Optimal performance of COD removal during aqueous treatment of alazine and gesaprim commercial herbicides by direct and inverse neural network

Y. El. Hamzaoui [a], J. A. Hernández [b], S. Silva-Martínez [b,*], A. Bassam [c], A. Álvarez [b], C. Lizama-Bahena [a]

[a] Posgrado en Ingeniería y Ciencias Aplicadas FCQeI-CIICAp, Universidad Autónoma del Estado de Morelos (UAEM), Mexico
[b] Centro de Investigación en Ingeniaría y Ciencias Aplicadas (CIICAp), UAEM. Av. Universidad No. 1001, Col Chamilpa, C.P. 62209, Cuernavaca, Morelos, México
[c] Posgrado en Ingeniería del Centro de Investigación en Energía, Universidad Nacional Autónoma de México, Privada Xochicalco s/n, 62580, Temixco, Morelos, México

ARTICLE INFO

ABSTRACT

A direct and inverse artificial neural network (ANN and ANNi) approach was developed to predict the chemical oxygen demand (COD) removal during the degradation of alazine and gesaprim commercial herbicides under various experimental conditions. The configuration 9–9–1 (9 inputs, 9 hidden and 1 output neurons) presented an excellent agreement ($R^2 = 0.9913$) between experimental and simulated COD value considering the hyperbolic tangent sigmoid and linear transfer function in the hidden layer and output layer. The sensitivity analysis showed that all studied input variables (reaction time, pH, herbicide concentration, contaminant, US ultrasound, UV light intensity, $[TiO_2]_o$, $[K_2S_2O_8]_o$, and SR solar radiation) have strong effect on the degradation of the commercial herbicide in terms of COD removal. In addition, reaction time is the most influential parameter with relative importance of 33.49%, followed by initial herbicide concentration. COD optimal performance was carried out by inverting artificial neural network. Now, ANNi could calculate the optimal unknown parameter (reaction time) to obtain a COD required. Very low percentage of error and short computing makes this methodology attractive to be applied to the on-line control of Advanced Oxidation Process (AOP) over the degradation of commercial herbicide.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

The excessive application of fertilizers and pesticides in the agriculture activity has contributed to water contamination. Herbicides represent the main class of pesticides and belong to the persistent organic pollutants because of their low biodegradability. Herbicides undergo the biotic and abiotic process in soil giving rise to the formation of complex metabolites as degradation products. The metabolites may pose potential human health problems because of their presence in groundwater [1] at relatively high concentrations [2]. This situation has given rise to environmental concerns around the world among the scientific community.

Nowadays, the use of herbicides in the agriculture has increased notoriously throughout the world, particularly during the past two decades and has consequently, led to increasing concern about the environmental fate of these substances because of their relatively slow rate of decomposition [3]. The alazine (comprised by alachlor and atrazine) and gesaprim (atrazine) commercial herbicides have been degraded by the combination of sonolysis and photocatalysis (sonophotocatalysis) using an ultrasound source of 20 kHz under

UV light [4]. It was reported that the photodegradation of these commercial herbicides was enhanced using ultrasound in the presence of TiO_2 catalyst with very high decomposition yields of the active compounds reaching practically a complete mineralization in both commercial herbicides. Furthermore, degradation profiles were recorded by measuring the concentration present in the alazine (alachlor and atrazine) and gesaprim (atrazine) by HPLC as a function of irradiation time (sound and/or light). Over 80% of chemical oxygen demand abatement was attained for both herbicides with sonophotocatalysis at 150 min of irradiation time.

The degradation of alazine and gesaprim commercial herbicides by sonophotocatalysis process 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 sonophotocatalytic degradation involving radical species. Since the process depends on several factors, the modeling of these processes involves many problems, exhibit nonlinear behaviors, which are difficult to describe by linear mathematical model *i.e.* dealing with a multivariate system. It is evident that these problems cannot be solved by simple linear multivariate correlation. However, the developments in artificial neural networks make them possible to be used in complex system modeling.

Artificial neural networks (ANNs) are now commonly used in many research areas of chemistry and represent a set of methods that may be useful in predicting water quality using water treatment parameters [5,6]. Unlike traditional statistical and differential equation approaches, ANNs are considered to be a powerful data modeling tool as it can capture and implicitly represent complex relationships with many variables, such as the input/output variables. Basically, the advantages of neural networks are that they are able to represent both linear and nonlinear relationships and are ingenious to learn the relationships directly from data used for training the network [7,8]. ANN 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 and sonophotocatalytic systems [9–11].

The application of ANN analysis to solve the environmental engineering problems has been the subject of numerous review articles. Gob and others [5], and Moraes and others [6] have reported kinetic modeling of the photochemical water treatment process. The article published by Lek and Guegan [11] describes the application of ANN as a tool in ecological modeling. Cinar and others [12] have determined the interrelationship and response of process variables involved in water treatment plants. They have analyzed the system behavior of a full-scale activated sludge wastewater treatment plant by using Kohonen self organizing feature map neural network. Wen and Vassiliadis [13] have proposed an automatic control system for the operation of the wastewater treatment process by applying hybrid artificial intelligence techniques in real-time control. In addition ANN was applied successfully in biological wastewater treatment [14–17]. Recently, ANN was used for the modeling of the photooxidative decolorization of BB3 [18]. ANN was also used for the modeling of the dye removal by advanced oxidation process [19–21] and for predicting the biochemical oxygen demand as an indicator of river pollution [7]. However, few studies on application of ANN in the advanced oxidation process (AOPs) have been reported [22–26].

Recently, to solve the environmental engineering problems the application of neural networks continues to expand. The present investigation discusses two main ideas, first the use of a multilayer feed-forward neural network model to predict the chemical oxygen demand removal during the degradation of alazine and gesaprim commercial herbicides using titanium dioxide suspensions and the tandem process of sonophotocatalysis under UV light in the advanced oxidation process. On the other hand, in many cases, when an optimal output is required, the optimal input parameters are unknown, for this reason. We found that the inverse artificial neural network is a fundamental strategy to calculate the optimal operation condition. Therefore, a sensitivity analysis was applied to show which parameters have the most influence on COD removal in order to optimize them by means of inverse neural network (ANNi).

2. Materials and methods

2.1. Chemicals

Alazine (30/18 LM), comprised by alachlor, atrazine and formulating agents. Gesaprim (90 GDA) contains atrazine and formulating agents. These herbicides were directly purchased from Syngenta Crop Protection Inc. (USA). TiO_2 (Degussa P25) and H_2SO_4 were analytical grade (Sigma-Aldrich). All chemicals were used as received without further purification. Distilled water was provided by Baxter México S.A.

2.2. Herbicide degradation experiments

The experimental set up used in this work has been previously described in detail elsewhere [4]. A series of photodegradation experiments for each herbicide was performed employing a photochemical reactor operating in a recirculating mode using a volume of 250 ml and a flow rate of $5.63 \, l \, min^{-1}$ (Fig. 1). The photochemical reactor consisted of a jacketed ultrasonic cell ($150 \, cm^3$) containing the ultrasonic probe (500 W, 20 kHz, Cole Parmer). The ultrasonic cell was temperature controlled with water recirculation and a UV lamp (15 W, 352 nm, Cole Parmer). Samples were withdrawn at different degradation time intervals to analyze the concentration of atrazine and alachlor by HPLC.

During sampling, care was taken to withdraw a volume of less than 10% of the total volume. Samples were filtered as collected prior the analysis. Chemical Oxygen Demand (COD) was analyzed using standard methods and standard tubes [27].

2.3. Artificial neural network

The neurons are grouped into distinct layers and interconnected according to a given architecture. As in nature, the network's function is determined largely by the connections between elements (neurons), each connection between two neurons has a weight coefficient attached to it. The standard network structure for an approximation function is the multiple-layer perception (or feed-forward network).

The feed-forward network often has one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple-layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. The linear output layer lets the network produce values outside the -1 to $+1$ range [21]. For the network, the appropriate notation is used in two-layer networks [28].

