



Electrocoagulation efficiency and energy consumption probing by artificial intelligent approaches

Afshin Maleki, Hiua Daraei*, Behzad Shahmoradi, Somaye Razee, Nahideh Ghobadi

Kurdistan Environmental Health Research Center, Kurdistan University of Medical Sciences, Sanandaj, Iran
Tel. +98 871 613 1504; Fax: +98 871 662 5131; email: hiua.daraei@gmail.com

Received 4 April 2012; Accepted 27 March 2013

ABSTRACT

Color removal efficiency (CR%) and energy consumption (EnC) of Electrocoagulation (EC) were investigated using synthetic wastewater, containing disperses like orange 25 dye (DO25). Five operational parameters including initial pH (pH_0) (2, 5.5, and 9), initial dye concentration (C_0) (20, 60, and 100 mg L^{-1}), applied voltage (V_{EC}) (10, 20, and 30 V), initial electrolyte concentration (C_S) (0, 1.5, and 3 g L^{-1}), and treatment time (t_{EC}) (0, 0.5, 5, 10, 15, 25, 35, and 50 min) were probed as more effective operational parameters of EC. Combined design of experiments (DOE) was designed and experiments were conducted in accordance with it. The experimental data were obtained in a laboratory through a handmade batch reactor. The achieved CR% (0–99.9) and EnC (0–69.4 wh) were gained under experimental conditions. The optimum value of C_0 was almost 20 ppm (minimum range). Two optimum clusters could be discriminated for other four parameters. First group was corresponded to conditions with $\text{pH}_0=9$ (maximum value of range), $C_S=0.7\text{--}1.1$ (g/lit), $V_{\text{EC}}=10$ V (minimum of range), and $t_{\text{EC}}=1$ min. Second group was corresponded to the conditions with $\text{pH}_0=6.8$ (except two cases), $C_S=1.1\text{--}2$ (g/lit), $V_{\text{EC}}=10\text{--}15.2$ V, and $t_{\text{EC}}=49.4\text{--}50$ min. The data was used for model building by employing two more popular models in this study: reduced quadratic multiple regression model (SMLR) and artificial neural network (ANN). Further statistical tests were applied to exhibit models' goodness and to compare the models. Based on statistical comparison, ANN models obviously outperformed SMLR models. Finally, multi objective optimization of CR% and EnC was carried out using genetic algorithm (GA) over the outperformed ANN models. The optimization procedure causes nondominated optimal points, which gave an insight into optimal operating conditions of the EC.

Keywords: Design of experiments (DOE); Artificial neural networks (ANNs); Genetic algorithm multi-objective optimization; Color removal; Electrocoagulation (EC)

1. Introduction

Dyes in wastewaters are a serious environmental concern [1–3]. Synthetic azo dyes comprise more than half of all dyes production [4]. Because of their toxicity, their release into the environment has caused a

lot of problems [5]. Owing to the complex structures of azo dyes, biological, physical, and chemical treatments of dye effluents are inefficient [6]. In recent years, electrochemical-based treatment methods such as electro-oxidation and electrocoagulation (EC) have drawn great attention [7–9]. EC has been gaining popularity due to simple and easy-to-install equipment,

*Corresponding author.

negligible start up time, compact treatment facility, and minimal amount of hazardous sludge [10].

EC involves the *in situ* generation of coagulants through electrolytic oxidation of an appropriate sacrificial anode [7]. It has been successfully used for decades to treat the wastewaters of various sources [11]. Although there have been a remarkable amount of studies on EC technique for wastewater treatment, large-scale applications of this technology have been relatively few. One possible reason is the energy demand of the EC [12].

In the EC process, some factors, such as initial pH (pH_0), initial dye concentration (C_0), applied voltage (V_{EC}), initial electrolyte concentration (C_S), and treatment time (t_{EC}), influence the process efficiency and energy consumption (EnC). The process efficiency may be increased by the optimization of these factors [13]. In conventional multifactor experiments, optimization is usually carried out by varying a single factor while keeping all the other factors fixed at a specific set of conditions [14]. This method is time consuming and incapable of effective optimization [13]. Recently, response surface methodology and full factorial design (FFD) have been employed to optimize and to understand the performance of complex systems, especially for EC [15,16].

In “*statistical*” methods, the order in which the predictor variables are entered into (or taken out of) the model is determined according to the strength of their correlation with the criterion variable. In fact, there are several versions of this method including forward selection, backward selection, and stepwise selection; stepwise is the most sophisticated one [17,18]. Recently, reduced quadratic multiple regression models have been employed in this kind of experiments [13,15,19–21]. Based on our best information, stepwise has not been applied as a reduction tool in these kinds of studies in spite of its advantages.

Artificial neural network (ANN) is parallel computational procedure consisting of highly interconnected processing elements groups named neurons [22]. Owing to their inherent nature to model and learn “*complexities*”, ANNs have found wide applications in various areas of wastewater treatment [23–26]. ANN has been recently used for color removal efficiency (CR%) and EnC modeling in EC [20,27].

Genetic algorithms (GA) are adaptive heuristic search algorithms based on the evolutionary ideas of natural selection and genetic. They belong to the larger class of evolutionary algorithms, which generate solutions to optimize problems by carrying out stochastic transformations inspired by natural evolution, such as inheritance, mutation, selection, and crossover [28–31].

To make a long story short, the main objectives of the study were to optimize the EC process to maximize CR% and minimize EnC, regarding the effective operational factors C_0 , pH_0 , V_{EC} , C_S , and t_{EC} using GA. SMLR and ANN methodologies were used to construct models. Further, both CR% and EnC were subjected to multiple objective optimizations using GA approach.

2. Materials and methods

2.1. Materials

The DO25 dye was purchased from Alvansabet Co., Iran. The chemical structure and some characteristics of this dye are shown in Table 1. Synthetic wastewater was prepared by dissolving the dye in distilled water. The pH_0 of the solutions was adjusted using NaOH (1M) and H_2SO_4 (1M) (Merck, Germany). The C_S was adjusted to desired value using NaCl (Merck, Germany) [19].

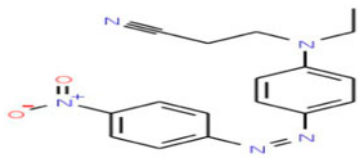
The EC system consisted of a glass ($12 \times 12 \times 21$ cm) cubic reactor, 400-rpm mixer, DC power supply (the high stability, reliability, and low-noise DC adjustable power supply RXN-303D-II, Zhaoxin Electronic Tech. Co.), and two aluminum electrodes. Fig. 1 shows the lab-scale batch experimental setup of EC unit. The cathode and anode were made of aluminum sheets ($4 \times 5 \times 0.1$ cm) and the immersed surface area of each electrode was 40 cm^2 . They were placed vertically and dipped in 1.5-L aqueous dye solutions. The distance between electrodes was fixed at 1 cm.

