

Application of Artificial Neural Network for Predicting Iron Concentration in the Location of Al-Wahda Water Treatment Plant in Baghdad City

Dr. Nawar Omran Ali Al-Musawi

Instructor

Engineering College-Baghdad University

E-mail:nawaromran@yahoo.com

ABSTRACT

Iron is one of the abundant elements on earth that is an essential element for humans and may be a troublesome element in water supplies. In this research an ANN model was developed to predict iron concentrations in the location of Al-Wahda water treatment plant in Baghdad city by water quality assessment of iron concentrations at seven WTPs up stream Tigris River. SPSS software was used to build the ANN model. The input data were iron concentrations in the raw water for the period 2004-2011. The results indicated the best model predicted Iron concentrations at Al-Wahda WTP with a coefficient of determination 0.9142. The model used one hidden layer with two nodes and the testing error was 0.834. The ANN model could be used to predict future iron concentrations as the results from the verification of the ANN model for years 2012 and 2013 indicated good accuracy with a coefficient of determination $R^2 = 0.8965$.

Key Words: Iron, ANN model, Tigris River, Water treatment plant, Al-Wahda.

استخدام شبكة الخلايا العصبية لتوقع تركيز ايون الحديد في محطة الوحدة لمعالجة المياه في مدينة بغداد

د.نوار عمران علي الموسوي

مدرس

كلية الهندسة-جامعة بغداد

الخلاصة

يعتبر الحديد احد العناصر المهمة على الارض وذو اهمية اساسية للانسان وقد يسبب مشاكل كيميائية في مياه الشرب, في هذا البحث تم بناء نموذج رياضي بالاعتماد على مبدا الخلايا العصبية في تقدير تركيز الحديد في موقع محطة الوحدة لمعالجة مياه الشرب وذلك عن طريق تقييم نوعية مياه نهر دجلة بالاعتماد على عنصر الحديد في مواقع المحطات السبعة لمعالجة مياه الشرب في مدينة بغداد وقد استخدم البرنامج الاحصائي SPSS في بناء النموذج الرياضي. تراكيز الحديد المستخدمة في النموذج كانت للفترة بين 2004-2011 وقد اظهرت النتائج من النموذج ان معامل التحديد هو 0.9142 وقد استخدم النموذج طبقة مخفية واحدة مع خليتين ونسبة خطأ قيمتها 0.834.

اظهرت النتائج دقة النموذج الرياضي من حيث التقارب بين القيم الحقيقية والقيم المتوقعة في المستقبل وعند استخدام تراكيز الحديد للعامين 2012 و 2013 كانت النتائج ذات تقارب جيد وبمعامل تحديد 0.8965.

الكلمات الرئيسية: الحديد, نموذج شبكة الخلايا العصبية, نهر دجلة, محطة معالجة المياه, الوحدة.

1. INTRODUCTION

Water quality of different sources and quality changes is a subject of ongoing concern. Consequently there is a need for effective methods for modeling water quality parameters in surface water to control pollution and apply necessary managements. Different modeling approaches are applied to analyze water resources, Artificial Neural Network (ANN) is one of these approaches, and the advantages of ANNs are that they are able to represent both linear and nonlinear relationships of water resources data. Also ANN was found suitable for prediction purposes as they produce accurate results. **Chi et al., 2006** applied ANNs for the classification and prediction of water quality of Yangtze River in China using DO, COD, NH₃ and pH values for the period from January 2003 to September 2005. The results showed severe pollution is reached in this river and many efforts should be taken for pollution control. **Musavi and Golabi, 2008.** selected an ANN model for water quality simulation of Karoo River in Iran. Several water quality parameters were chosen CO₃, HCO₃, SO₄, Cl, Na, Ca, Mg, K, Ec, TDS and SAR. The results showed that the ANN model was able to predict water quality in the river very successfully with more than 90% accuracy.

