Journal of Hazardous Materials 179 (2010) 127–134

Contents lists available at ScienceDirect

Journal of Hazardous Materials

journal homepage: www.elsevier.com/locate/jhazmat

# The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process

Emad S. Elmolla[a,*], Malay Chaudhuri[a], Mohamed Meselhy Eltoukhy[b]

[a] Dept. of Civil Engineering, Universiti Teknologi PETRONAS (UTP), Bandar Seri Iskandar, 31750 Tronoh, Perak, Malaysia
[b] Dept. of Electical and Elctronics Engineering, Universiti Teknologi PETRONAS (UTP), Bandar Seri Iskandar, 31750 Tronoh, Perak, Malaysia

ARTICLE INFO

*Article history:*
Received 16 October 2009
Received in revised form 11 January 2010
Accepted 21 February 2010
Available online 1 March 2010

*Keywords:*
Artificial neural networks
Antibiotics
Amoxicillin
Ampicillin
Cloxacillin
Fenton process

ABSTRACT

The study examined the implementation of artificial neural network (ANN) for the prediction and simulation of antibiotic degradation in aqueous solution by the Fenton process. A three-layer backpropagation neural network was optimized to predict and simulate the degradation of amoxicillin, ampicillin and cloxacillin in aqueous solution in terms of COD removal. The configuration of the backpropagation neural network giving the smallest mean square error (MSE) was three-layer ANN with tangent sigmoid transfer function (*tansig*) at hidden layer with 14 neurons, linear transfer function (*purelin*) at output layer and Levenberg–Marquardt backpropagation training algorithm (LMA). ANN predicted results are very close to the experimental results with correlation coefficient ($R^2$) of 0.997 and MSE 0.000376. The sensitivity analysis showed that all studied variables (reaction time, $H_2O_2$/COD molar ratio, $H_2O_2/Fe^{2+}$ molar ratio, pH and antibiotics concentration) have strong effect on antibiotic degradation in terms of COD removal. In addition, $H_2O_2/Fe^{2+}$ molar ratio is the most influential parameter with relative importance of 25.8%. The results showed that neural network modeling could effectively predict and simulate the behavior of the Fenton process.



## 1. Introduction

Antibiotics are hazardous contaminants in the aquatic environment because of their adverse effects on aquatic life and humans. Problem that may be created by the presence of antibiotics at low concentration in the environment is the development of antibiotic resistant bacteria [1]. Advanced oxidation processes (AOPs) have proved to be highly effective for the removal of most of the pollutants in wastewaters [2]. Application of advanced oxidation processes for degradation of amoxicillin, ampicillin and cloxacillin antibiotics has been reported [3–7]. Oxidation with Fenton's reagent is based on ferrous ions, hydrogen peroxide and hydroxyl radicals produced by the catalytic decomposition of hydrogen peroxide in acidic solution [8]. Treating of wastewater by AOPs is quite complex, since the process is influenced by several factors. Due to complexity of the process, it is difficult to be modeled and simulated using conventional mathematical modeling. Artificial neural networks (ANNs) are now used in many areas of science and engineering and considered as promising tool because of their simplicity towards simulation, prediction and modelling [9]. The advantages of ANN are that the mathematical description of the phenomena involved in the process is not required, less time is required for model development than the traditional mathematical models and prediction ability with limited numbers of experiments [10]. Application of ANN to solve environmental engineering problems has been reported in many articles. ANNs were applied in biological wastewater treatment [11–26] and physicochemical wastewater treatment [27–30]. However, few studies on applications of ANN in advanced oxidation processes (AOPs) have been reported [31–33].

The present work investigated the implementation of ANN for the prediction of antibiotic degradation in terms of COD removal by the Fenton process. The ANN modelling outputs were compared with the experimental data.

## 2. Materials and methods

### 2.1. Chemicals and antibiotics

Hydrogen peroxide (30%, w/w) and ferrous sulphate heptahydrate ($FeSO_4{\cdot}7H_2O$) were purchased from R & M Marketing, Essex, U.K. Analytical grade of amoxicillin (AMX) and ampicillin (AMP) were purchased from Sigma and cloxacillin (CLX) from Fluka to construct HPLC analytical curves for the determination and quantification of the antibiotics. Amoxicillin (AMX), ampicillin (AMP) and cloxacillin (CLX) used to prepare antibiotic aqueous solution

* Corresponding author. Tel.: +60 14 9047313.
*E-mail addresses:* em_civil@yahoo.com, emadsoliman3@gmail.com (E.S. Elmolla).

0304-3894/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhazmat.2010.02.068

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

| Backpropagation (BP) algorithm | Function | Mean square error (MSE) | Epoch | Correlation coefficient ($R^2$) | Best linear equation |
|---|---|---|---|---|---|
| Levenberg-Marquardt backpropagation | trainlm | 0.00824866 | 30 | 0.994 | $y = 0.995X + 0.407$ |
| Scaled conjugate gradient backpropagation | trainscg | 0.0167993 | 99 | 0.988 | $y = 0.986X + 0.928$ |
| BFGS quasi-Newton backpropagation | trainbfg | 0.018824 | 55 | 0.987 | $y = 0.989X + 0.839$ |
| One step secant backpropagation | trainoss | 0.0306701 | 29 | 0.974 | $y = 0.958X + 2.59$ |
| Batch gradient descent | traingd | 0.486095 | 100 | 0.703 | $y = 0.387X + 33$ |
| Variable learning rate back propagation | traingdx | 0.449448 | 22 | 0.781 | $y = 0.405X + 30$ |
| Batch gradient descent with momentum | traingdm | 0.508201 | 100 | 0.718 | $y = 0.363X + 34.5$ |
| Fletcher–Reeves conjugate gradient backpropagation | traincgf | 0.0275731 | 25 | 0.979 | $y = 1.02X - 0.874$ |
| Polak–Ribi'ere conjugate gradient backpropagation | traincgp | 0.0175631 | 100 | 0.987 | $y = 0.982X + 1.23$ |
| Powell–Beale conjugate gradient backpropagation | traincgb | 0.0203964 | 37 | 0.985 | $y = 0.963X + 2.09$ |

were obtained from a commercial source (Farmaniage Company). The commercial products were used as received without any further purification. Sodium hydroxide and sulfuric acid were purchased from HACH Company, USA.

