Comparison between Linear and Non-linear ANN Models for Predicting Water Quality Parameters at Tigris River

Prof. Dr. Rafa Hashim Al Suhili
Department of Civil Engineering
College of Engineering
University of Baghdad

Zainab Jaber Mohammed
Department of Water Resources Engineering
College of Engineering
University of Baghdad
Email:zainab_jaber_86@yahoo.com

ABSTRACT

In this research, Artificial Neural Networks (ANNs) technique was applied in an attempt to predict the water levels and some of the water quality parameters at Tigris River in Wasit Government for five different sites. These predictions are useful in the planning, management, evaluation of the water resources in the area. Spatial data along a river system or area at different locations in a catchment area usually have missing measurements, hence an accurate prediction. model to fill these missing values is essential.

The selected sites for water quality data prediction were Sewera, Numania , Kut u/s, Kut d/s, Garaf observation sites. In these five sites models were built for prediction of the water level and water quality parameters. the following (Biological Oxygen Demand(BOD_5), Phosphate,(PO_4) Sulfate(SO_4), Nitrate(NO_3), Calcium(Ca), Magnesium(Mg), Total Hardness(TH), Potassium(K), Sodium (Na), Chloride (CL), Total Dissolved Solids (TDS), Electric conductivity (EC), Alkalinity(ALK)).

The ANN models tried herein were the Multisite- Multivariate ANN models (5-sites, 14 variables), five models were built, one for each of the five stations as the missing data station. The linear ANN (traditional) models fail to make the prediction of all variables with high correlation coefficient simultaneously. Hence a non- linear input ANN model was developed herein and believed to be a new modification in ANN modeling. It was found that the ANNs have the ability to predict water level and water quality parameters at all the sites with a good degree of accuracy, the range of correlation coefficients obtained are (12.9%-97.2%) for linear models, while for this model with Non-linear terms, The range of correlation coefficients obtained is (71.8%-99.6%).

Key Word: Artificial Neural Network, Water level, The water quality parameters, Tigris River, Non- Linear Model.

المقارنة بين نموذج الشبكة العصبية الأصطناعية ذو المدخلات الخطية و اللاخطية لتخمين معاملات نوعية المياه في نهر دجلة

أ.د. رافع هاشم السهيلي
قسم الهندسة المدنية
جامعة بغداد / كلية الهندسة

زينب جابر محمد
قسم هندسة الموارد المائية
جامعة بغداد / كلية الهندسة

الخلاصة

في هذا البحثِ طُبّقتْ تقنية الشبكات العصبية الإصطناعية (ANNs) في محاولةِ لتَخمين مستوياتِ المياه والبعض مِنْ معايير نوعية المياه في نهرِ دجلةِ في محافظةِ واسط لخمسة مواقعِ مختلفةِ. هذه التخمينات مفيدة في التخطيط و الإدارة و تقييم مصادرِ المياه في المنطقةِ. أَن البيانات المكانية على طول النظام نهري أَو مواقع المنطقةِ المختلفةِ في المنطقةّ لَها قياساتُ مفقودُ عادة، لِذلك فأن من الضروري بناء نموذج تخمينِ دقيقِ لمَلْئ هذه القِيَمِ المفقودةِ.

المواقع المختارة لتنبؤِ بياناتِ معايير نوعية المياه و منسوب الماء هي(الصويرة و النعمانية و الكوت u /s و الكوت d /s والغراف). في هذه المواقعَ الخمسة بُنِيتْ نماذج تخمين لمنسوب الماءو معايير نوعية المياه. معايير نوعية المياه تمثلت في (أحتياج الأوكسجينِ الحيويِ و فوسفات و كبريتات و نترات و كالسيوم و مغنيسيوم و عسرة المياه الكليّة و بوتاسيوم و صوديوم و كلوريد و المواد الصلبة الكلية و التوصيلية الكهربائية و القاعدية).
أستخدام ANN في هذا البحث كان كنموذج متعدد المواقع متعدد المتغيراتُ (5 – مواقع و 14 متغيّر)، حيث تم بناء خمس نماذج واحد لكُلّ مِنْ المواقع الخمس حيث تم أعتبارموقع واحد مفقود لكل نموذج. نماذج ANN الخطيّة (التقليدية) فشلت في تخمين كُلّ المتغيّرات بمعاملِ الإرتباطِ عالي بشكل آني. لذلك تم تطوير نموذج ANN ذو المدخلات اللاخطية للحصول على معاملات أرتباط عالية. و قد تم أثبات قابلية (ANN) لتخمين مستوى الماء و معايير نوعية المياه في كلِ المواقعِ مَع درجة جيدة مِنْ الدقةِ،و كان مدى معاملاتِ الإرتباطِ (12.9%-97.2%) للنموذج الخطي، أما للنموذج اللاخطّي فكانَ مدى معاملاتِ الإرتباطِ (71.8%-99.6%).

الكلمات الرئيسية: الشبكة العصبية الأصطناعية, منسوب الماء, معايير نوعية المياه, نهر دجلة, النموذج اللاخطي.

1. INTRODUCTION

In the Middle East water resources are limited, and are currently decreasing, whilst the demand for water consumption is increasing for reasons such as the expanding industrialization and population growth. This combination makes the distribution of this limited resource very difficult and potentially difficult, **Serwan, 1997**.

In order to utilize the available water resources efficiently, data collection and management such as water levels, discharges and water quality parameters are necessary to conduct this efficient management. A usual practice in Iraq is the missing of some data required, hence models for estimating these data are necessary for proper management of these resources.

