

Prediction of azo dye decolorization by UV/H₂O₂ using artificial neural networks

A. Aleboyeh*, M.B. Kasiri, M.E. Olya, H. Aleboyeh

*Laboratoire Génie des Procédés Traitement des Effluents, Ecole Nationale Supérieure de Chimie de Mulhouse,
Université de Haute Alsace, 3 rue Alfred Werner, 68093 Mulhouse, France*

Received 9 March 2007; received in revised form 21 May 2007; accepted 22 May 2007

Available online 8 June 2007

Abstract

An artificial neural network model was developed to predict the photochemical decolorization of C.I. Acid Orange 7 solution by a combination of UV and hydrogen peroxide. The initial concentrations of dye and hydrogen peroxide, the pH of the solution and time of UV irradiation were employed as input to the network; the output of the network was decolorization efficiency. The data used in this study were obtained from our previous papers. The multilayer feed-forward networks were trained by 114 sets of input–output patterns using a backpropagation algorithm; a three-layered network with eight neurons in the hidden layer gave optimal results. The model gave good predictions of high correlation coefficient ($R^2 = 0.996$). As expected, the initial concentration of H₂O₂ with a relative importance of 48.89%, appeared to be the most influential parameter in the decolorization process.

© 2007 Published by Elsevier Ltd.

Keywords: Artificial neural networks; Advanced oxidation process; C.I. Acid Orange 7; UV/H₂O₂; Decolorization; Modelling

1. Introduction

Wastewater from the dye industry often contains significant amounts of nonbiodegradable organic compounds. The elimination of these contaminants is necessary before a biological treatment of these effluents. Combined UV radiation and hydrogen peroxide oxidation has been applied successfully in advanced oxidation processes (AOPs) to treat different pollutants in water [1–4]. The mechanism of dye destruction in AOPs is based on the formation of hydroxyl radicals ($\cdot\text{OH}$), with an oxidation potential of 2.80 V [5], that can oxidize a broad range of organic compounds. In comparison with other AOPs, i.e. Fenton, O₃, UV/O₃, UV/TiO₂, etc., the photolysis of hydrogen peroxide shows some advantages such as the complete miscibility of H₂O₂ with water, the stability and

commercial availability of hydrogen peroxide, the absence of phase transfer problems and lower investment costs. It can be carried out under ambient conditions and may lead to complete mineralization of organic carbon into CO₂ [6]. Unfortunately, the molar absorptivity of H₂O₂ is low in the readily accessible UV region. The absorption spectrum exhibits a slow and steady rise from 400 nm and, at 254 nm, the molar absorptivity is about 19 l mol^{−1} cm^{−1} [7].

Wastewater treatment by applying AOPs processes is, in general, quite complex. This is caused by the complexity of solving the equations that involve the radiant energy balance, the spatial distribution of the absorbed radiation, mass transfer, and the mechanisms of a photochemical or photocatalytic degradation involving radical species. Since the process depends on several factors, the modelling 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.

Artificial neural networks (ANNs) are now commonly used in many areas of chemistry and they represent a set of methods

* Corresponding author. Tel.: +33 389336800; fax: +33 389336805.

E-mail addresses: azam.aleboyeh@uha.fr (A. Aleboyeh), masoud.bagherzadeh-kasiri@uha.fr (M.B. Kasiri), mohammad.olya@uha.fr (M.E. Olya), hamid.aleboyeh@uha.fr (H. Aleboyeh).

