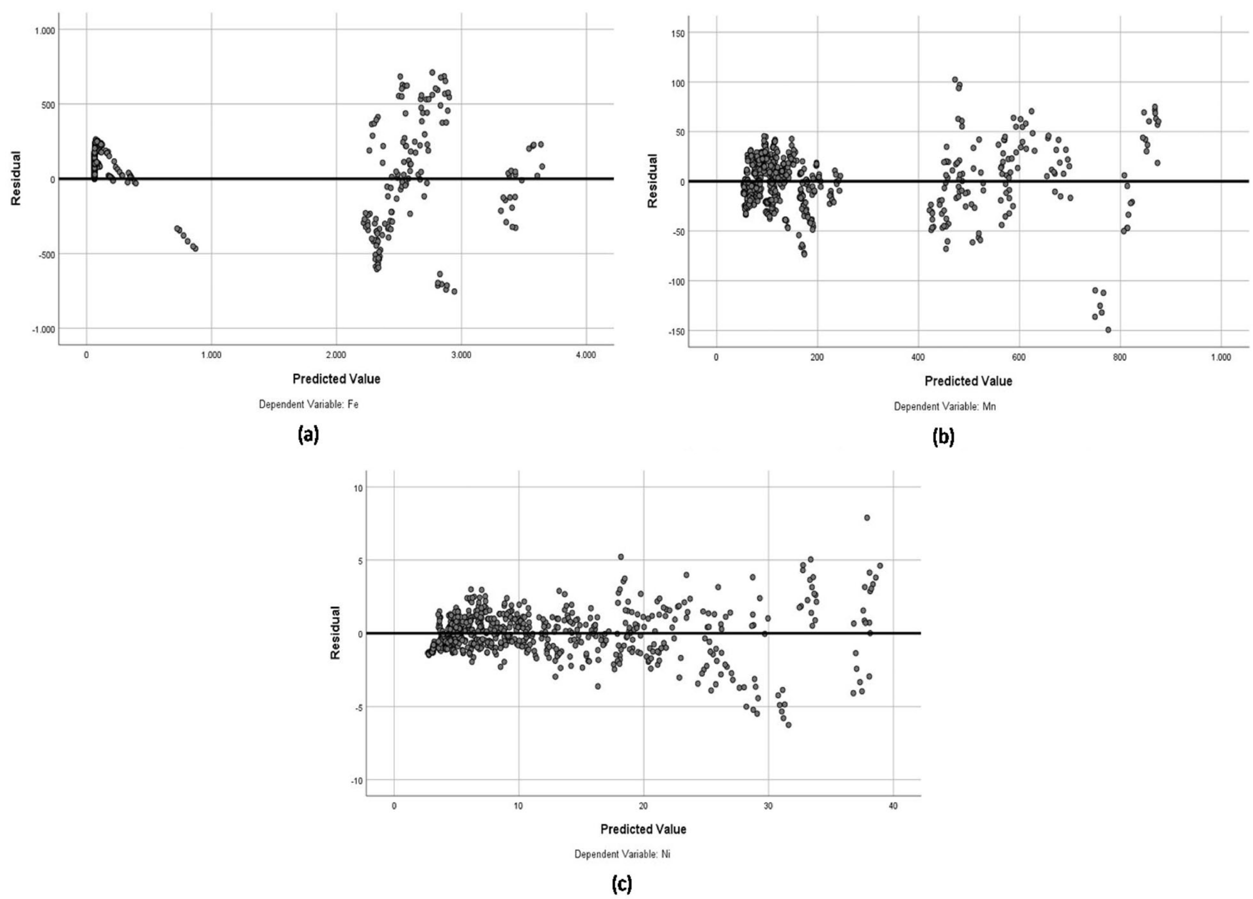

Figure 3. Residual values in the prediction of Fe (a), Mn (b), Ni (c) amounts in mg kg^{-1}.

Table 4. Independent variable importance.

	Fe		Mn		Ni	
	Importance	Normalized importance (%)	Importance	Normalized importance (%)	Importance	Normalized importance (%)
Plant parts	0.042	11.5	0.049	10.1	0.143	13.2
Location	0.271	73.7	0.171	34.9	0.098	19.2
Zn	0.320	87.0	0.491	100.0	0.742	100.0
Cu	0.367	100.0	0.288	58.7	0.017	2.3

Table 5. Parameter estimates for Fe, Mn, and Ni.