Mozejkod and Gniot 2008 developed an ANN model for time series modeling of total phosphorous concentration in Odra River in Poland. The variables used in the model were temperature, pH, NO₃, DO, BOD, COD, SO₄, Cl₂, and SS. The models prediction matched reasonably with the observations. The correlation coefficient R was 0.865 and mean absolute error MAE was 0.024 mgP/dm³. **Ali et al., (2009)** proposed ANN models for the prediction of TDS, EC and Turbidity in Johor River in Malaysia. The models predicted water quality parameters with high accuracy as the mean absolute error was 10% for the different models. In **2010, Vesna et al.,** developed an ANN model to predict DO in Gruza Reservoir in Serbia. The most effective inputs for the model were the pH and temperature, while Cl and total phosphate were found to be the least effective parameters. The correlation coefficient using the most effective inputs (pH, Temp., NO₃, NO₂, NH₃, conductivity, Fe, Mn) was 0.8478. **Hosseini and Ehsan, (2011)** presented empirical multilayer ANN model to estimate water quality indices BOD and DO, in Marad Big River in Iran. The input data for the ANN model were EC, TDS, SS, Turbidity, Na, HCO₃, NO₃, NH₃ and PO₄. The models are capable to capture long-term trends observed for DO and BOD both in time and space. DO model had a correlation coefficient of 0.972 and for BOD 0.937.

In this research an ANN model is to be developed to predict iron concentrations in Tigris River in Baghdad city. The primary sources of iron in water are from natural geologic sources, also iron based materials, such as cast iron and galvanized steel pipes that have been widely used in water distribution systems. 0.3 mg/L can cause water to turn a reddish brown color **Illinois Department of Public Health, 2010.**

The presence of iron and manganese in water cause threats to industrial and municipal water supplies, formation of scales, and blockage of water pipes thus leading to economic losses **Sodamade and Pearse, 2013.** Excessive ingested iron can cause excessive levels of iron in the blood that can damage the cells of the gastrointestinal tract, preventing them from regulating iron absorption. Humans experience iron toxicity above 20 milligrams of iron for every kilogram of mass, and 60 milligrams per kilogram is a lethal dose. Sullivan was the first to propose and continued to

reiterate that iron levels play a major role in producing atherosclerosis. **Sieliuchi et al., 2010** stated that women, who have a reduced iron load, have strong protection against atherosclerosis, compared to men in the same age group. They suggested that iron can be involved in Alzheimer's disease. The important mechanism is the interaction of iron and cholesterol in promoting oxidative damage, causing both atherosclerosis and neuro degeneration. Excessive iron in water can stain clothing and can give a metallic taste to water or to food. Furthermore, iron deposits can build up in pressure tanks, water heaters, and pipelines, reducing the quantity and pressure of the water supply **OVIVO, 2010**.

In this research an Artificial Neural Network was developed to predict iron concentration in Tigris River at the location of Al-Wahda water treatment plant in Baghdad city through water quality assessment of iron parameter upstream (locations of seven water treatment plants upstream).

2. MATERIALS AND METHODS

2.1 Study Area Description

Tigris river water is considered the only source of potable water for the city of Baghdad, and the river divides the city into right (Karkh) and left (Risafa) sides with a flow direction from north to south. The study area within Baghdad City is located in the Mesopotamian alluvial plain between latitudes $33^{\circ}14'$ - $33^{\circ}25'$ N and longitudes $44^{\circ}31'$ - $44^{\circ}17'$ E, 30.5 to 34.85 m above mean sea level. The area is characterized by arid to semi-arid climate with dry hot summers and cold winters; the mean annual rainfall is about 151.8 mm **Al-Adili, 1998**. In Baghdad city, a tremendous increase in freshwater demand is required due to the rapid growth in population and accelerated industrialization. The quality of the flowing water is affected by the effluent discharges from various uncontrolled sources as domestic, industries, agriculture along the downstream stretch.

2.2 Data Collection and Analysis

The data used in this paper were provided from Mayoralty of Baghdad for the period from January 2004 to December 2013 which represented the monthly average values for iron concentration in the flowing water. These data were at the locations of the water treatment plant on the Tigris River. There are eight water treatment plants (WTPs) from the north to the south of Baghdad city, Al-Karkh, East Tigris, Al Wathba, Al Karama, Al Qadisias, Al Dora, Al Rasheed and Al Wahda WTPs as shown in **Fig.1**. All water treatment plants in Iraq are designed as conventional plants. This treatment process does not significantly affect the concentrations of the dissolved constituents. Therefore it's important to monitor iron concentrations of the influent to these treatment plants to provide treated water with permissible iron concentrations to the consumer.