that may be useful in solving such problems [8–11]. The first results of neural network application were published half a century ago (1943) by McCulloch and Pitts [12]. In particular their flexibility and their ability to model highly nonlinear phenomena make them a candidate method in multivariate calibration. ANNs are promising alternative modelling tools that do not require the mathematical description of the phenomena involved in the process, and might therefore prove useful in simulating and up-scaling complex photochemical systems. Their ability to recognize and reproduce cause–effect relationships through training, for multiple input–output systems, makes them efficient to represent even the most complex systems [9]. Success in obtaining a reliable and robust network depends strongly on the choice of the process variables involved as well as the available set of data and the domain used for training purposes [13]. There are several types of artificial neural networks. Two popular ANNs are (i) multilayer feed-forward neural network trained by backpropagation algorithm that is widely used, and (ii) Kohonen self-organizing mapping [14]. The application of ANN analysis to solve the environmental engineering problems has been the subject of numerous review articles. Gob et al. [8] have reported kinetic modelling of the photochemical water treatment process. The article published by Lek and Guegan [14] describes the application of ANNs as a tool in ecological modelling. Cinar et al. [15] have determined the interrelationship and response of process variables involved in water treatment plants. They have analyzed the system behaviour of a full-scale activated sludge wastewater treatment plant by using Kohonen self-organizing feature maps neural network. Wen and Vassiliadis [16] have proposed an automatic control system for the operation of the wastewater treatment process by applying hybrid artificial intelligence techniques in real-time control. Recently, ANNs were used for the modelling of dye removal by advanced oxidation process [17,18].

The present investigation discusses the use of a multilayer feed-forward neural network model to predict the colour removal efficiency of C.I. Acid Orange 7 (AO7) dye solution by UV/H₂O₂ process.

2. Materials and methods

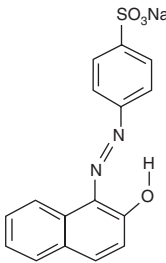
2.1. Chemicals

4-(2-Hydroxy-1-naphthylazo) benzene sulfonic acid sodium salt, C.I. Acid Orange 7 (90%), was obtained from Aldrich Chemical. The dye solutions were prepared by dissolving a defined quantity of the dye in deionised water. The characteristics of the dye (AO7) are shown in Table 1. Hydrogen peroxide (30%, w/w) was obtained from Prolabo.

2.2. Photolysis experiments

The photochemical set-up used in this work has been previously described in detail [19]. Briefly, it consists of a 2500 ml reservoir and a flow-through annular photochemical

Table 1
Chemical structure and characteristics of C.I. Acid Orange 7

Chemical structure	Colour index number	λ_{\max} (nm)	$\epsilon_{\lambda_{\max}}$ (l mol ⁻¹ cm ⁻¹)	Chemical class	M_w (g mol ⁻¹)
	15,510	483	19.43×10^3	Monoazo	350

reactor equipped with a low mercury 15 W lamp. The radiant flux of this lamp was determined by means of the chemical actinometer hydrogen peroxide [7]. The incident photon flux was 6.1×10^{-6} einstein s⁻¹. The solution of the dye was adjusted to the desired pH by addition of HCl or NaOH. The different initial H₂O₂ dosages were added to the dye solution to prepare the reaction mixtures. The dye solution was exposed to UV radiation. At regular time intervals, samples were taken and analyzed by JASCO (V-530) UV–vis spectrophotometer to determine the decolorization efficiency. The concentration of the residual dye in solution was calculated by Beer–Lambert law using the optical density and molar extinction observed at the characteristic wavelength ($\lambda = 483$ nm). The efficiency of colour removal is defined by the following expression:

$$\text{Decolorization (\%)} = \left(1 - \frac{C}{C_0}\right) \times 100 \quad (1)$$

where C_0 and C are the dye concentrations at time 0 and t , respectively.

2.3. Artificial neural networks (ANNs)

ANNs are directly inspired from the biology of the human brain, where billions of neurons are interconnected to process a variety of complex information. Accordingly, a computational neural network consists of simple processing units called neuron. Each network consists of artificial neurons grouped into layers and put in relation to each other by parallel connections. The strength of these interconnections is determined by the weight associated with them. For every ANN, the first layer constitutes the input layer (independent variables) and the last one forms the output layer (dependent variables). Between them, one or more neurons layers called hidden layers can be located.

