

Civil Engineering Journal

Vol. 4, No. 12, December, 2018



Artificial Neural Network Model for the Prediction of Groundwater Quality

Basim H. Khudair^{a*}, Mustafa M. Jasim^b, Awatif S. Alsaqqar^c

^a Assistant Professor, Department of Civil Engineering, University of Baghdad, Baghdad, Iraq.
^b Assistant Lecturer, Department of Civil Engineering, University of Baghdad, Baghdad, Iraq.
^c Assistant Professor, Department of Civil Engineering, Uruk University, Baghdad, Iraq.
Received 29 July 2018; Accepted 20 November 2018

Abstract

The present article delves into the examination of groundwater quality, based on WQI, for drinking purposes in Baghdad City. Further, for carrying out the investigation, the data was collected from the Ministry of Water Resources of Baghdad, which represents water samples drawn from 114 wells in Al-Karkh and Al-Rusafa sides of Baghdad city. With the aim of further determining WQI, four water parameters such as (i) pH, (ii) Chloride (Cl), (iii) Sulfate (SO4), and (iv) Total dissolved solids (TDS), were taken into consideration. According to the computed WQI, the distribution of the groundwater samples, with respect to their quality classes such as excellent, good, poor, very poor and unfit for human drinking purpose, was found to be 14.9 %, 39.5 %, 22.8 %, 6.1 %, and 16.7 %, respectively. Additionally, to anticipate changes in groundwater WQI, IBM® SPSS® Statistics 19 software (SPSS) was used to develop an artificial neural network model (ANNM). With the application of this ANNM model, the results obtained illustrated high prediction efficiency, as the sum of squares error functions (for training and testing samples) and coefficient of determination (R²), were found to be (0.038 and 0.005) and 0.973, respectively. However, the parameters pH and Cl influenced model prediction significantly, thereby becoming crucial factors in the anticipation carried out by using ANNM model.

Keywords: Assessment and Prediction; Groundwater Quality; Human Health; Water Quality Index; Artificial Neural Network Model.

1. Introduction

For sustaining life, safe drinking water is one of the fundamental needs of every human on earth, due to which it must be available to all people in an adequate amount via safe and accessible ways of water supply. The increased demand for clean drinking water draws attention towards the management of groundwater quality. Especially in developing countries, where the issues regarding the accessibility of clean water have become acute, the groundwater quality assessment is imperative. Approximately, one-third of the world's populations have been found to use groundwater for drinking. With an added advantage of being extracted at several places, the groundwater makes the transportation via pipe redundant. Generally, water having a constant composition is hygienically reliable, due to which the water can be distributed occasionally, without any treatment. However, simple and cheap treatment (e.g. disinfection) is often inevitable. On the other hand, alteration in the groundwater quality caused due to natural substances, in addition to, anthropogenic activities in the surrounding soil can show its repercussions on the public health if left untreated [16]. Recently, in Iraq, the flow rates of surface waters has decreased significantly, owing to the policy of the surrounding countries of the source, which in turn has led to the shortage of water required for irrigation and drinking. Therefore, there is an urgent need to evaluate the groundwater quality for ensuring appropriate use of groundwater resources for

doi http://dx.doi.org/10.28991/cej-03091212

© Authors retain all copyrights.

^{*} Corresponding author: dr.basimal-obaidy@coeng.uobaghdad.edu.iq

> This is an open access article under the CC-BY license (https://creativecommons.org/licenses/by/4.0/).

obtaining clean drinking water In the present investigation, quality of drinking groundwater available in rural areas in Tabriz, Iran, was studied [14]. For comparing and evaluating water quality, multivariate statistical techniques, including principal component analysis (PCA) and hierarchical cluster analysis (CA), were employed [17].

Further, the result obtained illustrated the presence of high salt concentrations in the eastern and northern parts of the studied area, as the water was described to be hard water, with the level of CaCO3 reaching up to 476 mg/L. On the contrary, in the western parts, arsenic contamination was observed (69 μ g/L), with the mercury being above the standard limits. Additionally, suitability of groundwater for different purposes in Dun valley, central Nepal, was also evaluated. Apart from the nitrate concentration, which was present above the permissible limit (>50 mg/L) in 53 % of the samples, all other physicochemical parameters were within the permissible limits. According to the groundwater quality index (GWQI), all the samples were found to be suitable for agricultural and industrial purposes [4]. Nevertheless, due to high nitrate concentration, about 50 % of samples were unfit for drinking purpose. [23] Further, the groundwater quality was assessed for irrigation in the coastal area (Skhirat region) Morocco. From the results, Na and Ca cations along with Cl and SO4 anions was found to be present predominantly. The study also reveals that irrigation water has high salinization risk, while alkalinization risk ranges from low to medium. In other words, these waters have a high risk of getting toxified by chloride ions. In a nutshell, groundwater in the Skhirat region was established to be of good quality and is highly suitable for irrigation [19].

