Turk. J. Fish.& Aquat. Sci. 26(12), TRJFAS30206
https://doi.org/10.4194/TRJFAS30206

RESEARCH PAPER

Turkish Journal of FISHERIES and AQUATIC SCIENCES

The Prediction of Water Quality Application in a Dam Reservoir

Aslıhan Katip[1,*], Asifa Anwar[2]

[1]Bursa Uludag University, Faculty of Engineering, Environmental Engineering Department, Bursa/Türkiye.
[2]Division of Sustainable Development, College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha/ Qatar.

How to Cite

Katip, A., Anwar, A. (2026). The Prediction of Water Quality Application in a Dam Reservoir. *Turkish Journal of Fisheries and Aquatic Sciences*, *26(12), TRJFAS30206.* https://doi.org/10.4194/TRJFAS30206

Article History
Received 28 January 2026
Accepted 06 July 2026
First Online 09 July 2026

Corresponding Author
E-mail: aballi@uludag.edu.tr

Keywords
Artificial neural networks
Drinking water dam
Training algorithms
Water quality modeling

Abstract

Artificial neural networks are important management tools that can be used to manage water pollution. In this study, artificial neural networks (ANNs) were used to predict the water quality of Doğancı dam. The water quality data of the Doğancı dam, from 1990 to 2020, comprised water temperature, alkalinity, turbidity, pH, and concentrations of suspended solids, dissolved oxygen, arsenic, manganese, and iron. Feed-forward neural networks with one input, one hidden, and an output layer were chosen, and the number of nodes was examined. The results indicated that the model incorporating water temperature, alkalinity, pH, and concentration of dissolved oxygen as inputs and trace elements (arsenic, manganese, and iron) as outputs, performed better, having a sufficiently higher correlation-R (between 0.994 and 0.999 for the entire dataset), lower root mean square error-RMSE (ranging from 0.01 to 0.03), and higher Nash-Sutcliffe efficiency-NSE (up to 0.999). Levenberg-Marquardt and resilient backpropagation training algorithms were also compared, with Levenberg-Marquardt's outperforming in the correlation and error metrics for the models tested. The findings provide a useful framework for predicting trace metals in drinking-water reservoirs using long-term monitoring data. This approach supports water-quality assessment, early-warning monitoring, and public-health-oriented management.

Introduction

Dams are crucial in providing the human population with safe and clean drinking water. However, various elements, including industrial and agricultural runoff, atmospheric deposition, and erosion, may interact and affect the water quality in dams. Therefore, monitoring and maintaining the water quality of dams becomes vital to ensure the security and sustainability of water supplies (Pule et al., 2017).

To predict water quality in water resources and reduce laboratory analysis costs, machine learning applications such as artificial neural networks (ANN), random forest (RF), support vector machines (SVM), hybrid models, k-nearest neighbors, extreme gradient boosting (XGBoost), and Long Short-Term Memory (LSTM) are used (Sham, et al., 2025; Chukwuemeka, et al., 2026), and many studies have been conducted on this subject in recent years.

In different studies, to determine the most suitable approach for classifying groundwater quality, three commonly used machine learning algorithms, including decision tree (DT), random forest (RF), and XGBoost, were compared. The XGBoost model has been found to be more effective in identifying key parameters controlling pollution by capturing nonlinear interactions of data in groundwater (Vu, et. al., 2026). In predicting groundwater quality, hybrid models Hybrid Light Gradient Boosting Machine with Particle Swarm Optimization (LightGBM-PSO) and Genetic Algorithm

with Random Forest (GA-RF) demonstrated the highest accuracy, achieving 93%, 95%, and 99% accuracy rates for Classes 0, 1, and 2, respectively. The GA-RF model demonstrated strong prediction performance with an R^2 of 0.99, an RMSE of 0.01, and a MAE of 0.02 (Nafouanti, et al., 2026). Naive Bayes and Deep Neural Network (DNN) performance in determining the parameters responsible for arsenic dissolution in groundwater aquifers was found to be highly accurate (Kumar and Pati, 2023). The ability to explain processes modeled with machine learning (ML) models has begun to increase in recent years. Parameters affecting arsenic mobility in groundwater and streams were determined using novel boosting algorithms, and a relationship was established between low pH, low electrical conductivity, low total dissolved solids, and low arsenic water. These relationships provided new information to improve the explanatory power of the process modeled by Machine Learning (ML) models (Ibrahim, et. al., 2022).

In dams, chlorophyll-a values were estimated using artificial neural networks with 4 years of data on phosphate, nitrate, dissolved oxygen, suspended solids, alkalinity, pH, water temperature, electrical conductivity, and Sechi depth (Karul et al., 2000); electrical conductivity and dissolved oxygen values in feed and discharge waters were estimated (Elhatip and Kömür, 2008); and dissolved oxygen and suspended solids concentrations were predicted using pH, conductivity, dissolved oxygen, water temperature, and stored water volume as inputs. The coefficient of determination values for the training set ranged from 0.995 to 0.998, and for the test set from 0.994 to 0.996 (Vicente et al., 2012). In the Feitsui Reservoir in Taiwan, water temperature, pH, conductivity, turbidity, suspended solids, and ammonium nitrogen concentration were selected as inputs, and total hardness and total alkalinity were selected as outputs. The performance of the models was evaluated using the mean absolute error, root mean square error, and correlation coefficient, which were measured and calculated through model simulation (Chen and Liu, 2014). In Köylü Lake, monthly average values of temperature, pH, electrical conductivity, and dissolved oxygen concentration were used as inputs, while ammonium nitrogen, total nitrogen, phosphorus, chlorophyll-a, and coliform concentrations were used as outputs to draw conclusions about eutrophication. The low error (0.046) and high correlation value (1) between the measured and estimated parameters revealed that the lake is eutrophic (Özçelik, 2015). Water quality in Taiwan's dam reservoirs was modeled over 10 years using artificial neural networks, support vector machines, regression trees, and linear regression. The study showed that the artificial neural network model was more accurate than other hybrid meta-heuristic regression models (Chou et al., 2018). Artificial neural network modeling is one of the popular and reliable methods to model water quality (Park et al., 2020). Based on a range of input variables, including weather, water flow rates, and land usage, ANNs can be utilized in dams to forecast water quality parameters and identify probable causes of pollution (Zhu et al., 2022).

This study aimed to use an artificial neural network (specifically a feed-forward neural network) to identify a relation between certain water quality parameters (water temperature, alkalinity, turbidity, pH, suspended, dissolved oxygen concentrations, arsenic, manganese, and iron) of the Doğancı dam. Also, to determine which of the listed water quality parameters can affect the water quality in Doğancı dam, based on the performance evaluation parameters. Nevertheless, it should also be noted that ANN models are primarily predictive tools. This states that the identified relationships should be interpreted as exploratory, model-based associations rather than direct cause-and-effect relationships (Schneider and Edelsbrunner, 2013).

