Desalination and Water Treatment
www.deswater.com
doi: 10.5004/dwt.2017.20235

64 (2017) 54–63
February

Evolutionary prediction of electrocoagulation efficiency and energy consumption probing

Azam Akhbari[a], Hossein Bonakdari[b,*], Isa Ebtehaj[b]

[a]*Department of Civil Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur, Malaysia, email: aazam.akhbari@yahoo.com*
[b]*Department of Civil Engineering, Razi University, Kermanshah, Iran, Phone: +98 833 427 4537, Fax: +98 833 428 3264, emails: bonakdari@yahoo.com (H. Bonakdari), isa.ebtehaj@gmail.com (I. Ebtehaj)*

Received 31 May 2016; Accepted 21 September 2016

ABSTRACT

This case study focuses on the color removal efficiency (CR%) and energy consumption (EnC) of electrocoagulation (EC) using synthetic wastewater containing disperses like orange 25 dye. Process parameters including initial pH (pH_0), initial dye concentration (C_0), applied voltage (V_{EC}), initial electrolyte concentration (C_S) and treatment time (t_{EC}) were found to be the more effective EC operational parameters to attain maximum decolorization efficiency. In order to investigate the effect of the independent variables on dye removal and determine the optimum condition, gene expression programming (GEP) was used, and the results were compared with the reduced quadratic multiple regression model (SMLR) method. The results indicate that the proposed model predicted the CR% with MARE of 17.28 and RMSE of 6.24, and EnC with MARE of 54.876 and RMSE of 5.33. This model can make more accurate predictions than the SMLR equations. The GEP technique presents two simple equations for predicting CR% and EnC for practical engineering. Also, two different explicit expressions are presented to estimate CR% and EnC as an alternative tool in practical. Moreover, a partial derivative sensitivity analysis was used to indicate the trend of each parameter in the proposed models.

Keywords: Artificial intelligence; Color removal; Electrocoagulation; Energy consumption; Sensitivity analysis

1. Introduction

Wastewater from the dye industry has serious, negative effects on aquatic ecosystems and human health, raising wide concerns and also causing problems for conventional biological wastewater treatment plants due to the various organic and inorganic chemical compounds involved. The colored and toxic wastewater released into the ecosystem undergoes chemical as well as biological changes. It also consumes dissolved oxygen from streams, and it is a dramatic source of esthetic pollution and perturbation for aquatic life. An estimated 50,000 tons of dye are discharged from the dyeing and coloration industries every year.

Dyeing wastewater is conventionally treated by various methods like adsorption, precipitation, chemical degradation, advanced oxidation processes, biodegradation and chemical coagulation. Despite the widespread application of these methods, they have some disadvantages [1,2].

For example, biological methods are time consuming and often ineffective in removing dyes, which are highly structured polymers with low biodegradability. Biological means cannot be applied to most textile wastewater due to the majority of commercial dyes toxicity to the organisms found in the process [3,4]. Activated carbon adsorption is associated with a costly and difficult regeneration process as well as high waste disposal cost [5]. Chemical coagulation causes extra pollution due to the undesired reactions in treated water, and it produces large amounts of sludge [1]. Chemical degradation by oxidative agents such as chlorine is the most important and effective method, but it produces some very toxic products like organochlorine compounds [6]. Advanced oxidation processes including ozonation, UV and ozone–UV combined oxidation, photocatalysis, Fenton

* Corresponding author.

reaction and ultrasonic oxidation are not economically feasible [5]. Furthermore, these methods are often expensive, and treatment efficiency is inadequate because of the large variability in textile wastewater compositions [1].

In recent years, strict environmental regulations have called for new processes to attain efficient and adequate treatment of various industrial wastewaters at relatively low operating costs. In this regard, the electrocoagulation (EC) process has attracted a great deal of attention for treating industrial wastewater owing to its versatility and environmental compatibility. This method is characterized by simple equipment, easy operation, a reduced reactive retention period, less or no equipment for adding chemicals and lower amounts of precipitate or sludge, which sediments rapidly. EC has been proven to be an efficient method for wastewater treatment. It has been tested successfully for treating municipal wastewater [7], textile wastewater [8], poultry manure wastewater [9], landfill leachate [10] and rose processing wastewater [11].

The artificial neural network (ANN) technique has recently been widely utilized in several disciplines including ecological and environmental engineering [12]. Classical polynomial regression techniques are employed to establish explicit parametric relationships between variables, but are often limited in applicability by the need to satisfy predefined fitness functions. Thus, a need has gradually emerged in contemporary metamodeling domains to combine the inherent efficiency, robustness and speed of ANNs with the clarity of the explicit analytical expressions of polynomial regression.

Gene expression programming (GEP) overcomes such limitations and provides closed-form analytical expressions for parametric evaluation and analysis [13,14]. Given its inherent search structure based on the evolution principle, GEP is not constrained by topology selection or the iteration algorithms of the ANN technique. Furthermore, in contrast to being obscured by complex weight matrices as in ANNs, the evolved model responses are explicit analytical functions of simpler and more apt mathematical operators conducive to the problem under study [15–17]. Despite the prediction capability of artificial intelligence-based techniques, very few GEP applications have been reported in recent literature with focus on the color removal efficiency (CR%) and energy consumption (EnC) of EC.

Therefore, the aim of the present work is development of GEP as a strong tool not only for estimating models but also for presenting certain relationships in practical instances accordingly. First, the effective parameters are recognized and introduced different models to survey the effect of each parameter in CR% and EnC prediction. Also, the results of proposed models are compared with existing method. Moreover, to study the trend of each parameter in proposed models, the partial derivative sensitivity analysis is employed.

