Environmental Science and Pollution Research (2020) 27:42495–42512
https://doi.org/10.1007/s11356-020-10156-w

RESEARCH ARTICLE

Application of artificial neural networks to predict the heavy metal contamination in the Bartin River

Handan Ucun Ozel[1] · Betul Tuba Gemici[1] · Ercan Gemici[2] · Halil Baris Ozel[3] · Mehmet Cetin[4] · Hakan Sevik[5]

Received: 28 March 2020 / Accepted: 15 July 2020 / Published online: 24 July 2020

Abstract
In this study, copper (Cu), iron (Fe), zinc (Zn), manganese (Mn), nickel (Ni), and lead (Pb) analyses were performed, and the results were modelled by artificial neural networks (ANN) and adaptive neuro-fuzzy inference system (ANFIS). Samples were taken from 3 stations selected on the Bartin River for 1 year between December 2012 and December 2013. Radial basis neural network (RBANN), multilayer perceptron (MLP) neural networks models, and adaptive neuro-fuzzy inference system (ANFIS) were applied to the data in order to predict the heavy metal concentrations. As a result of the study, the RMSE and MAE values of all the heavy metal models were found to have very low error values during the test phase, and it was found that the models created using MLP had R^2 values higher than 0.77 during the test phase; the test phase R^2 values of the models using RBN method were found to be ranging between 0.773 and 0.989, and the test phase R^2 value of the ANFIS model was higher than 0.80. If sorted from the best model to the worst by taking the MAE and RMSE values into consideration based on the test evaluation results, according to the heavy metal types, where all of the MLP, RBN, and ANFIS models were generally approximate to each other, RBN was successful for Cu, Zn, and Mn, while MLP model was successful for Ni and ANFIS model for Fe and Pb. According to the results, it can be inferred that the heavy metal contents can be estimated approximately with artificial intelligence models and relatively easy-to-measure parameters; it will be possible to detect heavy metals which are harmful to the viability of the rivers, both quickly and economically.

Keywords ANN · River · ANFIS model · Heavy metal · Contamination · Bartin River

Introduction

Today, water streams such as rivers are subject to intensive pollution due to the growth of industrial activities and the rapid increase in population. Chemical pollution of rivers in particular is one of the biggest threats to human health and aquatic ecosystems (Magdaleno et al. 2014). Heavy metals are discharged into the rivers as a result of natural processes and anthropogenic activities. Many metals such as Cu, Fe, Mn, Ni, and Zn are essential for human life and are beneficial to health at low concentrations, but higher concentrations can be toxic and contaminate the environment, and may endanger the whole ecosystem in return (Shanbehzadeh et al. 2014; Prabu 2009; Ozel et al. 2019; Alizamir and Sobhanardakani 2017). However, metals such as As, Cd, Cr, Hg, Pb, and Sn have no known basic functions in living organisms and are toxic even at low concentrations (Alizamir and Sobhanardakani 2017). The most notable property of heavy metals is bioaccumulation and their long biological half-lives (Alizamir and Sobhanardakani 2017, 2016; Sevik et al. 2019a, b; Cetin 2019; Cetin et al. 2019). This is why heavy metals are infamously toxic and hazardous (Alizamir and Sobhanardakani 2017; Hoha et al. 2014; Turkyilmaz et al. 2018a, b; Cetin et al. 2020; Sevik et al. 2020a, b, c; Alaqouri et al. 2020). These heavy metals enter the human body from the food

Responsible Editor: Marcus Schulz

✉ Mehmet Cetin
mcetin@kastamonu.edu.tr

1 Faculty of Engineering, Architecture and Design, Department of Environmental Engineering, Bartin University, Bartin, Turkey

2 Faculty of Engineering, Architecture and Design, Department of Civil Engineering, Bartin University, Bartin, Turkey

3 Faculty of Forestry, Department of Forest Engineering, Bartin University, Bartin, Turkey

4 Faculty of Engineering and Architecture, Department of Landscape Architecture, Kastamonu University, Kastamonu, Turkey

5 Faculty of Engineering and Architecture, Department of Environmental Engineering, Kastamonu University, Kastamonu, Turkey

chain, and can cause both acute and chronic effects in humans (Dokmeci 2017; Aricak et al. 2019; Aricak et al. 2020; Sevik et al. 2020a, b, c). For this reason, a thorough investigation on the heavy metal contamination of water resources (rivers, lakes, and groundwater) and the determination of pollutant sources have gained a lot of attention worldwide (Elzwayie et al. 2017; Alizamir and Sobhanardakani 2017; Lu et al. 2019).

However, monitoring the pollution levels of rivers by using direct measurement methods requires a lot of labor, equipment, and time. For a good management of water resources, river pollution should be well defined, monitored, and modelled. Modelling is an alternative solution to reduce water quality monitoring cost. In general, predictions produced by the models are based on limited sampling information of pollutant concentrations (Nhantumbo et al. 2018). Artificial neural networks (ANN) is an essential calculation method to predict and model complex relationships between data, even without any significant relationship between parameters. Compared with traditional methods, ANN allows inaccurate or incomplete data and approximate results, and they are less sensitive to outliers. However, when modelling studies were examined, ANFIS was preferred for its ability to establish reliable models quickly without the need for expert knowledge with. ANN and ANFIS have recently been successfully applied in environmental sciences such as water resources, water quality, and hydrologic time series (Singh et al. 2009; Sarkar and Pandey 2015; Küçükerdem et al. 2019).

There are publications on rivers trying to predict daily flow data using artificial neural networks, ANFIS, and multiple non-linear regression models, and that ANFIS model gives better results than other models (Altunkaynak and Başakın 2018; Anusree and Varghese 2016).

In many studies, ANN predicts heavy metal concentrations (As, Pb Cu, Fe, Mn, Zn, etc.) in both rivers and groundwater. ANN, multilayer sensor (MLP), backpropagation neural network (BPNN), general regression neural network (GRNN), and multiple linear regression (MLR) models are used. It was established that the ANN model was an efficient application for heavy metal forecasting. Many researchers have reported that the results were suitable for predicting heavy metals in a fast and cost-effective manner (Rooki et al. 2011; Abdolmaleki et al. 2013; Altunkaynak and Başakın 2018; Alizamir and Sobhanardakani 2016; Anusree and Varghese 2016; Alizamir and Sobhanardakani (2017). Also in numerous studies carried out in different regions, some results regarding the use of ANN in monitoring water quality were obtained (Arhami et al. 2013; Azid et al. 2014; Chelani et al. 2002; Venkatramanan et al. 2017).

Many studies have researched heavy metal concentrations in dissolved and river sediments (Swietlicka et al. 2017; Ozel et al. 2019; Lu et al. 2019). It has been revealed in studies that physicochemical and environmental factors may affect heavy metal concentrations. Therefore, pH, COD, SS, etc. parameters are used as input data for the prediction of heavy metal concentrations (Lu et al. 2019). Heavy metal concentrations and other environmental indices should be constantly monitored to comprehensively understand the behavior and associated risks of heavy metals in aquatic environments; however, it should be reiterated that regular and integrated monitoring is time-consuming and costly, and integrated monitoring systems that take into account both heavy metals and other relevant environmental indices are often lacking (Lu et al. 2019).

Because of the danger they pose, heavy metals should be investigated in drinking and utility waters. The Bartin River, the subject of this study, feeds underground waters used as drinking water. The river is also used for irrigation purposes and constitutes the recreation area of the city. Bartin River has an important place in the city. The determination of the quality of the water resources is important for all areas today, where the water resources are rapidly depleted and quality water is lacking. Heavy metals have a high toxicity and are thought to be found in large amounts in the Bartin River. In our country, four different categories were determined for the quality of surface waters and quality classes were created. Accordingly, the 1st class represents the high-quality water group, the 2nd class represents less polluted, the 3rd class contaminated, and the 4th class highly contaminated water group. High water quality limit value and little contaminated water quality value in our country is set between 20 and 50 μg/L for Cu, 300 and 1000 μg/L for Fe, and 200 and 500 μg/L for Zn. It was determined as 100 μg/L and 500 μg/L for Mn, 20 μg/L and 50 μg/L for Ni, and 10 μg/L and 20 μg/L for Pb, respectively (Official Gazette 2012)

The aim of this study is to estimate certain parameters by using artificial neural networks and adaptive neuro-fuzzy inference system (ANFIS) models, by determining the quality of water resources, and to develop methods to reduce the number of parameters to be analyzed. Estimating heavy metals in Bartin River requires more special analysis than the parameters that are frequently analyzed by local governments. For this purpose, samples were taken from 3 stations selected on the Bartin River for 1 year between December 2012 and December 2013. Some of the heavy metals, such as copper (Cu), iron (Fe), zinc (Zn), manganese (Mn), nickel (Ni), and lead (Pb) were subjected to analyses, and the obtained values were modelled by using artificial neural networks and adaptive neuro-fuzzy inference system.

Materials and methods

Study area

Surface water samples were collected from 3 sampling sites throughout the Bartin River. Figure 1 shows the location of

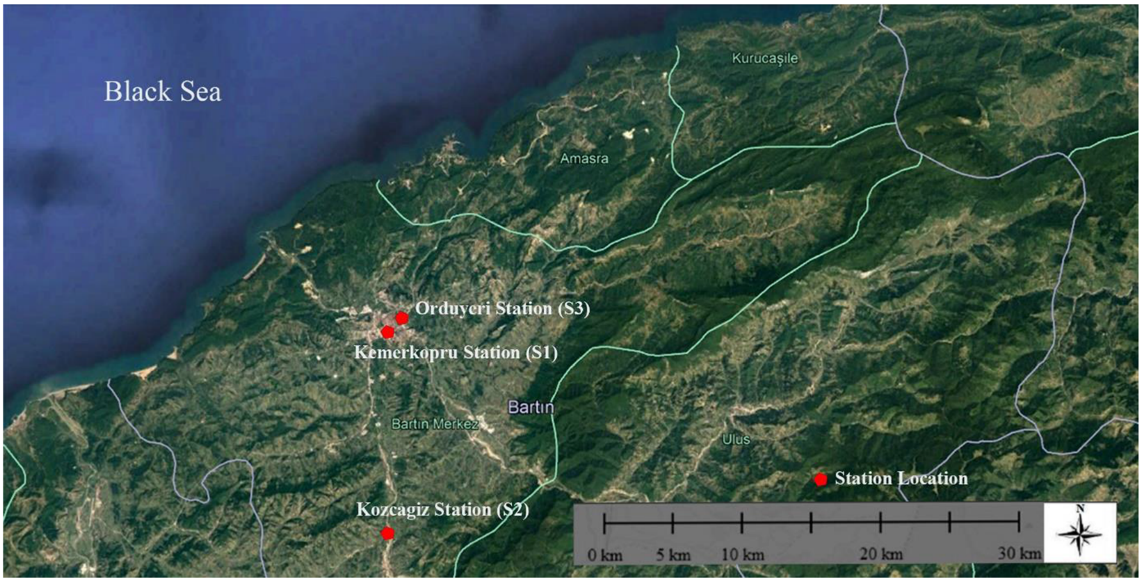


Fig. 1 The locations of the samples taken from the Bartin River

the samples taken from the Bartin River. S1, S2, and S3 refer to the locations of Kemerkopru, Kozcagiz, and Orduyeri, respectively. These stations are considered the best reflecting points of heavy metal pollution caused by domestic and industrial activity on the Bartin River.

