



# Deciding Heavy Metal Levels in Soil Based on Various Ecological Information through Artificial Intelligence Modeling

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

To cite this article: Murat Sari, Tahir Cosgun, Ibrahim Ertugrul Yalcin, Mahmut Taner & Ibrahim Ilker Ozyigit (2022) Deciding Heavy Metal Levels in Soil Based on Various Ecological Information through Artificial Intelligence Modeling, Applied Artificial Intelligence, 36:1, 2014189, DOI: 10.1080/08839514.2021.2014189

To link to this article: <https://doi.org/10.1080/08839514.2021.2014189>



© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 11 Dec 2021.



Submit your article to this journal [↗](#)



Article views: 1552



View related articles [↗](#)








View Crossmark data [↗](#)



Citing articles: 2 View citing articles [↗](#)

# Deciding Heavy Metal Levels in Soil Based on Various Ecological Information through Artificial Intelligence Modeling

Murat Sari <sup>a</sup>, Tahir Cosgun <sup>a,b</sup>, Ibrahim Ertugrul Yalcin <sup>c</sup>, Mahmut Taner <sup>d</sup>,  
and Ibrahim Ilker Ozyigit <sup>e,f</sup>

<sup>a</sup>Faculty of Arts & Sciences, Department of Mathematics, Yildiz Technical University, Istanbul, Turkey;

<sup>b</sup>Faculty of Arts & Sciences, Department of Mathematics, Amasya University, Amasya, Turkey; <sup>c</sup>Faculty of Engineering and Natural Sciences, Department of Civil Engineering, Bahcesehir University, Istanbul, Turkey; <sup>d</sup>Faculty of Engineering and Natural Sciences, Department of Mathematics, Bahcesehir University, Istanbul, Turkey; <sup>e</sup>Faculty of Arts & Sciences, Department of Biology, Marmara University, Istanbul, Turkey; <sup>f</sup>Faculty of Sciences, Department of Biology, Kyrgyz-Turkish Manas University, Bishkek, Kyrgyzstan

## ABSTRACT

The aim of this paper is to decide on heavy metal levels based on ecological parameters by effectively eliminating common disadvantages such as high cost and serious time-consuming laboratory procedures via an effective artificial intelligence approach. Therefore, this study is hinged on an artificial intelligence technique, ANN, because of its low cost and high accuracy in overcoming the mentioned limitations and obstacles in the determination process of the amounts of elements. The ANNs have thus been employed to determine essential heavy metals, such as Fe, Mn, and Zn depending on Ca, K, and Mg concentrations of soil samples obtained from different altitudes in Mount Ida. To the best knowledge of the authors, this is the first study in the literature in which altitude was considered as a parameter in the prediction of nutrient heavy metals. The computed relative errors are significantly low for each of the considered elements (Fe, Mn, and Zn); and are found to be between 1.0–4.1%, 1.0–4.2%, 1.5–7.1%, respectively, for the training, testing, and holdout data. The findings indicate that the relative errors could still be decreased further by assuming the altitude as a factor variable.

## ARTICLE HISTORY



Received 22 June 2021

Revised 25 November 2021

Accepted 30 November 2021

## Introduction

The amount of nutrient elements in the soil is of great importance for plants since the supplies and variety of these elements provide plenty and richness of the vegetation. In this respect, the Mount Ida, which has become a national park because of the abundance of vegetation and the biodiversity it provides therein, contains environmentally significant examples of various plant flora.

**CONTACT** Murat Sari  [seismic.sari@gmail.com](mailto:seismic.sari@gmail.com)  Faculty of Arts & Sciences, Department of Mathematics, Yildiz Technical University, 34220 Istanbul, Turkey

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The analytical elemental analysis instruments utilized to determine the amount of the elements in soil and plants have various difficulties and limitations.

The presence of nutrient elements in the soil is a necessity for plants (Arefi, Kafi, and Khazaei 2013). Elements that are characterized as macro elements such as nitrogen (N), phosphorus (P), calcium (Ca), potassium (K), and magnesium (Mg), whose rates of appearance in the soil are higher than many other elements, contribute greatly to the plant diversity in the habitat (Can et al. 2021). Considering the importance of macro elements, it is necessary to know how much they are in the soil (Yahia, García-Solís, and Celis 2019).

Determination of the amounts of elements in soil and plants is traditionally performed with elemental analysis instruments such as Atomic Absorption Spectrophotometer (AAS), Inductively Coupled Plasma Mass Spectrometer (ICP-MS), and Inductively Coupled Plasma Optical Emission Spectroscopy (ICP-OES) (Alam, Ahmed, and Howladar 2020; Dan-Badjo et al. 2019; Sayo, Kiratu, and Nyamato 2020; Yalcin et al. 2020). In addition to the conventional analytical methods, various mathematical models, and artificial intelligence techniques, including ANNs, are also utilized regarding the prediction of the amounts of stated elements (Marković et al. 2016; Zhang et al. 2020; Zhou et al. 2015).

Environmental components such as air, water, and soil are pollutant receptors of large amounts from multiple sources, and therefore they can be used to study the origin and properties of the pollution (Alengebawy et al. 2021; Emenike et al. 2020). Environmental factors such as air, water, and soil are susceptible to contamination due to various pollutant effluents and substances, and consequently, are prone to create severe ambiguities in risk assessment, decision making, and management processes (Suarez-Paba, Cruz, and Munoz 2020). Soil pollution is accepted as an important risk indicator especially in woodlands since it paves the way to severe dangers for health and the environment (Hossen et al. 2021; Jiang et al. 2021; Wu et al. 2021). Elemental analysis, among the different soil quality indicators, comes to the forefront as a traditional but efficient method to assess the extent of soil contamination. However, experimentally measuring the correlation between soil properties and elemental concentration of pollutants are both time-consuming and costly. Moreover, it seems impossible to simulate all naturally occurring variants with the highest accuracy.

In the use of elemental analysis instruments that are accepted as traditional methods, there are a lot of difficulties such as costs arising from the excessive use of consumables such as acid and pure gas, the need for providing appropriate temperature, humidity, cleaning conditions in the environment where the device is located, the device operator costs based on elapsed analysis times for each element, and deduction limits (Pearce et al. 2004; Vera et al. 2021). However, the mathematical modeling tools,

especially artificial intelligence techniques, outstand as important alternatives to save time and costs. In this respect, there are studies in the literature reporting that ANNs could make more accurate predictions as compared to the traditional methods (Inakollu et al. 2009; Mojid, Hossain, and Ashraf 2019; Olawoyin et al. 2013; Shadrin et al. 2020). The existing studies in the literature on element analysis include mathematical models regarding long time periods like the effects of environmental pollution in the long run mostly from a temporal viewpoint (Bui, Bui, and Nguyen 2021; Jia, Dong, and Du 2020; Jiang, Nan, and Yang 2013; Li et al. 2021; Oonk and Spijker 2015). To the best knowledge of the authors, the ANNs have not been used by taking into account the spatial-data viewpoint as in this study before to obtain the amounts of ingredients like heavy metals in soil. The most prominent aspect of this research is to assure that it is possible to obtain cost-effective results as compared to the derived results by using the conventional device-based techniques by performing fewer amount of elemental analysis via an artificial intelligence technique based on ANNs.

