



# Hybrid connectionist model determines CO<sub>2</sub>–oil swelling factor

Mohammad Ali Ahmadi<sup>1</sup> · Sohrab Zendehboudi<sup>1</sup> · Lesley A. James<sup>1</sup>

Received: 17 October 2017 / Published online: 26 April 2018  
© The Author(s) 2018

## Abstract

In-depth understanding of interactions between crude oil and CO<sub>2</sub> provides insight into the CO<sub>2</sub>-based enhanced oil recovery (EOR) process design and simulation. When CO<sub>2</sub> contacts crude oil, the dissolution process takes place. This phenomenon results in the oil swelling, which depends on the temperature, pressure, and composition of the oil. The residual oil saturation in a CO<sub>2</sub>-based EOR process is inversely proportional to the oil swelling factor. Hence, it is important to estimate this influential parameter with high precision. The current study suggests the predictive model based on the least-squares support vector machine (LS-SVM) to calculate the CO<sub>2</sub>-oil swelling factor. A genetic algorithm is used to optimize hyperparameters ( $\gamma$  and  $\sigma^2$ ) of the LS-SVM model. This model showed a high coefficient of determination ( $R^2 = 0.9953$ ) and a low value for the mean-squared error (MSE = 0.0003) based on the available experimental data while estimating the CO<sub>2</sub>-oil swelling factor. It was found that LS-SVM is a straightforward and accurate method to determine the CO<sub>2</sub>-oil swelling factor with negligible uncertainty. This method can be incorporated in commercial reservoir simulators to include the effect of the CO<sub>2</sub>-oil swelling factor when adequate experimental data are not available.

**Keywords** CO<sub>2</sub> injection · CO<sub>2</sub> swelling · Genetic algorithm · Predictive model · Least-squares support vector machine

## 1 Introduction

Due to the growing concern about global warming and the ongoing demand for energy resources, CO<sub>2</sub>-based enhanced oil recovery (EOR) methods have been attracting both the scientific and industrial interests. When CO<sub>2</sub> is injected into depleted oil reservoirs, different mechanisms contribute to oil production (Farajzadeh et al. 2009; Godec et al. 2013; Kuznetsova and Kvamme 2002; Ma et al. 2016). These mechanisms depend on the operational conditions and oil composition. The most common oil production mechanisms in CO<sub>2</sub>-based EOR methods are oil viscosity reduction, oil swelling, condensation, vaporization, and interfacial tension (IFT) reduction (Abedini and Torabi 2014; Ahmadi et al. 2015; Bachu 2016; Czarnota

et al. 2017; Farajzadeh et al. 2009; Li et al. 2013b, 2015; Shelton et al. 2016; Yang et al. 2012). Reducing the level of CO<sub>2</sub> emissions in the atmosphere by the use of geological CO<sub>2</sub> storage in depleted oil reservoirs as well as its role in the oil recovery processes highlights the importance of further studies of CO<sub>2</sub> injection operations and the corresponding PVT behavior (Ahmadi et al. 2016a, b; Bachu 2016; Davis et al. 2010; Jamali and Etehadtavakkol 2017; Kim and Santamarina 2014; Li et al. 2015; Li and Fan 2015; Liu and Wilcox 2012; Luo et al. 2013; Orr et al. 1982; Sell et al. 2012; Shelton et al. 2016; Yang et al. 2012; Yu et al. 2015; zeinali Hasanvand et al. 2013).

According to Rojas and Ali (1986) and Tunio et al. (2011), there are four effective mechanisms contributing to oil production using CO<sub>2</sub>-enhanced oil recovery strategies including (1) oil viscosity reduction, (2) oil swelling, (3) oil and water density reduction, and (4) vaporization and extraction of a portion of oil. It is clear that when CO<sub>2</sub> is dissolved in the oil phase, the oil swells and its viscosity reduces. Hence, the variation in the swelling factor allows CO<sub>2</sub> to substantially expand oil, which eventually improves the oil displacement and recovery (Perera et al. 2016). The immiscible CO<sub>2</sub>-EOR technique is dominated by the oil swelling phenomenon and oil viscosity reduction. The degree of oil swelling and oil viscosity change are

Edited by Yan-Hua Sun

✉ Mohammad Ali Ahmadi  
m.a.ahmadi@mun.ca

✉ Sohrab Zendehboudi  
szendehboudi@mun.ca

<sup>1</sup> Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL A1C 5S7, Canada

dependent on different parameters including CO<sub>2</sub> solubility in oil, pressure, temperature, and API degree of oil samples. CO<sub>2</sub> solubility is generally considered as the most significant factor that influences the efficiency of CO<sub>2</sub>-based EOR techniques, particularly under low-pressure conditions. For instance, this mechanism was confirmed through implementation of pilot-scale tests in Turkey (Bagci 2007; Issever and Topkaya 1998; Perera et al. 2016).

Experimental investigations and numerical reservoir simulations on binary systems including hydrocarbons and CO<sub>2</sub> were conducted to improve the hydrocarbon recovery (Bachu 2016; Bessières et al. 2001; Diep et al. 1998; Do and Pinczewski 1991; Fukai et al. 2016; Jamali and Etehdadvakkol 2017; Kim and Santamarina 2014; Kiran et al. 1996; Kwak and Kim 2017; Li et al. 2013a, 2015; Li and Fan 2015; Luo et al. 2007, 2013; Lv et al. 2015; Mulliken and Sandler 1980; Shelton et al. 2016; Yang and Gu 2005). Most of these studies investigated the oil swelling effect primarily as a result of CO<sub>2</sub> dissolution in the light fractions of oil. Bessières et al. (2001) and Kiran et al. (1996) examined the variation in the volume of several CO<sub>2</sub>-alkane systems. They concluded that the excess volume follows a sigmoidal change with the composition/concentration of CO<sub>2</sub>. The oil swelling effect was determined by the volume swelling coefficient defined by Yang and Gu (2005) and Yang et al. (2012). These investigations reveal that with an increase in the pressure (and consequently the solubility of CO<sub>2</sub> in oil), the volume swelling coefficient of the oil phase increases. Yang et al. (2012) studied the behavior of oil swelling through qualitative analysis of the dispersion of CO<sub>2</sub> in oil. Experiments at reservoir conditions (high temperature, high pressure, and live oil composition) are, however, challenging. A swelling/extraction experiment is a well-known technique to record composition and reservoir fluid volume changes due to CO<sub>2</sub> dissolution in reservoir oil at a given temperature. Swelling experiments are typically carried out in a high-pressure-resistant visual PVT cell with a constant volume, which is first filled with a specific volume of dead or stock-tank oil (Tsau et al. 2010). Depending on the number of steps required to reach the desired pressure, CO<sub>2</sub> is injected gradually to achieve a proper pressure increase. The main assumption of the swelling experiment is neglecting the vaporization of intermediate components of oil into the CO<sub>2</sub> phase until reaching the minimum miscibility pressure. The oil volume change owing to the swelling effect at each pressure step is recorded and the amount of CO<sub>2</sub> dissolved in the oil is measured. An increase in the pressure results in vaporization of a part of oil components, and the oil-rich phase shrinks. It should be noted that the phase behavior of the CO<sub>2</sub> and oil system can be visually detected using a swelling test. Various parameters including the

bubble point pressure, solubility of CO<sub>2</sub>, and swelling factor are usually employed to tune the equation of state (EOS) for the phase behavior modeling purposes (Tsau et al. 2010). Different sizes of visual PVT cells can be utilized for swelling experiments; these cells include 140 mL (Hand and Pinczewski 1990), 170 mL (Harmon and Grigg 1988), 190 mL (Orr et al. 1981), and 450 mL. Holm and Josendal (1982) recommended that a volume of 30% of the cell volume should be considered as the sample size for the swelling test. Therefore, the proper volume range is 40–100 mL of the crude oil sample to perform the swelling tests using the corresponding PVT cells. The most important issue with the sampling size is the time needed to achieve an equilibrium condition after each pressure change. The mixing process of large volumes of gas and oil at a given pressure seems to be another major concern in such a swelling test.

Thomas and Monger-McClure (1991) studied the effect of the CO<sub>2</sub>-oil swelling factor on oil recovery from light oil reservoirs using cyclic CO<sub>2</sub> injection. They correlated the oil incremental value to the CO<sub>2</sub>-oil swelling factor. Based on the results, an increase in the CO<sub>2</sub>-oil swelling factor led to an increase in the amount of produced oil (Thomas and Monger-McClure 1991).

Dong et al. (2001) determined the CO<sub>2</sub>-oil swelling factor by comparing the measured densities of the dead oil sample, reservoir live oil, and mixture of CO<sub>2</sub> and reservoir oil. Ghedan (2009) claimed that at high CO<sub>2</sub> concentrations the CO<sub>2</sub>-oil swelling factor will be 1.25–1.6; in most of the cases, the CO<sub>2</sub> content should be greater than 50%. Ning et al. (2011) carried out several multiple contact experiments (MCEs) to figure out the contribution of oil swelling as well as reduction in oil viscosity to the oil production from Alaska North Slope viscous oil. Heidaryan and Moghadasi (2012) investigated the influence of swelling and viscosity reduction on oil production using both experimental and theoretical methods. Based on their research outcome, they concluded that the optimum value of the CO<sub>2</sub>-oil swelling factor should be 1.7 to reach the maximum oil production from the reservoir (Heidaryan and Moghadasi 2012).

