International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2016): 79.57 | Impact Factor (2015): 6.391

Mathematical Models for Predicting of Organic and Inorganic Pollutants in Diyala River Using Analysis Neural Network

Dr. Nawar O. A. Al-Musawi

Assistant Professor, Civil Engineering Department, College of Engineering/ University of Baghdad, Iraq

Abstract: ***Diyala river is the most important tributaries in Iraq, this river suffering from pollution, therefore, this research aimed to predict organic pollutants that represented by biological oxygen demand BOD, and inorganic pollutants that represented by total dissolved solids TDS for Diyala river in Iraq, the data used in this research were collected for the period from 2011-2016 for the last station in the river known as D17, before the river meeting Tigris river in Baghdad city. Analysis Neural Network ANN was used in order to find the mathematical models, the parameters used to predict BOD were seven parameters EC, Alk, Cl, K, TH, NO_3, DO, after removing the less importance parameters. While the parameters that used to predict TDS were fourteen parameters pH, DO, BOD, PO4, NO_3,Ca, Mg, TH, K, Na, SO_4,Cl, EC, Alk. The results indicated that the best correlation coefficient is 86.5% for BOD, and the most important parameter is Chloride Cl, and the best correlation coefficient is 95.4% for TDS and the most important parameters are total hardness TH and electrical conductivity EC, according to direct relation between these parameters and TDS.***

Keywords: Diyala river, BOD, TDS, ANN.

1. Introduction

Iraq is part of the Middle East and North Africa which is known as MENA region. Iraq greatly relies in its water resources on the Tigris and Euphrates Rivers. It was considered rich in its water resources till 1970s. After that problems due to water scarcity appeared, it is expected that water shortage problems will be more serious. **[1].** This scarcity cause increases in water pollution. To conquer this problem, it is important to prevent our water resources from pollution in order to provide water pleasant and safe to drink. Baghdad city has two rivers, Tigris River as the main river and Diyala River in boundary of Baghdad City eastern of Baghdad as shown in Figure 1. The surface water in Baghdad is suffering from effect of conservative pollutants.

In Diyala province most streams and rivers suffer from an increase in environmental pollution rates resulting from numerous factors namely the negligence of citizens and the infractions of many service administrations that pump their waste directly into the rivers. one of the main rivers in the province is Diyala River stretching through more than 150 km. The river springs from Iran and it is considered to be one of Tigris important feeders. **[2]**

Diyala river was studied by many researches, artificial neural networks approaches were used for two case studies which were Diyala River and Adhim River northern Iraq. Different training algorithms and different artificial neural networks such as Levenburg Marqudat LMNN , Scaled conjugate gradient SCGNN , radial basis function networks RBNN and generalized regression networks GRNN were selected in modeling and generation of synthetic stream flow for the mentioned case studies. The performance of applied networks were determined according to well-known test parameters R 2, E nash, Rbias ,MAPE, MAE. It has been found in his study that Levenburg Marqudat is faster and have better performance than Scaled conjugate gradient algorithm in training operation while the radial basis networks and generalized regression networks presented the best performance among all kinds of networks. [3]

Another study included application and comparison of artificial neural network, for three case studies which were Diyala River, Adhim and Elkhzer Rivers northern Iraq. Different training algorithms and different artificial neural networks such as Levenburg Marqudat LMNN, Scaled conjugate gradient SCGNN, radial basis function networks RBNN and generalized regression networks GRNN were selected in modeling and generation of synthetic stream flow for the mentioned case studies. Two other methods were also applied to the mentioned case studies which are support vector machine SVM and adaptive neuro fuzzy interference ANFIS model which integrates both neural networks and fuzzy logic principles. A comprehensive comparison between the applied models was done to determine the best performance for each case study . The performance of applied networks and models were determined according to well known test parameters R^2 , E nash, Rbias ,MAPE, MAE. It has been found in this study that Levenburg Marqudat is faster and have better performance than Scaled conjugate gradient algorithm in training operation while the radial basis networks and generalized regression networks presented the best performance among all kinds of networks for Diyala and Adhim rivers while the best performance for Elkhazer river was only by radial basis function networks.[4]

Diyala River and Al-Rustumiya Wastewater Treatment Plant (WWTP) were studied, for Diyala river the parameters (BOD, SO4, PO4 & Cl) were 100% violations and that maximum level of pollution. While Al-Rustumiya Wastewater Treatment Plant having very high level of violation for the years (2005, 2006 & 2007) and the values were between (55% - 87%). **[2]**