Since they involve an adaptable and flexible structure consisting of layers and neurons, ANNs can identify and classify complex nonlinear relationships between input and output datasets. Therefore, ANN is a powerful technology for modeling the complex “input-output” relationship (Chen et al. 2019). ANN is also a technique with flexible mathematical structure that could be developed by inspiration from biological neural networks (Qaderi et al. 2017). The flexibility and adaptivity of ANNs to use in many areas enable them as the first option for modeling many scientific problems. Recently, ANNs, which have been used successfully toward the solution of various problems, have increased their popularity progressively in environmental problems, since they could be fruitful tools of the trade to estimate the features or behaviors encountered in environmental problems (Fangfang, Alagumalai, and Mahian 2021; Guo and Wang 2021; Kassem, Gokcekus, and Maliha 2021; Nourani et al. 2020; Shams et al. 2021; Wolski and Kruk 2020). In recent years, the performance of ANN models has increased tremendously with the development of computer hardware and software technologies and has attracted great attention regarding the construction of various prediction models. Highly complicated features or complex nonlinear relationships regarding a dataset could be enlightened by increasing the depth of a neural network and the number of neurons included in the hidden layers on a reasonable level and sound basis. Since it is a data-driven approach, the ANN technique does not require troublesome mathematical computations such as the computation of a Jacobian matrix, numerical integration, and so on to handle sophisticated scientific problems. Therefore, ANNs have been utilized successfully to deal with a lot of problems in a wide variety of disciplines such as biomechanics (Sari and

Cetiner 2009), climatology (Shahmansouri et al. 2021), financial and economic modeling (Sermpinis et al. 2021), environmental pollution (Maleki et al. 2019), medical applications (Zhang et al. 2018), and composition of music via computers (Abu Doush and Sawalha 2020; Briot 2021). In this respect, ANNs can be used to make the necessary estimates in a variety of research areas, and the flexibility of the number of dependent and independent variables is one of the most outstanding aspects. In the determination process of the optimization model, ANNs, which is inspired by biological neural networks, are taking place in the prediction tools and provide accurate responses in various fields such as determination of plant element concentration, estimation of water and soil pollution, determination of suitable agricultural lands, determination of irrigation type and amount.

In soil chemistry, while macro elements such as Ca, K, Mg, N, P, and S are allowed to be present in comparatively high concentrations for plant vitality, heavy metals are considered as a special element group having toxic effects on plants in the case of high concentrations (Karahan et al. 2020; Uchimiya et al. 2020). Although Cu, Fe, Mn, Mo, and Zn are essential heavy metals for plant growth, their excess amount in the soil above a certain concentration may cause toxic effects for plants (Jothimani, Arulbalachandran, and Yasmin 2017; Turan, Ozdemir, and Demir 2020).

Mount Kaz or Mount Ida is a mountain range located between Canakkale and Balikesir provinces and in the north of Edremit Bay. The mountain, called Kaz Mountain or Kaz Mountains, extends largely on the Biga Peninsula. Kaz Mountains consist of Mount Dede in the west, Mount Kaz in the middle with three hills (Babadag in the north, Karatas hill in the middle, Sarikiz hill in the south), Mount Eybek in the east, and Mounts Gorgen and Kocakatan in the northeast. The deep valleys and canyons located on Mount Ida and extended in the north–south direction exhibit a rich potential regarding flora and fauna (Efe et al. 2015, 2014; Ozyigit et al. 2015). Especially, the biological diversity housed by the vegetation constitutes one of the most remarkable values of the national park.

This study mainly aims to discover concentrations of essential heavy metals, such as Fe, Mn, and Zn based on Ca, K, and Mg concentrations of soil samples obtained from different altitudes in the Mount Ida.

## **Material and Methods**

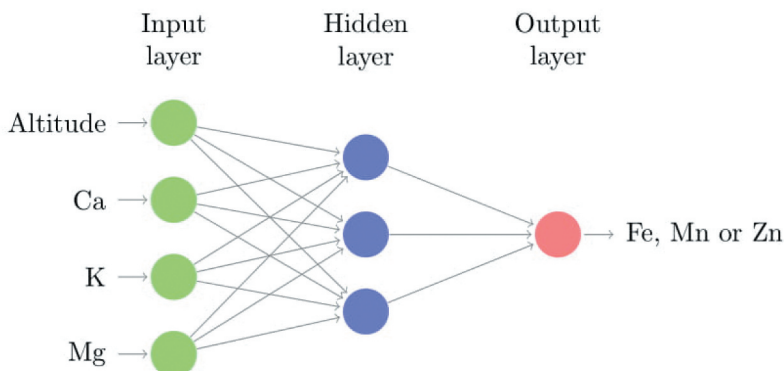
### ***Study Area, Elemental Analysis, and Data Structure***

In this study, soil samples were collected from the Mount Ida national park road first at an altitude of 20 m, then at every 100 m up to 1600 m, 1621 m, 1658 m, and 1678 m along the road to Sarikiz Hill in the national park. Fifteen

samples have been collected from each of the mentioned altitudes. These soil samples weighing 500 g were taken from the designated locations 20 cm below the surface and brought to the laboratory environment in sterile bags. They have been dried in an oven at 80°C for 48 hours in glass petri dishes. Then, it was made ready for weighing by passing it through a steel sieve with a 2 mm pore diameter. Soil samples weighed in the range of 0.200–0.250 g were transferred to Teflon vessels then 6 ml 65% HNO<sub>3</sub> (Merck), 3 ml 37% HCl (Merck) and 2 ml 48% HF (Merck) was added. The digestion process of the prepared samples was carried out by using the Berghof-MSW2 brand-model microwave device. After the process, the samples were transferred to 50 ml sterile falcon tubes using ultra-pure water by filtering with a blue band Whatman filter. The total volume was completed to 50 ml (Ozyigit et al. 2015; Yalcin et al. 2020). The concentrations of Ca, Fe, K, Mg, Mn, and Zn elements were determined as mg kg<sup>-1</sup> dry weight by the PerkinElmer-Optima 7000DV inductively coupled plasma optical emission spectroscopy (ICP-OES) device.

