World Journal of Environmental Biosciences
All Rights Reserved WJES © 2014
Available Online at: **www.environmentaljournals.org**

Volume 6, Supplementary: 26-39

ISSN 2277- 8047

Analysis of parameters BOD, DO, nitrite and nitrate for Jajrood River and its modeling by Artificial Neural Networks (ANN)

Alireza Parsa[1], Dr. Seyed Ahmad Mirbagheri[2], Dr. Seyedeh Hoda Rahmati[3]

[1]Water Resources Environmental Engineering Graduate Student, School of Environment and Energy, Islamic Azad University Science and Research Branch of Tehran.
[2]Professor and Chief of Department of Water Resources Environmental Engineering, Faculty of Environment and Energy, Islamic Azad University Science and Research Branch of Tehran.
[3]Assistant Professor of Environmental Engineering, Water Resources, Faculty of Environment and Energy, Islamic Azad University Science and Research Branch of Tehran.

ABSTRACT

Jajrood River is one of the most important rivers in Tehran province, supplying the freshwater for more than 30% of Tehran population. So, ensuring its water quality and environmental impacts is very important. It is performed by computer modeling, as a powerful managerial tool, for simulation of quality parameters like BOD, DO, nitrite and nitrate. First it is necessary to determine a suitable span of the river for sampling and studying. The objective of this study was to investigate the river parameters like BOD, DO, nitrite and nitrate and its modeling by regression and ANN. The span selected ranged from GarmabDarreh village to the beginning of Latian dam with a 30 km length. The main problem was found to be microbial and organic contaminations. Quality simulation of Jajrood River during four months (Ordibehesht, Mordad, Aban and Bahman) was performed by regression and ANN to model the four parameters of nitrate, nitrite, biochemical oxygen demand (BOD) and dissolved oxygen, for which SPSS and MATLAB were used. The results show that the model could do the simulation satisfactorily. Also, due to the accelerated aeration, Jajrood River has a suitable self purification capacity along the way. Based on the results of the DO and BOD parameters, it was found that these parameters are in good status even in the arid summer months. The results for other parameters are revealed in the corresponding chapters accordingly.

Key words: *Jajrood River, modeling, regression, neural network, nitrate, BOD, DO, Latian Dam, quality parameters*

Corresponding author: Seyyed Ahmad Mirbagheri

INTRODUCTION

Rivers and flowing waters have been in the human's interest from ancient times. Oftentimes, human civilizations, industrial complexes and agricultural lands were made near rivers to utilize water supplies. By using more of the water supplies and unnatural manipulations, water quality of the rivers have been declined. In this light, assessment, analysis and interpretation of water quality data make the data usable in different areas and also, make available appropriate managerial methodologies to reduce the water pollutions and move toward an acceptable quality [7]. So, various studies were made in the analytical areas in 50s and 60s. On the whole, however, the analytical approach is complicated and needs simplifying assumptions, especially related to modeling the transmission issues. Thereafter, attempts were made aiming at numerical methods at the same of computer technology development. Finite difference method and finite element method, along with the time series provided a major collection of numerical models. Qualitative models are able to predict the concentration variations of different parameters in the aqueous medium. So, they are aimed at providing solutions to eliminate the imbalance in the aqueous medium of the river [3].

Materials and methods

Jajrood basin, 1892 km wide, is located on the eastern/northeastern part of Tehran with a 35.25 to 36 ° altitude and a 51.23 to 52.6 ° longitude [1]. The main division of the river, 140 km long, originates from the Kharsang mountains on the northeastern Tehran, actually belonging to the southern part of Central Alborz Mountains. It runs towards the south and after Fasham and Oshan Villages it diverts its way toward southeast and along the Rudak and Lashkarak it pours into the Latian Lake and arrives at Jajrood area after passing Latian Dam. In the climatic sense, Jajrood is a moderate, semi-arid area with an average precipitation of 250 ml. a part of the basin, with an area of 692km2, which is higher than the Latian Dam supplies a great part of Jajrood Rivers water. Another part with a 1200 km2 area has a lesser role in supplying the River's water due to less precipitation [3]. Fig.1 shows the divisions of Latian Dam area and its villages. [6]

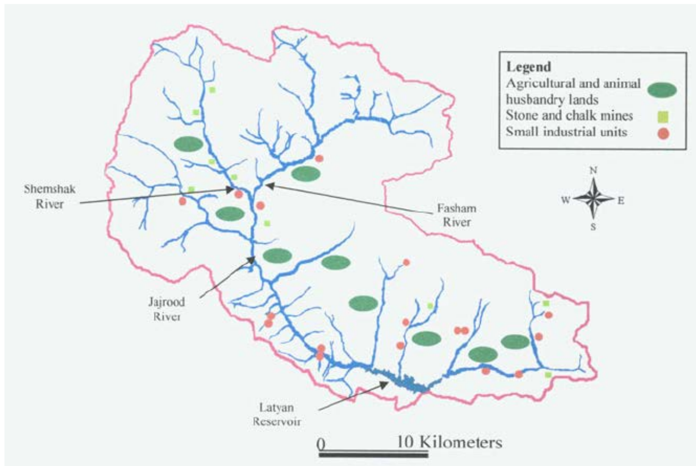


Figure 1: The divisions of Latian Dam area

Artificial Neural Networks (ANN)
Artificial networks are mathematical models or forms of algorithm calculation by taking the biological neuron systems into account. In practical terms, neural networks are non-statistical modeling tools used for modeling the complex relation between input and output [8]. Two or more neurons may combine in a layer and a network may be constructed from such a layer. Fig.2 shows a single layer network with R inputs and S neurons. Fig. 3 shows a three layer network with segregated information based on layers.

