

Prediction the Performance of Full Scale Wastewater Treatment Plant with A-B Process Using Artificial Neural Network and Genetic Algorithm

Farzaneh Mohammadi¹, Zeynab Yavari², Farideh Mohammadi³, Somaye Rahimi⁴

¹Department of Environmental Health Engineering, School of Health, Isfahan University of Medical Sciences, Isfahan, Iran, ²Department of Environmental Health Engineering, School of Abarkouh Paramedical, Genetic and Environmental Adventures Research Center, Shahid Sadoughi University of Medical Sciences, Yazd, Iran, ³Department of Textile Engineering, Isfahan University of Technology, Isfahan, Iran, ⁴Department of Environmental Health, Firoozabad Branch, Islamic Azad University, Firoozabad, Iran

Abstract

Aim: A reliable model for wastewater treatment plants (WWTPs) is essential to provide a tool for predicting their performance and to form a basis for controlling the operation of the process also; this would minimize the operation costs. In recent years, computer-based methods have been applied to many areas of environmental issues. Artificial neural networks (ANNs) and genetic algorithm (GA) techniques could be applied for modeling WWTP processes, owing to their high accuracy, adequacy, and quite promising applications in engineering. **Materials and Methods:** This study applied multilayer feed forward back propagation neural network and GA to predict and optimize the performance of the second phase of the Isfahan North WWTP. Experimental results, which demonstrated the performance of the plant over 6 years were applied for modeling. **Results:** A three-layer neural network was developed as a predictive model and the network was trained with Levenberg–Marquardt algorithm. The chemical oxygen demand (COD), biochemical oxygen demand (BOD), total suspended solids (TSS), total kjeldahl nitrogen (TKN), and total phosphorus (TP) were introduced as the model input and output. Neural network performance was evaluated with correlation coefficient R and least mean square error. Proposed model demonstrated the high consistency of the results of modeling and experiments. GA achieved the best value of input parameters as 324.36, 457.37, 359.11, 60.09, and 14.15 mg/l for BOD, COD, TSS, TKN, TP, respectively. **Conclusion:** ANN and GA combination provides powerful analysis tool for modeling and optimization of nonlinear relationships between the parameters in WWTPs and could be used for proper design and operation of the WWTPs.

Keywords: Artificial neural network, genetic algorithm, modeling, wastewater

INTRODUCTION

In recent years, computer simulations are widely utilized in solving environmental problems. Proper operation and control of the process in a biological wastewater treatment plant (WWTP) is very difficult, because of variations in composition, strength and discharge of raw wastewater and the complex nature of the treatment processes.^[1] Features of wastewater in each area differ depending on people's lifestyle, such that the operation of WWTP in each region depends on the experience and knowledge of engineers of that community. Moreover, the lack of adequate information about the appropriate range of input variables, affects the output quality of treatment plant.^[2] In traditional modeling techniques that were employed for biological processes, mass and energy

balance equations with the equations of microbial growth, substrate utilization, and the formation of products were applied. However, in microbial reactions and environmental interactions, there are nonlinear and time variable relationships with complex nature. Using these methods to predict the operation variables in WWTPs are time-consuming and create barriers to optimal process control.^[3]

Address for correspondence: Dr. Zeynab Yavari,
School of Abarkouh Paramedical, Yazd, Iran.
E-mail: zyavari2412@gmail.com

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

How to cite this article: Mohammadi F, Yavari Z, Mohammadi F, Rahimi S. Prediction the performance of full scale wastewater treatment plant with A-B process using artificial neural network and genetic algorithm. *Int J Env Health Eng* 2022;11:18.

Received: 01-12-2020, **Revised:** 26-12-2020,
Accepted: 28-01-2021, **Published:** 13-12-2022

Access this article online

Quick Response Code:



Website:
www.ijehe.org

DOI:
10.4103/ijehe.ijehe_52_20

Artificial neural network (ANN) technique is a good alternative in compared with traditional models because this method has a high potential in the interpretation of nonlinear and complex relations between variables. Over the past decade, ANN models have been introduced in different fields of environmental engineering.^[4,5]

There are a number of key variables in wastewater treatment and by studying them; the performance of the process can be understood. These variables include chemical oxygen demand (COD), biochemical oxygen demand (BOD), total suspended solids (TSS) using performance modeling; it was found that the neural network is a powerful tool for predicting the performance of WWTPs for the removal of these variables.^[6]

In this study, multilayer ANN model was designed to predict the performance of Isfahan North WWTP for the removal of the main variables BOD, COD, TSS, total kjeldahl nitrogen (TKN) and total phosphorus (TP) in the A-B process which is a two-stage activated sludge process. The 5-year data were employed for modeling. The model was tested for the last year of operation. By utilizing the results of this modeling, the plant operators will be able to predict the quality of the effluent based on input wastewater characteristics. Moreover, the optimized values of the input parameters for the best performance of the WWTP were determined using a genetic algorithm (GA).