Through a systematic research work, Sugai et al. (2014) experimentally determined oil swelling factors in porous media using two different types of micromodels (e.g., fine beads and coarse beads). They investigated the effect of interfacial area on the oil swelling and CO<sub>2</sub>-oil swelling factors. They used a digital camera to take images to determine the amount of oil trapped in the micromodels at different times. They obtained the swelling factor from the tests after a constant saturation degree in the porous systems was confirmed. In addition, they employed an oil-CO<sub>2</sub> simple contact model in a visual cell to determine CO<sub>2</sub>-oil swelling factors at different pressures via utilizing

a digital camera and an image processing method. They compared CO<sub>2</sub>–oil swelling factors from both types of the experiments to decide what other parameters should be taken into account to further improve the accuracy and reliability of the existing approach. According to the experimental results, they concluded that an increase in the interfacial area results in increasing the oil swelling. In other words, the swelling factor in the case of the fine bead micromodel was larger than that in the coarse bead micromodel due to an increase in the interfacial area (Sugai et al. 2014). Or et al. (2016) experimentally investigated the contribution of CO<sub>2</sub>–oil swelling and viscosity reduction to the oil recovery through implementation of CO<sub>2</sub> gas foaming in heavy oil reservoirs. It was concluded that CO<sub>2</sub> foam swelling increases with an increase in the pressure drawdown in a well. Also, further swelling of foamy oil can mobilize the residual oil towards the producer well, especially in the immobilized zone (Or et al. 2016).

Habibi et al. (2017) carried out experiments on CO<sub>2</sub>–oil systems to determine the interaction between CO<sub>2</sub> and oil in tight rock samples. They conducted constant composition experiments (CCEs) to determine the CO<sub>2</sub>–oil swelling factor and other measurable fluid and thermodynamic characteristics. Also, they performed CO<sub>2</sub> cyclic injection experiments to determine the amount of oil recovery. The CO<sub>2</sub>–oil swelling factor in their study was defined as “the volume of the oil after CO<sub>2</sub> injection divided by the volume of the oil before CO<sub>2</sub> injection into the cell.” In their experiments, increasing CO<sub>2</sub> concentration from 48.4% to 71.1% resulted in an increase in the CO<sub>2</sub>–oil swelling factor from 1.21 to 1.39, respectively. According to their experimental data, the oil swelling and expansion, CO<sub>2</sub> dissolution into the oil, and CO<sub>2</sub> diffusion into core samples are the main mechanisms contributing to the oil production (Habibi et al. 2017).

There are a few studies that have developed a reliable correlation or a deterministic model for predicting CO<sub>2</sub>–oil swelling factors. Welker (1963) proposed a very simple correlation to estimate the CO<sub>2</sub>–oil swelling factor. Their correlation suffers from the lack of applicability, particularly for light and intermediate crude oil samples. Simon and Graue (1965) developed a graphical method to determine the oil swelling factor. Their method was developed based on limited data samples from heavy crudes. Chung et al. (1988) proposed a simple correlation to estimate the oil swelling factor for CO<sub>2</sub>/heavy crude oil systems. Emera and Sarma (2006) developed a correlation to forecast the oil swelling factor for both light and heavy crude oils. However, they utilized a limited number of data points while developing their correlation. Table 1 demonstrates a summary of correlations and models to calculate the CO<sub>2</sub>–oil swelling factor.

Vapnik (1998) proposed the support vector machine (SVM) as an extended version of conventional artificial intelligent tools. SVM is a practical method which has been widely used for classification, regression, and pattern recognition (Cortes and Vapnik 1995). The principle idea of SVM is to transform the nonlinear input space to a higher-dimensional feature space to find a hyperplane via nonlinear mapping (Baylar et al. 2009; Cortes and Vapnik 1995). It is based on the statistical learning theory (SLT) and structural risk minimization (SRM) concepts (Mehdizadeh and Movagharnejad 2011). SVM tools obtain the solution via solving the quadratic programming (QP); the SVM always results in a global optimum solution, unlike other regression techniques such as neural networks, as the QP problem is a convex function (Vong et al. 2006). However, it suffers from computational burden.

The LS-SVM has not been used to model the CO<sub>2</sub>–oil swelling factor in the literature, to the best of our knowledge. This study employs the least-squares support vector machine (LS-SVM) paradigm, as a hybridized version of the original SVM method, to calculate the CO<sub>2</sub>–oil swelling factor. A genetic algorithm (GA) is utilized as an optimization technique to optimize the hyperparameters of the LS-SVM model. Through the comprehensive literature review, extensive experimental data are used for model development and validation.

## 2 Theory

### 2.1 Least-squares support vector machine (LS-SVM)

Suykens and Vandewalle (1999) proposed the least-squares support vector machine (LS-SVM) model as an alternate formulation of the SVM regression. The LS-SVM enjoys similar advantages as SVM. Also, it requires solving only a set of linear equations instead of a quadratic programming (QP) problem, which is computationally less demanding.

Given the training set  $\{x_k, y_k\}$ ,  $k = 1, 2, \dots, N$ , where  $x_k \in \mathbb{R}^n$  is the  $k$ th input data in the input space and  $y_k \in \mathbb{R}$  represent the output variable for the given input variable (i.e.,  $x_k$ ) and  $N$  refers to the number of the training samples. Using a nonlinear function  $\varphi(\cdot)$ , which maps the training set in the input space to a high (and possibly infinite)-dimensional space, the following regression model is constructed:

$$y = \omega^T \varphi(x) + b \quad \text{with} \quad \omega \in \mathbb{R}^n, \quad b \in \mathbb{R}, \quad (1)$$

$$\varphi(\cdot) \in \mathbb{R}^n \rightarrow \mathbb{R}^{n_h}, \quad n_h \rightarrow \infty$$

in which,  $\omega$  denotes the weight vector and  $b$  signifies a bias term. Note that the superscript “ $n$ ” refers to the dimension of data space and “ $n_h$ ” is attributed to the higher-

**Table 1** Correlations and models for calculating CO<sub>2</sub>–oil swelling factor (SF)

Correlation	Considerations/limitations	References
$SF = 1.0 + \frac{0.35(\text{solubility}(\text{scf}/\text{bbl}))}{1000}$	Developed for oils at $T = 80$ °F and $20^\circ$ API < oil gravity < $40^\circ$ API	Welker (1963)
Graphical correlation: This model is a function of CO <sub>2</sub> solubility, oil molecular weight ( $MW$ ), and oil density at 60 °F. Not recommended for high-pressure ranges	$P < 2300$ psi $110$ °F < $T < 250$ °F $12^\circ$ API < oil gravity < $33^\circ$ API	Simon and Graue (1965)
$SF = \frac{\rho_1}{\rho - S}$ $S = \text{CO}_2$ solubility, g/cm <sup>3</sup> $\rho =$ oil density without CO <sub>2</sub> at the same temperature and 1 atm pressure, g/cm <sup>3</sup> $\rho_1 =$ solution density, g/cm <sup>3</sup>	$16.89^\circ$ API oil gravity $75$ °F < $T < 200$ °F $14.7$ psi < $P < 5014.7$ psi	Chung et al. (1988)
For $MW > 300$ $SF = 1 + 0.3302Y - 0.8417Y^2 + 1.5804Y^3 - 1.074Y^4 - 0.0318Y^5 + 0.21755Y^6$	$23$ °C < $T < 121.1$ °C $0.1$ MPa < $P < 27.4$ MPa $12^\circ$ API < oil gravity < $37^\circ$ API	Emera and Sarma (2006)
For $MW < 300$ $SF = 1 + 0.48411Y - 0.9928Y^2 + 1.6019Y^3 - 1.2773Y^4 + 0.48267Y^5 - 0.06671Y^6$		
$Y = 1000 \times \left( \left( \frac{\gamma}{MW} \right) \times \text{sol}(\text{mole fraction})^2 \right)^{\exp\left(\frac{\gamma}{MW}\right)}$		

dimensional feature space (Vong et al. 2006). When the LS-SVM is applied, a new optimization case will be generated. The implemented strategy deals with the following optimization problem:

$$\min_{\omega, b, e} \mathcal{J}(\omega, e) = \frac{1}{2} \omega^T \omega + \frac{1}{2} \gamma \sum_{k=1}^N e_k^2 \quad (2)$$

subject to the following equality constraint:

$$y_k = \omega^T \varphi(x_k) + b + e_k \quad k = 1, 2, \dots, N \quad (3)$$

where  $\gamma$  represents the regularization parameter, which compromises between the model's complexity and the training error (Mehdizadeh and Movagharnjad 2011), and  $e_k$  is the regression error. The Lagrangian is constructed as follows in order to find the solution of the un-constrained optimization problem:

$$\mathcal{L}(\omega, b, e, \alpha) = \mathcal{J}(\omega, e) - \sum_{k=1}^N \alpha_k \{ \omega^T \varphi(x_k) + b + e_k - y_k \} \quad (4)$$

where  $\alpha_k$  stands for the Lagrange multiplier or support value. To attain the solution of Eq. (4), differentiating the equation with respect to  $\omega, b, e_k, \alpha_k$  gives:

$$\frac{\partial \mathcal{L}(\omega, b, e, \alpha)}{\partial \omega} = 0 \rightarrow \omega = \sum_{k=1}^N \alpha_k \varphi(x_k) \quad (5)$$

$$\frac{\partial \mathcal{L}(\omega, b, e, \alpha)}{\partial b} = 0 \rightarrow \sum_{k=1}^N \alpha_k = 0 \quad (6)$$

$$\frac{\partial \mathcal{L}(\omega, b, e, \alpha)}{\partial e_k} = 0 \rightarrow \alpha_k = \gamma e_k, \quad k = 1, \dots, N \quad (7)$$

$$\frac{\partial \mathcal{L}(\omega, b, e, \alpha)}{\partial \alpha_k} = 0 \rightarrow y_k = \varphi(x_k) \omega^T + b + e_k, \quad k = 1, 2, \dots, N \quad (8)$$

After removing the variables  $\omega$  and  $e$ , one acquires the Karush–Kuhn–Tucker system as follows:

$$\begin{bmatrix} 0 & 1_v^T \\ 1_v & \Omega + \gamma^{-1}I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (9)$$

In Eq. (9),  $y = [y_1 \dots y_N]^T$ ,  $1_N = [1 \dots 1]^T$ ,  $\alpha = [\alpha_1 \dots \alpha_N]^T$ ,  $I$  is an identity matrix, and  $\Omega_{kl} = \varphi(x_k)^T \cdot \varphi(x_l) = K(x_k, x_l) \forall k, l = 1, 2, \dots, N$ .  $K(x_k, x_l)$  is the kernel function and must meet Mercer's condition (Li et al. 2008). Three typical choices for the kernel function are:

- $K(x, x_k) = x_k^T x$
- $K(x, x_k) = (\tau + x_k^T x)^d$
- $K(x, x_k) = \exp(-x - x_k^2 / \sigma^2)$

The resulting formulation of LS-SVM model for function estimation becomes:

$$y(x) = \sum_{k=1}^N \alpha_k K(x, x_k) + b \quad (10)$$

where  $\tau$  refers to the slope,  $d$  stands for the polynomial degree,  $\sigma^2$  is the kernel sample variance, and  $(b, \alpha)$  represents the solution to the linear system of equations shown in Eq. (9).

