ISSN 2522-6967

Artificial Neural Network Modelling of Total Dissolved Solid (Elzawia city – Libya, as a case study)

Lubna S. Ben Taher

Assistant Lecturer, Civil Engineering Department, University of Benghazi, Libya.
Corresponding author; e-mail: lina7981@yahoo.com

ABSTRACT

In this study Mathematical, and statistical methods to simulate an aquifer water quality parameters have been considered. It is necessary to measured water quality for groundwater zonation maps. Therefore, more accurate maps produced more essential role in the discussionmanagement. Water samples were collected from thirty wells at Elzawi city located around 45km west of Tripoli and analysed for water quality parameters including: EC(Electrical Conductivity), TDS(Total Dissolved Solids), Ca (Calcium), Mg (Magnesium), and pH using standard methods. The wells location andwater level at wells observed by Global positioning system (GPS) type Garmin GPS 12XL. An artificial neural network (ANN) models were investigated to predict the TDS in water of Elzawi city wells. The input variables were the wells longitude, latitude, EC, Ca, Mg, and pH while the TDS in water was the output. The Levenberg–Marquardt (LM) algorithm and Back propagation used for training of the feed forward ANN. The ANN models performance were compared using the coefficient of determination(E), mean absolute percentage error (MAPE %), and 95% confidence interval(CI95%).It has been a good agreement between actual data and the ANN outputs for training, validation and testing data sets at the forth model (ANN4) while based on all inputs . ANN4 model performed superior to the other models in predicting TDS with high E=0.94 and lowest MAPE= 5.6% and the result average within the range of observed 95CI%. Also the ANN models could be successfully applied and provide high accuracy and reliability for water quality parameters forecasting.

Keywords:*Artificial Neural Networks, Total Dissolved Solid, Zawia city, Water Quality*.

1. Introduction

Water is the source and continuity of life. Water supply in developing countries has a major role in agriculture and industries. Elzawia city located at the sea costa –Libya as Figure.1. Libya is located greatly in the dry or semi-dry climatic conditions and shortage of water, the extraction and use of groundwater has been the main of water supply in the country. Vast increase in the utilization of underground water resources has led to an annual decline of groundwater table, which has been growing in the past two decades. Excessive extraction of groundwater resources in dropped down changes its quality. So it is necessary to measure it to manage the decision making entails a comprehensive project that should be provided based on the data and the maps from the surface of groundwater as well as zonation maps of the groundwater quality. Therefore, the more accurate maps result produced more essential role in the management because it is the basis for correct management and operation of groundwater extraction, assessment of the groundwater quality and quantity of zoning map changes [1, 3]. Much research has been done for underground water quality parameters, here are some studies in this field to be mentioned:

Nemati et al. [2] they studied of Total Dissolved Solid (TDS) time series reported in their paper using local water quality parameters of Calcium (Ca), Chloride (Cl), Magnesium (Mg), Sodium (Na), Bicarbonate (HCO3), Sulfate (SO4), and water discharge (Q) for a set of recorded data in Simineh River at Dashband gauging station during 1993-2011. Also employed the Garson equation to assess the relative importance of the variables to determine appropriate input combinations. The ANNs with different numbers of neurons in the hidden layer were constructed, and model performance estimated by means of several indicators, including data sequence, scatter diagrams, and quantitative measures of RMSE, MAE, and R^2. The modeling results indicated that reasonable prediction accuracy was achieved for the ANNs model.

Moasheri et al. [1], in their study, spatial distribution of the parameter TDS with two models of Kriging and Kriging-ANN model was examined. The correlation coefficient for the model Kriging = 0.684 and for model Kriging-ANN was obtained = 0.96 the results showed a significant superiority of the combined model Kriging-ANN. Besides the advantage prediction, it makes possible spatial to predict TDS distribution with access to parameters such as PH and EC.