2.2. Experimental design

Combined design of CCD for C_0 , pH_0 , V_{EC} , and C_S and FFD for t_{EC} were used to design the experiments using Minitab 14. Combined design of experiments was selected due to the complex kinetics of the EC reaction that makes essential to collect more complete t_{EC} data [23,32,33]. Furthermore, in this kind of study, only providing each run is time consuming and reducing its number is crucial.

Sixteen cube points, eight axial points, and one center point were chosen as experimental points for four C_0 , pH_0 , V_{EC} , and C_S operational parameters at three levels of each parameter using MINITAB 14. The runs were conducted in a randomized manner to guard against systematic bias. In general, in these 25 experiments, eight levels of t_{EC} (0, 0.5, 5, 10, 15, 25, 35, and 50 min) were determined as t_{EC} . All the selected experimental conditions can be seen in Table 2.

Table 1
The chemical properties of DO25

Chemical name	C.I. Disperse Orange 25
Molecular formula	$C_{17}H_{17}N_5O_2$
Molecular structure	
Molecular weight	323.35
CAS number	31482-56-1

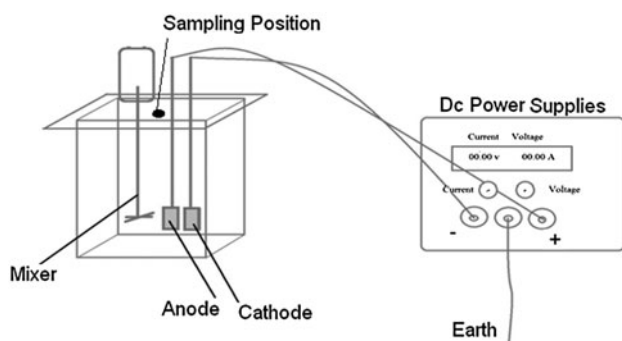


Fig. 1. The lab-scale batch experimental setup of EC unit.

2.3. EC process and data collection

In each run, 1.5-L solution containing DO25 was decanted into the EC reactor. Voltage was adjusted to the desired value, and then electrolysis was started. In each eight determined t_{EC} , 10-ml sample was extracted at the same position of EC vessels using 10-ml pipettes. Each sample was centrifuged (10 min at 3,000 rpm). The CR% was calculated for the decanted solution. A digital amperemeter and voltmeter incorporated in the power supply were applied for electrolysis current (I_{EC}) and V_{EC} monitoring during EC.

The color of solution was evaluated using PG T80⁺ spectrophotometer in the UV–Vis range (200–800 nm) at the λ_{max} . The λ_{max} for DO25 is 424 nm in the aqueous solution at natural pH. Because of matrix effects on the λ_{max} , it was extracted experimentally from the zero time sample spectra in each run. Tolerance of less than ± 10 nm was seen for the samples with different initial conditions. After EC progress, the CR% was calculated for samples using Eq. (1):

$$CR \% = (1 - A/A_0) \times 100 \quad (1)$$

where A_0 and A are absorptions of solution before and after EC, respectively.

The calibration curve was plotted for DO25 to assess the dynamic range of concentration (Fig. 2).

The standard plot illustrates that at least in the range of 0–200 ppm concentration, the Beer–Lambert law is valid; it is two times wider than the study range of concentration. In this range of concentration, it can be said:

$$\begin{aligned} A &= abc \rightarrow ((A_1 - A_2)/A_1) \times 100 \\ &= ((C_1 - C_2)/C_1) \times 100 \rightarrow \text{Color removal \%} \\ &= \text{Dye removal \%} \end{aligned}$$

The EnC was calculated at each condition as follows [19]:

$$\text{EnC (wh)} = V_{EC} \times I_{EC} \times t_{EC} \quad (2)$$

where V_{EC} is the applied voltage (V), I_{EC} is the average current of electrolysis (Ampere), and t_{EC} is the treatment time (hour).

2.4. Methodology of modeling

We used 200 data of CR% and EnC together with corresponding experimental conditions as a data-set. The five operational parameters were applied to generate all quadratic and linear inputs of models, while the CR% and EnC were considered as dependent variable. Data-set was randomly divided into three parts; 60% as a training set, 20% as a validation set, and 20% as testing set [18,34]. Two different SMLR and ANN models were constructed based on the same data-sets for both CR% and EnC separately.

Stepwise multiple regressions were applied to develop SMLR model using the aforementioned subsets, except the test and validation sets that were merged to use as an external test set together. The quadratic interactions of independent variables have been considered to improve the linear model efficiency. The more effective inputs and interactions were selected using stepwise algorithm [18,34].