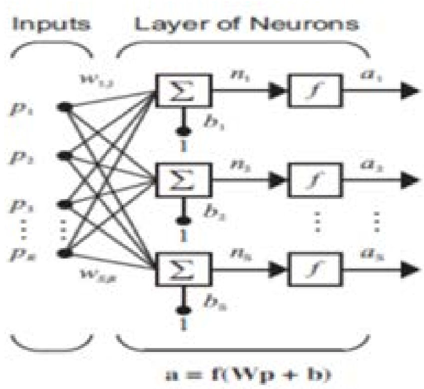
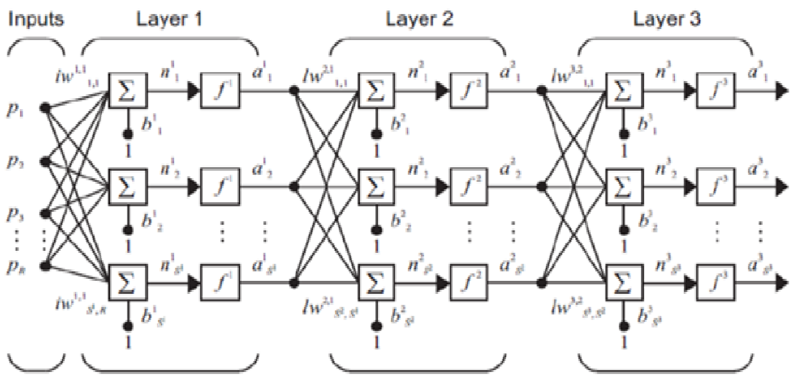


Figure 2: Single layer network with R inputs and S neurons [5]



$$a^1 = f^1(IW^{1,1}p+b^1) \qquad a^2 = f^2(LW^{2,1}a^1+b^2) \qquad a^3 = f^3(LW^{3,2}a^2+b^3)$$

$$a^3 = f^3(LW^{3,2}f^2(LW^{2,1}f^1(IW^{1,1}p+b^1)+b^2)+b^3)$$

Figure 3: 3-layer network with segregated data [5]

Selection and preprocessing of data

Before we begin, it is necessary to perform some preliminary tasks, specifically determination of a time period for quality sampling of river water and delimitation of a length along the river for sampling and studying. The modeling in the present paper aims at the 4 parameters of BOD, DO, nitrite and nitrate by intelligent neural networks. We used a 1-year period between spring and winter, 1394. A month in every season was considered as index. In order to simulate the water quality of the river by calibration and verification, water quality parameters are needed. So, the parameters have to be measured at some points of the river. Quality parameters for simulation of Jajrood water quality and model calibration for the river are: BOD, DO, nitrite and nitrate. Quality parameters were provided by the Tehran Province water and Wastewater. A limited number of stations were selected which were representative of the river specifications in the selected interval. So, the Station No.8 was selected.

Table 1: Quality parameters of Jajrood River taken from sampling stations for Ordibehesht

After Lashkarak Bridge	**After Zard-Band**	**After Roodak**	**After Haji-Abad**	**After Oshsan**	**Do-Ab**	**After - Zayegan**	**Mixture from Golabdar and Abnik**	**Parameter**
9	8	8	10	10	8	9	9	Temperature
306	306	303	270	261	287	287	279	Electric Conductance
8.2	8.8	8.4	8.5	8.6	8.6	8.8	8.2	Dissolved Oxygen
1	1.3	1	2	1.6	1.8	2	1	Biologically required oxygen
5.6	5.7	5.7	9.1	9	4.1	8.2	8.1	Nitrate
0.24	0.21	0.19	0.23	0.37	0.24	0.06	0.06	Inorganic phosphate
120	120	116	114	110	116	126	122	Total alkalinity
8.01	7.91	7.91	7.66	7.66	7.18	8.1	8.04	PH
0.37	0.35	0.31	0.47	0.43	0.4	0.12	0.39	Ammonia

Table 2: Quality parameters of Jajrood River taken from sampling stations for Mordad

After Lashkarak Bridge	After Zard-Band	After Roodak	After Haji-Abad	After Oshsan	Do-Ab	After - Zayegan	Mixture from Golabdar and Abnik	Parameter
16	14	14	13	13	13	14	13	Temperature
420	435	429	406	407	392	335	361	Electric Conductance
8.7	7.8	8.1	8.7	8.4	8.4	7.5	7.2	Dissolved Oxygen
1.5	2.1	1.7	4	1.9	3.6	1	3.3	Biologically required oxygen
5.7	5.2	5.7	8.5	8.5	7.9	6.8	6.6	Nitrate
0.02	0.02	0.04	0.08	0.05	0.06	0.07	0.06	Inorganic phosphate
150	162	164	162	162	162	144	152	Total alkalinity
7.96	7.84	7.8	7.98	7.99	7.8	8.03	7.94	PH
0.07	0	0.05	0.15	0.15	0.13	0.05	0.06	Ammonia

Table 3: Quality parameters of Jajrood River taken from sampling stations for Aban

After Lashkarak Bridge	After Zard-Band	After Roodak	After Haji-Abad	After Oshsan	Do-Ab	After - Zayegan	Mixture from Golabdar and Abnik	Parameter
10	10	10	9	8	7	5	4	Temperature
400	407	402	402	399	394	302	300	Electric Conductance
9.2	8.9	8.6	8.6	8.4	8.3	8.8	9	Dissolved Oxygen
1.5	1	1	1.4	2	1.3	2.1	2.2	Biologically required oxygen
4	4.2	4.7	7	7	6.9	3.4	3.6	Nitrate
0.04	0.02	0.02	0.05	0.06	0.02	0.03	0.03	Inorganic phosphate
147	153	149	159	153	155	127	129	Total alkalinity
8.47	8.43	8.4	8.38	8.27	8.25	8.25	8.3	PH
0	0.08	0.08	0	0	0	0	0	Ammonia

Table 4: Quality parameters of Jajrood River taken from sampling stations for Bahman

After Lashkarak Bridge	After Zard-Band	After Roodak	After Haji-Abad	After Oshsan	Do-Ab	After - Zayegan	Mixture from Golabdar and Abnik	Parameter

