



Water Quality Assessment with Artificial Neural Network Models: Performance Comparison Between SMN, MLP and PS-ANN Methodologies

Hakan Isık¹ · Tamer Akkan²

Received: 27 January 2024 / Accepted: 26 May 2024
© The Author(s) 2024

Abstract

Identifying and measuring potential sources of pollution is essential for water management and pollution control. Using a range of artificial intelligence models to analyze water quality (WQ) is one of the most effective techniques for estimating water quality index (WQI). In this context, machine learning-based models are introduced to predict the WQ factors of Southeastern Black Sea Basin. The data comprising monthly samples of different WQ factors were collected for 12 months at eight locations of the Türkiye region in Southeastern Black Sea. The traditional evaluation with WQI of surface water was calculated as average (i.e. good WQ). Single multiplicative neuron (SMN) model, multilayer perceptron (MLP) and pi-sigma artificial neural networks (PS-ANNs) were used to predict WQI, and the accuracy of the proposed algorithms were compared. SMN model and PS-ANNs were used for WQ prediction modeling for the first time in the literature. According to the results obtained from the proposed ANN models, it was found to provide a highly reliable modeling approach that allows capturing the nonlinear structure of complex time series and thus to generate more accurate predictions. The results of the analyses demonstrate the applicability of the proposed pi-sigma model instead of using other computational methods to predict WQ both in particular and other surface water resources in general.

Keywords Water management · Water quality prediction · Artificial neural networks · Pi-sigma neuron model · Sustainability

1 Introduction

The significance of water to the natural ecosystem is indisputable. However, the deterioration of water resources has long been a consequence of both natural phenomena and anthropocene activities. In the global panorama, the deterioration of water quality (WQ) represents a primary threat to public health [1]. One of the most widely accepted methods for detecting biological problems associated with aquatic environments is the assessment of WQ. The monitoring of key ecological indicators, which exhibit variations due to seasonal and temporal differences, is a valuable tool for investigating contamination rates [2–4]. Many studies conducted to determine the importance of water pollution have reported

that water resources play a critical role in protecting the quality of water resources and ensuring human health and the sustainability of ecosystems [5–7]. The deterioration of freshwater quality is a consequence of anthropogenic activities that contaminate aquatic habitats. The resulting impact on human health is a problem related to ecosystem degradation, management, and monitoring [8]. Currently, there is no fixed number and type of parameters for assessing WQ. Numerous studies have been conducted to assess WQ, including parameters grouped into indices such as pH, dissolved oxygen, turbidity, and metals. [9, 10]. One of the most efficacious instruments for disclosing information pertaining to WQ is the utilization of the term “water quality index” (WQI), which is a rating component that elucidates the collective impact of disparate WQ parameters. WQI provides a single number that expresses the overall WQ at a given time and place based on various WQ parameters [11].

The fact that artificial neural networks (ANNs) work based on data, adapt to data, and learn by establishing a link between input and output makes them different from many other methods. With these features, ANNs have become a frequently

✉ Tamer Akkan
biyoloji@yahoo.com

¹ Division of Biology, Graduate School of Natural and Applied Sciences, University of Giresun, Giresun, Türkiye

² Department of Biology, Faculty of Arts and Science, University of Giresun, Giresun, Türkiye



preferred method in the fields of classification, pattern recognition and prediction [12]. WQ modeling has been developed to solve WQ problems with advanced computing using artificial intelligence (AI) techniques. ANNs have helped monitor WQ systems by predicting changes in WQ [13]. Artificial intelligence methods can significantly reduce the costs of water supply and sanitation systems and help ensure compliance with drinking and wastewater treatment quality [14]. ANN models have been used to solve many environmental problems. For example, a number of studies have used ANN models to predict water resource variables such as water quantity and WQ in river systems [15]. Previous studies [16–19] have applied different machine learning algorithms to analyze ground or surface WQ.

Estimation of WQ plays an important role in environmental monitoring, ecosystem sustainability and aquaculture. Conventional estimation methods cannot predict the linear and non-stationary state of good WQ [20]. Also, traditional methods for calculating WQI are time-consuming and have errors in subdirectory derivatives [21]. The WQI has been adopted as a universal indicator that comprehensively represents the WQ status of the surface water body. Common traditional water quality indices (WQIs) suffer from limitations such as “shadowing” and “uncertainty” [22].

In addition, the accuracy of traditional WQ estimation methods is generally low, and the estimation results have significant autocorrelation [23]. Therefore, the traditional linear forecasting model cannot fully reflect its changing regulation and cannot accurately predict WQ. Reasons for needing mathematical techniques and models that can model and predict WQ efficiently increase in data scale and difficulties in interpreting land use, pollutant loading and disposal, WQ and ecosystem effects [24].

Techniques such as ANNs have gained importance to overcome the problems experienced in traditional WQIs [22]. WQ monitoring plays a vital role in protecting water resources, environmental management and decision-making. Artificial intelligence (AI), based on machine learning techniques, has been widely used in recent years to evaluate and classify WQ [25].

Regarding the simulation and prediction of WQ, adaptive boosting (Adaboost) [26], gradient boosting (GBM) [27], extreme gradient boosting (XGBoost) [28], decision tree (DT) [29, 30], extra trees (ExT) [31], random forest (RF) [26], multilayer perceptron (MLP) [32], radial basis function (RBF) [33], deep feed-forward neural network (DFNN) [34] and convolutional neural network (CNN) [17] have been reported. Among these techniques, the multilayer perceptron (MLP) model has mostly outperformed the others in precision and accuracy [35] and is the most widely used architecture [36]. MLP is a feed-forward ANN model that maps input data sets to appropriate outputs. It uses a supervised learning technique that includes the use of back propagation

algorithm [37]. Although there are many ML algorithms, researchers still face problems such as which ML techniques should be applied or which is the most suitable for a particular problem [38]. It is very important to minimize the impact of basin-based variability and complexity and nonlinearity in water quality determination. A robust and flexible model is also needed to improve the accuracy of WQI estimation [21].

As can be seen from the above applied models, many techniques can be applied for WQ estimation. The motivation of our study is the use of single multiplicative neuron model and pi-sigma artificial neural networks, which have not been used before in the modeling of the WQI and allow effective modeling. For this reason, it is necessary to first determine the drinking WQ variables and calculate with a good WQI. Then, it would be more useful and logical to use neural network modeling to map the relationship between them. Due to various problems such as time-consuming and low accuracy of current WQ calculation methods from the studies, it has encouraged to carry out a study with an ANN-based approach. In this research, the WQ of surface water data from southeastern Black Sea Basin was made with the help of artificial neural network (ANN) modeling. The contributions of this study to the literature are as follows.

- The WQ of surface water bodies was monitoring, and WQIs were calculated.
- Single multiplicative neuron model and pi-sigma artificial neural networks are firstly used for modeling of WQ.
- Six different optimal artificial neural networks are proposed for modeling and estimating the surface WQI.
- Pi-sigma artificial neural network was determined as the best artificial neural network model for modeling the surface WQI.

2 Material and Methods

2.1 Study Area

For this study, surface water samples collected from eight different stations for 12 months from the rivers flowing into the southeastern Black Sea were used as raw data to predict WQ using ANN (Fig. 1).

2.2 Physicochemical Analysis

In this study, monthly samples were brought to the laboratory and analyzed by applying standard methods. The samples were taken in accordance with the “Water Pollution Control Regulation Sampling and Analysis Methods Communiqué”

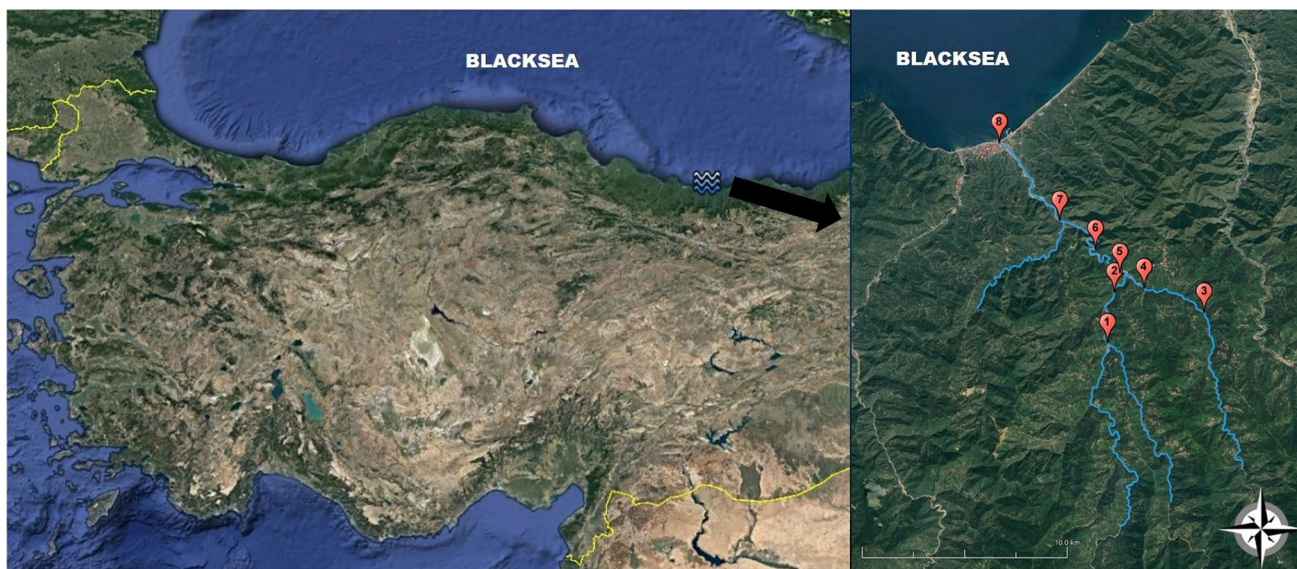


Fig. 1 Study location (adapted from Google Earth)

and were transported to the laboratory environment by keeping them in glass bottles at a temperature of + 4 °C, away from light. After the sample containers were shaken, they were immersed about 15 cm below the water surface and filled. Water temperature (WT), pH, dissolved oxygen (DO), electrical conductivity (EC), and total dissolved solids (TDS) were measured with the YSI 556 MPS. Turbidity (NTU) was measured in the field with the help of WTW brand land-type devices. The other variables for WQ, such as total ammonium nitrogen (TAN), total phosphorus (TP), and biological oxygen demand (BOD) were analyzed on the same day in the laboratory.

2.3 Statistical Analysis

In this paper, statistical analyses were conducted using the SPSS 17.0 for Windows software. The data set utilized in the analyses was collected seasonally and consisted of 96 observations. Descriptive statistics (mean, standard deviation, minimum, and maximum) of the data set were obtained along with statistical calculations.

2.4 Calculation of Water Quality Index (WQI)

WQI is defined as a rating reflecting the combined effect of different WQ parameters. The WQI is an indicator of the quality of any groundwater reservoir in the form of a single number representing a combination of different WQ parameters [39]. The drinking WQI was calculated using the method previously recommended by various research groups [40, 41].

The first WQI calculation method represents the weight attributed to each parameter and is assigned for each parameter from the PCA results based on the eigenvalues for each major component and factor loading and represents the relative importance of each WQ parameter for drinking purposes [42]. Based on these mentioned techniques, WQI was investigated for the water samples. The concentration of each measured parameter is C_i (mg/L) and the standard drinking water values allowed by the surface S_i . The whole equation showing drinking WQ is shown below [43, 44].

$$WQI_1 = \sum \left[W_i x \left(\frac{C_i}{S_i} \right) \right] \times 100 \tag{1}$$

In the second WQI calculation method, the following steps were followed. In the first step, each of the eight parameters was assigned a weight (w_i) between 1–5 (minimum to highest impact on WQ) based on their impact on health and their importance to the overall quality of drinking water (Table 1). Second, the relative weight (W_i) of the parameters (where n is the number of parameters) is calculated using the following equation [45]:

$$W_i = \frac{w_i}{\sum_{i=1}^n W_i} \tag{2}$$

Third, the quality rating scale (q_i) was calculated using equality (3). Determined concentration of each used parameter (C_i) and standard safe limit values (S_i) is as follows:

$$q_i = \frac{C_i}{S_i} \times 100 \tag{3}$$

Table 1 Relative weight of the parameters [44]

Parameters	Weight (wi)	Relative weight (Wi)
WT	4.0	0.11
EC	3.0	0.09
TDS	5.0	0.14
DO	5.0	0.14
Turbidity	5.0	0.14
TP	4.0	0.11
TAN	4.0	0.11
BOD	5.0	0.14
	$\sum 35$	$\sum 1$

Finally, the WQI_2 was calculated using Eq. (4).

$$\sum_{i=1}^n W_i \times q_i \tag{4}$$

WQI_2 values calculated at the end of both applications are classified into five categories as follows: WQ is excellent when $0 \leq WQI < 50$; WQ is good when $50 \leq WQI < 100$; when $100 \leq WQI < 200$, the WQ is poor; if $200 \leq WQI < 300$, the WQ is very bad; if $WQI \geq 300$, the water is not drinkable [46, 47].

