Journal of Environmental Chemical Engineering 2 (2014) 1672–1683

Contents lists available at ScienceDirect

Journal of Environmental Chemical Engineering

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

Assessment of competitive dye removal using a reliable method

Jafar Abdi [a], Dariush Bastani [a], Javad Abdi [b], Niyaz Mohammad Mahmoodi [c], Amin Shokrollahi [a], Amir H. Mohammadi [d,e,*]

[a] Department of Chemical and Petroleum Engineering, Sharif University of Technology, Tehran, Iran
[b] Department of Chemical Engineering, Sahand University of Technology, Tabriz, Iran
[c] Department of Environmental Research, Institute for Color Science and Technology, Tehran, Iran
[d] Institut de Recherche en Génie Chimique et Pétrolier (IRGCP), Paris Cedex, France
[e] Thermodynamics Research Unit, School of Engineering, University of KwaZulu-Natal, Howard College Campus, King George V Avenue, Durban 4041, South Africa

ARTICLE INFO

Article history:
Received 16 April 2014
Accepted 4 June 2014
Available online 12 June 2014

Keywords:
Dye removal
Ternary system
Support vector machine
Leverage approach
Selectivity analysis
Competitive adsorption

ABSTRACT

In this study, a reliable and predictive model namely, least-squares support vector machine (LS-SVM) was developed to predict dye removal efficiency. Four LS-SVM models have been developed and tested using more than 630 series of experimental data which were obtained from our previous paper. These data consist of adsorbate type, adsorbent dosage, initial dye concentration, salt, absorbance time and dye removal efficiency. Direct Red 31 (DR31), Direct Green 6 (DG6) and Acid Blue (AB92) were used as a model dyes. The results show that the developed model is more accurate and reliable with the average absolute relative deviation of 0.678%, 0.877%, 0.581% and 0.978% for single systems and ternary system, respectively and correlation coefficients close to unity for all systems. Additionally, it is demonstrated that the proposed method is capable of simulating the actual physical trend of the dye removal efficiency with variation of adsorbent dosage and initial dye concentration in single and ternary systems. Eventually, the Leverage approach, in which the statistical Hat matrix, Williams plot, and the residuals of the model results lead to identification of the likely outliers, has been carried out. Fortunately, all the experimental data seem to be reliable except five in single systems. Therefore, the developed model could be reliable for prediction of the dye removal efficiency in its applicability domain. Selectivity analysis showed that GPN had selective removal of DG6.

Introduction

Great number of industries such as plastics, textile, paper and cosmetics use dyes in order to tincture their products. These molecules are common water contaminants and they may be regularly found a lot in industrial wastewaters. Textile plants, particularly those involved in finishing processes, consume major amount of dyes and they are main source of water pollution. Before a biological treatment of these effluents, the removal of these contaminants is essential but getting rid of these colored waste-waters makes a huge problem for the industry as well as a menace to the environment. There are many processes to remove dyes from colored effluents such as adsorption, precipitation, chemical degradation, photo degradation, biodegradation, chemical coagulation and electrocoagulation [1]. Elimination of contaminants using adsorption process from wastewater is the effective and simple method. Great deals of adsorbents such as agricultural wastes, natural compounds, activated carbon, etc. were used as adsorbents [2–6]. Polymeric materials have been appearing as potential alternative to traditional adsorbents. Capacity of these materials is very high to remove variety of organic and inorganic pollutants [7,8]. Several polymeric adsorbents have been utilized to remove different dyes [9–12]. Environmental protection attempts and developments in the technology have resulted in sever discharge standards. To obtain the best control and management, new concepts including effective operation and design should be improved and comprehended. Hence, a high quality representative model can supply a desirable solution in the process control and helps to illustrate the real process performance and to develop a continual control strategy for this kind of technologies. Since the process relies on various factors, the modeling of these processes involves many problems, i.e. we are dealing with a multivariate system. It is obvious that these problems cannot be solved by simple linear multivariate correlation.

Mathematical correlations, especially artificial neural network (ANN), have shown encouraging results in comparison with other

* Corresponding author at: Institut de Recherche en Génie Chimique et Pétrolier (IRGCP), Paris Cedex, France. Tel.: +33 164694970; fax: +33 164694968.
E-mail addresses: a.h.m@irgcp.fr, amir_h_mohammadi@yahoo.com (A.H. Mohammadi).

http://dx.doi.org/10.1016/j.jece.2014.06.002

currently published models for anticipation of the dye removal systems [13–16]. These models have many adjustable parameters containing of weights and biases which may lead to unreliability of these. The present investigation discusses the use of least square supported vector machine (LS-SVM) model [17] to predict the color removal efficiency of C.I.Direct Red 31 (DR31), C.I.Acid Blue 92 (AB92) and C.I.Direct Green 6 (DG6) dyes solution by Gemini polymeric Nano architecture (GPN) from single and ternary systems in adsorption process. To improve this model, 630 data sets have been used from dye removal experimental tests for single and ternary systems. Statistical and graphical error analyses have been performed to confirm the accuracy of the model. The results show that the developed model provides predictions in excellent agreement with the experimental data. Furthermore, it is demonstrated that the proposed model is capable of simulating the actual physical trend of the dye removal with alternation of absorbent dosage and initial concentration of dyes. The statistical method has been used for investigation of any existing suspected dye degradation data. The algorithm is carried out on the basis of the Leverage approach, in which the statistical Hat matrix, Williams plot, and the residuals of the model results lead to recognition of the possible outliers. In a successful manner, all the experimental data appear to be reliable except two. Thus, the developed model could be reliable for prediction of the dye removal in its applicability domain.