MATERIALS AND METHODS

Artificial neural network theory

ANN is an information processing system that is inspired of the brain's nerve cell function. The development of the neural network is to compute the output variables of a process with the aid of input variables and calculations within the network.^[7]

The neural network is trained to perform a specific task. For training, the output values predicted by the model are compared with the target vector, which are the experimental results. Correcting connections values (weights) between the elements is carried out by calculating the error value. This will continue as long as the good predictive ability of the network is achieved and the model outputs have high compliance with the targeted vector. These concepts are shown in Figure 1a.

Trained neural networks are utilized to perform complex functions in various areas, such as pattern recognition, identification, classification, speech, vision, control systems and to solve complex problems that are difficult for traditional computers and humans. There are different types of learning algorithms. One of the most common classes of training algorithms for feed forward neural networks is called back propagation (BP).^[5] The main component of neural networks is neuron or node as shown in Figure 1b.

Input and output variables are shown with a_1 to a_n and O_j , respectively. Each node has received many signals and neuron that are investigated to make an output signal. W_{ij} to W_{nj} are weights given to the inputs of each neuron. Weights are adaptive factors within the network that defines the intensity of the input signal to the neuron. Each input (a_1, a_2, \dots, a_n) is multiplied by

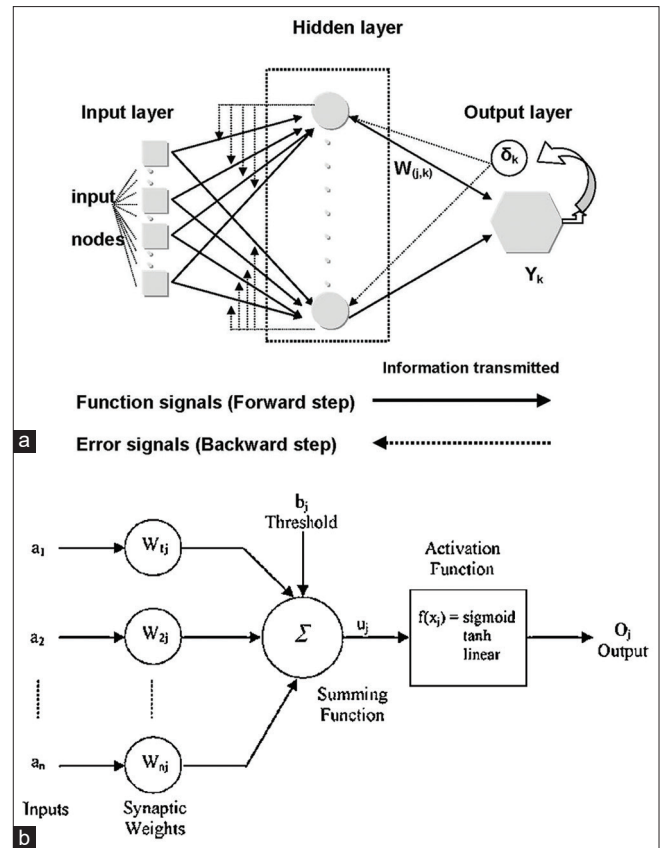


Figure 1: Neural networks training structure (a); the structure of a neural network neuron (b)

their respective weight (W_1, W_2, \dots, W_n) and all of these values are added together ($a_1 \times W_1 + a_2 \times W_2 + \dots + a_n \times W_n$); then the transfer function estimates output signal. Another input to node is b_j , which is called internal threshold or bias. This is a randomly selected value and will make outputs to be more accurate (Eq. 1).

$$u_j = \sum_{i=1}^n (W_{ij} * a_i) + b_j \quad (1)$$

The final output of nodes (O_j) is obtained with a mathematical operation on u_j . This is called transfer function. A transfer function change u_j into O_j by a linear or non-linear behavior. Three types of common transfer function are expressed in Eqs. (2-4).^[8]

- Sigmoid transfer function (Eq. 2)

$$f(x) = \frac{1}{1 + \exp^{-x}} \quad 0 \leq f(x) \leq 1 \quad (2)$$

- Hyperbolic tangent transfer function (Eq. 3)

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

- Linear transfer function (Eq. 4)

$$f(x) = x - \infty < f(x) < +\infty \quad (4)$$

In this study, Neural Network Toolbox of MATLAB R2013a software was used for modeling.

