Desalination and Water Treatment
www.deswater.com
1944-3994/1944-3986 © 2012 Desalination Publications. All rights reserved
doi: 10/5004/dwt.2012.2847

Taylor & Francis
Taylor & Francis Group

40 (2012) 174–182
February

Modeling Fentonic advanced oxidation process decolorization of Direct Red 16 using artificial neural network technique

Javad Saien[a,*], Ali Reza Soleymani[a], Hossein Bayat[b]

[a]*Department of Applied Chemistry, Bu-Ali Sina University, Hamedan 65174, Iran*
Tel./Fax: +98 811 8257407; email: jsaien@yahoo.com
[b]*Department of Soil Science Faculty of Agriculture, Bu-Ali Sina University, Hamedan 65174, Iran*

Received 24 March 2011; Accepted 13 September 2011

ABSTRACT

The present work has focused on the modeling of C.I. Direct Red 16 (DR16) decolorization using Fentonic reagents in a batch reactor. The reactor was equipped with an air bubbling for mixing and a water-flow coil for temperature regulating. Dye concentration was analyzed by measuring its absorbance at λ_{max} = 526 nm. An artificial neural network (ANN) model was developed to predict the behavior of the process. Six operational parameters and decolorization efficiency were employed as inputs and output of the network, respectively. A three layer feed-forward network with back-propagation algorithm was developed. Application of 10 neurons in the hidden layer and 300 iterations for the network calibration prevents overfitting by the model. The K-fold cross-validation method was employed for performance evaluation of the developed ANN model. The results showed high correlation coefficient (R^2 = 0.9984) and low mean square error (MSE = 1.56×10^{-4}) for testing data. Sensitivity analysis indicates the order of operational parameters relative importance on the network response as: pH ≈ time > $[H_2O_2]$ > [Fe(II)] > $[DR16]_0$ > temperature.

Keywords: Fenton process; Direct Red 16; ANN modeling; Feed forward; Cross-validation; Sensitivity analysis

1. Introduction

Artificial neural networks (ANNs) are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the connections between elements largely determine the network function. They consist of a large number of processing elements with their interconnections [1]. ANNs learn by examples of data inputs and outputs presented to them so that the subtle functional relationships among the data are captured, even if the underlying relationships are unknown or the physical meaning is difficult to explain [2]. In contrast to most traditional empirical and statistical methods, ANNs do not need prior knowledge about the nature of the relationships among the data. This is one of the main benefits of ANNs when compared with most empirical and statistical methods [3]. Feed-forward ANNs are most often used to map input-output relationships [4]. A diagram of a multilayer feed-forward ANN is given in Fig. 1. This figure shows that these ANNs are organized in layers that contain neurons (also called nodes). The number of neurons in input and output layers corresponds to the number of input and output variables, *J* and *L*, respectively. The number of "hidden" neurons, *K*, can be chosen freely, and determines the complexity that can be modeled [5].

*Corresponding author.

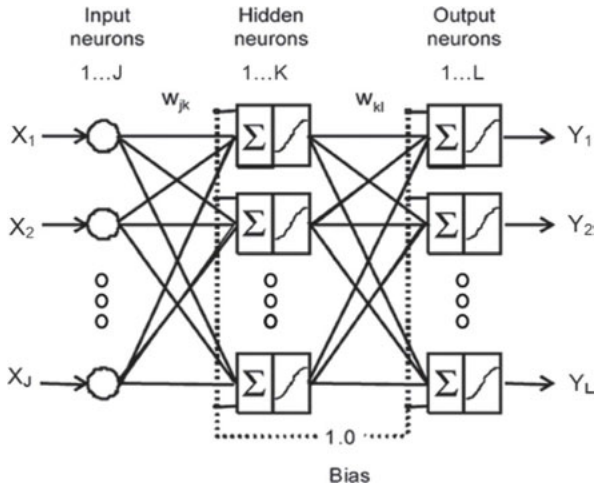

Fig. 1. General overview of a feed-forward artificial neural network [16].

For treatment of pollutants, advanced oxidation processes (AOPs) have been proposed and employed since the 1990s. In the so-called AOPs, highly reactive species, mainly hydroxyl radicals, are used as primary oxidants which are a highly reactive and non-selective oxidant, able to oxidize organics [6,7]. It has been documented that between 1% and 20% of the total world production of azo dyes is lost during the dyeing processes and providing major wastewater pollution in wastewaters [8]. The presence of even small amounts of dyes (below 1.0 mg l–1) is clearly visible and influences the water quality considerably [9,10]. Among AOPs, oxidation with Fenton's reagent has been considered to be a promising and attractive treatment technology for decolorization and effective degradation of textile dyes [11–13]. In addition, it has been applied to a wide range of organic pollutants, because of its simplicity and ease of implementation, running under mild conditions of temperature and pressure [14]. The Fenton reaction involves several reactions, which can be described by Eqs. (1)–(6) [15]:

$$Fe(II) + H_2O_2 \rightarrow Fe(III) + {}^{\bullet}OH + {}^{-}OH \quad (1)$$

$$Fe(II) + {}^{\bullet}OH \rightarrow Fe(III) + {}^{-}OH \quad (2)$$

$${}^{\bullet}OH + \text{organics} \rightarrow \text{products} \quad (3)$$

$${}^{\bullet}OH + H_2O_2 \rightarrow H_2O + HO_2^{\bullet} \quad (4)$$

$${}^{\bullet}OH + {}^{\bullet}OH \rightarrow H_2O_2 \quad (5)$$

$$Fe(III) + H_2O_2 \rightarrow Fe(II) + H^+ + HO_2^{\bullet} \quad (6)$$

These reactions show that hydrogen peroxide may be consumed when it reacts with Fe(II), as shown in Eq. (1), producing hydroxyl radicals that will degrade organic compounds through Eq. (3). This catalytic reaction is propagated from Fe(II) regeneration, which mainly takes place as a result of the reduction of Fe(III) species with H_2O_2 [16].