ANN was chosen for its ability to generalize results from unseen data and well-suited in modeling dynamic systems on a real-time basis. These properties of ANN are suitable for predication of missing water levels and water quality parameters. ANN has been used in water resources engineering over the last decade, **Rosmina et al., 2007.**

Consequently, ANN usage does not presuppose a detailed understanding of a catchment's physical characteristics, nor does it require extensive data preprocessing .

ANNs provide an appealing solution to the problem of relating input and output variables in complex systems. This has led to the application of ANNs in many fields, including financial management, manufacturing, control systems, design, environmental science and pattern recognition. Despite these advantages, the field has become characterized by a lack of consistency of approach and poor modeling practice. This situation has arisen because the choice of network type, training method(s) and data handling technique(s) has often been undertaken in unsystematic ways by neurohydrologists. The ANN is conceived to imitate the functioning of the human brain by acquiring knowledge through a learning process and finding optimum weights for the different connections between the individual nerve cells**, Saman, 2010.**

River water quality is a significant concern in many countries, considering agricultural and drinking consumptions. Therefore, prediction of water quality parameters, as the main water quality condition is a necessary tool for water resources planning and management, **Niroobakhsh et al, 2012.**

The main objective of this study was to develop a group of Multi-Site Multi -Variables ANN models for predicting missing Water levels and water quality parameters in different observation stations for the Tigris River through Wasit Government and develop a non- linear input artificial neural network model as a modification and compare its results with the traditional linear input model, to investigate whether this modification enhance the performance of the model.

2. HYPOTHESIS OF RESEARCH

The ANN has been used herein to model the water levels and water quality parameters in the Tigris River through Wasit Government depending on a series of such data in selected observed stations in the case study to predict the water levels and water quality parameters in missing points from measured points. As mentioned above this model will be used to mimic the cross-variable cross- site relationship between the observed data. It is believed that the model in this case will adopt weighting matrices that account for these inter-relationships. In addition to that, the usual practice and theory of ANN, is to use a linear weighted sum of the input variables. The non-linearity in these models is only represented by the activation functions of the hidden and output layers. For multi- variables multi- sites modeling, the ANN models may fail in predicting all the variables with the required accuracy simultaneously. Hence, in this research a modification is to be adopted and tested to represent the effect of adding non- linear terms to the weighted sum of the input variables. A technique that expected to improve the performance of the models.

3. ARTIFICIAL NEURAL NETWORK

The human brain provides proof of the existence of massive neural networks that can succeed at those cognitive, perceptual, and control tasks in which humans are successful. The brain is capable of computationally demanding perceptual acts. Artificial neural networks (ANN) have been developed as generalizations of mathematical models of biological nervous systems. A first wave of interest in neural networks (also known as connectionist models or parallel distributed processing) emerged after the introduction of simplified neurons.

The basic processing elements of neural networks are called artificial neurons, or simply neurons or nodes. In a simplified mathematical model of the neuron, the effects of the synapses are represented by connection weights that modulate the effect of the associated input signals, and the nonlinear characteristic exhibited by neurons is represented by a transfer function. The neuron impulse is then computed as the weighted sum of the input signals, transformed by the transfer function. The learning capability of an artificial neuron is achieved by adjusting the weights in accordance to the chosen learning algorithm **, Abraham 2004.**

4. CASE STUDY SITE

The case study is the reach of Tigris River that passes through Wasit
Government. Wasit Government is located in the southern part of the central region of Iraq between longitudes 44.52393°, 46.64164° and latitudes 31.93202°, 33.49615°.
bordered to the north by the capital Baghdad and the north-east by the Government of Diyala, Republic of Iran from the east and the south-east Government of Maysan from the south by the Government of Thi Qar. From the west Governments of Babil, Diwaniyah. Kut is the capital city of Wasit Government, which is located (180 km) south of the capital Baghdad.

The studied area is the Tigris River reach passing through Wasit Government; it is bounded by the coordinate: 47.7437°E, 36.48050°N and 60.7030°E, 36.13808°N

The Tigris River at this reach has the branches: Dujaili, and Garaf, Shatt Al-Shatra and Shatt heresy. The area has a climate transition between the Mediterranean climate and desert climate hot and dry, few rainfalls and high temperature, and start temperatures rise in March, peaking in July and August.

Fig.1 shows the map of Wasit Government. **Fig. 2** shows the schematic diagram of the observation stations of water quality parameters.

The data that used in producing the ANN models was recorded on a monthly mean basis during one year (2011) to observe the water levels and water quality parameters.

5. ARTIFICIAL NEURAL NETWORKS MODELING OF THE CASE STUDY

Multi- Variable Multi- Sites Model

This model was applied for predicting multi variables in many sites. It is abbreviated as MNPS for one missing site where Nis the number of missing site, N=1 and P for prediction, S mean the sites of the missing variables, S= A(Sewera site), B(Numania site), C(Kut u/s site), D(Kut d/s site) or E(Garaf site), **Fig. 2** show the distribution of the observed sites on the Tigris River.

Multi- Variables Multi- Sites Model, ANN Linear Model

These models had been built to predict fourteen variables simultaneously. These variables are (Biological Oxygen Demand, Phosphate, Sulfate, Nitrate, Calcium, Magnesium, Total Hardness, Potassium, Sodium, chloride, Total Dissolved Solids, Electric conductivity, Alkalinity).These models were built for the five sites.

