

PREDICTION OF CO₂ EQUILIBRIUM MOISTURE CONTENT USING LEAST SQUARES SUPPORT VECTOR MACHINES ALGORITHM

Mohammad M. Ghiasi¹, Jafar Abdi², Meysam Bahadori³, Moonyong Lee⁴, Alireza Bahadori⁵

¹ National Iranian Gas Company (NIGC), South Pars Gas Complex (SPGC), Asaluyeh, Iran

² Department of Chemical and Petroleum Engineering, Sharif University of Technology, Tehran, Iran

³ National Iranian Drilling Company, Department of Waste Management, Ahwaz, Iran.

⁴ School of Chemical Engineering, Yeungnam University, Gyeongsan, Republic of Korea

⁵ Southern Cross University, School of Environment, Science and Engineering, Lismore, NSW, Australia

Received December 25, 2015; Accepted February 16, 2016

Abstract

The burning of fossil fuels such as gasoline, coal, oil, natural gas in combustion reactions results in the production of carbon dioxide. The phase behavior of the carbon dioxide + water system is complex topic. Unlike methane, CO₂ exhibits a minimum in the water content. These minima cannot be predicted by existing methods accurately. In this communication, two mathematical-based procedures have been proposed for accurate computation of CO₂ water content for temperatures between 273.15 and 348.15 K and the pressure range between 0.5 and 21 MPa.

The first is based on least squares support vector machine (LSSVM) algorithm and the second applies multilayer perceptron (MLP) artificial neural network (ANN). Furthermore, the constants of the previously developed empirical correlation have been re-tuned. Statistical error analysis has been utilized to evaluate the adequacy and accuracy of the novel models and empirical correlation. It was found that the predictions of the presented intelligent models and the empirical correlation are in excellent agreement with reported data in the literature with the absolute average relative deviations percent (%AARDs) of generally less than 0.9 % and R^2 of generally greater than 0.999. However, using the LSSVM model contributes to obtain slightly better results.

Keywords: Carbon dioxide; water content; predictive model; least squares support vector machine; artificial neural network; correlation.

1. Introduction

There has been a serious concern about global warming caused by the emission of carbon dioxide (CO₂) from various sources. Current focus across the world is centered on carbon capture and storage (CCS) technology, and trial operations are carried out in several places. Accurate estimation of carbon dioxide water content is a critical problem in the emerging technologies for disposing of carbon dioxide by carbon sequestration technologies. In the transmission of carbon dioxide further condensation of water is problematic. Such engineering applications as enhanced oil recovery, CO₂ storage in deep saline aquifers, and natural gas sweetening represent the importance of knowing the precise values of CO₂ water content [1]. From economic point of view, this phenomenon, i.e. water condensation, is capable of increasing the pressure drop in the line and often leads to corrosion problems.

Different properties of the CO₂+ water system have been measured experimentally by some authors [2-4]. Valuable studies on CO₂+ water system have been performed by Wiebe and Gaddy [5-7]. Included in their works the water content of CO₂-rich phase has been measured for pressures and temperatures up to 71 MPa, and 100 °C, respectively. In other works, the

data of phase equilibrium for CO₂+ water system have been reported by Gillespie *et al.* [8], Coan and King [9], King *et al.* [10], Müller *et al.* [11], Zawisza and Malesinska [12], and Song and Kobayashi [13]. The CO₂ water content data have been measured by Gillespie *et al.* [8] for pressures from 0.69 to about 2.07 MPa at two temperature levels, 76° and 93.3°C. The data reported by Coan and King [9] covers CO₂ water content for pressures less than 5.17 MPa and temperatures up to 100 °C. The VLE next to the critical region and at a pressure equal to 20.7 MPa have been measured by King *et al.*[10]. Müller *et al.* [11] measured CO₂+water VLE for pressures less than 2.35 MPa and temperatures between 100 and 200°C. Zawisza and Malesinska [12] measured two aqueous dew points at 100°C. For temperatures more than 104°C, Song and Kobayashi [13] reported some data for the CO₂ water content.

The phase behavior of the carbon dioxide + water system is generally complex as that of the hydrogen sulfide + water system [14]. However, for the range of low temperature a CO₂-rich liquid phase is not faced; it only happens for temperatures less than about 32°C. On the other hand, the water content of CO₂ does exhibit a minimum. One of the most complex substances for water content calculation is carbon dioxide. Unlike methane, CO₂ demonstrates a minimum in the water content. As will be illustrated, these minima cannot be predicted by the relatively simple methods.

The main objective of this work is to utilize two Machine Learning approaches namely MLP-ANN and LS-SVM for accurate prediction of CO₂ water content. To this end, a total number of 461 data points reported by Wiebe, and Gaddy [5] have been gathered. The collected database covers wide range of temperatures and pressures. In this study, the coefficients of the previously published empirical correlation by Bahadori [15] are also retuned to obtain better results. By employing some statistical parameters the capability of both MLP-ANN and LS-SVM models in estimating the water content of CO₂ has been evaluated. The predictions of the novel intelligent models are also compared to the estimations of the modified Bahadori [15] mathematical expression for CO₂ water content calculation.

2. Support Vector Machine (SVM)

The SVM is a new and supervised machine learning technique based on the statistical learning theory [16-17]. It has been studied extensively for both classification and regression analysis [18-26]. The SVM algorithm constructs a separating hyper-surface in the input space. This process is carried out as follows [17,22,27-29]:

- 1) It maps the input patterns into a higher dimensional feature space by way of nonlinear mapping.
- 2) Create a separating hyper-plane with maximum margin.

Regard a given training samples $((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n))$ with input data $x_i \in R^n$ and output data $y_i \in R$ with class labels -1 , 1 for classes 1 and 2 respectively. If these data sample is linearly separable in the feature space, then the SVM approximate the function as [17,29]:

$$y = w^T \Phi(x) + b \tag{1}$$

where $\Phi(x)$ denotes the nonlinear function that maps x from low dimensional space into n -dimensional feature space; w and b are weight vector and bias terms. When the data of the two classes are separable, one can say [17,29]:

$$\begin{cases} w^T \Phi(x_k) + b \geq +1, & \text{if } y_k = +1 \\ w^T \Phi(x_k) + b \leq -1, & \text{if } y_k = -1 \end{cases} \tag{2}$$

Which is equivalent to [17,29]:

$$y_k [w^T \Phi(x_k) + b] \geq +1, \quad k = 1, 2, \dots, N \tag{3}$$

The expansion of linear SVMs to non-separable case was also made by Cortes and Vapnik in 1995 [27]. Fundamentally, it is performed by presenting additional slack variables into Eq. (3) as follows [17,29]:

$$y_k[w^T\Phi(x_k)+b]\geq 1-\zeta_k, \quad k=1,2,\dots,N \quad (4)$$

$$\zeta_k \geq 0 \quad \text{for } k=1,\dots,N$$

The generalized optimal separating hyper-plane is determined by the vector w that minimizes the functional [17,29]:

$$\Phi(w,\zeta) = \frac{1}{2}w^T w + \frac{C}{2} \sum_{i=1}^N \zeta_i^p \quad (5)$$

subject to the following limitations:

$$y_k[w^T\Phi(x_k)+b]\geq 1-\zeta_k, \quad k=1,2,\dots,N \quad (6)$$

where C is a positive real constant that determines the tradeoff between the maximum margin and the minimum classification error [17,21,29]. In the conventional SVM, optimal separating hyper plane is achieved by solving the above quadratic programming problem.

The solution to the optimization problem of Eq. (5) under the constraints of Eq.(6) is given by the saddle point of the Lagrangian [30],

$$\Phi(w,b,\alpha,\zeta,\beta) = \frac{1}{2}w^T w + \frac{C}{2} \sum_{i=1}^N \zeta_i - \sum_{i=1}^N \alpha_i (y_i [w^T x_i + b] - 1 + \zeta_i) - \sum_{j=1}^N \beta_j \zeta_j \quad (7)$$

where α, β are the Lagrange multipliers [17,29].

Least square SVM (LS-SVM) is the improved version of SVM which has been recently developed by Suykens and Vandewalle [17] for decreasing the SVM model complication and its development. In LS-SVM algorithm solution is obtained by solving a linear set of equations, instead of solving a quadratic programming problem involved by standard SVM [17,19-20,29].