Table 2
CR% and EnC of all 25 runs and 7 different sampling times of each run

Run No.	pH ₀	C ₀	V _{EC}	C _S	t _{EC}	EnC (kwh)	CR%	Run No.	pH ₀	C ₀	V _{EC}	C _S	t _{EC}	EnC (kwh)	CR%
1	2	20	30	0	0.5	0.065	65.0	14	2	20	10	3	0.5	0.100	0.5
1	2	20	30	0	5	0.650	90.5	14	2	20	10	3	5	1.000	4.0
1	2	20	30	0	10	1.300	92.3	14	2	20	10	3	10	2.000	5.1
1	2	20	30	0	15	1.950	96.4	14	2	20	10	3	15	3.000	7.1
1	2	20	30	0	25	3.250	97.7	14	2	20	10	3	25	5.000	13.1
1	2	20	30	0	35	4.550	98.6	14	2	20	10	3	35	7.000	97.5
1	2	20	30	0	50	6.500	98.6	14	2	20	10	3	50	10.000	99.5
2	2	60	20	1.5	0.5	0.205	4.3	15	5.5	60	20	0	0.5	0.002	0.2
2	2	60	20	1.5	5	2.050	30.2	15	5.5	60	20	0	5	0.017	1.6
2	2	60	20	1.5	10	4.100	30.7	15	5.5	60	20	0	10	0.033	5.5
2	2	60	20	1.5	15	6.150	43.4	15	5.5	60	20	0	15	0.050	5.7
2	2	60	20	1.5	25	10.250	57.0	15	5.5	60	20	0	25	0.083	5.9
2	2	60	20	1.5	35	14.350	87.3	15	5.5	60	20	0	35	0.117	10.0
2	2	60	20	1.5	50	20.500	89.1	15	5.5	60	20	0	50	0.167	10.5
3	5.5	100	20	1.5	0.5	0.210	7.8	16	9	100	10	0	0.5	0.001	0.6
3	5.5	100	20	1.5	5	2.100	98.4	16	9	100	10	0	5	0.008	2.6
3	5.5	100	20	1.5	10	4.200	98.5	16	9	100	10	0	10	0.017	3.0
3	5.5	100	20	1.5	15	6.300	98.7	16	9	100	10	0	15	0.025	3.4
3	5.5	100	20	1.5	25	10.500	98.8	16	9	100	10	0	25	0.042	6.3
3	5.5	100	20	1.5	35	14.700	98.9	16	9	100	10	0	35	0.058	6.7
3	5.5	100	20	1.5	50	21.000	99.3	16	9	100	10	0	50	0.083	11.9
4	9	100	10	3	0.5	0.083	6.7	17	9	20	30	0	0.5	0.008	0.5
4	9	100	10	3	5	0.825	99.3	17	9	20	30	0	5	0.075	0.9
4	9	100	10	3	10	1.650	99.5	17	9	20	30	0	10	0.150	1.4
4	9	100	10	3	15	2.475	99.6	17	9	20	30	0	15	0.225	2.8
4	9	100	10	3	25	4.125	99.7	17	9	20	30	0	25	0.375	16.4
4	9	100	10	3	35	5.775	99.8	17	9	20	30	0	35	0.525	16.8
4	9	100	10	3	50	8.250	99.9	17	9	20	30	0	50	0.750	98.6
5	9	100	30	3	0.5	0.760	18.2	18	5.5	60	20	3	0.5	0.435	8.4
5	9	100	30	3	5	7.600	96.3	18	5.5	60	20	3	5	4.350	98.4
5	9	100	30	3	10	15.200	97.0	18	5.5	60	20	3	10	8.700	98.6
5	9	100	30	3	15	22.800	98.0	18	5.5	60	20	3	15	13.050	98.8
5	9	100	30	3	25	38.000	99.0	18	5.5	60	20	3	25	21.750	99.0
5	9	100	30	3	35	53.200	99.1	18	5.5	60	20	3	35	30.450	99.1
5	9	100	30	3	50	76.000	99.5	18	5.5	60	20	3	50	43.500	99.5
6	5.5	60	30	1.5	0.5	0.530	13.2	19	9	20	10	3	0.5	0.102	1.0
6	5.5	60	30	1.5	5	5.300	98.4	19	9	20	10	3	5	1.017	96.4
6	5.5	60	30	1.5	10	10.600	98.8	19	9	20	10	3	10	2.033	96.9
6	5.5	60	30	1.5	15	15.900	98.9	19	9	20	10	3	15	3.050	97.4
6	5.5	60	30	1.5	25	26.500	99.3	19	9	20	10	3	25	5.083	97.9
6	5.5	60	30	1.5	35	37.100	99.6	19	9	20	10	3	35	7.117	98.4
6	5.5	60	30	1.5	50	53.000	99.8	19	9	20	10	3	50	10.167	99.0
7	2	100	30	3	0.5	0.743	62.2	20	2	20	10	0	0.5	0.013	10.9
7	2	100	30	3	5	7.425	99.4	20	2	20	10	0	5	0.125	60.0
7	2	100	30	3	10	14.850	99.6	20	2	20	10	0	10	0.250	90.9
7	2	100	30	3	15	22.275	99.7	20	2	20	10	0	15	0.375	91.5
7	2	100	30	3	25	37.125	99.8	20	2	20	10	0	25	0.625	92.1
7	2	100	30	3	35	51.975	99.9	20	2	20	10	0	35	0.875	92.7

(Continued)

Table 2 (Continued)

Run No.	pH ₀	C ₀	V _{EC}	C _S	t _{EC}	EnC (kwh)	CR%	Run No.	pH ₀	C ₀	V _{EC}	C _S	t _{EC}	EnC (kwh)	CR%
7	2	100	30	3	50	74.250	99.9	20	2	20	10	0	50	1.250	93.3
8	5.5	20	20	1.5	0.5	0.212	96.1	21	5.5	60	10	1.5	0.5	0.048	5.2
8	5.5	20	20	1.5	5	2.117	96.6	21	5.5	60	10	1.5	5	0.483	98.8
8	5.5	20	20	1.5	10	4.233	97.1	21	5.5	60	10	1.5	10	0.967	99.0
8	5.5	20	20	1.5	15	6.350	97.6	21	5.5	60	10	1.5	15	1.450	99.2
8	5.5	20	20	1.5	25	10.583	98.1	21	5.5	60	10	1.5	25	2.417	99.3
8	5.5	20	20	1.5	35	14.817	98.6	21	5.5	60	10	1.5	35	3.383	99.5
8	5.5	20	20	1.5	50	21.167	99.0	21	5.5	60	10	1.5	50	4.833	99.7
9	9	60	20	1.5	0.5	0.192	98.5	22	9	100	30	0	0.5	0.033	9.1
9	9	60	20	1.5	5	1.917	98.6	22	9	100	30	0	5	0.325	9.5
9	9	60	20	1.5	10	3.833	98.8	22	9	100	30	0	10	0.650	11.1
9	9	60	20	1.5	15	5.750	99.3	22	9	100	30	0	15	0.975	12.8
9	9	60	20	1.5	25	9.583	99.5	22	9	100	30	0	25	1.625	13.8
9	9	60	20	1.5	35	13.417	99.7	22	9	100	30	0	35	2.275	22.3
9	9	60	20	1.5	50	19.167	99.8	22	9	100	30	0	50	3.250	24.6
10	9	20	10	0	0.5	0.001	5.2	23	2	100	30	0	0.5	0.023	5.3
10	9	20	10	0	5	0.008	5.7	23	2	100	30	0	5	0.225	34.7
10	9	20	10	0	10	0.017	6.1	23	2	100	30	0	10	0.450	93.7
10	9	20	10	0	15	0.025	6.6	23	2	100	30	0	15	0.675	95.2
10	9	20	10	0	25	0.042	7.1	23	2	100	30	0	25	1.125	96.8
10	9	20	10	0	35	0.058	7.5	23	2	100	30	0	35	1.575	97.6
10	9	20	10	0	50	0.083	8.0	23	2	100	30	0	50	2.250	97.8
11	2	20	30	3	0.5	0.675	9.4	24	2	100	10	0	0.5	0.033	0.2
11	2	20	30	3	5	6.750	12.8	24	2	100	10	0	5	0.333	7.7
11	2	20	30	3	10	13.500	24.6	24	2	100	10	0	10	0.667	51.3
11	2	20	30	3	15	20.250	39.9	24	2	100	10	0	15	1.000	51.6
11	2	20	30	3	25	33.750	95.6	24	2	100	10	0	25	1.667	73.2
11	2	20	30	3	35	47.250	97.5	24	2	100	10	0	35	2.333	81.8
11	2	20	30	3	50	67.500	98.5	24	2	100	10	0	50	3.333	90.2
12	5.5	60	20	1.5	0.5	0.168	6.9	25	2	100	10	3	0.5	0.033	0.4
12	5.5	60	20	1.5	5	1.683	98.4	25	2	100	10	3	5	0.333	43.8
12	5.5	60	20	1.5	10	3.367	98.5	25	2	100	10	3	10	0.667	97.1
12	5.5	60	20	1.5	15	5.050	98.7	25	2	100	10	3	15	1.000	98.5
12	5.5	60	20	1.5	25	8.417	98.9	25	2	100	10	3	25	1.667	99.1
12	5.5	60	20	1.5	35	11.783	99.1	25	2	100	10	3	35	2.333	99.3
12	5.5	60	20	1.5	50	16.833	99.3	25	2	100	10	3	50	3.333	99.4
13	9	20	30	3	0.5	0.550	95.9								
13	9	20	30	3	5	5.500	96.4								
13	9	20	30	3	10	11.000	96.8								
13	9	20	30	3	15	16.500	97.3								
13	9	20	30	3	25	27.500	97.7								
13	9	20	30	3	35	38.500	98.2								
13	9	20	30	3	50	55.000	99.1								

