

# Forecasting Oxygen Demand in Treatment Plant Using Artificial Neural Networks

Moein Gheytaspour<sup>1\*</sup>, Omid Habibzadeh Bigdarvish<sup>2</sup>

<sup>1</sup>Department of Civil Engineering, College of Engineering, Islamic Azad University, Shahr-e-Qods Branch, Tehran, Iran

<sup>2</sup>Department of Civil Engineering, College of Engineering, University of Texas at Arlington, Texas, USA

**Abstract**— Modeling the wastewater treatment plant is difficult due to nonlinear properties of most of its different processes. Due to the increasing concerns over environmental effects of treatment plants considering the poor operation, fluctuations in process variables and problems of linear analyses, algorithms developed using artificial intelligence methods such as artificial neural networks have attracted a great deal of attention. In this research, first using regression analysis, the parameters of biological oxygen demand, chemical oxygen demand, and pH of the input wastewater were chosen as input parameter among other different parameters. Next, using error analysis, the best topology of neural networks was chosen for prediction. The results revealed that multilayer perception network with the sigmoid tangent training function, with one hidden layer in the input and output as well as 10 training nodes with regression coefficient of 0.92 is the best choice. The regression coefficients obtained from the predictions indicate that neural networked are well able to predict the performance of the wastewater treatment plant in Yazd.

**Keywords**—Yazd treatment plant, chemical oxygen demand, neural networks, sigmoid tangent.

## I. INTRODUCTION

To better and more efficiently control the performance of wastewater treatment plants, one can use a powerful mathematical tool, which is the basis for the recorded information associated with some essential parameters of wastewater during operational periods of the treatment plant. Traditional modeling which was used for biological processes and linear programming was based on writing equations for the rate of growth of microorganisms and formation of products (Zhao et al., 2008; Nahvi et al., 2018; Segre et al., 2002; Akram et al., 2013; Karlebach et al., 2008). However, as microbiological reactions are nonlinear and time-dependent, and have a complex nature. Such modelings had several limitations (Ginn et al., 1995; Xing & Pignatello, 1996; Jacques et al., 2018; Nguyen et al.,

2017; Daghighi, 2017). Use of methods capable of predicting the performance and efficiency of wastewater treatment plants especially based on quantitative and qualitative changes of the input wastewater is of great importance. Based on these predictions, the operator can adopt the necessary measures before incidence of problem, thereby applying suitable control and operation, preventing discharge of contaminated wastewater into the environment, returning the wastewater to the beginning of the treatment line (Eggen et al., 2014; Moles et al., 2003; Daghighi et al., 2017; Mahadevan et al., 2003). Artificial neural networks (ANNs) are among the methods for advanced statistical prediction, which have found extensive uses across all scientific fields. Neural network belongs to advanced statistical methods capable of predicting nonlinear and complex relations between inputs and outputs. Thus, to control the performance of wastewater treatment plants better and more efficiently, one can use a powerful mathematical tool, where the basis of the recorded information is related to some major parameters of wastewater during the treatment plant operational periods. Various research has been conducted about prediction with neural networks in different subjects. Haghiri et al (2018) modeled Wastewater Treatment Plant in Ardabil using ANN. In this study an optimal multi-objective design of neural networks was performed for modeling and optimizing the coagulator materials used in water treatment in Water Treatment Plant in Ardabil province, this research has been welcomed significantly by the Water Treatment Plant in Ardabil province. Han and Qiao (2012) employed ANN for predicting the efficiency of treatment plant. In this research work, ANN modeling was evaluated for predicting the efficiency of this treatment plant, which is equipped with active sludge system with diffusion aeration. Cao et al (2008) utilized ANN for predicting the changes in parameters of an anaerobic system. To optimize the weights of the ANN, they used multi-population genetic algorithm. Mjalli et al (2007) employed ANN models to predict the values of BOD, COD, and TSS parameters of the

wastewater of Doha treatment plant with active sludge system. They found that ANN enjoys a very high accuracy in predicting and estimating the operational parameters of wastewater treatment plants.

**Case Study**

The wastewater treatment plant in Yazd is located in the northwest of the town, 15 km ways from it, close to Yazd Wastewater Pond. It has an area of about 250 hectares. Based on these studies, the treatment plant has been designed in two stages. Its construction operations began in July 2012 and were exploited in the first half of 2017. The covered population in the year of origin (2012) was 24900 people, while in the goal year (2026), it will be 60000 citizens.