Similarly, the quality of groundwater in a rural community in North Central of Nigeria was determined. The results demonstrated the contamination of water samples by NO3, Mg, and Pb, with concentrations exceeding the WHO permissible limits [8]. Mainly, anthropogenic sources such as domestic wastewater and poor waste disposal to natural sources such as mineral dissolution from the clayey aquifer lead to high concentrations of these contaminants in water samples. Thus, the groundwater becomes highly acidic with further being unsuitable for consumption. Additionally, microbial contamination by Enterobactererogenes and Escherichia coli was also detected [17]. For evaluating the groundwater quality, the water samples were also collected, in pre-monsoon and post-monsoon seasons, from Andhra Pradesh, India [15]. According to the total dissolved solids classification, the majority of the groundwater samples collected were drawn, from the areas which require groundwater for drinking as well as for irrigation purposes. Upon evaluating, 85 and 89% of the samples were categorized, as very hard water, which further requires softening processes for becoming suitable for domestic use. On the other hand, 48 and 42 % of the samples contain sodium (200 mg/L), while 34 and 41 % of the samples have nitrate content (50 mg/L), both being above the permissible limits, making the samples unfit for human consumption. However, the average water quality index illustrated that the majority of the samples are good with further qualifying to be used as drinking water. Likewise, various irrigation indices demonstrated the samples to be reliable with acceptable application in agricultural activities. Moreover, the LSI and RSI values classify the groundwater of the area as very aggressive with being prone towards substantial corrosion [20].

Considering the peri-urban agglomeration, the situation of the groundwater quality of Ranchi city, India, was estimated. The samples from this area, consisted of highly toxic metals viz., As, Ni, Mn, and Se, with their concentrations (ppm) falling in the range of 0.0 to 200, 0.0 to 80, 0.0 to 4200, and 30 to 140, respectively. According to the WHO standards for drinking purposes, these values are above acceptable limits. In this context, the Water Quality Index (WQI) further describes the poor water quality of 80% of sampling locations, followed by very poor water quality of 7 %, with 4 % being unfit for drinking. Only 9 %, of all samples, have good water quality. Proceeding further, the groundwater quality of Al-Khadhimiya city of Iraq [9], collected from 13 wells during four seasons, was assessed using several quality parameters. From the obtained results, the quality of groundwater was found to be unfit for human drinking purpose, where the major cause of groundwater quality deterioration was; turbidity, hardness, Cl, SO4, Ca, and Mg present in the water [1]. Further, for the data collected from 29 wells located in Zubair, Safwan, and Um-Qasr of Basrah governorate, WQI was applied, whose value varied from poor to unfit making the groundwater of the studied area unsuitable for human drinking purpose. Therefore, the current research aims to evaluate the quality of groundwater for drinking purposes based on Water Quality Index (WQI) in Baghdad City, followed by the development of an artificial neural network model using IBM® SPSS® Statistics 19 software (SPSS) [3], for predicting the changes in groundwater WQI [7]. Figure 1 describes the schematic diagram of research methodology.

2. Material and Method

2.1. Case Study Description

The case study was carried out in Baghdad city, located in the Middle East region of Iraq between the latitude of 33°20′26.09″ in the north and longitude of 44°24′3.17″ of the east with an estimated area of 204.2 km². Further, according to the availability of groundwater in the area, the samples were collected, from 114 wells with depths varying from (0.2-7) m. The wells selected for sample collection were distributed, over different districts in Al-Karkh and Al-Rusafa sides of Baghdad City. Table 1 represents the possibility of laboratory testing of the chemical parameters related to groundwater.

Location	pH	Cl (mg/L)	$SO_4 \left(mg/L \right)$	TDS (mg/L)
Al- Rasafa	-			
Minimum	6.4	125	120	550
Maximum	11.3	21000	12000	30000
Average	7.45	3290	3632	8215
Al-Karkh				
Minimum	6.7	125	420	26300
Maximum	8.6	24000	7100	650
Average	7.35	1818	2289	4733
Testing standard	BS 1377 Part 3:1990 [5]			

