Multi-objective evolutionary polynomial regression-based prediction of energy consumption probing

Hossein Bonakdari, Isa Ebtehaj and Azam Akhbari

ABSTRACT

Electrocoagulation (EC) is employed to investigate the energy consumption (EnC) of synthetic wastewater. In order to find the best process conditions, the influence of various parameters including initial pH, initial dye concentration, applied voltage, initial electrolyte concentration, and treatment time are investigated in this study. EnC is considered the main criterion of process evaluation in investigating the effect of the independent variables on the EC process and determining the optimum condition. Evolutionary polynomial regression is combined with a multi-objective genetic algorithm (EPR-MOGA) to present a new, simple and accurate equation for estimating EnC to overcome existing method weaknesses. To survey the influence of the effective variables, six different input combinations are considered. According to the results, EPR-MOGA Model 1 is the most accurate compared to other models, as it has the lowest error indices in predicting *EnC* ($MARE = 0.35$, $RMSE = 2.33$, $SI = 0.23$ and $R^2 = 0.98$). A comparison of EPR-MOGA with reduced quadratic multiple regression methods in terms of feasibility confirms that EPR-MOGA is an effective alternative method. Moreover, the partial derivative sensitivity analysis method is employed to analyze the *EnC* variation trend according to input variables.

Key words | dye removal, electrocoagulation, energy consumption, multiple regression

Hossein Bonakdari (corresponding author)
Isa Ebtehaj
Department of Civil Engineering,
Razi University,
Kermanshah,
Iran
and
Water and Wastewater Research Center,
Razi University,
Kermanshah,
Iran
E-mail: *bonakdari@yahoo.com*

Azam Akhbari
Department of Civil Engineering, Faculty of Engineering,
University of Malaya,
Kuala Lumpur,
Malaysia

INTRODUCTION

The large quantity of aqueous waste generated by dyeing industries has become a significant environmental problem. Several industries, such as leather, textile, dyestuffs, tanning, plastics, rubber, and cosmetics, use huge amounts of synthetic dyes and pigments (Baldikova *et al.* 2015). The disposal of colored textile wastewater into the environment without efficient treatment poses a threat of serious damage to aquatic life. Dye wastewater can be treated by biological means, chemical coagulation, activated carbon adsorption, ultrafiltration, ozonation, and electrocoagulation (EC) with electroflotation (Do & Chen 1994; Ghanbari *et al.* 1997; Ghanbari & Moradi 2015).

One of the methods developed to overcome the drawbacks of conventional water and wastewater treatment technologies is EC. EC is a potentially effective method for treating different kinds of wastewater with high removal efficiency (Golder *et al.* 2007). The EC process is a simple, reliable and cost-effective method of treating wastewater without the need for any additional chemicals and consequently, without causing secondary pollution. It also reduces the amount of sludge that requires disposal. The EC process has successfully been utilized for the treatment of wastewater, including electroplating wastewater (Ge *et al.* 2004; Escobar *et al.* 2006), laundry wastewater (Drouiche *et al.* 2009), and photovoltaic wastewater (Daneshvar *et al.* 2006). Meanwhile, the EC process has been widely used to decolorize various textile wastewaters (Giustolisi & Savic 2006; Kabdasli *et al.* 2009; Parama Kalyani *et al.* 2009).

Pollution parameters can be effectively removed via EC provided the operating conditions are carefully optimized and feasibility studies are conducted. Considering that the main operating costs associated with the EC treatment process are electrical energy requirements and sludge handling, it is very important to assess these two economic parameters before proposing EC as an alternative treatment option for dyehouse effluents. Until now, EC electrical energy requirements have been addressed in only a few studies (Arslan-Alaton *et al.* 2008). Investigating the influence of operating parameters is at the center of attention; hence the role of these parameters should be considered. Several researchers

doi: 10.2166/wst.2017.158

have investigated the effect of pH on the EC process using various types of wastewater. By taking into account the necessity for pH adjustment for usually extremely alkaline dyehouse effluents, it seems to make more sense to apply the EC process to high-strength, low-volume and slightly acidic effluents from polyamide dyeing operations with acid dyes and their respective dye auxiliaries (wetting agents, pH buffers). It has been established (Gao *et al.* 2005) that pH has a considerable effect on the efficiency of the EC process. Moreover, as observed by other investigators, the pH of the medium changes during the process. This change depends on the type of electrode material and initial pH and alkalinity. Also, it is known that electrical current density is the major operating variable directly affecting EC performance and operating costs. It is found that by increasing the applied current from appreciably improved decolorization, the addition of an electrolyte (Na_2SO_4 or NaCl) can enhance wastewater conductivity, decrease cell voltage and reduce electrical energy as well. In general, electrolytes are used to obtain conductivity in the EC process. Solution conductivity affects current efficiency, cell voltage and electrical energy consumption (EnC) in electrolytic cells. Also, current density is the most important parameter in the entire EC process. The current density determines not only the coagulant dosage rate but also the bubble production rate and size, and the floc growth, and it strongly influences the treatment efficiency of the EC process. The EnC increases with increasing current density due to an increase in ion production on the anode and cathode (Edris *et al.* 2008; Sridhar *et al.* 2014).

