A Robust Method to Predict Equilibrium and Kinetics of Sulfur and Nitrogen Compounds Adsorption from Liquid Fuel on Mesoporous Material

Mohammad Reza Khosravi-Nikou[1*], Ahmad Shariati[1], Mohammad Mohammadian[1], Ali Barati-Harooni[2], and Adel Najafi-Marghmaleki[2]

[1] Associate Professor, Department of Gas Engineering, Petroleum University of Technology, Ahwaz, Iran

[2] Associate Professor, Department of Gas Engineering, Petroleum University of Technology, Ahwaz, Iran

[3] M.S. Student, Department of Gas Engineering, Petroleum University of Technology, Ahwaz, Iran

[4] M.S. Student, Department of Petroleum Engineering, Petroleum University of Technology, Ahwaz, Iran

[5] M.S. Student, Department of Petroleum Engineering, Petroleum University of Technology, Ahwaz, Iran

Abstract

This study presents a robust and rigorous method based on intelligent models, namely radial basis function networks optimized by particle swarm optimization (PSO-RBF), multilayer perceptron neural networks (MLP-NNs), and adaptive neuro-fuzzy inference system optimized by particle swarm optimization methods (PSO-ANFIS), for predicting the equilibrium and kinetics of the adsorption of sulfur and nitrogen containing compounds from a liquid hydrocarbon model fuel on mesoporous materials. All the models were evaluated by the statistical and graphical methods. The predictions of the models were also compared with different kinetics and equilibrium models. The results showed that although all the models lead to accurate results, the PSO-ANFIS model represented the most reliable and dependable predictions with the correlation coefficient (R^2) of 0.99992 and average absolute relative deviation (AARD) of 0.039%. The developed models are also able to predict the experimental data with better precision and reliability compared to literature models.

Keywords: Adsorption, Denitrogenation, Desulfurization, Equilibrium and Kinetics Model, PSO-ANFIS,

1. Introduction

Due to the ever-growing population of the world and the increasing demand for energy, searching for new sources of fossil fuel is of great importance. However, these new sources of fossil fuel are usually contaminated by sulfur and nitrogen-containing compounds [1–3]. The presence of sulfur and nitrogen impurities in these fuels, which are used in combustion engines, leads to the formation of SO_x and NO_x compounds which are important air pollutants [4]. Due to stringent environmental regulations, ultra-low concentrations of sulfur and nitrogen are allowable in transportation fuels. Hence, desulfurization

---

* Corresponding author:
  Email: mr.khosravi@put.ac.ir

and denitrogenation of the fuels to meet the environmental regulations is of great importance [4–6]. Selective elimination of sulfur and nitrogen compounds from fuels by using adsorptive desulfurization and denitrogenation process is a popular and attractive method. It benefits from different advantages such as being environmentally-friendly, cheap, simple and highly efficient [7–9]. There are various experimental works in the literature concerning the use of different adsorbents for the selective removal of sulfur and nitrogen compounds from liquid fuels. These adsorbents include zeolite-based material, activated carbons, metal oxides, metal sulfides, silica-based materials and silica gel [4, 5, 8, 10–14]. However, literature review shows that there are few modeling works reported on the adsorptive desulfurization and denitrogenation of fuels [15–17]. Most of these works make use of static and dynamic adsorption isotherms to model their experimental data. However, these models should be tuned on each set of experimental data and on each type of adsorbent. Hence, developing more general and accurate models that can predict the experimental adsorption data for larger adsorbents with fewer and fixed optimization parameters is of the utmost importance. In recent years, soft computing methods and intelligent models have been increasingly used for modeling of nonlinear and complex relations in different engineering systems. Taking into account the great capability of these models for controlling and handling the nonlinear and complex relations in engineering systems, they can be utilized for modeling the performance of adsorptive desulfurization and denitrogenation processes. This study focuses on the application of these models for the prediction of equilibrium and kinetics of the adsorption of sulfur and nitrogen from a model fuel on mesoporous adsorbents.

2. Predictive tools

2.1. Multilayer perceptron neural networks (MLP-NNs)

The multilayer perceptron type networks contain three layers, namely the input layer, the hidden layer, and the output layer, in their structure. The hidden layer could be composed of either one or several layers. Each hidden layer contains specific processing units known as neurons. The number of these neurons is dependent on the size of the input and output data. There are various methods and techniques for evaluating the number of neurons and the number of hidden layers in the structure of these networks. These methods include the trial-and-error approach and intelligent methods [18]. Usually a cost function which is the mean square error (MSE) between target values and predicted data is utilized to evaluate the reliability and precision of the outcomes of these networks. In MLP-NN models and models which are branches of artificial neural networks (ANNs), the error is back propagated within the network and the optimum values of weight and bias terms are determined by using a certain number of iterations known as epochs [19, 20]. The number of iterations or epochs must be such that they prevent the model from over fitting or under-fitting problems. In the case of under-fitting of the network, the training phase of the network is insufficient and incomplete because of providing less training for the network. Moreover, in the case of overfitting, the model fails to learn during the training phase and just memorizes. In both cases, the model predicts the test data with a poor degree of accuracy.

2.2. Radial basis function neural networks (RBF-NNs)

RBF-NNs are a branch of artificial neural networks in which they hold three layers, namely the input layer, the hidden layer, and the output layer, in their structure. These networks contain just one hidden layer in their structure. There are certain units called weight terms the value of which greatly affects the performance of each RBF-NN. The values of these terms are adjusted through using a technique named gradient decent approach. The hidden layer contains a number of units called neurons or hidden units in which each neuron holds a specific activation function (radial basis function (RBF)) [21, 22]. In

dealing with large scale and highly nonlinear problems, it is very important to increase the number of hidden neurons in order to access the desired accuracy and robustness of the model. Several specific characteristics of RBF-NNs such as their adaptive structure, great ability to generalize the results, acceptable accuracy, fast convergence to the optimum solution, and independency of the output value on the adjusted weight terms make these networks popular and powerful tools for solving nonlinear and complex problems in different areas of science and engineering. The number of nodes or neurons which are located in the input layer is identical to the number of input parameters (the dimensions of input vector) [23]. The output layer linearly combines the outputs of neurons of the hidden layer. In other words, this layer projects the nonlinearity of the problem into a new linear space.

Considering the number of N input data points containing D independent input parameters, there is an input vector with dimensions $N \times D$ represented by $X^p = \left\{ x_i^p : i = 1, 2, ..., D, \ p=1,2,....,N \right\}$ in which the aim of modeling and interpolating by RBF-NN is to probe a function such as $f(x)$ which confirms the equality of $f(X^p)=t^p$, where t^p is the related output parameter. For every given input parameter, the RBF-NN utilizes its RBFs which are in the hidden layer and are in the form of $\phi \left(\left\| x - x^p \right\| \right)$, where ϕ indicates the class and type of RBF explained later. Hence, the output of such a function is dependent on its input value which is denoted by $\left\| x - x^p \right\|$, i.e. the Euclidean distance between data point x and the related center x^p [24]. The output of this projection is the linear weighted summation of the outputs of RBFs denoted by $f(x) = \sum_{p=1}^{N} w_p \phi \left(\left\| x - x^p \right\| \right)$. The aim is to allocate the weight terms (w_p) in a manner that function $f(x)$ is capable of projecting the input data into their corresponding output values. To put it another way, the purpose is to adjust weight terms such that this relation $f\left(x^q \right) = \sum_{p=1}^{N} w_p \phi \left(\left\| x^q - x^p \right\| \right) = t^q$ is valid. By allocating these weights appropriately, there will be a continuous and differentiable function $f(x)$ which accurately passes through the input data points and effectively predicts their output values. Currently various types of RBFs are utilized to perform in the neurons of the hidden layer; however, the most popular one which represents acceptable results and is widely used in various modeling problems is the Gaussian type RBF represented by below equation [25, 26]:

$$\phi(r) = \exp\left(-\frac{r^2}{2\sigma^2} \right)$$

(1)

where σ controls the smoothness of RBF function, and r is the distance between a data point like x and a center like x^p; both σ and r are positive terms. In this work, this type of RBF was utilized in the hidden neurons of the hidden layer.

2.3. Adaptive neuro-fuzzy inference system (ANFIS)

Fuzzy logic (FL) is a concept introduced by Zadeh [27] for the first time to allow modeling of complex and obscure systems. The performances of fuzzy logic and classical logic are different. While classical logic takes the problem as the crisp true and false systems, the fuzzy logic tries to solve and model the problem under consideration by defining fuzzy sets of true and false or membership degree of truth in fuzzy sets [28].

Fuzzy logic utilizes if-then fuzzy linguistic rules to deal with the problem through an alternative and qualitative approach instead of performing exact quantitative evaluations. The accurate and clear knowledge about the features of the problem under consideration is essential for performing such a qualitative modeling process. However, due to insufficient information about the details and nonlinearity of the problem and also the different aspects of human knowledge, it is not suitable to define and design fuzzy systems based on human knowledge. A helpful and powerful alternative method is using the abilities of ANNs since they are able to extract and gain the essential knowledge from the instant and online data bases. The patterns learned and obtained from ANNs could be used by fuzzy logic to construct proper linguistic if-then rules to describe these patterns. The membership functions (MFs) are modified and tuned via using the learning capability of ANNs [28–31]. Generally, fuzzy inference system (FIS) structures include two main types, namely Mamdani FIS and Takagi-Sugeno-Kang (TSK) FIS. The Mamdani FIS utilizes fuzzy if-then rules on a subjective basis and according to expert statements. However, the TSK FIS puts these rules in a framework and creates them through relations between the input and the output data. The subjective definition which is utilized by Mamdani FIS for the generation of rules may result in ambiguity and inaccuracy. ANFIS models, which use the TSK type FIS for modeling nonlinear and complex relations between input and output data, usually provide better results [28, 29, 32–34]. Taking into account a first order TSK type FIS, the following rules can be used to construct a fuzzy model for a system with two input parameters and one output parameter [28]:

Rule I:

$$IF \ \langle X_1 \ is \ A_1 \ and \ X_2 \ is \ B_1 \rangle \ THEN \ \langle f_1 = m_1 X_1 + n_1 X_2 + r_1 \rangle \tag{2}$$

Rule ΙΙ:

$$IF \ \langle X_1 \ is \ A_2 \ and \ X_2 \ is \ B_2 \rangle \ THEN \ \langle f_2 = m_2 X_1 + n_2 X_2 + r_2 \rangle \tag{3}$$

Rule III:

$$IF \ \langle X_1 \ is \ A_1 \ and \ X_2 \ is \ B_2 \rangle \ THEN \ \langle f_3 = m_3 X_1 + n_3 X_2 + r_3 \rangle \tag{4}$$

Rule IV:

$$IF \ \langle X_1 \ is \ A_2 \ and \ X_2 \ is \ B_1 \rangle \ THEN \ \langle f_4 = m_4 X_1 + n_4 X_2 + r_4 \rangle \tag{5}$$

where, X_1 and X_2 are input parameters, $A_{1,2}$ and $B_{1,2}$ are fuzzy sets for X_1 and X_2, and Y is the corresponding output. The statements in IF parts are known as the antecedent, and the conditions in THEN parts are called the consequence.

2.4. Particle swarm optimization (PSO)

This algorithm is one of the newest optimization algorithms introduced by Eberhart and Kennedy [35]. The origin of this algorithm is attributed to random motions of several animals such as bird flocking or fish schooling. The operation of PSO is similar to genetic algorithm (GA). PSO generates an initial

population of random solutions in which each solution or each member of this population has a random location and velocity in the multidimensional search domain.

