

# Efficient estimation of hydrolyzed polyacrylamide (HPAM) solution viscosity for enhanced oil recovery process by polymer flooding

Alireza Rostami<sup>1,\*</sup>, Mahdi Kalantari-Meybodi<sup>2</sup>, Masoud Karimi<sup>1</sup>, Afshin Tatar<sup>3</sup>, and Amir H. Mohammadi<sup>4,\*</sup>

<sup>1</sup> Department of Petroleum Engineering, Petroleum University of Technology (PUT), Ahwaz, Iran

<sup>2</sup> Young Researchers and Elites Club, Meybod Branch, Islamic Azad University, Meybod, Iran

<sup>3</sup> Young Researchers and Elites Club, Ahwaz Branch, Islamic Azad University, Ahwaz, Iran

<sup>4</sup> Discipline of Chemical Engineering, School of Engineering, University of KwaZulu-Natal, Howard College Campus, King George V Avenue, Durban 4041, South Africa

Received: 18 September 2017 / Accepted: 22 February 2018

**Abstract.** Polymers applications have been progressively increased in sciences and engineering including chemistry, pharmacology science, and chemical and petroleum engineering due to their attractive properties. Amongst the all types of polymers, partially Hydrolyzed Polyacrylamide (HPAM) is one of the widely used polymers especially in chemistry, and chemical and petroleum engineering. Capability of solution viscosity increment of HPAM is the key parameter in its successful applications; thus, the viscosity of HPAM solution must be determined in any study. Experimental measurement of HPAM solution viscosity is time-consuming and can be expensive for elevated conditions of temperatures and pressures, which is not desirable for engineering computations. In this communication, Multilayer Perceptron neural network (MLP), Least Squares Support Vector Machine approach optimized with Coupled Simulated Annealing (CSA-LSSVM), Radial Basis Function neural network optimized with Genetic Algorithm (GA-RBF), Adaptive Neuro Fuzzy Inference System coupled with Conjugate Hybrid Particle Swarm Optimization (CHPSO-ANFIS) approach, and Committee Machine Intelligent System (CMIS) were used to model the viscosity of HPAM solutions. Then, the accuracy and reliability of the developed models in this study were investigated through graphical and statistical analyses, trend prediction capability, outlier detection, and sensitivity analysis. As a result, it has been found that the MLP and CMIS models give the most reliable results with determination coefficients ( $R^2$ ) more than 0.98 and Average Absolute Relative Deviations (AARD) less than 4.0%. Finally, the suggested models in this study can be applied for efficient estimation of aqueous solutions of HPAM polymer in simulation of polymer flooding into oil reservoirs.

## 1 Introduction

Polymer flooding is one of the popular technologies for Enhanced Oil Recovery (EOR) that is utilized to monitor the mobility ratio of displacing phase to the displaced fluid, which is a function of viscosity and relative permeability [1]. In details, addition of polymer modifies the properties of injected water via increasing viscosity and decreasing relative permeability of water even at low concentrations of polymer leading to a decrease in mobility ratio [2]. It means that for a longer period of time polymer solution could effectively sweep the oil in a piston-like manner through the porous media before occurring viscous fingering or early

breakthrough. As a consequence, more stable displacement and higher sweep efficiency could be achieved during this process [2,3].

One of the extensively employed polymers in petroleum engineering especially in EOR process partially hydrolyzed polyacrylamides (HPAMs) due to their moderately good solubility in water and reasonable economic [4]. As a definition, HPAMs are synthesized from acrylamide monomers; thus, they will have negatively charged linear chain macromolecules, in which some of their monomers are hydrolyzed by an electrolyte [5]. The large viscosities of HPAM solutions are caused by repulsion forces amongst the negatively charged polymer chains, which are dependent to chain extension and degree of acrylamide hydrolysis [6].

The salt concentration and hardness have paramount impacts on the rheological properties of polymers. In this regards, when a monovalent cationic solution like  $\text{Na}^+\text{Cl}^-$

\* Corresponding authors: [amir\\_h\\_mohammadi@yahoo.com](mailto:amir_h_mohammadi@yahoo.com) (A. H. Mohammadi) & [alireza.rostami.put2014@gmail.com](mailto:alireza.rostami.put2014@gmail.com) (A. Rostami)

is in contact with polymer, electrostatic attraction forces are created between the aforementioned polymer chains of negative charges and  $\text{Na}^+$  which lead to the so-called phenomenon of polymer coiling up [7,8]. As a result, the polymers might precipitate and thus the solution viscosity could be decreased. In the presence of divalent cations like  $\text{Ca}^{2+}$ ,  $\text{Ba}^{2+}$ , and  $\text{Sr}^{2+}$ , the processes of degradation and precipitation of the polymer will severely happen, which lead to crucial decreases in the viscosity of HPAM solution compared with the impact of monovalent cations [8,9].

In addition to high salinity environment, polymer degradation is the other major concern taking place at harsh reservoir conditions. Based on the origin of degradation, it can be comprised of mechanical, chemical, and biological processes [10,11]. Mechanical degradation occurs when polymer solution is injected at high flowrates or it passes through the porous media of high permeability; in this manner, the HPAM solution viscosity will decrease [12]. Biological degradation frequently happens when we are dealing with biopolymers such as Xanthan. Moreover, chemical degradation takes place via oxidation reactions and accelerates at high temperatures of porous media [10,11]. It is worth mentioning that the impact of temperature is not restricted to the chemical degradation, it also affects its hydrolysis degree causing precipitation of polymers. The other issues associated with the application of polymers in porous media are the hydrodynamic retention [13,14] and macromolecular adsorption [1]. From operational point of view, injection of highly viscous fluid through the wellbore into the reservoir leads to an extreme pressure drop. Therefore, polymers as non-Newtonian fluids exhibiting shear thinning behavior are employed to preclude this phenomenon.

A large number of investigations have been conducted to study the impact of polymer concentrations, shear rate, temperature, salt concentration, HPAM hydrolysis degree, and hardness on the solution viscosity as the main rheological property of HPAM solutions [5,7,15–23]. Some researchers have undertaken a number of experimental investigations to examine the performance of Hydrophobically Associating Polyacrylamide (HAPAM) [24] and HPAM/Cr(III) [25] to improve the performance of HPAM components in EOR. Apart from the experimental measurements, empirical and semi-empirical methodologies, different approaches based on mechanical models [26], continuum theories [27] and molecular theories [28,29] have been utilized to calculate and also to apprehend some of the detected laboratory phenomena.

The experimental measurements of HPAM solution viscosity are valuable; however, they are expensive and time-consuming to be conducted. Additionally, conducting experiments covering wide ranges of data are not normally feasible in practice. Furthermore, numerical models normally solve Partial Differential Equations (PDEs) and use different iterative approaches for converging to the target solution [30,31]. Physical and mathematical natures of these models require using some assumptions for the purpose of simplification. These assumptions introduce some errors in the results. In addition, numerical models are not always user-friendly and need advanced knowledge of physics of the model and advanced mathematics. The

developed empirically derived equations in literature were mostly established on the basis of limited range of data at a specific condition [30,31]. The other shortcoming of the existing correlations is that they do not take into account the effects of the all input variables on the model output; thereby, their application to other conditions may lead to noticeable uncertainties in output estimation. Consequently, it is crucially important to develop robust modeling techniques on the basis of the most comprehensive database existing in literature.

Soft computation technique is a robust computer technology which could handle and optimize the highly non-linear engineering problems. The Artificial Neural Network (ANN)-based algorithms are the early versions of soft computational methods, which have been proven to produce precise estimations; however, the main deficiency of ANN-based approaches is non-reproducibility of results, which is comparatively caused by the stopping criteria variation and networks random initialization [32]. To solve this issue, improved versions of soft computations namely, Support Vector Machine (SVM)-based schemes as a supervised machine learning method, have been presented. The chief advantages of SVM-based schemes over classical algorithms are fewer adjustable parameters, less probable over-fitting problems, no earlier requirement for determination of the network topology, satisfactory generalization performance, and no needs for selection of the hidden nodes quantity [32]. In last years, several investigators have utilized the SVM-based algorithms in a wide range of petroleum and chemical engineering as a robust tool successfully [30,33–37]. In addition to the above-mentioned SVM framework modeling, Adaptive Neuro Fuzzy Inference System (ANFIS) is another strong approach for precise estimation of different industrial and engineering goals. In this methodology, by combination of both fuzzy logic and neural network based systems the disadvantages of both systems will be overcome. The successful usages of ANFIS modeling have been witnessed in the open literature through several researches [38–41].

In this work, the innovative advantages of MLP neural network, GA-RBF neural network, CSA-LSSVM, CHPSO-ANFIS and CMIS modeling approaches have been used to model the viscosity of HPAM solution over a wide range of operational conditions. Based on the authors' knowledge, there is no report on modeling of the HPAM solution viscosities by the aforementioned approaches in the open literature. These methodologies provide an accurate and reliable estimation for achieving the target parameter. The available empirical correlations in the literature are not as accurate as the proposed smart techniques in this study. Moreover, the empirical correlations have a lot of tuning parameters to be determined; although, the above-mentioned smart tools applied in this study use the least numbers of tuning parameters which make them superior to the existing models. The higher the numbers of the model parameters cause more overfitting of the model to the experimental data. Graphical and statistical analyses have also been done on the results to check the validity of the proposed models. In addition, trend estimation capability of the proposed models was checked and confirmed. Finally, outlier detection and sensitivity analyses were done to check the validity of the proposed models.