### Neural Network Design and Setup

The ANNs are robust metaheuristic computational tools designed by inspiring the activities of neurons. The general structure of ANNs consists of the input layer, hidden layer(s), and output layer. The input layer contains neurons, so-called covariates, or independent variables, while the output variable contains neurons called dependent variables. Although the number of hidden layers and the number of artificial neurons therein are not definite values, both the number of independent variables and dependent variables is predetermined based on the studied problem. As is the case in Figure 1, there is no connection between the neurons of the same layer, but usually, the neurons of neighboring layers are densely interconnected via synapsis-like



**Figure 1.** The standard architecture of the network diagram.

**Table 1.** Case processing summary.

		Fe		Mn		Zn	
		N	Percent	N	Percent	N	Percent
Sample	Training	214	71.3%	204	68.0%	203	67.7%
	Testing	56	18.7%	76	25.3%	68	22.7%
	Holdout	30	10.0%	20	6.7%	29	9.7%
Valid		300	100.0%	300	100.0%	300	100.0%
Excluded		0		0		0	
Total		300		300		300	

structures. Thereby, the data collected from the environment is utilized to train the neural network to predict the continuous values of the dependent variable.

The data collected in a total of 300 samples are divided into three subgroups called training data, testing data, and holdout to prevent overfitting and to make reasonable predictions (Table 1). The supervised learning starts with random weights and determines the weights that will be applied to the current task. Back-propagation neural networks are typical supervised learning procedures, employing the back error procedure for learning. The utilized back-propagation algorithm to train the neural networks is endowed with gradient descent to minimize the error via updating the weights. SPSS 27 has been utilized to reach the desired data-driven model. During the computations, the method is chosen to be neural networks, the submodule is preferred as the multi-layer perceptron (MLP), the number of hidden layers is taken to be 1, the number of neurons or units in the hidden layer is taken to be 3, and the hyperbolic tangent function and the sigmoid function are utilized as the activation function for the hidden layer and the output layer, respectively.

The inner structure of the MLP, as indicated in Table 2, is constructed by attributing 4 input variables where Altitude, Ca, Mg, and K are set as independent variables in each case during the calculations. In addition, the number of the output variable adjusted as 1 that is represented individually by Fe, Mn, and Zn. The distribution of the training, testing and holdout data are

**Table 2.** Network information.

Input Layer	Covariates	1	Altitude
		2	Ca
		3	K
		4	Mg
Hidden Layer(s)	Number of Units <sup>a</sup>	4	
	Rescaling Method for Covariates		Normalized
	Number of Hidden Layers	1	
	Number of Units in Hidden Layer <sup>a</sup>	3	
Output Layer	Activation Function		Hyperbolic tangent
	Dependent Variables	1	Fe, Mn, or Zn
	Number of Units	1	
	Rescaling Method for Scale Dependents		Normalized
	Activation Function		Sigmoid
	Error Function		Sum of Squares

<sup>a</sup>.Excluding the bias unit



determined as 70%, 20%, and 10% for each cases Fe, Mn, and Zn. Regarding the structure of the MLP, it is important to note that the number of units in the input layer is seen to be 4 in Table 2. Both the dependent and independent variables are rescaled through a normalization process before the analysis via standardization. Different MLPs have been trained and separate analyses have been carried out for each dependent variable as it is understood from Figure 2–3 and the quantitative results in Table 3–5.

The learning rate of the algorithm is set to be 0.4 initially with a momentum of 0.9. The squared error function has been utilized during the operation of the gradient descent algorithm. The algorithm consumes all the data once and continues to use the recorded data until reaching one of the stopping criteria such as the maximum number of epochs, a predetermined number of consecutive steps with no decrease in error, maximum ratio of the relative errors, and so on.

## Results and Discussion

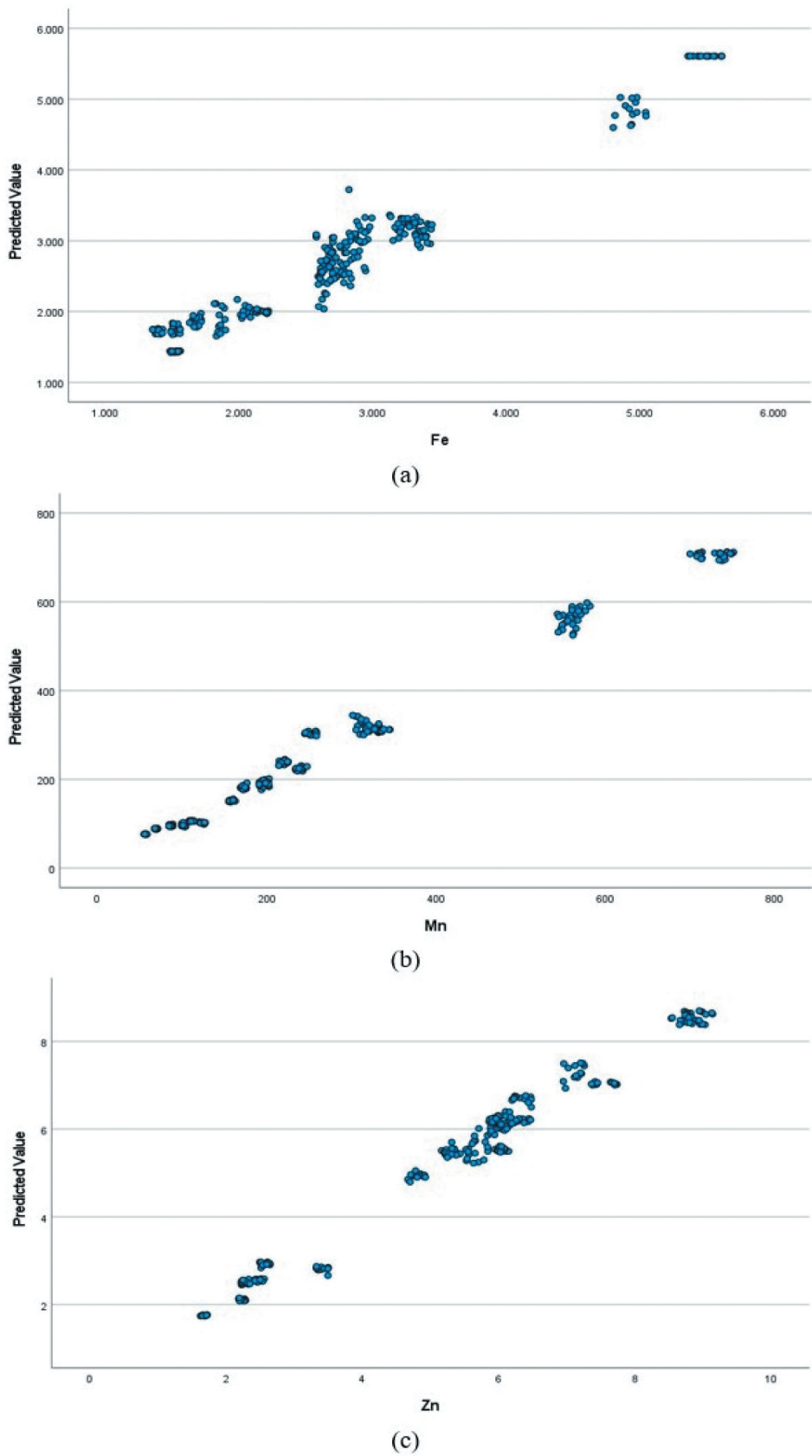
When classifications or predictions based on data-driven mathematical models are concerned without noticing whether there is causality between the dependent and independent variables or not ANNs are powerful tools. The current study focuses on the prediction regarding the amount of essential heavy metals like Fe, Mn, or Zn in the soil of Mount Ida by considering the altitude in addition to the macro elements like Ca, K, and Mg. Although various methods such as artificial neural networks, machine learning, and even deep learning have been used in the literature to predict element levels, so far, altitude has not been considered in any study (Bagheri, Bazvand, and Ehteshami 2017; Bhagat et al. 2021; Hanandeh, Mahdi, and Imtiaz 2021; Lu et al. 2019; Shadrin et al. 2020). A supervised learning procedure, ANNs, are adopted to achieve this aim. It has been seen that the ANNs can be preferred as an important alternative in determining the soil element concentrations as a time-saving and low-cost method. The current method could be employed to make predictions with higher accuracy as compared to the traditional analytical methods, in most cases. The current approach can be regarded as a prediction model based on the real-world data, and comparative results with statistical methods in the literature are not included in this paper. In addition, validation of the present approach has been performed via splitting the data into three parts and using one of them for validation. Thereby, the required validation has been observed by utilizing the real-world data that is not used in the training or testing stages. Besides, to the best knowledge of the authors, this study that utilizes the ANNs to forecast the amount of essential heavy metals by associating them with the altitude is a pioneering work in the literature. In addition, the importance of this ANN approach can be better understood when its possible impacts on industrial areas such as



