Evolving artificial intelligence techniques to model the hydrate-based desalination process of produced water
Maryam Sadi, Hajar Fakharian, Hamid Ganji and Majid Kakavand

ABSTRACT
In this study, two artificial intelligence models based on an adaptive neuro-fuzzy inference system (ANFIS) and a support vector machine (SVM) technique have been successfully developed to predict the desalination efficiency of produced water through a hydrate-based desalination treatment process. A genetic algorithm as an evolutionary optimization method has been used to determine the optimal values of SVM model coefficients. To this end, compressed natural gas and CO₂ hydrate formation experiments were carried out, and the desalination efficiency of produced water was measured and utilized for model training and validation. After model development, graphical and statistical analysis approaches have been applied to evaluate the performance of suggested models by a comparison of model predictions with measured experimental data. For the ANFIS model, the coefficient of determination ($R^2$) and average absolute relative error (AARE) are 0.9927 and 0.58%, respectively. The values of AARE and $R^2$ for the SVM model are obtained 0.35% and 0.9985, respectively. These statistical criteria confirm excellent accuracy and robustness of intelligent models in predicting the desalination efficiency of produced water through the hydrate-based desalination treatment process. Furthermore, the Leverage statistical technique has been carried out to define the outliers. The obtained results demonstrate that all experimental data are reliable and both ANFIS and SVM models are statistically valid.

Key words | adaptive neuro-fuzzy inference system, artificial intelligence techniques, desalination efficiency, gas hydrate, produced water, support vector machine

NOMENCLATURE

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>Premise parameter of the membership function</td>
</tr>
<tr>
<td>$A$</td>
<td>Linguistic value of fuzzy sets</td>
</tr>
<tr>
<td>AARE</td>
<td>Average absolute relative error</td>
</tr>
<tr>
<td>ANFIS</td>
<td>Adaptive neuro-fuzzy inference system</td>
</tr>
<tr>
<td>$b$</td>
<td>Premise parameter of the membership function</td>
</tr>
<tr>
<td>$B$</td>
<td>Linguistic value of fuzzy sets</td>
</tr>
<tr>
<td>$B_s$</td>
<td>Bias term in SVM formulation</td>
</tr>
<tr>
<td>$c$</td>
<td>Premise parameter of the membership function</td>
</tr>
<tr>
<td>$C$</td>
<td>Error penalty parameter</td>
</tr>
<tr>
<td>CNG</td>
<td>Compressed natural gas</td>
</tr>
<tr>
<td>$d$</td>
<td>Premise parameter of the membership function</td>
</tr>
<tr>
<td>$d_p$</td>
<td>Parameter of the kernel function</td>
</tr>
<tr>
<td>$D$</td>
<td>Training data set</td>
</tr>
<tr>
<td>$EC$</td>
<td>Electrical conductivity, mS/cm</td>
</tr>
<tr>
<td>$f$</td>
<td>Approximation function</td>
</tr>
<tr>
<td>$GA$</td>
<td>Genetic algorithm</td>
</tr>
<tr>
<td>$H$</td>
<td>Hat matrix</td>
</tr>
<tr>
<td>$H^*$</td>
<td>Critical Leverage value</td>
</tr>
<tr>
<td>$k$</td>
<td>Kernel function</td>
</tr>
<tr>
<td>$L$</td>
<td>Prediction error loss function</td>
</tr>
<tr>
<td>$m$</td>
<td>Dimension of input parameters</td>
</tr>
</tbody>
</table>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY-NC-ND 4.0), which permits copying and redistribution for non-commercial purposes with no derivatives, provided the original work is properly cited (http://creativecommons.org/licenses/by-nc-nd/4.0/).

doi: 10.2166/wrd.2019.024
INTRODUCTION

Produced water is brine wastewater which is produced during oil and gas process (Veil et al. 2004). This water contains different impurities, including salts, mineral ions, heavy metals, organics, hydrocarbons and other contaminants (Abousnina et al. 2015). The management of produced water is highly important and can play an essential role for preventing environmental pollution. The most commercial methods for desalination are distillation, membrane and reverse osmosis processes (Khawaji et al. 2012).

Although distillation is a common technology, it is too expensive due to the consumption of a large amount of heat energy for the vaporization of water. The application of reverse osmosis in the reuse of the municipal wastewater effluent has dramatically increased over the past decade and can compete with thermal distillation that uses membranes and pressure for water desalination (Zhu et al. 2018).