The operating cost is a very important economic parameter in the EC process. Operating costs involve the costs of chemicals, electrodes and EnC as well as labor, maintenance, sludge disposal and fixed costs. Energy, electrode and chemical costs are taken into account as major cost items in calculating operating costs (Bayramoglu *et al.* 2004; Sridhar *et al.* 2014). Therefore, a technically effective process must be economically feasible with regard to its electrical EnC and practical applicability to environmental problems. Therefore, EnC under optimum conditions is of great importance.

Evolutionary polynomial regression (EPR) is a data-driven hybrid technique that benefits from a combination of the effectiveness of genetic programming and numerical regression for developing simple and easily interpretable mathematical model expressions (Fonseca & Fleming 1993). The EPR approach overcomes some drawbacks of other modeling approaches, such as physical-based models and black-box data-driven models. The former can be difficult to construct due to underlying mechanisms that may not always be fully understood or the need for much data that is sometimes difficult to measure in the field. The latter, for example artificial neural networks, are very effective in reproducing any database related to a certain observed phenomenon but also carry overwhelming problems, like model structure identification, overfitting to training data, and the inability to exploit physical insight about the phenomenon at stake. EPR can overcome these problems by means of an explicit model expression for the system under observation (Deb *et al.* 2002; Maleki *et al.* 2014). Real-world optimization problems usually involve several conflicting objectives. The main aim in multi-objective optimization is to select the best tradeoffs among the conflicting objectives. Multi-objective evolutionary algorithms are regarded as promising methods for dealing with multi-objective optimization problems (MOPs) owing to their ability to generate populations of solutions for efficiently approximating a diverse set of optimal solutions in a single run (Deb *et al.* 2002). Therefore, developing an efficient optimization method to deal with expensive simulation-based MOPs by providing good accuracy and fewer function evaluations simultaneously is a crucial contemporary challenge.

The main goal of this study is to investigate the evolutionary polynomial regression–multi-objective genetic algorithm (EPR-MOGA) method in predicting EnC probing. For this purpose, the effective variables for predicting EnC probing are identified. Subsequently, to investigate the effect of each variable, six different input combinations are examined. All input combinations are modeled with EPR-MOGA and a simple equation is presented for each one. After selecting the best input combination, the results of EPR-MOGA are compared with reduced quadratic multiple regression. A partial derivative sensitivity analysis (PDSA) is also done to analyze the EnC variation trend of the proposed equation.

EXPERIMENTAL DATA

To estimate the EnC of EC, experimental data of Maleki *et al.* (2014) were utilized. Maleki *et al.* (2014) conducted experiments in an EC system consisting of a cubic ($12 \times 12 \times 21$ cm) glass reactor, 400 rpm mixer, DC power supply (high stability and reliability, and low-noise DC adjustable power supply RXN-303D-II, Zhaoxin Electronic Tech. Co.), and two aluminum electrodes (see Maleki *et al.* (2014) for more details). Table 1 represents the range of parameters that affect EnC in the Maleki *et al.* study.

Table 1 | Range of data in Maleki *et al.* (2014) study

Parameters	Range
Initial pH (pH_0)	2–9
Initial dye concentration (C_0), mg/l	8–100
Applied voltage (V_{EC}), V	10–30
Initial electrolyte concentration (C_S), mg/l	0–3
Treatment time (t_{EC}), min	0.5–50
Energy consumption (EnC), Wh	0.001–76

To present a new model for *EnC* prediction, 173 samples (Maleki *et al.* 2014) are used in this study. The samples regard different values of input parameters, namely initial pH, initial dye concentration, applied voltage, initial electrolyte concentration and treatment time.

The EPR-MOGA results presented in this paper are compared with the stepwise multiple linear regression (SMLR) results presented by Maleki *et al.* (2014). The comparison indicates that EPR-MOGA outperforms SMLR. The weakness of other artificial intelligence-based methods is the lack of a specific equation for use in future studies, whereas EPR-MOGA yields an expression. Therefore, the proposed method is more accurate than other artificial intelligence methods such as artificial neural networks that do not present specific equations (e.g. multiple linear regression), and can thus be used for future studies.

Moreover, the number of data in artificial intelligence methods such as EPR-MOGA is different and a good model should contain less than 100 samples (e.g. Rezania *et al.* 2008; Ahangar-Asr *et al.* 2011; Fiore *et al.* 2012). Therefore, 173 samples are sufficient to find an appropriate model in this study. These data were collected from Maleki *et al.* (2014).

EPR

EPR is a data-driven hybrid technique, which combines the effectiveness of genetic programming with numerical regression to develop simple and easily interpretable mathematical model expressions (Giustolisi & Savic 2006). The EPR approach overcomes some drawbacks of other modeling approaches, such as physical-based models and black-box data-driven models. The former can be difficult to construct due to underlying mechanisms that may not always be fully understood, or owing to the need for much data that is sometimes difficult to measure in the field. The latter, for example artificial neural networks, are very effective in reproducing any database related to some observed phenomenon but bring some overwhelming problems like model structure identification, overfitting to training data, and the inability to exploit physical insight about the phenomenon at stake. EPR can overcome these problems by means of an explicit model expression for the system under observation. EPR-MOGA can be defined as a nonlinear global stepwise regression for symbolic data modeling.