	Predicted			
	Hidden layer			Output layer
Predictor				
Input layer				
(Bias)	−0.783	−0.172	0.125	
Location	0.740	−0.026	1.382	
Plant parts	−0.178	−0.047	0.244	
Zn	0.612	−0.216	1.595	
Cu	0.192	0.189	0.463	
Hidden layer				
(Bias)				−1.361
H(1:1)				1.854
H(1:2)				2.083
H(1:3)				1.178
Predictor				
Input layer				
(Bias)	−0.279	0.080	0.298	
Location	0.362	0.015	−0.506	
Plant parts	0.348	0.038	0.099	
Zn	0.105	−0.192	−0.876	
Cu	−0.396	0.566	0.133	
Hidden layer				
(Bias)				−0.620
H(1:1)				−0.260
H(1:2)				0.764
H(1:3)				−1.110
Predictor				
Input layer				
(Bias)	−0.279	0.080	0.298	
Location	0.362	0.015	−0.506	
Plant parts	0.348	0.038	0.099	
Zn	0.105	−0.192	−0.876	
Cu	−0.396	0.566	0.133	
Hidden layer				
(Bias)				−0.620
H(1:1)				−0.260
H(1:2)				0.764
H(1:3)				−1.110

is computed to be 1.304 to forecast the Ni level. In addition, the RMSE is obtained as 0.049 and 0.045, respectively, for the training and testing data. The investigation has displayed a powerful relationship between Ni level and Zn level.

Apart from the emphasized results above, if the plant parts and location are considered as factor variables instead of independent variables during the analysis then the relative errors, in Table 6, could be diminished to values between 0.07 and 2.90%.

As far as the method of the paper is concerned, an increasing interest has been observed in the literature on multi-layer perceptron (MLP) so far. The MLP was successfully used in research to predict heavy metal levels in soil based on various parameters, especially altitude (Sari *et al.* 2021). This architecture was also utilized to identify the type and level of pollution (Denisov 2017). In addition to all these, the construction of an MLP-based model regarding the spatial distribution of heavy metals such as Cr and Mn in the topsoil is not far from attention (Baglaeva *et al.* 2021). In their studies, by partitioning the raw data, it was tried to discover the training dataset that minimizes the RMSE, which is the most appropriate training dataset. In addition to the multiple linear regression (MLR) model, we again come across the MLP model in predicting the sulfur dioxide concentration in the air separately (Shams *et al.* 2021). Again, the results were compared with the real data and the MLP model seems to be more accurate than the regression model, as the R2 and RMSE values are quite reasonable and convincing. Previously, a different ANN architecture, radial basis functions (RBF), together with the MLP was utilized to monitor green macroalgae population in coastal areas by tracking the heavy metal concentrations such as Cd, Cu, Ni, Zn, Mn, Pb, Na, Ca, K, and Mg (Zbikowski 2011). The same models, the MLP and RBF, were also effectively used by Falamaki (2013) to predict the soil distribution coefficient of nickel. In the corresponding

Table 6. Model summary for the case that plant part and location are factor variables.

Training	Sum of squares error	3.318
	Average overall relative error	0.018
	Relative error for scale dependents	
	Fe	0.012
	Mn	0.016
	Ni	0.029
	Stopping rule used	10 Consecutive step(s) with no decrease in error[a]
	Training time	0:00:00.06
Testing	Sum of squares error	1.005
	Average overall relative error	0.017
	Relative error for scale dependents	
	Fe	0.011
	Mn	0.014
	Ni	0.029
Holdout	Average overall relative error	0.014
	Relative error for scale dependents	
	Fe	0.007
	Mn	0.014
	Ni	0.022

[a]Error computations are based on the testing sample.

work, the advantages of the ANN were observed once again and also stated that using 80% or 90% of the data, for training purposes, produces a more effective neural network model than the use of 100% of the collected data. To sum up, this and similar prediction models have been applied in different regions, in various fields of science and even for different elements, and it has been seen that they continue to produce alternatives or guide scientists. In this manner, the successful and effective application of artificial intelligence methods to a wide range of science has led us to an attempt to predict the essential heavy metals Fe, Mn, and Ni.

One of the unmissable points to note when closing this section is that the required amount of time to predict the necessary values of dependent variables throughout the computation is approximately 0.06 seconds as could be figured out from Tables 3 and 6. This is one of the most distinguished aspects of modeling with the aid of the ANNs. Thereby, the ANNs enable us also to reduce both the required amount of labor and the CPU time drastically as compared to the conventional and chemical methods.