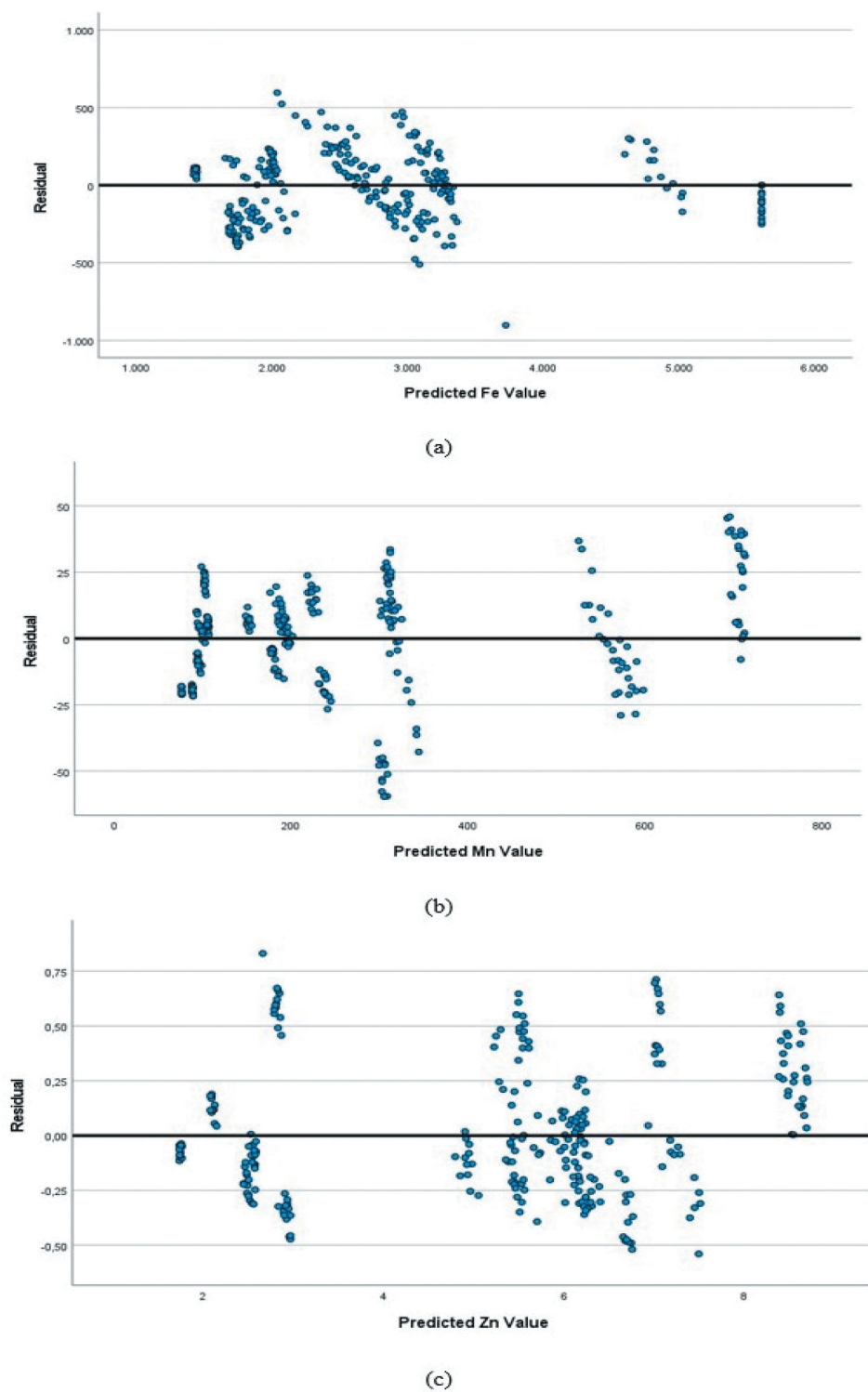
smart farming, precision agriculture, and mining are taken into account. With this study, it has been shown that the ANN technique is an effective alternative method to reach further element analysis results by doing fewer experiments in soil element content studies.