The number of neurons in the input and output layers is given respectively by the number of input and output variables in the process under investigation. In this work, a feed-forward is proposed, the input layer consists of nine variables (reaction time, pH, herbicide concentration, contaminant, US ultrasound, UV light intensity, $[TiO_2]_o$, $[K_2S_2O_8]_o$, and SR solar radiation), and the output layer contains one variable (COD). The optimal number of neurons in the hidden layer(s) n_s is difficult to specify, and depends on the type and complexity of the task. This number is usually determined iteratively. Each neuron in the hidden layer has a bias b (threshold), which is added to the weighted inputs to form the neuron n (Eq.(1)). This sum, n, is the argument of the transfer function f.

$$n_1 = Wi_{(1,1)}In_1 + Wi_{(1,2)}In_2 + \ldots\ldots + Wi_{(1,k)}In_k + b_1 \tag{1}$$

The coefficients associated with the hidden layer are grouped into matrices Wi (weights) and b1 (biases). The output layer computes the weighted sum of the signals provided by the hidden layer, and the associated coefficients are grouped into matrices W_o and b2. Using the matrix notation, the network output can be given by (Eq.(2)):

$$Out = g(Wo \times f(Wi \times In + b1) + b2) \tag{2}$$

Hidden layer neurons may use any differentiable transfer function to generate their output. In this work, a hyperbolic tangent sigmoid transfer function (TANSIG) on hidden layer with nine neurons and a linear transfer function (PURELIN) on output layer were used for f and g, respectively [29]. The system adjusts the weights of the internal connections to minimize errors between the network output and target output, which can be summarized as follows: At first take a group of random numbers as the initial
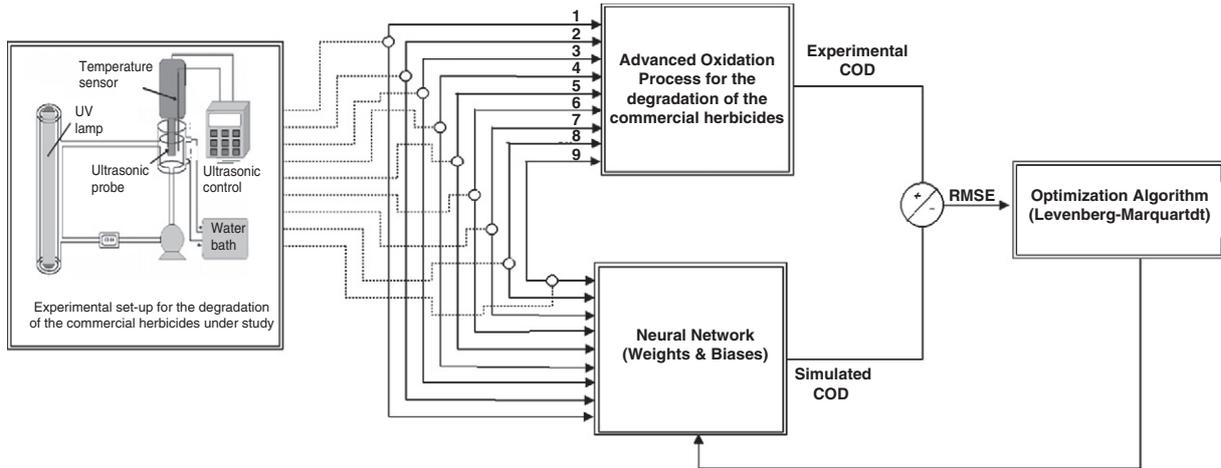
Fig. 1. Recurrent network architecture to the COD values and the procedure used for neural network learning.

values of the weights W and bias b, then compute the output of all neurons layers by layer, starting with the input layer, using the following program:

$$f = TANSIG(W_i{}^*In_k + b1_s) \tag{3}$$

$$f = \frac{2}{1 + exp[-2{}^*(W_i{}^*In_k + b1_s)]} - 1 \tag{4}$$

$$g = PURELIN(W_o{}^*f + b2_l) \tag{5}$$

$$g = Out_l \tag{6}$$

If considering the transfer functions, in the account that, the (Eq. (2)), may be expressed as follows:

$$Out_l = PURELIN\{W_o \times [TANSIG(W_i \times In_k + b1_s)] + b2_l\} \tag{7}$$

where s is the number of neurons in the hidden layer $(S = 9)$, k is the number of neurons in the input layer $(K = 9)$, l is the number of neurons in output layer $(l = 1)$, W_i, W_o and $b1_s$, $b2_l$ are weights and biases, respectively. Eq. (3) is not complex because it is made up of a simple arithmetic operation. Therefore, it can be used for on-line estimation application for industrial processes. In this work, multi-layer feed-forward ANN with one hidden layer was used for all data sets. Database sets were obtained from Bahena and others [4]. The ANN was trained using the backpropagation algorithm. All calculations were carried out with Matlab mathematical software with the ANN toolbox.

2.4. Neural network learning

A learning (or training) algorithm is defined as a procedure that consists of adjusting the coefficients (weights and biases) of a network, to minimize an error function (usually a quadratic one) between the network outputs, for a given set of inputs, and the correct (already known) outputs. If smooth nonlinearities are used, the gradient of the error function can be computed by the classical backpropagation procedure [30]. To determine the best back-propagation training algorithm, ten backpropagation algorithms were studied. In addition, five neurons were used in the hidden layer for all backpropagation algorithms. Table 1 shows a compar-

ison of different backpropagation training algorithms. Levenberg–Marquardt backpropagation training algorithm could have smaller mean square error ($RMSE$), on the other hand, we found training with Levenberg–Marquardt algorithm can run smoothly in computer with lower expanded memory specification (EMS), and the training time is quickly, than the other backpropagation algorithms. Because, the Levenberg–Marquardt algorithm was designed to approach second order training speed without having to compute the Hessian matrix. When the performance function has the form of a sum of squares (as is typical in training feed-forward networks), then the Hessian matrix can be approximated as:

$$H = J^T J \tag{8}$$

And the gradient can be computed as:

$$g = J^T e \tag{9}$$

where J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, and e is a vector of network errors. The Jacobian matrix can be computed through a standard backpropagation technique that is much less complex than computing the Hessian matrix. The Levenberg–Marquardt algorithm uses this approximation to the Hessian matrix in the following Newton like update:

$$X_{k+1} = X_k - [J^T J + \mu I]^{-1} J^T e \tag{10}$$

When the scalar μ is zero, this is just Newton's method, using the approximate Hessian matrix. When μ is large, this becomes gradient descent with a small step size. Newton's method is faster and more accurate near an error minimum, so the aim is to shift toward Newton's method as quickly as possible [31,32], thus μ is decreased after each successful step (reduction in performance function) and is increased only when a tentative step would increase the performance function. In this context, the performance function is always reduced at each iteration of the algorithm [33]. So, for this motivation, the Levenberg–Marquardt algorithm was considered the training algorithm in the present study.

However, the performance of the ANN model was statistically measured by the root mean square error (RMSE) and regression

Table 1
Comparison of 10 backpropagation algorithms with 5 neurons in the hidden layer.

Backpropagation algorithm	Function	Root mean square error (RMSE)	Epoch	Correlation coefficient (R^2)	Best linear equation
Levenberg–Marquardt backpropagation	trainlm	0.00235001	1000	0.990	$Y = 0.990X + 0.306$
Batch gradient descent	traingd	0.01657930	2000	0.988	$Y = 0.986X + 0.927$
Batch gradient descent with momentum	traingdm	0.01982300	2000	0.987	$Y = 0.988X + 0.837$
Polak–Ribiere conjugate gradient backpropagation	traincgp	0.03267010	2000	0.979	$Y = 0.957X + 2.53$
Scaled conjugate gradient backpropagation	trainscg	0.48619601	2000	0.974	$Y = 1.020X - 0.783$
BFGS quasi-Newton backpropagation	trainbfg	0.44944900	2000	0.971	$Y = 0.982X + 1.23$
Powell–Beale conjugate gradient backpropagation	traincgb	0.50820200	2000	0.965	$Y = 0.960X + 2.03$
One step secant backpropagation	trainoss	0.02753301	2000	0.782	$Y = 0.617X + 45.3$
Fletcher–Reeves conjugate gradient backpropagation	traincgf	0.01756320	2000	0.725	$Y = 0.425X + 34.8$
Variable learning rate backpropagation	traingdx	0.02039630	2000	0.718	$Y = 0.386X + 38$

coefficient, which are calculated with the experimental values and network predictions. These calculations are used as a criterion for model adequacy (see Fig. 1), obtained as follows:

$$RMSE = \sqrt{\frac{\left(\sum_{q=1}^{Q}\left(y_{q,pred}-y_{q,exp}\right)^2\right)}{Q}} \tag{11}$$

$$R^2 = 1 - \frac{\sum_{q=1}^{Q}\left(y_{q,pred}-y_{q,exp}\right)^2}{\sum_{q=1}^{Q}\left(y_{q,exp}-y_m\right)^2} \tag{12}$$

where Q is the number of data points, $y_{q,\,pred}$ is the network prediction, $y_{q,\,exp}$ is the experimental response, y_m is the average of actual values and q is an index of data.

Consequently, RMSE was used as the error function which measures the performance of the network. Therefore, the network having minimum RMSE and maximum R^2 was selected the best ANN model.