**II. METHODS AND MATERIALS**

An ANN is an idea which has been inspired by biological nervous system for information processing. ANN processes information like the brain. The key element in this idea is the new structure of information processing system. This

system consists of numerous highly interlinked neurons, which operate together for solving a problem. ANNs, like humans, learn through examples. An ANN is adjusted for performing a certain task such as pattern recognition and information grouping along a learning process. In biological systems, learning occurs with adjustments in synaptic links among the nerves. This method also applies to ANNs (Dawson & Wilby, 2001). Fig. 1 represents a general schema of ANNs. Based on the figure, every network consists of an input layer, one or several middle layer, and one output layer. Each layers is composed of some neurons. Every neuron receives the information from an input series, which act as dendrites in real neurons. After processing, that neuron delivers it to its output which indeed plays the role of synapses of a nervous cell. The output of this neuron is then used as the input for the next neuron. The number of neurons in the input and output layers is determined in proportion with the problem for which the network is used. Regarding the middle layer, the number of neurons is specified by the user through trial and error.

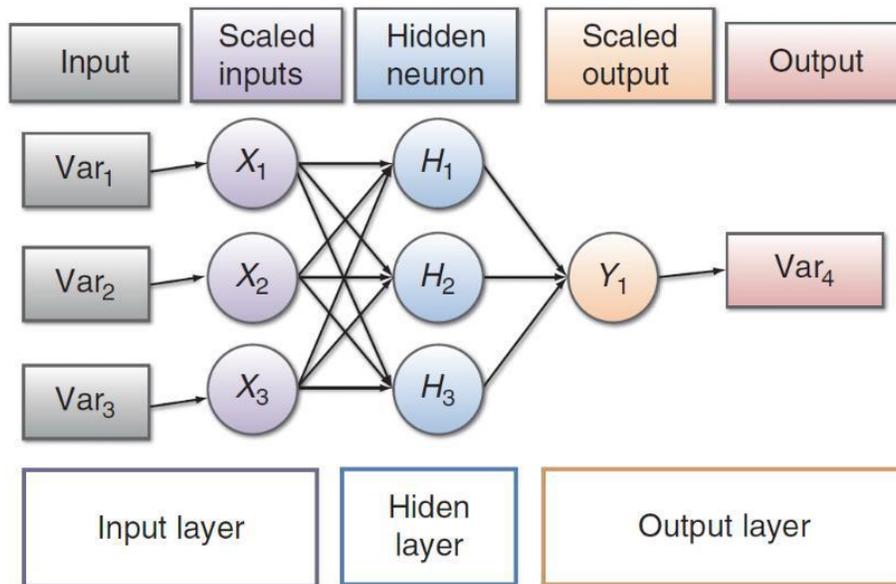


Fig.1: Example of a Simple ANN Network with Three Inputs, One Hidden Layer, Three Neurons, One Output Layer, and One Output Variable (Reinoso, 2017)

**III. RESULTS AND DISCUSSION**

**Input data**

The output data in this research is chemical oxygen demand (COD), while the input data include the treatment plant’s input flow rate ( $Q_{in}$ ), its output flow rate ( $Q_{out}$ ), the biological oxygen demand of the treatment plant’s input wastewater ( $BOD_{5in}$ ), the chemical oxygen demand of the treatment plant’s input wastewater ( $COD_{in}$ ), pH of the treatment plant’s input wastewater ( $pH_{in}$ ), the plant’s output

wastewater ( $pH_{out}$ ), and total suspended solids of the plant’s input wastewater ( $TSS_{in}$ ). In the stepwise multiple linear model, the most effective parameters are incorporated in the model. According to the regression coefficient, three equations were estimated for the COD of the output wastewater as follows:

Model 1, with regression coefficient 0.619,

$$COD_{out} = 11.091 + 0.071BOD_{in}$$

Model 2, with regression coefficient 0.668,

$$COD_{out} = 3.294 + 0.08BOD_{in} + 0.038COD_{in}$$

Model3, with regression coefficient 0.638,  $COD_{out} =$

$$93.163 + 0.069BOD_{in} + 0.031COD_{in} - 12.017PH_{in}$$

The significance and impact factor of each effective parameter in the above equations have been provided in the following table:

