Environ Monit Assess (2023) 195:536
https://doi.org/10.1007/s10661-023-11050-x

Forecasting contamination in an ecosystem based on a network model

Murat Sari · Ibrahim Ertugrul Yalcin · Mahmut Taner · Tahir Cosgun · Ibrahim Ilker Ozyigit

Received: 18 November 2022 / Accepted: 23 February 2023

Abstract This paper aims to predict heavy metal pollution based on ecological factors with a new approach, using artificial neural networks (ANNs), by significantly removing typical obstacles like time-consuming laboratory procedures and high implementation costs. Pollution prediction is crucial for the safety of all living things, for sustainable development, and for policymakers to make the right decisions. This study focuses on predicting heavy metal contamination in an ecosystem at a significantly lower cost because pollution assessment still primarily relies on conventional methods, which are recognized to have disadvantages. To accomplish this, the data collected for 800 plant and soil materials have been utilized in the production of an ANN. This research is the first to use an ANN to predict pollution very accurately and has found the network models to be very suitable systemic tools for modelling in pollution data analysis. The findings appear are promising to be very illuminating and pioneering for scientists, conservationists, and governments to swiftly and optimally develop their appropriate work programs to leave a functioning ecosystem for all living things. It has been observed that the relative errors calculated for each of the polluting heavy metals for training, testing, and holdout data are significantly low.

Keywords Environmental pollution · Cadmium · Chromium · Lead · Neural network model

M. Sari (✉)
Istanbul Technical University, Faculty of Science and Letters, Mathematical Engineering, 34469 Istanbul, Türkiye
e-mail: muratsari@itu.edu.tr

I. E. Yalcin
Bahcesehir University, Faculty of Engineering and Natural Sciences, Department of Civil Engineering, 34353 Istanbul, Türkiye

M. Taner
Istanbul Gelisim University, Department of Web Design and Development, 34310 Istanbul, Türkiye

T. Cosgun
Amasya University, Faculty of Arts & Sciences, Department of Mathematics, 05100 Amasya, Türkiye

I. I. Ozyigit
Marmara University, Faculty of Science, Department of Biology, 34722 Istanbul, Türkiye

Introduction

The majority of issues that occur in nature are frequently guided by mathematical models. One of the most impressive instruments that scientists, businesses, governments, and even the entire world attentively study and follow is the development of effective mathematical models that describe environmental mechanisms, including pollution modelling. Mathematical modelling has been considered a crucial instrument to deal with these problems arising in various fields of science. In this context, advances in the development of computers, algorithmic advancements, and software products have made life much easier in solving complex problems of life. The use of mathematical models is also the inevitable case for problems encountered in environmental pollution. Pollution prediction plays an important role for the academic community and all other stakeholders to influence environmental and scientific decisions (Cabaneros et al., 2017; Crouse et al., 2009; Peng et al., 2016). In particular, many researchers have focused on environmental pollution studies, and many methods have been designed to identify pollution for this purpose (Corradini et al., 2019; Sun et al., 2020; Yalcin et al., 2020). Predicting pollution is indispensable for living things, as there is a serious link between pollution, quality of life, sustainability of life, and the decisions of lawmakers (MacKerron & Mourato, 2009; Shah et al., 2018).

Numerous issues have been solved with ANNs in a variety of fields, including medicine (Niel & Bastard, 2019), biomechanics (Sari & Cetiner, 2009), accounting (Hsieh, 2011), music (Sim et al., 2014), financial and economical modelling (Cen & Wang, 2018), and control systems (Beccali et al., 2018). ANNs are the top choice for modelling numerous problems due to their adaptability and propensity to be used in many fields. Since they estimate the patterns seen in contemporary environmental issues, ANNs, which have been employed successfully to solve a variety of problems, have gained popularity in handling environmental issues (Ahmed et al., 2019; Nourani et al., 2020; Wolski & Kruk, 2020).

Environmental pollution, particularly the atmospheric one, is one of the most important issues for living beings. Pollution by toxic metals such as cadmium, chromium, copper, and lead has increased since the industrial revolution (Hocaoglu-Ozyigit & Genc, 2020; Mitra et al., 2022). It has been shown that chromium levels in both natural resources and plants have increased significantly in recent years, especially due to the rapid development in industry and the lack of environmental awareness (Favi et al., 2022; Ozyigit et al., 2016; Vijayakumar et al., 2022). The fact that Cd is a non-biodegradable element, especially easily accumulated in plant tissues, and therefore its toxic effect even at relatively low concentrations is a worldwide concern. In recent years, lead levels in soil materials and plants have been shown to increase significantly due to the rapid development in industry and lack of consideration for the environment (Greipsson et al., 2022; Pemberthy et al., 2021; Rasool et al., 2021). Despite the remarkable studies conducted with traditional methods in the literature (Cristaldi et al., 2020; Likus-Cieslik et al., 2020; Pulscher et al., 2020; Serbula et al., 2012), as far as the authors know, there is no study in the literature that predicts the presence and extent of environmental pollution using intelligence modelling techniques.

Although there are valuable studies on the detection of pollution in the literature, there is no study that deals with this issue based on intelligence approaches. The ANN model used here is a specific type of neural networking model that is modified to multi-layer perceptrons. Since the nature and severity of environmental pollution have a direct effect on living organisms, the development of an ANN model to be able to make predictions about pollution may be of benefit. Pollution prediction has a crucial role in both environmental engineering and the rest of science since it enables the academic community through making significant estimations and give more accurate decisions (Cabaneros et al., 2017; Crouse et al., 2009; Peng et al., 2016). The relationship between the metabolic processes of living things and the type and severity of environmental pollution can be successfully analyzed through mathematical models (Krstic et al., 2018; Jaskulak et al., 2020; Vaseashta et al., 2021; Sari et al., 2022).

Various bioenvironmental problems stemming from environmental pollution have been dealt with by different techniques. Experimental analysis of pollution is usually quite costly. Biosensors (Guo et al., 2018), mass spectrometry (Matsui et al., 2020), optical atomic spectroscopy (Yalcin et al., 2020), nanostructures (Cui et al., 2020), and sensor networks (Luo & Yang, 2019) are typically used to evaluate pollution (Gill et al., 2012). Although they have a lot of advantages and plenty of fruitful applications; these approaches have various disadvantages such as high cost, difficulty using equipment, time consumption, and usage restriction. The modelling of a process encountered in a particular cross-section of nature also allows researchers to do forecasting without high-cost and labor-intensive laboratory experiments. Heavy metal pollution, which is one of the most important factors affecting the natural balance, has a critical importance in terms of environmental pollution.