2.5. ANN model development

Since we mention previously, the input variables to ANN were the reaction time, pH, herbicide concentration, contaminant, US ultrasound, UV light intensity, $[TiO_2]_o$, $[K_2S_2O_8]_o$, SR solar radiation and the chemical oxygen demand was the experimental response or output variable. The characteristics of input and output variables are shown in Table 2.

The topology of an artificial neural network 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 [20].

In order to determine the optimum number of neurons in the hidden layer, a series of topologies was used, in which the number of neurons was varied from 1 to 10. 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 (RMSE) 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 RMSE 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 [34].

Experimental database provided by Bahena and others [4], consists of different COD values, obtained from the photochemical reactor for the degradation of the commercial herbicides under study. The experimental data set was obtained at different parameter process: Reaction time (0–400), pH of the solution (1–5), herbicide concentration (0.1540–0.3090 mM), contaminant (0.1, 0.3, 0.5, 0.7, and 0.9 for each contaminant, respectively (5 contaminants: alachor, chlorobromuron, atrazine, alazine, and gesaprim)), US ultrasound (0–20 kHz), UV light intensity (0–352 nm), initial concentration of TiO_2 (0–300 mg/l), initial concentration of $K_2S_2O_8$ (0–13 mM) and solar radiation intensity (0–820 W/m^2). After 2 h from start up data were collected for 4 h. The experiments were carried out at different initial conditions with at least two replicates. Thus, a database of 275 samples was obtained. These data were sufficient to train and test the ANN model. A summary of the operating parameters is shown in Table 2. The data sets were divided into training, validation and test subsets, each of which contains 138, 69 and 69 samples, respectively. The validation and test sets, for the evaluation of the validation and modeling power of the nets, were randomly selected from the experimental data. Since the transfer function used in the hidden layer was sigmoid, all samples must be normalized in the range of 0.1–0.9 [10]. So, all the input data sets X_i (from the training, validation and test sets) were scaled to a new value x_i as follows:

$$x_i = 0.8\left(\frac{X_i - X_{min}}{X_{max} - X_{min}}\right) + 0.1 \tag{13}$$

The final topology was obtained after 10^5 runs of 1000 iterations start from random initial weights. For each runs, it was computed the network error versus the number of neurons in the hidden layer. Fig. 2 illustrates the network error versus the number of neurons in the hidden layer. It was found that the network performance stabilized after inclusion of nine neurons on hidden layer (9–9–1). So, based on the approximation of RMSE function, a number of neurons in the hidden layer equal to nine, and a three layered feed-forward backpropagation neural network were used for modeling the process as depicted in Fig. 3.

Table 2
Characteristics of input and output variables to the ANN model.

Variable input layer	Range
Reaction time (min)	0–480
pH	1–5
Initial concentration of herbicide (mM)	0.1540–0.3090
Contaminant	0.1–0.9
US Ultrasound (Khz)	0–20
UV light intensity (nm)	0–352
$[TiO_2]_o$ (mg/L)	0–300
$[K_2S_2O_8]_o$ (mM)	0–13
SR solar radiation (W/m^2)	0–820
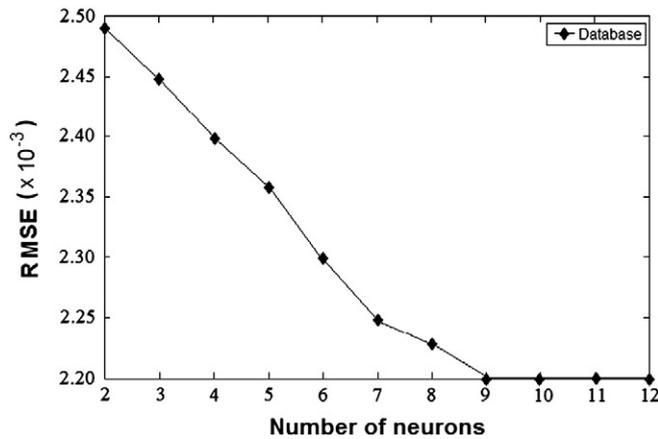
Fig. 2. Effect of the number of neurons in the hidden layer on the performance of the neural network.

3. Results and discussion

3.1. Proposed neural network model

A neural network with nine neurons in the hidden layer (involving 100 coefficients: 90 weights and 10 biases) was found to be efficient in predicting the chemical oxygen demand. Fig. 4 presents a comparison between the experimental and simulated of the COD values using all data available.

Experimental (COD_{EXP}) and simulated (COD_{ANN}) data were compared satisfactorily through a linear regression model ($COD_{ANN} = a + b\,COD_{EXP}$) obtaining a regression coefficient $R^2 = 0.9913$. According to Verma and others [35,36] to satisfy the statistical test of intercept and slope, upper and lower value of the intercepts must contain zero and upper and lower value of the slope must contain one.

Table 3 shows the limits for test indicators, with slope containing the one and with the intercept containing zero. Consequently, the proposed model passed the test with 99% confidence level. This test with information above guarantees that ANN model has a satisfactory level of confidence.

Table 4 gives the obtained parameters (Wi, Wo, $b1$, and $b2$) of the best fit for nine neurons in the hidden layer. These parameters are used in the proposed model to simulate the COD values. Consequently, the proposed ANN model follows Eq. (14):

$$COD = \sum_{s=1}^{S} \left[W_{o(1,s)} \left(\frac{2}{1 + exp\left(-2\left(\sum_{k=1}^{K} \left(W_{i(s,,k)} In_{(k)} \right) + b1_{(s)} \right) \right)} - 1 \right) \right] + b2_{(l)} \tag{14}$$
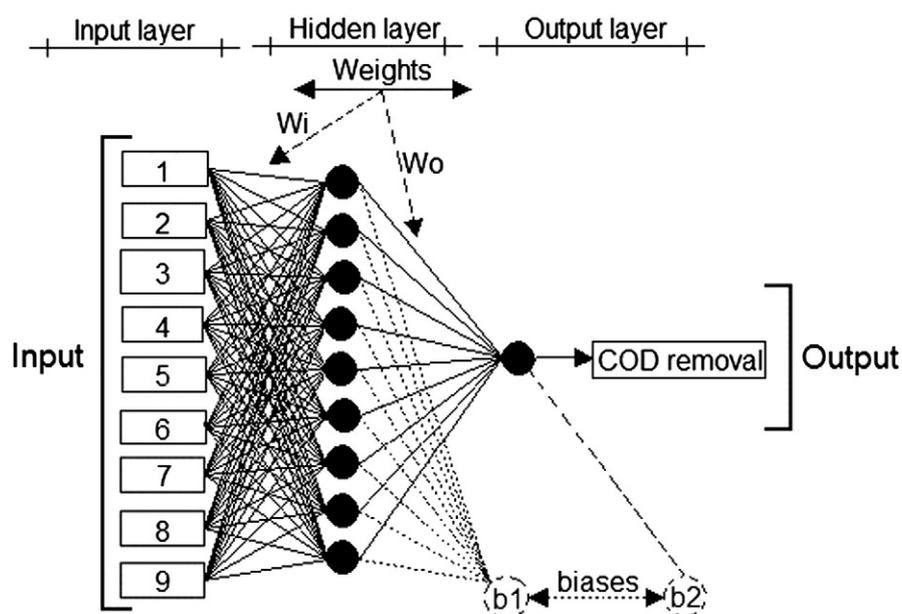
Fig. 3. Model for the prediction of COD values.
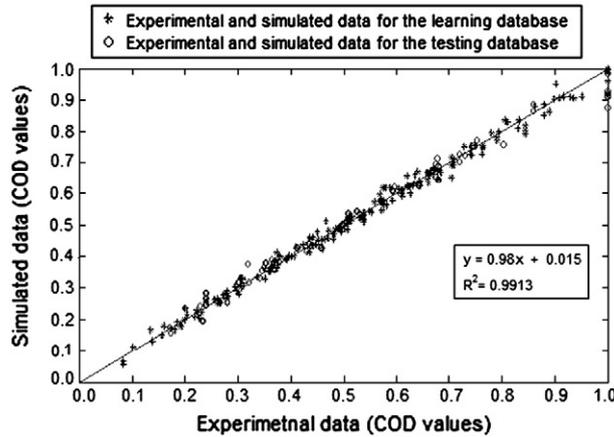
Fig. 4. Comparison of experimental and ANN-predicted values of Chemical Oxygen Demand (COD).

where s is the number of neurons in the hidden layer (S = 9), k is the number of the input (K = 9), and W and b are weight and bias, respectively.