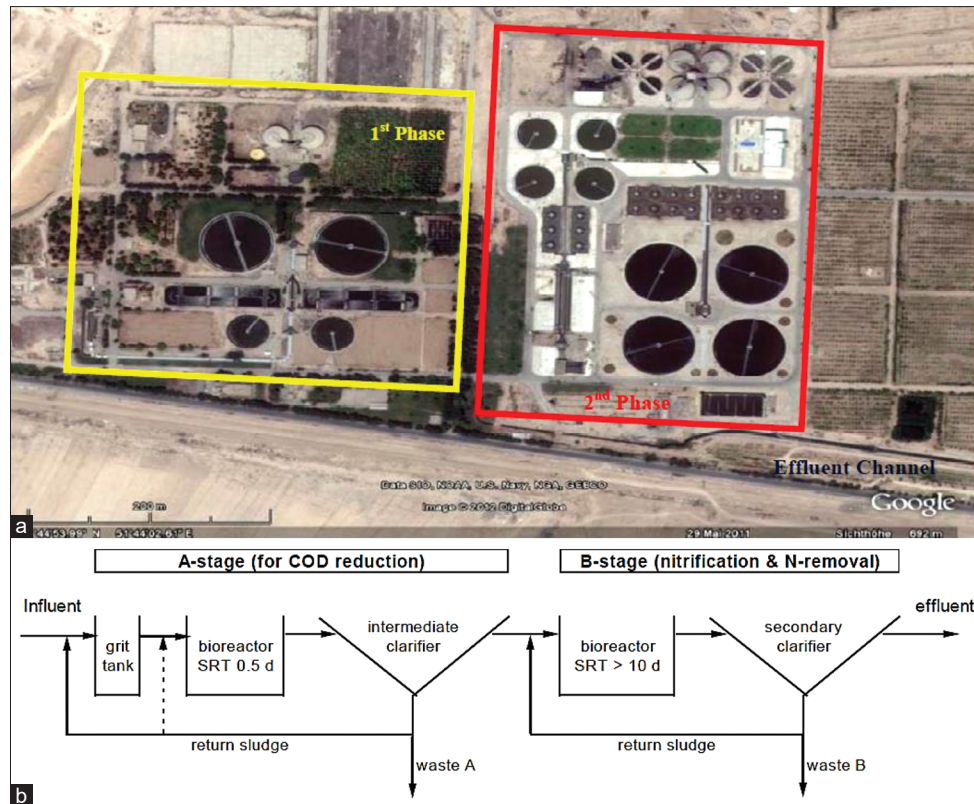


Figure 2: Aerial view of the 1st and 2nd phase of wastewater treatment plant Isfahan North (a), Flow diagram of the wastewater treatment plant second phase using A-B process (b)

Theory of genetic algorithm

GA is a heuristic search that mimics the process of natural selection. The development of this method is based on Darwin's theory of evolution. At first, GAs were developed by John Holland as a tool for finding the best solutions for complex issues. This procedure is implemented by creating a population of chromosomes. Thereafter, the people in the population are put at the control of the evolution process. Each evolutionary step is called a generation. Each member of the population is evaluated based on a series of predefined quality criteria. This is done by a fitness function. For the next generation, people are selected based on their competence. Therefore, the more qualified individuals are selected several times for the new population and will have more chances to reproduce, whereas less qualified people are faded, and have no chance to be selected and reproduce. Therefore, the more deserving chromosomes are increased until one of the stopping criteria is fulfilled. At this stage, the best solution is achieved.^[9,10]

In GA, a chromosome is often represented as a simple string and is a set of parameters that define a proposed solution to the problem the GA is trying to solve. Fitness function is a function that evaluates the quality of a chromosome as a response to the problem. GA component selection is done so that the chromosomes can be directed to evolution. The recombination is a process where the competent chromosomes called parent are selected and the next generation are produced from them. The major components of the recombination

are crossover and mutation. Crossover is performed via the exchange of parts of chromosome of two parents such that one or two new members called children are produced. Mutation acts on a single chromosome and with changes in chromosome components, the new members are produced. The stopping criteria are determined based on the number of produced generations or the desired competencies.^[11] In this study, the MATLAB R2013a Toolbox software was used to optimize the process.