1	1	1	3	2	3	3	3	Temperature
412	412	416	455	452	453	326	326	Electric Conductance
10.2	10.2	9.9	9.9	10.1	9.4	9.6	9.3	Dissolved Oxygen
1	1	1.2	1.5	2.2	1	1.6	1	Biologically required oxygen
6.5	6.7	6.8	10.4	10.2	9.8	8.7	6.8	Nitrate
0.03	0.05	0.04	0.03	0.02	0.01	0.03	0.02	Inorganic phosphate
155	159	159	153	153	151	139	143	Total alkalinity
8.41	8.38	8.32	8.62	8.42	8.51	8.49	8.36	PH
0	0	0	0.19	0.21	0.09	0.11	0.1	Ammonia

Results and discussion

According to the data taken from the Jajrood River during 4 out of 8 months, the parameters BOD, DO, nitrate and nitrite were simulated by ANN. Six parameters of temperature, electric conductance, nitrate, inorganic phosphate, total alkalinity, pH, and ammonia were taken as independent variable and the criteria BOD and DO were taken as dependent variables. In order to simulate the amount of nitrate, 5 parameters of temperature, electric conductance, inorganic phosphate, total alkalinity, pH, and ammonia were used as dependent variables. 70% of data were used for training and 30% for testing the network. A Free-forward backprop type of network was used as default to make possible the accessibility and manipulation of the network for neurons and layers. Training function of TRINLM and learning function of LEARNGDM were chosen by trial and error based on the network proposal and the type of data used. Error type MSE and two layers were used for training. According to the software recommendation, 10 neurons were used on every layer and the transfer function TANSIG was selected by trial and error. 1000 iterations and the least errors were chosen to implement the network. The results of simulations were obtained as follows. The benchmark R2 and error RMSE were used for verification. The quality parameters measured at 8 stations of the Jajrood River was evaluated as recorded in the table.

Results show that River's water level increases from the spring to the end of the year and then, it decreases. It should be noted from the data that a significant increase has been made in the dissolved oxygen in Bahman compared to other months. Also, in Mordad, the biologically required dissolved energy has risen to maximum. Also, inorganic phosphate is higher in Ordibehesht than all other months in all stations. Total alkalinity is lower in Ordibehesht than other months. However it has not changed in the next three months. The ammonia in the River is orders of magnitude higher in Ordibehesht than other months.

1. Simulation of the DO

The results of network training are as follows. In Fig.1 the results are simulated and plotted for DO parameter variations by 6 training sessions. In this figure, obtaining the best solution with least error is considered. Since the error criterion for the network designing phase was set on MSE, the software has obtained the best solution with this error.

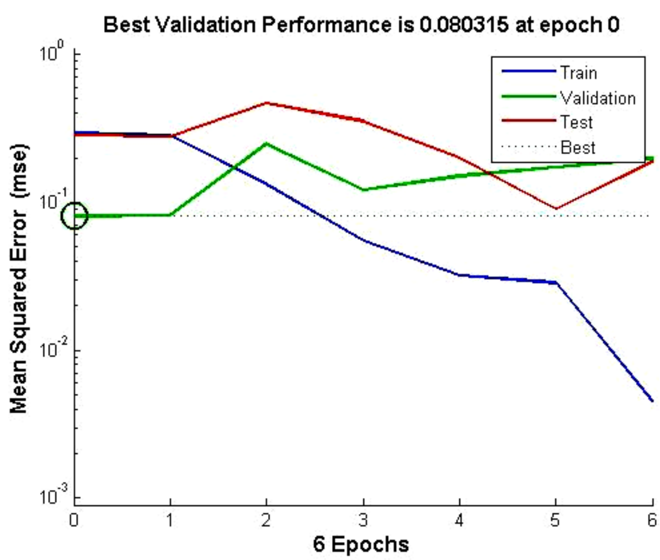


Figure 1: DO parameter variations in simulation by ANN

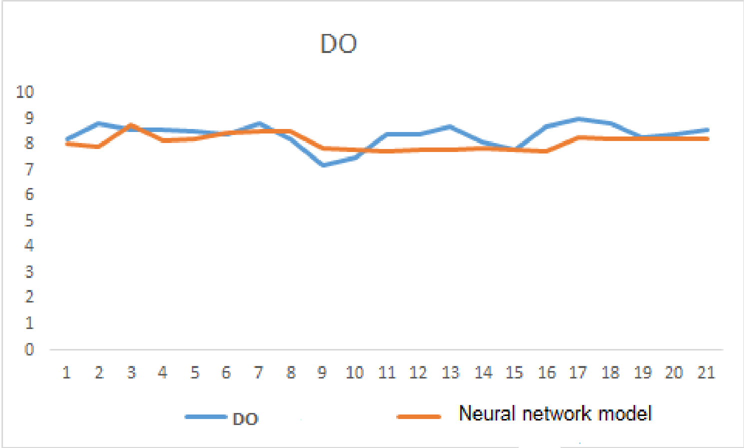


Figure 2: DO parameter variations in the training phase

From Fig. 2 it may be inferred that the results for DO parameter in the training session have reasonable accuracy. The best result has been obtained in this type of network for every iteration. So, it may be observed that using 2 layers with 10 neurons may give an appropriate result.

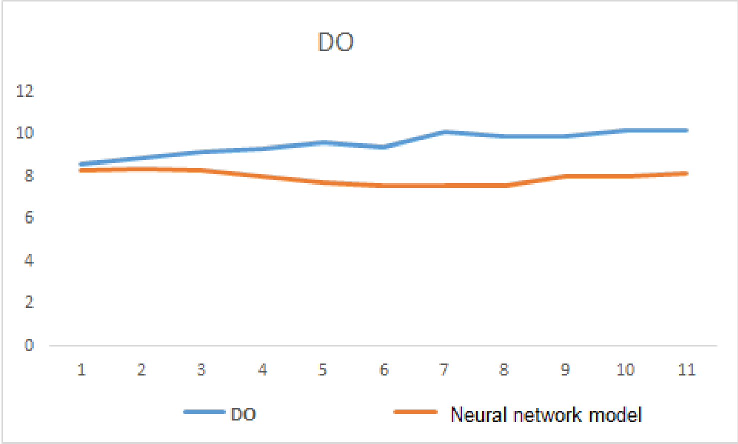


Figure 3: DO parameter variations in the network test phase

The results show that the test phase has a lesser accuracy than the training phase. A major cause for this is small size of dataset. Since this study has been performed in only 4 months of the year, we could only base our results on these 4 sets of data from 8 stations. The correlation between observatory and calculative data for DO parameter is as follows, which is not appropriate.