The following (Eq. (15)) gives COD removal during the degradation of alazine and gesaprim commercial herbicides with weights and biases. However, before to develop COD's equation, we have to denote the following variables:

$V_1 = \text{Reaction time}$
$V_2 = pH$
$V_3 = Herbicide\ concentration$
$V_4 = Conta\ minat$
$V_5 = US\ Ultrasound$
$V_6 = UV\ Light\ Intensity$
$V_7 = [TiO_2]_o$
$V_8 = [K_2S_2O_8]_o$
$V_9 = SR\ Solar\ radiation$

$$COD = 2\left[\frac{W_{o(1,1)}}{1+e^{X1}} + \frac{W_{o(1,2)}}{1+e^{X2}} + \frac{W_{o(1,3)}}{1+e^{X3}} + \frac{W_{o(1,4)}}{1+e^{X4}} + \frac{W_{o(1,5)}}{1+e^{X5}} + \frac{W_{o(1,6)}}{1+e^{X6}} + \frac{W_{o(1,7)}}{1+e^{X7}} + \frac{W_{o(1,8)}}{1+e^{X8}} + \frac{W_{o(1,9)}}{1+e^{X9}}\right]\dots$$
$$-\left(W_{o(1,1)} + W_{o(1,2)} + W_{o(1,3)} + W_{o(1,4)} + W_{o(1,5)} + W_{o(1,6)} + W_{o(1,7)} + W_{o(1,8)} + W_{o(1,9)}\right) + b2_{(1)} \tag{15}$$

where

$$X1 = -2\begin{pmatrix} W_{i(1,1)}V_1 + W_{i(1,2)}V_2 + W_{i(1,3)}V_3\dots \\ + W_{i(1,4)}V_4 + W_{i(1,5)}V_5 + W_{i(1,6)}V_6\dots \\ + W_{i(1,7)}V_7 + W_{i(1,8)}V_8 + W_{i(1,9)}V_9 + b1_{(1)} \end{pmatrix} \tag{16}$$

$$X2 = -2\begin{pmatrix} W_{i(2,1)}V_1 + W_{i(2,2)}V_2 + W_{i(2,3)}V_3\dots \\ + W_{i(2,4)}V_4 + W_{i(2,5)}V_5 + W_{i(2,6)}V_6\dots \\ + W_{i(2,7)}V_7 + W_{i(2,8)}V_8 + W_{i(2,9)}V_9 + b1_{(2)} \end{pmatrix} \tag{17}$$

$$X3 = -2\begin{pmatrix} W_{i(3,1)}V_1 + W_{i(3,2)}V_2 + W_{i(3,3)}V_3\dots \\ + W_{i(3,4)}V_4 + W_{i(3,5)}V_5 + W_{i(3,6)}V_6\dots \\ + W_{i(3,7)}V_7 + W_{i(3,8)}V_8 + W_{i(3,9)}V_9 + b1_{(3)} \end{pmatrix} \tag{18}$$

$$X4 = -2\begin{pmatrix} W_{i(4,1)}V_1 + W_{i(4,2)}V_2 + W_{i(4,3)}V_3\dots \\ + W_{i(4,4)}V_4 + W_{i(4,5)}V_5 + W_{i(4,6)}V_6\dots \\ + W_{i(4,7)}V_7 + W_{i(4,8)}V_8 + W_{i(4,9)}V_9 + b1_{(4)} \end{pmatrix} \tag{19}$$

Table 3
Intercept and slope statistical test.

COD	
(Chemical Oxygen Demand)	
a_{lower}	a_{upper}
−0.0008	0.0304
b_{lower}	b_{upper}
0.9503	1.0011

Table 4
Weights and biases for the ANN model.

Wi (s,k)	−0.3166 (1,1)	0.2236 (1,2)	7.2842 (1,3)	2.0881 (1,4)	1.811 (1,5)	3.3471 (1,6)	−1.0447 (1,7)	1.7596 (1,8)	2.0459 (1,9)
	−0.5515 (2,1)	−0.9564 (2,2)	−3.8864 (2,3)	2.0139 (2,4)	−2.867 (2,5)	−1.2945 (2,6)	−0.45 (2,7)	−6.4122 (2,8)	2.1782 (2,9)
	1.4092 (3,1)	−8.0291 (3,2)	−12.0342 (3,3)	15.7374 (3,4)	−3.313 (3,5)	−0.2425 (3,6)	−3.6828 (3,7)	11.6702 (3,8)	−3.6556 (3,9)
	−0.7383 (4,1)	−0.1696 (4,2)	−3.2518 (4,3)	1.2309 (4,4)	−0.3279 (4,5)	−1.0313 (4,6)	−0.541 (4,7)	−2.8458 (4,8)	1.6688 (4,9)
	50.505 (5,1)	−0.1108 (5,2)	6.7242 (5,3)	1.7719 (5,4)	−1.2484 (5,5)	−3.1512 (5,6)	1.6575 (5,7)	−1.4106 (5,8)	−3.0969 (5,9)
	−1.3825 (6,1)	−0.4662 (6,2)	−5.9643 (6,3)	5.8058 (6,4)	0.9334 (6,5)	−2.8553 (6,6)	−1.6653 (6,7)	−3.5912 (6,8)	2.8908 (6,9)
	−29.0201 (7,1)	0.7569 (7,2)	0.4108 (7,3)	0.2687 (7,4)	−0.086 (7,5)	−0.6303 (7,6)	1.0245 (7,7)	−0.7546 (7,8)	−7.7876 (7,9)
	−3.7356 (8,1)	0.0012 (8,2)	0.9109 (8,3)	−0.6673 (8,4)	−0.7864 (8,5)	−1.2263 (8,6)	1.2131 (8,7)	−1.3882 (8,8)	−0.1462 (8,9)
	−5.5039 (9,1)	1.1252 (9,2)	3.23 (9,3)	−2.7131 (9,4)	−0.7353 (9,5)	0.2545 (9,6)	1.9287 (9,7)	−3.0638 (9,8)	1.139 (9,9)
	0.6031	−5.9361	0.6256	7.0453	−6.087	−0.6846	5.127	1.7649	−0.6662
Wo (1,s)	0.6031 (1,1)	−5.9361 (1,2)	0.6256 (1,3)	7.0453 (1,4)	−6.087 (1,5)	−0.6846 (1,6)	5.127 (1,7)	1.7649 (1,8)	−0.6662 (1,9)
b1 (s)	−5.251 (1)								
	5.6355 (2)								
	−11.9996 (3)								
	2.6748 (4)								
	−2.4115 (5)								
	1.2615 (6)								
	1.4984 (7)								
	−0.0617 (8)								
	−0.6785 (9)								
b2 (L)	12.9598 (1)								

$$X5 = -2 \begin{pmatrix} W_{i(5,1)}V_1 + W_{i(5,2)}V_2 + W_{i(5,3)}V_3 ... \\ + W_{i(5,4)}V_4 + W_{i(5,5)}V_5 + W_{i(5,6)}V_6 ... \\ + W_{i(5,7)}V_7 + W_{i(5,8)}V_8 + W_{i(5,9)}V_9 + b1_{(5)} \end{pmatrix} \tag{20}$$

$$X6 = -2 \begin{pmatrix} W_{i(6,1)}V_1 + W_{i(6,2)}V_2 + W_{i(6,3)}V_3 ... \\ + W_{i(6,4)}V_4 + W_{i(6,5)}V_5 + W_{i(6,6)}V_6 ... \\ + W_{i(6,7)}V_7 + W_{i(6,8)}V_8 + W_{i(6,9)}V_9 + b1_{(6)} \end{pmatrix} \tag{21}$$

$$X7 = -2 \begin{pmatrix} W_{i(7,1)}V_1 + W_{i(7,2)}V_2 + W_{i(7,3)}V_3 ... \\ + W_{i(7,4)}V_4 + W_{i(7,5)}V_5 + W_{i(7,6)}V_6 ... \\ + W_{i(7,7)}V_7 + W_{i(7,8)}V_8 + W_{i(7,9)}V_9 + b1_{(7)} \end{pmatrix} \tag{22}$$

$$X8 = -2 \begin{pmatrix} W_{i(8,1)}V_1 + W_{i(8,2)}V_2 + W_{i(8,3)}V_3 .... \\ + W_{i(8,4)}V_4 + W_{i(8,5)}V_5 + W_{i(8,6)}V_6 ... \\ + W_{i(8,7)}V_7 + W_{i(8,8)}V_8 + W_{i(8,9)}V_9 + b1_{(8)} \end{pmatrix} \tag{23}$$

$$X9 = -2 \begin{pmatrix} W_{i(9,1)}V_1 + W_{i(9,2)}V_2 + W_{i(9,3)}V_3 ... \\ + W_{i(9,4)}V_4 + W_{i(9,5)}V_5 + W_{i(9,6)}V_6 ... \\ + W_{i(9,7)}V_7 + W_{i(9,8)}V_8 + W_{i(9,9)}V_9 + b1_{(9)} \end{pmatrix} \tag{24}$$