The members of population which are solutions to the corresponding optimization problem move around in the search domain to find the optimum solution. PSO utilizes an iterative procedure for finding the optimum solution. Throughout this process, the velocities and locations of particles are continuously modified and updated in the search domain. The basis for this modification of locations and velocities is the previous best condition of the particle itself or its neighborhood (cognitive behavior) and either the best condition of total particles (social behavior). The operation of PSO can be explained mathematically as given below [36–38]:

$$v_i^{j+1} = w^i v_i^j + c_1 \alpha_1^j \left(Pb_i^j - x_i^j \right) + c_2 \alpha_2^j \left(Gb_i^j - x_i^j \right) \tag{6}$$

$$x_i^{j+1} = x_i^j + v_i^{j+1} \tag{7}$$

In above expressions, v_i^j and x_i^j are the velocity and the location of ith particle in jth epoch. Pb_i represents the best condition of each particle in the population, and Gb_i is the global best condition which exists in the entire population; w is a parameter named inertia weight, and $\alpha_{1,2}$ represents acceleration constants; $c_{1,2}$ are vectors that contain uniformly distributed random numbers.

The purpose of using inertia weight in Equation (**6** is to prevent over fitting of PSO at higher iterations. This term was used by Shi and Eberhart [39] for the first time. As the optimization procedure proceeds, the decrease in the value of w results in a shift from an exploratory search to an exploitative one, which provides faster convergence to the global minima. The decrease in value of w from $w_{initial}$ to w_{final} with epoch number is expressed by:

$$w^j = \frac{\left(j_{max} - j \right) \times \left(w_{initial} - w_{final} \right)}{j_{max}} + w_{final} \tag{8}$$

where, j_{max} represents the maximum number of performed epochs. Setting $w_{initial}$ at 0.9 and w_{final} at 0.4 usually provide a good and appropriate starting point [38].

3. Results and discussion

3.1. Data acquisition

In order to develop a reliable and effective model, accurate and dependable experimental data need to be utilized. In this paper, in order to evaluate the performance of the models developed for predicting denitrogenation and desulfurization of the model fuel, 420 data points, including 260 experimental kinetic adsorption data points and 160 experimental equilibrium adsorption data points were employed. The data points used were from two published works by the co-authors [2, 40]. The detailed explanation and representation of measurement of the data points are explained in these published works. The data points include kinetic and equilibrium adsorption data of a model fuel containing quinoline, carbazole, benzothiophene (BT), and dibenzothiophene (DBT) into five types of mesostructured adsorbents, namely aluminosilicate mesostructure (MSU-S), iron-, cerium-, chromium-, and copper-modified aluminosilicate mesostructures (Fe-MSU-S, Ce-MSU-S, Cr-MSU-S, and Cu-MSU-S). The input parameters of the developed models were the Brunauer-Emmett-Teller (BET) surface area of the

adsorbent (m^2/g), average pore diameter of the adsorbent (Å), molecular weight of compounds of the model fuel (quinoline, carbazole, BT, and DBT) (g/mol), initial concentration of compounds of the model fuel (ppm), and time (min). The adsorption of nitrogen- and sulfur-containing compounds of the model fuel (mg/g) was also the output parameter of the developed models. Table 1 and Table 2 show some details about the utilized data.

Table 1

Textural parameters of the adsorbents [2, 45]

Adsorbent	BET surface area (m^2/g)	Average pore diameter (Å)
MSU-S	677	24.29
Cr-MSU-S	365.32	30.55
Fe-MSU-S	400	29.097
Ce-MSU-S	406.33	20.61
Cu-MSU-S	391.24	27.29

Table 2

Detailed analysis of the utilized data [2, 45]

System	Molecular weight (g/mol)	Time (min)	Adsorption (mg/g)
Quinoline/adsorbent	129.16	10-600	5.36-51.57
Carbazole/adsorbent	167.206	10-600	6.74-26.94
BT/adsorbent	134.196	10-600	5.36-29.69
DBT/adsorbent	184.26	10-600	2.73-33.63

3.2. Model development

In the present work, MATLAB® 2014a was implemented to develop MLP-NN, PSO-RBF, and PSO-ANFIS models. Before the construction of the models, the data points were separated into two subsets of training and testing data on a random basis. For the training set, 80% of the data points is utilized, and the remaining 20% of the data points is used as the testing dataset. The detailed process of the development of different models is explained below.

a. MLP-ANN model

Despite this fact that an MLP-NN model can include one or several hidden layers in its structure, it was proved on a mathematical basis that any nonlinear function can be predicted by MLP-NNs where there is only one hidden layer in their structure [41]. Taking this fact into account and to reduce the computational time, in the current study, an MLP-NN model with only one hidden layer was utilized.

Since the modeling problem involves five input parameters and one output parameter, the constructed network has five neurons in the input layer and one neuron in its output layer. In order to characterize the MLP-NN model with the best accuracy, the number of located neurons in the hidden layer was changed from 4 to 25, and the performance of the network was monitored. The MSE value between the

target data and MLP-NN predictions for the testing, training, and the whole data points in each epoch (iteration) versus the number of neurons of the hidden layer is depicted in Figure 1. It is evident from this figure that the MLP-NN model with 23 neurons in its hidden layer provides the lowest value of MSE for all the data, meaning that this network provides the best performance compared to other tested networks.
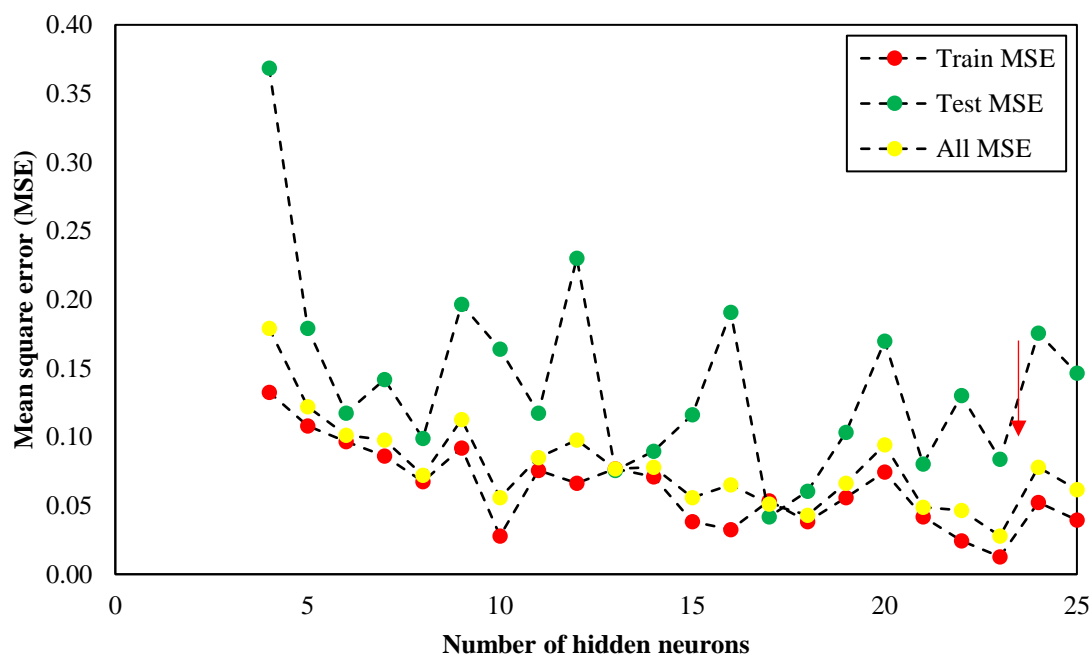

Figure 1

The change in the MSE of different MLP-NN networks against the number of hidden neurons.

b. PSO-RBF model

There are two parameters, namely the maximum number of neurons (MNNs) and spread, in the structure of RBF networks the value of which drastically influences the accuracy and performance of the RBF networks. Hence, to obtain an accurate and effective network for the estimation of the experimental data, the optimization ability of PSO was utilized to determine the optimum values of these parameters.

At the first stage, a population size of 50 random solutions was generated. In the next step, the members of the population were classified according to their cost function, which is the MSE value between each solution and the corresponding target values. Afterward, the PSO operators were utilized to construct a new population of better selected solutions. The optimum values of MNN and the spread obtained after 30 generations were 54 and 5 respectively. The monitored and recorded performance of the genetic algorithm during convergence to the optimum values is displayed in Figure 2.

c. PSO-ANFIS model

In order to develop the ANFIS model, initially a first order FIS was constructed by using the training data and genfis2 function of MATLAB. Then, the structure of the constructed FIS is trained by PSO method to achieve a better and more precise performance. This method assumes the problem of training the FIS structure as an optimization problem in which the tuning parameters of this optimization problem are the parameters of the corresponding MFs.

The performance of the PSO method during the training of the FIS is shown in Figure 3. It is clear that the total MSE value of PSO-ANFIS model is effectively lowered to values below the initial MSE of the ANFIS model, which confirms that the training of the ANFIS structure with the PSO method enhances the accuracy and effectiveness of this model.
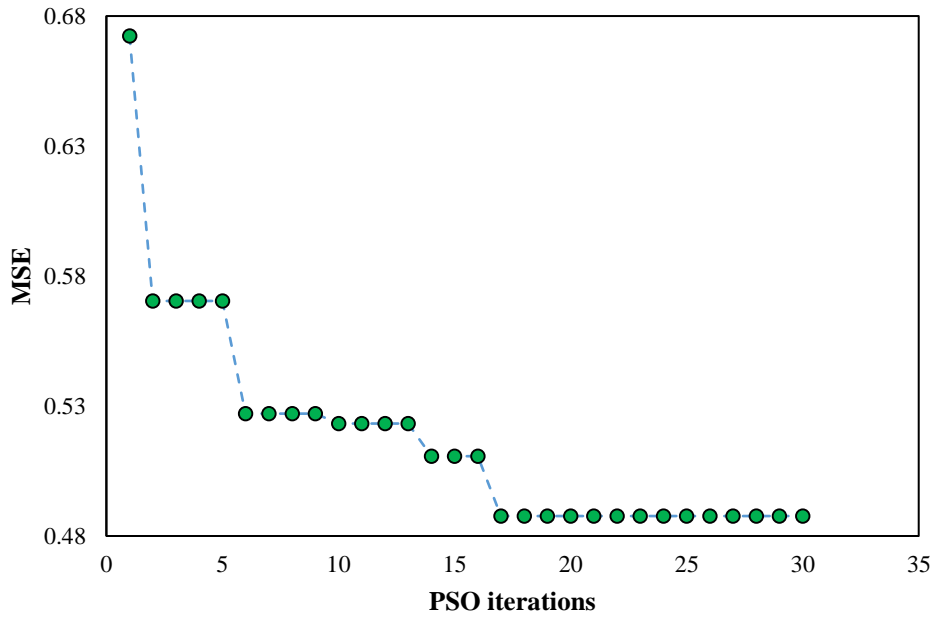

Figure 2

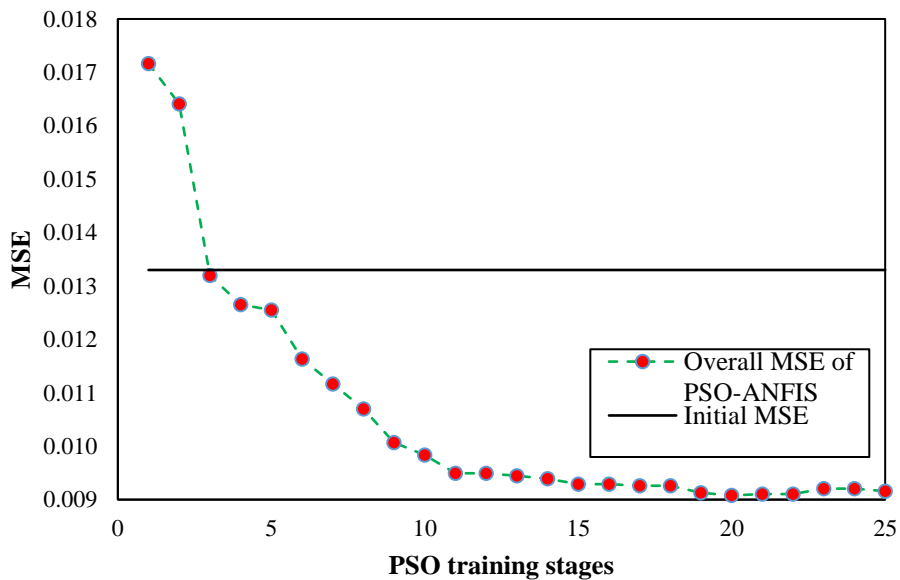Performance of PSO method during the optimization of the parameters of RBF-NN model.

