World Applied Sciences Journal 4 (5): 646-654, 2008
ISSN 1818-4952


# Application of Artificial Neural Network to Predict Total Dissolved Solid in Achechay River Basin

[1]S. Kanani, [2]G. Asadollahfardi and [2]A. Ghanbari

[1]Tarbiat Moallem Univeristy, Tehran, Iran
[2]Departpmet of Civil Engineering, Tarbiat Moallem University, Tehran, Iran

**Abstract:** River water salinity is a significant concern in many countries, considering agricultural and drinking usages. Therefore, prediction of amount of salinity is a necessary tool for planning and management of water resources. Since Achechay River Basin in East Azerbaijan province in Iran passes through saline zones, use of the water for irrigation has become problematic. In this regard, prediction of future salinity of Vaniar station in Achechay river basin was studied using Artificial Neural Network (ANN) with a month time delay as predictor, considering the effect of discharge with 24 hours time delay and Total Dissolved Solid (TDS), TDS monthly mean data and daily mean discharge for thirty years are considered as inputs for the ANN and TDS is the output of the models. Multi Layer Perceptron (MLP) and Input Delay Neural Network (IDNN) methods were applied to the data. The results of the study showed that predictions of river salinity using Artificial Neural Network are reasonable, suitable and of acceptable accuracy. Hence, prediction of water salinity by ANN may be useful for water quality planning and management.

**Key words:** Salinity prediction . Achechay River . Artificial neural network

## INTRODUCTION

Lack of water resources and optimum management have been two recent challenges of water resources engineering. Population growth, decrease of useable water resources, improvements in lifestyle, growing rate of consumption, climate change and several other parameters have caused useable water to be a significant problem for future. Economic and efficient use of water resources and its management have an increasingly significant role. Prediction of water salinity is one of the methods which have been recently considered for management of water resources. The predictions can be used for water resources planning and management in case they are of acceptable accuracy. There are two methodologies for prediction of salinity, like other water quality parameters,; first, precise study of different processes which can affect water salinity and developing statistical or deterministic models according to the obtained information.. Second, developing Data Driven Models using information and collected data; In the latter technique, relationship between input and output data can be found using input data, but still physical understanding of phenomena is significant for having suitable input data for model, however, it is not necessary to simulate complicate process. ANN is one of the Data Driven Models which has recently been applied as a tool for modeling complicated processes.

American Society of Civil Engineering (ASCE) research committee reported that ANN can be applied for various hydrological branches and they reached the conclusion that ANN is able to simulate many of complicated nonlinear processes (3) (4). Wenrui and Simon [1] applied ANN for assessment of variation of water salinity in Apalachicola River in Florida considering wind, ebb and flow and amount of influents freshwater. The result shows ANN is able to model non linear water salinity very well considering three mentioned parameters (10). Holger *et al*. [2] used ANN for modeling cyan bacteria Anabaena in Murray River in South Australia (11).

Achechay River is located in the North-West of Iran and it originates from Sabalan Mountain basin. The river passes through Vaniar Valley, located in the North of Tabriz City and then water of Sahand mountain basin enters the river. Finally, water of the river enters Oromeyh Lake (1).

Some of tributaries of Achechay River have high water salinity because the river passes through salty basins such as Markid, Karchay, Sabzinehcha and Tazahkan streams (3).Vaniar Dam is located at East longitude 46°,8’,23’’ and North latitude 38°,8’, 59’’on Achechay river. The purpose of constructed dam was to

