academicJournals

Vol. 8(38), pp. 1896-1905, 11 October, 2013
DOI 10.5897/SRE12.727
ISSN 1992-2248 © 2013 Academic Journals
http://www.academicjournals.org/SRE

Scientific Research and Essays

Full Length Research Paper

Modeling the performance of upflow anaerobic filter (UAF) reactors treating paper-mill wastewater using neural networks

Turan Yilmaz

Environmental Engineering Department, Cukurova University, 01330, Balcali, Adana/Turkey.

Accepted 12 March, 2013

This study investigated the predictive ability of neural networks in the estimation of methane yield (MY) and effluent substrate (Se as mg/L COD) produced by two anaerobic filters, one mesophilic (35°C) and one thermophilic (55°C), which were operated to treat paper-mill wastewater at varying organic loadings. An artificial neural network (ANN) architecture was optimized to obtain a three-layer neural network, composed of three inputs, namely hydraulic retention time (HRT), organic loading rate (OLR), and influent substrate (Si as mg/L COD), six hidden neurons and one output neuron, Se or MY. Stover-Kincannon model and Multi-linear regression (MLR) technique was also used for data analysis and to compare the prediction capability. Four statistical criteria also used for comparison were mean square error (MSE), mean absolute error (MAE), mean absolute relative error (MARE), and determination coefficient (R^2). The results showed that ANN approach predicted the performance of the anaerobic filters better than both Stover-Kincannon model and MLR technique.

Key words: Anaerobic digestion, mesophilic, thermophilic, paper-mill wastewater, neural networks.

INTRODUCTION

Many kinetic and mechanistic models have been proposed to understand the treatment process of anaerobic reactors including linear models such as Monod and Stover-Kincannon and non-linear models such as axial dispersion model (Singhal et al., 1998) and dynamic model (Wu and Hickey, 1997). However, several researchers pointed out that none of the mechanistic models able to completely explain or predict the performance of an upflow anaerobic sludge blanket (UASB) reactor treating industrial or domestic wastewater under various input conditions (Sinha et al., 2002; Mu and Yu, 2007; Singh et al., 2010). Therefore, new approaches are required for the estimation of the treatment performance of anaerobic reactors. One of the relatively new computational tools is *Artificial Neural Network* (ANN), which has already been applied to a wide spectrum of problems in a variety of fields, such as finance (Budcema and Sacco, 2000), medicine (Papik et al., 1998), physics (Fang and Wu, 2007), geology (Yuanyou et al., 1997), hydrology (ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000a; ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000b), and environmental engineering (Yilmaz et al., 2010).

Among the anaerobic treatment processes, upflow anaerobic filter reactors (UAFR) are gaining more popularity with reduced probability of sludge bulking and flotation while providing degradation rates similar to other anaerobic treatment processes. After an extensive literature search, it has been realized that there is no

E-mail: tyilmaz@cu.edu.tr.

literature information on ANN application to model the performance of UAF reactor treating paper-mill wastewater while they have been applied to UASB reactors (Sinha et al., 2002; Mu and Yu, 2007; Singh et al., 2010). The performance of mesophilic and thermophilic UAF reactors treating paper-mill wastewater was evaluated and modeled using a well-known Stover-Kincannon model in our previous study (Yilmaz et al., 2008). In order to further understand the methane producing mesophilic and thermophilic reactors, it is essential to develop new models to quantitatively describe the reactor performance. Furthermore, accurate estimation of methane yield (MY) and effluent substrate (Se as mg/L COD) will lead to time conservation and cost reduction in the operation instead of measuring these parameters repeatedly.

Therefore, the objective of this study is to establish and apply an ANN model to simulate the performance of UAF reactor. Multi-linear regression (MLR) technique was also used for data analysis and to compare the prediction ability of the developed models. Mean square error (MSE), mean absolute error (MAE), mean absolute relative error (MARE), and determination coefficient (R^2) were used as statistical means for comparison between the modeling approaches.

THEORETICAL BACKGROUND OF THE METHODS

Multi-linear regression (MLR)

The relationship between a dependent variable and independent variables can be linearly constructed by MLR method, where the dependent variable, y, is regarded as a linear function of ***p*** number of independent variables, x_1, x_2,..., x_p. In this case, the linear relationship can be described with the following equation:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ...+ \beta_p x_p + \varepsilon \qquad (1)$$

where ε denotes the residual, which is a normally distributed random variable with a mean of zero. The regression coefficients, β_0, β_1, β_2,…, β_p, are computed for the lowest sum of squares of differences between the predicted and observed values (Bayazıt and Oguz, 1998).

Artificial neural network (ANN) model

McCulloch and Pitts (1943) were the first to introduce artificial neural networks (ANNs) as modeling tools and the interest in these tools grew until Minsky and Papert (1969) demonstrated the training difficulties and low effectiveness in training. It was Rumelhart et al. (1986) who rediscover a calibration algorithm that could be used to train networks for sufficient sizes while simplifying the complexities for practical purposes. Since then, research with the focus on ANNs has expanded resulting in the evolution of a number of different network types, training algorithms and tools.

When sufficient data is provided for a given complexity, ANNs can be easily trained to model any relationship between a series of independent and dependent variables. This flexibility has enabled ANNs to be considered as a set of global predictors and to be usefully applied to a wide variety of problems that are difficult to understand, define, and quantify. In the context of this paper, ANNs are trained to represent the relationship between a range of operating parameters of OLR, HRT, and Si as input parameters and associated output parameters of MY and Se. There is no need for the modeler in this case to fully define the intermediate relationships between the input and output variables that the ANN identifies during the 'learning process'.