### 2.2. Analytical methods

Chemical oxygen demand (COD) was determined according to the Standard Methods [34]. If the sample contained hydrogen peroxide ($H_2O_2$), to reduce interference in COD determination pH was increased to above 10 to decompose hydrogen peroxide to oxygen and water [35–37]. pH was measured using a pH meter (HACH sension 4) and a pH probe (HACH platinum series pH electrode model 51910, HACH Company, USA). Biodegradability was measured by 5-day biochemical oxygen demand ($BOD_5$) test according to the Standard Methods [34].

### 2.3. Antibiotic aqueous solution

Antibiotics aqueous solution was prepared by dissolving the specific amounts of amoxicillin (AMX), ampicillin (AMP) and cloxacillin (CLX) in distilled water. It was prepared weekly and stored at 4 °C.

### 2.4. Experimental procedure

Batch experiments were conducted in a 600 mL Pyrex reactor with 500 mL of the antibiotic aqueous solution. The required amount of iron in the form of $FeSO_4{\cdot}7H_2O$ was added to the aqueous solution and mixed by a magnetic stirrer to ensure complete homogeneity during reaction. Thereafter, necessary amount of hydrogen peroxide was added to the mixture simultaneously with pH adjustment to the required value using $H_2SO_4$ or NaOH. The time at which hydrogen peroxide was added to the solution was considered the beginning of the experiment. Samples were taken at pre-selected time intervals using a syringe. The samples were then filtered through 0.45 μm membrane filter and tested for chemical oxygen demand (COD).

### 2.5. Artificial neural network (ANN)

Artificial neural networks are known for their ability of learning, simulation and prediction of data. The inspiration of using neural network came from the biology of human brain [29,31]. Disadvantage of artificial neural network is its "black box" nature. The individual relations between the input variables and the output variables are not developed by engineering judgment so that the model tends to be a black box. Further there is greater computational burden and proneness to overfitting, and the sample size has to be large [38]. The network consists of numerous individual processing units called neurons and commonly interconnected in a variety of structures. The strength of these interconnections is determined by the weight associated with neurons. The multilayer feed-forward net is a parallel interconnected structure consisting of input layer and includes independent variables, number of hidden layers and output layer [32].

In this study, a three-layered backpropagation neural network with tangent sigmoid transfer function (tansig) at hidden layer and a linear transfer function (purelin) at output layer was used. The backpropagation algorithm was used for network training.

Neural Network Toolbox V4.0 of MATLAB mathematical software was used for COD removal prediction. Data sets (120 experimental sets) were obtained from our previous study [3] and were divided into input matrix [p] and target matrix [t]. The input variables were reaction time ($t$), $H_2O_2$/COD molar ratio, $H_2O_2/Fe^{2+}$ molar ratio, pH and antibiotic concentration. The corresponding COD removal percent was used as a target. To ensure that all variables in the input data are important, principal component analysis (PCA) was performed as an effective procedure for the determination of input parameters. It was observed that all input variables were important. The data sets were divided into training (one half), validation (one fourth) and test (one fourth) subsets, each of which contained 60, 30 and 30 samples, respectively.

## 3. Results and discussion

### 3.1. Selection of backpropagation training algorithm

To determine the best backpropagation (BP) training algorithm, ten BP algorithms were studied. Tangent sigmoid transfer function (tansig) at hidden layer and a linear transfer function (purelin) at output layer were used. In addition, 5 neurons were used in the hidden layer as initial value for all BP algorithms. Table 1 shows a comparison of different backpropagation (BP) training algorithms. Levenberg–Marquardt backpropagation algorithm (LMA) was able to have smaller mean square error (MSE) compared to other backpropagation algorithms. So, LMA was considered the training algorithm in the present study.

### 3.2. Optimization of neurons number

The optimum number of neurons was determined based on the minimum value of MSE of the training and prediction set [28]. The optimization was done by using LMA as a training algorithm and varying neuron number in the range 1–20. Fig. 1 shows the relationship between number of neurons and MSE. MSE was 0.302316 when one neuron was used and decreased to 0.000376 when 14 neurons were used. Increasing of neurons more than 14 did not significantly decrease MSE. Hence, 14 neurons were selected as the best number of neurons. Fig. 2 shows the optimized neural network structure. It has three-layer ANN, with tangent sigmoid transfer function (tansig) at hidden layer with 14 neurons and linear transfer function (purelin) at output layer.

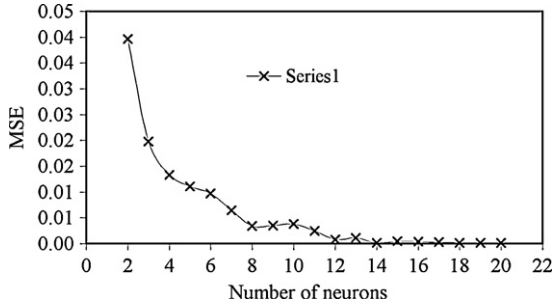


**Fig. 1.** Relationship between number of neurons and MSE.

### 3.3. Test and validation of the model

The data sets were used to feed the optimized network in order to test and validate the model. Fig. 3 shows a comparison between experimental COD removal values and predicted values using the neural network model. The figure contains two lines, one is the perfect fit $y = X$ (predicted data = experimental data) and the other is the best fit indicated by a solid line with best liner equation $y = (0.999)X + 0.116$, correlation coefficient ($R^2$) 0.997 and MSE 0.000376. This agrees well with the correlation coefficient reported in the literature—a correlation coefficient of 0.985 for prediction of nitrogen oxides removal by $TiO_2$ photocatalysis [39], 0.998 for prediction of methyl tert-butyl ether (MTBE) by UV/$H_2O_2$ process [40], 0.966 for prediction of polyvinyl alcohol degradation in aqueous solution by the photo-Fenton process [41], 0.995 for removal of humic substances from the aqueous solutions by ozonation [42] and 0.98 for decolouration of Acid Orange 52 dye by UV/$H_2O_2$ process [43].

