International Journal of Design & Nature and Ecodynamics
Vol. 20, No. 1, January, 2025, pp. 187-196
Journal homepage: http://iieta.org/journals/ijdne

Applications of Artificial Intelligence Methods for Irrigation Water Quality Index: Review

Mariam Hamad Abd, Yasir Al-Ani*, Mohammed Falah Allawi

Department of Dams & Water Resources Engineering, Faculty of Engineering, University of Anbar, Ramadi 31001, Iraq

Corresponding Author Email: aniyaser@uoanbar.edu.iq

Copyright: ©2025 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.18280/ijdne.200120

Received: 3 December 2024
Revised: 8 January 2025
Accepted: 15 January 2025
Available online: 31 January 2025

Keywords:
artificial neural networks, deep learning, irrigation water quality index, machine learning, random forest, sustainable water resources

ABSTRACT

The irrigation water quality index (IWQI) is utilized to quantify the suitability of water for growing crops. Irrigation water quality is important because it affects soil properties, plant growth, and agricultural production. The IWQI is calculated to reduce complex water quality data to a single number index to make it easier for decision-makers, researchers, and farmers to evaluate whether the water is appropriate for applying irrigation. When determining the IWQI, several parameters must be taken into consideration, including sodium absorption ratio (SAR), electrical conductivity (EC), pH levels, and concentrations of sulfates, chlorides, bicarbonates, and heavy metals that may be toxic to crops. The IWQI is calculated using these parameters, where a higher IWQI score indicates that irrigation is appropriate due to the water quality. This study reviewed previous studies that discussed artificial intelligence (AI) algorithms from 2016 to 2024 to predict the IWQI. Researchers have turned to artificial intelligence techniques to estimate and predict IWQI instead of traditional methods such as linear regression. Traditional methods have limitations, including a reliance on large sample sizes to achieve high accuracy and reduce errors. Also, large sample sizes require laboratory tests that demand more time, effort and cost. Small sample sizes, on the other hand, often result in inaccurate outcomes, making them unreliable. In addition, traditional methods cannot handle missing or non-linear data and lack the ability to learn from new data for improved accuracy. As for artificial intelligence (AI) algorithms, significant amounts of data are collected in real-time using geographic information systems, remote sensing devices, or other automated systems. These data are processed faster, more accurately, and efficiently, and complex patterns and relationships that may need to be clarified using traditional methods are identified.

1. INTRODUCTION

Irrigation water is essential for the growth of crops, especially in arid or semi-arid areas. It provides essential moisture for the development of crops and prolongs the growing seasons. Irrigation water is usually used when natural water sources and rainfall are insufficient to meet the water needs of crops. The purpose of IWQI assessment is to ensure adequate water quality to improve crop productivity, maintain soil health and reduce environmental risks. The source of irrigation water can be groundwater, surface water, springs, wells, rivers, lakes, and reservoirs, or from another source, such as treated sewage or desalinated water. In addition, farmers must choose a suitable irrigation water source and regularly assess irrigation water quality because it is a critical factor affecting crop production and soil health. Poor irrigation water quality leads to reduced crop growth, soil deterioration, and the accumulation of salts that cause soil salinization, making it less fertile, and heavy metals that are harmful to crop growth, making it unsafe for consumption. Groundwater is the backbone of protecting the aquatic and nutritional environment, particularly in dry or semi-arid regions. At the same time, the increased construction of dams and the diversion of surface water have reduced groundwater recharge and lowered groundwater levels [1]. Groundwater is currently the focus of water resources researchers [2]. Low salinity, low concentrations of toxic elements, appropriate pH levels, and the absence of chemicals and biological pollutants characterize good irrigation water quality. Researchers have developed many irrigation water quality indicators for sustainable agricultural development using artificial intelligence algorithms instead of traditional methods that consume more time, effort and money. The limitations of traditional methods have forced researchers to rely on alternative methods, including artificial intelligence techniques. These limitations include the inability of traditional methods to deal with complex or non-linear data, lengthy data analysis times, inaccurate results, and difficulties in managing missing data. In addition, they consume time, effort, and cost because they require a large sample size and data analysis in advanced laboratories. Researchers [3-7] applied traditional methods such as Additive Regression (AR), linear regression (LR), autoregressive moving average (ARAMA), multilinear regression (MLR), and (ARIMA) and found that these models are not able to capture randomness because they deal with nonlinearity. On this basis, researchers conducted studies on artificial intelligence models because they can overcome the shortcomings and limitations of traditional methods. Artificial intelligence models handle linear, non-linear, complex and missing data, learn and develop from new data, and offer high accuracy while consuming less time, effort and cost. Their speed in analyzing data and reducing errors has enabled multiple applications in environmental engineering and water resources engineering. When a problem occurs and decision makers must take quick action regarding water use, check its quality, and know the type of water suitable for each plant, AI models are effective, their results are accurate, highly efficient, and give results in a short time, effort, and cost when data is available. They can deal with complex, non-linear data, and missing data, and they can learn and develop. Traditional methods are effective, but they require large samples that require laboratory tests that take more time, effort, and cost, and they only deal with linear data.

A study assessed groundwater quality characteristics to obtain suitable and sustainable water in the Alnekheeb Basin in Anbar, Iraq. The researcher [8] used two machine learning models, namely the RBF-NN and PNN. The models were developed to predict groundwater parameters represented by sodium absorption ratio (SAR), salinity, and water hardness, where many water parameters were used as inputs, such as Cl, HCO_3, NO_3, Ca, Na, Mg, SO_4, and CO_3. Results have shown the supremacy of the PNN model in all predictions of groundwater parameters compared to RBF-NN. The main objective of the research is to review the studies that have estimated and predicted the water quality index for irrigation purposes using artificial intelligence techniques.

This research mainly focused on studies that modelled water levels for irrigation using artificial intelligence (AI) algorithms and finding the best models according to different performance evaluation criteria. Therefore, three objectives of this research paper can be summarized:

- A preliminary evaluation of the real irrigation water quality was made by calculating water quality indicators.
- Using AI techniques to forecast the IWQ index.
- Some studies compared AI models and traditional methods models, and some studies compared AI models with each other to find the exact values of the IWQ index for sustainable water supply management.

2. ARTIFICIAL INTELLIGENCE APPLICATIONS FOR IWQI PREDICTION

To predict the IWQI, irrigation water quality parameters such as salinity expressed by EC, SAR, pH, MAR, nitrate levels, carbonates, bicarbonates, etc., must be determined, which affect crop production and soil health. AI models could be applied to predict the IWQ index by using data on irrigation water quality factors. These models can assist in accurately assessing the quality and levels of irrigation water demand and the capacity to decrease water wastage. AI models have been relied upon by researchers over the past twenty years because they are able to model linear and nonlinear (complex) relationships between parameters [9].

The process of predicting IWQI using AI first requires collecting data that contains water quality parameters such as sodium absorption ratio, pH, electrical conductivity, magnesium absorption ratio, anions and cations. Then the data is divided into two parts, one for training and the other for testing, as most researchers divide the data into 70% for training and 30% for testing. When determining the problem and the purpose of the study, whether it is for the purpose of prediction or improving water quality, the appropriate model is chosen according to the type of data and according to the problem and the purpose of the study, or several comparative work models are applied between the models to find out the best model for the study by calculating the statistical indicators for each model such as the correlation coefficient (R^2), the mean square error (MSE), the mean absolute error (MAE), and the root mean square error (RMSE).