Another attractive process for water treatment is freezing, which is still under development and has not been applied commercially (Han et al. 2017).

The hydrate-based desalination process was introduced in the 1940s and has been fully considered since 1960 (Veil et al. 2004; Khawaji et al. 2008; Abousnina et al. 2015). Hydrate formation can eliminate water impurities because of the hydrate chemical structure. When hydrate is formed, clean water is obtained after dissociating the hydrate particles (Cha & Seol 2006). Gas hydrate technology is still under development, but after commercialization, it can be an inexpensive option for the conventional membrane and thermal desalination processes which are applied for water treatment (Ghalavand et al. 2015).

Only a few studies have been focused on the application of hydrate formation for the desalination of produced water (Cha & Seol 2013; Fakharian et al. 2017a, 2017b). Cha & Seol (2013) proposed cyclopentane and cyclohexane as secondary fluids.
hydrate guests to reduce the temperature in a desalination process of produced water. Fakharian et al. (2017a, 2017b) introduced CO₂ and compressed natural gas (CNG) hydrate formers for reducing the salinity of produced water through the hydrate-based desalination treatment process.

Due to the complexity of water treatment processes and the limitation of traditional physics-based models (Heddam et al. 2012; Wei 2013), artificial intelligence techniques and data-driven approaches have been proposed in recent years to simulate the water treatment process (Heddam et al. 2012; Araromi et al. 2018; Nadiri et al. 2018). The main advantage of artificial intelligence-based models is the capability of these techniques for modeling the complicated processes without any need for detailed information from the studied systems (Sadi & Shahrabadi 2018).

Wei (2013) proposed an improved dynamic neural network model to predict the total suspended solids as one of the major water pollutants that cause deterioration of the water quality. Jing et al. (2018) utilized a fuzzy inference system technique to study the removal of polycyclic aromatic hydrocarbons from offshore produced water through an ozonation process. Araromi et al. (2018) applied the adaptive neuro-fuzzy inference system (ANFIS) and the generalized linear model technique to simulate a biological wastewater treatment process.

As mentioned, intelligent approaches have been implemented to simulate the water treatment process, but to the best of our knowledge, there is no report on the application of artificial intelligence methods to model the hydrate-based desalination treatment process of produced water.

In the present study, two intelligent models based on the ANFIS and the support vector machine (SVM) technique have been successfully developed to predict the desalination efficiency of produced water through the hydrate-based desalination treatment process as a function of the initial salinity of produced water and gas hydrate equilibrium pressure. Due to the flexibility, simplicity and self-adaption capability of evolutionary optimization methods, the genetic algorithm (GA) has been coupled by a SVM approach to determine optimum values of model parameters. Salt removal efficiency has been measured experimentally for different produced water samples in the presence of CNG and CO₂ as a hydrate former and the measured data have been applied for model training and validation. Then, the reliability of developed models in the prediction of desalination efficiency is evaluated using both graphical and statistical error analysis techniques. Furthermore, the Leverage statistical method is employed for outlier detection purpose.

**EXPERIMENTAL SECTION**

**Experimental setup and procedure**

The experimental setup used in this research has been discussed in detail in Fakharian et al. (2017a, 2017b). This setup contains a 300 cm³ reactor that is placed in a cooling medium to control the temperature. The temperature and pressure of the reactor were measured using a thermocouple and a pressure transducer, respectively. A computer system with the suitable data acquisition software was used to record and collect experimental data during the time. The concentrated salty water was drained from the reactor and stored in a container.

Two different hydrate formers, CO₂ and CNG, were separately injected to the reactor which contained different produced water samples to start hydrate formation. The reactor pressure and temperature were recorded to monitor a hydrate formation trend. After hydrate formation, the concentrated saline water was drained and the produced hydrate was washed and filtered to increase salt removal efficiency and finally the hydrate was decomposed to produce fresh water.

Electrical conductivity (EC) of the initial water sample and desalinated water produced from hydrate dissociation was measured using a conductivity meter.

**Experimental data**

In the present study, a total of 120 input–output experimental data points have been measured and applied to develop intelligent model structures. For this purpose, 75% of data points which were selected randomly have been used as a training data set to train the model network and calculate the best values of model parameters. Also, 25% of data points have been utilized as a testing subset for model validation and evaluating the generalization ability of the proposed structure. It should be noted that the random dividing of experimental data points into training and testing subsets has been carried out several times to prevent from
the local accumulation of implemented data and obtain a homogeneous distribution (Ghorbani et al. 2016; Sadi & Shahrabadi 2018).

