Geosciences Journal
Vol. 20, No. 2, p. 259 − 271, April 2016
DOI 10.1007/s12303-015-0033-2

Modeling the permeability of heterogeneous oil reservoirs using a robust method

Arash Kamari* — *Thermodynamics Research Unit, School of Engineering, University of KwaZulu-Natal, Howard College Campus, King George V Avenue, Durban 4041, South Africa*

Farzaneh Moeini, Mohammad-Javad Shamsoddini-Moghadam, Seyed-Ali Hosseini — *Department of Chemical and Petroleum Engineering, Sharif University of Technology, Azadi St., Tehran 11155, Iran*

Amir H. Mohammadi — *Thermodynamics Research Unit, School of Engineering, University of KwaZulu-Natal, Howard College Campus, King George V Avenue, Durban 4041, South Africa*
Institut de Recherche en Génie Chimique et Pétrolier (IRGCP), Paris Cedex, France
Département de Génie des Mines, de la Métallurgie et des Matériaux, Faculté des Sciences et de Génie, Université Laval, Québec (QC) G1V 0A6, Canada

Abdolhossein Hemmati-Sarapardeh — *Young Researchers and Elites Club, Kerman Branch, Islamic Azad University, Kerman, Iran*

ABSTRACT: Permeability as a fundamental reservoir property plays a key role in reserve estimation, numerical reservoir simulation, reservoir engineering calculations, drilling planning, and mapping reservoir quality. In heterogeneous reservoir, due to complexity, natural heterogeneity, non-uniformity, and non-linearity in parameters, prediction of permeability is not straightforward. To ease this problem, a novel mathematical robust model has been proposed to predict the permeability in heterogeneous carbonate reservoirs. To this end, a fairly new soft computing method, namely least square support vector machine (LSSVM) modeling optimized with coupled simulated annealing (CSA) optimization technique was utilized. Statistical and graphical error analyses have been employed separately to evaluate the accuracy and reliability of the proposed model. Furthermore, this model performance has been compared with a newly developed multilayer perceptron artificial neural network (MLP-ANN) model. The obtained results have shown the more robustness, efficiency and reliability of the proposed CSA-LSSVM model in comparison with the developed MLP-ANN model for the prediction of permeability in heterogeneous carbonate reservoirs. Estimations were found to be within acceptable agreement with the actual field data of permeability, with a root mean square error of approximately 0.42 for CSA-LSSVM model in testing phase, and a R-squared value of 0.98. Additionally, these error parameters for MLP-ANN are 0.68 and 0.89 in testing stage, respectively.

Key words: permeability estimation, heterogeneous reservoir, least square support vector machine, coupled simulated annealing, artificial neural network

1. INTRODUCTION

Permeability parameter plays a key role in reservoir characterization, and represents and controls the relationship between pores and production rate. In addition, it is a required reservoir property for reserve estimation, numerical reservoir simulation, injection and production calculations, reservoir engineering calculations, mapping reservoir quality, and drilling planning (Al-Anazi and Gates, 2010a). In heterogeneous reservoirs, problems related to reservoir characterization are difficult due to complexity, natural heterogeneity, non-uniformity and non-linearity in parameters such as permeability.

Normally, well log technology and in-place testing (well testing) as well as experimental measurement are three methods to determine the permeability (Wong et al., 2000; Karimpouli et al., 2010). Generally, the aforementioned methods are time consuming, expensive and tedious, which makes the use of computing methods more attractive. Hence, a more robust and easier-to-use computational approach should be developed and proposed.

In recent years, intelligent/smart methods have been progressively employed petroleum and chemical calculations (Mohaghegh et al., 1994; Ghiasi et al., 2013; Hosseinzadeh and Hemmati-Sarapardeh, 2014; Kamari et al., 2014a; Kamari et al., 2014b; Kamari et al., 2014c; Nejatian et al., 2014; Talebi et al., 2014; Zendehboudi et al., 2014; Esfahani et al., 2015; Fathinasab et al., 2015; Kamari et al., 2015a; Kamari et al., 2015b). The artificial neural networks (ANNs) are able to solve complex nonlinear and classification problems, and they can perform prediction and generalization rapidly once trained (Gharbi, 1997). In the presence of a small size of dataset, ANN technique may lead to an overfitting problem during training/learning phase, which potentially consequences a poor performance for capability, applicability and generalization model (Al-Anazi and Gates, 2012). Although ANN has demonstrated some successful applications to estimation of permeability (Bhatt and Helle, 2002; Saemi et al., 2007; Tahmasebi and Hezarkhani, 2012), the basic training/learning

*Corresponding author: arash.kamari@yahoo.com

mathematical algorithm has been planned to overcome the problems of approximately large sample sizes. Hence, for a given small size of dataset, extensive experiments with several different training/learning methods are required to perform an accurate regression by ANN model (Kaviani et al., 2008).

In recent years, one of the most important smart techniques, namely support vector machine (SVM), have rapidly gained much popularity due to their excellent performance and have become well known in solving complex classification and regression problems (Al-Anazi and Gates, 2012). The SVM technique has found many applications in various fields of science and engineering including but not limited to porosity and permeability estimation from well log data, lithology identification, pattern recognition in medical science, speech and text detection, etc. (Li et al., 2000; Choisy and Belaid, 2001; Gao et al., 2001; Ma et al., 2001; Van Gestel et al., 2001)

El-Sebakhy (2009) used support vector machines for predicting the PVT properties of crude oil systems and solved most of the existing neural networks drawbacks. Furthermore, a comparative study was carried out to compare support vector machines regression performance with the one of the neural networks, nonlinear regression, and different empirical correlation techniques. The results indicated that the performance of support vector machines is accurate and outperforms most of the published correlations.

Al-Anzi and Gates (2010a) applied a nonlinear SVM technique in a highly heterogeneous sandstone reservoir to classify electrofacies and predict permeability distributions. The results showed that the SVM method yields comparable or superior classification of the lithology, and estimates the permeability better than the neural network methods. Moreover, comparison between log-based and core-based clustering revealed that permeability prediction based on core-based clustering was slightly better than that of the log based clustering. Al-Anzi and Gates (2012) evaluated the capability of support vector regression (SVR) for prediction of porosity and permeability in a heterogeneous sandstone reservoir under the effect of small sample size. Performance of SVR model was compared to the multilayer perception (MLP) neural network. The results demonstrated that SVR yields consistently better predictions of the porosity and permeability with small sample size than the MLP method.

Chamkalani et al. (2013) developed a new scaling equation for determining asphaltene precipitation quantitatively. They used least square support vector machines/regression (LSSVM/LSSVR) to build a nonlinear model. The results showed that the proposed LSSVM algorithm is highly satisfactory.

In this study, more than 700 data points collected from well log data of a heterogeneous carbonate oil reservoir located in Saudi Arabia have been applied. Then, to develop the proposed model, the dataset has been divided into two subsets of training (80% for construction of the model) and testing (20% for evaluation of the model performance and accuracy). Least square support vector machine (LSSVM) has been utilized in this study to construct nonlinear modeling. Besides, Coupled Simulated Annealing (CSA) optimization technique has been employed for tuning the LSSVM parameters. Here, it is worthwhile to note that the LSSVM technique has not so far been implemented for forecasting the permeability of heterogeneous oil reservoirs. Additionally, adequacy and reliability of the model has been evaluated through statistical and graphical error analyses, and finally the obtained results by the CSA-LSSVM model have been compared to an MLP-ANN model's outputs.

