Reservoir operation by Artificial Neural Networks
-A case study of Haditha reservoir, Iraq-

Dr. Thair sh. Khayyun
Assistant Professor
University of Technology
Building & Construction Eng. Dept.

Dr. Ayad S. Mustafa
Assistant Professor
Al-Anbar University
Civil Eng. Dept.

ABSTRACT

Water reservoir systems are complex, need systematic study and optimization techniques to assist its planning and management decisions. Reservoir operation is based on heuristic procedures, rule curves and to subjective judgments by the operators.

Artificial neural networks (ANN) are one of the widely used modeling techniques which can approximate a non-linear relationship between input and output data sets without considering physical processes and the corresponding equations of the system. Also, ANN model is much faster than a physically based model .

In this study, ANN model was applied for Haditha reservoir which locates on Euphrates River. The data set has a record length of 23 years covering (1986-2008).The best convergence was achieved for the combination of inflow (I_t), evaporation (E_{t-1}), reservoir storage (S_t) and outflow (O_t) with error tolerance, learning rate, momentum rate, number of cycles and number of hidden layers as 0.001, 0.2, 0.8,50000 and 7 respectively. The coefficient of correlation was 0.965. The results of ANN models for the training, testing and validation were compared with the observed data. The predicted values from the neural networks matched the measured values very well. The application of ANN technique assists the reservoir operation decision and future updating.

Keywords: ANN, inflow, evaporation, storage, outflow, Haditha reservoir.

INTRODUCTION

In the field of hydrology, besides stochastic method in the previous time, ANN techniques have been intensively developed and applied to the field (Dolling and Varas 2002).

Applying optimization techniques for reservoir operation is not a new idea. Various techniques have been applied in an attempt to improve the efficiency of reservoirs operation. These techniques include Linear Programming (LP);

Nonlinear Programming (NLP); Dynamic Programming (DP); Stochastic Dynamic Programming (SDP); and Neural Networks (Long 2006).

At the stage of the planning optimization of the reservoir, size is a very important subject. Before the construction of a dam, a procedure is assigned for determining the optimum size of the reservoir. This procedure is called the operation study of a dam. In operation study the aim is to optimize the reservoir volume (that is storage) for abstracting sufficient amount of water from the dam reservoir. The monthly inflow of the reservoir is the main data series. It is better to have a data record length as long as possible. In addition to the monthly inflows, the monthly evaporation losses are another main data (Ismail 2008). This data is the mean value of observed depths of the evaporated water in the reservoir area. Evaporation losses are measured for each month of the year and the statistical mean value of each month is found. Evaporation losses are in millimeters because they are measured as the depth of the evaporated water from a pan. Third important data is the amount of water that is planned to be taken from the reservoir. This is called the monthly demand. It may be for drinking or irrigation purposes or it may be for hydropower generation.

Many researches have applied ANN to model different complex hydrological processes. The ANN methods have good generalization efficiency and are commonly used in practical hydrologic projects. Even when there are missing data values, the ANN methods can be applied to aid in the completion of missing hydrological records (Shahram and Huang 2010).

In order to improve performance, ANN models need to be developed in a systematic manner. The process of optimizing the connection weights is known as (training) or (learning). This is equivalent to the parameter estimation phase in conventional statistical models. Stopping criteria are used to decide when to stop the training process. They determine whether the model has been optimally or sub-optimally trained. Training can be stopped, when the training error reaches a sufficiently small value or when no or slight changes in the training error occur.

When, the training and stopping criteria of the model have been successfully accomplished, the performance of the trained model should be validated. The purpose of the model validation is to ensure that the model has the ability to generalize within the limits set by the training data. The root means squared error, RMSE, the coefficient of correlation, r, and the mean absolute error, MAE, are the main criteria that are used to evaluate the prediction performance of ANN models.

The coefficient of correlation is a measure that is used to determine the relative correlation and the goodness-of-fit between the predicted and observed data. The RMSE is the most popular measure of error and has the advantage that large errors receive much greater attention than small errors (Hecht 1990). This paper presents the results from a study on the application of the ANN to forecast the next month's reservoir outflow.