Trace element predictions using ANNs have mostly been made in shorter time periods (Li, et. al., 2020; Lu, et. al., 2019; Ucun, et. al., 2020) for groundwater, rivers, and natural lakes (Al-Yaari, et. al, 2022; Martuza, et. al., 2024; Lu, et. al., 2019; Mohammadi, et. al., 2022). In contrast, the present study applies a conventional feed-forward ANN framework to a long-term (1990-2020) drinking-water reservoir dataset for predicting arsenic, manganese, and iron concentrations. These trace metals are important because arsenic is carcinogenic, while iron and manganese can create treatment challenges and potential health-related concerns in drinking-water systems (WHO, 2022). Therefore, the contribution of this work lies primarily in the application of ANN modeling to a long-term drinking-water monitoring dataset and in the evaluation of trace metal prediction within a drinking-water reservoir context. By focusing on trace metals that are relevant to both water treatment and public health, this study provides additional evidence regarding the applicability of ANN-based approaches for supporting reservoir water-quality assessment and management. The value of the current study is therefore mainly application-oriented and data-driven, with relevance to drinking-water reservoirs where similar long-term monitoring datasets are available.

Materials and Methods

Study Zone

The current study was conducted to predict the water quality of Doğancı dam using artificial neural networks. Doğancı Dam is located in Bursa, Turkey. Bursa, situated southeast of the sea of Marmara, is bounded by Bilecik and Adapazarı to the east, Kocaeli, Yalova, Istanbul, and the Sea of Marmara to the north, Kutahya to the south, and Balıkesir to the west. It is 155 meters above sea level. Bursa is confronted by the mild temperature of the Marmara Sea to the north and the harsh climate of Uludağ to the south. The city's hottest months are August and September, while the coldest

are February and March. Based on the last 52 years of monitoring, the average annual precipitation is 706 mm, and the relative humidity is approximately 69% (BUSKİ, 2023).

Doğancı, Nilufer, and the Çınarcık dams are all located in Bursa. Doğancı Dam, located on the Nilüfer Stream, is one of the most important dams that provide drinking water to the people of Bursa (Figure 1). The Doğancı dam is 164 feet in height and 140 feet in length. The surface area of Doğancı dam's reservoir is 1.5 square kilometers (0.58 square miles), and it could hold 15.5 million cubic meters of water (12,570 acre-feet) (BUSKI, 2023). The General Directorate of State Hydraulic Works owns and operates the dam (Bursa Metropolitan Municipality, 2023).

It was observed that the Doğancı dam's water quality declined, especially due to the presence of some trace metals (Fural et al., 2022). In addition, high polymeric aromatic hydrocarbon (PAH) concentrations were determined in the investigations made in the Nilufer stream over which the Doğancı dam was built (Karaca and Tasdemir, 2014).

Data Collection

The data for water quality parameters of Doğancı dam were obtained from the Water and Sewage Authority of Bursa (BUSKI). The data set was created by the authors by transferring the data on paper to electronic media. This data included parameters such as water temperature, alkalinity, turbidity, pH, concentrations of suspended, dissolved oxygen, arsenic, manganese, and iron. The data had monthly average values for the period between 1990 to 2020. Therefore, the ANN modeling was done for the defined period. Some statistical characteristics of water quality data of the Doğancı dam are given in Table 1.

Artificial Neural Networks

ANNs are information processing systems developed by drawing inspiration from the working principles of the biological nervous system in the human brain. ANNs possess learning, generalization, and decision-making capabilities, which are used in

Figure 1. Doğancı dam (acquired from Google Maps on 3rd February 2023 and edited)

modeling nonlinear relationships (Chou et al., 2018). These networks consist of some interconnected nodes (likewise to brain nerves), termed neurons. These neurons can process and transmit the information to the output layer using algorithms or mathematical operations. The neurons are connected through weighted connections in the network. These neurons work as independent elements and are coupled with a transfer function that defines how the weighted sum of the inputs is translated to an output value. The neurons in the hidden layer, or the output neurons, receive several weighted input signals from each unit in the layer underneath it (Kukreja et al., 2016).

As depicted in Figure 2, the inputs are associated with a specific weight at the beginning of a particular neuron, i.e., each input is multiplied by a separate weight. All the multiplied values are then summed up with the bias values, which are then transmitted to the outputs through a transfer function, also known in other terms as an activation function. Output can be one or more than one. The information is transferred to each output. Bias is an additional parameter added to increase the prediction power of the ANN. It allows each neuron to have a baseline activation level. The working of an artificial neuron in a neural network can be defined by Equation 1.

$$y = T\left(\sum x_i w_i + b\right) \quad (1)$$

Where y is the output, T is the transfer function, x is the input, w is the weight, and b is the bias value (Krenker et al., 2011).

Feed-forward Neural Networks

A feed-forward neural network is an artificial neural network in which the data flows from input to output in a forward or single direction without any back loops. Otherwise, the number of layers, the number of connections within artificial neurons, and the type of transfer function to be used are all unrestrained (Krenker et al., 2011). In other words, the selection choice for all these parameters is open and can be changed based on the accuracy of the results. A feed-forward neural network is a multilayer perceptron with input, hidden, and output layers, as shown in Figure 3.

The circles in Figure 3 represent the neurons. The information enters the input layer and then goes through the hidden layer to the output layer. A hidden layer is necessary for an artificial neural network as it performs a bulk processing of the input data and allows the ANN to determine non-linear relationships between inputs. In the hidden layer, a set of weights and biases are applied to the inputs to produce an output for each neuron in the layer.

Table 1. Statistical values for certain Doğancı dam's water quality parameters (1990-2020)

Parameters (Monthly Average Values)	Average± Standard Deviation	Min-Max	USEPA Standards[a] (USEPA, 2009)	WHO Standards[b] (WHO, 2022)	Turkish Standards (TS266, 2005)
pH	8.047±0.226	7.443-8.451	6.5-8.5	6.5-8.0	6.5-9.5
Turbidity (NTU)	6.207±7.312	0.388-85.752	1	5	1
Water Temperature (°C)	12.881±4.367	5.214-23.206	-	-	-
Suspended Solids (mg/L)	4.816±5.791	0.000-38.375	-	-	-
Dissolved oxygen (mg/L)	5.036±2.830	0.800-9.912	-	-	-
Alkalinity (mg/L)	175.050±25.070	134.000-242.000	-	-	-
Iron (mg/L)	0.043±0.030	0.007-0.175	0.3	0.3	0.2
Manganese (mg/L)	0.047±0.037	0.013-0.288	0.05	0.1	0.05
Arsenic (mg/L)	0.005±0.001	0.003-0.006	0.01	0.01	0.01

a,b USEPA stands for the United States Environmental Protection Agency, and WHO is abbreviated for World Health Organization.