Wastewater treatment by applying AOPs is, in general, quite complex. Since the process depends on several factors, the modeling of these processes involves many problems, i.e., we are dealing with a multivariate system. It is evident that these problems cannot be solved by simple linear multivariate correlation. In this case applying ANNs can be a good alternative. Following are some examples of recent done works in the case of ANN modeling of AOPs in water and wastewater treatments.

Monteagudo et al. [17] applied ANN for prediction of the values of kinetic degradation rate constants for homogenous ferrioxalate-assisted solar photo-Fenton degradation of Orange II aqueous solutions. Duran et al. [18] described the application of ANN for simulation of photo-Fenton degradation of Reactive Blue 4 in water. Khataee and Mirzajani [19] used ANN for modeling of key factors in the UV/$S_2O_8^{2-}$oxidation of C. I. Basic Blue 3. In another case, Salari et al. [20] have applied ANNs for modeling of the treatment of wastewater contaminated with methyl *tert*-butyl ether (MTBE) by UV/H_2O_2 process.

Modeling of a Fentonic decolorization process is a rather new study and not much information is available in literatures. The oxidation efficiency of the Fenton's process for a given wastewater or parent substrate concentration depends on several variables, namely pH of the reaction medium, temperature, time of the reaction (for a batch system), hydrogen peroxide dosage and amount of the Fentonic catalysts. The present research involves developing a three layered feed-forward neural network model to predict the effect of the above operational parameters (as input variables) on the decolorization efficiency (DE) of C.I. Direct Red 16 (DR16) as a model pollutant from aqueous solution by Fentonic AOP treatment. The K-fold cross validation method is applied for assessment the reliability of the desired ANN and using the results, sensitivity analysis is done for indicating the relative importance of each operational parameter on the ANN responses.

2. Experimental

2.1. Materials

All materials were used as received without further purification. The azo dye, DR16, $C_{26}H_{17}N_5Na_2O_8S_2$ (C.I. 27680, MW 637.26) was purchased from Alvan Sabet company, 99% pure. Fig. 2 displays the structure of this dye. H_2O_2 (30%, w/w), $FeSO_4{\cdot}7H_2O$, $K_2S_2O_8$, H_2SO_4 and

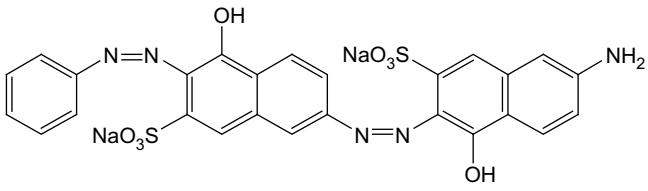

Fig. 2. Structure of the dye molecule.

NaOH were all purchased from Merk company. Distillated water was used throughout this study.

2.2. Reactor structure

Experiments were performed in a stainless steel rectangular cubic vessel with 3.3 l capacity. A micro-air compressor bubbled air through a distributor from the reactor bottom into the solution for well content mixing. The reactor was equipped with a stainless steel water-flow coil for regulating the temperature by means of an external circulating flow of a thermostat bath (OPTIMA-740, Japan) with an accuracy of ±0.1°C.

2.3. Procedure and analysis

In order to perform each run, 3 l of the solution containing initial concentration of 30 mg l^{-1} of the dye (about 5×10^{-5} M) which is within the range of typical concentration in textile waste-waters [21] was prepared and transferred into the reactor. The pH was adjusted to the desired values (2, 2.45, 3, 3.5, 4, 4.35 and 6) by means of a pH meter (Denver, UB-10) using dilute H_2SO_4 or NaOH solutions. After adjusting temperature (which maintained constant during all experiments), appropriate amount of Fentonic reagents [Fe(II) and H_2O_2] were added. Samples (4.0 ml) were taken at regular times (0, 4, 6, 10, 15, 20, 25 and 30 min). The dye concentration was analyzed using a spectrophotometer (Perkin Elmer, 55 OSE) by measuring the absorbance at λ_{max}= 526 nm, and using the appropriate calibration curve. DE in the Fentonic AOP treatment can be calculated as:

$$DE = \frac{[DR16]_0 - [DR16]_t}{[DR16]_0} \quad (7)$$

where $[DR16]_0$ and $[DR16]_t$ are the initial and the appropriate concentration at any time, t, respectively. Because the reaction was continued after sampling, measuring the absorbance were done in the less than 15 s.

2.4. ANN strategy

Since three-layer feed-forward networks can potentially learn virtually any input-output relationship, this performance was developed in this study. The network consists of an input layer, a hidden layer and an output layer. The input variables to the network were initial concentration of the dye, pH, H_2O_2 initial dosage, Fe(II) initial dosage, temperature and time of reaction. The DE was the experimental result and introduced to the network as output variable (network response). The number of neurons in the hidden layer and network calibration iteration are chosen as design variables in the network structure development. Choosing the proper algorithm and transfer function are vital jobs in designing a desirable ANN model; otherwise the output result will not be reliable. Various functions such as "poslin", "tansig" and "satlin" were tested as transfer functions in the hidden layer; also "purelin" and "tansig" functions were applied in this order in the output layer. Different algorithms were used to train the network. The optimum numbers of hidden layer neurons and calibration iteration were found using simple and commonly practiced trial and error method in all the applications herein. MATLAB software (version 2008a) was used in the work and all programs were run on a personal computer.

3. Results and discussion

3.1. Finding the best training and transfer functions

Experimental data sets at different operational conditions were used to train and test the ANN model. Different operational parameters ranges (input/output variables) are given in Table 1.