The architecture of the neural network consists of four independent variables Altitude, Ca, K, Mg in the input layer. The hidden layer comprises three neurons as depicted in [Figure 1](#). The output layer contains a single unit that represents Fe, Mn, or Zn in separate cases. Input and output data are normalized before starting the analysis. Then, the data is distributed into three categories for the purpose of training, testing, and holdout. These partitions of the data for three different micro elements (Fe, Mn, or Zn) can be found in [Table 1](#) in detail. The training part is utilized for the supervision of MLP while the testing or validation part is exploited to avoid overfitting. The holdout data is spared to control the accuracy of the prediction. The hidden layer made use of the hyperbolic tangent function while the output layer used the logistic sigmoid function as the activation function since the sigmoid functions are widely utilized in prediction models concerning mathematical biology and chemistry. The audience can refer to the work of Mustafa et al. (2021) for a specific example. The sum of squares error is adopted as the error function, and each of the number of consecutive steps without decreasing in the error, the maximum training time, the maximum number of epochs, the error change or the relative error change in the consecutive steps are introduced as stopping criteria.

Firstly, the predicted vs actual Fe values have been plotted in [Figure 2\(a\)](#) for the whole dataset, and the results accumulate in an acceptable and reasonable neighborhood of the perfect prediction line. The residuals have been depicted in [Figure 3\(a\)](#) for the predicted Fe values, and the results are shown in the figure again quite plausible when the actual values are considered. The quantitative results in [Table 3](#) show that the relative errors have been found to be 4.1%, 4.2%, and 7.1% for the training, testing, and holdout data, respectively. Since the actual values are sufficiently higher than zero so that the relative errors suffice to take a full grasp regarding the effectiveness of the current technique. Moreover, it is important to point out that the holdout dataset has not been included in the analysis during the training and testing stages. Besides, the sum of squares error is computed as 1.070 for the training set and 0.0131 for the testing set that seems to be considerably small values. Also, the RMSEs have been computed to be .071 and .075, respectively, for the training and testing data. Another important finding is the importance of the altitude among the independent variables. The conclusion that Altitude seems to be the most important independent variable among all



**Figure 2.** Predicted vs actual Fe (a), Mn (b), Zn (c) amounts in  $\text{mg kg}^{-1}$ .



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

**Table 3.** Model summary.

		Fe	Mn	Zn
Training	Sum of Squares Error	1.070	.081	.155
	RMSE	.071	.012	.028
	Relative Error	.041	.010	.020
	Stopping Rule Used	Relative change in training error criterion (.000001) achieved	Maximum number of epochs (1000) exceeded	Maximum number of epochs (1000) exceeded
	Training Time	0:00:01.16	0:00:00.10	0:00:00.09
Testing	Sum of Squares Error	.311	.032	.063
	RMSE	.075	.021	.030
	Relative Error	.042	.010	.022
Holdout	Relative Error	.071	.015	.034
Dependent Variable: Fe, Mn, or Zn				

**Table 4.** Independent variable importance.

Fe			Mn		Zn	
	Importance	Normalized Importance	Importance	Normalized Importance	Importance	Normalized Importance
Altitude	.328	100.0%	.205	58.2%	.322	99.2%
Ca	.302	92.0%	.352	100.0%	.187	57.6%
K	.115	35.2%	.335	95.3%	.325	100.0%
Mg	.255	77.8%	.109	30.9%	.166	51.1%

independent variables can be inferred from [Table 4](#). Also, the parameter estimation demonstrated in [Table 5](#) provides a sound base regarding this inference and is utterly persuasive.

Secondly, the predicted vs the actual Mn values and the residuals are illustrated in [Figures 2\(b\) and 3\(b\)](#), respectively. The results in [Figure 2\(b\)](#) exhibit an almost perfect match while the residual values seem fully acceptable in a harmony with the predicted vs actual values. According to the findings in [Table 3](#), the relative errors has been computed as 1.0%, 1.0%, 1.5% for the training, testing, and holdout data, respectively. The sum of squares errors is found to be 0.081 and 0.032 for the training and testing data, respectively. In addition, the RMSEs were computed to be .012 and .021, respectively, for the training and testing data. As reported by the results concerning the parameter estimation in [Tables 4 and 5](#), Ca and K are the foremost important independent variables in the prediction of Mn.

Finally, the predicted vs true Zn values are plotted in [Figure 2\(c\)](#), and the residual plot is supplied in [Figure 3\(c\)](#). The power of the ANN-based prediction method could be observed in the former one since the actual vs the predicted values concentrate in a tiny subset of the confidence region. Furthermore, the corresponding relative errors are found to be 2.0%, 2.2%, and 3.4%, respectively, for the training, testing, and holdout data while the sum of squares errors is computed to be 0.155 and 0.063 for the training and testing datasets. Moreover, the RMSEs have been computed to be .028 and .030

**Table 5.** Parameter estimates for Fe, Mn, and Zn.

		Predicted			
		Hidden Layer			Output Layer
Predictor		H(1:1)	H(1:2)	H(1:3)	Fe
Input Layer	(Bias)	.511	.941	−1.155	
	Altitude	1.869	−2.007	3.285	
	Ca	.237	1.963	−1.877	
	K	.022	−.379	.648	
	Mg	−4.267	−4.313	−1.667	
Hidden Layer	(Bias)				2.115
	H(1:1)				−3.261
	H(1:2)				2.828
	H(1:3)				3.614
Predictor		H(1:1)	H(1:2)	H(1:3)	Mn
Input Layer	(Bias)	.865	2.182	−4.448	
	Altitude	1.841	−2.984	1.062	
	Ca	−2.381	.781	5.578	
	K	−.091	.927	2.816	
	Mg	.518	−.908	−1.298	
Hidden Layer	(Bias)				−2.041
	H(1:1)				3.007
	H(1:2)				1.903
	H(1:3)				2.730
Predictor		H(1:1)	H(1:2)	H(1:3)	Zn
Input Layer	(Bias)	1.985	−.794	1.828	
	Altitude	−2.766	1.357	−5.497	
	Ca	−1.077	−4.024	1.237	
	K	2.094	.920	.444	
	Mg	−.639	.845	.123	
Hidden Layer	(Bias)				−2.266
	H(1:1)				4.165
	H(1:2)				−1.630
	H(1:3)				−2.858

**Table 6.** Model summary for the case that altitude is a factor variable.

		Fe	Mn	Zn
Training	Sum of Squares Error	.118	.026	.022
	Relative Error	.018	.005	.004
	Stopping Rule Used	Maximum number of epochs (100) exceeded	Maximum number of epochs (100) exceeded	10 consecutive step(s) with no decrease in error <sup>a</sup>
	Training Time	0:00:00.03	0:00:00.02	0:00:00.01
	Sum of Squares Error	.031	.021	.015
Testing	Relative Error	.031	.006	.005
Holdout	Relative Error	.022	.003	.005
Dependent Variable: Fe, Mn, or Zn				

for the training and testing data, respectively. According to the results of variable importance analysis, the importance of the independent variables K and Altitude seem the highest ones among the independent variable importance's via the parameter estimations in [Tables 4 and 5](#).