Table.1: the results obtained from multiple linear model for COD of the output wastewater

Model	Parameter	Impact factor	Significance level
1	BOD <sub>5in</sub>	0.071	0.001
2	BOD <sub>5in</sub>	0.08	0.003
	COD <sub>in</sub>	0.038	0.001
3	BOD <sub>5in</sub>	0.069	0.002
	COD <sub>in</sub>	0.031	0.014
	pH <sub>in</sub>	-12.017	0.039

According to Table 1 and in Model 3, the BOD of the input wastewater into the treatment plant has the maximum impact on the output. Next, using the three mentioned parameters in Model 3, prediction was performed.

**Error analysis**

To obtain the best topology, the criterion used is error and regression coefficient between observational and computational data. For this purpose, the Eq. 1 have been used.

$$r = \frac{\sum_i (x_i - \bar{x})(d_i - \bar{d})}{\sqrt{\sum_i (d_i - \bar{d})^2} \sqrt{\sum_i (x_i - \bar{x})^2}}$$

Eq. 1

where r is the regression coefficient between the observational data (x) and computational data (d), N

represents the number of data,  $\bar{d}$  and  $\bar{x}$  are the mean values of the computational and observational data, respectively. r value is always between -1 and 1. The closer its value to 1, the more complete the correlation is between the two variables in a direct fashion.

$$MSE = \frac{\sum_{j=0}^P \sum_{i=0}^N (d_{ij} - y_{ij})^2}{N.P}$$

Eq. 2

where MSE is the mean squared error, P represents the number of elements being processed in the output, N shows the number of data in the data series,  $y_{ij}$  indicates the network output for sample i in the elements being processed j, and  $d_{ij}$  is the output related to the prediction parameter for sample i in the element being processed j.

**Neural network design**

Network design means selecting the number of training nodes, number of hidden layers, and type of training function. This is because the experience of this research proved that these factors affect the prediction accuracy. In this research, for prediction out of 3, 10, 25, 50, 100, 500, and 1000 training nodes, 1, 2, and 3 hidden layers in the input and output as well as sigmoid function, sigmoid tangent, linear sigmoid tangent, and linear sigmoid functions have been used. The utilized network is multilayer perceptron and the training rule is momentum. A total of 84 models were specified for this research.

In order to select the best scenario, error analysis method was used. For this purpose, the network was trained by different scenarios and the prediction error was calculated. The number of hidden layers, number of nodes, and type of training function with the minimum error were considered as the best options. Based on the obtained results, the effect of these factors on the prediction accuracy is very considerable. The following diagrams reveal the regression coefficients between predicted as well as observational data and the number of different layers and nodes for the four functions utilized.