All algorithms and transfer functions may not be applicable for all processes. The appropriate training algorithm and transfer function are very sensitive parameters for network design. If the model does not comply with the experimental results, then the output value from water treatment plant will be haphazard and it will be difficult to control the plant. It has been reported that 8–11 hidden layer neurons produce minimum value of mean square error (MSE) [22]. So, all models at this step were calibrated and tested with 10 neurons, but with the best training and transfer

Table 1
The range of input and output parameters

Parameter	Range
Input layer	
pH	2–6
$[Fe(II)]_0$ (mg l^{-1})	0.5–4
$[H_2O_2]_0$ (mg l^{-1})	20–1000
$[DR16]_0$ (mg l^{-1})	10–50
T (°C)	20–40
Reaction time (min)	0–30
Output layer	
DE	0–1

function the number of neurons in hidden layer will be optimized later. The literature offers little guidance in selecting the size of training and testing data. In our case for best validation of designed network reliability, the available data set (232 data) was randomized and then divided into the training and testing sets equally. Before training, it is often useful to scale the inputs and targets so that they always fall within a specified range. Hence, all input data (X_i) were normalized in the 0–1 range as:

$$x_i = \frac{X_i - X_{min}}{X_{max} - X_{min}} \quad (8)$$

where X_i, X_{min} and X_{max} refer to the each, lowest and highest value of the ith input variables, respectively.

Thirty various model with different algorithms and transfer functions examined are given in Table 2. Each network run was repeated three times to avoid random correlation due to the random initialization of the weights. Also the table shows the correlation coefficient and MSE parameters for training and testing sets. Considering the results reveals that the model number 2 with the Levenberg–Marquardt (LM) back-propagation algorithm and tansig function in the both hidden and output layers gives most satisfactory results (lowest MSE value of 2.3×10^{-6}, and the exact R^2 value of 1.000) for training set and a completely appreciable MSE and R^2 values (3.3×10^{-3} and 0.966) for testing set. Therefore this model can be selected among all for more development; however, models with numbers of 7, 8, 25, 26 and 30 give satisfactory results as well.

Table 2
Summary of trial and error method used for degradation efficiency by the developed ANN model

Back–propagation algorithm	Training function	Transfer function		R^2 (Train)	MSE × 10^3 (Train)	R^2 (Test)	MSE × 10^3 (Test)	Model number
		Hidden layer	Output layer					
Levenberg–Marquardt	trainlm	tansig	purelin	0.9973	0.2653	0.3773	62.51	1
			tansig	1.000	0.0023	0.9663	3.320	2
		poslin	purelin	0.9658	3.229	0.8294	15.65	3
			tansig	0.9937	0.6035	0.9353	6.106	4
		satlin	purelin	0.9694	2.8924	0.7562	20.78	5
			tansig	0.9729	2.4814	0.7631	18.86	6
Polak–Ribiére conjugate gradient	traincgp	tansig	purelin	0.9956	0.4311	0.9648	3.489	7
			tansig	0.9985	0.1491	0.9924	0.7023	8
		poslin	purelin	0.9767	2.2274	0.9083	9.026	9
			tansig	0.9762	2.2936	0.9467	5.363	10
		satlin	purelin	0.9846	1.4715	0.9082	8.621	11
			tansig	0.9945	0.5188	0.931	6.034	12
One step secant	trainoss	tansig	purelin	0.9888	1.0852	0.9214	7.661	13
			tansig	0.9967	0.3126	0.7146	25.46	14
		poslin	purelin	0.9751	2.3724	0.7862	18.82	15
			tansig	0.9965	0.3422	0.9866	1.292	16
		purelin	purelin	0.9757	2.3143	0.9162	7.860	17
			tansig	0.9632	3.3029	0.6803	26.64	18
Resilient	trainrp	tansig	purelin	0.988	1.1516	0.8972	10.08	19
			tansig	0.9892	1.0160	0.9508	4.487	20
		poslin	purelin	0.9607	3.7021	0.8702	12.09	21
			tansig	0.8639	11.9326	0.8167	16.98	22
		purelin	purelin	0.978	2.0995	0.8973	9.611	23
			tansig	0.9435	5.0068	0.8246	14.61	24
Scaled Conjugate Gradient	trainscg	tansig	purelin	0.9954	0.4432	0.9656	3.356	25
			tansig	0.9954	0.4432	0.984	1.518	26
		poslin	purelin	0.9805	1.8625	0.9369	6.864	27
			tansig	0.9973	0.2615	0.9842	1.563	28
		purelin	purelin	0.9774	2.1491	0.9234	7.311	29
			tansig	0.9941	0.5623	0.9745	2.431	30

3.2. Optimization of neuron number and calibration iteration

Finding the optimum number of network calibration iteration (training epoch) and number of neuron in the hidden layer is an important step in the network designing. The "universal function approximator" capabilities of many ANNs can also cause problems. A common issue is the problem of "overfitting" that is caused by the usually large number of free coefficients inside an ANN. A different pattern arises when an independent data set is evaluated simultaneously. After a similar steep descent during the first few numbers of iterations, the error starts to rise again; indicating a worsening performance of the ANN. Clearly, the optimization of the ANN is going away after a number of calibration iterations or hidden layer neurons, probably because the optimization is trying to minimize the objective function on noise and artifacts in the calibration (training) data. This adversely affects the generality of the model being implemented by the ANN and leads to an increasing error for the testing data set. It is therefore prudent not to execute many numbers of calibration iterations and neurons (in hidden layer) while training an ANN and limit the allowable numbers to a preset maximum (as optimum). A good approach is the one where training and testing data sets are evaluated simultaneously. The set of ANN coefficients where the testing error reaches to its lowest value is taken as the most optimal and used for practical applications [23].