The number of input and output neurons effectively represents the number of variables used in the prediction and the number of variables to be predicted, respectively. The hidden layers act as feature detectors and, in theory, there can be more than one hidden layer. The universal approximation theory, however, suggests that a network with a single hidden layer

with a sufficiently large number of neurons can interpret any input–output structure [9,11,13]. The number of neurons in the hidden layer is determined by the desired accuracy in the neural predictions. Hence, it may be considered as a parameter for the neural net design. In the feed-forward neural net, all the neurons of a particular layer are connected to all the neurons of the layer next to it. The input layer of neurons acts as a distributor and the input to this layer is directly transmitted to the hidden layer. The inputs to hidden and output layers are calculated by performing a weighted summation of all the inputs received from the preceding layer. The weighted sum of the inputs is transferred to the hidden neurons, where it is transformed using an activation function. The output of hidden neurons, in turn, acts as inputs to output neurons where it undergoes another transformation. The most widely used transfer function for the input and hidden layers is the sigmoid transfer function and is given by

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

The linear activation function (Eq. (3)) is used as the output layer activation function [20].

$$f(x) = x \quad (3)$$

Other types of nets that have been used in chemical engineering applications include recurrent networks. They are similar to feed-forward neural nets, but also include time-delayed feedback or recycle [9].

In this work, multilayer feed-forward ANN with one hidden layer was used. For all data sets, sigmoid activation function in the hidden layer and a linear transfer function in the output node were used. Data sets were obtained from our previous paper [19]. The ANN was trained using the backpropagation algorithm. All calculations were carried out with Matlab mathematical software with the ANN toolbox.

3. Results and discussion

3.1. The effect of pH of the dye solution on colour removal efficiency

Since the structure and characteristics of the dyes are not constant at different pHs, the effect of pH on the decolorization efficiency in the pH range between 2 and 9 was studied. Dye solutions with different initial pHs were examined to monitor the evaluation of the absorbancy with time. The decolorization efficiency of the AO7 solution increased from pH 2 to 3.5 and then remained almost constant up to pH 5.5. After this value, the efficiency decreased as the pH increased toward 9 (Fig. 1).

3.2. The effect of H₂O₂ dosage on the colour removal efficiency

Hydrogen peroxide concentration is an important parameter for the degradation of the dye in the UV/H₂O₂ photoreactor. Indeed, the OH free radicals produced upon photolysis of

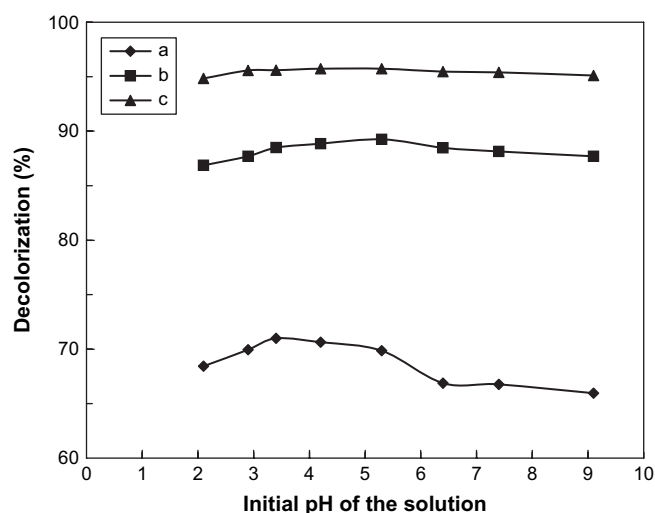
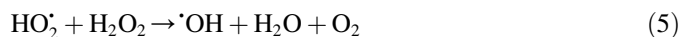


Fig. 1. The effect of the initial pH of the solution on colour removal efficiency. (Initial concentration of the dye = 5×10^{-5} M; initial concentration of H₂O₂ = 48.9 mM; (a) irradiation time = 4 min; (b) irradiation time = 8 min; (c) irradiation time = 12 min.)

H₂O₂ can react with dye molecules, but also with an excess of H₂O₂. Excess amount of hydrogen peroxide and high concentration of OH radical result in competitive reactions, an inhibitory effect on the decolorization. OH radicals are prone to react or to recombine according to the following scheme [4]:



Reactions (4) and (7) are $\cdot\text{OH}$ consuming, so the probability of oxidation of the substrate decreases. Therefore, an important step in the optimization of the method is the determination of the adequate amount of H₂O₂, to avoid an excess amount of reagent that would postpone the decolorization. The results showed in Fig. 2 point out the negative effect of an insufficient or excessive amount of H₂O₂ on the decolorization efficiency. An optimum value of about 30 mM for the hydrogen peroxide concentration was obtained when the initial concentration of the dye was 5×10^{-5} M, and the initial pH was about 5. At low concentration, H₂O₂ cannot generate enough hydroxyl radicals and the oxidation rate is logically slow. Furthermore, most of the free radicals are directly consumed by the dye. In the presence of high concentration of H₂O₂, it could be expected that more OH radicals would be produced. However, these radicals preferably react with the excess of H₂O₂. This undesirable reaction competes with the degradation of the dye.