In the literature, some comprehensive descriptions of the SVM are available (Borges 1998; Suykens and Vandewalle 1999; Vapnik 1998). The theory of LS-SVM is systematically explained by a number of researchers (Suykens and Vandewalle 1999; Suykens et al. 2002). Also, Liu et al. (2005a, b, 2007) provide a detailed comparison of the SVM and LS-SVM methods.

### 2.2 Genetic algorithm

Genetic algorithm (GA) is a stochastic method to solve optimization problems involving a fitness criterion, survival of the fittest, and different genetic operators, including crossover and mutation to satisfy a pre-defined fitness quantity, resembling the Darwinian evolution by natural selection (Niazi et al. 2008). The significant feature of the GAs and the other similar evolutionary algorithms is that they are derivative-free. The stochastic nature of the algorithm with dynamic evaluation of the fitness function brings a powerful systematic random search engine. This approach is an alternative to derivative-based methods to deal with problems in which the fitness function is non-differentiable, discontinuous, highly nonlinear, with multiple local optima, or stochastic (Reihanian et al. 2011).

### 3 Data gathering

Extensive data points for the CO<sub>2</sub>–oil swelling factor have been extracted from the literature (Abedini et al. 2014; Chung et al. 1988; Mosavat et al. 2014; Tsau et al. 2010; Wei et al. 2017). The statistical parameters for these data samples are reported in Table 2. As it is clear from Table 2, the data samples contain a broad range of crude oils from heavy oils to extra-light oil samples. The collected data also cover a wide range of temperature, pressure, and CO<sub>2</sub> solubility.

**Table 2** Statistical parameters of the data points (Abedini et al. 2014; Chung et al. 1988; Mosavat et al. 2014; Tsau et al. 2010; Wei et al. 2017) used for developing LS-SVM model

Parameter	Minimum	Maximum	Average
Oil gravity, API degrees	16.8	46.11	32.8
Temperature $T$ , °F	68	200	109.5
Pressure $P$ , psi	14.7	4100	1187.6
CO <sub>2</sub> concentration (mole fraction)	0	0.86	0.525

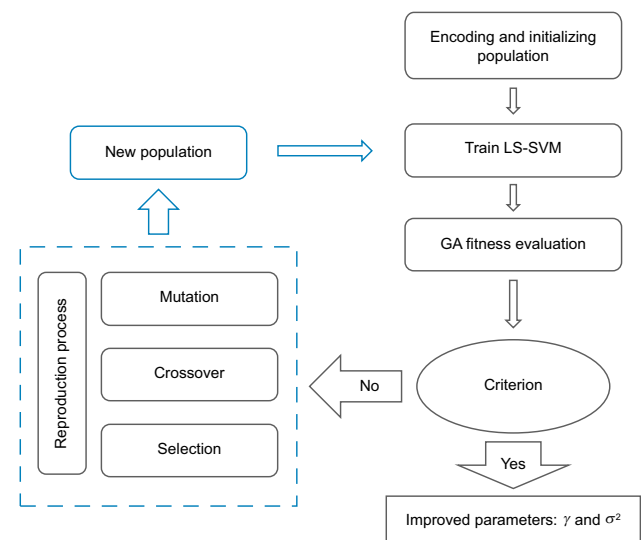
## 4 Methodology

In this paper, four parameters are considered as input variables to the LS-SVM model. These parameters are (1) CO<sub>2</sub> concentration in oil (mole fraction of CO<sub>2</sub>), (2) pressure, (3) temperature, and (4) the oil API degree. The output variable from the LS-SVM model is the CO<sub>2</sub>–oil swelling factor.

A total number of 225 data samples were extracted from the literature to develop our LS-SVM model to estimate the CO<sub>2</sub>–oil swelling factor. The data samples were divided into two data sets. The first set (also called the training data series) contained 80% of the total data points to construct the LS-SVM model. The second set of data contained 20% of the entire data points employed to validate the LS-SVM model.

The radial basis function (RBF) was selected because of its promising performance and simplicity as it only contains one adjustable parameter and has been successfully applied (Ahmadi 2015; Keerthi and Lin 2003; Reihanian et al. 2011). In the model development using LS-SVM with the RBF kernel function, according to Eqs. (9) and (10), the optimization of  $\gamma$  and  $\sigma^2$  is a crucial task. It was confirmed that the optimal magnitudes of these two vital parameters are required to better design a LS-SVM model towards greater precision and generalization (Vong et al. 2006).

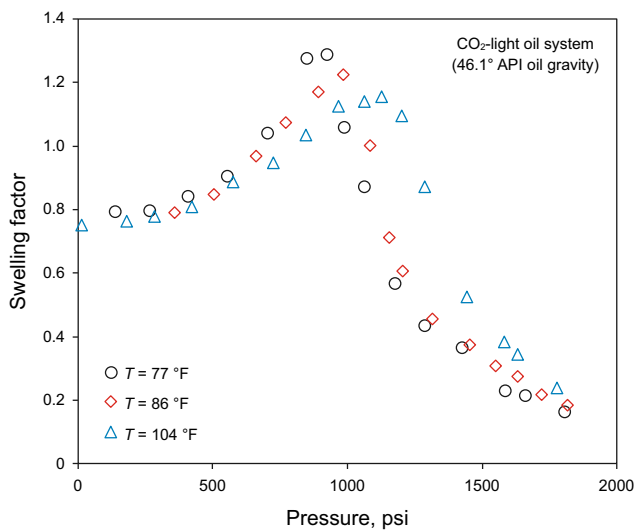
According to Ahmadi and Ebadi (2014), Ahmadi et al. (2014a, b), and Fazeli et al. (2013), the application of non-population-based optimization methods such as simulated annealing and Levenberg–Marquardt (LM) is not recommended due to their drawback in handling the nonlinearity in SVM methods. GA was applied in this research study to optimize the parameters of LS-SVM ( $\gamma$  and  $\sigma^2$ ) and the



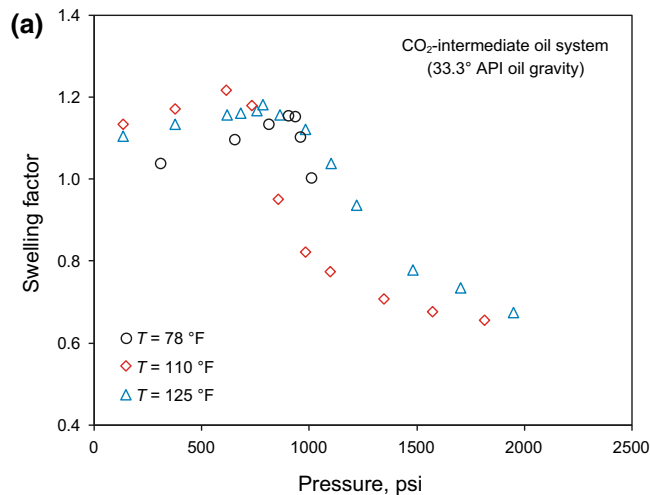
**Fig. 1** Flowchart of hyperparameters selection based on GA



average absolute relative deviation (AARD). The flowchart for the hyperparameter optimization using a GA algorithm is depicted in Fig. 1. The optimization procedure was repeated several times to attain the most plausible solution corresponding to the global optimum of the fitness function. As a result, values of  $\sigma^2$  and  $\gamma$  were obtained: 0.268829 and 33.4091, respectively.



**Fig. 2** Swelling factor of CO<sub>2</sub>–light oil system versus pressure at various temperatures (Abedini et al. 2014; Chung et al. 1988; Mosavat et al. 2014; Tsau et al. 2010; Wei et al. 2017), light oil with an API gravity of 46.11°



## 5 Results and discussion

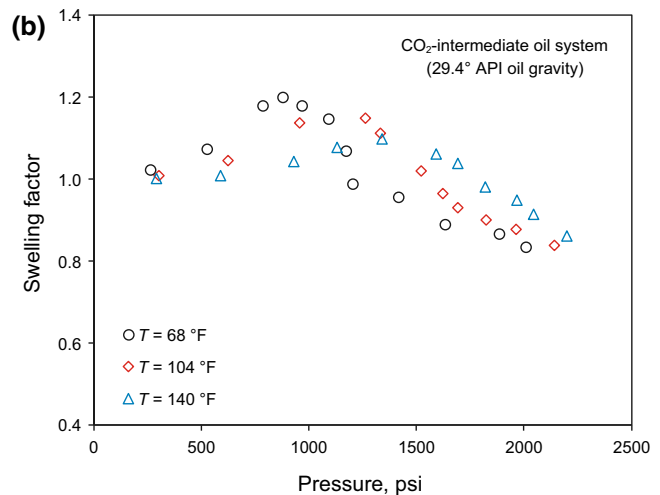
This study presents a new deterministic approach to obtain the swelling factor with higher accuracy. The oil swelling factor for the system of CO<sub>2</sub> and light oil versus pressure at different temperatures is demonstrated in Fig. 2. The variations of the oil swelling factor with pressure at various temperatures are shown in Figs. 3 and 4 for intermediate and heavy oil samples, respectively.