These models can then be verified on a test set and gaps can be filled in test datasets by using one selected model. Because of the pseudo-polynomial formulations achievable by EPR-MOGA, it requires fewer parameters to be estimated, which in turn requires shorter time series for training. Another advantage of the EPR-MOGA approach is the ability to choose objective functions pertaining to accuracy and parsimony (Barca *et al.* 2015).

The expressions obtained with the EPR method generally contain different terms that are multiplied by different coefficients (e.g. polynomials) as follows:

$$\widehat{Y} = a_0 + \sum_{j=1}^{m} (X_1)^{\mathbf{ES}(j,1)} \cdot \ldots \cdot (X_k)^{\mathbf{ES}(i,1)} \cdot f(((X_1)^{\mathbf{ES}(j,1)} \cdot \ldots \cdot (X_1)^{\mathbf{ES}(j,k)})) \quad (1)$$

where m, X_i and $\widehat{Y}$ are the maximum amounts of additive terms, input variables, and output variables; the f function is determined by the user; and $\mathbf{ES}((i,j))$ are chosen from a set of ***EX*** applicants provided by the user. A genetic algorithm is used to select the $\mathbf{ES}(j,i)$ exponents among the values presented in ***EX***. In fact, an integer coding of possible exponents $\mathbf{ES}(j,i)$ is assumed to achieve nonlinear relationships. If $\mathbf{ES}(j,i) = 0$ and the set of exponents includes zero, the related input is deleted from the final expression. Therefore, versatile and flexible structures can be expressed to generate patterns in data similar to Equation (1) (Giustolisi & Savic 2006).

In the development strategy of the model proposed by EPR, the final expression is linear according to the a_j coefficient. Parameter estimation is solved as an inverse linear problem to guarantee a unique relationship between each model parameter and its structure. In terms of numerical regression strategy, EPR, similar to the artificial neural network method, makes nonlinear mapping between data although the parameters are estimated with some constants and nonlinear regression (Haykin 1999). These features prevent overfitting, especially when the data number is low. In addition, the initial assumption in the mathematical structure is that the number of parameters and functions can serve as the initial hypothesis for the model structure. More details on the EPR model were presented by Giustolisi & Savic (2006). The most important advance in the basic EPR algorithm related to the multi-objective optimization strategy is achieved using genetic algorithms (EPR-MOGA), whereby model generation accuracy and feasible structures in the presented models are optimized simultaneously (Laucelli & Giustolisi 2011).

Maximizing the parsimony of the presented formulas leads to an easy physical concept of the final proposed expression and a comprehensive description of the phenomenon. The user determines the exploration space in the EPR-MOGA technique in terms of a basic mathematical expression structure, the maximum number of additive terms (m), exponent candidate $\boldsymbol{EX}$ series and the number of candidate explanatory variables (k). The model is probed using optimized MOGA (Giustolisi & Savic 2009), which is based on the Pareto dominance criteria (Van Veldhuizen & Lamont 2000). The EPR-MOGA technique is used to probe the space of m formula components using two or three objects: (i) model accuracy maximization, (ii) additive term amount minimization, and (iii) the minimized amount of actually utilized model inputs. The two latter objects indicate savings in the model. Finally, EPR-MOGA provides a set of optimal results (i.e. Pareto front) that consider the tradeoff between model complexity and accuracy.

ENC MODELING

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 EnC estimation, a functional relationship is provided as follows:

$$EnC = f(pH_0, C_0, V_{EC}, C_S, t_{EC}) \tag{2}$$

The methodology applied in this study entails dividing the dataset into two groups: training and testing. From the combinations employed in this study, Equation (1) was selected and modeled using EPR-MOGA. Then to examine the effect of each parameter, different models were proposed as follows:

EPR-MOGA (1): $EnC = f(pH_0, C_0, V_{EC}, C_S, t_{EC})$
EPR-MOGA (2): $EnC = f(pH_0, C_0, V_{EC}, C_S)$
EPR-MOGA (3): $EnC = f(pH_0, C_0, V_{EC}, t_{EC})$
EPR-MOGA (4): $EnC = f(pH_0, C_0, C_S, t_{EC})$
EPR-MOGA (5): $EnC = f(pH_0, V_{EC}, C_S, t_{EC})$
EPR-MOGA (6): $EnC = f(C_0, V_{EC}, C_S, t_{EC})$

By developing EPR-MOGA, various equations are proposed for all six input combinations as presented in Table 2.

Maleki et al. (2014) used the SMLR methodology for EnC after neglecting the non-significant terms (Bhatti et al.