There are various network types and training algorithms described and used in the literature. However, this paper focuses on only the Multi-Layer Perceptron (MLP) approach, which is the most referred ANN tool. Figure 1 provides an overview of the MLP structure. MLP employs three layers of neurons - an input layer, a hidden layer, and an output layer - with each neuron consisting of a number of inputs provided from outside the network or the previous layer and a number of outputs leading to the subsequent layer or out of the network. The output response is computed by neurons based on the weighted sum of all inputs according to an activation function. The data flows from external inputs, which are transmitted through the hidden layer, to the output layer from which the external outputs are obtained. The training of the network is then continued by adjusting the weights that connect the neurons using Levenberg-Marquardt algorithm. Interested readers are suggested to refer to neural network texts for more detailed coverage such as Haykin (1998).

MATERIALS AND METHODS

The data used for training and testing the MLR and ANN models were obtained from two laboratory-scale UAF reactors, one mesophilic (35°C) and one thermophilic (55°C) which were operated to treat paper-mill wastewater under varying organic loading rates. The paper-mill wastewater used in the study was obtained from a local paper mill plant, which processes and recycles scrap and waste paper collected from various sources. These paper materials are mostly collected from the rubbish bins or collection sites operated by governmental offices and bureaus for recycling. The characteristics of the paper-mill wastewater used herein are given in Table 1. Detailed information about the reactors and analytical methods can be found in our previous paper (Yilmaz et al., 2008).

The MLR technique described in this study was employed using MINITAB software program, whereas, ANN model was employed in MATLAB v.7.6 and run under the Microsoft Windows XP environment. Predicted results by the MLR and ANN models from the present study were compared with real measurements in the

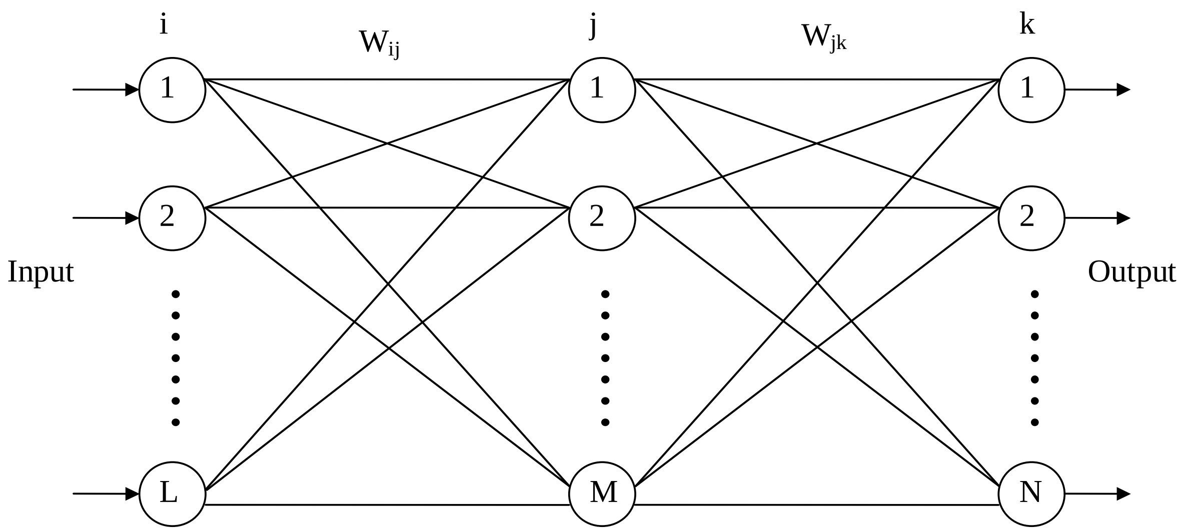

Figure 1. Multi-Layer Perceptron (MLP) structure.

Table 1. Characteristics of paper-mill wastewater.

Parameter	Concentrations (mg/L)
COD	1972 – 3536 (2701 ± 641)
SCOD	1032 – 2342 (1335 ± 427)
Total nitrogen	4.8 – 3.7
BOD_5	1058 – 1489 (1201 ± 304)
Alkalinity (as $CaCO_3$)	350
TSS	916 – 3100 (1201 ± 304)
Suspended solids after settlement	205
Total hardness (as $CaCO_3$)	500
Ca^{2+}	150
Mg^{2+}	30.45
Boron	0.27
pH	7.6 – 7.0 (7.45 ± 0.29)

lab-scale reactors with respect to statistical mean square error (MSE), mean absolute error (MAE), mean absolute relative error (MARE), and determination coefficient (R^2). MSE, MAE and MARE were found by the following equations:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(E_i^m - E_i^p\right)^2 \qquad (2)$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|E_i^m - E_i^p\right| \qquad (3)$$

$$\mathrm{MARE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{E_i^m - E_i^p}{E_i^m}\right| \times 100 \qquad (4)$$

In Equations 2 to 4, E_i^m and E_i^p denote the measured and predicted output parameters, respectively, and *N* is the total number of data.

RESULTS AND DISCUSSION

Application of multi-linear regression (MLR) and artificial neural network (ANN)

A total of 114 experimentally determined data sets were used with equal numbers of data for mesophilic and thermophilic digesters in this study. One of the most complex tasks is the selection of variables to be used in modeling the system. An important essential variable is the organic loading rate (OLR), which is well correlated with the Hydraulic Retention Time (HRT) (R^2=0.97) and the methane yield (MY) (R^2=0.89). The substrate concentration (Si as mg/L COD) in the influent is the most important variable to consider as it will influence the substrate concentration (Se) in the effluent. The determination coefficient, R^2, was equal to 0.79 between Se and Si. By taking the above considerations into account, organic loading rate (OLR), hydraulic retention time (HRT), and influent substrate (Si) were the three variable sets analysed as input variables along with two corresponding output variables, that is, methane yield (MY) and effluent substrate (Se).