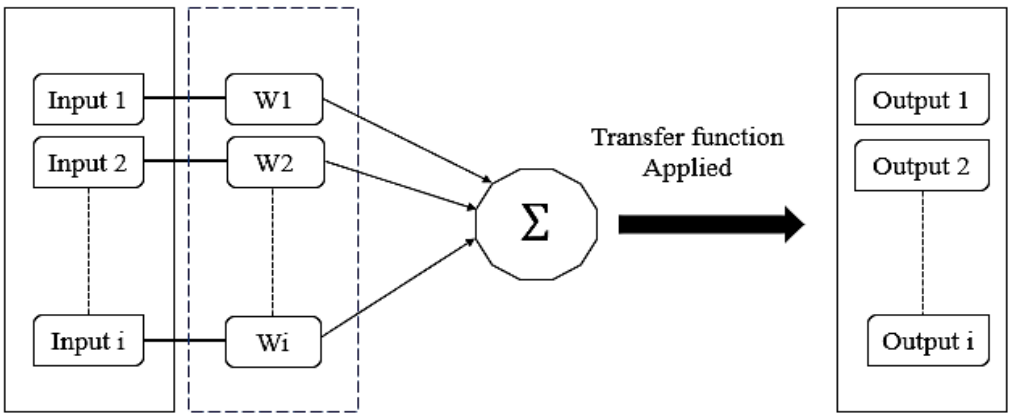

Figure 2. Working of a particular artificial neuron

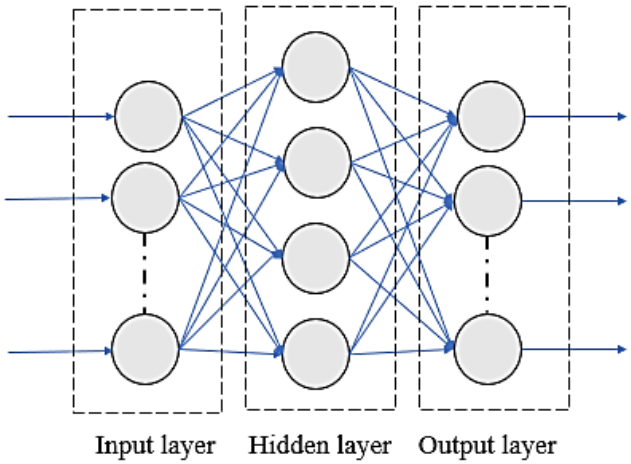

Figure 3. Structure of a feed-forward neural network

Training Algorithms

Two training algorithms were used in the current research, i.e., Levenberg-Marquardt and resilient backpropagation, to analyze which algorithm gives comparatively reliable predictions and displays a good correlation between input and output variables.

Levenberg-Marquardt's (LM) algorithm aims to minimize the sum of the squared error functions (Aldrich, 2002).

$$\mathrm{SSE} = \frac{1}{2}\sum(\mathrm{e_i})^2 - (2)$$

In Equation 2, e_i is the error associated with a particular vector element. The difference between the previous weight error and the new weight error should be as small as possible. It is a combination of the steepest descent and the Gauss-Newton method. The LM algorithm adjusts a parameter, the damping factor, between the two listed methods to control the step size of the updates to parameter values. A smaller damping factor leads to larger step sizes. LM has become a typical strategy for non-linear least-squares problems or where the residuals (differences between predicted and actual values) are highly non-linear (Lourakis, 2005).

Conversely, the resilient backpropagation algorithm (Rprop) is a first-order technique that enables guided batch learning in a multilayered perceptron. The fundamental premise of Rprop is to eliminate the negative impact of the partial error derivative size on the weight step. As a result, just the sign of the derivative indicates the direction of the weight update. The magnitude of the weight shift is defined solely by batch learning (Mushgil et al., 2015).

Application of Feed-forward Neural Networks

The current study used a feed-forward neural network (Asadollahfardi et al., 2017; Özçelik, 2015) to evaluate two different models. The analysis was done using MATLAB (R2022a version). Model 1 used water temperature, turbidity, and pH of Doğancı dam as input parameters, while dissolved oxygen, alkalinity, and the concentration of suspended solids, arsenic, manganese, and iron were the outputs. Likewise, model 2 was defined to have water temperature, alkalinity, pH, and concentration of dissolved oxygen as input parameters. The outputs of model 2 were chosen to be the concentrations of trace metals, i.e., arsenic, manganese, and iron. Prior to training the artificial neural network models, the dataset was preprocessed to improve model reliability and prediction stability. Missing values in both input and output datasets were handled using linear interpolation followed by nearest-neighbor replacement to remove any remaining undefined values. This preprocessing step ensured complete datasets for model development and eliminated computational instability during training. The data had very few missing values. Linear interpolation was chosen because it is an appropriate technique for handling time-dependent datasets (Katip and Anwar, 2025). To improve convergence and avoid dominance of large numerical values, all datasets were normalized between 0 and 1 using the Min-Max normalization technique (mapminmax). Normalization also enhanced the robustness, stability, and reproducibility of the developed ANN models by reducing sensitivity to variable scaling and improving the consistency of network learning performance (Chen, et. al., 2020). The dataset was then divided into training, validation, and testing subsets using a 70%-15%-15% random split configuration. The validation stage was included to reduce overfitting and improve generalization performance during network optimization. 70% of the data was used for training, while 30% was used for testing and validation (Chen and Liu, 2014; Zhang et al., 2015). A trial-based approach was used to determine the number of nodes in the hidden layer, with ten nodes providing the robust performance among the tested configurations. Some studies on water quality prediction in surface waters using artificial neural networks have shown that the best performance is obtained with a model structure containing 10 hidden neurons (Vijay and Kamaraj, 2021; Rizal, et. al., 2022).

Although this approach has been commonly used in ANN model development, it may be computationally time-consuming, which represents a limitation (Hettiarachchi et al., 2008; Ibnu et al., 2019). The hidden layer was processed through the sigmoid activation function, and a linear function was applied in the output layer (Karul et al., 2000; Özçelik, 2015). The epochs were set to 1000. The initial weight change was set to 0.07 because of the stability of the obtained results. LM and Rprop algorithms were used for training, and the modeling results were compared using correlation (R), root mean square error (RMSE), mean absolute error (MAE) and Nash-Sutcliffe efficiency (NSE) coefficient values.

Correlation is a statistical metric that describes the strength and direction of a relationship between two variables. The correlation coefficient might be between -1 and 1. A correlation coefficient of 1 shows a complete positive correlation, in which as one measure rises, the other rises linearly. A correlation coefficient of -1 denotes a perfect negative correlation, in which as one variable increases, the other variable drops linearly. A correlation coefficient of 0 shows that the two variables have no linear relationship (Asuero et al., 2006). On the other hand, RMSE and MAE are commonly used statistical parameters for evaluating the accuracy and predictive performance of models. MAE measures the average absolute difference between predicted and actual values, providing a straightforward indication of the average prediction error. Simultaneously, RMSE is calculated as the square root of the average squared difference between predicted and observed values, making it more sensitive to larger errors. Lower values of both RMSE and MAE indicate better model performance and higher prediction accuracy. MAE provides an easily interpretable measure of average error magnitude. However, RMSE places greater weight on large deviations and is therefore useful for identifying models with significant prediction errors. Together, these parameters provide a comprehensive assessment of model reliability and accuracy (Chai and Draxler, 2014; McCuen, 2016). Lastly, NSE coefficient is also a statistical parameter that measures how well the model reproduces the variability of the actual data. NSE values closer to 1 indicate a perfect performance of the model and high prediction accuracy, while values closer to or below zero indicate poor predictive capability (Dambe and Smit, 2021).