2. DATA GATHERING

As a result, huge amount of oil exists in carbonate reservoirs which normally have complex structures including matrix and fractures (Alizadeh et al., 2013; Hashemi-Kiasari et al., 2014; Kamari et al., 2015c). In other words, there are rock complexities at various scales, i.e., faults, fissures, vugs, microfractures poorly interconnected matrix pore structure, which make them attractive to more study. Furthermore, studying such reservoirs in terms of reservoir properties like permeability is required because the production rate of sandstone reservoir is decreasing. Therefore, a large databank (more than 700 data points) was provided from various wells of a real naturally fractured reservoir. The current case study is located in one of offshore reservoirs in Saudi Arabia. According to the obtained geological setting and the available data, this reservoir can be classified as a structural and heterogonous reservoir. The high variation in the data indicates that due to heterogeneity in this reservoir the traditional methods are not able to estimate the petrophysical parameters effectively. In this oil field, a lot of wells have been drilled in which several petrophysical parameters have been acquired. But, all of them are not suitable for this study. Therefore, among all of the available data, we selected the following parameters: total porosity (PHIT), gamma ray (GR), sonic compression transit time (DT), thermal neutron apparent porosity (NPHI), bulk density (RHOB), and deep Induction log (ILD). The descriptive statistics of these parameters are given in Table 1.

Furthermore, in order to indicate the impacts of the well log databank on the permeability clearer, permeability data points as a function of various input parameters are sketched in Figure 1. Moreover, calculated correlation coefficients for each well log variable against core permeability are summarized in Table 1. This table well reveals that PHIT has the highest correlation coefficient and GR has the lowest one. The collected database includes a wide range of permeability from 0.0001 to 9.929 D. This further indicates the high degree of heterogeneity of reservoir formations. A total of 702 core measurements for permeability and their corresponding well logging responses were available for model training and testing. A part of actual field data of permeability used in this study is presented in Appendix A.

Table 1. Descriptive statistics of well log data set

	PHIT	GR	DT	NPHI	RHOB	ILD
Min.	5.7	7.643	64.4	0.213	1.873	0.1682
Mean	26.7388	27.89966	93.2405	0.297066	2.233562	73.18516
Max.	36.3	83.0234	109.3	0.484	2.6822	2000
Var.	43.18466	240.1747	21.91796	0.001742	0.011873	86732.95
Corr.	0.698522	–0.62445	–0.10943	–0.34806	–0.53328	0.368705

var., variance; corr., correlation coefficient; min., minimum value; max., maximum value; mean, the average for each attribute.

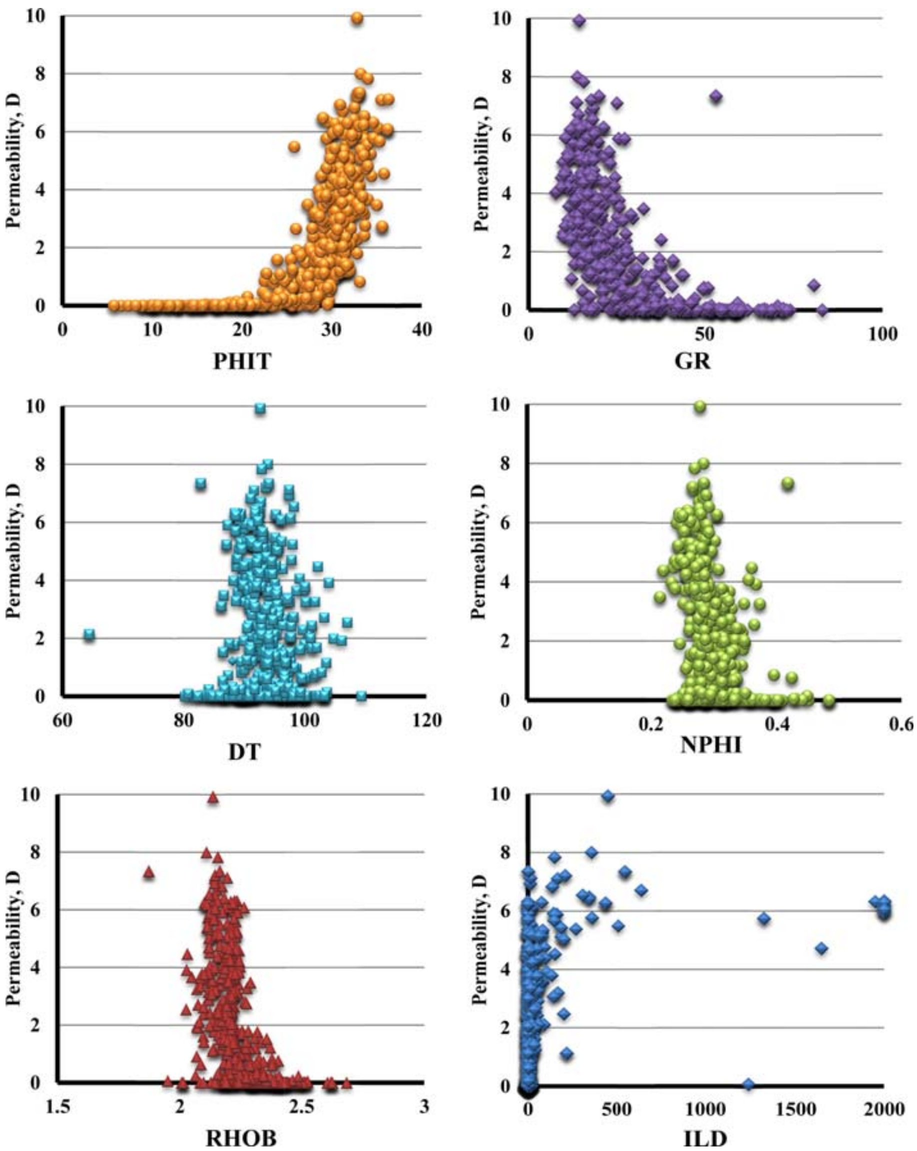

Fig. 1. Permeability as a function of input parameters.

3. MODEL DEVELOPMENT

3.1. Artificial Neural Network

A simple artificial neural network is an artificial intelligence (IA) technology which has various applications in robotic, electronics, financial, medical, and economic as well as oil and gas industry (Tahmasebi and Hezarkhani, 2012). ANNs simulate the work of the human brain and nervous system. Actually, originated by the biological nervous system, they effort to imitate the learning activities of human's brain (Mohaghegh et al., 1994; Balan et al., 1995; Mohaghegh et al., 1995). ANN is an especially efficient mathematical scheme to approximate any function with finite number of discontinuities by learning the relationships between input and output vectors (Ganguly, 2003; Laugier and Richon, 2003).

ANNs have the capability to identify complex problems quickly with a high degree of accuracy, and they are not prejudiced in their calculations and analysis (Ramgulam, 2006). Additionally, ANNs are efficient tools at predicting non-linear properties. Neural networks form a wide category of computer algorithms that solve several kinds of problems such as pattern classification, functions approximation, pattern completion, pattern association, filtering, optimization, and automatic control (Mohaghegh, 2000).

A multilayer ANN consists of different layers. The first and the last layers are the input and output layers, respectively. The intermediate layer, which is exactly between these two layers, connects them indirectly. This layer is called the hidden layer. It should be mentioned that the hidden layer can be more than one layer. How to connect neurons in a neural network makes the type of network. There are different types of neural network. The two general types are feed-forward and backward. Multi-layer perception is the most popular feed-forward network (Saeedi et al., 2007; Nowroozi et al., 2009).

The training phase consists of estimating weights that minimize deviations between network outputs and real data (Al-Anazi and Gates, 2010b). Training data are used for the determination of optimum values of weights and biases of the model, while testing data are used for checking the model performance. The MLP uses a back-propagation (BP) method for training the network which in essence is a type of supervised learning methods (Tahmasebi and Hezarkhani, 2012).