Analysis of the sample

The samples from the Bartin River were collected on a monthly basis for a whole year between December 2012 and December 2013. Sample analyses were determined by measuring the dissolved metal concentration in Bartin River. Before taking the sample, the sample bottles were rinsed several times from river water before being filled from the river water. After the samples were brought to the laboratory, concentrated nitric acid was added. The samples were then filtered through a filter paper. The amount of heavy metals in water samples was determined by using Atomic Absorption Spectrophotometer. Standard solutions in 5 different concentrations were prepared to obtain the calibration curve before the AAS analysis. This was done by dilution from standard stock solution at a concentration of 1000 ppm. Cathode lamps were used as radiation source for Cu, Mn, Ni, Pb, Zn, and Fe metals and the fuel was air acetylene. All samples were analyzed in triplicates. River water pH, temperature (T), and EC values were measured using a multi-parameter measuring device (Hach HQ40D) with on-site measurements at these stations. COD, BOD, and SS analyses were performed in the laboratory and all parameters were analyzed according to standard methods (APHA 1985)

Artificial intelligence models

In this section, the structures and model architectures used for ANN and ANFIS models are summarized, and the structure and the function types of ANN and ANFIS, input and output parameters used, and the training and testing phases are explained.

Multilayer perceptron model

The most widely used ANN model is multilayer perceptron (MLP). It is a flexible method which has simple neurons, consisting of an input layer, one or more hidden layers, and an output layer. The sensor calculates a single output by creating combinations of linear relationships according to input weights and transfer functions. It is generally suggested to use one hidden layer within the network, since more hidden layers cause local minimum problems. As shown in Eq. 1, the mathematical formula of the MLP model can be created.

$$Y = F\left(\sum_{j=1}^{m} W_{kj}.F\left(n_j\right) + B_k\right) = F\left(\sum_{j=1}^{m} W_{kj}.F\left(\sum_{i=1}^{n} W_{ji}X_i + B_j\right) + B_k\right) \quad (1)$$

Here, W_{kj} refers to the weights between the hidden layer and the input layer; W_{ji} refers to the weights between the hidden layer and the output layer; X_i to the input variables; m to the number of neurons in the hidden layer; n to the number of neurons in the input layer; B_j and B_k to the bias values of the neurons in the hidden layer and the output layer; F_j and F_k to the transfer functions in the hidden layer and the output layer; and Y to the output value (Messikh et al. 2017; Ghritlahre and Prasad 2018). An architectural representation of MLP consisting of a single hidden layer with three neurons, two input layers, and an output layer is given in Fig. 2.

MLP model setup is as follows respectively: selection of input data, determination of transfer function, normalization of the data, selection of MLP architecture, selection of training algorithm, determination of model performance criteria, and selection of stop criterion. Successful representation of events with complex behavior by ANN model depends on the selection of training data and input parameters correctly. In the training of ANN models, it is stated that 70–80% of the data set should be reserved (Ghritlahre and Prasad 2018; Assi et al. 2018; Fissa et al. 2019).

Selection of model inputs can be determined by testing due to the complexity of the event, and here, while increasing the number of inputs increases the complexity of the network, it may also affect the model success in both positively and negatively. Generally, the number of inputs in most ANN models is selected between 2 and 6. The transfer functions are used in order to simulate the event by using the input and output parameters. One of the most widely used transfer functions is the Log-Sigmoid transfer function (Logsig) shown in Fig. 3a. This function generates outputs within the range of 0 to 1 in response to the Logsig input data given in Eq. 2. On the other hand, the tangent sigmoid transfer function (Tansig), which is similar to Logsig, generates outputs within the range of − 1 to 1. The graph of Tansig, of which function is given in Eq. 3, is shown in Fig. 3b. Even though most of the real models have non-linear input and output parameters, in the cases close to linearity, the Purelin transfer function given in Fig. 3c can represent the event in the best way (Abdi-Khanghah et al. 2018).

$$a = \mathrm{Logsig}(z) = \frac{1}{1+e^{-z}} \tag{2}$$

$$a = \mathrm{Tansig}(z) = \frac{2}{1+e^{-2z}}-1 \tag{3}$$

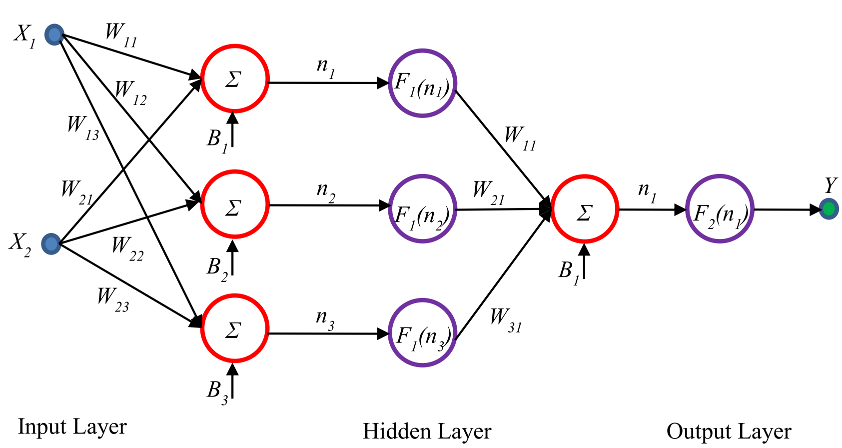
Normalization is a form of data processing used in data mining. Normalization is necessary when the values of the parameters differ greatly. By normalization, the values inside the data set are scaled down to a small range determined within the range of 0.0 to 1.0. The most commonly used normalization method in ANN models is the min-max method. Min-max normalization performs a linear transformation on the original data (Hang et al. 2019).

In the MLP models within the scope of this study, the normalization range for the data was selected to be between 0.2 and 0.8. The type of transfer functions and the number of neurons in the hidden layer were found by trial and error as recommended in the literature. In all of the MLP models, the training algorithm selected was Levenberg-Marquart, the revolutions per minute as 500, and the mean square error as MSE = 10^{-7}, like the target error.

Radial basis function network model

Despite resembling the MLP model fundamentally, radial basis function network (RBN) models have some critically distinguishing differences. As in MLP, the input layer, hidden layer, and output layer form the basic architectural structure in RBN. Although unlike MLP, RBN has weights only between the hidden layer and the output layer. These weights can be found by least squares method. However, the most important difference distinguishing RBN from MLP occurs in the hidden layer. Here, there are both structural and functional differences. While the reason for the structural difference is the number of neurons and the training algorithm, the functional difference is because of the fact that the neurons in the hidden layer contain radial functions. The Gaussian function is mainly used here as the activation function, and the function in the Gaussian structure is given in Eq. 4 and Fig. 4 (Asgharnia et al. 2019; Han et al. 2019).

Fig. 2 Architectural representation of a multilayer perceptron model (2, 3, 1)



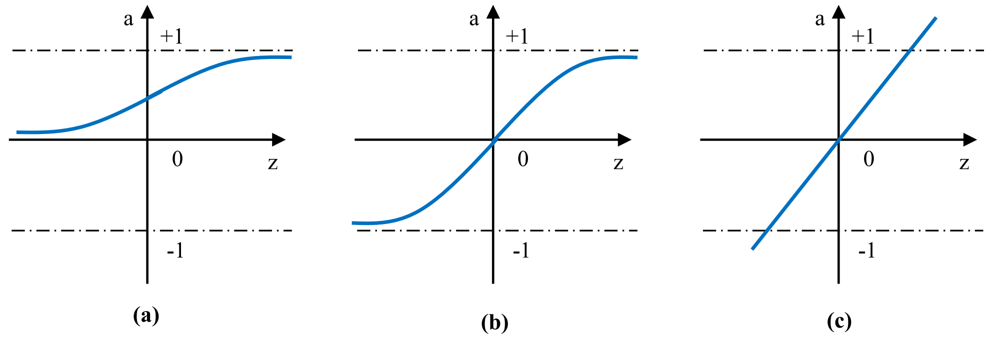


Fig. 3 Transfer functions. **a** Log-Sigmoid transfer function. **b** Tan-Sigmoid transfer function. **c** Purelin transfer function

$$f(x) = \exp\left(-\frac{\|x-c_i\|^2}{\sigma_i^2}\right), \quad i = 1, 2, \ldots.., k \tag{4}$$

Here, x refers to the input parameter, c to the distance from the center of the Gaussian function, and σ to the spread parameter. As can be seen from the function, the spread parameter significantly affects the network performance.

Unlike MLP, there is no parameter learning in the RBN method. Linear adjustment of the weights is performed for the bases of the radial function. In this way, a rapid convergence occurs without local minimums. The RBN theory is described in detail in Haykin (1998). In this study, the optimum number of neurons in the hidden layer and the spread numbers in return to this were determined by trial and error method.

Adaptive neuro-fuzzy inference system model

Adaptive neuro-fuzzy inference system (ANFIS), which uses gradient descent–based optimization methods in order to set the membership function parameters, is composed of three basic parts which are fuzzy rule base, membership functions defining the fuzzy sets of fuzzy rules, and reasoning mechanism. ANFIS, which emerges with the joint use of ANN and fuzzy logic, functions with the principle of optimizing the rule base and membership function values in order to model the systems, whose input and output values are known, with fuzzy logic. Optimization process occurs with the use of ANN's learning methods. By using ANN, learning ability is built to the fuzzy systems which normally do not have learning ability. In general, a hybrid learning algorithm is used. In order to apply the ANFIS method, generally a data set based on input-output is required. The model established depending on the number and type of membership function selected is formed using a learning algorithm. ANFIS forms a fuzzy model in the Takagi-Sugeno structure, which means that a fuzzy model is obtained from the inputs in the Mamdani (triangle, trapezoid, gauss, etc.) structure and the outputs in the function structure. If the fuzzy value generated by the method is

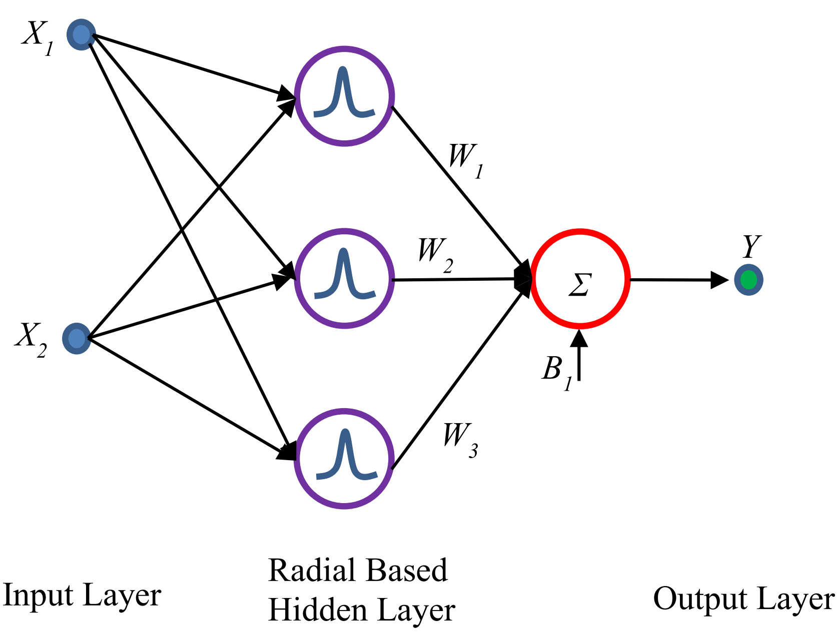


Fig. 4 Architectural representation of a radial basis function network model

negative, it uses a set of rules. The fuzzy rules of an ANFIS, which contains two inputs and two rules with input variables of x_1 and x_2, are as follows.