Finding the optimum number of iteration for the network calibration and the number of neurons in the hidden layer was done which the results are depicted in Figs. 3 and 4 respectively. As these figures show, MSE for training set with increasing of the calibration iteration or the number of neurons in hidden layer is reduced continuously, while MSE for testing set reaches to a minimum value at the 300th iteration (Fig. 3) and using 10 neurons in the hidden layer (Fig. 4) and then

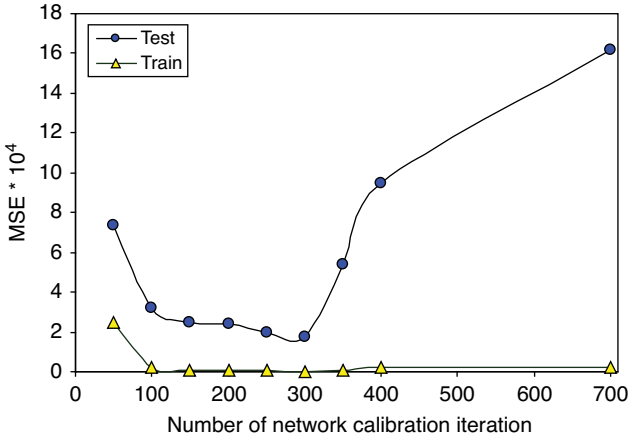

Fig. 3. Effect of iteration on MSE for training and testing data sets using 10 neurons in the hidden layer.

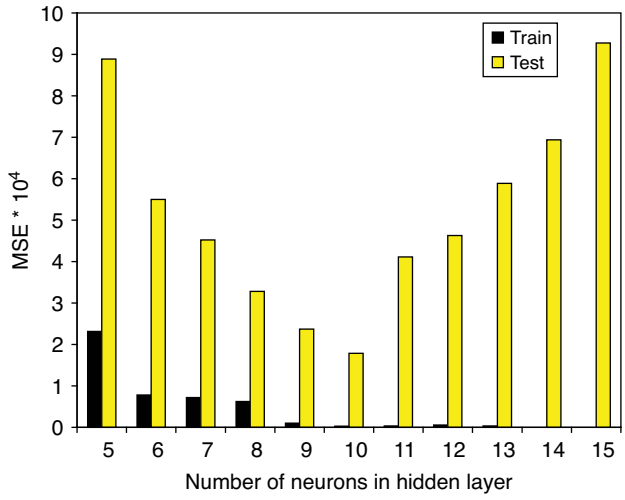

Fig. 4. Effect of the neuron numbers on MSE for training and testing data sets at 300th calibration iteration.

is increased due to the occurrence of overfitting. Therefore, the best simulation of the Fentonic AOP treatment can be performed using a three layer feed forward ANN model compose of 6, 10 and 1 neurons in the input, hidden and output layers, i.e., 6:10:1 structure. In the case of network components; the LM algorithm and tansig as both training and transfer functions are the best. Also best testing results are attributed to the 300th calibration iteration. A schematic diagram of the developed ANN model is presented in Fig. 5.

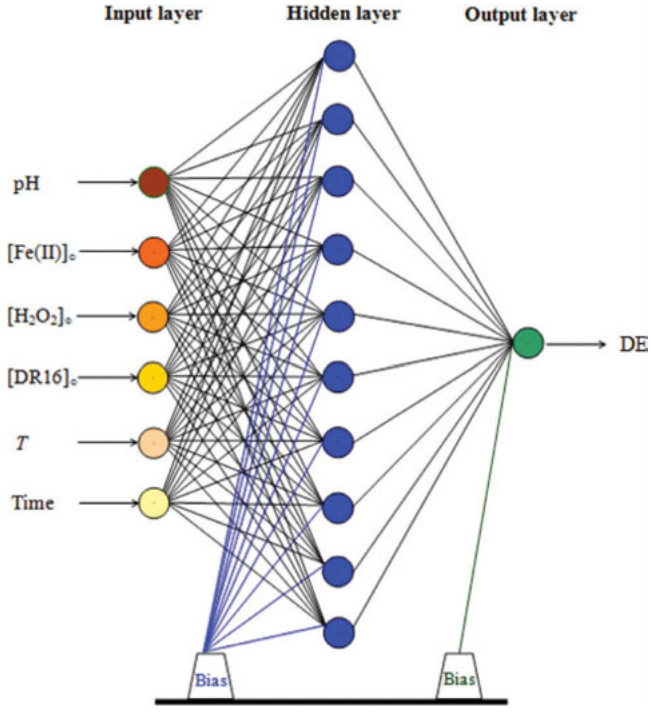

Fig. 5. A schematic view of the developed ANN.

3.3. Validation of the developed ANN using K-fold method

The cross-validation is a heuristic approach to balance complexity and accuracy [24]. In this research, the K-fold cross-validation method was applied to evaluate the performance of the developed ANN in the simulation process. In our case, the total data set are subdivided into 23 roughly equal-sized parts, and then repeated the modeling process 23 times leaving one section out each time for validation purposes. Average accuracy of the estimation was expressed by correlation coefficient and MSE parameters. Fig. 6 shows a comparison between experimental and predicted output values for the test series data using the adopted ANN model. R^2 and MSE values are 0.9984 and 1.56×10^{-4}, which indicate goodness of the model performance.

Figs. 7–10 show the ability of the designed ANN in simulation of the Fentonic AOP treatment via a comparison between predicted and experimental values of the output variable (DE) as a function of solution pH, initial dosages of Fe(II) and H_2O_2, temperature and initial concentration of the DR16 respectively. These results confirm that the ANN model could predict the Fentonic oxidative DE adequately well, within the experimental conditions adopted in the model fitting.