## 2 Modeling approaches

### 2.1 Least Squares Support Vector Machines (LSSVM)

Support Vector Machines (SVM) have been used as a robust tool for many function approximations in a wide range of engineering applications [42–44]. Generally, ordinary SVM method uses the quadratic programming subjected to the inequality constraints to solve the function approximation problems, in which its convergence is achieved very slowly. This scheme consumes a great deal of time and memory [45], which causes it to be applicable for small problems dealing with a small number of input data points. A new type of SVM, called Least Squares SVM (LSSVM) was developed by Suykens and Vandewalle [46] in which it was tried to reduce the complexity of the process and increase its convergence speed. This was achieved by replacing the inequality constraints in the ordinary SVM by the equality constraints in LSSVM, which can be solved more rapidly by an iterative process [45–47]. This causes LSSVM to be more desirable for problems dealing with a large number of experimental data and in cases where the time and memory are limited. As mentioned earlier, the final goal of LSSVM is to find the optimum separating hyper plane. If the hyper plane creating vector is called as  $w$ , the LSSVM will try to minimize the below cost function as follows [45]:

$$J(w,e) = \frac{1}{2}w^2 + \frac{1}{2}\gamma \sum_{i=1}^N e_i^2, \quad (1)$$

subjected to the following linear constraints:

$$y_k = w^T g(x_k) + b + e_k, \quad k = 1, \dots, N, \quad (2)$$

where  $e$  is error variable,  $\mu \geq 0$  is regularization constant,  $g(x)$  is the mapping function,  $w$  and  $b$  are weight vectors and bias terms, respectively, and superscript  $t$  denotes the transpose of the weight matrix. Coupling the two equations will result in the following equation [45]:

$$L_{\text{LSSVM}} = \frac{1}{2}w^2 + \frac{1}{2}\gamma \sum_{k=1}^N e_k^2 - \sum_{k=1}^N \beta_k \{w^T g(x_k) + b + e_k\}, \quad (3)$$

in which  $\beta_k$  are Lagrangian multipliers. Based on Lagrangian multipliers, initial conditions can be derived as follows:

$$\left\{ \begin{array}{l} \sum_{k=1}^N \beta_k = 0 \\ w = \sum_{k=1}^N \beta_k g(x_k) = 0 \\ \{w^T g(x_k) + b + e_k\} - y_k = 0 \\ \beta_k = \gamma e_k \end{array} \right. \quad k = 1, \dots, N. \quad (4)$$

In a linear regression problem, in which dependent and independent variables can be linearly separable, the LSSVM equation will be as follows [45]:

$$y = \sum_{k=1}^N \beta_k x_k^T x + b. \quad (5)$$

For using the last equation for the nonlinear problems, kernel functions must be used. By introducing kernel functions to Eq. (5), the below equation will be resulted [45]:

$$y = \sum_{k=1}^N \beta_k K(x_k \cdot x) + b, \quad (6)$$

where  $K(x_k \cdot x)$  is the kernel function that is the product of  $g(x)$  and  $g(x_k)$  in feasible margin as it is provided as follows [30,45]:

$$K(x_k \cdot x) = g(x)g(x_k), \quad k = 1, \dots, N. \quad (7)$$

Gaussian Radial Basis kernel Function (RBF) has been more attractive to the researchers among the other ones which is defined as follows [42]:

$$K(x_k \cdot x) = \exp\left(\frac{-\|x - x_k\|^2}{\sigma^2}\right), \quad k = 1, \dots, N, \quad (8)$$

where  $\sigma^2$  is squared bandwidth. In this study, this parameter is optimized using an external optimization technique namely, Coupled Simulated Annealing (CSA). More information about CSA is presented in open literature [41].

### 2.2 Neuro Fuzzy Inference System (ANFIS)

For the first time, Jang and Chuen-Tsai [48] proposed a novel version of intelligent-based approaches by combining Fuzzy Inference Systems (FIS) with Artificial Neural Network (ANN). By conducting this methodology, the capability of both ANN and FIS would be enhanced by achieving advantages of both fuzzy systems and neural network simultaneously. After that, Takagi and Sugeno [49] applied ANFIS mathematical strategy by extending to the highly nonlinear and complex problems. For this goal, two different training algorithms of hybrid and back propagation techniques were used. In back propagation scheme, the resulting parameters and their premise values are kept constant via gradient descent method; however, back propagation algorithm modifies the output parameters by the approach of least squared error approximation [50]. A typical ANFIS structure has five hidden layers of 1–5. There are some modifiable and constant nodes in this structure, which perform as rules and Membership Function (MF). Takagi and Sugeno [49] proposed two rules for development of ANFIS model as followings:

Rule 1:

$$\text{If } x_1 = A_1 \text{ and } x_2 = B_1, \text{ then } f_1 = m_1 x_1 + n_1 x_2 + r_1. \quad (9)$$

Rule 2:

$$\text{If } x_1 = A_2 \text{ and } x_2 = B_2, \text{ then } f_1 = m_1x_1 + n_1x_2 + r_1, \quad (10)$$

where,  $A_1, A_2, B_1$  and  $B_2$  are linguistic labels, and  $m_1, m_2, n_1, n_2, r_1,$  and  $r_2$  are the consequent parameters.

In addition, five hidden layers including fuzzification, rule, normalization, defuzzification, and output layers are applied in ANFIS strategy [39]. In fuzzification layer, the input data are converted to the linguistic expressions. In rule layer, the truthfulness of the constructed expressions in previous part will be determined. In next layer, the firing strength taken from the antecedent layers undergoes normalization process. A linguistic term for the output parameter called consequent layer is provided in the fourth layer. At final stage, the above-mentioned rules for one output are gathered as individual numerical outputs using an averaging method [38]. It is notable that ANFIS method can be optimized through the Conjugate Hybrid Particle Swarm Optimization (CHPSO). More information about CHPSO algorithm is presented in the open literature [41,51].

### 2.3 Multilayer Perceptron (MLP) neural network

MLP is a special structure of neural systems with three different types of input, hidden and output layers. There are some neurons exist in each layer. The number of neurons in hidden layers can be optimized using trial and error methods or an intelligent approach like least squares method. System errors will back propagated through the network and weights and bias iteratively estimated using epoch iteration system [52]. Undertraining and over training of the network quietly dependent on number of epochs in the network. This method will poorly forecast output result of test data set in the network.

### 2.4 Radial Basis Function (RBF) neural network

This method predicts different functions iteratively and has the same application like MLP method. The internal structure of this network has only three layers which are capable to highly noisy input data [53]. The internal structure of this method is simpler than MLP which uses feed forward structure with supervise training technique. In this framework, a nonlinear transformation is applied on the weighting vector of hidden layer through activation function  $\phi(r)$ . This network has three major properties as follows [54]:

- It can predict multivariable continuous function with the good accuracy interval if enough units present.
- The network uses linear unknown coefficient which somehow improves its optimal prediction capability.
- The solution obtained with this method will be the best solution according to minimization of cost function and control oscillation around best solution.

There are some differences between two previous mentioned neural network systems of MLP and RBF which include [55]:

- MLP structure is more complicated in comparison with RBF method.

- Training process of RBF is easier than MLP due to its simple structure of three layers with only one hidden layer.
- MLP network performs global approximation on input data and the outputs are estimated by neurons, while in RBF networks, the local performance can be obtained on input data with certain units determination of outputs.
- Discrimination process of these two artificial systems have shown some differences. In MLP networks, cluster classification is performed using hyper surfaces, while in RBF hyper spheres are used for this purpose.

It is notable that RBF network can be optimized through the Genetic Algorithm (GA) technique. More information about GA strategy is presented in the open literature [41].

### 2.5 Committee Machine Intelligent System (CMIS)

Committee Machine (CM) was introduced by Nilsson [56] for the first time as a way for doing supervised learning tasks. CM has a parallel structure, in which different experts run simultaneously and the results of these experts are combined together to find the better solution rather than the solution of individual experts [57,58]. CMs are categorized in two distinct categories including static and dynamic structures which are completely defined and expressed in literature [57]. In this study, ensemble averaging, which is one of the members of static structures, is used. After the creation of intelligent network of experts, it is essential to find a method for combining the results of individuals to find the optimum solution [59–61], which can be the linear averaging using the weighted average [62,63] that was used in this study to increase the contribution of the better expert's result in the final solution. Note that in this study a modified version of weighted average with a bias term for better fitness was used. In this work, the genetic algorithm, which is a very applicable method in optimum solution finding problems was used and its details has been discussed previously. In this study, Mean Square Error values (MSE) between the CMIS results and experimental values were used as the objective function for the genetic algorithm. The MSE can be defined as below:

$$MSE_{CMIS} = \frac{1}{n} \sum_{i=1}^n \left( \left( \sum_{j=1}^m w_j y_i^{\text{Expert}_j} + w_0 \right) - y_i^{\text{Experimental}} \right)^2, \quad (11)$$

in which  $n$  and  $m$  are the number of data and the number of experts,  $w$  is the weight of expert,  $y$  is the value of the HPAM solution viscosity, and  $w_0$  is the bias term.

## 3 Data gathering

Accuracy and reliability of a model is completely dependent on the accuracy and comprehensiveness of the data set used for its developing and testing [64–75]. For this reason, a comprehensive data set, covering a wide range of operating conditions, was gathered from the confident open

**Table 1.** Minimum, maximum and average values of the used parameters in each data set.

	Minimum			Maximum			Average		
	Train	Test	All	Train	Test	All	Train	Test	All
Hydrolysis (%)	25	25	25	30	30	30	29.30	29.36	29.31
MW (Dalton)	8	8	8	20	20	20	15.01	14.46	14.92
HPAM concentration (ppm)	0.5	0.5	0.5	3	3	3	2.84	2.87	2.85
T (K)	298	298	298	363	363	363	308.19	307.42	308.06
Salt concentration (ppm)	0	0	0	3	3	3	1.3984	1.3782	1.3980
Hardness ( $10^{-3}$ ppm)	0	0	0	3	3	3	0.47	0.47	0.47
Shear rate ( $s^{-1}$ )	1	1	1	10 000	10 000	10 000	463.41	461.51	463.03
HPAM solution viscosity (mPa.s)	0.4344	0.4344	0.4344	2.9174	2.9174	2.9174	1.3922	1.5423	1.4179

literature. Based on the literature, the most important parameters affecting the viscosity of HPAM solution (mPa.s) are degree of hydrolyzes (%), molecular weight of the polymer (Dalton), polymer concentration (ppm), temperature (K), salt concentration (ppm), water hardness ( $10^{-3}$  ppm), and shear rate condition ( $s^{-1}$ ), under which the HPAM solution viscosity is measured [3,5,21,76]. Thus, the prediction function was considered as below:

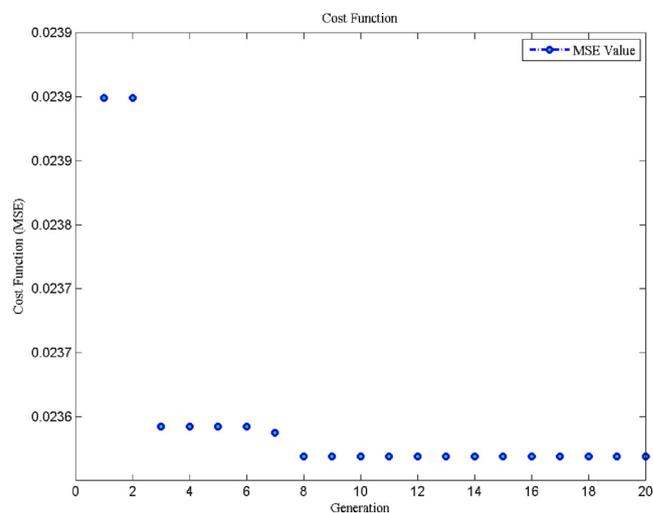
HPAM solution viscosity = f (degree of hydrolysis, HPAM molecular weight, HPAM concentration, temperature, salt concentration, water hardness, shear rate) (12)

The final data set used for developing and testing the models contained 403 data points, which were gathered from the most accurate and confident literature [5,21,76]. In a model development process, the data set is needed to be divided into two sets of data for development and testing of the models. For this reason, the gathered data set was randomly divided into two data sets, one containing 80% of data points namely train data set and the other containing the remaining 20% of data points namely test data set used for training and testing the models, respectively. The train data set was used for building the models and the test data set was used for testing the validity, generality, and overfitting and under-fitting problems detection of models. The proposed model will be valid and reliable if the statistical parameters of the model for train and test data sets are nearly the same, which show that the model is well-designed. It is very crucial that the range of the applied database in each of the mentioned data set must be nearly the same. The minimum, maximum and average of the parameters in each of the data sets are given in Table 1. As shown in this table, the minimum, maximum, and average of the data in each of the data sets are nearly the same which are consistent with the main data set confirming the appropriate division of the data set.