Experimental data and model development

Experimental database

The data used in this study were obtained from our previous paper [18]. To improve this model, 630 data sets have been used from dye removal experimental tests for single and ternary systems. Parameters of adsorbate type, adsorbent dosage, dye concentration, salt and absorbance time were used as inputs to LS-SVM model to predict dye removal efficiency (i.e., output). Table 1 summarizes the input/output variables of the models as well as their domains. As it can be seen, employed data bank covers a wide range of experimental conditions.

The type of adsorbent used in this study is GPN. This polymer as a novel adsorbent was synthesized in our previous work [18]. The structure of this polymer is shown in Fig. 1. The dye solutions were prepared by dissolving a defined quantity of the dye in distilled water. The characteristics of the dyes are shown in Table 2. Procedures of dye adsorption were done by mixing of GPN containing 250 mL of a dye solution (300 mg/L). The effects of adsorbent dosage dye concentration and salt were investigated on dye removal from single (sin.) and ternary (ter.) systems. At the end of the adsorption experiments, the solution samples were centrifuged, and the dye concentration was determined [18].

The maximum adsorption capacities (Q_0) were 1250, 1428 and 1000 mg/g for the adsorption of Direct Red 31, Direct Green 6 and Acid Blue 92, respectively. It is obvious that the adsorption capacity of GPN is much better than most of the other polymeric adsorbents reported currently (Table 3). The large values of Q_0 could belong to the strong adsorption affinity of GPN toward the three dyes, which was caused by the unique surface and charge neutralization.

Table 1
Ranges of the data were used for developed models.

	Variable	Range/type
Input	Adsorbate type	DR31, AB92, DG6
	Salt	$NaHCO_3$, Na_2CO_3, Na_2SO_4
	Adsorbent dosage	0.4–2 g/L
	Dye concentration	0.3–0.6 g/L
	Absorbance time	0–60 min
Output	Dye removal percent	0–100%

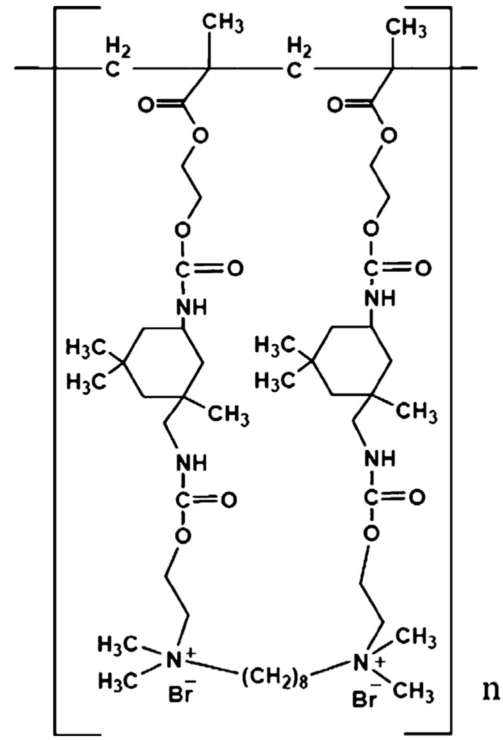

Fig. 1. The structure of absorbent (GPN) [18].

Model development

Broad family of algorithmic approaches are subsumed under category of machine learning techniques aiming to find reliable predictions through inferring from a set of experimental measurements which are based on a coherent statistical foundation. Machine learning techniques, for example, artificial neural networks (ANN) have already been applied in several engineering problems and have generally proven to yield highly accurate models [19–21]. However, ANN-based models may possess some deficiencies in reproducing the obtained results, mainly due to the fact that network parameters are randomly initialized and variations of stopping criteria during optimization processes may also impair the quality of model predictions [17]. On the other hand, support vector machine (SVM) has been proved to be an established and powerful tool employed in solving complex engineering problems [17,22]. The concept of SVM was first introduced by Vapnik [23] in 2000 as a supervised learning algorithm. Since then, it has been applied to several nonlinear classification and function approximation problems [24,25]. SVM has a number of distinct advantages compared to traditional learning methods based on ANN [17]:

- Possibility of convergence to the global optima is more than other ones in the SVM-based methods.
- Normally, these methods result in a solution that can be swiftly earned using standard algorithms.
- The topology of the network in the SVM-based methods is appointed after the training process ends, which there is no need to determine it before.
- Over-fitting problems occur less than the ANN models.