3.2. Least Square Support Vector Machine

Support vector machine is one of the most effective and consistent strategies developed from machine learning principles (Suykens and Vandewalle, 1999; Eslamimanesh et al., 2012). An SVM works as a promising tool for a category of relevant supervised learning methods which can be used not only for analyzing data and recognizing patterns, but also for regression analysis. Based on SVM fundamentals any function f(x) can be regressed and rewritten as below (Suykens et al., 2002b):

$$f(x) = w^T\varphi(x) + b, \tag{1}$$

where w^T, $\varphi(x)$ and b are the transposed output layer vector, the feature map and the bias, respectively. Moreover, x is a vector of dimension n. The following cost function proposed by Vapnik should be minimized to find w and b (Suykens et al., 2002b):

$$\text{Cost function} = \frac{1}{2}w^T + c\sum_{k=1}^{N}(\xi_k - \xi_k^*), \tag{2}$$

subjected to the following constraints:

$$\{y_k - w^T\varphi(x_k) - b \le \varepsilon + \xi_k, \quad k = 1, 2, \ldots, N,$$

$$\{w^T\varphi(x_k) + b + -y_k \le \varepsilon + \xi_k^*, \quad k = 1, 2, \ldots, N,$$

$$\{\xi_k, \xi_k^* \ge 0, \qquad k = 1, 2, \ldots, N, \tag{3}$$

where x_k and y_k are k^{th} input and output data points, respectively. The ε stands for the fixed precision of the function approximation, and the ξ_k and ξ_k are slack variables. It should be mentioned that selecting a small ε in order to increase the accuracy of the model might cause some of the data points to lie beyond the ε precision, and this issue may result in infeasible solution. Therefore, slack parameters should be used to determine the allowed margin of error. The amount of the deviation from the desired ε is determined by the tuning parameters of the SVM (the $c > 0$ in Eq. 2). To minimize the cost function defined in Equation (2) along with its constraints presented in Equation (3), the Lagrangian for this problem should be used as follows (Suykens et al., 2002b):

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

$$-\varepsilon\sum_{k=1}^{N}(a_k - a_k^*) + \sum_{k=1}^{N}y_k(a_k - a_k^*), \tag{4}$$

$$\sum_{k=1}^{N}(a_k - a_k^*) = 0, a_k, a_k^* \in [0, c], \tag{4a}$$

$$K(x_k, x_1) = \varphi(x_k)^T\varphi(x_l), k = 1, 2, \ldots, N, \tag{4b}$$

where a_k and a_k^* are Lagrangian multipliers. Lastly, the final form of the SVM is obtained as follows:

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

One should solve a quadratic programming problem in order to solve the above problem and find a_k, a_k^*, and b, which is very difficult and computationally expensive. Afterward, Suykens and Vandewalle (Suykens and Vandewalle, 1999; Pelckmans et al., 2002) modified the SVM to LSSVM and reformulated it as below (Suykens et al., 2002b):

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

subjected to the following constants (for $k = 1,\ldots,N$):

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

where γ is tuning parameter in LSSVM model, and e_k is the error variable. The Lagrangian for this problem is as below:

$$L(w, b, e, a) = \frac{1}{2}w^Tw + \frac{1}{2}\gamma\sum_{k=1}^{N}e_k^2$$

$$-\sum_{k=1}^{N}a_k(w^T\varphi(x_k) + b + e_k - y_k), \tag{8}$$

where a_k stands for Lagrangian multipliers. The derivatives of Equation (8), should be equated to zero to obtain the solution. Thus, the following equations have to be solved:

$$\{\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{n} a_k \varphi(x_k),$$
$$\{\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{n} a_k = 0,$$
$$\{\frac{\partial L}{\partial e_k} = 0 \rightarrow a_k = \gamma e_k,\ \mathrm{k} = 1,2,...,\ N. \tag{9}$$

As can be seen from Equation (9), there are 2N + 2, equations and 2N + 2 unknowns (a_k, e_k, w, and b). The solution to the system of equations defined in Equation (9) provides the parameters of LSSVM.

The LSSVM has a tuning parameter γ, as stated earlier. Both SVM and LSSVM are kernel-based methods, hence the parameters of the kernel function are considered as other tuning parameters. In this study the widely used RBF kernel function was used (Fayazi et al., 2013; Hemmati-Sarapardeh et al., 2013; Kamari et al., 2013a, 2013b, 2015), which is as follows:

$$K(x, x_k) = \exp\left(-\frac{\|x_k - x\|^2}{\sigma^2}\right). \tag{10}$$

The other tuning parameter is σ^2. Hence, two tuning parameters exist in LSSVM algorithm with RBF kernel function, which should be obtained by minimization of the deviation of the LSSVM model from experimental values (Suykens et al., 2002b). The mean square error (MSE) of the obtained results of the LSSVM algorithm is measured as follows:

$$\mathrm{MSE} = \frac{\sum_{i=1}^{n}(K_{red./pred_i} - K_{exp_i})^2}{n}, \tag{11}$$

where K is the permeability, *rep./pred.* and *exp.* denote the represented/predicted, and experimental permeability, respectively, and n shows the number of data points from the initial population. In this study, the LSSVM algorithm developed by Suykens and Vandewalle (1999) was used.

As pointed out in the previous section, LSSVM modifies the inequality constraints of Equation (3) to the equality constraints of Equation (7) and it is a superiority compared to the traditional SVM. The parameters of the LSSVM can easily be obtained by solving the system of equations presented in Equation (9) instead of solving a nonlinear quadratic programming (Suykens et al., 2002b).

4. RESULTS AND DISSCUSSION

The available permeability dataset was randomly divided into two sub-datasets of the "Training" and "Testing". Normally, the training set is used for constructing the model and the testing set is utilized to investigate the model performance and accuracy. About 80% of the main dataset was randomly chosen for the training purpose and generating the model and the remained 20% was used for testing. It should be noted that for the MLP-ANN model 80% of the main dataset was used as the training set, 5% as the "Validation" set to optimize the model parameters, and the remained 15% was considered as the testing set. The data points should be accumulated homogeneously in the domain of each sub-dataset to result in an appropriate distribution (Gharagheizi et al., 2011).

In the next step, in order to predict permeability of the heterogeneous reservoir, two distinct models were developed using CSA-LSSVM and MLP-ANN strategies. The same input parameters have been chosen to generate these two models consisting of total porosity (PHIT), gamma ray (GR), sonic compression transit time (DT), thermal neutron apparent porosity (NPHI), bulk density (RHOB), and deep Induction log (ILD).

CSA has been employed to determine the optimum values of the LSSVM parameters, namely γ and σ^2. The optimized values of LSSVM parameters are 1.004 and 18.0604 for σ^2 and γ, respectively. The number of reported digits for LSSVM parameters (σ^2 and γ) are generally acquired by means of sensitivity analysis of the overall error of the optimization approach.

In a parallel study, we used the same dataset to develop a MLP-ANN model in order to visualize, investigate and compare the obtained results of LSSVM and MLP-ANN models. In the MLP-ANN, the structure of the network was designed based on the input layer, single hidden layer, and output layer. The tanh-axon was selected as the transfer function, and Levenberg-Marquardt back propagation was employed in all training steps. The size of the hidden layer is measured by the number of processing neurons which must be optimized. The parameters associated with the ANN-MLP model were optimized through a testing-cross validation phases on the available data set. In each run of generated training and testing set, the values of RMSE and R^2 were observed to select the optimum network.