The ANN technique has been applied successfully for reliable data processing as it is a more adjustable and assumption-free procedure. As a general modelling tool, the ANN enables the forecasting process for many environmental models. ANN tools learn the hidden relationships between independent and dependent variables with the help of known ones. This article mainly aims to investigate the effective applicability of the ANN technique to predict environmental pollution levels, especially Cd, Cr, and Pb pollution, based on a range of plant and soil materials.

Material and methods

Fabaceae family member *Robinia pseudoacacia*, the black locust, is a medium-sized deciduous tree and has an invasive nature worldwide, especially in temperate regions (Martin, 2019). In addition to the provinces of Istanbul and Kocaeli, municipalities in most countries of the world have been making visual planting in urban areas, especially in parks, gardens, and roadsides. It is also planted in industrial sites that discharge heavy metals and other toxins into the environment since it is a powerful biomonitor (Tzvetkova & Petkova, 2015). Therefore, *R. pseudoacacia* plants can be significantly exposed to highway traffic and heavy industrial metal pollution. Four locations in Istanbul and one in Kocaeli were used to gather plant and soil samples for the current study. Five stations were identified according to pollution levels. The first station is Dilovasi, a heavy industry zone located within the provincial borders of Kocaeli. The 2nd, 3rd, and 4th stations were determined by traffic density within the borders of Istanbul. The 5th station was chosen as the Prince Islands without traffic.

The provinces of Istanbul and Kocaeli were used to collect washed and unwashed leaf samples, stem samples, and bark samples of the *R. pseudoacacia* plant. For 24 h, the stems, barks, and other insulated plant components were dried in an oven at 80 °C. A total of 8 ml of 65% HN03 (Merck) was then added to each Teflon container after each sample of plant components weighed 0.2 g. With the aid of a shovel, 500 g of soil samples was taken at intervals of 10 cm, dried in the ambient atmosphere, and sieved using a 2-mm sieve. Then, 0.250 g of each soil sample was mixed with 6 ml HNO3, 3 ml HCl, and 2 ml HF (Merck). The samples were mineralized in a microwave oven (Berghof-MWS2) at 145 °C for 5 min, 165 °C for 5 min, and 175 °C for 20 min. After cooling to room temperature, the samples were filtered through a 1–2-m filter paper in falcon tubes and filled to 50 ml with distilled deionized water (Yalcin et al., 2020). Inductively coupled plasma optical emission spectrophotometer (ICP-OES, PerkinElmer-Optima 7000 DV) standard solutions were generated using 1000 ppm multi-element stock solutions (Merck). ICP-OES was used to calculate the concentrations of Cd, Cr, and Pb. Eight hundred distinct pieces of data from samples taken in the Turkish provinces of Istanbul and Kocaeli were used in this investigation. For the data utilized in this investigation, the location, a plant component (including the soil), and heavy metal values for Cr and Pb (mg kg^{-1}) were considered as inputs, while heavy metal values for Cd, Cr, and Pb (mg kg^{-1}) were considered as outputs.

MLPN design and configuration

This part of the study aims to model the effect of soil and plant materials on environmental pollution and thus estimate the Cd pollution for input variable location, plant part, Cr, and Pb. ANNs are sophisticated computational algorithms that can solve complex problems by mimicking biological neuron processes. The standard ANN model, as shown in Fig. 1a, is made up of three interconnected layers: input, hidden, and output. The number of hidden layers in the ANN model can be chosen at random. A neuron in one layer is connected to all other neurons in the next layer, but no neurons in the same layer are connected (see Fig. 1a). The ANN model includes information from the environment in the input layer of the mechanism, information processing is in the hidden layer,

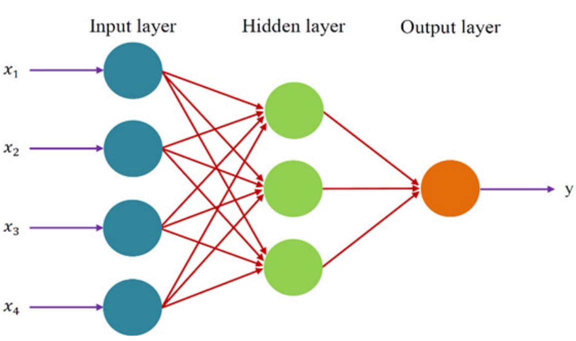

Fig. 1 Network diagram

and continuous values are calculated at the output layer. The ANN model consists of two modes of operation, training and usage. A data collection of inputs and outputs is used to train the ANN model on how to estimate output variables. Starting with weights that are chosen at random, supervised learning establishes the weights that will be used for the current task. The error function of the network must be reduced for a given training set made up of *k*-ordered pairs of *n* inputs and *m* outputs, also referred to as the input and output patterns. The back-error learning process is typically implemented using supervised learning techniques like back propagation neural networks. The ANN model was created and validated using the multi-layer perceptron (MLP) module. To minimize error, the MLP networks are trained using a BP technique that uses gradient descent to update weights.

The dataset was divided into three subgroups: training (70%), testing (20%), and holdout (10%). The testing data has never been given to the network models and has mainly been utilized to analyze the network performance. The weights are found taking into account the training data and then the model is created. Then, the errors of the model are computed by taking into consideration the testing part of the data, and over-training is also prevented during the training phase. The ANN model is validated by utilizing the holdout data. For both the hidden and output layers, the sigmoid function was used as an activation function. The optimal solution in the gradient descent optimization process can be reached by using either stack mode or online mode. The online training algorithm continually receives a record and the weights are updated until one of the stop rules is encountered. The process continues with the recycling of data records when all records are utilized once and none of the stop rules are followed. When the output layer is constructed with the sigmoid activation function, the squared error function is used to reduce error.

The weighted activation in the input variables is collected only by a single processing unit and this is converted into a total activation function and transferred to the resulting output variable. Therefore, in the ANNs, information processing is usually composed of units that transmit inputs to the output, and then the weights of the links are arranged as inputs for other units. Learning in the corresponding systems can be defined as the total adjustment of the weights to convert the output of the network to the required output, without being exposed to any changes in the network structure. It is taken into account that information is managed by the optimized link weights between processing units in the entire artificial system when learning is fully realized.