2. Materials and methods

2.1. Source of data set

To estimate the CR% and EnC of EC, Maleki et al.'s [18] experimental data were utilized. Maleki et al. [18] conducted experiments in an EC system consisting of a glass (12 × 12 × 21 cm) cubic reactor, 400 rpm mixer, DC power supply (high stability and reliability, and low-noise DC adjustable power supply RXN-303D-II, Zhaoxin Communications Industrial Co., Ltd., China) and two aluminum electrodes. The cathode and anode were made of aluminum sheets (4 × 5 × 0.1 cm), and the immersed surface area of each electrode was 40 cm^2. The electrodes were placed vertically and dipped in 1.5 L aqueous dye solution. The distance between electrodes was fixed at 1 cm. Table 1 represents the range of data used in Maleki et al.'s [18] tests. The parameters that affect CR% and EnC are as follows: initial pH (pH_0), initial dye concentration (C_0), applied voltage (V_{EC}), initial electrolyte concentration (C_S) and treatment time (t_{EC}).

Table 1
Range of data in Maleki et al.'s [18] study

Run No.	1–25
pH_0	2–9
C_0	8–100
V_{EC}	10–30
C_S	0–3
t_{EC}	0.5–50
EnC	0.001–76
CR%	0.2–99.9

2.2. Gene expression programming (GEP)

Evolutionary algorithms are problem-solving techniques proposed based on the Darwinian evolutionary theory. By using natural selection and a search among a population of solutions, an evolutionary algorithm performs a selection process to select the best solution. The selected solution is a possible acceptable solution as an individual of the population. In each repetition of an evolutionary algorithm, a competitive selection occurs with the lowest accuracy answer being eliminated from the assessment process of the fitness value, which indicates the quality of an individual solution to the problem. GEP is an evolutionary method that was introduced by Ferreira [13]. The most important feature of this method is facilitating chromosomes to be expression trees (ET). GEP is a developed model of genetic programming (GP) that exhibits much greater accuracy than GP [14]. In GEP, complex equations with simple linear structures and fixed lengths are called chromosomes, which are encoded. Chromosomes include linear strings of fixed length that may cover one or more different genes. Each gene has a head and a tail. The gene tail contains a junk sequence of terminals that enable gene modification by each genetic operator with no restrictions. Therefore, the gene tail is of considerable importance in genes. The head (h) length is user-defined according to the problem type, but the tail (t) length is associated with the head length, and the variable number (n) is calculated as $t = h(n - 1) + 1$. In addition, the function set for each problem should be determined prior to modeling. Such functions are fundamental to evolution in GEP as they allow modification with no limitation among similar genes or multiple genes in chromosomes with more than one gene. According to the initial problem set, GEP distributes terminals and functions randomly among chromosome genes. The initial population, which is generated randomly, is known as the parent. The purpose of creating parents is to achieve offspring via high genetic operator performance. Each individual uses its own genetic information to help create a new offspring that has a greater chance of survival. In the evolution process of a function, natural selection is based on the offspring equation search that produces less estimation error. The fitness function value employed in this study for program f_i is calculated with the following equation:

$$f_i = \frac{1000}{1+\mathrm{RRSE}_j} \tag{1}$$

where RRSE is the root relative square error for the ith offspring. The fitness function ranges between 0 and 1,000 (worst and best fitness, respectively). The RRSE is calculated as follows:

$$\mathrm{RRSE}_j = \sqrt{\frac{\sum_{i=1}^{n}\left(P_{ij}-O_i\right)^2}{\sum_{i=1}^{n}\left(O_i-\bar{O}\right)^2}} \tag{2}$$

where P_{ij} is the value predicted by program j; O_i is the observed value for fitness i; $\bar{O}$ is the average of all observed values; and n is the number of samples. The GEP parameters are presented in Table 2. JEdit, an open source software package, was used for the implementation of GEP models [19].

Table 2
GEP parameters

Parameter	CR%	EnC
Number of generations	300,000	300,000
Number of chromosomes	200	200
Number of genes	3	3
Head size	12	12
Function set	+,−,×,/, −, X2, 3Rt, 4Rt, 5Rt, sin, cos, Tan, Logi2	+,−,×,/, −, X2, X3, X5, Atan, cos, sin, 4rt, Gau2, Pow
Linking function	Addition	Addition
Mutation rate	0.084	0.01
Inversion rate	0.15	0.15
IS transposition rate	0.15	0.15
RIS transposition rate	0.15	0.15
Gene transposition rate	0.15	0.15
One-point recombination rate	0.30	0.30
Two-point recombination rate	0.35	0.30
Gene recombination rate	0.15	0.15

2.3. GEP models

The effective EC operational parameters are generally the initial pH (pH_0), initial dye concentration (C_0), applied voltage (V_{EC}), initial electrolyte concentration (C_S) and treatment time (t_{EC}). Therefore, to consider all these parameters in estimating the CR% and EnC, two functional relationships are provided as follows:

$$\mathrm{CR}(\%) = f(\mathrm{pH}_0, C_0, V_{EC}, C_S, t_{EC}) \tag{3}$$

$$\mathrm{EnC} = f(\mathrm{pH}_0, C_0, V_{EC}, C_S, t_{EC}) \tag{4}$$

For both CR% and EnC parameters, the models are as follows:

$$\mathrm{GEP(1)}:\mathrm{CR}(\%) \text{ or } \mathrm{EnC} = f(\mathrm{pH}_0, C_0, V_{EC}, C_S, t_{EC})$$

$$\mathrm{GEP(2)}:\mathrm{CR}(\%) \text{ or } \mathrm{EnC} = f(\mathrm{pH}_0, C_0, V_{EC}, C_S)$$

$$\mathrm{GEP(3)}:\mathrm{CR}(\%) \text{ or } \mathrm{EnC} = f(\mathrm{pH}_0, C_0, V_{EC}, t_{EC})$$

$$\mathrm{GEP(4)}:\mathrm{CR}(\%) \text{ or } \mathrm{EnC} = f(\mathrm{pH}_0, C_0, C_S, t_{EC})$$

$$\mathrm{GEP(5)}:\mathrm{CR}(\%) \text{ or } \mathrm{EnC} = f(\mathrm{pH}_0, V_{EC}, C_S, t_{EC})$$

$$\mathrm{GEP(6)}:\mathrm{CR}(\%) \text{ or } \mathrm{EnC} = f(C_0, V_{EC}, C_S, t_{EC})$$

