Kufa Journal of Engineering
Vol. 17, No. 3, July 2026, P.P. 623-636
Article history: Received 29 June 2025, last revised 16 December 2025, accepted 10 January 2026

MODELING OF GROUNDWATER QUALITY INDEX USING MULTILAYER PERCEPTRON (MLP) NEURAL NETWORK IN AL-MAHAWIL DISTRICT, IRAQ

Fatima Abed Alabass Alhusaini[1], Kadhim Naief Kadhim[2], and Hadeel Ali Abdulhussein Al Saleh[3]

[1] Department of Civil Engineering, College of Engineering, University of Babylon, Babylon, Iraq, Emal: 3000fati3000@gmail.com.

[2] Department of Civil Engineering, College of Engineering, University of Babylon, Babylon, Iraq, Emal: eng.kadhim.naief@uobabylon.edu.iq.

[3] Department of Civil Engineering, College of Engineering, University of Babylon, Babylon, Iraq, Emal:eng.hadeel.ali@uobabylon.edu.iq.

https://doi.org/10.30572/2018/KJE/170335

ABSTRACT

Monitoring water quality is essential for environmental protection and sustainable water resource management. This study employed a Multilayer Perceptron (MLP) neural network to model the Groundwater Quality Index (GWQI) based on key physicochemical parameters. The model was developed using MATLAB’s Neural Network Toolbox (nftool), with data divided into training, validation, and testing sets. Performance was evaluated using MSE, RMSE, MAE, and the correlation coefficient (R). The MLP model achieved good predictive accuracy:(MSE = 2.3, R ≈ 1.0) in training;(MSE = 7.6983, R = 0.99992) in validation; and (MSE = 15.8, R = 0.997) in testing. These results confirm the model’s ability to capture nonlinear relationships and generalize well without overfitting. The stability of the model is supported by its error distribution and training convergence. A comparison with previous studies shows improved performance, reinforcing the potential of neural networks in predicting water quality and supporting data-driven environmental decision-making.

KEYWORDS

Groundwater Quality Index (GWQI); Artificial Intelligence; Multilayer Perceptron (MLP); Neural Networks; Prediction; MATLAB.

1. INTRODUCTION

Groundwater is an essential form of fresh water, particularly in arid and semi-arid areas, which are at a high risk of groundwater pollution due to natural and human factors (Gemail & Abd-Elaty,2023; Al Saleh et al.,2024). Lack of control in agricultural activities, industrial effluents, and pollution of wastewater are among the threats to groundwater quality in Iraq, hence a threat to sustainable development in Iraq (Kurwadkar& Nakarmi, 2020). GWQI is the most popular indicator that entails the assessment of groundwater suitability regarding drinking water, combining several compound parameters into a unique composite score. It uses the weighted arithmetic index approach grounded on WHO drinking levels, which can give an overall evaluation of groundwater quality. Nevertheless, the paper points out the unavoidable subjectivity of the traditional indices, such as GWQI. It suggests further developing highly predictive models, such as artificial intelligence, to provide predictions of the groundwater quality (Nadiri et al.,2022) AI techniques have become leading in recent years in groundwater management. The methods have become popular due to being cost-effective and time-saving. They enhance the decision-making process by providing correct analytics modeling, data monitoring, and integration (Borgohain et al., 2025). The multilayer neural networks (MLPs) are among the different artificial intelligence methods employed in groundwater quality research studies, as they have proven to be effective in nonlinear relationships that are difficult to establish between chemical and physical water properties. As opposed to other linear models, MLPs are capable of learning the complex interactions of variables, including pH, electrical conductivity, salinity, and major ions that are frequently not simple in linear patterns. This has enabled them to be a good predictive tool in predicting the quality index of groundwater with high accuracy (Al Saleh, 2021; Miller, et al.,2023). Despite the efficiency of the MLP-based models in the estimation of GWQI, their drawbacks have been experienced in Iraq. The quality of the input data, as well as the selection of the architecture of the model, has generally been overlooked when developing the past works on the accuracy of the predictions (Ibrahim, et al.,2020). The recent modeling on groundwater quality has been intensively applied in machine learning methodology under varying climatic conditions to use the reliable and accurate predictions, like support vector regression (SVR), especially because of its stability under small datasets. Nevertheless, several studies have suggested that SVR falls short of completely capturing the nonlinear relations among physical and chemical variables, particularly when the quantity of data is huge, which is typically necessary to consider the Groundwater Quality Index (GWQI) with proper accuracy. Conversely, the neural networks (MLP) are more flexible in the modeling of complex and nonlinear relationships and also in adapting to large and heterogeneous datasets of the environment (Müller et al., 2021). Thus, an MLP architecture is selected as the main predictive instrument of study since it was shown to be superior to the traditional algorithms in the same circumstances as in the field of research. This gap is bridged by the proposed study that develops an MLP model to forecast GWQI, based on actual measurements of groundwater in the Mahawil region in Iraq. The proposed research is a study to assess the efficiency, precision, and consistency of the model in the comparison of measured and predicted values of the Global Water Quality Index (GWQI). It helps to deliver smart and statistically informed forecasts for making timely decisions and to elevate the incorporation of artificial intelligence approaches in assessing the environment and managing water resources sustainably across Iraq.

