Desalination and Water Treatment
www.deswater.com
doi: 10.1080/19443994.2014.902336

54 (2015) 2496–2503
May

Taylor & Francis
Taylor & Francis Group

Modeling nitrate concentrations in a moving bed sequencing batch biofilm reactor using an artificial neural network technique

Hakan Dulkadiroglu[a,*], Galip Seckin[b], Derin Orhon[a]

[a]*ENVIS Energy and Environmental Systems R&D Ltd, ITU Ayazağa Yerleşkesi, Ari Teknokent, Ari 1 Binasi, No: 16, 34469 Maslak, Istanbul, Turkey, Tel. +90 212 286 03 03; Fax: +90 212 286 03 00; email: hdulkadiroglu@isnet.net.tr*
[b]*Faculty of Engineering and Architecture, Environmental Engineering Department, Cukurova University, Balcali, Adana, Turkey*

Received 13 November 2013; Accepted 1 March 2014

ABSTRACT

In this study, the performance data of a moving-bed sequencing batch biofilm reactor (MBSBBR) treating synthetic wastewater were simulated using multi-layer perceptron neural-network technique. Multi-linear regression (MLR) technique is also used for a comparison. The performance of MBSBBR was evaluated using these models for a set of experimental results obtained from a model reactor operated with different cycle times and temperatures. The experimental data were retrieved from a previous reported work. Operational time, temperature, ammonium nitrogen, and pH were used as inputs for modeling, whereas nitrate concentration was the output variable. The results of the models were compared using statistical criteria, such as mean square error, mean absolute error, mean absolute relative error, and determination coefficient (R^2). The results showed that the multi-layer perceptron neural-network produced more accurate results than those of MLR, although the latter gave reasonable results.

Keywords: Artificial neural network; Biodegradation; Modeling; Multi-layer perceptron; Moving bed sequencing batch biofilm reactor; Nitrification

1. Introduction

Suspended growth and biofilm systems, such as different activated sludge and biofilter configurations, although widely used as successful biological treatment schemes for domestic and industrial wastewater, have a number of inherent limitations. In this context, moving bed biofilm reactors (MBBR), holding carrier elements freely moving in the reactor, have been developed as one of the most attractive hybrid systems [1,2]. Sequencing batch reactor (SBR), is another highly successful biological treatment alternative, widely studied in the last two decades [3–5]. Recently, it was suggested that MBBRs could be operated in a sequencing batch mode, in order to benefit from the advantages of both processes.

Full understanding of system performance in biological systems is only possible by means of an accurate interpretation of the complex set of biochemical reactions taking place in the reactor. This requires modeling, which defines process kinetics and stoichiometry for selected model components [6,7]. Model calibration of experimental data and assessment of process kinetics is now successfully applied to suspended growth systems including SBRs [8] and membrane bioreactors [9]. Similar modeling approaches

*Corresponding author.

have also been attempted for moving bed biofilm reactors: Plattes et al. [10] proposed a model including biofilm attachment and detachment using Activated Sludge Model No.1 (ASM1) and Monod kinetics for the dynamic simulation of a pilot-scale moving bed bioreactor. Sin et al. [11] defined a mathematical model integrating hydraulics, biofilm, and microbial conversion processes to evaluate nitrification in a moving bed biofilter. A similar model was also proposed by Lin [12] and calibrated with the experimental data derived from a pilot-scale moving-fixed bed biofilm reactor for the kinetics and nitrogen and carbon removal. Ferrai et al. [13] utilized respirometric techniques for the evaluation of kinetic and stoichiometric coefficients for a similar reactor operated with domestic sewage. While these approaches were claimed to be successful for the specific cases selected for investigation, the microbial structure of the moving bed biofilm reactor is too complex to be described in a single mathematical model, mainly because of severe identifiability problems involved in model calibration [14]. Therefore, there is a significant need for using auxiliary techniques, such as the artificial neural network (ANN) technique adopted in the study to interpret system performance [15].

