Modeling the removal of Sunfix Red S3B from aqueous solution by electrocoagulation process using artificial neural network

HA BUI MANH*

Department of Environmental Sciences, Sai Gon university, 273 An Duong Vuong St., District 5, Ho Chi Minh 700000, Vietnam

(Received 8 January, revised 15 March, accepted 25 March 2016)

Abstract: This study presents an application of artificial neural networks (ANNs) to predict the dye removal efficiency (color and chemical oxygen demand (COD) value) of electrocoagulation process from Sunfix Red S3B aqueous solution. The Bayesian regulation algorithm was applied to train the networks with experimental data including five factors: pH, current density, sulphate concentration, initial dye concentration (IDC) and electrolysis time. The predicting performance of the ANN models was validated through the low root mean square error value (9.844 %), mean absolute percentage error (13.776 %) and the high determination coefficient value (0.836). Garson’s algorithm, connection weight method and neural interpretation diagram were also used to study the influence of input variables on dye removal efficiency. For decolorization, determined most effective inputs are current density, electrolysis time and initial pH, while COD removal is found to be strongly affected by initial dye concentration and sulphate concentration. Through these steps, we ANN’s robustness in modeling and analysis of electrocoagulation process is demonstrated.

Keywords: color removal; Garson’s algorithm; sensitivity analysis; textile wastewater.

INTRODUCTION

Textile wastewater usually has characteristic intensive color, suspended solids and salinity. In addition, it usually contains a large amount of bioreistant organic contaminants, which have strong toxic impact on microbes.1,2 In the last few years, various modern methods such as ozonation, sonication, ion exchange, membrane filtration, etc. have been successfully applied for color removal of dye-containing wastewater.3-5 Nonetheless, most of these techniques are limited by engineering requirements, costs or difficulties in operation. Hence, they could...
not be employed for treatment of actual dyeing wastewater. Recently, research has focused on electrocoagulation, which completely removes color from dyeing wastewater.6,7

In electrocoagulation (EC), direct electrical current is used to dissolve metal electrode (mostly iron and aluminum) forming a range of coagulant species and metal hydroxides. Thus formed compounds possess adsorption efficiency 100 times greater than pre-precipitated hydroxides in conventional coagulation process.8 There are three proposed mechanisms for dye removal in EC (iron anode): precipitation, floatation and adsorption, as illustrated in Fig. S-1 of the Supplementary material to this paper. Generally, the overall reaction is a synergistic effect of the three removal mechanisms and could be shifted depending on operating parameters (pH, current density, electrolysis time, etc.), treatment conditions (monopolar, bipolar, sample volume, etc.) and type of textile dyes (reactive, basic, acidic, etc.).6,8,9 Therefore, it is not easy to explain the mechanisms in EC process.

Several authors tried to explain the phenomenological principle in this process. However, strong basic knowledge of EC including completed mathematical formulas is still required.10–12 Accordingly, we need a fast and reliable modeling approach to reveal the complicated relationship between variables in EC process. Many scientists have shown recently that artificial neural networks (ANNs) are robust and excellent tools for modeling water treatment processes.13 This method enables understanding of the process without actually simulating the physical and chemical laws governing the system.14 Therefore, ANN has attracted a great deal of attention. However, few reports have been published on modeling of an EC process applied to remove pollutants from dye-containing wastewaters using ANN. Thus, the focus of the present study was to assess the potential of EC technique for removal of Sunfix Red S3B from aqueous solutions, monitored through color and chemical oxygen demand (COD) removal efficiencies for variable parameters such as: electrolysis time, initial pH, initial dye concentration, electric current and Na₂SO₄ concentration. Herein, we also implement a model network, which is capable to investigate effectively the relationship between the parameters and the dye removal efficiencies of EC process.