3.3. ANN model development

In this study, one layer of hidden neurons was used. The net used was feed-forward neural network trained by backpropagation algorithm. The input variables to feed-forward neural

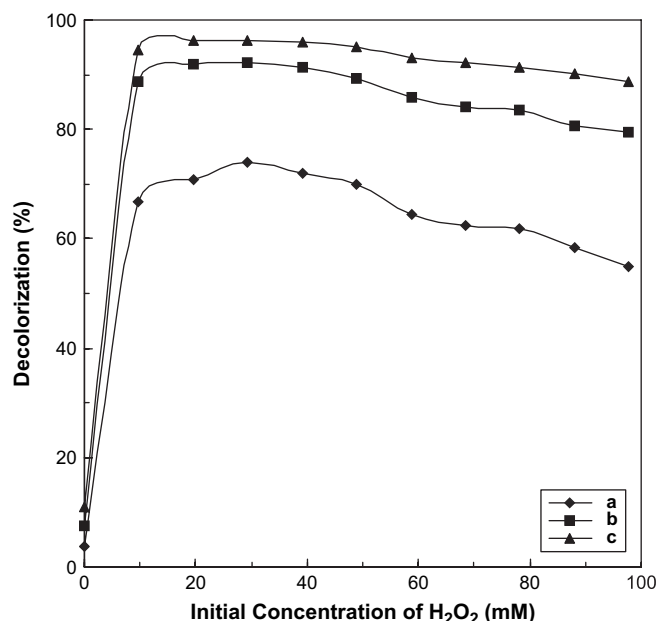


Fig. 2. The effect of H₂O₂ dosage on colour removal efficiency. (Initial concentration of the dye = 5×10^{-5} M; initial pH of the solution = 5.25; (a) irradiation time = 4 min; (b) irradiation time = 8 min; (c) irradiation time = 12 min.)

network were the initial concentrations of the dye and H₂O₂, initial pH and reaction time. The colour removal efficiency was the experimental response or output variable and was calculated by Eq. (1). The characteristics of input and output variables are shown in Table 2.

The topology of an artificial neural network (ANN) is determined by the number of layers, the number of nodes in each layer and the nature of the transfer functions. Optimization of ANN topology is probably the most important step in the development of a model [18].

In order to determine the optimum number of hidden nodes, a series of topologies was used, in which the number of nodes was varied from 2 to 10. The mean square error (MSE) was used as the error function. MSE measures the performance of the network according to the following equation:

$$\text{MSE} = \frac{\sum_{i=1}^N (y_{i,\text{pred}} - y_{i,\text{exp}})^2}{N} \quad (8)$$

Table 2
Characteristics of input and output variables

Variable	Range
<i>Input layer</i>	
Initial dye concentration (mM)	$2 \times 10^{-2} - 15 \times 10^{-2}$
Initial concentration of H ₂ O ₂ (mM)	0–97.80
Initial pH	2.15–9.05
Reaction time (min)	0–15
<i>Output layer</i>	
Decolorization (%)	0–100

where N is the number of data points, $y_{i,\text{pred}}$ is the network prediction, $y_{i,\text{exp}}$ is the experimental response and i is an index of data.

All ANNs were trained using the backpropagation algorithm (*scaled conjugate gradient algorithm*). Network training is a process by which the connection weight and bias on the ANN are adapted through a continuous process of simulation by the environment in which the network is embedded. The primary goal of training is to minimize the error function (MSE) by searching for a set of connection weights and biases that causes the ANN to produce outputs that are equal or close to target values. In other words, the backpropagation algorithm minimizes the MSE between the observed and the predicted output in the output layer, through two phases. In the forward phase, the external input information signals at the input neurons which are propagated forward to compute the output information signal at the output neuron. In the backward phase, modifications to the connection strengths are made, based on the basis of the difference in the predicted and observed information signals at the output neuron [20].