2.3 ANN Model

A typical neural network consists of a large number of elements known as nodes. Each node is connected to other nodes by links with associated weights. These weights represent information used by the net to solve the case to be solved. The nodes are arranged in a number of layers. The first layer is the input layer where the inputs (dependent variables) are applied to the net. The last layer is the output layer where the outputs (independent) are extracted. The layers between the input and

output are the hidden layers (one or more) with a number of nodes as proceeding elements. Many types of ANN are known, the popular is the multiple layer perceptron network. Important issues to design this network are the number of hidden layers and number of nodes they contain **Mozejko and Gniot, 2008**. A training process is carried out on a set of the input data as these data are divided into training and testing data. The learning process is based on the training data and testing data is used to verify the performance of the trained ANN. During the training process the weights of the input nodes are adjusted by several trails until the desired error function is obtained. The ANN is trained by minimizing this error in search space of weights **Chi et al., 2006**.

The SPSS software (Statistical Procedure for Social Science, version 20) was used to build the ANN model in this study. The model comprised seven nodes in the input layer which represent the influent iron concentration to the water treatment plants namely (Alkarkh, EastTigris, Wathba, Karama, Qadisia, Dora, and Rasheed), the output layer represents influent Iron concentration at Wahda water treatment plant. Many trials were performed, where in each run the software parameters were changed for selecting the best ANN model according to the highest correlation coefficient and the smallest testing error.

3. RESULTS AND DISCUSSION

3.1 Iron Concentration in Tigris River

The period under study was from 2004-2013, the data were collected are tabulated in **Table 1** in order to view the variation of iron concentration in raw water with respect to time and distance. **Fig. 2** and **3** indicated that the maximum iron concentration was 10.04 ppm at year 2011 in East Tigris WPT (upstream plant) and the minimum was 0.33 ppm in 2005 at Al Karama WTP (at the middle reach).

The main sources of iron in Tigris River are **Alsaqqar et al., 2015**:

- 1-The outlets of combined and storm sewers that discharge their wastewater directly in to the river.
- 2-The disposal of effluents from some industries: Al Taji gas factory, wool and textile factories upstream Al Karama WTP, vegetable oil factory, detergent factory and the cement factory near Al Wahda WTP.
- 3-Effluents from Al Dora refinery upstream Al Dora WTP and the effluent from Al Dora power plant which uses large amounts of water in the cooling towers.
- 4-There are several raw water pumping stations that provide irrigation water to the nearby lands. At the end of the working day, the flow is reversed in the pipes for cleaning the system from the sediments, this may lead to the corrosion of the pipes and cleaning water may contain iron concentrations.
- 5-Iron may precipitate in the coagulation and flocculation process in the WTPs, so high concentrations of iron are found in the settled sludge which is discharged into the river.

3.2 ANN Model for iron in Tigris River

The ANN model developed from the SPSS software is summarized in **Table. 2**. From the case processing summary the data (total of 70) were divided to 68.6% training, 15.7% testing and 15.7% holdout. The input data were from 2004 to 2011 which

represented the monthly average values for iron concentrations at the seven WTPs upstream to Wahda WTP.

The architecture of the ANN model is shown in the parameter estimates where two nodes are found in the hidden layer, H(1:1) and H(1:2) as shown in **Fig. 4**. In this section the weights of the input parameters from each hidden node is calculated.

The best ANN model representing iron concentrations at Al Wahda WTP had a coefficient of determination $R^2 = 0.9142$ as shown in **Fig. 5** and testing error 0.834 as shown in **Table 3**. While **Table 4** shows the parameters estimates. The importance of the independent variables is shown in **Table 5**, where Al Karama WTP had the highest importance of 30.52% and Al Rasheed WTP had 26.16% which affected the predicted values of Iron at Al Wahda WTP.

3.3 Verification of the Model

In order to verify the accuracy of the ANN model for predicting the iron concentrations at Al Wahda WPT the recorded data (observed) for years 2012 and 2013 were plotted against the predicted values (from the ANN model) as shown in **Fig. 6**. The model gave good accuracy with a coefficient of determination $R^2 = 0.8965$.

CONCLUSIONS

1- Iron concentrations in Tigris River are fluctuated.

2-The developed ANN model was suitable in predicting iron concentration in Tigris River with a coefficient of determination $R^2 = 0.914$. The model used one hidden layer with two nodes and the testing error was 0.834.

3- From the verification of the ANN model for years 2012 and 2013, the model gave good accuracy with a coefficient of determination $R^2 = 0.8965$ for predicting Iron concentrations in Tigris River.

