



## Research Article

# On the Evaluation of Rhamnolipid Biosurfactant Adsorption Performance on Amberlite XAD-2 Using Machine Learning Techniques

Fengqin Chen,<sup>1</sup> Jinbo Huang,<sup>2</sup> Xianjun Wu,<sup>3</sup> Xiaoli Wu <sup>4</sup> and Arash Arabmarkadeh <sup>5,6</sup>

<sup>1</sup>Inspection Department, Maoming People's Hospital, Maoming Guangdong 525000, China

<sup>2</sup>Logistics Department, Maoming People's Hospital, Maoming Guangdong 525000, China

<sup>3</sup>School of Computer, Guangdong University of Petrochemical Technology, Maoming Guangdong 525000, China

<sup>4</sup>Burn Department of Maoming People's Hospital, Maoming Guangdong 525000, China

<sup>5</sup>Biotechnology Group, Faculty of Chemical Engineering, Tarbiat Modares University, P.O. Box 14115-143, Tehran, Iran

<sup>6</sup>Microbial Biotechnology Department, Agricultural Biotechnology Research Institute of Iran (ABRII), Agricultural Research Education and Extension Organization (AREEO), Karaj, Iran

Correspondence should be addressed to Xiaoli Wu; [xiaoliwu\\_6859@21cn.com](mailto:xiaoliwu_6859@21cn.com)  
and Arash Arabmarkadeh; [arash.arabmarkade@modares.ac.ir](mailto:arash.arabmarkade@modares.ac.ir)

Received 25 January 2021; Revised 6 March 2021; Accepted 31 March 2021; Published 27 April 2021

Academic Editor: Alireza Baghban

Copyright © 2021 Fengqin Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Biosurfactants are a series of organic compounds that are composed of two parts, hydrophobic and hydrophilic, and since they have properties such as less toxicity and biodegradation, they are widely used in the food industry. Important applications include healthy products, oil recycling, and biological refining. In this research, to calculate the curves of rhamnolipid adsorption compared to Amberlite XAD-2, the least-squares vector machine algorithm has been used. Then, the obtained model is formed by 204 adsorption data points. Various graphical and statistical approaches are applied to ensure the correctness of the model output. The findings of this study are compared with studies that have used artificial neural network (ANN) and data group management method (GMDH) models. The model used in this study has a lower percentage of absolute mean deviation than ANN and GMDH models, which is estimated to be 1.71%. The least-squares support vector machine (LSSVM) is very valuable for investigating the breakthrough curve of rhamnolipid, and it can also be used to help chemists working on biosurfactants. Moreover, our graphical interface program can assist everyone to determine easily the curves of rhamnolipid adsorption on Amberlite XAD-2.

## 1. Introduction

As mentioned above, biosurfactants are organic compounds that are produced by microorganisms and consist of two parts: hydrophilic and hydrophobic. They are often produced by bacteria on living surfaces. One of the reasons for attracting many industrial applications to biosurfactants is due to their amphiphilic properties. Among the usable and outstanding capabilities of biosurfactants used in various industries such as mines, fertilizers, petrochemicals, and petroleum, we can mention the environmental degradability and reduction of surface tension between interstitial and low

toxicity. The reduction of the interfacial tension is due to the increase in the solubility of hydrophilic molecules when using biosurfactants. Capabilities such as surface modification and interfacial tension have made surfactants attractive to the industry. Rhamnolipids (RLs) are the most studied type of biosurfactants. According to the literature, rhamnolipids can reduce water surface tension by about 60% [1–3] for different concentrations of RL 50–65 mg/L. The production of RLs usually involves a final product from a dilute solution contaminated with undesirable impurities. There are several ways to increase the concentration and eliminate contaminants in which the adsorption process is widely studied.

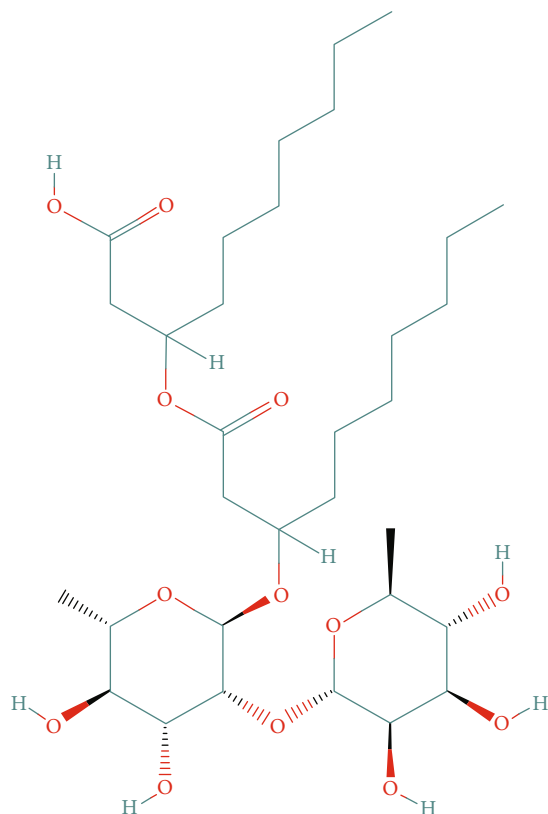


FIGURE 1: Chemical structure of the first identified rhamnolipid, symbolized as Rha-Rha-C<sub>10</sub>-C<sub>10</sub>.

In this research, activated carbon is used for adsorbent in the process. The breakthrough curve of a packed column is a very significant attribute of this system. As a result, determining such curves will be useful for optimizing and understanding the performance of the column. To model the adsorption phenomena, the mass balance in liquid and solid phases is evaluated. It may also include modeling the porous and liquid film resistance and also axial dispersion. Finally, with a suitable software package, a set of differential equations can be solved.

Ill conditions and uncertainty in differential equations make using conventional mathematical models not suitable. Intelligent models have to be a powerful tool in solving process modeling problems. To predict the optimized targets, in various fields such as petroleum and gas fields, methods such as SVM, ANN, group method data manipulation (GMDH), fuzzy logic system, and adaptive fuzzy neural inference system can be used. Interactions between AI neurons are achieved by connecting different units. Artificial neurons' interactions are achieved by connecting different units. Each weighted output is related to the sum of the output from the previous synaptic weight layer, and then it is used as an input for a specific neuron. Backpropagation ANNs are extensively applied, as they have shown to be a capable and powerful tool [4]. The GMDH model is a type of backpropagation ANN that was proposed by Ivakhnenko [5]. Darwin's theory of selection inspired this approach. The prominent feature of this method is the internal process of the elements [6–8].

