

Feed Forward Artificial Neural Network Model to Estimate the TPH Removal Efficiency in Soil Washing Process

Hossein Jafari Mansoorian^{a,b}, Mostafa Karimae^{c,d}, Mahdi Hadi^e, Elaheh Jame Porazmey^f, Farzan Barati^g, Mansour Baziar^{d*}

^aEnvironmental Health Engineering Research Center, Department of Environmental Health Engineering, School of Health, Kerman University of Medical Sciences, Kerman, Iran.

^bYoung Researchers and Elite Clubs, Hamedan Branch, Islamic Azad University, Hamedan, Iran.

^cDepartment of Environmental Health Engineering, Aradan School of Health and Paramedicine, Semnan University of Medical Science, Semnan, Iran.

^dDepartment of Environmental Health Engineering, School of Health, Tehran University of Medical Science, Tehran, Iran.

^eCenter for Water Quality Research (CWQR), Institute for Environmental Research (IER), Tehran University of Medical Sciences, Tehran, Iran.

^fResearch Center for Environmental Pollutants, Qom University of Medical Sciences, Qom, Iran.

^gDepartment of Mechanics, Faculty of Engineering, Hamedan Branch, Islamic Azad University, Hamedan, Iran.

*Correspondence should be addressed to Dr. Mansour Baziar, Email: baziar.ehe@gmail.com

A-R-T-I-C-L-E-I-N-F-O

Article Notes:

Received: Jun 5, 2016

Received in revised form:
Dec 28, 2016

Accepted: Dec 29, 2016

Available Online: Jan 1, 2017

Keywords:

Artificial neural network,
Modeling
TPH
Surfactant
Soil contamination
Iran

A-B-S-T-R-A-C-T

Background & Aims of the Study: A feed forward artificial neural network (FFANN) was developed to predict the efficiency of total petroleum hydrocarbon (TPH) removal from a contaminated soil, using soil washing process with Tween 80. The main objective of this study was to assess the performance of developed FFANN model for the estimation of TPH removal.

Materials and Methods: Several independent repressors including pH, shaking speed, surfactant concentration and contact time were used to describe the removal of TPH as a dependent variable in a FFANN model. 85% of data set observations were used for training the model and remaining 15% were used for model testing, approximately. The performance of the model was compared with linear regression and assessed, using Root of Mean Square Error (RMSE) as goodness-of-fit measure

Results: For the prediction of TPH removal efficiency, a FANN model with a three-hidden-layer structure of 4-3-1 and a learning rate of 0.01 showed the best predictive results. The RMSE and R² for the training and testing steps of the model were obtained to be 2.596, 0.966, 10.70 and 0.78, respectively.

Conclusion: For about 80% of the TPH removal efficiency can be described by the assessed regressors the developed model. Thus, focusing on the optimization of soil washing process regarding to shaking speed, contact time, surfactant concentration and pH can improve the TPH removal performance from polluted soils. The results of this study could be the basis for the application of FANN for the assessment of soil washing process and the control of petroleum hydrocarbon emission into the environments.

Please cite this article as: Jafari Mansooriana H, Karimae M, Hadi M, Jame Porazmey E, Barati F, Baziar M. Feed Forward Artificial Neural Network Model to Estimate the TPH Removal Efficiency in Soil Washing Process. Arch Hyg Sci 2017;6(1):96-104.

Background

Soil contamination with total petroleum hydrocarbon (TPH) compounds originated from

the entrance of chemical compounds to the natural soil environments that originally come from crude oil (1). TPHs have recognized carcinogenic effects and (or) other serious disorders on human health (2-4). The

penetration of these contaminants into the groundwater resources can lead to water quality impairment which will impose special and expensive treatment technologies for remediation. One commonly used technique is soil washing by surfactants, which involves physical-chemical processes for the removal of soil contaminations (5). This process can be used as a portable technology in the polluted sites and meaningfully declined the proportion of polluted soil (6). Non-ionic surfactants such as Tween 80 and Brij 35 are the most approved surfactants that are used in soil washing technique for the removal of PAHs, TPHs, PCBs and heavy metal compounds (Ahn et al. 2009; Villa et al. 2010). These surfactants have low tendency to flocculate clay particles in soil which are cost-effective and biodegradable.