Results and Discussion

By referring to Table 1, it can be seen that the pH of the water of Doğancı dam does not differ much throughout the study period, i.e., it varies between 7.4 and 8. However, there was a noticeable difference between the maximum and minimum values for turbidity, water temperature, concentration of suspended solids, dissolved oxygen, and alkalinity, which was also indicated by their high standard deviation values. The difference could be due to seasonal changes in Bursa. If looking at the values for trace elements, the average value of manganese (0.047 mg/L) touched the boundary defined by USEPA and Turkish standards (0.05 mg/L), with the maximum value exceeding these limits. The range and the average values for iron were acceptable. Arsenic concentration was also within limits and it did not pose a health risk level.

Table 2 gives the values for the statistical parameters against each model using different training algorithms. The FNN structure in the mentioned table indicates the number of neurons or nodes in the input, hidden, and output layers, i.e., the first model had three inputs, ten hidden neurons, and six outputs. The second model had four inputs, ten hidden, and three output nodes. A neural network is related to the number of neurons in its hidden layers. A small number of neurons results in a poor fit, while a large number of neurons results in a better fit. This can cause the fitness curve to fluctuate significantly, even when the trained data is properly matched. The scientifically recommended number of hidden nodes is directly proportional to the number of input and output neurons (Beale, et al., 1992; Moghaddasi et al., 2017). There is no exact number to determine the hidden layer. However, node numbers of 5 and 10, as taken from the literature, have been tried (Vijay and Kamaraj, 2021). In this study, the number of neurons in the hidden layer was taken as 10, and its effect on performance was observed to be positive.

Table 2. Statistical analysis of the results obtained from FNN modeling

	Inputs (Monthly average values)	Outputs (Monthly average values)	FNN Structure	Checks			
				R	RMSE	MAE	NSE
1	• Water temperature (℃) • pH • Turbidity (NTU)	• Alkalinity (mg/L) • Suspended solids (mg/L) • Dissolved oxygen (mg/L) • Arsenic (mg/L) • Manganese (mg/L) • Iron (mg/L)	3-10-6	RProp: 0.70516 LM: 0.72166	RProp: 0.21316 LM: 0.20811	RProp: 0.17335 LM: 0.16685	RProp: 0.28063 LM: 0.31769
2	• Water temperature (℃) • pH • Alkalinity (mg/L) • Dissolved oxygen (mg/L)	• Arsenic (mg/L) • Manganese (mg/L) • Iron (mg/L)	4-10-3	RProp: 0.99454 LM: 0.99928	RProp: 0.03416 LM: 0.01242	RProp: 0.02662 LM: 0.00673	RProp: 0.99003 LM: 0.99898

The results showed that the second model indicated a good correlation between input and output values, with R varying between 0.994 and 0.999 for the entire data, considering both training algorithms, i.e., LM and RProp. In terms of prediction error, model 1 exhibited comparatively higher error values, with RMSE reaching up to 0.21. In contrast, model 2 showed substantially lower RMSE values, ranging approximately from 0.01 to 0.03, indicating a closer agreement between the predicted and observed values. A similar trend was observed for MAE, with model 2 consistently yielding lower values. The minimum MAE was approximately 0.007 using the Levenberg-Marquardt algorithm. Furthermore, the NSE values also confirmed the superior performance of model 2. Model 1 produced relatively low NSE values, i.e., between 0.280 and 0.318, indicating limited ability to reproduce the variability of the observed data. In contrast, model 2 achieved very high NSE values of 0.990 and 0.999, suggesting a higher agreement between predicted and observed values and confirming its stronger model efficiency. Therefore, based on the combined evaluation of RMSE, MAE, R, and NSE coefficient, model 2 can be considered more reliable and effective for predicting the water quality of Doğancı Dam.

The regression plots presented in Figure 4 further demonstrate the superior predictive performance of model 2 compared to model 1. In the regression analysis, a model with predictions closely aligned to the target values is indicated by data points distributed near the diagonal reference line (Y=T). For model 1, trained using both the Rprop and LM algorithms (Figures 4a and 4b), the regression plots exhibit a relatively scattered distribution of data points around the regression line, with lower correlation coefficients of approximately 0.705 and 0.722, respectively. This dispersion indicates weaker agreement between predicted and actual values, suggesting the presence of larger prediction

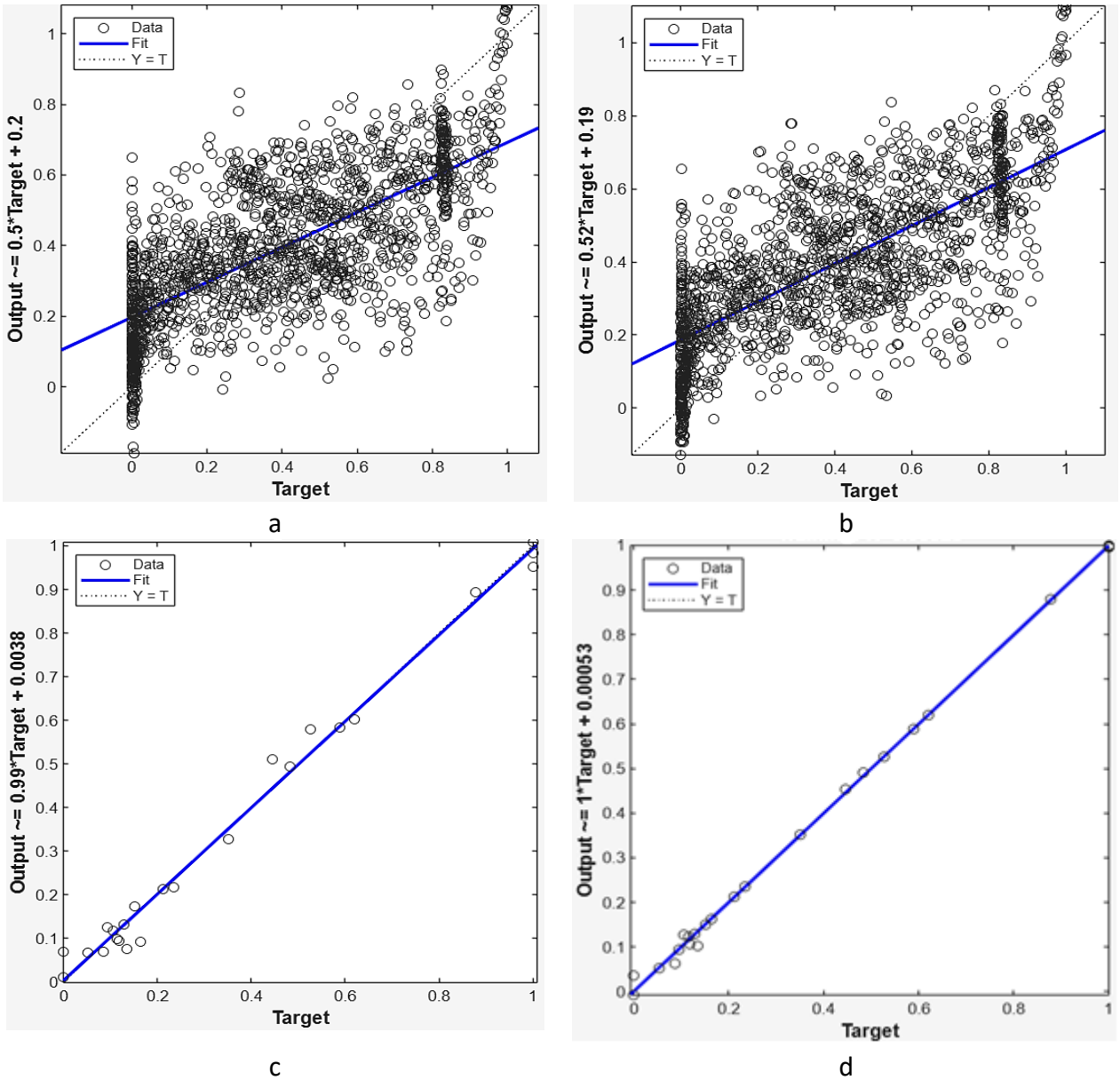

Figure 4. Regression plots obtained from neural networking using Rprop training algorithm (a- Model 1 and c- Model 2) and LM training algorithm (b- Model 1 and d- Model 2).