Table 1. Chemical parameters of groundwater samples in Baghdad city

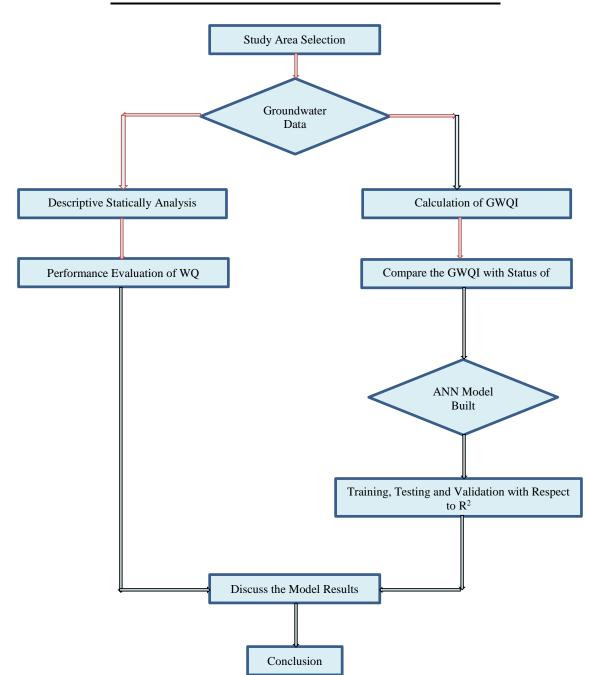


Figure 1. Schematic diagram of research methodology

2.2. Groundwater Quality Index (GWQI)

Upon evaluating several quality parameters of water, groundwater quality index (GWQI) provides a single number, which signifies the overall groundwater quality at a certain time and location. In this way, WQI turns complex data into a simple indicator of groundwater quality, with the intention of making it usable and understandable by the public [22, 13]. WQI have been employed in several studies related to groundwater quality assessment, for classifying the suitability of groundwater for various purposes [17, 20]. In the present study, the limits for drinking water of Iraqi standard specification [10] and weighted arithmetic index method [6] have been used, for calculating WQI [13].

• Proportionality constant (K) was calculated using the following equation:

$$K = \frac{1}{\sum_{j=0}^{n} \frac{1}{S_i}}$$
(1)

Where, S_i and n denotes the standard value for the ith parameter and number of parameters respectively.

• For calculating the unit weight of the ith parameter (Wi) Equation 2 was used (shown below):

$$W_i = \frac{K}{S_i}$$
(2)

• To calculate sub-index or rating of the ith parameter (Qi), a number that reflects the ith parameter relative value in polluted groundwater corresponding to its allowable standard value, Equation 3 (shown below) was employed:

$$Q_i = \frac{100 V_i}{S_i}$$
(3)

Where, V_i represents the ith parameter monitored value. While, the sub index for pH (Q_{pH}) was calculated according to the expression:

$$Q_{pH} = 100 \frac{[V_i - S]}{[S_i - S]} \text{ if } pH > 7$$
 (4)

$$Q_{pH} = 100 \frac{[S - V_i]}{[S_i - S]} \text{ if } pH < 7$$
 (5)

Where, S is equal to 7, which is the neutral value of pH.

• WQI was calculated using Equation 6:

$$WQI = \frac{\sum_{j=1}^{n} Qi W_{i}}{\sum_{j=1}^{n} W_{i}}$$
(6)

Table 2 displays the classes of water quality based on the computed WQI using a scale ranging from excellent to unfit for human drinking. The table reveals that 39 % of the tested samples correspond to good water quality, while more than 44 % falls in the range of poor to unfit for human use. Table 3 depicts the standard limits of water parameters along with their unit weights. Table 4 describes an example for the calculation of WQI for the well Al-Watheq Sq. in Al-Rusafa side of Baghdad city.

Table 2. WQI and status of water quality [6]

WQI value	Water quality status	Groundwater samples %
0-25	Excellent	14.91228
26-50	Good	39.47368
51-75	Poor	22.80702
76-100	Very poor	6.140351
>100	Unfit for human drinking purpose	16.66667

Parameters	Limits (S _i) ICOSQ,2009	$1/\mathbf{S_i}$	Κ	Unit Weight (W _i)
pH	8.5	o.117647		0.949
Cl (mg/L)	350	0.00285		0.023
SO ₄ (mg/L)	400	0.0025	8.065	0.02
TDS (mg/L)	1000	0.001		0.008
		$\Sigma = 0.123997$		$\Sigma = 1$

Table 3. Standard limits for quality parameters and unit weights

Parameters	Data (Vi)	$\mathbf{Q}_{\mathbf{i}}$	W_iQ_i
pН	7.1	6.667	6.327
Cl (mg/L)	750	214.286	4.929
SO ₄ (mg/L)	2475	618.750	12.375
TDS (mg/L)	4600	460.000	3.680
	WQI		$\Sigma = 27.310$
	Class		Good

2.3. Artificial Neural Network Model (ANNM)

In a manner similar to the human brain operation where neurons receive input signals and produce output signals, an output data could be predicted, using a network of artificial neurons, from input data [12]. Further, for modelling water quality, investigation of the mathematical relationship between the independent variables (chemical parameters of water quality) and the dependent variable (WQI), was carried out by training and testing the model from past data. Moreover, to utilize the gained knowledge from the input data, for future WQI predictions, it was stored in the network.