After the construction of the best SMLR models for both CR% and EnC, the effective parameters were selected regarding their presence or absence in SMLR models. Next, the ANN models were constructed based on these effective parameters. The quadratic

interaction addition is not necessary because of non-linear comprehensive nature of ANN model [16,19].

We used back propagation algorithm in this study as it is very fast and can be employed quite easily [35]. The training set should be used to adjust the

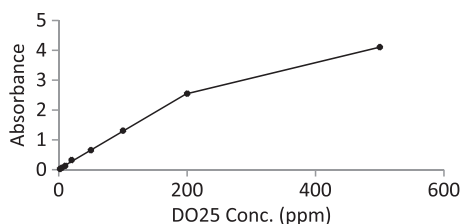


Fig. 2. The calibration curve for DO25.

parameters of the models; testing set was used to calculate its estimation power, and validation set to prevent over-train [34]. One of the most important factors in a neural network model configuration is to determine the number of layers to be used and the number of neurons in the layers. The number of input and output neurons is equal to the number of input and output parameters, respectively. The number of hidden layers and nodes is usually determined via a trial and error procedure [34].

Finally, the consistency of the models was revealed by tests quantified with predictive Q^2 and R^2 [18,34]. Another validation analysis of the comparison of ANN with other conventional methods is RMSE (Root-Mean-Square Error) as an indicator of reliability or accuracy of the models [18].

2.5. Multi-objective optimization

Simultaneous optimization of CR% and EnC falls in the field of multi-objective optimization [16,19]. There is no matchless solution to this type of optimization problem, but a set of mathematically equally solutions known as Pareto optimal solutions [16,19]. GA toolbox in MATLAB was used for generating the Pareto optimal solutions for CR% and EnC using “gamultiobj” function [19]. A MATLAB function, using ANN and SMLR models as the inputs, were written for creating a fitness function for the multi-objective optimization problem. The CR% component to be maximized was negated in the vector valued fitness function, since “gamultiobj” minimizes all the objectives. Experimental ranges were placed as bounds on the five inputs. The following algorithm options were set [16,19]: Selection Function: “Tournament of size two”, Crossover Fraction: “0.75with scattered crossover function”, Mutation Function: “Adaptive feasible mutation function”, Direction for migration: “Forward with migration fraction set to 0.2”, Distance measure function: “distance crowding”, Generations: “50”, Population Size: “100”, Stall-time Limit: “100”, Population size: “80”, and finally, The weighted average change in the fitness function value over “50”generations was used as the criteria for stopping the algorithm [19].

3. Results and discussion

3.1. EC process

The whole 200 CR% and EnC obtained in the all experiments conditions are given in Table 2. The effect of each operational parameter can be seen in this table. As a first result, a glance over this table approves the statement that different levels of experimental parameters result in different CR% and EnC.

3.2. SMLR models

The two SMLR models were separately developed using SMLR methodology for CR% and EnC after neglecting nonsignificant terms [16,19]. Both models are given in Eqs. (3) and (4) for the CR% and EnC, respectively.

$$\begin{aligned} \text{CR \%} = & 19.8554 + 4.1203(t_{\text{EC}}) - 0.0296(\text{pH}_0)(C_0) \\ & + 3.6542(\text{pH}_0)(C_S) - 0.1034(\text{pH}_0)(t_{\text{EC}}) \\ & + 0.1338(C_0)(C_S) + 0.2138(C_S)(t_{\text{EC}}) \\ & - 0.3291(\text{pH}_0)^2 + 0.0176(V_{\text{EC}})^2 \\ & - 6.7969(C_S)^2 - 0.0535(t_{\text{EC}})^2 \end{aligned} \quad (3)$$

$$\begin{aligned} \text{EnC} = & 5.7837 - 6.3669(C_S) - 0.7365(t_{\text{EC}}) \\ & + 0.3223(V_{\text{EC}})(C_S) + 0.0287(V_{\text{EC}})(t_{\text{EC}}) \\ & + 0.2689(C_S)(t_{\text{EC}}) - 0.0106(V_{\text{EC}})^2 \\ & + 0.0045(t_{\text{EC}})^2 \end{aligned} \quad (4)$$

Furthermore, statistical characteristics for both SMLR models have been presented in Table 3.

Based on unbiased standardized coefficients presented in Table 3, among linear parameters, t_{EC} and among quadratic parameters, $(C_S)^2$ and $(\text{pH}_0)(C_S)$ are the most important parameters in CR% prediction. The most important parameters in EnC prediction are C_S and t_{EC} . Moreover, it can be seen that the whole 5 parameters are effective on CR% when only three t_{EC} , V_{EC} , and C_S are effective on EnC that is in good agreement with previous reported studies [16,19]. Table 3 and Fig. 3 indicates that the MLR model does not have good predictability for CR% due to complex mechanism of EC. It demonstrates new interest in using more powerful modeling approach, especially ANN model [19].

It is notable that in this study, in addition to the five mentioned parameters, four more parameters consisting of electric current, final pH, and initial and final conductivity were measured and were examined in the models as input variables.