R1: if x_1 is A_1 and x_2 is B_1, then $f_1 = p_1x_1 + q_1x_2 + r_1$
R2: if x_1 is A_2 and x_2 is B_2, then $f_2 = p_2x_1 + q_2x_2 + r_2$

Here, A_i and B_i are the fuzzy sets corresponding to x_1 and x_2, respectively, and p_i, q_i, and r_i are linear outcome parameters. The ANFIS architecture is constructed by specifying the parameters in a way to minimize the difference between the output of the entire network and the target value, in other words the error. The adaptive neural network structure formed by ANFIS is shown in Fig. 5. Six layers and their functions within the ANFIS structure are given below (Jang 1993; Yilmaz and Kaynar 2011).

Layer 0: This layer showing the ANFIS inputs are generally deemed outside of main ANFIS structure.

Layer 1: This is the fuzzification layer where input values are divided into fuzzy sets and where the membership degrees are created by using membership functions. Here, the membership function gives the input variable a degree of membership from 0 to 1, and this value is the fuzzy value of the net input. Each node in the fuzzification layer corresponds to an adaptive membership function given in Eq. 5.

$$O_{1,i} = \mu A_j(x) \tag{5}$$

Here, $O_{1,\,i}$ i. is the output of the node and $\mu A_j(x)$ A_j is the output of the membership functions of the fuzzy set. Although there are various membership functions available, Gaussian and generalized Bell functions are the most commonly used ones.

Layer 2: It is named as the rule layer, and in line with a fuzzy rule for each node, the calculated value is exported by making a new calculation according to the output of the previous layer. For example, the output of the i. node can be calculated with Eq. 6.

$$O_{2,i} = w_i = \mu A_i(x_i)\mu B_i(x_2) \tag{6}$$

Here, $O_{2,\,i}$ i. is named as the ignition level of the rule, and $\mu A_i(x)$ and $\mu B_i(x)$ are the outcome of membership values.

Layer 3: This is defined as the normalization layer. It takes the nodes coming from the rule layer as the input value and calculates the normalized ignition level of j. node with Eq. 7, by proportioning the ignition level of the j. rule to the total ignition level of all rules.

$$O_{3,j} = \overline{w_j} = \frac{w_j}{\sum_{i=1}^{R} w_i} \tag{7}$$

Here, w_i i. refers to ignition level of the rule and R to the total number of rules.

Layer 4: This layer, which is related to updatable parameters, is an adaptive layer like Layer 2. The weighted outcome values of a rule given on each node are calculated. The linear function of this layer, namely clarification layer, is given with Eq. 8.

$$O_{4,i} = \overline{w_i}f_i = \overline{w_i}(p_ix_i + q_ix_2 + r_i) \tag{8}$$

Here, $O_{4,\,i}$ is the outcome of the i. node in this layer and p_i, q_i, and r_i are the coefficients of the functions updated at the optimization stage.

Layer 5: It calculates the actual outcome value of the system according to Eq. 9, by collecting all the signals coming from Layer 4.

$$O_{5,i} = \sum_i \overline{w_i}f_i = \frac{\sum_i w_if_i}{\sum_i w_i} \tag{9}$$

Here, $O_{5,\,i}$ is the last obtained outcome value and $\overline{w_i}f_i$ is the i. node outcome of Layer 4.

For the optimization of rule parameters in ANFIS, both standard error recirculation algorithm and hybrid learning algorithm based on gradient descent and least-squares method

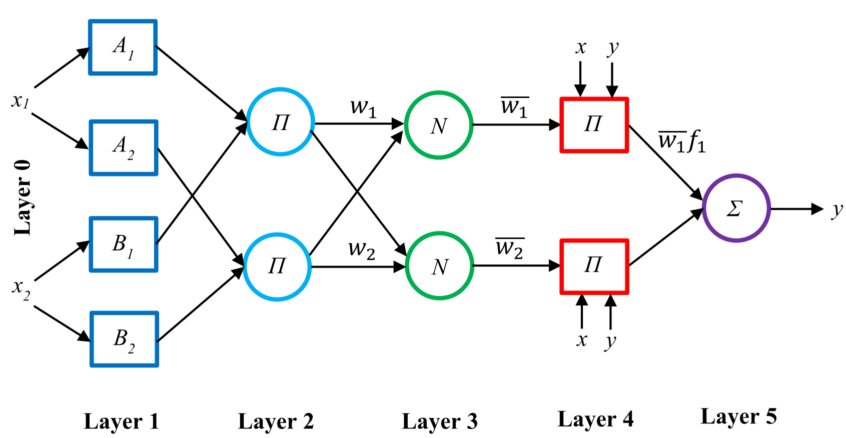


Fig. 5 Architectural representation of an ANFIS network model

can be used (Jang 1993; Yilmaz and Kaynar 2011; Rajab 2019).

Preparation of data set and performance measurement

The measured input parameters for modelling were T, pH, EC, COD, BOD, and SS. The data measured in this section are divided into two categories as training data and test data in order to train the models and measure the model performances. Seventy-five percent of the total data was allocated for training of models and 25% for performance measurement, completely randomly, in a way to have no statistically significant difference between the data. Statistical values of the data are given in Table 1.

The model performances were defined according to the principle that test data, which were completely independent of the data, whose models had been used in the training, being approximate to the observed values. In order to test this proximity, root mean square error (RMSE), mean absolute error (MAE), and determination coefficient (R^2) parameters were used. Here, R^2 is used in order to determine the linear relationship between observed and predicted models, whereas MAE and RMSE are used to measure model errors. MAE, RMSE, and R^2values are calculated using Eqs. 10–12, respectively (Guo et al. 2019).

$$\mathrm{MAE} = \frac{1}{n}\sum\left|y_{\mathrm{observed}} - y_{\mathrm{predicted}}\right| \tag{10}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum\left(y_{\mathrm{observed}} - y_{\mathrm{predicted}}\right)^2} \tag{11}$$

$$\mathrm{R}^2 = 1 - \frac{\sum\left(y_{\mathrm{observed}} - y_{\mathrm{predicted}}\right)^2}{\sum\left(y_{\mathrm{observed}} - y_{\mathrm{mean}}\right)^2} \tag{12}$$

Table 1 Statistics for data sets

		Input parameters						Output parameters					
Statistic	Data type	Temp (°C)	pH	EC (µS/cm)	COD (mg/L)	BOD (mg/L)	SS (mg/L)	Cu (mg/L)	Fe (mg/L)	Zn (mg/L)	Mn (mg/L)	Ni (mg/L)	Pb (mg/L)
Mean	Training	16.20	8.00	666.41	18.69	9.91	37.42	0.00952	0.82389	0.18894	0.04744	0.01033	0.00786
	Test	15.88	8.04	656.44	18.90	10.52	41.11	0.00944	0.85956	0.19056	0.04800	0.00911	0.00785
	Total	16.12	8.01	663.92	18.74	10.06	38.34	0.00950	0.83281	0.18935	0.04758	0.01003	0.00786
Std. Deviation	Training	7.04	0.52	208.49	4.84	4.17	9.57	0.00362	0.10106	0.06679	0.01652	0.00550	0.00155
	Test	5.15	0.51	192.32	3.41	3.45	7.31	0.00411	0.06676	0.06314	0.01610	0.00534	0.00158
	Total	6.62	0.52	204.62	4.53	4.01	9.20	0.00375	0.09493	0.06590	0.01642	0.00548	0.00156
Maximum	Training	28.50	8.80	1027.00	28.20	17.90	49.70	0.01500	0.94800	0.31800	0.07800	0.01800	0.01030
	Test	21.60	8.72	977.00	23.90	15.20	48.80	0.01600	0.95600	0.26300	0.06700	0.01700	0.01012
	Total	28.50	8.80	1027.00	28.20	17.90	49.70	0.01600	0.95600	0.31800	0.07800	0.01800	0.01030
Minimum	Training	6.50	7.25	427.00	10.80	3.50	18.70	0.00000	0.56600	0.09100	0.02300	0.00000	0.00390
	Test	6.30	7.12	476.00	12.60	3.60	29.90	0.00400	0.73200	0.09300	0.02100	0.00400	0.00563
	Total	6.30	7.12	427.00	10.80	3.50	18.70	0.00000	0.56600	0.09100	0.02100	0.00000	0.00390
Skewness	Training	0.32	− 0.38	0.71	0.10	0.26	− 0.59	− 0.65	− 0.78	0.58	0.12	− 0.39	− 0.44
	Test	− 1.05	− 0.76	0.91	− 0.40	− 0.69	− 0.48	0.25	− 0.50	0.01	− 0.73	0.59	− 0.09
	Total	0.19	− 0.45	0.72	0.03	0.09	− 0.64	− 0.37	− 0.87	0.45	− 0.06	− 0.17	− 0.34
Kurtosis	Training	− 1.20	− 1.51	− 1.25	− 0.83	− 0.92	− 0.87	0.39	0.03	− 0.91	− 1.30	− 0.93	− 0.10
	Test	− 0.14	− 0.37	− 1.16	− 0.60	− 0.08	− 1.65	− 1.24	− 0.24	− 1.99	− 0.65	− 1.78	− 1.73
	Total	− 1.01	− 1.32	− 1.23	− 0.71	− 0.91	− 0.69	− 0.22	0.37	− 1.10	− 1.22	− 1.23	− 0.52

Here, n refers to the number of data, $y_{observed}$ to the observed heavy metal value (mg/L), and y_{mean} to the mean value of observed heavy metal (mg/L)'dir.

Results and discussion

Heavy metal concentrations in samples from Bartin River produced varied values. According to the results of heavy metal analysis carried out in the water samples taken, it was determined that there was contamination at various rates with heavy metals at all stations. Cu was below the 1st water quality limit values for all stations. Fe is in the 2nd quality class for all stations. The values obtained for Zn were determined as 1st for S1 and S3 station, and 2nd for S2 station. In terms of Mn, values below 1st water quality class were obtained at all stations. For Ni and Pb, they were determined in the 1st quality class in all stations. If a general evaluation is made for all heavy metals considering all the stations, the water taken from the S1 station falls under the category of water quality that is less dirty in terms of Fe and Zn, and clean in terms of other metals. S2 station has similar characteristics with Kemerkopru (S1). The S3 station is in the category of water quality, which is less polluted in terms of Fe and in terms of other metals. As a result, pollution by heavy metal ions shows variable values.

Studies on the determination of heavy metal concentrations in water resources have revealed that physicochemical and environmental factors may affect heavy metal concentrations. Therefore, pH, EC, COD, SS, etc. parameters are used as input data for the prediction of heavy metal concentrations in the studies (Rooki et al. 2011; Bayatzadeh Fard et al. 2017; Nhantumbo et al. 2018; Lu et al. 2019; Agah and Soleimanpourmoghadam 2020).

In other studies, it was insufficient to use the input parameters such as pH, temperature, and COD alone in estimating metal concentration with ANN and ANFIS models; however, it is stated that, if used together, the success of model estimates significantly increases (Bayatzadeh Fard et al. 2017; Nhantumbo et al. 2018; Lu et al. 2019). While these three input data provide sufficient model performance in the estimation of some metals, the addition of EC, BOD, and SS values as input data in some metals estimation may increase or decrease the model success (Lu et al. 2019). Lu et al. (2019) show that the most effective input parameters used in their study for the prediction of metal concentration are pH, temperature, and COD.