2.3.1. Model Description

Usually, the artificial neurons of the network are being expanded to different layers, while showing a connection with each other via a weight, determined when the error of the sample data between the actual and predicted outputs is minimum [18]. Figure 2 illustrates a simple explanation of the network layers, in addition to, its mathematical function in which (f) is the activation function.

2.3.2. Model Validation

For measuring the difference between observed and predicted values, one can use the sum of the square error function. When the error function is low, the prediction efficiency of ANNM was found to be high, because the network tries to minimize it during training (IBM® SPSS® Statistics 19 User Gide). Alternatively, Khan et al. (2010) explored another statistical parameter R2, the coefficient of determination. The values of R2 vary from 0.0 to 1.0, thereby further indicating that higher the R2 value, the better the ANNM fits the input data [11].

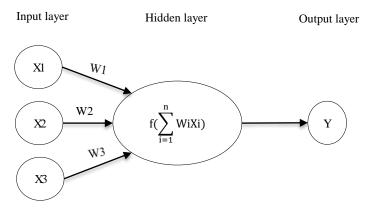


Figure 2. Structure of the neural networks

3. Results and Discussion

3.1. Groundwater Quality Parameters

For determining the quality of groundwater, the chemical parameters such as pH, Cl, SO4, and TDS, were analyzed.

3.1.1. PH

PH being the most vital parameters of water quality indicates the presence of alkaline or acidic material in the groundwater. In Al-Rasafa side, the pH value ranges from 6.4 to 11.3, with the highest pH values of 11.3, 10.5, and 9.6 in Karadat Maryam, Al-Naser, and Al-Nahda, respectively (see Figure 3). Thus, the pH of the area under investigation seems to be within the desirable limits specified for drinking water.

3.1.2. Chloride (Cl)

Generally, Cl extensively disseminates in all types of rocks. Further, high concentrations of Cl in groundwater samples implies to the existence of high organic content in it, thereby making it unsuitable for drinking and livestock watering [9]. For many regions in Baghdad City, the chloride concentration reaches a value of 24000 mg/L in Al-Karkh side, which is far away from the standard value of 350 mg/L (see Figure 4).

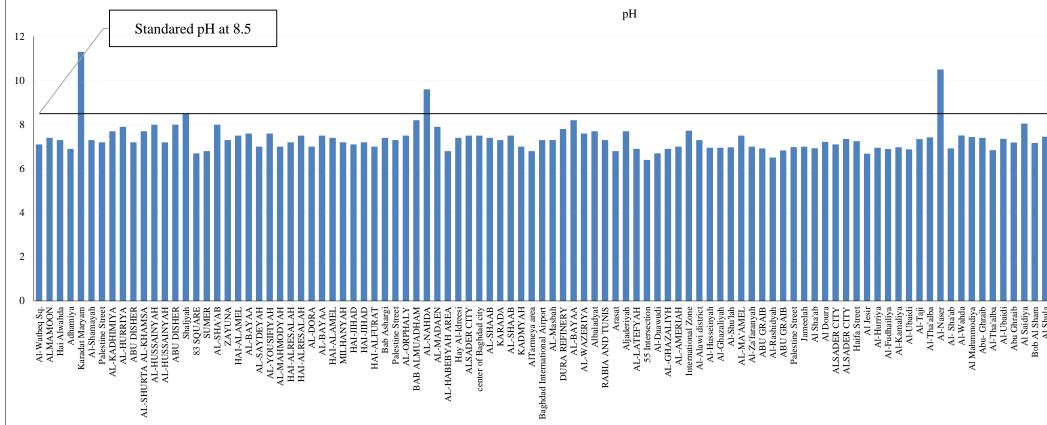
3.1.3. Sulfate (SO₄)

Natural water consists of approximately 50 mg/l of sulfate. Further, WHO, 2011 has mentioned that drinking water having a high concentration of sulfate may lead to corrosion in pipe networks used for its distribution accompanying a noticeable taste. Commonly, the recommended value of sulfate for human consumption is 400 mg/l. The present study illustrates that the concentrations of sulfate in groundwater reach a maximum value of 12000 mg/L in A-Rasafa side, due to the gypsum content present in the soil [2]. In this direction, Figure 5 depicts the majority of the groundwater samples having high SO4 concentration.