2.5 Artificial Neural Networks (ANN)

ANN, a concept taken from the human mind, is a widely used forecasting method [48]. The neural networks main characteristic is their learning capacity [49]. The concept of an artificial neural network (ANN) is a computational process that attempts to describe and generate a mapping from a multivariate data set as layers of input neurons to one or more hidden layers and an output layer containing one or more output neurons [50]. This makes long calculations and complex conditions easier understandable.

Similarly, assessments of the WQI include complex computations and transformation. In the WQI calculation for each formula, the parameters are formulated with different values or ranges of values for each formula. Therefore, alternatively, an ANN has been used to model the WQI of rivers [32]. Also, WQI determination is also made easier and quicker [51].

The usefulness of ANNs spans many different fields of study, including water resources engineering, hydrogeology, and environmental geology. ANNs are widely used in modeling groundwater and surface water resources [52]. Generally, obtaining optimal ANN models for WQ prediction involves simultaneous training of several networks with a wide variety of neuron counts, input and output activation functions, and training optimization algorithms; finally, the most suitable ANN model is selected for interpretation [53].

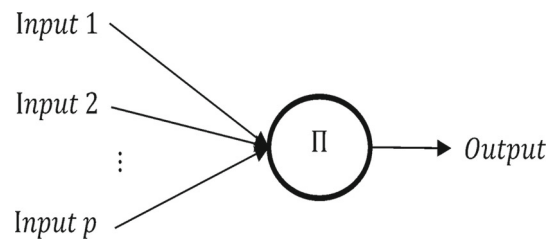


Fig. 2 Architecture of the single multiplicative artificial neural network

Before implementing any data intelligence model, several steps must be implemented to ensure the success of the forecast. One of the most important steps is to determine the best input combination of the model [54]. In this study, single multiplicative neuron model artificial neural network, pi-sigma artificial neural network models and multilayer perceptron were used to predict the result of WQI. In recent years, multilayer perceptron neural networks have been used in many studies for modeling WQ. In this study, we focused on the performance of the multiplicative neuron model artificial neural network and pi-sigma artificial neural network, which can produce more successful prediction results than multilayer perceptron. Multiplicative neuron model artificial neural network and pi-sigma artificial neural network are firstly applied for modeling WQ in the literature.

2.6 Single Multiplicative Neuron Model Artificial Neural Network

The single multiplicative artificial neural network was firstly proposed by [55]. Although the single multiplicative artificial neural network has a simple architectural structure, it is an artificial neural network that can solve complex nonlinear problems. Since it can work with a much less number of neurons and therefore fewer parameters than a multilayer perceptron, its generalization ability is stronger than a multilayer perceptron. The architecture of the single multiplicative artificial neural network is given in Fig. 2.

In the single multiplicative artificial neural network, a nonlinear transformation of the product of the linear transformations of the inputs is calculated as the output. Calculation of the outputs of the single multiplicative artificial neural network with p inputs is carried out with the following formulas:

$$\text{net} = \prod_{i=1}^p (w_i \times \text{input}_i + b_i) \tag{5}$$

$$\text{output} = \frac{1}{1 + \exp(-\text{net})} \tag{6}$$

The ANN contains a total of $2p$ weights and biases values. Training this neural network is a problem of estimating $2p$ parameters. The objective function in the optimization problem can be used as the sum of squares of error. The

optimization problem is expressed as follows:

$$\min_{\{w_1, \dots, w_p, b_1, \dots, b_p\}} \sum_{j=1}^n (\text{output}_j - \text{target}_j)^2 \tag{7}$$

In formula (7), n represents the number of learning examples. The solution to the optimization problem provides the set of parameter values that will enable the network to produce outputs closest to the targets. The optimization problem given by formula (7) can be solved by nonlinear least squares methods as well as by artificial intelligence optimization algorithms such as genetic algorithm and particle swarm optimization. It is well-known that particle swarm optimization can produce very successful results in the training of ANNs. In this study, a particle swarm optimization training algorithm that imitates swarm intelligence developed for numerical optimization problems is preferred. The training algorithm based on PSO is given below in steps.

Algorithm 1.

The training algorithm based on PSO for a single multiplicative artificial neural network [56].

Step 1. The parameters of the training processes are determined.

- pn* : The swarm size or particle number
- c*₁: Social coefficient
- c*₂: Cognitive coefficient
- w* : Inertia weight
- ε : The error tolerance for relative error difference
- maxitr* : The maximum number of iterations.

The counters are initialized. The restarting strategy counter (*rsc*) and early stopping counter (*esc*) are initially taken as zero.

Step 2. The initial positions and velocities are randomly generated as follows:

$$X_{i,j}^{(0)} \sim \text{Uniform}(0,1) \tag{8}$$

$$V_{i,j}^{(0)} \sim \text{Uniform}(-1,1) \tag{9}$$

$X_{i,j}^{(k)}$ is *j*th position value of *i*th particle of the population at *k*th iteration. The positions of an element in a population correspond to the weights and biases values of the neural network, and there are $2p$.

Step 3. The fitness function values are calculated for each swarm member. The fitness function is selected as the sum of square errors.

$$\text{SSE} = \sum_{j=1}^n (\text{output}_j - \text{target}_j)^2 \tag{10}$$

Step 4. According to the calculated fitness function values, the best element (X_{best}^k) in the population is determined as *gbest* and its fitness value ($\text{SSE}_{\text{best}}^k$) is saved. Moreover, the

Pbest matrix is constituted as a memory for each particle in the swarm.

Step 5. A new swarm is created by replacing the positions of all elements in the swarm with the following equation:

$$V_{i,j}^{(k+1)} = wV_{i,j}^{(k+1)} + c_1r_1(P\text{best}_{i,j}^{(k)} - X_{i,j}^{(k)}) + c_2r_2(X\text{best}_j^{(k)} - X_{i,j}^{(k)}) \tag{11}$$

$$X_{i,j}^{(k+1)} = X_{i,j}^{(k)} + V_{i,j}^{(k+1)} \tag{12}$$

Step 6. The fitness function values are calculated for each swarm member by using Eq. (11). According to the calculated fitness function values, the best element (X_{best}^k) in the population and its fitness value ($\text{SSE}_{\text{best}}^k$) are obtained and compared with $\text{SSE}_{\text{best}}^{k-1}$. If $\text{SSE}_{\text{best}}^k > \text{SSE}_{\text{best}}^{k-1}$, then $\text{SSE}_{\text{best}}^k = \text{SSE}_{\text{best}}^{k-1}$.

Step 7. The restarting strategy counter (*rsc* = *rsc* + 1) is increased, and its value is checked. If the *rsc* > *limit1*, then all positions are regenerated by using (8) and (9), and the *rsc* is taken as zero.

Step 8. The early stopping rule is checked. The *esc* counter is increased depending on the following condition:

$$\text{esc} = \begin{cases} \text{esc} + 1, & \text{if } \frac{\text{SSE}_{\text{best}}^{(k)} - \text{SSE}_{\text{best}}^{(k-1)}}{\text{SSE}_{\text{best}}^{(k)}} < \varepsilon \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

If *esc* > *limit2* is satisfied, the algorithm is stopped otherwise go to Step 5.

2.7 Pi-Sigma Artificial Neural Network

Pi-sigma neural network was firstly proposed by [57]. Pi-sigma artificial neural networks are a high-order type of neural network that is effectively used in time series prediction problems. This model is a sophisticated ANN approach that successfully captures complex structures in time-varying data [58]. Pi-sigma neural network is a high-order neural network proposed as an alternative to multilayer perceptron. Pi-sigma neural network is a type of neural network that uses additive and multiplicative aggregation functions together and performs well in solving prediction problems. Unlike the multilayer perceptron in the pi-sigma artificial neural network, the connection weights between the hidden layer and the input layer are taken as constant and one. Therefore, the number of parameters is less than the multilayer perceptron in the pi-sigma artificial neural network. The architecture of the pi-sigma neural network is given in Fig. 3.

The output of the pi-sigma neural network, which includes *p* inputs and *m* hidden layer units, is calculated with the

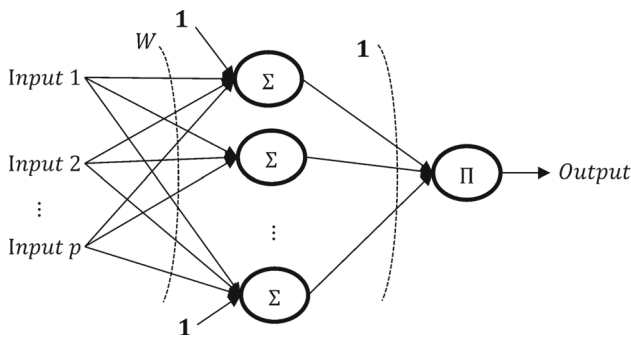


Fig. 3 Architecture of the pi-sigma artificial neural network

following equations in a few steps.

$$net_j = \sum_{i=1}^p w_{i,j} input_i + b_j, \quad j = 1, 2, \dots, m \quad (14)$$

$$output = f(\prod_{j=1}^m f(net_j)) \quad (15)$$

Here, f is the activation function, which is chosen as the logistic function given below in this study.

$$f(net) = \frac{1}{1 + \exp(-net)} \quad (16)$$

The ANN contains a total of $p \times m + m$ weights and biases values. Training this neural network is a problem of estimating $(w_{i,j}, b_j), i = 1, 2, \dots, p, j = 1, 2 \dots, m$ parameters. The objective function in the optimization problem can be used as the sum of squares of error. The optimization problem is expressed as follows.

$$\min_{\{w_{i,j}, b_j / i=1, 2, \dots, p, j=1, 2, \dots, m\}} \sum_{j=1}^n (output_j - target_j)^2 \quad (17)$$

Pi-sigma neural network can be trained with particle swarm optimization using Algorithm 1. In the algorithm, (17) is used instead of (3) as a fitness function. Also, the number of positions of the particles will be $p \times m + m$, not $2 \times p$.

2.8 Multilayer Perceptron Artificial Neural Network

The multilayer perceptron is the most commonly used type of neural network. The multilayer perceptron consists of the input layer, hidden layers and output layers. In the literature, the multilayer perceptron containing a single hidden layer has successfully solved many problems (Fig. 4).