3.2. Validation of the proposed ANN model

Figs. 5 and 6 depict the ability of the models to predict the chemical oxygen demand of alazine and gesaprim commercial herbicides at different parameters. These figures compare the simulated results with the experimental data for the test database. It can be seen that the model succeeded in predicting the experimental results. As expected, a total COD abatement over 93% and 86% was accomplished by the model for the degradation of Gesaprim and Alazine, respectively, in agreement with the experimental data obtained with the sonophotocatalytic process which appeared to be the most influential degradation method. This shows the importance of the artificial neural network to simulate the chemical oxygen demand of alazine and gesaprim commercial herbicides.
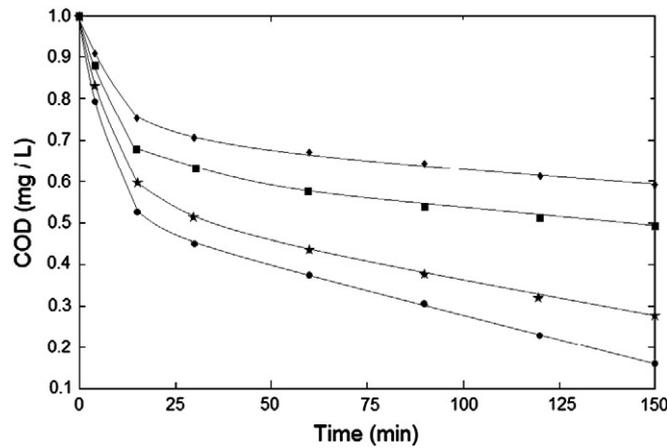
Fig. 5. Experimental data and simulated curve generated with the proposed model of COD abatement of the gesaprim commercial herbicide. ◊UV photolysis, *Sonolyisis, +Photocatalysis, ○Sonophotocatalysis, and the Continuous line is the prediction.

This model is not complex because simulation is realized by simple arithmetic operation, and therefore, it can be used for on-line estimation to predict degradation parameters during the treatment of the commercial herbicides. Thus, the network was tested and validated by comparing its predicted output values with the experimental data using an independent set of data (as shown in Figs. 5 and 6).

3.3. Sensitivity analysis

In order to assess the relative importance of the input variables, two evaluation processes were used. The first one was based on the neural net weight matrix and Garson equation [37]. He proposed an equation based on the partitioning of connection weights:

$$
I_j = \frac{\sum\limits_{m=1}^{Nh} \left(\left(\frac{\left| W_{jm}^{ih} \right|}{\sum\limits_{k=1}^{Ni} \left| W_{km}^{ih} \right|} \right) \times \left| W_{mn}^{ho} \right| \right)}{\sum\limits_{k=1}^{Ni} \left\{ \sum\limits_{m=1}^{Nh} \left(\frac{\left| W_{km}^{ih} \right|}{\sum\limits_{k=1}^{Ni} \left| W_{km}^{ih} \right|} \right) \times \left| W_{mn}^{ho} \right| \right\}}
\tag{25}
$$

where, I_j is the relative importance of the *jth* input variable on the output variable, N_i and N_h are the number of input and hidden neurons, respectively and W is connection weight, the superscripts '*i*', '*h*' and '*o*' refer to input, hidden and output neurons, respectively. Note that the numerator in Eq. (25) describes the sums of absolute products of weights for each input. However, the denominator in Eq. (25), represents the
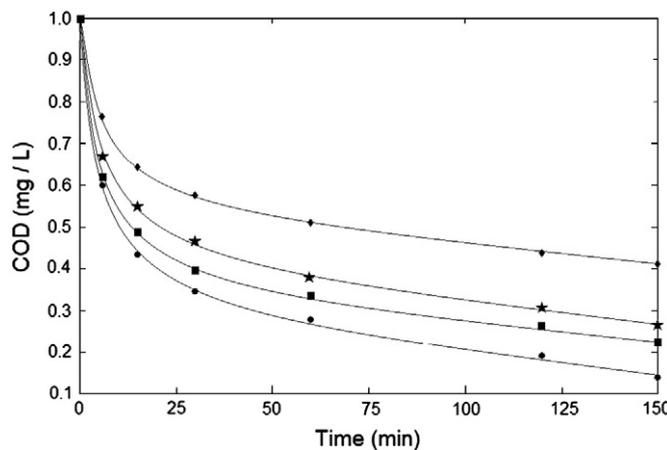
Fig. 6. Experimental data and simulated curve generated with the proposed model of COD abatement of the alazine commercial herbicide. ◊UV photolysis, *Sonolyisis, +Photocatalysis, ○Sonophotocatalysis, and the Continuous line is the prediction.

Table 5
Relative importance of input variables.

Input variable	Importance %
Reaction time	33.49
pH	2.18
Herbicide concentration	15.66
Contaminate	7.34
US ultrasound	5.05
UV light intensity	6.12
$[TiO2]_o$	3.92
$[K_2S_2O_8]_o$	15.34
SR solar radiation	10.91
Total	100

sum of all the weights feeding into hidden unit, taking the absolute values. Thus Eq. (25) can be expressed as follows: For the relative importance of reaction time:

$$\frac{\frac{|W_{i(1,1)}|}{|W_{i(1,1)}| + |W_{i(1,2)}| + \dots + |W_{i(1,9)}|} \times |W_{o(1,1)}| + \dots + \frac{|W_{i(9,1)}|}{|W_{i(9,1)}| + |W_{i(9,2)}| + \dots + |W_{i(9,9)}|} \times |W_{o(9,1)}|}{\frac{|W_{i(1,1)}|}{|W_{i(1,1)}| + \dots + |W_{i(1,9)}|} \times |W_{o(1,1)}| + \dots + \frac{|W_{i(1,2)}|}{|W_{i(1,1)}| + \dots |W_{i(1,9)}|} \times |W_{o(9,1)}| + \dots + \frac{|W_{i(1,8)}|}{|W_{i(1,1)}| + \dots + |W_{i(1,9)}|} \times |W_{o(1,1)}| + \dots + \frac{|W_{i(9,9)}|}{|W_{i(9,1)}| + \dots + |W_{i(9,9)}|} \times |W_{o(1,9)}|}$$

$$(26)$$

In the same way, we could find the relative importance about others inputs parameters. However, in this case it is not done because it would be enlarged this article very much.

Table 5 and Fig. 7 show the relative importance of the input variables calculated by Eq. (25). All variables have strong effect on herbicide degradation in terms of COD removal. However, as expected, the reaction time with a relative importance of 33.49% appeared to be the most influential parameter in the degradation process.

3.4. Optimal performance by mean of ANNi

According to ANN model (Eq.(15)), it is possible to simulate the chemical oxygen demand (COD) removal during the degradation of alazine and gesaprim commercial herbicide, when input parameters are well known. Since we found, that the reaction time is the most influential parameter. Therefore, it is important to know in this process, what optimal reaction time is needed for a required COD. Consequently, we developed a strategy to estimate the optimal reaction time in the degradation process from the inverse artificial neural network (ANNi). The proposed method (ANNi) inverts the artificial neural network (Eq.(35)). Then, we have the following equation that calculates COD removal during the degradation process of alazine and gesaprim.

The key information (optimal performance) for the chemical oxygen demand (COD) removal during the degradation of alazine and gesaprim commercial herbicide, when controlling the required output is to know the optimal input parameters. An inverted ANN can be considered [38,39] as a model based method of supervisory control, the control action in which the unknown input parameters are obtained by solving an on-line optimization problem for the desired output.
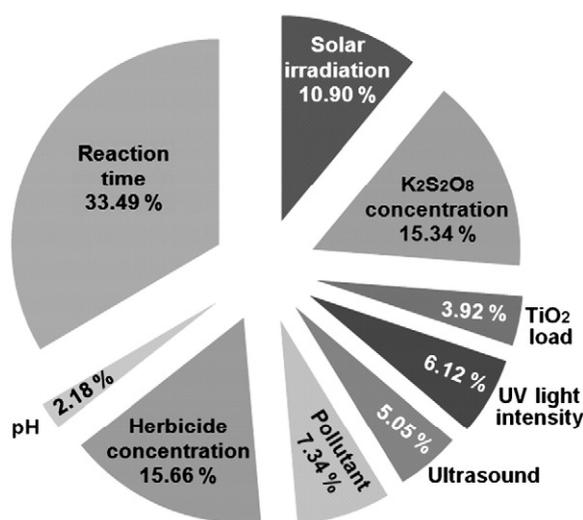
Fig. 7. Relative importance (%) of input variables on the value of chemical oxygen demand.