3.1.4. Total Dissolved Solids (TDS)

In humans, gastrointestinal irritation can occur due to high TDS concentrations. Additionally, consuming water with high TDS for long-term may lead to several heart diseases and kidney stones. Anthropogenic sources such as solid waste dumping, domestic sewage, and agricultural activities, may be one of the basic reason for such high values of TDS. In most of the groundwater samples, high TDS values were measured, with an average of 8215 mg/L in Al-Rasafa and 4733 mg/L in Al-Karkh sides. Further, Figure 6 represents the variation in TDS in groundwater samples in Baghdad City. Moreover, the figure reveals that in the case of the majority of the samples, the TDS value exceeds the permissible limit of 1000 mg/L, mentioned for drinking water.





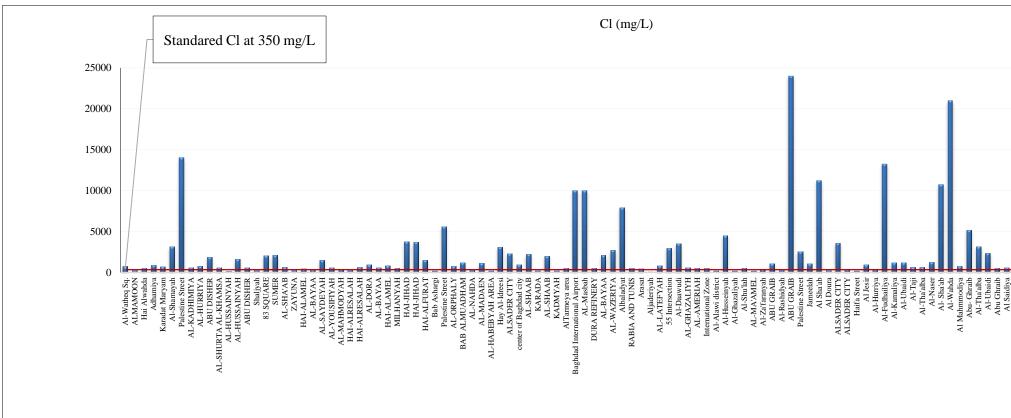
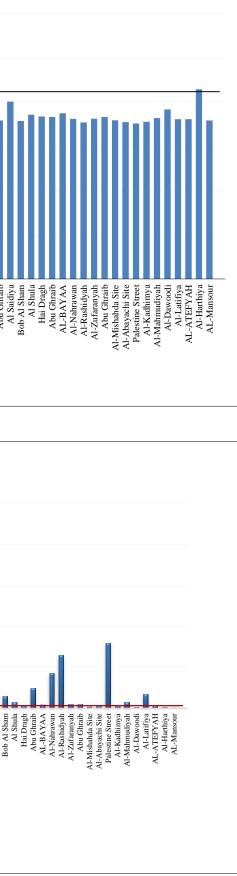
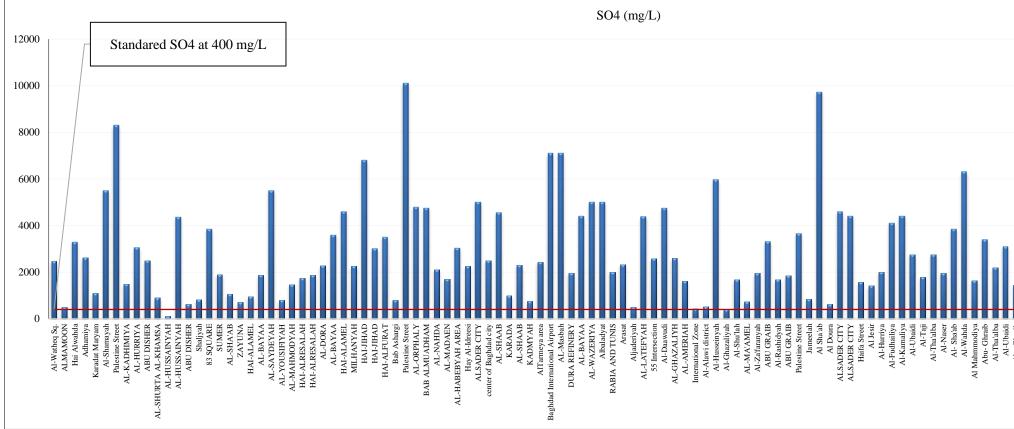
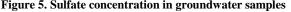