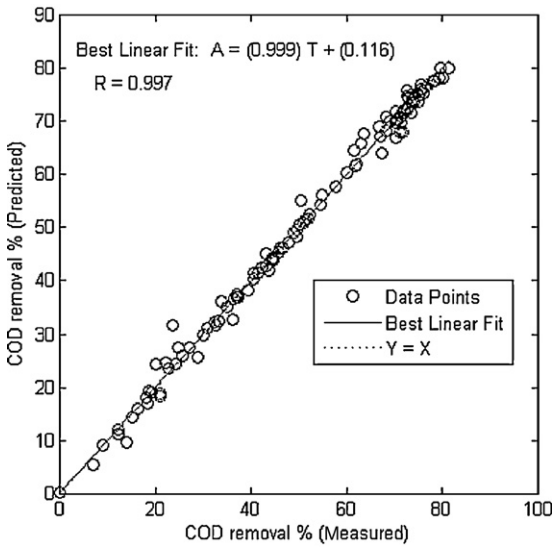


**Fig. 3.** Comparison between predicted and experimental values of the output.

### 3.4. 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 [31,44]. He proposed an equation based on the partitioning of connection weights:

$$Ij = \frac{\sum_{m=1}^{m=Nh}\left(\left(\left|W_{jm}^{ih}\right| / \sum_{k=1}^{Ni}\left|W_{km}^{ih}\right|\right) \times \left|W_{mn}^{ho}\right|\right)}{\sum_{k=1}^{k=Ni}\left\{\sum_{m=1}^{m=Nh}\left(\left|W_{km}^{ih}\right| / \sum_{k=1}^{Ni}\left|W_{km}^{ih}\right|\right) \times \left|W_{mn}^{ho}\right|\right\}} \quad (1)$$

where, $I_j$ is the relative importance of the $j$th input variable on the output variable, $N_i$ and $N_h$ are the number of input and hidden neu-

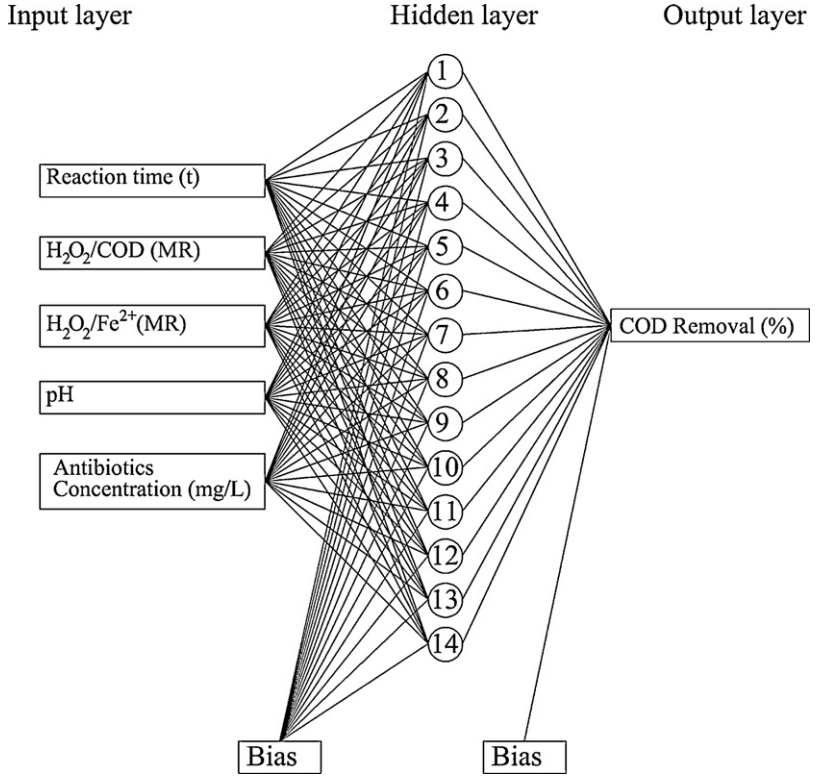


**Fig. 2.** Optimized ANN structure.

**Table 2**
Weight matrix, weights between input and hidden layers (W1) and weights between hidden and output layers (W2).

| Neuron | W1 | | | | | W2 |
|---|---|---|---|---|---|---|
| | Input variables | | | | | Output |
| | Time | $H_2O_2$/COD | $H_2O_2/Fe^{2+}$ | pH | Antibiotics concentration | COD removal % |
| 1 | 0.8869 | 0.0855 | −0.7863 | −0.2422 | 0.3197 | 0.7394 |
| 2 | 0.1251 | −2.0537 | −0.2143 | 0.8923 | −2.1553 | 1.9441 |
| 3 | 0.1996 | −0.0919 | −0.8751 | 0.2114 | 0.0137 | −1.4281 |
| 4 | 0.0048 | 0.6453 | 0.0236 | −0.786 | −0.4838 | 0.836 |
| 5 | 1.1772 | 0.84 | 0.8969 | 1.0508 | 0.045 | −0.7664 |
| 6 | −0.6675 | −1.1887 | 1.4041 | 0.7599 | 0.0555 | −0.5634 |
| 7 | −0.8075 | −1.0106 | 0.7954 | −1.0555 | 0.5573 | −0.9707 |
| 8 | −0.6322 | −0.2504 | 0.7846 | 0.4784 | −0.5846 | −1.0221 |
| 9 | 0.4747 | 0.2399 | 0.1719 | 0.6281 | −0.2614 | −1.0734 |
| 10 | −0.875 | 0.4465 | −0.0579 | 0.4996 | 0.9659 | −0.1611 |
| 11 | −0.9162 | −0.4413 | −1.73 | −1.4311 | 0.1075 | −1.4267 |
| 12 | 0.4797 | −0.0523 | −1.0736 | 0.2493 | 0.2117 | −1.018 |
| 13 | 0.3521 | 0.028 | 0.8121 | −0.5173 | −0.4002 | −0.3725 |
| 14 | 0.5326 | 1.355 | 0.4631 | 1.0192 | 1.5174 | 1.3226 |

**Table 3**
Relative importance of input variables.

| Input variable | Importance % |
|---|---|
| $H_2O_2/Fe^{2+}$ | 25.8 |
| pH | 22.1 |
| $H_2O_2$/COD | 18.2 |
| Time | 17.1 |
| Antibiotic concentration | 16.8 |
| Total | 100 |

rons, respectively and W is connection weight, 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.