The most important of these models are:

2.1 Artificial neural networks

The neural network contains a group of nerve cells connected to each other to form neural networks, as each layer contains cells connected to the cells of the layer that follows it without any connection between the cells of the layer itself [10]. Each network contains a different number of neurons located in each layer of the neural network. The input layer is where the data is processed and transferred to the next layer, which is the hidden layer, in which the complex relationships between patterns are represented and calculated. Then the output layer interacts with this process to find an estimate and predict the required value, as shown in Figure 1, which illustrates the ANN layers, the first layer represents the input layer that absorbs the raw data and converts it into a suitable structure, and the second layer represents the hidden layer, which can be one or more hidden layers. It is the layer in which the calculations, data processing, and error reduction are performed to give results with improved accuracy, and the last layer represents the output layer through which the final results are obtained. Thus, the neural network process can be described by finding the sum of the product of each input. In its weight value, then the sum is processed using a non-linear transfer function, then the final result is found for all cells of the neural network. The most commonly used transfer function is the S-shaped sigmoid curve [11]. When developing accurate models for artificial neural networks, it is necessary to separate the data into training and test groups. The training data is used to find the relationship between the expected and observed outputs [12]. Cross-validation closely monitors the training process and test data through which the performance of the neural network is evaluated [13]. The task of determining the quantity of layers that are concealed and nodes is based on trial and error when modeling (ANN). There is a general rule to find the hidden layer in which the sum of the training set contains more samples than the synaptic weights set. Training is stopped if the validation error rises after a certain number of iterations, in which case biases and weights are returned at the minimal validation error [14]. Artificial neural networks are used for future prediction, classification, and pattern recognition tasks [15]. In addition, study [16] estimate and predict the groundwater quality index using three ANN algorithms, where he took 16 water quality parameters collected from 47 wells and springs in Andimeshk in Iran such as turbidity, EC, pH, TH, Mg, F, Cr, Cu, Mn, Fe, and TDS. The models included artificial neural networks with Bayesian regularization, early-stopping artificial neural networks, and artificial neural networks with ensemble averages. It was the first study on AI to forecast the groundwater quality index. This prediction can decrease the time, effort, and potential of calculation errors. The results indicated good forecasting of the groundwater quality index, as the ANN model with

Bayesian regularization had a higher prediction performance than the rest of the models.

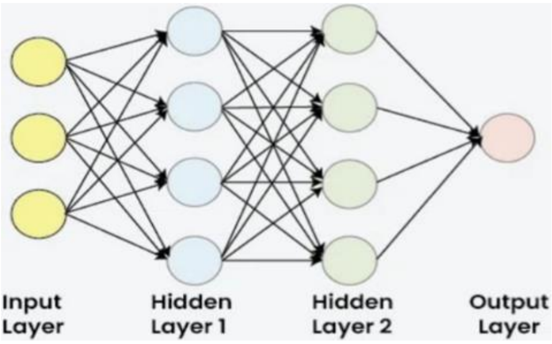

Figure 1. Artificial neural network algorithm

Vasanthi and Kumar [17] collected 64 water samples from Parakai Lake, which is a sewage discharged untreated water supplying irrigation areas in South Kanyakumari, where the lake is fed by the Pazhayar River in India, to forecast the water quality index using (ANN) algorithm and (MLR) model. Total salt, dissolved sodium ratio, boron, sodium, total dissolved solids, residual sodium carbonate, and sodium absorption ratio were input parameters based on improving irrigation water quality indicators. The water quality index is the output layer. The study showed that due to its mild pollution, the water is only fit for agriculture, i.e., drinking and other life-sustaining activities are not appropriate uses for it. The main reason for the pollution was the increased concentration of sodium. This study proved that the (ANN) model gives higher accuracy and better performance with a correlation coefficient (R^2=0.9907) than the (MLR) model as (R^2=0.7908) when predicting (IWQI) [17].

Gaya et al. [18] collected samples from the Yamuna River in India and calculated four parameters such as (PH, NH_4, DO, BOD) to estimate and predict the water quality index using a traditional model such as MLR with AI models, such as ANN, ANFIS, which gave superior performance in estimating and predicting the water quality index by 10% compared to the MLR model when estimating and predicting the water quality index. In addition, ANN performed better than ANFIS.

Yıldız and Karakuş [19] collected 32 water samples from Yukari Basin in Central Anatolia to estimate and predict water quality for surface irrigation purposes using ANN model and multiple regression analysis (MRA) model based on some water quality parameters such as (IWQI, SAR, Na%, PI). The results showed that the water samples were suitable for irrigation purposes and that the ANN model with (R^2=0.92) had better results than the MRA model with (R^2=0.6) when estimating and predicting (IWQI).

Abdel-Fattah et al. [20] relied on data collected for three years to estimate and predict the water quality index for irrigation purposes for the Bahr El-Baqar Drain in Egypt, based on some parameters such as (K, Na, Mg, Ca, SO_4, EC), where the prediction process was carried out using the (ANN) model with the (ARIMA) model to estimate the water quality index. Results indicated that the water quality index of Bahr El-Baqar Drain for irrigation purposes ranged from 46.15 (demonstrating superior water quality for irrigation) to 80.82 (showing that there is enough water for agriculture) during training and also during testing and verification, the ANN model gave a high prediction value in IWQI modelling when the (R) was more significant than 0.98. At the same time, the ARIMA model gave a coefficient of determination (R^2=0.23), and thus, according to the study, the ANN model has higher reliability with ten hidden neurons selected for it.

Ubah et al. [21] conducted research for irrigation purposes aiming to predict the water quality index of Ele River Nnewi, Anambra State, for one year using ANN by forecasting four water quality indicators, including pH, EC, Na, and TDS, at four distinct sites. After the modeling process and estimation and prediction of water indicators, the results showed that EC, TDS, and Na gave a higher scale than the standards of the Food and Agriculture Organization adopted in this region for the dry season's irrigation needs. As for the pH, it gave a regular scale and was within the standards of the Food and Agriculture Organization.

M'nassri et al. [22] took 49 groundwater samples and 15 parameters, for example, pH, (EC), (TDS), (Na%), (SAR), (MAR), (PI), (RSC), and essential ions, including (Na) (Mg), (Cl), (K), (Mg), (Ca), (HCO_3), and (SO_4). In a semi-arid Ouled Chamkeh Plain basin area, Sidi El Hani in the Mahdia region in Tunisia uses (ANN) and (MLR) models to forecast the irrigation water quality index. Results proved the excellent execution of ANN with (R^2=0.92, RMSE=1.02 MAE=0.9) and MLR with (R^2=0.81, RMSE=1.2, MAE=1.63) models regarding predicted and actual water quality.

Kandil et al. [23] conducted a study in a desert area with limited water resources with groundwater used extensively for irrigation in the Nile Valley in Egypt, where he created an ANN based on seventy-seven samples representing water quality indicators for the suitability of groundwater use such as (specific toxins, salinity, infiltration and various effects) based on several water quality parameters as inputs to the ANN model including Cl, Na, EC, SAR and HCO_3. The ANN model produced encouraging results for classifying irrigation water's quality and assessing its appropriateness. Results indicated that these indicators (HCO_3 and SAR) have a minimal effect on (IWQI) and can be dispensed with when calculating its value without a noticeable impact on the precision of the model's execution for monitoring and evaluating (IWQ).