**Corresponding Author:** Mr. S. Kanani, Tarbiat Moallem Univeristy, Tehran, Iran

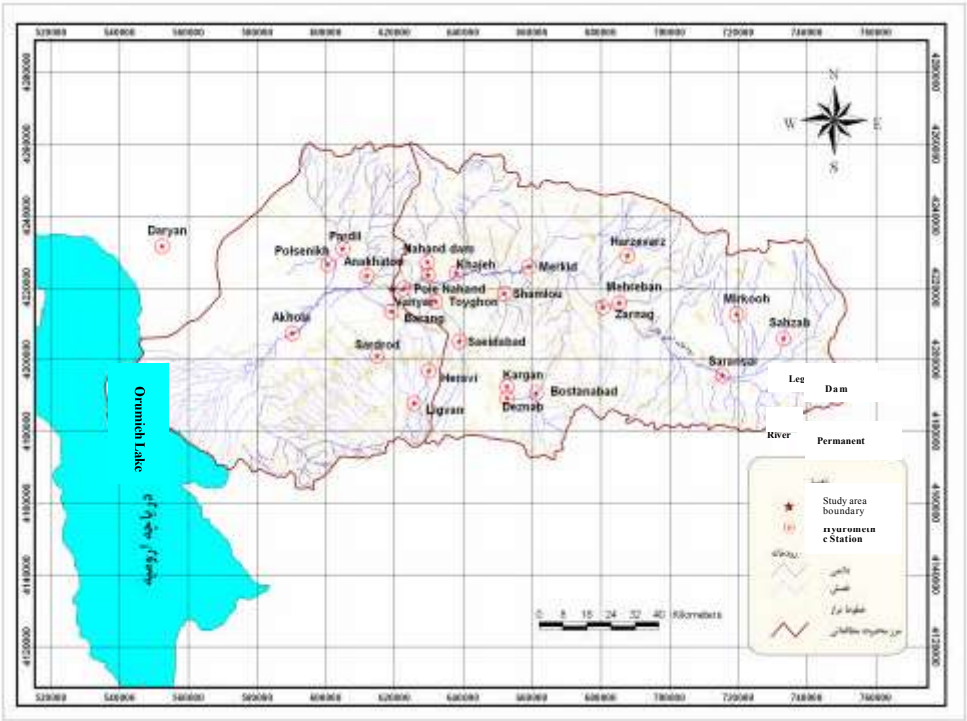



Fig. 1: Achechay River Basin

Table 1: Some characteristics of Achechay river basin (1)

| | |
|---|---|
| Basin area up to Vaniar dam | 7723.0 km$^2$ |
| Annual mean rainfall | 388.0 mm |
| Minimum Altitude from sea level | 1458.0 m |
| Maximum Altitude from sea level | 3882.0 m |
| Total amount of annual water | 409.0 mcm |
| Withdrawal upstream from annual water | 108.6 mcm |
| Vaniar dam annual evaporation | 1503.0 |

supply water for irrigation of 40000 hectares of Sofeyan, Azarshahr and Tabriz cities. Figure 1 and Table 1 show the situation of Vaniar dam and the basin characteristics respectively. The Problem which Azerbaijan water authorities encounter is water salinity control in Vaniar dam for irrigation usage. In this regard, having enough information about future salinity is vital for planning and management.

The first objective of the study is to develop models with thirty year previous data of monthly TDS and hourly discharge in Vaniar station at Vaniar Dam, using ANN models and to compare the results. Second, prediction of TDS parameter from developed models which may be useful in planning and management of water salinity in the area.

## THEORY OF ARTIFICIAL NEURAL NETWORK

Considering natural Neural and its components, scientists developed artificial Neural which is the smallest processing unit of an ANN. An artificial Neural consists of three components including weighting (W), bias (b) and transfer function (f).These three components are unique for each Neural. Figure 2 shows schematic of artificial Neural. In the figure p and a are input and output of a Neural, respectively. Parameter n is called net input, which is input of transfer function and it is built according to input p and Neural parameters. Mentioned artificia l Neural can be modeled by the following equations.

$$n = wp+b \quad (1)$$

$$a = f(n) = f(wp+b) \quad (2)$$

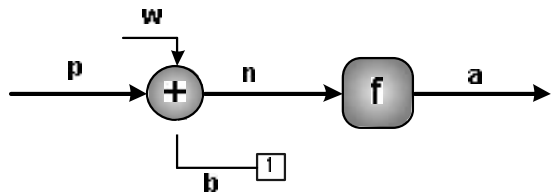
In Neural instruction process, w and b (Neural parameters) change until the best approximation for an output member corresponding to the input member is obtained. Weight of Neural determines the rate of p effect on "a" and parameter "b" causes Neural to be transformed to sub-space of bias input space. There are some types of transfer functions some of which are as follows:

- Linear, transfer function
- Hard-limit transfer function
- Log-Sigmoid transfer function

- Tan-Sigmoid transfer function and
- Tan-Hyperbolic transfer function

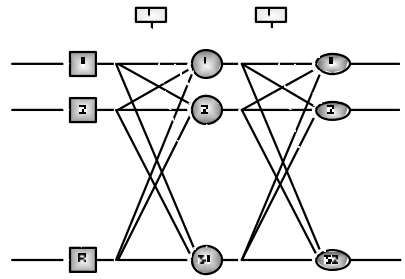
Generally, there are two types of ANN models as follows:

- Static model
- Dynamic model

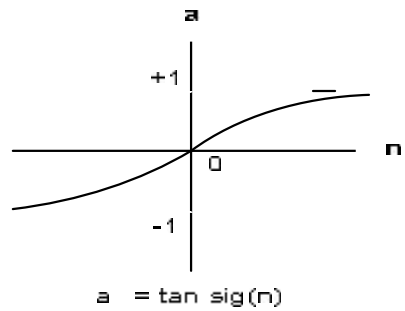



## STATIC ANN MODELS

In static models, time is not considered and outputs of network at any time depend on the inputs at the same time. the static model applied in this research is MLP.According to Universe Approximator, each MLP with a sigmoid hidden layer and a linear output layer is able to predict each complicated parameter if numbers of Neural in hidden layer are selected precisely (12). Figure 3 shows schematic of MLP with a hidden layer



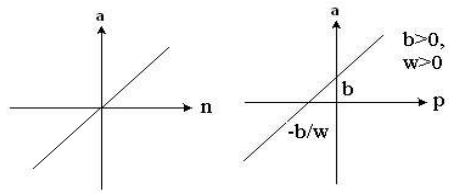
According to the mentioned theory, all ANN's which are applied in this research are MLP with a hidden layer, tangent sigmoid transfer function and linear layer outputs. Figure 4 and 5 show schematic tangent- sigmoid transfer function and linear transfer function for output layers respectively. Number of Neurals in hidden layers for each model may be obtained using trial and error. MLP with a hidden layer, tangent sigmoid transfer function and linear layer outputs can be modeled by the following equations:



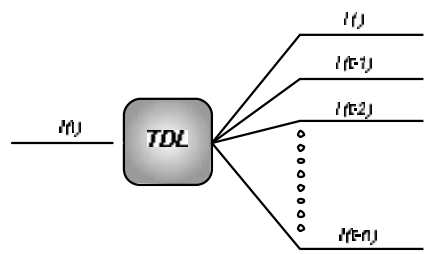

$$a^1_j(t) = F\{\Sigma^R_{i=1} w^1_{j,i} p_i(t)+b^1_j\} \quad (3)$$

$$a^2_k(t) = G\ \{\Sigma^{s1}_{j=1} w^2_{k,j}\ a_J^1(t)+b^2_k\} \quad (4)$$

In equations 3 and 4, R is the number of input vector components, $S_1$ and $S_2$ are number of Neurals in hidden and output layers, respectively. P is input vector. $W^1$, $w^2$ are weighting matrix in hidden and output layers, respectively. $b^1$, $b^2$ are bias vectors in hidden and output layers, respectively. G and F are Neural transfer functions in hidden and output layers.




**Dynamic ANN models:** The inputs of Dynamic TNN are the same as static model; the difference is that the effect of past periods is considered in this model. There are several methods by which Static model can be turned to dynamic. One of them is Time Delay Neural Network (TDNN) operators. A TDNN operator receives input signal and keeps it for a time step and in next time step, the input signal emerges as an output result. By connecting N series of TDNN operator, Tapped Delay Line (TDL) will be obtained. The output is a vector with N+1 Components. The N+1 components include the inputs in the current time step and N time steps before.