As opposed to conventional knowledge systems, where required rules need to be clearly defined, ANNs are attractive tools in the use of pattern recognition and classification systems. The input layer here consists of 4 real factors, location, plant part, and two of the heavy metals, which are thought to be related to the output of interest. Notice that for the three cases that occur, two of the three heavy metals are chosen alternately among themselves. The neural network model considered in the estimation of the effects of environmental parameters on the Cd, Cr, and Pb levels is shown in Fig. 1a–c. The hidden layer has been understood to be optimum and highly satisfactory with four processing units taken based on trial and error. As mentioned in the following section, there is a substantial link between the input and output factors. The data model consisted of location, plant part, Cr, and Pb, as well as the Cd parameter collected from 800 samples. The four input parameters have been used to separately investigate the effect on each of the three metals (Cd, Cr, or Pb). The existence of relationships between the parameters is reported in the section below.

Results and discussion

Case 1

This part of the study aims for researchers to use *R. pseudoacacia* as a biomonitor at different pollution levels and calculate Cd heavy metal values using Cr and Pb heavy metal data in their studies without the need to read them. By using the ANN model, finding the Cd heavy metal concentration including the plant parts in mg kg^{-1} unit with the specified sensitivity will significantly reduce the analysis times. Information on the data considered in the production of the neural network model is exhibited in Table 1. The neurons used in each of the input (4 neurons), hidden (3 neurons), and output (1 neuron) layers and four input variables are given in Table 2. The hidden layer of the network architecture requires 3 nodes; however, the output layer only needs 1 node to encode the output variable. The most commonly used activation

Table 1 Case processing summary

		Case 1		Case 2		Case 3	
		N	Percent	*N*	Percent	*N*	Percent
Sample	Training	561	70.1%	561	70.1%	537	67.1%
	Testing	156	19.5%	156	19.5%	173	21.6%
	Holdout	83	10.4%	83	10.4%	90	11.3%
Valid		800	100.0%	800	100.0%	800	100.0%
Excluded		0		0		0	
Total		800		800		800	

functions are the rectified linear unit, the sigmoid, and the hyperbolic tangent functions. Since the use of the sigmoid function produced slightly better results than the others during the preparatory observations, the sigmoid function has been considered for both hidden and output layers in this study.

A very important model has been discovered between the four inputs and the output, Cd, as shown in Tables 3–5. The tables exhibit the predicted results of the pollution based on environmental parameters compared to the actual evaluations. Note that the derived results from the testing data are very significant which means that the performance from the training data is very close to the performance from the testing data. As seen in Figs. 2 and 3, it turns out that the predicted values are in very good agreement with the actual values. The sensitivity level of the model parameters is shared for each in Fig. 4a. The model has been created using relative errors for the training, testing, and holdout phases calculated as 0.019, 0.020, and 0.014, respectively. The result in which the Cd factor was estimated from the ANN model revealed that the proposed model agreed with realistic behaviors of nature.

As with almost all pollutions, there is a strong correlation between environmental pollution and heavy metals in the literature (Aslam et al., 2020; Wang et al., 2020). For this reason, it cannot be denied that a quick and accurate assessment of Cd pollution is essential for living things, economies, and even political decisions. Therefore, the present part of the study focuses on the generation of the model that deals with estimating the effects of exogenous variables on Cd pollution. It is understood that the predictability of the models (Tables 3–5) regarding the output evaluations is quite high. Note that each exogenous variable of the ANN model has varying influence levels, as seen in Table 5. Explanatory and satisfactory information

Table 2 Network information

			Case 1	Case 2	Case 3
Input layer	Covariates	1	Location	Location	Location
		2	Plant part	Plant part	Plant part
		3	Cr	Cd	Cd
		4	Pb	Pb	Cr
	Number of units		4	4	4
	Rescaling method for covariates		Normalized	Normalized	Normalized
Hidden layer(s)	Number of hidden layers		1	1	1
	Number of units in hidden layer 1		3	3	3
	Activation function		Sigmoid	Sigmoid	Sigmoid
Output layer	Dependent variables		Cd	Cr	Pb
	Number of units		1	1	1
	Rescaling method for scale dependents		Normalized	Normalized	Normalized
	Activation function		Sigmoid	Sigmoid	Sigmoid
	Error function		Sum of squares	Sum of squares	Sum of squares

Table 3 Parameter estimates

	Case 1—predicted				
	Predictor	Hidden layer 1			Output layer
		H (1:1)	H (1:2)	H (1:3)	Cd
Input layer	(Bias)	1.108	−0.349	5.413	
	Location	−1.022	−0.197	0.261	
	Plant part	−4.409	−3.914	−3.457	
	Cr	−3.618	−2.185	−2.084	
	Pb	−3.523	−3.001	−1.792	
Hidden layer 1	(Bias)				3.362
	H (1:1)				−2.787
	H (1:2)				0.764
	H (1:3)				−6.171

	Case 2—predicted				
	Predictor	Hidden layer 1			Output layer
		H (1:1)	H (1:2)	H (1:3)	Cr
Input layer	(Bias)	1.564	-0.549	3.678	
	Location	−0.580	−2.067	−2.798	
	Plant part	−1.406	−2.967	1.400	
	Cd	−3.984	−2.048	−1.579	
	Pb	−3.484	−2.270	−1.665	
Hidden layer 1	(Bias)				3.275
	H (1:1)				−3.703
	H (1:2)				0.731
	H (1:3)				−4.045

	Case 3—predicted				
	Predictor	Hidden layer 1			Output layer
		H (1:1)	H (1:2)	H (1:3)	Pb
Input layer	(Bias)	0.360	0.772	0.219	
	Location	−1.463	−0.689	0.122	
	Plant part	−2.374	−2.831	−0.158	
	Cd	2.353	−3.128	−0.516	
	Cr	2.392	−3.527	−1.270	
Hidden layer 1	(Bias)				−0.942
	H (1:1)				4.905
	H (1:2)				−5.150
	H (1:3)				−1.356

about the results of the different stages in the network model is presented in Table 4. Pay enough attention to the generation of synaptic weights by the training example, even if the current dataset is divided into three data categories. The synaptic weights correspond to the strength of the ties between the two nodes like the attempt of two different brain cells to communicate through synapses. The ability of the model to predict pollution has been successfully demonstrated based on a significantly small error of 0.278 (see Table 4).

Case 2

It is aimed to model the effect of soil and plant materials on environmental pollution and thus predict the Cr pollution for input variables; location, plant part, Cd, and Pb in this case. Predicting the Cr heavy metal concentration with the required sensitivity with the help of ANN significantly eliminates the need for long analysis times and high experimental costs. Table 1 lists the data that were used to build the network model and provides information on them. The model architecture is used here the same as in case 1 (see Table 2).