## 4 Models development

### 4.1 LSSVM

In this study, the LSSVM modeling approach developed by Pelckmans *et al.* [77] and Suykens *et al.* [78] was used for modeling of HPAM solution viscosity. In the CSA-LSSVM



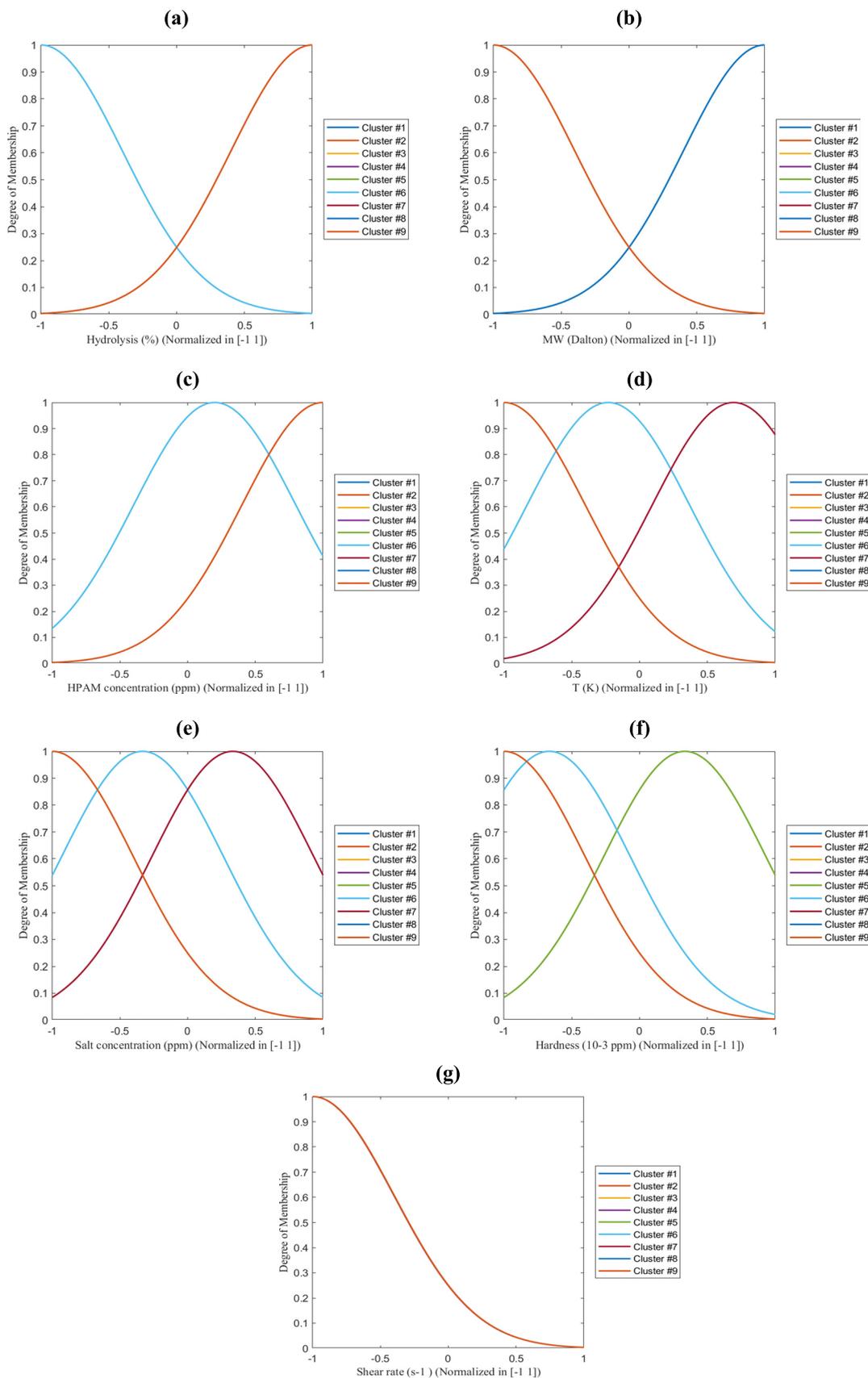
**Fig. 1.** The convergence of GA to the optimum value regarding the MSE of the developed ANFIS.

model development, the radial basis kernel function, which is the most widely used kernel function [42,79], was used. In addition, MSE parameter between model results and experimental values was used as the objective function. For optimization of parameters of the LSSVM model, CSA was used which resulted the optimum values of 22.97105 and 0.000361 for  $\gamma$  and  $\sigma$ , respectively.

### 4.2 CHPSO-ANFIS

For development of CHPSO-ANFIS model, after preparing the train and test data, the first step is to create an initial FIS. There are three different methods provided by MATLAB. The first method, namely, genfis1 uses grid partitioning. This method consumes a great deal of time and memory. The second method, namely, genfis2, which was utilized in this study enjoys from subtractive clustering method. The third method which is none of concern of this study, namely, genfis3 generates an FIS using Fuzzy C-Means (FCM) clustering by extracting a set of rules that models the data behavior.

As mentioned earlier, genfis2 benefits from subtractive clustering method to generate a Sugeno-type initial FIS. In this method, using the subclass function, the number of



**Fig. 2.** The MFs for a) degree of hydrolysis, b) HPAM molecular weight, c) HPAM concentration, d) temperature, e) salt concentration, f) water hardness, and g) shear rate for the initial FIS.

rules and antecedent membership functions are determined. After that, each rule's consequent equations are determined using least squares estimation. After that, the least squares estimation is utilized to determine each rule's consequent equations. These membership functions include a set of fuzzy rules to cover the feature space. In this method, the most important parameter is the radius of influence. This parameter spans between 0 and 1. Lower values of this parameter result in better performance of the FIS; however, it makes the FIS more complex and training stage consumes more time and memory. In this study, this parameter is assumed to be a value between 0.7 and 1. It is possible to find the optimum value by trial and error; however, it is more sophisticated to use an optimization method to determine the optimum value. In this study, GA was utilized to determine the optimum value of the radius of influence such that the generated FIS has acceptable accuracy and do not make the FIS too complex for training stage. For the data set used in this study, the optimum value for the radius influence was determined to be 0.8485. The convergence of the radius influence is depicted in Figure 1. In this figure, the vertical axis is the value of the MSE and the horizontal axis is the number of generation regarding the GA optimization method.

With this value, 9 rules were generated. The membership functions are depicted in Figure 2. As shown in Figure 2, all the values of the input parameters are normalized between  $-1$  and  $1$ . In this study, we have 7 input parameters; thus, there are correspondingly 7 subfigures showing the membership functions. The corresponding rules in indexed format are as follows:

```

1 1 1 1 1 1 1, 1 (1) : 1
2 2 2 2 2 2 2, 2 (1) : 1
3 3 3 3 3 3 3, 3 (1) : 1
4 4 4 4 4 4 4, 4 (1) : 1
5 5 5 5 5 5 5, 5 (1) : 1
6 6 6 6 6 6 6, 6 (1) : 1
7 7 7 7 7 7 7, 7 (1) : 1
8 8 8 8 8 8 8, 8 (1) : 1
9 9 9 9 9 9 9, 9 (1) : 1

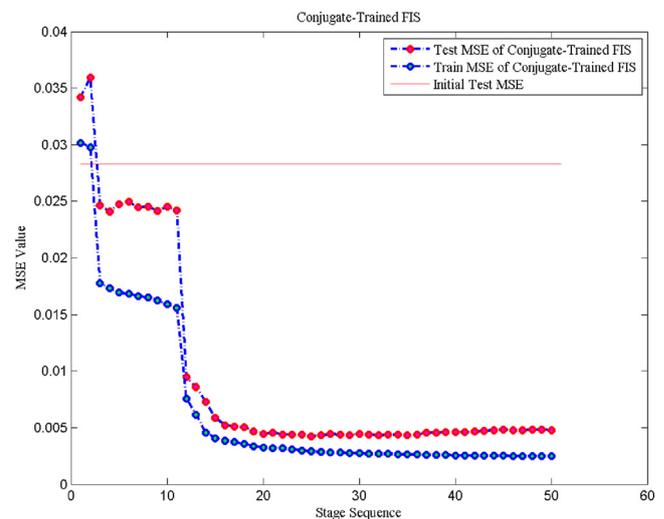
```

Now, it is time to tune the initial FIS to reach the best solution, i.e. the minimum MSE between the target and output values. MATLAB provides two methods to train the FIS, namely, back propagation and hybrid method. Another method to train the initial FIS is using optimization methods. In this case, the MF parameters are regarded as the tuning parameters. In this study, a population-based optimization algorithm, namely Particle Swarm Optimization (PSO) algorithm was utilized. PSO considers the MFs parameters as the tuning parameters and the MSE between the target and ANFIS output as the cost function. This method continues searching the best tuning parameters in search space until reaching to the stopping criteria.