Calculation of the outputs of a multilayer perceptron with p -input and m hidden layer units is performed with the following formula. Here, f is the logistic activation function in (16)

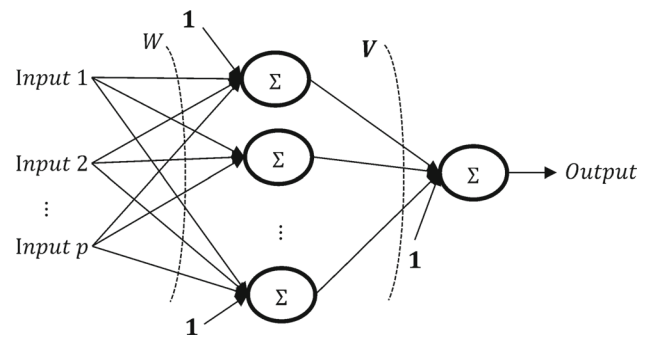


Fig. 4 Architecture of the multilayer perceptron artificial neural network

$$net_j = \sum_{i=1}^p w_{i,j} input_i + b_j, \quad j = 1, 2, \dots, m \quad (18)$$

$$o_j = f(\prod_{j=1}^m f(net_j)), \quad j = 1, 2, \dots, m \quad (19)$$

$$output = f(\sum_{j=1}^n v_j o_j + b) \quad (20)$$

The ANN contains a total of $p \times m + 2m + 1$ weights and biases values. Training this neural network is a problem of estimating $(w_{i,j}, v_j, b_j, b), i = 1, 2, \dots, p, j = 1, 2 \dots, m$ parameters. The objective function in the optimization problem can be used as the sum of squares of error. The optimization problem is expressed as follows:

$$\min_{\{w_{i,j}, b_j, v_j, b / i=1, 2, \dots, p, j=1, 2, \dots, m\}} \sum_{j=1}^n (output_j - target_j)^2 \quad (21)$$

Multilayer perceptron artificial neural network can be trained with particle swarm optimization using Algorithm 1. In the algorithm, (21) is used instead of (3) as a fitness function. Also, the number of positions of the particles will be $p \times m + 2m + 1$.

Furthermore, to obtain the best model, the predicted values of the models were compared with each other using the mean square error (RMSE) and mean absolute percentage error (MAPE) criteria. RMSE and MAPE were calculated using the equations given in (22) and (23), respectively. Model training, statistical analysis of parameters, calculation of correlation coefficients, and error analysis mainly implemented on MATLAB 2018b.

RMSE and MAPE values, whose formulas are given below, are given as error criteria.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (output_t - target_t)^2} \quad (22)$$

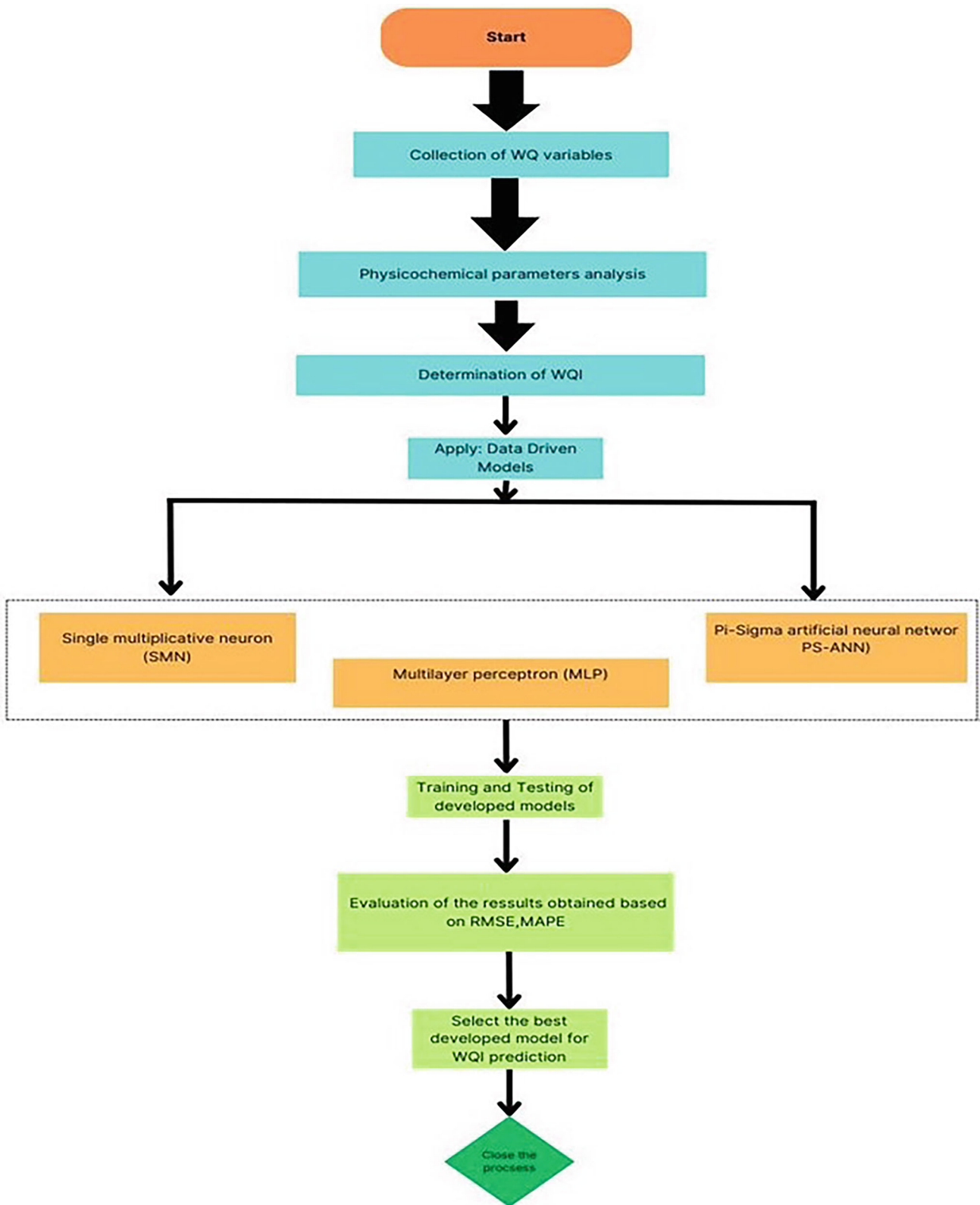


Fig. 5 A basic flowchart of the methodology for estimating WQI

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{\text{output}_t - \text{target}_t}{\text{target}_t} \right| \quad (23)$$

3 Results

In our study, ANN methods were used to predict the WQI results calculated using WQ parameters. For this purpose, the prediction performance results of single multiplicative neuron model artificial neural network, pi-sigma artificial neural network and multilayer perceptron artificial neural network models, which are statistical tools, were compared using MATLAB to predict the WQI value obtained in the study. Coefficient of determination root-mean-square error (RMSE) and mean absolute percentage error (MAPE) were used to evaluate the quality of the models. By comparing the validation values, it is aimed to select the application with the best prediction performance for WQI.

The framework of the proposed methodology consists of four stages: data collection from the study area, analysis of physicochemical parameters, calculation of WQI, application of ANN models and evaluation of model performances, as shown in Fig. 5.

3.1 Analysis of the Data Set for Traditional Evaluation

The physicochemical variables values obtained from the surface water samples were determined. Average values of physicochemical values are given in Table 2. The mean physicochemical values obtained in our study were determined for WT: 14.05, pH: 7.51, EC: 230.51 $\mu\text{S cm}^{-1}$, TDS: 0.1694 mg/L, DO: 10.07 mg/L, Turbidity: 27.75 NTU, TP: 0.061 mg/L, TAN: 0.413 mg/L, and BOD: 4.61 mg/L (Fig. 6). Also, Table 2 provides a detailed statistical overview of the data set and factors affecting WQ.

The WQI₁ value was between 45 and 199 (minimum and maximum), according to the first calculation method we applied, which was calculated based on the average concentrations for eight different water parameters (Table 3). The average WQI value was 72%, and good WQ was determined. However, unsuitable WQ was also found at different times in the stations.

The WQI₂, which we applied as the second method, was evaluated to reveal the drinking WQ of Gelevera Stream. The calculated WQI value was between 55 and 235 (minimum and maximum), and the average value was 86%, and it was determined that it was in the good water category [44]. However, bad water and unsuitable WQ were also encountered at different time intervals on the basis of stations.

Although WQI₂ model results require weight assignment for calculation, your results are as accurate as WQI₁. It therefore means that the weight assignment made by the authors in this study is correct and justified.

3.2 Proposed Modeling of Water Quality Indices by Using Artificial Neural Networks

For the surface WQ indices calculated with three different methods were modeled with single multiplicative neuron model artificial neural network, pi-sigma artificial neural network models and multilayer perceptron. Single multiplicative neuron model artificial neural network and pi-sigma artificial neural network models are firstly applied for modeling WQ in this paper in the literature. In practice, the inputs of the ANN are the variables used in the calculation of the WQI.

The target value is the previously calculated WQI value. Eight measurements of WQI values for all stations in the last month were used as test data and other data as training data. The data set is randomly divided to two sets as training and test set. The learning algorithm is based on particle swarm optimization. The all ANN applications were conducted in MATLAB by using notebook with Intel(R) Core (TM) i5-4210U CPU 2.40 GHz processor. Optimum neural network architectures and estimations of weights and biases obtained for both WQ indices are given in Figs. 7, 8, 9, 10, 11, 12.

In Table 4, the test set performance of the optimum ANNs used in the modeling of the WQI₁ is given.

When Table 4 is examined, it is seen that the pi-sigma artificial neural network given in Fig. 8 has the lowest RMSE and MAPE values in the modeling of WQI₁.

In Table 5, the test set performance of the optimum artificial neural networks used in the modeling of the WQI₂ is given.

When Table 5 is examined, it is seen that the pi-sigma artificial neural network given in Fig. 11 has the lowest RMSE and MAPE values in the modeling of WQI₂.

4 Discussion

4.1 Traditional Evaluation

The results of the analyses of the physicochemical parameters of the river water provide us with a comprehensive understanding of the WQ [59]. The effect of temperature on the rate of chemical reaction with respect to water is a very important factor because of its effect on aquatic life and beneficial use of water. In our study, the overall temperature average was determined as 14.05 °C. According to the Surface Water Quality Regulation (2015), the WQ of studied samples is first class. The pH level generally obtained in the