EXPERIMENTAL

Materials and instruments

Reactive dye stock. The commercial Sunfix Red S3B (SRS) was supplied by Oh-Young (a Korean company). Its molecular structure and UV–Vis absorption spectra are given in Fig. S-2 (Supplementary material). The dye stock solution (1000 mg L⁻¹) was prepared by completely dissolving 1.000 g dye powder into 1000 mL of distilled hot water, and the solution was diluted to appropriate concentrations (30–140 mg L⁻¹) before the experiment.
**Electrocoagulation reactor**

The experiments were performed in a bipolar batch reactor (Fig. S-3 of the Supplementary material), using four iron electrodes (Fe: 99.25% – CT2), dimensions of 146 mm\(\times\)150 mm\(\times\)4 mm, connected in parallel (bipolar mode). Only the outer electrodes were connected to the custom made digital DC power supply (220 V and 30 A) whose polarity automatically switch every 30 s. The electrocoagulator was made of Plexiglas, dimensions 160 mm\(\times\)160 mm\(\times\)250 mm including an Imhoff sedimentation tank. The active area and the distance of each electrode were 19.6 cm\(^2\) (140 mm\(\times\)140 mm) and 15 mm, respectively.

**Procedure and water analysis**

For each run, 4.0 L of the constant dye solution with specific concentration (30–140 mg L\(^{-1}\)) were placed into the electrocoagulator. The pH of the contents was adjusted to the needed pH value (3–12) by NaOH or HCl (0.05 N) solutions. Afterwards, Na\(_2\)SO\(_4\) was added for desired sulphate concentration (500–2500 mg L\(^{-1}\)). Finally, the current density was adjusted to the desired value (43.3, 65.0, 86.6, 108.3 and 130 A m\(^{-2}\)). After 30 min settling at the end of each run, 50 mL of the treated sample were collected, filtered and determined color removal efficiency by the reducing absorbance at 542 nm (\(\lambda_{\text{max}}\)) using a spectrophotometer (Thermo UV–Vis Genesys 10) and chemical oxygen demand (COD) removal efficiency by standard methods for the examination of water and wastewater.\(^{15}\) Finally, to avoid electrode passivation at the end of each run, the electrodes were washed thoroughly in water and dipped in 0.5 M HCl solution for at least 15 min to dissolve any oxide from electrode surface. Each analysis was performed in duplicate and the results presented are the average value ± standard deviations.

**Experimental design**

At first, the experiments were conducted with “one factor at a time” design to estimate the range values of the five factors or variables (31 assays). Afterwards, 19 experiments were designed using fractional factorial with 3 central points and two levels of operating variables (Table S-I of the Supplementary material) to generate information about the behaviour of EC treatment. Next, a full central composite design with two blocks (31 experiments) was performed to promote variability in the data factors as shown in Table S-II, Supplementary material. Finally, 5 additional tests were conducted to increase variability in the initial dye concentration factor (IDC) from 30 to 140 mg L\(^{-1}\) while keeping initial pH at 11, electrolysis time at 4 min, current density at 65.0 A m\(^{-2}\), and sulphate concentration at 1200 mg L\(^{-1}\). Therefore, 86 experiments were designed with descriptive statistics shown in Table S-III and distribution histogram of the factors shown in Fig. S-IV (Supplementary material). The Minitab software (version 16.2.0) was used for the statistical design of experiments.

**Artificial neural network method**

A three-layer feed-forward back-propagation network (Fig. 1) was chosen to model the relation of COD and color removal efficiencies to five independent variables with surveyed value, as mentioned in previous section.