One Missing Site Model (M1)

These models were built considering one missing site in every model and predicted its water level and water quality parameters from the other observed sites, five models have been built one for each site. In these models the number of predicted variables are fourteen. Many trials had been conducted to build the model using the ANN modeling, incorporated in the SPSS-17 software (Neural Network-Multilayer perceptron). Among these many trials the best one for each model is shown in **Table 1.** These trials include the change of the selected parameters of the ANN model, a method usually used to get good correlation coefficients. These parameters are the percentage division of the data into training, testing and verification (holdout) sub-sample, learning rate, number of hidden nodes, and type of activation functions of both the hidden layer and output layer. The problem here when this techniques is used for multivariate multisite prediction ANN models is that in a certain trial one may get high correlation coefficients for some of the output variables and low correlation coefficient for the other output variables, since in this case the number of output variables is large. Changing the ANN parameters selection results in increasing some of the correlation coefficients, while decreasing the others. **Table1**. shows the problem clearly, for example for the first station, the correlation coefficient are high enough for Ca, Mg, Na, CL , TDS, EC and ALK (> 0.8) according to ,**Smith criteria , 1986.** While for the other variables it is not accepted, For purpose of illustration the TDS value predicted for Sewera station using M1PA , **Fig.3** shows the comparison of the predicted values of TDS against the observed of Sewera station. **Fig.4** gives the time series of TDS observed and predicted.

Non- Linear Input Patterns FFBP Artificial Neural Networks

Since the Linear input Multi- Variables Multi- Sites model, indicates that it is not possible to get high correlation coefficients for all the variables simultaneously. Hence, it is believed that the reason may be due to the effect of a non-linear relationship between the corresponding variables of low correlations with their corresponding values in the other sites. This is not modeled in the ANN traditional method, hence a modification was made here in to solve this problem , as shown below:

6-DEVELOPMENT OF THE ANN LINEAR MODEL FOR THE CASE STUDY

In order to illustrate the modification of the ANN model tried herein, a brief discussion of the methodology of traditional ANN modeling is presented below The Artificial Neural Network modeling in engineering is well known by now as a relatively new modeling new modeling compared with regression methods for relating a set of output (dependent) variables to a set of input (independent) variables

Theories and literature on ANN modeling focus on the capability of such model to represent the non-linearity exist in many system dependent independent relationships.

Investigation of the methodology of ANN modeling reveals the process in the following steps referring to the three- layers model, shown in **Fig.5.**

A-Feed- forward process

This process was presented by Al- Suhaili by the following sub-steps.

1.For a given set of input variables vector$[x]_{n\times1}$ the vector$[z_{in}]_{p,1}$ of the inputs to the hidden nodes of the hidden layer is estimated using the following equation:

$$[z_{in}]_{p,1}= [v_{\circ\ bias}]_{p,1}+ [v]_{n\times p}{}^{T}\ [x]_{n\times1}\text{eq.} \quad (1)$$

Where:

$[v_{\circ\ bias}]_{p,1}$: is the bias weight vector of the hidden layer.

$[v\ ]_{n\times p}$: is the weight matrix modeling relating input and hidden layers.

2. Find the output vector of the hidden nodes $[Z]_{p\times1}$using:

$$[z]_{p\times1}=F_1[z_{in}]_{p\times1} \quad \text{eq.} \quad (2)$$

Where F_1: is the hidden layer activation function.

3. Find the input vector to the output layer nodes, $[y_{in}]_{m\times1}$using:

$$[y_{in}]_{m\times1}= [w_{\circ\ bias}]_{m\times1}+ [w]_{p\times m}{}^{T}\ [z]_{p\times1} \quad (3)$$

Where:

$[w_{\circ\ bias}]_{m\times1}$: is the bias weight vector of the output layer.

$[w]_{p\times m}$: is the weight matrix relating hidden and output layers.

4. Finding the output vector $[y]_{m\times1}$using

$$[y]_{m\times1}=F_2[y_{in}]_{m\times1} \quad (4)$$

B-Back-propagation process

The Back-Propagation Process, utilizes the error estimated from the difference in output vector $y_{m\times 1}$estimated by eq. (4) above from the real observed output vector $[t]_{m\times 1}$, to find the correction in the weight vectors and matrix $v_{\circ\, bias}$, $w_{\circ\, bias}$,$v_{n\times p}$ and $w_{p\times m}$, which usually assumed zeros at the first step in application using certain algorithm such as gradient decent method . Equations for corrections in weight are available in text books or papers related to the subject.

Investigating the steps in the Feed- forward Process above prevails that the summation weighted blocks in the hidden and output layers are all a linear weighted summation process for example:

$$z_{in(1)}= x_1\ v_{11}+ x_2\ v_2+x_3\ v_{31}+\ldots\ldots\ldots\ldots x_n\ v_{n1}+v_{\circ bias(1)} \quad (5)$$
$$y_{in(1)}= z_1\ w_{11}+ z_2\ w_2+z_3\ w_{31}+\ldots\ldots\ldots\ldots z_n\ w_{n1}+w_{\circ bias(1)}. \quad (6)$$

The only non-linearity existing in the model could be observed is due to a non-linear activation functions F1and F2 in eq. (5) and (6) respectively. If both F1and F2are of linear (identity) type the process in fully linear.

In order to simulate the non-linear relationship between the dependent and independent variables, the following pre-step is suggested, as a modification .This pre-step is illustrated first by an example of n=3.

