Modeling the Water Quality of River by Using Artificial Neural Networks (Case Study, Dez River, Iran)

B. Karami[1], M. Golabi[2] and K. N. Dhumal[3]

[1] Ph. D. Student of Environmental Sciences, Department of Environmental Sciences. University of Pune, India.
Email: karamibehnam@yahoo.com
[2] Ph. D. of Irrigation and Drainage, Khuzestan Water and Power Authority
Email: mona_golabi@yahoo.com
[3] Prof of Department of Environmental Sciences. University of Pune, India.

Abstract

Water is a vital resource for all aspects of human and ecosystem survival and health. Thus, its quality is also important. Water quality refers to the composition of a water sample. The interpretation of data may be difficult and lengthy. Evaluations of water quality parameters are necessary to enhance the performance of an assessment operation and to develop better water resources management and plan. New approaches such as Artificial Intelligence techniques have proven their ability and applicability for simulating and modeling various physical phenomena in the water engineering field. In additional, Artificial Neural Network (ANN) captures the embedded spatial and unsteady behavior in the investigated problem using its architecture and nonlinear nature compared with the other classical modeling techniques. To achieve goals of this research program, the Dez River in Iran is selected to assess the capability of ANNs for water quality simulation. This river is located in Khuzestan province in the southwest of the Iran. Several water quality variables including flow discharge, HCO_3^-, CO_3^{2-}, SO_4^{2-}, Cl^-, Na^+, Mg^{2+}, Ca^{2+} and K^+ were used to simulate electrical conductivity (EC), total dissolve solid (TDS) and sodium adsorption ratio (SAR). Data from 2000 to 2010 at monitoring stations including Dezful, Abshirin and Bamdezh has been used. Seventy percent of data is used for training the ANNs and thirty percent of available data is used for validation purposes. Qnet2000 software is selected for modeling purposes in the present research. Results show that Qnet2000 is able to simulate water quality of the Dez River very successfully with 90% accuracy.

Key words: Water Quality, Artificial Neural Networks, Qnet2000, Dez River.

Introduction

Water quality directly affects virtually all water uses. Fish survival, diversity and growth; recreational activities such as swimming and boating, municipal, industrial, and private water supplies, agricultural uses such as irrigation and livestock watering, waste disposal, and general aesthetics-all are affected by the physical, chemical, biological, and microbiological conditions that exist in watercourses and in subsurface aquifers. Water quality impairment often triggers conflict in a watershed, simply because degraded water quality means that desired uses are not possible or are safe (Heathcote, 1998).

Nowadays, water resources are the main basis for stable development around the world. Therefore, quality and quantity of these resources are so important. Khuzestan Province in south western Iran relies heavily on water from Karoun, Karkheh, Dez, Jarahi and Zohreh rivers. Among these rivers, Dez is one of the most important waterway rivers in the country, which provides water for many cities, villages, agricultural projects and industrial factories.

Therefore, the flow discharge of the river is decreased constantly. Decreasing discharge of Dez River and use of the river as receiving water for several industries in the basin have caused deterioration of the river water quality. Responsible development of the river basin in the future will prove to be more crucial than it is today. So, knowledge about future changes, simulation and forecasting water quality of Dez River, industry, population and development different project are so important. In recent years, several deterministic water quality models such as; QUAL2E, WATEVAL, MIKE11, HEC5Q and WASP have been developed to investigate about waterways.

Artificial Neural Networks (AANs) are one type of intelligent artificial system which was simulated from living neural cells. They can learn new mechanisms related to everyday phenomena. They are able to remember their knowledge like human brain and they can apply this knowledge to the same phenomena to predict their variation. The backgrounds of Artificial Neural Networks go back to 19^{th} and 20^{th} century. In that period of time some studies in physics, psychology and physiology were performed by scientists such as, Ivan Pavolov.

The new approach on Artificial Neural Networks began with Warren Mc Culloch and Walter Pitts's investigations. They showed that neural networks can calculate each logical and arithmetical function. The first applications of neural networks were introduction of Perceptron Network by Frank Rosenblant (1958) and delineation Adelin comparative Neural Network by Bernurd Widrow in 1960.

With the advent of new computer technology, the use of ANNs increased dramatically. (Menhaj, 2000).