Table 2 | Equations obtained by EPR-MOGA for calculating *EnC*

Model no.	Equations	Eq. no.
1	$EnC = 8.5156 \times \left(\frac{V_{EC}t_{EC}}{pH_0^2 C_0^2}\right) + 5.5463E-4 \times V_{ES}^2 C_S^{0.5} t_{EC} + 4.1254E-5 \times C_0^{0.5} V_{EC}^2 C_S^{0.5} t_{EC}$	(3)
2	$EnC = 1.0734 \times \left(\frac{V_{EC}^{1.5}}{pH_0^{1.5} C_0}\right) + 4.7369E-2 \times V_{EC}^{1.5} C_S^{0.5} + 2.2147E-3 \times pH_0^{0.5} V_{EC}^2 C_S^{0.5}$	(4)
3	$EnC = 0.24059 \times \left(\frac{V_{EC}t_{EC}^2}{pH_0^2 C_0^2}\right) + 1.415E-2 \times V_{EC}^{1.5} t_{EC}^{1.5} 2.292E-3 \times pH_0^{0.5} V_{EC} t_{EC}$	(5)
4	$EnC = 7.6889 \times \left(\frac{t_{EC}^2}{pH_0^2 C_0^2}\right) + 1.2121 \times C_S^{0.5} t_{EC}^{0.5} + 4.9741E-4 \times pH_0 \times C_S \times t_{EC}^4$	(6)
5	$EnC = 1.5837E-4 \times \left(\frac{V_{EC}^2 t_{EC}}{pH_0^{0.5}}\right) + 2.1034E-3 \times V_{EC}^{1.5} C_S t_{EC}^{0.5} + 6.9144E-4 \times V_{EC}^2 C_S^{0.5} t_{EC}$	(7)
6	$EnC = 3.9866E-4 \times V_{EC}^{1.5} t_{EC} + 5.9118E-4 \times V_{EC}^2 C_S^{0.5} t_{EC} + 2.0217E-5 \times C_0^{0.5} V_{EC}^2 C_S t_{EC}$	(8)

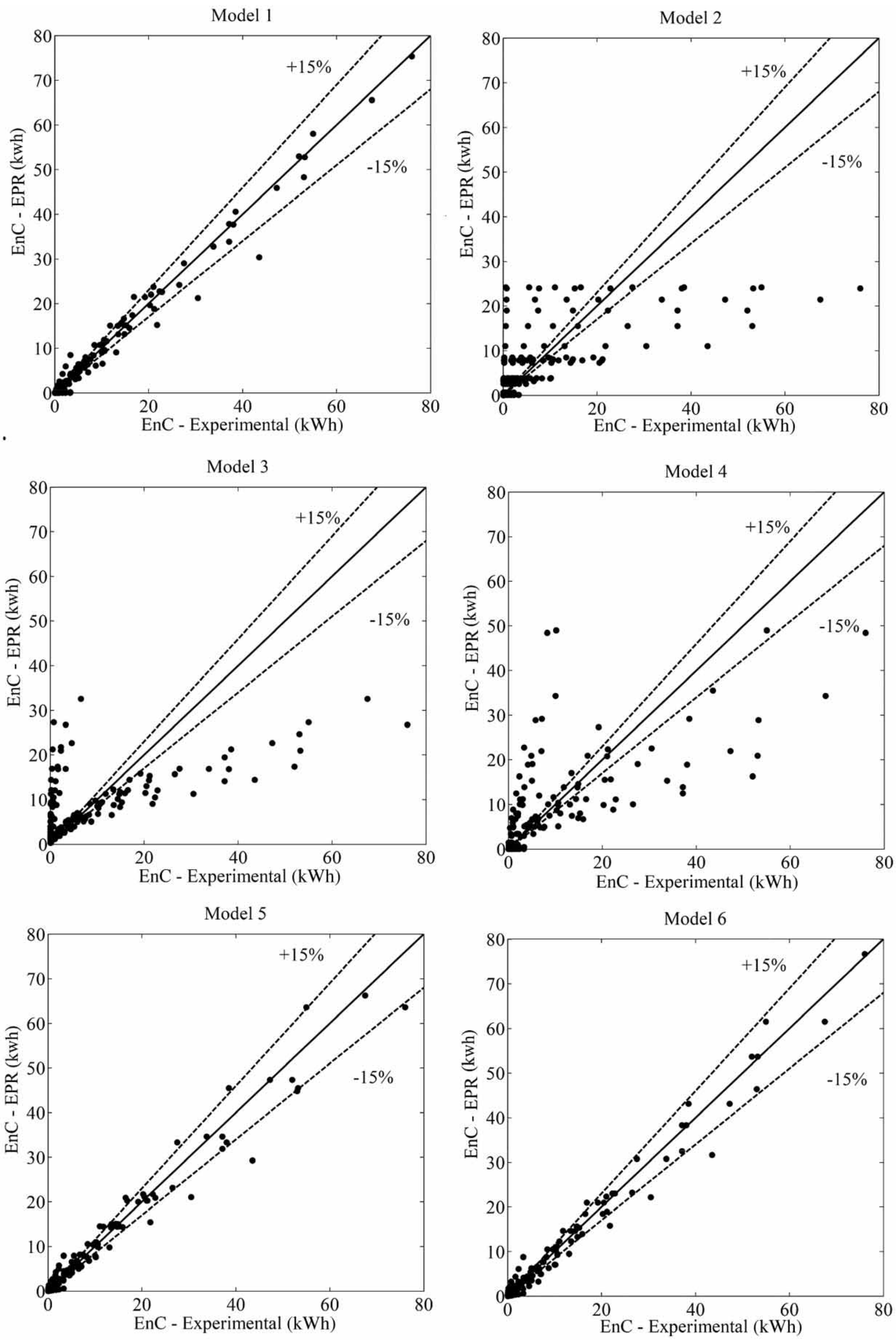

Figure 1 | EPR-MOGA performance evaluation in *EnC* prediction.