Table 3
Statistical characteristics of SMLR models of CR% and EnC

Model	CR%			EnC (wh)		
	Standardized coefficient	<i>t</i> -value	<i>p</i> -value	Standardized coefficient	<i>t</i> -value	<i>p</i> -value
Constant	7.70	2.57	0.001	1.49	3.89	<0.001
C_S	–	–	–	0.86	–7.35	<0.001
t_{EC}	0.64	6.36	<0.001	0.11	–6.50	<0.001
$(pH_0)(C_0)$	0.01	–2.11	0.04	–	–	–
$(pH_0)(C_S)$	0.54	6.75	<0.001	–	–	–
$(pH_0)(t_{EC})$	0.05	–2.03	0.04	–	–	–
$(C_0)(C_S)$	0.04	3.02	<0.001	–	–	–
$(C_S)(t_{EC})$	0.11	1.84	0.07	–	–	–
$(V_{EC})(C_S)$	–	–	–	0.037	8.57	<0.001
$(V_{EC})(t_{EC})$	–	–	–	0.003	9.42	<0.001
$(C_S)(t_{EC})$	–	–	–	0.021	12.55	<0.001
$(pH_0)^2$	0.15	–2.19	0.03	–	–	–
$(V_{EC})^2$	0.01	2.41	0.02	0.002	–4.59	<0.001
$(C_S)^2$	1.29	–5.24	<0.001	–	–	–
$(t_{EC})^2$	0.01	–5.054	<0.001	0.001	2.38	0.02
Data set	Train (120 data)		Test (80 data)	Same train set		Same test set
<i>F</i> value	21.134		–	120.412		–
R^2	0.660		0.647	0.882		0.871
Q^2	0.659		0.644	0.881		0.852
RMSE	25.7		26.9	4.73		5.24

In modeling studies, it is important to construct a model with least input parameters because this will lead to a simple and interpretable model. Therefore, in order to reduce the input parameters' number, SMLR was applied as common method.

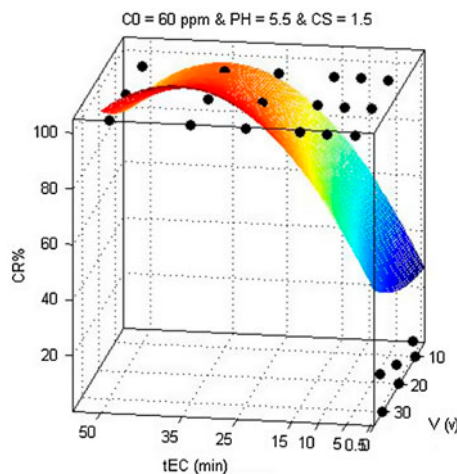


Fig. 3. Sample plot of the MLR-predicted values of CR% (colored surface) vs. experimental data of CR% (black dots), where $C_0 = 60$, $pH_0 = 5.5$, and $C_S = 1.5$.

As mentioned in the introduction section, SMLR is commonly used regression method, which is proposed to evaluate only a small number of subsets by either adding or deleting variables at a time, according to specific conditions. The number of remaining variables in the model is assigned based on the levels of significance assumed for inclusion and exclusion of variables from the model. In this study, the SMLR method neglects the additional parameters statistically.

Neglecting electrical current and conductivity by SMLR procedure is in good agreement with the fact that in EC studies, sometimes the voltage is considered as an operational parameter and sometimes electrical current or the related parameters like current density. This rises from Ohm's law; we know that if two parameters of ohm's equation were given, then the third one can be calculated from the equation. It means that by using two of these three parameters, the third one is excessive in modeling and causes the statistic problems like co-linearity and over-train. Then, when the combination of voltage, salt concentration, and pH were used as input variables, the electrical current is excessive. In addition, based on the conductivity definition, it can be said that when salt concentration and pH were used as input variables, the conductivity is excessive in the same manner.

3.3. ANN models

The best ANN models were separately constructed for CR% and EnC. In the case of CR%, the whole five parameters were considered as inputs of model, one hidden layer with 5 neurons and 0.18 learning rate were applied. The EnC model was constructed based on the three effective parameters and nine neurons for the only one hidden layer. The “*tansig*” transfer function was selected for input and hidden layer and “*purelin*” for output in both models [18,34]. Once the networks were trained, the weights and bias of each neuron and layer were saved in the ANN model. Then, they were used to estimate the test set. The (5:5:1) ANN for CR% and (3:9:1) ANN for EnC were trained using 120 data of the train set by the back propagation algorithm. The parameters of ANN models for CR% and EnC are shown in Table 4 and Table 5, respectively.

In order to show the predictability of ANN model of CR%, the scatter diagram of experimental CR% against ANN predicted CR% are presented in Fig. 4.

Nevertheless, Fig. 5 shows the samples of ANN model predictions and their corresponding experimental results. In following, the more statistical parameters of both ANN models are presented in Table 6. The comparison of Fig. 3 with Fig. 5 and Table 3 with Table 6 clearly shows that ANN models outperformed SMLR models. Therefore, ANN models would be used in optimization procedure by GA.

3.4. GA multi-objective optimization

Table 7 presents the Pareto optimal solutions. Each point on the Pareto set is associated with a set of decision variables. Moreover, the input decision variables corresponding to each of the Pareto optimal solutions were tabulated in Table 7.

Table 4
Network weights and biases of the ANN model for CR%

Neuron	Input layer to hidden layer weights					
	pH ₀	C ₀	V _{EC}	C _S	t _{EC}	Bias
n1	−0.9237	−0.47599	−0.07184	0.47857	6.4448	6.152
n2	−3.0781	−0.03258	0.00841	−4.057	−0.02157	−0.25904
n3	1.5493	0.3717	−0.55966	−0.62826	−1.2593	−0.37744
n4	4.3701	−0.08383	0.10222	−1.0587	0.32538	1.7183
n5	0.79223	0.27297	0.11829	−1.6336	0.6002	2.2133
	Hidden layer to output layer weights					
	n1	n2	n3	n4	n5	Bias
Output	3.4537	−4.9083	−1.135	−1.148	6.6809	−6.0168

n: neuron or processing elements.

Interpretation of Table 7 was used to illustrate and discriminate effect of each parameters investigated. As seen from Table 7, the optimum value of C₀ in all Pareto optimal solutions is almost 20 ppm (minimum of range). It means that more C₀ always causes less CR% and more EnC and it resulted in the selection of minimum of C₀ by GA in all conditions as an optimum.

However, the interpretation for the other four parameters is not as easy as C₀. Two optimum clusters can be discriminated in Table 7. One group with less CR% than 95% having corresponding optimum EnC values less than 0.1 wh and second group with CR% more than 95% having a corresponding optimum EnC 12.2–134.7.

These two groups of optimum conditions are corresponded to two different recognizable operational parameters. First group was corresponded to conditions with pH₀=9 (maximum value of range), C_S=0.7–1.1 (g L^{−1}), V_{EC}=10 V (minimum of range) and t_{EC}=1 min. Second group was corresponded to the conditions with pH₀=6.8 (except two cases), C_S=1.1–2 (g L^{−1}), V_{EC}=10–15.2 V, and t_{EC}=49.4–50 min.