In general, the proceedings of development neural networks have been divided to three stages:

1) Study on development of neural neuron and determination restrictive factors by Minsky and Papert.
2) Finding and generalization back propagation algorithm by Rummelhart and Mcland.
3) Evaluation restrictive of neural networks and comparison with other methods such as, genetic algorithm and phase theory (Dawson and Wilby, 2001).

Each neural network consist of three layers; input layer with dependent variable, output layer with independent variable and one or more hidden layer with proceeding element. Network is trained for some existing data and if this did with suitable iteration and transfer function can be expected to analyze phenomenon and give suitable results later. Schematic of an Artificial Neural Network is shown in Figure 1.

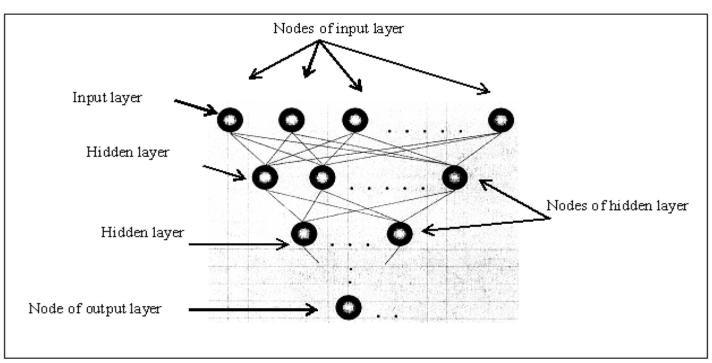

Figure 1. Components of an Artificial Neural Network

Recently, applications of ANNs in the areas of water engineering, ecological sciences, and environmental sciences have been reported since the beginning of the 1990s. Artificial Neural Networks have been used in most branches of science. In recent years, ANNs have been used intensively for prediction and forecasting in a number of water-related areas, including water resource studies (Liong et al., 1999, 2001; Muttil and Chau, 2006; El-Shafie et al., 2008), oceanography (Makarynskyy, 2004), and environmental science (Grubert, 2003).

The use of data-driven techniques for modeling the quality of both freshwater (Chen and Mynett, 2003) and seawater (Lee et al., 2000, 2003) has met with success in the past decade. Reckhow (1999) studied Bayesian probability network models for guiding decision making regarding water quality in the Neuse River in North Carolina. Bowers (2000) developed a model to predict suspended solids conceder local precipitation, stream flow rates and turbidity as input. Hatim (2008) employed an ANN approach using six variables for the initial prediction of suspended solids in the stream at Mamasin dam. Wen and Lee (1998) applied a neural network approach to multi objective optimization for water quality management in a river basin. Neural networks have been the base for several researches such as neural networks assisting in water quality modeling (Daniell and Wundke, 1993), application of Artificial Neural Network for water quality management (Zaheer and Bai, 2003), forecasting salinity using neural network and multivariate time series model (Maier and Daudy, 1994), river flow modeling using Artificial Neural Networks (Ozgur, 2004) and application of Artificial Neural Network in stage-discharge relationships (Bhattahcharya and Solomatine, 2000).

This paper demonstrates the application of ANNs to model the values of selected river water quality parameters. In addition, objective of this study is to determine the most and the least effective water parameters on EC, TDS and SAR of Dez river, Iran.

Materials and Methods

Dez river is one of the most important river of Iran. Junction of Bakhtiari, Zalaki, Sezar, Tireh and Marbareh river form Dez river. Dez river from south of Lorestan province and north of Khuzestan province convey to the Karoun river.

In this study, Qnet2000 is used to simulate water quality of the Dez river. Data used in this paper includes flow discharge, HCO3-, CO32-, SO42-, Cl-, Na+, Mg2+, Ca2+, K+, EC, TDS and SAR for 10 years from 2000 to 2010. 70% of data is used for training the ANNs and 30% of available data is used for validations.

Data was measured several days of each month. The average of measured data was divided on the maximum of each series to achieve relative quantity of the selected parameters, which were dimensionless between 0-1 to speed up the calculation. Flow discharge, HCO3-, CO32-, SO42-, Cl-, Na+, Mg2+, Ca2+ and K+ were used as independent variables. EC, TDS and SAR were considered as dependent variables to be simulated.