As described in the literature, ANNs are similar to conventional statistical models in the sense that model parameters (for example, connection weights) are adjusted in the model calibration phase (training) so as to minimize the error between the model outputs and the corresponding measured values for a particular data set (the training set). Like all models, ANNs also perform best when they do not extrapolate beyond the range of the data used for calibration (Minns and Hall, 1996; Tokar

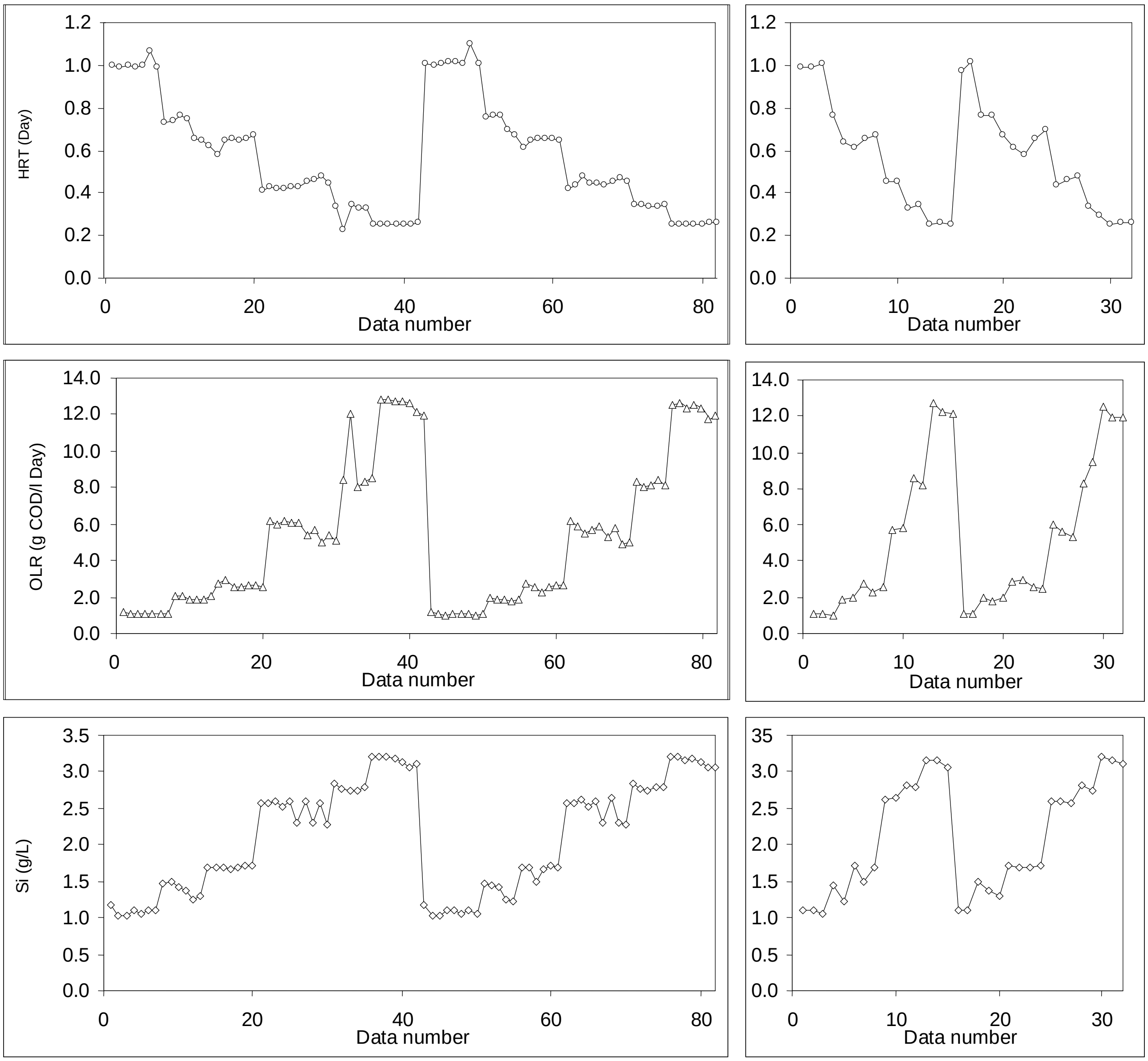

Figure 2. Sets of input variables used for training and validating the models for methane producing UAF reactor: (a) HRT for training; (b) HRT for testing; (c) OLR for training; (d) OLR for testing; (e) Si for training; and (f) Si for testing.

and Johnson, 1999). Thus, the purpose of ANNs is to non-linearly interpolate within high-dimensional space between the data used for calibration. Therefore, a separate validation set is needed to ensure that the model can generalize within the range of the data used for calibration. It is common practice to divide the available data into two subsets; a training set, to construct the neural network model, and an independent validation set to estimate the model performance in a deployed environment (Maier and Dandy, 2000). For this purpose, the operational data set was randomly split into two subsets, with 82 data sets corresponding to 75% of the whole data being used for training and 32 data sets corresponding to 25% of the data for testing the performance of the artificial neural network. The subsets were chosen randomly in order to not jeopardize the determination of the trend between the independent variables and the dependent variable. As long as the randomization technique was applied, the results proved to be similar based on our various trials. One important aspect here is to make sure that the minimum and maximum of the testing data set falls within the minimum and maximum of the training data set. Sets of input variables used to train and test the models are shown in Figure 2.

In the current work, the MLR and ANN techniques were firstly applied to the training dataset. Using MLR technique, the following equations were found to offer the

Table 2. Minimum and maximum values of the input and output parameters.