Figure 3

Performance of the PSO method during the training of the ANFIS model.

Figure 4 represents the structure of the created PSO-ANFIS model. The final MFs after the completion of the training process are depicted in Figure 5. According to these figures, the developed PSO-ANFIS model uses eight rules and has eight membership functions for each input parameter.
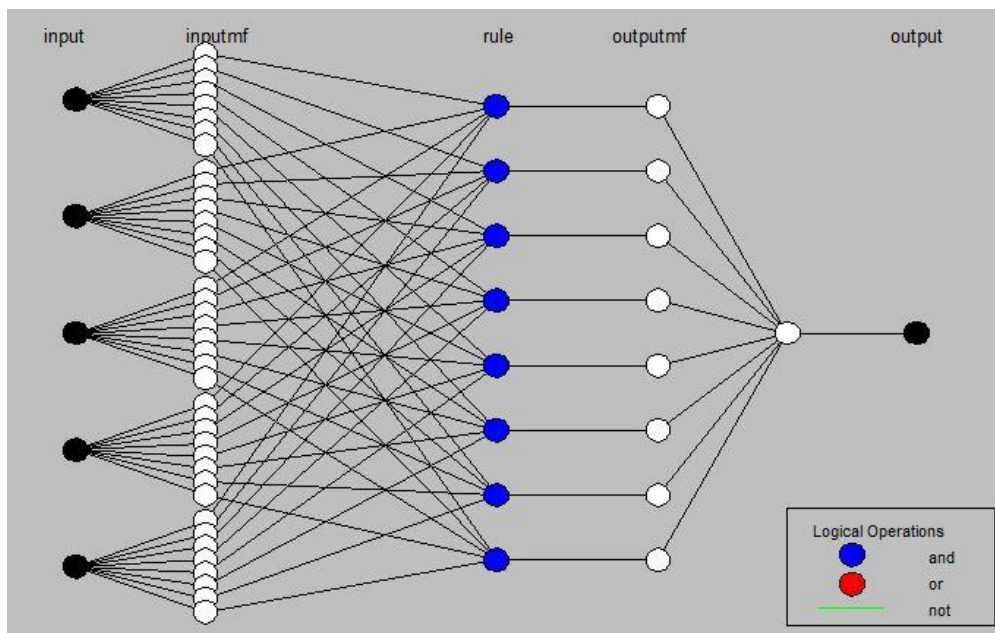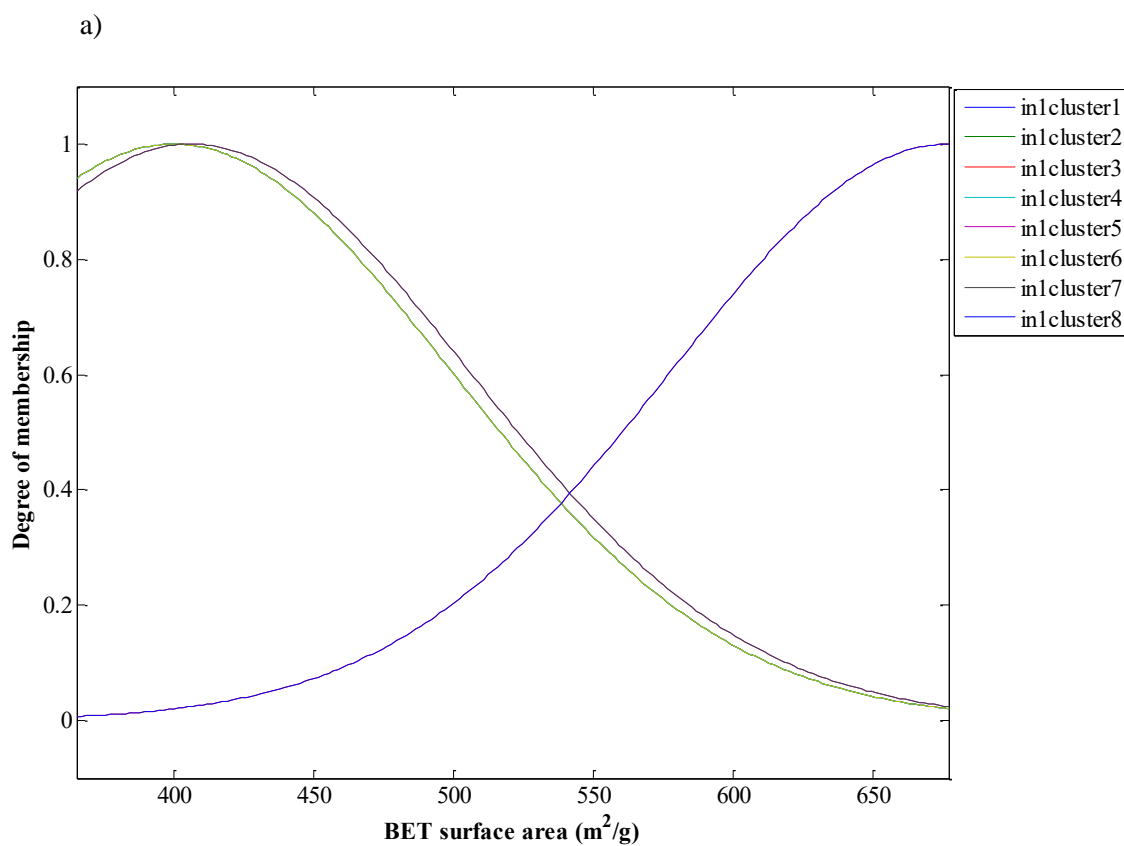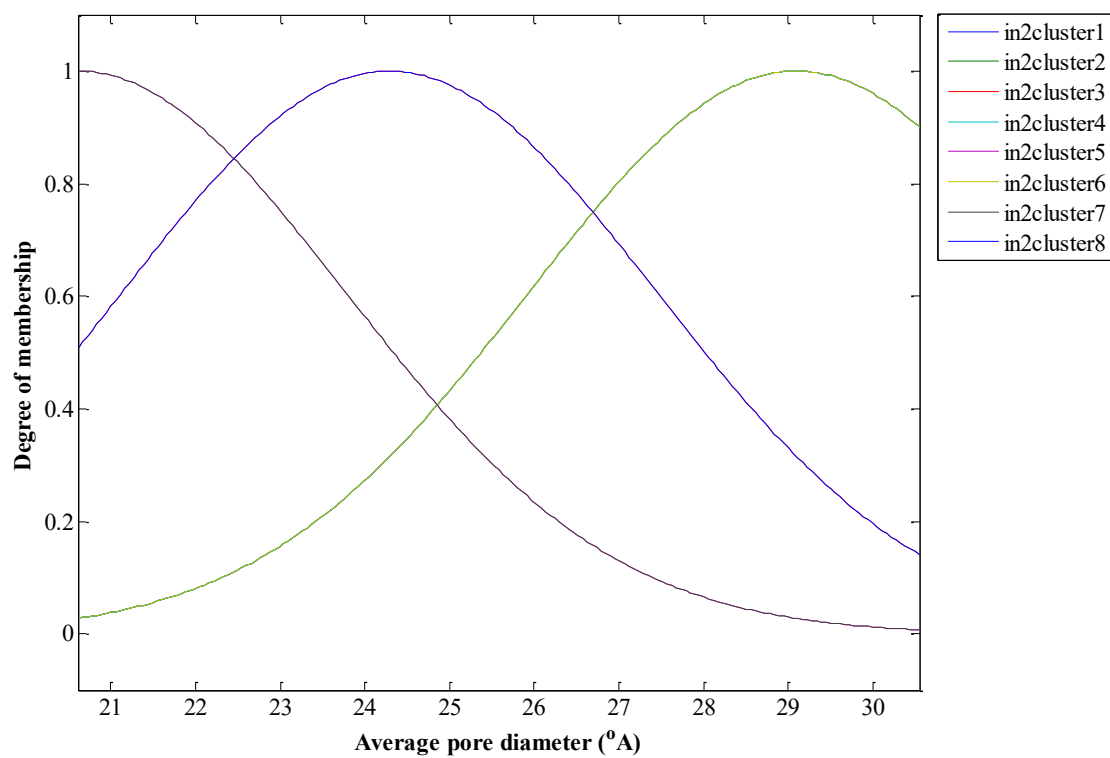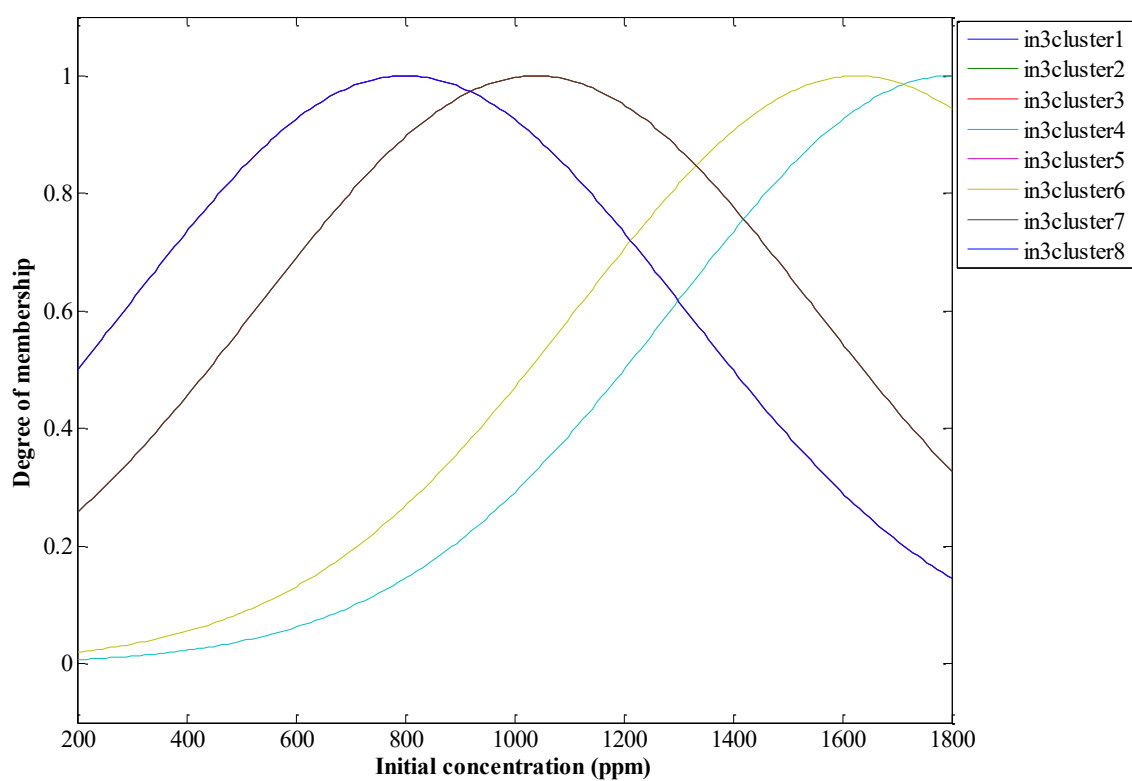
Figure 4