In comparison to SVM, the LS-SVM is trained by minimizing the cost function which is explained as follow [17,29]:

$$\Phi(w,\zeta) = \frac{1}{2}w^T w + \frac{C}{2} \sum_{i=1}^N \zeta_i^2 \quad (8)$$

subject to the equality constraints:

$$y_i[w^T\Phi(x_i)+b]=1-\zeta_i, \quad i=1,2,\dots,N \quad (9)$$

In the LS-SVM, one works with equality instead of inequality constraints. Therefore, the optimal solution can be obtained by solving a set of linear equations instead of solving a quadratic programming problem [17,29]. To derive the dual problem for LS-SVM non-linear classification problem, the Lagrange function is defined as [17,29]:

$$L(w,b,\zeta;\alpha) = \frac{1}{2}w^T w + \frac{C}{2} \sum_{i=1}^N \zeta_i^2 - \sum_{i=1}^N \alpha_i \{(y_i [w^T \Phi(x_i) + b] - 1 + \zeta_i)\} \quad (10)$$

where α_i values are Lagrange multipliers, which is positive or negative due to LS-SVM formulation. The conditions for optimality of function yield [19]:

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^N \alpha_i y_i \Phi(x_i) \\ \frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^N \alpha_i y_i = 0 \\ \frac{\partial L}{\partial \zeta_i} = 0 \Rightarrow \alpha_i = \gamma \zeta_i, \quad i = 1, 2, \dots, N \\ \frac{\partial L}{\partial \alpha_k} = 0 \Rightarrow y_i [w^T \Phi(x_i) + b] = 1 - \zeta \end{cases} \quad (11)$$

By defining $Z^T = [\Phi(x_1)^T y_1; \dots; \Phi(x_N)^T y_N]$, $Y = [y_1; \dots; y_N]$, $1_v = [1; \dots; 1]$, $\zeta = [\zeta_1; \dots; \zeta_N]$, $\alpha = [\alpha_1; \dots; \alpha_N]$ and eliminating w and ζ , then the optimization problem is transformed into the following form [17,29]:

$$\begin{bmatrix} 0 & 1_N^T \\ 1_N & \Omega + \gamma^{-1} I_N \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \quad (12)$$

where I_N is an $N \times N$ identity matrix, and $\Omega \in R^{N \times N}$ is the kernel matrix defined by $\Omega_{ij} = \Phi(x_i) \Phi(x_j) = K(x_i, x_j)$ (13)

There are many kernel function including linear, polynomial, spline, radial basis function (RBF), polynomial, sigmoid and so on for LS-SVM [31-32]. However, most widely used kernel functions are RBF (Eq. (14)) and polynomial (Eq. (15)).

$$K(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / \sigma^2) \quad (14)$$

$$K(x_i, x_j) = (1 + x_i^T x_j / c)^\gamma \quad (15)$$

where σ^2 is the squared variance of the Gaussian function and γ is the polynomial degree, which should be optimized by the user, to obtain the support vector.

This work utilized the RBF [18-20,24,33] kernel function, a common function that is useful in LS-SVM modeling studies. Coupled Simulating Annealing (CSA) [34] was Also employed for optimized hyper-parameters of LS-SVM model. In this work, we have used the LS-SVM algorithm developed by Packman *et al.* [76] and Suykens and Vandewalle [17].

3. Artificial Neural Network (ANN)

As machine learning approach, ANNs are widely employed for pattern identification, classification, and also prediction [35-37]. ANNs are developed on the foundation of nervous system of the human brain [38]. Detailed information about development and history of ANNs has been demonstrated elsewhere [39-72]. By applying a number of input-output training arrangements from given data sets, ANNs find mathematical relationships (linear or nonlinear) [43-44]. The basic construction blocks of an ANN are known as neurons or processing units. Figure 1 graphically illustrates an artificial neuron [45-46]. The neuron m shown in Figure 1 could be mathematically represented as follows [45-46]:

$$r_m = \sum_{i=1}^n (w_{mi} x_i + b_m) \quad (16)$$

$$y_m = F(r_m) \tag{17}$$

where x_1, x_2, \dots, x_n are the input signals; $w_{m1}, w_{m2}, \dots, w_{mn}$ are the synaptic weights; r_m is the linear combiner output; b_m is the bias term; f is activation function; and y_m is output signal of the neuron.

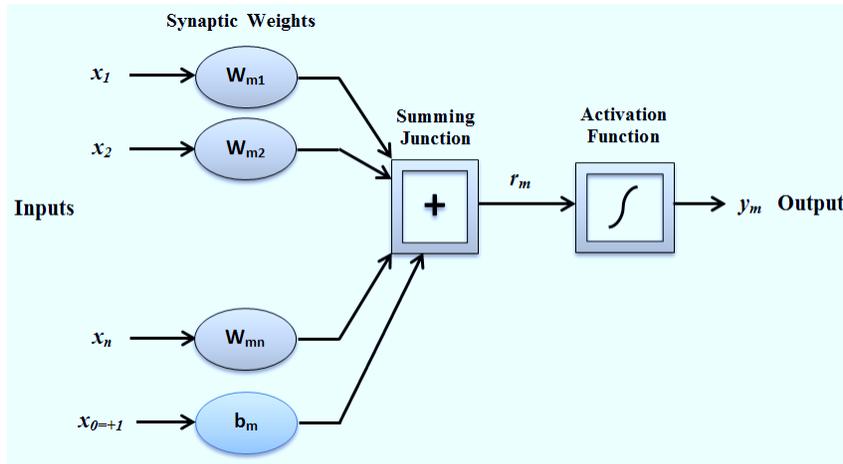


Figure 1. Typical model of an artificial neuron [45-46]

In the recurrent neural networks (RNNs), the connections between processing units build a directed cycle. Unlike RNNs, in feed-forward neural networks (FFNNs) information progresses in just one direction. It is believed that FFNNs are effective and reliable non-linear classifier identifiers [37]. Various kinds of FFNNs such as MLP networks, radial basis function networks (RBFNs), and functional link networks (FLNs) are available [37]. One or more layers namely hidden layer(s) are existed between input and output layers of a MLP network. Figure 2 shows a typical MLP with I input branching nodes, H hidden neurons, and O output neurons [45]. The power of MLP networks results from their capability to exert nonlinear functions. Various activation functions such as tan-sigmoid, log-sigmoid or threshold transfer functions could be used to introduce nonlinearity into the MLP networks.

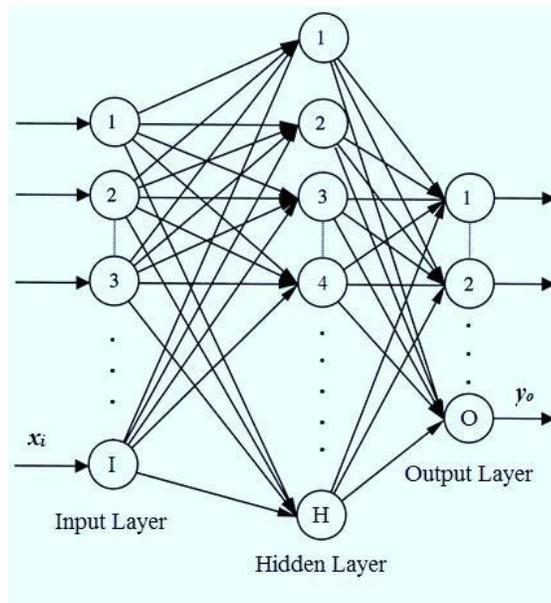


Figure 2. A typical 3-layer MLP [45-46]

With the aim of tuning the values of weights in MLP network, back-propagation (BP) learning algorithm is employed. Consequently, the differences between the network predictions and reported values will be minimized. Among available training algorithm, the well-known Levenberg–Marquardt (LM) method [47-48] is used in this work to train BP.

4. Development of intelligent models

This section presents the procedures of developing LS-SVM model and MLP-ANN model for predicting CO₂ water content. First, all the data were normalized between -1 and +1 to make uniform the domain of variables. Later, these values were altered to their original values. The next step in data pre-processing was assumed to be division of the collected database into three data sets including the "Training" set, the "Validation" set and the "Test" set. As it demonstrated elsewhere [33-35], the training set is utilized to generate the model structure, the validation set is used for optimization of the model and the test set is used to investigate the prediction capability and validity of the proposed model. The division of database into three subdata sets is performed randomly. For this end, 70%, 15% and 15% of the main data set were accidentally selected for the "Training" set, the "Optimization" set, and the "Test" set, respectively [35]. After doing these steps, LS-SVM algorithm has been used for predicting CO₂ water content at both low and high pressures. In the case of MLP-ANN, the final step is defined as sizing the network structure. A well-known three-layer MLP network has been utilized for CO₂ water content estimation [49-51]. The input layer consists of two nodes (temperature and pressure) and the output layer includes one neuron (CO₂ water content). In this work, the mean squared error (MSE) and regression R-value are employed as criteria for evaluating the performance of the MLP networks having different number of hidden neurons. The following equation represents the MSE:

$$MSE = \frac{1}{n} \sum_{i=1}^n (t_i - o_i)^2 \quad (18)$$

A linear transfer function has been used for the output layer. The hidden layer uses a log-sigmoid transfer function, as defined below, producing outputs between 0 and 1.

$$f(x) = \frac{1}{1 + e^{-s}} \quad (19)$$

5. Error analysis

5.1. Applied statistical parameters

To validate the preciseness and performance of the proposed models some statistical parameters including average percent relative error (APRE), average absolute percent relative error (AAPRE), root mean square error (RMSE) and R-squared (R^2) have been utilized. Definitions and equations of those parameters are given below:

A. APRE. It measures the relative deviation from the experimental data, defined by:

$$\%APRE = \frac{100}{N} \sum_i^N \frac{(pred.(i) - exp.(i))}{exp.(i)} \quad (20)$$