RMSE, MAE, and R^2 values of the training and test data of MLP, RBN, and ANFIS models are given in Tables 2, 3, and 4, respectively. The best results for MLP models were obtained from the models with 6 inputs for Cu prediction; with 5 inputs without any SS for Fe; with 4 inputs without BOD and SS variables for Zn; with 3 inputs with temperature, Ph, and EC for Mn and Ni heavy metals; and with 3 inputs with temperature, EC, and COD input variables for Pb. The number of neurons in the hidden layer was tested between 1 and 30, and in all good models, the number of neurons in the hidden layer was found to be 10 or less.

As can be observed in Table 2, the RMSE and MAE values of all heavy metal models created by using MLP were found to be acceptably low, and R^2 value was found to be above 0.77 during the test phase.

In heavy metal predictions, the models—where temperature, pH, EC, COD, BOD, and SS input variables were used—were found to be more successful in the models created by using RBN. The RBN model obtained without using BOD and SS variables has fewer errors for Ni metal only. For RBN models, the best values given in Table 3 were found by testing where number of spread values determining the width of the gauss curve was in the range of 0.1 to 3.0 and the number of cells in the hidden layer was in the range of 1 and 30.

In the models using RBN method, the test phase error values are close to MLP models with RMSE and MAE. R^2 values of the models in the test phase were also calculated in the range of 0.773 to 0.989.

Also, in ANFIS models, 6 input variables were introduced to the network by applying different input combinations. In ANFIS models, trimf, trapmf, pimf, gbellmf, gaussmf, gauss2mf, dsimf, and psimf were tested as the membership functions. The number of membership functions was applied in the range of 1 to 5. Hybrid and ANN algorithms were used for training the network and constant and linear were used as output-type membership functions.

As can be seen in Table 4, R^2 value of all ANFIS models is higher than 0.80 during the test phase, and the RMSE and MAE values indicate to low error values.

Actual values for Cu, Fe, Zn, Mn, Ni, and Pb heavy metals were compared with the estimates from RBN, ANFIS, and MLP models and R^2 values of each model were obtained. In order to test the models (Figs. 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 and 17), the graphs showing the distribution of the measured data and the values obtained from RBN, ANFIS, and MLP models are given in Fig. 6 for Cu; in Fig. 8 for Fe; in Fig. 10 for Zn; in Fig. 12 for Mn; in Fig. 14 for Ni; and in Fig. 16 for Pb. Scatter diagrams obtained from the data measured with the values estimated with RBN, ANFIS, and MLP models for the test data are shown in Fig. 7 for Cu; in Fig. 9 for Fe; in Fig. 11 for Zn; in Fig. 13 for Mn; in Fig. 15 for Ni; and in Fig. 17 for Pb.

Despite the variability depending on the model type, it was observed that the heavy metal content could be approximately predicted by using temperature, pH, EC, COD, BOD, and SS variables. Thus, by using artificial intelligence models and the aforementioned parameters, which are relatively easy-to-measure, it will be possible to detect the heavy metals, which are

Table 2 Performance of multilayer perceptron models

Inputs	Number of hidden layer nodes	Transfer function (hidden)	Output	Transfer function (output)	Training			Test		
					RMSE	MAE	R^2	RMSE	MAE	R^2
Temp., pH, EC, COD, BOD, SS	10	Logsig	Cu	Tansig	0.0000	0.0000	1.000	0.0015	0.0014	0.954
Temp., pH, EC, COD, BOD	8	Logsig	Fe	Tansig	0.0000	0.0000	1.000	0.1316	0.1029	0.842
Temp., pH, EC, COD	3	Logsig	Zn	Tansig	0.0056	0.0042	0.993	0.0165	0.0143	0.966
Temp., pH, EC	5	Logsig	Mn	Logsig	0.0031	0.0015	0.965	0.0110	0.0100	0.770
Temp., pH, EC	5	Logsig	Ni	Logsig	0.0000	0.0000	1.000	0.0024	0.0018	0.941
Temp., EC, COD	5	Logsig	pb	Tansig	0.0000	0.0000	0.999	0.0008	0.0006	0.805

harmful to living organisms inside the rivers, both quickly and economically. If sorted from the best model to the worst by taking the MAE and RMSE values into consideration based on the test evaluation results according to the heavy metal types, where all of the MLP, RBN, and ANFIS models were generally approximate to each other, RBN was successful for Cu, Zn, and Mn, while ANFIS model was successful for Fe and Pb, and MLP for Ni.

For the good management of water resources, river pollution should be well defined, monitored, and modelled. For that reason, ANN have recently been successfully applied in environmental sciences such as water resources, water quality, and hydrologic time series (Sarkar and Pandey 2015).

Elzwayie et al. (2017) used the radial basis function neural network algorithm in order to model the changes in the heavy metal content found in both polluted and non-polluted lake water bodies in tropical "Malaysia" and arid "Libya" depending on climatic and pollution conditions. The study was conducted under different climatic conditions, and the weekly records of physicochemical parameter data (e.g., pH, EC, WT, DO, TDS, TSS, CL, NO3, PO4, and SO4) and climatological parameters (e.g., air temperature, humidity, and precipitation) were used as input data. Three different scenarios were evaluated, and in general, the results obtained from all these scenarios showed a high level of accuracy

In a study conducted by Lu et al. (2019) on a potable water resource in the Taihu Region of China, the physicochemical parameters and the heavy metal concentrations of the ground water were measured. In the modelling carried out within the scope of the study, artificial neural networks (ANN) and support vector machine (SVM) models were used. Sensitivity analysis showed that simulated heavy metal concentrations were most sensitive to pH. Both ANN and SVM quick simulation models well simulated the particulate heavy metal concentrations with Nash-Sutcliffe efficiency coefficients > 0.8 most (Lu et al. 2019).

In a study conducted by Swietlicka et al. (2017) in Lublin Province of Poland, the Ag, As, Ba, Ca, Cd, Co, Cr, Cu, Fe, Hg, Mg, Mn, Ni, P, Pb, S, Sr, TOC (total organic carbon), V, and Zn contents in the lake and river sediments were evaluated. Artificial neural networks were then examined with respect to their ability to recognize and classify the data. Multilayer perceptron was used as the statistical model. Constructed models were able to give correct answers in 74% of cases, when classifying reservoir's area usage and 100% for the type of water body.

In their study, Nhantumbo et al. (2018) used both of the ANN and PPBM methods in order to predict the concentrations of major ions (Na+, K+, Mg2+, Ca2+, HCO3−, SO42−, Cl−, and NO3−) in the river water depending on pH,

Table 3 Performance of radial basis function network models

Inputs	Number of hidden layer nodes	Spread value	Output	Training			Test		
				RMSE	MAE	R^2	RMSE	MAE	R^2
Temp., pH, EC, COD, BOD, SS	15	0.6	Cu	0.0008	0.0006	0.954	0.0013	0.0011	0.920
Temp., pH, EC, COD, BOD	15	1.9	Fe	0.0408	0.0326	0.837	0.1281	0.0962	0.892
Temp., pH, EC, COD, BOD, SS	16	1.7	Zn	0.0140	0.0112	0.956	0.0115	0.0105	0.989
Temp., pH, EC, COD, BOD, SS	9	0.6	Mn	0.0080	0.0067	0.764	0.0083	0.0077	0.831
Temp., pH, EC, COD	18	0.7	Ni	0.0008	0.0006	0.977	0.0049	0.0044	0.773
Temp., pH, EC, COD, BOD, SS	13	2	Pb	0.0006	0.0005	0.859	0.0008	0.0007	0.829

Table 4 Performance of adaptive neuro-fuzzy inference system models

Inputs	MF type	Number of MFs	Training type	Output	Output type	Training			Test		
						RMSE	MAE	R^2	RMSE	MAE	R^2
Temp.pHECCODBODSS	trimf trimf gaussmf gaussmf gaussmf gaussmf	3 3 2 2 2 2	Hybrid	Cu	Constant	0.0000	0.0000	1.000	0.0020	0.0017	0.879
Temp.pHCOD	trimf trimf gaussmf	2 3 3	ANN	Fe	Linear	0.0645	0.0538	0.597	0.0425	0.0342	0.813
Temp.pHECCODBODSS	gauss2mf gaussmf gauss2mf gauss2mf gaussmf gaussmf	2 2 2 2 3 2	Hybrid	Zn	Constant	0.0000	0.0000	1.000	0.0235	0.0204	0.896
Temp.pHECCODSS	gaussmf gbellmf pimf gaussmf gauss2mf	2 2 3 2 2	ANN	Mn	Constant	0.0048	0.0031	0.917	0.0115	0.0103	0.826
Temp.pHCODBOD	trapmf pimf pimf gaussmf	2 2 2 3	Hybrid	Ni	Linear	0.0000	0.0000	1.000	0.0049	0.0044	0.807
Temp.pHECCODBODSS	trimf gaussmf trimf gaussmf trimf trimf	2 2 2 2 2 2	ANN	Pb	Constant	0.0004	0.0003	0.931	0.0007	0.0006	0.834

alkalinity, and temperature, and they compared the two methods. As a result of the study, they reported that the ANN model provided better results than PPBM in most cases (Nhantumbo et al. 2018).

In their study, Isiyaka et al. (2019) studied fourteen water quality physicochemical parameters obtained from eight regions in 8 years, and they aimed to develop the best input combination for water quality modelling. They stated that

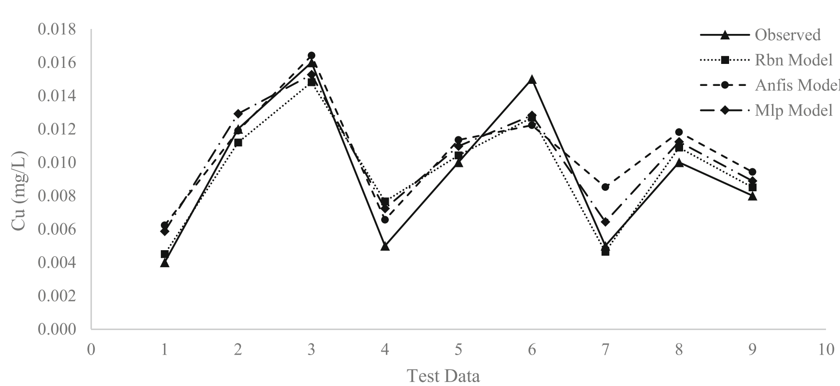


Fig. 6 Distribution of the observed Cu values and the predicted ones, using MLP, RBN, and ANFIS models for test data