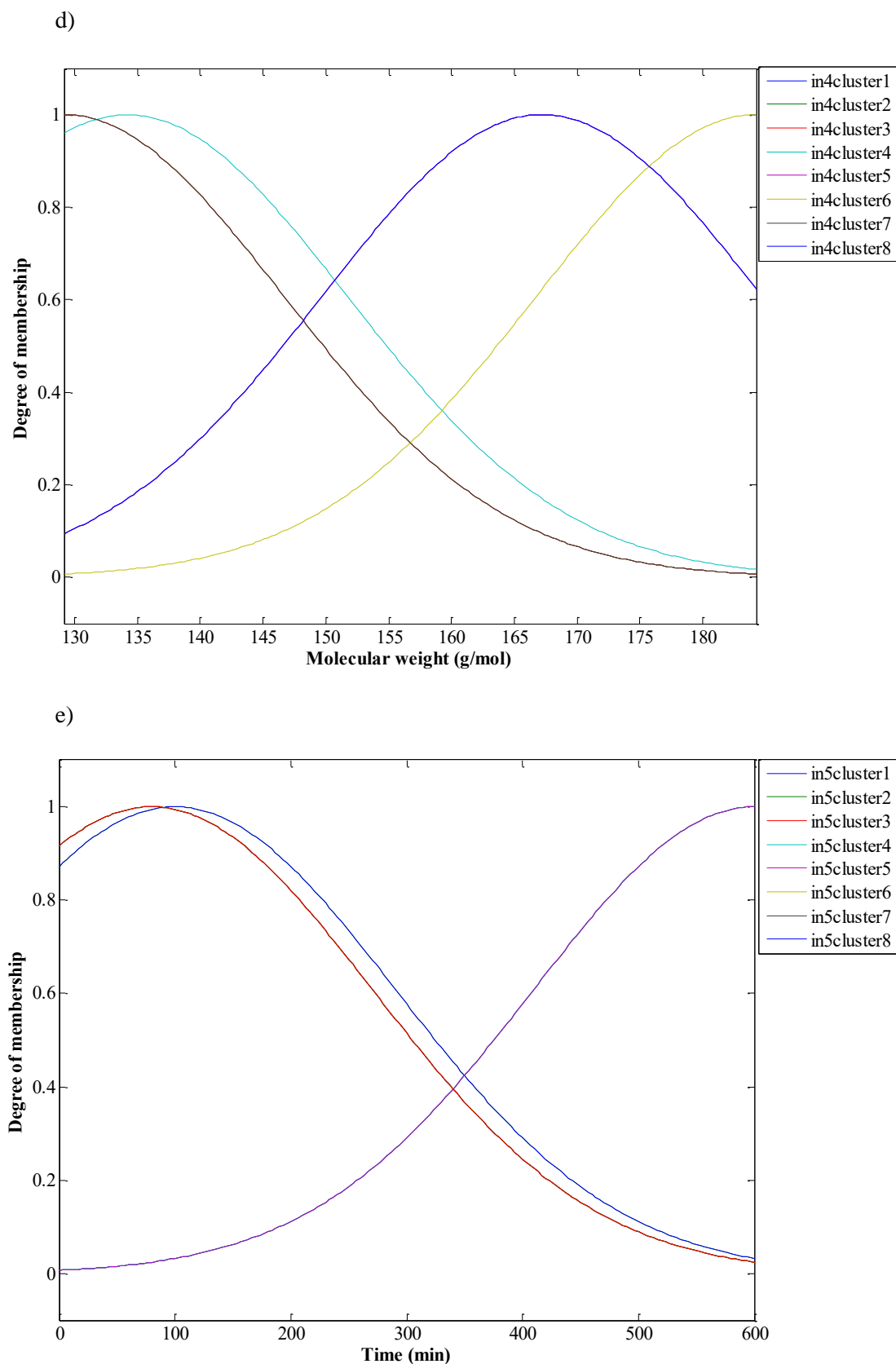The structure of the PSO-ANFIS model.

a)

b)

c)

d)

e)

Figure 5

The final MFs of the PSO-ANFIS model for different input parameters: a) best surface area; b) average pore diameter; c) Initial concentration; d) molecular weight; e) time.
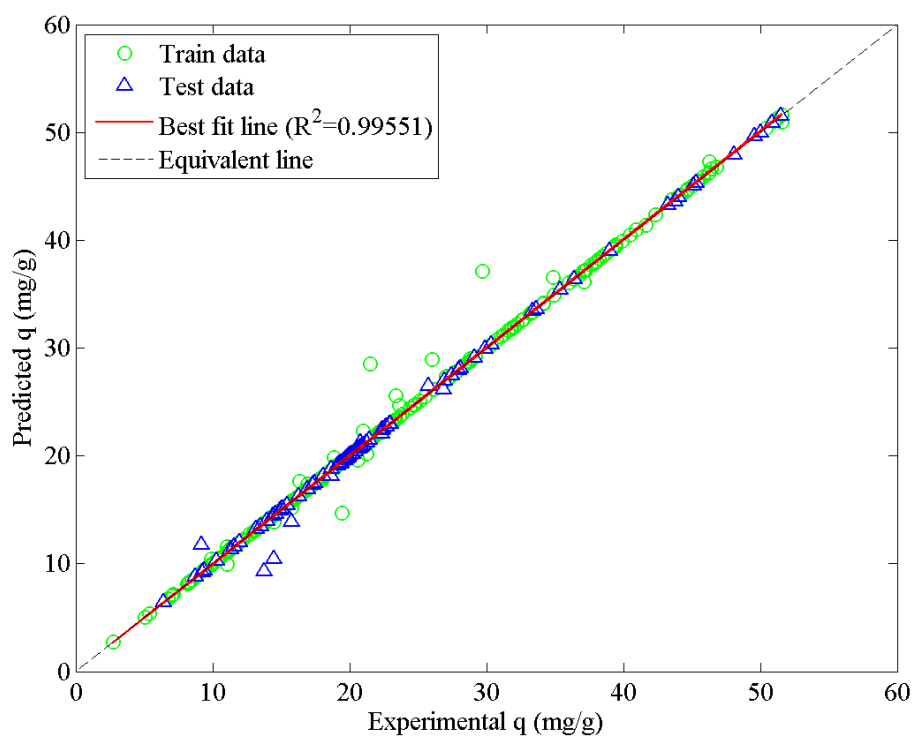
3.3. Evaluation of the developed models

The cross plot of the predicted adsorption data against the experimental data for three developed models is shown in Figure 6. This figure demonstrates that all the models are able to reproduce the experimental adsorption data accurately as is evident from the high-correlation coefficients (R^2) values of these models which are appropriately close to one. The R^2 values of the MLP-NN, PSO-RBF, and PSO-ANFIS models are 0.99967, 0.99551, and 0.99992 respectively, which confirms that although all the models have satisfactory performance, the accuracy of PSO-ANFIS model is better than that of the two other models.

The relative error plot of the predictions of the developed models is shown in Figure 7, which shows that the relative deviations of all the developed models are bounded in the span of ±10%. It is also clear that for the developed models, the relative errors are mostly concentrated near the zero-deviation line, which indicates the accuracy and the reliability of the proposed models. The deviations of the PSO-ANFIS model are less than those of the two other models, which further confirms the superiority of this model to the MLP-NN and PSO-RBF models.
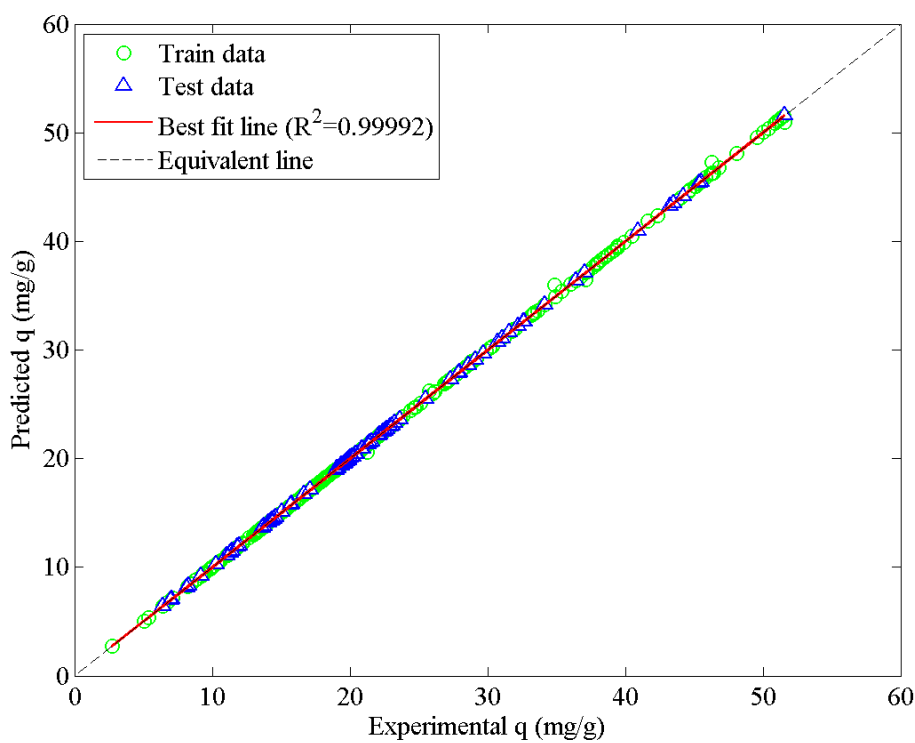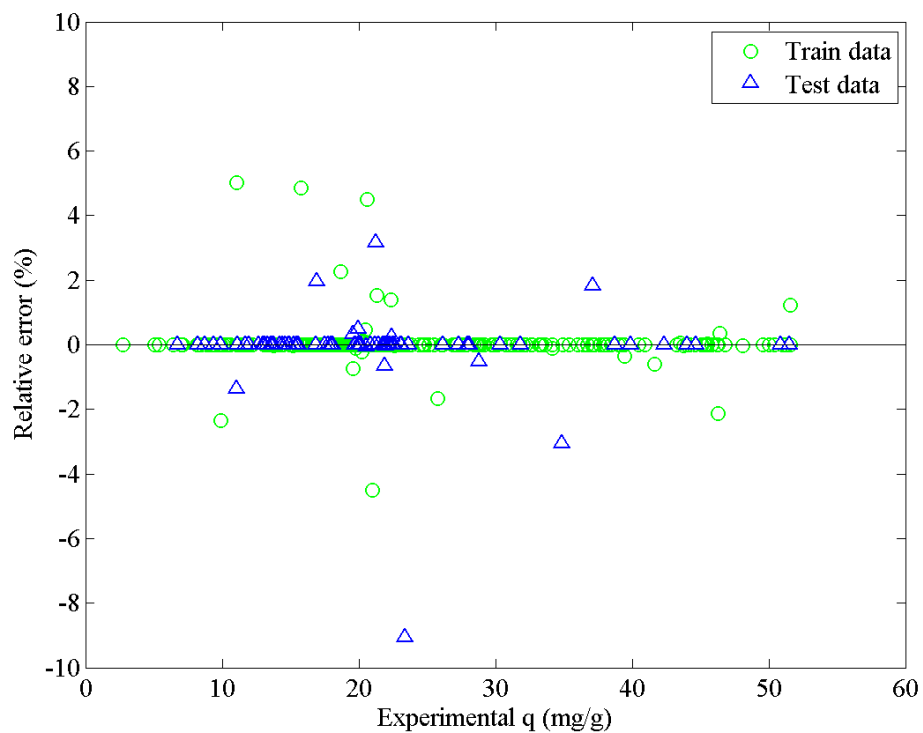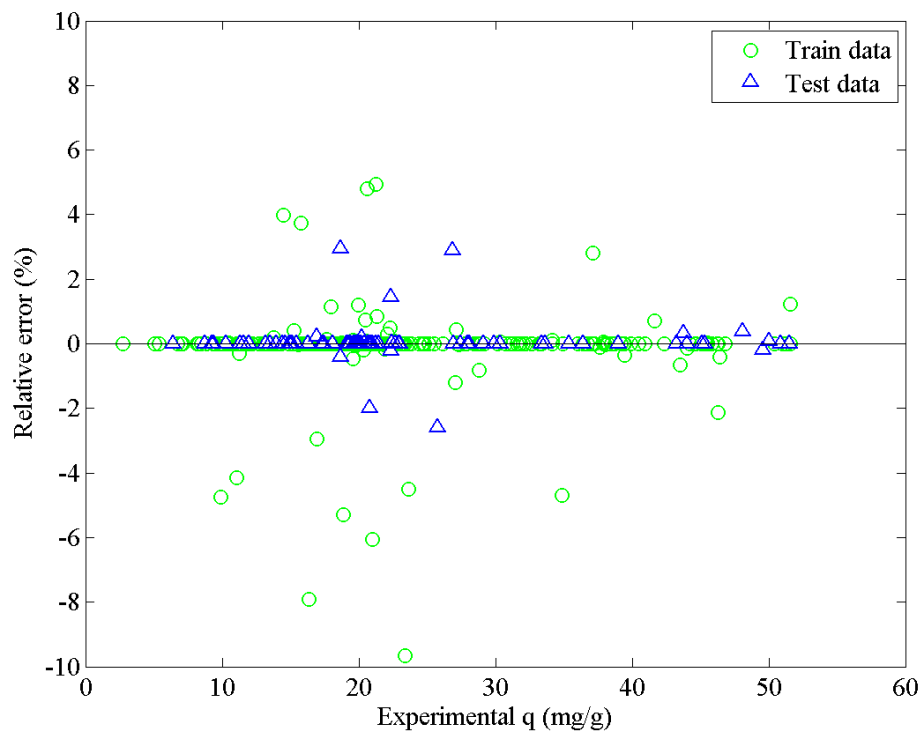
a)
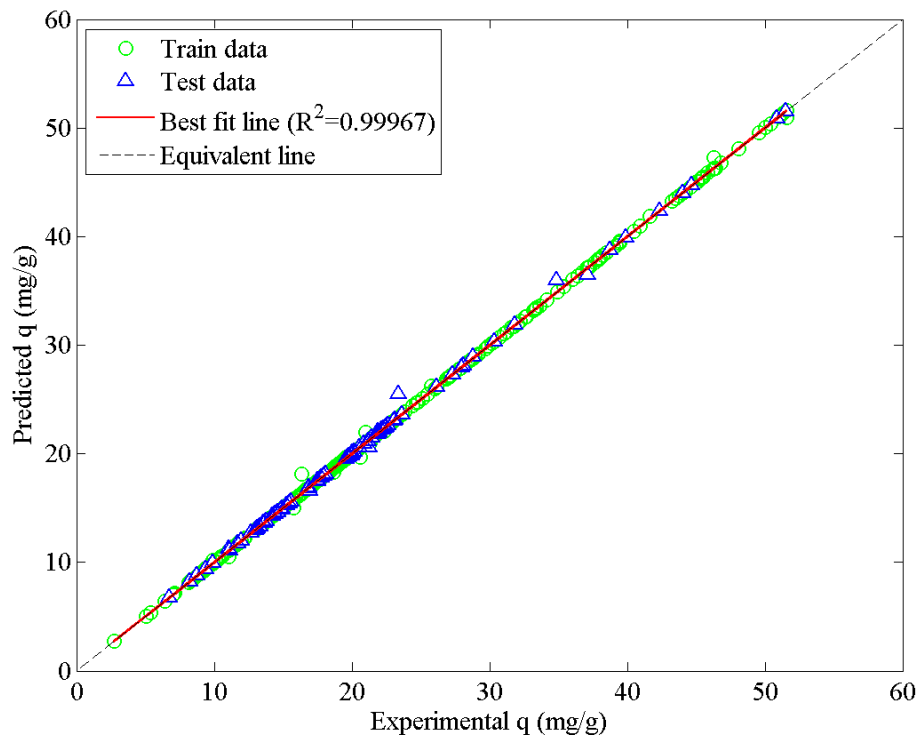
b)