Table 2 shows the weights between the artificial neurons produced by the ANN model used in this work. Table 3 shows the relative importance of the input variables calculated by Eq. (1). All variables have strong effect on antibiotic degradation in terms of COD removal. The $H_2O_2/Fe^{2+}$ molar ratio appears to be the most influential variable followed by pH, $H_2O_2$/COD molar ratio, reaction time and antibiotic concentration. The low relative importance of antibiotic concentration reveals that the selected $H_2O_2$/COD and $H_2O_2/Fe^{2+}$ molar ratios are valid for a wide range of wastewater strength.

The second evaluation process is based on the possible combination of variables [28]. Performance of the groups of one, two, three, four, and five variables were examined by the optimal ANN structure using the LMA with 14 hidden neurons. The input variables were $p_1$ (reaction time), $p_2$ ($H_2O_2$/COD molar ratio), $p_3$ ($H_2O_2/Fe^{2+}$ molar ratio), $p_4$ (pH), and $p_5$ (antibiotics concentration). Table 4

**Table 4**
Evaluation of possible combinations of input variables.

| Combination | Mean square error (MSE) | Epoch | Correlation coefficient ($R^2$) | Best linear equation |
|---|---|---|---|---|
| $p1$ | 365.889 | 6 | 0.315 | $y = 3.71X + 880$ |
| $p2$ | 276.46 | 8 | 0.599 | $y = 7.44X + 763$ |
| $p3$* | 270.141 | 10 | 0.616 | $y = 8.93X + 689$ |
| $p4$ | 378.575 | 7 | 0.395 | $y = 3.15X + 991$ |
| $p5$ | 404.727 | 12 | 0.284 | $y = 1.7X + 953$ |
| $p1 + p2$ | 0.500941 | 7 | 0.538 | $y = 0.409X + 29.2$ |
| $p1 + p3$ | 0.451707 | 8 | 0.649 | $y = 0.452X + 25.9$ |
| $p1 + p4$ | 0.65364 | 9 | 0.451 | $y = 0.32X + 31.8$ |
| $p1 + p5$ | 0.714965 | 6 | 0.391 | $y = 0.30X + 38$ |
| $p2 + p3$ | 0.415012 | 9 | 0.742 | $y = 0.528X + 25$ |
| $p2 + p4$ | 0.388861 | 5 | 0.764 | $y = 0.528X + 24.3$ |
| $p2 + p5$ | 0.552496 | 5 | 0.636 | $y = 0.405X + 32.1$ |
| $p3 + p4$* | 0.304122 | 9 | 0.848 | $y = 0.701X + 16.9$ |
| $p3 + p5$ | 0.571864 | 10 | 0.646 | $y = 0.509X + 23.5$ |
| $p4 + p5$ | 0.755573 | 5 | 0.487 | $y = 0.232X + 40.6$ |
| $p1 + p2 + p3$ | 0.313754 | 16 | 0.802 | $y = 0.642X + 18.1$ |
| $p1 + p2 + p4$ | 0.2901 | 14 | 0.825 | $y = 0.675X + 16.4$ |
| $p1 + p2 + p5$ | 0.453212 | 10 | 0.702 | $y = 0.675X + 25.2$ |
| $p1 + p3 + p4$ | 0.141262 | 25 | 0.873 | $y = 0.873X + 6.2$ |
| $p1 + p3 + p5$ | 0.43797 | 10 | 0.69 | $y = 0.57X + 21.1$ |
| $p1 + p4 + p5$ | 0.583005 | 16 | 0.528 | $y = 0.57X + 32.7$ |
| $p2 + p3 + p4$* | 0.117252 | 12 | 0.936 | $y = 0.849X + 9.37$ |
| $p2 + p3 + p5$ | 0.379122 | 47 | 0.77 | $y = 0.579X + 23.1$ |
| $p3 + p4 + p5$ | 0.300483 | 25 | 0.85 | $y = 0.695X + 17.1$ |
| $p1 + p2 + p3 + p4$* | 0.00278282 | 34 | 0.995 | $y = 0.997X + 0.402$ |
| $p1 + p2 + p3 + p5$ | 0.270749/0 | 25 | 0.818 | $y = 0.679X + 15.7$ |
| $p1 + p2 + p4 + p5$ | 0.264695 | 15 | 0.832 | $y = 0.682X + 15.8$ |
| $p1 + p3 + p4 + p5$ | 0.139748 | 15 | 0.912 | $y = 0.87X + 6.27$ |
| $p2 + p3 + p4 + p5$ | 0.113608 | 36 | 0.915 | $y = 0.862X + 8.92$ |
| $p1 + p2 + p3 + p4 + p5$* | 0.000376 | 20 | 0.997 | $y = 0.999X + 0.116$ |

* Best group performance.