2. Study Area

Diyala River is one of the important water resources and one of the important tributaries of Tigris River in Iraq, the draining area is reaching 32600 km^2. It covers a total distance of 445 km (275 miles). Diyala Weir is controlled the river, it controls floods and irrigates the area northeast of Baghdad City. The catchments of the river were divided into four parts; Derbendikhan, Upper Diyala, Middle Diyala and Lower Diyala, It finally feeds into the Tigris river below Baghdad City. **[5]**

3. ANN Structure

The ANN initial properties are the ability to learn, organize memory and consequent operation, eventually reaching the fault tolerance. A typical three-layered network has an input layer (I), a hidden layer (H) and an output layer (O). It's known that the increasing of the hidden layers number affects the complexity of the network and decreases the learning accuracy. Theoretical works have shown that a single hidden layer is sufficient for the ANNs to approximate to any complex nonlinear function. Meanwhile in 1999; Michael et al. recommended using one hidden layer to avoid increasing the complexity of the network. Indeed, many experimental results seem to confirm that one hidden layer may be enough for most forecasting problems. Each layer includes several neurons and the layers are interconnected by sets of correlation weights. The neurons produce outputs through transforming inputs comes from the initial inputs or the interconnections, by an adequate nonlinear transfer function. Sigmoid function is the common transfer function. **[6]**

4. Results and Discussion

In this research Artificial Neural Network was used in order to find mathematical models for predicting organic and inorganic pollutants in Diyala river, the data were collected for the period 2011-2016 for the last station on Diyala river which is known as D17 before Diyala river meeting Tigris river in Baghdad City. The data collected from the the Ministry of Water Resources in Baghdad [7]. Fourteen pollutants parameters were chosen pH, DO, TDS, PO_4, NO_3,Ca, Mg, TH, K, Na, SO_4,Cl, EC, Alk, to find the ANN model with the best correlation coefficient for the organic pollutants which is BOD concentration, the results indicated that the best correlation coefficient is 80%, while after removing the less importance parameters the correlation coefficient reach 86.5%, the architecture neural network for BOD is shown in figure 2, the parameters used are EC, Alk, Cl, K, TH, NO_3, DO, and the best correlation coefficient is 86.5% and the regression equation is y=31.32+0.71*x as shown in figure 3, while figure 4 shows the parameters importance, the most importance parameter is chloride Cl, the parameters estimates are shown in table 1, and the mathematical model for predicting BOD appears in equation 1.

$$Y=\frac{1}{1+\exp-(0.486+2.622\tanh X1+2.067\tanh X2-1.443\tanh X3)+1.8} \quad (1)$$

Where:
Y=BOD concentration
$X1= 4.22114 - (1.84*10^{-4})\ EC - (6.53*10^{-4})\ Alk - (4.24*10^{-3})\ Cl + .076\ K - (2.917*10^{-3})\ TH + 0.028\ NO_3 + 0.376\ DO$
$X2= -1.8311 + (3.07*10^{-5})\ EC - (2.74*10^{-4})\ Alk + (2.61*10^{-3})\ Cl - 0.04\ K + (6.19*10^{-4})\ TH - 0.0169\ NO_3 - 0.0775\ DO$
$X3= 3.5078 + (4.798*10^{-4})\ EC - (6.1598*10^{-3})\ Alk + (2.607*10^{-3}) + 0.033\ K - (5.345*10^{-3})\ TH - 0.0356\ NO_3 - 0.013\ DO$

The fourteen parameters also used to predict inorganic pollutants which is TDS concentration, figure 5 shows the architecture neural network for TDS, the results indicated that the best correlation coefficient is 95.4% and the regression equation is y=99.37+0.94*x as shown in figure 6. Figure 7 shows the parameters importance, the most importance parameters are total hardness TH and electrical conductivity EC, the parameters estimates are shown in table 2, and the mathematical model for predicting TDS appears in equation 2, Figures 8 and 9, shows the variation of observed and predicated data for both BOD and TDS respectively.

$$Y=\frac{1}{1+\exp-(-1.372-2.624\tanh X1+0.619\tanh X2+0.619\tanh X3+0.002\tanh X4)+555} \quad (2)$$