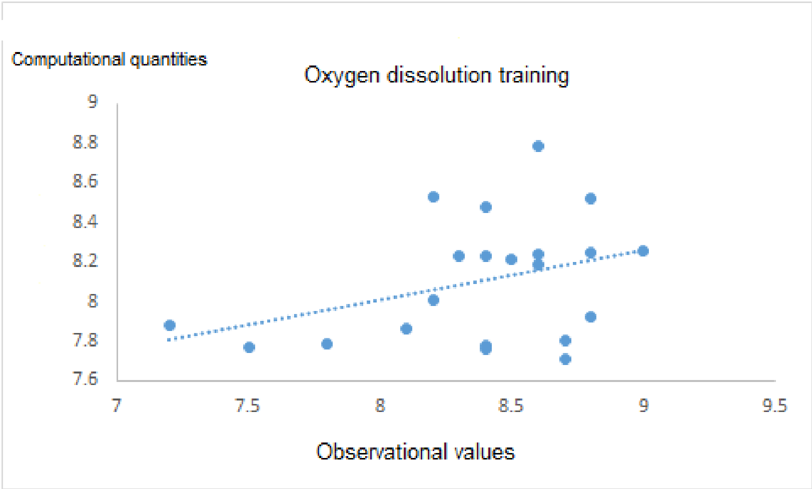


Figure 4: The correlation between observatory and calculative data for DO parameter in the network training phase

Also, for the test phase, we have:

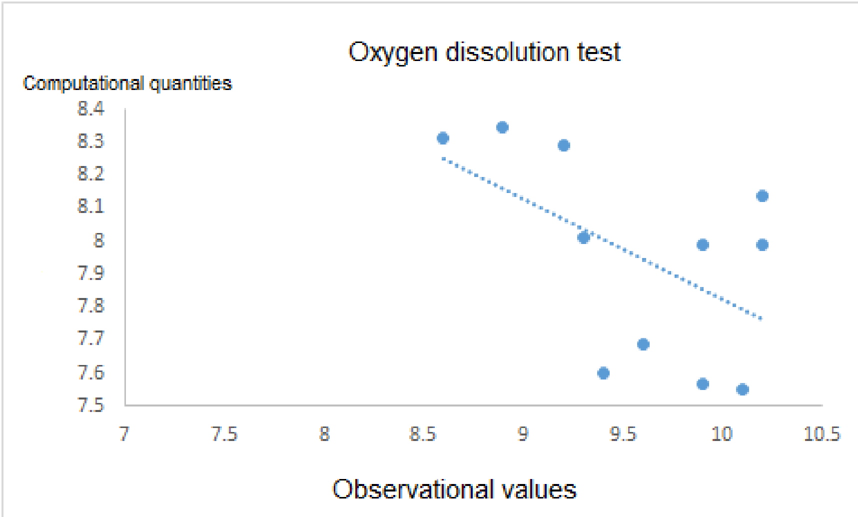


Figure 5: The correlation between observatory and calculative data for DO parameter in the network test phase

2. BOD simulation

Results from the network are as follows. In Fig. 1, variations of the parameter BOD with 18 training periods have been simulated and plotted by ANN. In this figure, we try to get the results with least amount of error. Since the error criterion for the network designing phase was set on MSE, the software has obtained the best solution with this error.