In this study, a combination of hybrid method (provided by MATLAB) and innovative PSO method was utilized. The method is called Conjugate Hybrid PSO ANFIS (CHPSO-ANFIS). In this method, there is a determined number of stages. At first stage, the initial FIS

**Table 2.** Parameters of the CHPSO-ANFIS training functions using the CHPSO method.

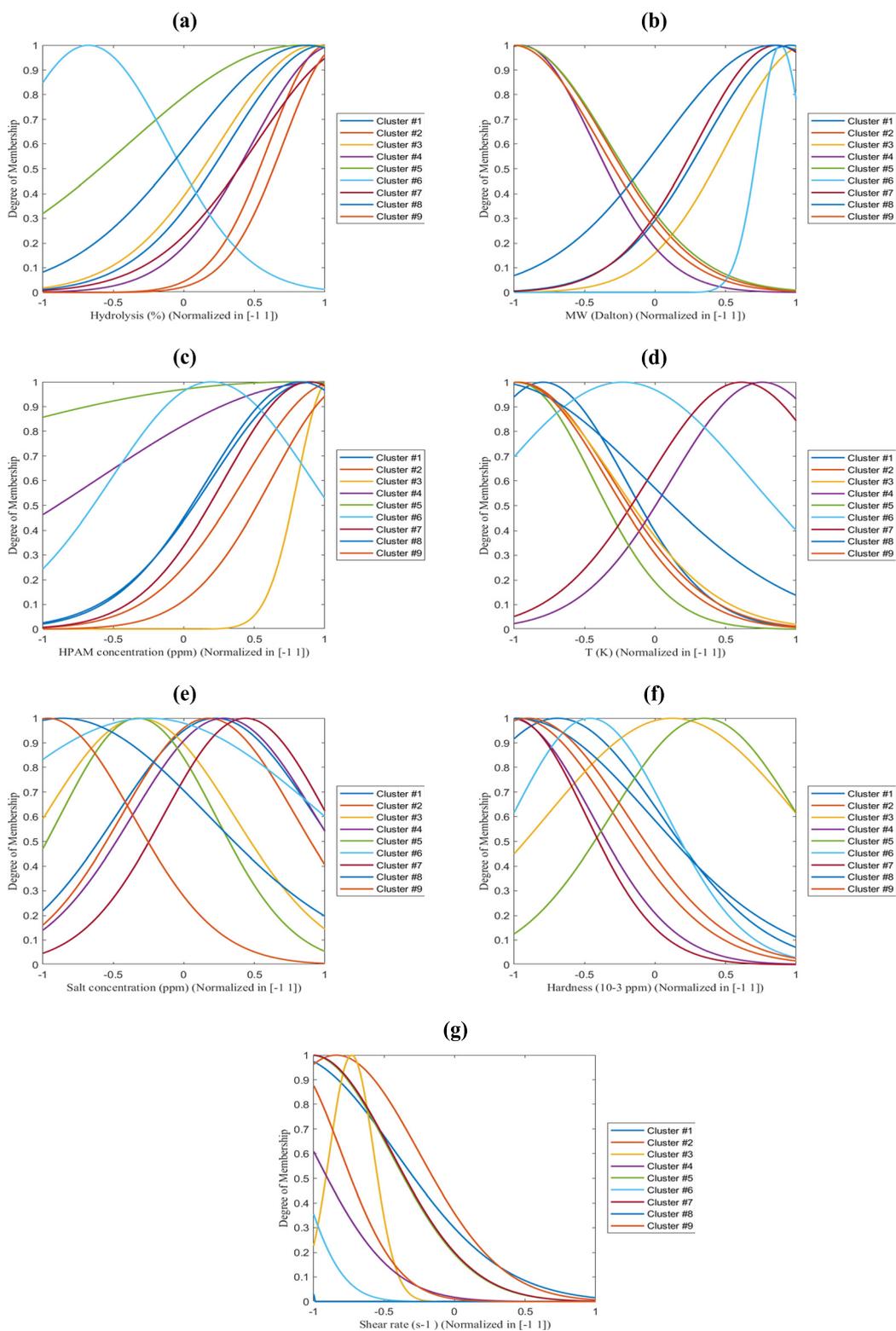
Parameter	Value
(Initial FIS) minimum value of influence radius	0.7
(Initial FIS) maximum value of influence radius	1
(Initial FIS) maximum number of GA iterations	20
(Initial FIS) number of GA initial population	30
(Hybrid) Training epoch number	10
(Hybrid) training error goal	0
(Hybrid) initial step size	0.01
(Hybrid) step size decrease rate	0.9
(Hybrid) step size increase rate	1.1
(Hybrid) number of each steps for the hybrid training	3
(PSO) Training epoch number	200
(PSO) number of initial population	100
(PSO) number of each steps for the PSO optimization training	1
(Conjugate) number of Stages	50



**Fig. 3.** The performance of the ANFIS training by CHPSO method.

undergo hybrid training for 3 successive times. Afterwards, the output FIS undergoes PSO method just for the first time. It should be noted that individual hybrid and PSO training steps are comprised of 10 and 200 epochs, respectively. After training by the PSO optimization method one stage is completed. The resulted tuned FIS is passed to the next stage. This process will be continued until reaching the stopping criteria. The parameters of the ANFIS training functions using the CHPSO method are listed in Table 2.

Figure 3 shows the performance of the CHPSO-ANFIS. The vertical and horizontal axes are cost function (i.e., the MSE between the target and output values) and the



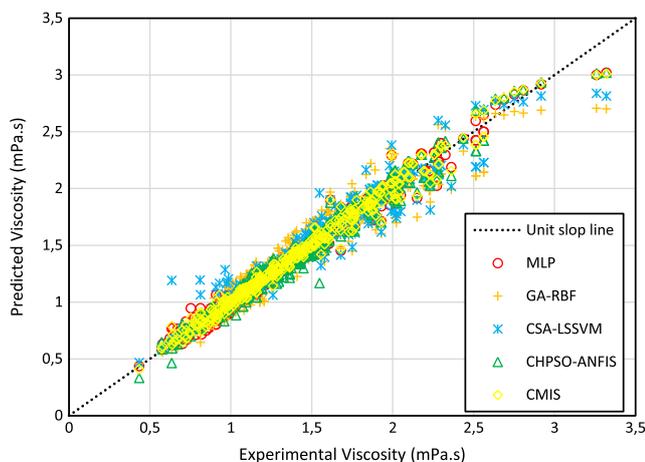
**Fig. 4.** The MFs for a) degree of hydrolysis, b) HPAM molecular weight, c) HPAM concentration, d) temperature, e) salt concentration, f) water hardness, and g) shear rate for the CHPSO-trained FIS.

number of stages, respectively. The overfitting problem is controlled during the training. The best stage is a stage that there is a balance between the MSE value for the train and test data sets. In this figure, the solid red line is the best

MSE of the initial FIS. As it is evident, there have been two great jumps in the MSE value for both train and test data sets at 3rd and 13th stages. However, the best stage was determined to be the 36<sup>th</sup> stage. After this stage, although

**Table 3.** Comparison of the static parameters of the initial and CHPSO trained FIS.

Predictor	Data Set	$R^2$	AARD	STD	RMSE	N
Initial FIS	Train Data	0.8365	11.33	0.1397	0.1986	322
	Test Data	0.7275	14.13	0.1712	0.2429	81
	All Data	0.8156	11.89	0.1476	0.2082	403
CHPSO trained FIS	Train Data	0.9773	4.06	0.0506	0.0740	322
	Test Data	0.9575	5.05	0.0676	0.0955	81
	All Data	0.9736	4.26	0.0544	0.0788	403

**Fig. 5.** Agreement between the estimated and experimental viscosities.

the MSE value for the training data sets continuously reduces, the MSE for the testing data sets little by little increases and this leads to overfitting problem. The final membership functions for each input data are illustrated in Figure 4. As it is obvious, the MFs will tremendously change after training. As it is evident, the MSE value for the CHPSO-ANFIS is much less than the MSE of the initial FIS. In order to have a better comparison, the statistical parameters of the initial FIS and the CHPSO trained FIS are listed in Table 3.

### 4.3 MLP

It has been demonstrated that the efficient estimation of each nonlinear function could be done by applying one hidden layer in MLP structure [80]. Thus, in this study, only one hidden layer was selected so as to reduce the calculation time. Moreover, the suggested MLP model has one neuron in its output layer and seven neurons in its input layer corresponding to one output parameter and seven input variables, respectively. By varying the neurons number from 4 to 25, the MSE value as an Objective Function (OF) was evaluated. In other words, the number of neurons with the lowest MSE value was selected. In this study, there are six neurons in hidden layers in which the best performance of the MLP would be achieved.

### 4.4 GA-RBF

The best performance of the RBF network will be accomplished when the tuning parameters, which control the accuracy of the model, are optimal values. These parameters are known as Maximum Neurons Number (MNN) and spread. For tuning these values, the GA technique was utilized. First, 60 arbitrary solution pairs were generated and categorized based on the value of MSE between the measured data and model estimates. The best values for MNN and spread parameters were obtained after 30 generations. As a result, the MNN and spread values are, respectively, 173 and 1.15612.

### 4.5 CMIS

For CMIS development, the GA method was employed to find the optimum weights of the model according to the following equation:

$$y_{\text{CMIS}}^i = ly_{\text{GA-RBF}}^i + my_{\text{CSA-LSSVM}}^i + ny_{\text{MLP}}^i + ry_{\text{CHPSO-ANFIS}}^i + s, \quad (13)$$

where,  $l$ ,  $m$ ,  $n$ ,  $r$  and  $s$  were found to be  $-0.031781$ ,  $0.201029$ ,  $0.641679$ ,  $0.2011969$ , and  $0.004969$ , respectively.