3.4. Description of chemical aspects

As Fig. 7 shows, DE is drastically decreased from 75.1% to only 5.0% after 30 min, when pH increases from 3.5 to 6.0. It is due to the formation of ferrous and ferric oxyhydroxides in pH values more than 4.0 which inhibits the reaction between Fe(II) and H_2O_2 [25] and therefore, small amounts of •OH radical is generated. Also, decreasing pH from 3.5 to 2.0 will cause a great reduction in degradation efficiency from 75.1% to 9.6%. The reason here is that when pH finds values less than 3.5, H_2O_2 molecules will react with excessive H^+ to form oxonium ion ($H_3O_2^+$) which is stable and can not react with Fe(II) to form radicals [26]. At the same time, the generated radicals may be scavenged by excess H^+ ions [25]. Similar trend of variation for DE with pH has been reported in a previous research for azo dye acid

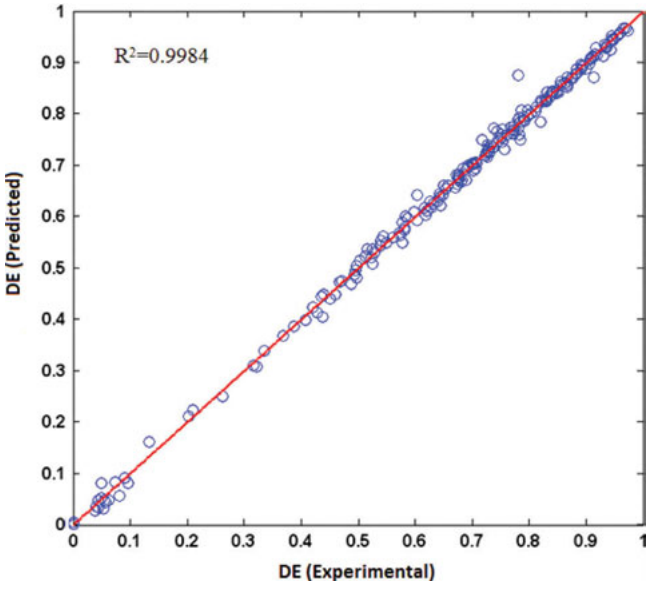

Fig. 6. Predicted output values for testing data sets using the adopted ANN model.

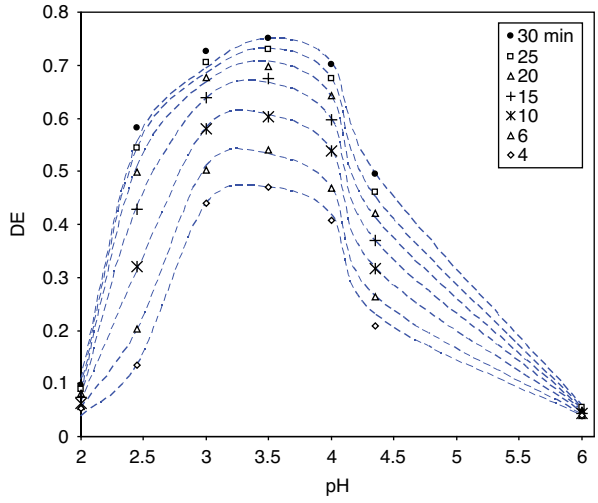

Fig. 7. Comparison between experimental (symbols) and the ANN predicted (dashed lines) values of DE as a function of pH at different times; $[DR16]_0 = 30$ mg l^{-1}, $[H_2O_2]_0 = 500$ mg l^{-1}, $[Fe(II)]_0 = 0.5$ mg l^{-1} and $T = 25$°C.

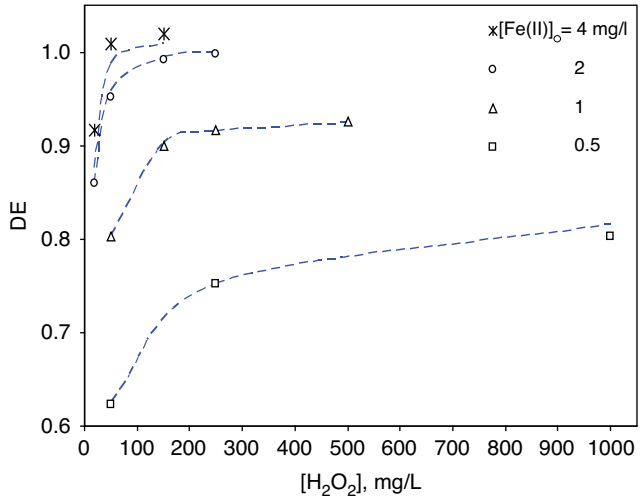

Fig. 8. Comparison between experimental (symbols) and the ANN predicted (dashed lines) values of DE as a function of initial concentration of H_2O_2 and $[Fe(II)]_0$; $[DR16]_0 = 30$ mg l^{-1}, pH = 6.0 and $T = 25$°C.

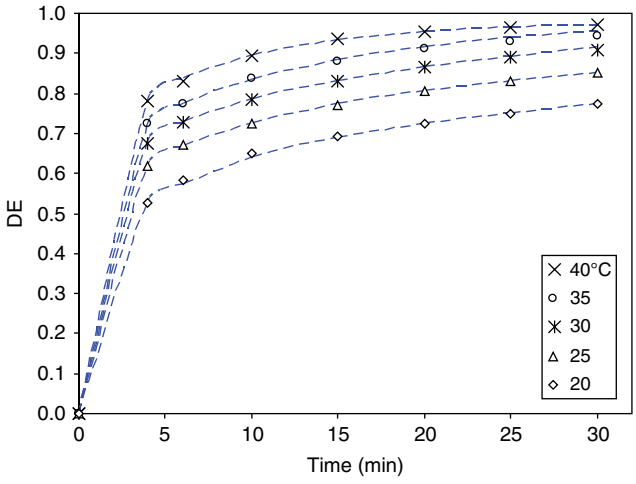

Fig. 9. Comparison between experimental (symbols) and the ANN predicted (dashed lines) values of DE as a function of reaction time for different temperatures; pH = 6.0, $[H_2O_2]_0$ = 150 mg l^{-1}, $[Fe(II)]_0$ = 1.0 mg l^{-1}.