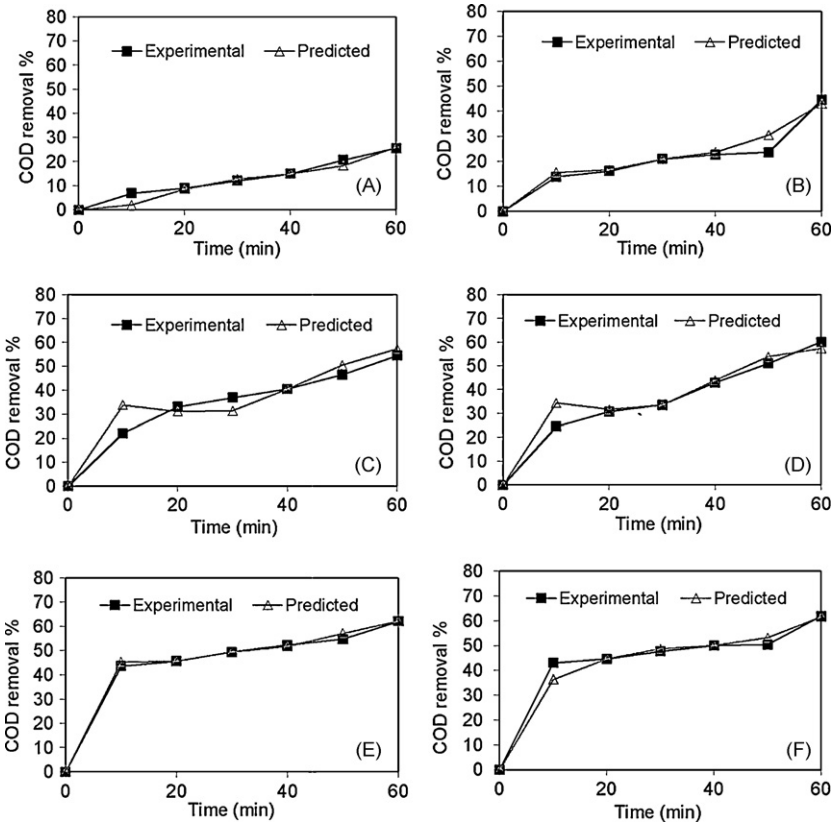


**Fig. 4.** Comparison between ANN output and experimental results at different $H_2O_2$/COD molar ratio: (A) 1.0, (B) 1.5, (C) 2.0, (D) 2.5, (E) 3.0 and (F) 3.5.

shows the results of the sensitivity analysis for different combinations of variables.

The sensitivity analysis showed that $p_3$ ($H_2O_2/Fe^{2+}$ molar ratio) was the most effective parameter among other variable in the group of one variable. The MSE (270.141) decreased up to 0.304122, which is the minimum value of the group of two variables when $p_4$ (pH) was used in combination with $p_3$. The MSE (0.304122) decreased up to 0.117252, which is the minimum value of the group of three variables when $p_2$ ($H_2O_2$/COD) was used in combination with $p_3$ and $p_4$. The MSE (0.117252) decreased up to 0.002782, which is the minimum value of the group of four variables when $p_1$ (reaction time) was used in combination with $p_3$, $p_4$ and $p_1$. Then, the MSE (0.002782) decreased up to 0.000376, which is the minimum value of the group of four variables when $p_5$ (antibiotics concentration) was used in combination with $p_3$, $p_4$, $p_1$ and $p_5$. The best group performances according to number of parameters are highlighted in Table 4. MSE values decreased as the number of variables in the group increased due to the contribution of all parameters (Table 4). It can be concluded that $H_2O_2/Fe^{2+}$ molar ratio is the most effective parameter. In addition, all variables have strong effect on antibiotic degradation in terms of COD removal and it agrees well with the sensitivity analysis using Garson equation.

### 3.5. Effect of $H_2O_2$/COD molar ratio

To examine the effect of $H_2O_2$/COD molar ratio on COD removal, initial $H_2O_2$ concentration was varied in the range 15–54 mM at constant initial COD 520 mg/L (16.25 mM). The corresponding $H_2O_2$/COD molar ratios were 1, 1.5, 2, 2.5, 3 and 3.5. Initial AMX, AMP and CLX concentrations were 104, 105 and 103 mg/L, respectively. The other operating conditions were fixed at pH 3 and $H_2O_2/Fe^{2+}$ molar ratio 50. Fig. 4A–F shows a comparison between the predicted and experimental values of COD removal at $H_2O_2$/COD molar ratio 1.0, 1.5, 2.0, 2.5, 3.0 and 3.5, respectively. The results show increase of COD removal at $H_2O_2$/COD molar ratio 1–3 and further increase in $H_2O_2$/COD molar ratio did not improve the COD removal. This may be due to scavenging of $OH^\bullet$ by $H_2O_2$ as in Reaction (R1) [45]. In terms of the relation between the experimental results and the predicted values of COD removal by the model, Fig. 4A–F shows that predicted values are in good agreement with the experimental results.

$$OH^\bullet + H_2O_2 \rightarrow H_2O + HO_2^\bullet \quad (R1)$$

### 3.6. Effect of $H_2O_2/Fe^{2+}$ molar ratio

In Fenton process, iron and hydrogen peroxide are two major chemicals determining the operation cost as well as efficiency. To examine the effect of $H_2O_2/Fe^2$ molar ratio on COD degradation, experiments were conducted at constant $H_2O_2$ concentration (46.87 mM) and varying $Fe^{2+}$ concentration in the range 0.32–24.3 mM. The corresponding $H_2O_2/Fe^{2+}$ molar ratios were in the range 2–150. Initial AMX, AMP and CLX concentration were 104, 105 and 103 mg/L, respectively. The operating conditions were pH 3, $H_2O_2$/COD molar ratio 3 and initial COD 520 mg/L (16.25 mM). Fig. 5A–F shows a comparison between the predicted and experimental values of COD removal at $H_2O_2/Fe^{2+}$ molar ratio 2, 5, 10, 20, 50 and 100, respectively. The results show that COD removal increased with decrease of $H_2O_2/Fe^{2+}$ molar ratio up to 10. Further decrease in $H_2O_2/Fe^{2+}$ molar ratio below 10 did not improve the degradation of antibiotics in terms of COD removal. This may be due to direct reaction of $OH^\bullet$ radical with metal ions at high concentration of $Fe^{2+}$ [46] as in Reaction (R2).

$$Fe^{2+} + HO^\bullet \rightarrow Fe^{3+} + HO^- \quad (R2)$$

In terms of the relation between the experimental results and the predicted values of COD removal by the model, Fig. 5A–E shows