2.4. Reduced quadratic multiple regression (SMLR) models

Maleki et al. [18] proposed two SMLR models separately for CR% and EnC after neglecting the non-significant terms [20]. The two models are given in Eqs. (5) and (6) for CR% and EnC, respectively:

$$\begin{aligned}\mathrm{CR}(\%) = {} & 19.8554 + 4.1203(t_{EC}) + 0.0296(\mathrm{pH}_0)(C_0) \\ & + 3.6542(\mathrm{pH}_0)(C_S) + 0.1034(\mathrm{pH}_0)(t_{EC}) \\ & + 0.1338(C_0)(C_S) + 0.2138(C_S)(t_{EC}) \\ & - 0.3291(\mathrm{pH}_0)^2 + 0.0176(V_{EC})^2 - 6.7969(C_S)^2 \\ & + 0.05.35(t_{EC})^2\end{aligned} \tag{5}$$

$$\begin{aligned}\mathrm{EnC} = {} & 5.7837 - 6.3669(C_S) - 0.7365(t_{EC}) \\ & + 0.3223(V_{EC})(C_S) + 0.0287(V_{EC})(t_{EC}) \\ & + 0.2689(C_S)(t_{EC}) - 0.0106(V_{EC})^2 + 0.0045(t_{EC})^2\end{aligned} \tag{6}$$

To check the generalization ability of proposed models, the k-fold cross validation is applied. The number of k is considered as 10 (McLachlan et al. [21]). In this technique, all data are randomly partitioned in 10 equal sub-samples. Among the 10 sub-samples, a single sub-sample is selected as validation, and the rest of the sub-samples (k – 1 = 9 sub-samples) are employed for training model. This cross validation is repeated 10 times until all sub-samples are considered for validation.

3. Results and discussion

In this section, the CR% and EnC results are estimated using the SMLR and GEP methods. For this purpose, the statistical indices applied are root mean square error (RMSE), mean absolute relative error (MARE), scatter index (SI) and BIAS. The equation of each statistical index is defined as follows:

$$R^2 = \left(\frac{\sum_{i=1}^{n} (O_i - \overline{O}_i).(P_i - \overline{P}_i)}{\sqrt{\sum_{i=1}^{n} (O_i - \overline{O}_i)^2 \sum_{i=1}^{n} (P_i - \overline{P}_i)^2}} \right)^2 \tag{7}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (O_i - P_i)^2} \tag{8}$$

$$\mathrm{MARE} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{|O_i - P_i|}{O_i} \right) \tag{9}$$

$$\mathrm{SI} = \frac{\mathrm{RMSE}}{\frac{1}{n} \sum_{i=1}^{n} (\bar{O}_i)} \tag{10}$$

$$\mathrm{BIAS} = \frac{\sum_{i=1}^{n} (P_i - O_i)}{n} \tag{11}$$

where O_i and P_i are the observed and predicted EnC, repectively; $\bar{O}_i$ and $\bar{P}_i$ are the average of observed and predicted value of EnC; and n is the number of parameters.

Due to the above-mentioned criteria not consider the average and variance of a model simultaneously, the Akaike information criterion (AIC) are utilized to compare the GEP model with existing model. This index characterizes a trade-off between variance and bias in model construction and offers a tool for model selection by considering the complexity and accuracy of model, simultaneously [22]. The AIC is calculated as follows:

$$\mathrm{AIC} = n.\log\left(\frac{1}{n} \sum_{i=1}^{n} (O_i - P_i)^2 \right) + 2K \tag{12}$$

where n and k are the number of data and estimated parameter in model, respectively.

Fig. 1 compares the performance of the GEP models in predicting CR%. It is evident that this model made underestimations and overestimations. It can be seen that all model estimations have a relative error of less than 15%, which indicates good model accuracy. Concerning Table 3, the SMLR method demonstrates weaker performance in CR% prediction (MARE = 122.09, RMSE = 24.13, SI = 0.35, BIAS = 0.23) compared with GEP, especially models 1 and 4 (MARE = 17.28 and 19.43; RMSE = 6.27 and 6.24; SI = 0.09 and 0.09; BIAS = 1.28 and 1.10, respectively). Therefore, it was found that a method is required that can increase model flexibility without significant effect on CR% prediction. Consequently, the GEP method was applied to resolve this matter. The estimations by the proposed method have less than 15% relative error; as a result, GEP is considered more accurate than SMLR. As seen from Table 3, the BIAS and SI indices of models 5 and 6 have the smallest values among the six GEP models. Models 1 and 4 predicted the CR% well. The figure also indicates that the GEP model with the highest R^2 (0.98) for CR% exhibits the highest accuracy.

The relationship using GEP to calculate CR% is as follows (Eq. (13)):

$$\begin{aligned} \mathrm{CR}(\%) = &\left(a\tan\left(t_{EC}{}^{5} \times \exp\left(-\left(\left(\mathrm{pH}_0 + \cos(C_0) \right)^2 \right) \right) \right) \right)^{10} \\ &+ \left(\mathrm{pH}_0 \left(t_{EC} + C_0 - C_0 \sin\left(66.09^{\mathrm{pH}_0} - C_S \right) \right) \right)^{0.25} \\ &+ \left(\left(\exp\left(-\left((V_{EC} - 30.52)^2 \right) \right) \right) \times \left(C_0 + 831379 \left(\exp\left(-\left((\mathrm{pH}_0 + t_{EC})^2 \right) \right) \right) \right) + t_{EC} \right)^{0.25} \end{aligned} \tag{13}$$