The number of neurons in hidden layer was varied from 1 to 17 to determine the optimal topology of the network. Each topology was repeated one hundred times to avoid random correlation due to random initialization of the weights and bias. This network was trained using the Bayesian regulation algorithm (trainbr) with a set of 76 experimental data points, randomly divided into 80% for training, 20% for validation. Then the resulting network was evaluated with another set of 10 experimental data points. Artificial neural network (ANN) toolbox of Matlab 8.1 mathematical software for Windows was used for training and testing the
network on a Pentium Intel core i3, 3.7 GHz computer. Performance of the ANN model can be evaluated using the root of mean square error (RMSE), mean absolute percentage error (MAPE) and the coefficient of determination ($R^2$). RMSE represents the errors associated with the model and could be calculated as:

$$RMSE = \left( \frac{1}{N} \sum_{i=1}^{N} (y_i - d_i)^2 \right)^{0.5}$$  \hspace{1cm} (1)

$R^2$ represents the degree of correlation between the measured and predicted values:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (d_i - y_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2}$$  \hspace{1cm} (2)

An alternative way to estimate accuracy between the measured and predicted values is through MAPE, represented as percentage of the error:

$$MAPE = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{y_i - d_i}{y_i} \right|$$  \hspace{1cm} (3)

where $N$ is the number of input values; $y_i$ and $d_i$ are the measured (actual) values and the predicted (desired) value obtained from ANN model corresponding to the $i^{th}$ input, respectively; and $\bar{y}$ is the average of measured values.

The relative importance of the input variables was also determined according to Garson’s algorithm (Eq. (4)) and the connection weight method (Eq. (5)) which could be illustrated using the neural interpretation diagram:

$$Q_{io} = \frac{k}{\sum_{j=1}^{k} \left| W_{ij} \right|} \sum_{j=1}^{k} \left( \left| W_{ij} \right| \sum_{i=1}^{n} \left( \sum_{j=1}^{k} \left| W_{ij} \right| \left| W_{oj} \right| \right) \right)^{-1}$$  \hspace{1cm} (4)
Herein, $Q_{io}$ is the relative significance of the $i^{th}$ input variable upon the output variable; $n$ and $h$ denote the number of input and hidden neurons, respectively. $W$ is the connection weight value, the subscripts “$i$” “$j$” and “$o$” represents the input, hidden and output neurons, respectively.

RESULTS AND DISCUSSION

Neural network topology

Since the problem of predicting dye removal efficiency in this study is a fitting problem, we adopted the three-layer feed-forward network, which has been proven the most effective topology for fitting problems. From these three layers, input layer and output layer were determined by experimental setup composed of five input neurons, which correspond to initial pH, electrolysis time, agitation speed, gum dosage and initial dye concentration (IDC), and two output neurons, which correspond to color and COD removal efficiencies. The hidden layer was then determined by 17 trials varying from one to 17 neurons as shown in Fig. 2 and Table S-IV (Supplementary material). Using less than 5 hidden neurons, we observed the sign of underfitting where coefficients of determination of both training set and validation set were low. On the contrary, using more than 8 hidden neurons, the signs of overfitting were observed; coefficient of training set determination was high, while for the validation set was low, i.e., the network lacked capability to predict untrained data. Therefore, the suitable hidden layer of ANN for the dye removal process was determined to have from 5 to 8 neurons. Such explanations are reasonable, and they have been verified experimentally by Al-Abri et al. and Oliveira-Esquerre et al. Considering explanations along with the results indicated in Fig. 2, the optimal value of 7 hidden neurons are chosen with $R^2_{total} = 0.836$, $RMSE_{total} = 9.844$ % and $MAPE_{total} = 13.776$ % (see Table S-IV, Supplementary material) for further investigation.

Neural network training

The most popular algorithm for training ANN is backpropagation Levenberg–Marquardt (L–M) which is very successful to model antibiotics, dye, metal or organic waste removal processes. However, like scaled conjugate gradient (SCG) algorithm, gradient descent with momentum and adaptive LR (GDX) algorithm and many other algorithms, L–M algorithm requires a validation set of data to decide whether to stop or continue to train the network in each run. This validation set usually costs 20 % of experimental data, and causes the trained networks to be much different between training runs, when provided a specific division of data to training, validation and test sets. Furthermore, thanks to the strength of statistical regularization, Bayesian regularization...
(BR) algorithm, which is used in the study, does not require a validation set, *i.e.*, 20% of the experimental data are added to the training set, which gives the result networks with almost the same performance across training runs on a specific division of data. Particularly, Fig. S-III (Supplementary material) shows variation of ANN performance across 10 training runs in the same configuration of seven-hidden neuron network and a specific division of data. BR algorithm had outperformed other algorithms by yielding determination coefficient higher and more stable compared to the other algorithms.