2. STUDY AREA

The present study took place in the Al-Mahawil district of the Babil Governorate of central Iraq. The region lies between 20°44.50°E and 25°32.20°N and 49°32.20°N with an estimated area of about 1,667 square kilometers, making up 33 percent of the total size of Babil Governorate. The climate of this region is defined as semi-arid, and the average rainfall is less than 150 mm per year. In summer, the rate of evaporation is great. The area is composed of silty clay, clayey silt, and fine- to medium-grained sand, which are geologically described as having medium to low permeability. Groundwater flows from the northwest to the southeast towards the Euphrates River. Natural factors, such as clay-rich deposits and evaporated minerals (gypsum and halite), contribute to high levels of salinity, sulfates, chlorides, and dissolved solids. (Al-Aarajyet al.,2024). Human activities, including intensive agriculture using chemical fertilizers and untreated industrial and domestic wastewater, further exacerbate water quality degradation. Groundwater is important for drinking water and other activities in the area due to the low surface water flow. (Chabuket al.,2022). Fig. 1 illustrates the geographical boundaries and location of the study area.

Fig. 1. Location of Al-Mahawil District in Iraq /Source: Google Earth)

3. METHODOLOGY

3.1. Groundwater sampling

After a careful examination of the well sites within the research area, samples were taken randomly during the winter rainy season as well as during the arid summer months, Fig.2. Most wells were at least 10 meters deep. The geographic coordinates of the wells were also recorded.

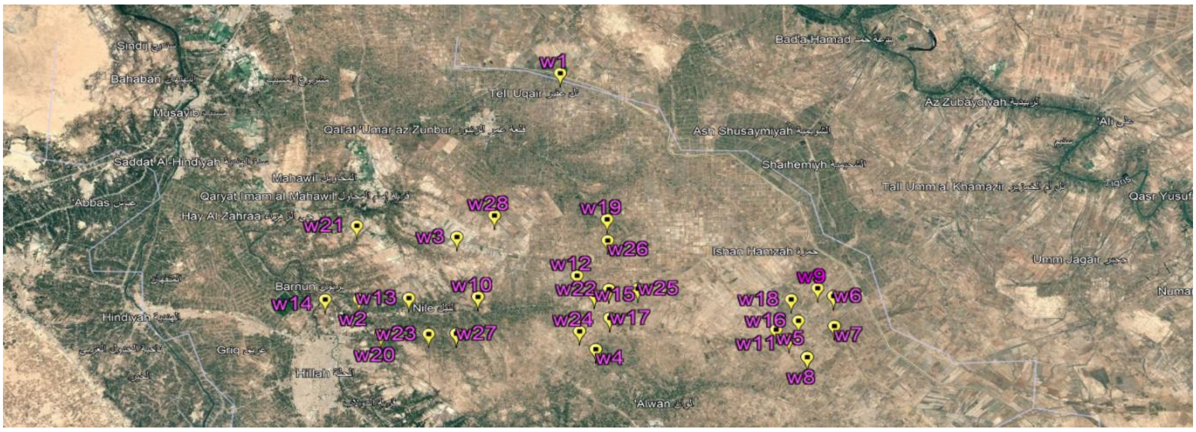

Fig. 2. Shows the spatial distribution of the samples collected /Source: Google Earth)

ten physico-chemical parameters were analyzed for 28 groundwater samples and the parameters analyzed are calcium (Ca), sodium (Na), magnesium (Mg), chloride (Cl), nitrate (NO_3), sulfate (SO_4), bicarbonate (HCO_3)). Electrical conductivity (EC), total dissolved solids (TDS) and hydrogen ion activity (Ph).

3.2. Groundwater quality index (GWQI)

Groundwater quality assessment is a vital step to ensure the sustainability of water resources (Karuppannan & Kawo, 2019). The Groundwater Quality Index (GWQI) has proven to be an effective tool for simplifying complex water quality data into a single value by integrating several chemical and physical parameters, and its calculation is as follows (Al-Saleh, 2021; Mohammed et al.,2023; Saadali et al.,2022).

1- Selecting 10 physiochemical parameters from 28 groundwater wells, which are: (Ca^{+2}),(Na^{+}), (Mg^{+2}), (Cl^{-}), (NO_3^{-}), (SO_4^{-2}), (HCO_3^{-}), (EC), (TDS), and (pH).

2-Initial weights are assigned to each criterion according to its relative importance in determining groundwater quality, and calculated the adjusted relative weight of each criterion using standard Eq.1 as in Table 1:

$$Wi = \frac{wi}{\sum_{1}^{n} wi} \qquad (1)$$

where wi: is the weight allocated to each parameter, and the Wi the relative weight used in subsequent steps.

Table 1 The weights and the relative weights of the physicochemical parameters.