The statistical details of empirical data, including input and target parameters which applied to develop intelligent models, are represented in Table 1. In this table, \( Pe \) denotes the gas hydrate equilibrium pressure at a constant temperature, which is obtained from the Pressure–Temperature diagram. It should be noted that \( Pe \) is considered as an input parameter to discriminate between different types of hydrate former gases (CO₂ or CNG).

### INTELLIGENT MODEL DEVELOPMENT

#### Data normalization

Data normalization which is considered as the basic type of data preprocessing method can be applied in data-based modeling techniques to unify the effect of different scale variables and increase the convergence speed of model training phase (Graf & Borer 2003). There are different types of data normalization methods, such as sigmoid normalization, zero-mean normalization, min-max normalization, soft-max normalization and decimal scaling.

In the present study, the min-max normalization approach has been utilized to normalize input data and corresponding output values using the following equation:

\[
X_{\text{norm}} = 2 \times \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} - 1
\]

where \( X \) is the actual data before normalization, \( X_{\text{max}} \) and \( X_{\text{min}} \) denote the maximum and minimum values of actual data and \( X_{\text{norm}} \) represents the normalized data.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial EC (mS/cm)</td>
<td>114</td>
<td>173</td>
<td>137.33</td>
</tr>
<tr>
<td>( Pe ) (psia)</td>
<td>142</td>
<td>202</td>
<td>172</td>
</tr>
<tr>
<td>Target parameter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desalination efficiency (%)</td>
<td>32.89</td>
<td>59.54</td>
<td>45.51</td>
</tr>
</tbody>
</table>

This normalization procedure reduces the effect of higher valued variables on the model output and increases the prediction accuracy by transforming the data set from \([X_{\text{min}}, X_{\text{max}}]\) range to \([-1, 1]\) scale.

#### Adaptive neuro-fuzzy inference system

The ANFIS firstly proposed by Jang (1993) is an intelligent modeling technique based on the first-order Takagi–Sugeno type of fuzzy system. In this method, the learning ability of artificial neural network and the knowledge principles of fuzzy logic are applied simultaneously to model nonlinear systems with a single target value.

For ANFIS model development, a training data set is utilized to fine-tune the fuzzy ‘if-then’ rules by using a neural network training concept. During this step to match model predictions with target values, the optimum type and the number of membership functions are specified. After training the ANFIS model, a testing data set is used to evaluate the generalization ability of the proposed structure.

For a first-order Takagi–Sugeno fuzzy system, a common set of fuzzy ‘if-then’ rules for a system with one output \( z \) and two input parameters \( x_1 \) and \( x_2 \) can be described as follows:

\[
\text{Rule 1: If } x_1 \text{ is } A_1 \text{ and } x_2 \text{ is } B_1 \text{ then } z_1 = p_1 x_1 + q_1 x_2 + r_1
\]

\[
\text{Rule 2: If } x_1 \text{ is } A_2 \text{ and } x_2 \text{ is } B_2 \text{ then } z_2 = p_2 x_1 + q_2 x_2 + r_2
\]

In the above equations, \( A_i \) and \( B_i \) are linguistic values of membership functions defined by fuzzy sets and \( p_i, q_i, \) and \( r_i \) represent consequent parameters of fuzzy ‘if-then’ rules. As observed, each fuzzy rule consists of two parts. The ‘if’ part of the fuzzy rule is known as premise section and the ‘then’ part is called consequent section.

As depicted in Figure 1, the ANFIS model has five layers with five different function types which produce a target value as a combination of input parameters. Some of these node functions have adjustable parameters which are calculated during network training. A brief description of ANFIS layer node functions is presented below:
First layer: Fuzzification
This layer has an adaptive node function which calculates a membership grade of the node input for a fuzzy set. The output of each node in the first layer \(O_{1,i}\) is a degree of membership value (between 0 and 1) in which the input variable \(x\) satisfies the membership function \(\mu_{Ai}\):

\[
O_{1,i} = \mu_{Ai}(x)
\]  