Mean-squared error (MSE) and coefficient of determination ( $R^2$ ) are employed in this statistical analysis as the performance evaluation criteria for the LS-SVM model in estimating the CO<sub>2</sub>–oil swelling factor. The expressions to obtain MSE and  $R^2$  are given below:

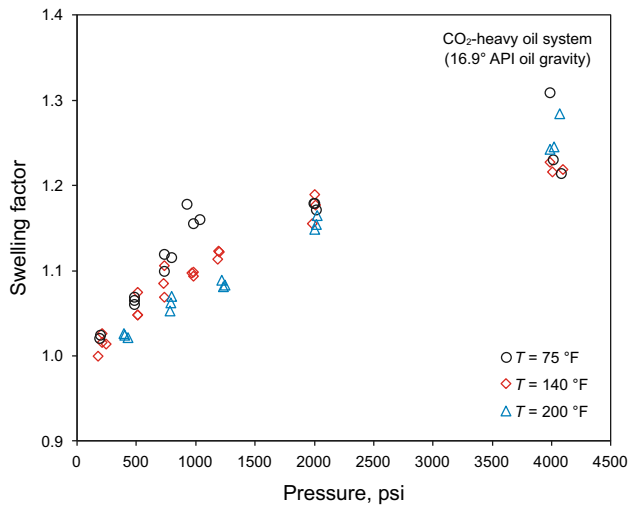
$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N \left( y_i^{\text{actual}} - y_i^{\text{predicted}} \right)^2 \quad (11)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N \left( y_i^{\text{actual}} - y_i^{\text{predicted}} \right)^2}{\sum_{i=1}^N \left( y_i^{\text{actual}} - \bar{y}^{\text{actual}} \right)^2} \quad (12)$$

where  $N$  represents the number of data points,  $y_i^{\text{actual}}$  denotes the  $i$ th observation (real data),  $y_i^{\text{predicted}}$  is the  $i$ th output from the model, and  $\bar{y}^{\text{actual}}$  signifies the average magnitudes of observations. The values of MSE and  $R^2$  are tabulated in Table 3 for training, testing, and overall data stages. The GA-LS-SVM predictions are satisfactory if  $R^2$  and MSE are close to 1 and 0, respectively. As can be seen in Table 3, these criteria were fulfilled.



**Fig. 3** Variations of swelling factor of CO<sub>2</sub>–intermediate oil system with pressure and temperature (Abedini et al. 2014; Chung et al. 1988; Mosavat et al. 2014; Tsau et al. 2010; Wei et al. 2017): **a** An oil with an API gravity of 33.3° and **b** An oil with an API gravity of 29.4°



**Fig. 4** Swelling factor of CO<sub>2</sub>-heavy oil system versus pressure at different temperatures (Abedini et al. 2014; Chung et al. 1988; Mosavat et al. 2014; Tsau et al. 2010; Wei et al. 2017), heavy oil with an API gravity of 16.9°

Figure 5 depicts a comparison between the experimental data for the CO<sub>2</sub>-oil swelling factor and the values estimated by the LS-SVM. Figure 5a shows a comparison between the estimated and experimental data in the training phase. Figure 5b demonstrates a comparison between the actual and predicted CO<sub>2</sub>-oil swelling factor behavior against the data index in the testing phase. As illustrated in Fig. 5, there is an excellent match between the oil swelling factor estimated from the LS-SVM method and those from experiments.

Figure 6 illustrates the regression plot between the CO<sub>2</sub>-oil swelling factor determined by LS-SVM model and the experimental data points. Figure 6a depicts the scatter plot for results obtained in the training phase of the LS-SVM model. As shown in Fig. 6a, the linear fit to data  $y = 0.9892x + 0.0103$  has a high correlation of coefficient ( $R^2 = 0.9944$ ), meaning that the training phase of the LS-SVM model is performed very well. The results achieved over the testing (validation) phase are displayed in the form of a scatter plot in Fig. 6b, based on the developed LS-SVM tool. As depicted in Fig. 6b, the high value of the correlation coefficient ( $R^2 = 0.9931$ ) between the predicted and experimental oil swelling factor shows the superior

performance of the LS-SVM model. Figure 6c illustrates the regression plot for the whole data set. The predicted swelling factor values are found to be scattered around the  $y = x$  line, indicating that the LS-SVM model that is optimized by GA predicts the swelling factor very well.

Figure 7 represents a comparison between the CO<sub>2</sub>-oil swelling factor determined by the LS-SVM model and the real data versus pressure at different temperatures. As shown in Fig. 7, the LS-SVM model follows the trend of experimental data points for an immediate oil of 29.4° API gravity. As the experimental data points show, at a constant pressure, the magnitude of swelling factor lowers with increasing the temperature. This behavior was confirmed by the LS-SVM model. This implies that the proposed LS-SVM model for determination of CO<sub>2</sub>-oil swelling factor is valid and acceptable in terms of technical and conceptual prospects.

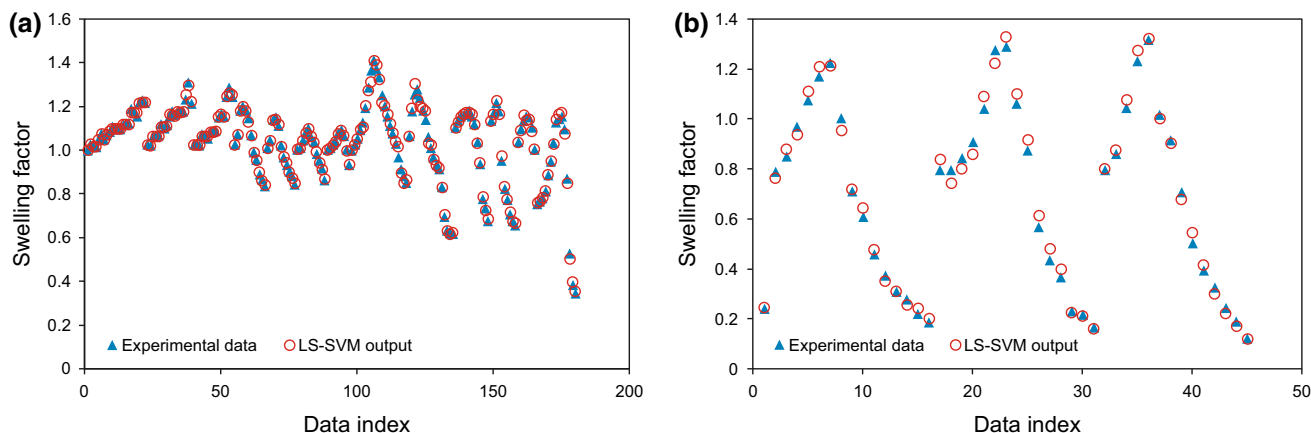
Figure 8 shows the relative error distribution for both the training and testing phases in developing the LS-SVM model. According to Fig. 8, the maximum relative error between the outputs of the LS-SVM model and the experimental CO<sub>2</sub>-oil swelling factors is within  $\pm 5\%$  for the training phase. Also, the maximum relative error between the CO<sub>2</sub>-oil swelling factor calculated by the LS-SVM model and experimental data points is within  $\pm 15\%$  for the testing phase.

Simon and Graue (1965) proposed a graphical method for determination of the CO<sub>2</sub>-oil swelling factor. In this method, the minimum value of the CO<sub>2</sub>-oil swelling factor is equal to 1 and the maximum value is equal to 1.38. Also, the Simon and Graue technique offers acceptable values for swelling factor within the limited ranges of API, temperature, and CO<sub>2</sub> solubility (Table 1). Hence, this graphical method is not able to provide reliable outputs over wide ranges of the input parameters. Figure 9 demonstrates the scatter plot of the results obtained by the graphical method proposed by Simon and Graue (1965) versus the experimental values of the CO<sub>2</sub>-oil swelling factor. As it is clear from Fig. 9, the linear fit line has a low correlation coefficient ( $R^2$ ). Also, the linear fit has a negative slope, concluding that the value of oil swelling factor at the lower boundary is overestimated.

**Table 3** Performance of GA-LS-SVM method with optimized parameters for prediction of swelling factor in terms of statistical parameters

Statistical parameters	Training data	Testing data	Overall data
MSE	0.00016	0.0009	0.0003
$R^2$	0.9944	0.9931	0.9953
Average absolute relative deviation (AARD) <sup>a</sup>	0.7918	4.549	1.5433
Maximum absolute error (MAE)	5.3403	5.4205	5.4205

$$^a \text{AARD} = \frac{\sum \left| \frac{y_{\text{actual}} - y_{\text{predicted}}}{y_{\text{actual}}} \right|}{n}$$



**Fig. 5** Comparison between estimated and measured swelling factors versus data index: **a** Training data and **b** Testing data

Figure 10 presents a comparison between the objective function values calculated by Emera and Sarma (2006) correlation and the real data of the CO<sub>2</sub>–oil swelling factor. Based on Figs. 9 and 10, the linear fit of the data obtained by Emera and Sarma (2006) correlation has a higher value of correlation of coefficient in comparison with the method proposed by Simon and Graue (1965). This is because the correlation introduced by Emera and Sarma (2006) was developed using a wider range of data points. However, this correlation still suffers from the common drawback for the most empirical correlations which can offer reliable outputs within limited ranges of input parameters (Table 1). As illustrated in Fig. 10, the Emera and Sarma (2006) correlation underestimates the magnitudes of the swelling factor in the middle range of the data.