2011; Safa & Bhatti 2011). The model is given in Equation (9).

$$EnC = 5.7837 - 6.3669(C_S) - 0.7365(t_{EC}) + 0.3223(V_{EC})(C_S) + 0.0287(V_{EC})(t_{EC}) + 0.2689(C_S)(t_{EC}) - 0.0106(V_{EC})^2 + 0.0045(t_{EC})^2 \quad (9)$$

RESULTS AND DISCUSSION

In this section, the EnC results are predicted using the EPR-MOGA technique. For this purpose, the statistical indices utilized are the determination coefficient (R^2), root mean square error ($RMSE$), mean absolute relative error ($MARE$), scatter index (SI) and $BIAS$, which are calculated as follows:

$$R^2 = \left(\frac{\sum_{i=1}^{n} (O_i - \overline{O_i}).(P_i - \overline{P_i})}{\sqrt{\sum_{i=1}^{n} (O_i - \overline{O_i})^2 \sum_{i=1}^{n} (P_i - \overline{P_i})^2}} \right)^2 \quad (10)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (O_i - P_i)^2} \quad (11)$$

$$MARE = \frac{1}{n}\sum_{i=1}^{n} \left(\frac{|O_i - P_i|}{O_i} \right) \quad (12)$$

$$SI = \frac{RMSE}{\frac{1}{n}\sum_{i=1}^{n} (\bar{O}_i)} \quad (13)$$

$$BIAS = \frac{\sum_{i=1}^{n} (P_i - O_i)}{n} \quad (14)$$

where O_i and P_i are the observed and predicted EnC respectively, $\overline{O_i}$ and $\overline{P_i}$ are the average observed and predicted EnC values, and n is the number of parameters.

Figure 1 shows the performance of EPR-MOGA with six different input combinations (Model 1 to Model 6) in EnC prediction. Model 1 considers all effective parameters in estimating EnC, which are presented as Equation (3). Table 3 presents the statistical indices for EnC modeling by EPR-MOGA. According to Figure 1, EPR-MOGA is highly capable of estimating EnC, as many estimated values have little difference from the observed values ($R^2 = 0.98$; $MARE = 0.35$; $RMSE = 2.33$; $SI = 0.23$). It can be seen that this model often estimates with less than 15% relative error, which indicates good model accuracy.

Table 3 | Statistical indices for *EnC* modeling by EPR-MOGA

EPR-MOGA	R^2	MARE	RMSE	SI	BIAS
Train					
Model 1	0.98	0.41	1.62	0.23	– 0.10
Model 2	0.41	6.48	9.16	1.30	– 0.02
Model 3	0.40	23.47	9.30	1.32	0.37
Model 4	0.56	2.16	7.90	1.12	0.04
Model 5	0.97	0.85	2.08	0.29	0.07
Model 6	0.98	1.21	1.81	0.26	– 0.06
Test					
Model 1	0.98	0.35	2.33	0.23	– 0.44
Model 2	0.54	3.37	12.59	1.22	– 3.30
Model 3	0.59	14.65	12.41	1.20	– 1.80
Model 4	0.42	1.66	12.77	1.24	– 0.47
Model 5	0.97	0.88	3.23	0.31	– 0.55
Model 6	0.98	1.20	2.39	0.23	– 0.38
SMLR	0.90	658.54	5.33	0.52	0.11

Model 1 makes underestimations and overestimations, while on average and especially for large errors, the EnC predictions are underestimated ($BIAS = -0.44$). Not using t_{EC} as an input parameter in EPR-MOGA EnC modeling leads to significantly reduced modeling accuracy. The EnC prediction results by Model 2 indicate low accuracy according to the statistical indices ($R^2 = 0.54$; $MARE = 0.3.3.7$; $RMSE = 12.59$; $SI = 1.22$). Models 3 and 4 do not use C_S and V_{EC} and are weaker than Model 2. Therefore, t_{EC}, V_{EC} and C_S should be used in predicting EnC. The most effective variable is the initial electrolyte concentration, with the largest MARE in the test stage ($MARE = 14.65$).

Unlike t_{EC}, C_S and V_{EC}, C_0 and pH_0 have low impact on EnC prediction by EPR-MOGA, whereby not using C_0 and pH_0 increases the relative errors by approximately three and four times (respectively) with Model 1, while the $MARE$ values of Models 2, 3 and 4 are more than 10 times greater than for Model 1. However, Models 5 and 6 in which C_0 and pH_0 were removed from the input combinations (respectively) do not present an acceptable level of accuracy.

Therefore, EPR-MOGA Model 1 with the lowest error index among all models is the most accurate in predicting the EnC ($MARE = 0.41$ and 0.35, $RMSE = 1.62$ and 2.33, $SI = 0.23$ and 0.23 and $R^2 = 0.98$ and 0.98 in training and testing, respectively). It can be seen that Model 1 has the highest compliance with the experimental values and is slightly different from the experimental model.