To process elements in a conventional ANN, log sigmoid, hard limit, linear, and tangent sigmoid transfer functions are considered. On the other hand, the GMDH method constructs simple polynomials, roughly predicting the targeted systems. In the next step, the complexity of the polynomials is further developed so that satisfactory models are achieved. [9, 10].

Due to the importance of predicting a trustworthy estimation of breakthrough curves, this research is aimed at predicting of breakthrough curves utilizing the LSSVM method for rhamnolipid (Figure 1) adsorption over Amberlite XAD-2. Furthermore, results are compared with those of ANN and GMDH models. The investigated model takes into account 204 data points in its network for adsorption over the Amberlite XAD-2. Various graphical and statistical methods are considered to evaluate the accuracy of this strategy.

## 2. Model Development

In the present research, the LSSVM strategy was applied to calculate the curve to achieve rhamnolipid uptake relative to the Amberlite XAD-2 model resulting in a more simplified way [11, 12]. SVM can be defined as a function as below:

$$f(x) = w^T \varphi(x) + b. \quad (1)$$

The parameters of the above expression are as follows:

$w^T$  denotes the transpose vector corresponding to the output layer.

$b$  and  $\varphi(x)$  represent the bias and the kernel function, respectively.

The input ( $x$ ) consists of  $N \times n$  dimension in which  $n$  and  $N$  are input parameters and some data points, respectively. The following cost function is optimized to evaluate  $w^T$  and  $b$  parameters [13]:

$$\text{cost function} = \frac{1}{2} w^T + C \sum_{k=1}^N (\xi_k - \xi_k^*), \quad (2)$$

which is constrained by

$$\begin{cases} y_k - w^T \varphi(x_k) - b \leq \varepsilon + \xi_k, k = 1, 2, \dots, N \\ w^T \varphi(x_k) + b - y_k \leq \varepsilon + \xi_k^*, k = 1, 2, \dots, N \\ \xi_k, \xi_k^* \geq 0 \end{cases} \quad (3)$$

$y_k$  is the  $k^{\text{th}}$  output while  $x_k$  is the  $k^{\text{th}}$  input.  $\varepsilon$  stands for the fixed precision of the estimation. Also, slack variables ( $\xi_k, \xi_k^*$ ) are dealing to determine the acceptable error margin. The below Lagrangian is applied to minimize the cost function:

$$L(a, a^*) = -\frac{1}{2} \sum_{k,l=1}^N (a_k - a_k^*)(a_l - a_l^*)K(x_k, x_l) - \varepsilon \sum_{k=1}^N (a_k - a_k^*) + \sum_{k=1}^N y_k (a_k - a_k^*),$$

$$\sum_{k=1}^N (a_k - a_k^*) = 0, a_k, a_k^* \in [0, c],$$

$$K(x_k, x_l) = \varphi(x_k)^T \varphi(x_l), k = 1, 2, \dots, N, \tag{4}$$

where  $a_k$  and  $a_k^*$  stand for Lagrangian multipliers. In the last step, the SVM is given below:

$$f(x) = \sum_{k,l=1}^N (a_k - a_k^*)K(x, x_k) + b. \tag{5}$$

Quadratic programming must be solved to determine the SVM parameters. The LSSVM eliminates deficiencies in the solving process of a quadratic programming problem [11, 12]. LSSVM uses the below equation in the process of model development:

$$\text{cost function} = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{k=1}^N e_k^2, \tag{6}$$

where

$\gamma$  denotes tuning parameter.

$e_k$  is the error variable.

The following constraints are applied to the cost function:

$$y_k = w^T \varphi(x_k) + b + e_k. \tag{7}$$

The Lagrangian of the LSSVM is expressed as

$$l(w, b, e, a) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{k=1}^N e_k^2 - \sum_{k=1}^N a_k (w^T \varphi(x_k) + b + e_k - y_k). \tag{8}$$

In the above phrase, the symbol  $a_k$  represents the Lagrangian multipliers. To optimize Eq. (8), its derivatives are set to zero, and as a result, the following equations are achieved:

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \implies w = \sum_{k=1}^N a_k \varphi(x_k), \\ \frac{\partial L}{\partial b} = 0 \implies \sum_{k=1}^N a_k = 0, \\ \frac{\partial L}{\partial e_k} = 0 \implies a_k = \gamma e_k, k = 1, 2, \dots, N, \\ \frac{\partial L}{\partial a_k} = 0 \implies w^T \varphi(x_k) + b + e_k - y_k = 0, k = 1, 2, \dots, N. \end{cases} \tag{9}$$

By solving the aforementioned equations, LSSVM parameters are obtained. LSSVM employs the kernel function in the same way that SVM strategy does. The most common applied kernel function is the radial basis function (RBF) which is given by

$$K(x, x_k) = \exp(-\|x_k - x\|^2 / \sigma^2), \tag{10}$$

where  $\sigma^2$  stands for the tuning parameter corresponding to the kernel function. As a result, two tuning parameters ( $\sigma^2$  and  $\gamma$ ) are adjustable. The last-mentioned parameters can be determined by minimizing the error between the predicted values and experimental ones through the application of mean square error (MSE):

$$MSE = \frac{1}{N} \sum_{k=1}^N (y_k^{\text{pred.}} - y_k^{\text{exp.}})^2, \tag{11}$$

where  $y$  is the output value, and exp. and pred. subscripts denote experimental and predicted values, respectively. Also, in this paper, we used the particle swarm optimization algorithm for the determination of these tuning parameters. A typical diagram of the proposed LSSVM approach has been shown in Figure 2.

The adjusted parameters are  $\gamma$  and  $\sigma^2$  in the LSSVM model and based on the identified cost function (Eq. (11)), and these parameters are optimally determined by optimization technique. The values of  $\gamma$  and  $\sigma^2$  in this study are 984523.52 and 0.246, respectively, through the PSO algorithm with swarm size and iteration of 80 and 1000, respectively.