B. AAPRE. It measures the relative absolute deviation from the experimental data, defined as below:

$$\%AAPRE = \frac{100}{N} \sum_i^N \left| \frac{(pred.(i) - exp.(i))}{exp.(i)} \right| \quad (21)$$

C. RMSE. It measures the data dispersion around zero deviation, defined by:

$$RMSE = \sqrt{\frac{\sum_i^N (pred.(i) - exp.(i))^2}{N}} \quad (22)$$

D. R^2 . It is a simple statistical parameter exhibits how good model matches the data. In fact, the closer the value of R^2 to 1, the better the model fits the data. It is defined as:

$$R^2 = 1 - \frac{\sum_i^N (pred.(i) - exp.(i))^2}{\sum_i^N (pred.(i) - average(exp.(i)))^2} \quad (23)$$

5.2. Basis of comparison

Bahadori [15] presented a simple-to-use Arrhenius-type asymptotic exponential function coupled with the Vandermonde matrix for estimation of CO₂ water content as a function of pressure and temperature. The final correlation is as follows:

$$\ln(W) = a + \frac{b}{P} + \frac{c}{P^2} + \frac{d}{P^3} \quad (24) \text{ where:}$$

$$a = A_1 + \frac{B_1}{T} + \frac{C_1}{T^2} + \frac{D_1}{T^3} \quad (25)$$

$$b = A_2 + \frac{B_2}{T} + \frac{C_2}{T^2} + \frac{D_2}{T^3} \quad (26)$$

$$c = A_3 + \frac{B_3}{T} + \frac{C_3}{T^2} + \frac{D_3}{T^3} \quad (27)$$

$$d = A_4 + \frac{B_4}{T} + \frac{C_4}{T^2} + \frac{D_4}{T^3} \quad (28)$$

in which A_i - D_i are coefficients.

In this study, the coefficients of Bahadori [15] correlation have been re-tuned to obtain better results. The obtained values for Bahadori [15] correlation are given in Table 1.

Table 1. Tuned coefficients used in equation 24-28 to correlate carbon dioxide-rich phase water content as a function of temperature

Coefficient	Value	Coefficient	Value
A_1	4.691678×10^2	A_3	4.825648×10^5
B_1	-4.437337×10^5	B_3	-4.610234×10^8
C_1	1.410149×10^8	C_3	1.464722×10^{11}
D_1	-1.500053×10^{10}	D_3	-1.547741×10^{13}
A_2	-3.200609×10^4	A_4	-1.790510×10^6
B_2	3.064587×10^7	B_4	1.708401×10^9
C_2	-9.764819×10^9	C_4	-5.419500×10^{11}
D_2	1.035334×10^{12}	D_4	5.716500×10^{13}

6. Results and discussion

6.1. Low pressure system

The proposed neural-based and LS-SVM models have been developed for temperatures between 273.15 and 348.15 K and pressures up to 6 MPa. The optimum values of γ and σ^2 as parameters of LS-SVM model are equal to 1.386e+05 and 0.372, respectively. Figures 3 and 4 demonstrate the predicted CO₂ water content using the presented LS-SVM model for low pressure system in comparison with reported experimental data. Relative errors between the proposed LS-SVM model for low pressure system and reported data in the literature are depicted in Figure 5, which is less than 4%.

In the case of MLP-ANN model, Figures 6 and 7 illustrate the effect of number of hidden neuron on performance of ANN model for the application of interest. As can be seen from Figures 6 and 7, employing more than 7 neurons in the hidden layer of ANN give no visible improvement in MSE and R-value. However, 2-12-1 topology will lead to obtain slightly better results. Cross plot for the presented MLP-ANN with 12 hidden neurons for low pressure system is shown in Figure 8. Figure 9 shows distribution of errors between the neural-based model estimations and reported data for low pressure system.

6.2. High pressure system

Proposed MLP-ANN and LS-SVM models for high pressure system are applicable for temperatures between 273.15 and 348.15 K as well as pressures between 6 and 21 MPa. By employing CSA optimization technique, 8.639e+04 and 0.490 were obtained for γ and σ^2 as parameters of LS-SVM model for high pressure system, respectively. The estimation capability of the presented LS-SVM model for high pressure system is graphically examined in Figures 10 to 12. As it is shown in Figures 10, the relative error is less than 0.05.

Figures 13 and 14 illustrate the MSE and R-value of the constructed MLP-ANN with 1 to 12 hidden neurons. As it shown in Figures 13 and 14, using hidden neurons equal to/ more than 8 for high pressure system give acceptable results. However, 2-12-1 topology gives slightly better predictions. Graphical representations of the accuracy of the presented MLP-ANN with 12 hidden neurons for high pressure system are provided in Figures 15 and 16.

6.3. Comparing the presented models

The statistical parameters including average relative error (ARE), squared correlation coefficient (R^2), average absolute relative error (AARE), and root mean square error (RMSE) for modified Bahadori correlation [15], presented MLP-ANN model with 12 hidden neurons, and proposed LS-SVM model for low pressure and high pressure systems have been tabulated in Table 2. Table 3 indicates the statistical parameters of the LS-SVM model for prediction of water content of carbon dioxide at low pressure. According to this table, the LS-SVM model gives an AARE of 0.499%, ARE of 0.05%, R^2 of 1, and RMSE of 0.018. Moreover, the main statistical parameters of the developed LS-SVM models at high pressure are reported in Table 4. In accord with this table, the improved model gives an AARE of 0.744%, ARE of -0.07%, R^2 of 0.999, and RMSE of 0.021. It is concluded that an excellent agreement between the prediction of developed model and the experimental data values.

Table 2. Statistical parameters of the developed LS-SVM and MLP-ANN models along with modified Bahadori [15] correlation to determine water content of carbon dioxide at low and high pressures

System	Models	APRE (%)	AAPRE (%)	R^2	RMSE
Low pressure	Modified Bahadori [15]	2.54	14.34	0.994	1.232
	MLP-ANN	-0.037	0.971	1	23.41
	LS-SVM	0.050	0.499	1	0.018
High pressure	Modified Bahadori [15]	3.373	20.33	0.920	0.550
	MLP-ANN	-0.093	0.908	0.999	25.04
	LS-SVM	-0.070	0.744	0.999	0.021

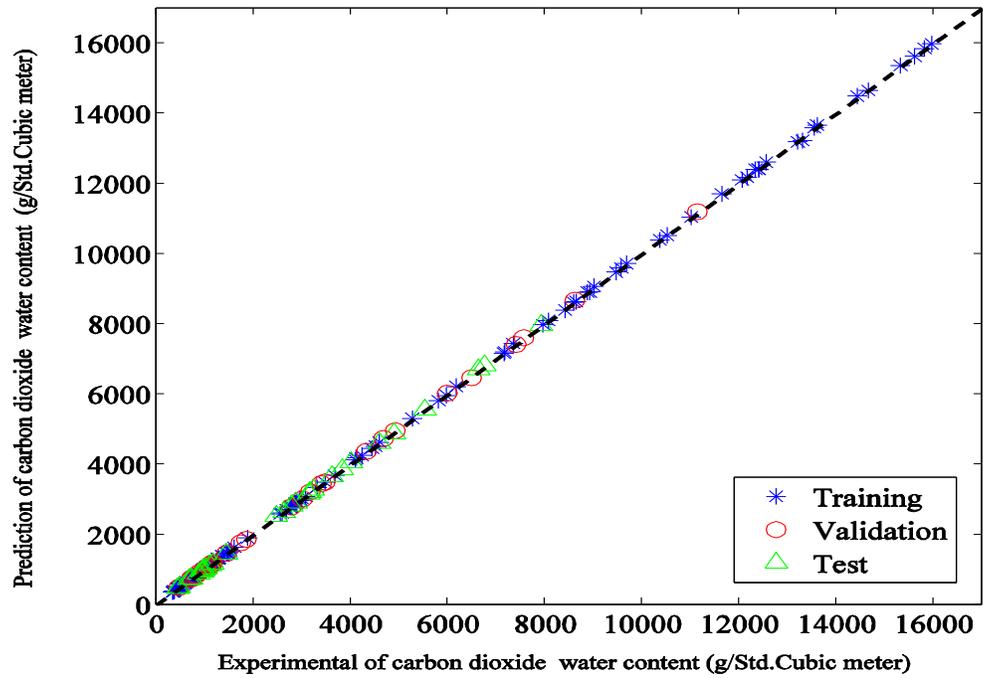


Figure 3. Predictions of the LS-SVM model for low pressure system versus reported data