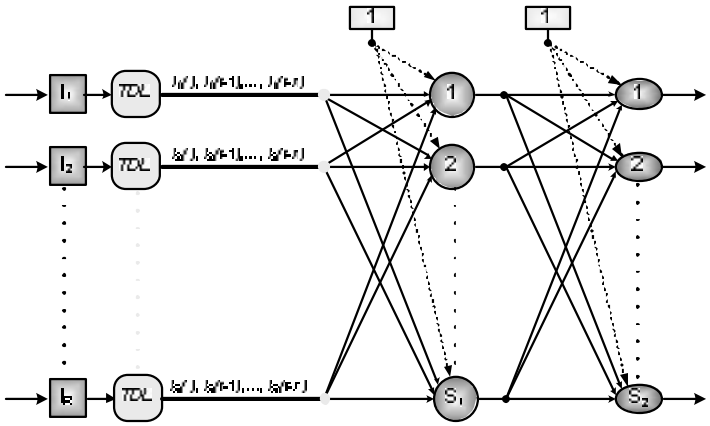



Fig. 7: A schematic of IDNN with a hidden layer

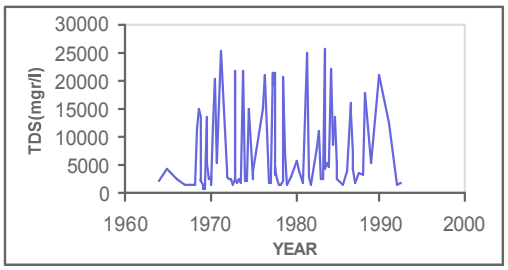



Fig. 8: Time series variations of TDS parameter in Vaniar station

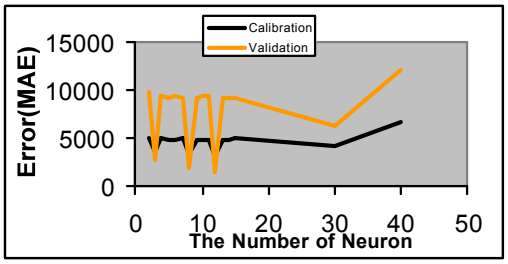



Fig. 9: Variations of RMSE error based on the number of neurons in IDNN network using flow input parameter with 24-hour time delay

As shown in Fig. 6, TDL results in inputs of present time step together with N time steps before. A TDNN is developed by Putting TDL in various layers of MLP. If TDL is placed in input layer, network is known as IDNN. The IDNN consists of two parts, the first part is memory which Saves latest information and the second part is the main body which processes information and predicts future. Figure 7 shows a schematic of IDNN with a hidden layer. Equations 5 show mathematical relations of IDNN.

$$a^1_j(t) = F\{\Sigma^D_{d=0}\Sigma^R_{i=1} w^1_{j,i,d} p_{i,d+1}(t)+b^1_j\} \qquad (5)$$

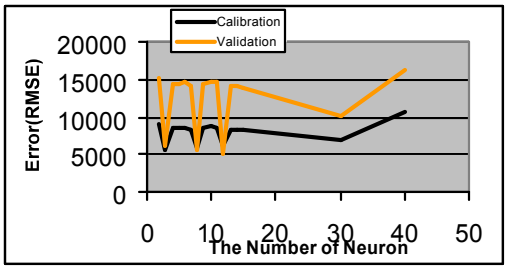



Fig. 10: Variations of MAE error based on the number of neurons in IDNN network using flow input parameter with 24-hour time delay

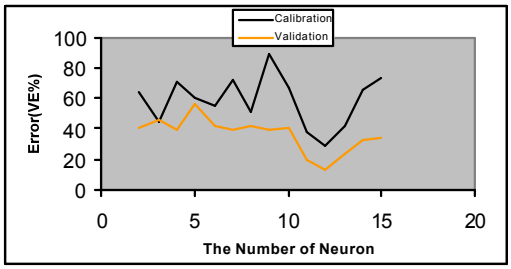



Fig. 11: Variations of VE error based on the number of neurons in IDNN network using flow input parameter with 1 month delay

D is the memory degree of Time Delay. The other terms have been previously described (4).For more detailed information related to Dynamic ANN, refer to (4) and (5).

For assessment of validity and workability of the models applied in this study three statistical methods are used as follows:

- Variance Error (VE)
- Root Mean Square Error (RMSE)
- Mean Absolute Error (MAE)

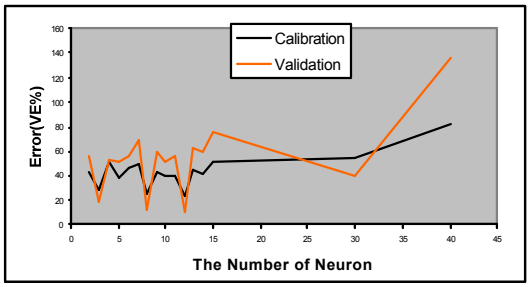



Fig. 12: Variations of VE error based on the number of neurons in IDNN network using flow input parameter with 24-hour delay