In previous works, a series of experiments were conducted using Moving Bed Sequencing Batch Biofilm Reactor (MBSBBR) in order to investigate its nitrification performance [16,17]. Dulkadiroglu et al. [17] measured pH, ammonium (NH_4-N in mg/L), oxidized nitrogen (NO_X-N in mg/L), temperature (T in ℃), operational time (OT in min), chemical oxygen demand (COD in mg/L), volatile suspended solid (VSS in mg/L), and total suspended solid (TSS in mg/L) in their previous work. Among these parameters, pH, T, OT, and NH_4-N are the known (input) parameters, whereas NO_X-N, COD, VSS, and TSS are unknown (output) parameters. There are two ways to determine these unknown parameters: (i) they can be measured as done by Dulkadiroglu et al. [17]; (ii) they can be modeled. Accurate estimation of these unknown (output) parameters will lead to time conservation and cost reduction in the operation instead of measuring these parameters repeatedly. For this purpose, in this study, a new ANN model is proposed in the estimation of NO_X-N concentrations in a MBSBBR.

There are several features in ANN that distinguish it from the empirical models: First, neural networks have flexible non-linear function mapping capability that can approximate any continuous measurable function with arbitrarily desired accuracy, whereas most of the commonly used empirical models do not have this property. Second, being non-parametric and data-driven, neural networks impose few prior assumptions on the underlying process from which data are generated. Also, high computation rate, learning ability through pattern presentation, prediction of unknown patterns, and flexibility affronts for noisy patterns are other advantages of using ANNs. Neural networks are composed of neurons as basic units. Each neuron receives input data and processes the input data and transforms them into output forms. The input forms may be pure data or the input results of other neurons, and the output forms may be the results of the final process or the input data of other neurons [18]. In current work, the output variable was the concentration of oxidized nitrogen (NO_X-N), while the input variables were the operational time (OT), temperature (T), ammonium nitrogen (NH_4-N), and pH. Because of lack of some data, the other parameters, such as COD, TSS, and VSS could not been modeled in current work. Multi-linear regression (MLR) technique is also used in the estimation of NO_X-N for a comparison. The predictive ability of the proposed models were assessed using four criteria, namely mean square error (MSE), mean absolute error (MAE), mean absolute relative error (MARE), and determination coefficient (R^2).

2. Materials and methods

2.1. Experimental

As experimental setup and procedure are given in detail elsewhere [17], they are just summarized herein.

The experiments were carried out in a lab-scale, cylindrical Plexiglas MBSBBR with a working volume of 10 L, placed into an incubator in order to operate at constant temperature. Wastewater was fed and discharged by peristaltic pumps controlled by timers. The react phase was selected as 420 min for 8 h cycle time and 300 min for 6 h cycle time. Fill- and draw-discharge phases were kept constant at 30 min each. MBSBBR was fed with a synthetic wastewater with COD concentration of 400 mg/L and ammonia nitrogen concentration of 40 mg NH_4-N/L, approximating domestic sewage characteristics. Macro- and micronutrients were added at necessary amounts. Reactor performance was monitored by NH_4-N and NO_X-N measurements. After an acclimation period of the biomass to loadings and temperatures, NH_4-N and NO_X-N variations within a given cycle were determined for each set of experiments and these in-cycle measurements were repeated at different times to have a reliable assessment of the system performance at the selected operation condition. NH_4-N experiments were performed as defined in Standard

Methods [19]. NO_X-N concentrations were determined using a ChemLab Autoanalyzer. Measurements indicated that NO_2-N concentrations were below the detection limits and negligible in all experiments. Therefore, it was safe to assume and express the total oxidized nitrogen (NO_X-N) levels as NO_3-N. pH was monitored using a Hanna HI8711E pH controller. All the analyses were conducted in triplicate and the average values were used for evaluation of the experiments. The variation among triplicate measurements usually remained below the analytical precision prescribed in the corresponding methods.