After completing the training and testing stages, as shown in Tables 3–5, a significantly strong relationship has been found between the four feasible factors and the Cr factor. The predicted results of pollution based on environmental parameters are presented in the corresponding tables in comparison with the actual assessments. The fact that the results from the testing data are very significant means that the performance from the training data is very close to the performance from the testing data. For the training, testing, and holdout stages, the relative errors of the Cr variable are 0.015, 0.017, and 0.012, respectively. As in case 1, when the Cr factor is predicted from the ANN model, it turns out that a very realistic and significant result is reached. It can be seen in Figs. 2, 3, and 4b that the estimated values have very good agreement with the actual ones.

Case 3

In this case, the Pb is the output to be determined by its values through the ANN model. So, the authors have aimed to estimate Pb heavy metal values in a very cost-effective and very accurate way by using Cd

Fig. 2 Predicted Cd value. **a** Cd, case 1; **b** Cr, case 2; **c** Pb, case 3

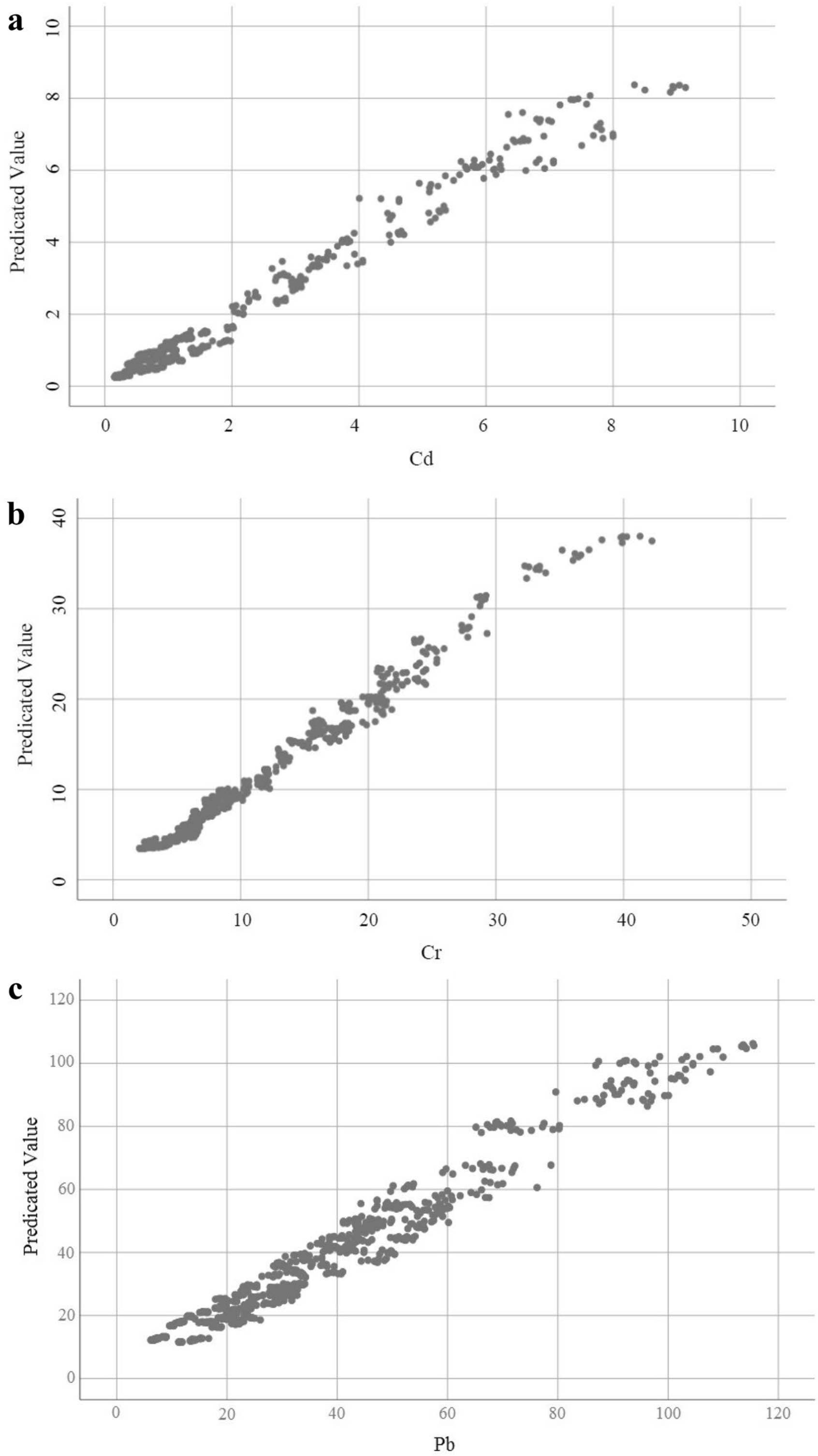

Fig. 3 Predicted residual value. **a** Cd, case 1; **b** Cr, case 2; **c** Pb, case 3