REFERENCES

- Ali Najah, Ahmed E., Othman A. K. and Othman J., 2009, Prediction of Johor Water Quality Parameters Using Artificial Neural Networks, European Journal of Scientific Research, Vol.28, No.3, pp 422-435.
- Alsaqqar, Awatif Soaded, A. Hashim, and A. Mahdi, 2015, Water Quality Index Assessment using GIS case study: Tigris river in Baghdad city, International Journal of Current Engineering and Technology, Vol.5, No.4, Aug 2015.
- S. Al-Adili, 1998, Geotechnical Evaluation of Baghdad Soil Subsidence and Their Treatment, Ph.D. Thesis, University of Baghdad, Iraq.
- Chi Zhou, Liang G., Haibing G. and Chuanyong P., 2006, **Pattern** Classification and Prediction of Water Quality by Neural Network with Particle Swarm Optimization, Proceedings of the 6th World Congress on Control and Automation June 21-23, Dalian, China.



- Hossein Banejad and Ehsan Olyaie, 2011, Application of an Artificial Neural Network Model to Rivers Quality Indexes Prediction Study, *Journal of American Science*, 7(1).
- Illinois Department of Public Health, 2010, Fact Sheet, Division of Environmental Health, Springfield. Updated Dec.
- Mayoralty of Baghdad, 2015, Baghdad Water Administration, Documented Data.
- Mozejko J. and Gniot R., 2008, Application of Neural Networks for the prediction of total Phosphorous Concentrations in Surface Waters, *Polish J. of Environmental Study*, Vol. 17, No. 3 pp 363-368.
- Musavi Jahromi S. H. and Golabi M., 2008, Application of Artificial Neural Networks in River Water Quality Modelling: Karoon River, Iran, *Journal of Applied Sciences* 8(12), pp 2324-2328.
- OVIVO, 2010, Municipal Drinking Water Removal, www.ovivowater.com, Tunisia.
- Sieliechi Joseph, Marie G., Joseph K. and Ion S, 2010, Effect of Water Treatment Residuals (Aluminum and Iron Ions) on Human Health and Drinking Water Distribution Systems, *International Journal of Conservation Science*, Vol. 1, Issue 3, July-September 2010, pp 175-182, www.ijcs.uaic.ro.
- Sodamade, G. A and Pearse, L. F., 2013, Removal of Iron and Manganese Using Rapid Sand Filtration, *NSE Technical Transaction*, Vol.47, No.3, July-September , Lagos – Nigeria.
- Vesna Rankovic, Jasna R., Ivana R., Aleksandar O. and Ljiljana C., 2010, Neural Network Modeling of Dissolved Oxygen in the Gruza Reservoir, Serbia, *Ecological Modeling* 221, pp 1239-1244.

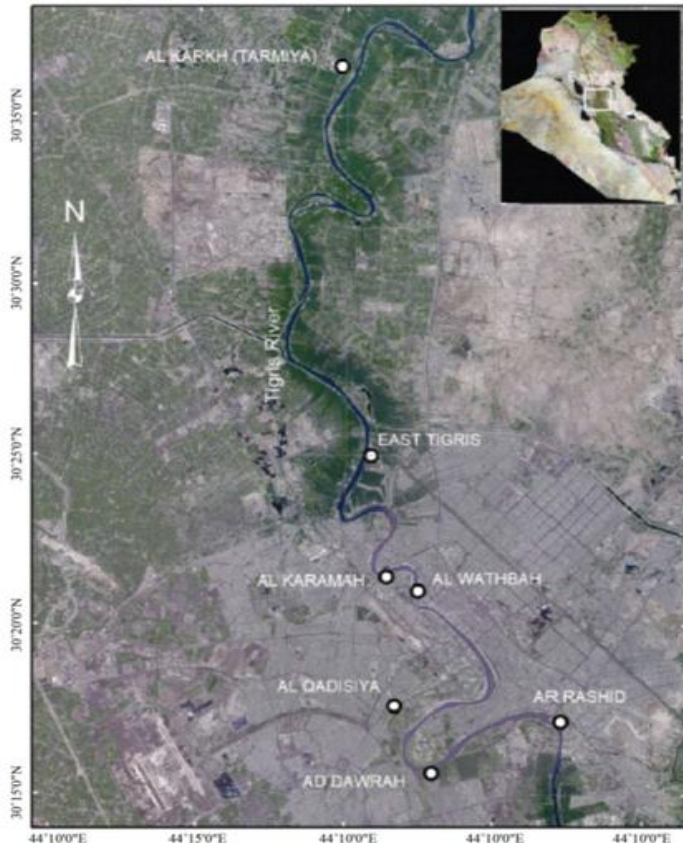


Figure 1 .Sampling locations across Tigris River, within Baghdad City. Location of the eight water treatment plants.