The common types of membership functions applied in the ANFIS technique are trapezoidal, generalized bell, Gaussian and triangular. The mathematical definitions of these parameterized membership functions are as follows:

\[
\text{Trapezoidal}(x; a, b, c, d) = \begin{cases} 
0 & x \leq a \\
\frac{x - a}{b - a} & a \leq x \leq b \\
1 & b \leq x \leq c \\
\frac{d - x}{d - c} & c \leq x \leq d \\
0 & d \leq x 
\end{cases}
\]

\[
\text{Generalized Bell}(x; a, b, c) = \frac{1}{1 + \frac{(x - c)^2}{a^2}}
\]

\[
\text{Gaussian}(x; \sigma, c) = \exp \left( -\frac{(x - c)^2}{2\sigma^2} \right)
\]

\[
\text{Triangular}(x; a, b, c) = \begin{cases} 
0 & x \leq a \\
\frac{x - a}{b - a} & a \leq x \leq b \\
\frac{c - x}{c - b} & b \leq x \leq c \\
0 & c \leq x 
\end{cases}
\]

where \(a, b, c, d\) and \(\sigma\) are adjustable parameters of the membership function which are known as premise parameters.

Second layer: If-then rules nodes
This layer has a fixed or non-adaptive node function labeled \(\Pi\) as 'Prod'. The output of the second layer \(O_{2,i}\), which represents the firing strength of the associated fuzzy 'if-then' rule \(w_i\), is calculated by multiplying all incoming signals:

\[
O_{2,i} = w_i = \Pi \mu_i = \mu_{Ai}(x) \ast \mu_{Bi}(x), \quad i = 1, 2
\]

where \(\mu_{Ai}\) and \(\mu_{Bi}\) denote the membership functions.

Third layer: Normalization
Similar to the previous layer, the third layer has a fixed node function labeled \(N\) as 'Norm'. The output of this layer \(O_{3,i}\), which called the normalized firing strength \(\tilde{w}_i\), is...
calculated as the ratio of the \(i\)th rule’s firing strength to the sum of all rules’ firing strengths:

\[
O_{3,i} = \frac{\bar{w}_i}{\sum \bar{w}_i} = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \tag{10}
\]

Fourth layer: Defuzzification

This layer has an adaptive node function. The output of the fourth layer \((O_4)\) is obtained by multiplying the normalized firing strength in the consequent section of fuzzy rule:

\[
O_{4,i} = \bar{w}_i z_i = \frac{w_i}{w_1 + w_2} (p_1 x_1 + q_1 x_2 + r_1), \quad i = 1, 2 \tag{11}
\]

Fifth layer: Summation

The last layer has a single fixed node labeled \(\Sigma\) as ‘Sum’. The overall output of the ANFIS model \((O_5)\) is computed in this layer by summation of all incoming signals from the previous layer:

\[
O_5 = \sum_i \bar{w}_i z_i = \sum_i \left( \frac{w_i}{w_1 + w_2} \right) z_i
\]

\[
= \frac{w_1}{w_1 + w_2} (p_1 x_1 + q_1 x_2 + r_1) + \frac{w_2}{w_1 + w_2} (p_2 x_1 + q_2 x_2 + r_2) \tag{12}
\]

For training the ANFIS structure and calculating the optimum values of unknown parameters, a hybrid learning algorithm consisting of two steps is applied. In the first step known as forward pass, the least square technique is utilized to identify the optimum values of the consequent parameters. In the second step which referred to as backward pass, the error rates are propagated backward to the first layer and the premise parameters are updated using the gradient descent technique.

Support vector machine

SVM introduced by Vapnik (1995) is a supervised learning approach originated from statistical learning theory (Vapnik 1998). The SVM technique which uses the statistical risk minimization principles is developed for both classification problems and regression tasks. The basic idea of SVM is fitting an approximation function as flat as possible to map data points from an input space into a feature space with a higher dimension (Smola & Scholkopf 2004).

Suppose a training data set as \(D = \{ (x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N) \}\), where \(x_i \in \mathbb{R}^m\) and \(y_i \in \mathbb{R}\), \(i = \{1, 2, \ldots, N\}\) stand for the input variables and corresponding output values, respectively; and \(N\) and \(m\) represent the number of training data points and the dimension of input parameters. For this data set, the SVM formulation can be expressed as follows:

\[
y = f(x) = \langle W, \varphi(x) \rangle + B_s \tag{13}
\]

In the above equation \(\langle \cdot, \cdot \rangle\) denotes dot product, \(W\) and \(B_s\) are the adjustable weight factor and the bias term, and \(\varphi(x)\) represents the mapping function which transfers input data \((x)\) from \(\mathbb{R}^m\) space into a higher dimensional feature space in order to estimate the model target \((\hat{y})\) as close as possible to its actual value \((y)\).