ANNs TECHNIQUE

The ANNs are a massively parallel-distributed information-processing system that has certain performance characteristics resembling biological neural networks of human brain. A typical ANN is shown in Figure 1 (Omid and Saeed 2005).

Each neural network consists of three layers: input, hidden and output layer and in every layer there are number of processors called (nodes).Each node is connected to other neurons with a directed link and a special weight. Neurons response is usually sent to other ones. A set of inputs in the form of input vector X is received by each unit and weights leading to the node form a weight vector W. The inner product of X and W is net and the output of the node is f(net) :

$$net = X \cdot W = \sum X_i . W_i \quad (1)$$

$$out = f(net) \quad (2)$$

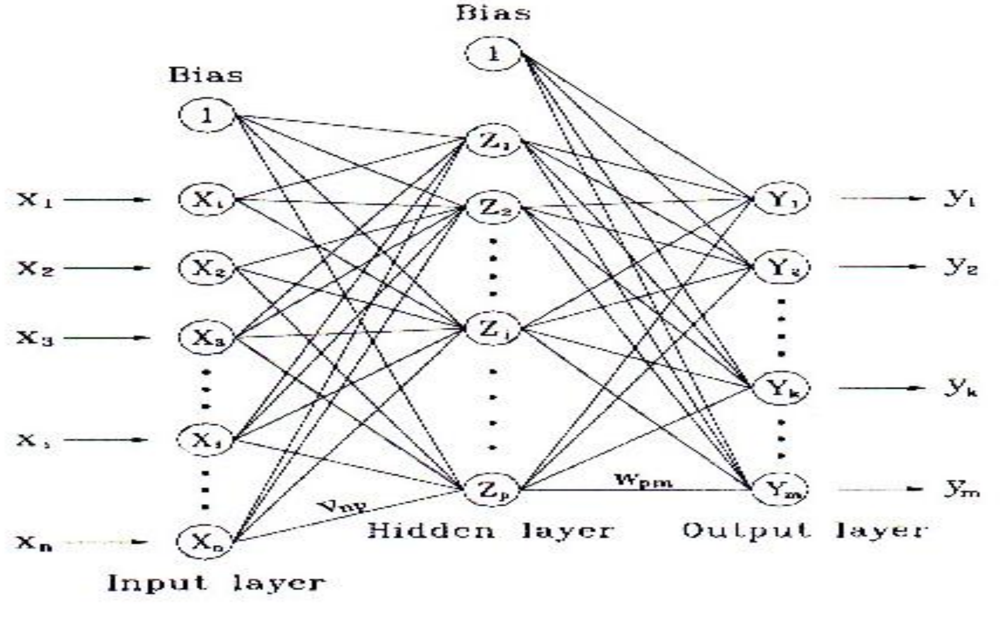


Figure 1: Typical Structure of ANN, (Omid and Saeed 2005).

f is called activation function whose functional form determines response of the node to the input signal it receives. Sigmoid function usually used in different applications, given as equation (1):

$$f(x) = \frac{1}{1+e^{-x}} \quad (3)$$

Training of networks is carried out in three steps: 1) Presenting training sets to input and output neurons, 2) Computation of the error of the network and back propagating it, and 3) Adjusting the weights in order to reduce the error. There are some learning rules based on back propagation algorithm in networks, and the most applicable one is Generalized Delta Rule. Weights are adjusted according to the following equation:

$$\Delta w_{ij}(n) = -\alpha.\frac{\partial E}{\partial w_{ij}} + \beta.\Delta w_{ij}(n-1) \tag{4}$$

Where, $\Delta w_{ij}(n)$ and $\Delta w_{ij}(n-1)$ are weight increments between node i and j during the n^{th} and $(n-1)^{th}$ pass, or epoch. In equation (4), α and β are learning rate and momentum rate respectively, and they are both useful for a better training process. Model validation is carried out to understand how a network is able to response to training set and to a new set to which the network hasn't faced to (testing set). Performance of a network is usually evaluated by some parameters, such as: 1- RMSE (Root Mean Square Error); 2- R (Correlation Coefficient), 3-e (Relative Error). All these parameters should be evaluated for both training and testing sets.