Parameters	Units	Weight (wi)	Relative Weight (Wi)
EC	µS/cm	5	0.15625
TDS	mg/L	5	0.15625
pH	/	1	0.03125
Na^{+1}	mg/L	3	0.09375
Mg^{+2}	mg/L	2	0.0625
Ca^{+2}	mg/L	4	0.125
Cl^{-1}	mg/L	4	0.125
SO_4^{-2}	mg/L	4	0.125
HCO_3^{-1}	mg/L	3	0.09375
NO_3	mg/L	1	0.03125

3- Calculation of the relative quality of each parameter (qi) is calculated using the Eq.2:

$$qi = \frac{ci}{si} \times 100 \tag{2}$$

Where: (ci) the measured concentration with, (si) the standard permissible limit×

4. Sub-index (SI) and GWQI were calculated by the following Eqs.3 and 4.

$$SI = Wi \times qi \tag{3}$$

$$GWQI = \sum_1^n SI \tag{4}$$

Water quality is classified based on the GWQI value into different levels (Dawood et al., 2018), as shown in Table 5.

3.3. Artificial Intelligence Using Multilayer Perceptron (MLP)

Artificial neural networks (ANNs) are computational models inspired by the structure of the human brain, where interconnected nodes or "neurons" enable learning from complex data patterns. These interconnected components collectively determine the model's ability to learn and generalize patterns from data (Qiu, 2023). In this study, the ANN model was implemented using MATLAB's programming environment efficient training and testing, The multilayer perceptron (MLP) has shown capability to handle nonlinear relationships. The model was trained using the" Levenberg-marquadt" algorithm, The model consists of an input layer containing 10 variables representing the chemical, physical of the sample, followed by a hidden layer containing 10 nodes, and then one output layer that predicts the groundwater quality index (GWQI). The nonlinear activation function "(tansig)" was used in the hidden layer to determine complex input-output relationships, while a linear function "(purelin)" was used in the output layer to generate continuous results. as Fig.3.

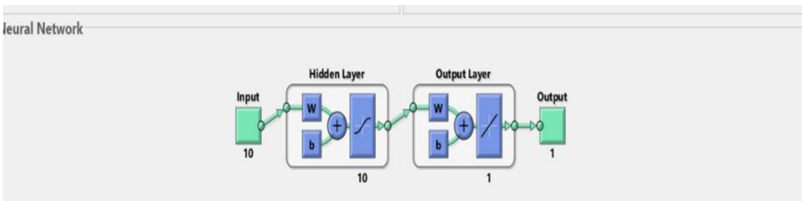

Fig .3. The multilayer neural model (MLP)

3.4. Performance Evaluation Metrics

A variety of statistical measures were used to evaluate the model performance, such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). The indicators are very informative of the prediction accuracy and generalization ability of the model (Kadhim, and Noor, 2018; Abdelhadi et al., 2022).

4. RESULTS AND DISCUSSION

4.1. General hydrochemistry

Physical and chemical characteristics of groundwater are some of the basic tenets in the establishment of groundwater characteristics. (Chabuket al.,2022). Analysis of the groundwater samples in the study area in a laboratory showed a distinct difference in the parameters (EC, TDS, pH, Na, Ca, Mg, HCO_3, Cl, SO_4, and NO_3), Table 2, and Fig.4.