To evaluate the proposed methods performance and compare it with the existing methods, the statistical indices used are calculated in Table 4. The GEP performance evaluation in predicting EnC is presented in Fig. 2. This figure shows the relative error values for EnC estimation using GEP and six different models. The majority of estimations with this method have less than 15% mean relative error. The quantitative GEP performance in EnC estimation with model 6 also shows that the R^2 and MARE are higher than the SMLR method, which confirms the adequately predicted values. Similar to models 5 and 6, model 1 estimated the EnC with good accuracy as well. The estimations by both methods proposed in this study are thoroughly examined. According to parameters C_0, V_{EC}, C_S and t_{EC} that were used as input combinations to predict EnC (Eq. (14)), the proposed SMLR method did not perform well in estimating (MARE = 658.54, RMSE = 5.33). Hence, this method exhibited lower accuracy than GEP models 1, 5 and 6. The GEP relationship to calculate EnC is as follows (Eq. (14)):

$$\begin{aligned} \mathrm{EnC} = &\sin\left(\tan(5.63 + V_{EC}) \right)^2 S \sin\left(C_S{}^2 \right)\left(C_S t_{EC} \right) \\ &+ \left(C_S \left(\sin\left(t_{EC}{}^{0.25} - 7.34 \right) \right)^2 \left(1/\left(1 + \exp(-(x+y)) \right) \right)^3 \right)^2 \\ &+ \tan\left(1/\left(1 + \exp\left(-\left(t_{EC} + \mathrm{Cos}(V_{EC}) + \tan(C_0 + 8.03) + \left((C_0 + V_{EC})^{1/3} \right) - 7.91 \right) \right) \right) \right) \end{aligned} \tag{14}$$

3.1. Sensitivity analysis

In this study, to analyze the pattern changes in the proposed relationships according to the input parameters considered for each, partial derivative sensitivity analysis is applied [23]. With this method, the partial derivative of the equation presented to each input parameter is calculated. Then the pattern changes for different input parameter values from which the derivative of each relationship has been calculated are studied. It is evident that the magnitude of the calculated partial derivative is directly related to its effect on the estimated result. The positive and negative values of a partial derivative show that increasing the input parameter value leads to a decrease or increase in the results, respectively.

Fig. 3 presents the partial derivative results of Eq. (13) for the parameters provided in this relationship. The partial derivatives for the V_{EC} and t_{EC} parameter values are positive. The increase (or decrease) in these two parameters leads to an increase (or decrease) in the estimated parameter CR% using Eq. (13). It is worth noting that the derivative of Eq. (13) has

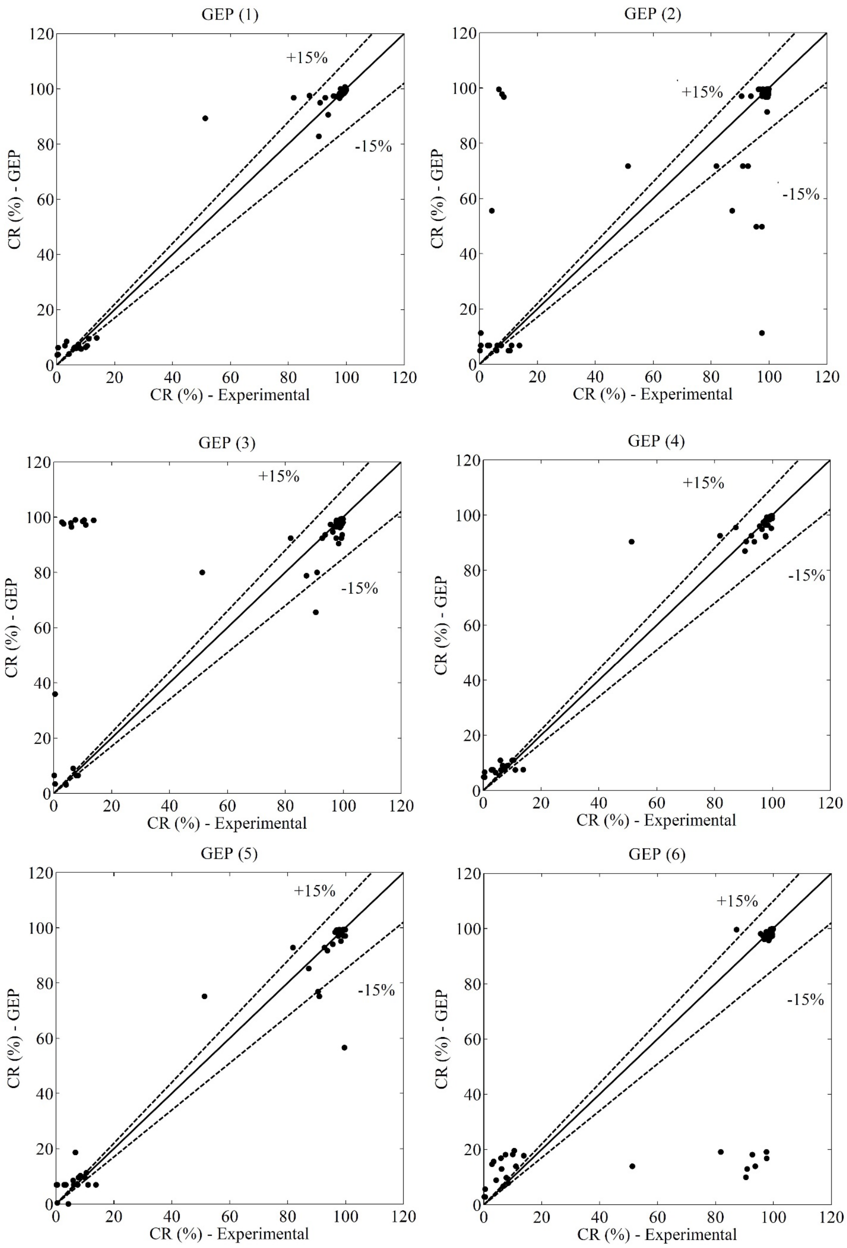

Fig. 1. GEP performance evaluation in CR% prediction.

