Thi-Qar University Journal for Engineering Sciences ISSN 2075 – 9764
http://www.doi.org/10.31663/tqujes.9.2.331(2018) Available at http://jeng.utq.edu.iq
Vol.9 No.2 (September 2018) engjrnal@jeng.utq.edu.iq

Daily Discharge Prediction Using Artificial Neural Networks (ANNs) For Al Gharraf River in Thi Qar Province, Iraq

Ahmed A. Dakheel[1]
[1]University of Thi-Qar, Thi Qar, Iraq.
Email: msc_ahmed@utq.edu.iq

Abstract

In the present study an Artificial Neural Networks (ANNs) model has been developed for Al Gharraf River in Thi Qar Province, Iraq. The modeled network is trained, validated and tested using daily discharge data pertaining to 3 years (January 2014 to January 2017) for four stations on the river Al Gharraf (Regulator II, Regulator III, Regulator IIII and Al Badaa). The number of hidden neurons is estimated according to trial and error procedure. The best model is selected according to based root mean square error (RMSE), mean absolute error (MAE) and coefficient of correlation (R). The results showed the optimum numbers of neuron in hidden layer is equal to 10 and indicate that the ANNs is effective technique for forecasting the river discharge, which are utmost essential to hydrologists around the globe.

Keywords: Artificial Neural Networks (ANNs), Regulator, Al Badaa, Al Gharraf River, Thi Qar Province, Iraq.

1. Introduction

The process of predicting river stream flow quickly and accurately is critical operation in flood forecasting. Prediction of stream flow are vital important for flood caution, operation of flood control purposed reservoir determination of river water potential, production of hydroelectric energy, allocation of domestic and irrigation water in drought seasons and navigation planning in rivers [1]. In spite of need to great amount of data and human effort to calibrate, validate and test the model physics based methods are useful to understand the entire of underlying process [2]. In the other hand, good advantage of forecasting models is that they just need limited amount of data, but their drawback is that they seem like a black box in usages and also need lengthy parameterization [3]. In the recent years, new techniques and algorithms have been applied as a powerful tool for modeling the problems of water resources. Artificial Neural Networks (ANNs) is one of them. ANNs have been used as successfully tool for solving many different types of hydrological problems [4].

Artificial Neural networks (ANNs) have been well suited and successfully applied in variety fields of hydrology including water resources [5]. ANNs techniques applied to hydrologic time series and forecasting have shown better performance than the classical techniques [6]. Studies have been reported in literature: Jain et. al., used Artificial Neural Networks (ANNs) methods to predict reservoir inflows in reservoir operation. They compared Artificial Neural Networks (ANNs) and Autoregressive Integrated Moving Average (ARIMA) models and concluded that ANNs yielded better result [7]. Zealand et. al., investigated the utility of Artificial Neural Networks (ANNs) for short term forecasting of stream flow [8]. Birikundavyi et. al., investigated the performance of Artificial Neural Networks (ANNs) method in prediction of daily stream flow. The results shown that ANNs method yielded better results than Autoregressive Moving Average (ARMA) models [9]. Huang et. al., compared Artificial Neural Networks (ANNs) and Autoregressive Integrated Moving Average (ARIMA) models in stream flow forecasting. The results showed that ANNs model is better than ARIMA model [10]. Muhammad et. al., have made a study on forecasting ground water contamination using Artificial Neural Networks (ANNs). The results obtained from the model were compared with actual values as well as the World Health Organization Standards [11]. Mafia et. al., studied the use of a Neural Network technique for the prediction of water quality parameters. The results demonstrated the ability of the appropriate ANNs models for the prediction of water quality parameters [12]. Mohsen and Zahra, studied on the application of Artificial Neural Networks (ANNs) for temperature forecasting. The results showed that Multi-Layer Perceptrons network has the minimum forecasting error and can be considered as a good method to model the short-term temperature forecasting systems [13]. Akhtar et. al., studied on river flow forecasting with ANNs using satellite observed precipitation preprocessed with flow length and travel time information in Ganges river basin. The ANNs showed its ability to forecast discharges 3 days ahead with an acceptable accuracy [14]. Najah et. al., carried out a study on prediction of Johor river water quality parameters using ANNs. The results showed that the proposed ANNs prediction model has a great potential to simulate and predict the total dissolved solids, electrical conductivity, and turbidity with absolute mean error 10% for different water bodies [15]. Ali H. Al-Aboodi, studied on prediction of Tigris river discharge in Baghdad City using artificial neural networks. Results indicated that the ANNs with Levenberg Marquradt back-propagation algorithm are a powerful tool for forecasting the river discharge for short term duration [16].

In this study, Daily prediction of Al Gharraf River flow discharge in Thi Qar Province by using Artificial Neural Networks (ANNs) approach is the main objective of this research, In addition to investigate the capability of ANNs model for predicting the discharge from water level by changes the temporal and spatial and to know the reasons for these changes.