Different statistical error analyses such as mean absolute error (MAE), coefficient of determination ( $R^2$ ), and root means square error (RMSE) are implicated to analyze the model's performance.

$$R^2 = 1 - \frac{\sum_{k=1}^N (y_k^{\text{exp.}} - y_k^{\text{pred.}})^2}{\sum_{k=1}^N (y_k^{\text{exp.}} - y_{\text{ave.}})^2},$$

$$\text{MAE} = \frac{\sum_{k=1}^N |y_k^{\text{pred.}} - y_k^{\text{exp.}}|}{N}, \tag{12}$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{k=1}^N (y_k^{\text{pred.}} - y_k^{\text{exp.}})^2}.$$

### 3. Identification of Outlying Experimental Data

The outlier is a set of data having a different behavior in comparison with the bulk of data. Finding outliers would improve the accuracy and reliability of a proposed model remarkably. To help to trace outliers, there are two procedures numerical and graphical procedures. One of the most powerful methods is the Leverage method in which the deviation of estimated values from the experimental ones is calculated. It also includes dealing with Hat matrix being made of experimental and predicted data. The equation below is used for calculating the Hat indices [14–16]:

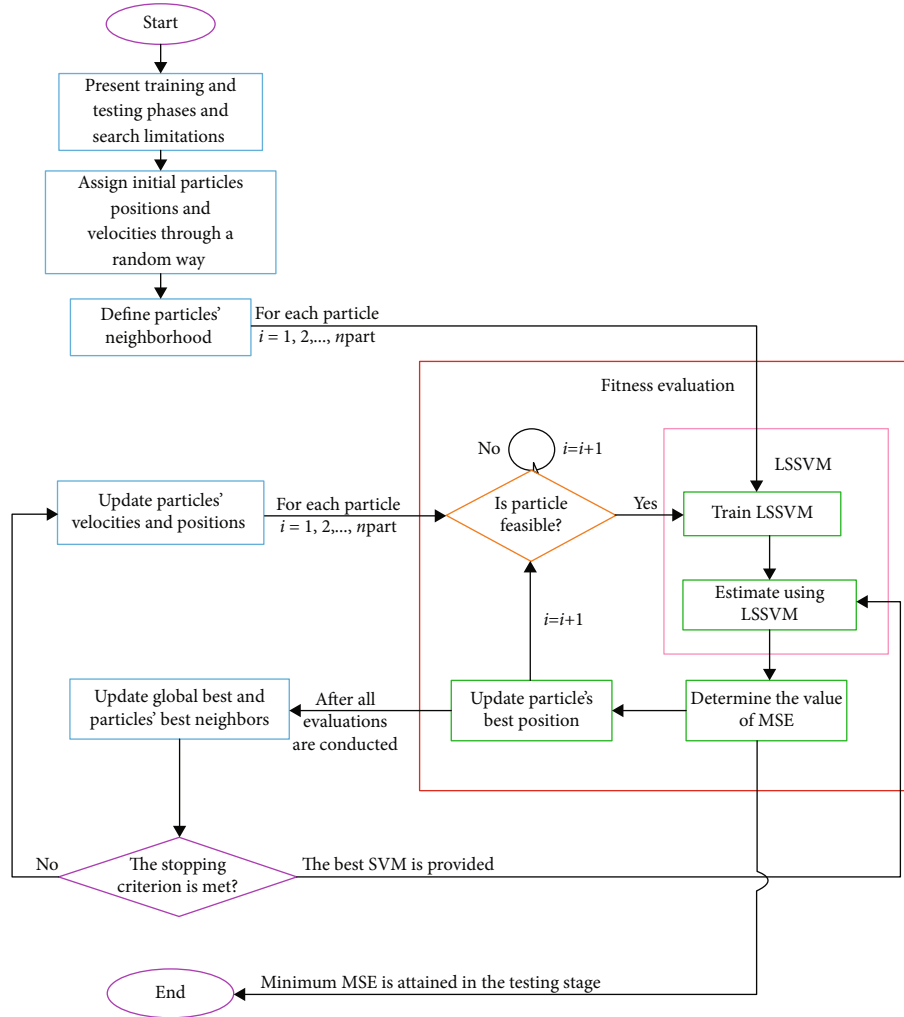


FIGURE 2: Schematic diagram of PSO-LSSVM strategy.

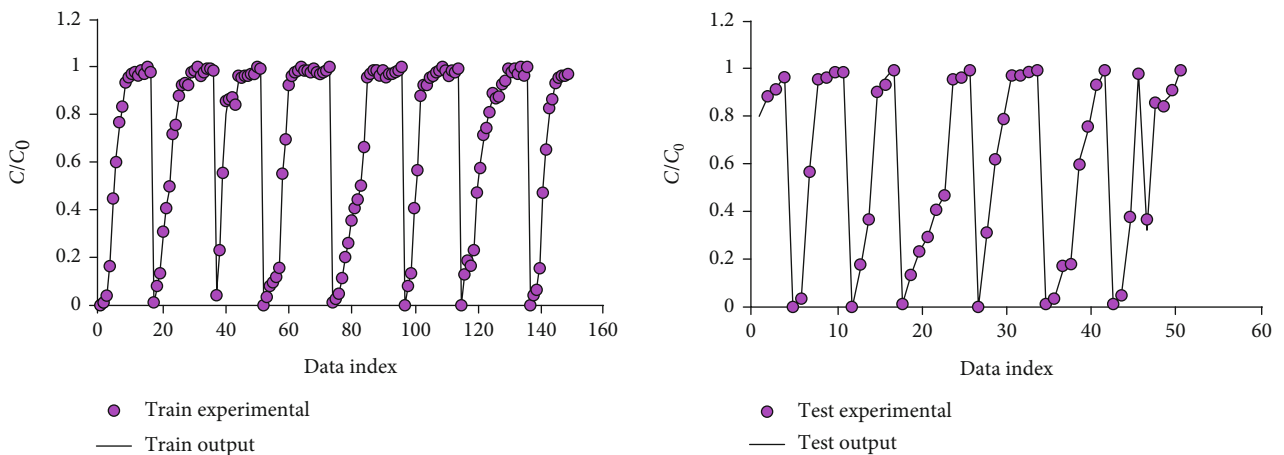


FIGURE 3: Plot of PSO-LSSVM model's prediction vs. experimental data at training and testing stages.

$$H = X(X^t X)^{-1} X^t. \quad (13)$$

$X (n \times x)$  is a matrix including  $n$  data and  $k$  parameters of the model, and  $t$  denotes the transpose matrix. The diagonal

values of the matrix ( $H$ ) are called  $H$  values.  $H$  values will aid in the detection process of the possible outliers, utilizing a Williams plot in which the relationship between standardized crossvalidated residuals ( $R$ ) and Hat indices is shown. The warning leverage is given as follows:

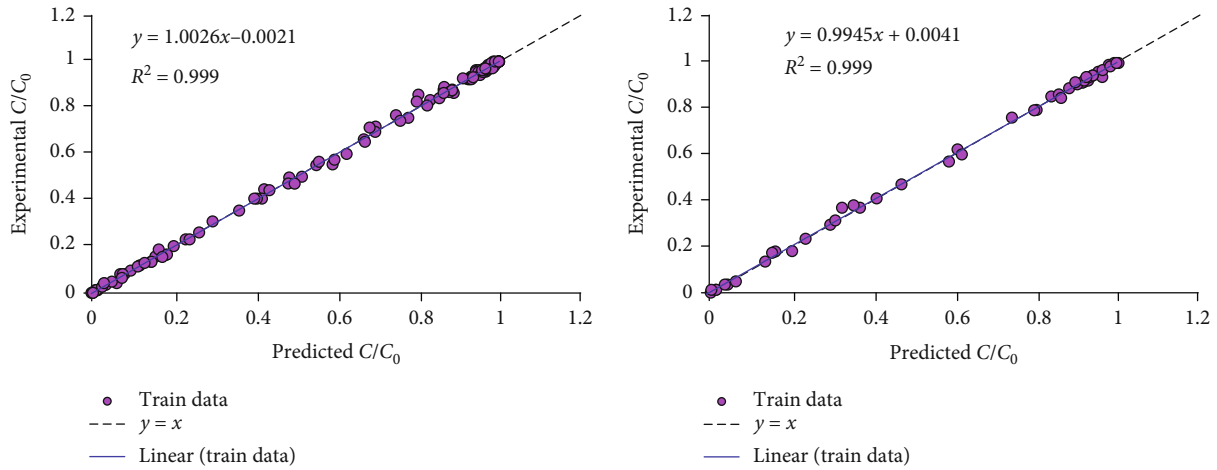


FIGURE 4: Regression plot of suggested PSO-LSSVM model at training and testing stages.

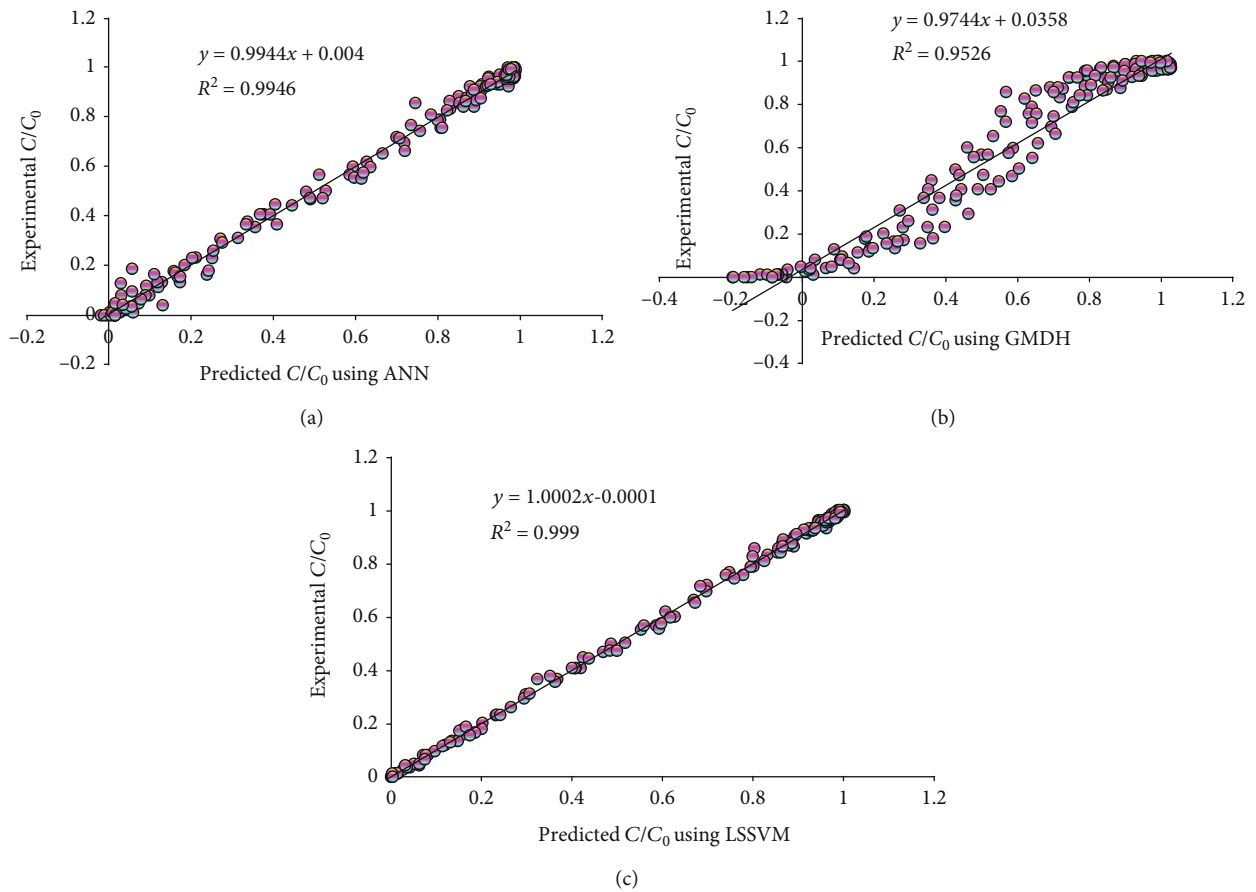


FIGURE 5: Cross plot of predictions of different models for total data points: (a) ANN, (b) GMDH, and (c) LSSVM.

$$H^* = 3 \frac{(f + 1)}{p} \tag{14}$$

$p$  and  $f$  stand for numbers of data points and model parameters, respectively.