Table 3
Statistical indices for CR% modeling using GEP

CR%	R^2	MARE	RMSE	SI	BIAS	AIC
GEP (1)	0.98	17.28	6.27	0.09	1.28	102.92
GEP (2)	0.57	113.62	28.35	0.42	1.56	171.06
GEP (3)	0.32	271.34	38.37	0.56	15.47	184.74
GEP (4)	0.98	19.43	6.24	0.09	1.10	102.70
GEP (5)	0.96	21.32	8.06	0.12	–0.39	114.26
GEP (6)	0.62	48.97	28.93	0.42	–9.11	171.98
SMLR	0.67	122.09	24.13	0.35	0.23	173.78

Table 4
Statistical indices for EnC modeling using GEP

EnC	R^2	MARE	RMSE	SI	BIAS	AIC
GEP (1)	0.98	106.13	2.50	0.24	–0.26	73.38
GEP (2)	0.55	264.31	12.87	1.25	–4.13	147.4
GEP (3)	0.59	1,282.62	11.37	1.10	–1.12	141.8
GEP (4)	0.48	150.07	12.10	1.17	–0.10	144.61
GEP (5)	0.98	89.21	2.25	0.22	0.15	68.63
GEP (6)	0.98	54.86	2.79	0.27	–0.44	78.34
SMLR	0.90	658.54	5.33	0.52	0.11	95.58

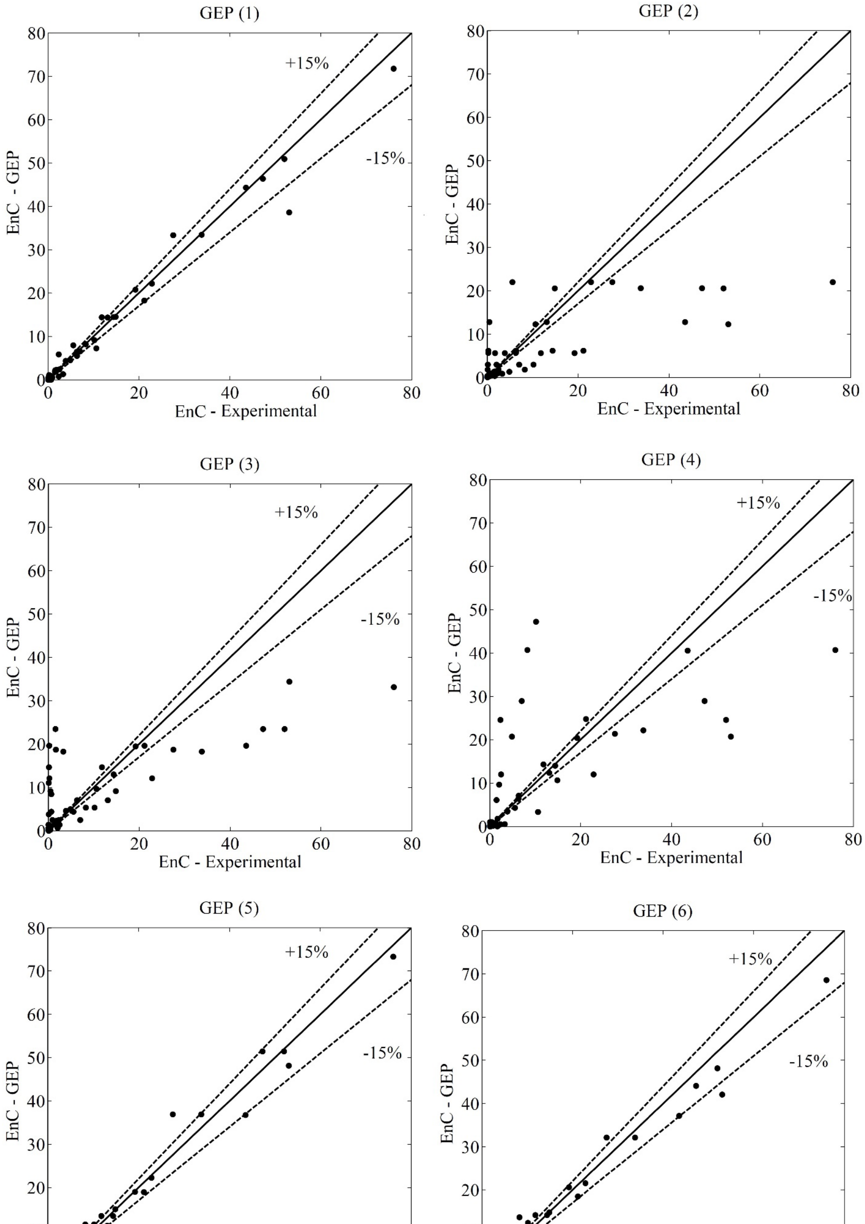

Fig. 2. GEP performance evaluation in predicting EnC.

a negative value for t_{EC} when t_{EC} = 0.5. Therefore, the pattern changes for t_{EC} do not lead to similar trends in CR% estimation. Due to the positive partial derivative value of Eq. (13) for variable C_S, changing this parameter has a direct relation with the CR% value obtained from Eq. (13). The partial derivative results regarding CR% compared with parameters pH_0 and C_0 are positive and negative. Therefore, by keeping all parameters fixed in this relationship and increasing or decreasing one of these two parameters, no clear trend in the results is observed. The maximum partial derivative value is related to the result obtained for pH_0. Therefore, it is concluded that Eq. (13) displays the most sensitivity to this parameter, and changing this parameter leads to significant changes in CR% results.