2. Artificial Neural Networks (ANNs)

The artificial neural networks are composed of a set of artificial neurons which are inspired by biological systems. An ANNs is a massively parallel-distributed information processing system that has certain performance characteristics resembling biological neural networks of the human brain [17]. The model of a neuron is represented in figure (1).The most efficient neural network training algorithm is the back propagation algorithm [18]. Back propagation algorithm is used in this paper. Back propagation (BP) is a gradient descent algorithm in which the gradient is computed for nonlinear multilayer networks. The ANNs parameters (weights and biases) can be adjusted to minimize the sum of the squares of the differences between the actual values and network output values [19].

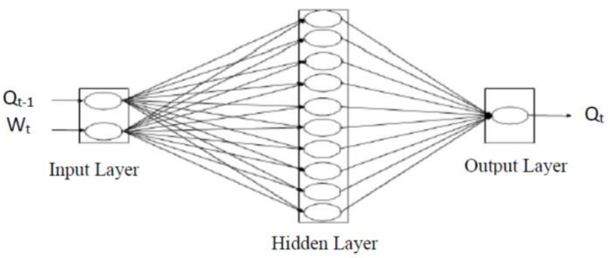

Figure 1 Architecture of multilayer perceptron of Artificial Neural Networks [20].

In any ANNs there are a number of data processing elements called neurons, which are grouped in layers. Neurons of the first layer which is called input layer receive the input vector and transfer the values to the next layer nodes or neurons across connections. This process is continued until the output layer is reached. ANNs can be categorized into two types according to the number of layers: single bi layer and multilayer networks and also can be categorized into feed forward and feed backward networks due to the direction of the information and processing [21]. During training, the contact weights are adjusted so as to reduce the squared difference between the desired output and the processing elements response. The optimal weights are the product of the inverse of the input autocorrelation matrix (R^{-1}) and the cross-correlation vector (P) between the input and the desired response. The analytical solution of this problem is equivalent to a search technique to obtain the minimum of the quadratic performance surface (wi), using gradient descent by adjusting the weights at each epoch [18]:

$$w_i(k+1)=w_i(k)-\eta\nabla Ji(k) \qquad \nabla Ji=\partial J\partial wi \tag{1}$$

Where:

η: coefficient of learning rate.

$\nabla(k)$: gradient vector of the performance surface at iteration (k) for the i^{th} input node.

Performance surface (J) is calculated by equation (2):

$$J=\Sigma_p(d_p-y_p)^2 \quad and \quad \min J\rightarrow w_{opt}=R^{-1}P \tag{2}$$

Where:

w_{op}: Optimal weight.

d_p: Target output.

y_p: Calculated output of the p^{th} output neuron.

Three layer feed forward neural network used in this study, which have been widely used for water resources modeling, because these layers are sufficient to generate arbitrarily complex output signals [22]. The output value of multilayer perception is calculated as follows [23]:

$$y_k = f_{\circ}[\sum_{i=1}^{M_N} W_{Kj}\,.\,f_h(\sum_{i=1}^{N_N} W_{ji}X_i + W_{j^\circ}) + W_{K^\circ}] \tag{3}$$

Where:

W_{ji}: A weight in the hidden layer connecting the i^{th} neuron in the input layer and the j^{th} neuron in the hidden layer.

W_{jo}: The bias for the j^{th} hidden neuron.

f_h: The activation functions of the hidden neuron.

W_{kj}: A weight in the output layer connecting the j^{th} neuron in the hidden layer and the k^{th} neuron in the output layer.

W_{ko}: The bias for the k^{th} output neuron.

f_o: The activation functions for the output neuron.

X_i: i^{th} input variable for input layer.

y_k: computed output variable.

M_N and N_N: The number of the neurons in the input and hidden layers, respectively.

3. Study Area and Data Set

The Province of Thi Qar is located at the heart of the southern part of Iraq and shares internal borders with provinces (Basrah, Muthanna, Qadissiya, Wassit and Missan) as shown in figure (2). The best part of the area of the Province is situated next to the Rivers of Euphrates and Al Gharraf. Al Gharraf River is one of the branches of the Tigris River. Al Gharraf River branches of the Tigris River at the Wassit Province and passes through the of Thi Qar Province in south of Iraq and the length of river is 140 Km. It is distinguished from the other branches of the Tigris River, the length and density of the population of the cities passing by where the population is estimated at 998,729 [24].

The data set utilized in this study was produced through monitoring of the discharge and water level of Al Gharraf River. Daily data was collected during the period of three years (January 2014 to January 2017). Four sampling stations (Regulator II, Regulator III, Regulator IIII and Al Badaa) are identified in figure (2) and Table (1), was obtained from the Iraqi Ministry of Water Resources. Table (2) and Table (3), show summary statistics of the raw data.

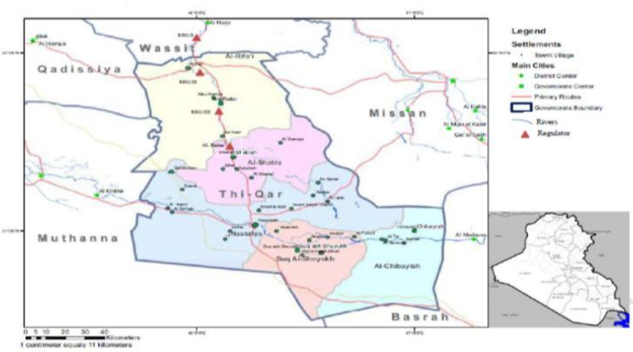

Figure 2 Location of study area in reference to map of Iraq [25].