Moasheri et al. [3], they obtained results from their research showed that estimated parameters of quality spatial distribution of sodium, calcium and magnesium with the optimized combination method with genetic algorithm provided more accurate results than the geo-statistical method in kriging. Water quality parameters that could be used to estimate the unknown values it could be used in the estimation of spatial and temporal distribution of quality parameters of groundwater of Kashan plain.

Sattari et al. [4], their research was motivated by determine an accurate and affordable method to predict EC and TDS; the research had demonstrated that TDS could be accurately predicted using only five parameters. Another motivation for the research was to evaluate suitability of data-mining algorithms to model relationships between parameters of the river water. The research had shown that the best scenario to estimate the TDS in water and EC involved a combination of the following parameters: Sodium, Mg, and Ca, Chloride, and Bicarbonate ion. The computational results showed no need to use the costly process of sampling hydro chemical parameters that affect river water quality. Rather, using fewer hydro chemical parameters had led to acceptable accuracy in estimating TDS and EC values.

Kheradpisheha et al. [5], their study's results, using artificial neural network with the back-propagation algorithms for modeling qualitative parameters of groundwater, such as Cl, EC, SO4, was accurate according to the chosen input parameters. They used back-propagation algorithm and obtained good results because they used a different water source or experienced the impact of other parameters. According to the highly expensive and time-consuming tests, these parameters could be modeled to

ISSN 2522-6967

estimate their range, a quick and cost-effective method for management practices, especially in emergency situations.

Salem and Alshergawi [6], studied the quality of groundwater of fifty one wells for drinking water from Alshati district and assessed for its suitability for drinking. Water samples were collected and analyzed for various physico-chemical parameters such as pH, Temperature (C°), Electrical Conductivity (EC), Total Hardness (TH), Total Dissolved Solids (TDS), Alkalinity (Alk), Chloride (Cl^-), Iron (Fe^{++}), Manganese (Mn^{++}), Calcium (Ca^{++}), Magnesium (Mg^{++}), Nitrate (NO_3^-), Sodium (Na^+), Potassium (K^+), Sulphate (SO_4^{--}) and Phosphate ($PO4^{--}$). The results revealed that some parameters of water samples were out of limit according to the WHO standards and Libyan standards for drinking water. The results shows rising of Fe^{++} and Mn^{++} in most of the wells studied, rising of Ca^{++} in water samples of wells of Al-Mansura and Abu-gadgud, rising of SO_4^{--}, TDS and EC in water samples of wells rising of Cl^- in water samples of Idri, Waanzarik, Taamasan, Al-Mansura and Al-kadra. NO_3^- was also rising in water samples of wells of Mahruga and Bergen.

Maedeh et al. [7], the results from their structures of different models of neural networks, was observed. The fifth model with least amount of data and, hence least number of tests to find out the different parameters, turns out to be the most cost-effective and involves lowest error, as regards TDS parameter prediction of Tehran groundwater, which in view of the inputs and the neural networks in models, the estimate thus obtained was remarkably and favorably high. In light of the model developed, a better future estimate and a more reliable forecast to enhance the quality and application of groundwater can be made via controlling the sulphide, chloride and sodium parameters in forecasting the TDS parameter.

Alshakel (2015), in his study some chemical analyses were carried out for about 30 samples of the Zawia city wells, to evaluation the concentration of total dissolved solids (TDS), Ca^+, Mg^+ Ions and the total hardness. The results of these analyses indicated to high concentration of (TDS), Ca^+, Mg^+ Ions and the total hardness in several samples .The source of these elements might from sea water intrusion because this area located beside the sea coast, or another source as sewage disposal.

Electrical conductivity and total dissolved solids are considered as important parameters in determining quality of drinking and agricultural water because they directly represent total salt concentration in the water. Increases in these parameter values indicate a reduction in water quality. In this study, estimation of the TDS in about thirty pumping well located at Elzawi city- Libya, was studied using the ANNs. The input variables were the well longitude, latitude, EC, Ca, Mg, and pH while the TDS in water was the output. The Levenberg–Marquardt (LM) algorithm was used to train ANN and back propagation used for the training of the feed forward ANN. The ANNs models performance were compared using the coefficient of determination (E), mean absolute percentage error (MAPE %), and 95% confidence interval.