Developing model for treatment processes may help to better understanding of the processes, better management of the system's problems and also, in many cases can help to estimate and improve the system's performance. Artificial neural network methodologies (ANN) have already been applied to the prediction of arsenic removal efficiency by modified Fenton (7), modeling of adsorption of Cu (II) from industrial leachate by pumice (Turan et al. 2011), prediction of wastewater treatment plant (8) and waste stabilization pond (9) performances, modeling and prediction of water treatment plant influent quality parameters (10). The results of these studies indicated that ANNs are easy to use and effective tool at the environmental modeling and prediction.

Aims of the study:

Therefore, the main objective of this study was to develop a feed forward artificial neural network (FFANN) model to estimate the removal of TPH based on several input parameters (i.e., shaking speed, contact time, surfactant concentration and pH) in soil washing process with Tween 80.

Materials & Methods

One of the most widely used artificial neural networks configurations in modeling environmental pollution removal is FFANN structure which was used in this work (11). On the basis of the experimental data collected from the previous work (12), in the present study, a FFANN network was developed as an artificial intelligence-based approach for modeling of soil washing process. Neurons and links are important components of a neural network structure, where neurons are known as the processing elements and links provide neurons connections. Optimization and identification of input and output variables are two important steps in developing an FFANN model (13). The schematic diagram of FFANN which was used in this work is shown in Fig. 1. Each neuron in a layer which receives weighted inputs from a previous layer and transmits its output to the neurons in the next layer. The summation of weighted input signals are calculated by Eq. (1) and this summation is transferred by a linear or nonlinear function called the activation function (14). The selection of the activation function can affect the network performance, significantly. Commonly used activation functions are the hyperbolic tangent function and sigmoid function. The type of activation function is also depends on the type of neural network to be designed. The sigmoid logistic, sigmoid tangent and hyperbolic tangent transfer functions are widely used and are given in Eqs. 2, 3 and 4, respectively. One of the characteristics of sigmoid logistic function is that it only produces a value between 0 and 1. For sigmoid tangent (tansig), which is an anti-symmetric activation function the output of each neuron is permitted to assume both positive and negative values in the interval between -1 and 1, in which case it is likely for its mean to be zero. This function is mathematically equivalent to hyperbolic tangent transfer function (tanh) but it differs in that it runs faster than function. If the network connectivity is large, back-propagation learning with anti-symmetric

activation functions can yield faster convergence than a similar process with non-symmetric activation functions such as sigmoid function (15). A linear function was also used the output layer. This arrangement of functions in modeling is common and can yield better results (16).

The results of network are compared with the actual observation results and the network error is calculated with Eq. (5). The training process continues until this error reaches an acceptable value (17).

- Summation function

$$Y_{net} = \sum_{i=1}^n (W_i \times X_i) + W_0 \quad (1)$$

Sigmoid transfer function $f(Y_{net}) = \frac{1}{1 + e^{-Y_{net}}}$

$$0 \leq f(x) \leq 1 \quad (2)$$

- Sigmoid tangent transfer function

$$f(Y_{net}) = \text{tan sig}(Y_{net}) = \frac{2}{(1 + e^{(-2 \times Y_{net})}) - 1}$$

$$-1 \leq f(x) \leq 1 \quad (3)$$

- Hyperbolic tangent transfer function

$$f(Y_{net}) = \text{tanh}(Y_{net}) = \frac{e^{Y_{net}} - e^{-Y_{net}}}{e^{Y_{net}} + e^{-Y_{net}}}$$

$$-1 \leq f(x) \leq 1 \quad (4)$$

- Error function

$$J_r = \frac{1}{2} \sum_{i=1}^m e_i^2 = \frac{1}{2} \sum_{i=1}^m (Y_i - O_i)^2 \quad (5)$$

Where, Y_i is the response of neuron i , $f(Y_{net})$ is the nonlinear activation function, Y_{net} is the summation of weighted inputs, X_i is the neuron input, W_i is weight coefficient of each neuron

input, W_0 is bias, J_r is the error function between observed value and network response, O_i is the observed value of neuron i (14).