Gautam et al. [12] performed a study in Pratapgarh district in southern Rajasthan, India, for consideration into whether groundwater is suitable for irrigation of agricultural lands using ANN algorithm to estimate and predict sodium hazards such as Na%, SAR, RSC, and Kelly ratio (KR). Some parameters (Ca, Mg, Na, K, HCO_3, CO_3, PH) were relied upon as inputs to the ANN model, which were calculated after collecting water samples from 76 wells. It was found that the (ANN) model gave an accurate and good prediction value with (R^2=1, IA=1, RMS=0.14, MBE=0.005), so it is a suitable model for predicting the groundwater quality index for irrigation purposes and producing good crops.

Gaagai et al. [24] used two AI models, ANN and GBR, to estimate and predict the quality of groundwater for irrigation purposes northeast of Ouled Djelal in the Doucen plain in Algeria. After examining 27 groundwater samples by traditional analytical methods, the parameters showed that $Cl^- > SO_4^{2-} > HCO_3^- > NO_3^-$ and $Ca^{2+} > Mg^{2+} > Na^+ > K^+$, which indicates that the water type is Ca-Cl because limestone predominates, sand, and clay minerals resulting from human activities, ion dissolution, and rock decomposition. In addition, the IWQI, Kelly index (KI), sodium absorption ratio (SAR), PI, Na%, and MH proved that 33% of the water samples are strongly restricted for crop irrigation, and 67% of the water samples represent medium too tight of an irrigation limitation, which indicates that crops have highly sensitive to salt. The findings indicated that the ANN model with (R^2=0.973, RMSE=2.492) for the training and (R^2=0.958, RMSE=2.175) for validation was the majority of precise forecaster among the rest of the models.

Gad et al. [25] collected surface water samples from 51 different locations across the Nile River in Egypt to estimate and predict water quality for irrigation purposes using the ANN with R^2=0.999 for calibration, R^2=0.945 for validation and PLSR with R^2=0.999 for calibration and validation models based on the parameters calculated from the water samples such as (Na%, SSP, SAR, MH, PI, PH, EC, T, TDS, Mg, K, Ca, NO_3, CO_3, SO_4, HCO_3). It was found that the models used performed well and were accurate in predicting and estimating water quality, as 98% of the water samples were from unrestricted irrigation, where crops that can tolerate salinity are not allowed to grow, and about 2% of the water samples are considered low restrictions in the use of crop irrigation.

Omeka [26] collected 21 samples of surface and groundwater in the southeastern section of Obubra in Rivers State in Nigeria to forecast the IWQ index using MLP-ANN and MLR models using input parameters such as (PH, EC, TH, TDS, Cl^-, HCO_3^-, SO_4^{2-}, NO_3^-, Ca^{2+}, Mg^{2+}, Na^+, K^+) and output parameters are (PI, PS, SAR, SSP, KR, MAR). The researcher also used a map of irrigation water quality through the weighted sum overlay method. The map revealed that 72.5% of the area was suitable for irrigation in the southeastern section, and 28.4% of Irrigation was not feasible in the region, found in isolated traces according to the MLP-ANN model's sensitivity analysis. The parameters (Cl^-, EC, SO_4^{2-}, PH, HCO_3^-) showed the most significant effect on the (IWQ) index for these. In addition, the results showed very few modelling errors in both models, proving the models' effectiveness. However, the MLP-ANN model gave higher accuracy with R^2=0.983 in predicting the IWQ index.

To find the IWQ index, Palabıyık and Akkan [27] took 14 variables from a sample obtained from the Aksu River in Turkey, which was observed in 5 different locations for a year to calculate some parameters such as turbidity, alkalinity, hardness, EC, TDS, DO, pH, total phosphorus (TP), total acid number (TAN), Na, K, NO_2, NO_3 and SO_4. They used AI algorithms including (MLP-ANN) and (MLR). They also used ML models like as NN, SVM, GPR, ensemble approach and decision tree. The results showed that 80% of the analyzed water gave a good quality index for general use, and 20% gave a poor-quality index and was unsuitable for general use. However, the river water was classified as specifically suitable for irrigation purposes.

2.2 Support vector machine

SVM is a machine learning model that basically classifies data into two groups by separating them with a single straight line called a hyperplane, as shown in Figure 2, which illustrates the separation of data using the SVM model and the selection of the best straight line (hyperplane) that separates the data from each other. The best straight line that separates the data can be known if the distance between the hyperplane and the closest data points is equal, and this distance is called the margin. This model also gives good solutions to convex and nonlinear optimization problems, as it excels over the ANN model, which is characterized by a local minimum and classifies data in a multi-class manner [28]. The best classification can be obtained when the margin, which represents the distance between the straight line separating the data and the nearest points of that data in each group, is equal [29, 30]. There are many uses for the SVM model such as predicting water quality index, lake and river discharge, and flood discharge. Therefore, the SVM model has achieved wide success in several fields [31]. The SVM model is a kernel-based technique for solving classification and regression problems, where the kernel type is chosen according to the data. The linear kernel is suitable if the data is linear, and the radial basis function kernel (RBF) and the sigmoid kernel are suitable in the case of non-linear data. In addition, we choose the kernel type according to the goal of the model, whether the goal is prediction or classification, as well as statistical indicators such as (R^2, MSE, RMSE, MAE) through which we choose the appropriate kernel type, where the values of the statistical indicators are found for all types of kernels, and we choose which kernel gave the best results for these indicators.

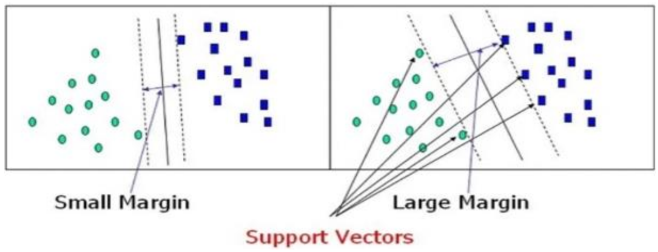

Figure 2. SVM methods

The input layer is nonlinear and transforms into a high-dimensional space [32]. The data close to the straight line called the hyperplane are called support vectors and are important for training the SVM model [33]. There are several researchers who have conducted studies to estimate and predict water quality index using SVM model, such as, Seyedzadeh et al. [34] took 1725 water samples, each sample was taken from a variable discharge according to pressure and temperature, in order to estimate and predict the emitter discharge for drip irrigation purposes using AI models such as (NF-SC, NF-FCM, ANN, LS-SVM). The results revealed the superiority of the (LS-SVM) model over the rest of the models used in terms of accuracy and better performance in prediction.

Shiri et al. [11] collected groundwater samples from 90 wells in Tabriz Plain in Iran. Some parameters were calculated from these samples such as (EC, TDS, SAR, SO_4) to predict the irrigation water quality index using different AI models such as (SVM, ANN, RF, MARS, ANFIS-SC, ANFIS-GP, SVM-FFA, ANN-FFA). The best prediction performance was obtained by the ANN-FFA and SVM-FFA models in both scenarios with a lower error rate.