Where:
Y=TDS concentration
$X1= 0.549 - (3.986*10^{-4})\ BOD - (4.726*10^{-3})\ PO_4 + (2.54788*10^{-3})\ NO_3 - (6.35*10^{-4})\ Ca - (2.33*10^{-3})\ Mg + (5.495*10^{-4})\ TH - (6.928*10^{-3})\ K - (2.043*10^{-4})\ Na + (2.9*10^{-4})\ SO_4 - (2.91*10^{-4})\ Cl + (1.197*10^{-4})\ EC + 92.329*10^{-4})\ Alk - 0.034\ pH + 0.016\ DO$
$X2= -0.931 + (1.213*10^{-3})\ BOD - 0.164\ PO_4 - 0.018\ NO_3 - (2.25*10^{-4})\ Ca - (7.606*10^{-4)}\ Mg + (5.02*10^{-5})\ TH + 0.015\ K + (8.94*10^{-4})\ Na - (4.363*10^{-4})\ SO_4 + 1.034*10^{-3}\ Cl - (8.003*10^{-6})\ EC - (1.13*10^{-3})\ Alk + 0.183\ pH - 0.042\ DO$
$X3= -0.75 + (1.5598*10^{-4})\ BOD - 0.057\ PO_4 - (2.107*10^{-3})\ NO_3 + (8.6*10^{-4})\ Ca - (3.6197*10^{-3})\ Mg - 5.117\ TH - 0.019\ K - (1.49*10^{-3})\ Na + (3.435*10^{-4})\ SO_4 - 6.879\ Cl + (1.232*10^{-4})\ EC - 2.146*10^{-3}\ Alk + 0.148\ pH - (7.47*10^{-3})\ DO$
$X4= 0.08798 + (1.63*10^{-3})\ BOD - 0.0198\ PO_4 + 0.0183\ NO_3 + (1.13*10^{-3})\ Ca + (1.739*10^{-3})\ Mg - (3.794*10^{-4})\ TH - 0.021\ K - (1.517*10^{-3})\ Na - (3.76*10^{-4})\ SO_4 - (752*10^{-5})\ Cl - (1.258*10^{-4})\ EC + (1.228*10^{-3})\ Alk + 0.048\ pH - 0.031\ DO$

5. Conclusion

From this research we can conclude that:
1) It is possible to predict organic pollutants in Diyala River, represented by BOD concentration using ANN model.
2) For BOD , the best correlation coefficient can be achieved after removing the less importance parameters, then the parameters that remained are EC, Alk, Cl, K, TH, NO_3, DO, the regression equation is

y=31.32+0.71*x and the correlation coefficient is 86.5%.

3) The most importance parameter is Chloride Cl, because it is effect on the concentration of BOD due to the high concentration of Chloride discharging in waste water.
4) It is possible to predict inorganic pollutants in Diyala River, represented by TDS concentration using ANN model.
5) For TDS, the fourteen parameters that mentioned before are used, the regression equation and the best correlation coefficient are y=99.37+0.94*x and 95.4% respectively.
6) The most importance parameters are total hardness TH and electrical conductivity EC, that is because these two parameters are related with TDS.

Figure 1: Diyala river map

Table 1: Parameter Estimates for BOD mathematical model

Predictor		Predicted			
		Hidden Layer 1			Output Layer
		H(1:1)	H(1:2)	H(1:3)	BOD
Input Layer	(Bias)	2.659	-1.501	1.292	
	EC	-.575	.096	1.499	
	Alk	-.143	-.060	-1.349	
	Cl	-2.258	1.391	1.387	
	K	1.279	-.701	.555	
	TH	-2.383	.506	-4.367	
	NO3	.581	-.345	-.727	
	DO	.438	-.902	-.154	
Hidden Layer 1	(Bias)				.486
	H(1:1)				2.622
	H(1:2)				2.067
	H(1:3)				-1.443

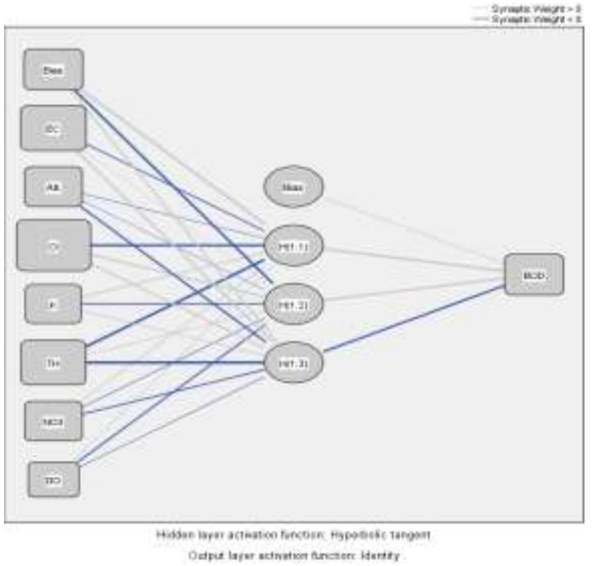


Figure 2: Architecture of the Artificial Neural Network for BOD Mathematical Model