STUDY AREA AND DATA

In this study, ANN model was applied for Haditha reservoir which locates on Euphrates River as shown in, Figure 2. Construction lasted between 1977 and 1986 and was a joint project by the Soviet Union and Iraq. Haditha dam is 9064 m long and 57 m high, with the hydropower station at 3310 m from the dams southern edge. The crest is at 154 m AMSL and 20 m wide. Haditha reservoir has maximum water storage of a capacity 8.3 cubic kilometers and maximum surface area of 500 square kilometers. In cross- section the dam consists of an asphaltic concrete cutoff wall at its core, followed by mealy dolomites, and a mixture of sand and gravel.

The ANN model inputs were the monthly inflow (I) in m^3/sec, evaporation losses (E) in m^3 and reservoir storages (S) in m^3. The output of the model was outflow (O) from the reservoir in m^3 /sec. The data set has a record length of 23 years covering between (1986- 2008).

ANN APPLICATION

Division of data

The data are randomly divided into three sets (training, testing and validation). In total, 80% of the data are used for training and 20% are used for validation. The training data are further divided into 70% for the training set and 30% for the testing set. These subsets are also divided in such away that they are statistically consistent and thus represent the same statistical population.

To examine how representative the training, testing and validation sets are with respect to each other, T-test and F-test are carried out. These results indicate that training, testing and validation sets are generally representative of a single population.

The data base used for the ANN model comprises a total 276 individual cases. Ranges of the data used for the input and output variable are summarized in Table 1.

Scaling of data

The input and output variables are pre-processed by scaling them between (0 and 1), to eliminate their dimensions and to ensure that all variables receive equal attention during training. The simple linear mapping of the variables extremes is adopted for scaling, as it is the most commonly used method (Maier and Dandy 2000). As part of this method, for each variable X with minimum and maximum values of X_{min} and X_{max} respectively, the scaled value X_n is calculated as follows:

$$X_n = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (5)$$

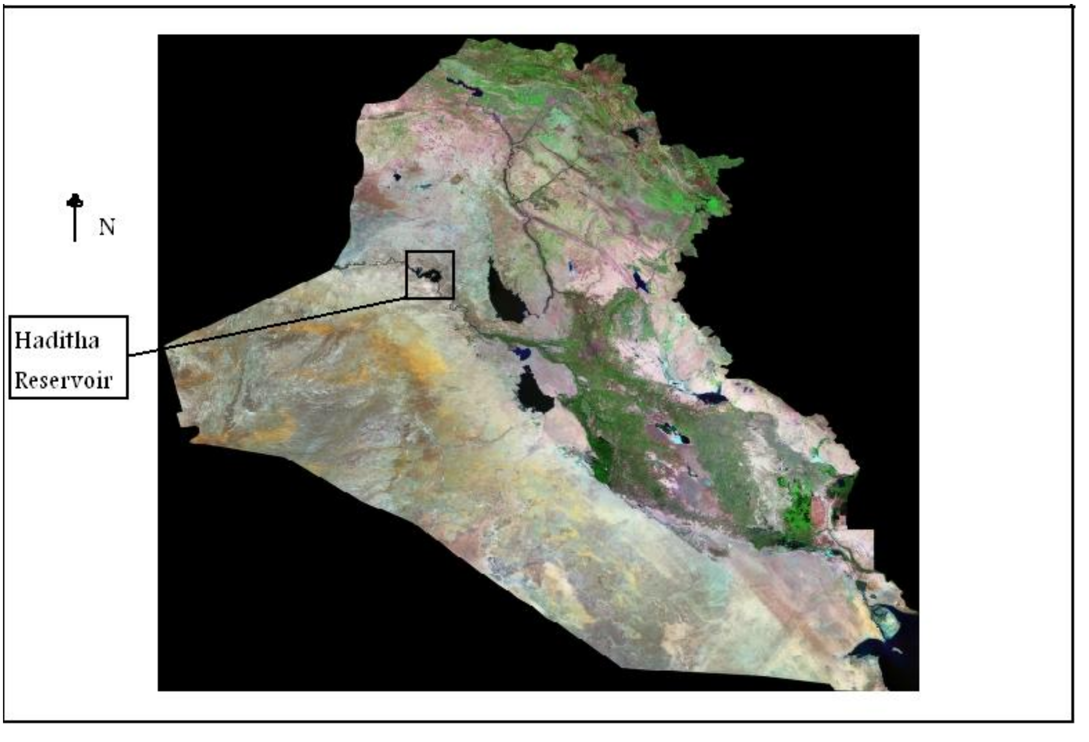


Figure 2: Study Area Map

Table 1 : Ranges of the data used for the ANN model

Model variables	**Minimum value**	**Maximum value**
Input flow, m^3/sec	153	2990
Evaporation, m^3	4994298	165000000
Storage, m^3	1860000000	8250000000
Output flow, m^3/sec	130	2464

Model architecture

The difficult task in the development of ANN models is determining the model architecture (the number of the hidden layer nodes and the values of learning and momentum rates).The number of nodes in the input and output layers are restricted by the number of model inputs and outputs, respectively. There is no direct and precise way of determining the best number of nodes in each hidden layer. A trial and error procedure for determining the number and connectivity of the hidden layer nodes can be used.

Using the default parameters of the ANN model, a number of networks with different numbers of hidden layer nodes is developed. The effect of the internal parameters controlling the back-propagation (momentum and learning rates) on model performance is investigated for the model with four hidden layer nodes. The effect of the learning and momentum rates on the model performance are shown in figures 3 and 4. It can be seen that the performance of the ANN model is relatively sensitive to learning rates in the range (0.10 to 0.40) then the prediction errors slightly increase to 140, Figure 3. Figure 4 shows the effect of the momentum rate on model performance. It can be seen that the performance of the ANN model is relatively sensitive to momentum rate value of (0.8). The optimum values for learning and momentum rate used is 0.20 and 0.80 respectively. Also, the network with (4) hidden layer nodes has the lowest prediction error for the training and validation tests. However, it is believed that network with 4 hidden layer nodes is considered optimal, Figure 5. The input and output variables for the above tests were at a time t.

RESULTS OF ANN

ANNs were used to derive and to develop models to predict the monthly outflow in m^3/sec from Haditha reservoir. The monthly data of three parameters (inflow (m^3/sec), evaporation (m^3), storage (m^3) and the outflow (m^3/sec) for the time period 1986-2008 were selected for this analysis.