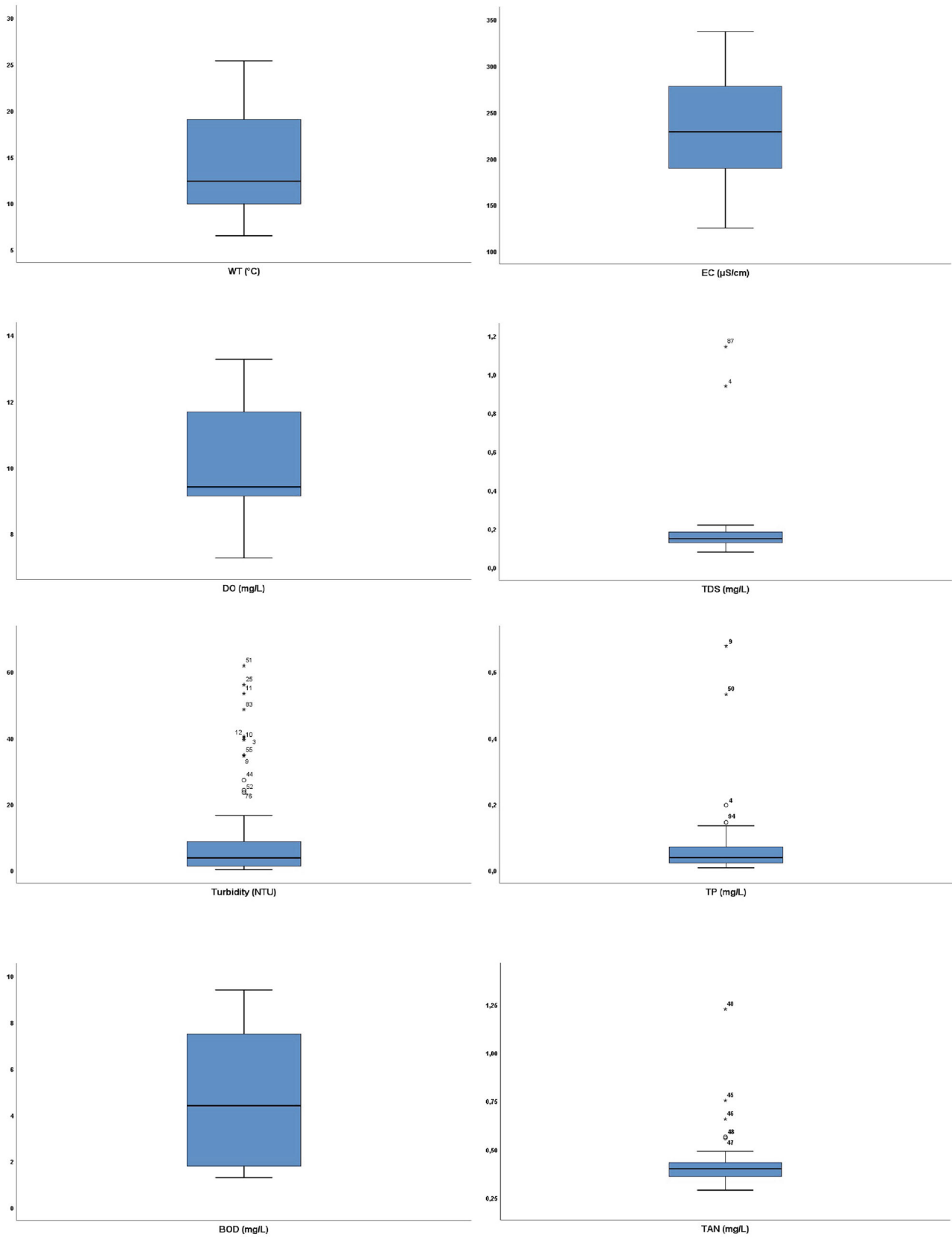


Fig. 6 Box plots of physicochemical variables of surface water samples

Table 2 Result of physicochemical variables of surface water samples

Parameters	Mean	Median	SE \pm	Min	Max	95% Confidence Interval		Skewness	Kurtosis	Coefficient of variation
						Lower	Upper			
WT ($^{\circ}$ C)	14.0547	12.365	0.52575	6.49	25.37	13.011	15.098	0.497	- 1.045	0.367
EC (μ Scm $^{-1}$)	230.510	229	5.55007	125.00	337.00	219.492	241.529	- 0.216	- 0.904	0.236
TDS (mg/L)	0.1694	0.1495	0.01361	0.08	1.14	0.1424	0.1964	6.236	41.429	0.787
DO (mg/L)	10.0763	9.41	0.15064	7.26	13.27	9.7772	10.375	0.661	- 0.601	0.146
pH	7.5151	7.505	0.01779	6.78	7.81	7.4798	7.550	- 0.979	3.262	0.023
Turbidity (NTU)	8.8486	3.75	1.35725	0.17	61.83	6.1542	11.543	2.396	5.288	1.503
TP (mg/L)	0.0616	0.039	0.00899	0.01	0.68	0.0438	0.0795	5.340	32.965	1.43
TAN (mg/L)	0.4139	0.399	0.01097	0.29	1.23	0.3921	0.4357	5.090	35.202	0.26
BOD (mg/L)	4.6167	4.4	0.27590	1.30	9.40	4.0689	5.1644	0.262	- 1.383	0.586

Table 3 Weights for different variables in surface water samples in this study

PC	Eigenvalue	Relative eigenvalue	Variable	Loading value	Relative loading value on same PC	Weight
1	3.02	0.552	WT	0.952	0.256	0.14
			EC	0.541	0.145	0.08
			DO	0.881	0.237	0.13
			TAN	0.410	0.110	0.14
			BOD	0.934	0.251	0.14
2	1.34	0.245	Turbidity	0.767	0.498	0.12
			TP	0.772	0.502	0.06
3	1.11	0.203	TDS	0.918	1.000	0.20

Table 4 Test performance results for modeling of WQI₁

Test Values	Methods		
	SMN-ANN	PS-ANN	MLP
82.929600	79.728915	81.734372	81.150364
88.484900	81.524285	89.514999	85.193944
91.526200	82.980763	88.398163	89.639241
77.050600	77.189467	75.979367	80.115037
77.791600	75.362140	79.132977	79.823172
73.844500	71.875387	72.540313	77.695892
67.114000	71.772131	70.196756	74.539965
68.398000	72.385872	70.596851	75.487315
RMSE	4.731800	1.976800	4.349000
MAPE	0.051200	0.023200	0.051500

Table 5 Test performance results for modeling of WQI₂

Test values	Methods		
	SMN-ANN	PS-ANN	MLP
83.000000	85.047173	83.510551	83.332535
88.000000	86.829469	86.054130	90.427683
92.000000	85.966027	87.972548	91.426646
77.000000	79.544201	78.117690	78.913552
78.000000	78.040712	77.282178	81.154644
74.000000	75.742839	75.151115	76.307917
67.000000	73.086176	71.118659	74.490094
68.000000	73.308245	71.984981	74.550512
RMSE	3.819400	2.650500	3.941500
MAPE	0.039900	0.028300	0.042800

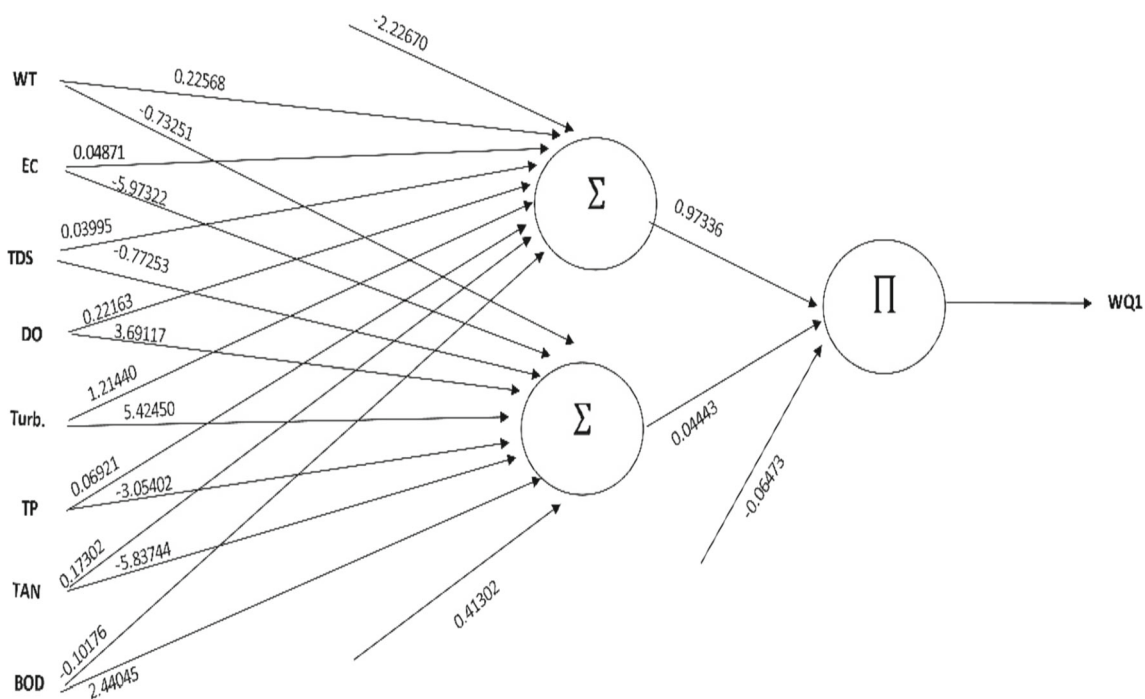


Fig. 7 Optimum multilayer perceptron artificial neural network for WQI₁

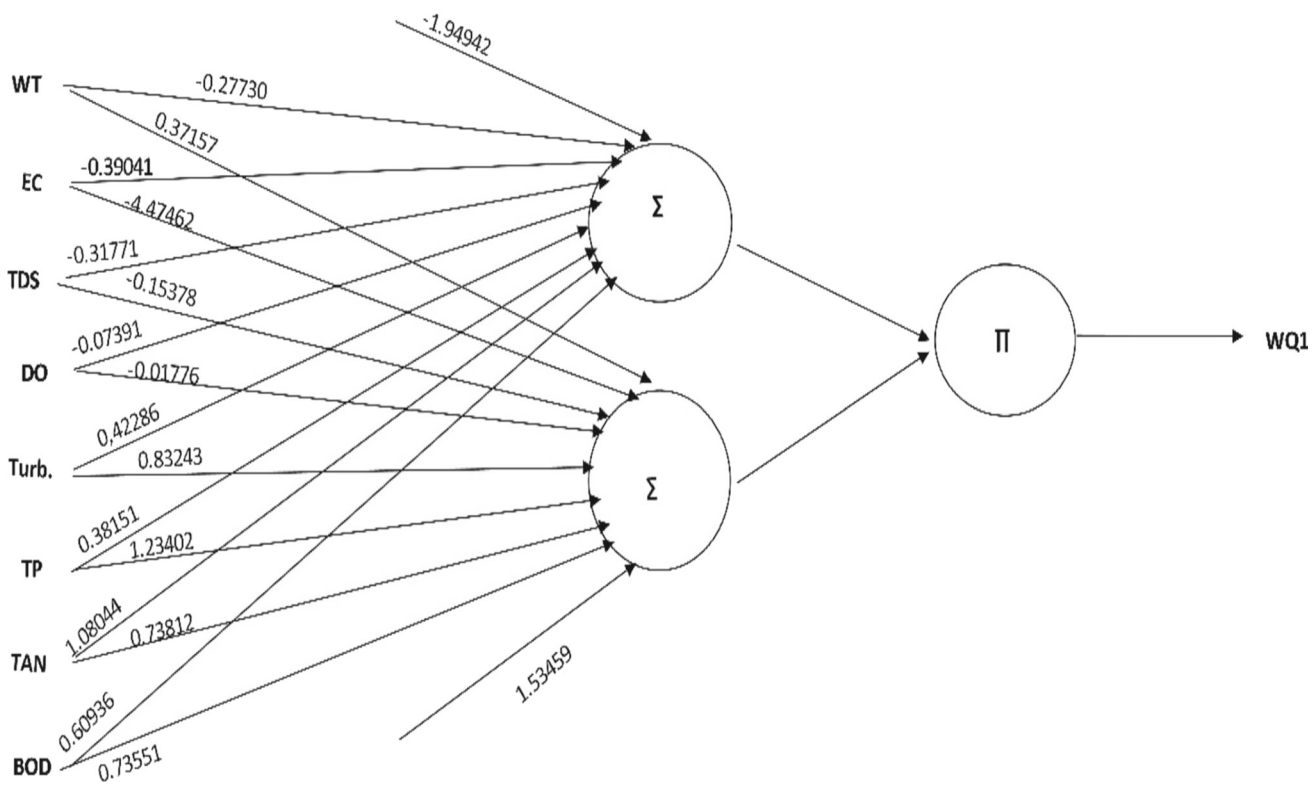


Fig. 8 Optimum pi-sigma artificial neural network for WQI₁

Fig. 9 Optimum single multiplicative neuron model artificial neural network for WQI₁