In each run, the training state usually varied within the first 100 epochs, then approached a stable state where ANN performance (evaluated by MSE value) did not improve any more.

As shown in Fig. 3, after 85 epochs, the error value (MSE) reached bottom, *i.e.*, the performance reached top (gradient reach minimum), and the \( \mu \) parameter (Mu) oscillated around the top value, *i.e.*, the algorithm looped around the error minimum. This implies that the observed values were very close to the prediction. In other words, the network training was good.

**Evaluation and interpretation of ANN model**

In the research field of water and wastewater treatment, most of the published papers evaluated the model predictability using only minimum value of RMSE (MSE) or maximum value of \( R^2 (R) \) in total set as suggested by Al-Abri *et al.*\(^{18}\) or Jing *et al.*\(^{23}\) This could result in the model fail to converge due to excessive neurons in the hidden layer, which is also called over-fitting. In such a case, the network learns the training data well, but has no ability to meet results for the test data.\(^{24}\) Therefore, in this study the performance of the optimal network (with 7 hidden neurons) was evaluated by three parameters in the validation set, coef-
Fig. 3. Number of epochs vs. performance parameters in training state: a) gradient, b) $Mu$ and c) mean squared error ($MSE$). The parameters at epoch 100 are 14.550, 0.005 and 82.597 for gradient, $Mu$ and $MSE$, respectively.

Figures (CC) 2016 SCS. All rights reserved.
Available on line at www.shd.org.rs/JSCS/
(CC) 2016 SCS. All rights reserved.
Available on line at www.shd.org.rs/JSCS/
Overall, the results show that the predicted outputs could fit measured outputs very well.

More than a “black box” that can predict data with good performance, analyzing the internal structure of the result network gives insight about the influence of each experimental factor (input) on removal efficiencies (outputs). Particularly, Garson’s algorithm and connection weights algorithm were used to evaluate the influence of each factor and the results are summarized in Table I. Further analysing ANN structure, neural interpretation diagram (Fig. 5), network weights and bias values (Table S-V of the Supplementary material) gave insight...
into the sophisticated contribution of each factor to the outputs and their correlation.

**TABLE I. Contribution of each input factor to output values in the electrocoagulation process**

<table>
<thead>
<tr>
<th>Output Factor</th>
<th>Garson’s algorithm, %</th>
<th>Rank$^a$</th>
<th>Connection weight</th>
<th>Rank$^a$</th>
<th>Influence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Decolorization</strong> $(O_1)$</td>
<td>Dye concentration $(I_1)$</td>
<td>15.48</td>
<td>4</td>
<td>−0.001</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>pH $(I_2)$</td>
<td>20.07</td>
<td>3</td>
<td>0.572</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Current density $(I_3)$</td>
<td>24.56</td>
<td>2</td>
<td>3.489</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Time $(I_4)$</td>
<td>25.89</td>
<td>1</td>
<td>1.743</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>$C_{\text{sulphate}}$ $(I_5)$</td>
<td>14.01</td>
<td>5</td>
<td>0.063</td>
<td>4</td>
</tr>
<tr>
<td><strong>COD removal</strong> $(O_2)$</td>
<td>Dye concentration $(I_1)$</td>
<td>23.09</td>
<td>2</td>
<td>0.863</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>pH $(I_2)$</td>
<td>20.82</td>
<td>3</td>
<td>−0.093</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Current density $(I_3)$</td>
<td>15.42</td>
<td>5</td>
<td>−0.159</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Time $(I_4)$</td>
<td>16.84</td>
<td>4</td>
<td>0.173</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>$C_{\text{sulphate}}$ $(I_5)$</td>
<td>23.81</td>
<td>1</td>
<td>0.773</td>
<td>2</td>
</tr>
</tbody>
</table>

$^a$Rank of the absolute values

Fig. 5. Neural interpretation diagram of dye removal using electrocoagulation.