2. Materials And Methodology

2-1 Study Area:

Elzawi city is located in around 45km west of Tripoli between North latitude 3244'53.160''N and East longitude 1243'23.160 see Figure. 1, and the population of it is about 186132. The average annual temperature in the region with 24.4 ºC, and the Annual rainfall 255 mm.

2-2 Data Analysis

A total of 30 water samples were collected from underground wells localities in El Zawia city the geographical location as longitude1243'41 to1244'74, latitude3246'25 to 3247'3. After collection, water samples were protected from direct sunlight and transported in a cooling box containing ice packs to the laboratory for analyses. Water samples were examined for physicochemical parameters (pH, Ec, Ca^{2+}, Mg^{2+} and TDS).Different maxima, minima, averages and deviations of standard have been calculated, the correlation between them and well latitude - longitude also see Table-1. Finally, the models designed by neural networks to predict the TDS parameter with different input combination. Figures 2 to 5 present the histogram and map distribution of the (pH, Ec, Ca^{2+}, $Mg2^+$, TDS) for the wells. The values presented in Table-2 indicate that the TDS is highly correlated with EC, Mg and Ca. noted that also there is a good correlated between the TDS, Mg, Ca, PH and Ec with the location of wells presented as the latitude and longitude presented as the latitude and longitude.

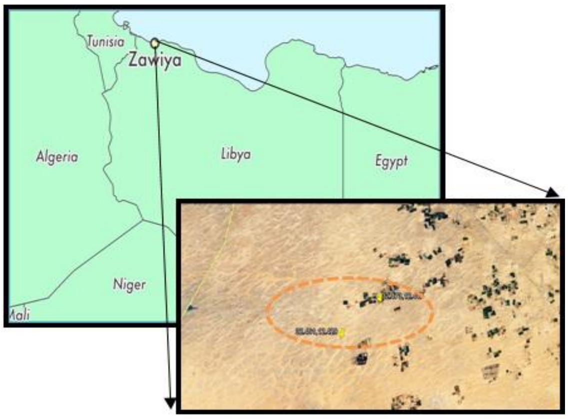

Fig. 1: Map showing El Zawia city and the places where the wells collected.

Table-1. Statistical Analysis of Field Measurement, El Zawia City Wells.

Variable	Mean	StDev	Minimum	Maximum
latitude	32.46	0.00312	32.461	32.473
longitude	12.43	0.00513	12.429	12.447
Mg^{2+} ppm	443.70	262.4	70.8	1318
Ca^{2+} ppm	504.3	207.4	140	1050
Ec Mc/cm	3040	1201	685	5905
PH	7.05	0.1874	6.68	7.44
TDs ppm	2049	1000	446	4900

Table-2.The correlation coefficient of Field Measurement, El Zawia City Wells

	latitude	longitude	Ec	PH	Mg^{2+}	Ca^{2+}
longitude	0.598					
Ec	0.528	0.313				
PH	-0.141	-0.249	-0.506			
Mg2+	0.612	0.277	0.906	-0.356		
Ca2+	0.324	0.045	0.873	-0.369	0.836	
TDs	0.524	0.228	0.932	-0.475	0.878	0.84

ISSN 2522-6967

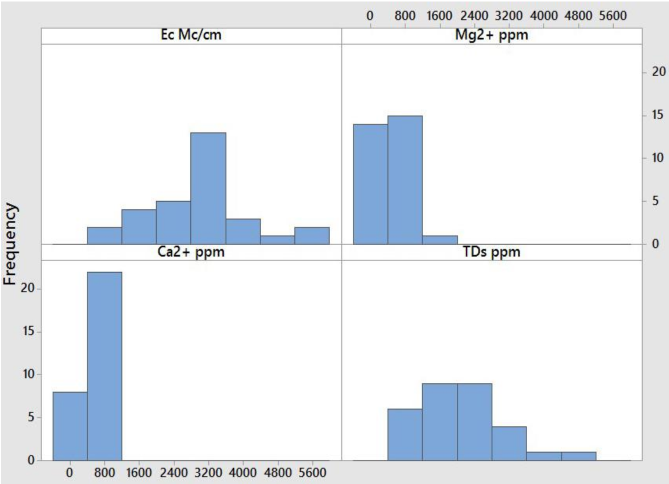

Fig.2: Histogram of Ec Mc/cm, Ca ppm, Mg ppm, and TDS ppm, El Zawia City Wells.