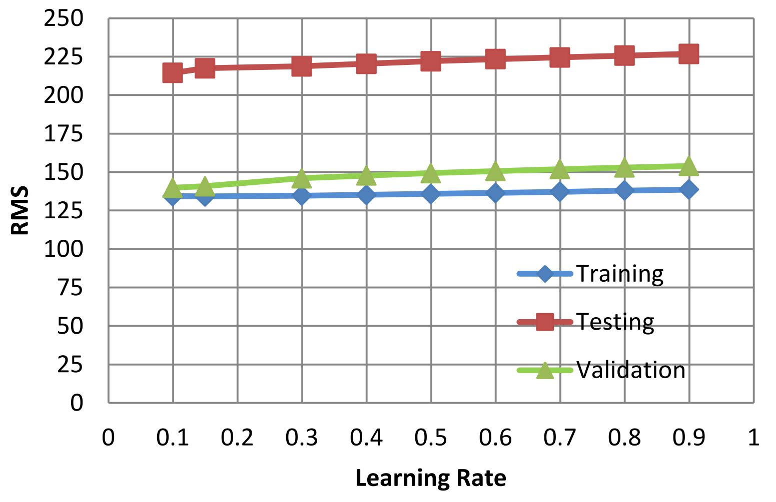


Figure 3 : Performance of the ANN Model with Different Learning Rates

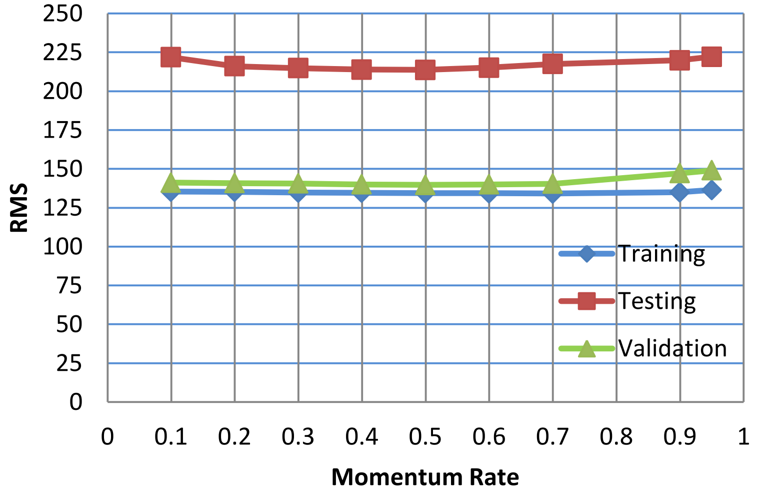


Figure 4 : Performance of the ANN Model with Different Momentum Rates

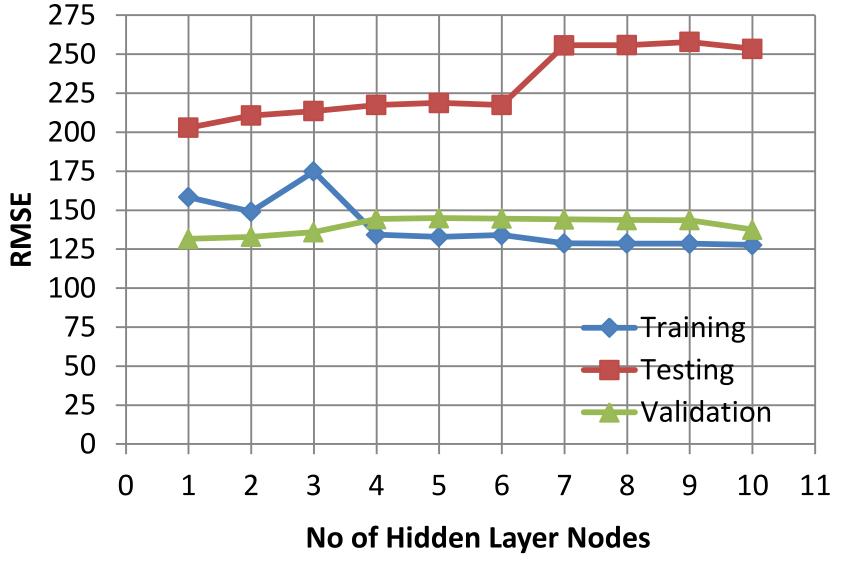


Figure 5 : Performance of the ANN Model with Different Hidden Layer Nodes.