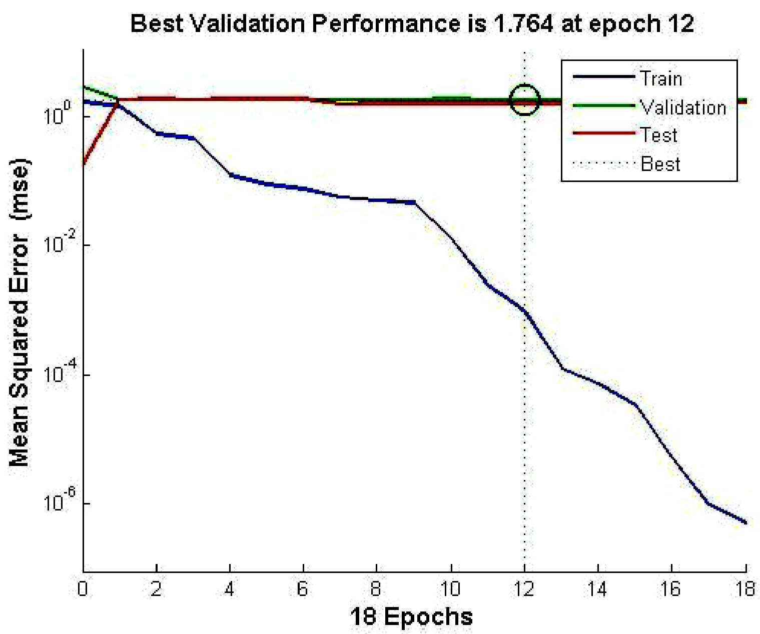


Figure 6: DO parameter variations in simulation by ANN

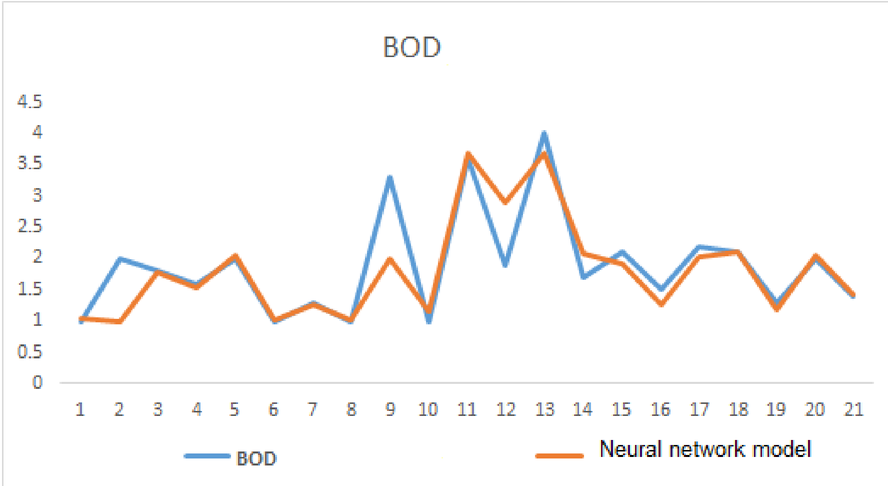


Figure 7: BOD parameter variations in the network training phase

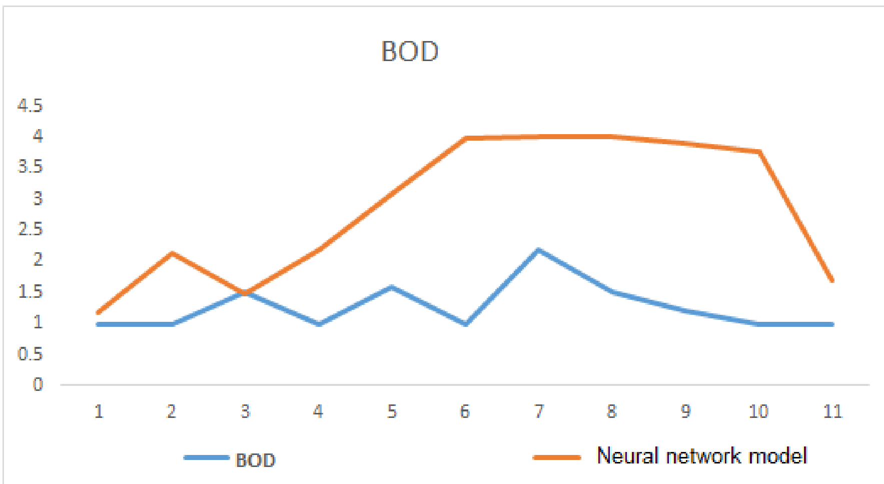


Figure 8: BOD parameter variations in the network test phase

In the training phase, the estimation of the neural network is highly consistent with the observatory data, but it is not acceptable in the test phase. However, in the analysis of the issue, the quantities of R^2 and error should be considered too. The correlation for the parameter BOD in test and training phases is as follows:

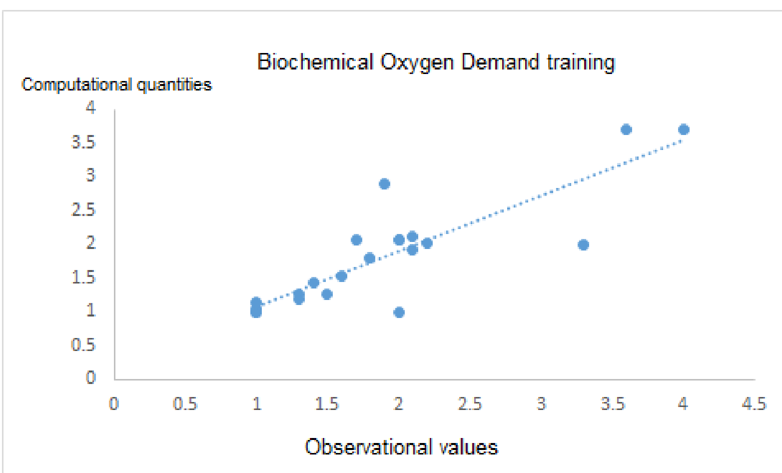


Figure 9: Correlation between observatory and calculative data for the parameter BOD in network training phase

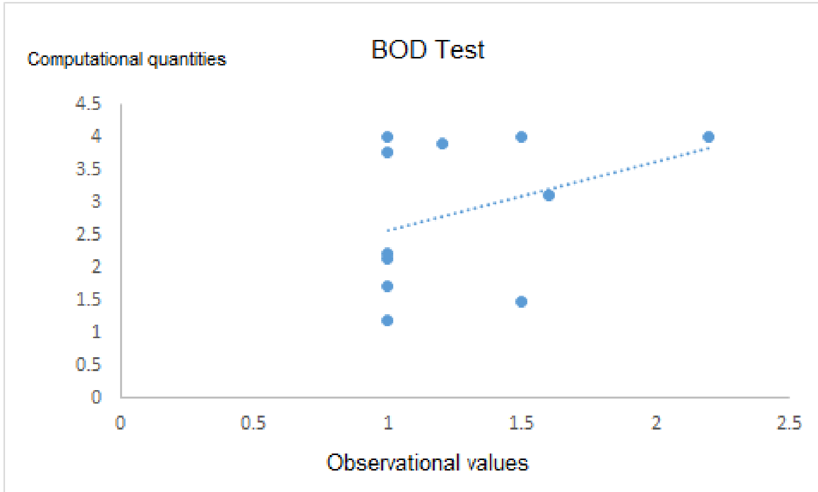


Figure 10: Correlation between observatory and calculative data for the parameter BOD in network test phase

In the training phase of the network, an appropriate correlation is observed based on the number of data for the BOD parameter. This is due to high capability of the chosen network for the number of layers and neurons. However, this is not the case for the test phase, in which the correlation is reduced and this is due to insufficient data.

3. Nitrate simulation

As discussed above, for nitrate simulation, 5 parameters are used which affect the amount of nitrate in the River. The neural network output for the parameter is as follows. In Fig.11 it is clear that from among 1000 training sessions for the network, the optimum result was obtained for the very first session.