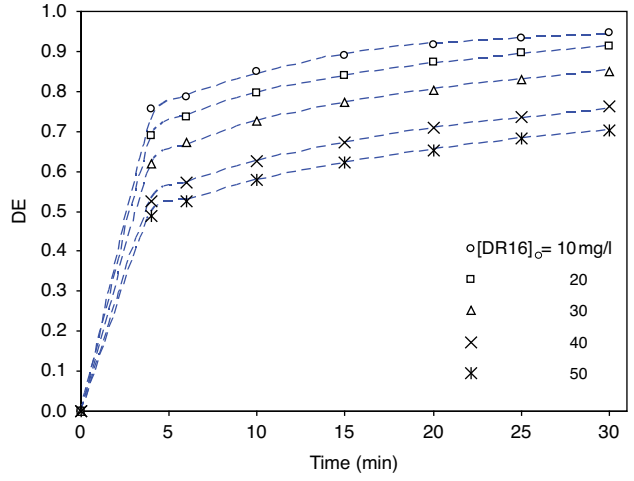

Fig. 10. Comparison between experimental (symbols) and the ANN predicted (dashed lines) values of DE as a function of reaction time for different dye initial concentration; pH = 6.0, $[H_2O_2]_0$ = 150 mg l^{-1}, $[Fe(II)]_0$ = 1.0 mg l^{-1} and T = 25°C.

black 1 [27]. So, the optimum pH of 3.5 is relevant. The solution pH during the Fenton process was measured and only a variation of 0.1 was appropriate during 30 min treatment.

The experimental data in Fig. 8 show the DE at different dosages of ferrous and hydrogen peroxide reagents under pH 3.5 and after 30 min of the reaction progress. It is observed that the DE is significantly influenced by H_2O_2 and Fe(II) dosages. At a constant amount of ferrous ions initial dosage, with increasing the initial dosage of hydrogen peroxide in the solution, the DE is raised up to a plateau region. This non–sensitive region is due to the hydroxyl radical consumption through various approaches, including the scavenging effects of H_2O_2 and recombination of •OH radicals [26,27]. Therefore, the threshold concentration of the plateau region should be used as an optimum dosage of H_2O_2. Meantime, as Fig. 8 demonstrates, for a constant initial amount of hydrogen peroxide, with increasing the ferrous ion initial dosage from 0.5 to 4.0 mg l^{-1}, better results will be obtained. This is, because Fe(II) plays the role of catalyst that start the decomposition of H_2O_2 to generate the very reactive •OH radicals in the process. The most enhancing effect on DE, hence with respect to the added amount of Fe(II), belongs to the 1.0 mg l^{-1}, after which, its effect is not so significant. Considering this fact, and the environmental privilege of applying low catalyst concentration in wastewater treatments, 1.0 mg l^{-1} of Fe(II) which is ten folds lower than the standard allowable limit for total iron content in water effluents [26,27] and 150 mg l^{-1} of H_2O_2, are considered as the most suitable amounts of Fentonic reagents for the process.

The empirical results shown in Fig. 9 indicate that the DE increases from 77.3% to 97.3% as a consequence of temperature increase from 20°C to 40°C after 30 min. As a case, 90% efficiency can be achieved at 40°C after the short time of 10 min. When the temperature rises, the reaction rate between hydrogen peroxide and any form of the ferrous/ferric ions will be increased. This in turn leads to accelerate hydroxyl radical generation rate.

The experimental results of tests about the effect of DR16 initial concentration on the DE by the Fenton process have been shown in Fig. 10. The DE of DR16 is inversely proportional to its initial concentration. A DE of 94.51% can be achieved after 30 min when the initial concentration of DR16 is 10 mg l^{-1} and increasing the initial concentration to 50 mg l^{-1} results in the reduction of the about 24%. This matter is expectable, since with increasing the dye initial concentration, total contaminant load in the media is raised while the inherence potential of the media for the generation of hydroxyl radicals has been kept constant.

3.5. The model sensitivity analysis

Neural networks are black-box models that allow little insight into the relations that are used to predict hydraulic variables. Except for trivially small networks, it is almost impossible to describe the models in short and easy-to-understand equations, thus making the practical implementation of ANNs is difficult. Despite the black-box nature of ANNs, something can be still learned from neural networks. It is possible to carry out sensitivity analysis with neural networks to indicate the influence of different input variables on the model's results [28]. The developed ANN in this work provided the weights listed in Table 3. The level of influence of each input

Table 3
Matrix of weights between input and hidden layers, W_1, and between hidden and output layers, W_2

Neuron	W_1							Bias 2 = 0.1275
	time	pH	T	$[Fe(II)]_0$	$[H_2O_2]_0$	$[DR16]_0$	Bias 1	W_2
1	2.1709	0.4778	0.0571	−0.6202	−1.9314	−0.1997	0.3201	2.7163
2	0.5944	0.4625	1.3088	2.1317	−0.2576	0.0019	−0.2789	−0.3046
3	0.7069	−1.7207	−0.0505	0.0318	−3.0658	0.3329	−1.9715	−1.1352
4	0.3773	2.6496	−1.0915	−1.1531	−0.3873	−1.3175	0.1740	0.4602
5	0.8351	−0.9378	0.3128	0.8860	−0.0168	0.7803	1.2201	1.2727
6	0.7826	0.4462	−0.1783	3.2150	−0.5210	−0.5954	2.4051	−0.2364
7	−3.9116	0.9074	−0.7920	−0.3276	1.2990	1.6294	−1.2979	0.6124
8	−0.3219	−3.6159	0.3560	1.3164	0.7004	0.0726	−0.1466	−1.3134
9	3.7239	−1.7447	−0.6690	−1.5578	−1.1615	−0.0111	1.2077	0.2209
10	−0.8647	−1.3693	0.2605	0.7563	−0.3747	0.4560	−0.8797	0.6389

variable concerning the modeling output variable can be obtained through the neural weight matrix.