Pre-step;

For n=3, we have the independent variablesx_1, x_2 and x_3, for a second degree model a pre-step matrix shown below could be used, for additional variables to be included.

$$\begin{array}{c|ccc} & x_1 & x_2 & x_3 \\ \hline x_1 & x_1^2 & x_1\ x_2 & x_1 x_3 \\ x_2 & x_2\ x_1 & x_2^2 & x_2 x_3 \\ x_3 & x_3\ x_1 & x_3 x_2 & x_3^2 \end{array} \quad (7)$$

Hence 6- variables could be added, in general$n^2 - n$, since the matrix is symmetrical.

Four the case of fourteen input variables of five sites the total number of variables that should be added is $n^2 - n$, with $n = 14 \times 5 = 70$, will be very large. Hence, a reduction is suggested by taking the diagonal variables only, i.e. neglecting the cross variables terms, leaving only the square of each variable, in this case the number of added terms is n ($n = 3$, in this example) as shown in eq.(7).

If we have one missing site is considered out of five, and fourteen variables in each site, then the total No. of the other site (with non- missing values). This number of added variables can be further reduced if only the variables of low correlation coefficients obtained by the traditional ANN modeling process is considered n_l , Hence the total No. of variables added are $n_l \times n_v$, where $n_v$are the No. of output variables, which are the square of each variable of n_l in each site. Furthermore, it may not be necessary to include the square values as additional variables in all of the other sites than the missing one. The selection of the sites where these additional variables should be included in a trial and error procedure.

Where 3-new variables could be added as

$$\begin{bmatrix} x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \end{bmatrix} \quad (8)$$

Then the weighted summation will include non-linear term as show in eq. (9) compared to Eq. (5).

$$z_{in(1)}= x_1\ v_{11}+ x_2\ v_{21}+x_3\ v_{31}+x_1^2\ v_{41}\ +x_2^2\ v_{51}+x_3^2\ v_{61}+v_{\circ bias(1)} \quad (9)$$

Similar equation can be written for z_{in1} toz_{inp}

7-APPLICATION OF MODIFICATION ANN MODEL (NON-LINEAR INPUTS MODEL)

The modification explained in section above was applied for the Multi- Variables Multi-Sites. For purpose of comparison. This model will be referred to here after as non-linear ANN, which the usual ANN model will be referred to as a linear model

For purpose of checking the ability of modification made here, the application was performed for Multi-Variables Multi-Sites (one missing site models (M1PS)

These models were built considering one missing site in every model and predicted their water level and water quality parameters from the other observed sites, using the Non- linear input model. **Table.2** shows the comparison of the correlation coefficient of each station between the linear and Non-linear models. It is clear that the obtained correlation of the Non-linear model is much higher than these of linear model

To illustrate the trial and error method that led to the final Non-Linear models obtained in **Table 2.** **Table 3.** shows the trials, in the first column the two low variables NO3, TH are square and added as additional variables in all of the other four stations. The second column shows these variables squared only for stations B and C, while the third columns it is squared only to the nearest station B. It is clearly shown that the third trial gave the highest correlation coefficients. For purpose of illustration the TDS value predicted for Sewera station using M1PA , **Fig.6** show the comparison of the predicted values of TDS against the observed of Sewera station. **Fig. 7** gives the time series of TDS observed and predicted.

8-CONCLUSION

1. The traditional ANN model for multisite multivariate prediction fails to give high correlation coefficients for all the predicted variables simultaneously. Changing the selected ANN model parameters such as data division, No. of hidden nodes in the hidden layer, learning rate and type of activation functions also not solve the problem. The result of these changes may increase the low correlation and still cannot observed high correlation for all the variables simultaneously
2. The modification adopted here to change the traditional ANN multisite multivariable linear ANN model to the Non- linear input pattern model had solved the problem of correlation coefficients. The results indicate the capability of the modified model to get high correlation for all the variables simultaneously
3. The selection of the added Non-linear variables is a process that should be done by trial and error procedure until getting the desired correlation coefficients. This process is similar to the process of step- wise regression in a multivariate regression analysis

REFERENCES

- Abraham, A. , 2004, *Meta-Learning Evolutionary Artificial Neural Networks*, Neurocomputing Journal, Vol. 56c, Elsevier Science, Netherlands,(1–38).)

- Rosmina Bustami, Nabil Bessaih, Charles Bong, and Suhaila 2007, *Artificial Neural Network for Precipitation and Water Level Predictions of Bedup River*. (Advance online publication: 17 November) IAENG International Journal of Computer Science, 34:2, IJCS_34_2_10.

- Serwan M. J.,1997, Baban ,*Potential Applications of Satellite Remote Sensing (SRS) and GIS in Maximizing the Use of Water Resources in the Middle East*: examining Iraq as a case study" IAHS Publ. no. 242.
- Smith, M. , 1993, *Neural Networks for Statistical Modeling*, Van Nostrand- Reinhold New York

- Saman.H. Al-Hamawandi, 2009, *Watershed Modeling Using Artificial Neural Networks*" P.HD. Thesis, Water Resources Eng. Dept. University of Baghdad