deviations and reduced model reliability. In comparison, model 2 (Figures 4c and 4d) showed an almost perfect linear relationship between predicted and target values. The data points were tightly clustered along the Y=T reference line with significantly higher correlation coefficients of approximately 0.994 and 0.999 for the Rprop and LM algorithms, respectively. The minimal scattering observed in these plots indicates very low prediction error and enhanced fitting capability of the neural network. Scientifically, this behaviour suggests that the selected input parameters in model 2 had a stronger and more consistent relationship with the output variables, enabling the network to learn the underlying patterns more effectively. Moreover, the LM-trained model 2 exhibited the highest regression coefficient and the least deviation from the ideal regression line, confirming its superior generalization ability and predictive accuracy. The correlation coefficient values of model 2 are closely comparable to the test set R values reported by (Vicente et al., 2012), which ranged from 0.994 to 0.996.

The results also correlated with the research's objective regarding the situation related to trace elements pollution in the Doğancı dam. It showed that the concentrations of trace elements in the dam could vary with the change in water temperature, pH, alkalinity, and dissolved oxygen concentration, as a high correlation was shown between them. This issue needs attention because an increase in the concentration of iron, manganese, and particularly arsenic can impact the public and be a root cause of different diseases ((Martuza, et. al., 2024; WHO, 2022). High R, low RMSE and MSE values could be observed in modeling systems where data is available for long term years, data quality is high, missing data is minimal, and variable selection and model architecture are correctly implemented. Long-term data includes seasonal variations, dry and wet periods, extreme events, and nonlinear events more comprehensively, thus forecasting performance improves. ANNs trained with little data produce high error (Banda and Kumarasamy, 2024; Zhi, et al., 2024; Mert and Kasapoğulları, 2024). In this study, 30 years of daily data were converted to monthly format for modeling. There were few missing data values. Furthermore, since the dam is located within a drinking water protection basin (BUSKİ, 2023), there are few pollution sources and therefore concentrations do not vary at very high levels. Therefore, the models were found to have high performance.

By comparing the current study with a few similar studies from previous literature (Table 3), it was observed that most of these studies used one input, one hidden, and one output layer in feed-forward neural networking. A similar structure was used in this study. Variation of the input and output numbers did not significantly change the correlation numbers. However, model structures with one output number were found to have fewer errors. The errors and correlation coefficients of the selected studies from the literature (as stated in Table 3) are for the complete dataset after testing, training, and validation. Earlier studies commonly predicted parameters such as chlorophyll-a, dissolved oxygen, conductivity, total dissolved solids, suspended solids, alkalinity, and eutrophication-related variables. In contrast, the present modeling focuses on the concentrations of arsenic, manganese, and iron, which are important from both water-treatment and public-health perspectives. The novelty of this study lies in estimating the amount of trace metals in drinking water dams based on a long-term data set. When compared with the literature from a statistical perspective, the developed models showed competitive and, in some cases, superior performance. Vicente et al., (2012) reported correlation coefficients ranging from 0.994 to 0.998 with error values between 0.097 and 0.300 for predicting oxidizability and total suspended solids. The correlation coefficient obtained from model 2 using the LM algorithm (R = 0.999) was higher than the range reported in the aforementioned study, while the corresponding RMSE and MAE values were also lower,

Table 3. Summary of a few studies from the literature on FNN usage in water quality estimation of a dam

Study No.	Inputs	Outputs	FNN Structure	Correlation coefficient	Error	Reference
1.	Calcium, magnesium, sodium, chloride, sulfate ions, and pH	Total dissolved solids	One input, one hidden, and one output layer	0.840-0.985	0.011-0.019	(Asadollahfardi et al., 2017)
2.	Suspended solids, total nitrogen, total phosphorus, chemical, and biological oxygen demand	Suspended solids, total nitrogen, and total phosphorus	One input, one hidden, and one output layer	0.659-0.712	2.699-3.120	(Kang et al., 2015)
3.	Chlorophyll-a, phosphate, nitrate-nitrogen, and ammonium nitrogen	Chlorophyll-a	One input, one hidden, and one output layer	0.710-0.740	-	(Park et al., 2015)
4.	pH, conductivity, dissolved oxygen, and water temperature	Oxidizability and total suspended solids	One input, two hidden, and one output layer	0.994-0.998	0.097-0.300	(Vicente et al., 2012)
5.	Suspended solids, total nitrogen, and water temperature	Conductivity and dissolved oxygen	One input, one hidden, and one output layer	-	4.720-6.460	(Elhatip and Kömür, 2008)
6.	Phosphates, nitrates, alkalinity, suspended solids, pH, water temperature, conductivity, dissolved oxygen, and sechi depth	Chlorophyll-a	One input, one hidden, and one output layer	0.950	-	(Karul et al., 2000)

indicating improved predictive accuracy and reduced model uncertainty. Similarly, Asadollahfardi et al. (2017) reported error values ranging from 0.011 to 0.019, which are comparable to the RMSE value obtained for model 2. In contrast, studies such as Kang et al. (2015) and Elhatip and Kömür (2008) reported considerably larger error ranges of 2.699-3.120 and 4.720-6.460, respectively, indicating comparatively lower prediction precision than the models studied in the present study.

The higher correlation coefficients and lower error values observed in model 2 indicate that the selected input parameters established a stronger relationship with the target water quality variables, allowing the neural network to learn the underlying trends more effectively. Furthermore, the superior agreement between predicted and actual values indicates that the LM algorithm provided better optimization and convergence characteristics for the developed FNN architecture. Overall, the statistical comparison with previously published studies confirms that model 2 exhibited highly reliable and accurate predictive performance for water quality assessment. Model 1 also demonstrated comparable predictive capability with previously reported studies from the literature. The RMSE values of model 1 (up to 0.21) remained lower than the larger error ranges reported by a few studies in Table 3. The findings indicate that model 1 can capture the relationship between the selected input and output parameters. However, the accuracy remains low as evaluated from statistical performance.

Lastly, an inference was made to determine the training algorithm that performed better based on R, RMSE, and MAE values for the models studied. For both developed models, the LM algorithm consistently demonstrated superior predictive performance compared to the Rprop algorithm. In model 1, the LM algorithm produced a slightly higher correlation coefficient (R=0.722) than Rprop (R=0.705), while simultaneously yielding lower RMSE and MAE values of 0.208 and 0.167, respectively, compared to 0.213 and 0.173 obtained using Rprop. Similarly, in model 2, LM achieved a higher correlation coefficient than Rprop, along with significantly lower RMSE and MAE values (as seen in Table 2). The NSE values also supported this trend, as LM produced higher NSE values than Rprop in both Model 1 and Model 2, indicating improved model efficiency and predictability.