Data collection

An ANN was employed for modeling in the second phase of Isfahan North WWTP. Isfahan North WWTP [Figure 2a] has two phases and its first phase has been exploited in 1987 with a nominal capacity of 400,000 people and nominal flow of 65,000 m³/day. The process of this phase is conventional activated sludge. The second phase of plant was started in 2008 with a nominal capacity of 800,000 people. The process of this phase is two-stage activated sludge (A-B) [Figure 2b], which has a pumping station unit, screening, grit chamber, first phase aeration and sedimentation (A), aeration and sedimentation of the second phase (B), and chlorination basin. Also in sludge treatment, three stages of thickening, digestion, and dewatering of sludge are considered. Measured values for the parameters of BOD, COD, TSS, TKN, and TP during 6 years (2012-2018) were collected at the treatment plant. The results of 5 years were used to build the optimal neural network model and the results of 1 year were used to simulate the model.

RESULTS

As mentioned before, the parameters of BOD, COD, TSS, TKN, and TP during 2012–2018 were taken from NWWTP operators. Given that for the modeling; the results of experiments conducted by operators were used, it is necessary to ensure normal distribution of data and outliers should be removed. Checking the normality of the data was done in SPSS software Version 26, (IBM Corp, Chicago, USA) and Kolmogorov–Smirnov test. The results are shown in Table 1 and Figure 3.

Results of 60 conducted trials over 5 years in the treatment plant were used to develop a neural network (ANN) model and the results of 12 trials in the past year were set aside for simulation at the end of the study. Features of network are shown in Table 2.

In FFBP network, the input and output matrix (5 × 60) with five variables; BOD, COD, TSS, TKN, and TP were considered. Seventy percent of the required data for network designing were allocated for network training and 15% for validation and 15% for network testing. Data selecting for the said cases was carried out randomly by software. Performance of neural network is assessed through the mean squared error (MSE) and correlation coefficient R , and their equations are shown in Eqs. (5, 6).

$$MSE = \frac{\sum_{i=1}^n (y_{obs,i} - y_{model,i})^2}{n} \quad (5)$$

$$R = \frac{\sum_{i=1}^n (y_{obs,i} - y_{obs,mean})(y_{model,i} - y_{model,mean})}{\sqrt{\sum_{i=1}^n (y_{obs,i} - y_{obs,mean})^2 (y_{model,i} - y_{model,mean})^2}} \quad (6)$$

In the equations above, n represents the number of data, $y_{model,i}$ and $y_{obs,i}$ are the output value predicted by the model and the measured output value, $y_{obs,mean}$ and $y_{model,mean}$ are average output values measured in the laboratory and average values predicted by the model.^[12]

BP Levenberg–Marquardt (LMBP) algorithm was used to train network. Figure 1a illustrates the stages of this algorithm. In general, in the function estimation problems with network parameters <100, the LMBP algorithm has high speed, performance and accuracy, because in many studies, this algorithm has reached the least error.^[13]

During the training, the output predicted by the model is

compared with the expected output and the MSE is calculated. If MSE is higher than the amount of error prescribed, it is redistributed from output to input and weights are modified accordingly. This is repeated until either the error reached the allowed limit or the maximum number of trial and error would be equal to the number of specified error. Weights modification in each trial is performed with the following equation (Eq. 7):

$$\Delta W(t) = -\eta MSE + \alpha \Delta W(t-1) \quad (7)$$

Where η is training rate and α is momentum coefficient that is both in the range of 0–1.

MSE is monitored during the training period. In the first phase of training, error will decline so that network reaches the minimum error and then by giving more data, the error is increased; in this stage, network’s training is stopped and the weights in the minimum error is returned.^[14]

The best neural network architecture is shown in Figure 4a. In addition, other results of MLPNN are observed in Figure 4b and c.

Designed network was applied for the simulation of data obtained in 2018, to evaluate the predictive ability of the ANN model for data that were not in the network designing in any way. The results of this simulation are shown in Figure 5.