Model parameter	Training data set		Testing data set	
	Min	Max	Min	Max
HRT(Day)	0.229	1.098	0.248	1.019
Si (g/l)	1.038	3.210	1.045	3.200
OLR (g COD/l day)	1.002	12.829	1.035	12.685
MY (Lmethane/gCOD)	0.130	0.350	0.180	0.330
Se	0.162	0.912	0.190	0.711

best statistical measures for a good fit of the training dataset:

$$MY = 0.1575 - 0.0356\ HRT + 0.0756\ Si - 0.0083\ OLR \quad (5)$$
$$Se = - 0.1822 + 0.0865\ HRT + 0.2954\ Si - 0.0092\ OLR \quad (6)$$

Before applying the ANN model to the observed data, the input and output values for the training step were normalized using the following equation:

$$a\frac{x_i - x_{\min}}{x_{\max} - x_{\min}} + b \quad (7)$$

where $x_{\min}$ and $x_{\max}$ denote the minimum and maximum of the input and output parameters provided in Table 2. There are no fixed rules for which standardization approach should be used for the scaling factors, "*a*" and "*b*", in particular circumstances, to which different values can be assigned (Dawson and Wilby, 1998). The values of *a* and *b* were taken as 0.6 and 0.2, respectively. The scaling factors were chosen as 0.2 and 0.8 in order to give the ANNs the flexibility to predict flows beyond the training range. Based on our trials, higher flexibility of the ANN model was obtained for these scaling factors. This means that the ANN model can be applied for some data out of the range of the training dataset.

In this study, the performance of various network models with different hidden layer neuron numbers was examined to choose an appropriate number of hidden layer neurons. Hence, one neuron was used in the hidden layer at the beginning of the process, and then the neuron number was gradually increased by adding one more neuron until no significant improvement is noted. The number of hidden nodes was determined using trial and error method. In the current study, the ANN was trained using Levenberg–Marquardt technique because this technique was proved to be more effective than the conventional gradient descent technique (Haykin, 1998). The ANN networks training were stopped after 50 epochs since the variation of error was too small after this epoch.

The tangent sigmoid, logarithmic sigmoid and pure linear transfer functions were tried as activation (transfer) functions for hidden and output layer neurons to determine the best network model. After trying various network structures and iteration numbers, the most appropriate results were obtained from the ANN (4, 6, 1) and ANN (4, 8, 1) models for estimating MY and Se, respectively.

In order to demonstrate the predictive ability of MLR and ANN models, experimental MY and Se values were plotted against their corresponding values obtained by training and testing in the modeling as shown in Figures 3 to 6. For the sake of better evaluation of the results, an equality line has been drawn. The determination coefficients (R^2) for MY (0.7618) and Se (0.8598) in training phase demonstrate that ANN learned the relationship between the input and output data better than MLR, which produced R^2 values of 0.5166 and 0.7616 for MY and Se, respectively. Similarly, in the testing phase, the determination coefficients for MY (0.8748) and Se (0.8622) suggest that the tested ANN model again obtained better estimations than those of MLR with R^2 values of 0.7986 and 0.8263 for MY and Se, respectively.

The inference above was supported by the MARE, MAE, and MSE values when both models were compared. For the quantitative evaluation of the comparison between the predicted MY and Se values using the MLR and ANN models and the measured values, MSE, MAE and MARE, are also provided as given in Table 3. The statistical values indicate that, in terms of MSE, MAE, and MARE, the ANN model performs slightly better than the MLR technique in the estimation of both MY and Se. For example, for MY, MARE values were 10,963 for MLR and 7,837 for ANN in training, whereas MARE values were 7,361 for MLR and 5,427 for ANN in testing.

For both training and testing phases, neither ANN nor MLR indicated systematic over- or under-prediction based on output variables. Therefore, both MLR and ANN were appropriate to predict the outputs of the UAF reactor although the latter produced slightly more accurate results.

Comparison of artificial neural network (ANN) and multi-linear regression (MLR) with Stover-Kincannon

The linearised Stover-Kincannon model for steady state

Table 3. MSE, MAE, MARE, and R^2 statistics of ANN and MLR models for both training and testing phases.

Parameter	Criteria	Training		Testing	
		MLR	ANN	MLR	ANN
MY	MARE	10.963	7.837	7.361	5.427
	MAE	0.025	0.018	0.019	0.014
	MSE	0.001	0.001	0.001	0.0003
	R^2	0.5166	0.7618	0.7986	0.8748
Se	MARE	15.745	10.879	14.843	10.080
	MAE	0.065	0.046	0.055	0.040
	MSE	0.009	0.005	0.006	0.004
	R^2	0.7616	0.8589	0.8263	0.8622