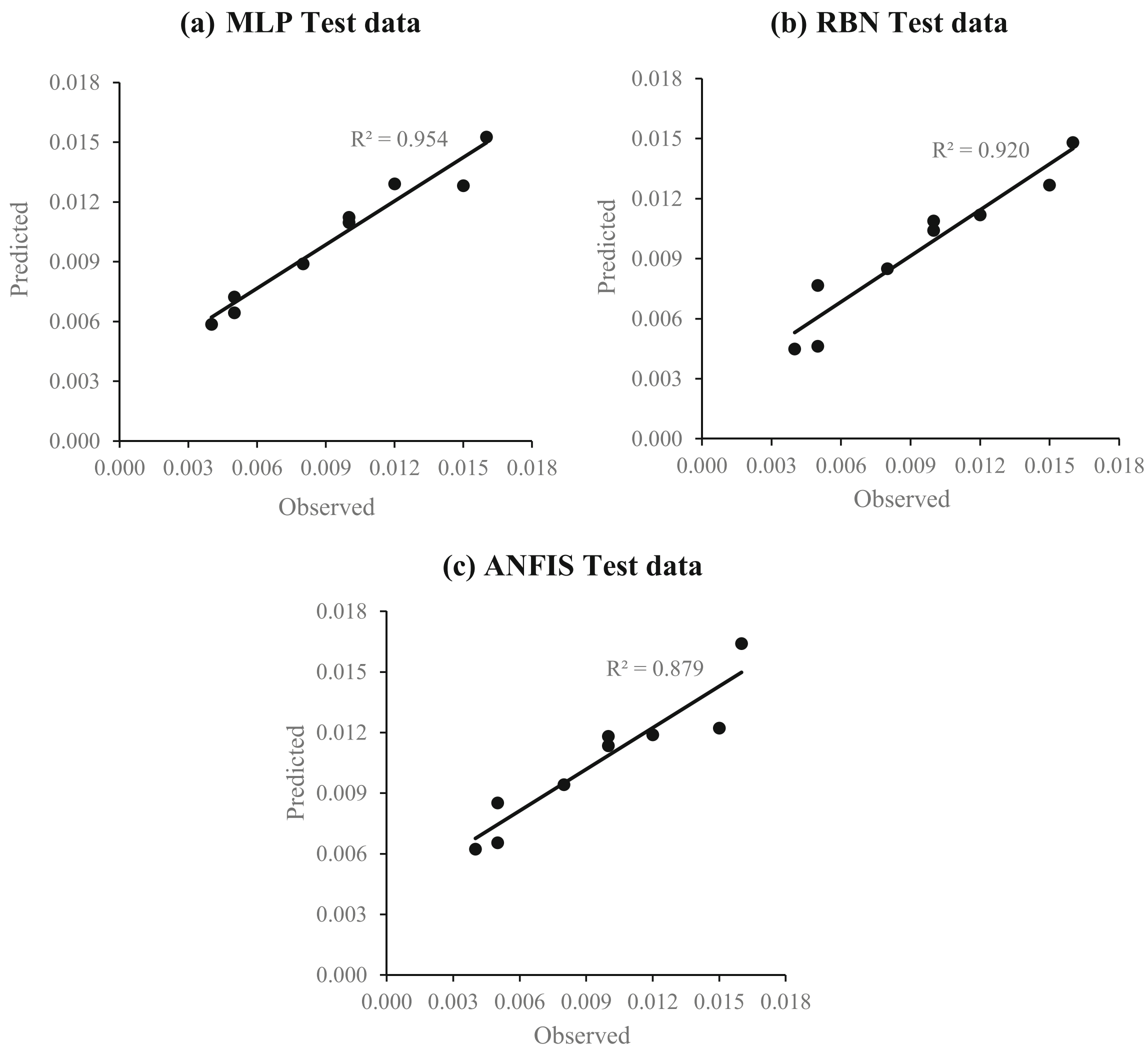
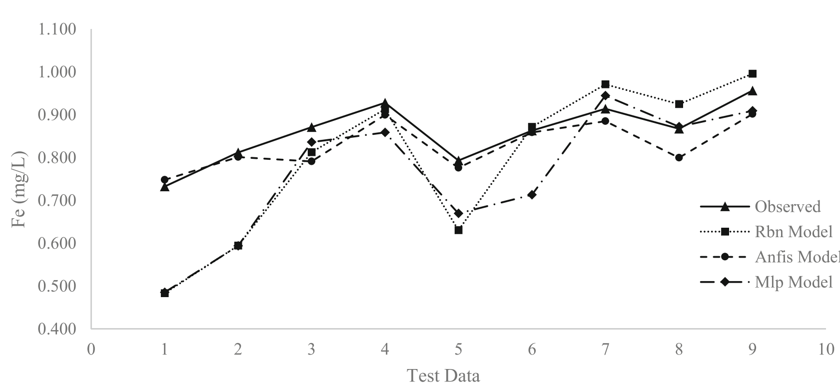


Fig. 7 **a**, **b**, and **c** The scatter plots of observed versus predicted Cu values in test data, using MLP, RBN, and ANFIS models, respectively

the study results were very satisfactory, and suggested that ANN should be used for monitoring water pollution (Isiyaka et al. 2019).

Najah et al. (2014) represent that performance of the ANFIS in experimenting and estimation of dissolved oxygen concentrations. Seifi and Riahi-Madvar (2019) are also studying ANFIS and ANN. They found that it improved. Antanasijević et al. (2013) in Serbia are using this modelling. Bansal and Ganesan (2019) are giving water management with ANN for getting good results for it. Ahamad et al. (2019), Venkatramanan et al. (2017), and Heddam et al. (2019) made qualifying water modelling with ANN. Above all of the studies show that ANN is getting improved.

Fig. 8 Distribution of the observed Fe values and the predicted ones, using MLP, RBN, and ANFIS models for test data



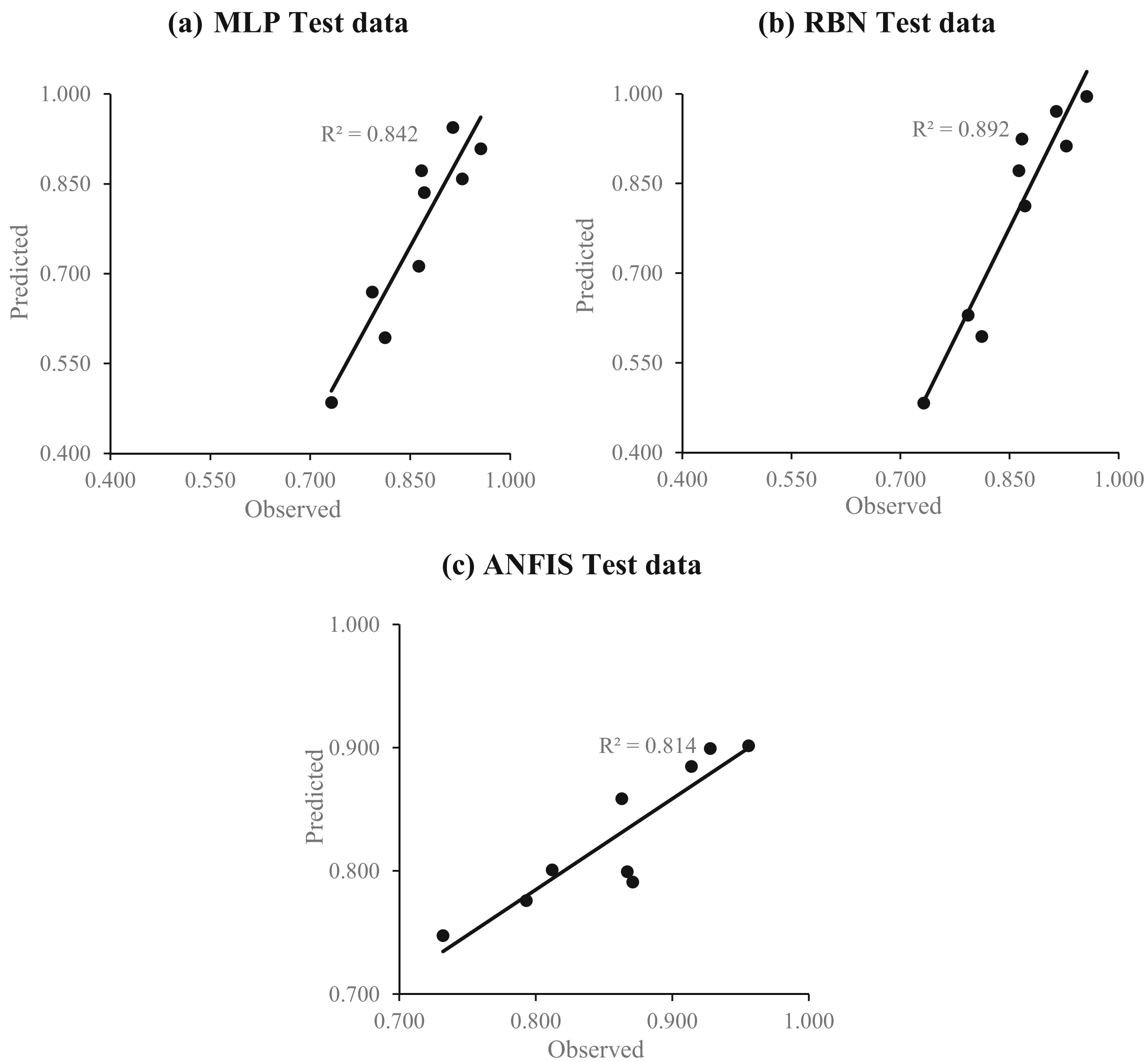


Fig. 9 **a**, **b**, and **c** The scatter plots of observed versus predicted Fe values in test data, using MLP, RBN, ANFIS models, respectively

Determination of sensitive input parameters in the estimation of metals was found by testing the pH, temperature, SS, EC, BOD, and COD values determined by river surface water measurements by combining one, two, three, four, five, and six together. Values that have a significant effect on model performance were chosen as sensitive parameters. Adding some parameters as input data did not improve model performance nor had little effect. It was determined that the parameters that did not increase the model performance were not related to the amount of heavy metal in the river. The error value between the measured and the results obtained from the model can remain

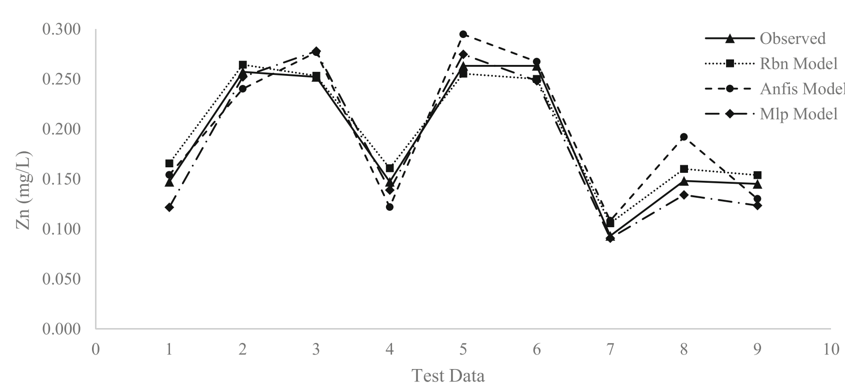


Fig. 10 Distribution of the observed Zn values and the predicted ones, using MLP, RBN, and ANFIS models for test data