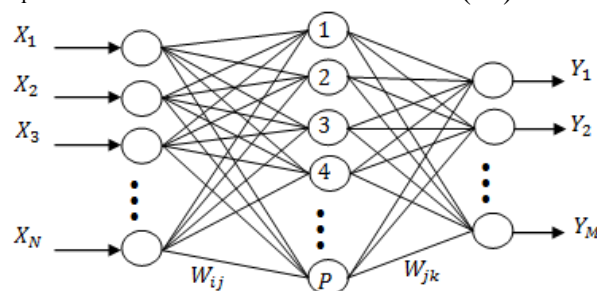


Figure 1) The schematic diagram of FFNN used in this study; P is the number of the hidden layer neurons

In this study for prediction of soil washing efficiency in removal of TPH from contaminated soil, data of pH_s, shaking speeds, surfactant concentrations and contact times are applied to the FFANN structure as repressors and TPH removal as the target. 85% of the data set was used for ANN model training and 15% for testing the model, approximately (18). As seen in Table 3, a random selection method was used for division of all data to training and testing sets. Thus, all data including input and output variables were scaled between -1 and 1. This action preserves the interpretation of the weights and prevents the numerical overflows (18). Eq. (6) shows the normalization approach used in the study.

$$S = (2 \times (X - X_{min} / X_{max} - X_{min})) - 1 \quad (6)$$

Where, X_{min} and X_{max} are the extreme values of variable X . Finally, in order to determine the errors in training and testing steps, all outputs were returned to the original scaling and were compared with experimental responses. MATLAB software was used for training and designing the feed forward neural network.

Results

The experimental investigations data were used to design the FFANN network. Summary of these data are presented in Table 1. Further

details of the network topography are summarized in Table 2.

Table 1) Summary of the experimental study results

Input parameters	Range	Removal range % (output parameters)	Best selected conditions for each parameter
Shaking speed (rpm)	100-250	43.5-70	250
Contact time (min)	10-120	38.5-75	120
Surfactant concentration (mg/L)	2-30	35-82	10
pH	2.5-9	61-72	7

Table 2) Neural network's properties

Training function	Performance function	Data division	Number of inputs	Number of outputs	Transfer functions	Number of neurons in 1 st layer	Number of neurons in 2 nd layer	Number of neurons in 3 rd layer	Number of epochs
Trainlm	MSE	Dividerand	4	1	Tansig	4	3	1	1000

The statistical parameters including mean (\bar{x}), standard deviation (S_x) and skewness coefficient (C_{sx}) for the training and testing data sets are given in Table 3. Table 4 shows

performance of FFNN for TPH removal estimations.

Table 5 shows the linear regression coefficients which are slopes and intercepts of the fitted lines.

Table 3) Statistical parameters for neural network input and output data

Variables	Training data set			Testing data set		
	\bar{x}	S_x	C_{sx}	\bar{x}	S_x	C_{sx}
Shaking speed (rpm)	233.03	41.41	-2.26	240	22.3607	-1.50
Contact time (min)	64.82	31.92	-0.43	81	32.86	-0.56
Surfactant concentration (mg/L)	11.14	13.09	2.66	7	2.73	0.40
pH	6.82	1.13	-1.78	5.90	1.94	-1.33
TPH removal (%)	57.31	14.33	-0.24	63.42	6.84	-0.31

Table 4) Performance of FFNN for removal percentage estimation

Output	Training step		Testing step	
	RMSE	RMSE	R ²	RMSE
TPH removal (%)	2.596	10.706	0.785	10.706

Table 5) Linear regression between neural network output and desired values

output	Training data set		Testing data set		All data set	
	a^1	b^2	a	b	a	b
Removal Percentage	1	7.89×10^{-4}	0.862	9.741	0.936	2.905

1: a is linear regression coefficient, 2: b is the Intercept

Figs. 2a, 3a and 4a show the real values and model computed values of TPH removal in testing, training and all data sets for the

structure of 4-3-1, respectively. In Figs. 2b, 3b and 4b, the TPH removal estimations by FFNN versus the observed values are shown.