Table 4 reports the maximum absolute error (MAE) and the average absolute relative deviation (AARD) for three different models based on the experimental data available for the CO<sub>2</sub>–oil swelling factor. The MAE of the LS-SVM model is lower, compared to the Emera and Sarma (2006) and Simon and Graue (1965) methods. This superior performance is attributed to the high predictive capability of the developed tool, proper procedure for the training phase, and careful selection of data samples. Using a broader range of data samples enables us to develop a more precise and reliable technique to calculate the CO<sub>2</sub>–oil swelling factor.

It should be noted the correlation proposed by Emera and Sarma (2006) is currently being used in the Computer Modelling Group (CMG) reservoir simulator package. It suggests that the LS-SVM strategy introduced in this research work can be included in the commercial reservoir simulators for various applications such as simulation of gas injection processes in the petroleum industry.

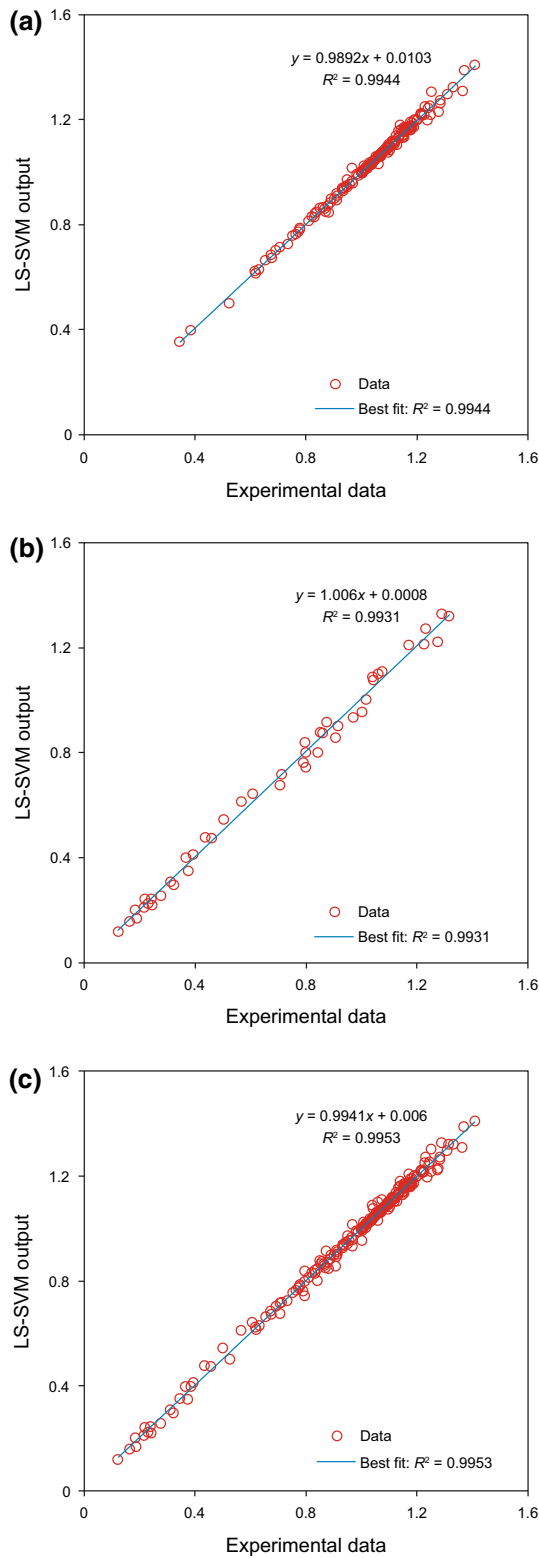
Appropriate statistical methods for identifying the applicability of the model are required for outlier detection. Recognition of outliers is to determine which data points

may differ from the bulk of the data present in the data bank under study (Gramatica 2007; Rousseeuw and Leroy 2005). For examining the capability of the LS-SVM model, the approach of Leverage Value Statistics has been carried out (Goodall 1993; Gramatica 2007). A graphical method (William plot) is used for outlier determination in this research work. The William plot depicts the standardized residual of the outputs versus corresponding hat ( $H$ ) values. Further details on the mathematical backgrounds and computational procedure of the William method can be found in the literature (Goodall 1993; Gramatica 2007; Rousseeuw and Leroy 2005). Figure 11 represents the William plot for the results obtained from the LS-SVM model while estimating the CO<sub>2</sub>–oil swelling factor. Having the majority of data points in the ranges of  $0 \leq H \leq 0.055$  and  $-3 \leq R \leq 3$  reveals that the LS-SVM model is convincing and reliable in terms of statistical criteria. In addition, it conveys the message that the entire data are located within the acceptable domains, again confirming the LS-SVM model offers accurate and satisfactory predictions.

Analysis of variance (ANOVA) was used to determine the relative importance of all the input parameters which are incorporated in this modeling strategy to develop the connectionist tool for estimation of CO<sub>2</sub>–oil swelling factor. The relative significance of the independent variables including API oil gravity, temperature, pressure, and CO<sub>2</sub> concentration (mole fraction) on the swelling factor is demonstrated in Fig. 12. As it is clear from the results, the most significant independent parameter is the API degree of the oil samples, temperature holds the second rank, and the CO<sub>2</sub> concentration exhibits the least impact on the target parameter.

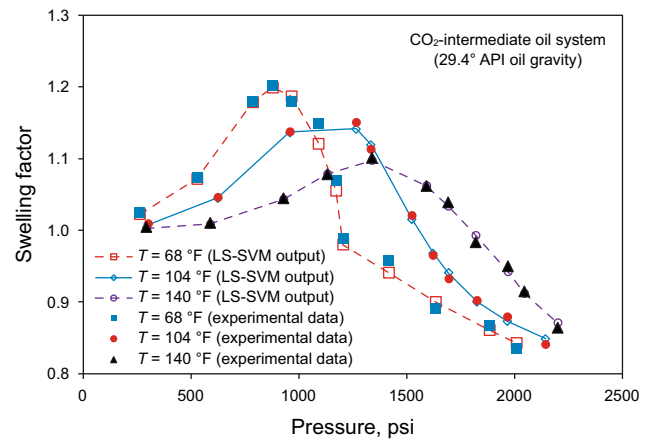
To show the effectiveness of the developed model for a real case, we consider sample AC with the composition reported in Table 5. A swelling test was performed on this sample with different mole fractions of CO<sub>2</sub>. As mentioned





**Fig. 6** Scatter plot of estimated and measured swelling factors: **a** Training data; **b** Testing data; and **c** All data

previously, one of the methods for swelling factor determination is using EOSs. Thus, the Peng–Robinson EOS as



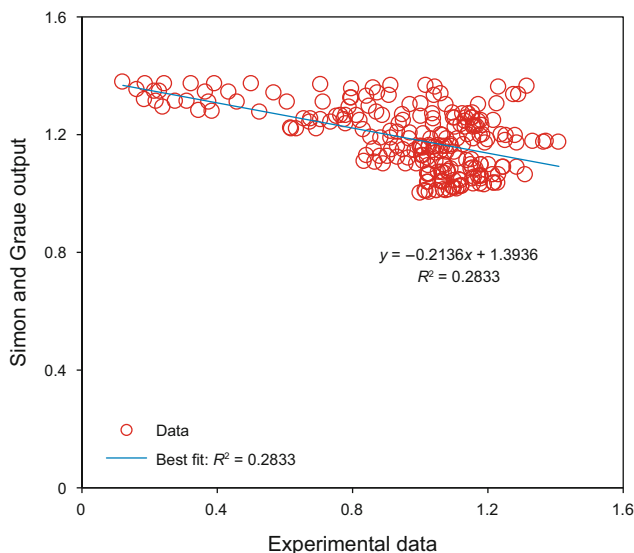
**Fig. 7** Comparison between calculated and measured swelling factors versus pressure at different temperatures for the CO<sub>2</sub>–intermediate oil system (The oil API gravity is 29.4°)



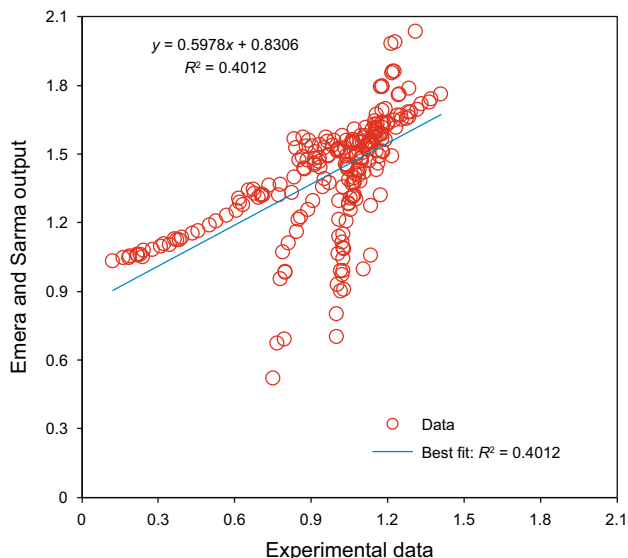
**Fig. 8** Relative error distribution of the estimated target variable (e.g., swelling factor)

a well-known and robust EOS was utilized to calculate the CO<sub>2</sub>–oil swelling factor of sample AC. Figure 13 displays a comparison between the outputs obtained from the LS-SVM model, Peng–Robinson EOS, and experimental data from a swelling test performed on sample AC. As illustrated in Fig. 13, both LS-SVM and Peng–Robinson EOS predict the CO<sub>2</sub>–oil swelling factor with reasonable accuracy. In this case, the LS-SVM underestimates swelling factor; however, using Peng–Robinson EOS results in overestimating the swelling factor.