Dimple et al. [35] selected five ML models, specifically random space (RSS), linear regression (LR), AR, REP Tree, and SVM to predict the IWQ index using the groundwater of the Nand Samand basin in Rajasthan, India. The models for machine learning were created and evaluated on six IWQ indices, namely Na%, sodium (SAR), Kelly ratio (KR), PI, MH, and sodium solubility ratio (SSP). Results indicated that the SVM model gave the most significant outcomes while testing among the rest of the models where (R^2=0.9364, 0.9618, 0.9588, 0.9819, 0.9547, 0.8903) and (RMSE=0.0662, 4.0568, 3.0168, 0.1113, 3.7046, 5.1066) and (MAE=0.042, 3.1999, 2.3584, 0.0726, 2.9603, 4.0582) for (KR, MH, SSP, SAR, %Na, PI), respectively. Therefore, artificial intelligence (AI) models can improve water quality features for irrigation and produce ideal crops.

Mokhtar et al. [36] used three AI algorithms (SVM, XGB, and RF) and four multiple regression models, including (SW stepwise regression, PCR, OLS, and PLS) to estimate and predict six IWQ indicators in Bahr El-Baqr in Egypt, such as RSC, SSP, SAR, Salinity Potential, Kelly's ratio KR, and PI. As for the model inputs, EC, Na^+, Ca^{+2}, and HCO_3^- were used. From the results of the RMSE and SI, the best performance in prediction was shown for the SW stepwise regression model by 21% and 0.03%, respectively, followed by PCR and PLS by 0.22% for RMSE and 0.21% for SAR, as for artificial intelligence models, the best model that gave good performance in prediction is SVM. According to the findings, the water source is unsuitable for irrigation and must be treated so that salt does not accumulate in the soil. The researcher proved that multiple regression and AI models can forecast and estimate the (IWQ) index.

Derdour et al. [37] selected samples from the driest and hottest regions of the world in the southwestern region of Adrar, Algeria, to forecast groundwater quality for use in irrigation. They collected 166 samples, each containing physical and chemical elements such as EC, temperature (T), cations such as Mg, Na, Ca, potassium (K), anions such as (SO_4), (Cl), (HCO_3), and pollution indicators such as (NO_3). The prediction process was carried out using artificial intelligence methods, including SVM and k-nearest neighbors (KNN) with five input parameters such as (SAR), (Na), (EC), chloride concentration (Cl), and bicarbonate level (HCO_3). The findings indicated that 57.23% of the groundwater samples were inappropriate for irrigating agricultural fields, 12.05% were of acceptable quality, 21.08% were adequate, and 9.64% were exceptional. Additionally, the findings demonstrated that the SVM and KNN models supply practical estimation and prediction in desert areas, especially SVM, as they gave a higher prediction accuracy.

Ibrahim et al. [38] collected 140 groundwater samples in the Western Desert area in El Kharga Oasis in Egypt and calculated some water quality parameters from these samples such as (PS, SSP, RSC, SAR, KI, PH, Na, T, Ca, K, EC, Mg, Cl, S, C, H, N) to estimate and predict water quality for irrigation purposes using AI models such as SVM and ANFIS. Results indicated that the ANFIS and SVM models could forecast the IWQ indicator with reasonable preciseness, as most of the groundwater samples were categorized for the purpose of irrigation.

2.3 Support vector regression

It is commonly called support vector regression (SVR) and is a support vector machine SVM model extension used in classification. SVR is used on regression tasks, meaning it estimates and predicts continuous values rather than group or class labels. The goal of SVR is to calculate a function whose result represents the expected inland IWQ index after inputting a set of water quality parameters [1]. There are studies to forecast the irrigation water quality index using (the SVR) model, including the deterioration of the IWQ index of the Nile River used for agricultural irrigation leads to geological pollution. Therefore, Elsayed et al. [39] conducted a study to forecast and estimate the (IWQ) index for irrigation purposes by collecting 110 samples from the surface water canal network for two years and measuring 21 physical and chemical parameters supported by multivariable models such as (SVR), (PCR) and (SMLR), in which the concentration of elemental ions appeared in the following form ($Ca^{2+} > Na^+ > Mg^{2+} > K^+$) and also (alkalinity > SO_4^{2-} > Cl^- > NO_3^- > F^-). As for the concentration of trace parameters, it appeared in the following form (B > Cr > Pb > Ni > Cu > Zn > Cd > Fe > Mn > B) and this indicates that the type of surface water is $Ca^{2+} > Mg^{2+} > Cl^- > SO_4^{2-}$ and $Ca^{2+} > Mg^{2+} > HCO_3^-$. This is due to silicate weathering pressure and the reverse ion exchange reaction. Six indicators of surface water quality were predicted, namely the (IWQI) which appeared in 82% of water samples and was considered a high group, and the remaining 18%, which represents an average group of surface water quality for irrigation purposes. As for the sodium percentage (Na%), it was 96% of water samples. The healthy category and 4% represent an acceptable category for irrigation. Furthermore, the indicators of SAR, Kelly index (KI), PI, and RSC showed that all samples were suitable for agricultural irrigation. Results indicated that the multivariable models gave strong performance and accurate prediction power for the (IWQ) index.

Eid et al. [40] took samples from 45 deep groundwater wells in Souf valley, the Algerian desert, to predict eight irrigation water quality indicators including SAR, Na%, SSP, PI, KI, potential salinity (PS), RSC, IWQI and some physical and chemical parameters including TDS, temperature (T), PH, EC, Na, potassium (K), Ca, Mg, NO_3, CO_3, HCO_3 and SO_4 using the support vector regression (SVR) and geographic information systems (GIS). The findings demonstrated that groundwater is usable for plants that tolerate salinity in large quantities because groundwater is classified from high restriction to medium restriction, as the SVR model was effective, reliable, and applicable when estimating all groundwater quality indicators for complex end layers for irrigation purposes.

2.4 Random forest

A model of machine learning that relies on decision trees [41]. It is mainly used to complete regression and classification tasks by building multiple decision trees as this model can deal with multiple data for this reason it has been used to predict water quality index and other applications that deal with large data and non-linear properties, as shown in Figure 3, which shows the clustering of decision trees in random forests, each tree works independently and specializes in a specific task to analyze data and produce the best prediction for the data. Random forest scheme is based on combining several decision trees, where each decision tree is a machine learning model based on decision-making principles, thus producing a forest of decision trees that work together to improve prediction accuracy, where each decision tree is trained on the available data and each tree gives a prediction value, then the predictions of the decision trees are combined to determine the final result. Building a random forest depends on two main parts: the number of trees that must be placed in the forest and the number of conditions that each node in the forest meets until the tree is developed [42, 43]. Increasing the accuracy of the (RF) model depends on the number and depth of trees, as increasing the number and depth of trees increases the accuracy and efficiency of the model while reducing potential errors, and increases the model's ability to learn and develop, but this requires more time for training. Despite this, researchers [44] conducted a study to find out the optimal number of trees in the forest, where he proved that sometimes increasing the number of trees in the forest leads to an increase in the calculation time only and does not affect the accuracy and performance of the model, because he analyzed 29 data with 128 trees and found that the difference in the results is very simple when using (256, 512, 1024, 2048, 4096) trees.