Fig. 4 shows the partial derivative results of Eq. (14) for the respective independent parameters provided to estimate the EnC parameter. The results in this figure represent an inverse relationship of the C_S variable with EnC, where by increasing this parameter value leads to a decreasing target parameter (EnC) value. The partial derivative value of Eq. (14) for C_S has the highest value among the other parameters, which indicates the importance of this parameter in EnC

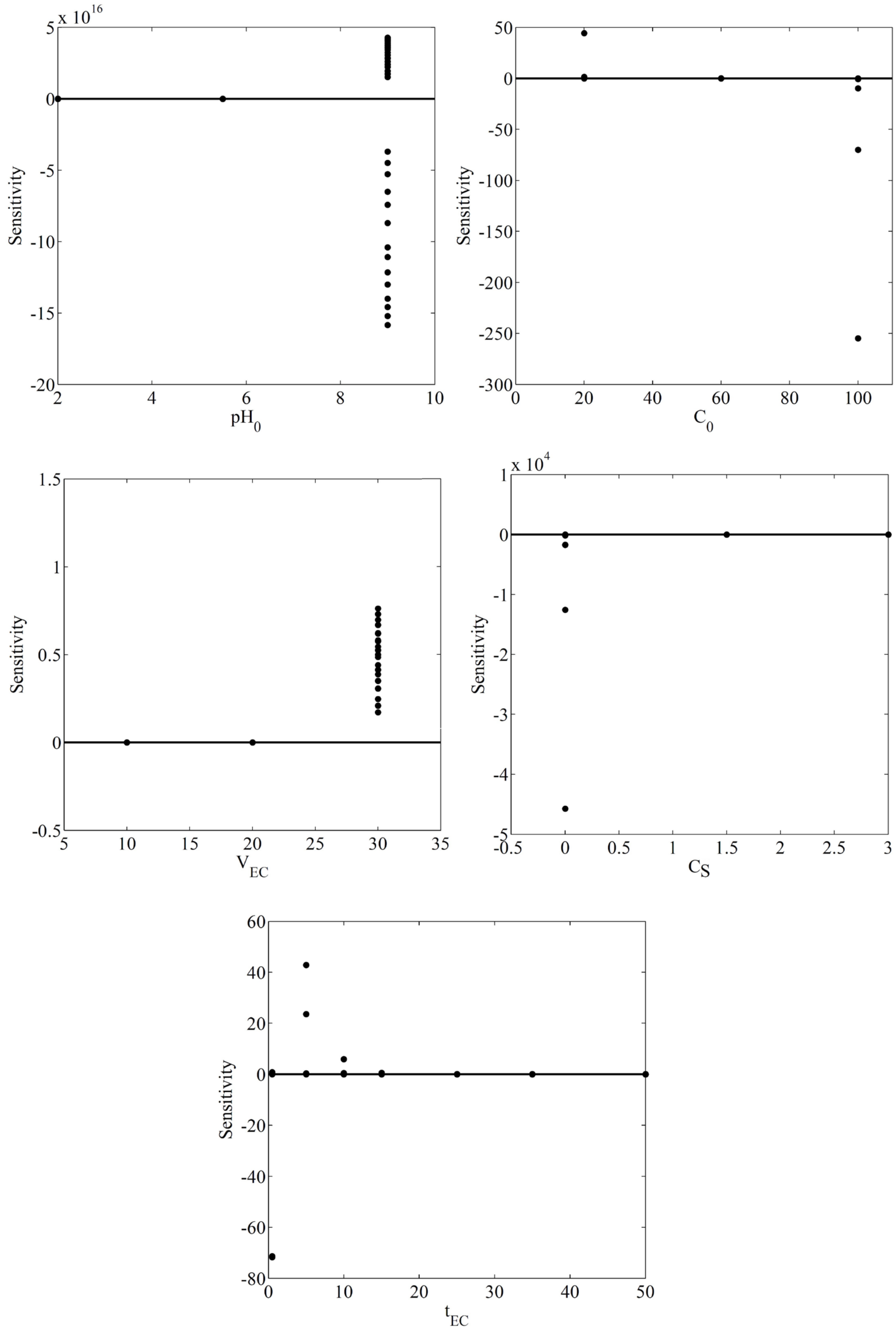

Fig. 3. Sensitivity analysis of input parameters to GEP (CR%).

estimation using Eq. (14). Unlike C_S, C_0 has a direct relation with EnC, since by increasing this parameter, EnC increases, but clearly, the values presented for C_0 and Cs are negligible regardless of sign. The partial derivative of Eq. (14) for parameters V_{EC} and t_{EC} does not have a constant trend, as the trend results of EnC obtained from Eq. (14) are not significant due to the increase in these two parameters.

The GEP model results were compared with the SMLR results for CR% and EnC prediction presented by Maleki et al. [18] and are plotted in Fig. 5. As seen in this figure, the CR% results are mostly close to the exact line for the GEP model, while the SMLR made overestimations and underestimations with over 15% relative error and MARE = 122.09. Unlike SMLR, the GEP model predicted most of the CR% with