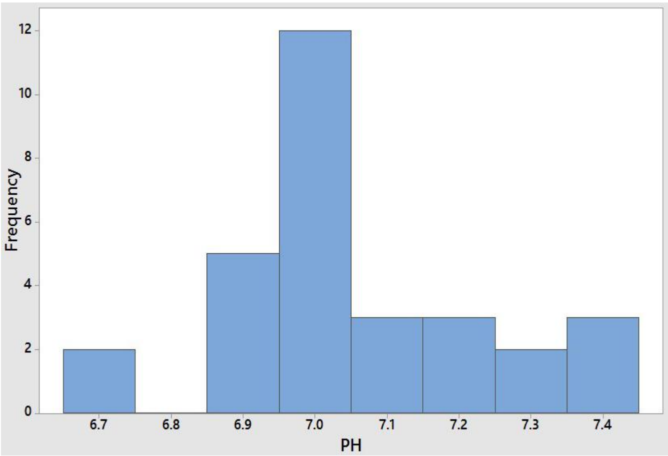

Fig.3: Histogram of the (PH)El Zawia City Wells.

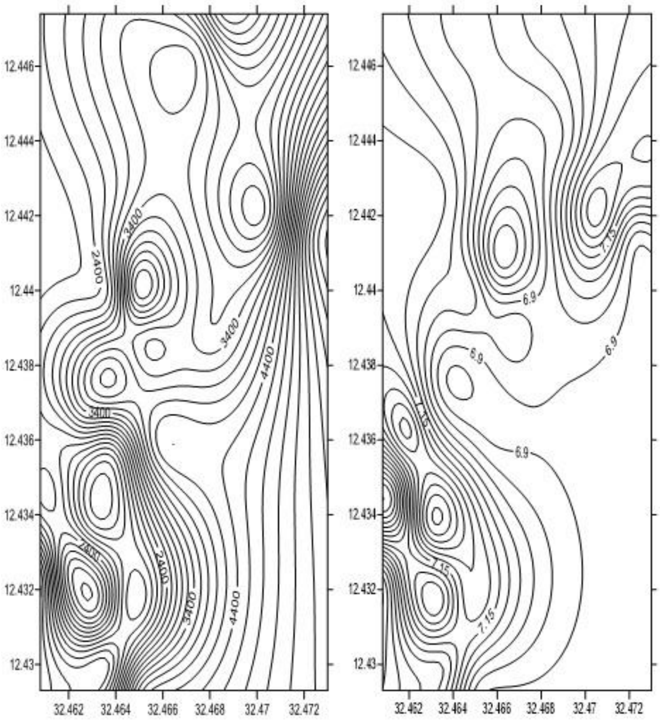

Fig.4: Map distribution of Ec Mc/cm and PH El Zawia City Wells.