Figure 3. pH Variation with location

Figure 4. Chloride concentration in in groundwater samples







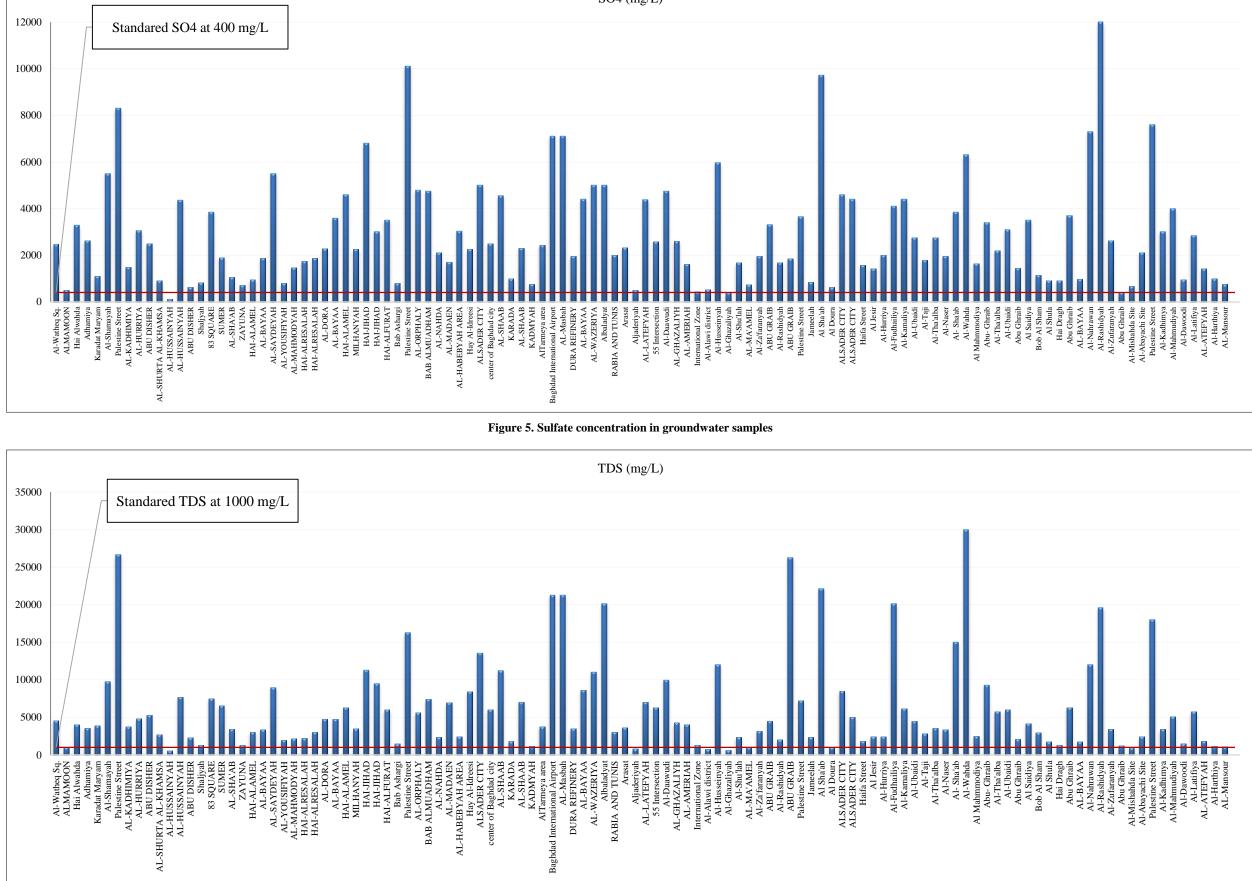


Figure 6. TDS Concentration of in groundwater samples

3.2 Artificial Neural Network Model (ANNM)

3.2.1. Model Training

For developing ANNM, the procedure used in this study is as follows:

3.2.1.1. Partition Dataset

Model training method was employed for calculating the model structure (connection weights and the numbers of hidden neurons) with the task being accomplished by partitioning the input data into training, testing, and holdout samples. Further, the set of data of the holdout sample were not considered, for building the model (IBM® SPSS® Statistics 19 User Gide). In this research, the best partition dataset depends on the sum of squares error for training and testing data as shown in Figure 7. This figure demonstrated that the best partition dataset was found to have values 65, 20, and 15 % corresponding to training, testing, and holdout data, respectively. Nevertheless, with the larger training sample, the value of the error function of the testing sample decreases until it reaches the best value at 0.142, followed by a subsequent increase.

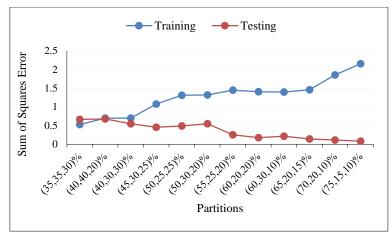


Figure 7. Sum of squares error with different partitions

3.2.1.2. Activation Function

The attainment of best partition dataset subsequently follows the investigation of the effect of changing the activation function of the output layer on the performance of the model. The automatic architecture tested the hyperbolic tangent and identified the sigmoid functions as shown in Figure 8. Among other functions, the sigmoid activation function was found to have the smallest error for both training and testing.