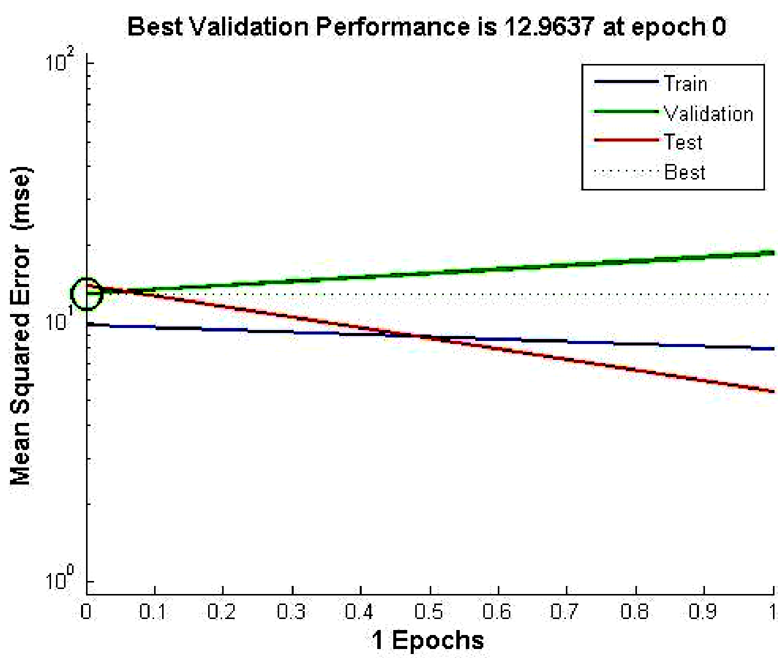


Figure 11: Nitrate parameter variations with simulation by the neural network

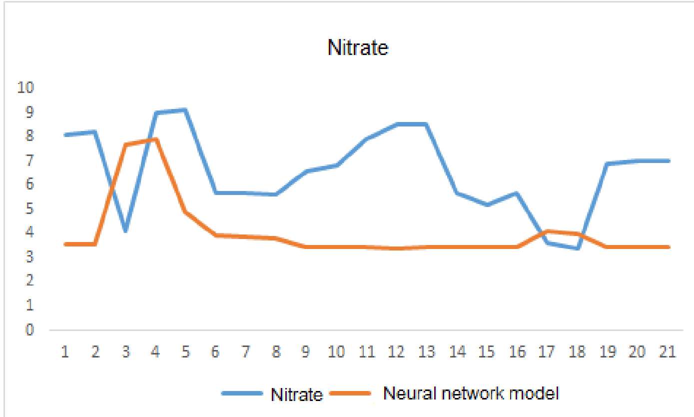


Figure 12: Nitrate parameter variations in the network training phase

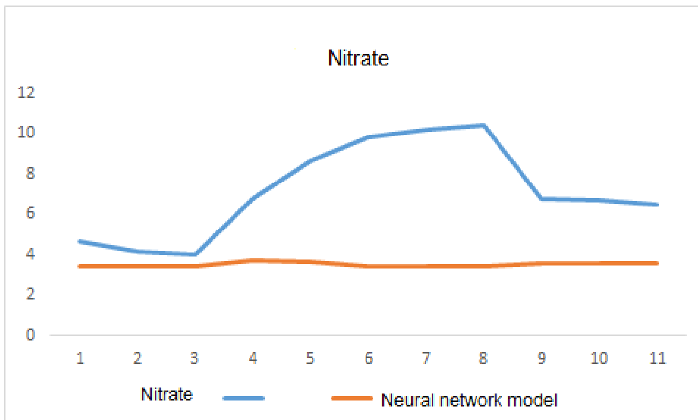


Figure 13: Nitrate parameter variations in the network test phase

In the beginning and ending training sessions, the simulation data are highly consistent with the observatory data. This is not the suitable in the intermediate sessions, however. It is only a result of small data set. This is more undesirable for network test. For the correlation graphs we have:

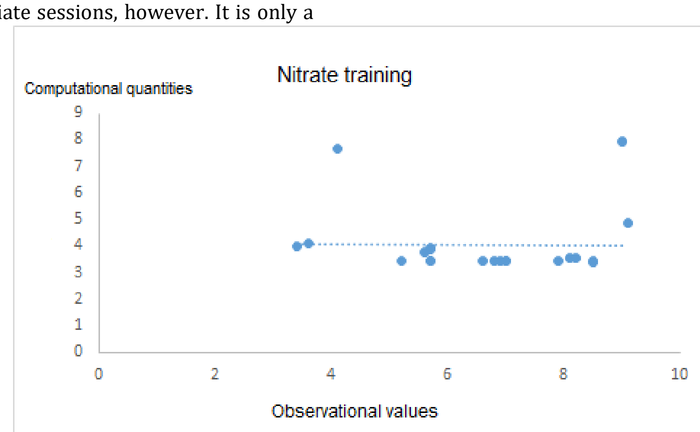


Figure 14: Correlation of observatory and calculative data for the nitrate parameter in the network training phase

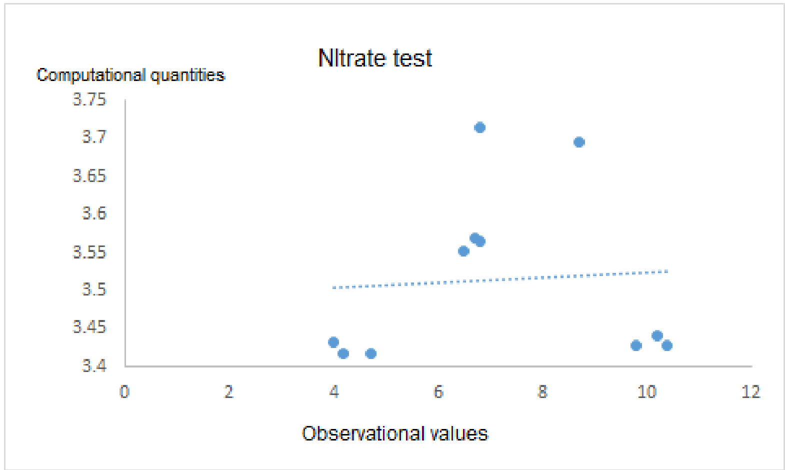


Figure 15: Correlation of observatory and calculative data for the nitrate parameter in the network test phase

The correlation values for nitrate show undesirable network estimation in regards with this parameter.

4. Nitrite simulation

Five parameters are used for nitrite simulation. The output of the neural network for this parameter is as follows. In Fig.16 it is clearly observed that from among 1000 training sessions, the optimum result was obtained by 9 sessions. The error criterion is MSE, which is also the solution criterion. Nitrite variations in Jajrood River are plotted in Fig.17 for the training phase and in Fig. 18 for the test phase.