The residual oil saturation, which directly corresponds to the oil recovery factor is inversely proportional to the swelling factor in CO<sub>2</sub>-based EOR processes. Hence, an accurate magnitude of the CO<sub>2</sub>–oil swelling factor increases the precision and reliability of the modeling and simulation studies, which are conducted to capture the main recovery mechanisms and to determine the production



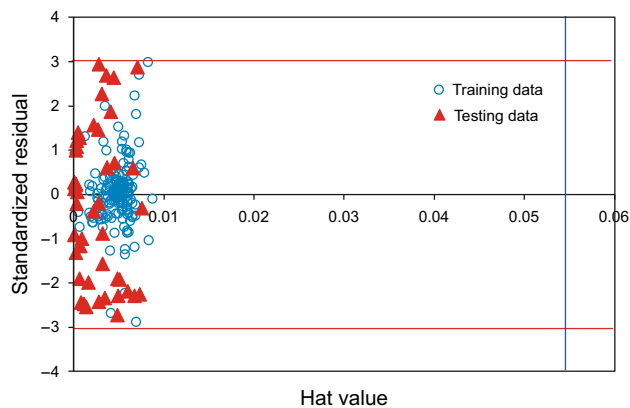
**Fig. 9** The Simon–Graue (1965) outputs versus measured swelling factors



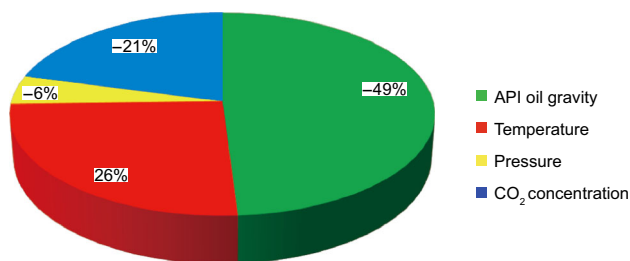
**Fig. 10** Scatter plot of values of the target parameter estimated by Emera and Sarma (2006) correlation and measured swelling factors

**Table 4** Maximum absolute error and average absolute relative deviation to indicate the difference between the predicted values and experimental data

Method	Maximum absolute error	Average absolute relative deviation
LS-SVM model	5.42	1.5433
Emera and Sarma method	91.2154	66.92
Simon and Graue method	125.70	56.87



**Fig. 11** Detection of the possible doubtful measured objective function and the applicability domain of the suggested approach for the CO<sub>2</sub>–oil swelling factor (The hat value (*H*) is 0.0555)



**Fig. 12** Relative importance of the independent variables affecting the swelling factor

performance of CO<sub>2</sub>–EOR strategies for both heavy oil and conventional oil reserves. The present study introduces an accurate and simple-to-use approach to calculate the CO<sub>2</sub>–oil swelling factor, which is an influential parameter throughout CO<sub>2</sub> injection operations. The precise value of this parameter helps engineers and researchers obtain the residual oil saturation and oil and water relative permeability curves with greater reliability for various oil reservoir development stages (e.g., optimization of operational conditions and economic analysis).

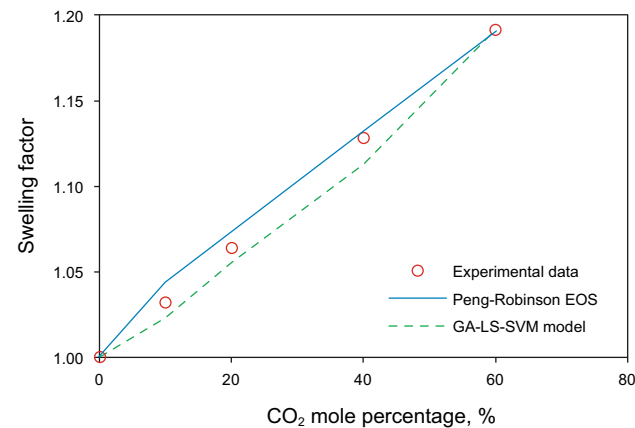
## 6 Conclusions

We used the least-squares support vector machine (LS-SVM) to estimate the CO<sub>2</sub>–oil swelling factor where the extensive experimental data were utilized. The genetic algorithm (GA) was employed to tune the model parameters. The following conclusions based on the research outputs are made:

- The feasibility and performance of the LS-SVM technique with a RBF kernel function were evaluated using the available experimental data on CO<sub>2</sub>–oil swelling factors.

**Table 5** Composition of oil sample AC

Component	Mole percentage, %	Molecular weight
N <sub>2</sub>	0.53	28.014
CO <sub>2</sub>	1.01	44.01
C <sub>1</sub>	45.305	16.043
C <sub>2</sub>	3.9	30.07
C <sub>3</sub>	1.39	44.097
<i>i</i> -C <sub>4</sub>	0.63	58.124
<i>n</i> -C <sub>4</sub>	0.81	58.124
<i>i</i> -C <sub>5</sub>	0.69	72.151
<i>n</i> -C <sub>5</sub>	0.41	72.151
C <sub>6</sub>	1.02	86.178
C <sub>7</sub>	4.22	96
C <sub>8</sub>	3.53	107
C <sub>9</sub>	3.5	121
C <sub>10</sub> –C <sub>12</sub>	8.011	145.496
C <sub>13</sub> –C <sub>15</sub>	6.521	186.218
C <sub>16</sub> –C <sub>18</sub>	4.84	227.455
C <sub>19</sub> –C <sub>23</sub>	3.672	283.28
C <sub>24</sub> –C <sub>30</sub>	3.203	370.644
C <sub>31</sub> –C <sub>37</sub>	2.232	469.117
C <sub>38</sub> –C <sub>46</sub>	1.906	579.203
C <sub>47</sub> –C <sub>58</sub>	1.489	722.448
C <sub>59</sub> –C <sub>80</sub>	1.179	940.536



**Fig. 13** Comparison between the LS-SVM method and Peng–Robinson EOS in terms of performance precision for swelling factor determination

- GA was implemented to determine the optimal extent of the model parameters; namely, regularization factor and variance used in the kernel function which were obtained to be:  $\gamma = 33.4091$  and  $\sigma^2 = 0.268829$ , respectively.
- The hybridized GA-LS-SVM provided excellent results in predicting the CO<sub>2</sub>–oil swelling factor. The

performance of the hybrid model was evaluated by  $R^2 = 0.9953$  and  $MSE = 0.0003$ , which reveal high accuracy and reliability of the developed model.

- The relative importance of input variables including API gravity of oil, temperature, pressure, and CO<sub>2</sub> concentration (mole fraction) on the CO<sub>2</sub>–oil swelling factor was investigated using a common statistical approach, ANOVA (analysis of variance). The API gravity of oil, temperature, pressure, and CO<sub>2</sub> content have the highest to the lowest impact on the objective function in the research study.
- The LS-SVM features high efficiency, excellent generalization, and routine computation methodology, which is suitable for classification and identification of nonlinear cases such as CO<sub>2</sub>–oil systems.

**Acknowledgements** The financial support from Natural Sciences and Engineering Research Council of Canada (NSERC), InnovateNL, and Statoil Canada is greatly appreciated.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

### Appendix: sample calculation

LS-SVM toolbox in MATLAB software is first installed. The next step is dragging and dropping the swelling.mat file in the MATLAB workspace. The following command should be then written in the MATLAB environment to calculate the CO<sub>2</sub>–oil swelling factor.

SF = simlssvm ({trainX', trainY', type, gam, sig2, 'RBF\_kernel', 'preprocess'}, {alpha, b}, [API T P Solubility (mole fraction)])

For example, consider the following data:

API = 29.4°

T = 176 °F

P = 2257 psia

CO<sub>2</sub> mole fraction = 0.66

SF = simlssvm ({trainX', trainY', type, gam, sig2, 'RBF\_kernel', 'preprocess'}, {alpha, b}, [29.4 176 2257 0.66])

The swelling factor forecasted by the LS-SVM model is equal to 0.9327, while the corresponding measured CO<sub>2</sub>–oil swelling factor is 0.9334, implying there is a very good agreement between the predicted and experimental values.