The following equation was proposed based on the contribution of connection weights [29]:

$$I_j = \frac{\sum_{m=1}^{m=N_h}\left(\left(\left|W_{jm}^{jh}\right|/\sum_{k=1}^{N_I}\left|W_{km}^{ih}\right|\right)\times\left|W_{mn}^{ho}\right|\right)}{\sum_{k=1}^{k=N_i}\left\{\sum_{m=1}^{m=N_h}\left(\left|W_{km}^{ih}\right|/\sum_{k=1}^{N_i}\left|W_{km}^{ih}\right|\right)\times\left|W_{mn}^{ho}\right|\right\}} \quad (9)$$

where I_j is the relative importance of the jth input variable on the output variable, N_i and N_h are the numbers of input and hidden neurons, respectively; W's are connection weights, the superscripts "i", "h" and "o" refer to input, hidden and output layers, respectively; and subscripts "k", "m" and "n" refer to input, hidden and output neurons, respectively. The relative importance of various variables as calculated by Eq. (9) is shown in Fig. 11.

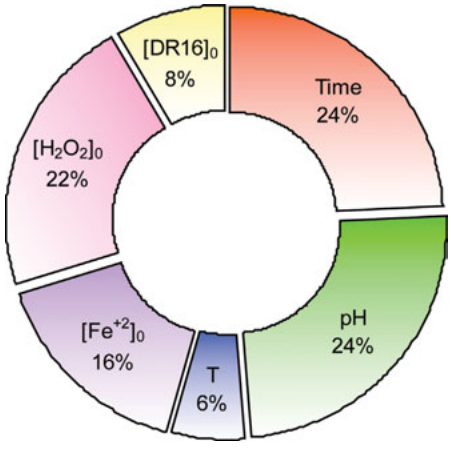

Fig. 11. Importance (%) of the input variables on the Fentonic DE.

As can be seen, under the used experimental conditions, all of the variables except the initial concentration of the dye and temperature have strong effects on the DE. However, the pH and reaction time with a relative importance of 24% appeared to be the most influential parameters in the Fentonic AOP treatment. The initial concentrations of H_2O_2 and Fe(II) find the subsequent orders.

Low sensitivity to the substrate initial concentration implies the inherence potential of the process in the case of higher contaminants loadings. Of course this matter should be further studied to find what extent the initial loading can be expanded. The low sensitivity to temperature, on the other hand, shows that the employed process has nice flexibility within the used temperature, i.e., under different ambient temperatures.

4. Conclusions

The performance of a Fentonic AOP on decolorization of DR16 aqueous solutions was studied, focusing on the influence of key operational factors. The optimum condition for a solution containing 30 mg l^{-1} of the dye can be introduced as: pH = 3.5; $[Fe(II)]_0$ = 1 mg l^{-1}; $[H_2O_2]_0$ = 150 mg l^{-1}; T = 40°C. Under these conditions after 10 min, about 90% of the dye was removed. The treatment process was modeled adequately well using a three-layered feed forward ANN based on LM back-propagation algorithm; including ten neurons in the hidden layer and with operation of tansig function in the both hidden and output layers. It was found that MSE for testing set of data reaches to its lowest value at 300th network calibration iteration. K-fold cross validation confirmed that under different conditions, the developed ANN can effectively reproduce experimental data and predict the behavior of the process.

Also, sensitivity analysis based upon the connection weights showed the importance of the input variables on the value of the DE in the order of: pH ≈ time > $[H_2O_2]_0$ > $[Fe(II)]_0$ > $[DR16]_0$ > temperature.

Acknowledgements

We express thanks to the university authorities for providing the financial support to carry out this work.

Symbols

DE	—	decolorization efficiency
t	—	reaction time (min)
T	—	temperature (°C)
x_i	—	ith normalized data
X_i	—	ith data
X_{min}	—	minimum of the data
X_{max}	—	maximum of the data
$y_{i,predic}$	—	ith predicted data
$y_{i,exp}$	—	ith experimental data
W_1	—	weight of the first layer of the ANN
W_2	—	weight of the second layer of the ANN
b1	—	biases of the first layer of the ANN
b2	—	bias of the second layer of the ANN
N_i	—	number of input neurons
N	—	number of total data
I_j	—	relative importance of jth input variable
N_h	—	number of hidden neurons