A general network (shown in Fig. 3) is constituted by TANSIG and PURELIN transfer function. Then, the output is given step by step procedure for ANNi will be presented bellow in order to avoid any ambiguity

$$COD = PURELIN\left(\sum_s \left\{ W_{o(1,s)} \cdot \left[TANSIG\left(W_{i(s,k)} \cdot In_{(k)} + b1_{(s)} \right) \right] \right\} + b2_{(1)} \right) \tag{27}$$

$$COD = \sum_s \left\{ W_{o(1,s)} \cdot \left[\frac{2}{1 + e^{-2 \cdot \left(b1_{(s)} + \sum_k W_{i(s,k)} \cdot In_{(k)} \right)}} - 1 \right] \right\} + b2_{(1)} \tag{28}$$

Eq. (28) can be expressed as Eq. (29). Then, we have:

$$COD = b2_{(1)} - \sum_s W_{o(1,s)} + \sum_s \left[\frac{2 \cdot W_{o(1,s)}}{1 + e^{-2 \cdot \left(b1_{(s)} + \sum_k W_{i(s,k)} \cdot In_{(k)} \right)}} \right] \tag{29}$$

At this step, we have obtained the function which has to be optimized to get the optimal input parameter(s) $In_{(k=x)}$:

$$Fun\left(In_{(x)} \right) = b2_{(1)} - \sum_s W_{o(1,s)} + \sum_s \left[\frac{2 \cdot W_{o(1,s)}}{1 + e^{-2 \cdot \left(W_{i(s,k)} \cdot In_{(x)} + \sum_{k \neq x} W_{i(s,k)} \cdot In_{(k)} + b1_{(s)} \right)}} \right] \tag{30}$$

where x is the reaction time value to be computed. It is important to note that the analytical solution with one neuron in the hidden layer neural model exists, and it has been already described [38]. Nevertheless, in the case that a proposed ANN model has more than one neuron in the hidden layer it is necessary to use an optimization method [38–41]. In this investigation, we used an optimization method to apply the inverse artificial neural network with nine neurons in the hidden layer. Optimization of the input parameter(s) is done by employing the Nelder–Mead simplex algorithms for unconstrained optimization of nonlinear functions [42,43]. The Nelder–Mead method attempts to minimize a multivariable objective nonlinear function using only function values, without any derivative information($f : \Re^n \rightarrow \Re$). The Nelder–Mead method is a membership of direct search methods, that doesn't use numerical o analytic gradient.

Two tests were performed with different data to optimize the reaction time (t) in different conditions to demonstrate the feasibility of this method about ANNi. However, the simulation outcomes were then compared with experimental data in order to check the accuracy of ANNi. This error is given by:

$$Err = 100 \frac{|Exp - Sim|}{Exp} \tag{31}$$

Case 1. A set of parameters are available for sonophotocatalysis process of alazine herbicide with $l = 1$, $s = 9$ and $k = 9$. The experimental conditions for this test, for a required output value COD = 0.138 mg/L, with input values: pH = 2.3, [Alazine]$_o$ = 0.203 mM, Alazine = 0.7, US Ultrasound = 20 kHz, UV light intensity = 352 nm, [TiO2]$_o$ = 200 mg/L, (t = ?) is the reaction time to be found:

$$\begin{aligned}
f(t) = -A &+ \frac{2W_{o(1,1)}}{1 + e^{(X_1 + 0.6332t)}} + \frac{2W_{o(1,2)}}{1 + e^{(X_2 + 1.103t)}} + \frac{2W_{o(1,3)}}{1 + e^{(X_3 - 2.8184t)}} \cdots \\
&+ \frac{2W_{o(1,4)}}{1 + e^{(X_4 + 1.4766t)}} + \frac{2W_{o(1,5)}}{1 + e^{(X_5 - 101.01t)}} + \frac{2W_{o(1,6)}}{1 + e^{(X_6 + 2.765t)}} \cdots \\
&+ \frac{2W_{o(1,7)}}{1 + e^{(X_7 + 58.0402t)}} + \frac{2W_{o(1,8)}}{1 + e^{(X_8 + 7.4712t)}} + \frac{2W_{o(1,9)}}{1 + e^{(X_9 + 11.0078t)}}
\end{aligned} \tag{32}$$

where

$$\begin{aligned}
A = COD - b2_{(1)} &+ W_{o(1,1)} + W_{o(1,2)} \cdots \\
&+ W_{o(1,3)} + W_{o(1,4)} + W_{o(1,5)} + W_{o(1,6)} \cdots \\
&+ W_{o(1,7)} + W_{o(1,8)} + W_{o(1,9)}
\end{aligned} \tag{33}$$

$$X_1 = -2 \begin{pmatrix} W_{i(1,2)}V_2 + W_{i(1,3)}V_3 + W_{i(1,4)}V_4 \cdots \\ + W_{i(1,5)}V_5 + W_{i(1,6)}V_6 + W_{i(1,7)}V_7 \cdots \\ + W_{i(1,8)}V_8 + W_{i(1,9)}V_9 + b1_{(1)} \end{pmatrix} \tag{34}$$

$$X_2 = -2 \begin{pmatrix} W_{i(2,2)}V_2 + W_{i(2,3)}V_3 + W_{i(2,4)}V_4 \cdots \\ + W_{i(2,5)}V_5 + W_{i(2,6)}V_6 + W_{i(2,7)}V_7 \cdots \\ + W_{i(2,8)}V_8 + W_{i(2,9)}V_9 + b1_{(2)} \end{pmatrix} \tag{35}$$

$$X_3 = -2 \begin{pmatrix} W_{i(3,2)}V_2 + W_{i(3,3)}V_3 + W_{i(3,4)}V_4 \cdots \\ + W_{i(3,5)}V_5 + W_{i(3,6)}V_6 + W_{i(3,7)}V_7 \cdots \\ + W_{i(3,8)}V_8 + W_{i(3,9)}V_9 + b1_{(3)} \end{pmatrix} \tag{36}$$

$$X_4 = -2 \begin{pmatrix} W_{i(4,2)}V_2 + W_{i(4,3)}V_3 + W_{i(4,4)}V_4 \ldots \\ + W_{i(4,5)}V_5 + W_{i(4,6)}V_6 + W_{i(4,7)}V_7 \ldots \\ + W_{i(4,8)}V_8 + W_{i(4,9)}V_9 + b1_{(4)} \end{pmatrix} \qquad (37)$$

$$X_5 = -2 \begin{pmatrix} W_{i(5,2)}V_2 + W_{i(5,3)}V_3 + W_{i(5,4)}V_4 \ldots \\ + W_{i(5,5)}V_5 + W_{i(5,6)}V_6 + W_{i(5,7)}V_7 \ldots \\ + W_{i(5,8)}V_8 + W_{i(5,9)}V_9 + b1_{(5)} \end{pmatrix} \qquad (38)$$

$$X_6 = -2 \begin{pmatrix} W_{i(6,2)}V_2 + W_{i(6,3)}V_3 + W_{i(6,4)}V_4 \ldots \\ + W_{i(6,5)}V_5 + W_{i(6,6)}V_6 + W_{i(6,7)}V_7 \ldots \\ + W_{i(6,8)}V_8 + W_{i(6,9)}V_9 + b1_{(6)} \end{pmatrix} \qquad (39)$$

$$X_7 = -2 \begin{pmatrix} W_{i(7,2)}V_2 + W_{i(7,3)}V_3 + W_{i(7,4)}V_4 \ldots \\ + W_{i(7,5)}V_5 + W_{i(7,6)}V_6 + W_{i(7,7)}V_7 \ldots \\ + W_{i(7,8)}V_8 + W_{i(7,9)}V_9 + b1_{(7)} \end{pmatrix} \qquad (40)$$

$$X_8 = -2 \begin{pmatrix} W_{i(8,2)}V_2 + W_{i(8,3)}V_3 + W_{i(8,4)}V_4 \ldots \\ + W_{i(8,5)}V_5 + W_{i(8,6)}V_6 + W_{i(8,7)}V_7 \ldots \\ + W_{i(8,8)}V_8 + W_{i(8,9)}V_9 + b1_{(8)} \end{pmatrix} \qquad (41)$$

$$X_9 = -2 \begin{pmatrix} W_{i(9,2)}V_2 + W_{i(9,3)}V_3 + W_{i(9,4)}V_4 \ldots \\ + W_{i(9,5)}V_5 + W_{i(9,6)}V_6 + W_{i(9,7)}V_7 \ldots \\ + W_{i(9,8)}V_8 + W_{i(9,9)}V_9 + b1_{(9)} \end{pmatrix} \qquad (42)$$

And with weights and biases in Table 4, we can calculate the optimum reaction time of the process for the required output, using Matlab software with the optimization Toolbox [29]. The simulated by ANNi outcome value (t) was 146.72 min, furthermore, by using Eq. (31) the calculated value has an error of 1.56% regarding the experimental result. Consequently, the COD error between the experimental and simulated by ANNi is 0.72%.