Table 1 Global positioning of the sampling points.

Station	Geographic Coordinate	
	Latitude	Longitude
Regulator II	32° 10' 50.3" N	46° 01' 21.7"E
Regulator III	31° 53' 00.3" N	46° 03' 26.7" E
Regulator IIII	31° 39' 30.3" N	46° 06'14.2" E
Al Badaa	31° 26' 44.3" N	46° 10' 33.8" E

Table 2 Summary statistics of the raw data for the discharge.

Station	Number of Observations data	Average (m^3/s)	Standard Deviation	Skewness Coefficient	Excess Kurtosis	Median (m^3/s)	Minimum (m^3/s)	Maximum (m^3/s)
Regulator II	1147	103.8457	21.7704	0.4726	2.1606	106	53	255
Regulator III	1147	92.1975	18.0838	0.1814	0.1389	92	40	168
Regulator IIII	1147	72.4838	14.3977	0.0654	-0.0995	73	9	118
Al Badaa	1147	21.1394	6.6004	0.3808	0.2552	21	6	45

Table 3 Summary statistics of the raw data for the water level.

Station	Number of Observations data	Average (m)	Standard Deviation	Skewness Coefficient	Excess Kurtosis	Median (m)	Minimum (m)	Maximum (m)
Regulator II	1147	11.0811	0.4194	-0.0170	-0.6064	11.1	10.1	12.2
Regulator III	1147	10.4018	0.3267	0.3139	-0.3679	10.37	9.6	11.4
Regulator IIII	1147	9.3522	0.2737	0.8650	1.1766	9.3	8.75	11.1
Al Badaa	1147	6.1033	0.4243	1.2489	2.1482	6	5.3	7.85

4. Methodology

The performance of different prediction models can be evaluated in terms of appropriate fit after each model structure is calibrated using the training, validation data set and the test data set. For each model (M1, M2, M3 and M4) of (Regulator II, Regulator III, Regulator IIII and Al Badaa), root mean squared error (RMSE) equation (4), mean absolute error (MAE) equation (5) and correlation coefficient (R) equation (6) were used as evaluation criteria [26].

$$\text{RMSE} = \left(\frac{\sum_{j=1}^{n}(Y_j - \hat{Y}_j)}{n}\right)^{1/2} \tag{4}$$

$$\text{MAE} = \frac{\sum_{j=1}^{n}|Y_j - \hat{Y}_j|}{n} \tag{5}$$

$$\text{R} = \frac{\sum_{j=1}^{n}[(Y_j - \bar{Y})(\hat{Y}_j - \bar{\hat{Y}})]}{\left[\sum_{j=1}^{n}(Y_j - \bar{Y})^2 \sum_{j=1}^{n}(\hat{Y}_j - \bar{\hat{Y}})^2\right]^{1/2}} \quad \& \quad R = \sqrt{R^2} \tag{6}$$

Where:

Y & $\hat{Y}$: The observed and estimated values respectively.

n: The number of observations.

$\bar{Y}$&$\bar{\hat{Y}}$: The mean of observed and estimated values.

Three layers feed forward network with sigmoid hidden neurons and linear output neurons are used in this research. A number of networks with different numbers of hidden layer nodes (1-20) and with different transfer functions were developed. The network is trained with Levenberg-Marquradt back-propagation algorithm. The data set is scaled by using mapminmax function according to this scale the range of the input lies inside the range ($-1 \le x \le 1$). Hence the total number of observations is 1147 samples for each model, these observations are divided into three statistically parts. 70% (803 samples) for training, these are presented to the network during training, and the network is adjusted according to its error. 15% (172 samples) is used as a validation part; these are used to measure network generalization, and to halt training when generalization stops improving. The last part of data set is testing part 15% (172 samples) these have no effect on training and so provide an independent measure of network performance during and after training. The early stopping method is selected to overcome over fitting problem. A trial and error procedure based on root mean square error (Eq.4), mean absolute error (Eq.5) and coefficient of correlation (Eq.6) are used to select the best network architecture and perform of ANNs for predicting the discharge of Al Gharaf River.

5. Results and Discussion

One preceding water level value is used for predicting the discharge value. Four models are adopted for selecting the best one. For each model are described as follow:

$$M1: Q_t= f (W_t) \quad (7)$$

$$M2: Q_t= f (W_t, Q_{t-1}) \quad (8)$$

$$M3: Q_t= f (W_t, W_{t-1}, Q_{t-1}) \quad (9)$$

$$M4: Q_t= f (W_t, W_{t-1}, W_{t-2}, Q_{t-1},Q_{t-2}) \quad (10)$$

Where:

Q_t : Discharge at a specified time.

Q_{t-1} & Q_{t-2}: Discharge at t-1 and t-2 respectively.

W_t: Water level at a specified time.

W_{t-1} &W_{t-2}: Water level at t-1 and t-2 respectively.