2.2. Modeling

ANN has the ability to learn from examples, recognize a pattern in a group of data, adapt solutions over time, and process information rapidly. The application of ANNs to issues related to wastewater treatment and water resources conservation is rapidly gaining popularity due to their immense power and potential in the mapping of non-linear system data. In the context of hydrological forecasting, recent studies have reported that ANN technique may offer a promising alternative for rainfall–runoff modeling [20], streamflow prediction [21,22], suspension of sediments [23], water resources [24], and reservoir inflow forecasting [25]. The variation in the characteristics of a water resource system may be non-linear and multivariate, and the variables involved may have complex interrelationships. For most cases, ANNs provide more reliable estimates for dependent variables of concern. The processes that involve several parameters are easily amenable to neuro-computing. Among the many ANN structures that have been studied, the most widely used network structure is the multilayer perceptron (MLP) network. An ANN consists of a number of data processing elements called neurons or nodes, which are grouped in layers. The input layer of neurons receives the input vector and transmits the information to the next layer with the help of cross-connections. In the current study, a MLP modeling technique was applied.

2.3. MLP neural-network model

A MLP distinguishes itself by the presence of one or more hidden layers, whose computation nodes are correspondingly called "hidden neurons of hidden units." The function of hidden neurons is to intervene between the external input and the network output in some useful manner. By adding one or more hidden layers, the network is enabled to extract higher order

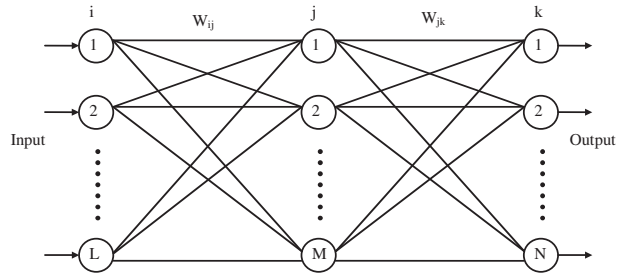

Fig. 1. MLP structure.

statistics. In a rather loose sense, the network acquires a global perspective despite its local connectivity due to the extra set of synaptic connections and the extra dimension of NN interconnections. Detailed theoretical information about MLP can be found in Haykin [26].

MLP network used in current study, is seen in Fig. 1. Index k is referred to the individual output layer neurons, the indices i and j refer to the input neurons and the hidden layer neurons, respectively. w_{ij} and w_{jk} represent the connection weights between the hidden-input layer and hidden-output layer, respectively. A hidden-layer neuron produces the following (Eq. (1)) as output;

$$h_j = f\left(\sum_{i=1}^{n} w_{ij}x_i + b_j\right) \tag{1}$$

while an output-layer neuron produces the following (Eq. (2)) as output;

$$y_k = f\left(\sum_{j=1}^{n} w_{jk}h_j + b_k\right) \tag{2}$$

where h_j is the output of the jth neuron in the hidden layer; x_i is the input of the ith neuron in the input layer; y_k is the output of the kth neuron in the output layer; b_j and b_k are the threshold values, also called the bias, associated with the hidden and output nodes, respectively; and f denotes the activation function. Each neuron multiplies every input by its interconnection weight, sums the product, and then passes the sum through a transfer function to produce its result. This transfer function is usually a steadily increasing S-shaped curve, called a sigmoid function.

The MLP can have more than one hidden layer. However, theoretical works have shown that a single hidden layer is sufficient for MLP to approximate any

complex non-linear function [27] Therefore, in this study, one-hidden-layer MLP is used. Throughout all MLP simulations, the adaptive learning rates are used to speed up training. The numbers of hidden layer neurons are found using simple trial-and-error method in all applications. The sigmoid and linear functions are used for the activation functions of the hidden and output nodes, respectively.

Some of the recent studies have reported that the performance of MLP was superior to conventional statistical and stochastic methods [22,23]. Multi-layer perceptions can get trapped in a local minimum when they try to find the global minimum of the error surface. Maier and Dandy [28] summarized the methods used in the published literature to overcome the local minima problem, such as training a number of networks, starting with different initial weights, an online training mode to help the network escape local minima, inclusion of the addition of a random noise, and employment of second-order schemes, such as Newton–Raphson and Levenberg–Marquardt algorithms, or global methods, such as stochastic gradient algorithms and simulated annealing. Other ANN methods, such as conjugate gradient algorithms, the radial basis function, the cascade correlation algorithm, and recurrent neural networks, were briefly explained in the report by the ASCE Task Committee on Application of ANNs in Hydrology [29].