To pursue our objective, statistical errors analysis, in which root mean square error (RMSE) and R-squared error (R^2) are employed, and graphical error analysis, in which crossplot is sketched, have been utilized. More details about statistical and graphical error analyses can be found elsewhere (Montgomery, 2008; Zendehboudi et al., 2011; Kamari et al., 2013a; Shafiei et al., 2013; Hemmati-Sarapardeh et al., 2014). Table 2 lists a comparison of the R-squared and root mean square errors for the estimation of permeability based on training and testing data sets for the LSSVM and MLP-ANN models. As a result, the table reports a R^2 of approximately 0.98 and 0.89 in testing phase for the LSSVM and MLP models, respectively. Furthermore, the LSSVM model has been developed with a RMSE of 0.30 and tested with the value of 0.42. On the other hand, the reported RMSE errors for the MLP-ANN model proposed in this study are 0.41 and 0.68 for training and testing stages, respectively. From the issues discussed above, it can be concluded that the developed LSSVM model is more accurate than the MLP-ANN model in both training and testing phases. In other words,

Table 2. Performance results of the proposed models

Error	MLP-Training	MLP-Testing	LSSVM-Training	LSSVM-Testing
RMSE	0.414048883	0.686396013	0.305114267	0.429597089
R^2	0.978118454	0.891354195	0.98207575	0.981200104

the results obtained in this study reveal that the LSSVM outperforms MLP-ANN for the estimation of permeabilities of heterogeneous oil reservoirs, although the differences are not large in training phase.

To better show the accuracy performance of the developed models, a graphical error analysis was conducted in this study. To this end, the values calculated by the CSA-LSSVM and MLP-ANN models have graphically been compared by two known plates called scatter diagram and point-to-point comparison curve. Additionally, the statistical error analysis in terms of RMSE and R-squared errors is illustrated graphically. The bar plots in Figure 2 present the calculated RMSE for the developed CSA-LSSVM and MLP-ANN models in both training and testing phases. In Figure 2, the results show that the LSSVM model outperforms the MLP-ANN in both training and testing phases. This means that the CSA-LSSVM is more able than the MLP-ANN model for the estimation of permeabilities of heterogeneous oil reservoir. Figure 3 demonstrates that the calculated R^2 for the MLP-ANN model has a value almost equal to the LSSVM model in training set. On the other hand, Figure 3 illustrates that the calculated R-squared in testing phase for the CSA-LSSVM is less than MLP-ANN model. This means that the permeability values estimated by CSA-LSSVM model are in more agreement with the actual field data of permeability. The obtained results indicated that overfitting problems are less probable in CSA-LSSVM method, and also it has proper generalization performance. It should be noted here that a R^2 magnitude greater than 0.9 generally shows a very satisfactory model performance; while a R^2 dimensions in the range of 0.8–0.9 expresses a good performance, and a value less than 0.8 indicates an unsatisfactory model performance, based on the statistical analysis (Ahmadi and Shadizadeh, 2012).

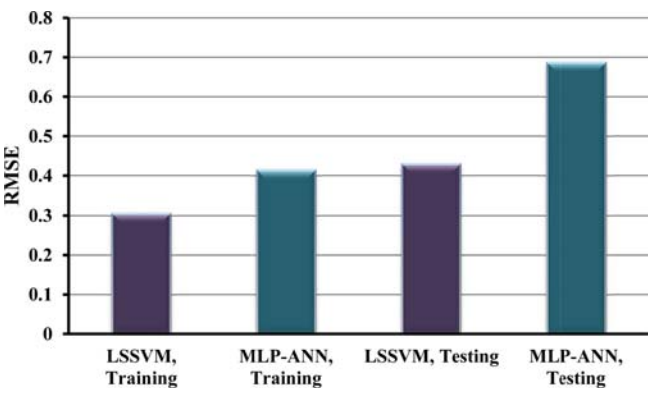

Fig. 2. Calculated RMSE for proposed models at training and testing phases.

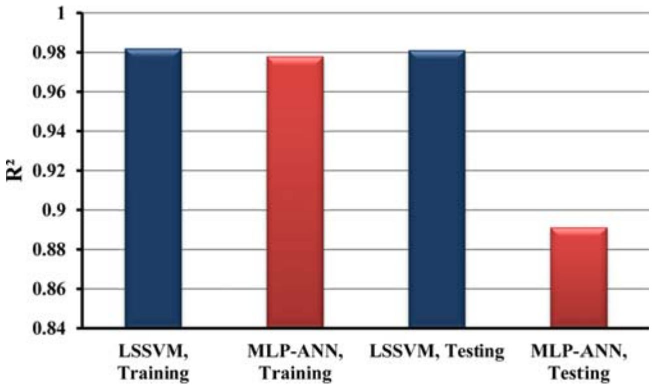

Fig. 3. Calculated R² for proposed models at training and testing phases.

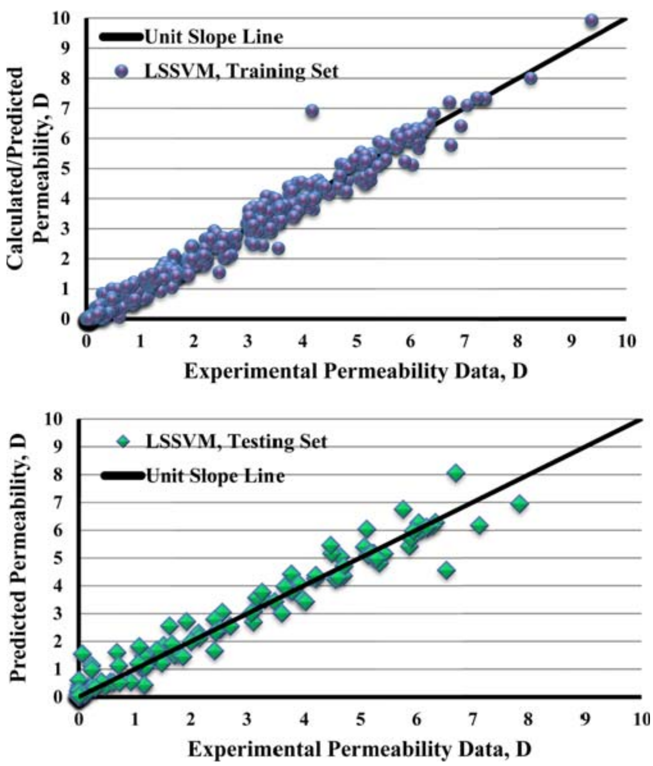

Fig. 4. Crossplot for LSSVM model.

Crossplots or scatter diagrams of training and testing datasets for the permeability values obtained by the CSA-LSSVM and MLP-ANN models against the actual field data are illustrated in Figures 4 and 5, respectively. As it is clear from the figures, the predicted values by the CSA-LSSVM model are matched better on the actual field data of permeability compared to the data estimated by the developed MLP-

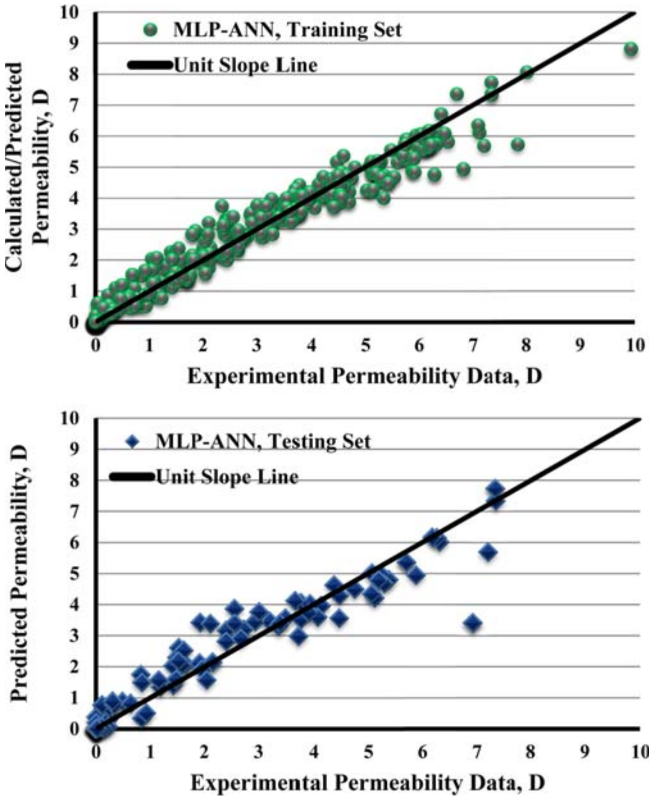

Fig. 5. Crossplot for MLP-ANN model.