3.2.1.3. Number of Hidden Neurons

The ability of the model to predict the output data gets affected by the number of hidden neurons. With a few hidden neurons, the model becomes incapable of learning sufficiently; thereby under-fitting of the model occurs. On the other hand, too many hidden neurons make the model lose its generalization ability leading to the over-fitting of the model [21]. Further, as indicated by Figure 9, 10.0 hidden neurons were the best number as they gave the smallest error of 0.038 and 0.005, for training and testing data, respectively.

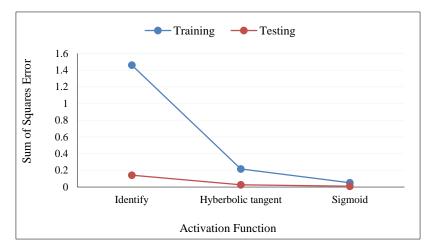


Figure 8. Sum of squares error with different activation functions

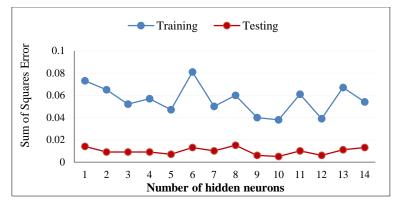


Figure 9. Sum of squares error with different hidden neurons number

3.2.2. Model Prediction

For rescaling input data, the ANNM was developed using the standardized method with the utilization of the scaled conjugate gradient as the optimization algorithm required for estimating the network weights. Further, by using the ANNM between the observed and predicted WQI values, a high value of the coefficient of determination (R2) of 0.973 was achieved (see Figure 10).

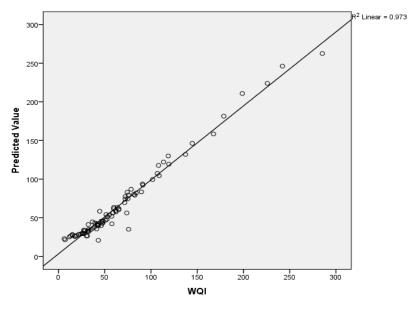


Figure 10. Predicted-by-observed chart for WQI

3.2.3. Independent Variable Importance

The significance of the independent variables lies in the measurement of the variation of the model (facilitated by independent variables) during the prediction of the output values. In this context, Figure 11 shows that pH exerts the highest effect on the method adopted by the model for the WQI prediction, followed by chloride, total dissolved solids, and sulfate.

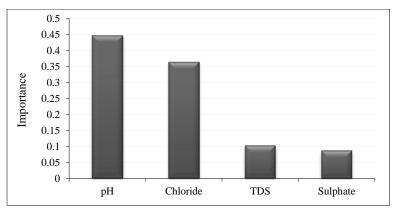


Figure 11. Independent variable importance chart

4. Conclusion

In the present work, for carrying out the analysis of the groundwater quality in Al-Karkh and Al-Rusafa sides of Baghdad city, determination of the WQI, required for establishment of the suitability of groundwater for drinking purposes, was carried out using four chemical water parameters. Upon evaluating the results most of the chemical parameters, for less than a quarter of water samples, exceeded the allowable standard limits of Iraq, which in turn makes it unfit for human drinking. Moreover, the presence of gypsum, solid waste dumping and other anthropogenic activities affected the groundwater quality of the area under investigation. With the aim of further predicting the groundwater quality in the future ANNM was developed. By using this model, the smallest values of error functions for training and testing samples and high R2 value were obtained. Thus, the model exhibits high prediction efficiency. However, chloride and pH, the most vital parameters, affected the model prediction.

5. Conflicts of Interest

The author declares no conflicts of interest.

6. References

[1] Abbas, Abdul Hussain A., Ammar S. Dawood, and Zainab M. Al-Hasan. "Evaluation of Groundwater Quality for Drinking Purpose in Basrah governorate By Using Application of Water Quality Index." Kufa Journal of Engineering 8, no. 1 (2017): 65-78.

[2] Ali, S. M. "Hydrological environmental Assessment of Baghdad area." PhD diss., Ph. D. thesis, University of Baghdad, College of Science. 245, 2012.

[3] Arbuckle, James L. "IBM SPSS Amos 19 user's guide." Crawfordville, FL: Amos Development Corporation 635 (2010).

[4] Assessment of water Quality Index of Groundwater in Al-Khadhimiya city. Iraqi Journal of Science, 58(4A):1898-1909.nterprises 65 (2004): 71-88. doi: 10.24996/ ijs.2017.58.4A.12.