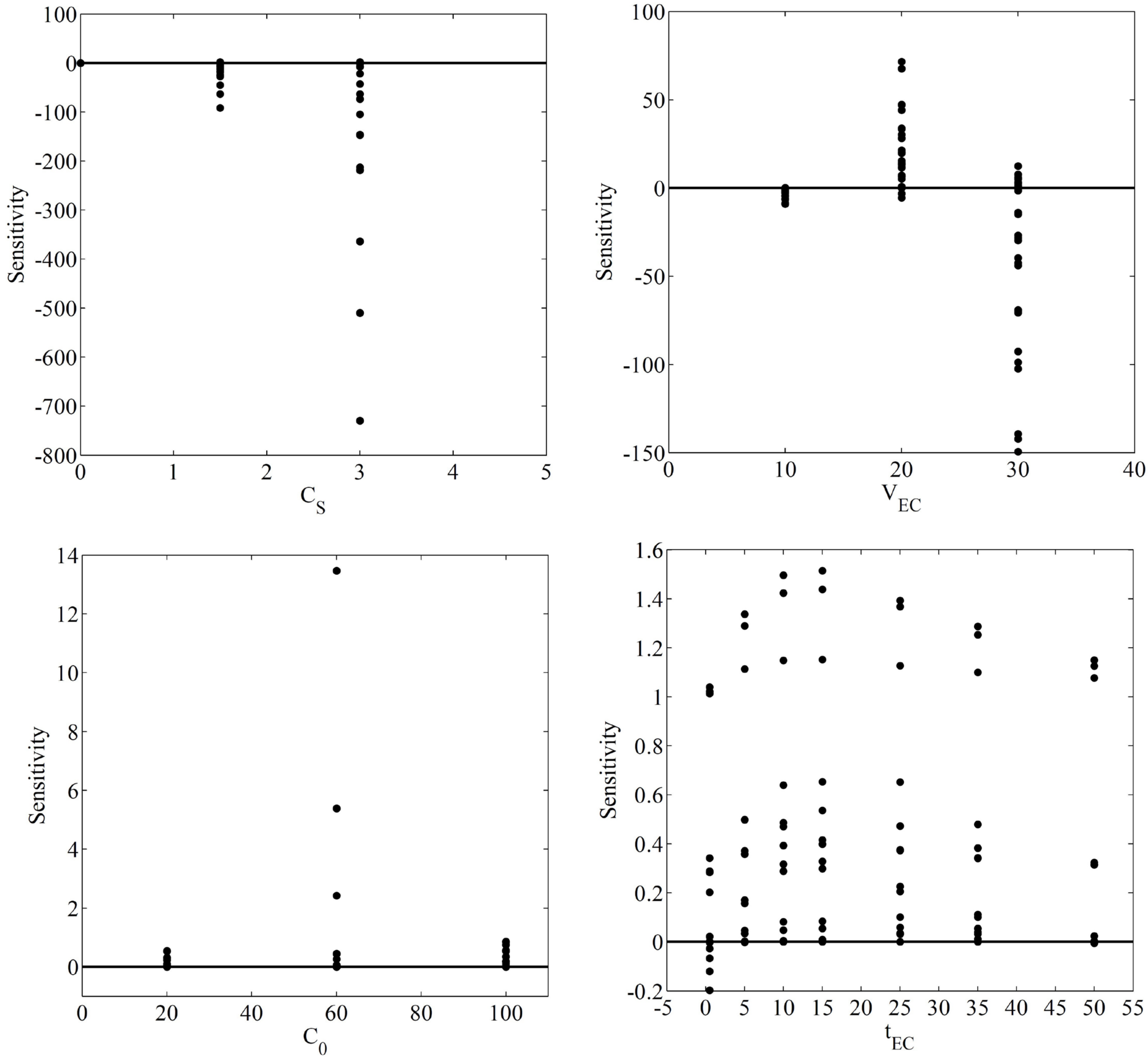

Fig. 4. Sensitivity analysis of input parameters to GEP (EnC).

less than 15% relative error and MARE of 48.97. The EnC values for the GEP and SMLR methods were overestimated and underestimated with 15% relative error. As a result, according to Fig. 5 and Tables 3 and 4, the GEP model outperformed SMLR in CR% and EnC prediction. In line with the explanations given, the equations proposed in this study outperform the equations suggested in previous studies. In addition, a comparison between the GEP and SMLR models by AIC demonstrates the superior performance of GEP in predicting the CR% (AIC (GEP) = 102.92; AIC (SMLR) = 173.78) and EnC (AIC (GEP) = 68.63; AIC (SMLR) = 95.58).

4. Conclusions

In this study, GEP and SMLR soft computing models were compared in terms of predicting the CR% and EnC of EC. The effect of different operational parameters on the CR% and EnC of EC of synthetic wastewater containing disperses like orange 25 dye was surveyed. The estimations by both methods proposed in this study are thoroughly examined. The SMLR method demonstrates weaker performance in CR% prediction (MARE = 122.09, RMSE = 24.13, SI = 0.35, BIAS = 0.23) compared with GEP, especially models 1 and 4 (MARE = 17.28 and 19.43; RMSE = 6.27 and 6.24; SI = 0.09 and 0.09; BIAS = 1.28 and 1.10, respectively). According to parameters C_0, V_{EC}, C_S and t_{EC} that were used as input combinations to predict EnC, the proposed SMLR method did not perform well in estimating models (MARE = 658.54, RMSE = 5.33). Hence, this method exhibited lower accuracy than GEP Models. Compared with SMLR, the GEP technique can provide simpler and more efficient solutions. To analyze the pattern changes in the proposed relationships according to

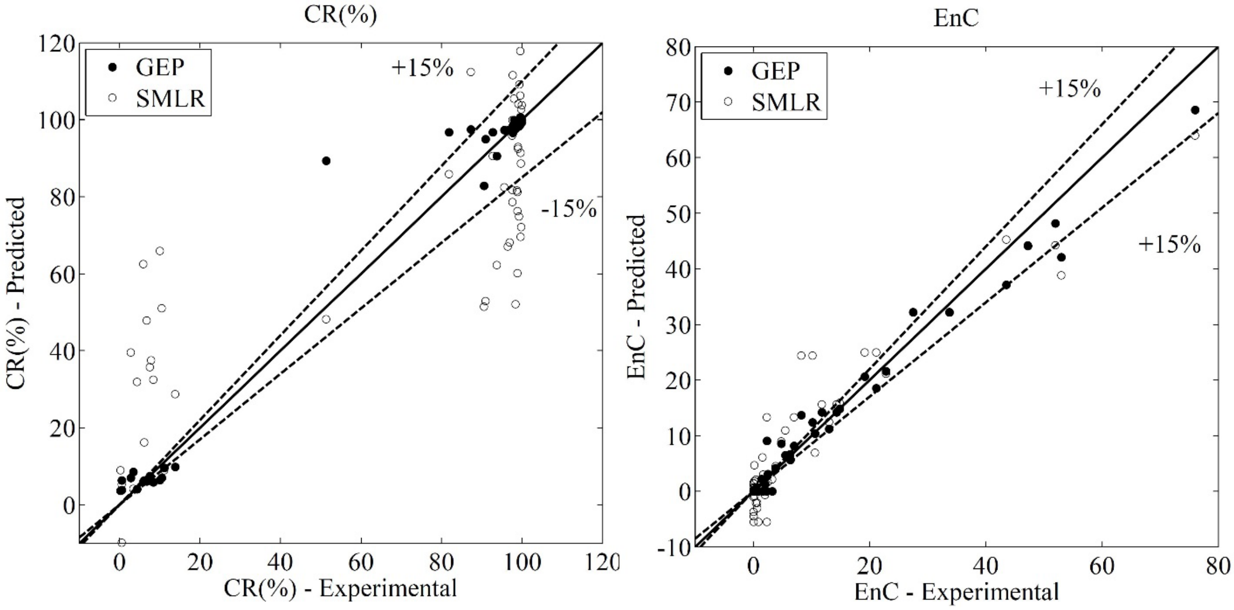

Fig. 5. Comparison of GEP and SMLR in predicting CR% and EnC.