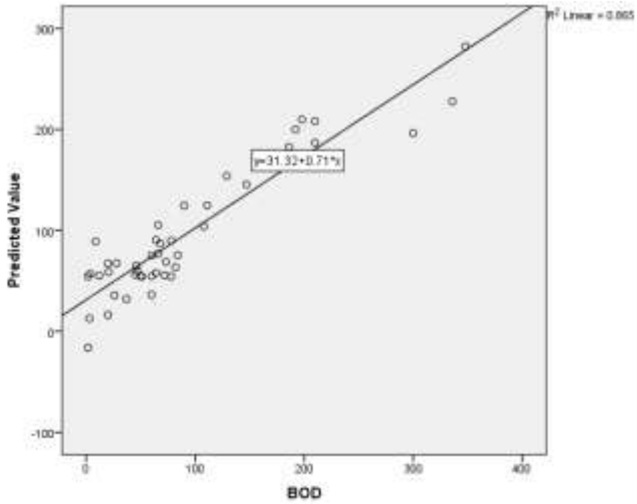


Figure 3: Predicted values versus dependent variable for BOD mathematical model

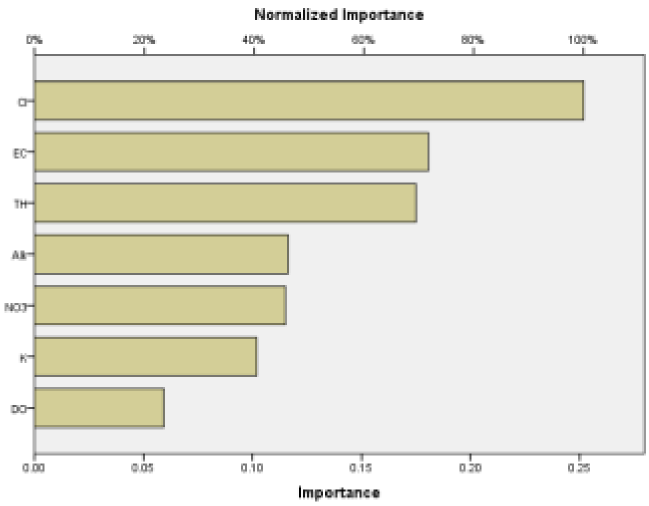


Figure 4: Parameters importance for BOD mathematical model

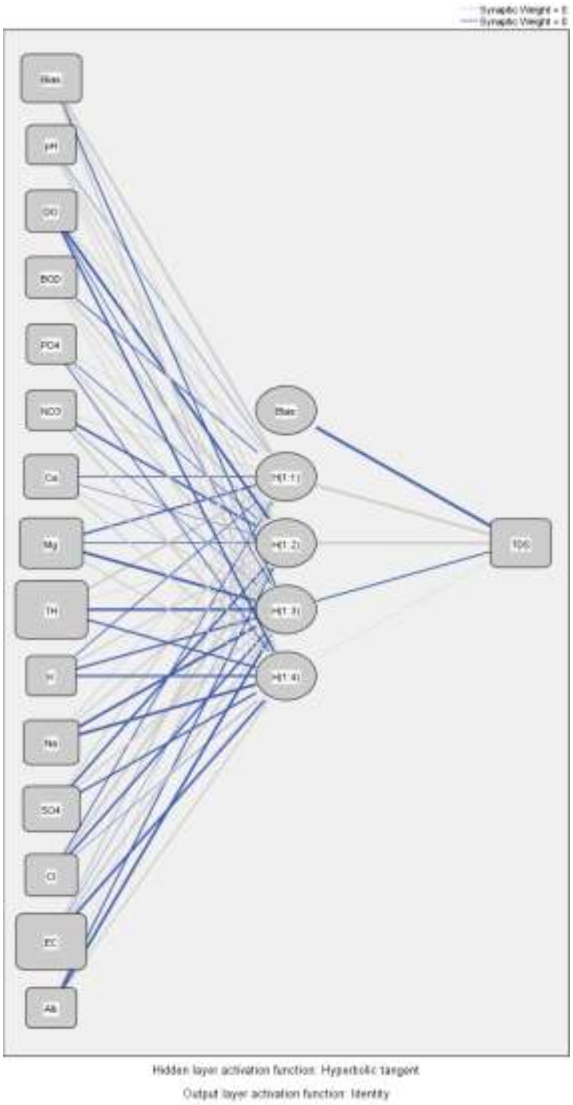


Figure 5: Architecture of the Artificial Neural Network for TDS Mathematical Model