2.4. Levenberg–Marquardt algorithm

In the present study, the Levenberg–Marquardt algorithm was employed because this algorithm is more powerful than the conventional gradient descent techniques [30]. The Levenberg–Marquardt algorithm is an approximation of Newton's method and is very efficient for training networks with up to a few hundred weights. Although the computational load of the Levenberg–Marquardt algorithm is greater than that of other techniques, this is compensated by the increased efficiency and much better precision in results. In many cases, the Levenberg–Marquardt algorithm was found to converge when other back-propagation techniques diverged [30].

2.5. Determination of an appropriate ANN model

Determining an appropriate architecture of a neural network for a particular problem is an important issue as the network topology directly affects its computational complexity and its generalization capability. The MLP model with one hidden layer can approximate any complex non-linear function provided a sufficient amount of hidden-layer neurons are available. Indeed, many experimental results seem to confirm that one hidden layer may be enough for most forecasting problems [31]. Therefore, in this study, one hidden-layer MLP model was used. Generally, the number of hidden-layer neurons is determined by a trial-and-error method. A common strategy for finding the optimum number of hidden-layer neurons is to start with a small number of neurons and increase their number, while monitoring the performance criteria until no significant improvement is observed [32].

2.6. MLR model

If it is assumed that the dependent variable Y is effected by m independent variables X_1, X_2, …, X_m and a linear equation is selected for the relation among them, the regression Eq. (3) of Y can be written as:

$$y = a + b_1x_1 + b_2x_2 + \ldots + b_mx_m \tag{3}$$

y in this equation shows the expected value of the variable Y, when the independent variables take the values $X_1 = x_1$, $X_2 = x_2$, …, and $X_m = x_m$.

The regression coefficients a, b_1, b_2, …, and b_m are evaluated, similar to simple regression, by minimizing the sum of the e_{yi} distances of observation points from the plane expressed by the regression Eq. (4):

$$\sum_{i=1}^{N} e_{yi}^2 = \sum_{i=1}^{N} (y_i - a - b_1x_{1i} - b_2x_{2i} - b_mx_{mi})^2 \tag{4}$$

In this study, the a coefficient was assumed to be 0, and the coefficients b_1, b_2, …, b_m were determined using least squares method.

3. Results and discussion

In this study, the dependent variable was the concentration of oxidized nitrogen (NO_X-N), while the independent variables were the operational time (OT), temperature (T), ammonium nitrogen (NH_4-N), and pH. The minimum and maximum values for the model variables are provided in Table 1. It should be noted that the synthetic wastewater was prepared to contain 40 mg N/L of ammonia nitrogen. However, the evaluation was based upon slightly lower values measured in the reactor volume at the start of the experiments. Furthermore, it is well known that nitrification is affected by changes in the temperature. In this context, the model evaluation was not limited to

Table 1
The minimum and maximum values of the input parameters

Parameters	Training data-set		Testing data-set	
	Min	Max	Min	Max
OT (minutes)	0	420	30	330
pH	7.09	7.63	7.18	7.51
NH_4-N (mg/L)	0	36	0.2	28.8
T (℃)	10	20	15	15