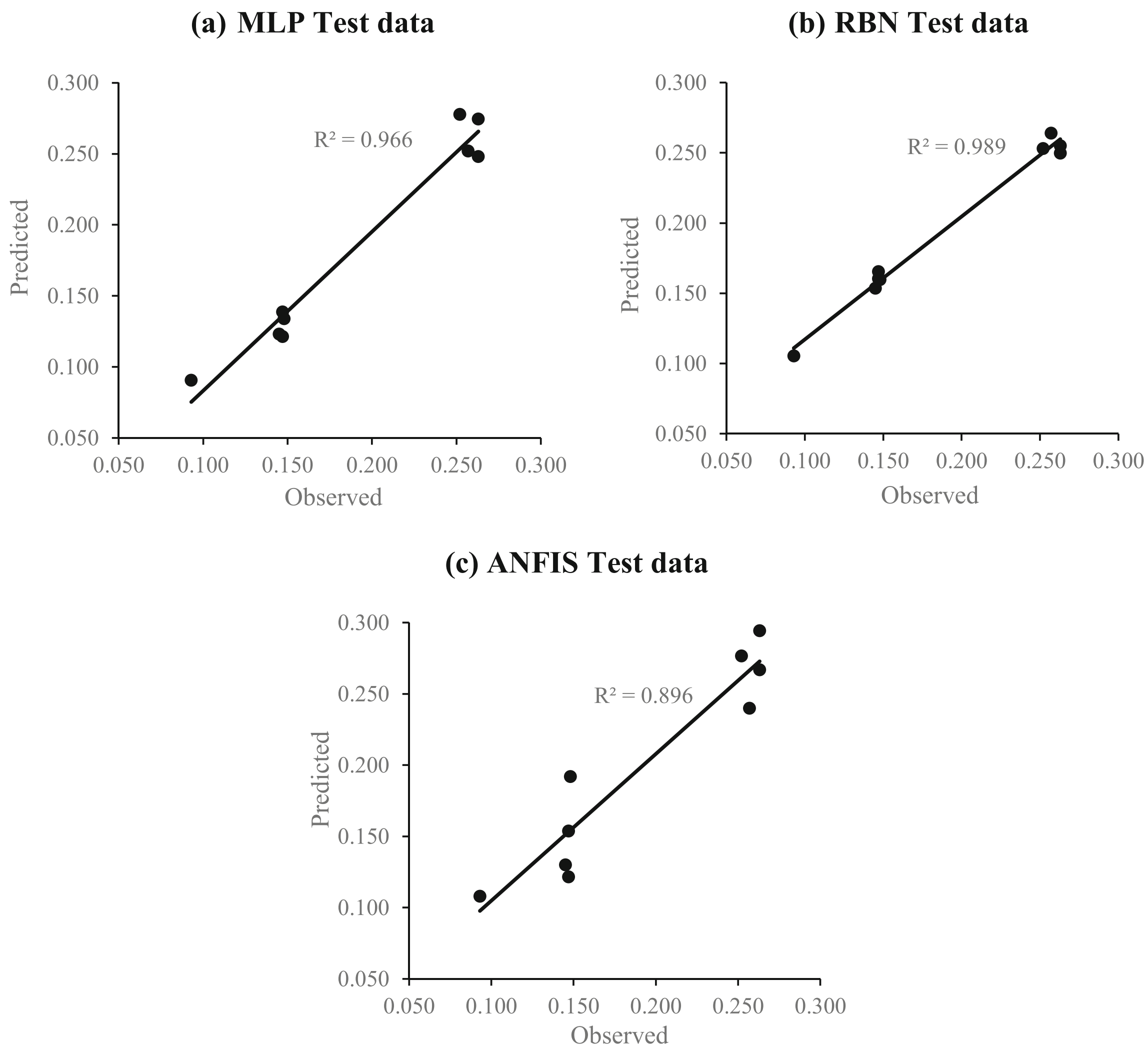


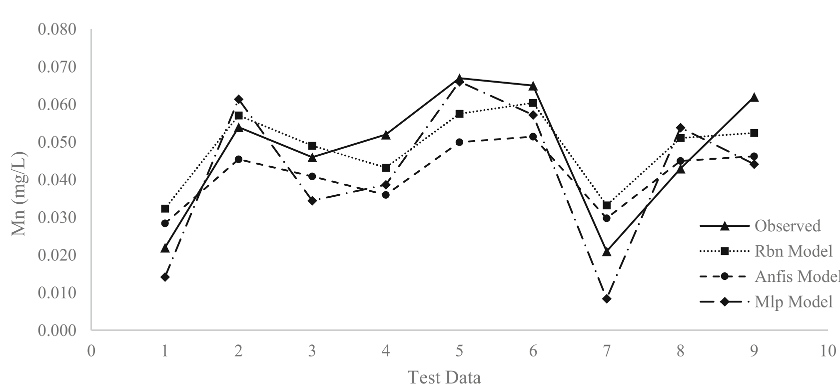
Fig. 11 **a**, **b**, and **c** The scatter plots of observed versus predicted Zn values in test data, using MLP, RBN, and ANFIS models, respectively

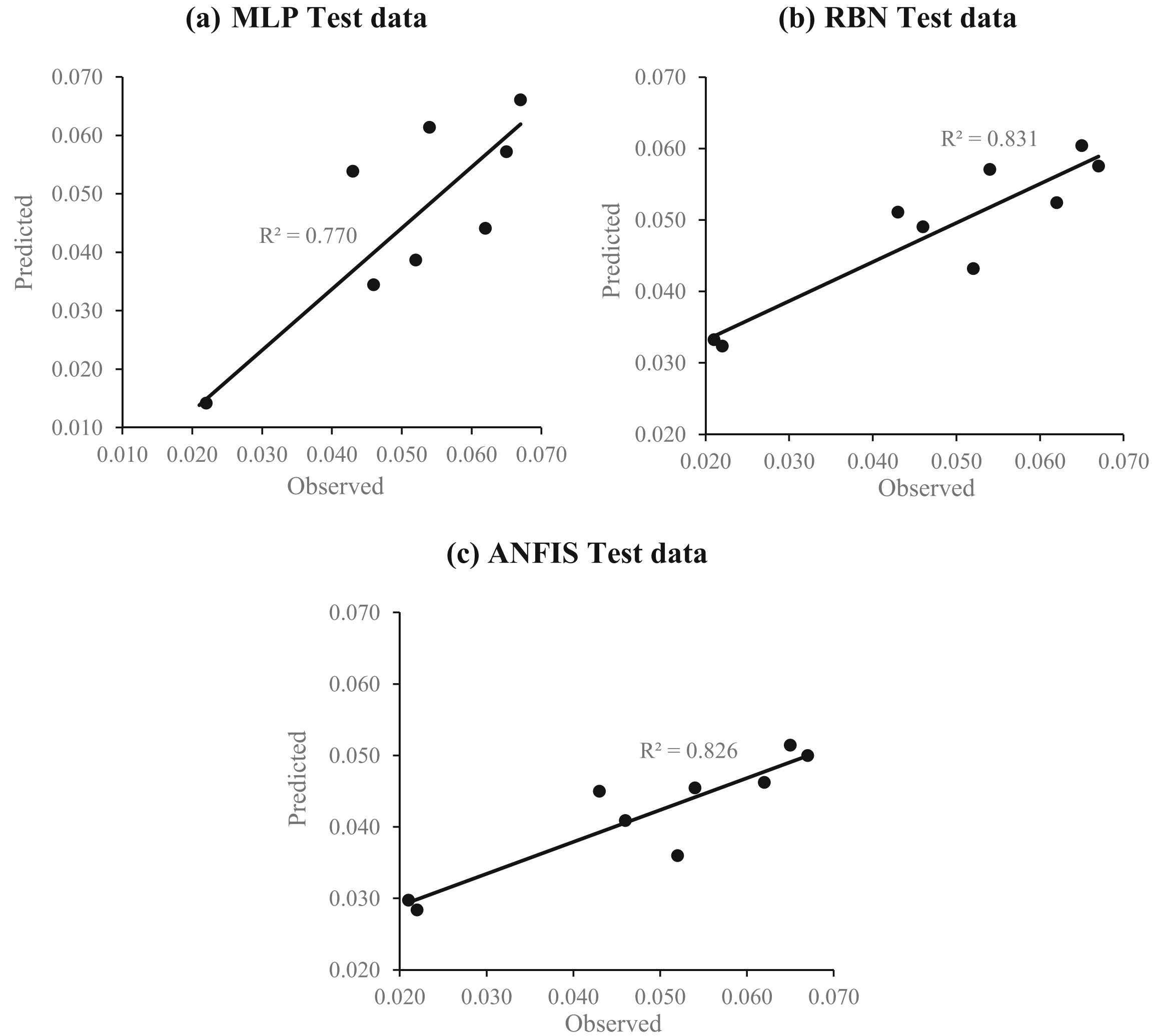
constant or change slightly, by increasing the number of input parameters. In this case, it is more appropriate not to use these input parameters, since unnecessary density will occur in the ANN network and the model will take longer to result. The sensitive parameters determined in this study are similar to the previous studies in the metal concentration estimation (Lu et al. 2019; Nhantumbo et al. 2018).

Conclusions

As was already established from previous literature and studies, monitoring the pollution level in water bodies is of paramount importance for the sustainable use of water resources. However, measurements carried out directly require great cost, time, and labor, whereas modelling offers an alternative

Fig. 12 Distribution of the observed Mn values and the predicted ones, using MLP, RBN, and ANFIS models for test data



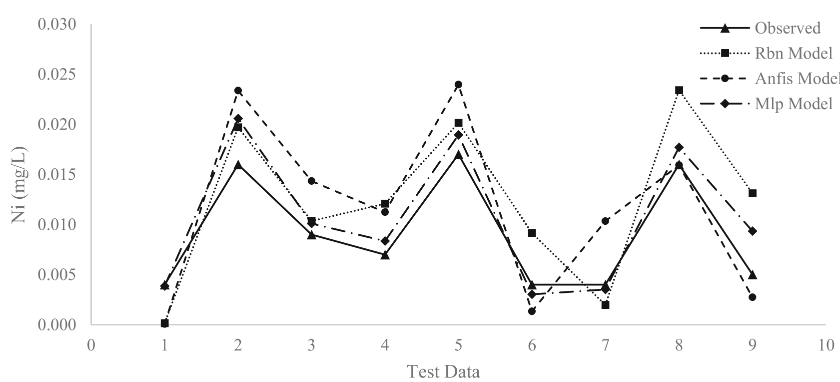


22

Fig. 13 **a**, **b**, and **c** The scatter plots of observed versus predicted Mn values in test data, using MLP, RBN, and ANFIS models, respectively

solution to reduce the cost of water quality monitoring. Generally, predictions are made based on limited sampling information of the pollutant concentration. Through these methods and predictions can be made with a very high level of reliability and a low margin of error. Thus, it is possible to save cost, time, and labor force significantly in monitoring water pollution. However, in order to use these methods, first the most suitable model for that region or resource should be determined, preliminary studies should be carried out, and the reliability level of the model should be identified.