During the SVM training step, the following cost function is minimized to obtain the optimum values of \(W\) and \(B_s\) as model parameters:

\[
\text{Cost function} = \frac{1}{2} ||W|| + \frac{C}{N_t} \sum_{i=1}^{N_t} L(f(x_i), y_i) \tag{14}
\]

where \(L\) is the prediction error loss function for training data points and \(C\) is a positive constant known as error penalty parameter which defines the trade-off between training error and model complexity.

Based on the statistical risk minimization concept, minimization of the above-mentioned cost function leads to the following convex quadratic optimization problem (Vapnik 1995):

\[
\min_{W, B_s, \xi_i} \frac{1}{2} ||W|| + C \sum_{i=1}^{N_t} (\xi_i + \xi_i^*)
\]

\[
y_i - \langle \langle W, \varphi(x_i) \rangle + B_s \rangle \leq \varepsilon + \xi_i
\]

Subject to:

\[
\xi_i > 0 \quad \xi_i^* > 0
\]

where \(\varepsilon\) is the maximum allowable deviation of the predicted value \((\hat{y})\) from actual data \((y)\) which is known as approximation precision, and \(\xi_i\) and \(\xi_i^*\) are the slack variables.
After utilizing Lagrange multipliers to the above-mentioned optimization problem, the final form of the regression function can be written as follows:

\[ f(x) = \sum_{i=1}^{N_i} (\alpha_i - \alpha_i^*) k(x_i, x) + B_s \]  

(17)

Here, \( \alpha_i \) and \( \alpha_i^* \) are Lagrange multipliers, and \( k(x_i, x) \) denotes the kernel function. The common types of kernel functions are polynomial, Gaussian or radial basis function (RBF) and sigmoid function which are defined by the following equations:

Polynomial: \( k(x_i, x) = (\langle x_i, x \rangle + 1)^{d_p} \)  

(18)

Gaussian or RBF: \( k(x_i, x) = \exp\left(-\frac{||x_i - x||^2}{2\gamma^2}\right) \)  

(19)

Sigmoid: \( k(x_i, x) = \tanh(\beta \langle x_i, x \rangle + \nu) \)  

(20)

where \( d_p, \gamma, \beta \) and \( \nu \) are the kernel function parameters which should be optimized along with two other parameters \((C \text{ and } \epsilon) \) at SVM training step using any appropriate optimization techniques, such as simulated annealing, GA and harmony search.

**Genetic algorithm**

The GA, firstly proposed by Holland (1975) and then developed by Goldberg (1989), is a metaheuristic optimization method which is inspired by Darwin's theory. This direct search method applies principles of the survival of the fittest to solve the complicated optimization problem.

The GA is an iterative process which starts by the random generation of an initial solution in the feasible region. After the creation of initial population, the fitness values of the generated individuals are calculated and the next generation members are selected in a reproduction step through a fitness-based procedure. The selected members which are known as parents are recombined in a crossover step to create two new offsprings as children population. To keep GA diversity, children members are randomly modified in a mutation step by a small probability known as mutation rate (Goldberg 1989).

These steps are repeated to create a new population at each generation until the satisfaction of the algorithm stopping criterion which can be defined as the maximum number of generation or a prespecified convergence value.

The GA parameters, applied in this study to optimize the coefficients of the kernel function and SVM formulation, are summarized in Table 2.

**RESULTS AND DISCUSSION**

**Evaluation of developed model performance**

In this study, ANFIS and SVM techniques have been applied to model the hydrate-based desalination treatment process and predict the desalination efficiency of produced water using gas hydrate. In the developed models, the initial salinity of produced water and gas hydrate equilibrium pressure have been selected as model inputs to determine desalination efficiency as a target value. Desalination efficiency \((\eta_{\text{DeSal}})\) is defined as follows:

\[ \eta_{\text{DeSal}}(\%) = \left(\frac{\text{EC}_i - \text{EC}_f}{\text{EC}_i}\right) \times 100 \]  

(21)

where \(\text{EC}_i\) and \(\text{EC}_f\) are the EC of initial and final brine solution (water produced from hydrate dissociation) which show produced water salinity before and after the desalination treatment process, respectively.