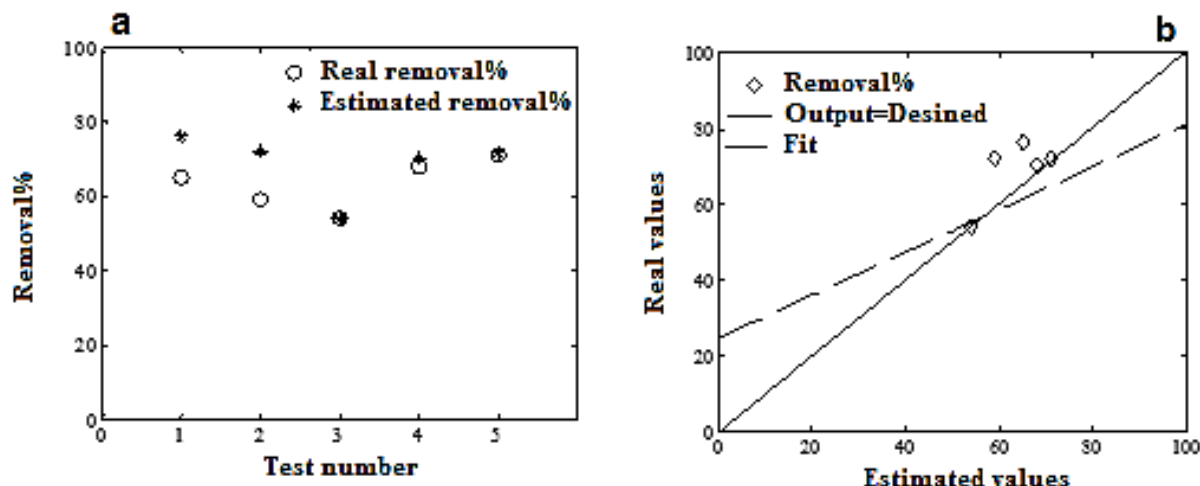


Figure 2) Removal percentage calculated by the neural network for testing data set versus the real values (a) and corresponding linear regression (b) for test data set testing step

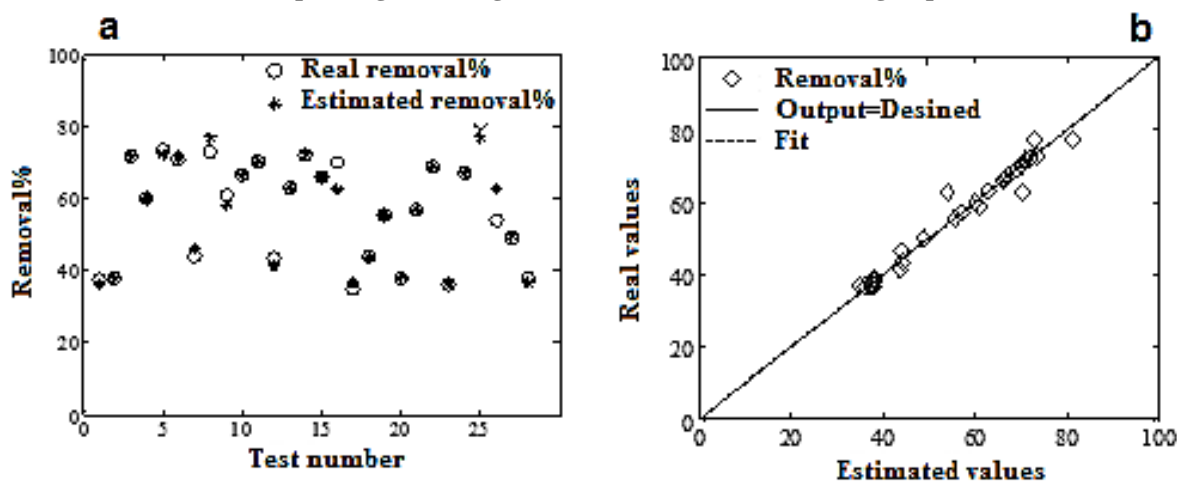


Figure 3) Removal percentage calculated by the neural network versus the real values (a) and corresponding linear regression (b) for train data set