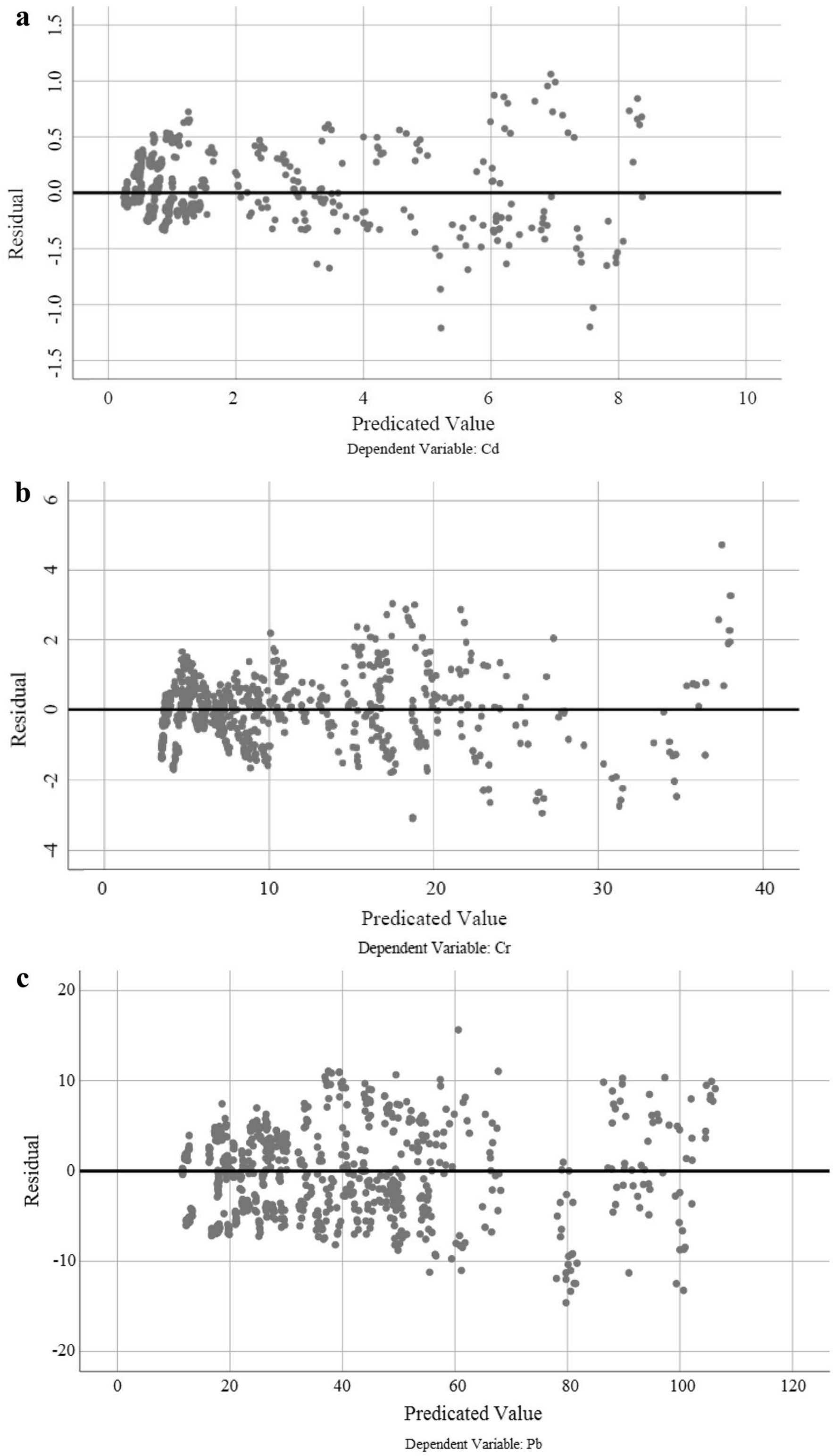

Fig. 4 Normalized importance. **a** Cd, case 1; **b** Cr, case 2; **c** Pb, case 3

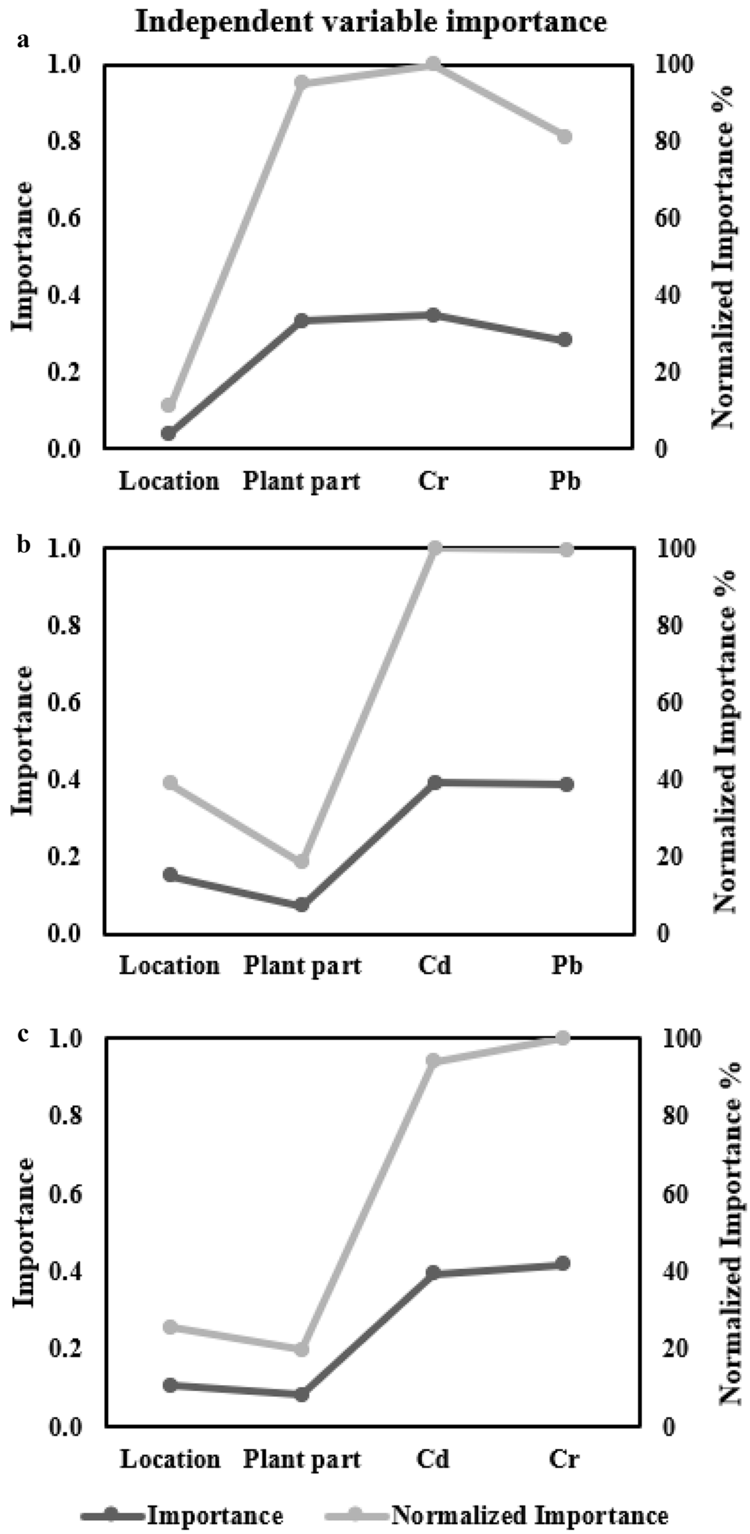

Table 4 Model summary

		Case 1	Case 2	Case 3
Training	Sum of squares error	0.278	0.168	0.583
	Relative error	0.019	0.015	0.042
	Stopping rule used	10 consecutive step(s) with no decrease in error[a]	10 consecutive step(s) with no decrease in error[a]	10 consecutive step(s) with no decrease in error[a]
	Training time	0:00:00:12	0:00:00:03	0:00:00:02
Testing	Sum of squares error	0.0465	0.044	0.199
	Relative error	0.020	0.017	0.048
Holdout	Relative error	0.014	0.012	0.024
Dependent variable		Cd	Cr	Pb