Indeed, although the produced results are quite satisfactory, the relative errors could be declined up to one-fifth of the received results. Particularly, the derived results in [Table 6](#) have been deduced by setting Altitude as a factor

variable instead of an independent variable, even for modest stopping criteria. At its core, the attribution of Altitude as a factor variable at the beginning of the training leads to a higher numerical accuracy since each separate value of Altitude is involved in the training process as a separate independent variable in that case.

## Conclusions

In this paper, efficient neural network models with an innovative perspective of ecological importance have been produced to determine the heavy metal (Fe, Mn, and Zn) level in soil by using local factors (altitude, Ca, K, and Mg). It is believed that the derived network models will assist in finding heavy metal levels in soil, assessing economic and environmental impacts in various local conditions, and making accurate decisions in improving desired environmental policies. In this context, heavy metal levels have been investigated based on the soil information obtained from the Mount Ida. Unlike conventional approaches, the ANN has then succeeded in investigating heavy metal levels, depending on the real data, in the soil through altitude and various soil parameters for the first time. This article has provided very informative and guiding results from the latest research through the derived neural network models for the ecosystem of interest, based on the relation between the altitude and the examined element levels for the first time in the literature. The simulation results can be viewed as an important indication of the superiority of the network algorithm over other conventional approaches. It has been concluded that these findings will be very beneficial in terms of environmental and all aspects in organizing suitable research plans for scientists in this field. There is no doubt that in modeling the behavior of such a wide variety of realistic problems involving a wide range of sciences, optimal assessment of heavy metal status for living things and therefore for agricultural economies and even political decisions in daily life is vital. The results obtained appear to be original and optimal and therefore it is a requirement for the interested readers to be more careful in modeling such problems. Although this study has been carried out to determine the heavy metal level in the Mount Ida, the fact that the same study can be realized very comfortably for any ecosystem is believed to have increased the importance of this study even higher. For further research, such studies can be done for plant element status in a more involving way, similar to the element status determined here in the soil.

## Disclosure statement

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

## ORCID

Murat Sari  <http://orcid.org/0000-0003-0508-2917>

Tahir Cosgun  <http://orcid.org/0000-0003-2970-0863>

Ibrahim Ertugrul Yalcin  <http://orcid.org/0000-0003-3140-7922>

Mahmut Taner  <http://orcid.org/0000-0002-2838-3651>

Ibrahim Ilker Ozyigit  <http://orcid.org/0000-0002-0825-5951>

## Informed consent

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

## References

- Abu Doush, I., and A. Sawalha. 2020. Automatic music composition using genetic algorithm and artificial neural networks. *Malaysian Journal of Computer Science* 33 (1):35–51.
- Alam, R., Z. Ahmed, and M. F. Howladar. 2020. Evaluation of heavy metal contamination in water, soil and plant around the open landfill site Mogla Bazar in Sylhet, Bangladesh. *Groundwater for Sustainable Development* 10:100311. doi:10.1016/j.gsd.2019.100311.
- Alengebawy, A., S. T. Abdelkhalek, S. R. Qureshi, and M. Q. Wang. 2021. Heavy metals and pesticides toxicity in agricultural soil and plants: Ecological risks and human health implications. *Toxics* 9 (3):42–85. doi:10.3390/toxics9030042.
- Arefi, I. H., M. Kafi, and H. Khazaei. 2013. Effect of different levels of nitrogen, phosphorus and potassium nutrient elements on yield, yield components and of Persian Shallot (*Allium altissimum*). *International Journal of Agriculture* 3 (3):516.
- Bagheri, M., A. Bazvand, and M. Ehteshami. 2017. Application of artificial intelligence for the management of landfill leachate penetration into groundwater. *And Assessment of Its Environmental Impacts, Journal of Cleaner Production* 149 (Pages):784–96.
- Bhagat, S. K., K. Pyrgaki, S. Q. Salih, T. Tiyyasha, U. Beyaztas, S. Shahid, and Z. M. Yaseen. 2021. Prediction of copper ions adsorption by attapulgite adsorbent using tuned-artificial intelligence model. *Chemosphere* 276:130162. doi:10.1016/j.chemosphere.2021.130162.
- Briot, J. P. 2021. From artificial neural networks to deep learning for music generation: History, concepts and trends. *Neural Computing and Applications* 33 (1):39–65. doi:10.1007/s00521-020-05399-0.
- Bui, X. N., H. B. Bui, and H. Nguyen (2021). A Review of Artificial Intelligence Applications in Mining and Geological Engineering. In *Proceedings of the International Conference on Innovations for Sustainable and Responsible Mining* (pp. 109–42). Springer, Cham.
- Can, H., I. I. Ozyigit, M. Can, A. Hocaoglu-Ozyigit, and I. E. Yalcin. 2021. Environment-Based Impairment in Mineral Nutrient Status and Heavy Metal Contents of Commonly Consumed Leafy Vegetables Marketed in Kyrgyzstan: A Case Study for Health Risk Assessment. *Biological Trace Element Research* 199 (3):1123–44. doi:10.1007/s12011-020-02208-6.
- Chen, M., U. Challita, W. Saad, C. Yin, and M. Debbah. 2019. Artificial neural networks-based machine learning for wireless networks: A tutorial. *IEEE Communications Surveys & Tutorials* 21 (4):3039–71. doi:10.1109/COMST.2019.2926625.
- Dan-Badjo, A. T., O. Z. Ibrahim, Y. Guéro, J. L. Morel, C. Feidt, and G. Echevarria. 2019. Impacts of artisanal gold mining on soil, water and plant contamination by trace elements at Komabangou, Western Niger. *Journal of Geochemical Exploration* 205:106328. doi:10.1016/j.gexplo.2019.06.010.