Table 1 .Variation of average annual iron concentration (ppm) in raw water with Respect to time and distance, Mayoralty of Baghdad, 2015

RAW	AL-KARKH	ESTTIGRIS	WATHBA	KARAMA	QADISIA	DORA	RASHEED	WAHDA	AVG
2004	1.00	1.13	1.13	1.54	1.58	1.42	2.14		1.42
2005	0.55	1.08	0.67	0.33	1.59	1.19	1.41		0.97
2006	1.70	1.92	2.88	0.84	2.79	2.93	2.56	2.11	2.22
2007	1.19	1.93	1.99	0.71	3.34	1.35	1.79	1.12	1.68
2008	0.50	1.25	0.89	0.32	1.37	0.66	0.70	0.66	0.79
2009	0.97	2.00	0.82	1.57	3.52	1.39	2.54	2.37	1.90
2010	0.84	1.81	0.63	2.61	2.46	2.69	1.21	1.23	1.68
2011	0.83	10.04	3.51	1.97	2.11	1.08	1.13	2.95	2.95
2012	0.63	2.72	1.97	1.51	1.58	0.50	0.89	0.80	1.33
2013	1.02	1.95	4.84	1.39	2.15	1.23	1.44	1.16	1.90
AVG	0.92	2.58	1.93	1.28	2.25	1.44	1.58	1.55	1.69

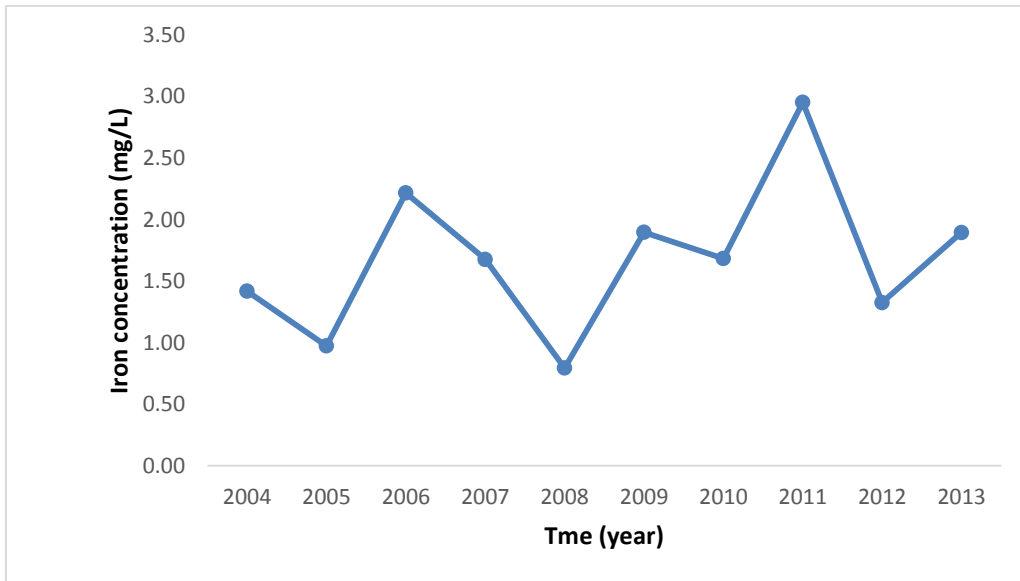


Figure 2. Variation of average annual iron concentrations (ppm) in raw water with respect to time