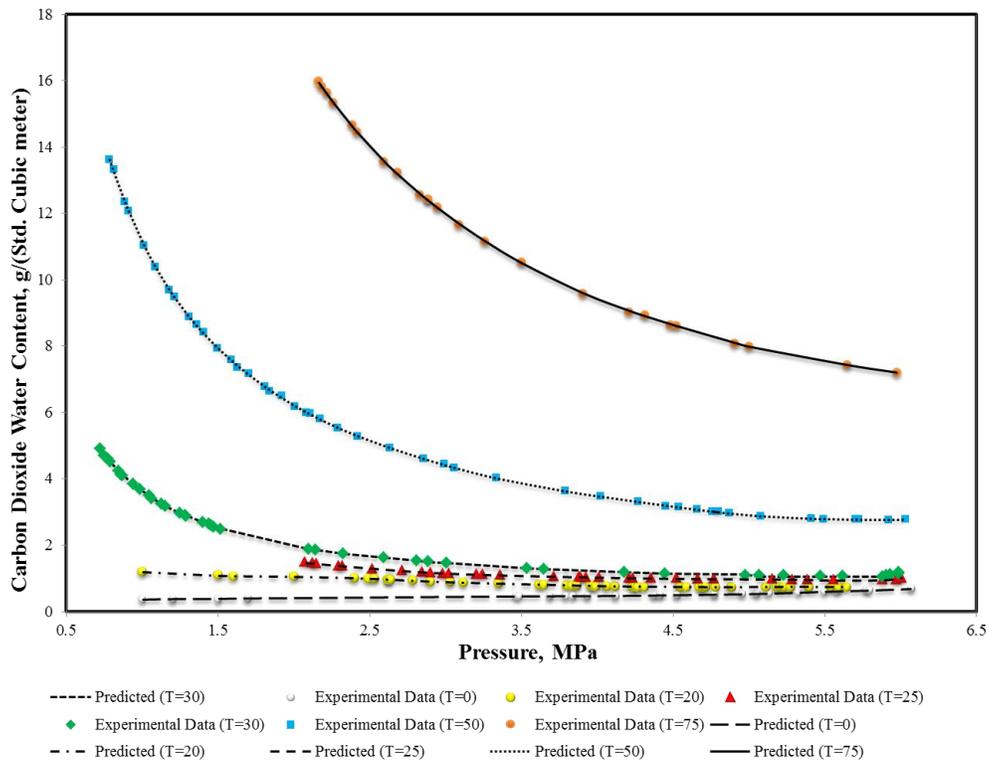


Figure 4. Predictions of the proposed LS-SVM model against the reported data for low pressure system

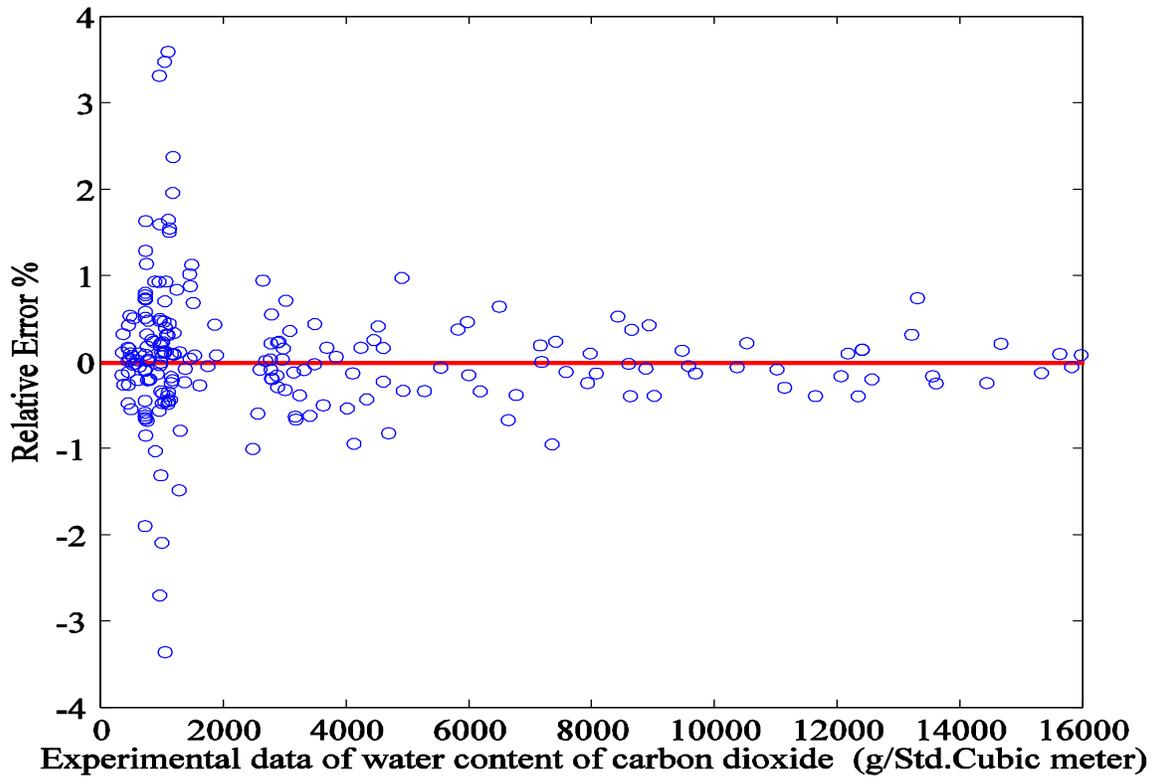


Figure 5. Deviations of predictions of the LS-SVM model for low pressure system from reported

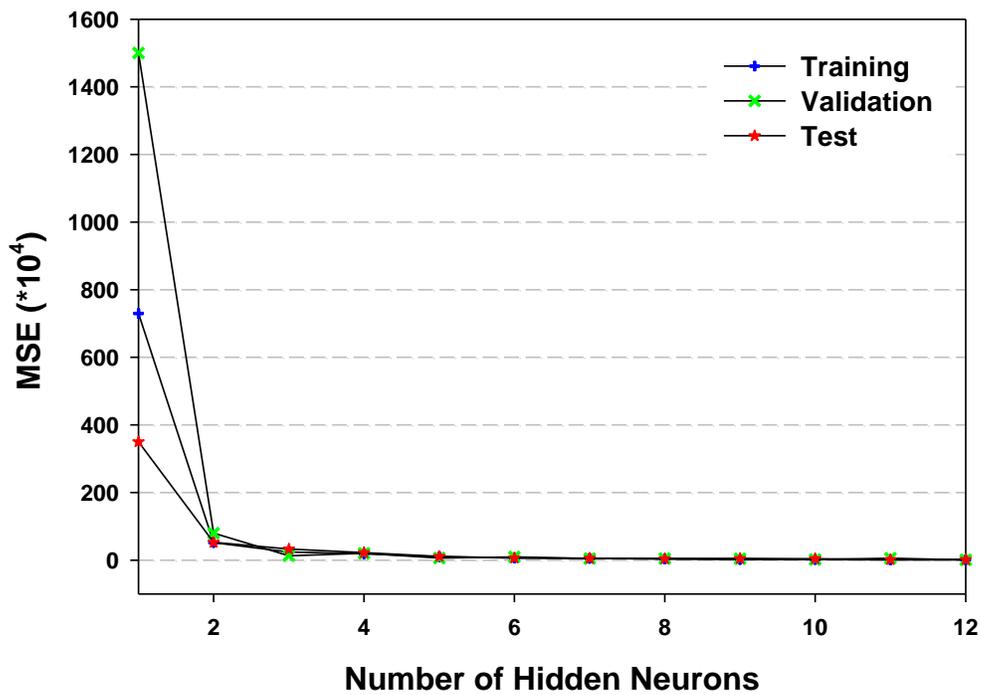


Figure 6. Estimation capability of various ANN structures for low pressure system

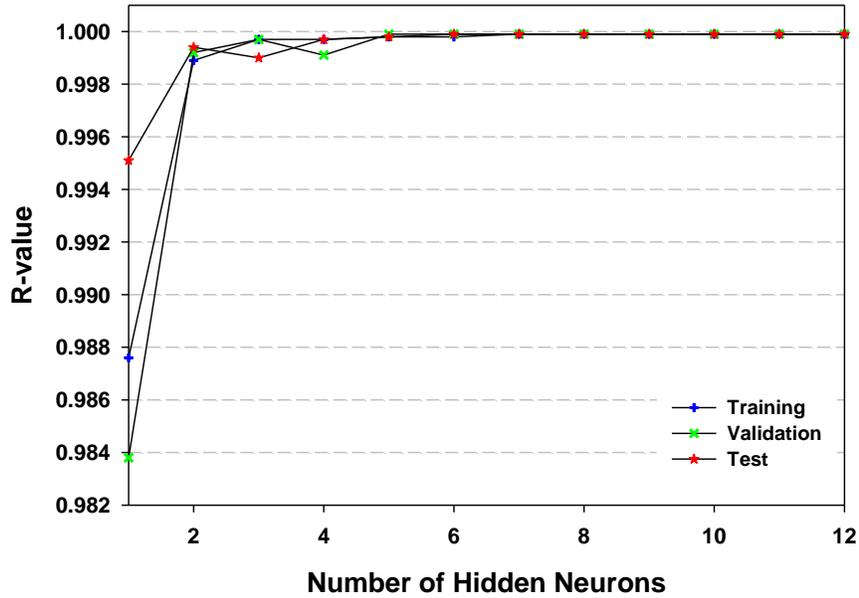


Figure 7. Correlation analysis between predictions of various ANN structure for low pressure system and reported data