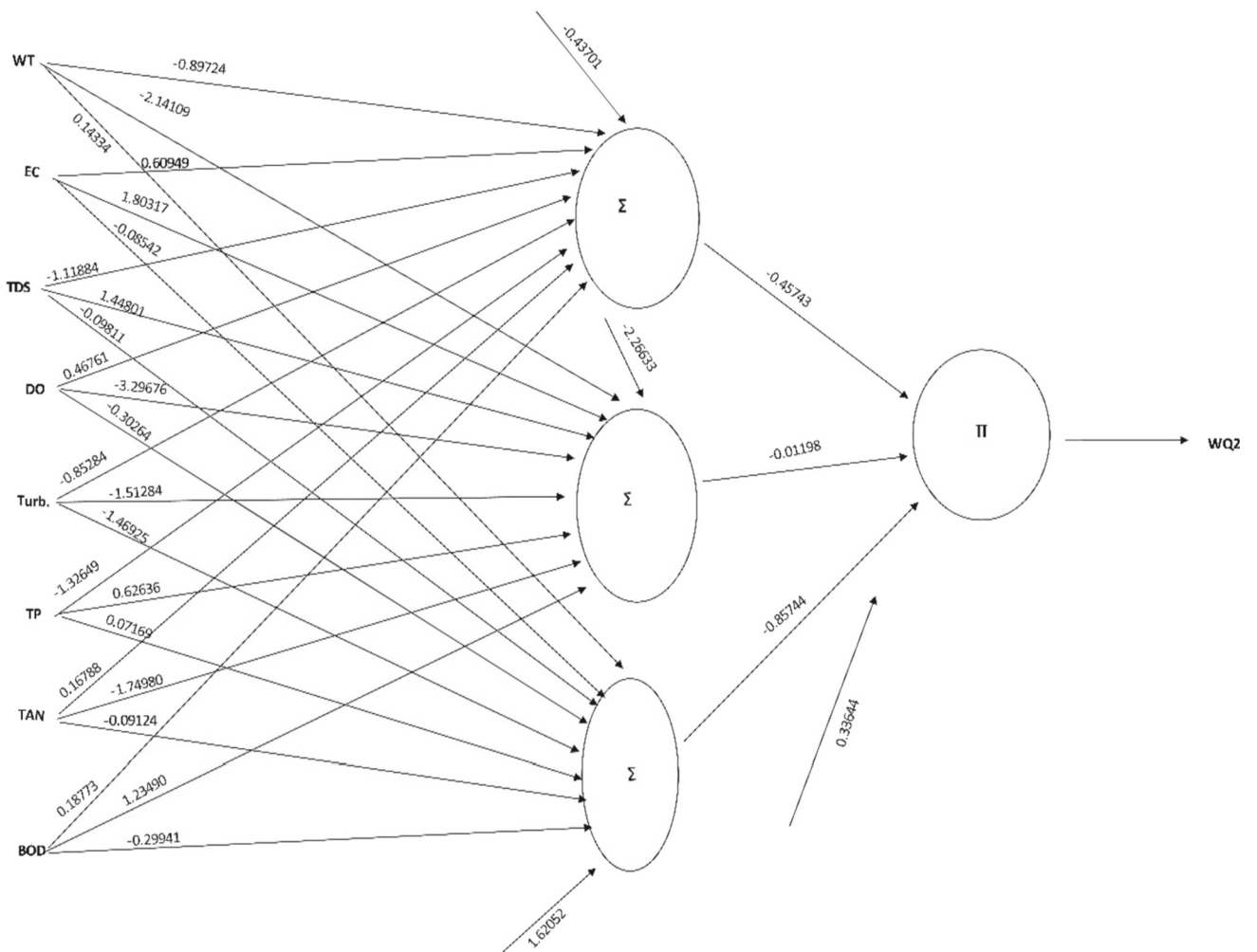
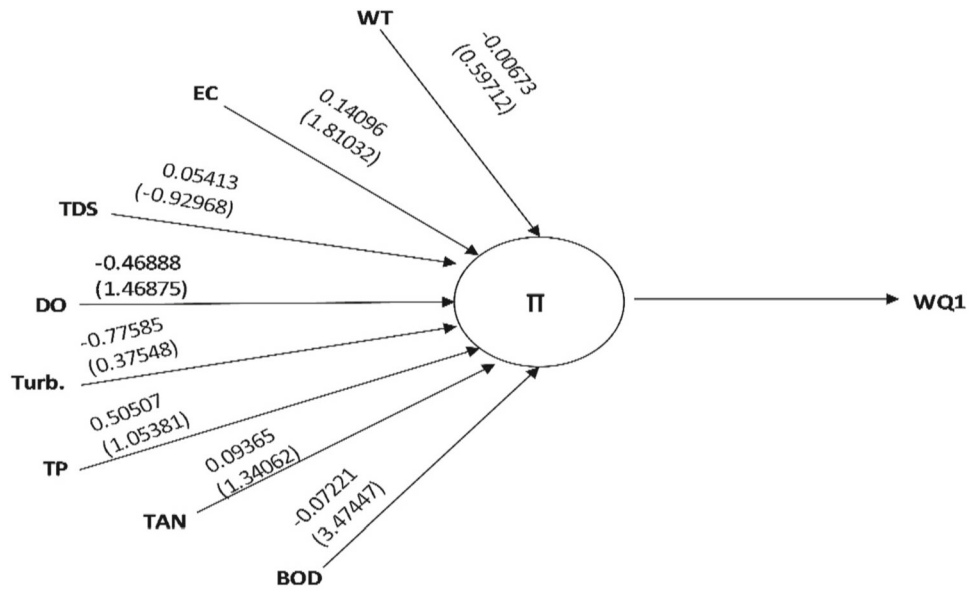


Fig. 10 Optimum multilayer perceptron artificial neural network for WQI₂

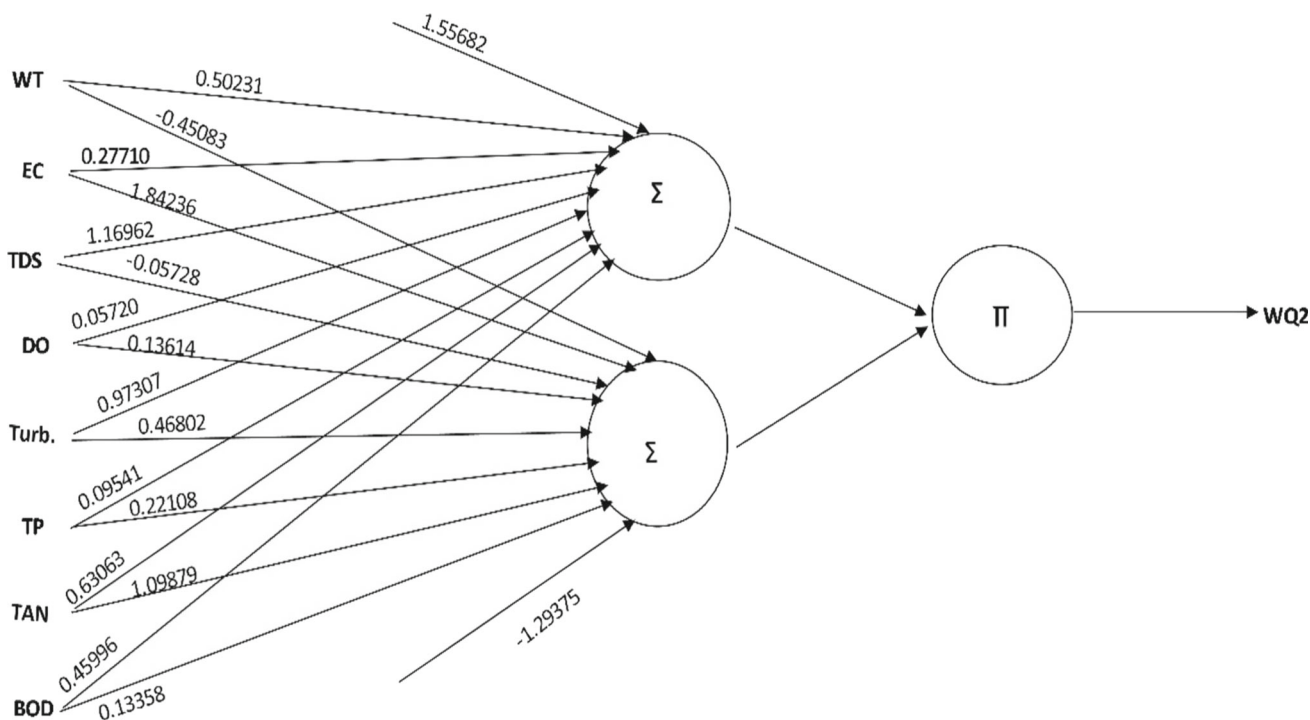
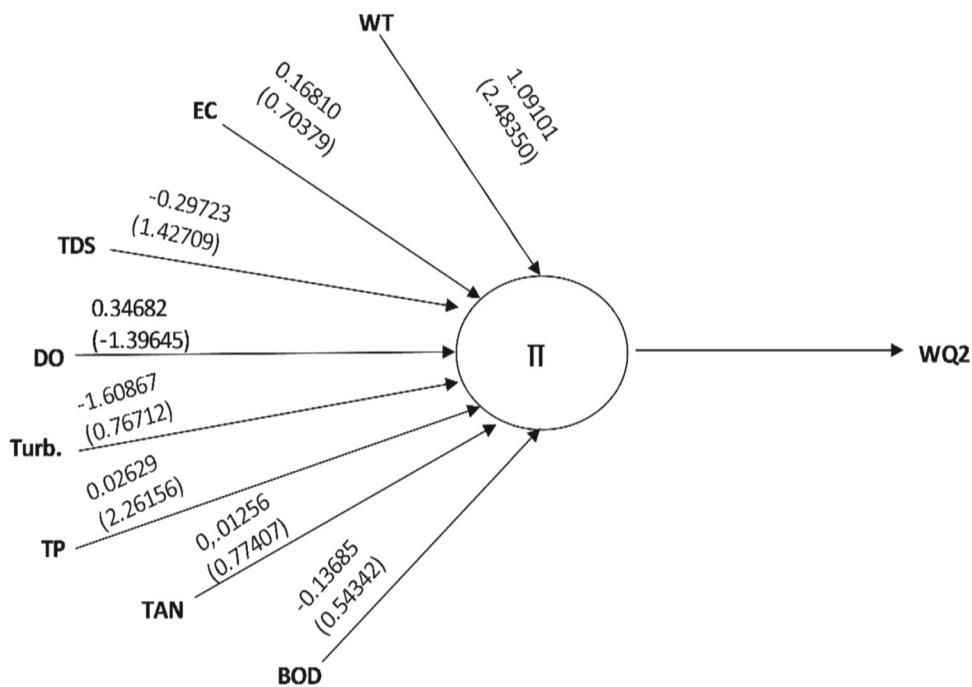


Fig. 11 Optimum pi-sigma artificial neural network for WQI₂

Fig. 12 Optimum single multiplicative neuron model artificial neural network for WQI₂



study is 7.51, and it is accepted as the first class according to the Surface Water Quality Regulation (2015). The normal pH range of surface waters is between 6.5 and 8.5 [60]. In the statistical analysis of the data, it was determined that the pH changes did not differ between the stations. As a result of the study, we can say that the water is basic. The overall EC level obtained in the study was 230.51 $\mu\text{S cm}^{-1}$, and it was included in the first class WQ according to the surface (2015).

It has been reported that the conductivity in river waters is a maximum of 1000 $\mu\text{S cm}^{-1}$ according to WHO standards [61].

The average TDS value obtained overall in the study was reported as 0.1694 (ppm). Since TDS values are below 1000 ppm in eight station sources that are the subject of this study, they are in the freshwater class [62].

The average DO level obtained in the study is 10.07 mg/L, and it is first class WQ according to the Surface WQ Regulation (2015). The amount of dissolved oxygen (DO) necessary for the life of aquatic living things should be ≥ 5 mg/L. Dissolved oxygen is one of the most important parameters of water pollution [63]. In the study, the average value of turbidity (NTU) measured from the stations was determined as 8.84.

The average TP level obtained in the study was determined as 0.061 mg/L, and it is included in the second class WQ according to Surface WQ Regulation (2015). Phosphate can enter the aquatic environment naturally, as well as from artificial fertilizers and industrial wastes. As a result of the excessive presence of phosphate in aquatic ecosystems, algae multiply and can cause odor and taste problems in the water [64]. Phosphorus is the most important element of eutrophication in the aquatic environment [65]. In the study, the average level of TAN obtained in general was determined as 0.413 mg/L. The sum of ammonium and ammonia ($\text{NH}_4^+ + \text{NH}_3$) is defined as total ammonia (TAN). High ammonia concentrations represent one of the most harmful among the different aquatic environmental problems affecting the fitness of aquatic organism in surface water ecosystems as well as oxygen depletion [66]. The average BOD level obtained in the study was determined as 4.61 mg/L. The BOD is the most important measure of organic pollution. BOD is the amount of oxygen required to stabilize decomposable organic materials by bacteria under aerobic conditions [67]. BOD values in receiving environments close to the discharge point of wastewater are more than 10 mg/L [68]. The main cause of pollution in Gelevera Stream was determined as uncontrolled domestic waste and waste from animal slaughterhouses [69, 70].