It clarifies that higher pH₀ in presence of lower C_S, V_{EC}, and t_{EC} causes saving energy. In addition, pH₀≈6.8, higher V_{EC}, C_S, and t_{EC} can result in maximizing CR%. The optimum pH₀=6–7 for EC has been reported frequently [19,21].

It is a famous statement that the solid precipitate of aluminum hydroxide is formed at pH 4–6. Solubility of aluminum hydroxide increases when the solution becomes either more acidic or alkali.

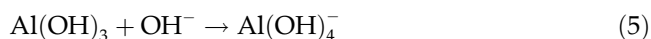


Table 5
Network weights and biases of the ANN model for EnC

Neuron	Input layer to hidden layer weights									
	V_{EC}	C_S				t_{EC}				Bias
n1	-0.96976	1.0665				2.762				2.8851
n2	-0.32997	2.0103				2.3174				2.3192
n3	0.4556	-2.4448				-1.845				-3.1891
n4	0.11071	2.7265				0.22363				1.2826
n5	-1.773	2.8924				0.59999				0.70601
n6	-0.6252	0.054423				-2.0213				-1.8544
n7	-0.51866	-0.7349				-0.8294				2.1073
n8	1.5587	-3.5239				-2.0161				0.464
n9	3.602	-1.3748				-0.26013				2.2255
Hidden layer to output layer weights										
Output	n1	n2	n3	n4	n5	n6	n7	n8	n9	Bias
	0.7027	0.47076	0.63391	1.0345	-0.55247	-0.74761	-1.455	-0.27083	0.34713	-1.5599

n: neuron or processing elements.

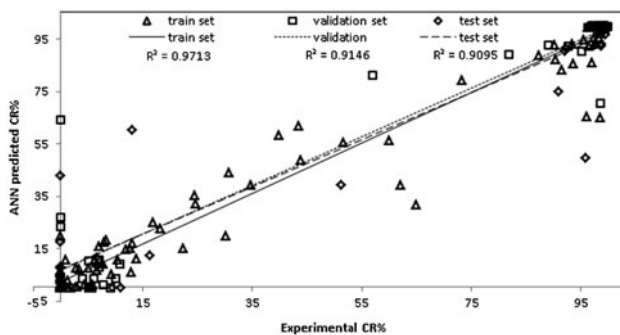


Fig. 4. Regression plot (actual vs. predicted) using four input variables, five processing elements in hidden layer, and two output variables using ANN model.

It is in good agreement with optimum value of pH_0 in the second group (6.8). But, at about pH_0 9 of the first group that is not in range, on one hand, it can be said that the prominence of this group rises from other parameters except the pH_0 and on the other hand, this group has optimum value of EC not CR%.

The other point that can be extracted from Table 7 is the fact that EC is efficient for dye removal until certain level of removal (not more). In the other hand, in this work as an example, to achieve CR% upper than 95%, the EnC rapidly increases. It is a good demonstration of new interest in combining EC with other methods like adsorption to reach complete removal and efficient EnC at the same time [36].

In this study, the final conductivity and pH were measured and compared with correspond initial

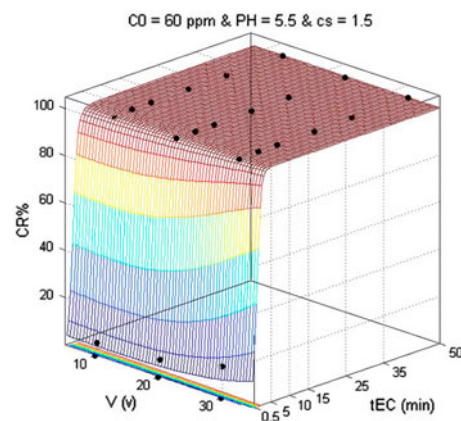


Fig. 5. Sample plot of the ANN-predicted values of CR% (colored surface) vs. experimental data of CR% (black dot), where $C_0 = 60$, $pH_0 = 5.5$, and $C_S = 1.5$.

values. The comparison of these values was demonstrated in Fig. 6. This figure demonstrates positive variation in pH in more runs. Positive variation rises from OH^- production during EC process. On the other hand, negative variation happens for runs with $pH = 10$. However, the continuous production of OH^- occurred during EC process, the final pH did not reach 10 even for the runs with initial $pH = 10$. It is in good agreement with the fact in this kind of EC studies, more OH^- produce $Al(OH)_4^-$ from $Al(OH)_3$ according to the chemical Eq. (6). Then, it can be said that $Al(OH)_3$ acts as basic buffer preventing the pH to be more than 9.

Table 6
Statistical characteristics of both ANN models for CR% and EnC

Model	CR%			EnC (wh)		
	Train set	Validation set	Test set	Train	Validation	Test
R^2	0.971	0.914	0.909	0.986	0.963	0.972
Q^2	0.971	0.912	0.908	0.943	0.960	0.970
RMSE	7.5	13.3	13.6	4.37	5.53	9.68

Table 7
Variables values corresponding to each of the Pareto optimal solutions

Solutions	pH ₀	C ₀	V _{EC}	C _S	t _{EC}	CR%	EnC	Solutions	pH ₀	C ₀	V _{EC}	C _S	t _{EC}	CR%	EnC
1	9.0	20.0	10.0	0.7	1.1	55.4	<0.1	15	9.0	20.0	10.0	0.7	1.0	64.4	<0.1
2	9.0	20.0	10.0	0.7	1.0	53.6	<0.1	16	6.9	20.2	15.2	2.0	50.0	99.9	134.7
3	6.8	20.0	14.4	2.0	49.7	99.9	93.0	17	6.8	20.0	14.6	2.0	50.0	99.9	100.0
4	7.2	20.0	10.0	1.8	49.3	99.9	33.1	18	6.9	20.0	15.1	2.0	49.9	99.9	125.0
5	9.0	20.0	10.0	0.8	1.0	85.6	<0.1	19	9.0	20.0	10.0	0.7	1.0	69.6	<0.1
6	6.8	20.0	12.2	2.0	49.9	99.9	49.2	20	6.9	20.2	14.8	2.0	50.0	99.9	108.5
7	9.0	20.1	10.0	0.7	1.1	71.6	<0.1	21	8.9	20.0	12.2	1.1	49.4	99.9	12.2
8	9.0	20.0	10.0	0.8	1.0	83.8	<0.1	22	9.0	20.1	12.5	1.2	1.0	94.8	<0.1
9	6.8	20.0	15.2	2.0	50.0	99.9	131.2	23	6.8	20.1	14.1	2.0	50.0	99.9	81.1
10	6.8	20.0	12.2	1.8	49.9	99.9	45.2	24	6.8	20.0	10.0	2.0	50.0	99.9	38.05
11	9.0	20.0	10.1	0.8	1.1	78.1	<0.1	25	9.0	20.0	10.0	0.7	1.0	53.6	<0.1
12	9.0	20.0	10.0	0.9	1.0	90.8	<0.1	26	9.0	20.0	12.5	1.2	49.4	99.9	21.9
13	6.8	20.0	14.9	2.0	50.0	99.9	118.2	27	8.8	20.1	12.4	1.1	49.4	99.9	14.5
14	6.8	20.0	12.8	2.0	50.0	99.9	55.9	28	6.8	20.0	13.8	2.0	50.0	99.9	74.1