Table 2: Parameter Estimates for TDS mathematical model

Predictor		Predicted				
		Hidden Layer 1				Output Layer
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	TDS
Input Layer	(Bias)	.630	.088	-.241	.353	
	pH	-.041	.223	.180	.059	
	DO	.190	-.484	-.087	-.364	
	BOD	-.138	.420	.054	.056	
	PO4	.030	-.104	.361	-.126	
	NO3	.052	-.368	-.043	.373	
	Ca	-.127	-.045	.172	.226	
	Mg	-.333	-.108	-.514	.247	
	TH	.449	.041	-.423	-.310	
	K	-.117	.255	-.328	-.355	
	Na	.065	.284	-.474	-.482	
	SO4	.247	-.371	.292	-.320	
	Cl	-.155	.550	-.366	-.040	
	EC	.374	-.025	.385	-.393	
	Alk	.051	-.247	-.470	.269	
Hidden Layer 1	(Bias)					-1.372
	H(1:1)					2.629
	H(1:2)					.619
	H(1:3)					-.223
	H(1:4)					.002

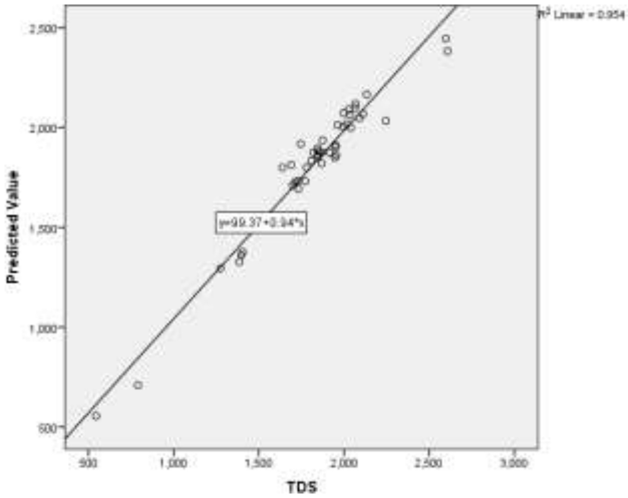


Figure 6: Predicted values versus dependent variable for TDS mathematical model

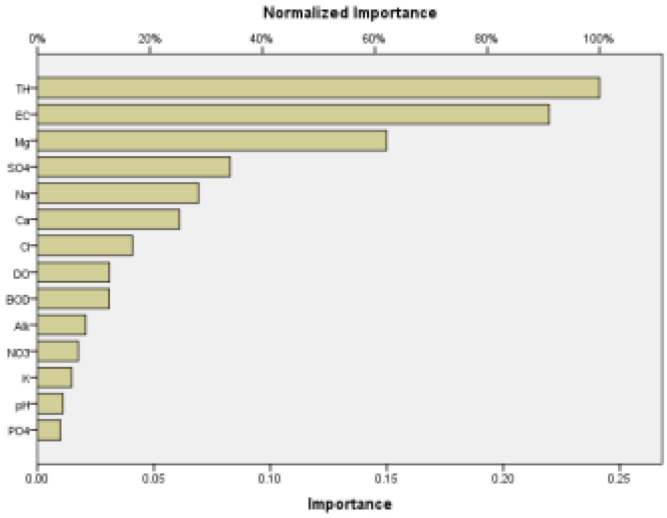


Figure 7: Parameters importance for TDS mathematical model

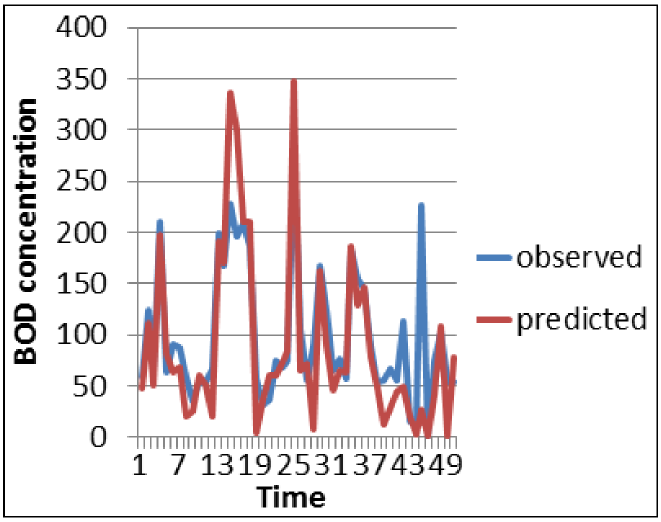


Figure 8: Variation of observed and predicted values for BOD concentrations

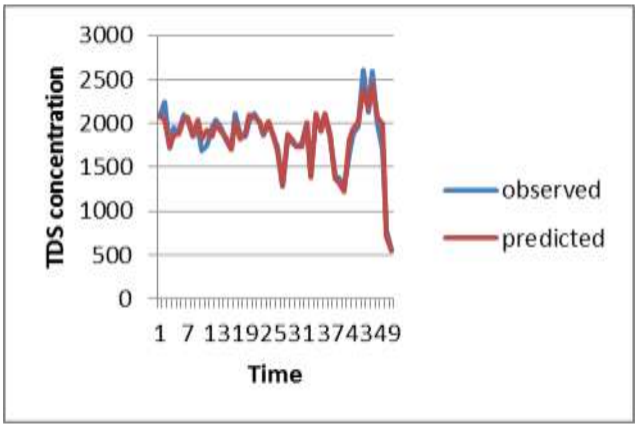


Figure 9: Variation of observed and predicted values for TDS concentrations

References

[1] **Nadhir Al-Ansari, Ammar A. Ali, Sven Knutsson, 2014,** Present Conditions and Future Challenges of Water Resources Problems in Iraq, Journal of Water Resource and Protection, 2014, 6, 1066-1098**,** Published Online September 2014 in SciRes. http://www.scirp.org/journal/jwarp http://dx.doi.org/10.4236/jwarp.2014.612102

[2] Allaa M. Aenab & S. K. Singh, 2014, Critical Assessment of River Water Quality and Wastewater Treatment Plant (WWTP), IOSR Journal Of Environmental Science, Toxicology And Food Technology (IOSR-JESTFT) e-ISSN: 2319-2402,p-ISSN: 2319-2399.Volume 8, Issue 2 Ver. IV (Mar-Apr. 2014), PP 44-48 www.iosrjournals.org.