Root mean square error (RMSE), mean absolute error (MAE) and coefficient of correlation (R) are used for evaluating the performance of models. The ANNs models were trained utilizing different numbers of neuron in the hidden layer, for training, testing and validation. The results showed the error values obtained for discharge when compared with observed data is shown in figure (3), for the validation data set. As evident from figure (3), the minimum error in discharge values has been obtained with 10 neurons in the hidden layer for all models. Table (4), shows the comparison of models for (Regulator II, Regulator III, Regulator IIII and Al Badaa) computed over the test dataset, with marked values corresponding with best performance according to the criteria in each column. Model number 2 for all stations is the best in performance in testing stages compared to the other models has the lowest RMSE and MAE values while has the highest R value.

The performance of artificial neural network models for prediction of flow discharge is demonstrated from figure (4) to figure (7), in the form of hydrograph and scatterplot. These Figures also show an analysis between the network outputs and the corresponding targets for the test dataset. Also these figures show the ability of artificial neural networks as a powerful tool to predicate the river discharge. The performance of the neural networks could be improves by using additional information that related to the variable under consideration such as precipitation rain.

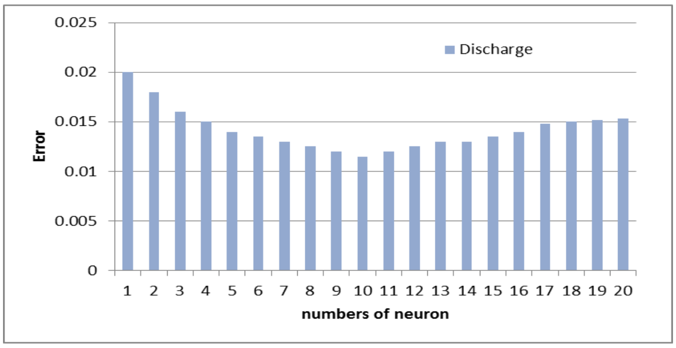

Figure 3 Variation of error values with hidden layer neurons for the validation data set.

Table 4 Performance parameters of the artificial neural network models in testing period with 10 nodes.

Stations	No. of models	RMSE	MAE	R
Regulator II	1	13.0027	9.0911	0.7200
	2	**7.3558**	**3.6503**	**0.9560**
	3	8.2972	4.2034	0.9556
	4	8.7773	4.0161	0.9501
Regulator III	1	9.3757	6.2450	0.8309
	2	**5.232**	**3.3724**	**0.9619**
	3	6.5718	3.5416	0.9513
	4	9.3817	3.9297	0.9383
Regulator IIII	1	6.9978	5.1564	0.7946
	2	**4.0167**	**3.1306**	**0.9296**
	3	4.0660	2.8737	0.9230
	4	5.1301	3.2822	0.9245
Al Badaa	1	3.2918	2.1979	0.8812
	2	**1.9183**	**1.0392**	**0.9703**
	3	2.2016	1.2303	0.9701
	4	1.9625	1.1227	0.9601