A reliable model would contain the majority of the predicted values by satisfying the following constraint:

$$R \in [-3, 3], 0 < H < H^* \tag{15}$$

Regardless of the value of  $H$ , if the value of  $R$  for a given

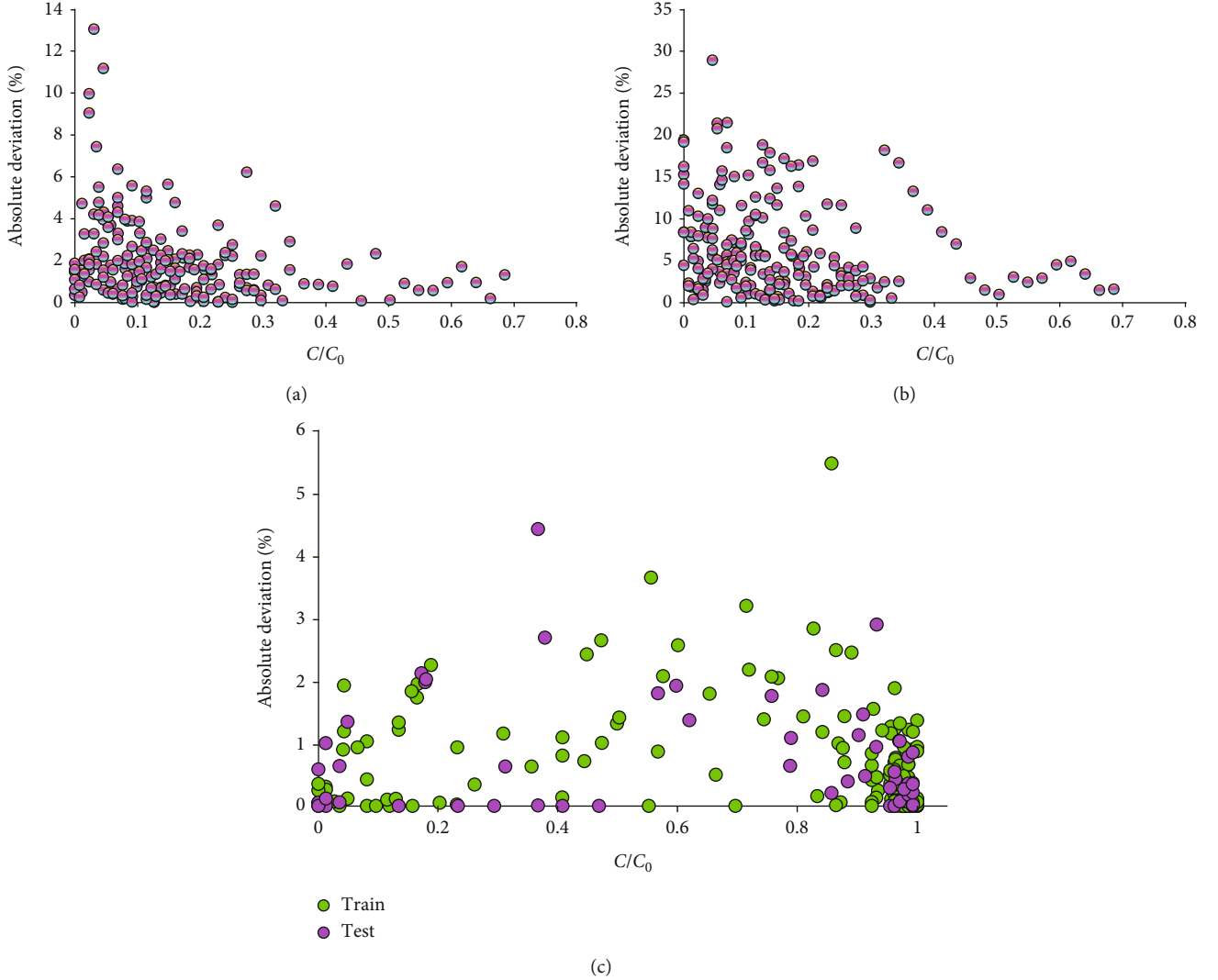


FIGURE 6: Absolute error for different models' outcomes: (a) ANN, (b) GMDH, and (c) LSSVM.

data is outside the above range, it is considered a possible candidate for being an outlier. The data in this paper are provided in Table S1, and this data set is taken from the previous paper [17]. As discussed, input parameters are initial rhamnolipid concentration, fixed bed height, flow velocity, and run time, while the ratio of final to initial concentration of rhamnolipid ( $C/C_0$ ) would be the output parameters ones. The 204 data points are divided into two categories: training and testing.

To create the LSSVM model, 75% of data points are considered as learning points, and the rest of them were used to examine the efficiency of the opposed model. Furthermore, data are normalized within the range of  $[-1,1]$  applying the equation below:

$$D_N = 2 \frac{D - D_{\min}}{D_{\max} - D_{\min}} - 1. \quad (16)$$

Here,  $D$  and  $D_N$  represent actual and normalized data

TABLE 1: Statistical parameters calculated for three models.

Analysis	LSSVM		ANN Total	GMDH Total
	Train	Test		
MSE	0.0001	0.0002	0.0001	0.0005
AAD	0.7293	0.8043	0.7481	1.9111
$R^2$	0.9990	0.9990	0.9990	0.9946
STD	0.0112	0.0122	0.0115	0.0271

points, respectively. Also,  $D_{\min}$  and  $D_{\max}$  stand for minimum and maximum values of data points, respectively.

**3.1. Evaluation of the model's Accuracy.** The predictive model's accuracy is investigated employing different graphical and statistical methods. Figure 3 represents experimental data points and model estimation by the proposed LSSVM method in the training and testing stages.

Figure 4 shows predicted values against experimental ones. The more it would be close to line  $Y = X$ , the more appropriate the prediction of the proposed model.