## 5 Results and discussion

### 5.1 Models validation

In this section, the accuracy and validity of the proposed models are investigated through statistical and graphical analyses. The graphical agreement between the estimated HPAM solution viscosity values using the proposed models and experimental values is shown in Figure 5. As it is shown in this figure, there exists a good agreement between the experimental viscosities and estimated values with high value of the determination coefficient of 0.9826 for CMIS model, respectively. This figure also shows that the CSA-LSSVM and GA-RBF models estimations are more scattered around the 45° line as compared with other models. In addition, it shows that CMIS model has a good estimation capability for lower HPAM solution viscosity values, which decrease slightly but not unreliably for higher viscosity values. For further investigation of the reliability of the proposed models, the relative error distributions of

**Table 4.** Statistical parameters of the proposed models.

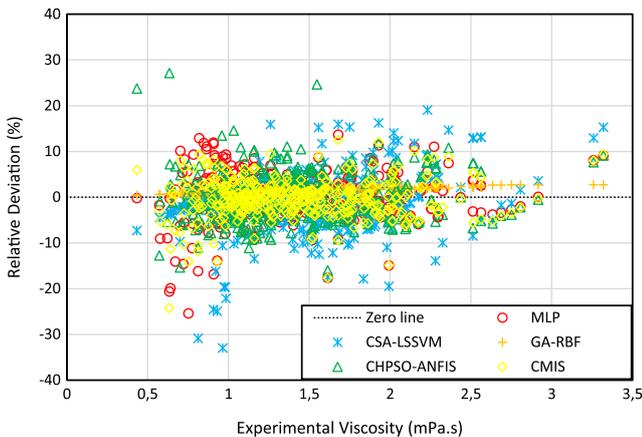
		R <sup>2</sup>	ARD	AARD	RMSE	min RD	max RD
MLP	Train	0.9841	-0.0508	3.48	0.0620	0.0125	25.43
	Test	0.9617	-0.1319	4.81	0.0901	0.1090	20.68
	All	0.9800	-0.0671	3.74	0.0685	0.0125	25.43
GA-RBF	Train	0.9285	-0.2548	6.73	0.1313	0.0232	19.87
	Test	0.8801	-3.4229	8.82	0.1626	0.2560	21.85
	All	0.9188	-0.8916	7.15	0.1382	0.0232	21.85
CSA-LSSVM	Train	0.9687	-0.4695	3.15	0.0877	0.0156	15.26
	Test	0.8501	-6.1658	10.64	0.1822	0.0996	87.45
	All	0.9460	-1.6144	4.66	0.1132	0.0156	87.45
CHPSO-ANFIS	Train	0.9773	-0.1700	4.06	0.0740	0.0038	23.73
	Test	0.9575	0.2008	5.05	0.0955	0.2521	27.15
	All	0.9736	-0.0955	4.26	0.0788	0.0038	27.15
CMIS	Train	0.9869	-0.0755	2.77	0.0562	0.0031	14.12
	Test	0.9627	-1.2070	4.19	0.0885	0.0570	24.26
	All	0.9826	-0.3029	3.05	0.0640	0.0031	24.26

$$\text{Determination coefficient: } R^2 = 1 - \frac{\sum_{i=1}^N (\lambda_{\text{Pred}}(i) - \lambda_{\text{Exp}}(i))^2}{\sum_{i=1}^N (\lambda_{\text{Pred}}(i) - \bar{\lambda}_{\text{Exp}})^2}$$

$$\text{Average Relative Deviation: } \%ARD = \frac{100}{N} \sum_{i=1}^N \frac{(\lambda_{\text{Pred}}(i) - \lambda_{\text{Exp}}(i))}{\lambda_{\text{Exp}}(i)}$$

$$\text{Average Absolute Relative Deviation: } \%AARD = \frac{100}{N} \sum_{i=1}^N \frac{|\lambda_{\text{Pred}}(i) - \lambda_{\text{Exp}}(i)|}{\lambda_{\text{Exp}}(i)}$$

$$\text{Root Mean Square Error: } RMSE = \left( \frac{\sum_{i=1}^N (\lambda_{\text{Pred}}(i) - \lambda_{\text{Exp}}(i))^2}{N} \right)^{0.5}$$



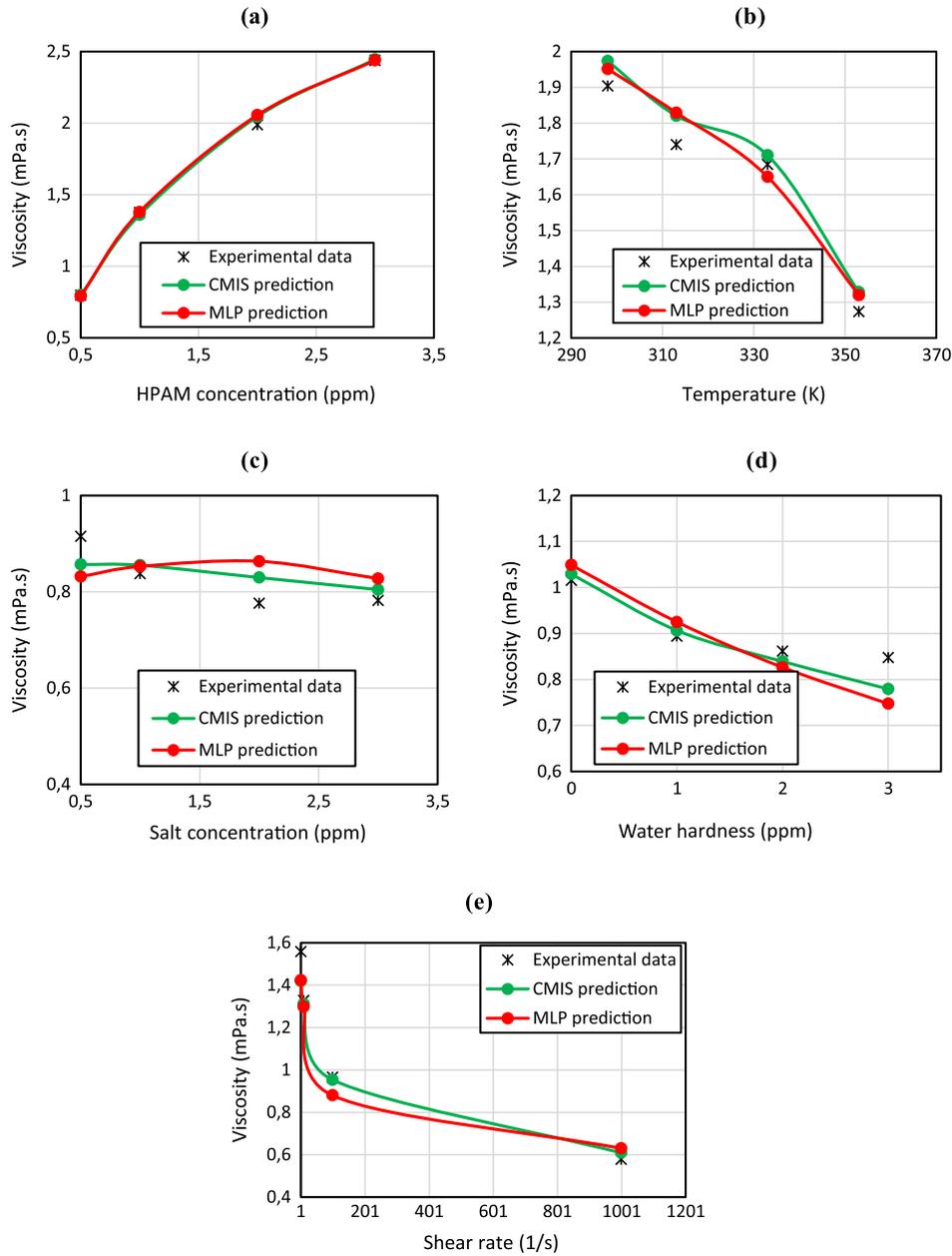
**Fig. 6.** Relative deviation of estimated values using both proposed models.

the models results against corresponding experimental values are shown in Figure 6. As shown in this figure, the data points for CMIS model are highly concentrated around the zero horizontal error line, which confirm the relative superiority of this model in comparison with the other proposed models.

The statistical investigation of the proposed models is shown in Table 4 through numerous statistical parameters including coefficient of correlation (R<sup>2</sup>), Average Relative Deviation (ARD), Average Absolute Relative Deviation (AARD), Root Mean Square Error (RMSE), minimum Relative Deviation (min RD) and maximum Relative Deviation (max RD) for all train, and test data sets. As known, the two most important parameters for evaluation of any model are AARD and R<sup>2</sup>. When the value of AARD is close to zero and R<sup>2</sup> value is near unity, it can be concluded that the suggested model is well-predictive and efficient. According to Table 4, it is obvious that the CMIS and GA-RBF models are, respectively, the most accurate and the least accurate techniques for prognostication of HPAM solution viscosity in this study. Moreover, MLP statistical quality measures show that it is the second most precise model after CMIS model.

### 5.2 Models trend estimation capability

One of the very important points in reliability and estimation capability of the models is the ability to estimate the true trend of viscosity changes by variations of different parameters. It is a critical point for examining the usefulness of any developed model. In this section, the

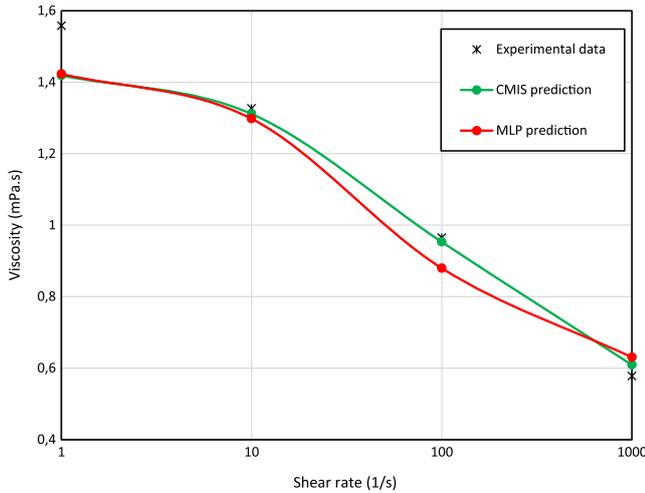


**Fig. 7.** Trend estimation of the CMIS and MLP models for comparison with experimental HPAM solution viscosity values versus: a) HPAM concentration, b) Temperature, c) Salt concentration, d) Water hardness, and e) Shear rate.

trend predictability of the best developed models (i.e., MLP and CMIS) is investigated. Figure 7 shows the trend estimation capability of CMIS and MLP models for changes with respect to the HPAM concentration, temperature, salt concentration, water hardness, and shear rate. As it is shown in these figures, trends have successfully been captured by both models; however, the proposed CMIS model is a slightly better tool rather than the proposed MLP model.

In dealing with polymer solutions, the behavior of the solution with respect to the shear rate is very important. For this reason, a model will be reliable if it can estimate

the HPAM solution viscosity variations with respect to shear rate. Figure 7e shows the estimation capability of the proposed MLP and CMIS models for shear rate changes. As it is shown in this figure, both proposed models have estimated the trend, successfully; however, the proposed CMIS model has better estimation performance. To deeper investigate, the Figure 7e was re-drawn with the logarithmic x-axis (shear rate) as it is shown in Figure 8. This figure clearly confirms the better performance of the proposed CMIS model rather than MLP model. For shear rates more than 10, CMIS model exhibits a decreasing and linear trend for estimating HPAM solution viscosity.