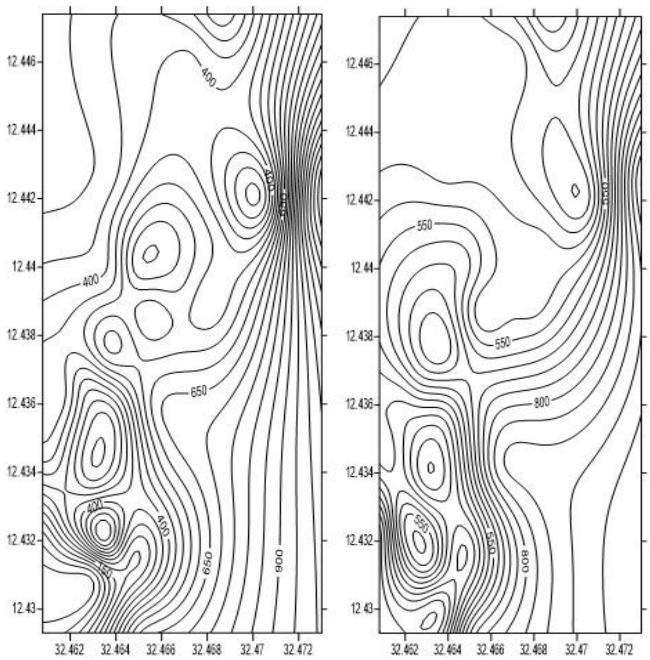

Fig.5. Map distribution of Mg^{+2} ppm and Ca^{+2}ppm, El Zawia City Wells.

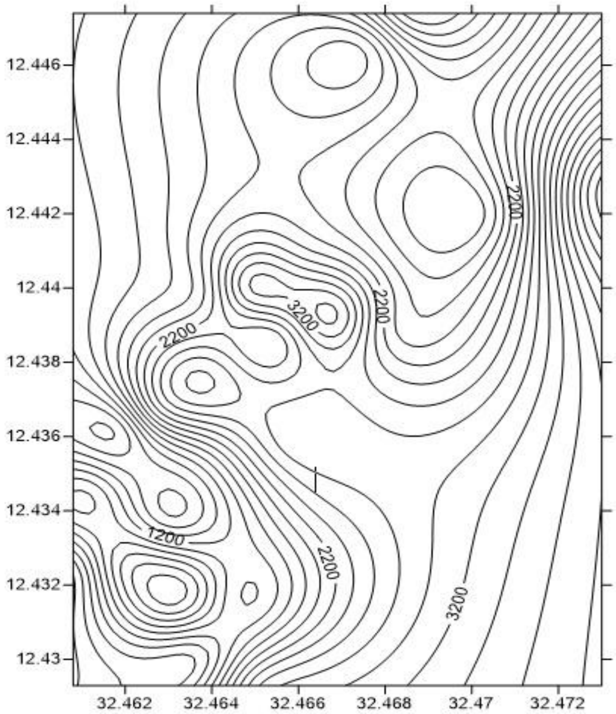

Fig.6: Map distribution ofTDS ppm, El Zawia City Wells.

3. Study Methodology

Artificial Neural Networks (ANN), are a form of computing inspired by the functioning of the brain and nervous system. Neural networks consist of a set of neurons or nodes arranged in layers, and in the case weighted inputs are used, these nodes provide suitable inputs by conversion functions. Each neuron in a layer is connected to all the neurons of the next layer but without any interconnection among the neurons in the same layer. The weight learned for each neuron in ANNs model remains internal, and therefore, their associations with physical systems are often overlooked (Nemati, 2014).

The feed forward ANN has been adopted in many environmental modeling studies because of its applicability to a variety of different problems. Noted that more than one hidden layer may require in feed forward networks because a three-layer network can generate arbitrarily complex decision regions. Also, the appropriate input vector to the ANN model can be identified according to the procedure of the modeler. Back propagation is the most popular algorithm used for the training of the feed forward ANN. An objective function that considers both the ANN's structure and error, minimizes a linear combination of the resulting ANN's squared errors, weights, and biases in order to develop a less complex model at the end of training the resulting network has good generalization qualities.

The Levenberg–Marquardt (LM) training algorithm is a trust region based method with a hyper-spherical trust region. This algorithm was implemented in this study using the Neural Network Toolbox of MATLAB, an example of developed structure of ANN with 6 inputs see in Figure 7.