First, 70% of data and two cases represented below were selected in the training stage.

1) Input layer with 9 nodes, one hidden layer with 9 node
2) Input layer with 9 nodes, one hidden layer with 10 node

For example, Figure 2 shows network designed of EC with one hidden layer with 10 nodes.

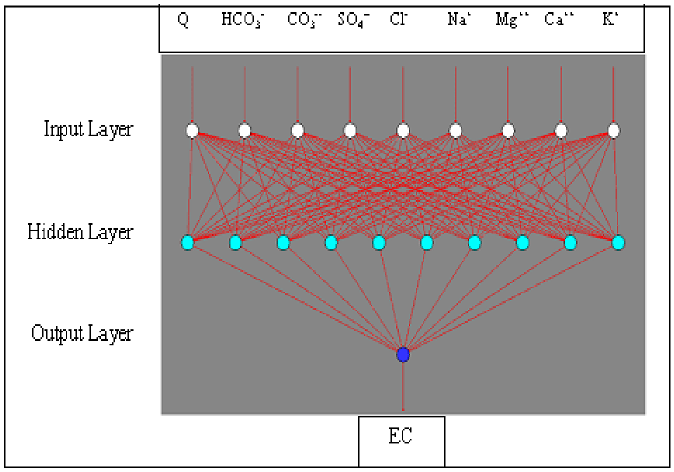

Figure 2. Network designed of EC with one hidden layer with 10 nodes

To chase the modeling purpose, different transfer function for hidden and output layers were selected. Also, the model was run with 10000 to 200000 of iteration. The purpose of these cases was comparison between measurement and simulated values, determination of suitable transfer function for hidden and output layer, suitable number of node of hidden layer in the training stage. This stage was selected the best case using statistical criteria. For that reason, model performance criteria such as RMSE and R2 were used.

Formulas for calculating RMSE and R2 are given as follows:

$$RMSE = \left[\frac{\sum_{i=1}^{n} (P_i - O_i)^2}{n} \right]^{1/2} \times \frac{100}{\overline{O}} \qquad (1)$$

$$R^2 = \frac{\sum_{i=1}^{n} \left(P_i - \overline{O}\right)^2}{\sum_{i=1}^{n} \left(O_i - \overline{O}\right)^2} \qquad (2)$$

Where Pi is the predicted data, Oi the measured data, $\overline{O}$ average of measurement data and n is the number of samples.

After the best transfer function for hidden and output layers, number of node of hidden layers and iteration was selected, the model with 30% of data was validated.

Finally, the performance of model for simulation of EC, TDS and SAR were evaluated.

Results and Discussion

The modeling process includes training and validation. Two statistical criteria have been used for comparing the models outputs: root mean square error (RMSR) and the coefficient of determination. Root mean square error equal to zero and coefficient of determination equal to one denote the best agreement between simulated and observed data. In each station Sigmoid, Gaussian, Hyperbolic tangent and Hyperbolic secant by way of transfer function for hidden and output layer, 10000-200000 iteration, one hidden layer with 9 and 10 nodes were tested and were determined that Gaussian and Hyperbolic tangent for hidden and output layer, 200000 iterations and 10 nodes in hidden layer were represented the best results. After that the effect input variation on output were evaluated. Table 1 includes the results of model in training at Dezful station with 10000 iterations, one hidden layer, 9 nodes and different transfer function.

Table 1. Comparison of using different function for TDS prediction in Dezful station

Hidden	Output layer	RMS	R^2	Hidden	Output layer	RMS	R^2

layer function	function			layer function	function		
Sigmoid	Sigmoid	0.0205	0.9839	Gaussian	Sigmoid	0.0188	0.9863
Sigmoid	Gaussian	0.3779	-0.2887	Gaussian	Gaussian	0.3779	0.4394
Sigmoid	Tanh(x)	0.0199	0.9846	Gaussian	Tanh(x)	0.0185	0.9868
Sigmoid	Sech(x)	0.0201	0.9843	Gaussian	Sech(x)	0.0188	0.9863
Tanh(x)	Sigmoid	0.0195	0.9853	Sech(x)	Sigmoid	0.0194	0.9855
Tanh(x)	Gaussian	0.3779	-0.3992	Sech(x)	Gaussian	0.3779	0.5767
Tanh(x)	Tanh(x)	0.0192	0.9858	Sech(x)	Tanh(x)	0.0188	0.9863
Tanh(x)	Sech(x)	0.0190	0.9859	Sech(x)	Sech(x)	0.0193	0.9855

The results of models for each station, maximum and minimum of effective input parameters in the training stage have been shown in Table 2 and Table 3.