15℃, which was recorded as the temperature of the experiments, but they also covered the range of 10–20℃. There are no acceptable rules to determine the optimum size of the training data-set. The networks are not very sensitive to the number of training data, but very sensitive to the number of testing data. Attempts at reducing the training data size resulted in poor generalization capabilities in the testing phase. A training sample and a test sample are typically required for building on ANN forecaster. The training sample is used for ANN model development and the test sample is adopted for evaluating the forecasting ability of the model. There is no general rule to the problem of division of the data into training and data-sets. Several factors, such as the problem of structure, the data type, and the size of the available data should also be considered in making the decision. It is critical to have both the training and test sets representative of the population or underlying mechanism. This has particular importance for time-series forecasting problems. Inappropriate separation of the training and test sets will affect the selection of optimal ANN structure and the evaluation of the forecasting performance [33]. In this study, these data were randomly divided into two independent parts. To overcome some extrapolation difficulties in prediction of extreme values, minimum and maximum values of parameters used in modeling were set in training data. Therefore, the available data-set was partitioned into a training set and a testing set with 75 and 25% of the available experimental measurements selected for training and testing phases, respectively. Maier and Dandy [34] pointed out that "It is common practice to split the available data into two sub-sets; a training set and an independent validation set." This process was frequently used in the related literature [35–38]. Before the training phase of the network, both input and output variables were normalized within the range of 0.1–0.9 as follows (Eq. (5)):

$$x_i = 0.8\frac{(x - x_{\min})}{(x_{\max} - x_{\min})} + 0.1 \tag{5}$$

where x_i is the normalized value of a certain parameter, x is the measured value for this parameter, $x_{\min}$ and $x_{\max}$ are the minimum and maximum values in the database for this parameter, respectively.

While training the ANN model, an overfitting problem may appear. The best way to avoid overfitting is to use lots of training data. For noise-free data, if there are at least five times as many training cases as there are weights in the network, we are unlikely to suffer from overfitting. The other way to avoid overfitting problem is to use different feed-forward neural network structures [39].

For all created neural networks, the general structure of input, one hidden and one output layer were used. In order to determine the optimal architecture, several neural networks were trained with different iteration numbers (epoch) and number of nodes in the hidden layer. In this study, the tangent sigmoid, logarithmic sigmoid, and pure linear transfer functions were tried as activation functions for hidden and output layer neurons to determine the best network model. Accordingly, the activation functions of the hidden and output layer were found using simple trial-and-error method in all the applications herein. When the log sig was applied, the inputs and the outputs were normalized within the range of 0–1. The most accurate estimations of the ANNs were obtained with logarithmic sigmoid transfer function. The best MLP results were obtained from the ANN (4, 6, 1) model using the logarithmic sigmoid activation functions for both hidden and output layer neurons, respectively.

The MSE, MARE, and R^2 (determination coefficient) values of ANNs for both training and testing phases are given in Table 2. The MSE, MAE, and MARE are defined as follows (Eqs. (6)–(8)):

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(Yi_{\text{observed}} - Yi_{\text{predicted}})^2 \tag{6}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|Yi_{\text{observed}} - Yi_{\text{predicted}}\right| \tag{7}$$

$$\mathrm{MARE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{Yi_{\text{observed}} - Yi_{\text{predicted}}}{Yi_{\text{observed}}}\right| \times 100 \tag{8}$$

Table 2
The training and testing performances of the MLP and MLR models

	Training phase				Testing phase			
Method	MARE (%)	MAE (m^3/s)	MSE $(m^3/s)^2$	R^2	MARE (%)	MAE (m^3/s)	MSE $(m^3/s)^2$	R^2
MLP	25.089	0.600	0.596	0.992	25.275	0.639	0.592	0.990
MLR	58.100	1.712	4.547	0.940	26.961	1.006	1.593	0.985

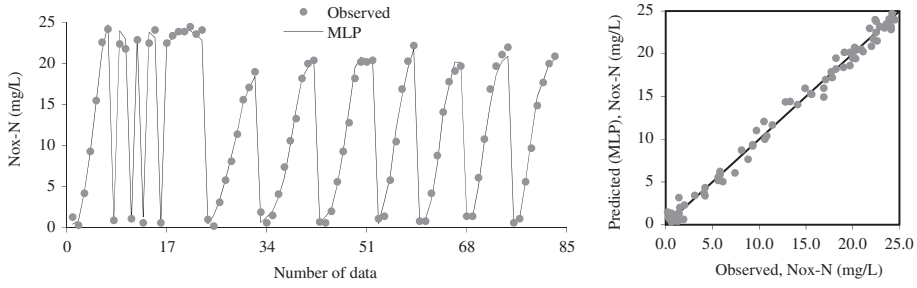

Fig. 2. Comparison between observed and predicted nitrate concentrations in training phase using MLP.