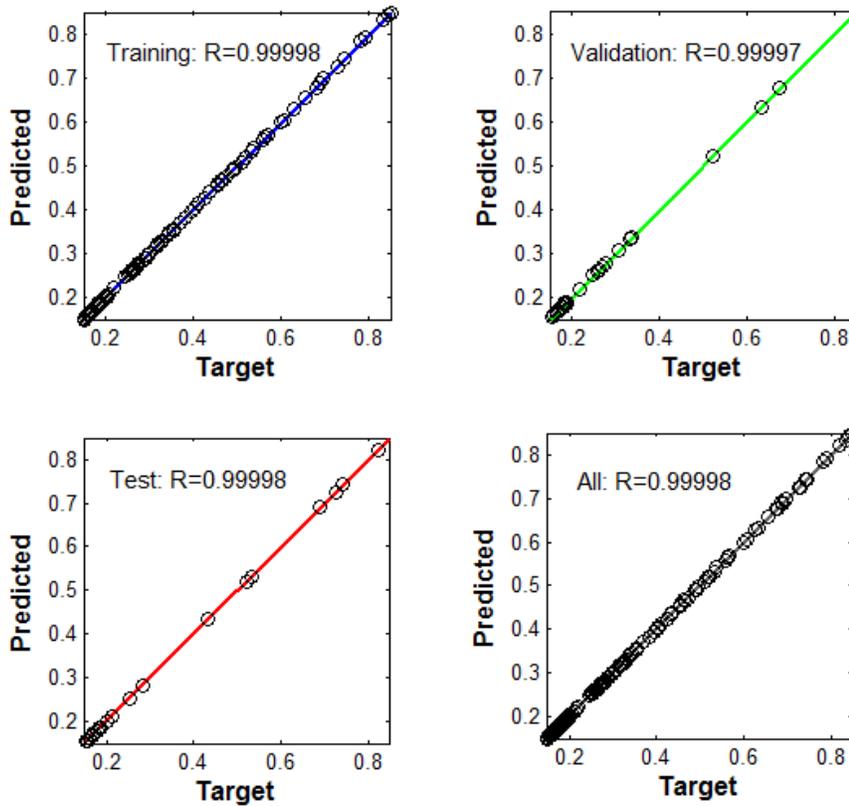


Figure 8. Predictions of the presented ANN model with 12 hidden neurons for low pressure system versus reported data

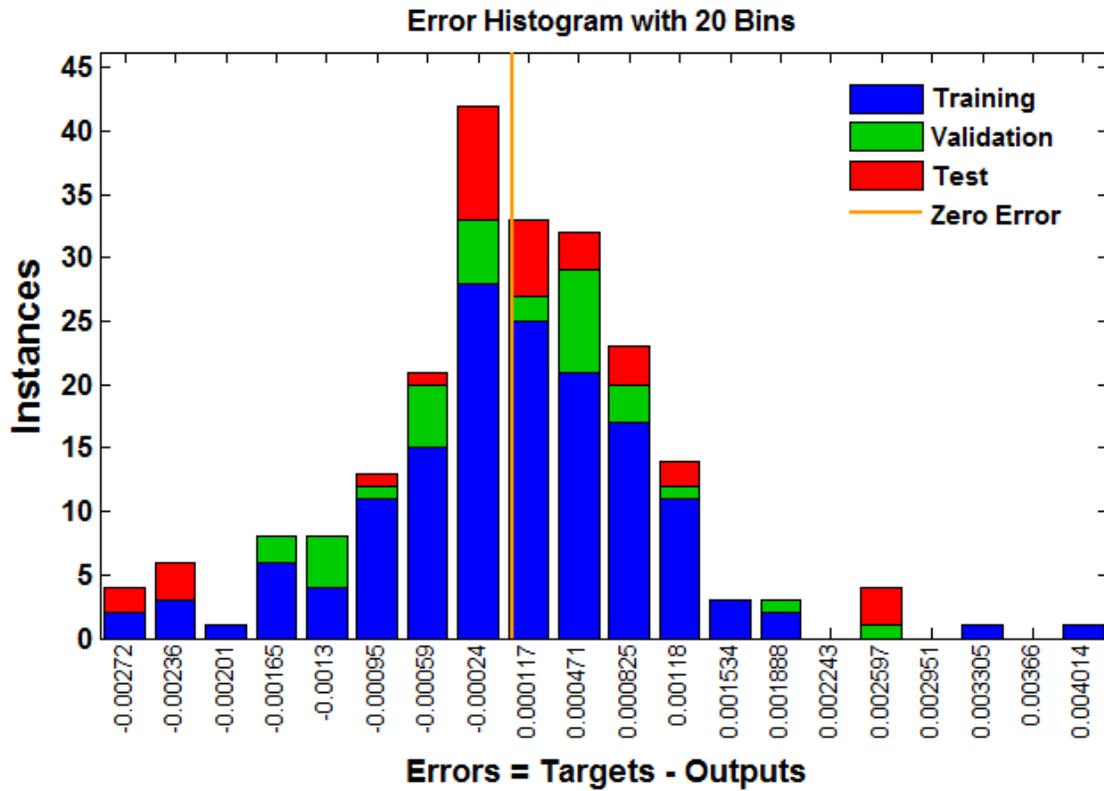


Figure 9. Histogram of error values of presented ANN model with 12 hidden neurons for low pressure system

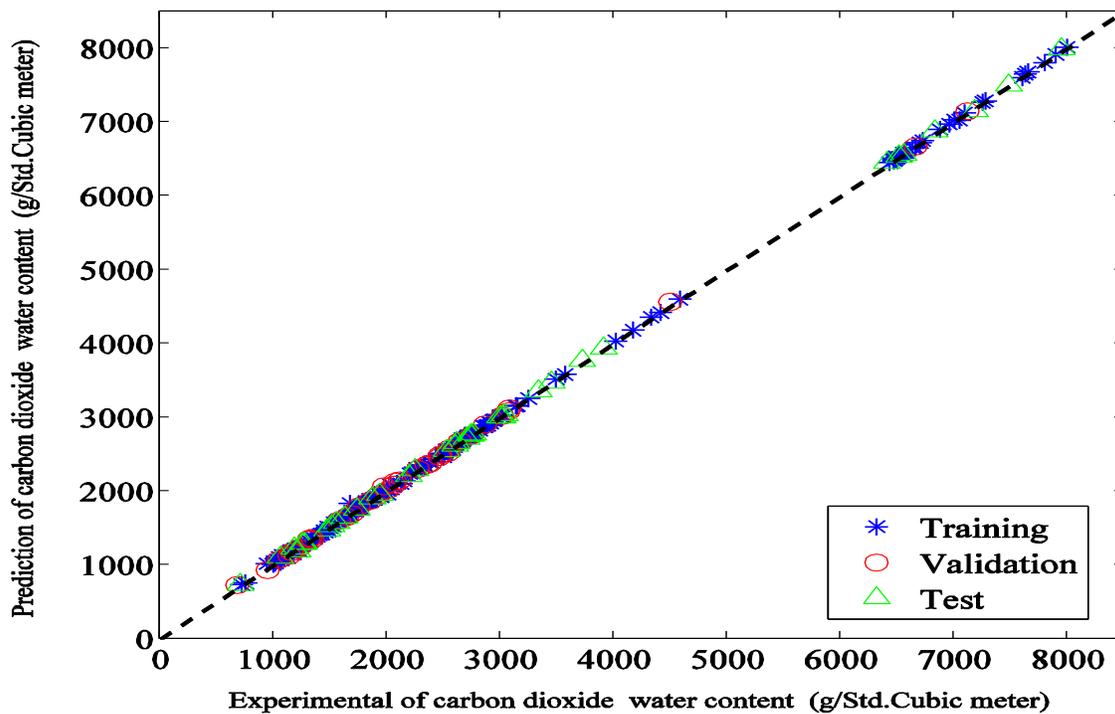


Figure 10. Predictions of the LS-SVM model for high pressure system versus reported data