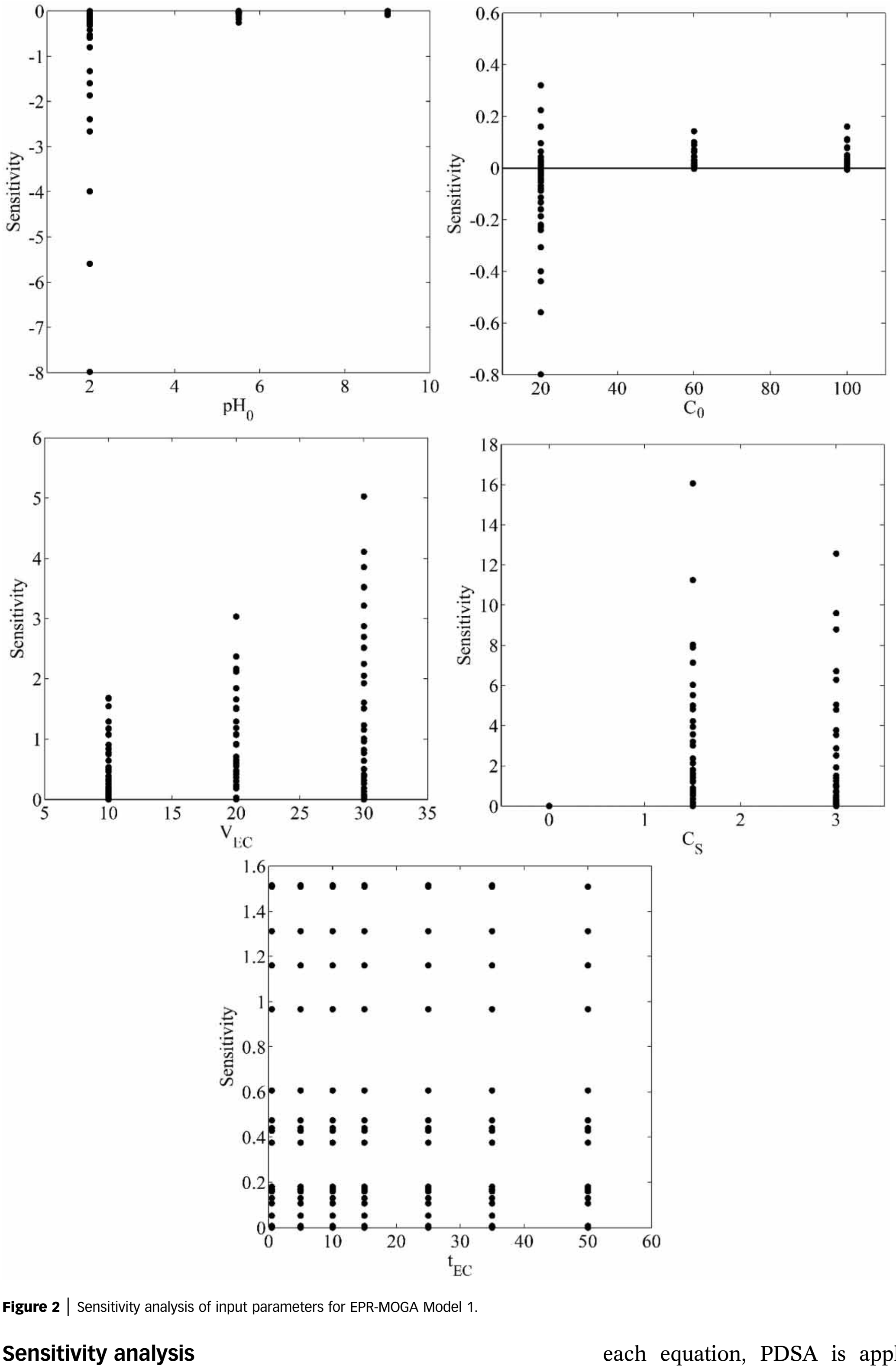

Figure 2 | Sensitivity analysis of input parameters for EPR-MOGA Model 1.

Sensitivity analysis

To analyze the trend changes in the proposed EPR-MOGA equation according to the input variables considered for each equation, PDSA is applied in this study (Ebtehaj *et al.* 2015). With this method, the partial derivative of the equation presented for each input variable is calculated. Then the trend changes are studied for different input variable values, from which the derivative of each equation has been calculated. It is clear that the extent of the partial derivative calculated is directly related to its effect on the predicted 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.

Figure 2 presents the partial derivative results of Equation (3) for the parameters provided in this relationship. The partial derivatives for the V_{EC}, t_{EC} and C_s parameter values are positive. The increase (or decrease) in these three parameters leads to an increase (or decrease) in the *EnC* estimated using Equation (3). It is clear that the derivative of Equation (3) has a negative pH_0 value. Therefore, the pH_0 trend changes do not result in similar *EnC* estimation trends. Due to the positive partial derivative value of Equation (3) for variables C_S, t_{EC} and V_{EC}, the change in these parameters is directly related to the *EnC* value attained from Equation (3). The partial derivative results concerning *EnC* compared to parameter 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 C_s. Therefore, it is concluded that Equation (3) displays the greatest sensitivity to this parameter, and changing this parameter leads to significant changes in *EnC* results.