Conclusions and recommendations

In this study, efficient artificial neural network models with an innovative perspective and an ecological importance have been produced to forecast the heavy metals levels, Fe, Mn, and Ni, in plants and soil utilizing estimate parameters such as location, plant parts, and concentrations of Zn and Cu. It is believed that the produced network models (MLPs) can facilitate to discover the essential heavy metal concentrations in plants and soils, evaluating economic and environmental impacts in various local circumstances, and supporting decision-making regarding the improvement of required environmental policies. In that respect, the metal levels have been scrutinized depending on plant and soil information gathered from different pollution levels. It has been concluded that the ANN method, unlike traditional approaches, can forecast the heavy metal levels in plants and soils through plant parts and various soil parameters in a very short time with a high rate of accuracy. It has been realized that this paper produces remarkably informative and illuminating results about the ecosystem of interest, with a unique, innovative and effective approach to the use of artificial intelligence techniques. Additionally, these techniques have been applied to obtain numerical data for more elements with the least possible number of laboratory analyses. When abiotic and biotic factors affect the elements to be analyzed, the concentrations of the elements to be predicted will be updated directly in line with these effects. The simulation results show that the ANN algorithm has significant advantages over traditional methods. It has been deduced that these findings are quite useful to scientists who deal with all aspects of environmental sciences or who prepare an appropriate research plan on this subject. Undoubtedly, in daily life, the optimal appraisal of heavy metal status regarding the understanding the behavior of significant realistic problems including a wide range of disciplines, is vitally important for living organisms, for agricultural economies and, even for the decision-making authorities. Even though this study has been accomplished to estimate the heavy metal levels in plants and soil at different pollution levels, it is considered that the applicability of the same study for any ecosystem in a very comfortable way increases the significance of this study further. In further reserach, such studies could be re-implemented, as like for the element status determined in plant and soil in this paper, only for plant or only for soil element status or for the wastewater element contents.

Ethical approval

This manuscript did not involve human or animal participants; therefore, informed consent was not collected.

Acknowledgment

Authors are thankful to Dr. M. E. Uras (Marmara University) for providing the data.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Murat Sari http://orcid.org/0000-0003-0508-2917
Ibrahim Ertugrul Yalcin http://orcid.org/0000-0003-3140-7922
Mahmut Taner http://orcid.org/0000-0002-2838-3651
Tahir Cosgun http://orcid.org/0000-0003-2970-0863
Ibrahim Ilker Ozyigit http://orcid.org/0000-0002-0825-5951

References

Alani RA, Nwude D, Abati A, Akinrinade OE. 2020. Heavy metal contamination of vegetables planted in Lagos soils and their potential pollution risks. Ife J Agricult. 32(3):12–30.

Arshad H, Saleem M, Pasha U, Sadaf S. 2022. Synthesis of *Aloe vera*-conjugated silver nanoparticles for use against multidrug-resistant microorganisms. Electron J Biotechnol. 55:55–64. doi:10.1016/j.ejbt.2021.11.003.

Baglaeva EM, Sergeev AP, Shichkin AV, Buevich AG. 2021. The extraction of the training subset for the spatial distribution modelling of the heavy metals in topsoil. CATENA. 207:105699. doi:10.1016/j.catena.2021.105699.

Cabaneros SMS, Calautit JKS, Hughes BR. 2017. Hybrid artificial neural network models for effective prediction and mitigation of urban roadside NO2 pollution. Energy Procedia. 142:3524–3530. doi:10.1016/j.egypro.2017.12.240.

Cristaldi A, Oliveri Conti G, Cosentino SL, Mauromicale G, Copat C, Grasso A, Zuccarello P, Fiore M, Restuccia C, Ferrante M. 2020. Phytoremediation potential of *Arundo donax* (Giant Reed) in contaminated soil by heavy metals. Environ Res. 185:109427. doi:10.1016/j.envres.2020.109427.

Crouse DL, Goldberg MS, Ross NA. 2009. A prediction-based approach to modelling temporal and spatial variability of traffic-related air pollution in Montreal, Canada. Atmos Environ. 43(32):5075–5084. doi:10.1016/j.atmosenv.2009.06.040.

Cui KP, Dai RR, Liu X, Weerasooriya R, Hong ZY, Chen X, Wu YC. 2020. New strategy for fabricating Cd (II) sensing electrochemical interface based on enhanced adsorption followed by redox processes: Ferro-cerium oxide nanocomposite as an example. J Alloys Compd. 829:154551. doi:10.1016/j.jallcom.2020.154551.

Denisov IA. 2017. Luciferase-based bioassay for rapid pollutants detection and classification by means of multilayer artificial neural networks. Sens Actuators, B. 242:653–657. doi:10.1016/j.snb.2016.11.071.