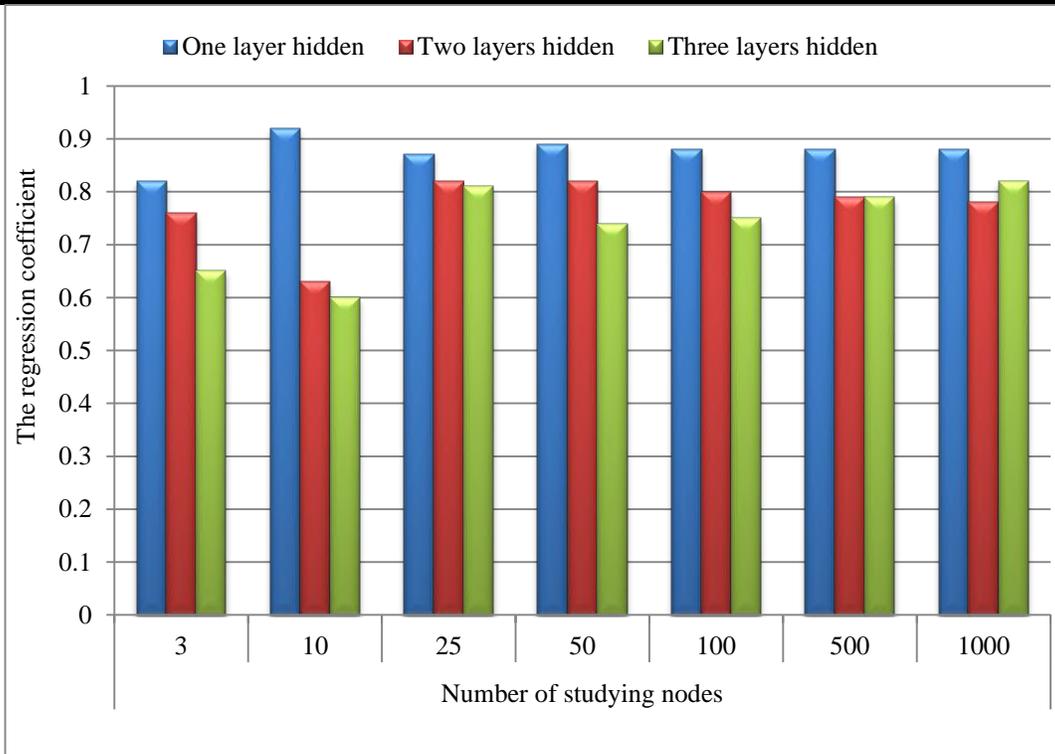


Fig.3: Comparing the prediction regression coefficient with different layers, sigmoid tangent function with different training nodes

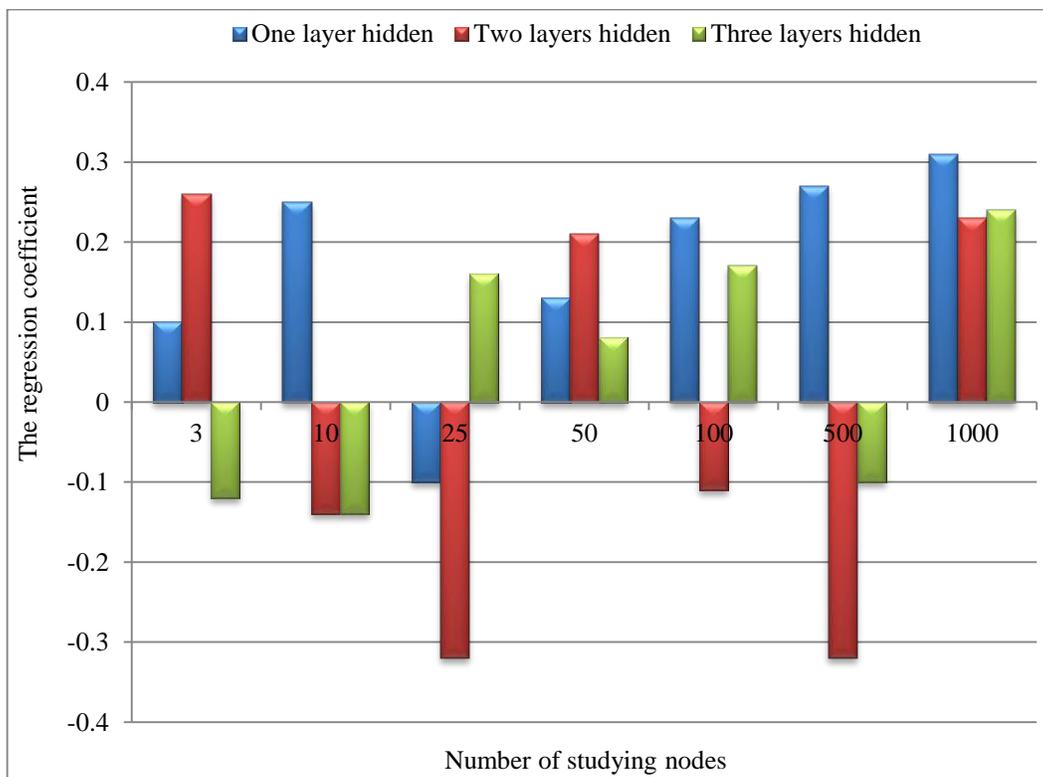


Fig. 4: Comparing the prediction regression coefficient with different layers, sigmoid function with different training nodes