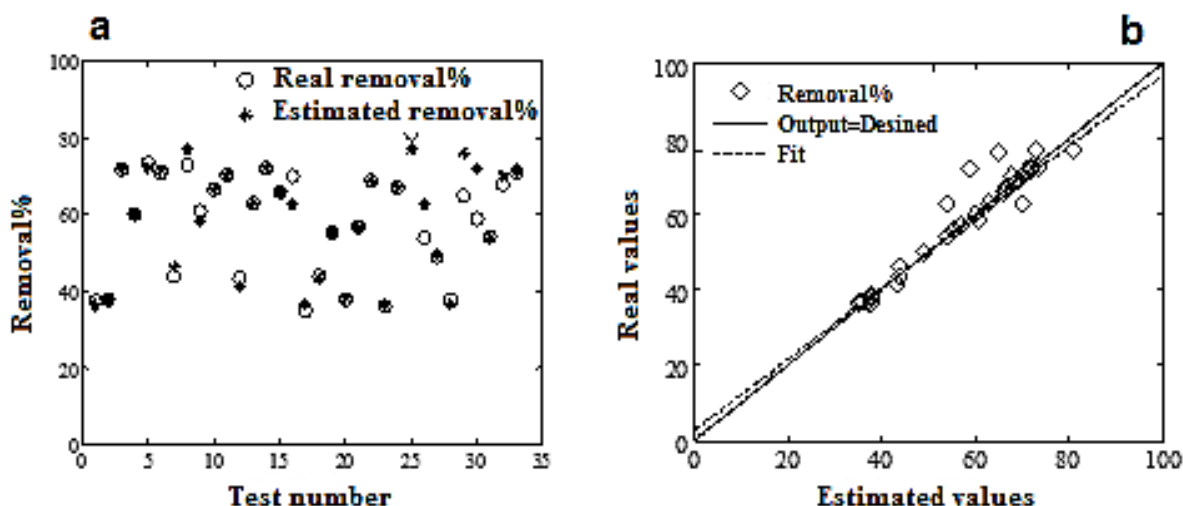


Figure 4) Removal percentage calculated by the neural network versus the real values (a) and corresponding linear regression (b) for all data set

Discussion

FFNN with a backpropagation learning algorithm (BPLA) used to compute the network parameters by an iterative Levenberg-Marquardt (LM) algorithm (14). The MATLAB trainlm was used for the optimization process. This algorithm has a highest learning speed and a high performance in comparison other optimization algorithms (14). Based on the selected network structure in this study, the training process was conducted to achieve a performance target of 1×10^{-3} for a maximum training epoch of 1000 with a learning rate of 0.01. This value was obtained after several trials and error runs (more than 100 iterations). It was found that this value insures stable fast learning.

Large standard deviation for shaking speed, contact time and removal percentage means that values in data set are farther away from the mean, on average. In the training phase, it is important to have a data set with considerably large standard deviation, because the normalized data are uniformly distributed and the network inputs do not have a compact data set (14). Performances of the neural network model for both training and testing data were

evaluated by coefficient of determination (R^2) and Root of Mean Square Error (RMSE) (9, 18).

$$RMSE = \sqrt{\left[\frac{\sum_{i=1}^N (X_d - X_s)^2}{N} \right]} \quad (7)$$

$$R^2 = \frac{\sum_{i=1}^N (X_d - \bar{X}_d)^2 - \sum_{i=1}^N (X_d - \bar{X}_s)^2}{\sum_{i=1}^N (X_d - \bar{X}_d)^2} \quad (8)$$

Where, is the estimated variable, is the desired variable and is the average of the desired variable. The Efficiency (E) evaluates model performance. The values of R^2 close to 1.0 indicate good model performance. RMSE evaluates the residual between estimated and desired variable. Whatever this criteria be close to zero then model represents the good fit.

To determine the performance of the developed network in removal estimation of TPH, the regression coefficients are precisely measured for training and testing data sets.