[3] **ChelengA.Arslan, 2015,** Performance of Different Artificial Neural Networks in Monthly Stream flow Forecasting for Diyala And Adhim Rivers Northern Iraq, Civil Engineering, Selcuk University, Konya TURKEY, *ISITES2015 Valencia –Spain.*

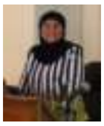
[4] Chelang A. Arslan , 2016, Performance of Artificial Neural Networks , Support Vector machine and Fuzzy logic networks ANFIS in Monthly Stream flow Forecasting for Diyala, Adhim and Elkhazer Rivers Northern Iraq**,** International Journal of Scientific & Engineering Research, Volume 7, Issue 6, June-2016 105 ISSN 2229-5518. **Journal of Water Resource and Protection, 2014, 6, 1066-1098**

[5] **Nahida H. Hamza, 2011,** Evaluation Of Water Quality Of Diyala River For Irrigation Purposes**,** Vol. 05, No. 02, pp. 82-98 , December 2012, Diyala Journal of Engineering Sciences, Iraq.

[6] **Abbas m. Abd 1, saad sh. Sammen, 2014,** Artificial Neural Network Model For Managing And Forecasting Water Reservoir Discharge (Hemren Reservoir As A Case Study), Diyala Journal of Engineering Sciences, Vol. 07, No. 04, pp. 132-143, December 2014.

[7] **Operation Department**, 2017, Documented Information, Ministry of Water Resources, Baghdad, Iraq.

Author Profile



Nawar O. A. Al-Musawi, Baghdad, Iraq , B.Sc. in Water Resources Engineering, M.Sc. In Environmental Engineering, Ph.D. In Environmental Engineering, from University of Baghdad, Assistant Professor in water quality, work as faculty member in Civil Engineering Department, College of Engineering, University of Baghdad, Iraq.