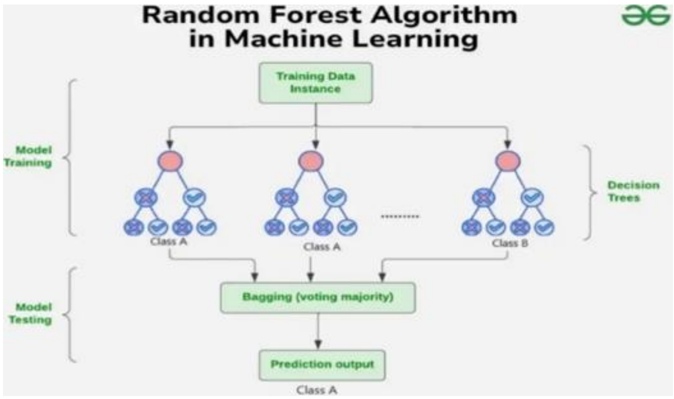

Figure 3. Random forest algorithm in machine learning

Several researchers forecast the IWQ index using random forest models. The researchers [45] started by collecting 520 inputs about 14 groundwater quality metrics in Morocco's Berrechid basin. To predict the potential parameters of salinity (PS), TDS, MAR, SAR, exchangeable sodium ratio (ESP), and RSC, where pH, EC, and temperature (T) were used as input indicators. They developed and evaluated four artificial intelligence models, including (ANN), random forests (RF), adaptive boosting models (Ada boost), and (SVR). The findings indicated that the effectiveness of (RF and AB) in prediction is higher than that of (SVR and ANN). For example, the statistical indicators of the TDS parameter when predicting it were (R^2=0.94, 0.96, 0.96, 0.99) and (RMSE=1270.68, 400.63, 343.92, 182) for the (SVR, ANN, RF, Ada boost) models respectively, and also the rest of the parameters (RSC, MAR, ESP, SAR, PS) gave better statistical indicators for the two models (RF & Ada boost), which made their performance better than the performance of the rest of the models used in this study. while at the global level, compared to RF and AB, ANN and SVR are less sensitive to input factors and more generalizable.

To forecast groundwater quality index values in the Illizi region, southeast Algeria. Kouadri et al. [28] used eight machine learning algorithms MLR, RF, MSP Tree (M5P), RSS, AR, ANN, SVR, and LWLR. The work was based on 114 samples collected from six groundwater layers simultaneously. The work was developed in two stages. The first stage aims to reduce the time consumed to calculate WQI, where all indicators were taken as inputs. The second stage aims to highlight the water quality variation when the necessary analyses in critical cases are unavailable, such as when sensitivity analysis is used to decrease all inputs. The RF model outperformed the other models, giving (R^2=0.9984, MAE=1.9942, RMSE=3.2488) for training and (R^2=0.9926, MAE=2.1563, RMSE=3.8228) for testing. These results were better than the results of the other models used in this study.

2.5 Gradient boosting machine

It is a model of ML models used for regression and classification, as it works to build models sequentially and corrects errors in previous models [46] by merging them into a single, powerful model. It works to build decision trees sequentially, unlike random forests that build decision trees independently, as the GBM model creates decision trees, with each tree concentrating on fixing the mistakes of the one before it. In addition, the GBM model gives high accuracy in performance. However, its training time is slow, as it builds decision trees sequentially, and it also improves a specific loss function by adding decision trees that work to reduce the remaining error in the previous model [47]. As shown in Figure 4, which describes the gradual assembly of a set of simple models into decision trees, each time a new decision tree is produced, it corrects the errors of the models in the previous tree to improve accuracy and give the best results.

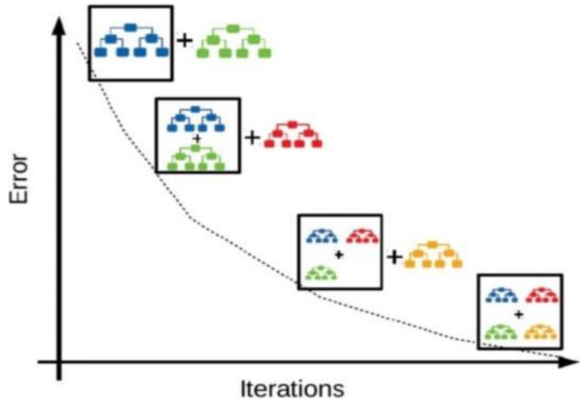

Figure 4. Gradient boosting algorithm in machine learning

Through this model, the IWQ index was predicted by several researchers such as, Raheja et al. [48] collected 392 groundwater samples in Haryana State, India, and measured 12 physical and chemical parameters such as (Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, Cl^-, SO_4^{2-}, NO_3^-, F-, PH, EC, TH) to predict water quality through two indicators: WQI and entropy water quality index (EWQI) using three machine learning models: Extreme gradient boosting (XGB), gradient boosting machine (GBM), and deep neural network (DNN). 294 samples were taken (75%) for training and 98 (25%) for testing. It was observed that (EC) is the most important input parameter in (WQI and EWQI) predictions, and pH is the least important in prediction. The findings indicated that the DNN model gave the best accuracy in forecasting in both indicators with a relatively lower error value.

To improve the prediction of water quality for irrigation systems in the Red River Delta, Vietnam, compared with traditional methods with less cost and time. Nguyen et al. [49] carried out research to calculate the surface water quality index efficiently and accurately for irrigation purposes using two ML models such as gradient boosting (GB) technique to select appropriate parameters for the models including (BOD), (DO), aluminum (NH_4), (PO_4), suspended solids (TSS), coliforms and turbidity. Machine learning models such as (GB) model, (XGBoost) model, and deep learning models such as LSTM model and RNN model were relied upon to estimate and predict the surface water quality index for irrigation purposes. The results indicated that (GB) model is more effective because it gave the highest coefficient of determination (R^2=0.96) during prediction, followed by (XGBoost) model which gave (R^2=0.89), followed by (LSTM) model which gave (R^2=0.85) and (RNN) model which gave (R^2=0.84), where the results indicated that ML models gave more effectiveness than DL models.

2.6 Deep learning

The DL model is a branch of artificial intelligence techniques containing deep multi-layer neural networks because they contain many hidden layers. Deep learning is effective and powerful, especially for tasks that depend on a large set of data, because its working principle is inspired by the structure and functions of the human brain, as it works to analyze and improve large amounts of data with high accuracy in less time, effort and cost than traditional statistical methods. Deep learning models have been widely used in engineering to estimate and predict the quality index of irrigation water based on water quality parameters such as ion concentration, acidity, salinity, hardness, pH, and other physical and chemical components. Deep learning models are selected according to the dataset relationships' complexity, size, and nature. The feedforward neural network (FNN) model is selected for regression tasks to predict IWQI, the convolutional neural network (CNN) model is selected for spatially structured data or data obtained from satellite images that affect water quality, and (RNN) model, also called (LSTM) networks, is selected for time-series data taking seasonal variations in water quality into account. The appropriate model is chosen in DL according to the type of data and the purpose of the study. The LSTM model is used if the data is time-series, as it can predict a large number of temporal data with high accuracy. As for the CNN model, it is used with data taken from several different locations, meaning that it deals with spatial data because it is suitable for detecting spatial patterns and changes. The LSTM model has a system that controls the parameters obtained and through which we get the final result of the prediction process [50]. The prediction process of the LSTM model is sequential, as this model excels in prediction and learning with high accuracy and good efficiency from sequential time data [51]. Therefore, the LSTM model, considered one of the deep learning models, is used to predict time-series data with a long-range [52]. The LSTM model contains memory blocks connected through layers. It also contains gates: the input gate, the output gate, and a gate between them called the forget gate [53]. As shown in Figure 5, which illustrates the main gates in neural networks for deep learning models, they are the input gate that controls the data signals entering the model, the forgetting gate that controls the forgetting memory, and the output gate that controls the data signals exiting the model. In addition, recent research used models containing two or more models called hybrid deep learning models, where these models outperformed any independent model capable of predicting time series data [54, 55]. Research has shown that deep learning models such as LSTM and RNN are good models for estimating and predicting IWQI with high accuracy and better performance [56-58]. Sewage water is polluted and contains various chemicals, so it is necessary to estimate and predict the quality of this water in the event of recycling and using it for irrigation purposes so that the pollutants and chemicals present in sewage water are not transferred to crops. In this way, Chen et al. [59] collected samples of contaminated wastewater to predict water quality index for irrigation purposes and find chemical oxygen demand (COD) through NIR technology. It must be subject to AI to solve the problems of large numbers of data and solve dynamic problems if they appear during work. An improved convolutional neural network model (CNN) was developed to deeply calibrate the data obtained from near-infrared rays. The model structure is shallow and built with a convolutional layer and one pool for each. The researcher used the decision tree algorithm in the pooling layer to find informational characteristics in a data-driven manner. The researcher collected 100 samples of sewage water in China. Eighty-three samples were taken, considered the most polluted, and divided into two parts, one containing 55 samples. For training and a part containing 28 samples to validate the model, the results showed a good fit of the decision tree algorithm with the shallow CNN and good accuracy prediction of the water pollution level through the rapid calculation of the COD values of wastewater samples.