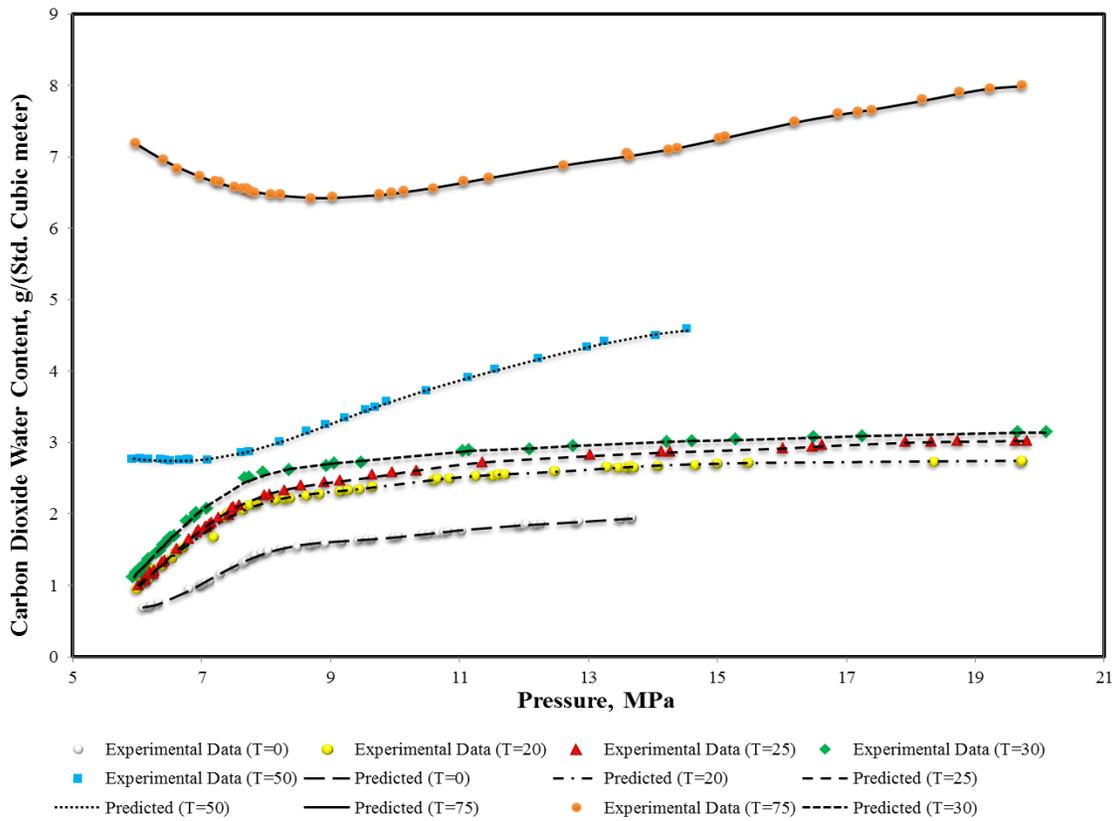


Figure 11. Predictions of the proposed LS-SVM model against the reported data for high pressure system

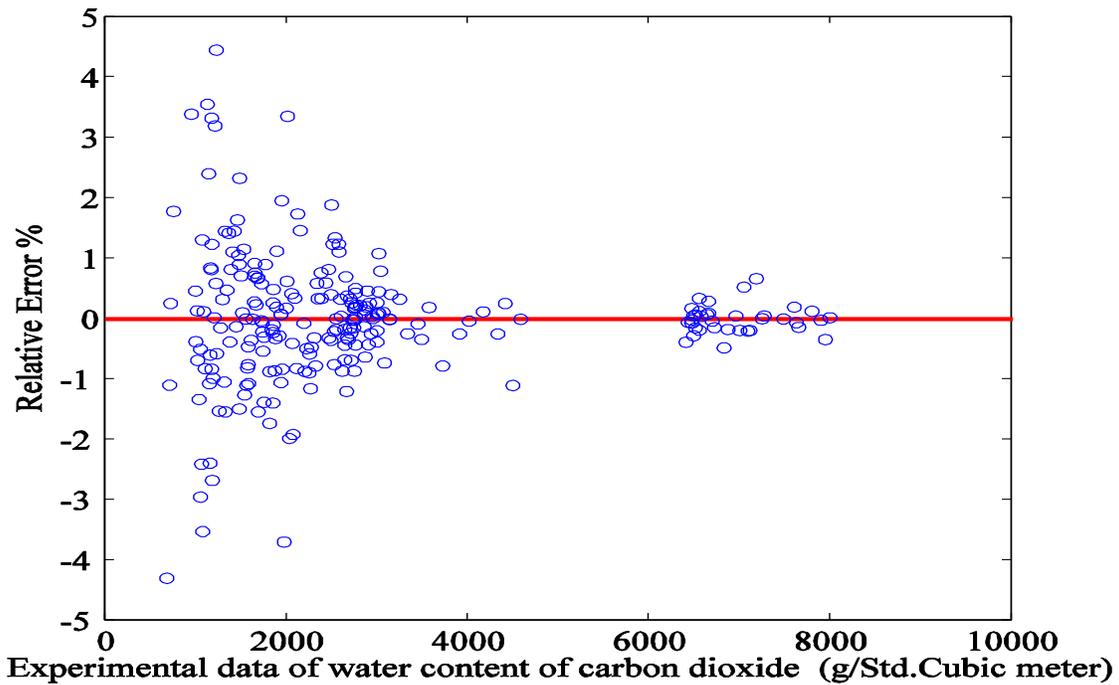


Figure 12. Deviations of predictions of the LS-SVM model for low pressure system from reported data

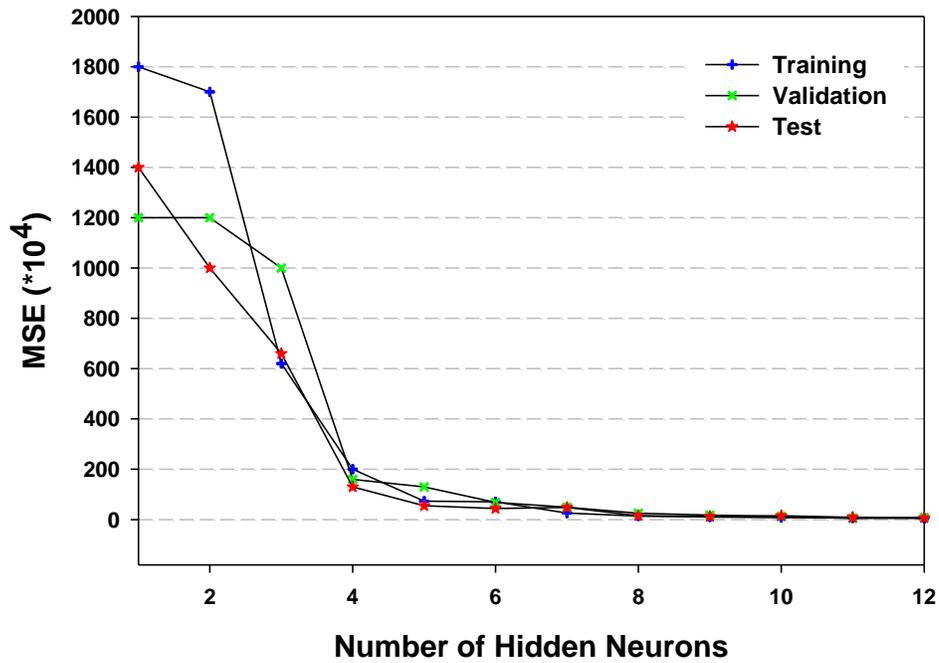


Figure 13. Estimation capability of various ANN structures for high pressure system

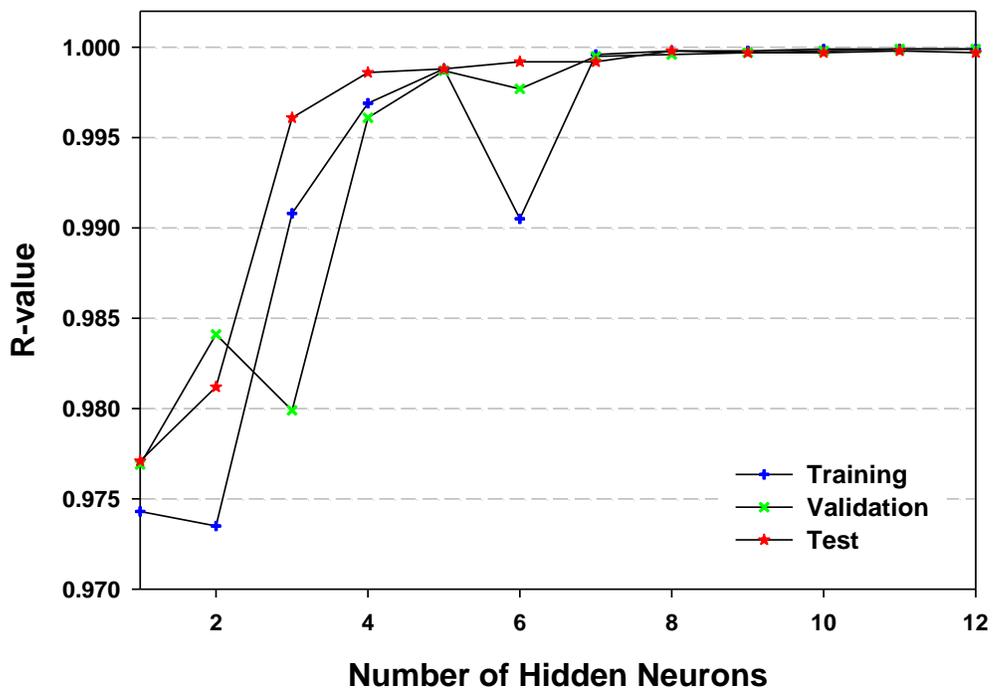


Figure 14. Correlation analysis between predictions of various ANN structure for high pressure system and reported data