[5] British Standard, B. S. "1377: Part 3: 1990." British Standard Methods of Test for Soils for Civil Engineering Purposes.

[6] Brown, Robert M., Nina I. McClelland, Rolf A. Deininger, and Michael F. O'Connor. "A Water Quality Index -Crashing the Psychological Barrier." Indicators of Environmental Quality (1972): 173–182. doi:10.1007/978-1-4684-1698-5_15.

[7] Chatterjee, Chinmoy, and M. Raziuddin. "Determination of Water Quality Index (WQI) of a degraded river in Asansol industrial area (West Bengal)." Nature, Environment and pollution technology 1, no. 2 (2002): 181-189.

[8] Edition, Fourth. "Guidelines for drinking-water quality." WHO chronicle 38, no. 4 (2011): 104-8.

[9] Egereonu, U. U. "Assessment of atmospheric aerosols from three satellite stations: Heavy metal pollutants." Journal of Association for the Advancement of Modeling and Simulation Techniques in EHussein, A. T., AL Kind, Y. G. and AL Ani, H. F. (2017).

[10] ICOSQC, I.C.O.f.S.a.Q.C. (2009). Standard Specification for Drinking Water (Vol. 417). Iraq: Iraqi Chronicle.

[11] Khan, Zafar, Tarek Zayed, and Osama Moselhi. "Structural Condition Assessment of Sewer Pipelines." Journal of Performance of Constructed Facilities 24, no. 2 (April 2010): 170–179. doi:10.1061/(asce)cf.1943-5509.0000081.

[12] Kley, G., and N. Caradot. "Review of sewer deterioration models." (2013).

[13] Mahmood, Amal A., Amal M. Eassa, H. Mohammed Muayad, and Y. Shubbar Israa. "Assessment of ground water quality at Basrah, Iraq by water quality index (WQI)." Journal of University of Babylon 21, no. 7 (2013): 2531-2543.

[14] Mosaferi, Mohammad, Mojtaba Pourakbar, Mohammad Shakerkhatibi, Esmaeil Fatehifar, and Mehdi Belvasi. "Quality modeling of drinking groundwater using GIS in rural communities, northwest of Iran." Journal of Environmental Health Science and Engineering 12, no. 1 (2014): 99. doi:10.1186/2052-336x-12-99.

[15] Munawar, Mohiuddin. "Limnological studies on freshwater ponds of Hyderabad-India I. The Biotope." Hydrobiologia 35, no. 1 (1970): 127-127. doi:10.1007/bf00143305.

[16] National Ground Water Association (NGWA). Groundwater Use in the United States of America. [Online] 2017.[Cited: May 10, 2016.] http://www.ngwa.org/Fundamentals/Documents/usa-groundwater-use-fact-sheet.pdf.

[17] Raju, N. Janardhana, Priyanka Patel, Deepa Gurung, Prahlad Ram, Wolfgang Gossel, and Peter Wycisk. "Geochemical assessment of groundwater quality in the Dun valley of central Nepal using chemometric method and geochemical modeling." Groundwater for Sustainable Development 1, no. 1-2 (2015): 135-145. doi:10.1016/j.gsd.2016.02.002.

[18] Salman, Baris. "Infrastructure management and deterioration risk assessment of wastewater collection systems." PhD diss., University of Cincinnati, 2010.

[19] Sojobi, Adebayo Olatunbosun. "Evaluation of groundwater quality in a rural community in North Central of Nigeria." Environmental monitoring and assessment 188, no. 3 (2016): 192. doi:10.1007/s10661-016-5149-y.

[20] Tirkey, Poonam, Tanushree Bhattacharya, Sukalyan Chakraborty, and Suman Baraik. "Assessment of groundwater quality and associated health risks: a case study of Ranchi city, Jharkhand, India." Groundwater for Sustainable Development 5 (2017): 85-100. doi:10.1016/j.gsd.2017.05.002.

[21] Tran, Huu Dung. "Investigation of deterioration models for stormwater pipe systems." PhD diss., Victoria University, 2007.

[22] Yogendra, K., and E. T. Puttaiah. "Determination of water quality index and suitability of an urban waterbody in Shimoga Town, Karnataka." In Proceedings of Taal2007: The 12th world lake conference, vol. 342, p. 346. 2008.

[23] Zouahri, Abdelmjid, Houria Dakak, Ahmed Douaik, Mhamed El Khadir, and Rachid Moussadek. "Evaluation of groundwater suitability for irrigation in the Skhirat region, Northwest of Morocco." Environmental monitoring and assessment 187, no. 1 (2015): 4184. doi:10.1007/s10661-014-4184-9.