Table 2. Predicted TDS, SAR and EC for different stations on Dez river, Iran (200000 Iterations and 10 nodes in hidden layer)

Stations	variables	RMS	R^2	Maximum Input effective parameter	Minimum Input effective parameter
Dezful	TDS	0.01265	0.9978	Na	CO_3
	EC	0.01050	0.9991	Cl	Q
	SAR	0.00114	0.9999	Na,Ca,Mg	Q
Abshirin	TDS	0.01264	0.9965	Na	Q
	EC	0.01301	0.9983	Cl	Q
	SAR	0.00113	0.9999	Na,Ca,Mg	CO_3
Bamdezh	TDS	0.01205	0.9939	K	Q
	EC	0.01234	0.9940	Cl	HCo_3
	SAR	0.00113	0.9999	Na,Ca,Mg	Q

Table 3. Results of ANN model validation in different stations on Dez river, Iran

Station name	Iterations	No. of hidden layer Nodes	R^2(TDS)	R^2(EC)	R^2(SAR)
Dezful	10000	9	0.90020	0.92003	0.94051
		10	0.91130	0.93000	0.94631
	200000	9	0.95421	0.96105	0.96870
		10	0.95670	0.97780	0.97500
Abshirin	10000	9	0.92500	0.94587	0.90100
		10	0.92961	0.95006	0.91235
	200000	9	0.94780	0.98610	0.96541
		10	0.94900	0.99000	0.96890
Bamdezh	10000	9	0.91783	0.90022	0.90282
		10	0.91936	0.90056	0.92878
	200000	9	0.93381	0.95910	0.91323
		10	0.97532	0.96566	0.93515

Overall, it seems ANN is able to predict several chemical water quality parameters of the Dez River very successfully. After determination of the best transfer function of hidden and output layer, suitable number of node in hidden layer and iteration number, EC, TDS and SAR in each station were predicted. Figure 3 compares the predicted TDS profile along the selected reach of the river.

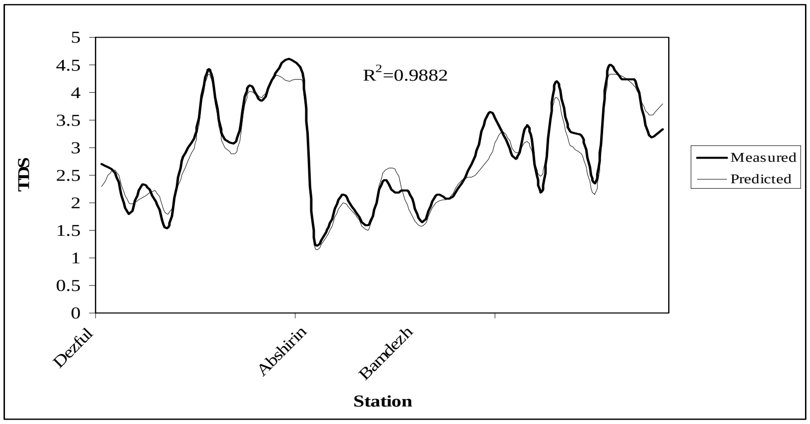

Figure 3. Comparison of predicted and measurement TDS along Dez river, Iran

It is shown a very good agreement between predicted and measured profiles with R^2= 0.99 which shows that ANN is capable of TDS simulation in the Dez river.

Figure 4 compares the predicted profile of the electrical conductivity of the Dez river water using ANN along the selected reach. The predicted profile is very close to the measured profile. The coefficient of determination is R^2= 0.97. It denotes that ANN can produce the TDS profile very successful.