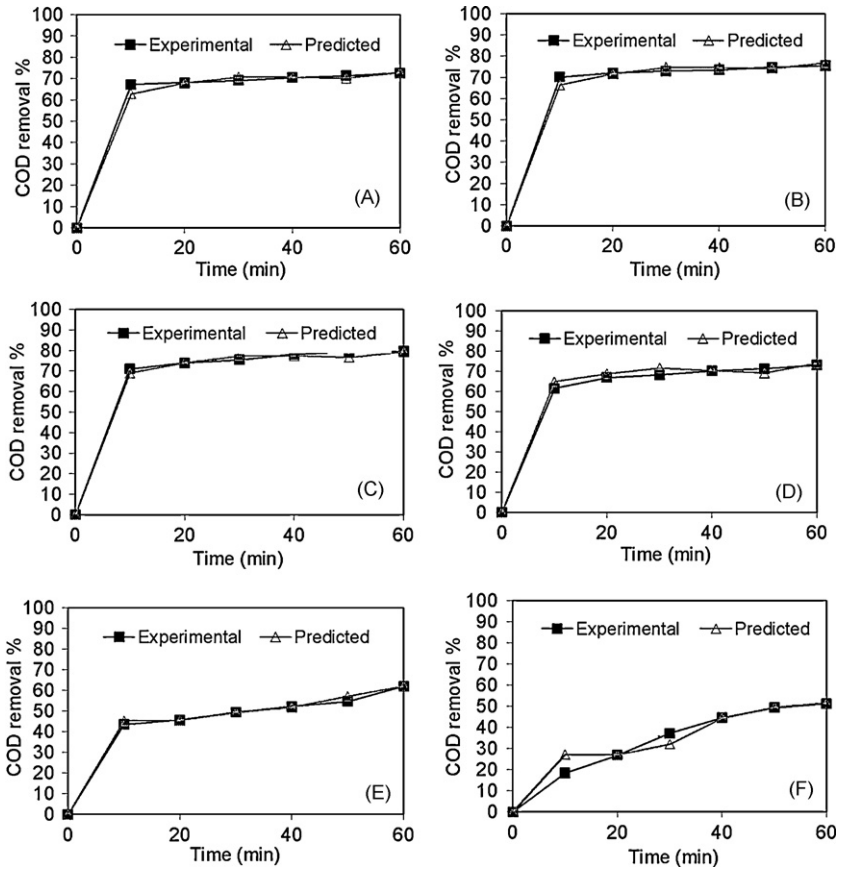


**Fig. 5.** Comparison between ANN output and experimental results at different $H_2O_2/Fe^{2+}$ molar ratio: (A) 2, (B) 5, (C) 10, (D) 20, (E) 50 and (F) 100.

that predicted values are in good agreement with the experimental results.

### 3.7. Effect of pH

The pH value influences the generation of hydroxyl radicals and hence the oxidation efficiency. To examine the effect of pH, experiments were conducted by varying the pH in the range 2–4. Initial AMX, AMP and CLX concentrations were 104, 105 and 103 mg/L, respectively. The operating conditions were $H_2O_2$/COD molar ratio 3, $H_2O_2/Fe^{2+}$ molar ratio 10 and initial COD 520 mg/L. Fig. 6A–D shows a comparison between the predicted and experimental values of the COD removal at pH 2, 2.5, 3 and 4, respectively.

The results show that pH significantly influences COD removal. Decrease in COD removal at pH higher than 3 may be due to the

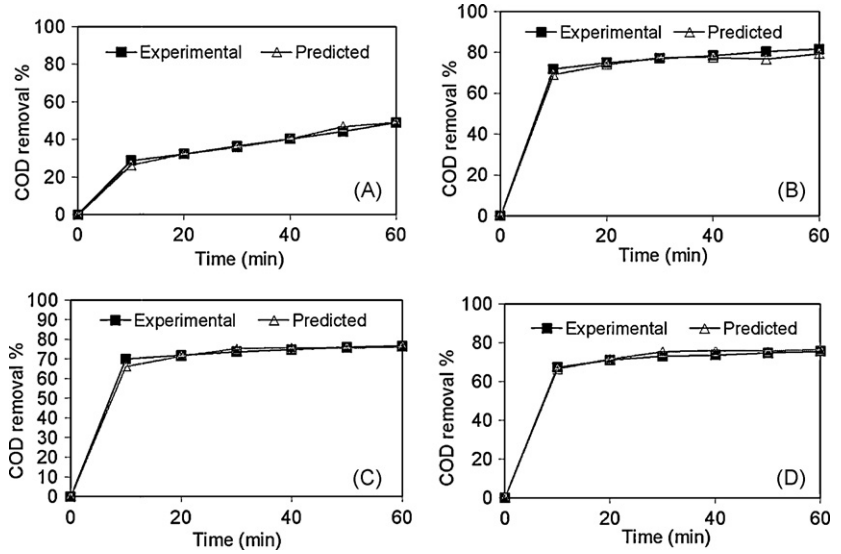


**Fig. 6.** Comparison between ANN output and experimental results at different pH: (A) 2, (B) 2.5, (C) 3 and (D) 4.

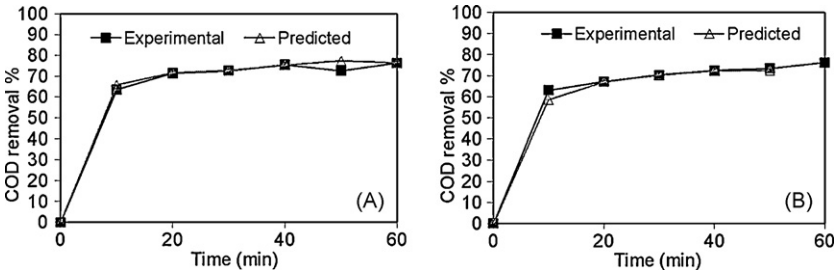


**Fig. 7.** Comparisons between ANN output and experimental results at different AMX, AMP and CLX concentration: (A) 100 and (B) 500 mg/L for each antibiotic.

decrease in dissolved iron and oxidation rate [45,47,48]. Further, hydrogen peroxide is stable at low pH probably because it gets solvated in the presence of high concentration of $H^+$ to form stable oxonium ion ($H_3O_2^+$), thus reducing substantially its reactivity with ferrous ion [49]. Therefore, amount of hydroxyl radicals would decrease at low pH, decreasing degradation of antibiotic intermediates. In terms of the relation between the experimental results and the predicted values of COD removal by the model, Fig. 6A–D shows that predicted values are in good agreement with the experimental results.