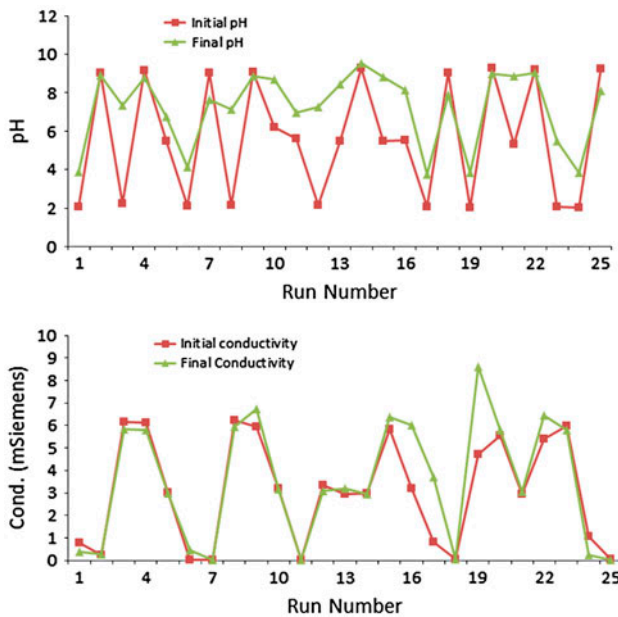


Fig. 6. Conductivity and pH variation during EC process.

In addition, figure ? shows the small positive and negative variation for conductivity. The positive

variation rises from ions species production like OH⁻, Al³⁺ and its hydroxyl species like Al(OH)₂⁺, and Al(OH)₂²⁺ during EC process. On the other hand, negative variation that happened for pH=2 samples (unless one sample) rises from the H⁺ removal from solution as an ion with highest conductivity.

4. Conclusion

The SMLR and ANN approaches were successfully used to construct a QSAR model. Further, statistical tests exhibit models' goodness. Based on statistical comparison, ANN model obviously outperformed SMLR model. Finally, multi-objective optimization of CR% and EnC was successfully carried out using genetic algorithm (GA) over the outperformed ANN models. The optimization procedure causes nondominated optimal points, which gave an insight into optimal operating conditions of the EC. The optimum value of C₀ in all Pareto optimal solutions was almost 20 ppm (minimum of range). It means that more C₀ always causes less CR% and more EnC and it resulted in the selection of minimum of C₀ by GA in all conditions as an optimum. Two optimum clusters can be discriminated regarding two

objective parameters. One group with less CR% than 95% having corresponding optimum EnC values less than 0.1 wh and second group with CR% more than 95% having a corresponding optimum EnC 12.2–134.7. First group was corresponded to conditions with $\text{pH}_0 = 9$ (maximum value of range), $C_S = 0.7\text{--}1.1$ (g L^{-1}), $V_{\text{EC}} = 10$ V (minimum of range), and $t_{\text{EC}} = 1$ min. Second group was corresponded to the conditions with $\text{pH}_0 = 6.8$ (except two cases), $C_S = 1.1\text{--}2$ (g L^{-1}), $V_{\text{EC}} = 10\text{--}15.2$ V, and $t_{\text{EC}} = 49.4\text{--}50$ min. In addition, the achieved CR% (0–99.9) and EnC (0–69.4 wh) approves two statements; first, the good efficiency of EC in color removal and second, the operational parameter's importance and influence.

Symbols

CR%	—	color removal efficiency, percent
pH_0	—	initial pH of dye solution
C_0	—	initial dye concentration, mg L^{-1}
V_{EC}	—	applied voltage, V
C_S	—	initial electrolyte (NaCl) concentration, g L^{-1}
t_{EC}	—	treatment time, min
EnC	—	energy consumption, watt-hour
A_0	—	initial absorption of dye solution
A	—	absorption of solution after EC
CCD	—	central composite design
FFD	—	full factorial design
DOE	—	design of experiment
ANN	—	artificial neural network
SMLR	—	stepwise multiple linear regression
GA	—	genetic algorithm
RSM	—	response surface methodology
DO25	—	Disperse Orange 25
RMSE	—	root mean square error