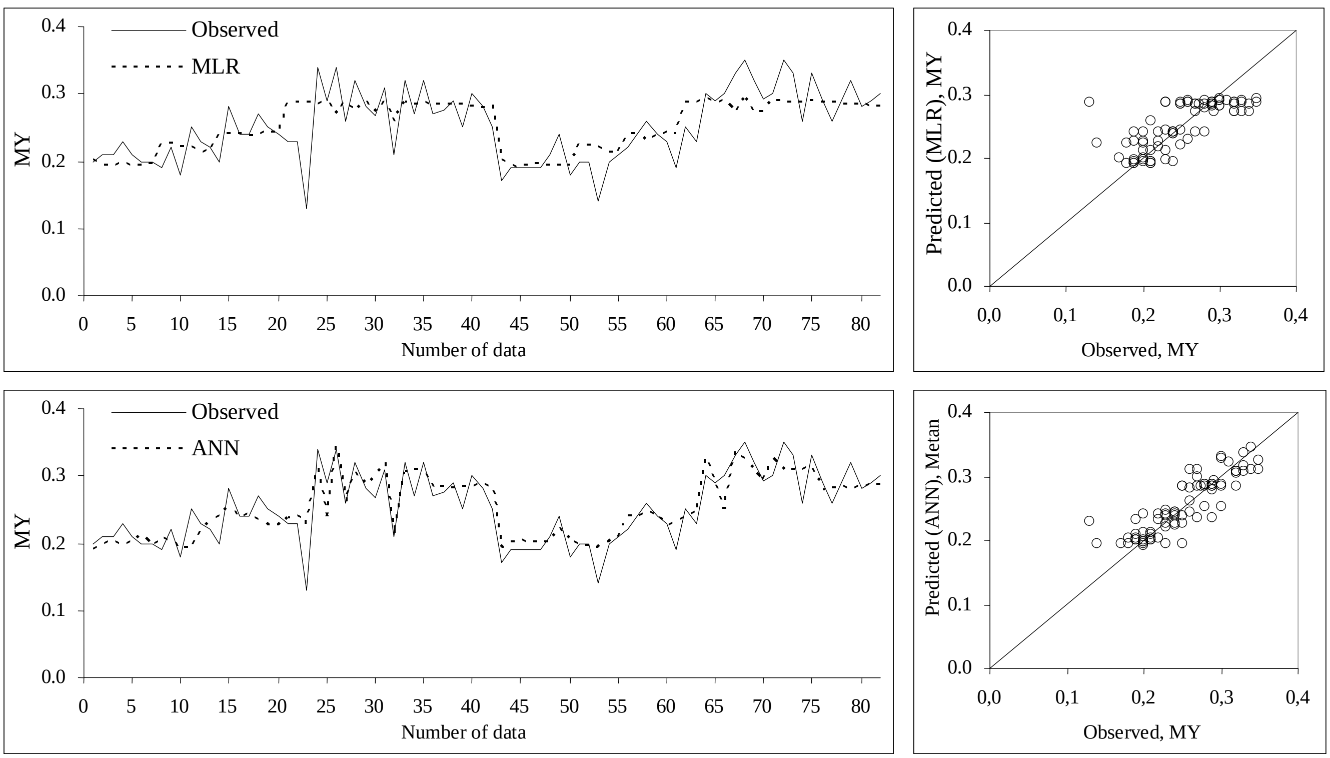

Figure 3. Plot of observed and predicted MY values by MLR and ANN for training phase.

conditions is given as follows (Tay et al., 1996):

$$\frac{V}{Q(Si-Se)} = \frac{K_B}{U_{max}}\left[\frac{V}{QSi}\right] + \frac{1}{U_{max}} \qquad (8)$$

Yilmaz et al. (2008) previously plotted V/Q(Si-Se) versus 1/OLR using the same data analyzed in this work and determined the following relationship equations with K_B/U_{max} as the slope and $1/U_{max}$ as the intercept for the mesophilic and thermophilic digesters, respectively:

$$\frac{V}{Q(Si-Se)} = 1.2081\left[\frac{1}{OLR}\right] + 0.0116 \qquad (9)$$

$$\frac{V}{Q(Si-Se)} = 1.2658\left[\frac{1}{OLR}\right] + 0.0015 \qquad (10)$$

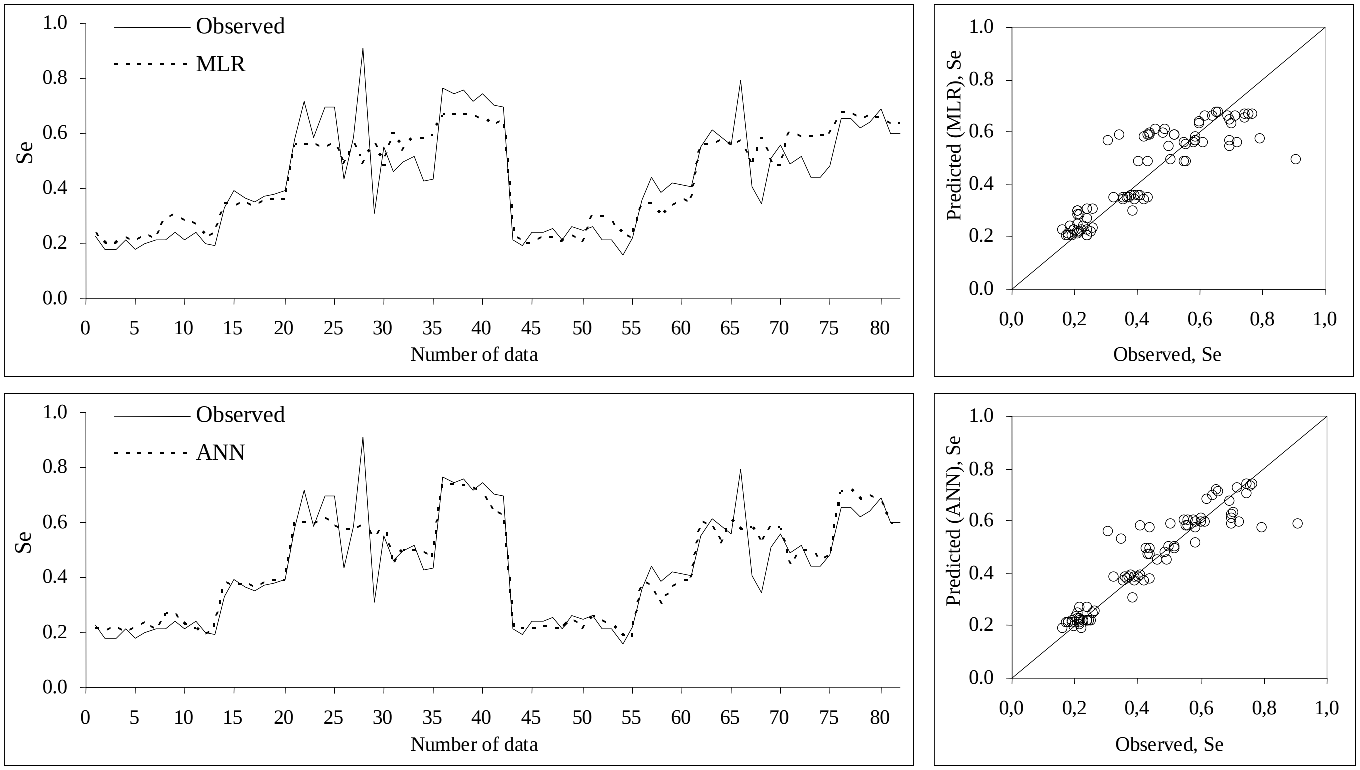

Figure 4. Plot of observed and predicted Se values by MLR and ANN for training phase.