c)

Figure 6

Cross plot of the predictions of the developed models against experimental adsorption data: a) MLP-NN; b) PSO-RBF; c) PSO-ANFIS
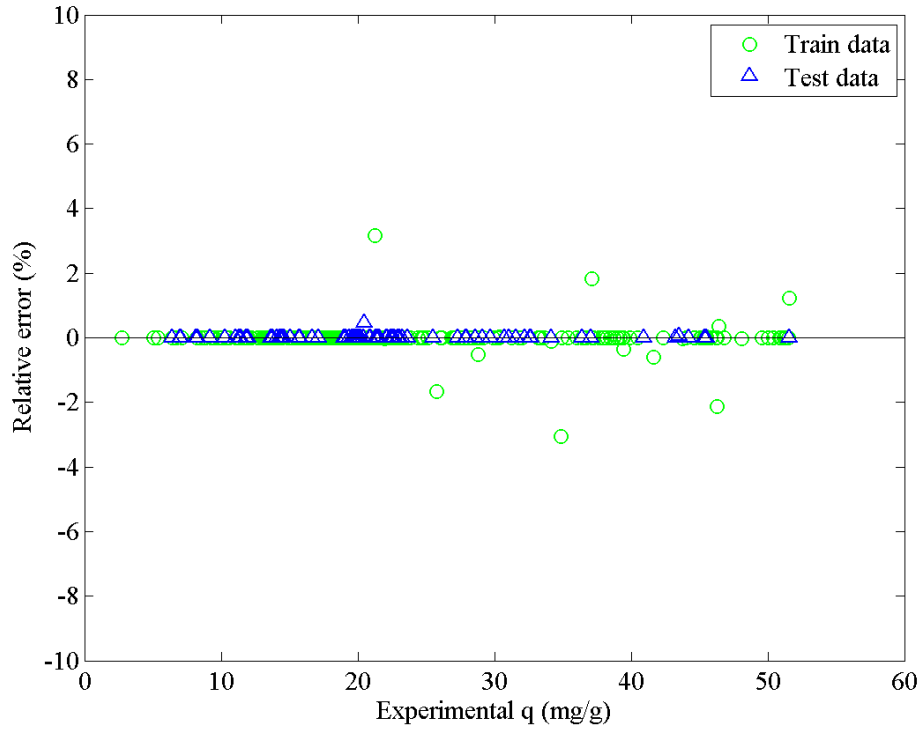
a)

b)

c)

Figure 7

Relative error plot of the predictions of the developed models against the experimental adsorption data: a) MLP-NN; b) PSO-RBF; c) PSO-ANFIS.

In order to further explore the accuracy of the developed models for the prediction of the experimental data, three statistical quality measure parameters, namely correlation coefficient (R^2), root mean square error (RMSE), and average absolute relative deviation (AARD%), were calculated for these models and are represented in Table 3. These parameters are defined below:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(q_{Pred}(i) - q_{Exp}(i)\right)^2}{\sum_{i=1}^{N}\left(q_{Pred}(i) - \overline{q}_{Exp}\right)^2} \tag{9}$$

$$AARD\% = \frac{100}{N}\sum_{i=1}^{N}\frac{\left(q_{Pred}(i) - q_{Exp}(i)\right)}{q_{Exp}(i)} \tag{10}$$

$$RMSE = \left(\frac{\sum_{i=1}^{N}(q_{Pred}(i) - q_{Exp}(i))^2}{N}\right)^{0.5} \tag{11}$$

where, $q_{Pred}(i)$ is the ith predicted adsorption value, $q_{Exp}(i)$ is the ith experimental adsorption value, and N is the number of data points. Table 3 shows that the developed models exhibit high R^2 values and low RMSE and AARD% values in training, testing, and the whole data points. According to this table, the PSO-ANFIS model exhibits better performance in comparison with the MLP-NN and PSO-RBF models. The accuracy of the MLP-NN model with AARD% and RMSE values of 0.171 and 0.192 respectively is also better than that of the PSO-RBF model, which gives AARD% and RMSE values of 0.759 and 0.714 respectively.

Table 3

Statistical parameters of the developed models.

Model	Data set	R^2	AARD%	RMSE
MLP-NN	Train data	0.9998	0.143	0.157
	Test data	0.9991	0.285	0.290
	Total data	0.9997	0.171	0.192
PSO-RBF	Train data	0.9956	0.592	0.698
	Test data	0.9954	1.428	0.774
	Total data	0.9955	0.759	0.714
PSO-ANFIS	Train data	0.99999	0.007	0.011
	Test data	0.99990	0.047	0.108
	Total data	0.99992	0.039	0.097

At this point the predictions of the developed models are compared with the literature data on equilibrium and kinetics. The predictions of the developed models for equilibrium adsorption data are compared with Langmuir [42] and Freundlich [43] adsorption models, and the estimation of the developed models for the kinetic adsorption data are compared with the pseudo first-order [44] and the pseudo second-order [44] models. R^2 values of the developed models, Langmuir model, and Freundlich model for different adsorbents are listed in Table 4. Similarly, R^2 values of these models for the kinetic data are presented in Table 5. Table 4 shows that the developed models exhibit acceptable levels of accuracy as they exhibit R^2 values higher than the Langmuir and Freundlich isotherms. Table 5 also confirms that the predictions of the developed models are better than those of the pseudo first-order and the pseudo second-order models as R^2 values of these models are higher compared to the pseudo first-order and the pseudo second-order models in different systems.

Table 4

Comparison between R^2 values of the developed models and the available equilibrium adsorption models.

| Adsorbent	Compounds	Langmuir			Freundlich			MLP-NN	PSO-RBF	PSO-ANFIS
		q_m (mg/g)	K_L (kg/mg)	R^2	$1/n$	K_F (mg$^{1-(1/n)}$. g^{-1}.kg$^{1/n}$)	R^2	R^2	R^2	R^2
MSU-S	Quinoline	67.56	0.00256	0.9991	0.6946	0.504	0.9927	0.9999	0.9999	0.9999
	Carbazole	36.23	0.00211	0.9967	0.6386	0.352	0.9921	0.9999	0.9999	0.9999
	Dibenzothiophene	19.6	0.00159	0.9933	0.472	0.454	0.9732	0.9999	0.9999	0.9999
	Benzothiophene	22.17	0.00132	0.9966	0.5037	0.381	0.9751	0.9999	0.9999	0.9999
Cr-MSU-S	Quinoline	50.251	0.007825	0.9887	0.5546	1.575	0.9988	0.9994	0.9999	0.9999
	Carbazole	33.557	0.003778	0.9829	0.5573	0.7095	0.9925	0.9997	0.9813	0.9999
	Dibenzothiophene	27.47	0.002135	0.9904	0.4225	0.9997	0.9455	0.9999	0.9994	0.9994
	Benzothiophene	28.98	0.00203	0.9821	0.4315	0.973	0.9442	0.9999	0.9996	0.9999
Fe-MSU-S	Quinoline	46.511	0.0203	0.9405	0.479	3.1556	0.9915	0.9999	0.9999	0.9999
	Carbazole	39.52	0.00303	0.9971	0.5984	0.5904	0.9972	0.9993	0.9956	0.9999
	Dibenzothiophene	24.39	0.00201	0.9946	0.4327	0.8106	0.9594	0.9934	0.9999	0.9999
	Benzothiophene	25.84	0.00199	0.9857	0.4484	0.766	0.9765	0.9999	0.9925	0.9999
Ce-MSU-S	Quinoline	29.32	0.00196	0.9959	0.4579	0.816	0.9727	0.9999	0.9999	0.9999
	Carbazole	32.79	0.00162	0.9949	0.5089	0.599	0.9852	0.9995	0.9998	0.9999
	Dibenzothiophene	50.76	0.00415	0.9721	0.6368	0.754	0.9963	0.9908	0.9731	0.9999
	Benzothiophene	36.50	0.00250	0.9986	0.5280	0.752	0.9931	0.9999	0.9896	0.9998
Cu-MSU-S	Quinoline	56.82	0.00134	0.9988	0.5787	0.602	0.9816	0.9999	0.9999	0.9999
	Carbazole	48.10	0.00136	0.9971	0.5422	0.645	0.9637	0.9999	0.9999	0.9999
	Dibenzothiophene	49.75	0.00836	0.9569	0.5781	1.448	0.9947	0.9999	0.9999	0.9999
	Benzothiophene	36.76	0.00316	0.9962	0.5823	0.616	0.9883	0.9999	0.9999	0.9999