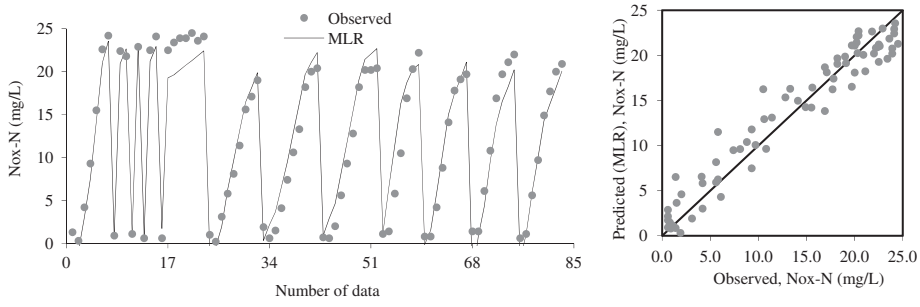

Fig. 3. Comparison between observed and predicted nitrate concentrations in training phase using MLR.

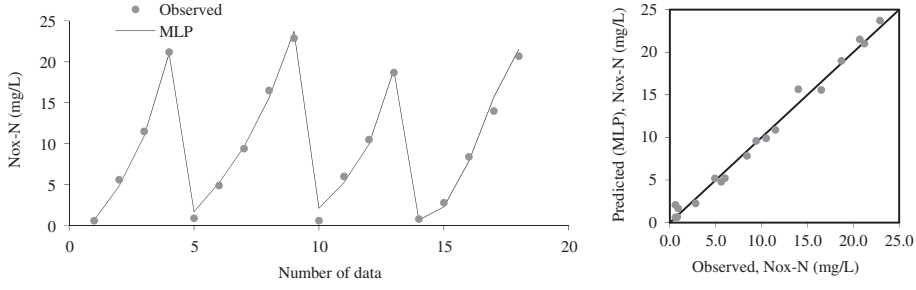

Fig. 4. Comparison between observed and predicted nitrate concentrations in testing phase using MLP.

In Eqs. (6)–(8), Y denotes nitrate concentrations and N is the total number of data.

The MLP and MLR models were trained, tested, and then the results were compared by means of MSE, MAE, MARE, and R^2 statistics as shown in Table 2. As seen from Table 2, both the MLP and the MLR have the capability of modeling nitrate concentration, although the former gave better results.

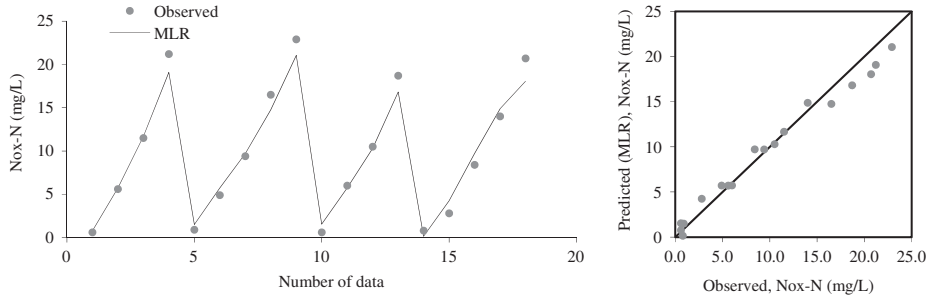

Fig. 5. Comparison between observed and predicted nitrate concentrations in testing phase using MLR.

The performances of the MLP and MLR models analyzed, herein, are shown in Figs. 2 and 3 for training and Figs. 4 and 5 for testing phases, respectively. As can be seen from these figures, the MLP produced highly more accurate results than those of MLR in the estimation of nitrate concentrations for training phase, whereas they gave similar results in the testing phase.