The results showed that a model with three layered structure including an input layer with 4

neurons, a hidden layer with 3 neurons and an output layer with 1 neurons (4-3-1) may provide the most of performance. To receive this structure, different structures of FFNN were evaluated. Some structures with one hidden layer containing 2, 3, 4, 5, 6, 8, 10, 12, and 14 neurons were also evaluated. Moreover, the performance of some structures with two hidden layers containing 2:2, 4:2, 6:4, 4:6, 8:8, 10:8, and 8:10 neurons in first and second hidden layer were also investigated.

As shown in Fig 2a, the maximum errors are related to the datasets 1 and 2. This is mainly due to the conjunction of selected backpropagation algorithm with gradient descent optimization in neural network training step (14). The principal restriction of this algorithm is converging the results into the local minima (14). Hence, derived weights from the network cannot make the minimum errors in some of train data sets. As shown in Table 4, the performance of FFNN for the estimation of removal percentage was analyzed by RMSE and R^2 statistics. The values of RMSE and R^2 for the selected structure of FFNN (4-3-1) in training and testing steps were found to be 2.596, 0.966 and 10.706, 0.785, respectively. These mean that the residual values between estimated and desired (observed) values for training and testing steps are 2.596 and 10.706%. In order to better determining the performance of the network, linear regression between neural network outputs and desired values was also drawn and the regression coefficients according to the following equation were calculated: where. It can be seen in Table 5; the linear regression coefficient for all data set was obtained 0.936 it means that only 0.064 errors in our results can be found. Thus, the present study showed that FFANN is able to model and predicts TPH removal based on the 4 input parameters data and with a small data set. However, due to inadequate data in the testing phase of FFANN model, it was not able to improve the results of this step. It seems that the training and testing of an FFANN with

small data set is difficult to guarantee an excellent predictive model, where only 28 and 5 data were available.

In recent years, utilization of computational tools for the assessment of pollutants has become increasingly valuable due to their capability to interpret integrated variable measurements. ANNs are considered as inexpensive techniques for the data interpretation and prediction. In a study conducted by Olawoyin et al (19) ANN was used for recognizing the spatial patterns in contaminated zones Niger Delta by integrating chemical, physical, ecotoxicological and toxicokinetic variables in the identification of pollution sources. In this study the physico-chemical properties (pH, TPH, BTEX, PAH, COD, SO_4 , PO_4 , NO_3 , and heavy metals) in the recipient environments were assessed in sediments, soils and water. The ANN was used as a powerful visualization tool to identify trends in the dataset. However in some studies some drawbacks of feedforward neural models (FFANN) have been highlighted for example when a large number of input/output variables are available the learning process for FFANN may be very time consuming. In these cases the use of a recurrent neural network model (RNNM) for the prediction are proposed. As a case study the prediction of the biodegradation profiles of hydrocarbons contained in an aged polluted soil was conducted with RNNM models by De la Torea-Sanchez et al. (20).

Although various types of neural network models have been developed, but the model is still considered as one of the most popular and widely-used network paradigm (11). FANN in this study provides acceptable findings regarding its valuable application in soil washing process optimization.

Conclusion

It was revealed that the soil washing process influencing parameters could be applied to the

prediction of the removal efficiency of TPH from the polluted soils using a FFANN model. In this paper following conclusions could be highlighted:

- A feed forward artificial neural network with input repressors including shaking speed, contact time, surfactant concentration and pH and a three-hidden-layer structure of 4-3-1 was provide the best performance for the estimation of TPH's removal efficiency (as the only output variable).
- The performance of model in the training and testing steps was found acceptable by using linear regression, Root of Mean Square Error (RMSE) and coefficient of determination (R^2) statistics. Results of the testing step as a critical step of model checking reveals that the FFANN model could provide reasonable predictions for the TPH removal efficiency (%) parameter.
- It is estimated that for about 80% of the TPH removal can be described by the assessed regressors developed model. Thus according to the findings of this study focusing on the optimization of soil washing process regarding to these four recognized control variables (shaking speed, contact time, surfactant concentration and pH) can improve the more performance of TPH removal efficiency from polluted soils.
- The results of this study could be the basis for the application of artificial neural network in the assessment of soil washing process and control of petroleum hydrocarbon emission into the environments.

Footnotes

Conflict of Interest:

The authors declared no conflict of interest.