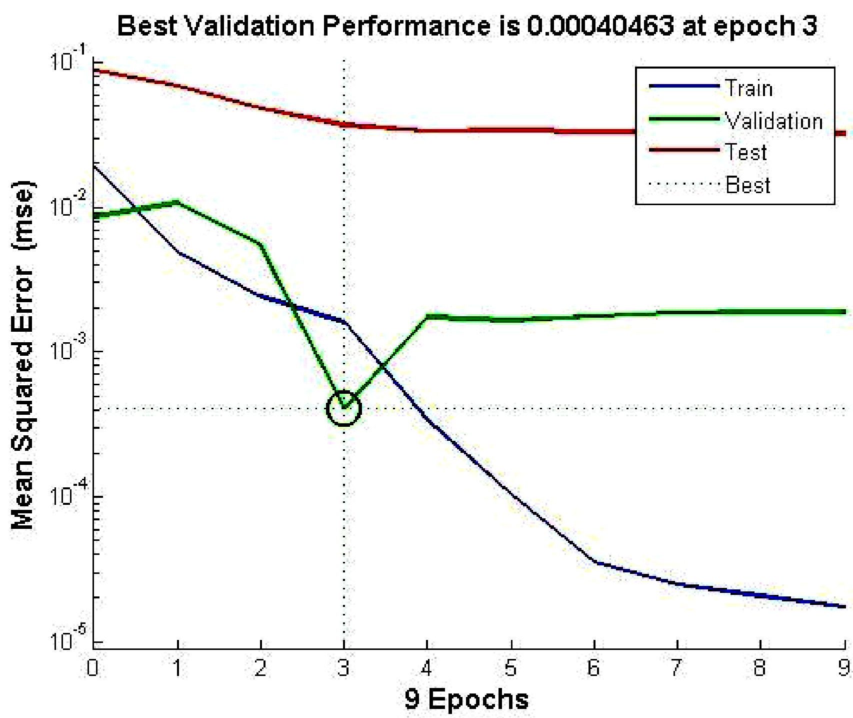


Figure 16: Nitrite parameter variations in the simulation by ANN

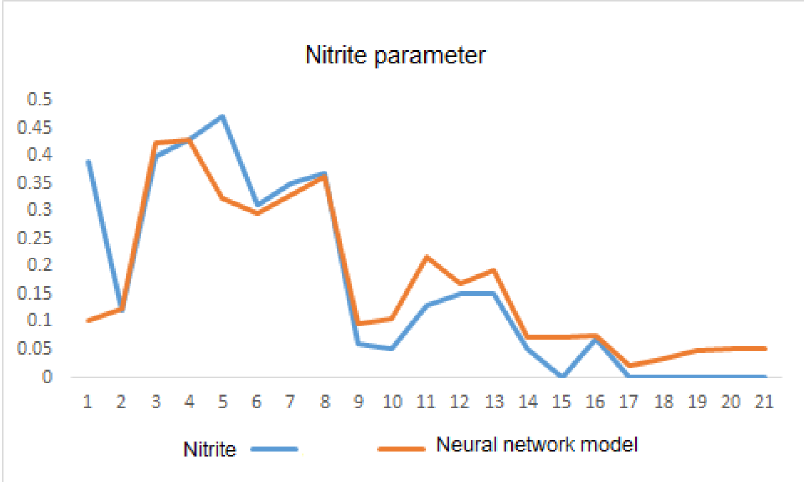


Figure 17: Nitrite parameter variations in the training phase

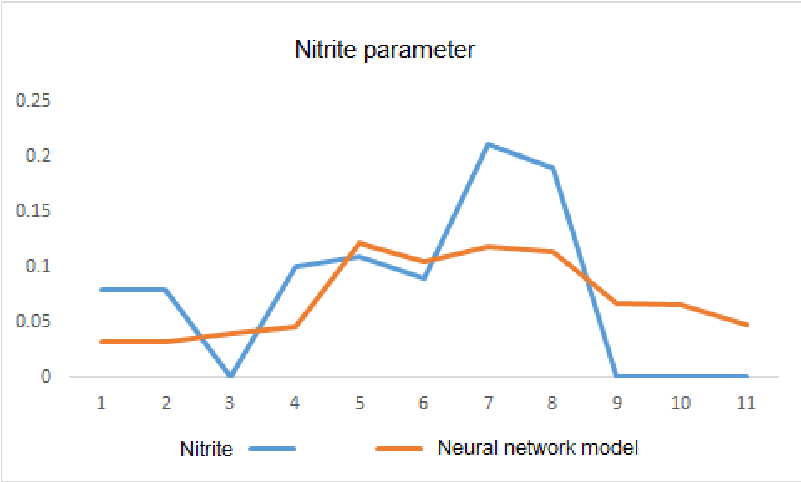


Figure 18: Nitrite parameter variations in the test phase

Nitrite variations and its estimation by the neural network show that in both training and testing sessions, the network resulted in a good estimation. Having more layers and more neurons per layer could give appropriate estimations despite insufficient data. In the test phase, too, simulation data correlation with the calculative data is good. Correlation graphs are as follows:

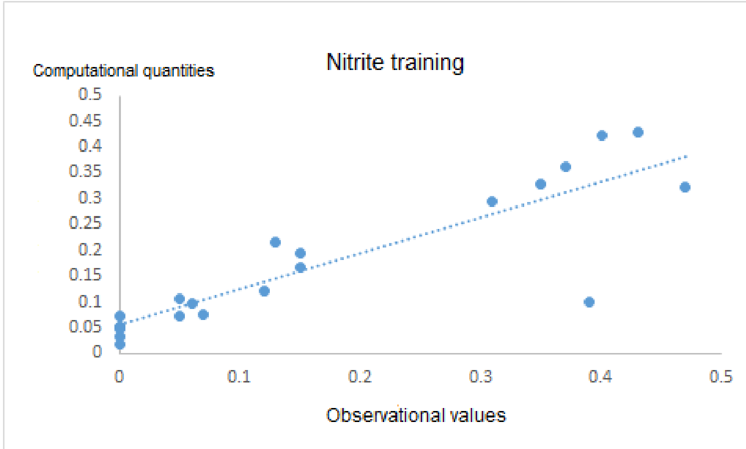


Figure 19: Correlation between observatory ad calculative data for nitrite parameter in the training phase