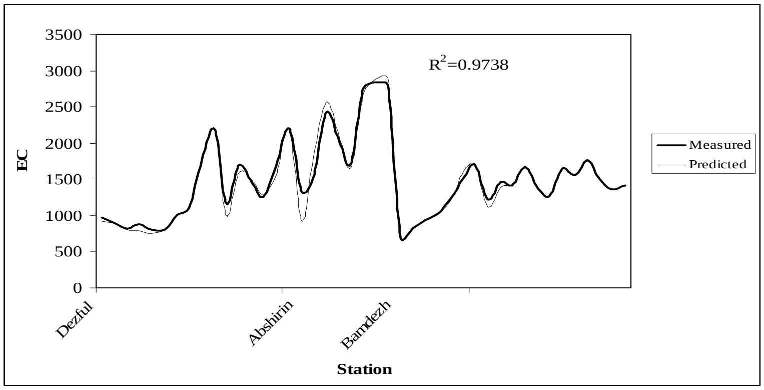

Figure 4. Comparison of predicted and measurement EC along Dez river, Iran

Finally, Figure 6 presents the comparison between predicted and measured SAR profiles. R^2 of almost 90% for this prediction shows good agreement. Overall, it seems ANN is able to predict several chemical water quality parameters of the Dez river very successfully.

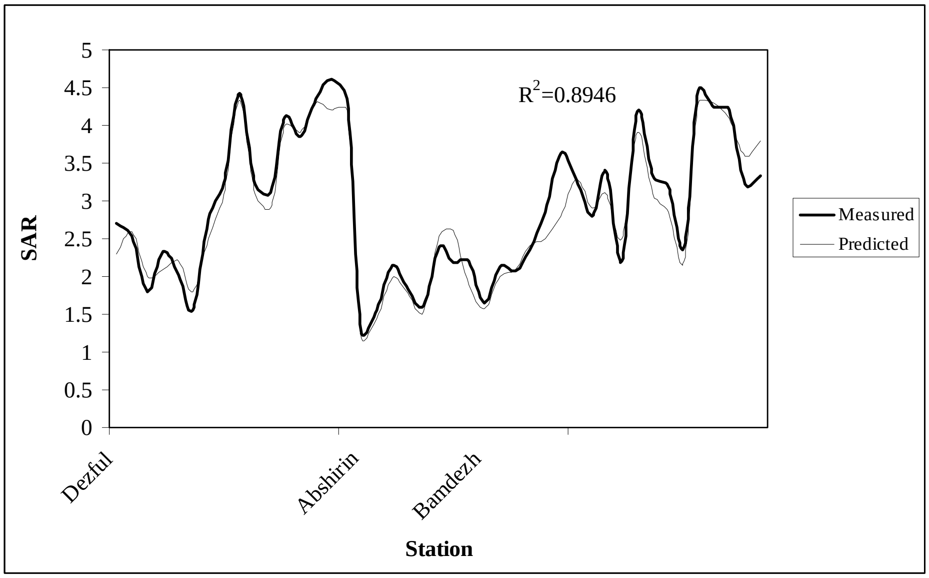

Figure 5. Comparison of predicted and measurement SAR along Dez river, Iran

Conclusion

This research presents a study on application of the Artificial Neural Network to water quality simulation in the river system. Using river data such as discharge, HCO3-, CO32-, SO42-, Cl-, Na+, Mg2+, Ca2+, K+, EC, TDS and SAR, Qnet2000 as an ANN has been trained. 70% of data have been used for training and 30% of data have been used for validation. In training and validation the coefficient of determination between predicted and measured results was more than 90%. Also, according to Figure 3-5 the coefficient of determination (R2) between measured and predicted water quality variables is approximately 90%. This shows the ability of ANNs in forecasting several water quality variables in the river systems.

Results show that the ANN model with Gaussian and Hyperbolic tangent functions for hidden and output layer with 10 nodes in hidden layer and 200000 iterations gave the best results. Also, in TDS simulation the most effective parameters were Na+, Ca2+, K+, SO42-and Cl- and the least effective input parameters were discharge and CO3--. In EC simulation the most effective input parameter were Ca2+, SO42-and Cl- and the least effective input parameters were discharge, CO32-, HCO3- , K+ and Mg2+. Also, in SAR simulation the most effective input parameters were Ca2+, Mg2+and Na+ and the least effective input parameters were discharge, CO32- and K+.