[a]Error computations are based on the testing sample

and Cr heavy metal data and other variables in this part of the study. While doing this, it will be possible to accurately predict the Pb by using the heavy metal concentrations of Cd and Cr, depending on our model, thus significantly reducing the analysis times. It is undeniable that the acquisition of two analysis data and the third analysis data for biomonitor studies contributes to a wider range, increasing the quality of publication and interpretation. Detailed information about the data taken into account in the creation of our neural network model is presented in Table 1. As in the previous cases, the neurons used in each layer and the four input variables are presented in Table 2. The hidden layer in the suggested network architecture requires three nodes, whereas the output layer only requires one node to encode the output variable. The sigmoid function is also supported for both the hidden and output layers.

A very important model has been discovered between the four inputs and the output, Pb, as shown in Tables 3–5. As for the prediction of the output factor Pb, the predicted pollution results based on environmental parameters compared to actual assessments are presented in the tables. The results produced reveal that the performance from the training data is in very good agreement with the performance from the testing data. The relative errors are 0.042, 0.048, and 0.024 for the three phases, training, testing, and holdout, respectively. The results of this statement, in which the Cd output has been estimated with the help of the ANN model, showed that the proposed model is in very good agreement with the realistic behavior of nature. It should be noted that the parameters in the data are the same here as before, except that the roles of the output, Pb, and one of the other metals are mutually altered. It is seen that the predicted values are in very good agreement with the actual values as seen in Figs. 2 and 3. The sensitivity level of the model variables is given for each in Fig. 4b.

According to research (Aslam et al., 2020; Chen et al., 2021; Wang et al., 2020; Yang et al., 2020), there is a strong relationship between environmental pollution and heavy metals. Therefore, it is thought that it is crucial to evaluate the heavy metal (Cd, Cr, or Pb) pollution quickly and efficiently for all living things, economies, and even political decisions. Therefore, this study has concentrated on estimating the effects of the input factors on heavy metal (Cr, Cd, or Pb) pollution. The most important finding obtained from all results is that there is a very significant relation between the input variables and the Cd, Cr, or Pb pollution. In the ANN model, it is seen that the predictability of the sample patterns (Tables 3–5) regarding the output evaluations is quite high with the varying degree of influence of each independent variable. Note that each independent variable in the ANN model has varying effect levels, as seen in Table 5. Quite satisfactory information about the results of training, testing, and holdout samples on the network model is summarized in Table 4. Even though the present data collection is divided into three kinds of data, the training example still results in synaptic weights. The table shows the sum of squares error for both the training and testing samples. This research intends to reveal and apply the attractive ability of the network model with a challenging study in environmental systems. For all cases, the backpropagation MLP has here adopted as an ANN model for the Cd, Cr, or Pb pollution model estimation.

Table 5 Independent variable importance

		Importance	Normalized importance
Case 1	Location	0.038	10.9%
	Plant part	0.331	95.1%
	Cr	0.348	100.0%
	Pb	0.282	81.0%
Case 2	Location	0.151	38.8%
	Plant part	0.072	18.4%
	Cd	0.390	100.0%
	Pb	0.387	99.3%
Case 3	Location	0.106	25.4%
	Plant part	0.082	19.7%
	Cd	0.393	93.9%
	Cr	0.418	100.0%

Many studies in the literature (Gupta et al., 2022; Shang & Xu, 2021; Yalcin et al., 2020) point to the necessity and importance of measuring environmental pollution. This idea encouraged us to estimate the relationship between the environmental parameters and the Cd, Cr, or Pb pollution. This research intends to reveal and apply the attractive ability of the network model with a challenging study in environmental systems. The backpropagation MLP has been accepted as the ANN model for the pollution model estimation.

Conclusions

In this paper, important neural network models have been created to predict the level of environmental pollution by producing an innovative perspective and finding, which is important for the development of environmental protection and sustainable development. It is highly anticipated that the generated network models will aid in determining the extent of environmental contamination processes, assessing the socioeconomic and environmental repercussions under various ecosystem situations, and helping to establish the necessary environmental management strategies. From this point of view, the pollution level of heavy metals based on plant and soil information from an ecosystem has been discovered here. At a remarkably low cost, unlike traditional approaches, the ANN model has therefore accomplished for the first time in analyzing heavy metal pollution levels through plant and soil parameters: location, plant parts, and any two of the heavy metals (Cd, Cr, or Pb). This study has reported the latest research results through the ANN modelling for environmental systems. The acquisition of two analysis data and the third analysis data will also contribute to increasing the quality of publication and interpretation with a wider spectrum for biomonitor studies. The simulation results have proven the superiority of the network algorithm over the other competitive traditional counterparts. The ANN has been seen to be very successful in optimizing pollution assessments. The results revealed that this model should be expected to provide undeniable savings in device consumable expenses such as nitrogen and argon gas use, acid usage, analysis, and storage tube costs. It has been concluded that findings are environmentally expected to be very useful for environmentalists and scientists in organizing proper work programs for their institutions. In modelling the behavior of such a variety of realistic problems that relate to a wide range of science, it is believed that it is important to evaluate heavy metal pollution in a fast and optimum way for all living things and hence economies and even political behaviors in daily life. The results produced are original, comprehensible, and optimal, and therefore it is unavoidable that interested readers will draw much more attention to modelling such challenging environmental problems. For future research, this study can be carried out for data based on more comprehensive variables. In the forthcoming works, more comprehensive and thus more illustrative results may be found with various datasets.

Acknowledgements The authors are thankful to Dr. M. E. Uras (Marmara University, Türkiye) for his technical support.

Author contribution Murat Sari: validation; writing—review and editing; supervision. Ibrahim Ertugrul Yalcin: conceptualization, methodology, investigation. Mahmut Taner: methodology; writing—original draft. Tahir Cosgun: formal analysis; writing—original draft. Ibrahim Ilker Ozyigit: validation, resources, visualization, supervision.