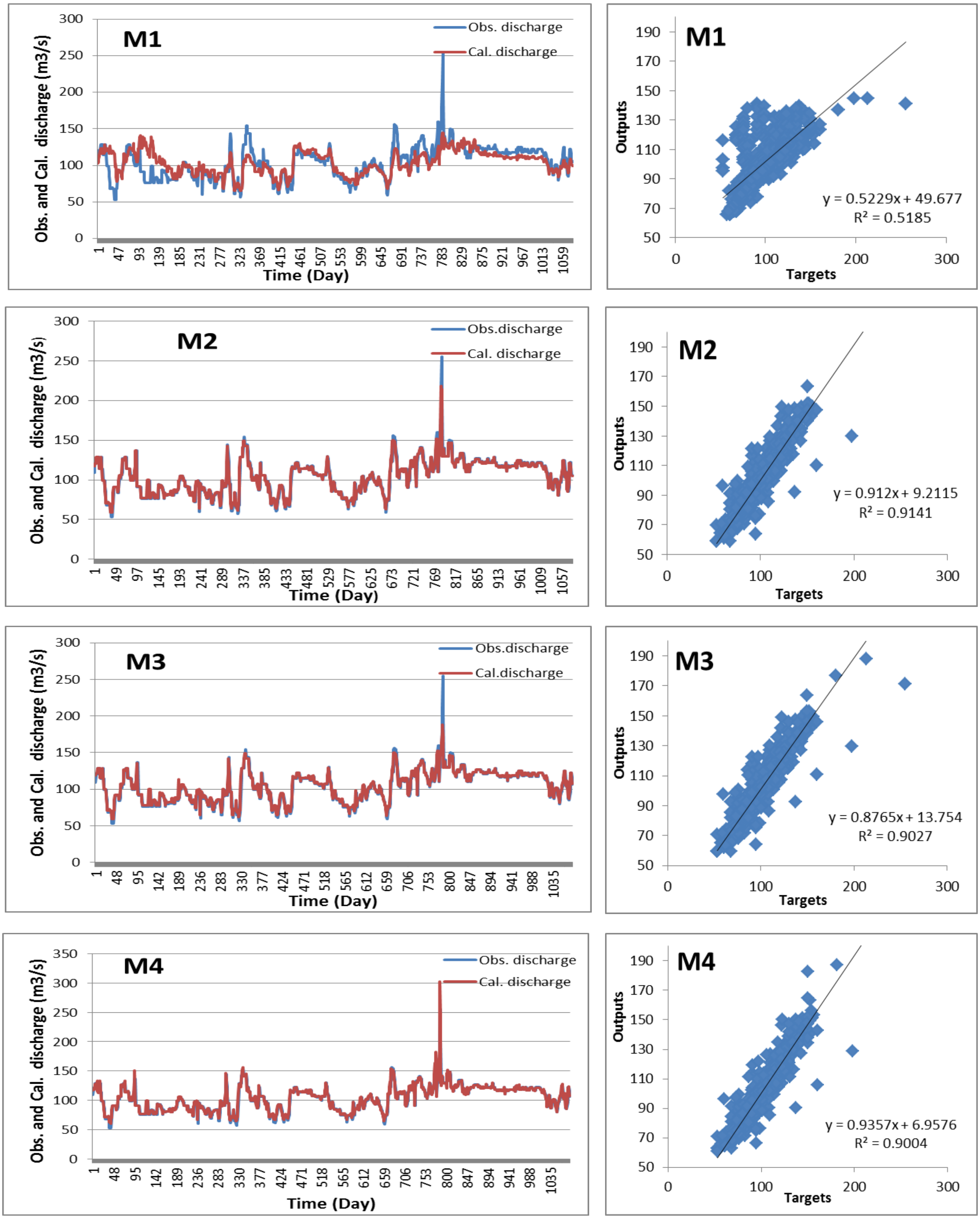

Figure 4 Comparison artificial neural network models for Regulator II.

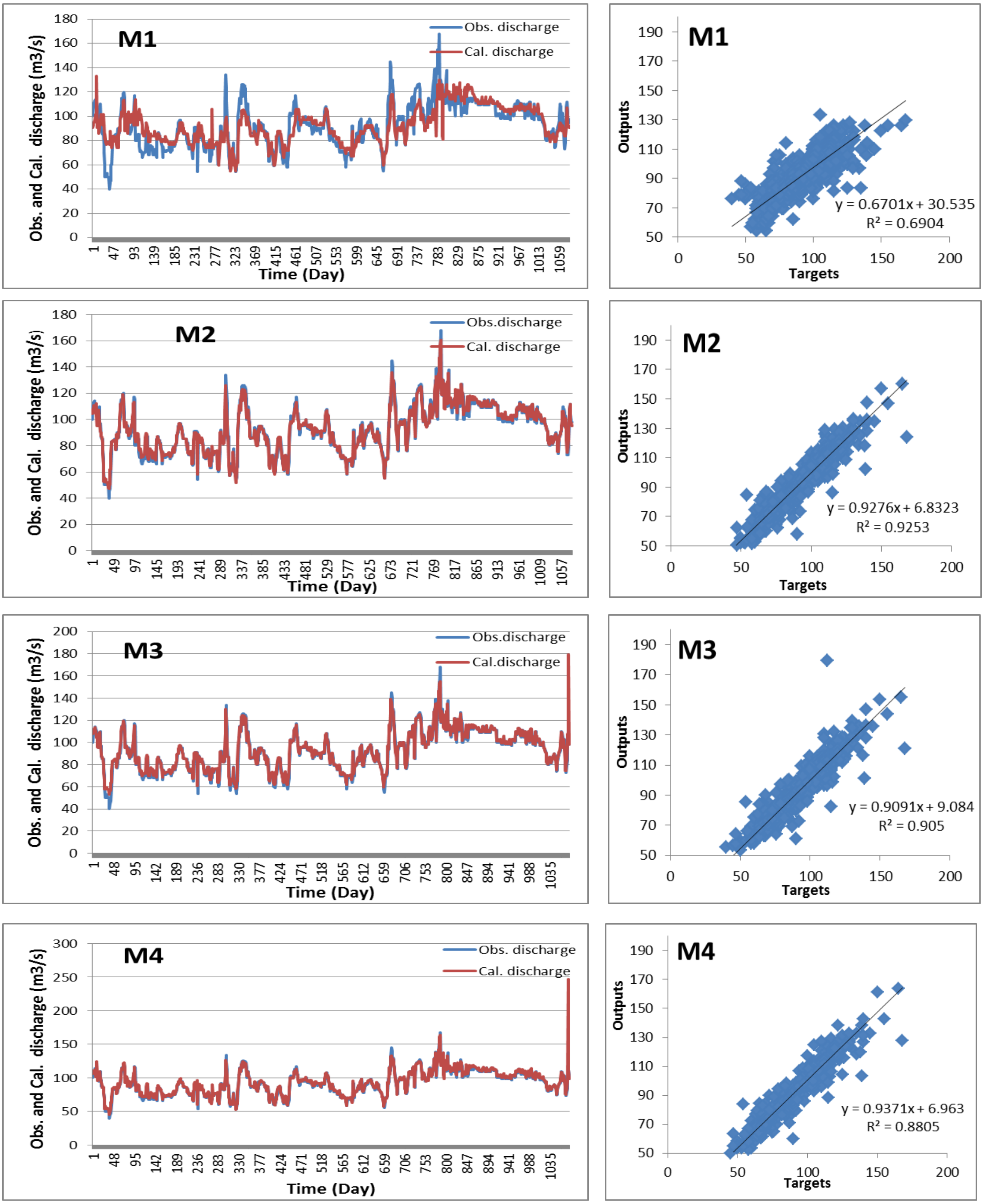

Figure 5 Comparison artificial neural network models for Regulator III.