Overall, it is shown that Qnet2000 is a very trustable ANN model for simulation of the water quality variables in the river systems.

Acknowledgements

This research was supported by the Khuzestan Water and Power Authority and the Office of Research and Irrigation and Drainage Networks Standard. The authors are grateful for their support.

References

- Bhattacharya, B., Solomatine, D. P., 2000. “Applied of Artificial Neural Network in Stage-Discharge Relationships”. Proceeding of 4th International Conference on Hydroinformation. Iowa City, USA., 23-27 July, pp: 1-7.
- Bowers, J.A., Shedrow, C.B., 2000. Predicting stream water quality using artificial neural networks. U.S. Department of Energy Report WSRC-MS-2000-00112, 7 pp.

- Chen, Q., Mynett, A.E., 2003. Integration of data mining techniques and heuristic knowledge in fuzzy logic modeling of eutrophication in Taihu Lake. Ecological Modeling 162 (1/2), 55–67.
- Daniell, T. M., Wundke, A. D., 1993. Neural Network – Assisting in Water Quality Modeling. Australian Institute of Engineers.
- Dawson, C. W., Wilby, R. L., 2001. Hydrological Modeling Using Artificial Neural Network. Progress in Physical Geograph., Vol. 1, No. 25, PP. 80-108.
- El-shafie,A.E. Noureldin, M.R. Taha and H. Basri., 2008. Neural network model for Nile River inflow forecasting analysis of historical inflow data. Journal of Applied Sciences 8 (24), 4487-4499.
- Grubert, J.P., 2003. Acid deposition in the eastern United States and neural network predictions for the future. Journal of Environmental Engineering and Science 2 (2), 99–109.
- Hatim Elhatip, M. Aydin Ko¨mu¨r., 2008. Evaluation of water quality parameters for the Mamasin dam in Aksaray City in the central Anatolian part of Turkey by means of artificial neural networks, Environ Geol, 53:1157–1164 DOI 10.1007/s00254-007-0705-y
- Heathcote, Isobel W., 1998. Integrated Watershed Management: Principles and Practices, John Wiley & Sons, Inc. New York.
- Lee, J.H.W., Huang, Y., Dickmen, M., Jayawardena, A.W., 2003. Neural network modeling of coastal algal blooms. Ecological Modeling 159, 179–201.
- Lee, J.H.W., Wong, K.T.M., Huang, Y., Jayawardena, A.W., 2000. A real time early warning and modeling system for red tides in Hong Kong. In: Wang, Z.Y. (Ed.), Proceedings of the Eighth International Symposium on Stochastic Hydraulics. Balkema, Beijing, pp. 659–669.
- Liong, S.Y., Lim, W.H., Paudyal, G., 1999. Real time river stage forecasting for flood Bangladesh: neural network approach”. Journal of Computing in Civil Engineering ASCE 14(1), 1 8.
- Makarynskyy, O., 2004. Improving wave predictions with artificial neural networks. Ocean Engineering 31, 709–724.
- Maier, H. R., Dandy, G. C., 1994. Forecasting Salinity Usinig Neural Network and Multivariate Time Series Models. Water Down, Adelaide, Australia.
- Menhaj, M., 2000. Principles of neural networks (Pesian). Amir-Kabir University Press.
- Muttil, N., Chau, K.W., 2006. Neural network and genetic programming for modeling coastal algal blooms. International Journal of Environment and Pollution 28 (3/4), 223–238.
- Ozgur, K., 2004. River Flow Modeling Using Artificial Neural Networks. Journal of Hydrologic Engineering., ASCE, Vol. 9, No. 1, PP. 60-63.
- Reckhow, K.H., 1999. Water quality prediction and probability network models. Canadian Journal of Fisheries and Aquatic Sciences 56, 1150–1158.
- Wen, C. G., Lee, C. S., 1998. A Neural Network Approach to Multiobjective Optimization for Water Quality Management in A River Basin. Water Resources Research, Vol. 34, No. 3, PP. 427-436.
- Zaheer, I., Bai, C. G., 2003. Application of Artificial Neural Network for Water Quality Management. Lowland Technology International, Vol. 5, No. 2, PP. 10-15