Table 5

Comparison between R^2 values of the developed models and the available kinetic adsorption models.

| Adsorbent	Compounds	Pseudo-first order			Pseudo-second order			MLP-NN	PSO-RBF	PSO-ANFIS
		q_e (mg/g)	k_{1ad} (1/min)	R^2	q_e (mg/g)	k_{2ad} (g/mg.min)	R^2	R^2	R^2	R^2
MSU-S	Quinoline	35.58	0.037	0.9764	39.682	0.00209	0.9951	0.9999	0.9999	0.9999
	Carbazole	8.658	0.027	0.8876	21.097	0.00419	0.9943	0.9999	0.9999	0.9999
	Dibenzothiophene	15.16	0.032	0.9659	15.923	0.00225	0.9791	0.9999	0.9999	0.9999
	Benzothiophene	11.6	0.026	0.9109	16.181	0.00462	0.9946	0.9999	0.9999	0.9999
Cr-MSU-S	Quinoline	31.06	0.0382	0.9842	47.39	0.00308	0.9983	0.9999	0.9999	0.9999
	Carbazole	11.06	0.0259	0.8933	22.83	0.00648	0.9984	0.9999	0.9843	0.9999
	Dibenzothiophene	9.452	0.0267	0.945	20.576	0.00651	0.9981	0.9973	0.9937	0.9973
	Benzothiophene	12.14	0.0296	0.958	22.8	0.0065	0.9987	0.9987	0.9987	0.9999
Fe-MSU-S	Quinoline	18.11	0.0314	0.9062	52.91	0.0039	0.9987	0.9999	0.9999	0.9999
	Carbazole	8.12	0.0265	0.8556	23.98	0.0075	0.9984	0.9999	0.9817	0.9999
	Dibenzothiophene	4.87	0.0269	0.8381	18.28	0.00741	0.9973	0.9999	0.9997	0.9999
	Benzothiophene	10.04	0.0297	0.9142	20.08	0.0074	0.9984	0.9925	0.9919	0.9999
Ce-MSU-S	Quinoline	10.04	0.0297	0.9142	21.5	0.00792	0.9984	0.9999	0.9999	0.9999
	Carbazole	7.78	0.0224	0.7795	23.81	0.01144	0.9994	0.9999	0.9751	0.9999
	Dibenzothiophene	15.27	0.0311	0.8812	40.485	0.00501	0.9988	0.9990	0.9976	0.9990
	Benzothiophene	6.94	0.0283	0.87	20.876	0.01073	0.9943	0.9721	0.9735	0.9999
Cu-MSU-S	Quinoline	12.606	0.0244	0.8015	33.89	0.0052	0.9987	0.9999	0.9999	0.9999
	Carbazole	10.558	0.0252	0.7872	29.41	0.0063	0.9986	0.9999	0.9999	0.9999
	Dibenzothiophene	16.6	0.0273	0.8133	47.39	0.0039	0.9984	0.9999	0.9999	0.9999
	Benzothiophene	7.135	0.0186	0.8443	22.62	0.0072	0.9986	0.9999	0.9999	0.9999

The plot of the experimental equilibrium adsorption data versus the initial concentration of adsorbent and the plot of the experimental kinetic adsorption data versus the time for the different models are illustrated in Figure 8 and Figure 9 respectively to provide a better visual comparison between these models. These figures also show that the developed models are better and more accurate than the literature models for capturing the trend of the experimental equilibrium and kinetic adsorption data in different systems. Another advantage of the developed models over other models is that the models proposed in the present work are able to predict the experimental equilibrium and kinetic data for all systems with a single trained structure and with a satisfactory degree of accuracy. However, the use of Langmuir and Freundlich models for equilibrium data and also the use of the pseudo first-order and the pseudo second-order model for kinetic data require fitting the related parameters of each model to the

experimental data for each system separately. Therefore, thanks to higher R^2 values and lower AARD% and RMSE, the models developed herein exhibit a higher generalization ability and better accuracy compared to the models available in the literature.
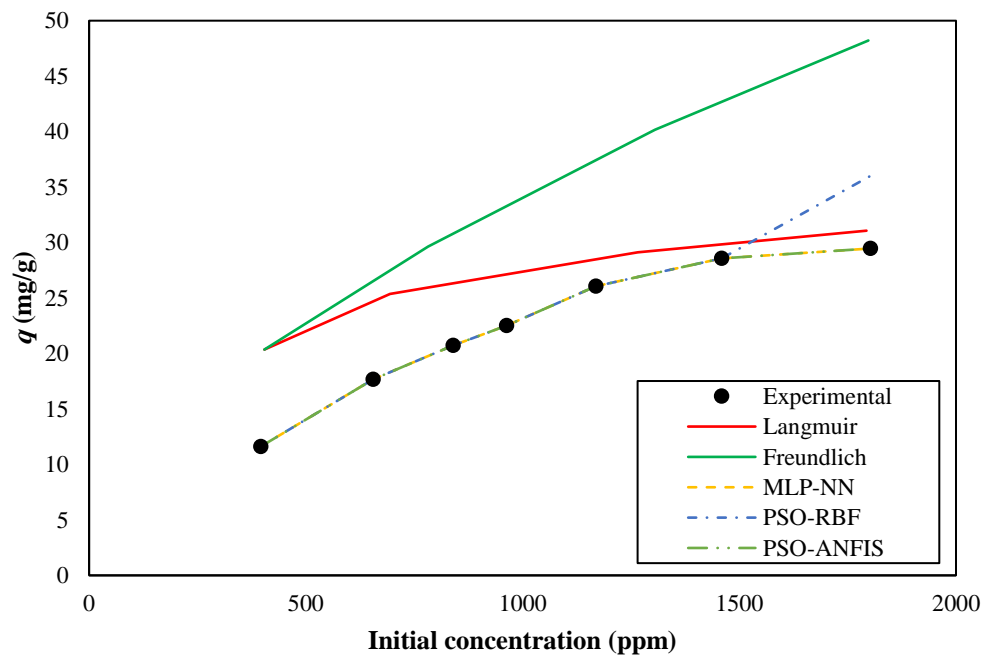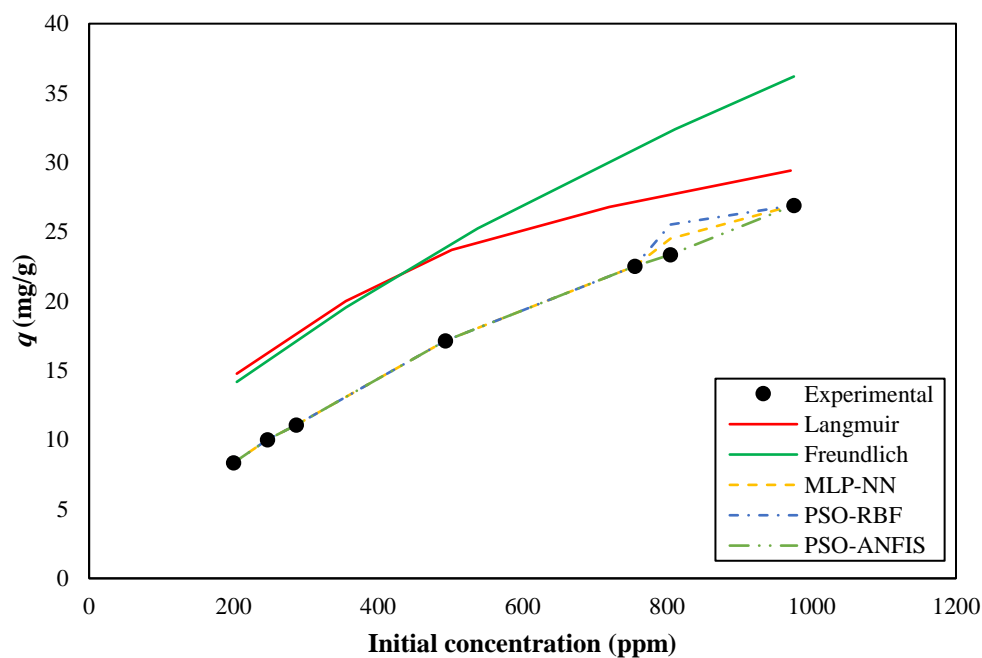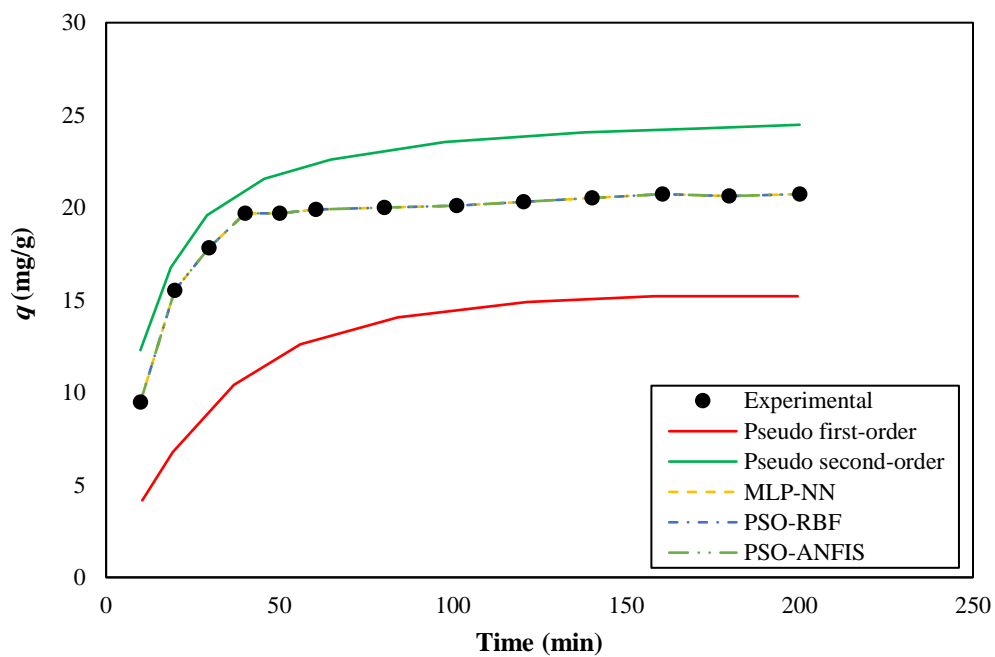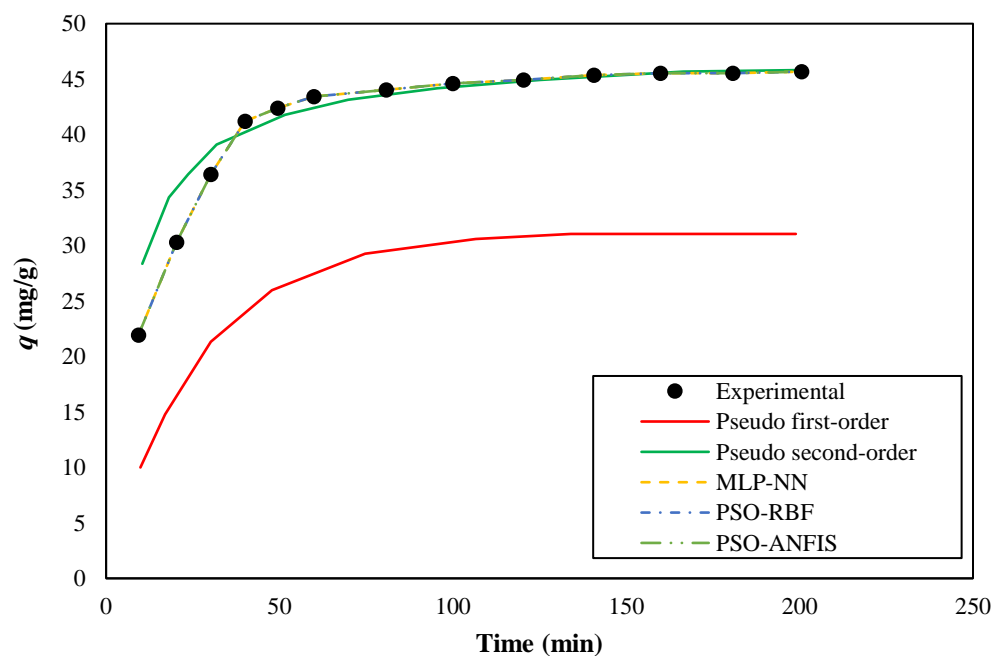
a)
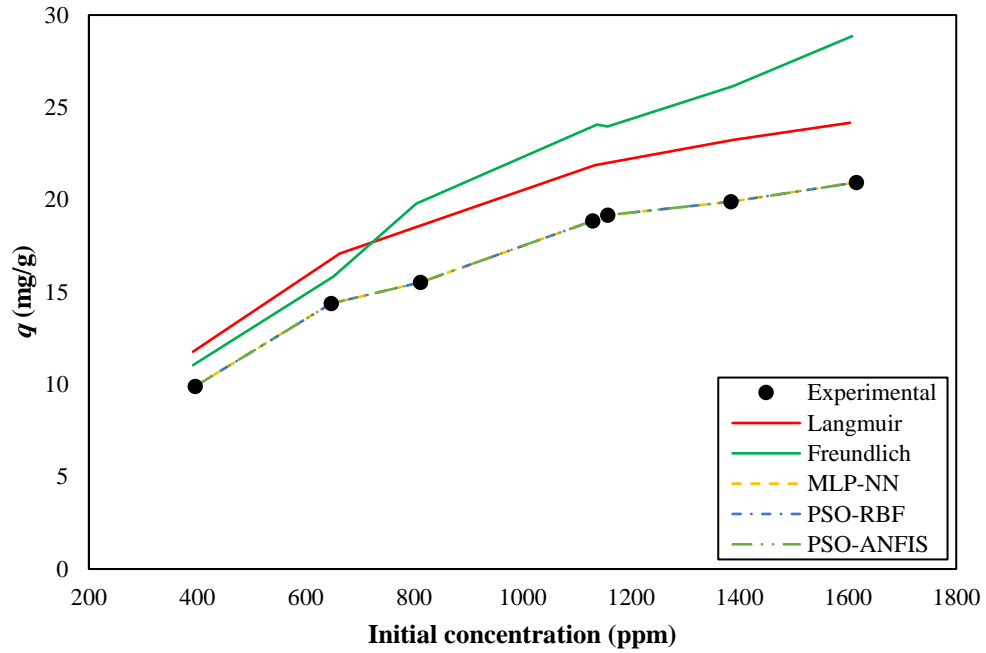
b)
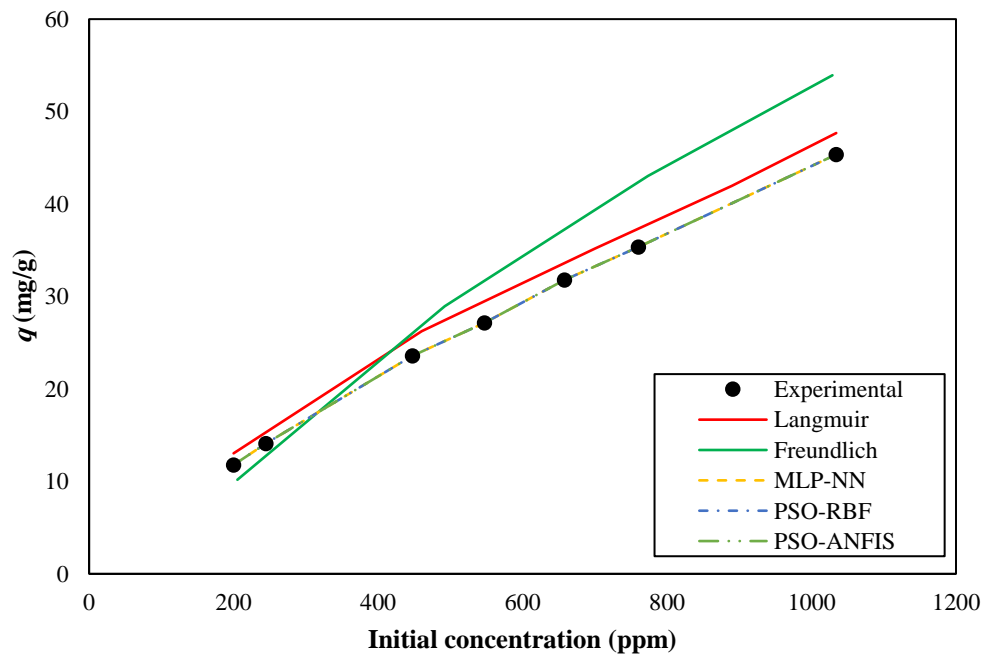
c)