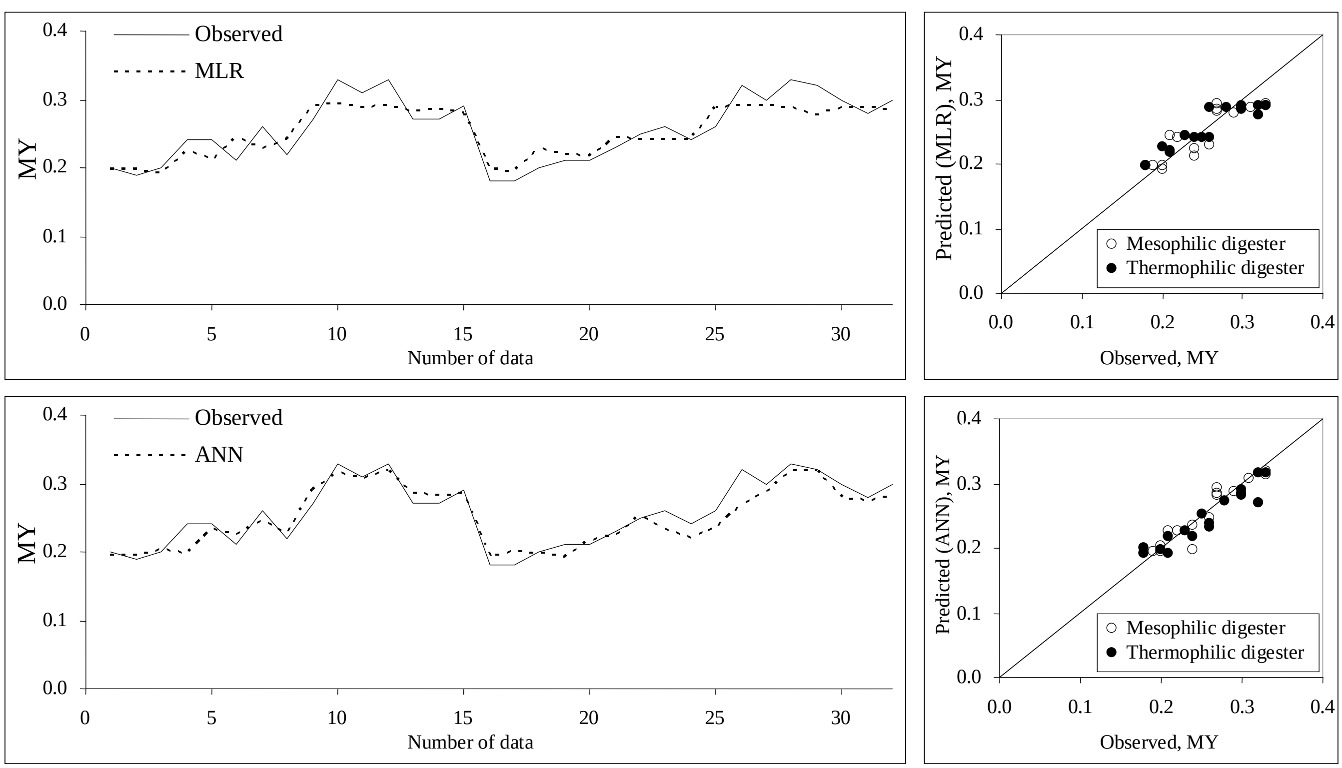

Figure 5. Plot of observed MY and predicted values by MLR and ANN for testing phase.