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval All authors have read, understood, and complied as applicable with the statement on "Ethical responsibilities of Authors" as found in the Instructions for Authors and are aware that with minor exceptions, no changes can be made to authorship once the paper is submitted.

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

Conflict of interest The authors declare no competing interests.

References

Ahmed, N., Islam, M. N., Tuba, A. S., Mahdy, M. R. C., & Sujauddin, M. (2019). Solving visual pollution with deep learning: A new nexus in environmental management. *Journal of Environmental Management, 248*, 109253. https://doi.org/10.1016/j.jenvman.2019.07.024

Aslam, S., Chan, M. W. H., Siddiqui, G., Boczkaj, G., Kazmi, S. J. H., & Kazmi, M. R. (2020). A comprehensive assessment of environmental pollution by means of heavy metal analysis for oysters' reefs at Hab River Delta, Balochistan, Pakistan. *Marine Pollution Bulletin, 153,* 110970. https://doi.org/10.1016/j.marpolbul.2020.110970

Beccali, M., Bonomolo, M., Ciulla, G., & Brano, V. L. (2018). Assessment of indoor illuminance and study on best photosensors' position for design and commissioning of Daylight Linked Control systems. A new method based on artificial neural networks. *Energy, 154*, 466–476. https://doi.org/10.1016/j.energy.2018.04.106

Cabaneros, S. M. S., Calautit, J. K. S., & Hughes, B. R. (2017). Hybrid artificial neural network models for effective prediction and mitigation of urban roadside NO2 pollution. *Energy Procedia., 142*, 3524–3530. https://doi.org/10.1016/j.egypro.2017.12.240

Cen, Z., & Wang, J. (2018). Forecasting neural network model with novel CID learning rate and EEMD algorithms on energy market. *Neurocomputing, 317*, 168–178. https://doi.org/10.1016/j.neucom.2018.08.021

Chen, Y. G., Huang, J. H., Luo, R., Ge, H. Z., Wołowicz, A., Wawrzkiewicz, M., & Chen, S. H. (2021). Impacts of heavy metals and medicinal crops on ecological systems, environmental pollution, cultivation, and production processes in China. *Ecotoxicology and Environmental Safety*, *219*, 112336.

Corradini, F., Meza, P., Eguiluz, R., Casado, F., Huerta-Lwanga, E., & Geissen, V. (2019). Evidence of microplastic accumulation in agricultural soils from sewage sludge disposal. *Science of the Total Environment., 671*, 411–420. https://doi.org/10.1016/j.scitotenv.2019.03.368

Cristaldi, A., Conti, G. O., Cosentino, S. L., Mauromicale, G., Copat, C., Grasso, A., & Ferrante, M. (2020). Phytoremediation potential of *Arundo donax* (giant reed) in contaminated soil by heavy metals. *Environmental Research,* 109427. https://doi.org/10.1016/j.envres.2020.109427

Crouse, D. L., Goldberg, M. S., & Ross, N. A. (2009). A prediction-based approach to modelling temporal and spatial variability of traffic-related air pollution in Montreal Canada. *Atmospheric Environment, 43*(32), 5075–5084. https://doi.org/10.1016/j.atmosenv.2009.06.040

Cui, K. P., Dai, R. R., Liu, X., Weerasooriya, R., Hong, Z. Y., Chen, X., & Wu, Y. C. (2020). New strategy for fabricating Cd (II) sensing electrochemical interface based on enhanced adsorption followed by redox processes: Ferrocerium oxide nanocomposite as an example. *Journal of Alloys and Compounds,* 154551. https://doi.org/10.1016/j.jallcom.2020.154551

Favi, C., Marconi, M., Mandolini, M., & Germani, M. (2022). Sustainable life cycle and energy management of discrete manufacturing plants in the industry 4.0 framework. *Applied Energy*, *312*, 118671.

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* Vol. 36 (p. 1–8). Orlando Florida.

Greipsson, S., McElroy, T., & Koether, M. (2022). Effects of supplementary nutrients (soil-nitrogen or foliar-iron) on switchgrass (*Panicum virgatum* L.) grown in Pb-contaminated soil. *Journal of Plant Nutrition*, 1–12.

Guo, Y., Liu, R., Liu, Y., Xiang, D., Liu, Y., Gui, W., & Zhu, G. (2018). A non-competitive surface plasmon resonance immunosensor for rapid detection of triazophos residue in environmental and agricultural samples. *Science of the Total Environment., 613*, 783–791. https://doi.org/10.1016/j.scitotenv.2017.09.157

Gupta, A., Kumar, M., Ghosh, P., & Thakur, I. S. (2022). Risk assessment of a municipal extended aeration activated sludge treatment plant using physico-chemical and in vitro bioassay analyses. *Environmental Technology & Innovation*, 102254.

Hocaoglu-Ozyigit, A., & Genc, B. N. (2020). Cadmium in plants, humans and the environment. *Frontiers in Life Sciences and Related Technologies, 1*(1), 12–21.

Hsieh, K. L. (2011). Employing a recommendation expert system based on mental accounting and artificial neural networks into mining business intelligence for study abroad's P/S recommendations. *Expert Systems with Applications., 38*(12), 14376–14381. https://doi.org/10.1016/j.eswa.2011.03.078

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. https://doi.org/10.1016/j.chemosphere.2020.126196

Krstic, V., Urosevic, T., & Pesovski, B. (2018). A review on adsorbents for treatment of water and wastewaters containing copper ions. *Chemical Engineering Science, 192*, 273–287.