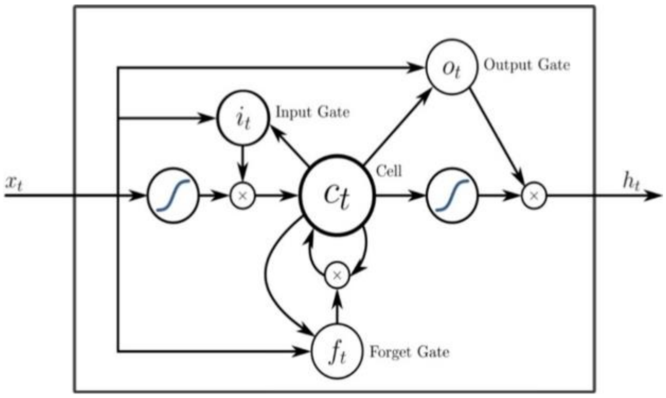

Figure 5. Deep learning algorithm

3. DISCUSSION

AI models are selected to predict IWQI to improve water quality, increase agricultural production, and reduce water waste. All of these models require dividing data into training data, testing data, data analysis, and improved prediction tasks and high accuracy. In addition, each model is selected to perform a specific task. For example, the ANN model gives high accuracy in prediction and can analyze large data and represents complex relationships between parameters. The SVM model solves classification and regression tasks and also works to predict with high accuracy as it can handle disparate data and works to separate them perfectly. The RF and GBM models deal with different and multiple data and solve classification and prediction problems with high accuracy. DL models are advanced models and give high accuracy in prediction and are used to analyze complex data. These advantages have made researchers in recent years use these models to estimate and predict IWQI for sustainable water resources management.

4. CONCLUSIONS

Previous studies on predicting irrigation water quality using AI techniques have proven practical and innovative. AI models, such as ML, DL, ANN, SVM, and gradient boosting machines, have proven highly accurate and efficient in estimating and predicting IWQ based on various parameters like pH, salinity, ion concentration, and other physical and chemical components. These AI-based techniques or models provide efficient processing of large data sets and scalable solutions in real time that traditional methods may not achieve. As AI techniques continue to develop, applying AI models to predict water quality will undoubtedly support or encourage decision-making processes, ensure sustainable agriculture practices, and improve water use. However, there is still a need for more research to address difficulties or challenges such as data availability, model generalization, and integration with other intelligent agriculture technologies, which confirms the realization of the full potential of AI techniques in this field. Many researchers have replaced traditional methods and relied on AI models because they can handle complex, non-linear and missing data and give results in less time, effort and cost than traditional methods that cannot handle this type of data as they require a large sample size and this sample needs laboratory tests that take time, effort and cost. Through previous studies, it was shown that AI models gave high prediction accuracy and efficiency, so they were adopted to estimate and predict IWQI as they can process and analyze large and complex amounts of data at high speed and their ability to achieve and improve high accuracy when predicting IWQI. These advantages of AI technologies help engineers and water managers make informed decisions about water use and thus increase the efficiency of water resources and improve water quality, which leads to increased agricultural crop production. AI models are used in multiple applications in water resources engineering, environment and agriculture to monitor water quality, improve agricultural productivity, know the health of crops, give warning in case of a problem, and direct water to plants as needed, thus reducing water waste. This provides integrated and sustainable management of water resources.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the College of Engineering, Department of Dams and Water Resources, University of Anbar, and to researcher Othman Abdulhameed Mahmood, who contributed specific information to the research.

REFERENCES

[1] Noori, R., Maghrebi, M., Mirchi, A., Tang, Q., Bhattarai, R., Sadegh, M., Noury, M., Haghighi, A.T., Kløve, B., Madani, K. (2021). Anthropogenic depletion of Iran's aquifers. Proceedings of the National Academy of Sciences, 118(25): e2024221118. https://doi.org/10.1073/pnas.2024221118

[2] Shihab, A., Al-Ani, Y., Mustafa, A.S. (2024). Assessment of quantity of a specified groundwater wells in Anbar governorate, Iraq. AIP Conference Proceedings, Al Anbar, Iraq, 3009(1). https://doi.org/10.1063/5.0190419

[3] Juahir, H., Zain, S.M., Toriman, M.E., Mokhtar, M., Man, H.C. (2004). A pplication of a rtificial neural Network models transportation. Langat River Basins is the most rapid urban area in Malaysia, 16(2): 42-55.

[4] Zare Abyaneh, H. (2014). Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters. Journal of Environmental Health Science and Engineering, 12: 1-8. https://doi.org/10.1186/2052-336X-12-40

[5] Kadam, A.K., Wagh, V.M., Muley, A.A., Umrikar, B.N., Sankhua, R.N. (2019). Prediction of water quality index using artificial neural network and multiple linear regression modelling approach in Shivganga River basin, India. Modeling Earth Systems and Environment, 5: 951-962. https://doi.org/10.1007/s40808-019-00581-3

[6] Tiwari, S., Babbar, R., Kaur, G. (2018). Performance evaluation of two ANFIS models for predicting water quality index of River Satluj (India). Advances in Civil Engineering, 2018(1): 8971079. https://doi.org/10.1155/2018/8971079

[7] Hoque, M.A., Apon, A.A., Hassan, M.A., Adhikary, S.K., Islam, M.A. (2024). Enhanced forecasting of groundwater level incorporating an exogenous variable: Evaluating conventional multivariate time series and artificial neural network models. Geographies, 5(1): 1. https://doi.org/10.3390/geographies5010001

[8] Allawi, M.F., Al-Ani, Y., Jalal, A.D., Ismael, Z.M., Sherif, M., El-Shafie, A. (2024). Groundwater quality parameters prediction based on data-driven models. Engineering Applications of Computational Fluid Mechanics, 18(1): 2364749. https://doi.org/10.1080/19942060.2024.2364749

[9] Shamshirband, S., Jafari Nodoushan, E., Adolf, J.E., Abdul Manaf, A., Mosavi, A., Chau, K.W. (2019). Ensemble models with uncertainty analysis for multi-day ahead forecasting of chlorophyll a concentration in coastal waters. Engineering Applications of Computational Fluid Mechanics, 13(1): 91-101. https://doi.org/10.1080/19942060.2018.1553742