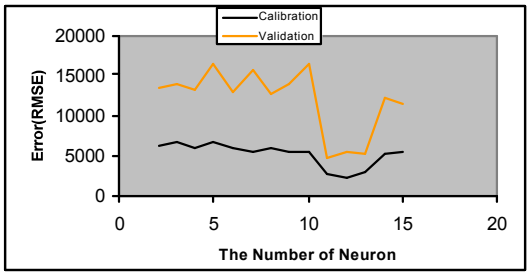



Fig. 13: Variations of MAE error based on the number of neurons in IDNN network using TDS input parameter with 1 month delay

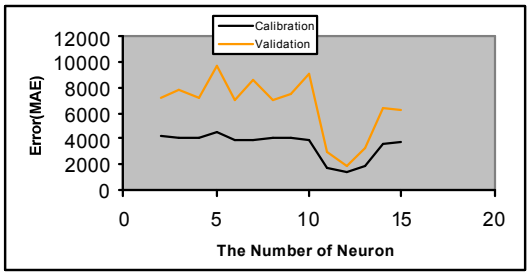



Fig. 14: Variations of RMSE error based on the number of neurons in IDNN network using TDS input parameter with 1 month delay

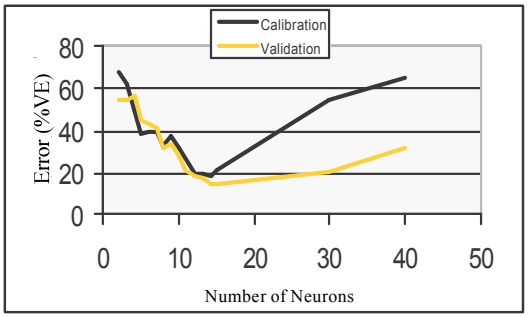



Fig. 15: Variations of VE error based on the number of neurons in MLP network using flow input parameter with 24-hour delay

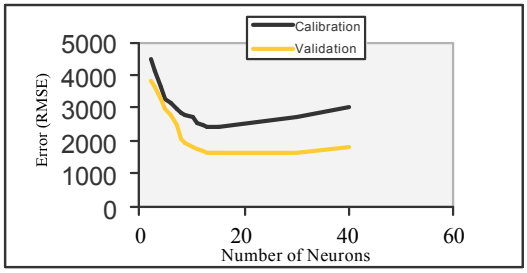



Fig. 16: Variations of RMSE error based on the number of neurons in MLP network using flow input parameter with 24-hour delay

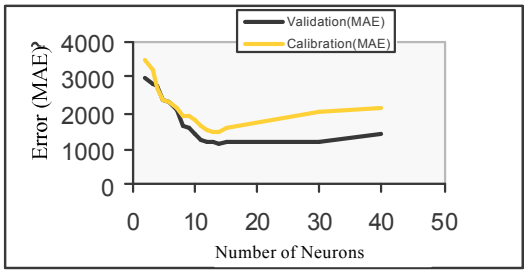



Fig. 17: Variations of MAE error based on the number of neurons in MLP network using flow input parameter with 24-hour delay

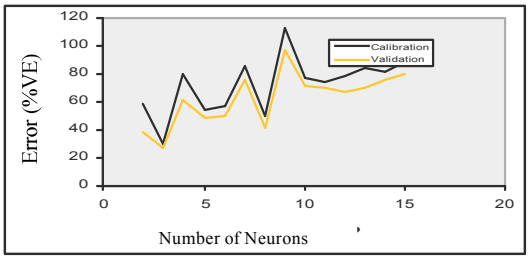



Fig. 18: Variations of VE error based on the number of neurons in MLP network using TDS input parameter with one month delay

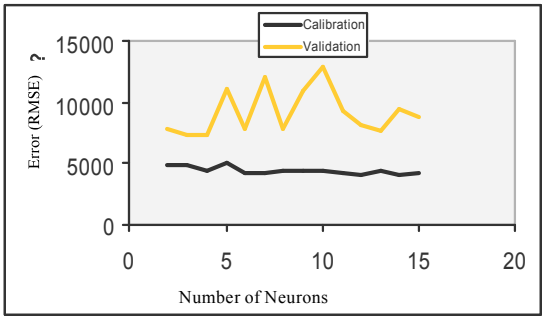



Fig. 19: Variations of RMSE error based on the number of neurons in MLP network using TDS input parameter with one month delay