d)

Figure 8

Comparison between predictions of the developed models and literature equilibrium models for different systems:
a) Ce-MSU-S/DBT system; b) Cr-MSU-S/quinolone; c) Cu-MSU-S/BT; d) Fe-MSU-S/Carbazole.
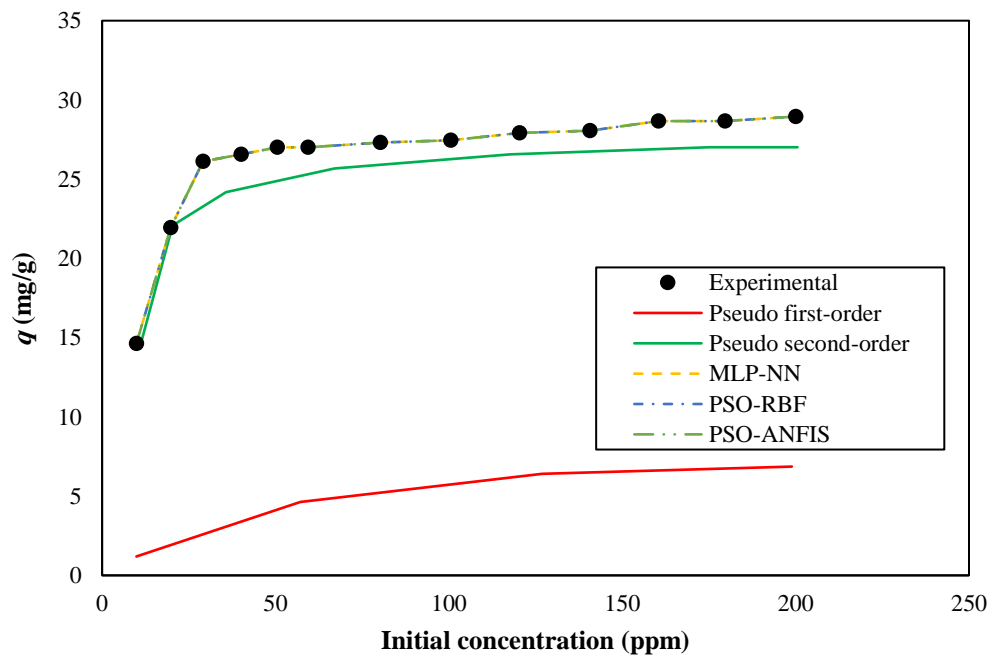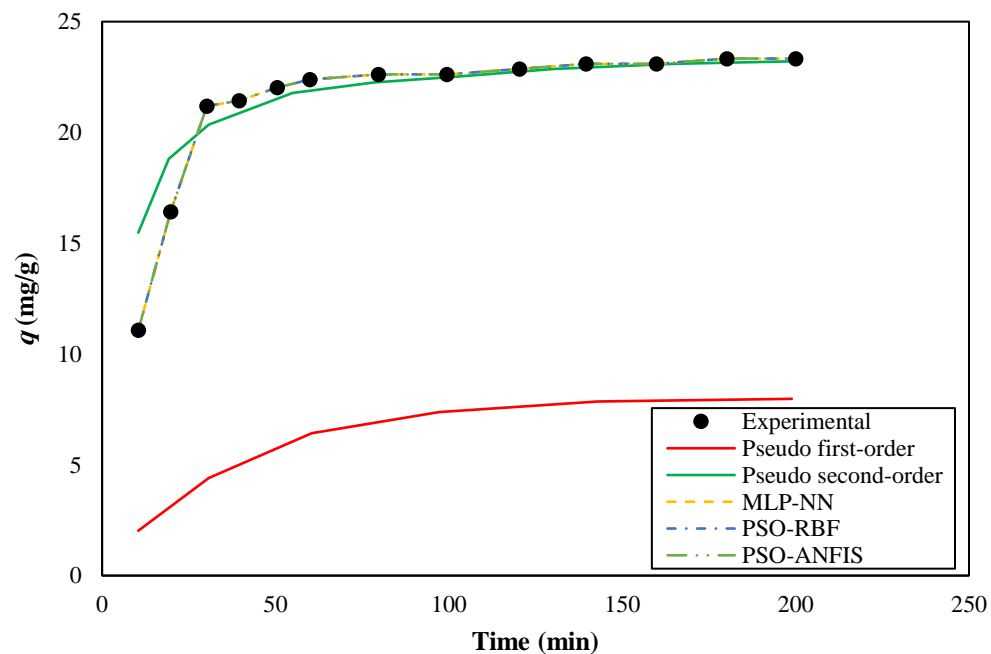
a)

b)

c)

d)

Figure 9

Comparison between predictions of the developed models and literature kinetic models for different systems: a) Ce-MSU-S/DBT system; b) Cr-MSU-S/quinolone; c) Cu-MSU-S/BT; d) Fe-MSU-S/Carbazole.

4. Conclusions

In this study, three intelligent models, namely PSO-RBF, MLP-NN, and PSO-ANFIS, were implemented to predict the denitrogenation and desulfurization of a model fuel in the presence of mesostructured adsorbents. The predictions of the developed models were evaluated through various approaches, and it was observed that all the models represented accurate and reliable predictions. R^2, AARD%, and RMSE values of the developed models were 0.99992, 0.039, and 0.097 respectively for PSO-ANFIS; 0.9997, 0.171, and 0.29 respectively for MLP-NN; and 0.9955, 0.759, and 0.714 respectively for PSO-RBF. However, the PSO-ANFIS model exhibited better results compared to the two other models. The predictions of the developed models were also compared with the literature kinetic and the equilibrium adsorption models, and it was observed that the developed models presented better predictions. It should be mentioned that each of the models developed in the present work exhibited some shortcomings. The main limitation of these models is that they are black box type models in nature and they could either underpredict or overpredict the experimental values which are outside their applicability domain. Another limitation is that in order to develop an accurate and dependable model, sufficient experimental data are needed for the training of these models. However, in the case of the availability of data, it is possible to develop accurate models for the prediction of experimental values by applying these techniques. The current models can be employed in areas where the accurate and rapid prediction of the desulfurization and denitrogenation of model fuels by mesoporous materials is required. Moreover, developing computing models for the prediction of desulfurization and denitrogenation of different materials in the presence of various adsorbents can be investigated to implement more general models in future works.