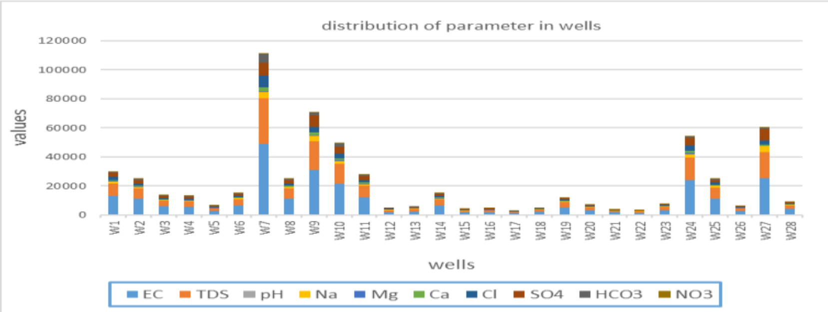

Fig .4. Distribution of parameters in wells

Table 2 Physicochemical parameters

w	EC	TDS	pH	Na	Mg	Ca	Cl	SO4	HCO3	NO3	X	Y	GWQI	pred.-mlp
W1	13060	8442	7.1	1266	544	826	2079	2876	624	1.2	44.694	32.786	669	666
W2	10980	7112	7.1	1179	357	654	1381	2296	1075	0.7	44.464	32.581	552	550
W3	5900	3833	7.2	570	161	335	664	1495	461	0.9	44.575	32.607	299	297
W4	5660	3674	7.2	610	184	350	715	1233	516	1.1	44.889	32.48	288	286
W5	2740	1780	7.2	262	117	186	414	532	233	1.2	44.926	32.498	141	141
W6	6550	4250	7.2	641	242	412	757	1488	557	0.1	44.991	32.533	330	329
W7	49000	31458	7.1	4202	304	3194	7618	9345	5724	0.4	44.991	32.533	2415	2414
W8	10970	7111	7.1	1019	333	607	1425	2449	1070	1.2	44.96	32.466	550	548
W9	30800	20100	7.1	3429	829	1619	3983	7970	1837	1.1	44.973	32.542	1581	1400
W10	21400	13780	7.1	1865	897	1198	2926	4996	2204	0.8	44.596	32.431	1096	1091
W11	12200	7900	7.1	1180	472	747	1566	2597	1143	0.8	44.94	32.487	614	611
W12	1907	1380	8.1	184	54	141	270	509	67	0.01	44.707	32.562	105	105
W13	2480	1700	8.4	246	97	106	319	580	200	0.01	44.52	32.646	128	128
W14	6510	4300	7.7	533	213	461	910	1500	411	5	44.427	32.578	332	328
W15	1749	1180	8.1	115	61	144	200	480	108	0.88	44.979	32.534	94	93
W16	1984	1320	8	115	77	140	210	520	98	0.5	44.951	32.507	101	103
W17	1313	885	7.8	87	47	11	138	379	73	0.6	44.952	32.514	65	65
W18	2120	1411	7.8	228	76	96	249	537	128	0.04	44.944	32.531	110	109
W19	5310	3500	8.1	598	122	261	950	850	311	0.2	44.466	32.623	261	269
W20	3030	2009	7.2	322	86	164	504	437	354	0.06	44.488	32.588	151	150
W21	1633	1120	8	115	38	100	180	320	95	0.1	44.886	32.577	78	84
W22	1459	986	8	93	53	128	145	379	119	2	44.541	32.546	78	81
W23	3300	2190	7.9	352	115	196	639	466	366	4	44.464	32.621	170	170
W24	23800	15800	7.8	2300	645	1604	3901	5043	1100	14.5	44.708	32.517	1210	1209
W25	10970	7800	7.7	1150	204	600	1617	1518	1196	0.2	44.774	32.546	528	527
W26	2630	1885	8.1	277	61	150	312	585	208	1.4	44.738	32.538	135	135
W27	25200	17890	7.7	4140	450	802	2841	7872	933	5.01	44.571	32.657	1369	1367
W28	3920	2620	7.6	288	166	301	638	911	180	30	44.617	32.63	208	121

Table 3 gives the descriptive statistics of the analyzed physicochemical parameters. The mean electrical conductivity turned out to be approximately 9591.96 µS/cm, which is higher than the recommended value of 1500 µS/cm by the World Health Organization (WHO), indicating that high levels of dissolved salts may be present. Otherwise, the mean TDS concentration (6336.29 mg/L) was higher than the allowable concentration (1000 mg/L), indicating that the water is likely influenced by possible sources of pollution, such as agricultural spills or salty geological structures (Chai, et al.,2020). The means of Na, Mg, Ca, Cl, SO_4, HCO_3, and NO_3 (977.357, 250.217, 554.835, 1341.107, 2148.66, 763.96, and 2.642 mg/L, respectively) were also found to have an average that was far higher than the recommended international values, indicating the possibility of geological effects or possible contamination of the groundwater sources (Rashid et al.,2023). Conversely, the average pH was 7.59, and it is in the acceptable range as per the international standard (8.5 maximum).

Table .3 Statistical characteristics of groundwater parameters compared to the limits of the (WHO) (Edition,2017) (Zhu, 2024).

Parameter	Unit	Min	Max	Mean	WHO Standard
EC	US/cm	1313	49000	9591.964	1500
TDS	mg/L	885	31458	6336.286	1000
pH	/	7.12	8.35	7.591071	8.5
Na	mg/L	87.4	4202	977.3571	200
Mg	mg/L	38	897	250.2179	150
Ca	mg/L	11.2	3194	554.8357	200
Cl	mg/L	138.4	7618	1341.107	250
SO_4	mg/L	320	9345	2148.661	250
HCO_3	mg/L	67.1	5724	763.9607	350
NO_3	mg/L	0.01	30	2.6425	50

Table 4 was tested using a Pearson correlation matrix to establish the intensity of the association between the various indices. The results indicated that various parameters were correlated strongly and positively, with the most notable correlation existing between the electrical conductivity (EC) and the total dissolved solids (TDS), with a correlation coefficient of r = 0.999, which was a direct relationship between the two parameters (Ahmad et al.,2021). There were also high correlations between EC and Cl (r = 0.989) as well as between Na and SO_4 (r = 0.988), which shows that these elements play an important role in the determination of the salinity of water (Andrade Foronda and Colinet, 2023). Conversely, the findings demonstrated that there was a high relationship between Ca and HCO 3 (r = 0.943), which could mean that carbonate rocks contributed to the enrichment of the groundwater with these ions (Zhu, 2024). Conversely, nitrate (NO_3^-) was found to have a low correlation with all the other parameters, including its association with EC (r = 0.038), which means it could have its sources in agricultural activity and not geochemical influences.