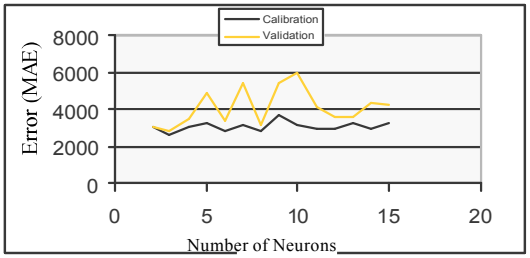



Fig. 20: Variations of MAE error based on the number of neurons in MLP network using TDS input parameter with one month delay

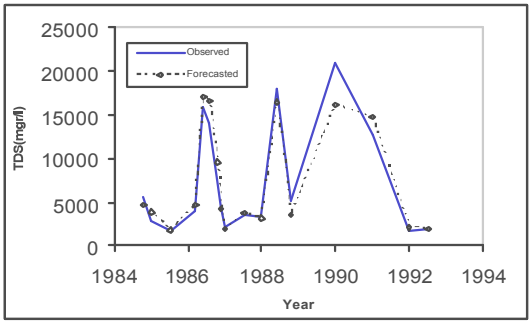



Fig. 21: Results of prediction of MLP network (Using 13 neurons in hidden layer) in Vaniar station using discharge parameter with 24 hours time delay

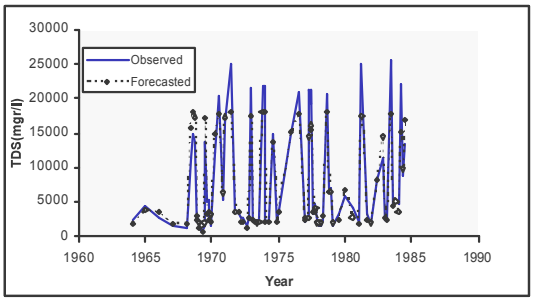



Fig. 22: Results of prediction of MLP network (Using 13 neurons in hidden layer) in Vaniar station using discharge parameter with 24 hours time delay in calibration stage in validation stage

The following equations are related to the mentioned statistical methods (9).

$$VE = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{Obs_t - For_t}{Obs_t}\right| \times 100 \qquad (6)$$

$$MAE = \frac{1}{T}\sum_{t=1}^{T}\left|Obs_t - For_t\right| \qquad (7)$$

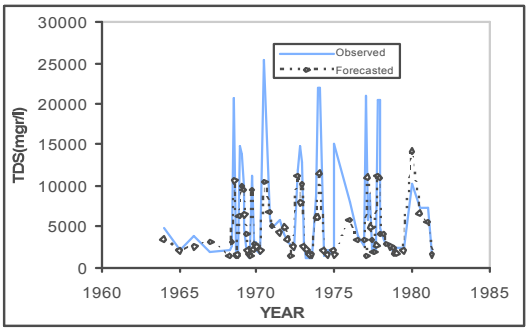



Fig. 23: Results of prediction of MLP network (Using 3 neurons in hidden layer) in Vaniar station using TDS parameter with one month time delay in validation stage

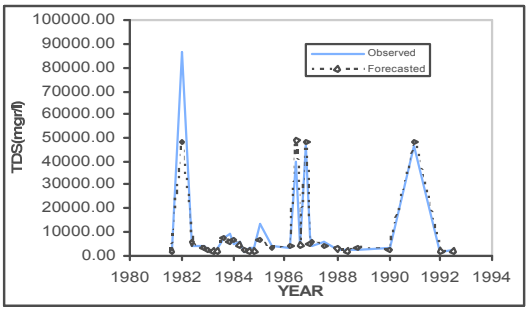



Fig. 24: Results of prediction of MLP network (Using 3 neurons in hidden layer) in Vaniar station using TDS parameter with one month time delay in calibration stage

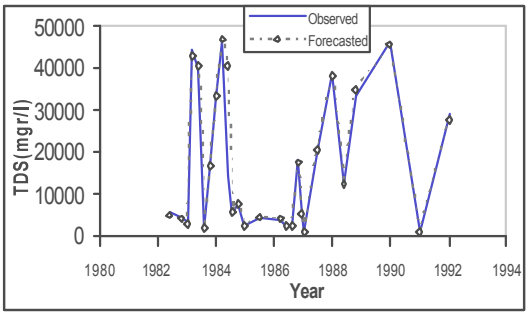



Fig. 25: Results of prediction of IDNN network (Using 12 neurons in hidden layer) in Vaniar station using discharge parameter with 24 hours time delay in validation stage