- Efe, R., A. Soykan, I. Curebal, and S. Sonmez. 2014. Geomorphologic and ecologic Fundamentals of the air quality in the vicinity of Kaz Mountains. *Procedia-Social and Behavioral Sciences* 120:335–46. doi:[10.1016/j.sbspro.2014.02.111](https://doi.org/10.1016/j.sbspro.2014.02.111).
- Efe, R., S. Sonmez, I. Curebal, and A. Soykan. 2015. Ecological conditions and vegetation of subalpine zone of Kaz Mountain (Mount Ida, NW Turkey). In *Climate Change Impacts on High-Altitude Ecosystems*, ed. Efe, Recep, Ozturk, Munir, and Atalay, 591–608. Cham: Springer.
- Emenike, P. C., I. T. Tenebe, J. B. Neris, D. O. Omole, O. Afolayan, C. U. Okeke, and I. K. Emenike. 2020. An integrated assessment of land-use change impact, seasonal variation of pollution indices and human health risk of selected toxic elements in sediments of River Atuwara, Nigeria. *Environmental Pollution* 265:114795.
- Fangfang, F., A. Alagumalai, and O. Mahian. 2021. Sustainable biodiesel production from waste cooking oil: ANN modeling and environmental factor assessment. *Sustainable Energy Technologies and Assessments* 46:101265. doi:[10.1016/j.seta.2021.101265](https://doi.org/10.1016/j.seta.2021.101265).
- Guo, X., and J. Wang. 2021. Projecting the sorption capacity of heavy metal ions onto microplastics in global aquatic environments using artificial neural networks. *Journal of Hazardous Materials* 402:123709. doi:[10.1016/j.jhazmat.2020.123709](https://doi.org/10.1016/j.jhazmat.2020.123709).
- Hanandeh, A., Z. Mahdi, and M. S. Imtiaz. 2021. Modelling of the adsorption of Pb, Cu and Ni ions from single and multi-component aqueous solutions by date seed derived biochar: Comparison of six machine learning approaches. *Environmental Research* 192:110338. doi:[10.1016/j.envres.2020.110338](https://doi.org/10.1016/j.envres.2020.110338).
- Hossen, M. A., A. I. H. Chowdhury, M. R. A. Mullick, and A. Hoque. 2021. Heavy metal pollution status and health risk assessment vicinity to Barapukuria coal mine area of Bangladesh. *Environmental Nanotechnology. Monitoring & Management* 100469: 1–14.
- Inakollu, P., T. Philip, A. K. Rai, F. Y. Yueh, and J. P. Singh. 2009. A comparative study of laser induced breakdown spectroscopy analysis for element concentrations in aluminum alloy using artificial neural networks and calibration methods. *Spectrochimica Acta Part B: Atomic Spectroscopy* 64 (1):99–104. doi:[10.1016/j.sab.2008.11.001](https://doi.org/10.1016/j.sab.2008.11.001).
- Jia, B., R. Dong, and J. Du. 2020. Ozone concentrations prediction in Lanzhou, China, using chaotic artificial neural network. *Chemometrics and Intelligent Laboratory Systems* 204:104098. doi:[10.1016/j.chemolab.2020.104098](https://doi.org/10.1016/j.chemolab.2020.104098).
- Jiang, Y., Z. Nan, and S. Yang. 2013. Risk assessment of water quality using Monte Carlo simulation and artificial neural network method. *Journal of Environmental Management* 122:130–36. doi:[10.1016/j.jenvman.2013.03.015](https://doi.org/10.1016/j.jenvman.2013.03.015).
- Jiang, Z., Z. Guo, C. Peng, X. Liu, Z. Zhou, and X. Xiao. 2021. Heavy metals in soils around non-ferrous smelteries in China: Status, health risks and control measures. *Environmental Pollution* 282:117038. doi:[10.1016/j.envpol.2021.117038](https://doi.org/10.1016/j.envpol.2021.117038).
- Jothimani, K., D. Arulbalachandran, and K. Yasmin. 2017. Amelioration of Environmental Stress for Sustainable Crop Productivity. In *Sustainable Agriculture towards Food Security* Dhanarajan, Arulbalachandran, 327–48. Singapore: Springer.
- Karahan, F., I. I. Ozyigit, I. A. Saracoglu, I. E. Yalcin, A. H. Ozyigit, and A. Ilcim. 2020. Heavy metal levels and mineral nutrient status in different parts of various medicinal plants collected from eastern Mediterranean region of Turkey. *Biological Trace Element Research* 197 (1):316–29. doi:[10.1007/s12011-019-01974-2](https://doi.org/10.1007/s12011-019-01974-2).
- Kassem, Y., H. Gokcekus, and M. R. M. Maliha. 2021. Identifying most influencing input parameters for predicting chloride concentration in groundwater using an ANN approach. *Environmental Earth Sciences* 80 (7):248. doi:[10.1007/s12665-021-09541-6](https://doi.org/10.1007/s12665-021-09541-6).

- Li, X., B. Yang, J. Yang, Y. Fan, X. Qian, and H. Li. 2021. Magnetic properties and its application in the prediction of potentially toxic elements in aquatic products by machine learning. *Science of the Total Environment* 783:147083. doi:10.1016/j.scitotenv.2021.147083.
- Lu, H., H. Li, T. Liu, Y. Fan, Y. Yuan, M. Xie, and X. Qian. 2019. Simulating heavy metal concentrations in an aquatic environment using artificial intelligence models and physico-chemical indexes. *Science of the Total Environment* 694:133591. doi:10.1016/j.scitotenv.2019.133591.
- Maleki, H., A. Sorooshian, G. Goudarzi, Z. Baboli, Y. T. Birgani, and M. Rahmati. 2019. Air pollution prediction by using an artificial neural network model. *Clean Technologies and Environmental Policy* 21 (6):1341–52. doi:10.1007/s10098-019-01709-w.
- Marković, J., M. Jović, I. Smičiklas, L. Pezo, M. Šljivić-Ivanović, A. Onjia, and A. Popović. 2016. Chemical speciation of metals in unpolluted soils of different types: Correlation with soil characteristics and an ANN modelling approach. *Journal of Geochemical Exploration* 165:71–80. doi:10.1016/j.gexplo.2016.03.004.
- Mojid, M. A., A. B. M. Z. Hossain, and M. A. Ashraf. 2019. Artificial neural network model to predict transport parameters of reactive solutes from basic soil properties. *Environmental Pollution* 255:113355. doi:10.1016/j.envpol.2019.113355.
- Mustafa, H. M., A. Mustapha, G. Hayder, and A. Salisu, “Applications of IoT and Artificial Intelligence in Water Quality Monitoring and Prediction: A Review,” 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2021, pp. 968–75.
- Nourani, V., H. Gokcekus, I. K. Umar, and H. Najafi. 2020. An emotional artificial neural network for prediction of vehicular traffic noise. *Science of the Total Environment* 707:136134. doi:10.1016/j.scitotenv.2019.136134.
- Olawoyin, R., A. Nieto, R. L. Grayson, F. Hardisty, and S. Oyewole. 2013. Application of artificial neural network (ANN)–self-organizing map (SOM) for the categorization of water, soil and sediment quality in petrochemical regions. *Expert Systems with Applications* 40 (9):3634–48. doi:10.1016/j.eswa.2012.12.069.
- Oonk, S., and J. Spijker. 2015. A supervised machine-learning approach towards geochemical predictive modelling in archaeology. *Journal of Archaeological Science* 59:80–88. doi:10.1016/j.jas.2015.04.002.
- Ozyigit, I. I., Z. Severoglu, R. Vatansever, and M. Ozturk. 2015. Soil–Plant Interactions in the High-Altitude Ecosystems: A Case Study from Kaz Dagi (Mount Ida), Turkey. In *Climate Change Impacts on High-Altitude Ecosystems*, ed. Ozturk, Munir, Hakeem, Khalid Rehman, Hanum, I. Faridah, and Efe, Recep, 343–60. Cham: Springer.
- Pearce, N. J., J. A. Westgate, W. T. Perkins, and S. J. Preece. 2004. The application of ICP-MS methods to tephrochronological problems. *Applied Geochemistry* 19 (3):289–322. doi:10.1016/S0883-2927(03)00153-7.
- Qaderi, F., and E. Babanezhad. 2017. Prediction of the groundwater remediation costs for drinking use based on quality of water resource, using artificial neural network. *Journal of Cleaner Production* 161:840–49. doi:10.1016/j.jclepro.2017.05.187.
- Sari, M., and B. G. Cetiner. 2009. Predicting effect of physical factors on tibial motion using artificial neural networks. *Expert Systems with Applications* 36 (6):9743–46. doi:10.1016/j.eswa.2009.02.030.
- Sayo, S., J. M. Kiratu, and G. S. Nyamato. 2020. Heavy metal concentrations in soil and vegetables irrigated with sewage effluent: A case study of Embu sewage treatment plant, Kenya. *Scientific African* 8:e00337. doi:10.1016/j.sciaf.2020.e00337.
- Sermpinis, G., A. Karathanasopoulos, R. Rosillo, D. de La Fuente. 2021. Neural networks in financial trading. *Annals of Operations Research*. 297(1–2):293–308. doi:10.1007/s10479-019-03144-y.