### 3.8. Effect of initial antibiotics concentration and reaction time

To observe the effect of initial antibiotic concentration, experiments were conducted by varying the initial concentration of AMX, AMP and CLX as 100, 250 and 500 mg/L for each antibiotic in the aqueous solution. The corresponding COD were 520, 1200 and 2440 mg/L. The operating conditions were $H_2O_2$/COD molar ratio 3, $H_2O_2/Fe^{2+}$ molar ratio 10 and pH 3. Fig. 7A and B shows a comparison between the predicted and experimental values of COD removal at AMX, AMP and CLX concentration 100 and 500 mg/L (data not shown for 250 mg/L) for each antibiotic. The results indicate marginal decrease in COD removal with increase of antibiotics concentration. A statistical analysis (one-way ANOVA) performed on the results at a 5% level of significance indicated no significant effect of antibiotic concentration on the COD degradation ($p$-value 0.917). It is also confirmed by sensitivity analysis in Section 3.4 that the antibiotics concentration is the lowest influential variable among the studied variables. This reveals that the selected COD/$H_2O_2/Fe^{2+}$ molar ratio (1:3:0.30) is optimum for this type of wastewater and suitable for a wide range of antibiotics wastewater concentration. In terms of the relation between the experimental results and the predicted values of COD removal by the model, Fig. 7A and B shows that the predicted values are in good agreement with the experimental results.

## 4. Conclusions

A three-layer backpropagation neural network was optimized to predict the degradation of amoxicillin, ampicillin and cloxacillin in aqueous solution in terms of COD removal. The configuration of the backpropagation neural network giving the smallest MSE was three-layer ANN with tangent sigmoid transfer function (tansig) at hidden layer with 14 neurons, linear transfer function (purelin) at output layer and Levenberg–Marquardt backpropagation training algorithm (LMA). ANN predicted results are very close to the experimental results with correlation coefficient ($R^2$) of 0.997 and MSE 0.000376. The sensitivity analysis showed that all studied variables (reaction time, $H_2O_2$/COD molar ratio, $H_2O_2/Fe^{2+}$ molar ratio, pH and antibiotic concentration) have strong effect on antibiotic degradation in terms of COD removal. In addition, $H_2O_2/Fe^{2+}$ molar ratio is the most influential parameter with relative importance of 25.8%. ANN results showed that neural network modeling could effectively simulate and predict the behavior of the process.

## Acknowledgement

The authors are thankful to the management and authorities of the Universiti Teknologi PETRONAS for providing facilities for this research.

## References

[1] M.V. Walter, J.W. Vennes, Occurrence of multiple-antibiotic resistant enteric bacteria in domestic sewage and oxidative lagoons, Appl. Environ. Microbiol. 50 (1985) 930–933.
[2] M. Pera-Titus, V. Garcıa-Molina, M.A. Banos, J. Gimĭenez, S. Esplugas, Degradation of chlorophenols by means of advanced oxidation processes: a general review, Appl. Catal. B 47 (2004) 219–256.
[3] E. Elmolla, M. Chaudhuri, Optimization of Fenton process for treatment of amoxicillin, ampicillin and cloxacillin antibiotics in aqueous solution, J. Hazard. Mater. 170 (2009) 666–672.
[4] E.S. Elmolla, M. Chaudhuri, Degradation of the antibiotics amoxicillin, ampicillin and cloxacillin in aqueous solution by the photo-Fenton process, J. Hazard. Mater. 172 (2009) 1476–1481.
[5] E.S. Elmolla, M. Chaudhuri, Degradation of amoxicillin, ampicillin and cloxacillin antibiotics in aqueous solution by the UV/ZnO photocatalytic process, J. Hazard. Mater. 173 (2010) 445–449.
[6] E. Elmolla, M. Chaudhuri, Improvement of biodegradability of synthetic amoxicillin wastewater by photo Fenton process, World Appl. Sci. J. 5 (2009) 53–58.
[7] E.S. Elmolla, M. Chaudhuri, Photocatalytic degradation of amoxicillin, ampicillin and cloxacillin antibiotics in aqueous solution using UV/$TiO_2$ and UV/$H_2O_2$/$TiO_2$ photocatalysis, Desalination 252 (2010) 46–52.
[8] E. Chamarro, A. Marco, S. Esplugas, Use of Fenton reagent to improve organic chemical biodegradability, Water Res. 35 (2001) 1047–1051.
[9] N. Prakash, S.A. Manikandan, L. Govindarajan, V. Vijayagopal, Prediction of biosorption efficiency for the removal of copper(II) using artificial neural networks, J. Hazard. Mater. 152 (2008) 1268–1275.
[10] V.K. Pareek, M.P. Brungs, A.A. Adesina, R. Sharma, Artificial neural network modeling of a multiphase photodegradation system, J. Photochem. Photobiol. A: Chem. 149 (2002) 139–146.
[11] M. Cote, B.P.A. Grandjean, P. Lessard, J. Thibault, Dynamic modeling of activated sludge process: improving prediction using neural networks, Water Res. 29 (1995) 995–1004.
[12] C.A. Gontarski, P.R. Rodrigues, M. Mori, L.F. Prenem, Simulation of an industrial wastewater treatment plant using artificial neural networks, Comput. Chem. Eng. 24 (2000) 1719–1723.
[13] T.Y. Pai, Y.P. Tsai, H.M. Lo, C.H. Tsai, C.Y. Lin, Grey and neural network prediction of suspended solids and chemical oxygen demand in hospital wastewater treatment plant effluent, Comput. Chem. Eng. 31 (2007) 1272–1281.
[14] T.Y. Pai, S.H. Chuang, H.H. Ho, L.F. Yu, H.C. Su, H.C. Hu, Predicting performance of grey and neural network in industrial effluent using online monitoring parameters, Process Biochem. 43 (2008) 199–205.
[15] D.S. Lee, M.W. Lee, S.H. Woo, Y.J. Kim, J.M. Park, Nonlinear dynamic partial least squares modeling of a full-scale biological wastewater treatment plant, Process Biochem. 41 (2006) 2050–2057.
[16] L. Luccarini, E. Porr, A. Spagni, P. Ratini, S. Grilli, S. Longhi, G. Bortone, Soft sensors for control of nitrogen and phosphorus removal from wastewaters by neural networks, Water Sci. Technol. 45 (2002) 101–107.
[17] P. Holubar, L. Zani, M. Hager, W. Froschl, Z. Radak, R. Braun, Advanced controlling of anaerobic digestion by means of hierarchical neural networks, Water Res. 36 (2002) 2582–2588.
[18] S.H. Hong, M.W. Lee, D.S. Lee, J.M. Park, Monitoring of sequencing batch reactor for nitrogen and phosphorus removal using neural networks, Biochem. Eng. J. 310 (2007) 365–370.
[19] S. Sinha, P. Bose, M. Jawed, S. John, V. Tare, Application of neural network for simulation of upflow anaerobic sludge blanket (UASB) reactor performance, Biotechnol. Bioeng. 77 (2002) 806–814.