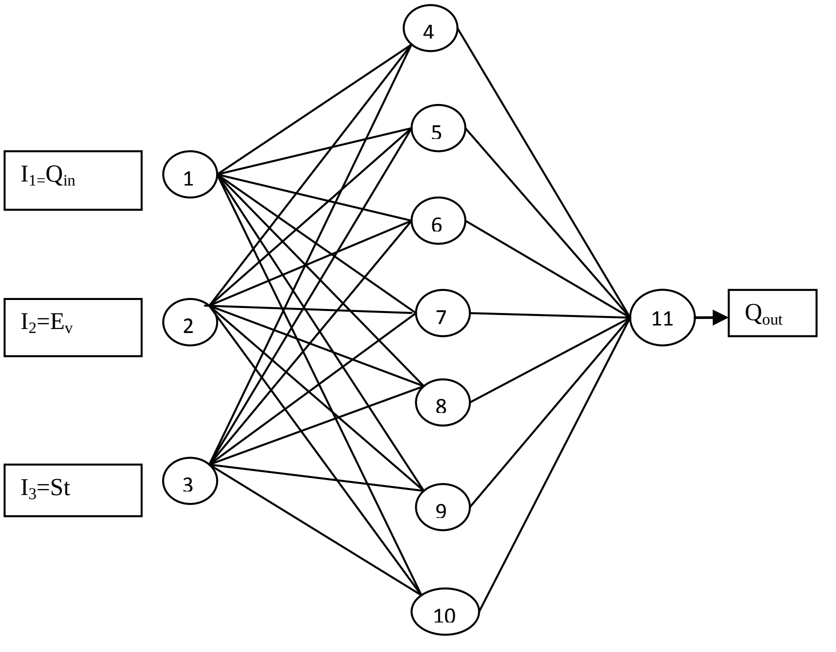
To get the optimize weights of the ANN model, six different combinations of input and output variables were trained. A number of trials are carried out for each of the six scenarios, using the parameters of the ANN model used (1-13 hidden layer nodes, learning rates 0.01-0.99 and momentum rates 0.01-0.99). The coefficient of correlation and the sum of squared errors and other statistical performance measures for the different network scenarios were computed. The structure which gives the highest coefficient of correlation and small sum of squared error was selected

The final structure of the optimal ANN model is shown in Figure 6.Different networks structures tested in order to determine the optimum number of hidden layers and the number of nodes in each. The best solution given by the ANN for the outflow composed of one input layer with 3 input variables, one hidden layer with 7 nodes and one output layer with one output variable..

The best convergence was achieved for the combination of inflow (I_t), evaporation (E_{t-1}), reservoir storage (S_t) and outflow (O_t) with error tolerance, learning rate, momentum rate, number of cycles and number of hidden layer nodes as 0.001, 0.2, 0.8, 50000 and 7 respectively. The coefficient of correlation was 0.965 for the training data, Table 2.

The results of ANN models for the training, testing and validation were compared with the observed data. The predicted values from the neural networks match the measured values very well. Figure 7, shows the trained results of predicted values by ANN vs. the measured values of output flow in m^3/sec from

Haditha reservoir. Figure 8, shows the measured verses predicted values of all data records for the release flow (outflow) in m^3/sec.



Input layer Hidden layer Output layer

Figure 6: Structure of the ANN optimal model.

Table 2: Results of ANN training for reservoir operation.
(number of cycles=50000)

Input combination	Error tolerance	Learning rate	Momentum rate	Number of hidden layer nodes	Coefficient of correlation	Root mean squared error
I(t), E(t), S(t)	0.001	0.2	0.8	4	0.949	134.3
I(t-1), E(t), S(t)	0.001	0.2	0.8	7	0.934	152.4
I(t), E(t-1), S(t)	0.001	0.2	0.8	7	0.965	111.3
I(t), E(t), S(t-1)	0.001	0.2	0.8	6	0.962	115.7
I(t-1),E(t-1), S(t)	0.001	0.2	0.8	10	0.938	146.8
I(t-1),E(t), S(t-1)	0.001	0.2	0.8	9	0.949	134.5