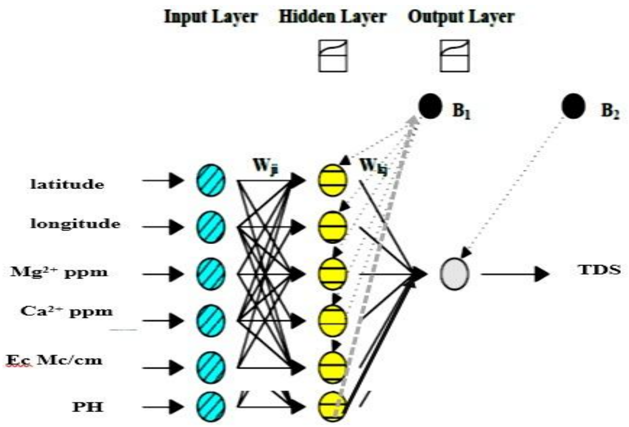

Fig.7: Developed Structure of ANN with 6 inputs.

In this study, several statistical parameters were used to evaluate the performance of predicted models, which were given by the following relations:

1- Mean absolute percentage error (MAPE %)

$$MAPE\% = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{TDSobs. - TDSpre.}{TDSobs.}\right| \quad - - - - - 1$$

2- 95% confidence limit (95CI%):Standard error of the mean is given as; $1.\ S_{X} =$

The quantity (TDS. - μ) / Sx has a t-distribution with n-1 degrees of freedom, And for 95% confidence limit

$$1.\ TDS. - 1.95\left(\frac{s}{\cdot}\right) < \mu < TDS. + 1.95\left(\frac{s}{\cdot}\right. \quad 2$$

The value on the left side of the inequality yields the lower limit, and on the right side yields the upper limit for the mean

3- *E*, Coefficient of Determination

$$E = 1 - \left(\frac{\sum_{i=1}^{n}\left(TDS.obs._i - TDS.pre._i\right)^2}{\sum_{i=1}^{n}\left(TDS.obs._i - \overline{TDS.obs.}\right)^2}\right) \quad - - - - - - - 3$$

Where:

n= number of data,

TDS.obs.= observed value,

TDS.pre.= predicted value,

S = standard deviation, and

1. $\overline{TDS.obs}$ = the average of the observed da

A better fit, with zero indicating MAPE% and high value of E. Coefficient of determination(E = 0 to 1) is calculated on the basis of the relationship between the predicted and observed mean deviations and it can show the correlation between the predicted and observed data. E is better suited to evaluate model than the square root of the correlation coefficient between the predicted and observed value (R^2).The probability of procedure produces an interval that contains the actual true parameter value is known as the Confidence Level and is generally chosen to be 95CI%. So the model if have a good performance well produce a results within the range of 95CI% of the mean observed data. The models are used to generate data which conserve the main statistical characteristics of the historical data. This is verified through comparing values of mean, of generated data with those of historical data.

4. Estimate the spatial distribution TDS by artificial neural networks

In this study, the water quality data of 30 wells located at El Zawia city were used. Concentrations of the parameters have been measured, and record consists of 5 parameters including: (Ca), (Mg), PH, (EC), and (TDS). In order to develop ANN models for prediction TDS, the data were divided into two groups randomly: training data, accounting for 70 percent and testing and validation data, making up 30 percent of the total data. The ANN models were trained using Bayesian Regularization (BR) and Levenberg–Marquardt (LM) algorithms. In ANN models the number of neurons in the hidden layer were found by a trial and error procedure. The activation functions used for the hidden and output layers were the 'logsig' and 'purelin' functions, respectively. Table- 3 showing the structure of ANN models according to the input combinations, moreover the models were improved by the accuracy with respect to MAPE%, E, and CI 95%. ANN (6, 20, 1) model indicates model having 6, 20 and 1 for the input, hidden and output layer, respectively and the data divided in to (20 values for model training, 5 values for model validation, and 5 values for model testing). Over all ANN models showing best prediction for all input combination in both test and validation periods. Figure 8 presented the comparison between the predicted and observed TDS data. The best architecture was obtained for ANN TDS models (ANN-4) had been selected based on minimum value of MAPE% and maximum value of E. The output from the best selected architecture for the ANN-4 model was validated using the testing data set. The objective of the validation process is to investigate the ability of the model to work with an independent data series that have not been used in training of the evaporation model.