In total, 228 experimental sets were used to feed the ANN structure. The data sets were divided into training, validation and test subsets, each of which contains 114, 57 and 57 samples, respectively. The validation and test sets, for the evaluation of the validation and modelling power of the nets, were randomly selected from the experimental data.

All samples were normalized in the 0–1 range. So, all of the data sets (X_i) (from the training, validation and test sets) were scaled to a new value x_i as follows:

$$x_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (9)$$

Each topology was repeated three times to avoid random correlation due to the random initialization of the weights. Fig. 3 illustrates the network error versus the number of neurons in the hidden layer. It could be seen that the network performance stabilized after inclusion of eight nodes on the hidden layer.

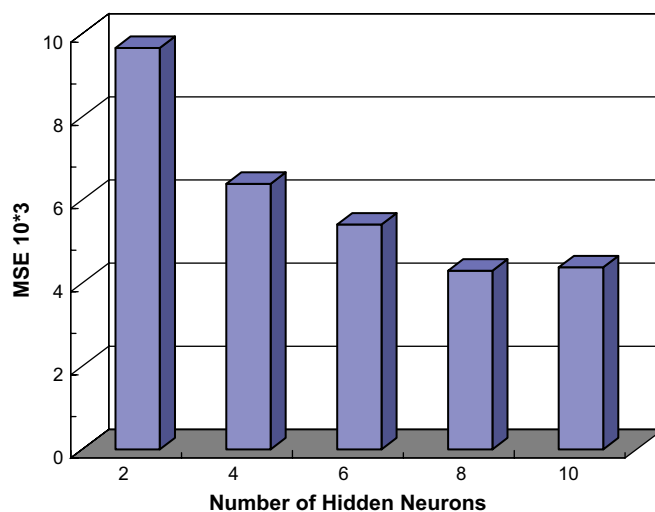


Fig. 3. Effect of the number of neurons in the hidden layer on the performance of the neural network.

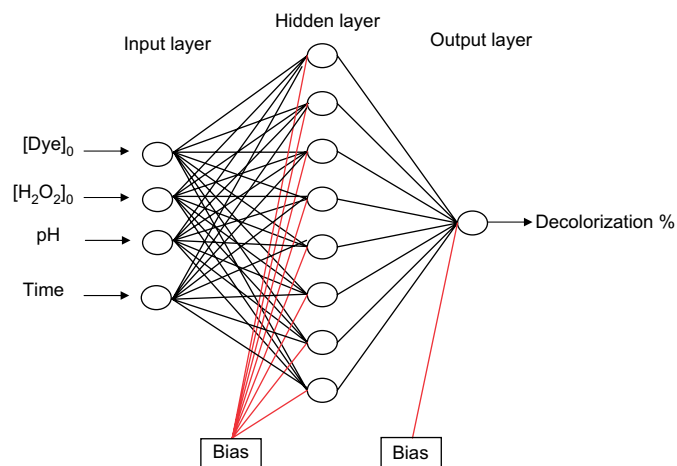


Fig. 4. ANN optimized structure.

So, based on the approximation of MSE function, a number of hidden neurons equal to eight was adopted and a three-layered feed-forward backpropagation neural network was used for the modelling of the process. (Fig. 4).

The network was tested and validated by comparing its predicted output values with the experimental ones using an independent set of data (test set).

3.4. Test of the fitted model

In order to calculate modelling errors, all of the outputs were performed in an inverse range scaling to return the predicted responses to their original scale and compare them with experimental responses. Fig. 5 shows a comparison between experimental values and predicted output variables using the adopted neural network model. We used two lines to show the success of the prediction. One is the perfect fit