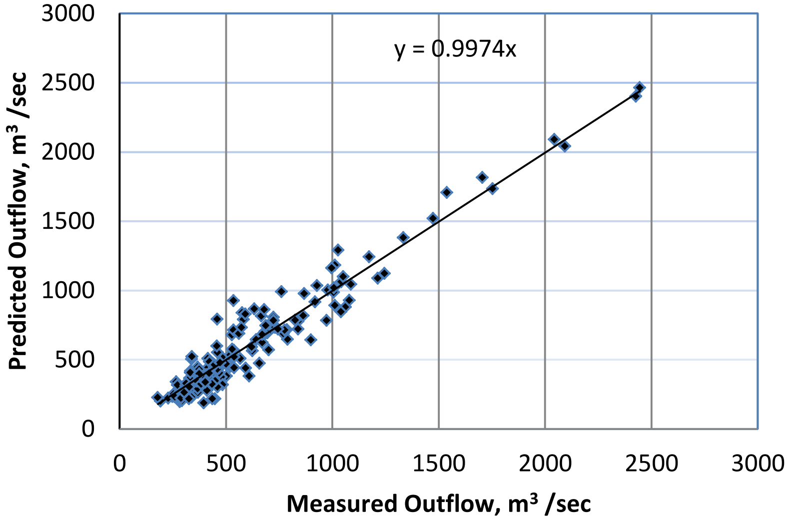


Figure 7: Training results of predicted vs. measured outflow (Haditha reservoir).

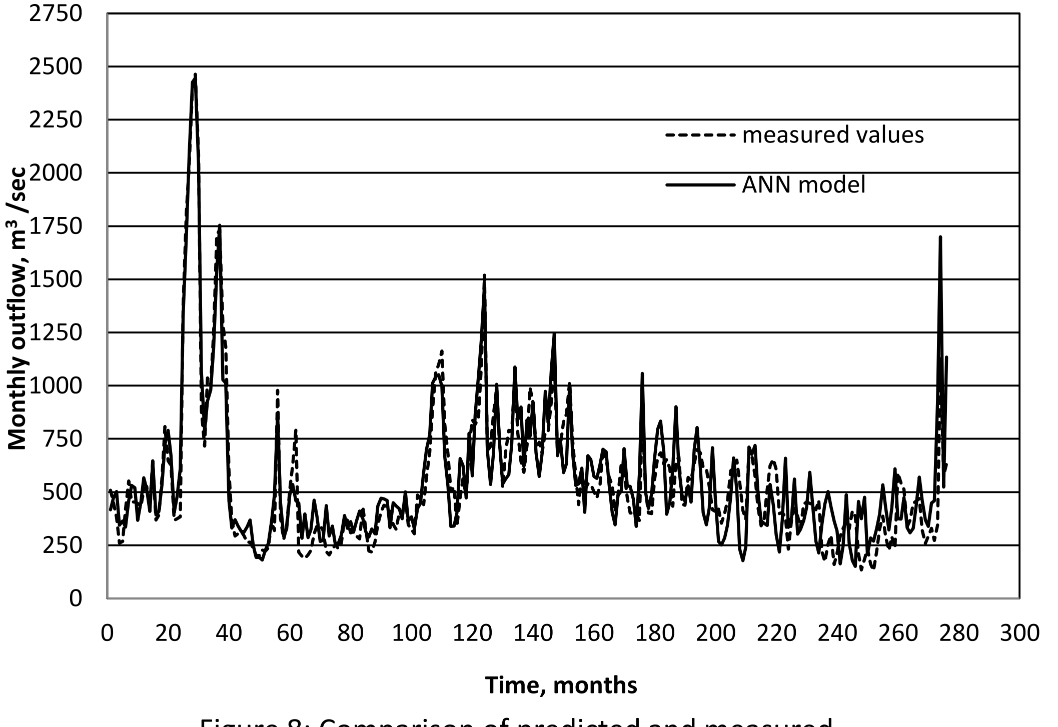


Figure 8: Comparison of predicted and measured values of outflow for all data records.

Using the connection weights and the threshold levels which obtained from ANN model (Table 3), the predicted outflow rate in m^3/sec from Haditha reservoir can be expressed as follows:

$$O_t = (2276/(1+e^{(-0.106+0.797\tanh(x1)+0.641\tanh(x2)-1.014\tanh(x3)+0.991\tanh(x4)+0.93\tanh(x5)+0.618\tanh(x6)-0.791\tanh(x7)})) + 188 \quad (6)$$