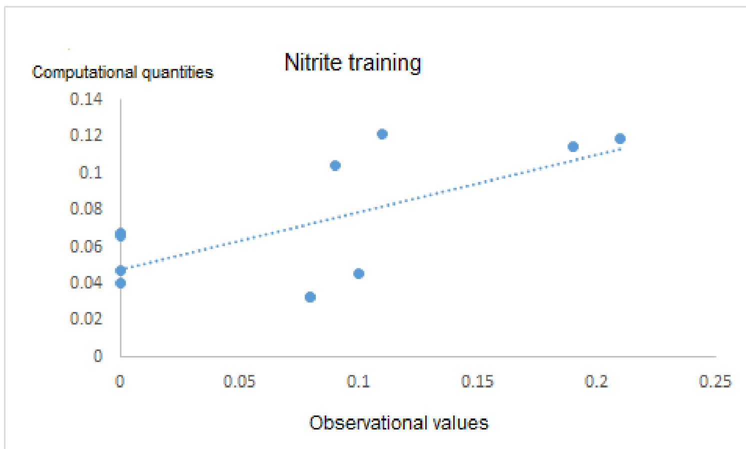


Figure 20: Correlation between observatory ad calculative data for nitrite parameter in the test phase

According to the correlation graphs, it is clearly observed that nitrite observatory and calculative data correlations are good in both testing and training sessions. For a better analysis, calculations of R2 value and the RMSE error are performed for all parameters in both training and testing sessions, reported as follows:

Table 5: Verification criteria for the neural network in different parameters

Nitrite		Nitrate		DO		BOD		Verification parameters and criteria
test	Training	test	Training	test	Training	test	Training	
0/41	0/77	0/0058	0/0002	0/28	0/12	0/13	0/72	R2
0/05	0/08	4/27	3/29	1/28	0/51	1/88	0/44	RMSE

According to Table 5, it is revealed that the ANN has its best results for the nitrite parameter with a correlation value of 0.77 for training and 0.41 for test phases and error values of 0.08 and 0.05, respectively. One reason behind this is having parameters like nitrate in the simulation. Addition or elimination of any kind of input for the neural network has a major effect on the final solution.

After nitrite, the ANN simulated biological oxygen with the R2 of 0.72 and 0.13, and an error of 0.44 for training and 1.88 for testing phases, respectively.

ANN is not acceptable for nitrate and DO. The main reason is insufficient data for the training session. However, the DOB and nitrite parameters were simulated with suitable accuracy. Using two layers consisting of 10 neurons for data processing on the one hand, and the nature of input data for the simulation on the other were the main causes of accuracy in these two parameters. A comparison with the literature shows the significance of data set size. For example, Misaghi & Mohammadi (1383) evaluated the quality of Zayande Rood water quality by ANN. Hey used the quality parameters of DO and BOD for the river between 1366 and 1380. Their results have a good correlation in terms of evaluation of observatory and calculative data. This is mainly caused by suitable data length for the implementation of neural network, while in the present study, the period of the study only consisted of 4 months.

In the analysis of DO and BOD by ANN in the Morad Beik Valley of Hamedanby Oliai et. al (1389), the results show acceptable applicability of multilayer perceptron neural network as a better technique for simulating the DO and BOD. Data from 6 stations during 12 months of a year were simulated using 10 parameters. Studies made by Konvar et. al (2009) showed that the input parameters and their combinations are capable of estimating these criteria.

Also, Noshady et. al (1386) simulated some of the quality parameters for Zayande Rood by ANN. They used the statistics data for 14 years from 9 stations. They used five parameters of chloride, bicarbonate, acidity, dissolved solid particles, and electric conductance. Their results had a high correlation and low error for simulation of every parameter.

Conclusions and future work

According to the results, the followings could be concluded. In this paper, a Free-forward backprop type of network with 2 layers and 10 neurons was used by trial and error which gave acceptable results for BOD and nitrite data. A major cause of inaccuracy in the parameter simulations has been insufficient size of data set for the training and testing phases. Considering the inappropriate network training because of small dataset and irrelevant inputs, the test (i.e. prediction) of the network will be erroneous.

Also, comparison with the literature shows that not only the time period (duration) is very significant for estimation and simulation of the neural network, but also, using input data with relevant nature is very effective on the output model. So, using any type of data may not improve the simulation results.

Recommendations:

According to the above, the following recommendations are made:

1. A longer time period should be used for estimation of the quality parameters
2. Simulation methods other than the ANN should be applied for the data, e.g. fuzzy-neural combinational system
3. Higher number of input parameters has to be applied. For example use may be made of debit, dissolved solid particles, bicarbonate, and chloride may be effective
4. Optimization techniques like genetic algorithm should be used to optimize the output data

REFERENCES

1) Ibrahimchi, Ahmed. (1994). River water quality mathematical models. Water and Development. The third number.
2) Eslamizadeh, Mohsen. (2006). Mathematical Modeling of Rivers in the TOC release. Master's Thesis, School of Environment and Energy, Science and Research Branch of Islamic Azad University.
3) Barati, Farnaz. (2013). Pollution Modeling of Rivers Tehran model QUAL2K. Master's Thesis, School of Environment and Energy, Science and Research Branch of Islamic Azad University.
4) Kalaghchian, Reza. (2008). Application of entropy in optimum design of monitoring systems rivers Case Study of Rivers. Applied research project research and basic studies Tehran Regional Water Company Research Committee.
5) Arbib, M.A., *The handbook of brain theory and neural networks*. 1995: The MIT press.
6) Poznanski, R.R., Biophysical neural networks: foundations of integrative neuroscience. 2001: Mary Ann Liebert.
7) Abrahart, R., P.E. Kneale, and L.M. See, *Neural networks for hydrological modeling*. 2004: CRC Press.