Fig. 14 Distribution of the observed Ni values and the predicted ones, using MLP, RBN, and ANFIS models for test data



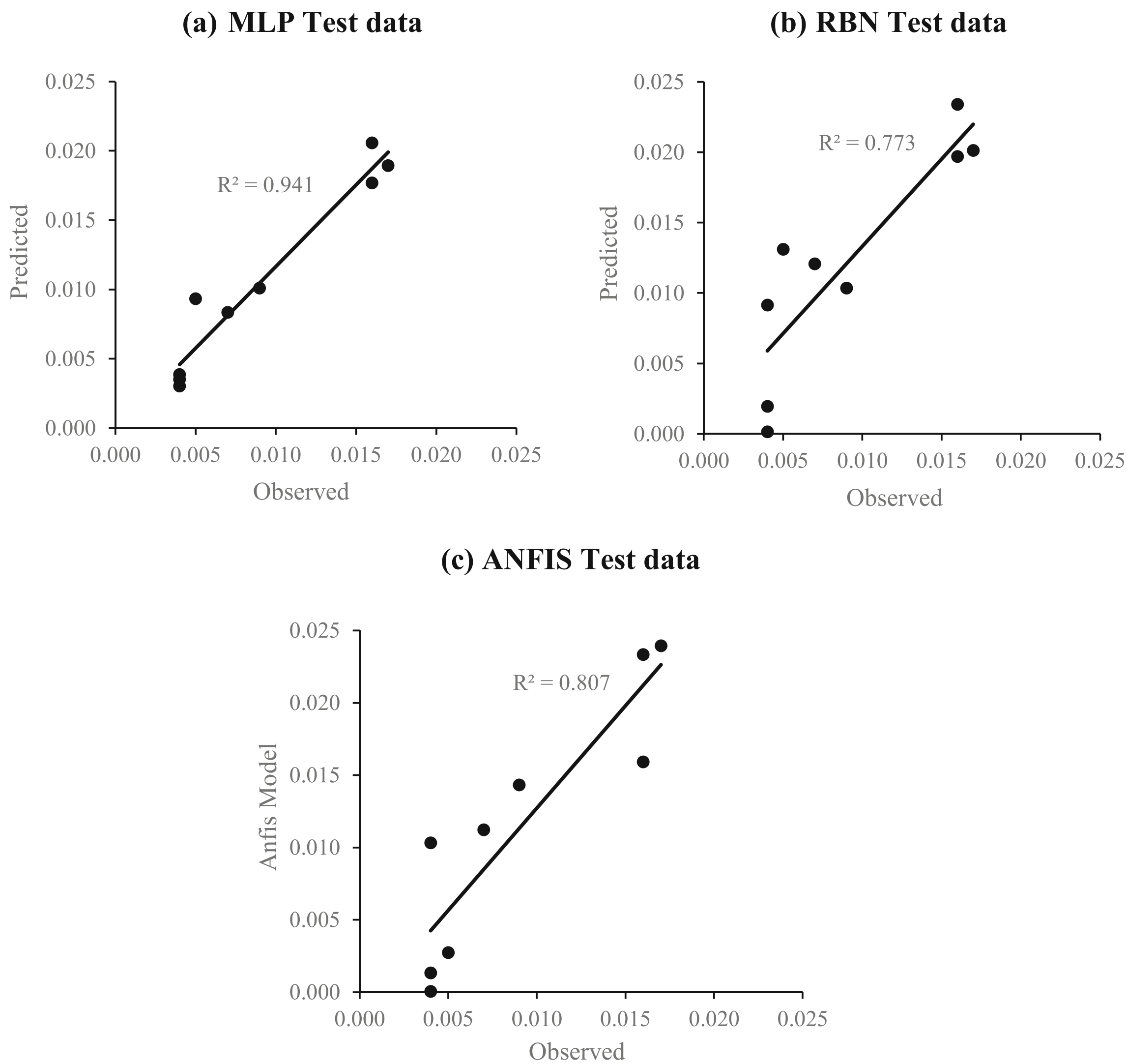


Fig. 15 **a**, **b**, and **c** The scatter plots of observed versus predicted Ni values in test data, using MLP, RBN, and ANFIS models, respectively

In this study, it was aimed to define the most suitable model for monitoring the pollution parameters in the Bartin River, and the most suitable model was identified as a result of the study. Similar studies may be conducted for different water resources (groundwater, lakes, rivers, etc.) in order to monitor the pollution levels of these water bodies. However, it is necessary to determine the most suitable model for each region and water resource by preliminary studies. In this way, the models with the highest reliability levels can be identified and used. For the good management of water resources, river pollution should be well defined, monitored, and modelled. It can be concluded that the employed procedure offers an easy-

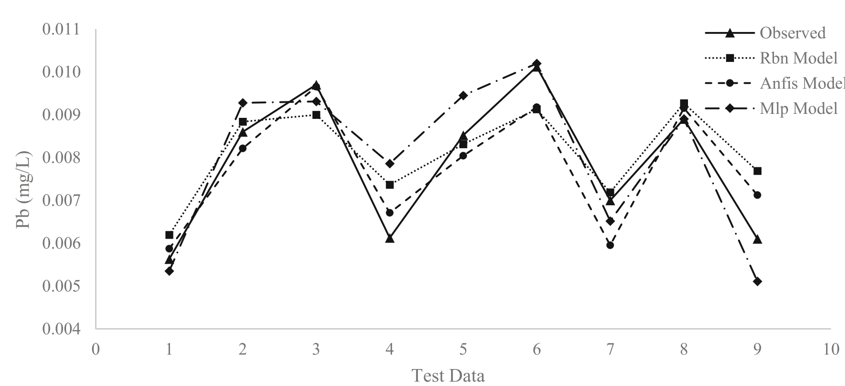
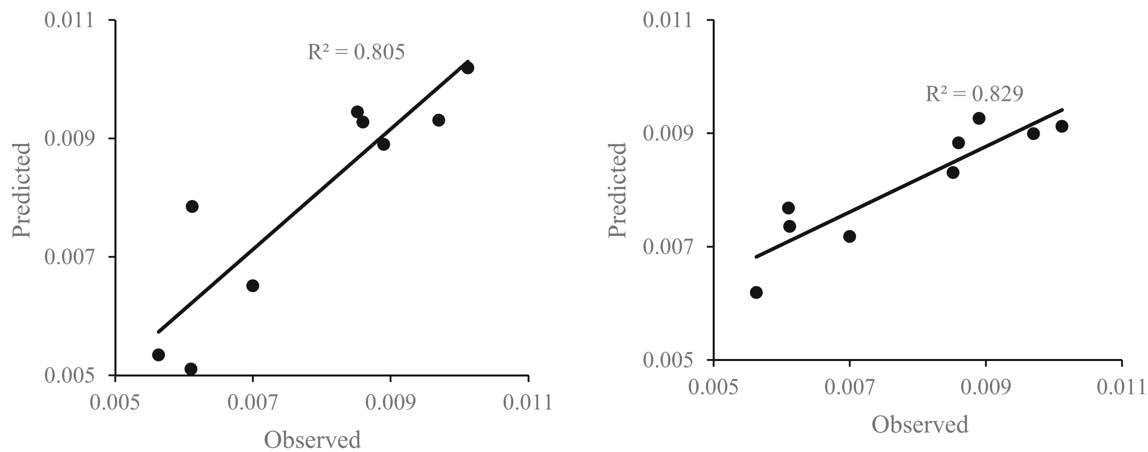


Fig. 16 Distribution of the observed Pb values and the predicted ones, using MLP, RBN, and ANFIS models for test data



(c) ANFIS Test data

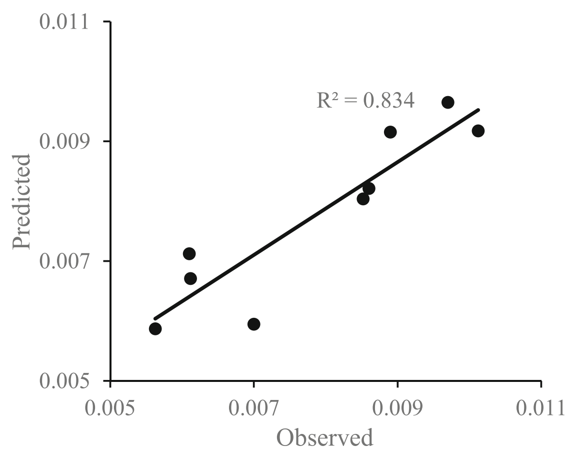


Fig. 17 **a**, **b**, and **c** the scatter plots of observed versus predicted Pb values in test data, using MLP, RBN, and ANFIS models, respectively

to-use and cost-effective alternative to forecast the presence of heavy metals in the river.

Compliance with ethical standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

Abdi-Khanghah M, Bemani A, Naserzadeh Z, Zhang Z (2018) Prediction of solubility of N-alkanes in supercritical CO_2 using RBF-ANN and MLP-ANN. J CO2 Uti 25:108–119 https://www.sciencedirect.com/science/article/abs/pii/S2212982017307291

Abdolmaleki AS, Ahangar AG, Soltani J (2013) Artificial neural network (ANN) approach for predicting Cu concentration in drinking water of Chahnimeh1 Reservoir in Sistan-Balochistan, Iran. Health Scope 2(1):31–38 https://www.sid.ir/en/Journal/ViewPaper.aspx?ID=437069

Agah A, Soleimanpourmoghadam N (2020) Design and implementation of heavy metal prediction in acid mine drainage using multi-output adaptive neuro-fuzzy inference systems (ANFIS) - a case study. Int J Min Geo-Eng 54-1:59–64. https://doi.org/10.22059/IJMGE.2019.278558.594794

Ahamad KU, Raj P, Barbhuiya NH, Deep A (2019) Surface water quality modeling by regression analysis and artificial neural network. In Advances in waste management (pp. 215-230). Springer, Singapore. https://link.springer.com/chapter/10.1007/978-981-13-0215-2_15

Alaqouri HAA, Genc CO, Aricak B, Kuzmina N, Menshikov S, Cetin M (2020) The possibility of using Scots pine needles as biomonitor in determination of heavy metal accumulation. Environ Sci Pollut Res 27:20273–20280. https://doi.org/10.1007/s11356-020-08449-1

Alizamir M, Sobhanardakani S (2016) Forecasting of heavy metals concentration in groundwater resources of Asadabad plain using Artificial Neural Network approach. Adv Environ Health Res 4(2):68–77 http://jaehr.muk.ac.ir/index.php/jaehr/article/view/article_40223.html

Alizamir M, Sobhanardakani S (2017) A comparison of performance of artificial neural networks for prediction of heavy metals concentration in groundwater resources of Toyserkan Plain. Adv Environ Health Res 4(1):11792 http://ajehe.umsha.ac.ir/Article/11792

Altunkaynak A, Başakın EE (2018) River flow estimation by using time series and comparison with different methods. Erzincan University Journal of Science Institute 11(1):92–101

Antanasijević D, Pocajt V, Povrenović D, Perić-Grujić A, Ristić M (2013) Modelling of dissolved oxygen content using artificial neural networks: Danube River, North Serbia, case study. Environ Sci Pollut Res 20(12):9006–9013 https://link.springer.com/article/10.1007/s11356-013-1876-6

Anusree K, Varghese K (2016) Streamflow prediction of Karuvannur River basin using ANFIS, ANN and MNLR models. Procedia Technology 24:101–108 https://www.sciencedirect.com/science/article/pii/S2212017316300986

APHA (1985) WEF, AWWA, 1985. Standard methods for the examination of water and wastewater, 16th ed. APHA, Washington, DC, USA.