The EPR-MOGA model results are compared with the SMLR results for *EnC* prediction presented by Maleki *et al.* (2014). The EPR-MOGA and SMLR model results are plotted in Figure 3. Here, the EPR-MOGA *EnC* results are mostly close to the exact line, while SMLR made over- and underestimations with over 15% relative error and

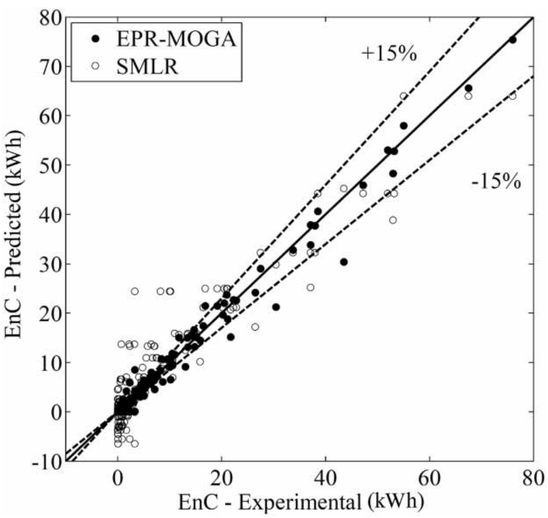

Figure 3 | Comparison of EPR-MOGA and SMLR in predicting EnC.

MARE of 658.54. Unlike SMLR, the EPR-MOGA model predicted most EnC values with less than 15% relative error and MARE of 0.41 and 0.35 in testing and training respectively. Consequently, according to Figure 3, the EPR-MOGA model outperformed SMLR in *EnC* prediction. In line with the explanations given, the equations proposed in this study perform better than the equations suggested in previous studies. In addition, a comparison between the EPR-MOGA and SMLR models demonstrates the superior performance of EPR-MOGA in predicting *EnC*.

CONCLUSION

In this study, multi-objective evolutionary polynomial regression was combined with a genetic algorithm (EPR-MOGA) for EnC and new model expressions were derived. The EPR-MOGA and SMLR soft computing models were compared in terms of predicting the *EnC* of EC. In order to find the best process condition, the influence of various parameters was investigated. It is evident from the results that the data predicted by EPR-MOGA are in good compliance with the experimental data. In training and testing modes, the *RMSE* values obtained for the superior model were 1.62 and 2.33 respectively, while the *SI* and *BIAS* values calculated were 0.23 and 0.23, and −0.1 and −0.44 respectively. The superior model's sensitivity to the input parameters was evaluated using PDSA. According to the sensitivity analysis results, with increasing C_S the sensitivity increased. Additionally, it was found that the EPR-MOGA model outperformed SMLR in making predictions, suggesting the inherent sensitivity and robustness of the model. A comparison of the EPR-MOGA and SMLR methods in terms of feasibility confirmed that EPR-MOGA is an effective alternative method.

REFERENCES

Ahangar-Asr, A., Faramarzi, A., Javadi, A. A. & Giustolisi, O. 2011 The modeling mechanical behavior of rubber concrete using evolutionary polynomial regression. *Engineering Computations* **28** (4), 492–507.

Arslan-Alaton, I., Kabdaslı, I. & Sahin, Y. 2008 Effect of operating parameters on the electrocoagulation of simulated acid dyebath effluent. *The Open Environmental & Biological Monitoring Journal* **1**, 1–7.

Baldikova, E., Safarikova, M. & Safarik, I. 2015 Organic dyes removal using magnetically modified rye straw. *Journal of Magnetism and Magnetic Materials* **380**, 181–185.

Barca, E., Berardi, L., Laucelli, D. B., Passarella, G. & Giustolisi, O. 2015 Evolutionary polynomial regression application for missing data handling in meteo-climatic gauging stations. *2015 European Regional Conference of The International Environmetrics Society, GRASPA 2015, June 15–16, Bari, Italy*.

Bayramoglu, M., Kobya, M., Can, O. T. & Sozbir, M. 2004 Operating cost analysis of electrocoagulation of textile dye wastewater. *Separation and Purification Technology* **37** (2), 117–125.

Bhatti, M. S., Kapoor, D., Kalia, R. K., Reddy, A. S. & Thukral, A. K. 2011 RSM and ANN modeling for electrocoagulation of copper from simulated wastewater: multi-objective optimization using genetic algorithm approach. *Desalination* **274** (1–3), 74–80.

Daneshvar, N., Oladegaragoze, A. & Djafarzadeh, N. 2006 Decolorization of basic dye solutions by electrocoagulation: an investigation of the effect of operational parameters. *Journal of Hazardous Materials* **129** (1–3), 116–122.

Deb, K., Pratap, A., Agarwal, S. & Meyarivan, T. 2002 A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transaction on Evolutionary Computation* **6** (2), 182–197.

Do, J. S. & Chen, M. L. 1994 Decolourization of dye-containing solutions by electrocoagulation. *Journal of Applied Electrochemistry* **24**, 785–790.