Results of 72 experiments performed at plant in 5 years were introduced as an input to the GA. Each chromosome in the GA contains five components; BOD, COD, TSS, TKN, and TP. The aim is to determine the optimum amount of input to the plant so that all output parameters are minimal. The fitness function in

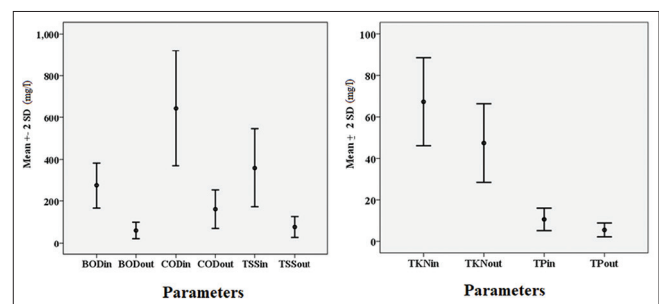


Figure 3: Error bar graph based on the error of standard deviation for input and output variables

Table 1: One-sample Kolmogorov-Smirnov test result

	One-sample Kolmogorov-Smirnov test									
	BOD _{in}	BOD _{out}	COD _{in}	COD _{out}	TSS _{in}	TSS _{out}	TKN _{in}	TKN _{out}	TP _{in}	TP _{out}
<i>n</i>	72	72	72	72	72	72	72	72	72	72
Mean (mg/l)	274.13	59.96	644.71	161.01	358.92	76.49	67.31	47.39	10.60	5.52
SD	54.11	19.52	137.42	45.69	93.26	24.77	10.61	9.48	2.70	1.67
Kolmogorov-Smirnov Z	0.876	0.807	0.727	0.707	0.817	0.630	0.641	1.210	1.109	1.192
Asymptotic significance (two-tailed)	0.426	0.533	0.666	0.700	0.517	0.823	0.806	0.107	0.171	0.117

SD: Standard deviation, BOD: Biochemical oxygen demand, COD: Chemical oxygen demand, TSS: Total suspended solids, TKN: Total kjeldahl nitrogen, TP: Total phosphorus

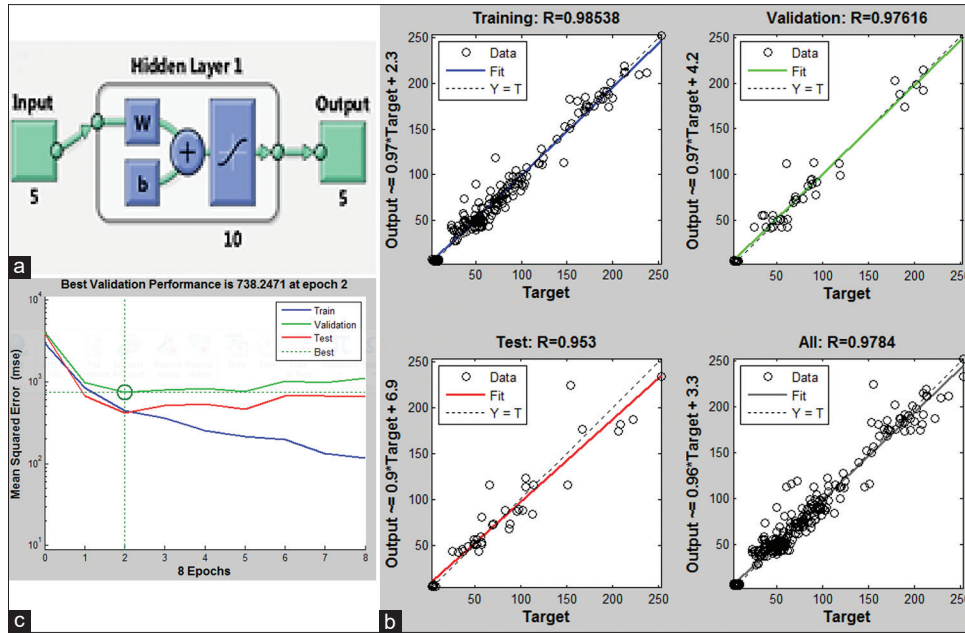


Figure 4: Designed neural network architecture (a), Regression graph (b), Network performance graph (c)