Likus-Cieslik, 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. *Environmental Pollution, 258,* 113559. https://doi.org/10.1016/j.envpol.2019.113559
Luo, X., & Yang, J. (2019). A survey on pollution monitoring using sensor networks in environment protection. *Journal of Sensors,* 2019. https://doi.org/10.1155/2019/6271206
MacKerron, G., & Mourato, S. (2009). Life satisfaction and air quality in London. *Ecological Economics., 68*(5), 1441–1453. https://doi.org/10.1016/j.ecolecon.2008.10.004
Martin, G. D. (2019). Addressing geographical bias: A review of *Robinia pseudoacacia* (black locust) in the Southern Hemisphere. *South African Journal of Botany., 125*, 481–492. https://doi.org/10.1016/j.sajb.2019.08.014
Matsui, K., Ishimura, T., Mattonai, M., Iwai, I., Watanabe, A., Teramae, N., & Watanabe, C. (2020). Identification algorithm for polymer mixtures based on Py-GC/MS and its application for microplastic analysis in environmental samples. *Journal of Analytical and Applied Pyrolysis,* 104834. https://doi.org/10.1016/j.jaap.2020.104834
Mitra, S., Chakraborty, A. J., Tareq, A. M., Emran, T. B., Nainu, F., Khusro, A., & Simal-Gandara, J. (2022). Impact of heavy metals on the environment and human health: Novel therapeutic insights to counter the toxicity. *Journal of King Saud University-Science*, 101865.
Niel, O., & Bastard, P. (2019). Artificial intelligence in nephrology: Core concepts, clinical applications, and perspectives. *American Journal of Kidney Diseases, 74*(6), 803–810. https://doi.org/10.1053/j.ajkd.2019.05.020
Nourani, V., Gokcekus, H., Umar, I. K., & Najafi, H. (2020). An emotional artificial neural network for prediction of vehicular traffic noise. *Science of The Total Environment, 707,* 136134. https://doi.org/10.1016/j.scitotenv.2019.136134
Ozyigit, I. I., Dogan, I., Igdelioglu, S., Filiz, E., Karadeniz, S., & Uzunova, Z. (2016). Screening of damage induced by lead (Pb) in rye (*Secale cereale* L.) a genetic and physiological approach. *Biotechnology & Biotechnological Equipment, 30*(3), 489–496. https://doi.org/10.1080/13102818.2016.1151378
Pemberthy, M. D., Padilla, Y., & Penuela, G. A. (2021). Assessment of chromium, lead and mercury in seawater and some fish species from the Gulf of Uraba, Colombian Caribbean: A baseline study. *Environmental Science and Pollution Research, 28*(13), 16804–16815.
Peng, C., Wang, M., & Chen, W. (2016). Modelling cadmium contamination in paddy soils under long-term remediation measures: Model development and stochastic simulations. *Environmental Pollution, 216*, 146–155. https://doi.org/10.1016/j.envpol.2016.05.038
Pulscher, L. A., Gray, R., McQuilty, R., Rose, K., Welbergen, J., & Phalen, D. N. (2020). Investigation into the utility of flying foxes as bioindicators for environmental metal pollution reveals evidence of diminished lead but significant cadmium exposure. *Chemosphere,* 126839. https://doi.org/10.1016/j.chemosphere.2020.126839
Rasool, B., Ramzani, P. M. A., Zubair, M., Khan, M. A., Lewińska, K., Turan, V., & Iqbal, M. (2021). Impacts of oxalic acid-activated phosphate rock and root-induced changes on Pb bioavailability in the rhizosphere and its distribution in mung bean plant. *Environmental Pollution*, *280*, 116903.
Sari, M., & Cetiner, B. G. (2009). Predicting effect of physical factors on tibial motion using artificial neural networks. *Expert Systems with Applications, 36*(6), 9743–9746. https://doi.org/10.1016/j.eswa.2009.02.030
Sari, M., Yalcin, I. E., Taner, M., Cosgun, T., & Ozyigit, I. I. (2022). An investigation on environmental pollution due to essential heavy metals: A prediction model through multilayer perceptrons. *International Journal of Phytoremediation*, 1–9. https://doi.org/10.1080/15226514.2022.2059056
Serbula, S. M., Miljkovic, D. D., Kovacevic, R. M., & Ilic, A. A. (2012). Assessment of airborne heavy metal pollution using plant parts and topsoil. *Ecotoxicology and Environmental Safety, 76*, 209–214. https://doi.org/10.1016/j.ecoenv.2011.10.009
Shah, P. J., Anagnostopoulos, T., Zaslavsky, A., & Behdad, S. (2018). A stochastic optimization framework for planning of waste collection and value recovery operations in smart and sustainable cities. *Waste Management, 78*, 104–114. https://doi.org/10.1016/j.wasman.2018.05.019
Shang, Y., & Xu, A. (2021). An evaluation of the impact of natural ecotourism on environmental pollution. *Environmental Science and Pollution Research, 28*(26), 33764–33770.
Sim, J. J., Tan, G. W. H., Wong, J. C., Ooi, K. B., & Hew, T. S. (2014). Understanding and predicting the motivators of mobile music acceptance–a multi-stage MRA-artificial neural network approach. *Telematics and Informatics, 31*(4), 569–584. https://doi.org/10.1016/j.tele.2013.11.005
Sun, C., Wang, D., Shen, X., Li, C., Liu, J., Lan, T., & Zhang, Y. (2020). Effects of biochar, compost and straw input on root exudation of maize (*Zea mays* L.): From function to morphology. Agriculture. *Ecosystems & Environment, 297,* 106952. https://doi.org/10.1016/j.agee.2020.106952
Tzvetkova, N., & Petkova, K. (2015). Bioaccumulation of heavy metals by the leaves of *Robinia pseudoacacia* as a bioindicator tree in industrial zones. *Journal of Environmental Biology., 36*(1), 59.
Vaseashta, A., Gevorgyan, G., Kavaz, D., Ivanov, O., Jawaid, M., & Vasović, D. (2021). Exposome, biomonitoring, assessment and data analytics to quantify universal water quality. In *Water Safety, Security and Sustainability* (pp. 67–114). Springer, Cham.
Vijayakumar, C. R., Balasubramani, D. P., & Azamathulla, H. M. (2022). Assessment of groundwater quality and human health risk associated with chromium exposure in the industrial area of Ranipet, Tamil Nadu, India. *Journal of Water, Sanitation and Hygiene for Development, 12*(1), 58–67.
Wang, S., Liu, J., Li, J., Xu, G., Qiu, J., & Chen, B. (2020). Environmental magnetic parameter characteristics as indicators of heavy metal pollution in the surface sediments off the Zhoushan Islands in the East China Sea. *Marine pollution bulletin, 150,* 110642. https://doi.org/10.1016/j.marpolbul.2019.110642
Wolski, G. J., & Kruk, A. (2020). Determination of plant communities based on bryophytes: The combined use of Kohonen artificial neural network and indicator species analysis. *Ecological Indicators, 113,* 106160. https://doi.org/10.1016/j.ecolind.2020.106160
Yalcin, I. E., Ozyigit, I. I., Dogan, I., Demir, G., & Yarci, C. (2020). Using the Turkish red pine tree to monitor heavy metal pollution. *Polish Journal Environmental Studies, 29*(5), 1–9. https://doi.org/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. *Ecotoxicology and Environmental Safety, 197,* 110622. https://doi.org/10.1016/j.ecoenv.2020.110622