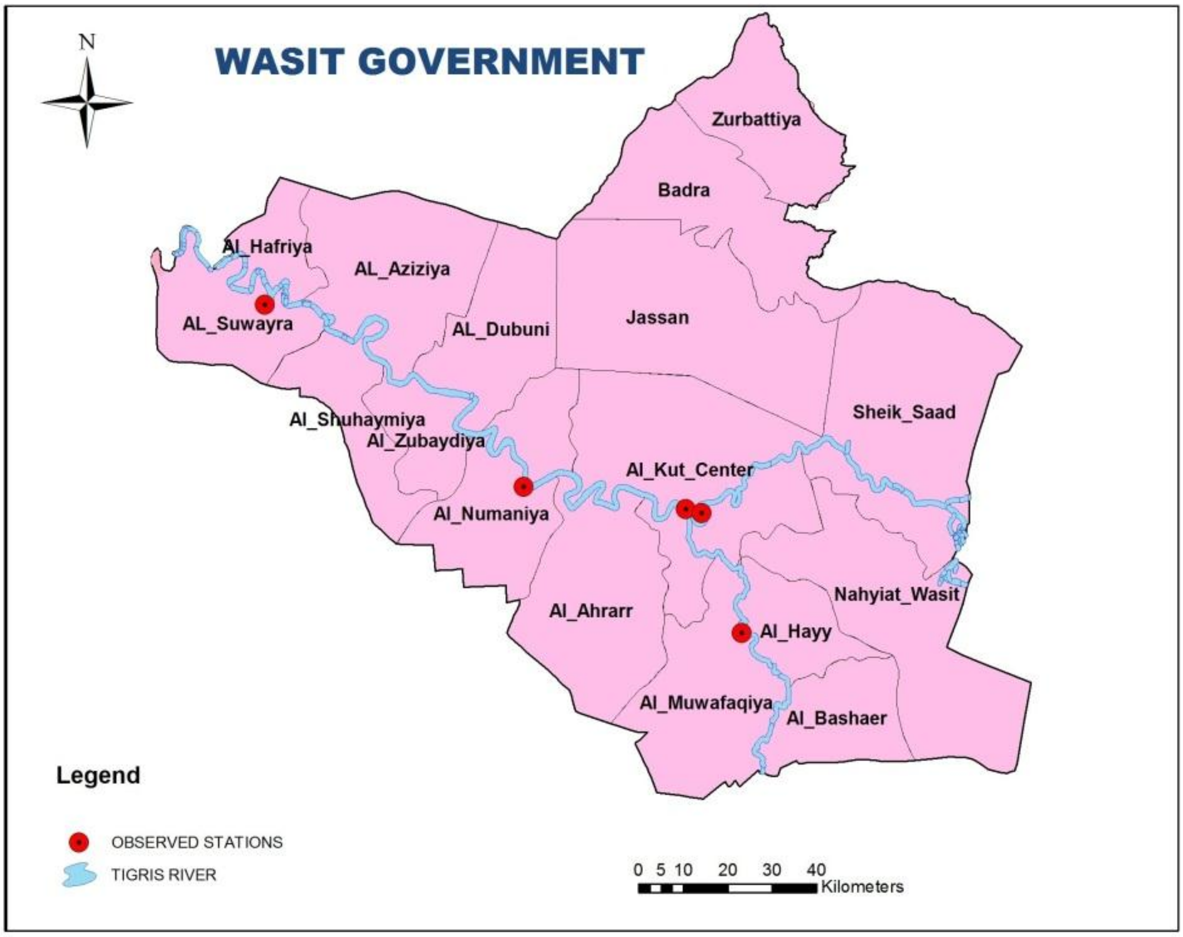

Figure 1. Map of Wasit Government. (Directorate of the Environment for Wasit Government).