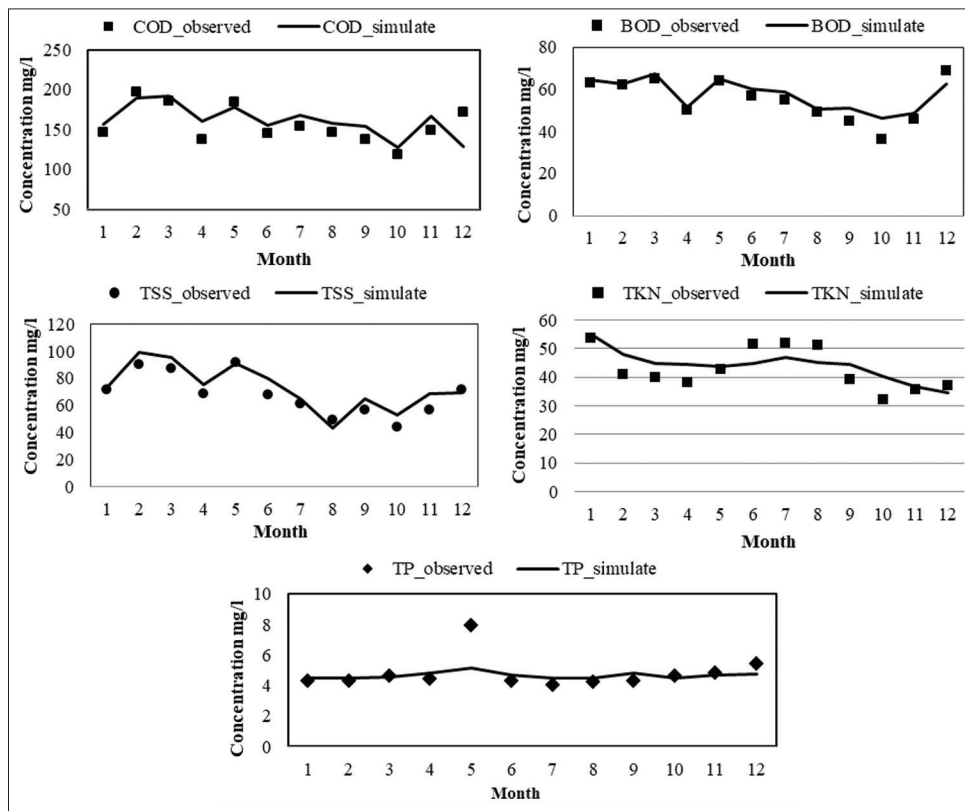


Figure 5: Compare the results of modeling and experiments in 2018

GA in the studied problem is neural network model developed to predict the values in the output of the plant. Other settings intended for GA are: Number of generation 100, the Rank scaling function, selection function of Stochastic uniform, the number of elite 2, Crossover fraction equal to 0.8, and the mutation function of constraint dependent and combination of

scattered function. Finally, GA was able to optimize the input parameters to minimize the output parameters. The results of GA are presented in Table 3.

In Figure 6a, the diagram of best and average of fitness values in each generation is presented. Figure 6b also represents the best fitness value in the present generation.

Table 2: Features of designated neural network

Parameters	Explanations and values
Network inputs	BOD _{in} , COD _{in} , TSS _{in} , TKN _{in} , TP _{in}
Network outputs	BOD _{out} , COD _{out} , TSS _{out} , TKN _{out} , TP _{out}
Network type	Multi-layer FFBP
Training function	Trainlm
Adaption learning function	Learnqdm
Performance function	MSE, R
Number of layers	3
Number of neurons in hidden layer	10
Transfer function	In hidden layer: tangent sigmoid In output layer: purelin

BOD: Biochemical oxygen demand, COD: Chemical oxygen demand, TSS: Total suspended solids, TKN: Total kjeldahl nitrogen, TP: Total phosphorus, FFBP: Feed-forward back-propagation, MSE: Mean squared error, R: Correlation coefficient

Table 3: Variables optimized by genetic algorithm

	BOD	COD	TSS	TKN	TP
Input (mg/l)	324.36	457.37	359.11	60.09	14.15
Output (mg/l)	19.52	70.11	27.52	41.50	4.75

BOD: Biochemical oxygen demand, COD: Chemical oxygen demand, TSS: Total suspended solids, TKN: Total kjeldahl nitrogen, TP: Total phosphorus