Diaconu M, Pavel LV, Hlihor R-M, Rosca M, Fertu DI, Lenz M, Corvini PX, Gavrilescu M. 2020. Characterization of heavy metal toxicity in some plants and microorganisms–a preliminary approach for environmental bioremediation. N Biotechnol. 56:130–139. doi:10.1016/j.nbt.2020.01.003.

Falamaki A. 2013. Artificial neural network application for predicting soil distribution coefficient of nickel. J Environ Radioact. 115:6–12. doi:10.1016/j.jenvrad.2012.06.008.

Fang T, Jiang T, Yang K, Li J, Liang Y, Zhao X, Cui K. 2021. Biomonitoring of heavy metal contamination with roadside trees from metropolitan area of Hefei. China. Environ Monitor Assess. 193(3):1–14.

Gill M, McKeever S, Gavaghan D. 2012. Modular mathematical modelling of biological systems. In: Proceedings of the 2012 Symposium on Theory of Modeling and Simulation-DEVS Integrative M&S Symposium. Orlando Florida. Vol. 36. p. 1–8.

Golui D, Datta SP, Dwivedi BS, Meena MC, Trivedi VK, Jaggi S, Bandyopadhyay KK. 2021. Assessing geoavailability of zinc, copper, nickel, lead and cadmium in polluted soils using short sequential extraction scheme, soil and sediment contamination. Soil Sediment Contamin. 30(1):74–91.

Guo Y, Liu R, Liu Y, Xiang D, Liu Y, Gui W, Li M, Zhu G. 2018. A non-competitive surface plasmon resonance immunosensor for rapid detection of triazophos residue in environmental and agricultural samples. Sci Total Environ. 613–614:783–791.

Jaskulak M, Grobelak A, Vandenbulcke F. 2020. Modelling assisted phytoremediation of soils contaminated with heavy metals–main opportunities, limitations, decision making and future prospects. Chemosphere. 249:126196.

Jeddi K, Fatnassi M, Chaieb M, Siddique KH. 2021. Tree species as a biomonitor of metal pollution in arid Mediterranean environments: case for arid southern Tunisia. Environ Sci Pollut Res Int. 28(22):28598–28605.

Jothimani K, Arulbalachandran D, Yasmin K. 2017. Amelioration of environmental stress for sustainable crop productivity. In: Dhanarajan A, editor. Sustainable agriculture towards food security. Singapore: Springer. p. 327–348.

Likus-Cieślik J, Socha J, Gruba P, Pietrzykowski M. 2020. The current state of environmental pollution with sulfur dioxide (SO_2) in Poland based on sulfur concentration in Scots pine needles. Environ Pollut. 258:113559. doi:10.1016/j.envpol.2019.113559.

Luo X, Yang J. 2019. A survey on pollution monitoring using sensor networks in environment protection. J Sens. 2019:1–11. doi:10.1155/2019/6271206.

Martin GD. 2019. Addressing geographical bias: a review of *Robinia pseudoacacia* (black locust) in the Southern Hemisphere. S Afr J Bot. 125:481–492. doi:10.1016/j.sajb.2019.08.014.

Matsui K, Ishimura T, Mattonai M, Iwai I, Watanabe A, Teramae N, Ohtani H, Watanabe C. 2020. Identification algorithm for polymer mixtures based on Py-GC/MS and its application for microplastic analysis in environmental samples. J Anal Appl Pyrolysis. 149:104834. doi:10.1016/j.jaap.2020.104834.

Moyo B, Matodzi V, Legodi MA, Pakade VE, Tavengwa NT. 2020. Determination of Cd, Mn and Ni accumulated in fruits, vegetables and soil in the Thohoyandou town area, South Africa. Water SA. 46(2):285–290. doi:10.17159/wsa/2020.v46.i2.8244.

Ozyigit II, Karahan F, Yalcin IE, Hocaoglu-Ozyigit A, Ilcim A. 2022. Heavy metals and trace elements detected in the leaves of medicinal plants collected in the southeast part of Turkey. Arab J Geosci. 15(1):1–21. doi:10.1007/s12517-021-09264-9.

Ozyigit II, Yalcin B, Turan S, Saracoglu IA, Karadeniz S, Yalcin IE, Demir G. 2018. Investigation of heavy metal level and mineral nutrient status in widely used medicinal plants' leaves in Turkey: Insights into health implications. Biol Trace Elem Res. 182(2):387–406.

Peng C, Wang M, Chen W. 2016. Modelling cadmium contamination in paddy soils under long-term remediation measures: Model development and stochastic simulations. Environ Pollut. 216:146–155. doi:10.1016/j.envpol.2016.05.038.