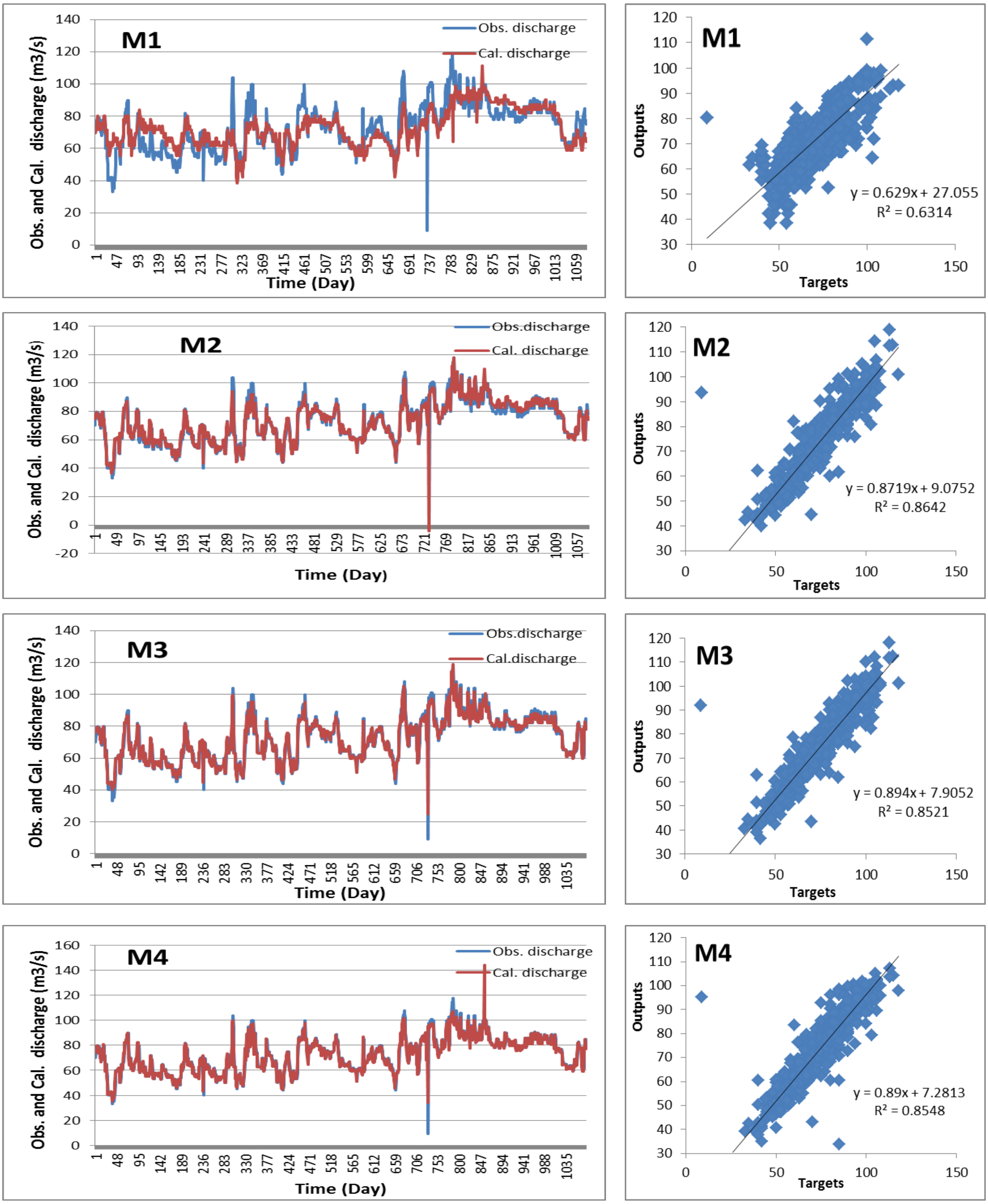

Figure 6 Comparison artificial neural network models for Regulator IIII.