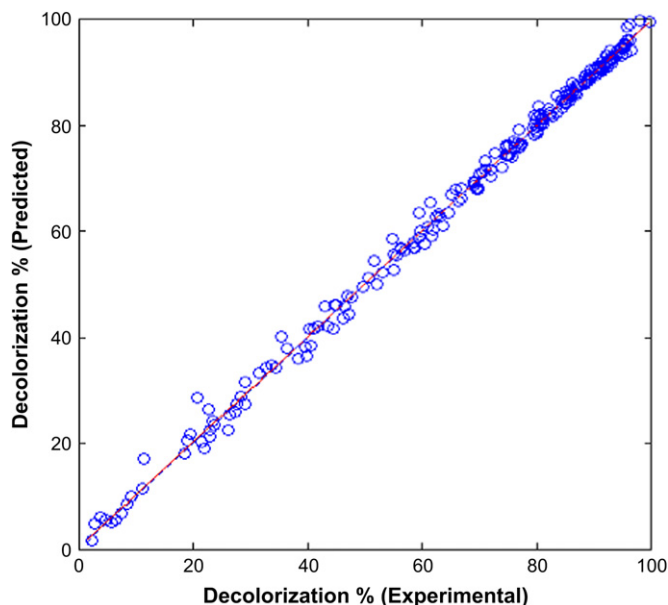


Fig. 5. Comparison between calculated and experimental values of the output.

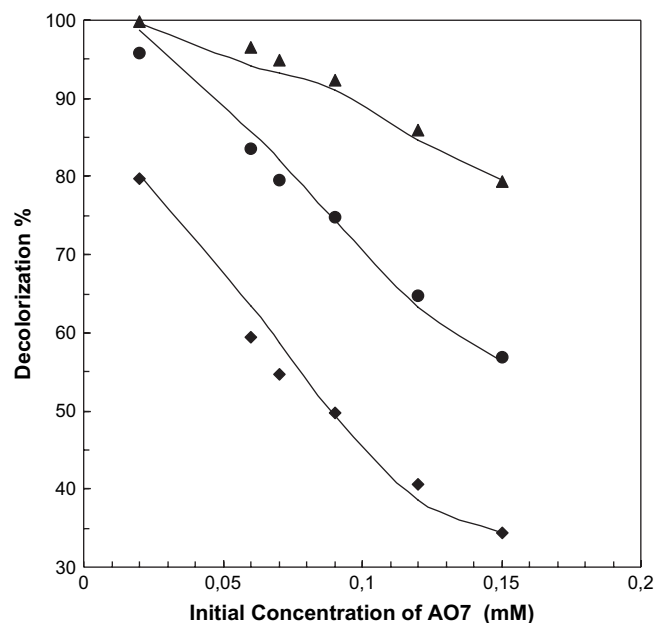


Fig. 6. Comparison between experimental and ANN predicted values of colour removal efficiency as a function of the initial concentration of the dye, using the neural network model. (Concentration of $H_2O_2 = 48.9$ mM, initial pH of the solution = 5.6, irradiation time: (▲) 12 min, (●) 8 min, (◆) 4 min – corresponding predicted curve.)

(predicted data = experimental data), on which all the data of an ideal model should lie. The other is the line that best fits the data of the scatter plot with equation $Y = ax + b$, and it is obtained with regression analysis based on the minimization of the squared errors. The correlation coefficient (R^2) value is 0.996.

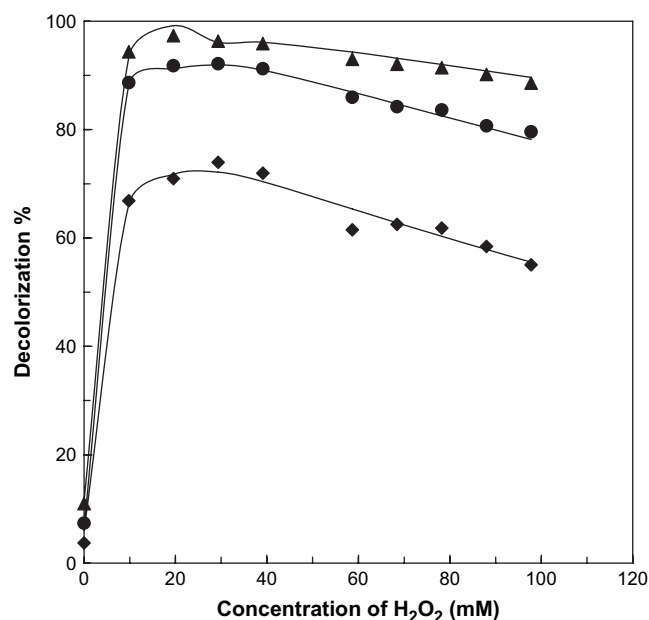


Fig. 7. Comparison between experimental and ANN predicted values of colour removal efficiency as a function of concentration of H_2O_2 , using the neural network model. (Initial concentration of the dye = 5×10^{-5} M, initial pH of the solution = 5.25, irradiation time: (▲) 12 min, (●) 8 min, (◆) 4 min – corresponding predicted curve.)