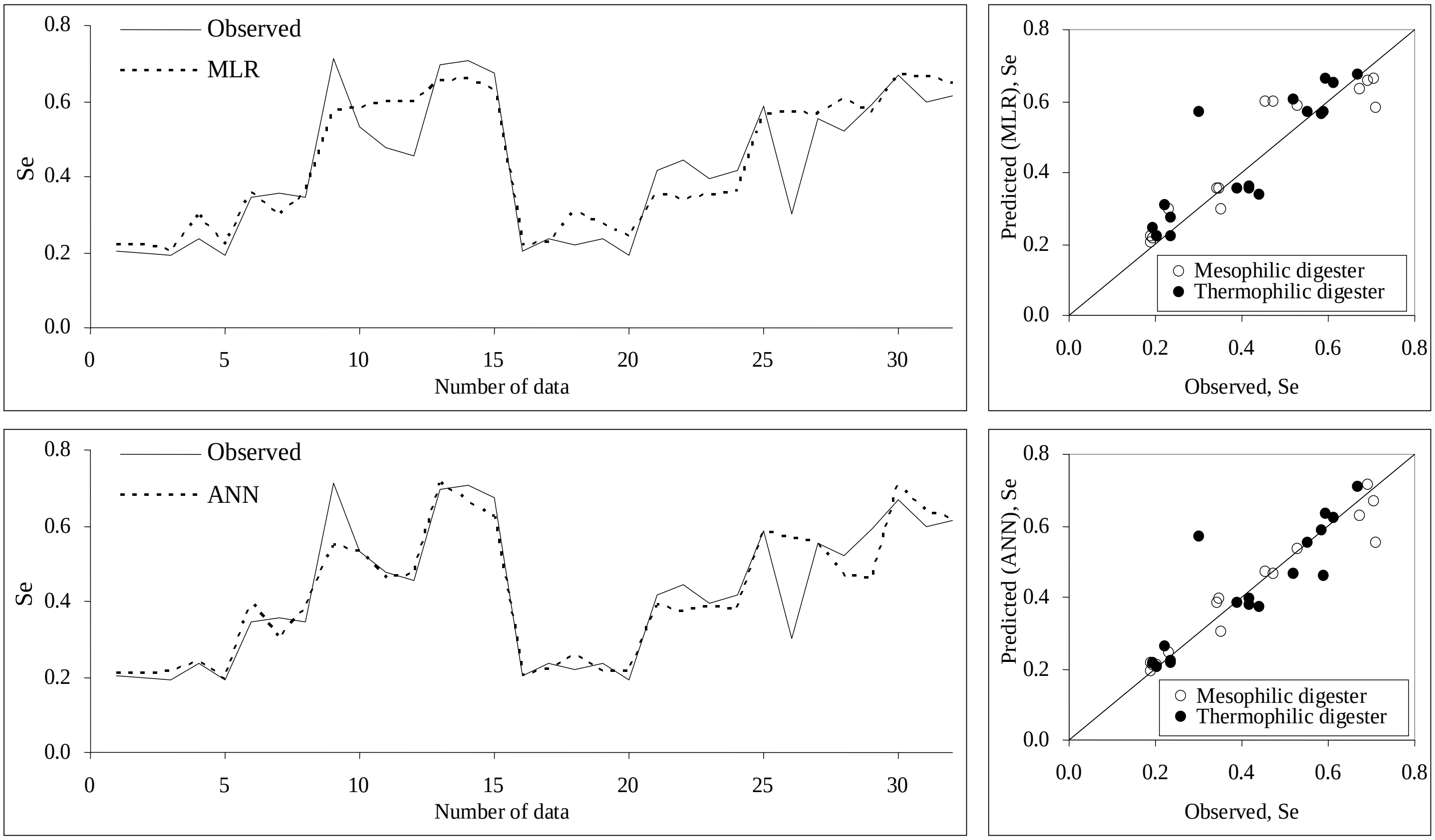

Figure 6. Plot of observed and predicted Se values by MLR and ANN for testing phase.

After some arrangements, Equations 9 and 10 determined by Yilmaz et al. (2008) can be expressed as follows for the mesophilic and thermophilic digesters, respectively:

$$Se = Si - \left[\frac{V\ OLR}{1.2081\ Q + 0.0116\ Q\ OLR}\right] \qquad (11)$$

$$Se = Si - \left[\frac{V\ OLR}{1.2658\ Q + 0.0015\ Q\ OLR}\right] \qquad (12)$$

Herein, Equations 11 and 12 developed based on Stover-Kincannon model were applied to the data obtained from the mesophilic and thermophilic digesters and were compared to the MLR and ANN models in terms of Se (Figure 7). Using the ANN model, coefficients of determination obtained for Se estimation were 0.8890 and 0.8404 for mesophilic and thermophilic digesters, respectively, suggesting that the ANN model is better in estimating the Se than those of Equations 11 and 12 based on Stover-Kincannon model, which produced R^2 of 0.7574 and 0.7835 for mesophilic and thermophilic digesters, respectively. It should be noted that the MLR technique was similar to Stover-Kincannon model in predicting Se where R^2 were 0.7894 and 0.7871, respectively.

Conclusions

This study demonstrates that it is feasible to apply ANN modeling technique to simulate the performance of UAF reactor producing methane under various operating conditions such as organic loading rate (OLR), hydraulic retention time (HRT), and influent substrate (Si). The results showed that Se and MY can be simulated well by both MLR and ANN models although the latter produced more accurate estimations than MLR. The MLR models proposed herein are rough approximations of nonlinear models with similar results to those of Stover-Kincannon. Therefore, MLR is suggested that it can be used for the preliminary analysis of UAF reactor as well as Stover-Kincannon instead of ANN.

The predictive ability of different neural network approaches is usually case dependent. It would be interesting to find out how some of the other models such as neuro-fuzzy inference approach, Bayesian neural networks, and radial basis neural networks would predict.