Table .4 Pearson correlation matrix analysis for the physiochemical parameters

Parameters	EC	TDS	pH	Na	Mg	Ca	Cl	SO4	HCO3	NO3
EC	1.00									
TDS	1.00	1.00								
pH	-0.45	-0.44	1.00							
Na	0.95	0.96	-0.39	1.00						
Mg	0.70	0.70	-0.54	0.70	1.00					
Ca	0.97	0.96	-0.45	0.86	0.64	1.00				
Cl	0.99	0.98	-0.42	0.91	0.65	0.99	1.00			
SO4	0.97	0.98	-0.43	0.99	0.75	0.90	0.93	1.00		
HCO3	0.91	0.90	-0.47	0.78	0.49	0.94	0.93	0.81	1.00	
NO3	0.04	0.04	0.08	0.04	0.10	0.06	0.06	0.06	-0.08	1.00

The Groundwater Quality Index (GWQI) was computed by the weighted calculation method to evaluate the quality of 28 of the samples. The GWQI values were categorized in accordance with the standards by the World Health Organization (WHO) (2017), as described in Table 5. into four distinct categories: 21% were classified within the good water category, 22% are in the poor water class, 14% are in the very poor water, and 43% of the samples are considered unsuitable water. The calculated groundwater quality index values ranged between 64.99 and 2415.3, reflecting a large variation in groundwater conditions throughout the study area as in Fig 5.

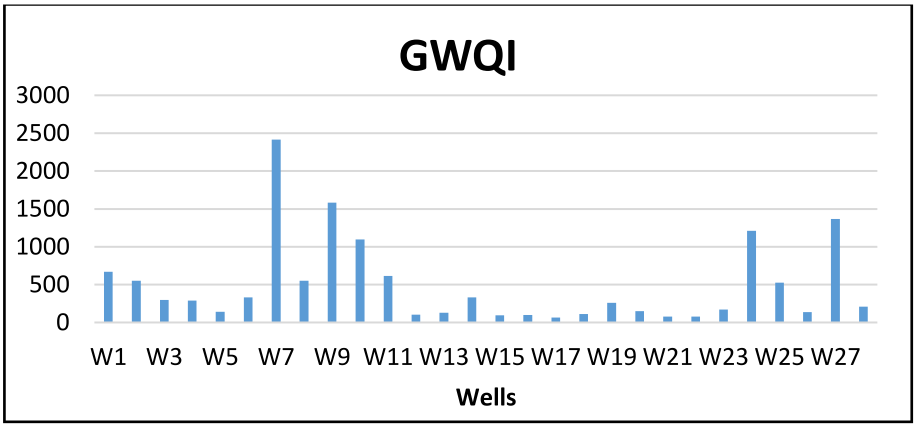

Fig .5. Distribution of GWQI in wells

Table .5. Classification of Groundwater Quality Based on GWQI (WHO, 2017)

GWQI Range	Water Quality Class	Suitability for Use
0 – 50	Excellent	Suitable for drinking without any risk
51 – 100	Good	Minor concern; suitable with minimal treatment
101 – 200	Poor	Not suitable for drinking without treatment
201 – 300	Very Poor	Serious concern; only for agricultural use
> 300	Unsuitable for Drinking	Not recommended for any direct use

4.2. Artificial intelligence models:

The neural model was trained using nftool in MATLAB environment, and the results showed high accuracy in predicting the GWQI based on the input dataset, The measured water specimens were divided into 70% for training, 10% for validation, and 10% for testing. as shown in Table 6. The training data, showed a perfect coefficient of determination $R^2 \approx 1.0$, indicating a perfect match between the actual and forecast values. The mean square error (MSE) = 2.3, RMSE = 1.52, and MAE = 1.22, reflecting the model's very strong performance during training.

Table .6 MLP performance through training and validation phases for WQI forecast

MLP Models	MSE	RMSE	MAE	R^2	Support
MLP training	2.3	1.52	1.22	≈1	20
MLP validation	7.698	2.775	2.22	0.999	4
MLP TEST	15.3	3.91	3.13	0.997	4

In the validation phase, the best performance was recorded in Epoch 5, with MSE = 7.6983, RMSE = 2.775, MAE = 1.2, and The value of the coefficient of determination (R^2) of 0.99992. These values indicate that model was able to generalize well to data not used in direct training, without showing clear signs of overlearning. In the testing phase, the best performance was recorded in Epoch 5, with MSE = 15.3, RMSE = 3.91, MAE = 3.13, and (R^2) of 0.997. This performance is a strong indicator of accurate results when applied to new data. Since $R^2 \approx 1$ was high in training, and R^2 was also high in validation and testing (e.g., 0.999 and 0.997), this is an excellent indication that the model is learning well and that overlearning did not occur (Zhu, 2024). The model's accuracy was analyzed through several graphs. Fig.6 shows a line graph of the variation in actual and predicted GWQI values across all wells, this figure demonstrates that the model is able to accurately track general trends, including peaks and troughs in values. Fig.7: A comparison between the actual and forecast values of the Water Quality Index (GWQI), showing the error between the two values. The results showed good agreement in most wells, with minimal differences between the actual and predicted values. Fig.8: The line between the predicted and actual values is drawn as a scatterplot with a linear regression line in it, which indicates the high consistency and reliability of the performance of the predictive model. Lastly, Fig. 9 shows the working of the Multi-Layer Perceptron Network (MLP) model at the training, validation, and testing stages with a series of evaluation indicators, namely, MSE, RMSE, MAE, and R^2, as well as the number of samples (Support) at each stage. It is observed that the very high accuracy of the model was attained with the value of the coefficient of determination R^2 in the training (approximately 1.0), validation (almost 1.0), and testing (almost 1.0) phases, indicating that the model has a high power to explain the variance in the data and make predictions with a very high level of accuracy. Despite a relative difference in the MSE and MAE values between the phases, these values remain within acceptable limits, especially considering that R^2 reflects a strong correlation between the predicted and actual values. This indicates that the model has good generalization capabilities and is able to handle unseen data.