Pulscher LA, Gray R, McQuilty R, Rose K, Welbergen J, Phalen DN. 2020. Investigation into the utility of flying foxes as bioindicators for environmental metal pollution reveals evidence of diminished lead but significant cadmium exposure. Chemosphere. 254:126839. doi:10.1016/j.chemosphere.2020.126839.

Sari M, Cetiner BG. 2009. Predicting effect of physical factors on tibial motion using artificial neural networks. Expert Syst Appl. 36(6):9743–9746. doi:10.1016/j.eswa.2009.02.030.

Sari M, Cosgun T, Yalcin IE, Taner M, Ozyigit II. 2021. Deciding heavy metal levels in soil based on various ecological information through artificial intelligence modeling. Appl Artif Intelligence. doi:10.1080/08839514.2021.2014189.

Sen S, Dutta A, Ponnala R, Kamila B, Baltrėnas P, Baltrėnaitė E, Dutta S. 2020. Removal of hexavalent chromium from synthetic wastewater using alginate immobilized cyanobacteria: experiment and mathematical modeling. Environ Eng Sci. 37(4):283–294. doi:10.1089/ees.2019.0035.

Serbula SM, Miljkovic DD, Kovacevic RM, Ilic AA. 2012. Assessment of airborne heavy metal pollution using plant parts and topsoil. Ecotoxicol Environ Saf. 76(2):209–214. doi:10.1016/j.ecoenv.2011.10.009.

Shams SR, Jahani A, Kalantary S, Moeinaddini M, Khorasani N. 2021. The evaluation on artificial neural networks (ANN) and multiple linear regressions (MLR) models for predicting SO2 concentration. Urban Clim. 37:100837. doi:10.1016/j.uclim.2021.100837.

Sinduja M, Sathya V, Maheswari M, Dhevagi P, Kalpana P, Dinesh GK, Prasad S. 2022. Evaluation and speciation of heavy metals in the soil of the sub urban region of southern India, soil and sediment contamination. Int J. 1–21. doi:10.1080/15320383.2022.2030298.

Sohrabi H, Majidi MR, Arbabzadeh O, Khaaki P, Pourmohammad S, Khataee A, Orooji Y. 2022. Recent advances in the highly sensitive determination of zearalenone residues in water and environmental resources with electrochemical biosensors. Environ Res. 204(Pt B): 112082. doi:10.1016/j.envres.2021.112082.

Turan O, Ozdemir H, Demir G. 2020. Deposition of heavy metals on coniferous tree leaves and soils near heavy urban traffic. Front Life Sci RT. 1(1):35–41.

Tzvetkova N, Petkova K. 2015. Bioaccumulation of heavy metals by the leaves of *Robinia pseudoacacia* as a bioindicator tree in industrial zones. J Environ Biol. 36(1):59.

Yalcin IE, Ozyigit II, Dogan I, Demir G, Yarci C. 2020. Using the Turkish red pine tree to monitor heavy metal pollution. Pol J Environ Stud. 29(5):3881–3889. doi:10.15244/pjoes/114505.

Yang Y, Zhang H, Wang S, Yang W, Ding J, Zhang Y. 2020. Variation in sperm morphology and performance in tree sparrow (*Passer montanus*) under long-term environmental heavy metal pollution. Ecotoxicol Environ Saf. 197:110622. doi:10.1016/j.ecoenv.2020.110622.

Yilmaz N, Ozyigit II, Demir G, Yalcin IE. 2015. Determination of phytoplankton density, and study of the variation of nutrients and heavy metals in the surface water of Riva Stream; one of the water sources of Istanbul, Turkey. Desalin Water Treat. 55(3):810–820. doi:10.1080/19443994.2014.952674.

Zbikowski R. 2011. Use of artificial neural networks to identify the origin of green macroalgae. Estuarine Coastal Shelf Sci. 94(2):138–143. doi:10.1016/j.ecss.2011.05.027..

Zhao S, Afgan MS, Zhu H, Gao X. 2022. Femtosecond laser filamentation-induced breakdown spectroscopy combined with chemometrics methods for soil heavy metal analysis. Optik. 251:168444. doi:10.1016/j.ijleo.2021.168444.

Zhao YH, Jia X, Wang WK, Liu T, Huang SP, Yang MY. 2016. Growth under elevated air temperature alters secondary metabolites in *Robinia pseudoacacia* L. seedlings in Cd-and Pb-contaminated soils. Sci Total Environ. 565:586–594.