Table 3

Matrix of weights, W1: weights between input and hidden layers; W2: weights between hidden and output layers

W1						W2	
Neuron	Variable				Bias	Neuron	Weight
	[Dye] ₀	[H ₂ O ₂] ₀	pH	Time			
1	0.4557	−1.7201	0.9567	−2.2077	1.3472	1	−4.4917
2	−1.5875	−0.1835	0.0449	−1.3063	−0.1366	2	1.3619
3	−0.3141	−0.1985	0.0487	−1.1296	−2.9870	3	−1.8950
4	−3.0995	9.4325	1.2324	−1.1858	2.8743	4	7.5475
5	0.8872	0.3068	−0.6494	−0.6329	−0.6773	5	−0.2181
6	−2.4871	6.8689	−1.0309	−1.1175	−2.6848	6	5.3248
7	0.0738	−2.9360	0.6401	0.1713	−1.7200	7	−3.2429
8	−0.5212	−0.3558	−0.4496	−0.4222	1.7893	8	−1.5154
						Bias	−1.4390

Also, Figs. 6 and 7 show a comparison between calculated and experimental values of the output variable (colour removal efficiency) as a function of the initial concentration of the dye and H₂O₂, respectively, using the neural network model.

These results confirm that the neural network model could effectively reproduce the experimental results.

The ANN used in this work provided the weights listed in Table 3. The weights are coefficients between the artificial neurons, which are analogous to synapse strengths between the axons and dendrites in real biological neurons. Therefore, each weight decides what proportion of the incoming signal will be transmitted into the neuron's body [21].

The neural net weight matrix can be used to assess the relative importance of the various input variables on the output variables. Garson [22] proposed an equation based on the partitioning of connection weights:

$$I_j = \frac{\sum_{m=1}^{m=N_h} \left(\left(\frac{|W_{jm}^{ih}|}{\sum_{k=1}^{N_i} |W_{km}^{ih}|} \right) \times |W_{mn}^{ho}| \right)}{\sum_{k=1}^{k=N_i} \left\{ \sum_{m=1}^{m=N_h} \left(\frac{|W_{km}^{ih}|}{\sum_{k=1}^{N_i} |W_{km}^{ih}|} \right) \times |W_{mn}^{ho}| \right\}} \quad (10)$$

where I_j is the relative importance of the j^{th} 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. (10) is shown in Table 4. As may be seen, all of the variables (initial concentration of the dye and H₂O₂, initial pH and reaction time) have strong effects on the decolorization efficiency. Therefore, none of the variables studied in this work

could have been neglected in the present analysis. However, as expected, the initial concentration of H₂O₂ with a relative importance of 48.89% appeared to be the most influential parameter in the decolorization process.

4. Conclusion

The performance of photochemical decolorization of C.I. Acid Orange 7 solution by UV/H₂O₂ process was successfully predicted by applying a three-layered neural network with eight neurons in the hidden layer, and using a backpropagation algorithm. Simulations based on the ANN model were performed in order to estimate the behaviour of the system under different conditions. All of the studied parameters in this work (initial concentration of the dye and H₂O₂, initial pH and reaction time) have considerable effects on the decolorization efficiency and, as expected, the initial concentration of H₂O₂ with a relative importance of 48.89%, appeared to be the most influential parameter in the decolorization process. The results of modelling confirmed that neural network modelling could effectively reproduce experimental data and predict the behaviour of the process.