References

[1] A. Celekli and F. Geyik, Artificial neural networks (ANN) approach for modeling of removal of Lanaset Red G on *Chara contraria*, Bioresour. Technol., 102 (2011) 5634–5638.
[2] T. Yilmaz, G. Seckin and A. Yuceer, Modeling of effluent COD in UAF reactor treating cyanide containing wastewater using artificial neural network approaches, Adv. Eng. Software, 41 (2010) 1005–1010.
[3] A.R. Khataee, M. Zarei, G. Dehghan, E. Ebadi and M. Pourhassan, Biotreatment of a triphenylmethane dye solution using a Xanthophyta alga: Modeling of key factors by neural network, J. Taiwan Inst. Chem. Eng., 42 (2011) 380–386.
[4] A.J. Maren, C.T. Harston and R.M. Pap, Handbook of Neural Computing Applications. Academic Press, San Diego, 1990.
[5] R. Hecht-Nielsen, Neurocomputing. Addison-Wesley Publishing Company, Reading, MA, 1990.
[6] J. Saien and A.R. Soleymani, Degradation and mineralization of Direct Blue 71 in a circulating upflow reactor by UV/TiO_2 process and employing a new method in kinetic study, J. Hazard. Mater., 144 (2007) 506–512.
[7] A. Bach, G. Zelmanov and R. Semiat, Wastewater mineralization using advanced oxidation process, Desalin. Water Treat., 6 (2009) 152–159.
[8] I.K. Konstantinou and T.A. Albanis, TiO_2-assisted photocatalytic degradation of azo dyes in aqueous solution: kinetic and mechanistic investigations: A review, Appl. Catal. B., 49 (2004) 1–14.
[9] J. Grzechulska and A.W. Moraviski, Photocatalytic decomposition of azo-dye acid black 1 in water over modified titanium dioxide, Appl. Catal. B., 36 (2002) 45–51.
[10] S. Ledakowicz, M. Solecka and R. Zylla, Biodegradation, decolorization and detoxification of textile wastewater enhanced by advanced oxidation processes, J.Biotechnol., 89 (2001) 175–184.
[11] J.H. Ramirez, C.A. Costa and L.M. Madeira, Experimental design to optimize the degradation of the synthetic dye Orange II using Fenton's reagent, Catal. Today, 107 (2005) 68–76.
[12] M. Qiu and C. Huang, A comparative study of degradation of the azo dye C.I. Acid Blue 9 by Fenton and photo-Fenton oxidation, Desalin. Water Treat., 24 (2010) 273–277.
[13] L. Gomathi Devi, S. Girish Kumar, K. Mohan Reddy and C. Munikrishnappa, Effect of various inorganic anions on the degradation of Congo Red, a di azo dye, by the photo-assisted Fenton process using zero-valent metallic iron as a catalyst, Desalin. Water Treat., 4 (2009) 294–305.
[14] R.J. Bigda, Consider Fenton's chemistry for wastewater treatment, Chem. Eng. Prog., 91 (1995) 62–66.
[15] N. Masomboon, C.W. Chen, J. Anotai and M.C. Lu, A statistical experimental design to determine o-toluidine degradation by the photo-Fenton process, Chem. Eng. J., 159 (2010) 116–122.
[16] Y.H. Huang, Y.F. Huang, P.S. Chang and C.Y. Chen, Comparative study of oxidation of dye-Reactive Black B by different advanced oxidation processes: Fenton, electro-Fenton and photo-Fenton, J. Hazard. Mater., 154 (2008) 655–662.
[17] J.M. Monteagudo, A. Duran and C. Lopez-Almodovar, Homogeneous ferrioxalate-assisted solar photo-Fenton degradation of Orange II aqueous solutions, Appl. Catal. B., 83 (2008) 46–55.
[18] A. Duran, J.M. Monteagudo and M. Mohedano, Neural networks simulation of photo-Fenton degradation of Reactive Blue 4, Appl. Catal. B., 65 (2006) 127–134.
[19] A.R. Khataee and O. Mirzajani, UV/peroxydisulfate oxidation of C. I. Basic Blue 3: Modeling of key factors by artificial neural network, Desalination, 251 (2010) 64–69.
[20] D. Salari, N. Daneshvar, F. Aghazadeh and A.R. Khataee, Application of artificial neural networks for modeling of the treatment of wastewater contaminated with methyl tert-butyl ether (MTBE) by UV/H_2O_2 process, J. Hazard. Mater., B 125 (2005) 205–210.
[21] N. Daneshvar, A. Aleboyeh and A.R. Khataee, The evaluation of electrical energy per order (EEo) for photooxidative decolorization of four textile dye solutions by the kinetic model, Chemosphere, 59 (2005) 761–767.
[22] S. Dutta, S.A. Parsons, C. Bhattacharjee, S. Bandhyopadhyay and S. Datta, Development of an artificial neural network model for adsorption and photocatalysis of reactive dye on TiO_2 surface, Expert Syst. Appl., 37 (2010) 8634–8638.
[23] K.V. Kumar and K. Porkodi, Modeling the solid–liquid adsorption processes using artificial neural networks trained by pseudo second order kinetics, Chem. Eng. J., 148 (2009) 20–25.
[24] P.I. Good, Resampling Methods: A Practical Guide to Data Analysis. Birkhäuser, Boston, 1999.
[25] X.R. Xu, Z.Y. Zhao, X.Y. Li and J.D. Gu, Chemical oxidative degradation of methyl tertbutyl ether in aqueous solution by Fenton's reagent, Chemosphere, 55 (2004) 73–79.
[26] N.K. Daud and B.H. Hameed, Decolorization of Acid Red 1 by Fenton-like process using rice husk ash-based catalysis, J. Hazard. Mater., 176 (2010) 938–944.
[27] J.H. Sun, S.P Sun, J.Y. Sun, R.X. Sun, L.P. Qiao, H.Q. Guo and M.H. Fan, Degradation of azo dye Acid Black 1 using low concentration iron of Fenton process facilitated by ultrasonic irradiation, Ultrason. Sonochem., 14 (2007) 761–766.
[28] M.G. Schaap and W. Bouten, Modeling water retention curves of sandy soils using neural networks, Water Resour. Res., 32 (1996) 3033–3040.
[29] G.D. Garson, Interpreting neural-network connection weights, AI Expert, 6 (1991) 47–51.