ANN model. Additionally, Figures 4 and 5 illustrate that the estimated data points by the CSA-LSSVM model are more distributed around the unit slope line than the developed MLP-ANN model. In other words, the distribution of data points predicted by MLP-ANN model in outside the unit slope line is more than the developed CSA-LSSVM model, which shows the superiority of LSSVM approach in predicting permeability of heterogeneous oil reservoirs. In a further comparison, the predicted and actual field data of permeability at both training and testing sets for CSA-LSSVM and MLP-ANN models is illustrated in Figures 6 and 7, respectively. As can be seen from Figures 6 and 7, the CSA-LSSVM model outputs have a better agreement with the actual field data of permeability than the MLP-ANN model results. This means that the estimated values by the developed CSA-LSSVM model are closer to the actual field data. In other words, this shows the high capability of the CSA-LSSVM approach in predicting permeabilities of heterogeneous oil reservoirs.

To show the importance degree (weights) of well log data on permeability estimated by LSSVM approach (accurate model), a sensitivity analysis has been performed in the current study. To this end, the relevancy factor (r) approach (Chen et al., 2014) is employed for evaluating the effect of well log data during estimation of permeability. The relevancy approach gives a clear image of effective variables for the estimation of permeability. The results of sensitivity analysis performed using relevancy approach are illustrated in Figure 8. As clear from the figure, the weights of well log data (most important input variables) in terms of their calculated relevancy factor for prediction of permeability has been shown graphically. Figure 8 indicates that PHIT, GR, and RHOB have highest influences on the permeabilities estimated by the CSA-LSSVM model, respectively. The results of sensitivity analysis conducted in

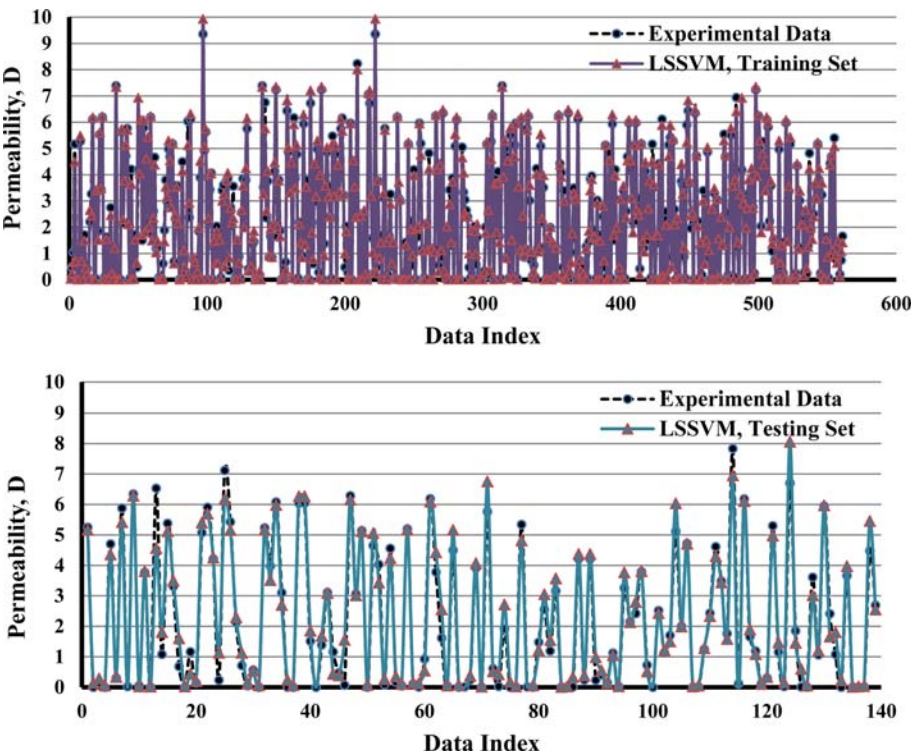

Fig. 6. Comparison between experimental data and estimated permeability by LSSVM model.

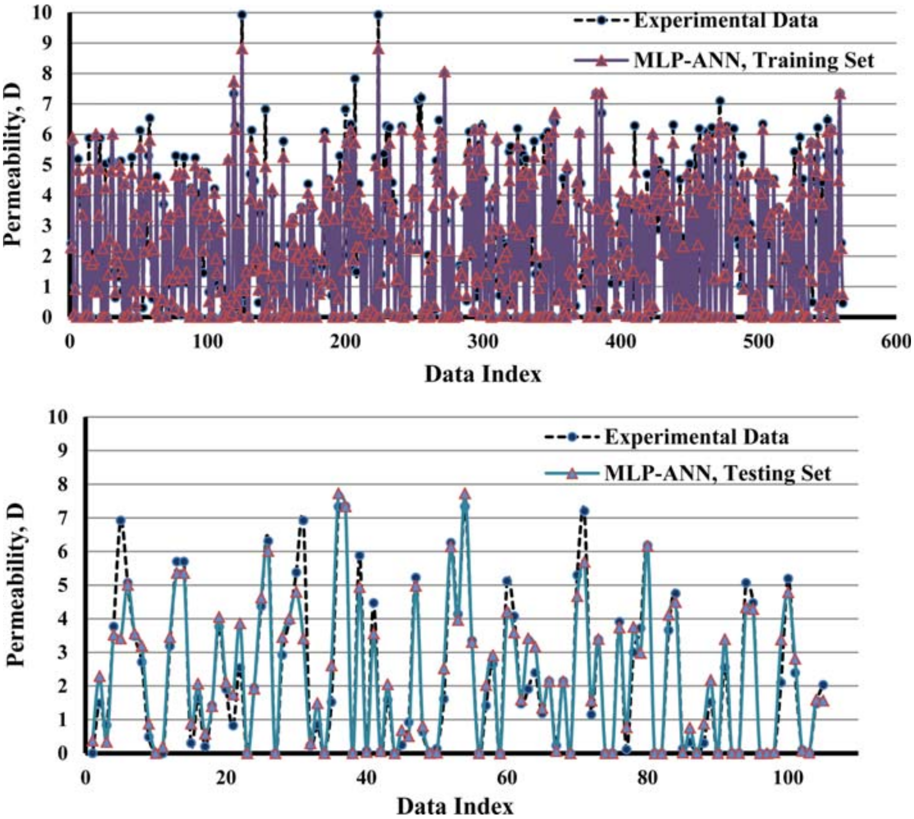

Fig. 7. Comparison between experimental data and estimated permeability by MLP-ANN model.

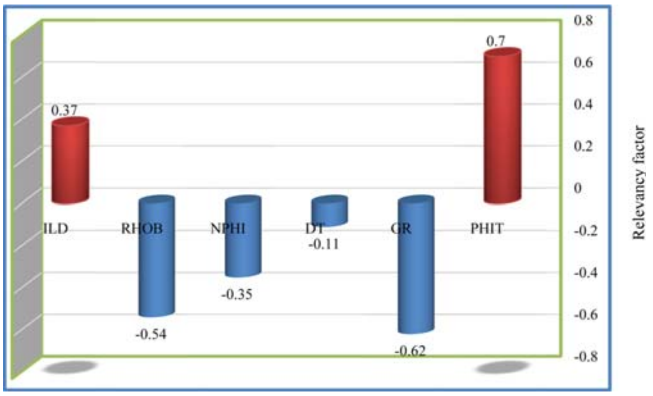

Fig. 8. The effects of well log data on the permeabilities estimated by CSA-LSSVM model.

this study are important to find the importance degree of well log data on the permeability for seismic studies in future.