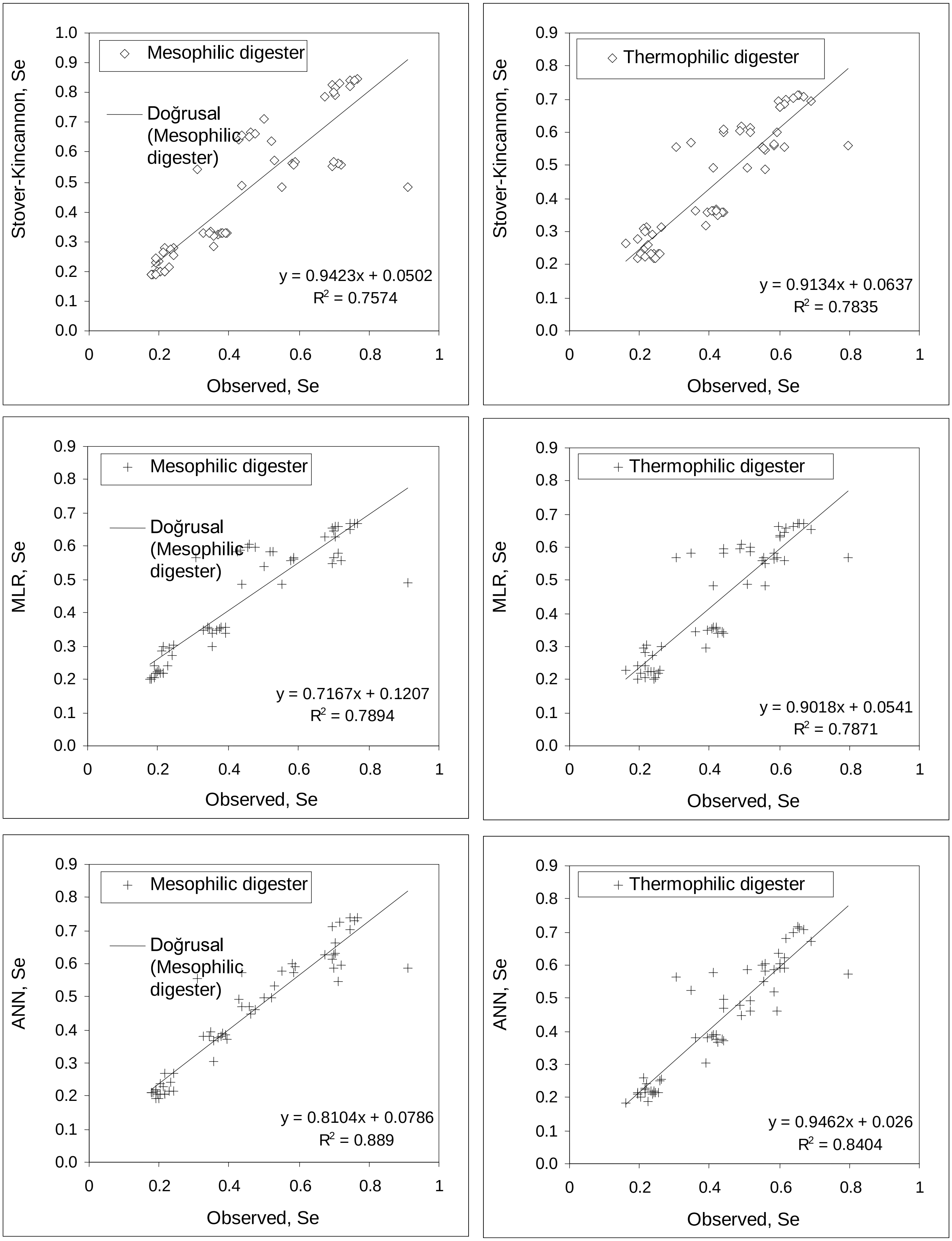

Figure 7. Plot of observed and predicted Se values by (a) Stover-Kincannon model, (b) MLR model, and (c) ANN model for mesophilic and thermophilic digesters.

REFERENCES

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000a). Artificial neural networks in hydrology I: preliminary concepts. J. Hydrol. Eng. ASCE 5(2):115-123.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000b). Artificial neural networks in hydrology II: hydrologic applications. J. Hydrol. Eng., ASCE 5(2):124-137.

Bayazıt M, Oguz B (1998). Probability and Statistics for Engineers, Birsen Publishing House, Istanbul, Turkey, p. 159.

Budcema M, Sacco PL (2000). Feedforward networks in financial predictions: the future that modifies the present. Expert Syst. Appl. 17(3):149-169.

Dawson WC, Wilby R (1998). An artificial neural network approach to rainfall-runoff modeling. Hydrol. Sci. J. 43(1):47-66.

Fang YC, Wu BW (2007). Neural network application for thermal image recognition on flow-resolution objects. J. Opt. A: Pure Appl. Opt. 9(2): 134-144.

Haykin S (1998). Neural Networks - A Comprehensive Foundation (2nd. ed.). Prentice-Hall, Upper Saddle River, NJ.

Maier HR, Dandy GC (2000). Neural networks for the prediction and forecasting of water resources variables: A review of modeling issues and applications. Environ. Mod. Soft. 15:101-124.

McCulloch WS, Pitts WH (1943). A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. 5:115-133.

Minns AW, Hall MJ (1996). Artificial neural networks as rainfall-runoff models. Hydrol. Sci. J. 41(3):399-417.

Minsky M, Papert S (1969). Perceptrons. MIT Press, Cambridge, MA.

Mu Y, Yu HQ (2007). Simulation of biological hydrogen production in aUASB reactor using neural network and genetic algorithm. Int. J. Hydrogen Energy 32:3308-3314.

Papik K, Molnar B, Schaefer R, Dombovari Z, Tulassay Z, Feher J (1998). Application of neural networks in medicine - A review. Med. Sci. Monitor 4(3):538-546.

Rumelhart D, Hinton G, Williams R (1986). Learning internal representations by error propagation, In Parallel Distributed Processing, chapter 8, MIT Press, Cambridge, MA.

Singh KP, Basant N, Malik A, Jain G (2010). Modeling the performance of up-flow anaerobic sludge blanket reactor based wastewater treatment plant using linear and nonlinear approaches-A case study. Anal. Chem. Acta. 658:1-11.

Singhal A, Gomes J, Praveen VV, Ramachandran KB (1998). Axial dispersion model for upflow anaerobic sludge blanket reactors. Biotechnol. Prog, 14:645-648.

Sinha S, Bose P, Jawed M, John S (2002). Application of neural network for simulation of upflow anaerobic sludge blanket (UASB) reactor performance. Biotechnol. Bioeng. 77:806-814.

Tay JH, Show KY, Jeyaseelan S (1996). Effects of media characteristics on performance of upflow anaerobics packed-bed reactors. J. Environ. Eng. 122:469-476.

Tokar SA, Johnson PA (1999). Rainfall-runoff modeling using artificial neural networks. J. Hydrol. Eng. 4(3):232-239.

Wu MM, Hickey RF (1997). Dynamic model for UASB reactor including reactor hydraulics, reaction, and diffusion. J. Environ. Eng-ASCE. 123:244-252.

Yilmaz T, Seckin G, Yuceer A (2010). Modeling of effluent COD in UAF reactor treating cyanide containing wastewater using artificial neural network approaches. Adv. Eng. Softw. doi:10.1016/j.advengsoft.2010.04.002.

Yilmaz T, Yuceer A, Basibuyuk M (2008). A comparison of the performance of mesophilic and thermophilic anaerobic filters treating paper-mill wastewater. Bioresour. Technol. 99(1):156-163.

Yuanyou X, Yanming X, Ruigeng Z (1997). An engineering geology evaluation method based on an artificial neural network and its application. Eng. Geol. 47(1-2):149-156.