References

1. ATSDR. Toxicological profile for total petroleum hydrocarbons (tph). US Department of Health and Human Services, Public Health Service Atlanta, GA, USA. 1999;
2. Haigh SD. A review of the interaction of surfactants with organic contaminants in soil. *Sci Total Environ* 1996;185(1):161-170.
3. Lizhong Z, Baoliang C, Jing W, Hongxin S. Pollution survey of polycyclic aromatic hydrocarbons in surface water of Hangzhou, China: *Chemosphere* 2004;56(11):1085-1095.
4. Iturbe R, Flores-Serrano R, Castro A, Flores C, Torres L. Subsoil TPH contamination in two oil pipeline pumping stations and one pipeline right-of-way in north Mexico. *J Environ Manag* 2010;91(11):2396-2402.
5. Khalladi R, Benhabiles O, Bentahar F, Moulai-Mostefa N. Surfactant remediation of diesel fuel polluted soil. *J Hazard Mater* 2009;164(2):1179-1184.
6. US EPA. Citizen's Guide to Soil Washing, 1996, EPA 542-F-01-008, May 2001. United States Environmental Protection Agency; 2001.
7. Jaafarzadeh N, Ahmadi M, Amiri H, Yassin MH, Martinez SS. Predicting Fenton modification of solid waste vegetable oil industry for arsenic removal using artificial neural networks. *J Taiwan Institute Cheml Eng* 2012;43(6):873-878.
8. Hamed MM, Khalafallah MG, Hassanien EA. Prediction of wastewater treatment plant performance using artificial neural networks. *Environ Model Softw* 2004;19(10):919-928.
9. Khodadadi M, Mesdaghinia A, Nasser S, Ghaneian MT, et al. Prediction of the waste stabilization pond performance using linear multiple regression and multi-layer perceptron neural network: a case study of Birjand,. *Iran Environ Health Eng Manage J* 2016;3(2). (Full Text in Persian)
10. Aydiner C, Demir I, Yildiz E. Modeling of flux decline in crossflow microfiltration using neural networks: the case of phosphate removal. *J Membr Sci* 2005;248(1):53-62.
11. Zhang G, Patuwo BE, Hu MY. Forecasting with artificial neural networks: The state of the art. *Int J Forecast* 1998;14(1):35-62.
12. Baziar M, Mehrasebi MR, Assadi A, Fazli MM, Maroosi M, Rahimi F. Efficiency of non-ionic surfactants-EDTA for treating TPH and heavy metals from contaminated soil. *J Environ Health Sci Eng* 2013;11(1):1.
13. Aleboyeh A, Kasiri M, Olya M, Aleboyeh H. Prediction of azo dye decolorization by UV/H₂O₂ using artificial neural networks. *Dyes and Pigments*. 2008;77(2):288-294.
14. Hagan MT, Demuth HB, Beale MH. *Neural network design*. Thomson Learning Stamford, CT; 1996.
15. LeCun YA, Bottou L, Orr GB, and Müller KR. Efficient backprop, in *Neural networks: Tricks of the trade*. New York: Springer; 2012. p. 9-48.
16. Sharma A, Sharma H, Sahoo P, Tripathi R, Meher L. ANN Based Modeling of Performance and

Emission Characteristics of Diesel Engine Fuelled with Polanga Biodiesel at Different Injection Pressures. *Int Energy J* 2015;15(2):57-72.

17. Looney CG. Pattern recognition using neural networks: theory and algorithms for engineers and scientists. New York: Oxford University Press, Inc; 1997.

18. Solaimany-Aminabad M, Maleki A, Hadi M. Application of artificial neural network (ANN) for the prediction of water treatment plant influent characteristics. *J Adv Environ Health Res* 2014;(2):89-100.

19. Olawoyin R., Nieto A, Grayson R.L, Hardisty F, Oyewole S. Application of artificial neural network (ANN)-self-organizing map (SOM) for the categorization of water, soil and sediment quality in petrochemical regions. *Expert Sys Appl* 2013;40(9):3634-3648.

20. De la Torre-Sánchez R, Baruch I, Barrera-Cortés J. Neural prediction of hydrocarbon degradation profiles developed in a biopile. *Expert Sys Appl* 2006;31(2):383-389.