[10] Kulisz, M., Kujawska, J., Przysucha, B., Cel, W. (2021). Forecasting water quality index in groundwater using artificial neural network. Energies, 14(18): 5875. https://doi.org/10.3390/en14185875

[11] Shiri, N., Shiri, J., Yaseen, Z.M., Kim, S., Chung, I.M., Nourani, V., Zounemat-Kermani, M. (2021). Development of artificial intelligence models for well groundwater quality simulation: Different modeling scenarios. Plos one, 16(5): e0251510. https://doi.org/10.1371/journal.pone.0251510

[12] Gautam, V.K., Pande, C.B., Moharir, K.N., Varade, A.M., Rane, N.L., Egbueri, J.C., Alshehri, F. (2023). Prediction of sodium hazard of irrigation purpose using artificial neural network modelling. Sustainability, 15(9): 7593. https://doi.org/10.3390/su15097593

[13] El-Shafie, A., Najah, A., Alsulami, H.M., Jahanbani, H. (2014). Optimized neural network prediction model for potential evapotranspiration utilizing ensemble procedure. Water Resources Management, 28: 947-967. https://doi.org/10.1007/s11269-014-0526-1

[14] Beale, M.H., Hagan, M.T., Demuth, H.B. (2010). Neural network toolbox. User's Guide, MathWorks, 2: 77-81. https://ge0mlib.com/papers/Books/04_Neural_Network_Toolbox_Reference.pdf.

[15] Mahmoud, O.A., Sulaiman, S.O., Al-Jumeily, D. (2023). Artificial neural network model for forecasting haditha reservoir inflow in the west of Iraq. 2023 16th International Conference on Developments in eSystems Engineering (DeSE), Istanbul, Turkiye, pp. 138-143. https://doi.org/10.1109/DeSE60595.2023.10468804

[16] Sakizadeh, M. (2016). Artificial intelligence for the prediction of water quality index in groundwater systems. Modeling Earth Systems and Environment, 2: 1-9. https://doi.org/10.1007/s40808-015-0063-9

[17] Vasanthi, S.S., Kumar, S.A. (2019). Application of artificial neural network techniques for predicting the water quality index in the Parakai Lake, Tamil Nadu,

India. Applied Ecology & Environmental Research, 17(2): 1947. https://doi.org/10.15666/aeer/1702_19471958
[18] Gaya, M.S., Abba, S.I., Abdu, A.M., Tukur, A.I., Saleh, M.A., Esmaili, P., Wahab, N.A. (2020). Estimation of water quality index using artificial intelligence approaches and multi-linear regression. IAES International Journal of Artificial Intelligence, 9(1): 126-134. https://doi.org/10.11591/ijai.v9.i1.pp126-134
[19] Yıldız, S., Karakuş, C.B. (2020). Estimation of irrigation water quality index with development of an optimum model: A case study. Environment, Development and Sustainability, 22: 4771-4786. https://doi.org/10.1007/s10668-019-00405-5
[20] Abdel-Fattah, M.K., Mokhtar, A., Abdo, A.I. (2021). Application of neural network and time series modeling to study the suitability of drain water quality for irrigation: A case study from Egypt. Environmental Science and Pollution Research, 28(1): 898-914. https://doi.org/10.1007/s11356-020-10543-3
[21] Ubah, J.I., Orakwe, L.C., Ogbu, K.N., Awu, J.I., Ahaneku, I.E., Chukwuma, E.C. (2021). Forecasting water quality parameters using artificial neural network for irrigation purposes. Scientific Reports, 11(1): 24438. https://doi.org/10.1038/s41598-021-04062-5
[22] M'nassri, S., El Amri, A., Nasri, N., Majdoub, R. (2022). Estimation of irrigation water quality index in a semi-arid environment using data-driven approach. Water Supply, 22(5): 5161-5175. https://doi.org/10.2166/ws.2022.157
[23] Kandil, N.M., Rayan, R., Sadek, M.A. (2023). Development of irrigation water quality index using artificial neural network. Journal of Scientific Research in Science, 40(1): 86-101. https://doi.org/10.21608/jsrs.2023.331805
[24] Gaagai, A., Aouissi, H.A., Bencedira, S., Hinge, G., Athamena, A., Heddam, S., Gad, M., Elsherbiny, O., Elsayed, S., Eid, M.H., Ibrahim, H. (2023). Application of water quality indices, machine learning approaches, and GIS to identify groundwater quality for irrigation purposes: A case study of Sahara Aquifer, Doucen Plain, Algeria. Water, 15(2): 289. https://doi.org/10.3390/w15020289
[25] Gad, M., Saleh, A.H., Hussein, H., Elsayed, S., Farouk, M. (2023). Water quality evaluation and prediction using irrigation indices, artificial neural networks, and partial least square regression models for the Nile River, Egypt. Water, 15(12): 2244. https://doi.org/10.3390/w15122244
[26] Omeka, M.E. (2024). Evaluation and prediction of irrigation water quality of an agricultural district, SE Nigeria: An integrated heuristic GIS-based and machine learning approach. Environmental Science and Pollution Research, 31(41): 54178-54203. https://doi.org/10.1007/s11356-022-25119-6
[27] Palabıyık, S., Akkan, T. (2024). Evaluation of water quality based on artificial intelligence: Performance of multilayer perceptron neural networks and multiple linear regression versus water quality indexes. Environment, Development and Sustainability, pp. 1-24. https://doi.org/10.1007/s10668-024-05075-6
[28] Kouadri, S., Elbeltagi, A., Islam, A.R.M.T., Kateb, S. (2021). Performance of machine learning methods in predicting water quality index based on irregular data set: Application on Illizi region (Algerian southeast). Applied Water Science, 11(12): 190. https://doi.org/10.1007/s13201-021-01528-9
[29] Benmahamed, Y., Kherif, O., Teguar, M., Boubakeur, A., Ghoneim, S.S. (2021). Accuracy improvement of transformer faults diagnostic based on DGA data using SVM-BA classifier. Energies, 14(10): 2970. https://doi.org/10.3390/en14102970
[30] Zhang, Y., Li, J., Fan, X., Liu, J., Zhang, H. (2020). Moisture prediction of transformer oil-immersed polymer insulation by applying a support vector machine combined with a genetic algorithm. Polymers, 12(7): 1579. https://doi.org/10.3390/polym12071579
[31] Mahmood, O.A., Sulaiman, S.O., Al-Jumeily, D. (2024). Forecasting for Haditha reservoir inflow in the West of Iraq using Support Vector Machine (SVM). PloS one, 19(9): e0308266. https://doi.org/10.1371/journal.pone.0308266
[32] Ehteram, M., Salih, S.Q., Yaseen, Z.M. (2020). Efficiency evaluation of reverse osmosis desalination plant using hybridized multilayer perceptron with particle swarm optimization. Environmental Science and Pollution Research, 27(13): 15278-15291. https://doi.org/10.1007/s11356-020-08023-9
[33] Avendaño-Valencia, L.D., Fassois, S.D. (2015). Natural vibration response based damage detection for an operating wind turbine via random coefficient linear parameter varying AR modelling. Journal of Physics: Conference Series, Ghent, Belgium, 628(1): 012073. https://doi.org/10.1088/1742-6596/628/1/012073
[34] Seyedzadeh, A., Maroufpoor, S., Maroufpoor, E., Shiri, J., Bozorg-Haddad, O., Gavazi, F. (2020). Artificial intelligence approach to estimate discharge of drip tape irrigation based on temperature and pressure. Agricultural Water Management, 228: 105905. https://doi.org/10.1016/j.agwat.2019.105905
[35] Dimple, D., Rajput, J., Al-Ansari, N., Elbeltagi, A. (2022). Predicting irrigation water quality indices based on data-Driven algorithms: Case study in semiarid environment. Journal of Chemistry, 2022(1): 4488446. https://doi.org/10.1155/2022/4488446
[36] Mokhtar, A., Elbeltagi, A., Gyasi-Agyei, Y., Al-Ansari, N., Abdel-Fattah, M.K. (2022). Prediction of irrigation water quality indices based on machine learning and regression models. Applied Water Science, 12(4): 76. https://doi.org/10.1007/s13201-022-01590-x
[37] Derdour, A., Abdo, H.G., Almohamad, H., Alodah, A., Al Dughairi, A.A., Ghoneim, S.S., Ali, E. (2023). Prediction of groundwater quality index using classification techniques in arid environments. Sustainability, 15(12): 9687. https://doi.org/10.3390/su15129687
[38] Ibrahim, H., Yaseen, Z.M., Scholz, M., Ali, M., et al. (2023). Evaluation and prediction of groundwater quality for irrigation using an integrated water quality indices, machine learning models and GIS approaches: A representative case study. Water, 15(4): 694. https://doi.org/10.3390/w15040694
[39] Elsayed, S., Hussein, H., Moghanm, F.S., Khedher, K.M., Eid, E.M., Gad, M. (2020). Application of irrigation water quality indices and multivariate statistical techniques for surface water quality assessments in the Northern Nile Delta, Egypt. Water, 12(12): 3300. https://doi.org/10.3390/w12123300
[40] Eid, M.H., Elbagory, M., Tamma, A.A., Gad, M., Elsayed, S., Hussein, H., Moghanm, F.S., Omara, A.E.,