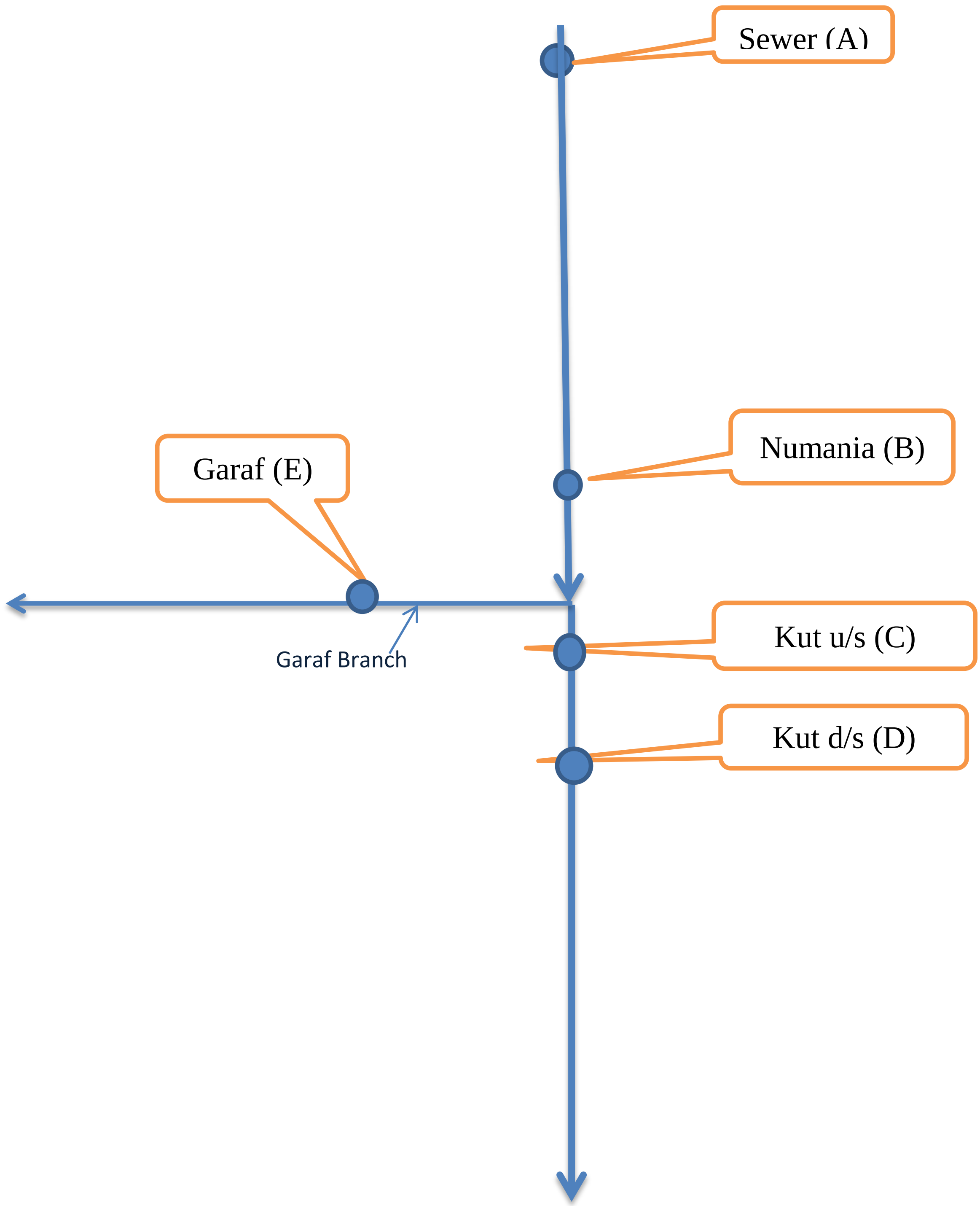

Figure 2. Schematic diagram of Tigris River with the observation sites for water quality parameters in Wasit Government

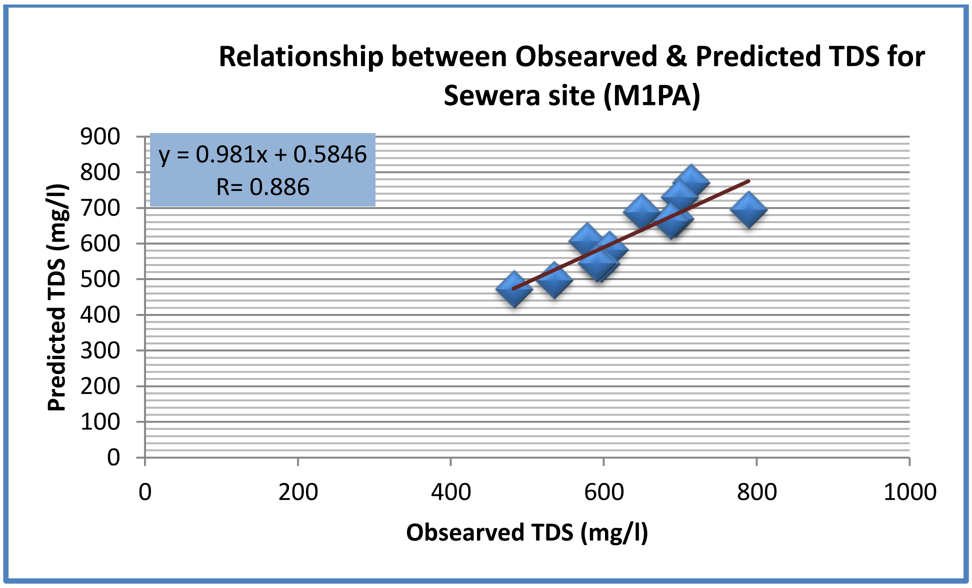

Figure 3. Relationship between observed and predicted TDS for Sewera site (M1PA-Linear Model).

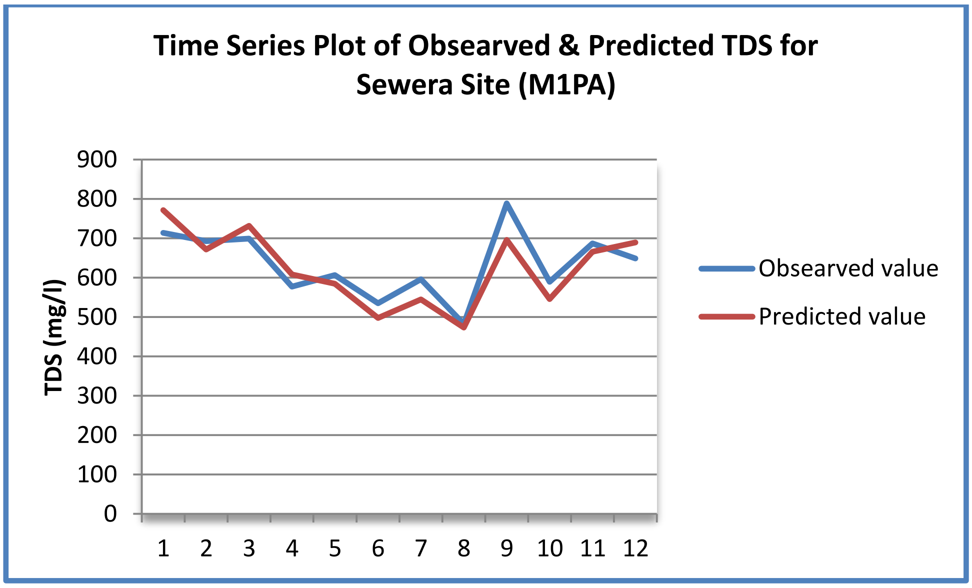

Figure 4. Time series plot of observed and predicted TDS for Sewera site (M1PA-Linear Model).