The reliability and precision of the proposed models have been evaluated using graphical representations and statistical analysis methods. For the graphical analysis of model performance, cross plot and relative error diagram have been plotted. Also, some statistical parameters, including coefficient of determination \((R^2)\), average absolute
Relative error (AARE) and mean square error (MSE), have been calculated. These statistical criteria are denoted as follows:

\[ R^2 = 1 - \frac{\sum_{i=1}^{n} (y_{\text{exp}} - y_{\text{cal}})^2}{\sum_{i=1}^{n} (y_{\text{exp}} - \bar{y}_{\text{exp}})^2} \]  

(22)

\[ \text{AARE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_{\text{exp}} - y_{\text{cal}}}{y_{\text{exp}}} \right| \times 100 \]  

(23)

\[ \text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_{\text{exp}} - y_{\text{cal}})^2 \]  

(24)

where \( y_{\text{exp}}, y_{\text{cal}}, \bar{y}_{\text{exp}} \) and \( n \) are experimental data, model predictions, a mean value of experimental data and the number of measured data points, respectively.

The optimum structure of the developed ANFIS model is demonstrated in Figure 2, and the characterization of this optimum architecture is summarized in Table 3. As observed, the triangular-type membership function with five linguistic levels for the initial salinity of produced water and three levels for gas hydrate equilibrium pressure is the best ANFIS model structure.

\begin{table}[h]
\centering
\caption{Characterization of the developed ANFIS model}
\begin{tabular}{|l|l|}
\hline
\textbf{Parameters} & \textbf{Value} \\
\hline
Membership function type & Triangular \\
Number of membership function for initial salinity of produced water & 5 \\
Number of membership function for gas hydrate equilibrium pressure & 3 \\
Number of fuzzy rules & 15 \\
Number of linear parameters & 45 \\
Number of nonlinear parameters & 24 \\
Output membership function type & Linear \\
Optimization method & Hybrid \\
Epoch number & 80 \\
\hline
\end{tabular}
\end{table}

It is obvious that the performance of the SVM model strongly depends on the values of adjustable parameters of SVM formulation as well as kernel function type and parameters. Due to the high accuracy and acceptable generalization ability of the Gaussian kernel function (Ghorbani et al. 2016), this type of kernel function has been used in the present study. The optimum values of SVM model parameters, including SVM formulation and...
kernel function parameters calculated by the GA technique, are summarized in Table 4.

The values of statistical parameters for developed ANFIS and SVM models are listed in Table 5. These results indicate the reliability and excellent precision of both intelligent models in predicting desalination efficiency through the hydrate-based desalination treatment process. Based on the reported values, it can be said that the SVM model has a relatively better performance than the ANFIS model in the prediction of desalination efficiency.

In addition to the statistical technique, the graphical error analysis is also used to verify the reliability of the proposed models. The related graphs are shown in Figures 3–5.

The cross plot of model predictions and experimental data for ANFIS and SVM models are demonstrated in Figures 3(a) and 3(b), respectively. As observed, all experimental data are accumulated around the diagonal line which confirms the accuracy and robustness of both intelligent models.

Moreover, experimental data and predicted values for ANFIS and SVM models are plotted versus data points in Figures 4(a) and 4(b), respectively. These figures indicate the excellent agreement between model predictions and experimental values for both training and testing subsets.

Furthermore, the relative deviations of modeling results from experimental values for ANFIS and SVM models are depicted in Figures 5(a) and 5(b). As is evident, the absolute values of maximum relative deviations between ANFIS model results and experimental data for training and testing subsets are 3.53 and 7.91%, respectively. These values for the SVM model are 3.36 and 4.67%, respectively. These results reveal that the developed models can precisely predict the desalination efficiency of produced water through the hydrate-based desalination process.