The WQI is considered one of the main factors for assessing drinking WQ. This index is the most commonly used evaluation of WQ for drinking purposes [39, 44, 45, 71]. Therefore, in calculating the first method WQI to estimate the suitability of the WQI of Türkiye region in Southeastern Black Sea for drinking purposes, the weights of each parameter (w_i) obtained on the basis of the FA/PCA results of eight physicochemical parameters were used for WQI_s (Table 3). Krishan et al., 2023 [72] reported that the value of WQI during pre-monsoon (PRM) and post-monsoon (POM) ranged between 72 and 3683 and 51 to 2451 at 50 sampling groundwater stations, respectively. Furthermore, scientists also classified two categories, “Very poor and poor water,” representing 64% of the PRM and 58% of the POM of the collected samples in India. A further drinking WQ assessment study in southwest China used the WQI and reported that the majority of groundwater samples (96%) had EWQI values lower than the WHO drinking standard, which is suitable for drinking purposes [73]. Similarly, through the WQI, Bomadi water was found to be unsuitable for drinking and

other domestic purposes and needs to be improved [74]. The present study showed that in WQI₁, the average value was 72 and the WQ was found to be “good”, and in WQI₂ the average value was 86, and in 82% “good water” 14% “bad water” and 4% “very bad water” for all stations.

4.2 Proposed Methodology for Evaluation

Water pollution is one of the most important environmental problems facing humanity. The main cause of this problem is largely due to the lack of the necessary conditions for forecasting, early warning and solving the problem in emergency situations. Therefore, the provision of an effective monitoring and early warning system to enable smart decision-making and WQ management will be the most important scientific and technological step to be taken [75].

Although the methods tested in this study have never been tried before in WQ assessment, they are original methods. Therefore, the literature is very limited and there is a lack of WQ assessment. The single multiplicative neuron model is a new automatic forecasting method based on a new input significance test [76]. The results of the single multiplicative neuron (SMN) model, which was applied for the first time in this study for the prediction of WQ, showed performance with RMSE and MAPE values (Table 4; WQI₁ for RMSE: 4.731800 and MAPE: 0.051200, Table 5; WQI₂ for RMSE: 3.819400 and MAPE: 0.039900, accuracy). Similarly, pi-sigma is another machine learning method that has been found to be underutilized in the assessment of water resources quality prediction. Pi-sigma was able to predict WQI with good performance, showing less error than MLP and SMN model. Based on the statistical parameters, the pi-sigma algorithm showed the highest predictive power for WQI prediction (Table 4; WQI₁ for RMSE: 1.976800 and MAPE: 0.023200, Table 5; WQI₂ for RMSE: 2.650500 and MAPE: 0.028300, accuracy).

Multilayer perceptron (MLP) model has been applied and validated by most researchers. MLP method was used to estimate the WQI and to find the WQ classification [77, 78, 79]. In the study using machine learning models to predict the WQI in La Buong River, it was found that the RMSE for MLP had a value of 0.132 [38]. It was found that the proposed MLP-ANN model (RMSE = 0.1984) for predicting WQ parameters such as dissolved oxygen using limiting data sets for Ganga River [80]. The results of applied MLP model to estimate the WQI in this study are given in tables above (Table 4; WQI₁ for RMSE: 4.349000 and MAPE: 0.051500, Table 5; WQI₂ for RMSE: 3.941500 and MAPE: 0.042800, accuracy).

As a result of this study, the comparison of the prediction performance of the model (pi-sigma) that gives the best prediction of WQI with other studies in the literature is given in Table 6.

Table 6 Best prediction of WQI in the literature

Country	Best predictive models	Timing of input data	Number of selected parameters	Scholar	MAPE	RMSE
Turkey	Pi-sigma	1 year	8	Current study	WQI ₁ :0.023,WQI ₂ :0.028	WQI ₁ :1.976,WQI ₂ :2.650
Tayvan	MLP	95 year	5	Latif et al. [78]	-	1.199
Iran	MLP	3 year	10	Singh et al. [81]	0.291	0.085
China	PSO-GA-BPNN	1 month	6	Yan et al. [82]	6.271	0.359
China	FNN	4 year	4	Zhou et al. [83]	10.9051	0.4634
India	GNN	-	5	Roy and Majumder [84]	13.60	10.24
Tayvan	ANNs Tiering	22 year	6	Chou et al. [85]	6.786	3.941
Saudi Arabia	NF	1 year	13	Manzar et al. [86]	13	-
India	Time series	4 year	9	Akkaraboyina and Raju [87]	5.15 & 3.48	7.31 & 5.82
Malaysia	SMWOA-LSSVM	19 year	6	Chia et al. [88]	5.62	3.93

ANN MLP = Multilayer perceptron ANN, PSNN = Pi-sigma neural network, PSO-GA-BPNN = BP neural network combined with particle swarm optimization and genetic algorithm, FNN = Hybrid fuzzy neural network, GNN = Generalized neural network, ANNs tiering model = Artificial neural networks (ANNs), support vector machines, classification and regression trees, and linear regression, NF = Neuro-fuzzy, SMWOA-LSSVM = Hybrid search space enhanced modified whale optimization algorithm—particle swarm optimization and genetic algorithm

From Table 6, it is seen that the pi-sigma model with eight input parameters (WT, EC, TDS, DO, Turb., TP, TAN and BOD) outperformed previously developed models such as independent (e.g., GNN and MLP-ANN) and hybrid (SMWOA-LSSVM, FNN and PSO-GA-BPNN). Therefore, pi-sigma has a very different architecture from other ANNs due to its use of both additive and multiplicative structure [81, 89–92], and more accurate prediction results are compared to other predictive models proposed in the literature.

5 Conclusion

The present study aimed on classifying WQ using machine learning techniques and proposed an intelligent real-time WQ assessment model. Modeling and predicting WQ using AI algorithms is crucial for environmental protection. Artificial intelligence methods can significantly reduce the costs of water supply and sanitation systems and help ensure compliance with drinking and wastewater treatment quality. Therefore, modeling and estimation of WQ to control water pollution are topics that require extensive research.

This article presented surface water sources of the Türkiye region in Southeastern Black Sea data to find out if machine learning could be useful for determining WQ class instead of WQI. The results of the three models were compared, and the pi-sigma model results (RMSE for WQI₁: 1.976800 and MAPE: 0.023200, RMSE for WQI₂: 2.650500 and MAPE: 0.028300, accuracy) showed a significant superiority over the other two models. In this study, the optimal ANN models are proposed for WQ assessment. This approach aims to avoid the disadvantages of traditional assessment methods such as computational time and complexity. According to the findings of the study, the proposed method provides high prediction accuracy. The findings do not mean that systematic modelling and optimization (SMN), multilayer perceptron (MLP) or particle swarm (PS) models will always outperform other conventional methods. However, in this particular case, the proposed PS-ANN model is suitable and highly efficient for WQ prediction. Moreover, it is presented that the models we have applied in the scope of the study are able to reflect well the nonlinear relationships associated with complex watershed processes. In this context, using machine learning and related technologies to develop innovative solutions to control water pollution, improve WQ and ensure watershed ecosystem security can make significant contributions.

The single multiplicative neuron (SMN) model, multilayer perceptron (MLP) and pi-sigma artificial neural network (PS-ANNs) algorithms, which were specifically selected and compared for water pollution control or WQ improvement, were predicted to provide more effective solutions for these issues. As a result of the comparison, the

PS-ANNs algorithm can process more specific types of data or predict certain WQ parameters more accurately. In addition, the difference of the algorithm we have implemented from the others may be that it works faster or more efficiently, which may provide an advantage in the field of application. The aim is to provide information for WQ management, such as the determination of WQ levels, to enable management to respond effectively and quickly in such situations.

Acknowledgements This study was produced from preliminary studies of the doctoral thesis prepared by Hakan ISIK. Hakan ISIK is supported by YÖK 100/2000 project scholarship for PhD studies. The authors are grateful to Prof. Dr. Erol EGRIOGLU and Prof. Dr. Eren BAS for their insightful comments and suggestions.

Author Contributions HI involved in investigation, visualization, original draft preparation, methodology, data curation, writing, and review and editing, TA involved in conceptualization, resources, investigation, formal analysis, visualization, methodology, data curation, validation, original draft preparation, writing, and review and editing.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK). This research was funded by the Technopolis organization Medinora Pharmaceutical Food Cosmetic Consultancy Service Trade Limited Company.

Data Availability The data sets used and/or analyzed during the current study are available from the corresponding author on reasonable request. The preview of present article has been published to a preprint server (<https://doi.org/10.21203/rs.3.rs-2170056/v1>).

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Rahman, M.M.; Howladar, M.F.; Hossain, M.A.; Shahidul Huq Muzemder, A.T.M.; Al Numanbakth, M.A.: Impact assessment of anthropogenic activities on water environment of Tillai River and its surroundings, Barapukuria Thermal Power Plant, Dinajpur, Bangladesh. *Groundw. Sustain. Dev.* **10**, 100310 (2020). <https://doi.org/10.1016/j.gsd.2019.100310>
- Mutlu, T.; Minaz, M.; Baytaşoğlu, H.; Gedik, K.: Microplastic pollution in stream sediments discharging from Türkiye's eastern Black sea basin. *Chemosphere.* 141496 (2024)
- Akkan, T.; Mutlu, T.; Eren, B.: Forecasting sea surface temperature with feed-forward artificial networks in combating the global climate change: the sample of Rize, Türkiye. *Ege J. Fish. Aquat. Sci.* **39**, 311–315 (2022)
- Mutlu, T.: Seasonal variation of trace elements and stable isotope ($\delta^{13}\text{C}$ and $\delta^{15}\text{N}$) values of commercial marine fish from the black sea and human health risk assessment. *Spectrosc. Lett.* **54**, 665–674 (2021)
- Mutlu, T.; Minaz, M.; Baytaşoğlu, H.; Gedik, K.: Monitoring of microplastic pollution in sediments along the Çoruh River Basin, NE Türkiye. *J. Contam. Hydrol.* **263**, 104334 (2024)
- Nong, X.; Shao, D.; Zhong, H.; Liang, J.: Evaluation of water quality in the South-to-North Water Diversion Project of China using the water quality index (WQI) method. *Water Res.* (2020). <https://doi.org/10.1016/j.watres.2020.115781>
- Burgan, H.I.; İçağa, Y.; Bostanoğlu, Y.; Kilit, M.: Akarçay Akarsuyu 2006–2011 Dönemi Su Kalite Eğilimi. *Pamukkale Univ. J. Eng. Sci.* **19**, 127–132 (2013). <https://doi.org/10.5505/pajes.2013.46855>
- Iqbal, J.; Shah, N.S.; Sayed, M.; Imran, M.; Muhammad, N.; Howari, F.M.; Alkhoori, S.A.; Khan, J.A.; Haq Khan, Z.U.; Bhatnagar, A.; Polychronopoulou, K.; Ismail, I.; Haija, M.A.: Synergistic effects of activated carbon and nano-zerovalent copper on the performance of hydroxyapatite-alginate beads for the removal of As^{3+} from aqueous solution. *J. Clean. Prod.* **235**, 875–886 (2019). <https://doi.org/10.1016/j.jclepro.2019.06.316>
- Custodio, M.; Peñaloza, R.; Chanamé, F.; Hinostroza-Martínez, J.L.; De la Cruz, H.: Water quality dynamics of the Cunas River in rural and urban areas in the central region of Peru. *Egypt. J. Aquat. Res.* **47**, 253–259 (2021). <https://doi.org/10.1016/j.ejar.2021.05.006>
- Othman, F.; UddinChowdhury, M.S.; WanJaafar, W.Z.; MohammadFaresh, E.M.; Shirazi, S.M.: Assessing risk and sources of heavy metals in a tropical river basin: a case study of the Selangor River, Malaysia. *Pol. J. Environ. Stud.* (2018). <https://doi.org/10.15244/pjoes/76309>
- Olowe, B.M.; Oluyeye, J.O.; Famurewa, O.: Drinking water quality assessment using water quality index in Ado-Ekiti and Environs, Nigeria. *Challenges Adv Chem Sci.* **1**, 132–147 (2021). <https://doi.org/10.9734/bpi/cacs/v1/1687D>
- Egrioglu, E.; Bas, E.; Karahasan, O.: Winsorized dendritic neuron model artificial neural network and a robust training algorithm with Tukey's biweight loss function based on particle swarm optimization. *Granul. Comput.* (2022). <https://doi.org/10.1007/s41066-022-00345-y>
- Gomolka, Z.; Twarog, B.; Zeslawska, E.; Lewicki, A.; Kwater, T.: Using artificial neural networks to solve the problem represented by BOD and DO indicators. *Water.* **10**, 4 (2018). <https://doi.org/10.3390/w10010004>
- Hmoud Al-Adhaileh, M.; Waselallah Alsaade, F.: Modelling and prediction of water quality by using artificial intelligence. *Sustainability.* **13**, 4259 (2021). <https://doi.org/10.3390/su13084259>
- Zhang, Y.; Gao, X.; Smith, K.; Inial, G.; Liu, S.; Conil, L.B.; Pan, B.: Integrating water quality and operation into prediction of water production in drinking water treatment plants by genetic algorithm enhanced artificial neural network. *Water Res.* **164**, 114888 (2019). <https://doi.org/10.1016/j.watres.2019.114888>
- Isiyaka, H.A.; Mustapha, A.; Juahir, H.; Phil-Eze, P.: Water quality modelling using artificial neural network and multivariate statistical techniques. *Model. Earth Syst. Environ.* **5**, 583–593 (2019). <https://doi.org/10.1007/s40808-018-0551-9>
- Prasad, D.V.V.; Venkataramana, L.Y.; Kumar, P.S.; Prasanna-medha, G.; Harshana, S.; Srividya, S.J.; Harrine, K.; Indraganti, S.: Analysis and prediction of water quality using deep learning and auto deep learning techniques. *Sci. Total. Environ.* **821**, 153311 (2022). <https://doi.org/10.1016/j.scitotenv.2022.153311>