Case 2. A set of parameters are available for photocatalysis process of gesaprim herbicide with: $l=1$, $s=9$ and $k=9$. The experimental conditions for this test, for a required output value COD = 0.138 mg/L, with input values: pH = 2.3, [Gesaprim]$_o$ = 0.193 mM, Gesaprim = 0.9, UV light intensity = 352 nm, [TiO2]$_o$ = 200 mg/L, (t = ?) is the reaction time to be found:

$$\begin{aligned} g(t) = &-B + \frac{2W_{o(1,1)}}{1 + e^{(X_1 + 0.6332t)}} + \frac{2W_{o(1,2)}}{1 + e^{(X_2 + 1.103t)}} + \frac{2W_{o(1,3)}}{1 + e^{(X_3 - 2.8184t)}} \cdots \\ &+ \frac{2W_{o(1,4)}}{1 + e^{(X_4 + 1.4766t)}} + \frac{2W_{o(1,5)}}{1 + e^{(X_5 - 101.01t)}} + \frac{2W_{o(1,6)}}{1 + e^{(X_6 + 2.765t)}} \cdots \\ &+ \frac{2W_{o(1,7)}}{1 + e^{(X_7 + 58.0402t)}} + \frac{2W_{o(1,8)}}{1 + e^{(X_8 + 7.4712t)}} + \frac{2W_{o(1,9)}}{1 + e^{(X_9 + 11.0078t)}} \end{aligned} \qquad (43)$$

where

$$\begin{aligned} B = &COD - b2_{(1)} + W_{o(1,1)} + W_{o(1,2)} \cdots \\ &+ W_{o(1,3)} + W_{o(1,4)} + W_{o(1,5)} + W_{o(1,6)} \cdots \\ &+ W_{o(1,7)} + W_{o(1,8)} + W_{o(1,9)} \end{aligned} \qquad (44)$$

$$X_1 = -2 \begin{pmatrix} W_{i(1,2)}V_2 + W_{i(1,3)}V_3 + W_{i(1,4)}V_4 \ldots \\ + W_{i(1,5)}V_5 + W_{i(1,6)}V_6 + W_{i(1,7)}V_7 \ldots \\ + W_{i(1,8)}V_8 + W_{i(1,9)}V_9 + b1_{(1)} \end{pmatrix} \qquad (45)$$

$$X_2 = -2 \begin{pmatrix} W_{i(2,2)}V_2 + W_{i(2,3)}V_3 + W_{i(2,4)}V_4 \ldots \\ + W_{i(2,5)}V_5 + W_{i(2,6)}V_6 + W_{i(2,7)}V_7 \ldots \\ + W_{i(2,8)}V_8 + W_{i(2,9)}V_9 + b1_{(2)} \end{pmatrix} \qquad (46)$$

$$X_3 = -2 \begin{pmatrix} W_{i(3,2)}V_2 + W_{i(3,3)}V_3 + W_{i(3,4)}V_4 \ldots \\ + W_{i(3,5)}V_5 + W_{i(3,6)}V_6 + W_{i(3,7)}V_7 \ldots \\ + W_{i(3,8)}V_8 + W_{i(3,9)}V_9 + b1_{(3)} \end{pmatrix} \qquad (47)$$

$$X_4 = -2 \begin{pmatrix} W_{i(4,2)}V_2 + W_{i(4,3)}V_3 + W_{i(4,4)}V_4 \ldots \\ + W_{i(4,5)}V_5 + W_{i(4,6)}V_6 + W_{i(4,7)}V_7 \ldots \\ + W_{i(4,8)}V_8 + W_{i(4,9)}V_9 + b1_{(4)} \end{pmatrix} \qquad (48)$$

$$X_5 = -2 \begin{pmatrix} W_{i(5,2)}V_2 + W_{i(5,3)}V_3 + W_{i(5,4)}V_4 \ldots \\ + W_{i(5,5)}V_5 + W_{i(5,6)}V_6 + W_{i(5,7)}V_7 \ldots \\ + W_{i(5,8)}V_8 + W_{i(5,9)}V_9 + b1_{(5)} \end{pmatrix} \qquad (49)$$

$$X_6 = -2 \begin{pmatrix} W_{i(6,2)}V_2 + W_{i(6,3)}V_3 + W_{i(6,4)}V_4... \\ + W_{i(6,5)}V_5 + W_{i(6,6)}V_6 + W_{i(6,7)}V_7... \\ + W_{i(6,8)}V_8 + W_{i(6,9)}V_9 + b1_{(6)} \end{pmatrix} \qquad (50)$$

$$X_7 = -2 \begin{pmatrix} W_{i(7,2)}V_2 + W_{i(7,3)}V_3 + W_{i(7,4)}V_4... \\ + W_{i(7,5)}V_5 + W_{i(7,6)}V_6 + W_{i(7,7)}V_7... \\ + W_{i(7,8)}V_8 + W_{i(7,9)}V_9 + b1_{(7)} \end{pmatrix} \qquad (51)$$

$$X_8 = -2 \begin{pmatrix} W_{i(8,2)}V_2 + W_{i(8,3)}V_3 + W_{i(8,4)}V_4... \\ + W_{i(8,5)}V_5 + W_{i(8,6)}V_6 + W_{i(8,7)}V_7... \\ + W_{i(8,8)}V_8 + W_{i(8,9)}V_9 + b1_{(8)} \end{pmatrix} \qquad (52)$$

$$X_9 = -2 \begin{pmatrix} W_{i(9,2)}V_2 + W_{i(9,3)}V_3 + W_{i(9,4)}V_4... \\ + W_{i(9,5)}V_5 + W_{i(9,6)}V_6 + W_{i(9,7)}V_7... \\ + W_{i(9,8)}V_8 + W_{i(9,9)}V_9 + b1_{(9)} \end{pmatrix} \qquad (53)$$

According to the weights and biases shown in Table 4, we can calculate the optimum reaction time of the process for the required output. The outcome value (t) was 119.67 min, furthermore, by using Eq. (31) the calculated value has an error of 0.18% with regard to experimental result. Consequently, the COD error between the experimental and simulated by ANNi is 3.94%.

4. Conclusion

The chemical oxygen demand parameter of alazine and gesaprim commercial herbicides during their photodegradation was successfully predicted by applying a three layered neural network with nine neurons in the hidden layer, and using backpropagation algorithm. Simulations based on the ANN model were performed in order to estimate the behavior of the system under different conditions. The results obtained by ANN model show high agreement with experimental results: very good correlation ($R^2 > 0.99$) and small error ($RMSE = 0.000259$).

Very high level of confidence for the ANN model was confirmed with the intercept and slope statistical test (99%). Moreover, this paper proposes, also, a methodology to calculate the optimum operating conditions from ANNi when it is required to obtain an optimum result (optimal performance). In this case, the chemical oxygen demand behavior was simulated by an artificial neural network. This model considers as well-known input parameters: reaction time, pH, herbicide concentration, contaminant, US ultrasound, UV light intensity, $[TiO_2]_o$, $[K_2S_2O_8]_o$ and solar radiation. According to the sensitivity analysis, we found that, the reaction time is the most influential parameter. Therefore, from an optimum chemical oxygen demand value as the output variable, and taking into account the above well-known input values excepting the reaction time, it is possible to calculate the optimal reaction time by the ANNi, considering the Nelder–Mead simplex method of optimization. Nevertheless, the mathematical validation of ANNi was carried out in Case 1 and Case 2.

In addition, once the optimum reaction time is calculated. It is possible to simulate the other output variable. Therefore, by this methodology, we are able to obtain any unknown input variable on-line.

Through this flexibility appears to be one of the main characteristics of ANNi system, enhancing its huge interest as a tool for engineering process. Indeed, it is very important to remark, that the elapsed time to calculate the optimum input parameter (reaction time) is minor than 0.3 s.

Hence, the best advantage about ANNi methodology is simple in structure and faster convergence to predict optimal parameters, and will be useful for the on-line optimal control in the framework of the advanced oxidation process.

Acknowledgements

References

[1] W.E. Pereira, C.E. Rostad, Occurrence distribution and transport of herbicides and their degradation products in the lower Mississippi River and its tributaries, Environmental Science & Technology 24 (9) (1990) 1400–1406.