the input parameters considered for each, partial derivative sensitivity analysis was applied. The maximum partial derivative value was related to the result obtained for pH_0. Therefore, it was concluded that Eq. (13) displays the most sensitivity to this parameter, and changing this parameter leads to significant changes in CR% results. The comparison of the GEP model with the SMLR results for CR% and EnC prediction showed that GEP model outperformed SMLR in CR% and EnC prediction. The overall results support the use of GEP as an alternative to more conventional methods of estimating the CR% and EnC of EC. It is evident from the result evaluation that the data predicted by GEP matched the experimental data with high overall accuracy, with a correlation coefficient (R^2) of 0.98 and RMSE within acceptable margins. Additionally, the GEP model results were compared with SMLR model results, and it was found that GEP performed better than SMLR in making predictions, suggesting the inherent sensitivity and robustness of the model. The current study results indicate that the proposed method can be used as an alternative in practical applications.

References

[1] 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.

[2] L. Du, Y. Yang, G. Li, S. Wang, X.M. Jia, Y. Zhao, Optimization of heavy metal containing dye Acid Black 172 decolorization by *Pseudomonas* sp DY1 using statistical designs, Int. Biodeterior. Biodegrad., 64 (2010) 566–573.

[3] N. Mohan, N. Balasubramanian, C.A. Basha, Electrochemical oxidation of textile wastewater and its reuse, J. Hazard. Mater., 147 (2007) 644–651.

[4] Z. Zaroual, M. Azzi, N. Saib, E. Chainet, Contribution to the study of electrocoagulation mechanism in basic textile effluent, J. Hazard. Mater., 131 (2006) 73–78.

[5] O.T. Can, M. Kobya, E. Demirbas, M. Bayramoglu, Treatment of the textile wastewater by combined electrocoagulation, Chemosphere, 62 (2006) 181–187.

[6] N. Daneshvar, A. Oladegaragoze, N. Djafarzadeh, Decolorization of basic dye solutions by electrocoagulation: an investigation of the effect of operational parameters, J. Hazard. Mater., 129 (2006) 116–122.

[7] A.K. Sharma, A.K.J. Chopra, Influence of operating conditions on the electrolytic treatment for the removal of color, TSS, hardness and alkalinity using Al-Al electrode combination, J. Appl. Nat. Sci., 6 (2014) 279–285.

[8] E. Bazrafshan, A.H. Mahvi, M.A. Zazouli, Textile wastewater treatment by electrocoagulation process using aluminum electrodes, Iran. J. Health Sci., 2 (2014) 16–29.

[9] K. Yetilmezsoy, F. Ilhan, Z. Sapci-Zengin, S. Sakar, M.T.J. Gonullu, Decolorization and COD reduction of UASB pretreated poultry manure wastewater by electrocoagulation process: a post-treatment study, J. Hazard. Mater., 162 (2009) 120–132.

[10] F. Ilhan, U. Kurt, O. Apaydin, M.T. Gonullu, Treatment of leachate by electrocoagulation using aluminum and iron electrodes, J. Hazard. Mater., 154 (2008) 381–389.

[11] Y. Avsar, U. Kurt, T. Gonullu, Comparison of classical chemical and electrochemical processes for treating rose processing wastewater, J. Hazard. Mater., 148 (2007) 340–345.

[12] B.Y. Tak, B.S. Tak, Y.J. Kim, Y.J. Park, Y.H. Yoon, G.H. Min, Optimization of color and COD removal from livestock wastewater by electrocoagulation process: application of Box–Behnken design (BBD), J. Ind. Eng. Chem., 28 (2015) 307–315.

[13] C. Ferreira, Gene Expression Programming in Problem Solving, 6th Online World Conference on Soft Computing in Industrial Applications, The University of Arizona, Tucson, AZ, USA, 2001.

[14] C. Ferreira, Gene expression programming: a new adaptive algorithm for solving problems, Complex Syst., 13 (2001) 87–129.

[15] A. Nazari, F. Pacheco Torgal, Modeling the compressive strength of geopolymeric binders by gene expression programming-GEP, Expert Syst. Appl., 40 (2013) 5427–5438.

[16] I. Ebtehaj, H. Bonakdari, Comparison of genetic algorithm and imperialist competitive algorithms in predicting bed load transport in clean pipe, Water Sci. Technol., 70 (2014) 759–766.

[17] I. Ebtehaj, H. Bonakdari, A.H. Zaji, H. Azimi, A. Sharifi, Gene expression programming to predict the discharge coefficient in rectangular side weirs, Appl. Soft Comput., 35 (2015) 1695–1701.

[18] A. Maleki, H. Daraei, B. Shahmoradi, S. Razee, N. Ghobadi, Electrocoagulation efficiency and energy consumption probing by artificial intelligent approaches, Desal. Wat. Treat., 52 (2014) 2400–2411.

[19] JEdit, 2011. Available from: http://sourceforge.net/projects/jedit
[20] T.H. Kim, C. Park, E.B. Shin, S. Kim, Decholorization of disperse and reactive dyes by continuous electro coagulation process, Desalination, 150 (2002) 165–175.
[21] G.J. McLachlan, K.A. Do, C. Ambroise, Analyzing Microarray Gene Expression Data, John Wiley & Sons, Inc., New York, USA, 2004.
[22] I. Ebtehaj, H. Bonakdari, A. Sharifi, Design criteria for sediment transport in sewers based on self-cleansing concept, J. Zhejiang Univ. Sci. A, 15 (2014) 914–924.
[23] I. Ebtehaj, H. Bonakdari, A.H. Zaji, H. Azimi, F. Khoshbin, GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs, Eng. Sci. Technol. Int. J., 18 (2015) 746–757.