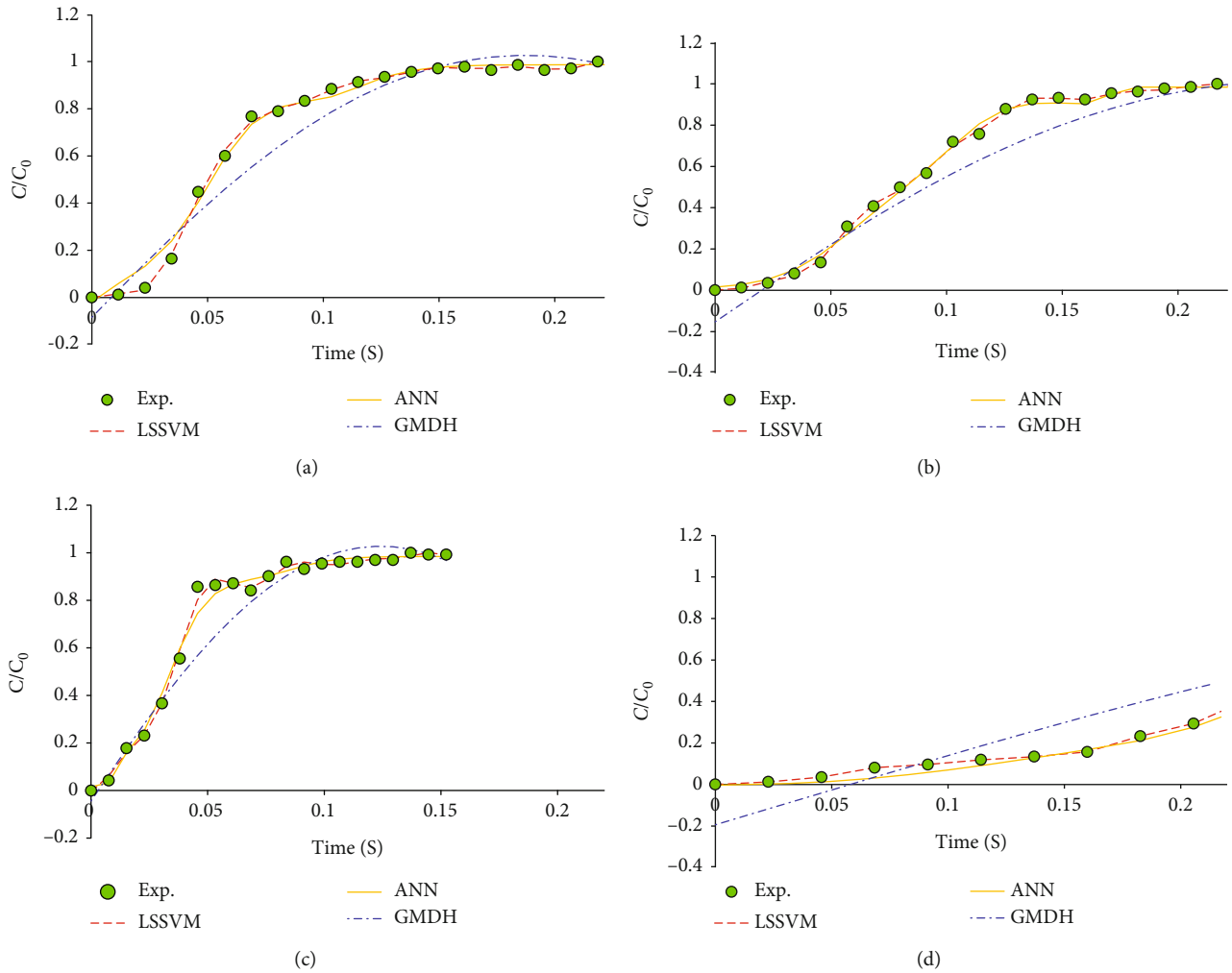


FIGURE 7: Predicted  $C/C_0$  as a function of time by the three models for (a)  $H_0 = 7, U = 160,$  and  $C_0 = 24,$  (b)  $H_0 = 11, U = 160,$  and  $C_0 = 8,$  (c)  $H_0 = 7, U = 240,$  and  $C_0 = 24,$  and (d)  $H_0 = 11, U = 80,$  and  $C_0 = 8.$

Results of the current study are compared to the LSSVM, ANN, and GMDH models [17]. To estimate the curvature of rhamnolipids on activated carbon, the structures of the proposed GMDH model are given below:

(i) First layer:

$$\text{Node 1 : } z_1 = 0.183x_1^2 + 0.149x_2^2 + 0.165x_1x_2 + -0.242x_1 - 0.048x_2 + 0.678$$

$$\text{Node 2 : } z_2 = 0.141x_2^2 - 1.6x_4^2 - 0.21x_2x_4 - 0.028x_2 + 2.76x_4 - 0.167$$

$$\text{Node 3 : } z_3 = 0.182x_3^2 - 1.606x_4^2 - 0.124x_3x_4 - 0.075x_3 + 2.778x_4 - 0.167$$

$$\text{Node 4 : } z_4 = -0.002x_1^2 - 1.64x_4^2 + 0.172x_1x_4 - 0.169x_1 - 2.684x_4 - 0.075 \quad (17)$$

(ii) Second layer:

$$\text{Node 1 : } w_1 = -0.253z_1^2 + 0.248z_2^2 - 0.319z_1z_2 - 0.05z_1 + 0.948z_2 - 0.044$$

$$\text{Node 2 : } w_2 = 4.923z_1^2 + 0.234z_3^2 - 0.057z_1z_3 - 5.738z_1 + 0.794z_3 + 1.676$$

$$\text{Node 3 : } w_3 = 2.915z_1^2 + 0.237z_4^2 - 0.535z_1z_4 - 2.733z_1 + 1.097z_4 + 0.577 \quad (18)$$

(iii) Third layer:

$$\text{Node 1 : } u_1 = -0.802w_1^2 - 1.248w_2^2 + 1.972w_1w_2 + 0.419w_1 + 0.666w_2 - 0.011$$

$$\text{Node 2 : } u_2 = 1.433w_2^2 + 1.571w_3^2 - 3.077w_2w_3 + 0.603w_2 + 0.48w_3 - 0.014 \quad (19)$$



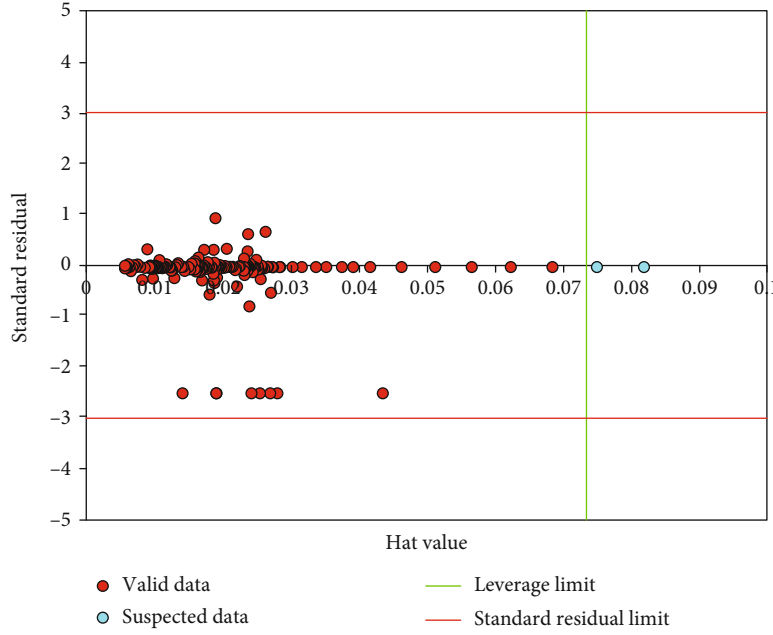


FIGURE 8: Diagnosis of the probable outlier data and applicability domain of the applied model.