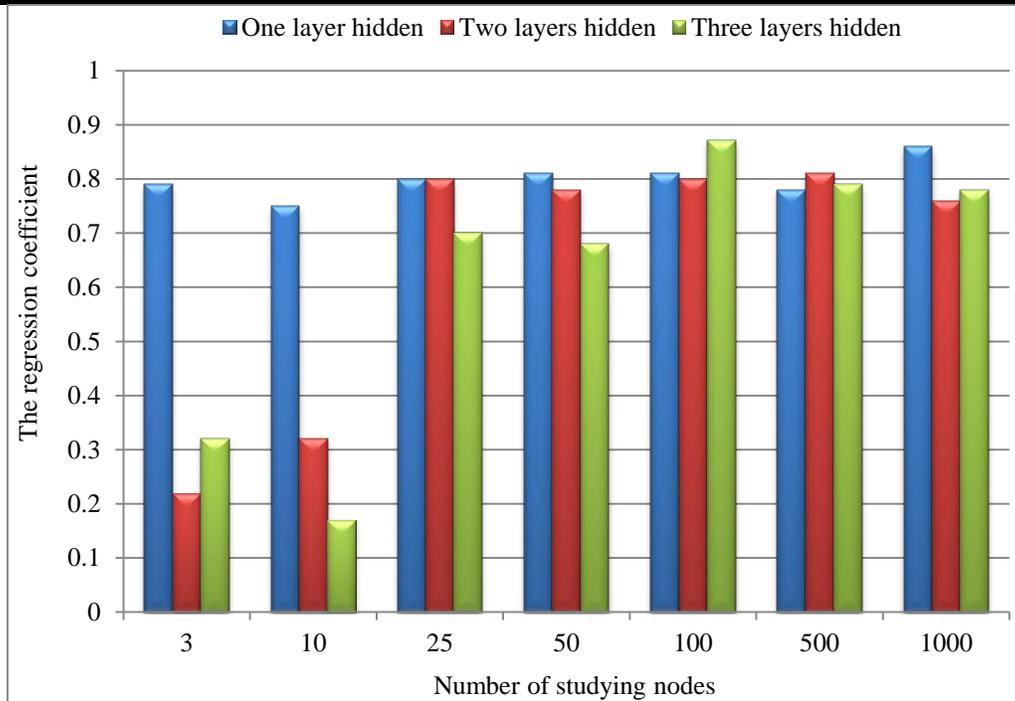


Fig. 5: Comparing the prediction regression coefficient with different layers, linear sigmoid tangent function with different training nodes

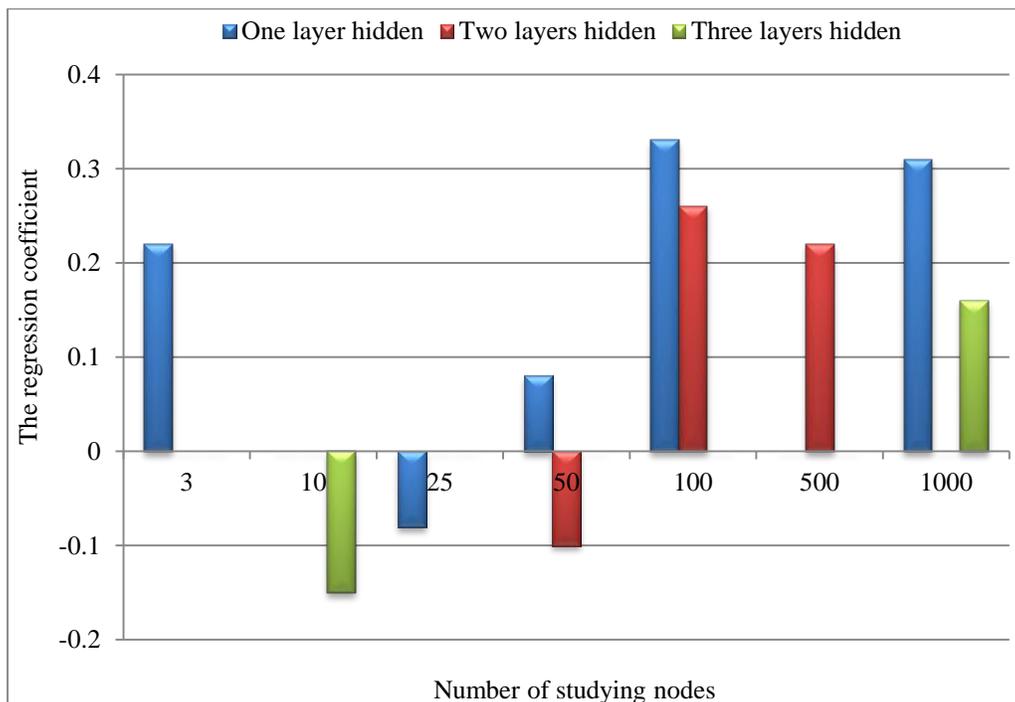


Fig.6: Comparing the prediction regression coefficient with different layers, linear sigmoid function with different training nodes