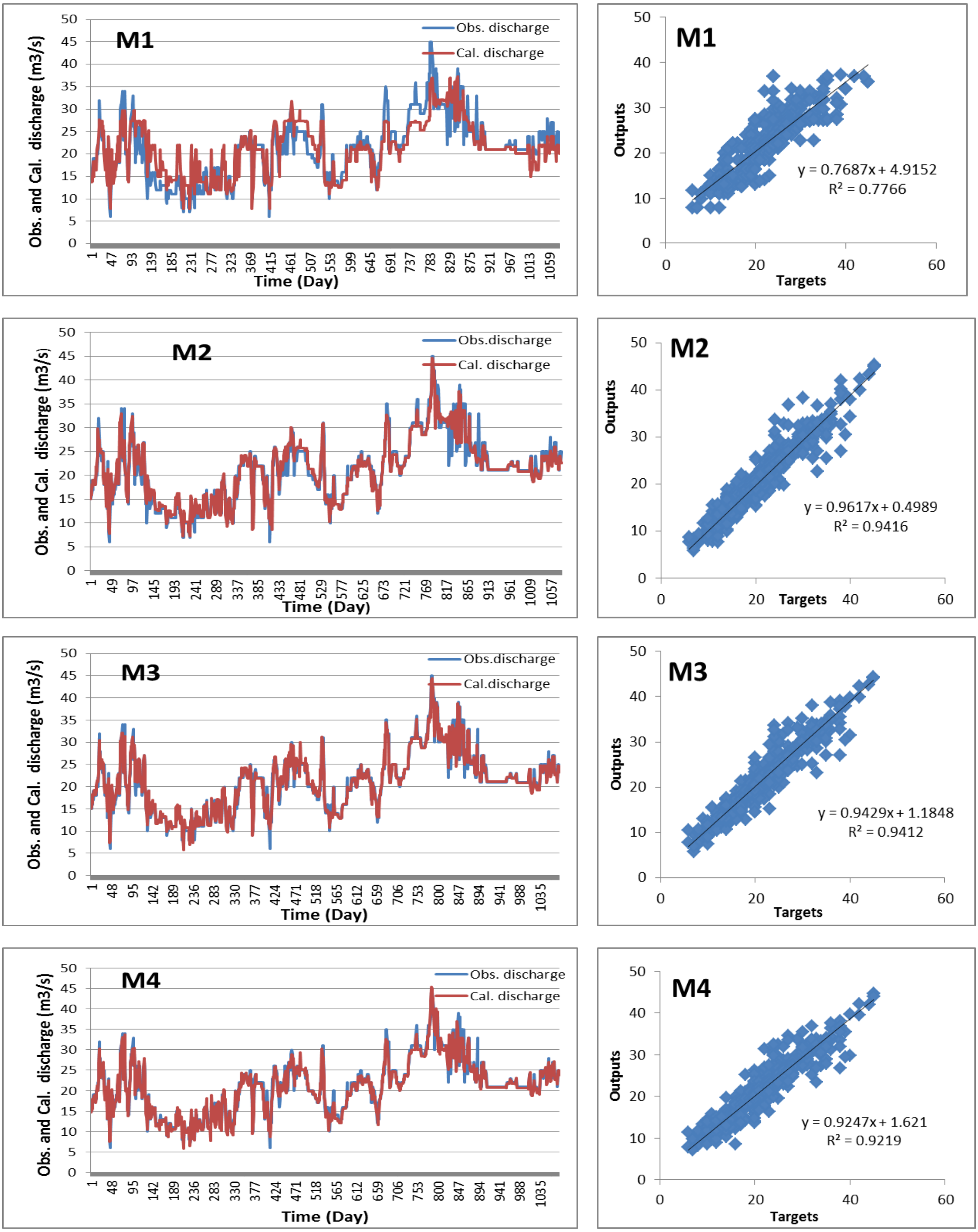

Figure 7 Comparison artificial neural network models for Regulator Al Badaa.

6. Conclusions

In this research, Artificial Neural Networks models were developed to predict daily discharge in the Al Garraf River of Thi Qar Province, south of Iraq. By using data included daily values of discharge and water level during three years. The data were spilt into three data sets for training, validation and test in the ratio 70:15:15, respectively. Three layers feed forward network with sigmoid hidden neurons and linear output neurons are used. The back-propagation algorithm gives a prescription for changing the weights in any feed forward network to learn a training vector of input-output pairs. The results showed that the artificial neural network with back-propagation algorithm is convincing technique for predicting the river discharge and the best numbers of neuron in hidden layer is equal to 10. But with few hidden neurons the network may be unable to learn the relationships amongst the data and the error will fail to fall below an acceptable level while large number of hidden neurons the network is able to correctly predict the data it has been trained but may result in overtraining. Also the results showed the efficiency of ANNs is begun to decrease when increasing the length of forecasting period. The study illustrates practical application of ANNs approaches, adequately combined with other frequently used tools in the context of water resources systems planning and management. Therefore, it can be concluded that the model a viable to be considered in future applications, competing with other classical techniques. The study recommends using hybrid systems developed from various artificial intelligence methods in order to get more accurate predictions.

7. References

[1]- Bayazıt M., Hidrolojik Modeller, İ.T.Ü. rektörlüğü, İstanbul. Birikundavyi S., Labib R., Trung H.and Rousselle J. “Performance of Neural Networks in Daily Stream flow Forecasting”, Journal of Hydrologic Engineering, 7(5),pp. 392-398, 2002.

[2]- Vanderkwaak J.E. and Loague K. “Hydrologic-response simulations for the R^{-5} catchment with a comprehensive physics-based model”, Water Resources Research 37 (4), pp.999, 2001.