## References

- Abedini A, Torabi F. On the CO<sub>2</sub> storage potential of cyclic CO<sub>2</sub> injection process for enhanced oil recovery. *Fuel*. 2014;124:14–27. <https://doi.org/10.1016/j.fuel.2014.01.084>.
- Abedini A, Mosavat N, Torabi F. Determination of minimum miscibility pressure of crude oil–CO<sub>2</sub> system by oil swelling/extraction test. *Energy Technol*. 2014;2(5):431–9. <https://doi.org/10.1002/ente.201400005>.
- Ahmadi MA. Connectionist approach estimates gas–oil relative permeability in petroleum reservoirs: application to reservoir simulation. *Fuel*. 2015;40:429–39. <https://doi.org/10.1016/j.fuel.2014.09.058>.
- Ahmadi MA, Ebadi M. Evolving smart approach for determination dew point pressure through condensate gas reservoirs. *Fuel*. 2014;117:1074–84. <https://doi.org/10.1016/j.fuel.2013.10.010>.
- Ahmadi MA, Ebadi M, Hosseini SM. Prediction breakthrough time of water coning in the fractured reservoirs by implementing low parameter support vector machine approach. *Fuel*. 2014a;117:579–89. <https://doi.org/10.1016/j.fuel.2013.09.071>.
- Ahmadi MA, Ebadi M, Marghmaleki PS, Fouladi MM. Evolving predictive model to determine condensate-to-gas ratio in retro-graded condensate gas reservoirs. *Fuel*. 2014b;124:241–57. <https://doi.org/10.1016/j.fuel.2014.01.073>.
- Ahmadi MA, zeinali Hasanvand M, Shokrolahzadeh S. Technical and economic feasibility study of flue gas injection in an Iranian oil field. *Petroleum*. 2015;1(3):217–22. <https://doi.org/10.1016/j.petlm.2015.07.010>.
- Ahmadi MA, Pouladi B, Barghi T. Numerical modeling of CO<sub>2</sub> injection scenarios in petroleum reservoirs: application to CO<sub>2</sub> sequestration and EOR. *J Nat Gas Sci Eng*. 2016a;30:38–49. <https://doi.org/10.1016/j.jngse.2016.01.038>.
- Ahmadi MA, zeinali Hasanvand M, Behbahani SS, Nourmohammad A, Vahidi A, Amiri M, et al. Effect of operational parameters on the performance of carbonated water injection: experimental and numerical modeling study. *J Supercrit Fluids*. 2016b;107:542–8. <https://doi.org/10.1016/j.supflu.2015.07.012>.
- Bachu S. Identification of oil reservoirs suitable for CO<sub>2</sub>-EOR and CO<sub>2</sub> storage (CCUS) using reserves databases, with application to Alberta, Canada. *Int J Greenh Gas Control*. 2016;44:152–65. <https://doi.org/10.1016/j.ijggc.2015.11.013>.
- Bagci A. Immiscible CO<sub>2</sub> flooding through horizontal wells. *Energy Sources Part A*. 2007;29(1):85–95. <https://doi.org/10.1080/009083190934022>.
- Baylar A, Hanbay D, Batan M. Application of least square support vector machines in the prediction of aeration performance of plunging overfall jets from weirs. *Expert Syst Appl*. 2009;36(4):8368–74. <https://doi.org/10.1016/j.eswa.2008.10.061>.
- Bessières D, Saint-Guirons H, Daridon JL. Volumetric behavior of decane + carbon dioxide at high pressures. Measurement and calculation. *J Chem Eng Data*. 2001;46(5):1136–9. <https://doi.org/10.1021/je010016k>.
- Burges CJ. A tutorial on support vector machines for pattern recognition. *Data Min Knowl Discov*. 1998;2(2):121–67. <https://doi.org/10.1023/A:100971592>.
- Chung FT, Jones RA, Nguyen HT. Measurements and correlations of the physical properties of CO<sub>2</sub>-heavy crude oil mixtures. *SPE Reserv Eng*. 1988;3(03):822–8. <https://doi.org/10.2118/15080-PA>.
- Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20(3):273–97. <https://doi.org/10.1007/BF00994018>.
- Czarnota R, Janiga D, Stopa J, Wojnarowski P. Determination of minimum miscibility pressure for CO<sub>2</sub> and oil system using acoustically monitored separator. *J CO<sub>2</sub> Util*. 2017;17:32–6. <https://doi.org/10.1016/j.jcou.2016.11.004>.
- Davis SJ, Caldeira K, Matthews HD. Future CO<sub>2</sub> emissions and climate change from existing energy infrastructure. *Science*. 2010;329(5997):1330–3. <https://doi.org/10.1126/science.1188566>.
- Diep P, Jordan KD, Johnson JK, Beckman EJ. CO<sub>2</sub>–fluorocarbon and CO<sub>2</sub>–hydrocarbon interactions from first-principles calculations. *J Phys Chem A*. 1998;102(12):2231–6. <https://doi.org/10.1021/jp9730306>.
- Do H, Pinczewski W. Diffusion controlled swelling of reservoir oil by direct contact with injection gas. *Chem Eng Sci*. 1991;46(5–6):1259–70. [https://doi.org/10.1016/0009-2509\(91\)85053-Z](https://doi.org/10.1016/0009-2509(91)85053-Z).
- Dong M, Huang S, Srivastava R. A laboratory study on near-miscible CO injection in Steelman reservoir. *J Can Pet Technol*. 2001. <https://doi.org/10.2118/01-02-05>.
- Emera M, Sarma H. A genetic algorithm-based model to predict co-oil physical properties for dead and live oil. In: Canadian international petroleum conference, 13–15 June, Calgary, Alberta; 2006. <https://doi.org/10.2118/2006-197>.
- Farajzadeh R, Andrianov A, Bruining H, Zitha PL. Comparative study of CO<sub>2</sub> and N<sub>2</sub> foams in porous media at low and high pressure–temperatures. *Ind Eng Chem Res*. 2009;48(9):4542–52. <https://doi.org/10.1021/ie801760u>.
- Fazeli H, Soleimani R, Ahmadi MA, Badrnezhad R, Mohammadi AH. Experimental study and modeling of ultrafiltration of refinery effluents using a hybrid intelligent approach. *Energy Fuels*. 2013;27(6):3523–37. <https://doi.org/10.1021/ef400179b>.
- Fukai I, Mishra S, Moody MA. Economic analysis of CO<sub>2</sub>-enhanced oil recovery in Ohio: implications for carbon capture, utilization, and storage in the Appalachian Basin region. *Int J Greenh Gas Control*. 2016;52:357–77. <https://doi.org/10.1016/j.ijggc.2016.07.015>.
- Ghedan S. Global laboratory experience of CO<sub>2</sub>-EOR flooding. In: SPE/EAGE reservoir characterization & simulation conference, 19–21 October, Abu Dhabi, UAE; 2009. <https://doi.org/10.2118/125581-MS>.
- Godec ML, Kuuskraa VA, Dipietro P. Opportunities for using anthropogenic CO<sub>2</sub> for enhanced oil recovery and CO<sub>2</sub> storage. *Energy Fuels*. 2013;27(8):4183–9. <https://doi.org/10.1021/ef302040u>.
- Goodall CR. 13 computation using the QR decomposition. *Handb Stat*. 1993;9:467–508. [https://doi.org/10.1016/S0169-7161\(05\)80137-3](https://doi.org/10.1016/S0169-7161(05)80137-3).
- Gramatica P. Principles of QSAR models validation: internal and external. *Mol Inform*. 2007;26(5):694–701. <https://doi.org/10.1002/qsar.200610151>.
- Habibi A, Yassin MR, Dehghanpour H, Bryan D. CO<sub>2</sub>-oil interactions in tight rocks: an experimental study. In: SPE unconventional resources conference, 15–16 February, Calgary, Alberta, Canada; 2017. <https://doi.org/10.2118/185047-MS>.
- Hand JL, Pinczewski WV. Interpretation of swelling/extraction tests. *SPE Reserv Eng*. 1990;5(04):595–600. <https://doi.org/10.2118/19471-PA>.
- Harmon RA, Grigg RB. Vapor-density measurement for estimating minimum miscibility pressure (includes associated papers 19118 and 19500). *SPE Reserv Eng*. 1988;3(04):1215–20. <https://doi.org/10.2118/15403-PA>.
- Heidaryan E, Moghadasi J. A laboratory investigation into carbon dioxide flooding by focusing on the viscosity and swelling factor changes. *Pet Sci Technol*. 2012;30(14):1441–52. <https://doi.org/10.1080/10916466.2010.482037>.
- Holm LW, Josendal VA. Effect of oil composition on miscible-type displacement by carbon dioxide. *SPE J*. 1982;22(01):87–98. <https://doi.org/10.2118/8814-PA>.