From the results obtained in this study, it is concluded that the CSA-LSSVM technique shows better capability for modeling nonlinear problems than the traditional neural networks so that this makes it more attractive for prediction of permeability of heterogeneous oil reservoirs. As a consequence, the ANN methodology has a poor performance for prediction targets in the presence of a small size of dataset because of its high number of adjustable parameters. Additionally, the models developed on the basis of ANN approach have the over-fitting problem so that they may lead to error in the outside ranges of data which have been developed based on. Furthermore, the number of adjustable parameters required to develop an ANN based model is many more than the LSSVM methodology. As a result, LSSVM methodology has only two adjustable parameters. Normally, the high number of adjustable parameters causes the over-fitting problem, and decreases the capability and reliability of modeling approaches in solving nonlinear problems. The advantage of LSSVM approach is that it does not need to use a large number of data points in order to optimize and achieve the best (optimal) condition for prediction. Whereas, other locally regression methods such as neural networks have a poor performance for prediction in the presence of a small size of dataset, as pointed out earlier. Additionally, priori determination of the network topology is not required in LSSVM approach and can be automatically determined as the training process. Moreover, the number of hidden nodes and hidden layers should not be determined in CSA-LSSVM model. Furthermore, this model has fewer adjustable parameters (typically two) compared to ANN methods. However, despite the attractive benefits in terms of time, cost, and mathematical point of views, the LSSVM approach has some potential disadvantages: 1) every data point of an existing database is contributing to the model developed and the relative importance of a data point is given by its support value; 2) the second problem is that it is well known that the utilize of a sum squared error cost function without regularization might lead to predictions which are less robust (Suykens et al., 2002a).

As a matter of fact, accurate determination of permeability is of great importance in issues related to fluid flow through porous media and reservoir engineering. Therefore, the models developed in this study can be applied precisely in calculations associated with the production rates and Darcy's equation. Furthermore, the LSSVM model has an acceptable capability to apply in reservoir and production engineering software. Additionally, the determination of permeability for reservoir candidate to apply enhanced oil recovery methods is a critical issues nowadays. Regarding the well log data, the model developed in the current study is able to determine the range of permeability before implanting different enhanced oil recovery in reservoirs candidates.

5. CONCLUSION

The least square support vector machine regression methodology as a supervised learning technique has been described in order to predict permeability distribution of a heterogeneous carbonate reservoir in this study. A hybrid LSSVM-CSA technique was used to optimize the model parameters. Moreover, an MLP-ANN model was constructed to compare the results and also to investigate advantages of the proposed LSSVM model. The accuracy and prediction capability of these models have been tested through statistical and graphical error analyses, and the superiority of the CSA-LSSVM model over the MLP-ANN model was shown. The results illustrated that SVR-based technique with the tuning parameters optimized by the CSA, can result in excellent generalization and can advantageously predict the permeability in heterogeneous carbonate reservoirs. The CSA-LSSVM model can easily be implemented in any reservoir simulation software and outperforms the traditional ANN models in permeability prediction.

REFERENCES

Ahmadi, M.A. and Shadizadeh, S.R., 2012, New approach for prediction of asphaltene precipitation due to natural depletion by using evolutionary algorithm concept. Fuel, 102, 716–723.

Al-Anazi, A. and Gates, I., 2010a, A support vector machine algorithm to classify lithofacies and model permeability in heterogeneous reservoirs. Engineering Geology, 114, 267–277.

Al-Anazi, A. and Gates, I., 2010b, Support vector regression for porosity prediction in a heterogeneous reservoir: A comparative study. Computers & Geosciences, 36, 1494–1503.

Al-Anazi, A. and Gates, I., 2012, Support vector regression to predict porosity and permeability: Effect of sample size. Computers & Geosciences, 39, 64–76.

Alizadeh, N., Mighani, S., Hashemi kiasari, H., Hemmati-Sarapardeh, A., and Kamari, A., 2003, Application of Fast-SAGD in Naturally Fractured Heavy Oil Reservoirs: A Case Study. Proceedings of the 18th Middle East Oil & Gas Show and Conference: Transforming the Energy Future (MEOS), Manama, March 10–13, 3 p. 1946–1953.

Balan, B., Mohaghegh, S., and Ameri, S., 1995, State-of-the-art in permeability determination from well log data: Part 1-A comparative study, model development. SPE, 30978, 17–21.

Bhatt, A. and Helle, H.B., 2002, Committee neural networks for porosity and permeability prediction from well logs. Geophysical Prospecting, 50, 645–660.

Chamkalani, A., Amani, M., Kiani, M.A., and Chamkalani, R., 2013, Assessment of asphaltene deposition due to titration technique. Fluid Phase Equilibria, 339, 72–80.

Chen, G., Fu, K., Liang, Z., Sema, T., Li, C., Tontiwachwuthikul, P., and Idem, R., 2014, The genetic algorithm based back propagation neural network for MMP prediction in CO_2-EOR process. Fuel, 126, 202–212.

Choisy, C. and Belaid, A., 2001, Handwriting recognition using local methods for normalization and global methods for recognition. Proceedings of 6th International Conference on Document Analysis and Recognition, Seattle, Sep. 10–13, p. 23–27.

El-Sebakhy, E.A., 2009, Forecasting PVT properties of crude oil systems based on support vector machines modeling scheme. Journal of Petroleum Science and Engineering, 64, 25–34.

Esfahani, S., Baselizadeh, S., and Hemmati-Sarapardeh, A., 2015, On determination of natural gas density: Least square support vector machine modeling approach. Journal of Natural Gas Science and Engineering, 22, 348–358.

Eslamimanesh, A., Gharagheizi, F., Illbeigi, M., Mohammadi, A.H., Fazlali, A., and Richon, D., 2012, Phase equilibrium modeling of clathrate hydrates of methane, carbon dioxide, nitrogen, and hydrogen + water soluble organic promoters using Support Vector Machine algorithm. Fluid Phase Equilibria, 316, 34–45.

Fathinasab, M., Ayatollahi, S., and Hemmati-Sarapardeh, A., 2015, A Rigorous Approach to Predict Nitrogen-Crude Oil Minimum Miscibility Pressure of Pure and Nitrogen Mixtures. Fluid Phase Equilibria, 399, 30–39.

Fayazi, A., Arabloo, M., Shokrollahi, A., Zargari, M.H., and Ghazanfari, M.H., 2013, State of the Art of Least Square Support Vector Machine for Accurate Determination of Natural Gas Viscosity. Industrial & Engineering Chemistry Research, 53, 945–958.

Ganguly, S., 2003, Prediction of VLE data using radial basis function network. Computers & chemical engineering, 27, 1445–1454.

Gao, D., Zhou, J., and Xin, L., 2001, SVM-based detection of moving vehicles for automatic traffic monitoring. Proceedings of Intelligent Transportation Systems, Oakland, Aug. 25–29, p. 745–749.

Gharagheizi, F., Eslamimanesh, A., Farjood, F., Mohammadi, A.H., and Richon, D., 2011, Solubility parameters of nonelectrolyte organic compounds: determination using quantitative structure-property relationship strategy. Industrial & Engineering Chemistry Research, 50, 11382–11395.

Gharbi, R., 1997, Estimating the isothermal compressibility coefficient of undersaturated Middle East crudes using neural networks. Energy & Fuels, 11, 372–378.

Ghiasi, M.M., Bahadori, A., Zendehboudi, S., Jamili, A., and Rezaei-Gomari, S., 2013, Novel methods predict equilibrium vapor methanol content during gas hydrate inhibition. Journal of Natural Gas Science and Engineering, 15, 69–75.

Hashemi-Kiasari, H., Hemmati-Sarapardeh, A., Mighani, S., Mohammadi, A.H., and Sedaee-Sola, B., 2014, Effect of operational parameters on SAGD performance in a dip heterogeneous fractured reservoir. Fuel, 122, 82–93.