Nomenclature

Abbreviations	
AARD%	Average absolute relative deviation
FIS	Fuzzy inference system
MF	Membership function
MLP-NN	Multilayer perceptron neural network
MSE	Mean square error
N	Number of data points
RBF	Radial basis function
RMSE	Root mean square error
Variables	
$\alpha_{1,2}$	Acceleration constants
$c_{1,2}$	Vectors contain random numbers
Gb_i	Global best condition
n	Constant of Kendall-Monroe model
I	Position of an input pattern in RBF model
Inp_k	kth input parameter
$Inp_{k,i}$	ith component of kth input parameter
$\phi(r)$	Nonlinear activation function
Pb_i	Best condition of each particle in population

$q_{Exp}(i)$	ith experimental adsorption value
$q_{Pred}(i)$	ith predicted adsorption value
σ	Parameter that controls the smoothness in activation function
r	Euclidian distance in RBF model
R^2	Correlation factor
v_j	Center of jth cluster in RBF model
v_i^j	Velocity of ith particle *in j*th epoch
w	Inertia weight
w_p	Weights vector
x_i^j	Location of ith particle in jth epoch

References

Ahmadi, M., Anvaripour, B., Khosravi-Nikou, M. R., and Mohammadian, M., Selective Denitrogenation of Model Fuel Through Iron- and Chromium-Modified Microporous Materials (MSU-S), Journal of Environmental Chemical Engineering, Vol. 5, p. 849–860, 2017.

Ahmadi, M. A. and Shadizadeh, S. R., New Approach for Prediction of Asphaltene Precipitation Due to Natural Depletion by Using Evolutionary Algorithm Concept, Fuel, Vol. 102, p. 716–723, 2012.

Chen, H., Wang, Y., Yang, F. H., and Yang, R. T., Desulfurization of High-sulfur Jet Fuel by Mesoporous Π-Complexation Adsorbents, Chemical Engineering Science, Vol. 64, p. 5240–5246, 2009.

Coulibaly, P. and Baldwin, C. K., Nonstationary Hydrological Time Series Forecasting Using Nonlinear Dynamic Methods, Journal of Hydrology, Vol. 307, p. 164–174, 2005.

Cybenko, G., Approximation by Superpositions of a Sigmoidal Function, Math, Control Signal Systems, Vol. 2, p. 303–314, 1989.

Esmaeili-Jaghdan, Z., Shariati, A., and Nikou, M. R. K., A Hybrid Smart Modeling Approach for Estimation of Pure Ionic Liquids Viscosity, Journal of Molecular Liquids, Vol. 222, p. 14–27, 2016.

Freundlich, H., Ueber Die Adsorption in Loesungen, Zeitschrift für Physikalische Chemie, Vol. 57, p. 385–470, 1907.

Gargiulo, N., Peluso, A., Aprea, P., Pepe, F., and Caputo, D., CO_2 Adsorption on Polyethylenimine-Functionalized SBA-15 Mesoporous Silica: Isotherms and Modeling, Journal of Chemical and Engineering Data, Vol. 59, p. 896–902, 2014.

Ghiasi, M. M., Arabloo, M., Mohammadi, A. H., and Barghi, T., Application of ANFIS Soft Computing Technique in Modeling the CO_2 Capture With MEA, DEA, and TEA Aqueous Solutions, International Journal of Greenhouse Gas Control, Vol. 49, p. 47–54, 2016.

Halali, M. A., Azari, V., Arabloo, M., Mohammadi, A. H., and Bahadori, A., Application of a Radial Basis Function Neural Network to Estimate Pressure Gradient in Water–oil Pipelines, Journal of the Taiwan Institute of Chemical Engineers, Vol. 58, p. 189–202, 2012.

Heidari, E., Sobati, M. A., and Movahedirad, S., Accurate Prediction of Nanofluid Viscosity Using a Multilayer Perceptron Artificial Neural Network (MLP-ANN), Chemometrics and Intelligent Laboratory Systems, Vol. 155, p. 73–85, 2016.

Kim, J. H., Ma, X., Zhou, A., and Song, C., Ultra-deep Desulfurization and Denitrogenation of Diesel Fuel by Selective Adsorption Over Three Different Adsorbents: A Study on Adsorptive Selectivity and Mechanism, Catalysis Today, Vol. 111, p. 74–83, 2006.

Koriakin, A., Ponvel, K. M., and Lee, C. H., Denitrogenation of Raw Diesel Fuel by Lithium-modified Mesoporous Silica, Chemical Engineering Journal, Vol. 162, p. 649–655, 2010.

Lagergren, S., Zur Theorie Der Sogenannten Adsorption Geloster Stoffe, Kungliga Svenska Vetenskapsakademiens, Handlingar, Vol. 24, p. 1–39, 1898.

Langmuir, I., The Constitution and Fundamental Properties of Solids and Liquids, Part I: Solids, Journal of the American Chemical Society, Vol. 38, p. 2221–2295, 1916.

Mello, M., Eić, M., Adsorption of Sulfur Dioxide From Pseudo Binary Mixtures on Hydrophobic Zeolites: Modeling of The Breakthrough Curves, Adsorption, Vol. 8, p. 279–289, 2015.

Mohammadian, M., Ahmadi, M., and Khosravi-Nikou, M. R., Adsorptive Desulfurization and Denitrogenation of Model Fuel by Mesoporous Adsorbents (MSU-S and Coo-MSU-S), Petroleum Science and Technology, Vol. 35, p. 608–614, 2017.

Mohammadian, M., Khosravi-Nikou, M. R., Shariati, A., and Aghajani, M., Model Fuel Desulfurization and Denitrogenation Using Copper and Cerium Modified Mesoporous Material (MSU-S) Through Adsorption Process, Clean Technologies and Environmental Policy, Vol. 20, p. 95–112, 2018.

Montazerolghaem, M., Rahimi, A., and Seyedeyn-azad, F., Equilibrium and Kinetic Modeling of Adsorptive Sulfur Removal from Gasoline by Synthesized Ce–Y Zeolite, Applied Surface Science, Vol. 257, p. 603–609, 2010.

Nasery, S., Barati-Harooni, A., Tatar, A., Najafi-Marghmaleki, A., and Mohammadi, A. H., Accurate Prediction of Solubility of Hydrogen in Heavy Oil Fractions, Journal of Molecular Liquids, Vol. 222, p. 933–943, 2016.

Najafi-Marghmaleki, A., Khosravi-Nikou, M. R., and Barati-Harooni, A., A New Model for Prediction of Binary Mixture of Ionic Liquids+ water Density Using Artificial Neural Network, Journal of Molecular Liquids, Vol. 220, p. 232–237, 2016.

Pawelec, B., Navarro, R. M., Campos, J. M., and Fierro, J. L., Towards Near Zero-sulfur Liquid Fuels: a Perspective Review, Catalysis Science and Technology, Vol. 3, p. 3376–3376, 2013.

Rashidi, S., Nikou, M. R. K., and Anvaripour, B., Adsorptive Desulfurization and Denitrogenation of Model Fuel Using HPW and Nio-HPW Modified Aluminosilicate Mesostructures, Microporous and Mesoporous Materials, Vol.211, p. 134–141. 2015.

Sabzevari, S. and Moosavi, M., Density Prediction of Liquid Alkali Metals and Their Mixtures Using an Artificial Neural Network Method Over the Whole Liquid Range, Fluid Phase Equilibria, Vol. 361, p. 135–142, 2014.

Santos, A. L., Reis, R. A., Rossa, V., Reis, M. M., Costa, A. L., Veloso, C. O., Henriques, C. A., Zotin, F. M., Paredes, M. L., and Silveira, E. B., Silica–alumina Impregnated with Cerium, Nickel, and

Molybdenum Oxides for Adsorption of Sulfur and Nitrogen Compounds from Diesel, Materials Letters, Vol. 83, p. 158–160, 2013.

Safari, H., Nekoeian, S., Shirdel, M. R., Ahmadi, H., Bahadori, A., and Zendehboudi, S., Assessing the Dynamic Viscosity of Na–K–Ca–Cl–H_2O Aqueous Solutions at High-Pressure and High-Temperature Conditions, Industrial and Engineering Chemistry Research, Vol. 53, p. 11488–11500, 2013.

Sarda, K., Bhandari, A., Pant, K., and Jain, S., Deep Desulfurization of Diesel Fuel by Selective Adsorption Over Ni/Al_2O_3 and Ni/ZSM-5 Extrudates, Fuel, Vol. 93, p. 86–91, 2012.

Srivastava, V. C., An Evaluation of Desulfurization Technologies for Sulfur Removal From Liquid Fuels, Rsc Advances, Vol. 2, p. 759–783, 2012.

Sun, B., Li, G. and Wang, X., Facile Synthesis of Microporous Carbon Through a Soft-template Pathway and Its Performance in Desulfurization and Denitrogenation, Journal of Natural Gas Chemistry, Vol. 19, p. 471–476, 2010.

Tatar, A., Barati-Harooni, A., Najafi-Marghmaleki, A., Mohebbi, A., Ghiasi, M. M., Mohammadi, A. H., and Hajinezhad, A., Comparison of Two Soft Computing Approaches for Predicting CO_2 Solubility in Aqueous Solution of Piperazine, International Journal of Greenhouse Gas Control, Vol. 53, p. 85–97, 2016.

Wang, L., Sun, B., Yang, F. H., and Yang, R. T., Effects of Aromatics on Desulfurization of Liquid Fuel By Π-Complexation and Carbon Adsorbents, Chemical Engineering Science, Vol. 73, p. 208–217, 2012.

Wen, J., Han, X., Lin, H., Zheng, Y., and Chu, W., A Critical Study on the Adsorption of Heterocyclic Sulfur and Nitrogen Compounds by Activated Carbon: Equilibrium, Kinetics and Thermodynamics, Chemical Engineering Journal, Vol. 164, p. 29–36, 2010.

Xia, C., Wang, J., and Mcmenemy, K., Short, Medium- and Long-term Load Forecasting Model and Virtual Load Forecaster Based on Radial Basis Function Neural Networks, International Journal of Electrical Power and Energy Systems, Vol. 32, p. 743–750, 2010.

Xu, X., Zhang, S., Li, P., and Shen, Y., Desulfurization of Jet-A Fuel in a Fixed-bed Reactor at Room Temperature and Ambient Pressure Using A Novel Selective Adsorbent, Fuel, Vol. 117, p. 499–508, 2014.

Xu, X., Zhang, S., Li, P., and Shen, Y., Equilibrium and Kinetics of Jet-A Fuel Desulfurization by Selective Adsorption at Room Temperatures, Fuel, Vol. 111, p. 172–179, 2013.

Zadeh, L. A., Fuzzy Sets, Information and Control, Vol. 8, p. 338–353, 1965.

Zamani, H. A., Rafiee-Taghanaki, S., Karimi, M., Arabloo, M., and Dadashi, A., Implementing ANFIS for Prediction of Reservoir Oil Solution Gas-oil Ratio, Journal of Natural Gas Science and Engineering, Vol. 25, p. 325–334, 2015.