4. Conclusions

The abilities of MLP neural network and MLR technique in the estimation of the nitrate concentration as the product of nitrification in the MBSBBR were assessed in this paper by comparing the results with observed concentrations of nitrate. From the results obtained, both the MLP with a Levenberg–Marquardt algorithm and MLR technique used in current study appear to be useful tools for prediction of the nitrate concentrations in the MBSBBR, although the former produced more accurate estimations. The results of this study show that ANN technique can be used in the design and operation of an MBSBBR system for prediction of optimum cycle time required for nitrification for a known influent NH_4-N concentration and different temperatures.

References

[1] H. Ødegaard, B. Rusten, T. Westrum, A new moving bed biofilm reactor—Applications and results, Water Sci. Technol. 29 (2004) 157–165.
[2] A. Tawfik, N. Badr, E. Taleb, W. El-Senousy, Sewage treatment in an up-flow anaerobic sponge reactor followed by moving bed biofilm reactor based on polyurethane carrier material, Desalin. Water Treat. 37 (2012) 350–358.
[3] P.A. Wilderer, R.L. Irvine, M.C. Goronsky, Sequencing Batch Reactor Technology, Scietific and Technical Report No. 10, IWA Publishing, London, 2001.
[4] N. Artan, P.A. Wilderer, D. Orhon, E. Morgenroth, N. Ozgur, The mechanism and design of sequencing batch reactor systems for nutrient removal—The state of the art, Water Sci. Technol. 43 (2001) 53–60.
[5] N. Artan, N.O. Yagci, S.R. Artan, D. Orhon, Design of sequencing batch reactors for biological nitrogen removal from high strength wastewaters, J. Environ. Sci. Health, Part A 38 (2003) 2125–2134.
[6] M. Henze, C.P.L. Grady, W. Gujer, G.v.R Marais, T. Matsuo, A general model for single sludge wastewater treatment systems, Water Res. 21 (1987) 742–748.
[7] O. Karahan-Gul, D. Orhon, M.C.M. van Loosdrecht, Modification of activated sludge Model No. 3 considering direct growth on primary substrate, Water Sci. Technol. 47 (2003) 219–225.
[8] S. Murat, G. Insel, N. Artan, D. Orhon, Effect of temperature on the nitrogen removal performance of a sequencing batch reactor trating tannery wastewater, Water Sci. Technol. 48 (2003) 319–326.
[9] M. Sarioglu, G. Insel, N. Artan, D. Orhon, Model evaluation of simultaneous nitrification and denitrification in a membrane bioreactor operated without an anoxic reactor, J. Membr. Sci. 337 (2009) 17–27.
[10] M. Plattes, E. Henry, P.M. Schosseler, A. Weidenhaupt, Modelling and dynamic simulation of a moving bed bioreactor for the treatment of municipal wastewater, Biochem. Eng. J. 32 (2006) 61–68.
[11] G. Sin, J. Weijma, H. Spanjers, I. Nopens, Dynamic model development and validation for a nitrifying moving bed biofilter: Effect of temperature and influent load on the performance, Process Biochem. 43 (2008) 384–397.
[12] Y.H. Lin, Kinetics of nitrogen and carbon removal in a moving-fixed bed biofilm reactor, Appl. Math. Modell. 32 (2008) 2360–2377.
[13] M. Ferrai, G. Guglielmi, G. Andreottola, Modelling respirometric tests for the assessment of kinetic and stoichiometric parameters on MBBR biofilm for municipal wastewater treatment, Environ. Modell. Software 25 (2010) 626–632.
[14] G. Insel, D. Orhon, P.A. Vanrolleghem, Identification and modelling of aerobic hydrolysis—application of optimal experimental design, J. Chem. Technol. Biotechnol. 78 (2003) 437–445.
[15] H. Liu, M. Huang, C.K. Yoo, A fuzzy neural network-based soft sensor for modeling nutrient removal mechanism in a full-scale wastewater treatment system, Desalin. Water Treat. 51 (2013) 6184–6193.
[16] H. Dulkadiroglu, D. Orhon, E.U. Cokgor, Performance of a moving bed sequencing batch reactor for carbon and nutrient removal, Fresenius Environ. Bull. 13 (2004) 1028–1032.