For decolorization, both algorithms comply that the three most influential factors are current density $(I_3)$, electrolysis time $(I_4)$ and initial pH $(I_2)$. Connection weight algorithm confirms that these factors have positive effect on decolorization, i.e., when their values increase, the decolorization efficiency increases. Significance of current density, electrolysis time and initial pH were also recorded for decoloring Basic Red 46 by EC in the research by Daneshvar et al.$^7$ For current density (or electrolysis time), their positive effects could be explained by the following: when the current density (or electrolysis time) increased, the amount of iron ion produced on the electrodes increased according to Faraday’s law. That led to increase in flocculant production and then improved color removal capacity. In the case of pH, its value determined the iron hydroxides formation in EC process.$^6$ If the initial pH value was higher, the more iron hyd-
roxides were formed, which increased the decolorization efficiency. The other two factors, initial dye concentration ($I_1$) and sulphate concentration ($I_5$), have minor effect on decolorization improvement.

For COD removal, both algorithms agree that initial dye concentration ($I_1$) and sulphate concentration ($I_5$) are strongly influential factors, while current density ($I_3$) is weakly influential, but they disagree on whether initial pH ($I_2$) or electrolysis time ($I_4$) have stronger influence. Plotting COD removal efficiency vs. initial pH and electrolysis time, as shown in Fig. 6, we observed that both factors actually have high degree of correlation. It showed that electrolysis time had positive influence on COD removal, as confirmed by connection weight algorithm’s results in Table I. It also showed the effect of initial pH on COD removal was complicated having both positive and negative trends and the complication increased with electrolysis time. Complication of initial pH was confirmed by the neural interpretation diagram (shown in Fig. 5), where the influence of $I_2$ on $O_2$ is contributed by both positive paths through $H_2$, $H_4$, $H_5$ and $H_6$, and negative paths through $H_1$, $H_3$ and $H_7$. The two algorithms have different resolutions for this complication. While Garson’s algorithm uses absolute values to disregard the sign of influences, i.e., negative influence is treated equivalent to the positive influence; the connection weight algorithm just lets them cancel each other out. Therefore, Garson’s algorithm concludes that $I_2$ has strong influence because it has both positive and negative influences while connection weight algorithm concludes it has weak influence in total. Furthermore, we should be careful about difference in result interpretation of the two algorithms.

Fig. 6. Response and contour plot of electrolysis time vs. sulphate concentration on the removal.
CONCLUSIONS

Electrocoagulation process (EC) was demonstrated to be a simple and efficient method for color removal of simulated wastewater from Sunfix Red S3B dye. The newly proposed ANN model successfully predicted the coagulation process with determination coefficient ($R^2$), the root mean square error (RMSE), and mean absolute percentage error (MAPE) found as 0.836, 9.844 % and 13.776 %, respectively. Based on the model and sensitivity analysis methods (Garson’s algorithm, connection weight method and neural interpretation diagram), it can be concluded that the efficiency of EC was highly dependent on current density, electrolysis time and initial pH for color removal. COD removal efficiency was highly dependent on IDC, sulphate concentration, and electrolysis time combined with initial pH. These results showed a potential of ANN as a robust tool for modeling coagulation process and for analyzing effects of each factor on dye removal process.

SUPPLEMENTARY MATERIAL

Details about soil treatment, materials, chemicals and are available electronically from http://www.sbd.org.rs/JSCS/ or from the corresponding author on request.