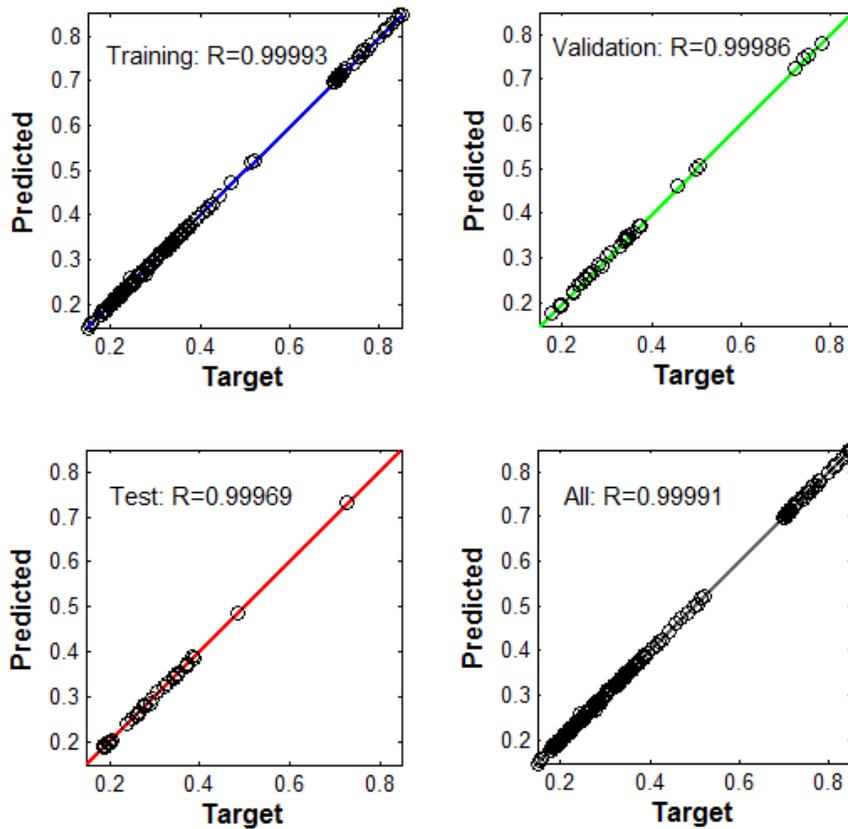


Figure 15. Predictions of the presented ANN model with 12 hidden neurons for high pressure system versus reported data

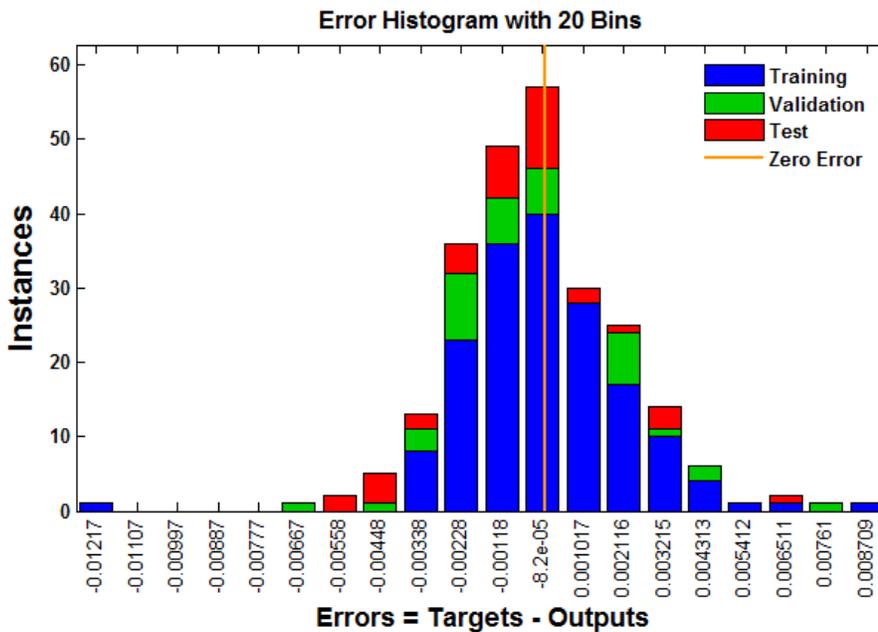


Figure 16. Histogram of error values of presented ANN model with 12 hidden neurons for high pressure system

Table 3. Statistical parameters of the developed LS-SVM model to determine water content of carbon dioxide at low pressure

Statistical parameter	Value	Statistical parameter	Value
Training set		Test set	
R^2	1	R^2	1
Average absolute percent relative error	0.394	Average absolute percent relative error	0.857
Average percent relative error	-0.005	Average percent relative error	0.137
Root mean square error	0.017	Root mean square error	0.019
n	152	n	32
Validation set		Total	
R^2	1	R^2	1
Average absolute percent relative error	0.636	Average absolute percent relative error	0.499
Average percent relative error	0.220	Average percent relative error	0.05
Root mean square error	0.018	Root mean square error	0.018
n	33	n	217

Table 4. Statistical parameters of the developed LS-SVM model to determine water content of carbon dioxide at high pressure

Statistical parameter	Value	Statistical parameter	Value
Training set		Test set	
R^2	0.999	R^2	0.999
Average absolute percent relative error	0.730	Average absolute percent relative error	0.606
Average percent relative error	-0.016	Average percent relative error	-0.176
Root mean square error	0.021	Root mean square error	0.019
n	171	n	36
Validation set		Total	
R^2	0.999	R^2	0.999
Average absolute percent relative error	0.945	Average absolute percent relative error	0.744
Average percent relative error	-0.227	Average percent relative error	-0.07
Root mean square error	0.022	Root mean square error	0.021
n	37	n	244

7. Conclusions

In this article, the capability of intelligent approaches including LS-SVM and a three-layer MLP network in modeling/predicting the water content of CO₂ have been evaluated. Furthermore, the constants of the correlation presented by Bahadori *et al.* [15] for estimation of CO₂ water content have been re-tuned to obtain better results. Construction of the LS-SVM and MLP-ANN models and modification of Bahadori *et al.* [15] correlation are based on reported data in the open literature for water+CO₂ system [5]. According to the error analysis results, it can be concluded that the proposed LS-SVM and MLP-ANN models and also modified Bahadori *et al.* [15] correlation give acceptable predictions. However, the developed LS-SVM and MLP-ANN models are more accurate than modified Bahadori *et al.* [15] correlation. It was illustrated that the built LS-SVM and MLP-ANN models are capable of simulating the actual physical trend of the CO₂ water content with variation of temperature and pressure. Nevertheless, implementation of LS-SVM algorithm for predicting CO₂ water content will contribute to obtain the best results. Two of the most important advantages of the LS-SVM model are that it uses only two simple parameters (γ and σ^2) and does not require high theoretical knowledge or

human experience during the training process. Thus, prior knowledge has not been utilized and the model training approach is based on the reported data only.

List of symbols

$K(x_i, x_j)$	kernel function
σ	width of kernel function
L	Lagrangian
α_i	Lagrange multipliers
Z_i	mole fraction of stream gas
I_N	$N \times N$ identity matrix
ζ	slack variable
w	weight vector
b	bias term
γ	Regularization constant
1_v	[1;1,...,1]
A^T	transpose of matrix A
x	input vector of network
y	output vector of network
Ω	Kernel matrix
Φ	map from input space into feature space
d	the polynomial degree
$C_{salt(s)}$	concentration of salt
P	Pressure, MPa
T	Temperature, K
W	water content of carbon dioxide-rich phase, $g/m^3(Std.)$
i	index
j	index
o	predicted value
r_m	linear combiner output
t	target value
w_{mn}	synaptic weight
x_n	input signal of neuron
y_m	neuron's output
n	number of data points

References

- [1] Salari H, Hassanzadeh H, Gerami S, Abedi, On estimating the water content of CO₂ in equilibrium with formation brine, Canadian International Petroleum Conference, Calgary, Alberta, Canada (2008).
- [2] Koolivand- Salooki M, Azarmehr A, Bahadori A, Kord Sh, Sharifi F, Alfkhan Sh. Prediction of Carbon Dioxide Water Content Using a Particle Swarm Optimization Artificial Neural Network Hybrid Model, *Petroleum & Coal* 57(5) 573-586, 2015.
- [3] Duan Z, Sun R. An improved model calculating CO₂ solubility in pure water and aqueous NaCl solutions from 273 to 533 K and from 0 to 2000 bar, *Chemical Geology*, 193 (2003) 257-271.
- [4] Diamond LW, Akinfiyev NN. Solubility of CO₂ in water from – 1.5 to 100° C and from 0.1 to 100 MPa: evaluation of literature data and thermodynamic modelling, *Fluid phase equilibria*, 208 (2003) 265-290.