[17] H. Dulkadiroglu, E.U. Cokgor, N. Artan, D. Orhon, The effect of temperature and sludge age on COD removal and nitrification in a moving bed sequencing batch reactor, Water Sci. Technol. 51 (2005) 95–103.
[18] S. Kim, H. Kim, Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling, J. Hydrol. (Amsterdam) 351 (2008) 299–317.
[19] L.S. Clesceri, A.E. Greenberg, A.D. Eaton, Standard Methods for the Examination of Water and Wastewater, 20th ed., American Public Health Association/American Water Works Association/Water Environment Federation, Washington, DC, 1998.
[20] G.-F. Lin, L.-H. Chen, A non-linear rainfall-runoff model using radial basis function network, J. Hydrol. 289 (2004) 1–8.
[21] H. Raman, N. Sunilkumar, Multivariate modelling of water resources time series using artificial neural networks, Hydrological Sci. J. 40 (1995) 145–163.
[22] O. Kisi, River flow modelling using artificial neural networks, J. Hydrologic Eng. 9 (2004) 60–63.
[23] O. Kisi, Multi-layer perceptions with Levenberg–Marquardt training algorithm for suspended sediment concentration prediction and estimation, Hydrological Sci. J. 49 (2004) 1025–1040.
[24] M. Cobaner, T. Haktanir, O. Kisi, Prediction of hydropower energy using ANN for the feasibility of hydropower plant installation to an existing irrigation dam, Water Resour. Manage. 22 (2008) 757–774.
[25] P. Coulibaly, M. Hache, V. Fortin, B. Bobeé, Improving daily reservoir inflow forecasts with model combination, J. Hydrologic Eng. 10 (2005) 91–99.
[26] S. Haykin, Neural networks—A comprehensive foundation, 2nd ed., Prentice—Hall, Upper Saddle River, NJ, 1998, pp. 26–32.
[27] H.R. Maier, G.C. Dandy, The use of artificial neural networks for the prediction of water quality parameters, Water Resour. Res. 32 (1996) 1013–1022.
[28] H.R. Maier, G.C. Dandy, Neural networks for the prediction and forecasting water resources variables: A review of modeling issues and applications, Environ. Modell. Software 15 (2000) 101–124.
[29] ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, Artificial neural networks in hydrology—I: Preliminary concepts, J. Hydrologic Eng. 5 (2000) 115–123.
[30] M.T. Hagan, M.B. Menhaj, Training feedforward networks with the Marquardt algorithm, IEEE Trans. Neural Networks 5 (1994) 989–993.
[31] P. Coulibaly, F. Anctil, B. Bobeé, Prévision hydrologique par réseaux de neurones artificiels: état de l'art [Hydrological forecasting using artificial neural network: State of the art], Can. J. Civil Eng. 26 (1999) 293–304.
[32] A.T.C. Goh, Back-propagation neural networks for modeling complex systems, Artifi. Intell. Eng. 9 (1995) 143–151.
[33] G. Zhang, B.E. Patuwo, M.Y. Hu, Forecasting with artificial neural networks, Int. J. Forecasting 14 (1998) 35–62.
[34] H.R. Maier, G.C. Dandy, Neural networks for the prediction and forecasting of water resources variables: A review of modeling issues and application, Environ. Modell. Software 15 (2000) 101–124.
[35] S.K. Jain, D. Das, D.K. Srivastava, Application of ANN for reservoir inflow prediction and operation, J. Water Resour. Plann. Manage. ASCE 125 (1999) 263–271.
[36] K.P. Sudheer, Knowledge extraction from trained neural network river flow models, J. Hydrologic Eng. 10 (2005) 264–269.
[37] O. Kisi, Suspended sediment estimation using neuro-fuzzy and neural network approaches. Hydrological Sci. J. 50 (2005) 683–696.
[38] M. Alp, H.K. Cigizoglu, Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data, Environ. Modell. Software 22 (2007) 2–13.
[39] K.P. Sudheer, A.K. Gosain, D. Mohana Rangan, S.M. Saheb, Modelling evaporation using an artificial neural network algorithm. Hydrological Processes 16 (2002) 3189–3202.