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(Obs_t - For_t)^2} \qquad (8)$$

In equations 6, 7 and 8 t parameters are defined as follows:

Table 2: Results of prediction of TDS in Vaniar station with number of hidden layer which caused minimum VE error

| | | Percentages VE error | | | |
|---|---|---|---|---|---|
| Type of network | Input parameters | Numbers of neural | Calibration | Numbers of neural | Validation |
| MLP | Discharge | 13 | 17.45 | 13 | 17.45 |
| | TDS | 3 | 30.78 | 3 | 17.39 |
| IDNN | Discharge | 12 | 22.69 | 12 | 9.88 |
| | TDS | 12 | 29.00 | 12 | 12.39 |

Table 3: Results of prediction of TDS in Vaniar station with number of hidden layer which caused minimum RMSE error

| | | RMSE (mg/l) error | | | |
|---|---|---|---|---|---|
| Types of network | Input parameters | Numbers of neural | Calibration | Numbers of neural | Validation |
| MLP | Discharge | (13,14) | 24.36 | 14 | 1635 |
| | TDS | 14 | 4082.00 | 4 | 7416 |
| IDNN | Discharge | 12 | 5767.00 | 3 | 5101 |
| | TDS | 12 | 2296.00 | 11 | 4761 |

Table 4: Results of prediction of TDS in Vaniar station with number of hidden layer which caused minimum MAE error

| | | MAE (mg/l) error | | | |
|---|---|---|---|---|---|
| Types of network | Input parameters | Numbers of neural | Calibration | Numbers of neural | Validation |
| MLP | Discharge | 14 | 1186 | 14 | 1114 |
| | TDS | 3 | 2575 | 3 | 2808 |
| IDNN | Discharge | 12 | 3114 | 12 | 1382 |
| | TDS | 12 | 1466 | 12 | 1937 |

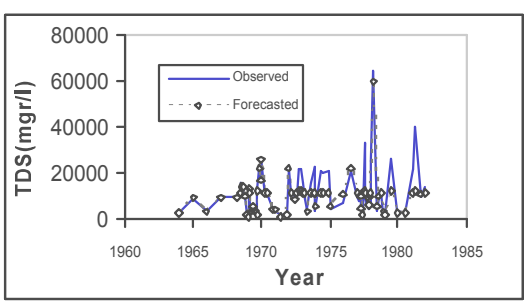



Fig. 26: Results of prediction of IDNN network (Using 12 neurons in hidden layer) in Vaniar station using discharge parameter with 24 hours time delay in calibration stage

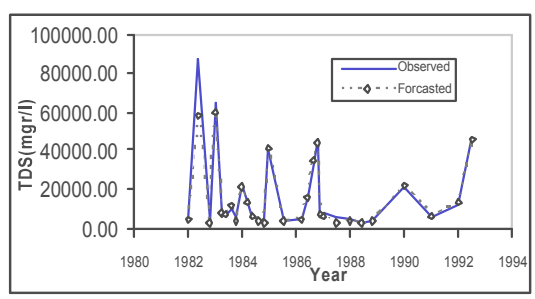



Fig. 27: Results of prediction of IDNN network (Using 12 neurons in hidden layer) in Vaniar station using TDS parameter with one month time delay h in validation stage

t: interrupted time
T: time series duration
$Obs_t$ observed parameter (as ex. TDS) 1=t=T
$For_t$ predicted parameter (as ex. TDS) 1=t=T

Also, R is the correlation coefficient which is applied to degree of validity of predictions data. The relation is defined as follows (9) :

$$R = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} \qquad (9)$$

In equation 9, $\bar{x}$ and $\bar{y}$ are means of x and y series. R shows relationship between the observed and predicted data. If relations are very strong R approaches one (9).