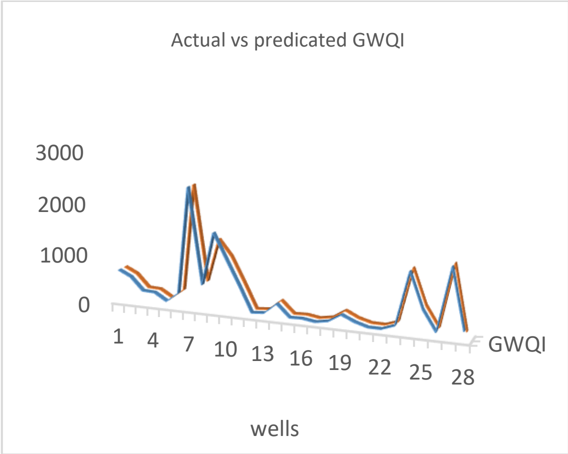

Fig .6. actual vs predicated GWQI

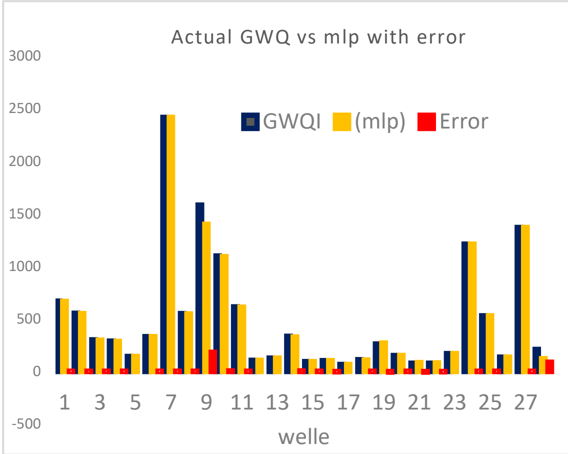

Fig .7. actual vs predicated GWQI with error

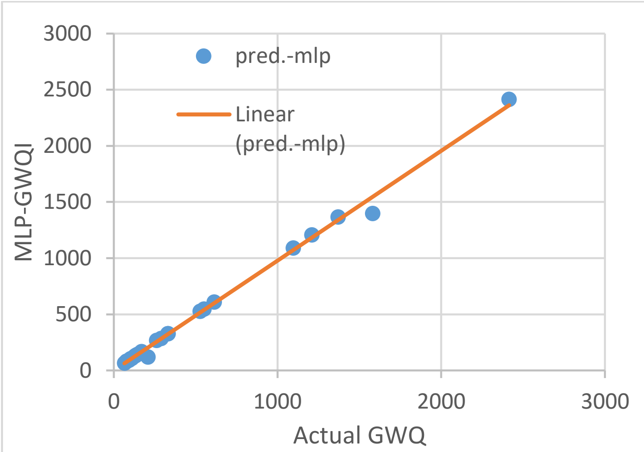

Fig .8. actual vs predicated GWQI

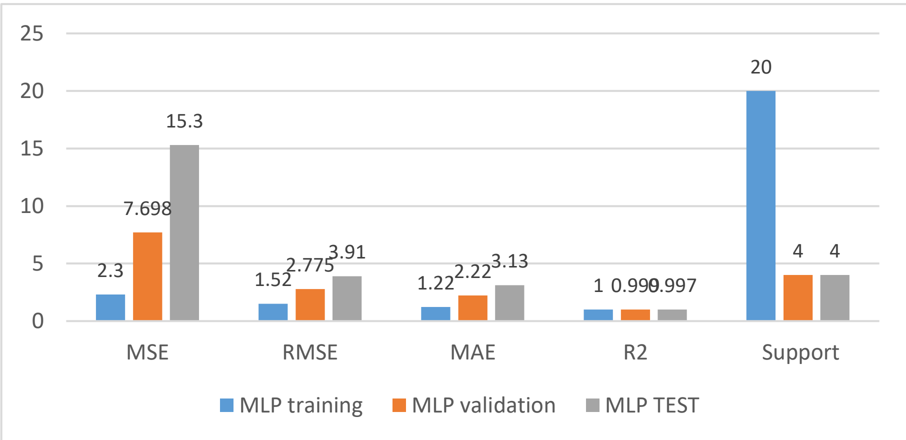

Fig .9. "Comparing the performance of an MLP model across training, validation, and testing phases"

"The patterns predicted by the Multilayer Neural Network (MLP) model for the groundwater quality index can be physically explained by the hydrogeological characteristics of the region. The high levels of trisodium phosphate (TDS), hydrocarbons (EC), sodium (Na), chlorine (Cl), calcium (Ca), magnesium (Mg), and sodium (SO_4) in various locations are in agreement with the natural dissolution of the evaporate minerals, especially gypsum and halite. These are the geological activities that cause an increment of the groundwater salinity (El Yousfi et al.,2022). Alterations in bicarbonate ion (HCO_3) and pH are determined by carbonate and buffering reactions, which are part of the characteristics of shallow groundwater due to the interaction of soil and water. Intensive agricultural practices cause an increase in nitrate (NO_3) levels, which are explained by the application of fertilizers along with intensive irrigation, which causes a leaching of nitrate into the shallow groundwater (Sanad et al.,2025).