Kovács, A., Péter, S. (2023). Evaluation of groundwater quality for irrigation in deep aquifers using multiple graphical and indexing approaches supported with machine learning models and GIS techniques, Souf Valley, Algeria. Water, 15(1): 182. https://doi.org/10.3390/w15010182

[41] Jin, Z.W., Shang, J.X., Zhu, Q.W., Ling, C., Xie, W., Qiang, B.H. (2020). RFRSF: Employee turnover prediction based on random forests and survival analysis. Web Information Systems Engineering—WISE 2020: 21st International Conference, Amsterdam, The Netherlands, Singapore, pp. 503-515. https://doi.org/10.1007/978-3-030-62008-0_35

[42] Pham, B.T., Bui, D.T., Prakash, I., Dholakia, M.B. (2017). Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. Catena, 149: 52-63. https://doi.org/10.1016/j.catena.2016.09.007

[43] Sihag, P., Mohsenzadeh Karimi, S., Angelaki, A. (2019). Random forest, M5P and regression analysis to estimate the field unsaturated hydraulic conductivity. Applied Water Science, 9(5): 129. https://doi.org/10.1007/s13201-019-1007-8

[44] Oshiro, T.M., Perez, P.S., Baranauskas, J.A. (2012). How many trees in a random forest? Machine Learning and Data Mining in Pattern Recognition: 8th International Conference, Berlin, Germany, pp. 154-168. https://doi.org/10.1007/978-3-642-31537-4_13

[45] El Bilali, A., Taleb, A., Brouziyne, Y. (2021). Groundwater quality forecasting using machine learning algorithms for irrigation purposes. Agricultural Water Management, 245: 106625. https://doi.org/10.1016/j.agwat.2020.106625

[46] Elith, J., Leathwick, J.R., Hastie, T. (2008). A working guide to boosted regression trees. Journal of animal ecology, 77(4): 802-813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

[47] Shin, Y., Kim, T., Hong, S., Lee, S., Lee, E., Hong, S.W., Lee, C.S., Kim, T.Y., Park, M.S., Park, J., Heo, T.Y. (2020). Prediction of chlorophyll-a concentrations in the Nakdong River using machine learning methods. Water, 12(6): 1822. https://doi.org/10.3390/w12061822

[48] Raheja, H., Goel, A., Pal, M. (2022). Prediction of groundwater quality indices using machine learning algorithms. Water Practice & Technology, 17(1): 336-351. https://doi.org/10.2166/wpt.2021.120

[49] Nguyen, D.P., Ha, H.D., Trinh, N.T., Nguyen, M.T. (2023). Application of artificial intelligence for forecasting surface quality index of irrigation systems in the Red River Delta, Vietnam. Environmental Systems Research, 12(1): 24. https://doi.org/10.1186/s40068-023-00307-6

[50] Yan, J.Z., Chen, X.Y., Yu, Y.C., Zhang, X.J. (2019). Application of a parallel particle swarm optimization-long short term memory model to improve water quality data. Water, 11(7): 1317. https://doi.org/10.3390/w11071317

[51] Zhang, Z., Kong, L.D., Zheng, L. (2018). Power-type varying-parameter RNN for solving TVQP problems: Design, analysis, and applications. IEEE Transactions on Neural Networks and Learning Systems, 30(8): 2419-2433. https://doi.org/10.1109/TNNLS.2018.2885042

[52] Bi, J., Lin, Y.Z., Dong, Q.X., Yuan, H.T., Zhou, M.C. (2021). Large-scale water quality prediction with integrated deep neural network. Information Sciences, 571: 191-205. https://doi.org/10.1016/j.ins.2021.04.057

[53] Barzegar, R., Aalami, M.T., Adamowski, J. (2020). Short-term water quality variable prediction using a hybrid CNN—LSTM deep learning model. Stochastic Environmental Research and Risk Assessment, 34(2): 415-433. https://doi.org/10.1007/s00477-020-01776-2

[54] Baek, S.S., Pyo, J., Chun, J.A. (2020). Prediction of water level and water quality using a CNN-LSTM combined deep learning approach. Water, 12(12): 3399. https://doi.org/10.3390/w12123399

[55] Sha, J., Li, X., Zhang, M., Wang, Z.L. (2021). Comparison of forecasting models for real-time monitoring of water quality parameters based on hybrid deep learning neural networks. Water, 13(11): 1547. https://doi.org/10.3390/w13111547

[56] Ye, Q.Q., Yang, X.Q., Chen, C.B., Wang, J.C. (2019). River water quality parameters prediction method based on LSTM-RNN model. 2019 Chinese control and decision conference (CCDC), Nanchang, China, pp. 3024-3028. https://doi.org/10.1109/CCDC.2019.8832885

[57] 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. (2020). Implementation of data intelligence models coupled with ensemble machine learning for prediction of water quality index. Environmental Science and Pollution Research, 27: 41524-41539. https://doi.org/10.1007/s11356-020-09689-x

[58] Aldhyani, T.H.H., Al-Yaari, M., Alkahtani, H., Maashi, M. (2020). Water quality prediction using artificial intelligence algorithms. Applied Bionics and Biomechanics, 2020(1): 6659314. https://doi.org/10.1155/2020/6659314

[59] Chen, H., Chen, A., Xu, L., Xie, H., Qiao, H., Lin, Q., Cai, K. (2020). A deep learning CNN architecture applied in smart near-infrared analysis of water pollution for agricultural irrigation resources. Agricultural Water Management, 240: 106303. https://doi.org/10.1016/j.agwat.2020.106303