- Issever K, Topkaya I. Use of carbon dioxide to enhanced heavy oil recovery. In: 7th Unitar international conference on heavy crude and tar sands, Beijing, China, 1998. p. 27–30.
- Jamali A, Etehadavakki A. CO<sub>2</sub> storage in residual oil zones: field-scale modeling and assessment. *Int J Greenh Gas Control*. 2017;56:102–15. <https://doi.org/10.1016/j.ijggc.2016.10.005>.
- Keerthi SS, Lin CJ. Asymptotic behaviors of support vector machines with Gaussian kernel. *Neural Comput*. 2003;15(7):1667–89. <https://doi.org/10.1162/089976603321891855>.
- Kim S, Santamarina JC. Engineered CO<sub>2</sub> injection: the use of surfactants for enhanced sweep efficiency. *Int J Greenh Gas Control*. 2014;20:324–32. <https://doi.org/10.1016/j.ijggc.2013.11.018>.
- Kiran E, Pöhler H, Xiong Y. Volumetric properties of pentane + carbon dioxide at high pressures. *J Chem Eng Data*. 1996;41(2):158–65. <https://doi.org/10.1021/je9501503>.
- Kuznetsova T, Kvamme B. Thermodynamic properties and interfacial tension of a model water–carbon dioxide system. *Phys Chem Chem Phys*. 2002;4(6):937–41. <https://doi.org/10.1039/B108726F>.
- Kwak DH, Kim JK. Techno-economic evaluation of CO<sub>2</sub> enhanced oil recovery (EOR) with the optimization of CO<sub>2</sub> supply. *Int J Greenh Gas Control*. 2017;58:169–84. <https://doi.org/10.1016/j.ijggc.2017.01.002>.
- Li CH, Zhu XJ, Cao GY, Sui S, Hu MR. Identification of the Hammerstein model of a PEMFC stack based on least squares support vector machines. *J Power Sources*. 2008;175(1):303–16. <https://doi.org/10.1016/j.jpowsour.2007.09.049>.
- Li J, Yao X, Liu M, Hu Z, Fan B. Quantitative structure–activity relationship study of acyl ureas as inhibitors of human liver glycogen phosphorylase using least squares support vector machines. *Chemometr Intell Lab Syst*. 2007;87(2):139–46. <https://doi.org/10.1016/j.chemolab.2006.11.004>.
- Li L, Khorsandi S, Johns RT, Dilmore RM. CO<sub>2</sub> enhanced oil recovery and storage using a gravity-enhanced process. *Int J Greenh Gas Control*. 2015;42:502–15. <https://doi.org/10.1016/j.ijggc.2015.09.006>.
- Li X, Fan X. Effect of CO<sub>2</sub> phase on contact angle in oil-wet and water-wet pores. *Int J Greenh Gas Control*. 2015;36:106–13. <https://doi.org/10.1016/j.ijggc.2015.02.017>.
- Li X, Li H, Yang D. Determination of multiphase boundaries and swelling factors of solvent(s)–CO<sub>2</sub>–heavy oil systems at high pressures and elevated temperatures. *Energy Fuels*. 2013a;27(3):1293–306. <https://doi.org/10.1021/ef301866e>.
- Li X, Ross DA, Trusler JM, Maitland GC, Boek ES. Molecular dynamics simulations of CO<sub>2</sub> and brine interfacial tension at high temperatures and pressures. *J Phys Chem B*. 2013b;117(18):5647–52. <https://doi.org/10.1021/jp309730m>.
- Liu H, Yao X, Zhang R, Liu M, Hu Z, Fan B. Accurate quantitative structure – property relationship model to predict the solubility of C<sub>60</sub> in various solvents based on a novel approach using a least-squares support vector machine. *J Phys Chem B*. 2005a;109(43):20565–71.
- Liu H, Yao X, Zhang R, Liu M, Hu Z, Fan B. Prediction of the tissue/blood partition coefficients of organic compounds based on the molecular structure using least-squares support vector machines. *J Comput Aided Mol Des*. 2005b;19(7):499–508. <https://doi.org/10.1007/s10822-005-9003-5>.
- Liu Y, Wilcox J. Molecular simulation studies of CO<sub>2</sub> adsorption by carbon model compounds for carbon capture and sequestration applications. *Environ Sci Technol*. 2012;47(1):95–101. <https://doi.org/10.1021/es3012029>.
- Luo P, Yang C, Gu Y. Enhanced solvent dissolution into in situ upgraded heavy oil under different pressures. *Fluid Phase Equilib*. 2007;252(1):143–51. <https://doi.org/10.1016/j.ijggc.2013.03.001>.
- Luo P, Er V, Freitag N, Huang S. Recharacterizing evolving fluid and PVT properties of Weyburn oil–CO<sub>2</sub> system. *Int J Greenh Gas Control*. 2013;16:S226–35.
- Lv G, Li Q, Wang S, Li X. Key techniques of reservoir engineering and injection–production process for CO<sub>2</sub> flooding in China’s SINOPEC Shengli Oilfield. *J CO<sub>2</sub> Util*. 2015;11:31–40. <https://doi.org/10.1016/j.jcou.2014.12.007>.
- Ma J, Wang X, Gao R, Zeng F, Huang C, Tontiwachwuthikul P, et al. Study of cyclic CO<sub>2</sub> injection for low-pressure light oil recovery under reservoir conditions. *Fuel*. 2016;174:296–306. <https://doi.org/10.1016/j.fuel.2016.02.017>.
- Mehdizadeh B, Movagharnjad K. A comparative study between LS-SVM method and semi empirical equations for modeling the solubility of different solutes in supercritical carbon dioxide. *Chem Eng Res Des*. 2011;89(11):2420–7. <https://doi.org/10.1016/j.cherd.2011.03.017>.
- Mosavat N, Abedini A, Torabi F. Phase behaviour of CO<sub>2</sub>–brine and CO<sub>2</sub>–oil systems for CO<sub>2</sub> storage and enhanced oil recovery: experimental studies. *Energy Proc*. 2014;63:5631–45. <https://doi.org/10.1016/j.egypro.2014.11.596>.
- Mulliken CA, Sandler SI. The prediction of CO<sub>2</sub> solubility and swelling factors for enhanced oil recovery. *Ind Eng Chem Process Des Dev*. 1980;19(4):709–11. <https://doi.org/10.1021/i260076a033>.
- Niazi A, Jameh-Bozorgi S, Nori-Shargh D. Prediction of toxicity of nitrobenzenes using ab initio and least squares support vector machines. *J Hazard Mater*. 2008;151(2):603–9. <https://doi.org/10.1016/j.jhazmat.2007.06.030>.
- Ning SX, Jhaveri BS, Jia N, Chambers B, Gao J. Viscosity reduction EOR with CO<sub>2</sub> & enriched CO<sub>2</sub> to improve recovery of Alaska North Slope viscous oils. In: SPE western north American region meeting, 7–11 May, Anchorage, Alaska, USA; 2011. <https://doi.org/10.2118/144358-MS>.
- Or C, Sasaki K, Sugai Y, Nakano M, Imai M. Swelling and viscosity reduction of heavy oil by CO<sub>2</sub>-gas foaming in immiscible condition. *SPE Reserv Eval Eng*. 2016;19(02):294–304. <https://doi.org/10.2118/179738-PA>.
- Orr FM Jr, Yu A, Lien C. Phase behavior of CO<sub>2</sub> and crude oil in low-temperature reservoirs. *SPE J*. 1981;21(04):480–92. <https://doi.org/10.2118/8813-PA>.
- Orr FM Jr, Heller JP, Taber JJ. Carbon dioxide flooding for enhanced oil recovery: promise and problems. *J Am Oil Chem Soc*. 1982;59(10):810A–7A. <https://doi.org/10.1007/BF02634446>.
- Perera MSA, Gamage RP, Rathnaweera TD, Ranathunga ASD, Koay A, Choi X. A review of CO<sub>2</sub>-enhanced oil recovery with a simulated sensitivity analysis. *Energies*. 2016;9(7):481. <https://doi.org/10.3390/en9070481>.
- Reihanian M, Asadollahpour S, Hajarpour S, Gheisari K. Application of neural network and genetic algorithm to powder metallurgy of pure iron. *Mater Des*. 2011;32(6):3183–8. <https://doi.org/10.1016/j.matdes.2011.02.049>.
- Rojas G, Ali S. Scaled model studies of carbon dioxide/brine injection strategies for heavy oil recovery from thin formations. *J Can Pet Technol*. 1986. <https://doi.org/10.2118/86-01-07>.
- Rousseeuw PJ, Leroy AM. Robust regression and outlier detection. Hoboken: Wiley; 2005. p. 589.
- Sell K, Enzmann F, Kersten M, Spangenberg E. Microtomographic quantification of hydraulic clay mineral displacement effects during a CO<sub>2</sub> sequestration experiment with saline aquifer sandstone. *Environ Sci Technol*. 2012;47(1):198–204. <https://doi.org/10.1021/es3013358>.
- Shelton JL, McIntosh JC, Hunt AG, Beebe TL, Parker AD, Warwick PD, et al. Determining CO<sub>2</sub> storage potential during miscible CO<sub>2</sub> enhanced oil recovery: noble gas and stable isotope tracers. *Int J Greenh Gas Control*. 2016;51:239–53. <https://doi.org/10.1016/j.ijggc.2016.05.008>.



- Simon R, Graue D. Generalized correlations for predicting solubility, swelling and viscosity behavior of CO<sub>2</sub>-crude oil systems. *J Pet Technol.* 1965;17(01):102–6. <https://doi.org/10.2118/917-PA>.
- Sugai Y, Babadagli T, Sasaki K. Consideration of an effect of interfacial area between oil and CO<sub>2</sub> on oil swelling. *J Pet Explor Prod Technol.* 2014;4(1):105–12. <https://doi.org/10.1007/s13202-013-0085-7>.
- Suykens JA, Vandewalle J. Least squares support vector machine classifiers. *Neural Process Lett.* 1999;9(3):293–300. <https://doi.org/10.1023/A:1018628609742>.
- Suykens JA, Van Gestel T, De Brabanter J. Least squares support vector machines. Singapore: World Scientific; 2002.
- Thomas G, Monger-McClure T. Feasibility of cyclic CO<sub>2</sub> injection for light-oil recovery. *SPE Reserv Eng.* 1991;6(02):179–84. <https://doi.org/10.2118/20208-PA>.
- Tsau JS, Bui LH, Willhite GP. Swelling/extraction test of a small sample size for phase behavior study, In: SPE improved oil recovery symposium, 24–28 April, Tulsa, Oklahoma, USA; 2010. <https://doi.org/10.2118/129728-MS>.
- Tunio SQ, Tunio AH, Ghirano NA, El Adawy ZM. Comparison of different enhanced oil recovery techniques for better oil productivity. *Int J Appl Sci Technol.* 2011;1(5):143–153.
- Vapnik V. *Statistical learning theory.* New York: Wiley; 1998.
- Vong CM, Wong PK, Li YP. Prediction of automotive engine power and torque using least squares support vector machines and Bayesian inference. *Eng Appl Artif Intell.* 2006;19(3):277–87. <https://doi.org/10.1016/j.engappai.2005.09.001>.
- Wei B, Gao H, Pu W, Zhao F, Li Y, Jin F, et al. Interactions and phase behaviors between oleic phase and CO<sub>2</sub> from swelling to miscibility in CO<sub>2</sub>-based enhanced oil recovery (EOR) process: a comprehensive visualization study. *J Mol Liq.* 2017;232:277–84. <https://doi.org/10.1016/j.molliq.2017.02.090>.
- Welker J. Physical properties of carbonated oils. *J Pet Technol.* 1963;15(8):873–6. <https://doi.org/10.2118/567-PA>.
- Yang C, Gu Y. A novel experimental technique for studying solvent mass transfer and oil swelling effect in the vapour extraction (VAPEX) process. In: Canadian international petroleum conference, 7–9 June, Calgary, Alberta; 2005. <https://doi.org/10.2118/2005-099>.
- Yang Z, Li M, Peng B, Lin M, Dong Z. Dispersion property of CO<sub>2</sub> in oil. 1. Volume expansion of CO<sub>2</sub> + alkane at near critical and supercritical condition of CO<sub>2</sub>. *J Chem Eng Data.* 2012;57(3):882–9. <https://doi.org/10.1021/je201114g>.
- Yu W, Lashgari HR, Wu K, Sepehrmoori K. CO<sub>2</sub> injection for enhanced oil recovery in Bakken tight oil reservoirs. *Fuel.* 2015;159:354–63. <https://doi.org/10.1016/j.fuel.2015.06.092>.
- zeinali Hasanvand M, Ahmadi MA, Shadizadeh SR, Behbahani R, Feyzi F. Geological storage of carbon dioxide by injection of carbonated water in an Iranian oil reservoir: a case study. *J Pet Sci Eng.* 2013;111:170–7. <https://doi.org/10.1016/j.petrol.2013.09.008>.