(iv) Genome expression:

$$\begin{aligned} \text{Node 1 : } \frac{C}{C_0} = & -0.802u_1^2 - 1.24u_2^2 - 1.972u_1u_2 \\ & + 0.419u_1 + 0.666u_2 - 0.011 \end{aligned} \quad (20)$$

$$\begin{aligned} \text{MSE} = & \left( \frac{\sum_{i=1}^N (C_{\text{Pred}}(i) - C_{\text{Exp}}(i))^2}{N} \right), \\ \text{STD} = & \sum_{i=1}^n \left( \frac{(C_{\text{Pred}}(i) - \bar{C}_{\text{Exp}}(i))^2}{N} \right)^{0.5}. \end{aligned} \quad (21)$$

The ANN model based on these four input variables as mentioned as follows:

- (1) input layer
- (2) hidden layer including six neurons
- (3) output layer

Figure 5 represents the cross plot of the aforementioned strategies. As explained, data points of the LSSVM model are closer to the line  $Y = X$ , than ANN and GMDH models. Also, the calculation of the determination coefficient shows that the proposed LSSVM approach is superior to ANN and GMDH in terms of accuracy.

Compared to ANN and GMDH models, the less relative error is observed in the proposed LSSVM model. Figure 6 indicates more reliability of the suggested LSSVM model.

Estimation accuracy is also investigated by applying the following statistical methods:

$$\begin{aligned} R^2 = & 1 - \frac{\sum_{i=1}^N (C_{\text{Pred}}(i) - C_{\text{Exp}}(i))^2}{\sum_{i=1}^N (C_{\text{Pred}}(i) - \bar{C}_{\text{Exp}}(i))^2}, \\ \% \text{AAD} = & \frac{100}{N} \sum_{i=1}^N |C_{\text{Pred}}(i) - C_{\text{Exp}}(i)|, \end{aligned}$$

Table 1 presents statistical values of the presented model compared with ANN and GMDH approaches showing the higher value of  $R^2$  and lower values of STD, AAD, and RMSE, and as a result, the LSSVM model possesses higher accuracy and reliability than others. The dependency of the  $(C/C_0)$  as an output parameter on input parameters is illustrated in Figure 7.

Four different conditions of  $H_0 = 7$ ,  $U = 160$ , and  $C_0 = 24$ , and  $H_0 = 11$ ,  $U = 160$ , and  $C_0 = 8$  and  $H_0 = 7$ ,  $U = 240$ , and  $C_0 = 24$  and  $H_0 = 11$ ,  $U = 80$ , and  $C_0 = 8$  were investigated to measure the prediction ability of LSSVM, ANN, and GMDH models for indicating that the LSSVM model acquired better estimation. As this figure shows, as time goes by, the ratio of  $C/C_0$  increases.

In the last part of this research, the leverage approach is applied to find outliers, employing the Hat matrix, Williams plot, and residuals. As discussed, Eq. (13) is used to calculate  $H$  values. Figure 8 also illustrates the Williams plot. All of the  $H$  is in the range  $[-3, +3]$ , and  $R$  is in the range  $[0, 0.08]$ , and then the accuracy of the proposed model is desirable and acceptable; so, the accuracy of the proposed model is satisfactory. There are only two of the data points that are outside of the applicable domain which is shown in the figure by a blue circle. As  $R$  values approach zero and  $H$  value reduces, the reliability of data points is increased [18–23].



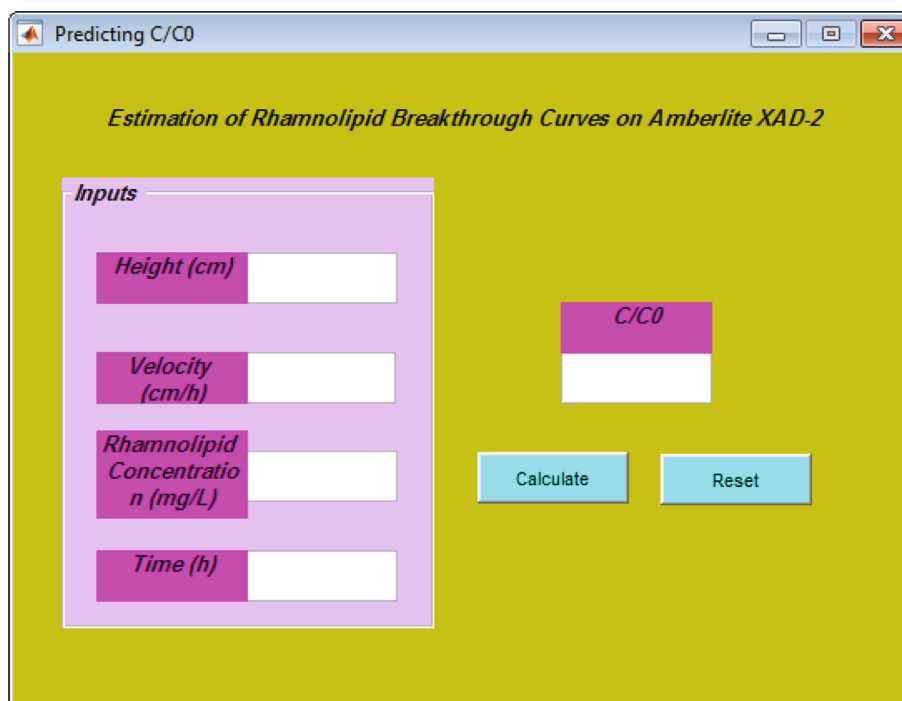


FIGURE 9: GUI version of the developed LSSVM model.

#### 4. Conclusion

Then LSSVM approach was employed to estimate breakthrough curves of rhamnolipid adsorption over Amberlite XAD-2 as a function of fixed bed height, flow velocity, runtime, and initial rhamnolipid concentration. The particle swarm optimization method was employed for the training process enhancing the accuracy of the proposed model. Various statistical and graphical methods were applied to evaluate the model's reliability showing that the AAD% value for adsorption over activated carbon was 0.75%. For ANN and GMDH models that were developed by Padilha et al., AAD% of activated carbon is reported to be 1.9% and 6.2%. Based on the above evidence, we can find that the proposed LSSVM model is more reliable for the process of predicting the breakthrough curves.

#### Appendix

##### Instructions of the Developed Program

A graphical user interface (GUI) version of the model is developed (Figure 9). The code is compiled to an Exe file which is given in the supplementary content. Matlab software must be installed before running the code. It starts by giving four parameters as input; then, to show the output result, it is enough to click on the calculate button.

##### Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

#### Conflicts of Interest

The authors declare that they have no conflicts of interest.

#### Acknowledgments

The research presented in this paper was supported by the Funds of High-level Hospital Construction Research Project of Maoming People's Hospital.

#### Supplementary Materials

Table S1: experimental data points used in this study. (*Supplementary materials*)

#### References

- [1] A. Abalos, A. Pinazo, M. Infante, M. Casals, F. García, and A. Manresa, "Physicochemical and antimicrobial properties of new rhamnolipids produced by *Pseudomonasaeruginosa* AT10 from soybean oil refinery wastes," *Langmuir*, vol. 17, no. 5, pp. 1367–1371, 2001.
- [2] A. A. Bodour and R. M. Miller-Maier, "Application of a modified drop-collapse technique for surfactant quantitation and screening of biosurfactant-producing microorganisms," *Journal of Microbiological Methods*, vol. 32, no. 3, pp. 273–280, 1998.
- [3] J.-Y. Wu, K.-L. Yeh, W.-B. Lu, C.-L. Lin, and J.-S. Chang, "Rhamnolipid production with indigenous *Pseudomonas aeruginosa* EM1 isolated from oil-contaminated site," *Bioresource Technology*, vol. 99, no. 5, pp. 1157–1164, 2008.
- [4] R. Hecht-Nielsen, "Theory of the backpropagation neural network," *Neural Networks*, vol. 1, pp. 445–448, 1988.

- [5] A. G. Ivakhnenko, "Polynomial theory of complex systems," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. -SMC-1, no. 4, pp. 364–378, 1971.
- [6] S. Atashrouz, M. Mozaffarian, and G. Pazuki, "Modeling the thermal conductivity of ionic liquids and ionanofluids based on a group method of data handling and modified Maxwell model," *Industrial & Engineering Chemistry Research*, vol. 54, no. 34, pp. 8600–8610, 2015.
- [7] S. Atashrouz, H. Mirshekar, A. Hemmati-Sarapardeh, M. K. Moraveji, and B. Nasernejad, "Implementation of soft computing approaches for prediction of physicochemical properties of ionic liquid mixtures," *Korean Journal of Chemical Engineering*, vol. 34, no. 2, pp. 425–439, 2017.
- [8] A. Dargahi-Zarandi, A. Hemmati-Sarapardeh, S. Hajirezaie, B. Dabir, and S. Atashrouz, "Modeling gas/vapor viscosity of hydrocarbon fluids using a hybrid GMDH-type neural network system," *Journal of Molecular Liquids*, vol. 236, pp. 162–171, 2017.
- [9] A. Hemmati-Sarapardeh, M. H. Ghazanfari, S. Ayatollahi, and M. Masihi, "Accurate determination of the CO<sub>2</sub>-crude oil minimum miscibility pressure of pure and impure CO<sub>2</sub>-streams: a robust modelling approach," *The Canadian Journal of Chemical Engineering*, vol. 94, no. 2, pp. 253–261, 2016.
- [10] E. Mohagheghian, H. Zafarian-Rigaki, Y. Motamedi-Ghahfarrokhi, and A. Hemmati-Sarapardeh, "Using an artificial neural network to predict carbon dioxide compressibility factor at high pressure and temperature," *Korean Journal of Chemical Engineering*, vol. 32, no. 10, pp. 2087–2096, 2015.
- [11] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300, 1999.
- [12] K. Pelckmans, J. A. K. Suykens, T. Van Gestel et al., *LS-SVMlab: a matlab/c toolbox for least squares support vector machines*, Tutorial, vol. 142, KULeuven-ESAT, Leuven, Belgium, 2002.
- [13] J. A. K. Suykens, T. Van Gestel, and J. De Brabanter, *Least Squares Support Vector Machines*, World Scientific, 2002.
- [14] M. E. Johnson, "Wiley series in probability and mathematical statistics," in *Multivariate Statistical Simulation*, pp. 231–235, John Wiley & Sons, Inc., 1969.
- [15] C. R. Goodall, "13 computation using the QR decomposition," *Handbook of Statistics*, vol. 9, pp. 467–508, 1993.
- [16] F. Gharagheizi, A. Eslamimanesh, M. Sattari, B. Tirandazi, A. H. Mohammadi, and D. Richon, "Evaluation of thermal conductivity of gases at atmospheric pressure through a corresponding states method," *Industrial & Engineering Chemistry Research*, vol. 51, no. 9, pp. 3844–3849, 2012.
- [17] C. E. . A. Padilha, C. A. . A. Padilha, D. F. . S. Souza, J. A. de Oliveira, G. R. de Macedo, and E. S. dos Santos, "Prediction of rhamnolipid breakthrough curves on activated carbon and Amberlite XAD-2 using artificial neural network and group method data handling models," *Journal of Molecular Liquids*, vol. 206, pp. 293–299, 2015.
- [18] A. Hemmati-Sarapardeh, R. Alipour-Yeganeh-Marand, A. Naseri et al., "Asphaltene precipitation due to natural depletion of reservoir: determination using a SARA fraction based intelligent model," *Fluid Phase Equilibria*, vol. 354, pp. 177–184, 2013.
- [19] A. Baghban, M. A. Ahmadi, B. Pouladi, and B. Amanna, "Phase equilibrium modeling of semi-clathrate hydrates of seven commonly gases in the presence of TBAB ionic liquid promoter based on a low parameter connectionist technique," *The Journal of Supercritical Fluids*, vol. 101, pp. 184–192, 2015.
- [20] A. Baghban, M. A. Ahmadi, and B. Hashemi Shahraki, "Prediction carbon dioxide solubility in presence of various ionic liquids using computational intelligence approaches," *The Journal of Supercritical Fluids*, vol. 98, pp. 50–64, 2015.
- [21] A. Baghban, A. Bahadori, A. H. Mohammadi, and A. Behbahaninia, "Prediction of CO<sub>2</sub> loading capacities of aqueous solutions of absorbents using different computational schemes," *International Journal of Greenhouse Gas Control*, vol. 57, pp. 143–161, 2017.
- [22] A. Baghban, M. Bahadori, J. Rozyn et al., "Estimation of air dew point temperature using computational intelligence schemes," *Applied Thermal Engineering*, vol. 93, pp. 1043–1052, 2016.
- [23] A. Baghban, A. H. Mohammadi, and M. S. Taleghani, "Rigorous modeling of CO<sub>2</sub> equilibrium absorption in ionic liquids," *International Journal of Greenhouse Gas Control*, vol. 58, pp. 19–41, 2017.