Based on the results of these diagrams, the extent of changes in error with the different number of hidden layers, number of nodes, and type of training function cannot be neglected. Comparing the results, it was found that 1 hidden

layer in the input, middle, and output along with 10 training nodes with sigmoid tangent function is the most suitable option for predicting COD parameter of Yazd treatment plant. Further, the above figures show that the sigmoid

tangent function has the minimum sensitivity to node change, while the sigmoid function is the most sensitive one. In addition, the sigmoid tangent function has the minimum sensitivity to changes in the number of layers, while the sigmoid function is most sensitive. Based on the

results, sigmoid tangent function is the best function, while the linear sigmoid is the worst function. The following figures demonstrate the training error analysis, validation, and test for the best neural network scenario.

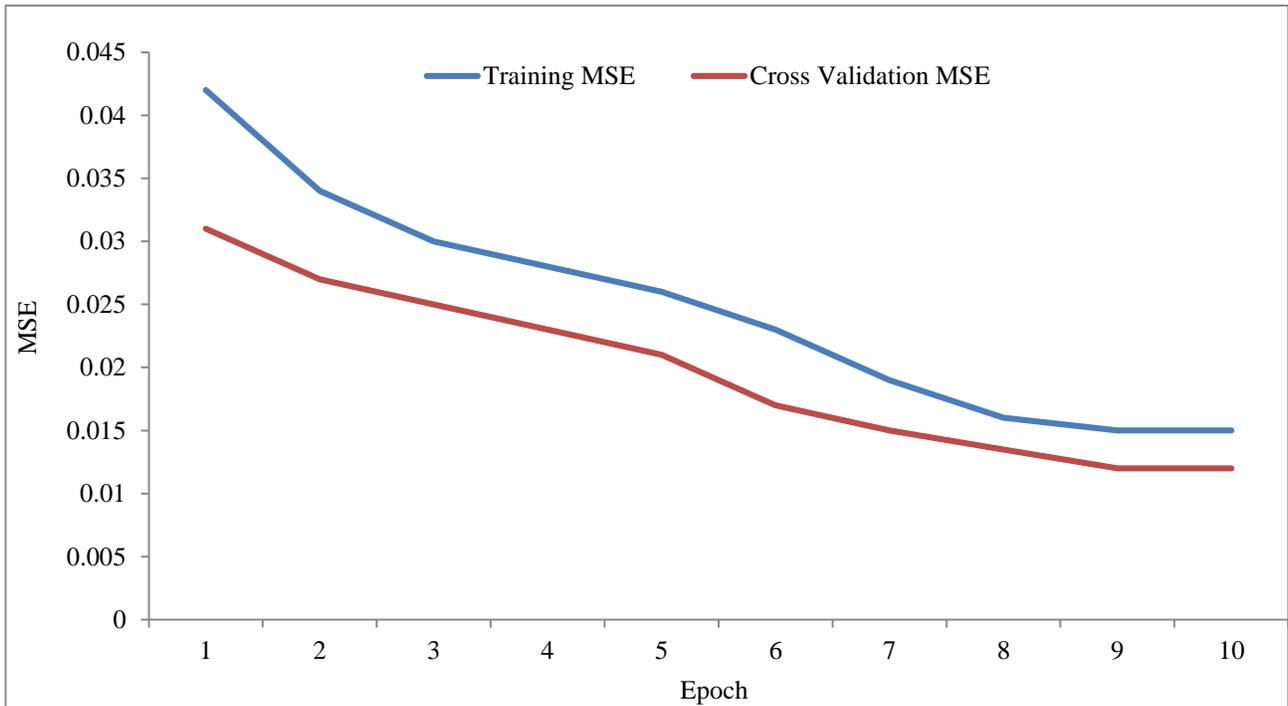


Fig.7: The training error and validation, sigmoid tangent function, one hidden layer and 10 training nodes