Looking into the models, it is found out that, forecasting the TDS parameter in the first and second models have a greater magnitude of error, as they have fewer input parameters. Therefore, this indicated the inefficiency of the simulating and training algorithms. In the third and the fourth models the error declines as the simulating and training algorithms were kept constant. It was also revealed that error dwindles to its minimum as the number of input neurons decreased and an extra layer was built into the fourth model. By contrast, the third and fourth model, consisting of 4 and 6 parameters, respectively result in

ISSN 2522-6967

acceptable error. Figurer 8 the distribution of the observed data (the vertical axis) and the predicted data (the horizontal axis) in the testing and training stage of the all models. It should be noted that the closer the data get to a one –to-one diagram, the more reliably the model evaluates the TDS proportion.

An examination of figure 8 shows that the sensitivity of different parameters in predicting the neural network of the four models was observed for (Ca), (Mg), (EC) parameters, and the latitude, longitude respectively. It goes through the Ec, latitude, and longitude of the wells, played a good pronounced role in predictions for TDS.

The spatial distribution of the parameter TDS with the ANN model was examined. The results show a significant superiority of the combined model ANN to the well latitude, longitude. The advantage prediction, it makes possible to predict spatially TDS distribution with access to other parameters such as EC.

Table-3. Error statistics for input combinations using ANN models in test and validation stage.

Input combinations	ANN Model architecture	MPE%	E	The average Predicted TDS. ppm	95% CI **Observed TDS. ppm**
Mg^{2+} ppm Ca^{2+} ppm	ANN1 (2,20,1)	17.6	0.88	1989.95	1675.6 –2422.5
PH Mg^{2+} ppm Ca^{2+} ppm	ANN 2 (3,20,1)	19.5	0.83	2080.44	
Ec Mc/cm PH Mg^{2+} ppm Ca^{2}+ ppm	ANN3 (4,20,1)	7.7	0.94	2098.10	
Latitude Longitude EcMc/cm PH Mg^{2+}ppm Ca^{2+}ppm	ANN 4 (7,20,1)	5.6	0.94	2028.97	

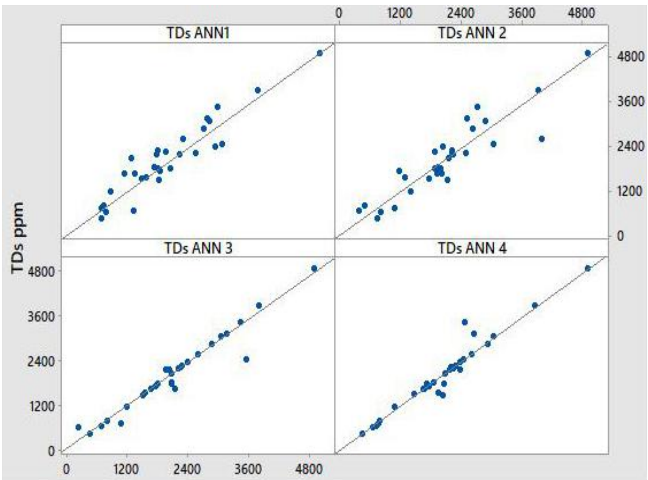

Fig.8: The observed and estimated TDS using ANN at El Zawia City Wells.