**Fig. 8.** Trend estimation of the proposed models for shear rate changes (logarithmic scale).

### 5.3 Outlier detection

In any modeling study, the accuracy and the reliability of the proposed model is completely dependent on the accuracy of used experimental data, as mentioned earlier [81]. The erroneous data lower the accuracy and applicability of the proposed models; thus, these data must be specified. Fortunately, these data normally behave in a different scheme rather than the bulk of data, which could be distinguished by using outlier detection methods. Normally, the outlier detection is done graphically through William's plot [42, 82–84]. This graphical approach needs standardized residual values and hat values. The outliers will be detected using two limits including leverage and residual limits. The hat values for each of the data points are the diagonal values of  $H$  matrix, which can be calculated as follows [82–84]:

$$H = X(X^t X)^{-1} X^t, \quad (14)$$

in which,  $X$  is a matrix containing  $m$  rows and  $n$  columns. The parameter  $m$  is the number of data points and  $n$  is the number of input parameters. The residual values are the difference between the estimated values using proposed model and experimental values. The standardized residual values can be calculated using Eq. (15) as follows [82,83]:

$$SR_i = \frac{\mu_{\text{exp}} - \mu_{\text{pred}}}{\sqrt{MSE(1 - H_i)}}, \quad (15)$$

in which  $SR_i$  is the standardized residual value of  $i$ -th data point,  $\mu_{\text{exp}}$  is the experimental HPAM solution viscosity,  $\mu_{\text{pred}}$  is the estimated HPAM solution viscosity using the proposed model, MSE is the mean square error, and  $H_i$  is the hat value of the  $i$ -th data point. The warning leverage value is normally denoted by  $H^*$  and from the leverage point of view, the data that applies in  $0 \leq h \leq H$ , in which  $h$  is the hat value of the data point. This limit is indicated by a red dashed line in William's plot shown in Figure 9. The  $H^*$  can be calculated

using the following equation:

$$H^* = \frac{3(n+1)}{m}, \quad (16)$$

in which  $n$  is the number of input parameters and  $m$  is the number of data point. In this study,  $H^*$  was calculated to be equal to 0.06045. The other limit, which is the residual limit, is normally selected as the radius of 3, which means the standard residual values to be between 3 and  $-3$ . This means that from the residual point of view, the data points with  $-3 \leq SR \leq 3$  are reliable for the developed model, which are shown by two green dashed lines in Figure 9. Considering both leverages and residual limits, the data points that are applied to  $0 \leq h \leq H$  and  $-3 \leq SR \leq 3$  are highly accurate and the model is increasingly reliable. These data points are shown by blue star in Figure 9.

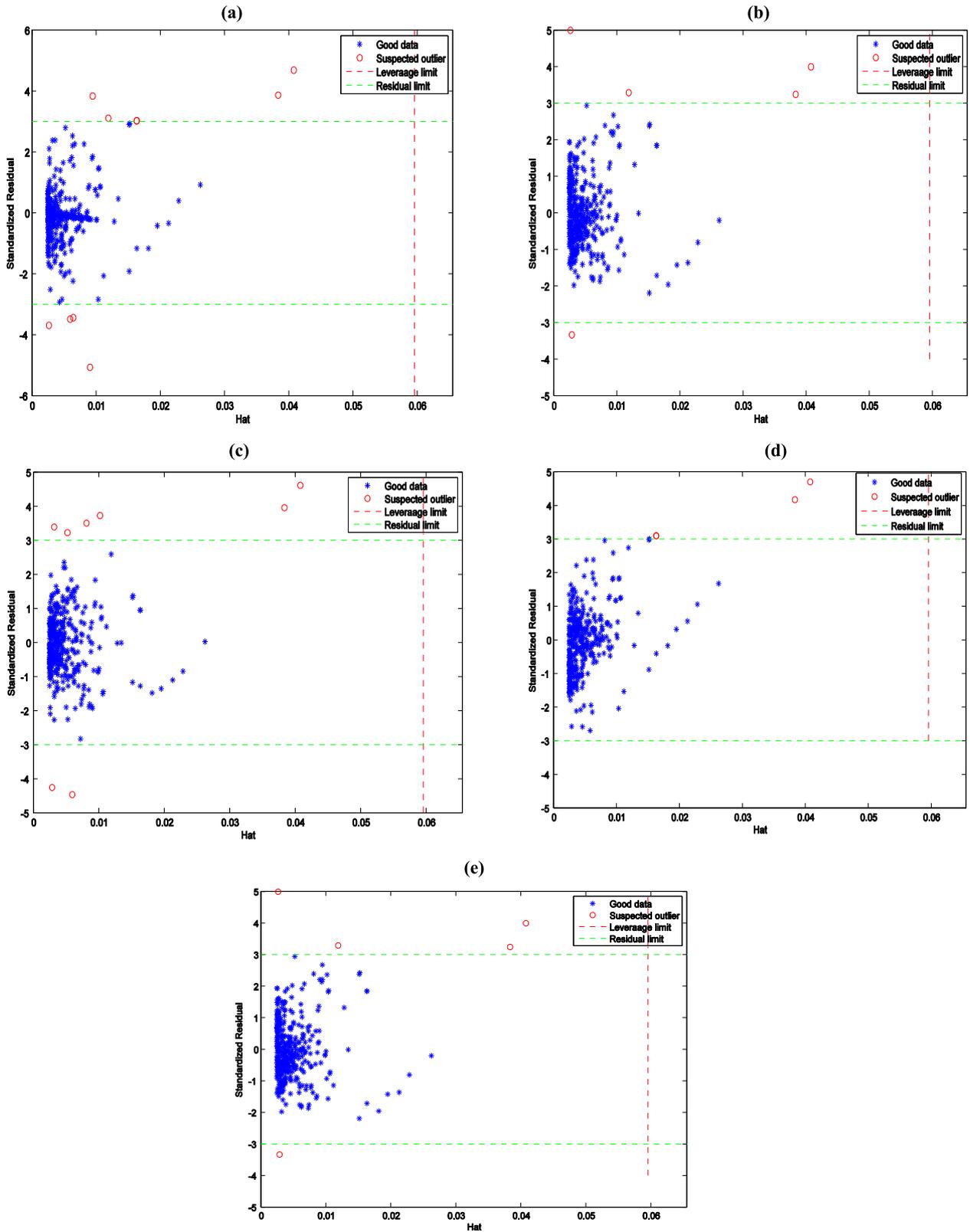
Figure 9 shows the William's plots for the proposed CSA-LSSVM, CHPSO-ANFIS, MLP, GA-RBF and CMIS models in this study over the entire database applied for modeling. This figure shows that only 2.23% of data (9 data points), 1.24% of data (5 data points), 1.99% of data (8 data points), 0.74% of data (3 data points) and 1.24% of data (5 data points) are suspected outliers for the proposed CSA-LSSVM, CHPSO-ANFIS, MLP, GA-RBF and CMIS models, respectively, which are very small and negligible; therefore, the major portion of the data are concentrated in the valid region, which are bounded by  $0 \leq h \leq H^*$  and  $-3 \leq SR \leq 3$ . This confirms almost always our modeling techniques in this study lie in valid range based on the pre-discussed outlier analysis presented in Figure 9.

### 5.4 Sensitivity analysis

A model can be reliable when the sensitivity of the model to the input parameters is the same as the sensitivity of the experimental values to them. This approach was first used by Chen et al. [85] for sensitivity analysis of models, which shows the degree and the sign of the effect of that parameter on the output values. In this approach, Equation (17) is used to calculate the relevancy factor of each of input parameters on the output estimation:

$$\text{Relevancy factor}(I_k, \mu) = \frac{\sum_{i=1}^n (I_{k,i} - I_k)(\mu_i - \bar{\mu})}{\sqrt{\sum_{i=1}^n (I_{k,i} - I_k)^2 (\mu_i - \bar{\mu})^2}}. \quad (17)$$

In the previous equation,  $\mu$  is the HPAM solution viscosity,  $I_k$  is the  $k$ -th input parameter,  $I_{k,i}$  is the  $i$ -th value of the  $k$ -th parameter, and  $I_k$  and  $\bar{\mu}$  are the average values of  $k$ -th input parameter and HPAM solution viscosity, respectively. Figure 10 shows the calculated normalized relevancy factors for all the input parameters affecting the experimental target values as well as the proposed CSA-LSSVM, CHPSO-ANFIS, MLP, RBF and CMIS models. This figure shows that the sensitivities of the proposed models are nearly the same as the sensitivities of the experimental values, which confirm the reliability of all the proposed models in this study. As it is obvious from this figure, the temperature is the most effective parameter on



**Fig. 9.** Outlier detection for the proposed models including: a) CSA-LSSVM, b) CHPSO-ANFIS, c) MLP, d) GA-RBF, and e) CMIS.

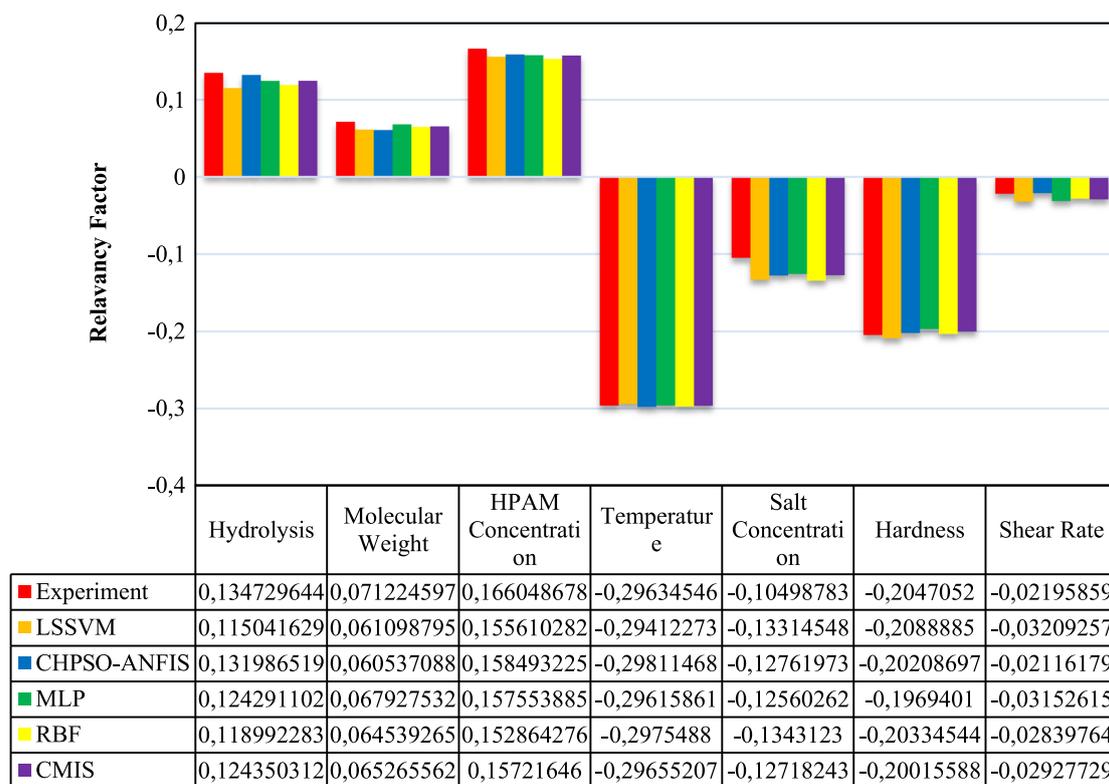


Fig. 10. Relevancy factor of input parameters.

the HPAM solution viscosity, and it has the negative effect which is normally expected for HPAM solution viscosity.