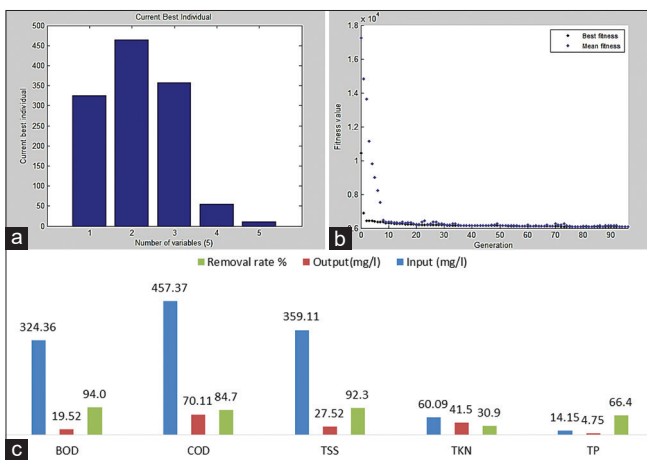


Figure 6: Output genetic algorithm graphs. (a) Fitness graph, (b) Best individuals, (c) Variables optimized by genetic algorithm

DISCUSSION

To determine the optimal network that has the lowest error and highest correlation coefficient, different networks are built with hidden layers and different neurons. As shown in Figure 4a, architecture of the best neural network that has the mentioned features with 3 layers. The number of neurons in hidden layer is equal to 10. As shown in Figure 4b, correlation coefficient in collections of training, validation, and testing are well consistent on target vectors with correlation coefficients >0.95. The correlation coefficient of the entire collection has been obtained as 0.98. Hence, the developed neural network has high ability to predict the performance of the plant.

In the graph of network performance [Figure 4c], MSE of the network starts from a great value and gradually decreases. This implies that the learning process of network has developed. The graph includes three lines because 60 input and target vectors have been divided randomly into three series of training (70%), validation (15%), and test (15%). Validation collection is employed in keeping the popularity of the network. Training process will continue until the network error is decreased on the validation network. Therefore, the least square error on validation graph has been identified. Hence, the over fit of the network on the training set is prevented.

Nasr *et al.*^[15] by developing a three-layer network and neuron number 10, 30, and 3, attained a correlation coefficient of 0.90 for the variables of BOD, COD, and TSS, respectively in WWTP of EL-AGAMY with SBR process. Mjalli *et al.*^[16] developed three separate three-layer networks with neuron number 20, 10, and 1 for variables of BOD, COD and TSS, respectively. The correlation coefficients of these parameters are 0.924, 0.924, and 0.839, respectively in treatment plant in Doha East.

After training of the network, it can be employed to simulate the new data. As earlier mentioned, treatment plant data for the years 2012–2017 were utilized to develop optimized neural network and experimental results of 2018 were considered for simulation. Therefore, the values of BOD, COD, TSS, TKN, and TP that were measured during 2018 were given to the optimized neural network as input. Predicted output value of these parameters in comparison with the measured output value in the WWTP is shown in Figure 5.

The methodology developed in this study is based on the combination of a GA and ANN for multiobjective optimization in existing WWTPs. According to the results, without having a large number of kinetic and stoichiometric coefficients that are needed in biochemical models such as ASMs, the WWTP performance could be predicted and optimized by having the main parameters of wastewater during some years and using a combination of multilayer neural network and GA. In this method, which wastewater treatment modeled using database methods, the characteristics of the input wastewater could even be optimized so that the treatment plant has the best performance. Apart from its applicability of biochemical models in the studied field, another constraint on the biochemical models to choose is that they should be able to solve multiobjective problems, which are complicated and time consuming. Adaptations of the ANN and GA algorithms were found to be the best solution to tackle such problems because of its working procedure, as well as the very little knowledge required about the problem to be solved.

CONCLUSION

The results of this study presented the appropriate ability of an ANN in predicting the performance of full-scale WWTP. The correlation coefficient was equal to 0.98 and introduces ANN model as reliable and accurate one. Conducted

simulation demonstrated the high consistency of the results of modeling and experiments. GA achieved the best value of input parameters; BOD, COD, TSS, TKN, TP, as 324.36, 457.37, 359.11, 60.09, and 14.15 mg/l, respectively. It seems ANN alongside GAs could be applied to upgrade operation of WWTP. The performance of the WWTP without having all kinetic and stoichiometric coefficients is also predictable and manageable using data based modeling.