<p>| Table 4 | Optimal values of SVM model parameters |</p>
<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>612.37</td>
</tr>
<tr>
<td>ε</td>
<td>0.0068</td>
</tr>
<tr>
<td>γ</td>
<td>0.5417</td>
</tr>
</tbody>
</table>

<p>| Table 5 | Statistical criteria for developed ANFIS and SVM models |</p>
<table>
<thead>
<tr>
<th>Parameters</th>
<th>ANFIS</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>R^2 (Train)</td>
<td>0.9980</td>
<td>0.9992</td>
</tr>
<tr>
<td>MSE (Train)</td>
<td>0.0944</td>
<td>0.0453</td>
</tr>
<tr>
<td>AARE (Train)</td>
<td>0.3278</td>
<td>0.2916</td>
</tr>
<tr>
<td>R^2 (Test)</td>
<td>0.9841</td>
<td>0.9966</td>
</tr>
<tr>
<td>MSE (Test)</td>
<td>0.6856</td>
<td>0.1617</td>
</tr>
<tr>
<td>AARE (Test)</td>
<td>1.1018</td>
<td>0.5023</td>
</tr>
<tr>
<td>R^2 (Total)</td>
<td>0.9927</td>
<td>0.9985</td>
</tr>
<tr>
<td>MSE (Total)</td>
<td>0.2412</td>
<td>0.0843</td>
</tr>
<tr>
<td>AARE (Total)</td>
<td>0.5845</td>
<td>0.3527</td>
</tr>
</tbody>
</table>

Figure 3 | Cross plot of model predictions and experimental data for the (a) ANFIS model and (b) the SVM model.
Validation of developed models

In this section, the trend analysis of developed models is conducted in order to study the validity of both ANFIS and SVM models. To this end, desalination efficiency as a function of initial EC of produced water which indicates water salinity for CO2 and CNG gas hydrate is demonstrated in Figures 6(a) and 6(b), respectively. As observed, by increasing the initial salinity of produced water, the desalination efficiency increases accordingly. Salts are gas hydrate formation inhibitors and therefore when the water salinity increases, the amount of hydrate will be diminished. This leads to a lower amount of entrapped salt between hydrate crystals, which finally increases desalination efficiency. Also, a relatively better prediction capability of the SVM model is observed in these figures.

Furthermore, these figures show that CO2 gas hydrate has a higher average desalination efficiency to that of natural gas. Based on the experiment’s results, hydrate formed with CO2 is well packed in contrast to the natural gas hydrate which was spongy. In fact, when the hydrate is more packed, less salt can be entrapped between the hydrate crystals, which leads to more water desalination.

Detection of suspected data

The accuracy of empirical data affects the reliability and performance of the developed model (Rousseeuw & Leroy 1987). Therefore, a Leverage approach has been applied in this study for the detection of outliers. In the Leverage method, the suspected data are graphically identified by drawing the Williams plot based on the calculated Hat.
matrix \((H)\) which indicates the residual values (i.e. the differences between model predictions and experimental data). The details of the Leverage technique can be found in the literature (Mohammadi et al. 2012).

The William plots of two developed models for the estimation of the desalination efficiency of produced water have been sketched in Figures 7(a) and 7(b). As can be seen, the whole data sets are in the ranges \(-3 \leq R \leq 3\) and \(0 \leq H \leq H^* = 0.075\) (critical Leverage value) which approves the reliability of the implemented data for model development. Therefore, it can be concluded that both ANFIS and SVM models are statistically acceptable and can precisely predict the desalination efficiency of produced water in the presence of gas hydrate.

**CONCLUSION**

In the present study, ANFIS and SVM techniques have been employed for the first time to model the hydrate-based desalination process of produced water. In the developed models, the initial salinity of produced water and gas hydrate equilibrium pressure have been selected as input parameters to predict the desalination efficiency of produced water.
produced water as a model target. To develop the model structure, salt removal efficiency for different produced water in the presence of CNG and CO₂ as a hydrate former has been measured experimentally. The GA as an evolutionary optimization technique has been applied to obtain the adjustable parameters of the SVM model. The reliability of the proposed models has been evaluated by comparing model predictions with measured experimental data and calculation of statistical parameters, such as $R^2$, MSE and AARE. The AARE and MSE of the ANFIS model in the prediction of the desalination efficiency of produced water are 0.58% and 0.2412, respectively, with a high $R^2$ value of 0.9927. The values of AARE, MSE and $R^2$ for the SVM technique are obtained 0.35%, 0.0843 and 0.9985, respectively. These statistical criteria indicate the excellent agreement of ANFIS and SVM model predictions with experimental data. Therefore, it can be concluded that although the SVM method has a relatively better performance than the ANFIS technique, both proposed intelligent models can be applied by high accuracy to simulate the hydrate-based desalination process. Finally, an outlier analysis on the basis of the Leverage technique has been performed to evaluate the accuracy of the implemented data. The obtained results show that all measured data are reliable and both developed models are statistically correct and acceptable.

REFERENCES