The validity of the predictions of the MLP model can be explained by this physical description, where the outputs of the model can be used to show that the hydrochemical processes that took place in the aquifer are reflected in the model. Al-Mahawil.

5. CONCLUSION:

1- The variability in the quality of water necessitates the use of composite indexes to determine the water quality in terms of the Groundwater Quality Index (GWQI) as developed by the World Health Organization (WHO).

2- The model showed good performance on the training set with low MSE, RMSE, and MAE, and a perfect coefficient of determination ($R^2 = 0.0$), which indicates the accuracy of learning and prediction, which is done using the available data.

3- The model was also found to be highly accurate during the validation phase, with an $R^2 = 0.999$, which means that it has high generalization and prediction capabilities using previously unknown data.

4- The test set showed that the model works well, and the coefficient of determination was very high ($R^2 = 0.999$), which is a value that indicates that the model is accurate in prediction when using a completely new set of data.

5- The network was not over trained, which implies that there was a balance between the learning phase using data as inputs and the generalization to unknown data.

6. LIMITATION

The research is restricted by a rather small number of groundwater samples in the transformer location, which raises questions about the accuracy and generalization of the GWQI model. The sampling range should be increased in future research to increase the accuracy, reliability, and statistical significance of the results of the model.

7. REFERENCES

Abdelhadi, A., Zainudin, S. and Sani, N.S. (2022) 'A regression model to predict key performance indicators in higher education enrollments', International Journal of Advanced Computer Science and Applications, 13(1) https://doi.org/10.14569/ijacsa.2022.0130156.

Ahmad, W., Zubair, A., Abbasi, H.N. and Nasir, M.I. (2021) 'Water study of physical, chemical and heavy metals parameters in River Indus and its tributaries, Sindh, Pakistan', Pakistan Journal of Scientific & Industrial Research Series A: Physical Sciences, 64(2), pp. 103-109. https://doi.org/10.52763/PJSIR.PHYS.SCI.64.2.2021.103.109.

Al-Aarajy, K.H., Sultan, M.A. and Hassan, Z.H. (2024) 'Determination of the groundwater potential zones in Babylon using remote sensing and GIS techniques', The Iraqi Geological Journal, 31 March, pp. 261-275. https://doi.org/10.46717/igj.57.1c.17ms-2024-3-29.

Al-Hessnawi, J.K., Jawad, Z.S. and Fadhil, F.A. (2018) 'Spatial analysis of natural properties in Mahaweel District', Journal of University of Babylon, 26(8).

Al-Saleh, H. (2021) 'Applying ANN-based PSO algorithm for the prediction of DO and PO4 in Al-Hillah River', International Review of Civil Engineering (IRECE)*, 12(6), pp. 371-381. https://doi.org/10.15866/irece.v12i6.20092.

Al Saleh, H.A. (2021) 'Comparison study of water quality produced from a conventional water treatment plant and a number of compact units', Journal of Ecological Engineering, 22(5), pp. 258-262. https://doi.org/10.12911/22998993/135864.

Al Saleh, H.A. and Nehaba, S.S. (2024) 'Evaluating groundwater vulnerability and assessing its quality for sustainable management', South African Journal of Chemical Engineering, 50, pp. 291-298. https://doi.org/10.1016/j.sajce.2024.09.001.

Andrade Foronda, D. and Colinet, G. (2023) 'Prediction of soil salinity/sodicity and salt-affected soil classes from soluble salt ions using machine learning algorithms', Soil Systems, 7(2), Article 47. https://doi.org/10.3390/soilsystems7020047.

Borgohain, L., Gogoi, M., Dutta, J. and Medhi, S. (2025) 'Emerging trends and technologies for conservation and sustainable approach in groundwater management', in Emerging Trends and Technologies in Water Management and Conservation*, pp. 175-202. https://doi.org/10.4018/979-8-3693-6920-3.ch006.

Chabuk, A., Jahad, U.A., Majdi, A., Isam, M., Al-Ansari, N., Majdi, H., Laue, J. and Abed, S.A. (2022) 'Creating the distribution map of groundwater for drinking uses using physio-chemical variables; Case study: Al-Hilla City, Iraq', Water, Air, & Soil Pollution, 233(6), 218. https://doi.org/10.1007/s11270-022-05660-3.

Chai, Y., Xiao, C., Li, M. and Liang, X. (2020) 'Hydrogeochemical characteristics and groundwater quality evaluation based on multivariate statistical analysis', Water, 12(10), Article 2792. https://doi.org/10.3390/W12102792.

Dawood, A.S., Jabbar, M.T., Al-Tameemi, H.J. and Baer, E.M.D. (2022) 'Application of water quality index and multivariate statistical techniques to assess and predict groundwater quality with aid of geographic information system', Journal of Ecological Engineering, 23(6), pp. 189-204. https://doi.org/10.12911/22998993/148195.