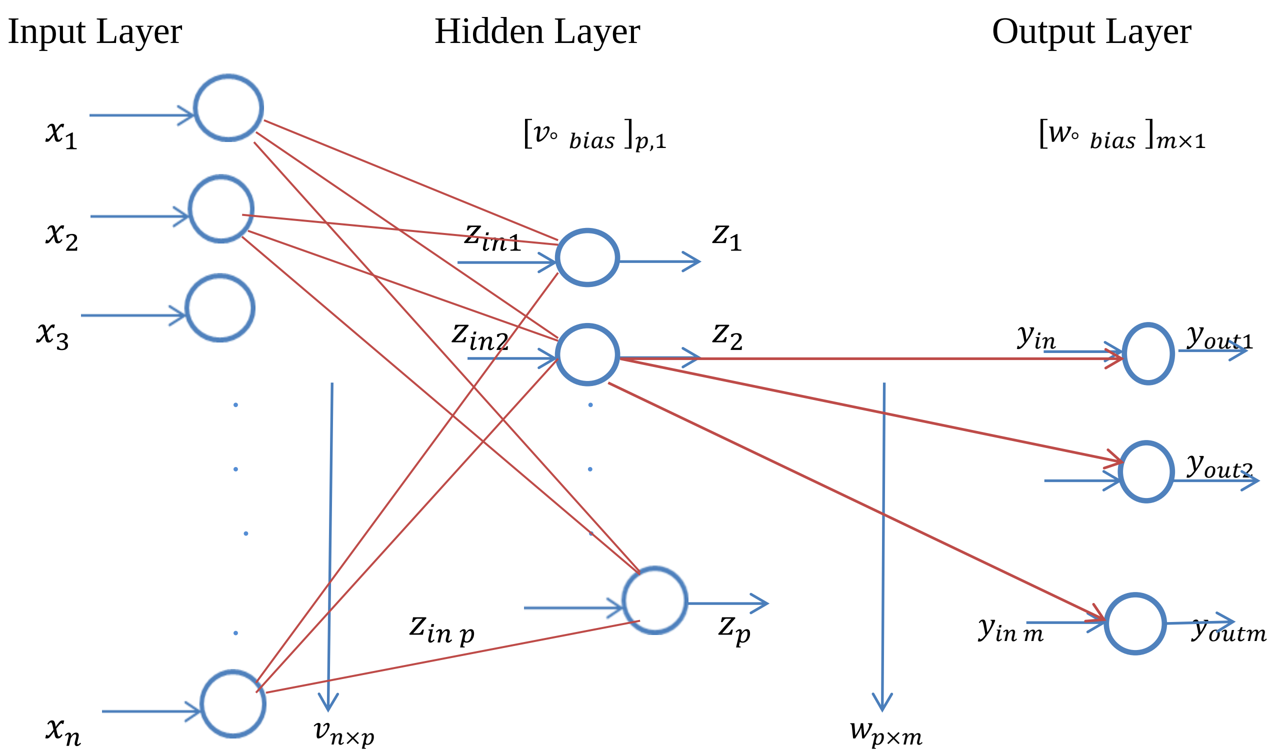

Figure 5. Three- layer standard ANN network model.

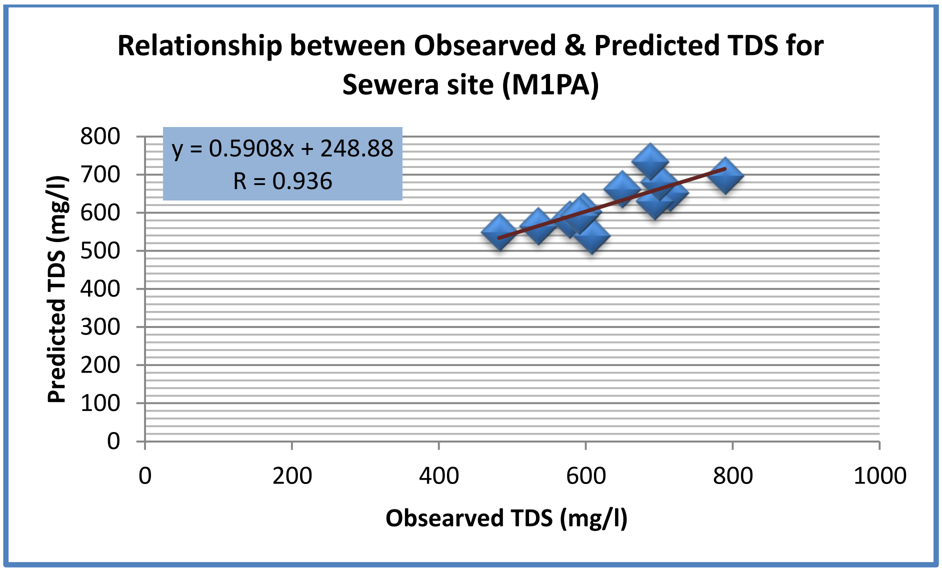

Figure 6. Relationship between observed and predicted TDS for Sewera site (M1PA-Nonlinear Model).