- Shadrin, D., M. Pukalchik, E. Kovaleva, and M. Fedorov. 2020. Artificial intelligence models to predict acute phytotoxicity in petroleum contaminated soils. *Ecotoxicology and Environmental Safety* 194:110410. doi:[10.1016/j.ecoenv.2020.110410](https://doi.org/10.1016/j.ecoenv.2020.110410).
- Shahmansouri, A. A., M. Yazdani, S. Ghanbari, H. A. Bengar, A. Jafari, and H. F. Ghatte. 2021. Artificial neural network model to predict the compressive strength of eco-friendly geopolymer concrete incorporating silica fume and natural zeolite. *Journal of Cleaner Production* 279:123697. doi:[10.1016/j.jclepro.2020.123697](https://doi.org/10.1016/j.jclepro.2020.123697).
- Shams, S. R., A. Jahani, S. Kalantary, M. Moeinaddini, and N. Khorasani. 2021. The evaluation on artificial neural networks (ANN) and multiple linear regressions (MLR) models for predicting SO<sub>2</sub> concentration. *Urban Climate* 37:100837. doi:[10.1016/j.uclim.2021.100837](https://doi.org/10.1016/j.uclim.2021.100837).
- Suarez-Paba, M. C., A. M. Cruz, and F. Munoz. 2020. Emerging Natech risk management in Colombia: A survey of governmental organizations. *Safety Science* 128:104777. doi:[10.1016/j.ssci.2020.104777](https://doi.org/10.1016/j.ssci.2020.104777).
- Turan, O., H. Ozdemir, and G. Demir. 2020. Deposition of heavy metals on coniferous tree leaves and soils near heavy urban traffic. *Frontiers in Life Sciences and Related Technologies* 1 (1):35–41.
- Uchimiya, M., D. Bannon, H. Nakanishi, M. B. McBride, M. A. Williams, and T. Yoshihara. 2020. Chemical speciation, plant uptake, and toxicity of heavy metals in agricultural soils. *Journal of Agricultural and Food Chemistry* 68 (46):12856–69. doi:[10.1021/acs.jafc.0c00183](https://doi.org/10.1021/acs.jafc.0c00183).
- Vera, J. B., M. C. Bisinoti, C. D. Amaral, and M. H. Gonzalez. 2021. ICP-quadrupole MS for accurate determination of chromium in environmental and food matrices. *Environmental Nanotechnology, Monitoring & Management* 15:100421. doi:[10.1016/j.enmm.2020.100421](https://doi.org/10.1016/j.enmm.2020.100421).
- Wolski, G. J., and A. Kruk. 2020. Determination of plant communities based on bryophytes: The combined use of Kohonen artificial neural network and indicator species analysis. *Ecological Indicators* 113:106160. doi:[10.1016/j.ecolind.2020.106160](https://doi.org/10.1016/j.ecolind.2020.106160).
- Wu, Q., W. Hu, H. Wang, P. Liu, X. Wang, and B. Huang. 2021. Spatial distribution, ecological risk and sources of heavy metals in soils from a typical economic development area, Southeastern China. *Science of the Total Environment* 780:146557. doi:[10.1016/j.scitotenv.2021.146557](https://doi.org/10.1016/j.scitotenv.2021.146557).
- Yahia, E. M., P. García-Solís, and M. E. M. Celis. 2019. Chapter 2 - Contribution of fruits and vegetables to human nutrition and health. In *Postharvest physiology and biochemistry of fruits and vegetables*, ed. E. M. Yahia, 19–45. Sawston, Cambridge, UK: Woodhead publishing.
- Yalcin, I. E., I. I. Ozyigit, I. Dogan, G. Demir, and C. Yarci. 2020. Using the Turkish red pine tree to monitor heavy metal pollution. *Polish Journal of Environmental Studies* 29 (5):3881–89. doi:[10.15244/pjoes/114505](https://doi.org/10.15244/pjoes/114505).
- Zhang, H., S. Yin, Y. Chen, S. Shao, J. Wu, M. Fan, C. Gao, and C. Gao. 2020. Machine learning-based source identification and spatial prediction of heavy metals in soil in a rapid urbanization area, eastern China. *Journal of Cleaner Production* 273:122858. doi:[10.1016/j.jclepro.2020.122858](https://doi.org/10.1016/j.jclepro.2020.122858).
- Zhang, Z., M. W. Beck, D. A. Winkler, B. Huang, W. Sibanda, and H. Goyal; written on behalf of AME Big-Data Clinical Trial Collaborative Group. 2018. Opening the black box of neural networks: Methods for interpreting neural network models in clinical applications. *Annals of Translational Medicine* 6 (11):216. doi:[10.21037/atm.2018.05.32](https://doi.org/10.21037/atm.2018.05.32).
- Zhou, P., Y. Zhao, Z. Zhao, and T. Chai. 2015. Source mapping and determining of soil contamination by heavy metals using statistical analysis, artificial neural network, and adaptive genetic algorithm. *Journal of Environmental Chemical Engineering* 3 (4):2569–79. doi:[10.1016/j.jece.2015.08.003](https://doi.org/10.1016/j.jece.2015.08.003).