Financial support and sponsorship

The authors wish to acknowledge to Vice Chancellery of Research of Isfahan University of Medical Sciences for the financial support, Research Project, #198151 and ethics code IR.MUI.RESEARCH.REC.1398.498. In addition, the authors are grateful to Isfahan North WWTP for their technical and logistic assistance during this work.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Falås P, Baillon-Dhumez A, Andersen HR, Ledin A, la Cour Jansen J. Suspended biofilm carrier and activated sludge removal of acidic pharmaceuticals. *Water Res* 2012;46:1167-75.
2. Bina B, Mohammadi F, Amin MM, Pourzamani HR, Yavari Z. Evaluation of the effects of AlkylPhenolic compounds on kinetic coefficients and biomass activity in MBBR by means of respirometric techniques. *Chin J Chem Eng* 2018;26:822-9.
3. Berardi C, Fibbi D, Coppini E, Renai L, Caprini C, Scordo CV, *et al.* Removal efficiency and mass balance of polycyclic aromatic hydrocarbons, phthalates, ethoxylated alkylphenols and alkylphenols in a mixed textile-domestic wastewater treatment plant. *Sci Total Environ* 2019;674:36-48.
4. Mohammadi N, Mirabedini SJ. Comparison of particle swarm optimization and backpropagation algorithms for training feedforward neural network. *J Math Comput Sci* 2018;12:113-23.
5. Tarentino AL, Maley F. A comparison of the substrate specificities of endo-beta-N-acetylglucosaminidases from *Streptomyces griseus* and *Diplococcus pneumoniae*. *Biochem Biophys Res Commun* 1975;67:455-62.
6. Niu G, Yi X, Chen C, Li X, Han D, Yan B, *et al.* A novel effluent quality predicting model based on genetic-deep belief network algorithm for cleaner production in a full-scale paper-making wastewater treatment. *J Clean Prod* 2020;265:121787. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S0959652620318345>.
7. Boeri C, Chiappa C, Galli F, De Berardinis V, Bardelli L, Carcano G, *et al.* Machine learning techniques in breast cancer prognosis prediction: A primary evaluation. *Cancer Med* 2020;9:3234-43.
8. Karimi H, Ghaedi M. Application of artificial neural network and genetic algorithm to modelling and optimization of removal of methylene blue using activated carbon. *Ind Eng Chem* 2014;20:2471-6.
9. Ding Y, Zhang W, Yu L, Lu K. The accuracy and efficiency of GA and PSO optimization schemes on estimating reaction kinetic parameters of biomass pyrolysis. *Energy* 2019;176:582-8.
10. Mohammadi F, Bina B, Amin MM, Pourzamani HR, Yavari Z, Shams MR. Evaluation of the effects of AlkylPhenolic compounds on kinetic parameters in a moving bed biofilm reactor. *Can J Chem Eng* 2018;96:1762-9.
11. Li Z, Zhang F, Luo X, Chang W, Cai Y, Zhong W, *et al.* Material removal mechanism of laser-assisted grinding of RB-SiC ceramics and process optimization. *Eur Ceram Soc* 2019;39:705-17.
12. Mohammadi F, Samaei MR, Azhdarpoor A, Teiri H, Badeenezhad A, Rostami S. Modelling and optimizing pyrene removal from the soil by phytoremediation using response surface methodology, artificial neural networks, and genetic algorithm. *Chemosphere* 2019;237:124486.
13. Mammadli S. Financial time series prediction using artificial neural network based on levenberg-marquardt algorithm. *Procedia Comput Sci* 2017;120:602-7.
14. Bahrami M, Akbari M, Bagherzadeh SA, Karimipour A, Afrand M, Goodarzi M. Develop 24 dissimilar ANNs by suitable architectures & training algorithms via sensitivity analysis to better statistical presentation: Measure MSEs between targets & ANN for Fe-CuO/Eg-Water nanofluid. *Phys A Stat Mech its Appl* 2019;519:159-68.
15. Nasr MS, Moustafa MA, Seif HA, El Kobrosy G. Application of artificial neural network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-EGYPT. *Alexandria Eng J* 2012;51:37-43.
16. Mjalli FS, Al-Asheh S, Alfadala HE. Use of artificial neural network black-box modelling for the prediction of wastewater treatment plants performance. *Environ Manage* 2007;83:329-38.