References

- [1] E. Rosales, M. Pazos, M.A. Sanroman, Comparative efficiencies of the decolorisation of leather dyes by enzymatic and electrochemical treatments, *Desalination* 278 (2011) 312–317.
- [2] M.S.M. Eldin, S. El-Sakka, M. El-Masry, I. Abdel-Gawad, S. Garybe, Removal of methylene blue dye from aqueous medium by nano poly acrylonitrile particles, *Desalin. Water Treat.* 44 (2012) 151–160.
- [3] M.M.A.B.D. El-Latif, A.M. Ibrahim, Removal of reactive dye from aqueous solutions by adsorption onto activated carbons prepared from oak sawdust, *Desalin. Water Treat.* 20 (2010) 102–113.
- [4] H. Kusic, D. Juretic, N. Koprivanac, V. Marin, A.L. Bozic, Photooxidation processes for an azo dye in aqueous media: Modeling of degradation kinetic and ecological parameters evaluation, *J. Hazard. Mater.* 185 (2011) 1558–1568.
- [5] H. Ma, M. Wang, R. Yang, W. Wang, J. Zhao, Z. Shen, S. Yao, Radiation degradation of Congo Red in aqueous solution, *Chemosphere* 68 (2007) 1098–1104.
- [6] E.S.Z. El-Ashtoukhy, N.K. Amin, Removal of acid green dye 50 from wastewater by anodic oxidation and electrocoagulation—A comparative study, *J. Hazard. Mater.* 179 (2010) 113–119.
- [7] M. Kobya, M. Bayramoglu, M. Eyvaz, Techno-economical evaluation of electrocoagulation for the textile wastewater using different electrode connections, *J. Hazard. Mater.* 148 (2007) 311–318.
- [8] D. Belhout, D. Ghernaout, S. Djeddar-Douakh, A. Kellil, Electrocoagulation of a raw water of Ghib Dam (Algeria) in batch using aluminium and iron electrodes, *Desalin. Water Treat.* 16 (2010) 1–9.
- [9] E.S.Z. El-Ashtoukhy, T. Zewail, N. Amin, Removal of heavy metal ions from aqueous solution by electrocoagulation using a horizontal expanded Al anode, *Desalin. Water Treat.* 20 (2010) 72–79.
- [10] G.H. Chen, Electrochemical technologies in wastewater treatment, *Sep. Purif. Technol.* 38 (2004) 11–41.
- [11] M. Kobya, O.T. Can, M. Bayramoglu, Treatment of textile wastewaters by electrocoagulation using iron and aluminum electrodes, *J. Hazard. Mater.* 100 (2003) 163–178.
- [12] C. Phalakornkule, P. Sukkasem, C. Mutchimsattha, Hydrogen recovery from the electrocoagulation treatment of dye-containing wastewater, *Int. J. Hydrogen Energy* 35 (2010) 10934–10943.
- [13] T. Olmez, The optimization of Cr(VI) reduction and removal by electrocoagulation using response surface methodology, *J. Hazard. Mater.* 162 (2009) 1371–1378.
- [14] J.P. Wang, Y.Z. Chen, X.W. Ge, H.Q. Yu, Optimization of coagulation-flocculation process for a paper-recycling wastewater treatment using response surface methodology, *Colloid Surf. A* 302 (2007) 204–210.
- [15] O. Prakash, M. Talat, S.H. Hasan, Response surface design for the optimization of enzymatic detection of mercury in aqueous solution using immobilized urease from vegetable waste, *J. Mol. Catal. B Enzym.* 56 (2009) 265–271.
- [16] Y. Safa, H.N. Bhatti, Adsorptive removal of direct textile dyes by low cost agricultural waste: Application of factorial design analysis, *Chem. Eng. J.* 167 (2011) 35–41.
- [17] G.S. Kapur, S. Mukherjee, A.S. Sarpal, S.K. Jain, Development of a ^{13}C -NMR spectroscopic method for estimation of heavy alkylated benzene (HAB) in industrial oils using stepwise multiple linear regression, *Lubr. Eng.* 54 (1998) 21–28.
- [18] H. Daraei, M. Irandoust, J. Ghasemi, A. Kurdian, QSPR probing of Na^+ complexation with 15-crown-5 ethers derivatives using artificial neural network and multiple linear regression, *J. Inclusion Phenom. Macro.* 72 (2012) 423–435.
- [19] M.S. Bhatti, D. Kapoor, R.K. Kalia, A.S. Reddy, A.K. Thukral, RSM and ANN modeling for electrocoagulation of copper from simulated wastewater: Multi objective optimization using genetic algorithm approach, *Desalination* 274 (2011) 74–80.
- [20] T. Olmez-Hanci, Z. Kartal, İ. Arslan-Alaton, Electrocoagulation of commercial naphthalene sulfonates: Process optimization and assessment of implementation potential, *J. Environ. Manage.* 99 (2012) 44–51.
- [21] M. Tir, N. Moulai-Mostefa, Optimization of oil removal from oily wastewater by electrocoagulation using response surface method, *J. Hazard. Mater.* 158 (2008) 107–115.
- [22] M.T. Hagan, H.B. Demuth, M.H. Beale, *Neural Network Design*, 1st ed., PWS Pub, Boston, MA, 1996.
- [23] D.M. Himmelblau, Applications of artificial neural networks in chemical engineering, *Korean J. Chem. Eng.* 17 (2000) 373–392.
- [24] A. Ames, E. Rogel-Hernandez, S.W. Lin, L.Z. Flores-López, J.R. Castro, F.T. Wakida, S. Melendez, H. Espinoza-Gomez, Prediction of metal ion rejection in electro-cross-flow ultrafiltration using an artificial neural network, *Desalin. Water Treat.* 36 (2011) 105–118.
- [25] J. Saien, A.R. Soleymani, H. Bayat, Modeling Fenton advanced oxidation process decolorization of Direct Red 16 using artificial neural network technique, *Desalin. Water Treat.* 40 (2012) 174–182.
- [26] G. Singh, J. Kandasamy, H. Shon, J. Cho, Measuring treatment effectiveness of urban wetland using hybrid water quality—Artificial neural network (ANN) model, *Desalin. Water Treat.* 32 (2011) 284–290.

- [27] B.K. Körbahti, K. Artut, C. Geçgel, A. Özer, Electrochemical decolorization of textile dyes and removal of metal ions from textile dye and metal ion binary mixtures, *Chem. Eng. J.* 173 (2011) 677–688.
- [28] S. Tsutsui, D.E. Goldberg, Search space boundary extension method in real-coded genetic algorithms, *Inf. Sci.* 133 (2001) 229–247.
- [29] A. Hasseine, A. Kabouche, A.H. Meniai, M. Korichi, Salting effect of NaCl and KCl on the liquid–liquid equilibria of water+ ethyl acetate+ ethanol system and interaction parameters estimation using the genetic algorithm, *Desalin. Water Treat.* 29 (2011) 47–55.
- [30] G. Sashi Kumar, A. Mahendra, A. Sanyal, G. Gouthaman, Genetic algorithm-based optimization of a multi-stage flash desalination plant, *Desalin. Water Treat.* 1 (2009) 88–106.
- [31] M. Tabesh, S. Hoomehr, Consumption management in water distribution systems by optimizing pressure reducing valves' settings using genetic algorithm, *Desalin. Water Treat.* 2 (2009) 95–100.
- [32] D. Ghernaout, A. Mariche, B. Ghernaout, A. Kellil, Electromagnetic treatment-doubled electrocoagulation of humic acid in continuous mode using response surface method for its optimisation and application on two surface waters, *Desalin. Water Treat.* 22 (2010) 311–329.
- [33] S.T. Ong, E.C. Khoo, P.S. Keng, S.L. Hii, S.L. Lee, Y.T. Hung, S.T. Ha, Plackett-Burman design and response surface methodological approach to optimize basic dyes removal using sugarcane bagasse, *Desalin. Water Treat.* 25 (2011) 310–318.
- [34] T. Hatami, M. Rahimi, H. Daraei, E. Heidaryan, A.A. Alsairafi, PRSV equation of state parameter modeling through artificial neural network and adaptive network-based fuzzy inference system, *Korean J. Chem. Eng.* 29 (2012) 657–667.
- [35] S.P. Niculescu, Artificial neural networks and genetic algorithms in QSAR, *J. Mol. Struct. (Thoechem)* 622 (2003) 71–83.
- [36] N.V. Narayanan, M. Ganesan, Use of adsorption using granular activated carbon (GAC) for the enhancement of removal of chromium from synthetic wastewater by electrocoagulation, *J. Hazard. Mater.* 161 (2009) 575–580.