Drouiche, N., Aoudj, S., Hecini, M., Ghaffour, N., Lounici, H. & Mameri, N. 2009 Study on the treatment of photovoltaic wastewater using electrocoagulation: fluoride removal with aluminum electrodes–characteristics of products. *Journal of Hazardous Materials* **169** (1–3), 65–69.

Ebtehaj, I., Bonakdari, H., Zaji, A. H., Azimi, H. & Khoshbin, F. 2015 GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs. *Engineering Science and Technology* **18** (4), 746–757.

Edris, B., Amir Hossein, M., Simin, N. & Ali Reza, M. 2008 Performance evaluation of electrocoagulation process for removal of chromium (VI) from synthetic chromium solutions using iron and aluminum electrodes. *Turkish Journal of Engineering and Environmental Sciences* **32**, 59–66.

Escobar, C., Soto-Salazar, C. & Toral, M. 2006 Optimization of the electrocoagulation process for the removal of copper, lead, and cadmium in natural waters and simulated wastewater. *Journal of Environmental Management* **81** (4), 384–391.

Fiore, A., Berardi, L. & Marano, G. C. 2012 Predicting torsional strength of RC beams by using evolutionary polynomial regression. *Advances in Engineering Software* **47** (1), 178–187.

Fonseca, C. M. & Fleming, P. J. 1993 Genetic algorithms for multiobjective optimization: formulation, discussion, and generalization. In: *Proceedings of the Fifth International Conference on Genetic Algorithms*, University of Illinois, Urbana-Champaign, USA.

Gao, P., Chen, X., Shen, F. & Chen, G. 2005 Removal of chromium (VI) from wastewater by combined electrocoagulation–electroflotation without a filter. *Separation and Purification Technology* **43**, 117–123.

Ge, J., Qu, J., Lei, P. & Liu, H. 2004 New bipolar electrocoagulation-electroflotation process for the treatment of laundry wastewater. *Separation and Purification Technology* **36** (1), 33–39.

Ghanbari, F. & Moradi, M. 2015 A comparative study of electrocoagulation, electrochemical Fenton, electro-Fenton and peroxide-coagulation for decolorization of real textile wastewater: electrical energy consumption and biodegradability improvement. *Journal of Environmental Chemical Engineering* **3** (1), 499–506.

Ghanbari, F., Moradi, M. & Manshouri, M. 1997 Textile wastewater decolorization by zero valent iron activated peroxomonosulfate: compared with zero valent copper. *Journal of Environmental Chemical Engineering* **2** (3), 1846–1851.

Giustolisi, O. & Savic, D. A. 2006 A symbolic data-driven technique based on evolutionary polynomial regression. *Journal of Hydroinformatics* **8** (3), 207–222.

Giustolisi, O. & Savic, D. A. 2009 Advances in data-driven analyses and modeling using EPR-MOGA. *Journal of Hydroinformatics* **11** (3–4), 225–236.

Golder, A. K., Samanta, A. N. & Ray, S. 2007 Removal of Cr^{3+} by electrocoagulation with multiple electrodes: bipolar and monopolar configurations. *Journal of Hazardous Materials* **141** (3), 653–661.

Haykin, S. 1999 *Neural Networks: A Comprehensive Foundation*, 2nd edn. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Kabdasli, I., Vardar, B., Arslan-Alaton, I. & Tunay, O. 2009 Effect of dye auxiliaries on color and COD removal from simulated reactive dye bath effluent by electrocoagulation. *Chemical Engineering Journal* **148** (1), 89–96.

Laucelli, D. & Giustolisi, O. 2011 Scour depth modeling by a multi-objective evolutionary paradigm. *Environmental Modelling & Software* **26** (4), 498–509.

Maleki, A., Daraei, H., Shahmoradi, B., Razee, S. & Ghobadi, N. 2014 Electrocoagulation efficiency and energy consumption probing by artificial intelligent approaches. *Desalination and Water Treatment* **52** (13–15), 2400–2411.

Parama Kalyani, K. S., Balasubramanian, N. & Srinivasakannan, C. 2009 Decolorization and COD reduction of paper industrial effluent using electro-coagulation. *Chemical Engineering Journal* **151** (1–3), 97–104.

Rezania, M., Javadi, A. A. & Giustolisi, O. 2008 An evolutionary-based data mining technique for assessment of civil engineering systems. *Engineering Computations* **25** (6), 500–517.

Safa, Y. & Bhatti, H. N. 2011 Adsorptive removal of direct textile dyes by low-cost agricultural waste: application of factorial design analysis. *Chemical Engineering Journal* **167** (11), 35–41.

Sridhar, R., Sivakumar, V. & Prakash Maran, J. 2014 Influence of operating parameters on the treatment of egg processing effluent by electrocoagulation process. *International Journal of Environmental Science and Technology* **11**, 1619–1630.

Van Veldhuizen, D. A. & Lamont, G. B. 2000 Multiobjective evolutionary algorithms analyzing the state-of-the-art. *Evolutionary Computation* **8** (2), 125–144.

First received 21 July 2016; accepted in revised form 2 March 2017. Available online 16 March 2017