References

- [1] Colonna GM, Caronna T, Marcandalli B. Oxidative degradation of dyes by ultraviolet radiation in presence of hydrogen peroxide. *Dyes and Pigments* 1999;41:211–20.
- [2] Neamtu M, Siminiceanu I, Yediler A, Kettrup A. Kinetics of decolorization and mineralization of reactive azo dyes in aqueous solution by UV/H₂O₂ oxidation. *Dyes and Pigments* 2002;53:93–9.
- [3] Muruganandham M, Swaminathan M. Photochemical oxidation of reactive azo dye with UV–H₂O₂ process. *Dyes and Pigments* 2004;62: 269–75.
- [4] Aleboyeh A, Aleboyeh H, Moussa Y. Critical effect of hydrogen peroxide in photochemical oxidative decolorization of dyes: Acid Orange 8, Acid Blue 74 and Methyl Orange. *Dyes and Pigments* 2003;57: 67–75.
- [5] Legrini O, Oliveros E, Braun AM. Photochemical processes for water treatment. *Chemical Reviews* 1993;93:671–98.
- [6] Galindo C, Jacques P, Kalt A. Total mineralization of an azo dye (Acid Orange 7) by UV/H₂O₂ oxidation. *Journal of Advanced Oxidation Technology* 1999;4:400–8.

Table 4

Relative importance of input variables on the value of decolorization efficiency

Input variable	Importance (%)
Initial concentration of the dye	18.50
Initial concentration of H ₂ O ₂	48.89
Initial pH	11.61
Reaction time	21.00
Total	100.00

- [7] Nicole I, Laat JD, Dore M, Duguet JP, Bonnel C. Use of UV radiation in water treatment: measurement of photonic flux by hydrogen peroxide actinometry. *Water Research* 1990;24:157–68.
- [8] Gob S, Oliveros E, Bossmann SH, Braun AM, Guardani R, Nascimento CAO. Modeling the kinetics of a photochemical water treatment process by means of artificial neural networks. *Chemical Engineering and Processing* 1999;38:373–82.
- [9] Pareek VK, Brungs MP, Adesina AA, Sharma R. Artificial neural network modeling of a multiphase photodegradation system. *Journal of Photochemistry and Photobiology A: Chemistry* 2002;149:139–46.
- [10] Stegemann JA, Buenfeld NR. Prediction of leachate pH for cement paste containing pure metal compounds. *Journal of Hazardous Materials* 2002;B90:169–88.
- [11] Moraes JEF, Quina Nascimeto CAO, Silva DN, Chiavone-Filho O. Treatment of saline waste water contaminated with hydrocarbons by the photo-Fenton process. *Environmental Science and Technology* 2004;38:1183–7.
- [12] McCulloch WW, Pitts W. A logical calculus of ideas imminent in nervous activity. *Bull Math Biophys* 1943;5:115–33.
- [13] Despange F, Massart DL. Neural networks in multivariate calibration. *Analyst* 1998;123:157–78.
- [14] Lek S, Guegan JF. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological Modelling* 1999;120:65–73.
- [15] Cinar S, Onay TT, Erdinciler A. Co-disposal alternatives of various municipal waste water treatment-plant sludge's with refuse. *Advances in Environmental Research* 2004;8:477–82.
- [16] Wen CH, Vassiliadis CA. Applying hybrid artificial intelligence techniques in waste water treatment. *Engineering Applications of Artificial Intelligence* 1998;11:685–705.
- [17] Daneshvar N, Khataee AR, Djafarzadeh N. The use of artificial neural networks (ANN) for modeling of decolorization of textile dye solution containing C.I. Basic Yellow 28 by electrocoagulation process. *Journal of Hazardous Materials* 2006;B 137:1788–95.
- [18] Duran A, Monteagudo JM. Solar photocatalytic degradation of reactive blue 4 using a Fresnel lens. *Water Research* 2007;41:690–8.
- [19] Aleboyeh A, Moussa Y, Aleboyeh H. Kinetics of oxidative decolourisation of Acid Orange 7 in water by ultraviolet radiation in the presence of hydrogen peroxide. *Separation and Purification Technology* 2005;43: 143–8.
- [20] Hamed MM, Khalafallah MG, Hassanien EA. Prediction of waste water treatment plant performance using artificial neural networks. *Environmental Modelling and Software* 2004;19:919–28.
- [21] Slokar YM, Zupan J, Marechal AML. The use of artificial neural network (ANN) for modeling of the H₂O₂/UV decoloration process: Part I. *Dyes and Pigments* 1999;42:123–35.
- [22] Garson GD. Interpreting neural-network connection weights. *AI Expert* 1991;6(7):47–51.