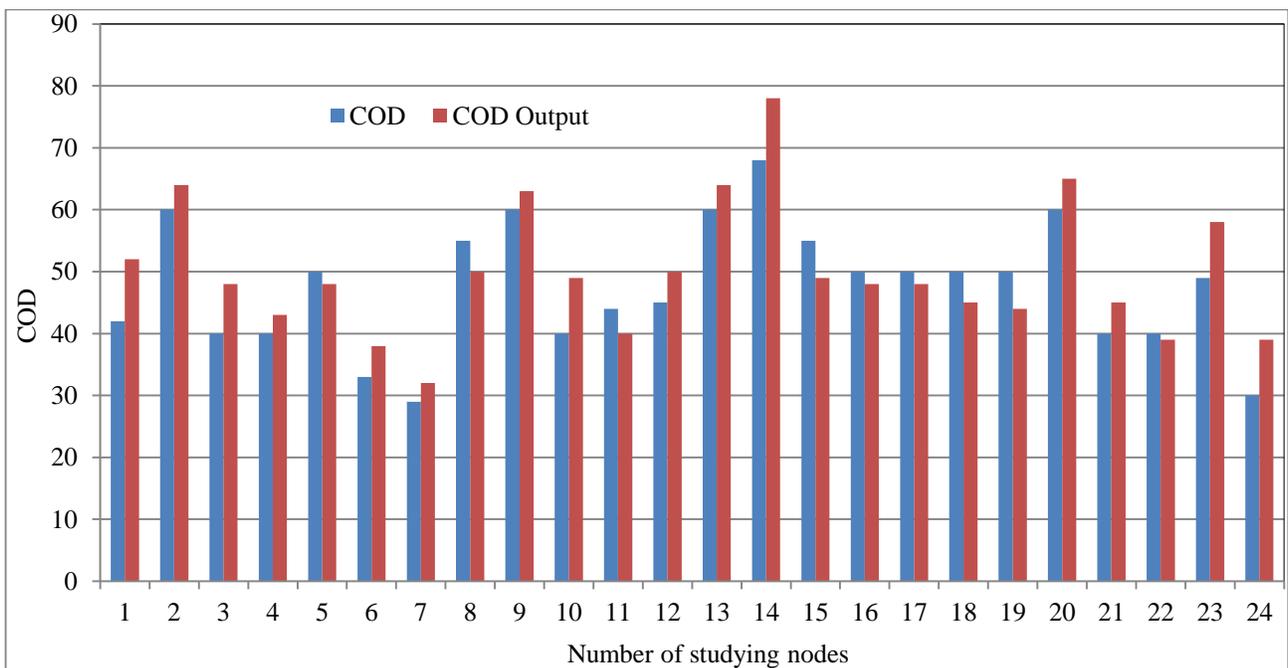


Fig.8: Comparing observational and prediction data, sigmoid tangent function, one hidden layer and 10 training nodes

Table.2: The predicted error analysis, sigmoid tangent function, one hidden layer and 10 training nodes

Reference	Result
MSE	18.95384
NMSE	0.5194133
MAE	4.271379
Min Abs Error	0.813998
Max Abs Error	6.014608
r	0.920182

The regression coefficient 0.92 indicates that neural networks can predict treatment plant performance regarding COD with a high accuracy. However, different researchers should not oversee the effect of different parameters for prediction. If suitable elements and factors are chosen, neural networks can predict COD with a good power. Otherwise, the results will be very unsatisfactory. The results of the research indicate that ANNs are a suitable instrument for predicting the performance of wastewater treatment plants.

#### IV. CONCLUSION

Use of ANNs can be very valuable for predicting the output COD of treatment plant. As the quantitative and qualitative properties of the input wastewater to the treatment plant, temperature, and other influential parameters in designing wastewater treatment plants are different with those of other treatment plants, it is not possible to employ a neural network trained in a certain treatment plant for other similar plants. It can be stated that a neural network in a treatment plant is considered acceptable when it is trained by correct and acceptable data of the same treatment plant. Further, all factors affecting the networks including number of nodes, number of layers, type of training function, type of network, and type of training rule should have also been investigated. The results of this research indicated that if all factors influencing the neural networks are examined, these networks are able to predict the chemical oxygen demand of the output wastewater leaving wastewater treatment plant in Yazd with a regression coefficient over 92%.