The consistently better statistics achieved suggest that LM was more effective in capturing the nonlinear relationships between the input and output parameters within the dataset. Scientifically, the enhanced performance of the LM algorithm can be attributed to its faster convergence capability and efficient optimization process, which combines the advantages of gradient descent and the Gauss-Newton method. This enables LM to minimize the error function more effectively, particularly for complex nonlinear modeling problems such as water quality prediction. Additionally, among the developed models and training approaches, model 2 trained using the LM algorithm exhibited the best predictive performance. Therefore, it can be deduced from the analysis that the Levenberg-Marquardt algorithm was more suitable for determining dam water quality, particularly the concentration of trace elements) using water temperature, pH, alkalinity, and dissolved oxygen concentration as inputs. In surface waters, increasing temperature decreases dissolved oxygen concentrations, which may create reducing conditions that enhance the solubility and mobility of certain metals. In contrast, increasing alkalinity and pH generally reduce metal solubility by forming insoluble hydroxide and carbonate species. Furthermore, higher turbidity can promote the adsorption of metal ions onto suspended particles, thereby decreasing dissolved metal concentrations in the water. However, mixing caused by water currents may disturb bottom sediments and release accumulated metals back into the water (Okbah et al., 2020). In artificial neural network modeling in ecological systems, selecting biochemically related parameters as output parameters increases the model performance (Singh et al., 2009). In both model structures, temperature and pH were used as common input parameters. Except for temperature and pH, two different model structures were created: model 1 to examine the relationship between turbidity and metals, and model 2 to examine the relationship between alkalinity, dissolved oxygen, and metals. Black box models detect patterns but cannot independently explain cause-and-effect relationships. Establishing causality would require additional experimental evidence, intervention-based analyses, or formal causal inference frameworks (Zhao et al., 2021). However, in groundwater and streams, machine learning models have been able to establish relationships between parameters affecting metal mobility and metal concentrations, thus it provided new information to increase the explanatory power of the modeled process (Ibrahim et. al., 2022). Since chemical relationships were considered when constructing the model structures, the interpretations also included discussions of the relationships between temperature, pH, dissolved oxygen, and trace metals, and attempts were made to approach meaningful inferences (Peters et al., 2017). Therefore, selecting the parameters used in the output layer in model 2 from the same type (only metals) increased the performance of the model. It was thought that the reason for model 2's higher performance might be due to the larger number of inputs. Examining Table 3, it was seen that models with a larger number of input parameters (Karul et al., 2000) and a larger number of hidden layers (Vicente et al., 2012) provide better performance.

Overall, the discussion confirms that the developed ANN framework provides reliable predictive performance for trace metal estimation in Doğancı Dam; its methodological limitations should also be acknowledged. The ANN models were evaluated using a single train-validation-test split, which is commonly

applied in ANN-based modeling (Rasmussen et al., 2003; Wu et al., 2013). Nonetheless, it could introduce bias that may possibly impact the predictive validation of the models. Additionally, although the validation subset was used during model development to reduce overfitting, the use of a single random split does not completely exclude the possibility that the reported performance was influenced by the selected data partition (Wu et al., 2012). Cross-validation could provide a more robust estimate of generalization performance by evaluating the model across multiple data splits (Hadjisolomou et al., 2021; Qiu, 2024). Therefore, future studies could apply k-fold cross-validation and systematic hyperparameter optimization, such as grid search, random search, or Bayesian optimization, to further confirm model stability and improve the reproducibility and generalizability of the ANN framework.

Conclusion

Two models with distinct feed-forward neural structures were evaluated, and their R, RMSE and MSE values were compared in this study. Two different training algorithms were also applied to each of the tested models in the research. The model with four inputs (water temperature, pH, alkalinity, and dissolved oxygen concentration), ten hidden nodes, and three outputs (concentrations of arsenic, manganese, and iron) had a stronger correlation coefficient for the entire dataset and a much lower RMSE and MSE values. Comprehensive preprocessing techniques including missing-value treatment and Min-Max normalization improved model stability and training consistency. Comparative analysis between the trainrp and trainlm learning algorithms demonstrated that the Levenberg-Marquardt algorithm was more appropriate for modeling the dam's water quality using the input variables provided. Regression analysis and residual analysis further verified the predictive capability and reliability of the proposed framework. In the applied model structures, it was determined that a larger number of hidden layer nodes and input parameters increased model performance, while a larger number of outputs also increased the number of errors. Furthermore, the use of long-term 30-year data and the narrow number of missing data also improved model performance.

The findings of this study raise the possibility of using artificial neural networks to predict or generate hypotheses for trace metal concentrations in drinking water reservoirs using long-term monitoring data. The results of this study are valuable for the development of sampling, monitoring, and protection strategies to manage the water quality of drinking water dams. Nevertheless, ecological relationships were not determined with this model and future studies should further validate the proposed ANN framework using cross-validation and systematic hyperparameter optimization.

Ethical Statement

Not applicable.

Funding Information

This work was not supported by any funding.

Author Contribution

A.K.: Conceptualization, formal analysis, investigation, methodology, resources, supervision, software, validation, writing- original draft, writing, review and editing.

A.A.: Investigation, methodology, software, validation, writing- original draft.

Conflict of Interest

The authors declare that they have no known competing financial or non-financial, professional, or personal conflicts that could have appeared to influence the work reported in this paper.

References

Aldrich, C. (2002). *Exploratory analysis of metallurgical process data with neural networks and related methods*. Elsevier.

Al-Yaari, M., Aldhyani, T.H.H., Rushd, S. (2022). Prediction of Arsenic Removal from Contaminated Water Using Artificial Neural Network Model, Applied Sciences, 12,999. https://doi.org/10.3390/app12020999

Asadollahfardi, G., Zangooei, H., Aria, S. H., & Danesh, E. (2017). Application of artificial neural networks to predict total dissolved solids at the Karaj Dam. *Environmental Quality Management, 26*(3), 55-72. https://doi.org/10.1002/tqem.21493

Asuero, A. G., Sayago, A., & González, A. (2006). The correlation coefficient: An overview. *Critical Reviews in Analytical Chemistry, 36*(1), 41-59. https://doi.org/10.1080/10408340500526766

Bayram, A. (2011). *Harşit çayı su kalitesinin mevsimsel değişiminin incelenmesi ve askı madde konsantrasyonunun yapay sinir ağları yöntemi ile tahmin edilmesi.* Karadeniz Teknik Univeristy, . Retrieved from http://acikerisim.ktu.edu.tr/jspui/handle/123456789/714

Banda, T. D., Kumarasamy, M. (2024). Artificial neural network (ANN)-based water quality index (WQI) for assessing spatiotemporal trends in surface water quality—A case study of South African river basins. Water, 16(11), 1485. https://doi.org/10.3390/w16111485