[20] G.M. Zeng, X.S. Qin, L. He, G.H. Huang, H.L. Liu, Y.P. Lin, A neural network predictive control system for paper mill wastewater treatment, Eng. Appl. Artif. Intell. 16 (2003) 121–129.
[21] N. Ren, Z. Chen, X. Wang, D. Hu, A. Wang, Optimized operational parameters of a pilot scale membrane bioreactor for high-strength organic wastewater treatment, Int. Biodeter. Biodegr. 56 (2005) 216–223.
[22] I.S. Baruch, P. Georgieva, J. Barrera-Cortes, S. Feyo de Azevedo, Adaptive recurrent neural network control of biological wastewater treatment, Int. J. Intell. Syst. 20 (2005) 173–193.
[23] S. Grieu, A. Traore, M. Polit, J. Colprim, Prediction of parameters characterizing the state of a pollution removal biologic process, Eng. Appl. Artif. Intell. 18 (2005) 559–573.
[24] Y.J. Kim, H. Bae, J.H. Ko, K.M. Poo, S. Kim, C.W. Kim, H.J. Woo, Rule-based fuzzy inference system for estimating the influent COD/N ratio and ammonia load to a sequencing batch reactor, Water Sci. Technol. 53 (2006) 199–207.
[25] I. Machón, H. López, J. Rodríguez-Iglesias, E. Marañón, I. Vázquez, Simulation of a coke wastewater nitrification process using a feed-forward neuronal net, Environ. Model. Softw. 22 (2007) 1382–1387.
[26] H. Moral, A. Aksoy, C.F. Gokcay, Modeling of the activated sludge process by using artificial neural networks with automated architecture screening, Comput. Chem. Eng. 32 (2008) 2471–2478.
[27] S. Aber, N. Daneshvar, S.M. Soroureddin, A. Chabok, K. Asadpour-Zeynali, Study of acid orange 7 removal from aqueous solutions by powdered activated carbon and modeling of experimental results by artificial neural network, Desalination 211 (2007) 87–95.
[28] K. Yetilmezsoy, S. Demirel, Artificial neural network (ANN) approach for modeling of Pb(II) adsorption from aqueous solution by Antep pistachio (*Pistacia Vera* L.) shells, J. Hazard. Mater. 153 (2008) 1288–1300.
[29] N. Prakash, S.A. Manikandan, L. Govindarajan, V. Vijayagopal, Prediction of biosorption efficiency for the removal of copper(II) using artificial neural networks, J. Hazard. Mater. (2008) 1268–1275.
[30] 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, J. Hazard. Mater. B 137 (2006) 1788–1795.
[31] A. Aleboyeh, M.B. Kasiri, M.E. Olya, H. Aleboyeh, Prediction of azo dye decolorization by UV/$H_2O_2$ using artificial neural networks, Dyes Pigments 77 (2008) 288–294.
[32] 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, J. Hazard. Mater. 157 (2008) 455–463.
[33] 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.
[34] APHA, AWWA, WPCF, Standard Methods for the Examination of Water and Wastewater, 18th ed., American Public Health Association/American Water Works Association/Water Pollution Control Federation, Washington, DC, USA, 1992.
[35] I. Talinli, G.K. Anderson, Interference of hydrogen peroxide on the standard COD test, Water Res. 26 (1992) 107–110.
[36] Y.W. Kang, M.J. Cho, K.Y. Hwang, Correction of hydrogen peroxide interference on standard chemical oxygen demand test, Water Res. 33 (1999) 1247–1251.
[37] http://www.H2O2.com.
[38] J.V. Tu, Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes, J. Clin. Epidemiol. 49 (1996) 1225–1231.
[39] F.L. Toma, S. Guessasma, D. Klein, G. Montavon, G. Bertrand, C. Coddet, Neural computation to predict $TiO_2$ photocatalytic efficiency for nitrogen oxides removal, J. Photochem. Photobiol. A: Chem. 165 (2004) 91–96.
[40] D. Salari, N. Daneshvar, F. Aghazadeh, 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. B125 (2005) 205–210.
[41] J.A. Giroto, R. Guardani, A.C.S.C. Teixeira, C.A.O. Nascimento, Study on the photo-Fenton degradation of polyvinyl alcohol in aqueous solution, Chem. Eng. Process. 45 (2006) 523–532.
[42] E. Oguz, 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, J. Hazard. Mater. 157 (2008) 455–463.
[43] O.L.C. Guimarães, M.H.d.R. Filho, A.F. Siqueira, H.J.I. Filho, M.B. Silva, Optimization of the AZO dyes decoloration process through neural networks: determination of the $H_2O_2$ addition critical point, Chem. Eng. J. 141 (2008) 35–41.
[44] G.D. Garson, Interpreting neural-network connection weights, AI Expert 6 (1991) 47–51.
[45] V. Kavitha, K. Palanivelu, Destruction of cresols by Fenton oxidation process, Water Res. 39 (2005) 3062–3072.
[46] J.M. Joseph, H. Destaillats, H.M. Hung, M.R. Hoffmann, The sonochemical degradation of azobenzene and related azo dyes: rate enhancements via Fenton's reactions, J. Phys. Chem. A 104 (2000) 301–307.
[47] M. Tamimi, S. Qourzal, N. Barka, A. Assabbane, Y. Ait-Ichou, Methomyl degradation in aqueous solutions by Fenton's reagent and the photo-Fenton system, Sep. Purif. Technol. 61 (2008) 103–108.
[48] X. Zhao, G. Yang, Y. Wang, X. Gao, Photochemical degradation of dimethyl phthalate by Fenton reagent, J. Photochem. Photobiol. A: Chem. 161 (2004) 215–220.
[49] B.G. Kwon, D.S. Lee, N. Kang, J. Yoon, Characteristics of P-chlorophenol oxidation by Fenton's reagent, Water Res. 33 (1999) 2110–2118.