- [5] Wiebe R, Gaddy V. Vapor phase composition of carbon dioxide-water mixtures at various temperatures and at pressures to 700 atmospheres, *Journal of the American Chemical Society*, 63 (1941) 475-477.
- [6] Wiebe R, Gaddy V. The solubility of carbon dioxide in water at various temperatures from 12 to 40 and at pressures to 500 atmospheres. critical phenomena, *Journal of the American Chemical Society*, 62 (1940) 815-817.
- [7] Wiebe R, Gaddy V. The solubility in water of carbon dioxide at 50, 75 and 100, at pressures to 700 atmospheres, *Journal of the American Chemical Society*, 61 (1939) 315-318.
- [8] Gillespie ., Owens ., Wilson G. Sour water equilibria extended to high temperatures and with inerts present, in: *AIChE Winter National Meeting, Atlanta, GA, 1984*.
- [9] King AD Jr, Coan C. Solubility of water in compressed carbon dioxide, nitrous oxide, and ethane. Evidence for hydration of carbon dioxide and nitrous oxide in the gas phase, *Journal of the American Chemical Society*, 93 (1971) 1857-1862.
- [10] King M, Mubarak A, Kim J, Bott T. The mutual solubilities of water with supercritical and liquid carbon dioxides, *The Journal of Supercritical Fluids*, 5 (1992) 296-302.
- [11] Müller G, Bender E, Maurer G. Das Dampf-Flüssigkeitsgleichgewicht des ternären Systems Ammoniak-Kohlendioxid-Wasser bei hohen Wassergehalten im Bereich zwischen 373 und 473 Kelvin, *Berichte der Bunsengesellschaft für Physikalische Chemie*, 92 (1988) 148-160.
- [12] Zawisza ., Malesinska B, Solubility of carbon dioxide in liquid water and of water in gaseous carbon dioxide in the range 0.2-5 MPa and at temperatures up to 473 K, *Journal of Chemical and Engineering Data*, 26 (1981) 388-391.
- [13] Song ., Kobayashi R, Water content of CO₂ in equilibrium with liquid water and/or hydrates, *SPE Formation Evaluation*, 2 (1987) 500-508.
- [14] Carroll JJ. The Water Content of Acid Gas and Sour Gas from 100 to 220F and Pressures to 10,000 Psia, in: *Presented at the 81st Annual GPA Convention, 2002*.
- [15] Bahadori A. A simple predictive tool for monitoring steam loss in traps, *Chemical Engineering Research and Design*, 90 (2012) 1896-1900.
- [16] Suykens JA, Vandewalle J. Least squares support vector machine classifiers, *Neural Processing Letters*, 9 (1999) 293-300.
- [17] Suykens JAK, Vandewalle J. Least Squares Support Vector Machine Classifiers, *Neural Processing Letters*, 9 (1999) 293-300.
- [18] 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 Systems with Applications*, 36 (2009) 8368-8374.
- [19] Shokrollahi A, Arabloo M, Gharagheizi F, Mohammadi AH. Intelligent model for prediction of CO₂ - Reservoir oil minimum miscibility pressure, *Fuel*, (in press) (2013).
- [20] Rafiee-Taghanaki S, Arabloo M, Chamkalani A, Amani M, Zargari MH, Adelzadeh MR. Implementation of SVM framework to estimate PVT properties of reservoir oil, *Fluid Phase Equilibria*, 346 (2013) 25-32.
- [21] Übeyli ED. Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals, *Expert Systems with Applications*, 37 (2010) 233-239.
- [22] Amendolia SR, Cossu G, Ganadu ML, Golosio B, Masala GL, Mura GM. A comparative study of K-Nearest Neighbour, Support Vector Machine and Multi-Layer Perceptron for Thalassemia screening, *Chemometrics and Intelligent Laboratory Systems*, 69 (2003) 13-20.
- [23] Chen TS, Chen J, Lin YC, Tsai YC, Kao YH, Wu K. A Novel Knowledge Protection Technique Base on Support Vector Machine Model for Anti-classification, in: M.

- Zhu (Ed.) *Electrical Engineering and Control*, Springer Berlin Heidelberg, 2011, pp. 517-524.
- [24] Chamkalani A, Amani M., Kiani MA, Chamkalani R. Assessment of asphaltene deposition due to titration technique, *Fluid Phase Equilibria*, 339 (2013) 72-80.
- [25] Arabloo M, Shokrollahi A, Gharagheizi F, Mohammadi AH. Toward a predictive model for estimating dew point pressure in gas condensate systems, *Fuel Processing Technology*, 116 (2013) 317-324.
- [26] Farasat A, Shokrollahi A, Arabloo M, Gharagheizi F, Mohammadi AH. Toward an intelligent approach for determination of saturation pressure of crude oil, *Fuel Processing Technology*, 115 (2013) 201-214.
- [27] Cortes C, Vapnik V. Support-vector networks, *Mach Learn*, 20 (1995) 273-297.
- [28] Bazzani A, Bevilacqua A, Bollini D, Brancaccio R, Campanini R, Lanconelli N, Riccardi A, Romani D. An SVM classifier to separate false signals from micro-calcifications in digital mammograms, *Physics in Medicine and Biology*, 46 (2001) 1651.
- [29] Suykens JAK, Gestel TV, Brabanter JD, Moor BD, Vandewalle J. *Least Squares Support Vector Machines*, World Scientific Pub. Co., Singapor, 2002.
- [30] Minoux M. *Mathematical Programming: Theory and Algorithms*, 1986.
- [31] Gunn SR. *Support Vector Machines for Classification and Regression*, in, University of Southampton, Faculty of Engineering, Science and Mathematics School of Electronics and Computer Science, 1998.
- [32] Muller KR, Mika S, Ratsch G, Tsuda K, Scholkopf B. An introduction to kernel-based learning algorithms, *Neural Networks, IEEE Transactions on*, 12 (2001) 181-201.
- [33] Deng S, Yeh TH. Applying least squares support vector machines to the airframe wing-box structural design cost estimation, *Expert Systems with Applications*, 37 (2010) 8417-8423.
- [34] Xavier-de-Souza S, Suykens JA, Vandewalle J, Bollé D. Coupled simulated annealing, *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 40 (2010) 320-335.
- [35] Bose NK, Liang P. *Neural network fundamentals with graphs, algorithms, and applications*, McGraw-Hill, Inc., 1996.
- [36] Kasabov NK. *Foundations of neural networks, fuzzy systems and knowledge engineering*, Marcel Alencar, 1996.
- [37] Looney CG. *Pattern recognition using neural networks: theory and algorithms for engineers and scientists*, Oxford University Press, Inc., 1997.
- [38] Piazza L, Scalabrin G, Marchi P, Richon D. Enhancement of the extended corresponding states techniques for thermodynamic modelling. I. Pure fluids, *International journal of refrigeration*, 29 (2006) 1182-1194.
- [39] Hebb DO. *The organization of behavior: A neuropsychological theory*, Psychology Press, 2002.
- [40] Rochester N, Holland J, Haibt L, Duda W. Tests on a cell assembly theory of the action of the brain, using a large digital computer, *Information Theory, IRE Transactions on*, 2 (1956) 80-93.
- [41] D.E. Rumelhart, J.L. McClelland, *Parallel distributed processing: explorations in the microstructure of cognition. Volume 1. Foundations*, (1986).
- [42] Yen JWGG, Polycarpou MM. *Advances in Neural Networks- ISNN 2012*.
- [43] Bahadori A, Maddahi M and Vuthaluru HB. Simple-to-use Predictive Tool for an Accurate Estimation of the Water Content of CO₂,. *SPE EUROPEC/EAGE Annual Conference and Exhibition*, 14-17 June 2010, Barcelona, Spain
- [44] James W. 1950. *The principles of psychology*, in, New York: Dover Publications, 1890.

- [45] Ghiasi MM, Bahadori A, Zendeboudi S, Jamili A, Rezaei-Gomari S. Novel methods predict equilibrium vapor methanol content during gas hydrate inhibition, *Journal of Natural Gas Science and Engineering*, 15 (2013) 69-75.
- [46] Ghiasi MM, Bahadori A, Zendeboudi S. Estimation of the water content of natural gas dried by solid calcium chloride dehydrator units, *Fuel*, 117, Part A (2014) 33-42.
- [47] Bahadori A, Vuthaluru HB. Estimation of performance of steam turbines using a simple predictive tool, *Applied Thermal Engineering*, 30 (2010) 1832-1838.
- [48] Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters, *Journal of the Society for Industrial & Applied Mathematics*, 11 (1963) 431-441.
- [49] Mohammadi AH, Richon D. Use of artificial neural networks for estimating water content of natural gases, *Industrial & engineering chemistry research*, 46 (2007) 1431-1438.
- [50] Zendeboudi S, Ahmadi MA, Bahadori A, Shafiei A, Babadagli T. A developed smart technique to predict minimum miscible pressure-or implications, *The Canadian Journal of Chemical Engineering* 2013, 91(7), 1325-1337.
- [51] Chouai A, Laugier S, Richon D. Modeling of thermodynamic properties using neural networks: Application to refrigerants, *Fluid Phase Equilibria*, 199 (2002) 53-62.

Corresponding authors: Email: Alireza.bahadori@scu.edu.au (A. Bahadori) TEL:+61 2 6626 9347, FAX:+61 2 6626 2680