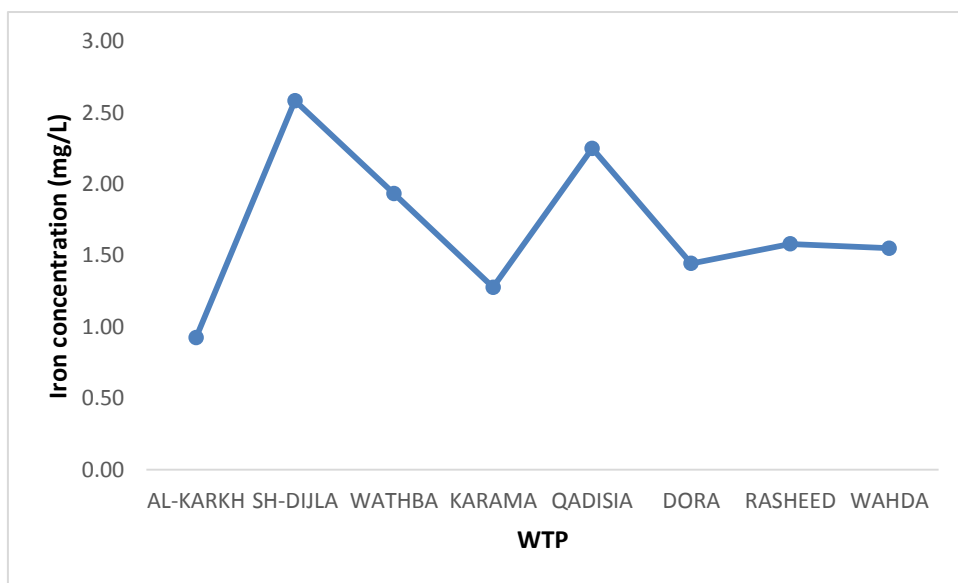


Figure 3. Variation of average annual iron concentrations (ppm) in raw water with respect to distance



Table 2. Case Processing Summary

		N	Percent
Sample	Training	48	68.6%
	Testing	11	15.7%
	Holdout	11	15.7%
Valid		70	100.0%
Excluded		0	
Total		70	

Table 3. Model Summary

Training	Sum of Squares Error	1.798
	Relative Error	.076
	Stopping Rule Used	1 consecutive step(s) with no decrease in error ^a
	Training Time	0:00:00.02
Testing	Sum of Squares Error	.697
	Relative Error	.834
Holdout	Relative Error	3.946

Dependent Variable: WAHDA

a. Error computations are based on the testing sample.

Table 4. Parameter Estimates

Predictor		Predicted		
		Hidden Layer 1		Output Layer
		H(1:1)	H(1:2)	WAHDA
Input Layer	(Bias)	-.325	.087	
	ALKARKH	.304	-.045	
	EASTTIGRIS	.226	.174	
	WATHBA	.082	.155	
	KARAMA	.777	.473	
	QADISIA	-.373	-.176	
	DORA	-.038	-.163	
	RASHEED	.968	-.249	
Hidden Layer 1	(Bias)			.417
	H(1:1)			1.207
	H(1:2)			-1.600

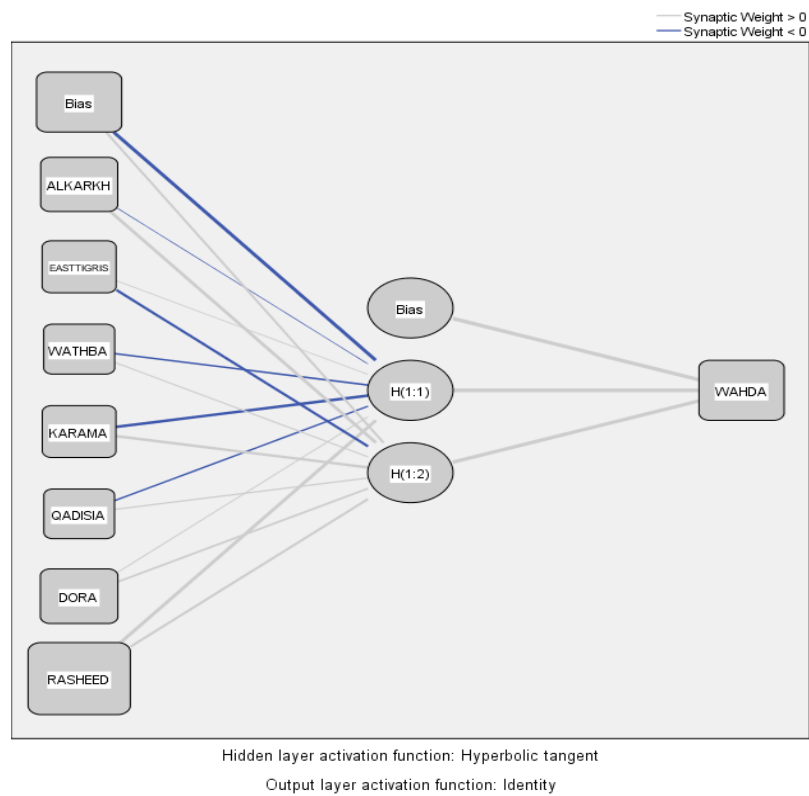


Figure 4. Architecture of the Artificial Neural Network Model

Table 5. Independent Variable Importance

WTP	Importance %
Alkarkh	7.06
Easttigris	10.15
Wathba	6.88
Karama	30.52
Qadisia	12.87
Dora	6.37
Rasheed	26.16