[3]- Wang Y. C., Chen S. T., Yu P. S. and Yang T.C. “Storm-even rainfall-runoff modeling approach for ungauged sites in Taiwan”, Hydrological Processes 22, pp. 4322–4330, 2008.

[4]- ASCE, task committee on application of artificial neural networks in Hydrology (Rao Govindaraju), artificial neural networks in Hydrology. II: Hydrologic applications, Journal of Hydrologic Engineering 5(2), pp.124-137, 2000.

[5]- Lohani A. K., Goel N. K. and Bhatia K. K. S. “Comparative study of neural network, fuzzy logic and linear transfer function techniques in daily rainfall-runoff modeling under different input domains”, Hydrological Process, 25, pp. 175–193, 2011.

[6]- Govindaraju R. S. and Rao A. R. “Artificial neural networks in hydrology”, Kluwer Academic Publishers, Dordrecht, 2000.

[7]- Jain S., Das A., and Srivastava D. “Application of ANNs for Reservoir in Flow Prediction and Operation”, Journal of Water Resources Planning and Management, 5(125), pp. 263-271, 1999.

[8]- Zealand C., Burn D. and Simonovic S. “Short term Stream flow Forecasting Using Artificial Neural Networks”, Journal of Hydrology, 214, pp. 32-48, 1999.

[9]- Birikundavyi S., Labib R., Trung H., and Rousselle J. “Performance of Neural Networks in Daily Stream flow Forecasting”, Journal of Hydrologic Engineering, 7(5), pp. 392- 398, 2002.

[10]- Huang W., Bing Xu B., and Hilton A. “Forecasting Flows in Apalachicola River Using Neural Networks”, Hydrological Processes, 18, pp.2545-2564, 2004.

[11]- Muhammad A. S., Ghumman A.R. and Usman G. “Forecasting groundwater contamination using Artificial Neural Networks”, International Conf. on Water Resources and Arid Environment, pp. 1-8, 2004.

[12]- Mafia, J. Diamantopoulou, Dimitris M., Papamichail and Vassilis Z. Antonopoulos. “The use of a Neural Network technique for the prediction of water quality parameters”, Operational Research An International Journal, 5(1), pp. 115-125, 2005.

[13]- Mohsen H. and Zahra M. “Application of Artificial Neural Networks for temperature forecasting”, World Academy of Science, Engineering and Technology, 28, pp. 275-279, 2007.

[14]- Akhtar, M.K., Corzo1, G.A., Van Andel S. J. and Jonoski A. “River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information: Case study of the Ganges river basin”, Hydrological. Earth Syst. Sci.,13, pp. 1607-1618, 2009.

[15]- Najah A., Elshafie A., Karim O. A. and Jaffar O. “Prediction of Johor River water quality parameters using Artificial Neural Networks”, European Journal of Scientific Research, 28(3), pp. 422- 435, 2009.

[16]- Ali H. Al Aboodi , “Prediction of Tigris river discharge in Baghdad City using Artificial Neural networks”, Kufa journal of Engineering (K.J.E), Iraq.,5 (2) , pp.107-116, 2014.

[17]- A. A. Masrur Ahmed, “Prediction of dissolved oxygen in Surma River by biochemical oxygen demand and chemical oxygen demand using the artificial neural networks (ANNs)”, Journal of King Saud University – Engineering Sciences 29, 151–158, 2017.

[18]- Haykin S. "Neural Networks, a Comprehensive Foundation", Second edition. Prentice-Hall, New Jersey, 1999.

[19]- Rankovic V., Radulovi J., Radojevic I., Ostojic A. and Comic L., “Neural network modeling of dissolved oxygen

in the Gruza reservoir, Serbia", Ecol. Model, 221, pp.1239–1244, 2010.

[20]- KHAIRI A. O., " Prediction Of Dissolved Oxygen In Tigris River By Water Temperature And Biological Oxygen Demand Using Artificial Neural Networks (ANNs)", Journal of University of Duhok, 20(1), pp 691-700, 2017.

[21]- Haddad O. B., F. Sharife and S. Alimohammadi," ANNs in river flow forecasting", Proceedings of the 6th WSEA Int. Conf. on Evolutionary Computing, Lisbon, Portugal, pp.316-324, 2005.

[22]- Lippmann, R.P., "An introduction to computing with neural nets". IEEE ASSP, 1987.

[23]- Nourani V. and Babakhani A., "Integration of artificial neural networks with radial basis function interpolation in earth fill dam seepage modeling", Journal Comput Civil Engineering 27(2), pp.183-195, 2013.

[24]- Sahir A., Mahsin A. and Muhammed T., "A seasonal study of some physical and chemical properties of the water and sediment of the Gharraf River", Journal of Thi Qar Science, 1(4), pp.72-80. 2009.

[25]- https://en.wikipedia.org/wiki/Districts_of_Iraq.

[26]- Ali H., Ahmed A. and Husham T., " Comparison of Data-Driven Modelling Techniques for Predicting River Flow in an Arid Region", International Journal of Applied Engineering Research, 12(11), pp. 2647-2655, 2017.