Hemmati-Sarapardeh, A., Alipour-Yeganeh-Marand, R., Naseri, A., Safiabadi, A., Gharagheizi, F., Ilani-Kashkouli, P., and Mohammadi, A.H., 2013, Asphaltene precipitation due to natural depletion of reservoir: Determination using a SARA fraction based intelligent model. Fluid Phase Equilibria, 354, 177–184.

Hemmati-Sarapardeh, A., Majidi, S.-M.-J., Mahmoudi, B., Ahmad Ramazani, S.A., and Mohammadi, A., 2014, Experimental measurement and modeling of saturated reservoir oil viscosity. Korean Journal of Chemical Engineering, 31, 1253–1264.

Hosseinzadeh, M. and Hemmati-Sarapardeh, A., 2014, Toward a predictive model for estimating viscosity of ternary mixtures containing ionic liquids. Journal of Molecular Liquids, 200, 340–348.

Kamari, A., Gharagheizi, F., Bahadori, A., Mohammadi, A.H., 2014, Rigorous Modeling for Prediction of Barium Sulfate (Barite) Deposition in Oilfield Brines. Fluid Phase Equilibria, 366, 117–126.

Kamari, A., Hemmati-Sarapardeh, A., Mirabbasi, S.-M., Nikookar, M., and Mohammadi, A.H., 2013a, Prediction of sour gas compressibility factor using an intelligent approach. Fuel Processing Technology, 116, 209–216.

Kamari, A., Khaksar-Manshad, A., Gharagheizi, F., Mohammadi, A.H., and Ashoori, S., 2013b, Robust Model for the Determination of Wax Deposition in Oil Systems. Industrial & Engineering Chemistry Research, 52, 15664–15672.

Kamari, A., Bahadori, A., Mohammadi, A.H., and Zendehboudi, S., 2014a, Evaluating the Unloading Gradient Pressure in Continuous Gas-lift Systems During Petroleum Production Operations. Petroleum Science and Technology, 32, 2961–2968.

Kamari, A., Mohammadi, A., Bahadori, A., and Zendehboudi, S., 2014b, A Reliable Model for Estimating the Wax Deposition Rate During Crude Oil Production and Processing. Petroleum Science and Technology, 32, 2837–2844.

Kamari, A., Mohammadi, A.H., Bahadori, A., and Zendehboudi, S., 2014c, Prediction of Air Specific Heat Ratios at Elevated Pressures Using a Novel Modeling Approach. Chemical Engineering & Technology, 37, 2047–2055.

Kamari, A., Safiri, A., and Mohammadi, A.H., 2015, A Compositional Model for Estimating Asphaltene Precipitation Conditions in Live Reservoir Oil Systems. Journal of Dispersion Science and Technology, 36, 301–309.

Kamari, A., Arabloo, M., Shokrollahi, A., Gharagheizi, F., and Mohammadi, A.H., 2015a, Rapid method to estimate the minimum miscibility pressure (MMP) in live reservoir oil systems during CO_2 flooding. Fuel, 153, 310–319.

Kamari, A., Bahadori, A., Mohammadi, A.H., and Zendehboudi, S., 2015b, New tools predict monoethylene glycol injection rate for natural gas hydrate inhibition. Journal of Loss Prevention in the Process Industries, 33, 222–231.

Kamari, A., Hemmati-Sarapardeh, A., Mohammadi, A.H., Hashemi-Kiasari, H., and Mohagheghian, E., 2015c, On the evaluation of Fast-SAGD process in naturally fractured heavy oil reservoir. Fuel, 143, 155–164.

Karimpouli, S., Fathianpour, N., and Roohi, J., 2010, A new approach to improve neural networks' algorithm in permeability prediction of petroleum reservoirs using supervised committee machine neural network (SCMNN). Journal of Petroleum Science and Engineering, 73, 227–232.

Kaviani, D., Bui, T., Jensen, J.L., and Hanks, C., 2008, The Application of Artificial Neural Networks With Small Data Sets: An Example for Analysis of Fracture Spacing in the Lisburne Formation Northeastern Alaska. SPE Reservoir Evaluation & Engineering, 11, 598–605.

Laugier, S. and Richon, D., 2003, Use of artificial neural networks for calculating derived thermodynamic quantities from volumetric property data. Fluid Phase Equilibria, 210, 247–255.

Li, Z., Weida, Z., and Licheng, J., 2000, Radar target recognition based on support vector machine. Proceedings of 5th International Conference on Signal Processing, Beijing, Aug. 21–25, p. 1453–1456.

Ma, C., Randolph, M.A., and Drish, J., 2001, A support vector machines-based rejection technique for speech recognition. Proceedings of Acoustics, Speech, and Signal, Salt Lake City, May 7–11, p. 381–384.

Mohaghegh, S., Arefi, R., Ameri, S., and Hefner, M.H., 1994, A Methodological Approach for Reservoir Heterogeneity Characterization Using Artificial Neural Networks. Proceedings of SPE Annual Technical Conference and Exhibition, New Orleans, Sep. 25–28, SPE 28394.

Mohaghegh, S., Arefi, R., Bilgesu, I., Ameri, S., and Rose, D., 1995, Design and development of an artificial neural network for estimation of formation permeability. SPE Computer Applications, 7, 151–154.

Mohaghegh, S., 2000, Virtual intelligence and its applications in petroleum engineering. Journal of Petroleum Technology. Distinguished Author Series, 52. http://dx.doi.org/10.2118/58046-JPT

Montgomery, D.C., 2008, Design and analysis of experiments (7th edition). John Wiley & Sons Inc., Hoboken, 656 p.

Nejatian, I., Kanani, M., Arabloo, M., Bahadori, A., and Zendehboudi, S., 2014, Prediction of natural gas flow through chokes using support vector machine algorithm. Journal of Natural Gas Science and Engineering, 18, 155–163.

Nowroozi, S., Ranjbar, M., Hashemipour, H., and Schaffie, M., 2009, Development of a neural fuzzy system for advanced prediction of dew point pressure in gas condensate reservoirs. Fuel Processing Technology, 90, 452–457.

Pelckmans, K., Suykens, J.A., Van Gestel, T., De Brabanter, J., Lukas, L., Hamers, B., De Moor, B., and Vandewalle, J., 2002, LS-SVMlab: a matlab/c toolbox for least squares support vector machines. Tutorial. KULeuven-ESAT. Leuven, Belgium, 8 p.

Ramgulam, A., 2006, Utilization of artificial neural networks in the optimization of history matching. M.Sc. Thesis, The Pennsylvania State University, University Park, 118 p.

Saeedi, A., Camarda, K.V., and Liang, J.-T., 2007, Using Neural Networks for Candidate Selection and Well Performance Prediction in Water-Shutoff Treatments Using Polymer Gels-A Field-Case Study. SPE Production & Operations, 22, 417–424.

Saemi, M., Ahmadi, M., and Varjani, A.Y., 2007, Design of neural networks using genetic algorithm for the permeability estimation of the reservoir. Journal of Petroleum Science and Engineering, 59, 97–105.

Shafiei, A., Dusseault, M.B., Zendehboudi, S., and Chatzis, I., 2013, A new screening tool for evaluation of steamflooding performance in Naturally Fractured Carbonate Reservoirs. Fuel, 108, 502–514.

Suykens, J.A. and Vandewalle, J., 1999, Least squares support vector machine classifiers. Neural processing letters, 9, 293–300.

Suykens, J.A., De Brabanter, J., Lukas, L., and Vandewalle, J., 2002a, Weighted least squares support vector machines: robustness and sparse approximation. Neurocomputing, 48, 85–105.

Suykens, J.A., Van Gestel, T., De Brabanter, J., De Moor, B., Vandewalle, J., Suykens, J., and Van Gestel, T., 2002b, Least squares support vector machines. World Scientific, Singapore, 308 p.