[2] E. Funari, L. Donati, D. Sandroni, M. Vighi, Pesticide levels in groundwater: value and limitations of monitoring, in: M. Vighi, E. Funari (Eds.), Pesticide Risk in Groundwater, Lewis Publishers, 1995, pp. 3–43.

[3] E. Evgenidou, K. Fytianos, Photodegradation of trazine herbicides in aqueous solutions and natural waters, Journal of Agricultural and Food Chemistry 50 (2002) 6423–6427.

[4] C. Lizama-Bahena, S. Silva-Martínez, D. Morales-Guzmán, M.R. Trejo-Hernández, Sonophotocatalytic degradation of alazine and gesaprim commercial herbicides in TiO₂ slurry, Chemosphere 71 (2008) 982–989.

[5] S. Gob, E. Oliveros, S.H. Bossmann, A.M. Braun, R. Guardani, C.A.O. Nascimento, Modeling the kinetics of a photochemical water treatment process by means of artificial neural networks, Chemical Engineering and Processing 38 (1999) 373–382.

[6] J.E.F. Moraes, F.H. Quina, C.A.O. Nascimeto, D.N. Silva, O. Chiavone-Filho, Treatment of saline waste water contaminated with hydrocarbons by the photo-Fenton process, Environmental Science & Technology 38 (2004) 1183–1187.

[7] A. Talib, Y.A. Hasan, N.N.A. Rahman, Predicting biochemical oxygen demand as indicator of river pollution using artificial neural networks, 18th World IMACS/ MODSIM Congress Cairns, Australia 13–17, 2009. http://mssanz.org.au/modsim09 (consulted February 2010).

[8] W.W. McCulloch, W.A. Pitts, A logical calculus of ideas imminent in nervous activity, The Bulletin of Mathematical Biophysics 5 (1943) 115–133.

[9] V.K. Pareek, M.P. Brungs, A.A. Adesina, R. Sharma, Artificial neural network modeling of a multiphase photodegradation system, Journal of Photochemistry and Photobiology A: Chemistry 149 (2002) 139–146.

[10] F. Despange, D.L. Massart, Neural networks in multivariate calibration, The Analyst (1998) 157–178.

[11] S. Lek, J.F. Guegan, Artificial neural networks as a tool in ecological modelling, an introduction, Ecological Modeling 120 (1999) 65–73.

[12] S. Cinar, T.T. Onay, A. Erdincler, Co-disposal alternatives of various municipal waste water treatment-plant sludge's with refuse, Advances in Environmental Research 8 (2004) 477–482.

[13] C.H. Wen, C.A. Vassiliadis, Applying hybrid artificial intelligence techniques in waste water treatment, Engineering Applications of Artificial Intelligence 11 (1998) 685–705.

[14] A.R. Khataee, M.B. Kasiri, Modeling of biological water and wastewater treatment processes using artificial neural networks, Clean-Soil, Air, Water Inpress, 2011.

[15] A.R. Khataee, M. Kasiri, Artificial neural networks modeling of contaminated water treatment by homogeneous and heterogeneous nanocatalysis, Journal of Molecular Catalysis A: Chemical Inpress (2010).

[16] A.R. Khataee, D. Salari, N. Daneshvar, F. Aghazadeh, Application of artificial neural networks for modeling of the treatment of wastewater contaminated with Methyl tertButyl Ether (MTBE) by UV/H2O2 process, Journal of Hazardous Materials 125 (1–3) (2005) 205–210.

[17] A.R. Khataee, L. Alidokht, A. Reyhanitabar, S. Oustan, Reductive removal of Cr(VI) by starch-stabilized FeO nanoparticles in aqueous solution, Desalination 270 (1–3) (2011) 105–110.

[18] A.R. Khataee, O. Mirzajani, UV/peroxydisulfate oxidation of C. I. Basic Blue 3: modeling of key factors by artificial neural network, Desalination 251 (2010) 64–69.

[19] N. Daneshvar, A.R. Khataee, N. Djafarzadeh, 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 137 (2006) 1788–1795.

[20] A. Duran, J.M. Monteagudo, Solar photocatalytic degradation of reactive blue 4 using a Fresnel lens, Water Research 41 (2007) 690–698.

[21] A. Aleboyeh, M.B. Kasiri, M.E. Olya, H. Aleboyeh, Prediction of azo dye decolorization by UV/H2O2 using artificial neural networks, Dyes and Pigments 77 (2) (2008) 288–294.

[22] E. Oguza, A. Tortum, B. Keskinler, Determination of the apparent rate constants of the degradation of humic substances by ozonation and modeling of the removal of humic substances from the aqueous solutions with neural network, Journal of Hazardous Materials 157 (2008) 455–463.

[23] A. Durán, J.M. Monteagudo, M. Mohedano, Neural networks simulation of photo-Fenton degradation of Reactive Blue 4, Appl. Catal. B: Environ. 65 (2006) 127–134.

[24] E.S. Elmolla, M. Chaudhuri, M.M. Eltoukhy, The Use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton Process, Journal of Hazardous Materials 179 (1–3) (2010) 127–134.

[25] A.R. Khataee, M. Zarei, M. Fathinia, M. Khobnasab Jafari, Photocatalytic degradation of an anthraquinone dye on immobilized TiO2 nanoparticle in a rectangular reactor: destruction pathway and response surface approach, Desalination 268 (1–3) (2011) 126–133.

[26] A.R. Khataee, M. Zarei, Photocatalysis of a dye solution using immobilized ZnO nanoparticles combined with photoelectrochemical process, Desalination Inpress, 2011.

[27] Hach Company, Water Analysis Handbook, Loveland, Colorado, 1992.

[28] Neural Networks in Computer Intelligence, in: F. Limin (Ed.), McGraw-Hill International Series in Computer Science, 1994.

[29] H. Demuth, M. Beale (Eds.), Neural Network Toolbox for Matlab — User's Guide Version 3, MathWorks, MA, 1998.

[30] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning internal representations by error propagation, Parallel Data Processing 1 (1986) 318–362.

[31] M.T. Hagan, M.B. Menhaj, Training feedforward networks with the Marquardt algorithm, IEEE Transactions on Neural Network 5 (6) (1994) 989–993.

[32] J.A. Hernández-Pérez, M.A. Garcia-Alavarado, G. Trystram, B. Heyd, Neural networks for the heat and mass transfer prediction during drying of cassava and mango, Innovative Food Science and Emerging Technology 5 (2004) 57–64.

[33] A.R. Khataee, M.B. Kasiri, Photocatalytic degradation of organic dyes in the presence of nanostructured titanium dioxide: influence of the chemical structure of dyes, Journal of Molecular Catalysis A: Chemical 328 (1–2) (2010) 8–26.

[34] M.M. Hamed, M.G. Khalafallah, E.A. Hassanien, Prediction of waste water treatment plant performance using artificial neural networks, Environmental Modelling and Software 19 (2004) 919–928.

[35] S.P. Verma, J. Andaverde, E. Santoyo, Application of the error propagation theory in estimates of static formation temperatures in geothermal and petroleum boreholes April, Heat Transfer in Components and Systems for Sustainable Energy Technologies: Heat-SET 5–7, 2005 Grenoble, France, 47 pp. 3659–3671.

[36] S.P. Verma, J. Andaverde, E. Santoyo, Application of the error propagation theory in estimates of static formation temperatures in geothermal and petroleum boreholes, Energy Conversion and Management 47 (2006) 3659–3671.

[37] G.D. Garson, Interpreting neural-network connection weights, AI Expert 6 (1991) 47–51.

[38] J.A. Hernández, Use of neural networks and neural network inverse in optimizing food processes, CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources 4 (61) (2009) 1–11.

[39] J.A. Hernández, Optimum operating conditions for heat and mass transfer in foodstuffs drying by means of neural network inverse, Food Control 20 (2009) 435–438.

[40] J.A. Hernández, B. Heyd, G. Trystram, Prediction of brightness and surface area kinetics during coffee roasting, Journal of food Engineering 89 (2008) 156–163.

[41] O. Cortés, G. Urquiza, J.A. Hernandez, Optimization of operating conditions for compressor performance by means of neural networks inverse, Applied Energy 86 (11) (2009) 2487–2493.

[42] J.A. Hernández, D. Juárez-Romero, L.I. Morales, J. Siqueiros, COP prediction for the integration of a water purification process in a heat transformer: with and without energy recycling, Desalination 219 (2008) 66–80.

[43] J.C. Lagarias, J.A. Reeds, M.H. Wright, P.E. Wright, Convergence properties of the Nelder–Mead simplex method in low dimensions, SIAM Journal of Optimization 9 (1998) 112–147.