El Yousfi, Y., Himi, M., El Ouarghi, H., Aqnouy, M., Benyoussef, S., Gueddari, H., Ait Hmeid, H., Alitane, A., Chaibi, M., Zahid, M., Essahlaoui, N., Hitouri, S., Essahlaoui, A. and Elaaraj, A. (2022) 'Assessment and prediction of the water quality index for the groundwater of the Ghiss-Nekkor (Al Hoceima, Northeastern Morocco)', Sustainability, 15(1), 402. https://doi.org/10.3390/su15010402.

Fu, X., Dong, Z., Gan, S., Wang, Z. and Wei, A. (2021) 'Groundwater quality evaluation for potable use and associated human health risk in Gaobeidian City, North China Plain', Journal of Chemistry, 2021(1), Article ID 3008567. https://doi.org/10.1155/2021/3008567.

Gemail, K.S. and Abd-Elaty, I. (2023) 'Unveiling the hidden depths: A review for understanding and managing groundwater contamination in arid regions', Groundwater Quality and Geochemistry in Arid and Semi-Arid Regions, 15 December, pp. 3-5. https://doi.org/10.1007/698_2023_1049.

Ibrahim, M.A., Mohammed-Ridha, M.J., Hussein, H.A. and Faisal, A.A. (2020) 'Artificial neural network modeling of the water quality index for the Euphrates river in Iraq', *The Iraqi Journal of Agricultural Science, 51(6), pp. 1572-1580. https://doi.org/10.36103/IJAS.V51I6.1184.

Kadhim, K.N. and Noor, A. (2018) 'Geospatial technology for groundwater quality parameters assessment in Al-Kifl District, Babylon-Iraq'.

Karuppannan, S. and Kawo, N.S. (2019) 'Groundwater quality assessment using geospatial techniques and WQI in north east of Adama Town, Oromia region, Ethiopia', Hydrosphere Analysis, 3(1), pp. 22-36. https://doi.org/10.21523/GCJ3.19030103.

Kurwadkar, S., Kanel, S.R. and Nakarmi, A. (2020) 'Groundwater pollution: Occurrence, detection, and remediation of organic and inorganic pollutants', Water Environment Research, 92(10), pp. 1659-1668. https://doi.org/10.1002/WER.1415.

Miller, T., Cembrowska-Lech, D., Kisiel, A. and Krzemińska, A. (2023) 'Advancing water quality monitoring through artificial neural networks: Present insights and future opportunities in scientific exploration', Inter Conf, 32(151), pp. 399-409. https://doi.org/10.51582/interconf.19-20.04.2023.041.

Mohammed, M.A., Khleel, N.A., Szabó, N.P. and Szűcs, P. (2023) 'Modeling of groundwater quality index by using artificial intelligence algorithms in northern Khartoum State, Sudan', Modeling Earth Systems and Environment, 9(2), pp. 2501-2516. https://doi.org/10.1007/s40808-022-01638-6.

Müller, J., Park, J., Sahu, R.K., Varadharajan, C., Arora, B., Faybishenko, B. and Agarwal, D. (2021) 'Surrogate optimization of deep neural networks for groundwater predictions', Journal of Global Optimization, 81(1), pp. 203-231. https://doi.org/10.1007/S10898-020-00912-0.

Nadiri, A.A., Barzegar, R., Sadeghfam, S. and Rostami, A.A. (2022) 'Developing a data-fused water quality index based on artificial intelligence models to mitigate conflicts between GQI and GWQI', Water, 14(19), 3185. https://doi.org/10.3390/w14193185.

Qiu, X. (2023) 'The application of artificial intelligence - artificial neural networks - in wastewater treatment', E3S Web of Conferences, 393, 03003. https://doi.org/10.1051/e3sconf/202339303003.

Rashid, A., Ayub, M., Ullah, Z., Ali, A., Sardar, T., Iqbal, J., Gao, X., Bundschuh, J., Li, C., Khattak, S.A. and Ali, L. (2023) 'Groundwater quality, health risk assessment, and source distribution of heavy metals contamination around chromite mines', International Journal of Environmental Research and Public Health, 20(3), Article 2113. https://doi.org/10.3390/ijerph20032113.

Saadali, R., Dadach, M., Bouguettaya, S. and Benjeddi, S. (2022) 'Evaluation of the groundwater quality in Amizour Plain (North Algeria) through the application of heavy metal pollution index', Advanced Research in Life Sciences, 6(1), pp. 40-47. https://doi.org/10.2478/arls-2022-0037.

Sanad, H., Moussadek, R., Zouahri, A., Oueld Lhaj, M., Mouhir, L. and Dakak, H. (2025) 'AI and machine learning-based spatial modeling of groundwater quality indices and hydrogeochemistry for accurate prediction of seawater intrusion and irrigation sustainability in coastal agroecosystems'. https://doi.org/10.21203/rs.3.rs-7705609/v1.

World Health Organization (2017) Guidelines for drinking-water quality: Fourth edition incorporating the first addendum. Geneva: World Health Organization.

Zhu, Z. (2024) 'Systematic optimization of overfitting problem in machine learning', Highlights in Science, Engineering and Technology, 111, pp. 353-359. https://doi.org/10.54097/3tkzrj84.