## 6 Conclusion

In this study, Multilayer Perceptron (MLP) neural network, Least Squares Support Vector Machine approach optimized with Coupled Simulated Annealing (CSA-LSSVM), Radial Basis Function neural network optimized with Genetic Algorithm (GA-RBF), Adaptive Neuro Fuzzy Inference System coupled with Conjugate Hybrid Particle Swarm Optimization (CHPSO-ANFIS) approach and Committee Machine Intelligent System (CMIS) were used to model the viscosity of HPAM solution over a wide range of operational conditions. The accuracy of the proposed models was investigated through statistical and graphical analyses, which show the good accuracy and reliability of the proposed models. As a result, it was found that the CMIS model gives the most accurate estimates for HPAM solution viscosity. Trend analysis of the CMIS and MLP methods as the best models confirms the large fitness of the proposed models estimations to the target solution viscosities. Outlier detection was executed through leverage values statistics; accordingly, the applicability of the developed models here with a very small number of outlier data points was confirmed. Based on the results of sensitivity analysis, it was demonstrated that the temperature is the most affecting parameter on HPAM solution viscosity estimation. Finally, it can be stated that the

results of this study are of significant usefulness for engineers and researchers working on polymer flooding projects of hydrocarbon reservoirs.

## References

- 1 Ait-Kadi A., Carreau P., Chauveteau G. (1987) Rheological properties of partially hydrolyzed polyacrylamide solutions, *J. Rheol.* (1978-present) **31**, 537–561
- 2 Sheng J. (2010) *Modern chemical enhanced oil recovery: theory and practice*, Gulf Professional Publishing, Houston, Texas, United States
- 3 Hashmet M.R., Onur M., Tan I.M. (2014) Empirical correlations for viscosity of polyacrylamide solutions with the effects of concentration, molecular weight and degree of hydrolysis of polymer, *J. Appl. Sci.* **14**, 1000
- 4 Lee K.E., Khan I., Morad N., Teng T.T., Poh B.T. (2012) Physicochemical and rheological properties of novel magnesium salt-polyacrylamide composite polymers, *J. Dispers. Sci. Technol.* **33**, 1284–1291
- 5 Hashmet M.R., Onur M., Tan I.M. (2014) Empirical correlations for viscosity of Polyacrylamide solutions with the effects of salinity and hardness, *J. Dispers. Sci. Technol.* **35**, 510–517
- 6 Garrouch A.A., Gharbi R.B. (1999) An empirical investigation of polymer flow in porous media, *Ind. Eng. Chem. Res.* **38**, 3564–3571
- 7 Levitt D., Pope G.A. (2008) Selection and screening of polymers for enhanced-oil recovery, in: *SPE Symposium on Improved Oil Recovery*, Society of Petroleum Engineers

- 8 Yen H.-Y., Yang M.-H. (2003) The effect of metal ions additives on the rheological behavior of polyacrylamide solution, *Polym. Test.* **22**, 389–393
- 9 Niu Y., Jian O., Zhu Z., Wang G., Sun G. (2001) Research on hydrophobically associating water-soluble polymer used for EOR, in: *SPE International Symposium on Oilfield Chemistry*, Society of Petroleum Engineers
- 10 Sorbie K.S. (1991) *Polymer-Improved Oil Recovery*, Blackie and Son Ltd, Glasgow and London
- 11 Seright R., Henrici B. (1990) Xanthan stability at elevated temperatures, *SPE Reserv. Eng.* **5**, 52–60.
- 12 Ghoniem S., Chauveteau G., Moan M., Wolff C. (1981) Mechanical degradation of semi-dilute polymer solutions in laminar flows, *Can. J. Chem. Eng.* **59**, 450–454
- 13 Reed R., Healy R., Shah D., Schechter R. (1977) *Improved Oil Recovery by Surfactant and Polymer Flooding*, in: Shah D.O., Schechter R.S. (eds), Academic Press Inc., New York, 383 p
- 14 Willhite G.P., Dominguez J.G. (1977) Mechanisms of polymer retention in porous media, in: Schechter R.S. (ed), *Improved Oil Recovery by Surfactant and Polymer Flooding*, Academic Press, Cambridge, Massachusetts, United States pp. 511–554
- 15 Fuoss R.M. (1948) Viscosity function for polyelectrolytes, *J. Polym. Sci.* **3**, 603–604
- 16 Doe P.H., Moradi-Araghi A., Shaw J.E., Stahl G.A. (1987) Development and evaluation of EOR polymers suitable for hostile environments part I: Copolymers of vinylpyrrolidone and acrylamide, *SPE Reserv. Eng.* **2**, 461–467
- 17 Gao C.H. (2011) Scientific research and field applications of polymer flooding in heavy oil recovery, *J. Pet. Explor. Prod. Technol.* **1**, 65–70
- 18 Seright R.S., Fan T., Wavrik K., Balaban R.D.C. (2011) New insights into polymer rheology in porous media, *SPE J.* **16**, 35–42
- 19 Ward J., Martin F.D. (1981) Prediction of viscosity for partially hydrolyzed polyacrylamide solutions in the presence of calcium and magnesium ions, *Soc. Pet. Eng. J.* **21**, 623–631
- 20 Rashidi M., Blokhuis A.M., Skauge A. (2011) Viscosity and retention of sulfonated polyacrylamide polymers at high temperature, *J. Appl. Polym. Sci.* **119**, 3623–3629
- 21 Hashmet M.R., Onur M., Tan I.M. (2014) Empirical correlations for viscosity of polyacrylamide solutions with the effects of temperature and shear rate. II, *J. Dispers. Sci. Technol.* **35**, 1685–1690
- 22 Gao C. (2013) Viscosity of partially hydrolyzed polyacrylamide under shearing and heat, *J. Pet. Explor. Prod. Technol.* **3**, 203–206
- 23 Gao C.H. (2014) Comprehensive correlations to calculate viscosity of partially hydrolyzed polyacrylamide, in: *SPE EOR Conference at Oil and Gas West Asia*, Society of Petroleum Engineers
- 24 Zhang P., Wang Y., Yang Y., Chen W., Bai S. (2015) Effective viscosity in porous media and applicable limitations for polymer flooding of an associative polymer, *Oil Gas Sci. Technol. - Rev. IFP Energ. Nowv.* **70**, 931–939
- 25 Haiyang Y., Yefei W., Jian Z., Peng L., Shenglong S. (2015) Dynamic gelation of HPAM/Cr(III) under shear in an agitator and porous media, *Oil Gas Sci. Technol. - Rev. IFP Energ. Nowv.* **70**, 941–949
- 26 Hinch E. (1977) Mechanical models of dilute polymer solutions in strong flows, *Phys. Fluids* (1958–1988) **20**, S22–S30
- 27 Gordon R., Schowalter W. (1972) Anisotropic fluid theory: a different approach to the dumbbell theory of dilute polymer solutions, *Trans. Soc. Rheol.* (1957–1977) **16**, 79–97
- 28 Bird R., Dotson P., Johnson N. (1980) Polymer solution rheology based on a finitely extensible bead–spring chain model, *J. Non-Newton. Fluid Mech.* **7**, 213–235
- 29 De Gennes P. (1974) Coil-stretch transition of dilute flexible polymers under ultrahigh velocity gradients, *J. Chem. Phys.* **60**, 5030–5042
- 30 Rostami A., Masoudi M., Ghaderi-Ardakani A., Arabloo M., Amani M. (2016) Effective thermal conductivity modeling of sandstones: SVM framework analysis, *Int. J. Thermophys.* **37**, 1–15
- 31 Meybodi M.K., Daryasafar A., Karimi M. (2016) Determination of hydrocarbon-water interfacial tension using a new empirical correlation, *Fluid Phase Equilib.* **415**, 42–50
- 32 Shokrollahi A., Arabloo M., Gharagheizi F., Mohammadi A. H. (2013) Intelligent model for prediction of CO<sub>2</sub>-reservoir oil minimum miscibility pressure, *Fuel* **112**, 375–384
- 33 Mesbah M., Soroush E., Azari V., Lee M., Bahadori A., Habibnia S. (2015) Vapor liquid equilibrium prediction of carbon dioxide and hydrocarbon systems using LSSVM algorithm, *J. Supercrit. Fluids* **97**, 256–267
- 34 Mesbah M., Soroush E., Shokrollahi A., Bahadori A. (2014) Prediction of phase equilibrium of CO<sub>2</sub>/cyclic compound binary mixtures using a rigorous modeling approach, *J. Supercrit. Fluids* **90**, 110–125
- 35 Rafiee-Taghanaki S., Arabloo M., Chamkalani A., Amani M., Zargari M.H., Adelzadeh M.R. (2013) Implementation of SVM framework to estimate PVT properties of reservoir oil, *Fluid Phase Equilib.* **346**, 25–32
- 36 Arabloo M., Shokrollahi A., Gharagheizi F., Mohammadi A. H. (2013) Toward a predictive model for estimating dew point pressure in gas condensate systems, *Fuel Process. Technol.* **116**, 317–324
- 37 Tatar A., Barati-Harooni A., Najafi-Marghmaleki A., Mohebbi A., Ghiasi M.M., Mohammadi A.H., Hajinezhad A. (2016) Comparison of two soft computing approaches for predicting CO<sub>2</sub> solubility in aqueous solution of piperazine, *Int. J. Greenh. Gas Control.* **53**, 85–97
- 38 Esmaeili-Jaghdan Z., Shariati A., Nikou M.R.K. (2016) A hybrid smart modeling approach for estimation of pure ionic liquids viscosity, *J. Mol. Liq.* **222**, 14–27
- 39 Ghiasi M.M., Arabloo M., Mohammadi A.H., Barghi T. (2016) Application of ANFIS soft computing technique in modeling the CO<sub>2</sub> capture with MEA, DEA, and TEA aqueous solutions, *Int. J. Greenh. Gas Control.* **49**, 47–54
- 40 Dadkhah M.R., Tatar A., Mohebbi A., Barati-Harooni A., Najafi-Marghmaleki A., Ghiasi M.M., Mohammadi A.H., Pourfayaz F. (2017) Prediction of solubility of solid compounds in supercritical CO<sub>2</sub> using a connectionist smart technique, *J. Supercrit. Fluids* **120**, 181–190
- 41 Tatar A., Barati A., Yarahmadi A., Najafi A., Lee M., Bahadori A. (2016) Prediction of carbon dioxide solubility in aqueous mixture of methyldiethanolamine and N-methylpyrrolidone using intelligent models, *Int. J. Greenh. Gas Control.* **47**, 122–136
- 42 Meybodi M.K., Shokrollahi A., Safari H., Lee M., Bahadori A. (2015) A computational intelligence scheme for prediction of interfacial tension between pure hydrocarbons and water, *Chem. Eng. Res. Des.* **95**, 79–92