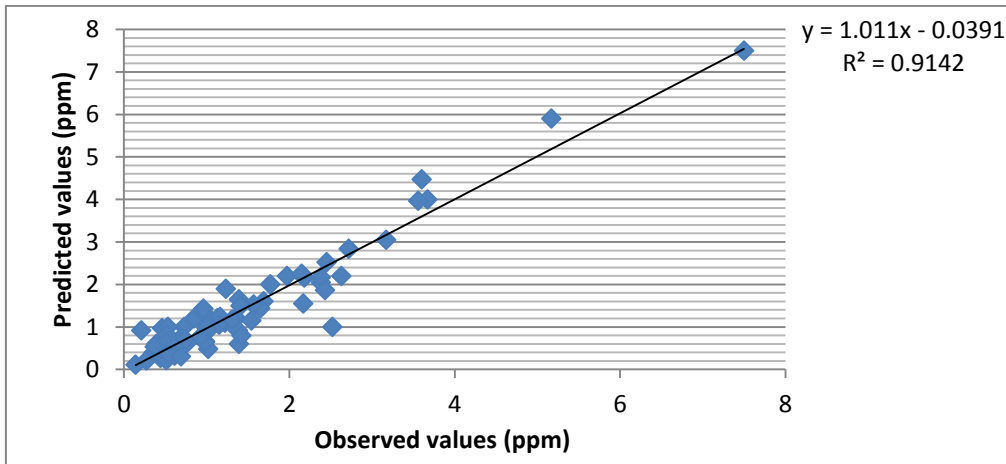


Figure 5. Coefficient of determination of the developed ANN model

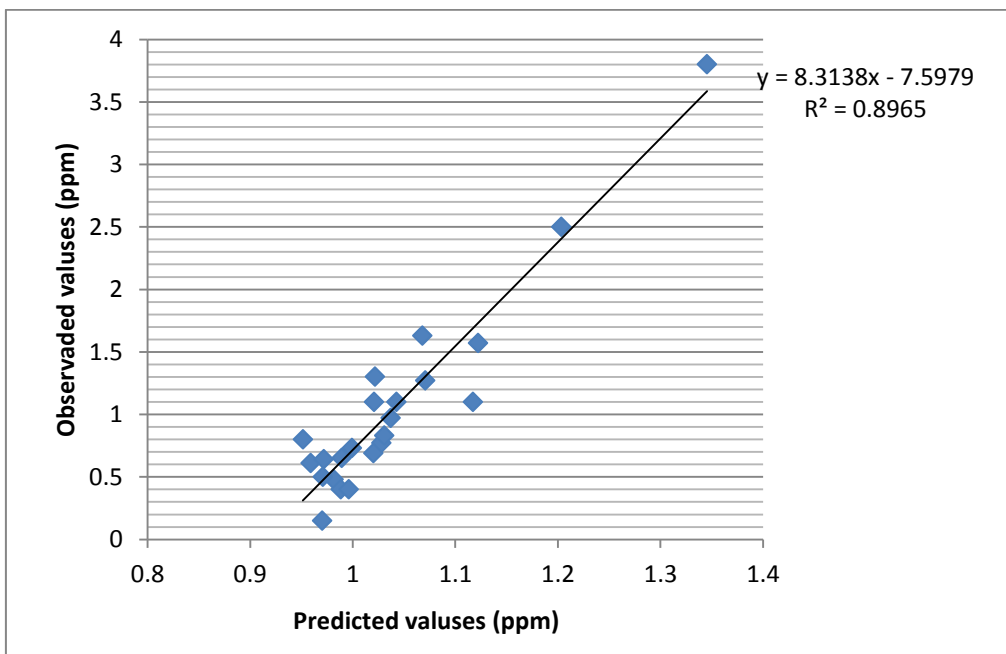


Figure 6. Verification of the ANN model for predicting Iron concentrations at Wahda WTP for years 2012-2013