#### REFERENCES

- [1] Akram, M., Azari, A., Nahvi, A., Bakhtiari, Z., & Safaee, H. D. (2013). Subsurface drainage in Khuzestan, Iran: environmentally revisited criteria. *Irrigation and Drainage*, 62(3), 306-314.
- [2] Cao, G., Li, M., & Mo, C. (2008, December). Prediction of the anaerobic systems based on neural network with multipopulation parallel genetic algorithm. In *Intelligent Information Technology Application*, 2008. IITA'08. *Second International Symposium on* (Vol. 2, pp. 947-951). IEEE.
- [3] Daghighi, A. (2017). Harmful Algae Bloom Prediction Model for Western Lake Erie Using Stepwise Multiple Regression and Genetic Programming, Electronic Thesis, available at: <https://etd.ohiolink.edu/> (last access: 22 February 2018).
- [4] Daghighi, A., Nahvi, A., & Kim, U. (2017). Optimal Cultivation Pattern to Increase Revenue and Reduce Water Use: Application of Linear Programming to Arjan Plain in Fars Province. *Agriculture*, 7(9), 73.
- [5] Dawson, C. W., & Wilby, R. L. (2001). Hydrological modelling using artificial neural networks. *Progress in physical Geography*, 25(1), 80-108.
- [6] Eggen, R. I., Hollender, J., Joss, A., Schärer, M., & Stamm, C. (2014). Reducing the discharge of micropollutants in the aquatic environment: the benefits of upgrading wastewater treatment plants. *Environmental science & technology*, 48(14), 7683-7689.
- [7] Ginn, T. R., Simmons, C. S., & Wood, B. D. (1995). Stochastic-Convective Transport with Nonlinear Reaction: Biodegradation With Microbial Growth. *Water Resources Research*, 31(11), 2689-2700.
- [8] Haghiri, S., Daghighi, A., & Moharramzadeh, S. (2018). Optimum coagulant forecasting by modeling jar test experiments using ANNs. *Drinking Water Engineering and Science*, 11(1), 1-8.
- [9] Han, H. G., & Qiao, J. F. (2012). Prediction of activated sludge bulking based on a self-organizing RBF neural network. *Journal of Process Control*, 22(6), 1103-1112.
- [10] Jacques, D., Šimůnek, J., Mallants, D., & van Genuchten, M. T. (2018). The HPx software for multicomponent reactive transport during variably-saturated flow: Recent developments and applications. *J. Hydrol. Hydromech*, 66(2), 211-226.
- [11] Karlebach, G., & Shamir, R. (2008). Modelling and analysis of gene regulatory networks. *Nature Reviews Molecular Cell Biology*, 9(10), 770.
- [12] Mahadevan, R., & Schilling, C. H. (2003). The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metabolic engineering*, 5(4), 264-276.
- [13] Mjalli, F. S., Al-Asheh, S., & Alfadala, H. E. (2007). Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance. *Journal of Environmental Management*, 83(3), 329-338.

- [14] Moles, C. G., Mendes, P., & Banga, J. R. (2003). Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome research*, 13(11), 2467-2474.
- [15] Nahvi, A., Daghighi, A., & Nazif, S. (2018). The environmental impact assessment of drainage systems: a case study of the Karun river sugarcane development project. *Archives of Agronomy and Soil Science*, 64(2), 185-195.
- [16] Nguyen, T. T., Yoon, W. J., Lee, N. Y., & Ju, H. (2017). Integration of a microfluidic polymerase chain reaction device and surface plasmon resonance fiber sensor into an inline all-in-one platform for pathogenic bacteria detection. *Sensors and Actuators B: Chemical*, 242, 1-8.
- [17] Reinoso, N. (2017). Forecasting Harmful Algal Blooms for Western Lake Erie using Data Driven Machine Learning Techniques. (Electronic Thesis or Dissertation). Retrieved from <https://etd.ohiolink.edu/>
- [18] Segre, D., Vitkup, D., & Church, G. M. (2002). Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences*, 99(23), 15112-15117.
- [19] Xing, B., & Pignatello, J. J. (1996). Time-dependent isotherm shape of organic compounds in soil organic matter: Implications for sorption mechanism. *Environmental toxicology and chemistry*, 15(8), 1282-1288.
- [20] Zhao, X. M., Wang, R. S., Chen, L., & Aihara, K. (2008). Uncovering signal transduction networks from high-throughput data by integer linear programming. *Nucleic acids research*, 36(9), e48-e48.