Table 5: R-Squared value between observed and prediction data in different networks with different input parameters

| Type of networks | Input parameters | R-squared value Validation | Calibration |
|---|---|---|---|
| MLP | discharge | 0.929 | 0.909 |
| | TDS | 0.859 | 0.622 |
| IDNN | discharge | 0.906 | 0.706 |
| | TDS | 0.949 | 0.891 |

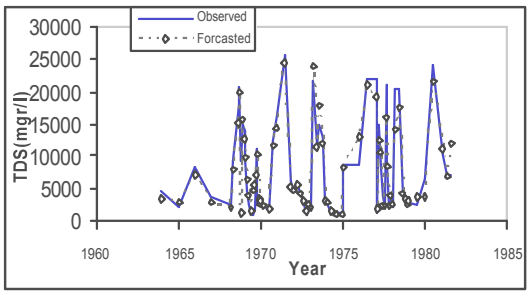



Fig. 28: Results of prediction of IDNN network (Using 12 neurons in hidden layer) in Vaniar station using TDS parameter with one month time delay in calibration stage

## RESULTS

As mentioned previously, the objective of this study is to determine workability of ANN in prediction of future salinity in Vaniar station in Achechay river basin (Vaniar dam), using discharge parameter for a 24-hour time delay and TDS with one- month time delay. MLP and IDNN techniques were applied to the data. The data were collected by Water Authority of Azerbaijan Province in Iran and consist of thirty-year monthly mean data of TDS and hourly mean data of discharge. Last ten years data were used for validation of the models. MATLAB Software version 7/1 and cased of Neural Network were applied to the data. Figure 8 shows time series variations of TDS parameter data in Vaniar station.

Regarding the theory of universal approximation, all the ANN applied in this study have relatively similar structures and the difference lies in the number of Neurals in hidden layer. Figure 9-20 show suitable number of hidden layers which gives the minimum error according to VE, RSME and MAE methods. Proper number of hidden layers for prediction of TDS with minimum error observed in the figures is summarized in Table 2-4.

The optimimum delay in IDNN dynamic network is 4, in other words, in addition to the current time step, the signals are effective on the salinity of current time step for 4 steps before (up to 5 days for flow and 5 months for TDS).

The results of predictions of artificial Neural network using flow parameter with a 24 hour delay and TDS with one month delay will be followed respectively. The prediction graphs are fist presented for the models and the table of errors will be followed for calibration and verification stages.

Figures 8-15 show the results of predictions for different networks with different input parameters and considering the number of hidden layers.

## CONCLUSIONS

In this study, MLP and IDNN artificial Neural network models were applied for prediction of water salinity in Achechay river basin (Vaniar station).IDNN results showed higher accuracy compared to MLP. In more than 50% of cases, both applied models predicted future salinity of water with high accuracy (above 80%). The results obtained from both models of Neural network showed their acceptable precision in prediction of salinity in the study area. The errors in these prediction models mainly originate from the shortness of time series used in calibration and verification stages. From the results of the studies it can be stated that using artificial Neural network for predicting the effect of flow and TDS parameters on salinity in the Basin of Apalachicola River can be considered as an effective option. the results of this study can be utilized in optimized management and planning of water resources of the study area,

## REFERENCES

1. Mortazavi, Vahid, 2004. Summary of studies on water resources of Achechayy basin. first Achechayy conference, Tabriz.
2. Ghanbari, Ali, Plan of Achechayy salinity control.
3. Ghanbari, Ali, 2004. Components of Achechayy plan. first Achechayy conference.
4. Manhaj, Mohammadbagher, 1998. Computational intelligence (Vol I), principles of artificial neural networks. Amirkabir University.
5. Manhaj, Mohammadbagher and Seyfipour, Navid, 1998. Computational intelligence (Vol II), application of artificial intelligence in control. Professor Hesabi Publications.
6. ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000a. Artificial Neural Networks in Hydrology: Preliminary Concepts. Part I, J. Hydrologic Engrg., ASCE, 5 (2): 115-123.

7. ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000b. Artificial Neural Networks in Hydrology: Hydrologic Applications. Part II, J. Hydrologic Engrg., ASCE, 5 (2): 123-137.
8. Hornik, K.M., M. Stinchcombe and H. White, 1989. Multilayer Feedforward Networks are Universal Approximators. Neural Networks, Vol: 2.
9. Kennedy, J.B.and A.D. Neville, 1964. Basic statistical methods for engineers and scientists, Happer International Edition.
10. Wenrui Huang and Simon Foo, 2001. Neural network modeling of salinity variation in Apalachicola River.
11. Holger, R. Maier, Greame and C. Dandy, 1998. Use of artificial neural networks for modeling cyanobacteria Anabaena spp. In the Rivar Murray, South Australia.
12. Sezin Tokar A. and Momcilo Markus, 1999. Precipation. Runoff modeling using artificial netural networks and conceptual models.