- 43 Taghvaei H., Amooie M.A., Hemmati-Sarapardeh A., Taghvaei H. (2016) A comprehensive study of phase equilibria in binary mixtures of carbon dioxide + alcohols: Application of a hybrid intelligent model (CSA-LSSVM), *J. Mol. Liq. Part A* **224**, 745–756
- 44 Atashrouz S., Mirshekar H., Hemmati-Sarapardeh A. (2017) A soft-computing technique for prediction of water activity in PEG solutions, *Colloid Polym. Sci.* **295**, 421–432
- 45 Haifeng W., Dejin H. (2005) Comparison of SVM and LS-SVM for regression, in: *2005 International Conference on Neural Networks and Brain*, Vol. 1, ICNN & B'05. IEEE, pp. 279–283
- 46 Suykens J.A.K., Vandewalle J. (1999) Least squares support vector machine classifiers, *Neural Process. Lett.* **9**, 293–300
- 47 Suykens J.A., Van Gestel T., De Brabanter J., De Moor B., Vandewalle J., Suykens J., Van Gestel T. (2002) *Least Squares Support Vector Machines*, World Scientific, Singapore.
- 48 Jang J.S.R., Chuen-Tsai S. (1995) Neuro-fuzzy modeling and control, *Proc. IEEE* **83**, 378–406
- 49 Takagi T., Sugeno M. (1985) Fuzzy identification of systems and its applications to modeling and control, *IEEE Trans. Syst. Man Cybern.* **15**, 116–132
- 50 Ziaee H., Hosseini S.M., Sharafpoor A., Fazavi M., Ghiasi M. M., Bahadori A. (2015) Prediction of solubility of carbon dioxide in different polymers using support vector machine algorithm, *J. Taiwan Inst. Chem. Eng.* **46**, 205–213
- 51 Rostami A., Arabloo M., Esmaeilzadeh S., Mohammadi A.H. (2018) On modeling of bitumen/n-tetradecane mixture viscosity: Application in solvent-assisted recovery method, *Asia-Pac. J. Chem. Eng.* **13**, e2152 <https://doi.org/2110.1002/apj.2152>
- 52 Haykin S., Network N. (2004) A comprehensive foundation, *Neural Netw.* **2**, 41
- 53 Sayahi T., Tatar A., Bahrami M. (2016) A RBF model for predicting the pool boiling behavior of nanofluids over a horizontal rod heater, *Int. J. Therm. Sci.* **99**, 180–194
- 54 Wilamowski B.M., Jaeger R.C. (1996) Implementation of RBF type networks by MLP networks, in: *Neural Networks, 1996, IEEE International Conference on IEEE*, pp. 1670–1675
- 55 Tatar A., Shokrollahi A., Mesbah M., Rashid S., Arabloo M., Bahadori A. (2013) Implementing radial basis function networks for modeling CO<sub>2</sub>-reservoir oil minimum miscibility pressure, *J. Nat. Gas Sci. Eng.* **15**, 82–92
- 56 Nilsson N.J. (1965) *Learning Machines: Foundations of Trainable Pattern-Classifying Systems*, McGraw-Hill, United States
- 57 Haykin S.S. (2001) *Neural Networks: A Comprehensive Foundation*, Tsinghua University Press, Beijing, China
- 58 Sharkey A.J.C. (1996) On combining artificial neural nets, *Connect. Sci.* **8**, 299–314
- 59 Genest C., Zidek J.V. (1986) Combining probability distributions: A critique and an annotated bibliography, *Stat. Sci.* **1**, 114–135
- 60 Xu L., Krzyzak A., Suen C.Y. (1992) Methods of combining multiple classifiers and their applications to handwriting recognition, *IEEE Trans. Syst. Man Cybern.* **22**, 418–435
- 61 Jacobs R.A. (1995) Methods for combining experts' probability assessments, *Neural Comput.* **7**, 867–888
- 62 Perrone M.P., Cooper L.N. (1992) When networks disagree: Ensemble methods for hybrid neural networks, in: *DTIC Document*
- 63 Hashem S. (1993) Approximating a function and its derivatives using MSE-optimal linear combinations of trained feedforward neural networks, in: *Proc. 1993 World Congress on Neural Networks, 1*, pp. 617–620
- 64 Rostami A., Arabloo M., Lee M., Bahadori A. (2018) Applying SVM framework for modeling of CO<sub>2</sub> solubility in oil during CO<sub>2</sub> flooding, *Fuel* **214**, 73–87
- 65 Rostami A., Ebadi H., Arabloo M., Meybodi M.K., Bahadori A. (2017) Toward genetic programming (GP) approach for estimation of hydrocarbon/water interfacial tension, *J. Mol. Liq.* **230**, 175–189
- 66 Rostami A., Arabloo M., Ebadi H. (2017) Genetic programming (GP) approach for prediction of supercritical CO<sub>2</sub> thermal conductivity, *Chem. Eng. Res. Des.* **122**, 164–175
- 67 Rostami A., Arabloo M., Joonaki E., Ghanaatian S., Youzband A.H. (2017) Fast estimation of supercritical CO<sub>2</sub> thermal conductivity by a supervised learning machine-implications for EOR, in: *79th EAGE Conference and Exhibition 2017*
- 68 Rostami A., Anbaz M.A., Gahrooei H.R.E., Arabloo M., Bahadori A. (2017) Accurate estimation of CO<sub>2</sub> adsorption on activated carbon with multi-layer feed-forward neural network (MLFNN) algorithm, *Egypt. J. Petrol.* doi:10.1016/j.ejpe.2017.01.003 (in press)
- 69 Kamari A., Pournik M., Rostami A., Amirlatif A., Mohammadi A.H. (2017) Characterizing the CO<sub>2</sub>-brine interfacial tension (IFT) using robust modeling approaches: A comparative study, *J. Mol. Liq.* **246**, 32–38
- 70 Rostami A., Arabloo M., Kamari A., Mohammadi A.H. (2017) Modeling of CO<sub>2</sub> solubility in crude oil during carbon dioxide enhanced oil recovery using gene expression programming, *Fuel* **210**, 768–782
- 71 Rostami A., Ebadi H. (2017) Toward gene expression programming for accurate prognostication of the critical oil flow rate through the choke: Correlation development, *Asia-Pac. J. Chem. Eng.* **12**, 884–893
- 72 Rostami A., Shokrollahi A. (2017) Accurate prediction of water dewpoint temperature in natural gas dehydrators using gene expression programming approach, *J. Mol. Liq.* **243**, 196–204
- 73 Bahari M., Rostami A., Joonaki E., Ali M. (2014) Investigation of a novel technique for decline curve analysis in comparison with the conventional models, *Int. J. Comput. Appl.* **98**, 1–11
- 74 Rostami A., Baghban A. (2018) Application of a supervised learning machine for accurate prognostication of higher heating values of solid wastes, *Energy Sources, Part A: Recovery, Util. Environ. Eff.* **40**, 558–564
- 75 Karkevandi-Talkhoonchek A., Rostami A., Hemmati-Sarapardeh A., Ahmadi M., Husein M.M., Dabir B. (2018) Modeling minimum miscibility pressure during pure and impure CO<sub>2</sub> flooding using hybrid of radial basis function neural network and evolutionary techniques, *Fuel* **220**, 270–282
- 76 Gao C. (2014) Empirical correlations for viscosity of partially hydrolyzed Polyacrylamide, *J. Pet. Explor. Prod. Technol.* **4**, 209–213
- 77 Pelckmans K., Suykens J., Gestel T., Brabanter J., Lukas L., Hamers B., Moor B., Vandewalle J. (2002) A Matlab/c toolbox for least square support vector machines, in: *ESAT-SCD-SISTA Technical Report*, 02-145

- 78 Suykens J.A., De Brabanter J., Lukas L., Vandewalle J. (2002) Weighted least squares support vector machines: Robustness and sparse approximation, *Neurocomputing* **48**, 85–105
- 79 Hemmati-Sarapardeh A., Shokrollahi A., Tatar A., Gharagheizi F., Mohammadi A.H., Naseri A. (2014) Reservoir oil viscosity determination using a rigorous approach, *Fuel* **116**, 39–48
- 80 Cybenko G. (1989) Approximation by superpositions of a sigmoidal function, *Math. Control. Signals Syst.* **2**, 303–314
- 81 Hemmati-Sarapardeh A., Ameli F., Dabir B., Ahmadi M., Mohammadi A.H. (2016) On the evaluation of asphaltene precipitation titration data: Modeling and data assessment, *Fluid Phase Equilib.* **415**, 88–100
- 82 Gramatica P. (2007) Principles of QSAR models validation: Internal and external, *QSAR Comb. Sci.* **26**, 694–701
- 83 Goodall C.R. (1993) Computation using the QR decomposition, in: *Handbook of Statistics*, Elsevier, Amsterdam, Netherland, pp. 467–508
- 84 Eslamimanesh A., Gharagheizi F., Mohammadi A.H., Richon D. (2011) Phase equilibrium modeling of structure H clathrate hydrates of methane + water “insoluble” hydrocarbon promoter using QSPR molecular approach, *J. Chem. Eng. Data* **56**, 3775–3793
- 85 Chen G., Fu K., Liang Z., Sema T., Li C., Tontiwachwuthikul P., Idem R. (2014) The genetic algorithm based back propagation neural network for MMP prediction in CO<sub>2</sub>-EOR process, *Fuel* **126**, 202–212