Arhami M, Kamali N, Rajabi MM (2013) Predicting hourly air pollutant levels using Artificial Neural Networks coupled with uncertainty analysis by Monte Carlo simulations. Environ Sci Pollut Res 20(7):4777–4789 https://link.springer.com/article/10.1007/s11356-012-1451-6

Aricak B, Cetin M, Erdem R, Sevik H, Cometen H (2019) The change of some heavy metal concentrations in Scotch pine (Pinus sylvestris) depending on traffic density, Organ and washing. Appl Ecol Env Res 17(3):6723–6734

Aricak B, Cetin M, Erdem R, Sevik H, Cometen H (2020) The usability of Scotch pine (Pinus sylvestris) as a biomonitor for traffic originated heavy metal concentrations in Turkey. Pol J Environ Stud 29(2):1051–1057

Asgharnia A, Jamali A, Shahnazi R, Maheri A (2019) Load mitigation of a class of 5-MW wind türbine with RBF neural network based fractional-order PID controller. ISA Transactions 96:272–286. https://doi.org/10.1016/j.isatra.2019.07.006

Assi KJ, Md NK, Ratrout NT, Aldosary AS (2018) Mode choice behavior of high school goers: evaluating logistic regression and MLP neural networks. Case Studies on Transport Policy 6:225–230

Azid A, Juahir H, Toriman ME, Kamarudin MKA, Saudi ASM, Hasnam CNC, Yamin M (2014) Prediction of the level of air pollution using principal component analysis and Artificial Neural Network techniques: a case study in Malaysia. Water Air Soil Pollut 225(8):1–14 https://link.springer.com/article/10.1007/s11270-014-2063-1

Bansal S, Ganesan G (2019) Advanced evaluation methodology for water quality assessment using artificial neural network approach. Water Resour Manage 33(9):3127–3141 https://link.springer.com/article/10.1007/s11269-019-02289-6

Bayatzadeh Fard Z, Ghadimi F, Fattahi H (2017) Use of artificial intelligence techniques to predict distribution of heavy metals in groundwater of Lakan lead-zinc mine in Iran. J Min Environ 8(1):35–48. https://doi.org/10.22044/jme.2016.592

Cetin M (2019) The effect of urban planning on urban formations determining bioclimatic comfort area's effect using satellitia imagines on air quality: a case study of Bursa city. Air Qual Atmos Health 12(10):1237–1249. https://doi.org/10.1007/s11869-019-00742-4

Cetin M, Adiguzel F, Gungor S, Kaya E, Sancar MC (2019) Evaluation of thermal climatic region areas in terms of building density in urban management and planning for Burdur, Turkey. Air Qual Atmos Health 12(9):1103–1112. https://doi.org/10.1007/s11869-019-00727-3

Cetin M, Sevik H, Cobanoglu O (2020) Ca, Cu, and Li in washed and unwashed specimens of needles, bark, and branches of the blue spruce (Picea pungens) in the city of Ankara. Environ Sci Pollut Res 27:21816–21825. https://doi.org/10.1007/s11356-020-08687-3

Chelani AB, Gajghate DG, Hasan MZ (2002) Prediction of ambient PM10 and toxic metals using Artificial Neural Networks. J Air Waste Manag Assoc 52(7):805–810 https://www.tandfonline.com/doi/abs/10.1080/10473289.2002.10470827

Dokmeci AH (2017) Evaluation of heavy metal pollution in the Ergene River Basin from a public health perspective. Turk J Public Health 15(3):212–221

Elzwayie A, El-Shafie A, Yaseen ZM, Afan HA, Allawi MF (2017) RBFNN-based model for heavy metal prediction for different climatic and pollution conditions. Neural Comput Applic 28(8):1991–2003 https://link.springer.com/article/10.1007/s00521-015-2174-7

Fissa MR, Lahiouel Y, Khaouane L, Hanini S (2019) QSPR estimation models of normak boiling point and relative liquid density of püre hydrocarbons using MLR and MLP-ANN methods. J Mol Graph Model 87:109–120 https://www.sciencedirect.com/science/article/abs/pii/S1093326318306673

Ghritlahre HK, Prasad RK (2018) Exergetic performance prediction of solar air heater using MLP, GRNN and RBF models of Artificial Neural Network technique. J Environ Manage 223:566–575 https://www.sciencedirect.com/science/article/pii/S0301479718306753

Guo H, Nguyen H, Vu DA, Bui XN (2019) Forecasting mining capital cost for open-pit mining projects based on Artificial Neural Network approach. Resources Policy 101474. https://doi.org/10.1016/j.resourpol.2019.101474.

Han S, Wang H, Tian Y, Christov N (2019) Time-delay estimation based computed torque control with robust adaptive RBF neural network compensator for a rehabilitation exoskeleton. ISA Transactions 97:171–181. https://doi.org/10.1016/j.isatra.2019.07.030

Hang C, Liu W, Wang P (2019) A method of barkhausen noise feature exctraction based on an adaptive threshold. Applied Sciences 9(15):2964–2974 https://www.mdpi.com/2076-3417/9/15/2964

Haykin S (1998) Neural networks: a comprehensive foundation, second ed. Prentice-Hall, Upper Saddle River, Nj, pp. 26-32.

Heddam S, Kisi O, Sebbar A, Houichi L, Djemili L (2019) Predicting water quality indicators from conventional and nonconventional water resources in Algeria Country: adaptive neuro-fuzzy inference systems versus artificial neural networks. https://link.springer.com/chapter/10.1007/698_2019_399

Hoha GV, Costachescu E, Leahu A, Pasarin B (2014) Heavy metals contamination levels in processed meat marketed in Romania. Environ Engin Manag J. 13(9):2411–2415 http://www.eemj.icpm.tuiasi.ro/pdfs/vol13/no9/full/36_335_Hoha_14.pdf

Isiyaka HA, Mustapha A, Juahir H, Phil-Eze P (2019) Water quality modelling using artificial neural network and multivariate statistical techniques. Model Earth Syst Environ 5(2):583–593 https://link.springer.com/article/10.1007/s40808-018-0551-9

Jang JR (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Transactions on Systems, Man, and Cybernetics 23:665–685 https://ieeexplore.ieee.org/abstract/document/256541/

Küçükerdem TS, Kilit M, Saplıoğlu K (2019) Determination of the number of clusters used in fuzzy inference systems by means of K-means and modeling of dam volume: Kestel Dam example. Pamukkale University Journal of Engineering Sciences 25(8):962–967

Lu H, Li H, Liu T, Fan Y, Yuan Y, Xie M, Qian X (2019) Simulating heavy metal concentrations in an aquatic environment using artificial intelligence models and physicochemical indexes. Sci Total Environ 694:133591 https://www.sciencedirect.com/science/article/pii/S0048969719335168

Magdaleno A, De Cabo L, Arreghini S, Salinas S (2014) Assessment of heavy metal contamination and water quality in an urban river from. Argentina 18(1):113–120

Messikh N, Bousba S, Bougdah N (2017) The use of a multilayer perceptron (MLP) FOR modelling the phenol removal by emulsion liquid membrane. J Environ Chem Engin 5:3483–3489 https://www.sciencedirect.com/science/article/abs/pii/S2213343717303044

Najah A, El-Shafie A, Karim OA, El-Shafie AH (2014) Performance of ANFIS versus MLP-NN dissolved oxygen prediction models in water quality monitoring. Environ Sci Pollut Res 21(3):1658–1670 https://link.springer.com/article/10.1007/s11356-013-2048-4

Nhantumbo C, Carvalho F, Uvo C, Larsson R, Larson M (2018) Applicability of a processes-based model and artificial neural networks to estimate the concentration of major ions in rivers. J

Geochem Explor 193:32–40 https://www.sciencedirect.com/science/article/pii/S0375674217302959

Official Gazette (2012) Surface water quality management regulation. T. C. Ministry of Forestry and Water Affairs Official Gazette Number: 28483. November, Ankara.

Ozel HU, Ozel HB, Cetin M, Sevik H, Gemici BT, Varol T (2019) Base alteration of some heavy metal concentrations on local and seasonal in Bartin River. Environ Monit Assess 191(9):594 https://link.springer.com/article/10.1007/s10661-019-7753-0

Prabu PC (2009) Impact of heavy metal contamination of Akaki river of Ethiopia on soil and metal toxicity on cultivated vegetable crops. Electronic Journal of Environmental, Agricultural and Food Chemistry8(9): 818–827.

Rajab S (2019) Handling interpretability issues in ANFIS using rule base simplifcation and constrained learning. Fuzzy Set Syst 368:36–58 https://www.sciencedirect.com/science/article/abs/pii/S0165011418309175

Rooki R, Doulati A, Aryafar A, Bani Asadi A (2011) Prediction of heavy metals in acid mine drainage using artificial neural network from the Shur River of the Sarcheshmeh porphyry copper mine, Southeast Iran. Environ Earth Sci 64:1303–1316 https://link.springer.com/article/10.1007/s12665-011-0948-5

Sarkar A, Pandey P (2015) River water quality modelling using artificial neural network technique. Aquatic Procedia 4:1070–1077 https://www.sciencedirect.com/science/article/pii/S2214241X15001364

Seifi A, Riahi-Madvar H (2019) Improving one-dimensional pollution dispersion modeling in rivers using ANFIS and ANN-based GA optimized models. Environ Sci Pollut Res 26(1):867–885 https://link.springer.com/content/pdf/10.1007/s11356-018-3613-7.pdf

Sevik H, Ozel HB, Cetin M, Ozel HU, Erdem T (2019a) Determination of changes in heavy metal accumulation depending on plant species, plant organism, and traffic density in some landscape plants. Air Qual Atmos Health 12(2):189–195 https://link.springer.com/article/10.1007/s11869-018-0641-x

Sevik H, Cetin M, Ozel HB, Pinar B (2019b) Determining toxic metal concentration changes in landscaping plants based on some factors. Air Qual Atmos Health 12(8):983–991 https://link.springer.com/article/10.1007%2Fs11869-019-00717-5

Sevik H, Cetin M, Ucun Ozel H, Ozel HB, Mossi MMM, Zeren CI (2020a) Determination of Pb and Mg accumulation in some of the landscape plants in shrub forms. Environ Sci Pollut Res 27(2):2423–2431. https://doi.org/10.1007/s11356-019-06895-0

Sevik H, Cetin M, Ozel HB, Akarsu H, Zeren Cetin I (2020b) Analyzing of usability of tree-rings as biomonitors for monitoring heavy metal accumulation in the atmosphere in urban area: a case study of cedar tree (Cedrus sp.). Environ Monit Assess 192:23. https://doi.org/10.1007/s10661-019-8010-2

Sevik H, Cetin M, Ozel HB, Ozel S, Zeren CI (2020c) Changes in heavy metal accumulation in some edible landscape plants depending on traffic density. Environ Monit Assess 192(78). https://doi.org/10.1007/s10661-019-8041-8

Shanbehzadeh S, Vahid Dastjerdi M, Hassanzadeh A, Kiyanizadeh T (2014) Heavy metals in water and sediment: a case study of Tembi River. J Environ Public Health 2014:28–34 https://www.hindawi.com/journals/jeph/2014/858720/

Singh KP, Basant A, Malik A, Jain G (2009) Artificial neural network modeling of the river water quality—a case study. Ecol Model 220(6):888–895 https://www.sciencedirect.com/science/article/abs/pii/S0304380009000283

Swietlicka I, Sujak A, Muszyński S, Świetlicki M (2017) The application of artificial neural networks to the problem of reservoir classification and land use determination on the basis of water sediment composition. Ecol Indic 72:759–765 https://www.sciencedirect.com/science/article/abs/pii/S1470160X1630543X

Turkyilmaz A, Sevik H, Cetin M (2018a) The use of perennial needles as bio-monitors for recently accumulated heavy metals. Landsc Ecol Eng 14(1):115–120. https://doi.org/10.1007/s11355-017-0335-9

Turkyilmaz A, Sevik H, Cetin M, Ahmaida Saleh EA (2018b) Changes in heavy metal accumulation depending on traffic density in some landscape plants. Pol J Environ Stud 27(5):2277–2284. https://doi.org/10.15244/pjoes/78620

Venkatramanan S, Chung SY, Selvam S, Son JH, Kim YJ (2017) Interrelationship between geochemical elements of sediment and groundwater at Samrak Park Delta of Nakdong River Basin in Korea: multivariate statistical analyses and Artificial Neural Network approaches. Environ Earth Sci 76(13):456 https://link.springer.com/article/10.1007/s12665-017-6795-2

Yilmaz I, Kaynar O (2011) Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. Expert Syst Appl 38(5):5958–5966 https://www.sciencedirect.com/science/article/abs/pii/S0957417410012649