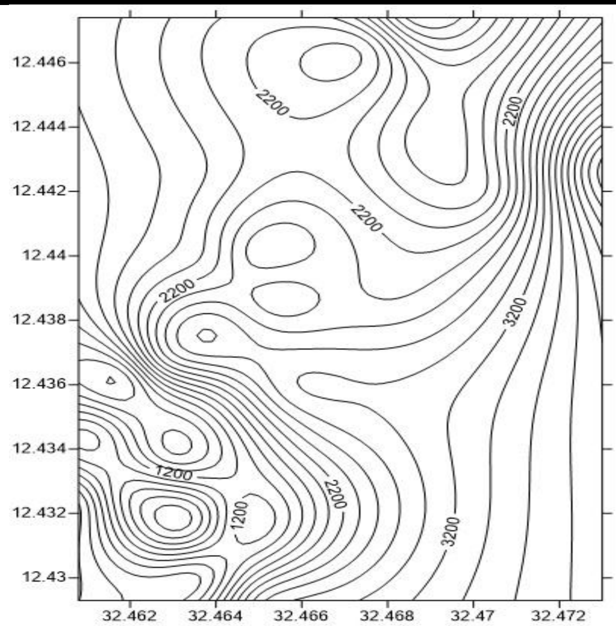

Fig.9: Map distribution ofpredicted TDS ANN4 El Zawia City Wells.

5. Conclusion

In this study, estimation of the TDS in thirty pumping well located at Elzawi city- Libya, was studied using the ANN models. The input variables were the well longitude, latitude, EC, Ca, Mg, and pH while the TDS the output. The Levenberg–Marquardt (LM) algorithm was used to train ANN and back propagation used for the training of the feed forward ANN. The ANN models performance were compared using the coefficient of determination (E), mean absolute percentage error (MAPE %), and 95% confidence interval.

The sensitivity of different parameters in predicting the neural network of the four models was observed for Ca, Mg, EC parameters, and the latitude, longitude respectively. It goes through the Ec, latitude, longitude of the wells, played a good pronounced role in predictions for TDS.

The models designed by neural networks to predict the TDS parameter with different input combination. There is a good correlation between the TDS, Mg, Ca, PH and Ec with the location of wells presented as the latitude and longitude.

ANN4 model performed superior to the other models in predicting TDS with high E=0.94 and lowest MAPE= 5.6% and have predicted mean within the range of observed 95CI%.

The results show a significant superiority of the combined model ANN of the well latitude, longitude. Besides the advantage prediction, it makes possible spatially prediction of TDS distribution with access to other parameters such as EC. This is a quick and cost-effective method for management practices, especially in emergency situations and makes.

References

[1] Moasheri S. A, KhammarG.A,Poornoori.Z,Beyranvand. Z,Soleimani M., 2013,Estimate the spatial distribution TDS the fusion method Geostatistics and artificial neural networks, International Journal of Agriculture and Crop Sciences., 7,410-420.

[2] Nemati S., Naghipour L, and Fazeli Fard MH. 2014,. Artificial Neural Network Modeling of Total Dissolved Solid in the Simineh River, Iran., J. Civil Eng. Urban., 4, 08-14.

[3] Moasheri S.A, RezapourO. M, Beyranvand .Z, Poornoori .Z ,2013, Estimating the spatial distribution ofgroundwater qualityparameters of Kashan plain with integration method of Geostatistics - Artificial Neural Network Optimized by Genetic-Algorithm ,International Journal of Agriculture and Crop Sciences.,20, 2434-2442.

[4] Sattari M. T, Joudi A. R, Kusiak.A,2016,Estimation of Water Quality Parameters With Data-Driven Model, http://dx.doi.org/10.5942/jawwa.2016.108.0012 Peer-Reviewed.

[5] Kheradpisheh. Z, Talebi .A, Rafati. L, Ghaneeian M.T, Ehrampoush M.H,Groundwater quality assessment using artificial neural network: A case study of Bahabad plain, Yazd, Iran, Desert Online at http://desert.ut.ac.ir, 20, 65-71.

[6] SalemM. A., AlshergawiM .I, 2013, Physico-chemical Evaluation of Drinking Water Quality in Alshati District of Libya,Journal Of Environmental Science, Toxicology And Food Technology (IOSR-JESTFT), 4, 46-51.

[7] MaedehP. A, Mehrdadi. N, Nabi Bidhendi1 G.R , Zare H. ,2013, Application of Artificial Neural Network to Predict Total Dissolved Solids Variations in Groundwater of Tehran Plain, Iran,International Journal of Environment and Sustainability, 2,45-51.