Where:

$$X_1 = 0.984 + 1.021*10^{-3} I_t + 3.665*10^{-8} E_{t-1} - 8.624*10^{-10} S_t$$
$$X_2 = 3.77 + 1.758*10^{-3} I_t - 1.344*10^{-8} E_{t-1} - 7.020*10^{-10} S_t$$
$$X_3 = -1.214 + 0.970*10^{-3} I_t + 0.305*10^{-8} E_{t-1} - 1.604*10^{-10} S_t$$
$$X_4 = 2.229 - 1.920*10^{-3} I_t - 2.402*10^{-8} E_{t-1} + 2.995*10^{-10} S_t$$
$$X_5 = 1.852 - 1.736*10^{-3} I_t + 1.954*10^{-8} E_{t-1} - 0.921*10^{-10} S_t$$
$$X_6 = -6.031 - 0.309*10^{-3} I_t - 0.924*10^{-8} E_{t-1} + 12.781*10^{-10} S_t$$
$$X_7 = 0.318 + 2.269*10^{-3} I_t + 0.478*10^{-8} E_{t-1} - 2.291*10^{-10} S_t$$

It should be noted that equation (6) is valid only for the range of values of (I, E, S and O) given in Table 1. This is due to the fact that ANN should be used only in interpolation and not extrapolation (Tokar and Johnson 1999).

Table 3: Weights and threshold levels for the ANN model.

Hidden layer nodes	W_{ji} (weight from node i in the input layer to node j in the hidden layer)							Hidden layer threshold θ_j
	i=1		i=2		i=3			
j=4	2.858		5.864		-5.209			-0.544
j=5	4.920		-2.150		-4.240			2.488
j=6	2.715		0.488		-0.969			-1.369
j=7	-5.374		-3.843		1.809			2.404
j=8	-4.858		3.126		-0.556			1.416
j=9	-0.864		-1.479		7.720			-3.311
j=10	6.352		0.765		-1.384			0.269
Output layer nodes	W_{ji} (weight from node i in the hidden layer to node j in the output layer)							Output layer threshold θ_j
	i=4	i=5	i=6	i=7	i=8	i=9	i=10	
j=11	-0.797	-0.641	1.014	-0.991	-0.930	-0.618	0.791	0.106

CONCLUSION

The ANN procedure to determine the outflow from Haditha reservoir-Iraq was developed in this study. Feed –forward layer are used and trained with the back propagation algorithm. The applicant of ANN technique to Haditha reservoir found that the ANN technique can accurately predict the monthly outflow (release flow) and assist the reservoir operation decision and future updating. The results indicate that the evaporation data in cubic meter for a time (t-1) has the most significant effect on the predicted of the monthly outflow. The combination of inflow (I_t), evaporation (E_{t-1}), reservoir storage (S_t) and outflow (O_t) was found to be the best for the reservoir operation. ANN performance is sensitive to the number of hidden layer nodes, momentum rates, learning rates and transfer functions. The credibility of predicted outflow from reservoir can be investigated graphically by translating the equation obtained from the ANN model into a set of design charts.

REFERENCES

Dolling, O.R. Varas, E.A. (2000)
Artificial Neural Networks for Stream Flow Prediction, Journal of Hydraulic Research, 40(5):545-554.

Hecht, R. (1990)
Neuro-Computing, Addison-Wesely Publishing Company, Reading.MA.

Ismail, K. (2008)
Reservoir Management Using Artificial Neural Networks, State Hydraulic Works, Istanbul, Turkey.

Long, L.N. (2006)
Optimizing Reservoir Operation, Institute of Environment and Resources, Technical University of Denmark, p.2-3.

Maier, H.R. and Dandy, G.C. (2000)
Neural Networks for the Prediction and Forecasting of Water Resources Variables, Environmental Modeling and Software, 15(2000), 101-124.

Omid, B.H. and Saeed, A. (2005)
Evaluation of ANN in Optimization Models of Hydropower Reservoir Operation, Ninth International Water Technology Conference, IWTC9 2005, Egypt.

Shahram, K.G. and Huang, Y.F. (2010)
Neural Networks for Forecasting Daily Reservoir Inflows, National Hydraulic Research Institute of Malaysia, Selangor, p 33-34.

Tokar, S.A. and Johnson, P.A. (1999)
Rainfall-Runoff Modeling Using Artificial Neural Networks, J. Hydrologic Eng., 4(3), 232-239.