18. Ibrahim, A.; Ismail, A.; Juahir, H.; Iliyasu, A.B.; Wailare, B.T.; Mukhtar, M.; Aminu, H.: Water quality modelling using principal component analysis and artificial neural network. *Mar. Pollut. Bull.* **187**, 114493 (2023)
19. Jayaraman, P.; Nagarajan, K.K.; Partheeban, P.; Krishnamurthy, V.: Critical review on water quality analysis using IoT and machine learning models. *Int. J. Inf. Manag. Data Insights.* **4**, 100210 (2024). <https://doi.org/10.1016/j.jjime.2023.100210>
20. Chen, Y.; Song, L.; Liu, Y.; Yang, L.; Li, D.: A review of the artificial neural network models for water quality prediction. *Appl. Sci.* **10**, 5776 (2020). <https://doi.org/10.3390/app10175776>
21. Sheikh Khozani, Z.; Iranmehr, M.; Wan Mohtar, W.H.M.: Improving water quality index prediction for water resources management plans in malaysia: application of machine learning techniques. *Geocarto Int.* (2022). <https://doi.org/10.1080/10106049.2022.2032388>
22. Nayak, J.G.; Patil, L.G.; Patki, V.K.: Artificial neural network based water quality index (WQI) for river Godavari (India). *Mater. Today Proc.* (2021). <https://doi.org/10.1016/j.matpr.2021.03.100>
23. Song, C.; Yao, L.; Hua, C.; Ni, Q.: A water quality prediction model based on variational mode decomposition and the least squares support vector machine optimized by the sparrow search algorithm (VMD-SSA-LSSVM) of the Yangtze River, China. *Environ. Monit. Assess.* **193**, 363 (2021). <https://doi.org/10.1007/s10661-021-09127-6>
24. Othman, F.; Alaaeldin, M.E.; Seyam, M.; Ahmed, A.N.; Teo, F.Y.; Ming Fai, C.; Afan, H.A.; Sherif, M.; Sefelnasr, A.; El-Shafie, A.: Efficient river water quality index prediction considering minimal number of inputs variables. *Eng. Appl. Comput. Fluid Mech.* **14**, 751–763 (2020). <https://doi.org/10.1080/19942060.2020.1760942>
25. Dilmi, S.; Ladjal, M.: A novel approach for water quality classification based on the integration of deep learning and feature extraction techniques. *Chemom. Intell. Lab. Syst.* **214**, 104329 (2021). <https://doi.org/10.1016/j.chemolab.2021.104329>
26. El Bilali, A.; Taleb, A.; Brouziyne, Y.: Groundwater quality forecasting using machine learning algorithms for irrigation purposes. *Agric. Water Manag.* **245**, 106625 (2021). <https://doi.org/10.1016/j.agwat.2020.106625>
27. Nayan, A.-A.; Kibria, M.G.; Rahman, Md.O.; Saha, J.: River water quality analysis and prediction using GBM. In: 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT), pp. 219–224 (2020)
28. Bedi, S.; Samal, A.; Ray, C.; Snow, D.: Comparative evaluation of machine learning models for groundwater quality assessment. *Environ. Monit. Assess.* **192**, 776 (2020). <https://doi.org/10.1007/s10661-020-08695-3>
29. Ahmed, M.; Mumtaz, R.; Hassan Zaidi, S.M.: Analysis of water quality indices and machine learning techniques for rating water pollution: a case study of Rawal Dam, Pakistan. *Water Suppl.* **21**, 3225–3250 (2021). <https://doi.org/10.2166/ws.2021.082>
30. Radhakrishnan, N.; Pillai, A.S.: Comparison of water quality classification models using machine learning. In: 2020 5th International Conference on Communication and Electronics Systems (ICCES), pp. 1183–1188 (2020)
31. Asadollah, S.B.H.S.; Sharafati, A.; Motta, D.; Yaseen, Z.M.: River water quality index prediction and uncertainty analysis: a comparative study of machine learning models. *J. Environ. Chem. Eng.* **9**, 104599 (2021). <https://doi.org/10.1016/j.jece.2020.104599>
32. Gazzaz, N.M.; Yusoff, M.K.; Aris, A.Z.; Juahir, H.; Ramli, M.F.: Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. *Mar. Pollut. Bull.* **64**, 2409–2420 (2012). <https://doi.org/10.1016/j.marpolbul.2012.08.005>
33. Hameed, M.; Sharqi, S.S.; Yaseen, Z.M.; Afan, H.A.; Hussain, A.; Elshafie, A.: Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia. *Neural Comput. Appl.* **28**, 893–905 (2017). <https://doi.org/10.1007/s00521-016-2404-7>
34. Bowes, B.D.; Wang, C.; Ercan, M.B.; Culver, T.B.; Beling, P.A.; Goodall, J.L.: Reinforcement learning-based real-time control of coastal urban stormwater systems to mitigate flooding and improve water quality. *Environ. Sci. Water Res. Technol.* **8**, 2065–2086 (2022). <https://doi.org/10.1039/D1EW00582K>
35. Castañeda-Miranda, A.; Castaño-Meneses, V.M.: Smart frost measurement for anti-disaster intelligent control in greenhouses via embedding IoT and hybrid AI methods. *Measurement* **164**, 108043 (2020). <https://doi.org/10.1016/j.measurement.2020.108043>
36. Sharaf El Din, E.; Zhang, Y.; Suliman, A.: Mapping concentrations of surface water quality parameters using a novel remote sensing and artificial intelligence framework. *Int. J. Remote Sens.* **38**, 1023–1042 (2017). <https://doi.org/10.1080/01431161.2016.1275056>
37. Gupta, T.K.; Raza, K.: Optimizing deep feedforward neural network architecture: a tabu search based approach. *Neural. Process. Lett.* **51**, 2855–2870 (2020). <https://doi.org/10.1007/s11063-020-10234-7>
38. Khoi, D.N.; Quan, N.T.; Linh, D.Q.; Nhi, P.T.T.; Thuy, N.T.D.: Using machine learning models for predicting the water quality index in the La Buong River. *Vietnam. Water.* **14**, 1552 (2022). <https://doi.org/10.3390/w14101552>
39. Gupta, A.N.; Kumar, D.; Singh, A.: Evaluation of water quality based on a machine learning algorithm and water quality index for mid gangetic region (South Bihar plain), India. *J. Geol. Soc. India.* **97**, 1063–1072 (2021). <https://doi.org/10.1007/s12594-021-1821-0>
40. Acharya, S.; Sharma, S.K.; Khandegar, V.: Assessment of ground-water quality by water quality indices for irrigation and drinking in South West Delhi, India. *Data Brief.* **18**, 2019–2028 (2018). <https://doi.org/10.1016/j.dib.2018.04.120>
41. Deshmukh, K.K.; Aher, S.P.: Assessment of the impact of municipal solid waste on groundwater quality near the Sangamner city using GIS approach. *Water Resour. Manag.* **30**, 2425–2443 (2016). <https://doi.org/10.1007/s11269-016-1299-5>
42. Wang, J.; Liu, G.; Liu, H.; Lam, P.K.S.: Multivariate statistical evaluation of dissolved trace elements and a water quality assessment in the middle reaches of Huaihe River, Anhui, China. *Sci. Total Environ.* **583**, 421–431 (2017). <https://doi.org/10.1016/j.scitotenv.2017.01.088>
43. Meng, Q.; Zhang, J.; Zhang, Z.; Wu, T.: Geochemistry of dissolved trace elements and heavy metals in the Dan River Drainage (China): distribution, sources, and water quality assessment. *Environ. Sci. Pollut. Res.* **23**, 8091–8103 (2016). <https://doi.org/10.1007/s11356-016-6074-x>
44. Şener, Ş.; Şener, E.; Davraz, A.: Evaluation of water quality using water quality index (WQI) method and GIS in Aksu River (SW-Turkey). *Sci. Total Environ.* **584–585**, 131–144 (2017). <https://doi.org/10.1016/j.scitotenv.2017.01.102>
45. Ismael, M.; Mokhtar, A.; Farooq, M.; Lü, X.: Assessing drinking water quality based on physical, chemical and microbial parameters in the Red Sea State, Sudan using a combination of water quality index and artificial neural network model. *Groundw. Sustain. Dev.* **14**, 100612 (2021). <https://doi.org/10.1016/j.gsd.2021.100612>
46. Deng, L.; Shahab, A.; Xiao, H.; Li, J.; Rad, S.; Jiang, J.; Guoyu, J.P.; Huang, H.; Li, X.; Ahmad, B.; Siddique, J.: Spatial and temporal variation of dissolved heavy metals in the Lijiang River, China: implication of rainstorm on drinking water quality. *Environ. Sci. Pollut. Res.* **28**, 68475–68486 (2021). <https://doi.org/10.1007/s11356-021-15383-3>
47. Xiao, J.; Wang, L.; Deng, L.; Jin, Z.: Characteristics, sources, water quality and health risk assessment of trace elements in river water and well water in the Chinese Loess Plateau. *Sci. Total Environ.*