BUSKI (2023). Barajlar. Doğancı Barajı. Retrieved from https://www.buski.gov.tr/Icerik/IcerikDetay?url=barajlar

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE). *Geoscientific Model Development Discussions, 7*(1), 1525-1534. https://doi.org/10.5194/gmdd-7-1525-2014

Chen, W. B., & Liu, W.C. (2014). Artificial neural network modeling of dissolved oxygen in reservoir. *Environmental Monitoring and Assessment, 186*(2), 1203-1217. https://doi.org/10.1007/s10661-013-3450-6

Chen, Y., Song, L., Liu, Y., Yang, L., & Li, D. (2020). "A Review of the Artificial Neural Network Models for Water Quality Prediction." *Applied Sciences*, 10(17), 5776. https://doi.org/10.3390/app10175776

Chou, J. S., Ho, C. C., & Hoang, H. S. (2018). Determining quality of water in reservoir using machine learning. *Ecological Informatics, 44*, 57-75. https://doi.org/10.1016/j.ecoinf.2018.01.005

Chukwuemeka, P., Imoni, O., Mogo, F. C., et al. (2026). A hybrid AI model integrating LSTM, XGBoost, and K-means for interpretable prediction and clustering of water quality in data-scarce regions. Discover Civil Engineering, 3, 24. https://doi.org/10.1007/s44290-026-00417-x

Dambe, N., & Smit, J. (2021). Application of low to medium resolution data for hydrological modeling in Malawi. In *Space Fostering African Societies: Developing the African Continent through Space, Part 2* (pp. 39-53). Cham: Springer International Publishing.

Elhatip, H., & Kömür, M. A. (2008). Evaluation of water quality parameters for the Mamasin dam in Aksaray City in the central Anatolian part of by means of artificial neural networks. *Environmental Geology, 53*, 1157-1164. https://doi.org/10.1007/s00254-007-0705-y

Fural, Ş., Kükrer, S., Cürebal, İ., & Aykır, D. (2022). Ecological degradation and non-carcinogenic health risks of potential toxic elements: a GIS-based spatial analysis for Doğancı Dam (). *Environmental Monitoring and Assessment, 194*(4), 269. https://doi.org/10.1007/s10661-022-09870-4

Hadjisolomou, E., Stefanidis, K., Herodotou, H., Michaelides, M., Papatheodorou, G., & Papastergiadou, E. (2021). Modelling freshwater eutrophication with limited limnological data using artificial neural networks. *Water*, *13*(11), 1590. 10.3390/w13111590

Hettiarachchi, Y. N., & Premaratne, H. L. (2008, September). A new parameter determining mechanism for radial basis neural networks. In *2008 Eighth International Conference on Hybrid Intelligent Systems* (pp. 787-792). IEEE. https://doi.org/10.1109/HIS.2008.150

Ibnu, C. R. M., Santoso, J., & Surendro, K. (2019, February). Determining the neural network topology: A review. In *Proceedings of the 2019 8th International Conference on Software and Computer Applications* (pp. 357-362). https://doi.org/10.1145/3316615.3316697

Ibrahim, B., Ewusi, A., Ahenkorah, I. (2022). Assessing the suitability of boosting machine-learning algorithms for classifying arsenic-contaminated waters: A novel model-explainable approach using SHapley Additive exPlanations. Water, 14(21), 3509. https://doi.org/10.3390/w14213509

Kalin, L., & Isik, S. (2010). Prediction of water quality parameters using an artificial neural networks model. Paper presented at the World Environmental and Water Resources Congress 2010: Challenges of Change, 3145-3153.

Kang, B., Do Kim, Y., Lee, J. M., & Kim, S. J. (2015). Hydro-environmental runoff projection under GCM scenario downscaled by Artificial Neural Network in the Namgang Dam watershed, Korea. *KSCE Journal of Civil Engineering, 19*(2), 434-445. https://doi.org/10.1007/s12205-015-0580-0

Katip, A., & Anwar, A. (2025). Simulating the Impacts of Climate Change on the Hydrology of Doğancı Dam in Bursa, Turkey, Using Feed-Forward Neural Networks. Sustainability, 17(14), 6273. https://doi.org/10.3390/su17146273

Karaca, G., & Tasdemir, Y. (2014). Temporal and spatial variations in PAH concentrations in the sediment from the Nilufer Creek in Bursa, *Journal of Environmental Science and Health, Part A, 49*(8), 900-912. https://doi.org/10.1080/10934529.2014.894330

Karul, C., Soyupak, S., Çilesiz, A. F., Akbay, N., & Germen, E. (2000). Case studies on the use of neural networks in eutrophication modeling. *Ecological Modelling, 134*(2-3), 145-152. https://doi.org/10.1016/S0304-3800(00)00360-4

Krenker, A., Bešter, J., & Kos, A. (2011). Introduction to the artificial neural networks. *Artificial Neural Networks: Methodological Advances and Biomedical Applications. InTech*, 1-18.

Kukreja, H., Bharath, N., Siddesh, C., & Kuldeep, S. (2016). An introduction to artificial neural network. *Internationa Journal of Advanced Research and Innovative Ideas in Education, 1*, 27-30.

Kumar, S., Pati, J. (2023). Machine learning approach for assessment of arsenic levels using physicochemical properties of water, soil, elevation, and land cover. Environmental Monitoring and Assessment, 195(6), 641. https://doi.org/10.1007/s10661-023-11231-8

Li, P., Hua, P., Gui, D. Niu, J., Pei, P., Zhang, J., Krebs, P. (2020). A comparative analysis of artificial neural networks and wavelet hybrid approaches to long-term toxic heavy metal prediction. *Sci Re* 10, 13439. https://doi.org/10.1038/s41598-020-70438-8

Lourakis, M. I. (2005). A brief description of the Levenberg-Marquardt algorithm implemented by levmar. *Foundation of Research and Technology, 4*(1), 1-6.

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. Science of the Total Environment, 694, 133591. https://doi.org/10.1016/j.scitotenv.2019.133591

Martuza, M. A., Shafiquzzaman, M., Haider, H., Ahsan, A., & Ahmed, A. T. (2024). *Predicting removal of arsenic from groundwater by iron based filters using deep neural network models*. Scientific Reports, 14, 26428. https://doi.org/10.1038/s41598-024-77536-5

McCuen, R. H. (2016). Assessment of hydrological and statistical significance. *Journal of Hydrologic Engineering*, *21*(4), 02516001.

Mert, B. K., Kasapoğulları, D. (2024). A case study of using artificial neural networks to predict heavy metal pollution in Lake Iznik. Environmental Monitoring and Assessment, 196(6): 586. https://doi.org/10.1007/s10661-024-12730-y

Moghaddasi, H., Ahmadzadeh, B., Rabiei, R., & Farahbakhsh, M. (2017). *Study on the efficiency of a multi-layer perceptron neural network based on the number of hidden layers and nodes for diagnosing coronary artery disease*. Jentashapir Journal of Health Research, 8(4), e63032. https://doi.org/10.5812/jjhr.63032

Mohammadi, B., Moazenzadeh, R., Christian, K., Duan, Z., Band, S. S., Mosavi, A., & Chau, K. W. (2022). Groundwater quality modeling with artificial intelligence techniques: A review. Environmental Science and Pollution Research, 29, 32329–32359. https://doi.org/10.1007/s11356-021-17838-z

Municipality, B. M. (Producer). (2023). Bursa's Population, Climate, and Geography. *Go to B*. Retrieved from

https://www.bursa.com.tr/en/sayfa/nufus-konum-iklim-ve-cografya-47/

Mushgil, H. M., Alani, H. A., & George, L. E. (2015). Comparison between resilient and standard back propagation algorithms efficiency in pattern recognition. *International Journal of Scientific and Engineering Research, 6*(3), 773-778.