Tahmasebi, P. and Hezarkhani, A., 2012, A fast and independent architecture of artificial neural network for permeability prediction. Journal of Petroleum Science and Engineering, 86, 118–126.

Talebi, R., Ghiasi, M.M., Talebi, H., Mohammadyian, M., Zendehboudi, S., Arabloo, M., and Bahadori, A., 2014, Application of soft computing approaches for modeling saturation pressure of

reservoir oils. Journal of Natural Gas Science and Engineering, 20, 8–15.
Van Gestel, T., Suykens, J.A., Baestaens, D.-E., Lambrechts, A., Lanckriet, G., Vandaele, B., De Moor, B., and Vandewalle, J., 2001, Financial time series prediction using least squares support vector machines within the evidence framework. IEEE Transactions on Neural Networks, 12, 809–821.
Wong, P.M., Jang, M., Cho, S., and Gedeon, T.D., 2000, Multiple permeability predictions using an observational learning algorithm. Computers & Geosciences, 26, 907–913.
Zendehboudi, S., Chatzis, I., Mohsenipour, A.A., and Elkamel, A., 2011, Dimensional analysis and scale-up of immiscible two-phase flow displacement in fractured porous media under controlled gravity drainage. Energy & Fuels, 25, 1731–1750.
Zendehboudi, S., Shafiei, A., Bahadori, A., James, L.A., Elkamel, A., and Lohi, A., 2014, Asphaltene precipitation and deposition in oil reservoirs–Technical aspects, experimental and hybrid neural network predictive tools. Chemical Engineering Research and Design, 92, 857–875.

Manuscript received February 26, 2015
Manuscript accepted May 27, 2015

APPENDIX

A part of the actual field databank utilized in the current study to propose the predictive models is summarized in Table A1. The table lists most of the effective parameters for the estimation of permeability of heterogeneous oil reservoir. The parameters are the total porosity, gamma ray, sonic compression transit time, thermal neutron apparent porosity, bulk density, and deep Induction log, as well as corresponding the permeability data.

Table A1. Well log data used for the estimation of permeability

Index	PHIT	GR	DT	NPHI	RHOB	ILD	Permeability
1	14.4	44.777	88.2	0.241	2.402	1.857	0.0061
2	19	12.8864	90.32	0.259	2.2064	0.293	0.0086
3	33.4	15.1676	90.16	0.2398	2.22124	0.3176	5.983749
4	31.2	13.961	89.66	0.2812	2.2462	0.2076	4.041334
5	28.6	26.665	91.75	0.282	2.131	143.125	3.067557
6	28.1	19.1954	91.5	0.2616	2.212	0.34	2.378969
7	31.2	17.274	101.1	0.319	2.19	8.164	1.659201
8	31.5	16.8878	93.6748	0.2668	2.1362	187.625	5.11862
9	8.9	58.957	88.84	0.3136	2.49	0.796	0.0001
10	21.5	23.872	96.4	0.319	2.193	4.609	0.054
11	30.2	14.3808	90.84	0.2852	2.2198	0.3268	3.161486
12	30.7	18.4414	92.2124	0.2614	2.145	2000	6.042994
13	29.6	27.502	94.6872	0.312	2.0978	34.9496	2.379101
14	32.2	11.71	96.3	0.282	2.232	26.719	4.390206
15	30.2	13.6856	91.86	0.2868	2.2048	0.3148	4.032048
16	30.8	17.2338	94.42	0.2884	2.1776	0.438	2.552121
17	24.2	12.0938	93.74	0.2548	2.2186	20.7282	1.076501
18	31	11.656	91.2	0.262	2.214	4.676	5.299903
19	20.6	57.7588	100.38	0.3414	2.314	2.3852	0.0056
20	25.5	57.6692	83.2	0.3206	2.523	0.3822	0.037945
21	13	45.597	96.3	0.335	2.32	1.799	0.0018
22	22.5	80.6332	96.58	0.3956	2.3236	2.216	0.8542
23	28.8	11.2358	89.66	0.2568	2.2132	0.3808	4.474302
24	25.2	40.612	89.5	0.273	2.377	0.409	0.8266
25	33	11.8444	89.24	0.2446	2.2186	0.4304	6.189636
26	30.3	22.727	93	0.285	2.113	185.25	5.428903
27	14.2	67.4222	100.14	0.388	2.3702	1.2296	0.0016
28	20.9	51.954	90.4	0.26	2.407	1.366	0.002
29	18.7	47.2568	97.88	0.3142	2.306	1.8418	0.0035
30	11.8	52.047	89.6	0.379	2.161	1.543	0.0006
31	30.1	17.9964	90.188	0.259	2.1318	2000	5.945203
32	26.4	26.9002	81.82	0.2794	2.5112	0.4544	0.0143
33	28.6	21.2044	97.22	0.2644	2.2364	0.46	0.845063
34	33.2	9.722	89.1	0.277	2.228	11.461	5.075784
35	31.2	14.1276	92.7372	0.2754	2.1504	636.6	6.701497
36	35.6	18.1234	98.08	0.3252	2.1924	0.316	2.731614
37	33.2	13.7828	93.85	0.2826	2.109	355.8	8
38	11.8	52.047	89.6	0.379	2.161	1.543	0.0006
39	32.8	14.2958	92.5996	0.2772	2.1358	448.9	9.929606
40	9.6	32.684	94.82	0.2852	2.292	2.527	0.0001
41	29.7	51.715	89.7	0.273	2.391	1.794	0.157314

Table A1. (continued)

42	29.2	7.643	91.1	0.271	2.206	10.703	4.022576
43	33.1	52.944	82.8	0.418	1.873	1.299	7.345
44	30.5	37.373	91.7	0.281	2.333	2.221	1.411469
45	30.3	22.727	93	0.285	2.113	185.25	5.428903
46	9.8	45.0748	91.36	0.2422	2.3658	2.3444	0.0006
47	29.7	22.726	94.2	0.313	2.198	13.227	1.476318
48	32.3	17.6922	89.96	0.27	2.1954	0.3724	4.496103
49	31.4	17.2864	91.36	0.218	2.1942	0.3492	4.379002
50	14.8	50.1168	99.88	0.3998	2.463	1.613	0.0002
51	29.9	20.8188	88.94	0.2752	2.2152	0.3648	3.571343
52	30.9	17.4846	94.975	0.2848	2.1608	82.0752	3.537902
53	17.1	71.524	91.6	0.288	2.434	1.005	0.0111
54	29.6	18.8316	89.38	0.2704	2.2024	0.3796	3.748102
55	31.9	14.1288	91.62	0.2612	2.2036	0.3646	5.196547
56	12.8	59.388	89.7	0.262	2.439	0.807	0.0036
57	29.4	13.5626	90.9	0.303	2.1998	0.2974	2.085143
58	26.2	36.973	95.2	0.276	2.245	3.162	0.1141
59	29.9	14.1722	91.12	0.2932	2.1926	0.3016	3.692202
60	33.4	25.53	92.5996	0.3116	2.09	20.797	3.770458
61	14.8	49.439	98.3	0.318	2.3488	1.6442	0.0027
62	25	41.045	93.6	0.353	2.311	1.394	0.151893
63	27.4	31.954	91.675	0.3118	2.2294	0.9352	0.4885
64	28.6	26.5026	91.875	0.2774	2.135	64.5128	2.099201
65	31.5	16.8878	93.6748	0.2668	2.1362	187.625	5.11862
66	32	19.262	64.4	0.298	2.168	7.828	2.144901
67	30.9	17.4846	94.975	0.2848	2.1608	82.0752	3.537902
68	29.5	17.1246	93.22	0.3194	2.181	0.1842	1.418501
69	29	17.2026	92.3372	0.2758	2.1462	342.4	6.470538
70	28.2	16.246	91.2	0.265	2.235	21.656	3.16916