- 650, 2004–2012 (2019). <https://doi.org/10.1016/j.scitotenv.2018.09.322>
48. Sihag, P.; Singh, B.; Sepah Vand, A.; Mehdipour, V.: Modeling the infiltration process with soft computing techniques. *ISH J. Hydraul. Eng.* **26**, 138–152 (2020). <https://doi.org/10.1080/09715010.2018.1464408>
 49. Sharma, N.; Zakaullah, Md.; Tiwari, H.; Kumar, D.: Runoff and sediment yield modeling using ANN and support vector machines: a case study from Nepal watershed. *Model. Earth Syst. Environ.* **1**, 23 (2015). <https://doi.org/10.1007/s40808-015-0027-0>
 50. Mohammadpour, R.; Shaharuddin, S.; Zakaria, N.A.; Ghani, A.A.; Vakili, M.; Chan, N.W.: Prediction of water quality index in free surface constructed wetlands. *Environ. Earth Sci.* **75**, 1–12 (2016)
 51. Yilma, M.; Kiflie, Z.; Windsperger, A.; Gessese, N.: Application of artificial neural network in water quality index prediction: a case study in Little Akaki River, Addis Ababa, Ethiopia. *Model. Earth Syst. Environ.* **4**, 175–187 (2018). <https://doi.org/10.1007/s40808-018-0437-x>
 52. Ucu Ozel, H.; Gemici, B.T.; Gemici, E.; Ozel, H.B.; Cetin, M.; Sevik, H.: Application of artificial neural networks to predict the heavy metal contamination in the Bartin River. *Environ. Sci. Pollut. Res.* **27**, 42495–42512 (2020). <https://doi.org/10.1007/s11356-020-10156-w>
 53. Egbueri, J.C.: Predicting and analysing the quality of water resources for industrial purposes using integrated data-intelligent algorithms. *Groundw. Sustain. Dev.* **18**, 100794 (2022). <https://doi.org/10.1016/j.gsd.2022.100794>
 54. Abba, S.I.; Pham, Q.B.; Saini, G.; Linh, N.T.T.; Ahmed, A.N.; Mohajane, M.; Khaledian, M.; Abdulkadir, R.A.; Bach, Q.-V.: Implementation of data intelligence models coupled with ensemble machine learning for prediction of water quality index. *Environ. Sci. Pollut. Res.* **27**, 41524–41539 (2020). <https://doi.org/10.1007/s11356-020-09689-x>
 55. Yadav, R.N.; Kalra, P.K.; John, J.: Time series prediction with single multiplicative neuron model. *Appl. Soft Comput.* **7**, 1157–1163 (2007). <https://doi.org/10.1016/j.asoc.2006.01.003>
 56. Bas, E.; Eğrioglu, E.: A new recurrent pi-sigma artificial neural network inspired by exponential smoothing feedback mechanism. *J. Forecast.* (2022)
 57. Shin, Y.; Ghosh, J.: The pi-sigma network: an efficient higher-order neural network for pattern classification and function approximation. In: *IJCNN-91-Seattle International Joint Conference on Neural Networks*. **1**, 13–18 (1991)
 58. Bas, E.; Eğrioglu, E.; Karahasan, O.: A Pi-Sigma artificial neural network based on sine cosine optimization algorithm. *Granul. Comput.* **7**, 813–820 (2022). <https://doi.org/10.1007/s41066-021-00297-9>
 59. Ewaid, S.H.; Abed, S.A.; Kadhum, S.A.: Predicting the Tigris River water quality within Baghdad, Iraq by using water quality index and regression analysis. *Environ Technol Innov.* **11**, 390–398 (2018). <https://doi.org/10.1016/j.eti.2018.06.013>
 60. Zhao, M.M.; Wang, S.; Chen, Y.; Wu, J.; Xue, L.; Fan, T.T.: Pollution status of the Yellow River tributaries in middle and lower reaches. *Sci. Total. Environ.* **722**, 137861 (2020). <https://doi.org/10.1016/j.scitotenv.2020.137861>
 61. Leong, S.S.; Ismail, J.; Denil, N.A.; Sarbini, S.R.; Wasli, W.; Debbie, A.: Microbiological and physicochemical water quality assessments of river water in an industrial region of the northwest Coast of Borneo. *Water*. **10**, 1648 (2018). <https://doi.org/10.3390/w10111648>
 62. Catroll, D.: Rain water as a chemical agent of geological process: a view. *USGS Water Supply*. **1533**, 18–20 (1962)
 63. Doğanay, E.: The evaluation of analysis methods to be used for monitoring of the water in our country within the scope of EU water framework directive from the point of physicochemical and chemical parameters. Master's Thesis, Republic of Türkiye Ministry of Forestry and Water Affairs, Ankara, Türkiye (2014)
 64. Güler, Ç.; Çobanoğlu, Z.: Water quality. Republic of Türkiye Ministry of Ministry of Health, Ankara, Türkiye (1997)
 65. Mackay, D.: Eutrophication of freshwaters, principles, problems and restoration, David Harper, Chapman and Hall, London, 1992. viii + 327pp. Price: £35.00. ISBN 0412 32970 0. *Aquatic Conservation: Marine and Freshwater Ecosystems*. **2**, 364–364 (1992). <https://doi.org/10.1002/aqc.3270020408>
 66. Egniew, N.; Renukdas, N.; Ramena, Y.; Yadav, A.K.; Kelly, A.M.; Lochmann, R.T.; Sinha, A.K.: Physiological insights into large-mouth bass (*Micropterus salmoides*) survival during long-term exposure to high environmental ammonia. *Aquat. Toxicol.* **207**, 72–82 (2019). <https://doi.org/10.1016/j.aquatox.2018.11.027>
 67. Khdary, N.H.; Gasim, A.E.; Muriyani, M.E.; Alshehrie, A.A.: Modeling distribution of selective ions in urban and rural areas using geographical information system. *J. Water Resour. Prot.* **07**, 516 (2015). <https://doi.org/10.4236/jwarp.2015.76041>
 68. Pulatsü, S.; Topçu, A.; Atay, D.: Water pollution and Control. Faculty of Agriculture, Ankara University (2014)
 69. Akkan, T.; Çolaker, F.: Determining the level of bacteriological pollution level in Gelevera Creek, Giresun. *J. Anat. Environ. Anim. Sci.* **5**, 691–695 (2020). <https://doi.org/10.35229/jaes.818132>
 70. Işık, H.; Akkan, T.: The global problem of the antibiotic and heavy metal resistance in aquatic resources, an examination of Gelevera Creek (Giresun), Turkey. *J. Anat. Environ. Anim. Sci.* **6**(3), 382–389 (2021). <https://doi.org/10.35229/jaes.960110>
 71. Akkan, T.; Yazicioglu, O.; Yazici, R.; Yilmaz, M.: Assessment of irrigation water quality of Turkey using multivariate statistical techniques and water quality index: Siddıklı Dam Lake. *Desalin. Water Treat.* **115**, 261 (2018)
 72. Krishan, G.; Kumar, M.; Rao, M.S.; Garg, R.; Yadav, B.K.; Kansal, M.L.; Singh, S.; Bradley, A.; Muste, M.; Sharma, L.M.: Integrated approach for the investigation of groundwater quality through hydrochemistry and water quality index (WQI). *Urban Climate*. **47**, 101383 (2023). <https://doi.org/10.1016/j.uclim.2022.101383>
 73. Zhang, Y.; Dai, Y.; Wang, Y.; Huang, X.; Xiao, Y.; Pei, Q.: Hydrochemistry, quality and potential health risk appraisal of nitrate enriched groundwater in the Nanchong area, southwestern China. *Sci. Total. Environ.* **784**, 147186 (2021). <https://doi.org/10.1016/j.scitotenv.2021.147186>
 74. Iwegbue, C.M.A.; Faran, T.K.; Iniaighe, P.O.; Ikpefan, J.O.; Tesi, G.O.; Nwajei, G.E.; Martincigh, B.S.: Water quality of Bomadi Creek in the Niger Delta of Nigeria: assessment of some physico-chemical properties, metal concentrations, and water quality index. *Appl. Water Sci.* **13**, 36 (2023). <https://doi.org/10.1007/s13201-022-01804-2>
 75. Wu, J.; Wang, Z.: A hybrid model for water quality prediction based on an artificial neural network, wavelet transform, and long short-term memory. *Water*. **14**, 610 (2022). <https://doi.org/10.3390/w14040610>
 76. Eğrioglu, E.; Bas, E.: A new automatic forecasting method based on a new input significance test of a single multiplicative neuron model artificial neural network. *Netw. Comput. Neural Syst.* **33**, 1–16 (2022). <https://doi.org/10.1080/0954898X.2022.2042609>
 77. Juna, A.; Umer, M.; Sadiq, S.; Karamti, H.; Eshawi, A.A.; Mohamed, A.; Ashraf, I.: Water quality prediction using KNN imputer and multilayer perceptron. *Water*. **14**, 2592 (2022). <https://doi.org/10.3390/w14172592>
 78. Latif, S.D.; Birima, A.H.; Ahmed, A.N.; Hatem, D.M.; Al-Ansari, N.; Fai, C.M.; El-Shafie, A.: Development of prediction model for phosphate in reservoir water system based machine learning algorithms. *Ain Shams Eng. J.* **13**, 101523 (2022). <https://doi.org/10.1016/j.asej.2021.06.009>

79. Palabiyik, S.; Akkan, T.; Ahmed, U.: Evaluation of water quality based on artificial intelligence: performance of multilayer perceptron neural networks and multiple linear regression versus water quality indexes. *Environ Dev Sustain.* (2024). <https://doi.org/10.1007/s10668-024-05075-6>
80. Bhardwaj, R.; Singh, R.K.: Water quality modeling of the river ganga in the Northern Region of India using the artificial neural network technique. *J. Water. Manag. Model.* (2022)
81. Singh, B.; Sihag, P.; Singh, V.P.; Sepahvand, A.; Singh, K.: Soft computing technique-based prediction of water quality index. *Water Supply.* **21**, 4015–4029 (2021). <https://doi.org/10.2166/ws.2021.157>
82. Yan, J.; Xu, Z.; Yu, Y.; Xu, H.; Gao, K.: Application of a hybrid optimized BP network model to estimate water quality parameters of Beihai Lake in Beijing. *Appl. Sci.* **9**, 1863 (2019). <https://doi.org/10.3390/app9091863>
83. Zhou, C.; Zhang, C.; Tian, D.; Wang, K.; Huang, M.; Liu, Y.: A software sensor model based on hybrid fuzzy neural network for rapid estimation water quality in Guangzhou section of Pearl River, China. *J. Environ. Sci. Health Part A* **53**, 91–98 (2018). <https://doi.org/10.1080/10934529.2017.1369815>
84. Roy, R.; Majumder, M.: A quick prediction of hardness from water quality parameters by artificial neural network. *Int. J. Environ. Sustain. Dev.* **17**, 247–257 (2018). <https://doi.org/10.1504/IJESD.2018.094037>
85. Chou, J.-S.; Ho, C.-C.; Hoang, H.-S.: Determining quality of water in reservoir using machine learning. *Ecol. Inform.* **44**, 57–75 (2018). <https://doi.org/10.1016/j.ecoinf.2018.01.005>
86. Manzar, M.S.; Benaafi, M.; Costache, R.; Alagha, O.; Mu'azu, N.D.; Zubair, M.; Abdullahi, J.; Abba, S.I.: New generation neuro-computing learning coupled with a hybrid neuro-fuzzy model for quantifying water quality index variable: a case study from Saudi Arabia *Ecol. Inf.* **70**, 101696 (2022). <https://doi.org/10.1016/j.ecoinf.2022.101696>
87. Akkaraboyina, M.K.; Raju, B.: A Comparative study of water quality indices of River Godavari. *Int. J. Eng. Res. Dev.* **2**, 29–34 (2012)
88. Chia, S.L.; Chia, M.Y.; Koo, C.H.; Huang, Y.F.: Integration of advanced optimization algorithms into least-square support vector machine (LSSVM) for water quality index prediction. *Water Supply.* **22**, 1951–1963 (2021). <https://doi.org/10.2166/ws.2021.303>
89. Yilmaz, O.; Bas, E.; Egrioglu, E.: The training of pi-sigma artificial neural networks with differential evolution algorithm for forecasting. *Comput. Econ.* **59**, 1699–1711 (2022). <https://doi.org/10.1007/s10614-020-10086-2>
90. Egrioglu, E.; Bas, E.: Modified Pi Sigma artificial neural networks for forecasting. *Granul. Comput.* **8**(1), 131–135 (2023). <https://doi.org/10.1007/s41066-022-00320-7>
91. Bas, E.; Egrioglu, E.; Kolemen, E.: A novel intuitionistic fuzzy time series method based on bootstrapped combined pi-sigma artificial neural network. *Eng. Appl. Artif. Intell.* **114**, 105030 (2022). <https://doi.org/10.1016/j.engappai.2022.105030>
92. Bas, E.; Egrioglu, E.; Tunc, T.: Multivariate Picture Fuzzy Time Series: New Definitions and a New Forecasting Method Based on Pi-Sigma Artificial Neural Network. *Comput. Econ.* **61**, 139–164 (2023). <https://doi.org/10.1007/s10614-021-10202-w>