Nafouanti, M. B., El Hmaidi, A., El Ouali, A., El Morhit, M., & others. (2026). Groundwater quality prediction using novel hybrid classification and regression models. Journal of Contaminant Hydrology, 277, 104834. https://doi.org/10.1016/j.jconhyd.2026.104834

Okbah, M. A., El-Gamal, M. M., & Sayad, H. H. (2020). Assessment of heavy metals pollution in water and sediments of Lake Manzala, Egypt. Environmental Monitoring and Assessment, 192, 1–15. https://doi.org/10.1007/s10661-020-8129-8

Özçelik, O. (2015). *Assessment and prediction of water quality parameters in Lake Köyceğiz using artificial neural network approach.* Middle East Technical University, . Retrieved from https://hdl.handle.net/11511/25382

Park, J., Kim, K. T., & Lee, W. H. (2020). Recent advances in information and communications technology (ICT) and sensor technology for monitoring water quality. *Water, 12*(2), 510. https://doi.org/10.3390/w12020510

Park, Y., Cho, K. H., Park, J., Cha, S. M., & Kim, J. H. (2015). Development of early-warning protocol for predicting chlorophyll-a concentration using machine learning models in freshwater and estuarine reservoirs, Korea. *Science of the Total Environment, 502*, 31-41. https://doi.org/10.1016/j.scitotenv.2014.09.005

Peters, J., Janzing, D., & Schölkopf, B. (2017). Elements of causal inference: Foundations and learning algorithms. The MIT Press. https://doi.org/10.7551/mitpress/10994.001.0001

Pule, M., Yahya, A., & Chuma, J. (2017). Wireless sensor networks: A survey on monitoring water quality. *Journal of Applied Research and Technology, 15*(6), 562-570. https://doi.org/10.1016/j.jart.2017.07.004

Qiu, J. (2024). An analysis of model evaluation with cross-validation: techniques, applications, and recent advances. *Advances in Economics, Management and Political Sciences*, *99*, 69-72. https://doi.org/10.54254/2754-1169/99/2024OX0213

Rasmussen, B., Wesley Hines, J., & Gribok, A. V. (2003, May). An applied comparison of the prediction intervals of common empirical modeling strategies. In *Proceedings of the 2003 Annual Maintenance and Reliability Conference, Knoxville, TN*.

Rizal, N. M. N., Hayder, G., Mnzool, M., Elnaim, B. M. E., Mohammed, A. O. Y., & Khayyat, M. M. (2022). Comparison between Regression Models, Support Vector Machine (SVM), and Artificial Neural Network (ANN) in River Water Quality Prediction. *Processes*, *10*(8), 1652. https://doi.org/10.3390/pr10081652

Schneider, M., & Edelsbrunner, P. (2013). Modelling for Prediction vs. Modelling for Understanding. *Frontline Learning Research*, *1*(2), 99-101. https://doi.org/10.14786/flr.v1i2.74

Sham, F. A. F., El-Shafie, A., Jaafar, W. Z. W., S, A., Sherif, M., Ahmed, A. N. (2025). Advances in AI-based rainfall forecasting: A comprehensive review of past, present, and future directions with intelligent data fusion and climate change models. Results in Engineering, 27, 105774. https://doi.org/10.1016/j.rineng.2025.105774

Singh, K. P., Basant, A., Malik, A., & Jain, G. (2009). Artificial neural network modeling of the river water quality—a case study. *Ecological modelling,* 220(6), 888-895. https://doi.org/10.1016/j.ecolmodel.2009.01.004

Türk Standardlari, E. (2005). *Türk Standardı TS 266 (Sular - İnsanî tüketim amaçlı sular)*. Retrieved from https://intweb.tse.org.tr/Standard/Standard/Standard.aspx?081118051115108051104119110104055047105102120088111043113104073100115103084043078075072076065089

Ucun Ozel, H., Gemici, B. T., Gemici, E., Ozel, H. B., Cetin, M., & Sevik, H. (2020). Application of artificial neural networks to predict the heavy metal contamination in the Bartin River. *Environmental Science and Pollution Research*, *27*(34), 42495-42512.

USEPA (Producer). (2009). National Primary Drinking Water Regulations. Retrieved from https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations

Vicente, H., Couto, C., Machado, J., Abelha, A., & Neves, J. (2012). Prediction of water quality parameters in a reservoir using artificial neural networks. *International Journal of Design, Nature and Ecodynamics, 7*(3), 310-319. https://doi.org/10.2495/DNE-V7-N3-310-319

Vijay, S., Kamaraj, K. (2021) Prediction of Water Quality Index in Drinking Water Distribution System Using Activation Functions Based Ann. Water Resour Manage 35, 535–553. https://doi.org/10.1007/s11269-020-02729-8

Vu, T. D., Nguyen, T. D., Pham, T. K. T., Berg, M., Pham, H. V. (2026). Evaluating groundwater quality in an arsenic-contaminated aquifer in the Red River Delta using machine learning: A case study in Van Phuc, Hanoi, Vietnam. Environmental Science: Advances, 5(4), 1027–1038. https://doi.org/10.1039/D5VA00368G

WHO. (2022). Drinking water quality guidelines. Retrieved from https://www.who.int/publications/i/item/9789240045064

Wu, W., May, R., Dandy, G. C., & Maier, H. R. (2012). A method for comparing data splitting approaches for developing hydrological ANN models.

Wu, W., May, R. J., Maier, H. R., & Dandy, G. C. (2013). A benchmarking approach for comparing data splitting methods for modeling water resources parameters using artificial neural networks. *Water Resources Research*, *49*(11), 7598-7614. https://doi.org/10.1002/2012WR012713

Zhang, Y., Huang, J. J., Chen, L., & Qi, L. (2015). Eutrophication forecasting and management by artificial neural network: a case study at Yuqiao Reservoir in North China. *Journal of Hydroinformatics, 17*(4), 679-695. https://doi.org/10.2166/hydro.2015.115

Zhao, Q., & Hastie, T. (2021). Causal interpretations of black-box models. *Journal of Business & Economic Statistics*, *39*(1), 272-281. https://doi.org/10.1080/07350015.2019.1624293

Zhi, W., Appling, A.P., Golden, H.E., Podgorski, J., Li, L. (2024). Deep learning for water quality. Nat Water 2, 228–241, https://doi.org/10.1038/s44221-024-00202-z

Zhu, M., Wang, J., Yang, X., Zhang, Y., Zhang, L., Ren, H., Wu, B., & Ye, L. (2022). A review of the application of machine learning in water quality evaluation. *Eco-Environment and Health*. https://doi.org/10.1016/j.eehl.2022.06.001