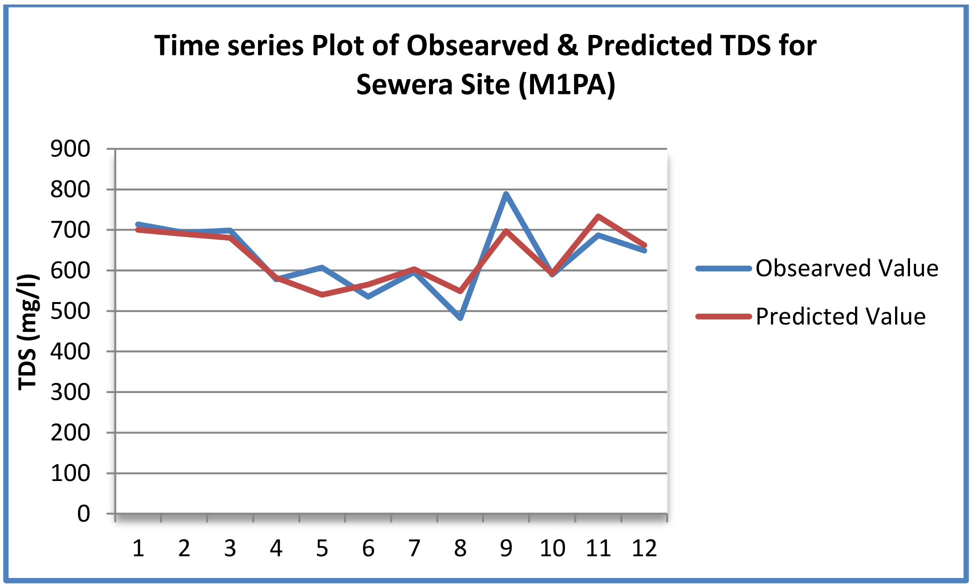

Figure 7. Time series plot of observed and predicted TDS for Sewera site (M1PA-Nonlinear Model).

Table 1. The correlation coefficient (R) for prediction the water level and water quality parameters one missing site model.

Parameters	M1PA	M1PB	M1PC	M1PD	M1PE
W.L	0.779	0.891	0.618	0.836	0.923
BOD5	0.597	0.392	0.853	0.935	0.842
PO4	0.681	0.59	0.768	0.917	0.945
NO3	0.320	0.771	0.854	0.842	0.551
Ca	0.915	0.858	0.929	0.965	0.900
Mg	0.811	0.657	0.774	0.733	0.754
TH	0.129	0.744	0.933	0.838	0.870
K	0.701	0.854	0.872	0.940	0.931
Na	0.819	0.609	0.934	0.855	0.933
SO4	0.582	0.657	0.933	0.456	0.902
CL	0.801	0.450	0.921	0.976	0.860
TDS	0.886	0.771	0.964	0.811	0.765
EC	0.839	0.591	0.903	0.884	0.818
ALK	0.846	0.764	0.923	0.972	0.888

Table 2. The comparison between correlation coefficient obtained by linear and non- linear ANN model.

Parameters	M1PA		M1PB		M1PC		M1PD		M1PE	
	Linear	Non-linear	Linear	Non-linear	Linear	Non-linear	Linear	Non-linear	Linear	Non-linear
W.L	0.779	0.718	0.891	0.761	0.618	0.909	0.836	0.986	0.923	0.932
BOD5	0.597	0.936	0.392	0.742	0.853	0.930	0.935	0.995	0.842	0.897
PO4	0.681	0.801	0.59	0.771	0.768	0.881	0.917	0.988	0.945	0.868
NO3	0.320	0.893	0.771	0.785	0.854	0.856	0.842	0.995	0.551	0.977
Ca	0.915	0.944	0.858	0.787	0.929	0.943	0.965	0.960	0.900	0.994
Mg	0.811	0.952	0.657	0.887	0.774	0.881	0.733	0.988	0.754	0.974
TH	0.129	0.938	0.744	0.868	0.933	0.984	0.838	0.985	0.870	0.985
K	0.701	0.914	0.854	0.792	0.872	0.898	0.940	0.972	0.931	0.987
Na	0.819	0.921	0.609	0.763	0.934	0.906	0.855	0.996	0.933	0.975
SO4	0.582	0.902	0.657	0.966	0.933	0.959	0.456	0.980	0.902	0.982
CL	0.801	0.921	0.450	0.765	0.921	0.984	0.976	0.987	0.860	0.991
TDS	0.886	0.936	0.771	0.893	0.964	0.993	0.811	0.980	0.765	0.963
EC	0.839	0.940	0.591	0.477	0.903	0.810	0.884	0.994	0.818	0.970
ALK	0.846	0.971	0.764	0.961	0.923	0.956	0.972	0.995	0.888	0.973

Table 3. The correlation coefficient for the trials for ANN-nonlinear model for predicting the parameters for Sewera site.

The variables selected	Station selected for adding squared variables		
NO3, TH	B,C, D, E	B, C	B
Parameters	M1PA,T1	M1PA,T2	M1PA,T3
W.L	0.850	0.469	0.718
BOD5	0.612	0.479	0.936
PO4	0.687	0.448	0.801
NO3	0.708	0.162	0.893
Ca	0.407	0.803	0.944
Mg	0.866	0.736	0.952
TH	0.732	0.854	0.938
K	0.854	0.844	0.914
Na	0.772	0.713	0.921
SO4	0.763	0.800	0.902
CL	0.720	0.869	0.921
TDS	0.968	0.650	0.936
EC	0.856	0.831	0.94
ALK	0.647	0.684	0.971