ELSEVIER

Fluid Phase Equilibria 152 (1998) 23–42

FLUID PHASE EQUILIBRIA

A new correlation for predicting hydrate formation conditions for various gas mixtures and inhibitors

Ahmed A. Elgibaly, Ali M. Elkamel

College of Engineering and Petroleum, University of Kuwait, P.O. Box 5969, Assafat 13060, Kuwait

Received 27 December 1997; accepted 22 May 1998

Abstract

In the last 50 years, several studies have been performed on the measurement and prediction of hydrate forming conditions for various gas mixtures and inhibitors. Yet, the correlations presented in the literature are not accurate enough and consider most of the time, simple pure gases only and their mixtures. In addition, some of these correlations are presented mainly in graphical form, thus making it difficult to use them within general computer packages for simulation and design. The purpose of this paper is to present a comprehensive neural network model for predicting hydrate formation conditions for various pure gases, gas mixtures, and different inhibitors. The model was trained using 2387 input–output patterns collected from different reliable sources. The predictions are compared to existing correlations and also to real experimental data. The neural network model enables the user to accurately predict hydrate formation conditions for a given gas mixture, without having to do costly experimental measurements. The relative importance of the temperature and the different components in the mixture has also been investigated. Finally, the use of the new model in an integrated control dosing system for preventing hydrate formation is discussed.

Keywords: Prediction; Neural networks; Natural gas; Hydrate formation; Hydrate inhibition

1. Introduction

Hydrate formation may be a common occurrence during oil and gas drilling and production operations. Hydrates can form in surface production equipment and in the tubing of gas wells. Experimental measurement of hydrate formation conditions for every specific gas composition is impractical. Therefore, accurate prediction of the hydrate formation conditions is important in many applications such as the determination of the limits to expansion of natural gas during throttling in valves, chokes and restrictions.

Various correlations have been presented in the literature for predicting the hydrate formation conditions. These correlations can be classified into five major methods.

0378-3812/98/$ - see front matter
PII: S0378-3812(98)00368-9

The first is the *K*-value method, which utilizes the vapor–solid equilibrium constants for predicting hydrate-forming conditions [1]. The method is based on the concept that hydrates are solid solutions. The hydrate forming conditions are predicted from empirically estimated vapor–solid equilibrium constants given by

$$K_{\mathrm{vs},i} = y_i / x_i \tag{1}$$

where y_i is the mole fraction of the ith hydrocarbon component in the gas phase considered on a water-free basis and x_i is the mole fraction of the same component in the solid phase on a water-free basis. The hydrate formation conditions should satisfy

$$\sum_{i=1}^{n} y_i / K_{\mathrm{vs},i} = 1 \tag{2}$$

Inclusion of non-hydrocarbon gases such as CO_2, N_2 and H_2S may cause inaccurate results. Mann et al. [2] presented new *K*-charts that cover a wide range of pressures and temperatures. These charts can be an alternate to the tentative charts constructed by Carson and Katz [1], which are not a function of structure or composition. The new charts are based on the statistical thermodynamic calculations.

The second method is the gas–gravity plot developed by Katz [3]. The plot relates the hydrate formation pressure and temperature with gas gravity defined as the apparent molecular weight of a gas mixture divided by that of air. This method is a simple graphical technique that may be useful for an initial estimate of hydrate formation conditions. The hydrate formation chart was generated from a limited amount of experimental data and a more substantial amount of calculations based on the *K*-value method. A statistical accuracy analysis reported by Sloan [4], showed that this method is not accurate. For the same gas gravity, different mixtures may lead to about 50% error in the predicted pressure. Over the last 50 years, enormous experimental data on hydrate formation conditions have been collected and hence a more accurate gas gravity chart can be developed. One purpose of this study is to develop such chart as will be explained later.

The third method consists of empirical correlations developed according to the following form by Holder et al. [5] and Makogon [6] for selected pure gases.

$$P = \exp(a + b/T) \tag{3}$$

where a and b are empirical coefficients that depend on the temperature range for each gas. For natural gas, another correlation was presented following a general equation based on the gas gravity by Makogon [6].

$$\ln P = 2.3026\beta + 0.1144(T + kT^2) \tag{4}$$

where $\beta = 2.681 - 3.811\ \gamma + 1.679\ \gamma^2$ and $k = -0.006 + 0.011\ \gamma + 0.011\ \gamma^2$.

Based on the fit to Katz [3] gas–gravity plot, Kobayashi et al. [7] developed the following empirical equation for hydrate-forming conditions of natural gases.

$$\begin{aligned} T = 1/\Big[& A_1 + A_2(\ln \gamma_g) + A_3(\ln P) + A_4(\ln \gamma_g)^2 + A_5(\ln \gamma_g)(\ln P) + A_6(\ln P)^2 + A_7(\ln \gamma_g)^3 \\ & + A_8(\ln \gamma_g)^2(\ln P) + A_9(\ln \gamma_g)(\ln P)^2 + A_{10}(\ln P)^3 + A_{11}(\ln \gamma_g)^4 + A_{12}(\ln \gamma_g)^3(\ln P) \\ & + A_{13}(\ln \gamma_g)^2(\ln P)^2 + A_{14}(\ln \gamma_g)(\ln P)^3 + A_{15}(\ln P)^4 \Big] \end{aligned} \tag{5}$$

where $A_1 = 2.7707715 \times 10^{-3}$, $A_2 = -2.782238 \times 10^{-3}$, $A_3 = -5.649288 \times 10^{-4}$, $A_4 = -1.298593 \times 10^{-3}$, $A_5 = 1.407119 \times 10^{-3}$, $A_6 = 1.785744 \times 10^{-4}$, $A_7 = 1.130284 \times 10^{-3}$, $A_8 = 5.9728235 \times 10^{-4}$, $A_9 = -2.3279181 \times 10^{-4}$, $A_{10} = -2.6840758 \times 10^{-5}$, $A_{11} = 4.6610555 \times 10^{3}$, $A_{12} = 5.5542412 \times 10^{-4}$, $A_{13} = -1.4727765 \times 10^{-5}$, $A_{14} = 1.3938082 \times 10^{-5}$, $A_{15} = 1.4885010 \times 10^{-6}$. The equation is not recommended for temperatures above 62°F, pressures above 1500 psia, and gas gravities above 0.9.

The fourth method involves the charts of permissible expansion that a natural gas can undergo without risk of hydrate formation. These charts were redrawn with the aid of the gas gravity charts using the Joule–Thomson cooling curves [8]. This method is more suitable for rough design of valves, chokes, and flow provers where gas expansion normally occurs downstream of these flow restrictions. The charts are limited to natural gases composed predominantly of methane. An average error of 10% should be expected with this method. Higher errors may result when significant amounts of non-hydrocarbons or hydrocarbons heavier than ethane are present.

The fifth method is based on a statistical thermodynamic approach developed by van der Waals and Platteeuw [9]. This approach accounts for the interactions between water molecules forming the crystal lattice and gas molecules. It has been modified by many investigators [10–15]. Recent studies [16–21] have been done to validate the model for other constituents such as inhibitors and to improve its predictive accuracy by modifying the involved parameters. Accuracy and limitations of this method have been discussed by Sloan [22]. The method cannot predict accurately the gas hydrate formation pressures in systems containing carbon dioxide and aqueous electrolyte solutions [23]. For pure component systems excluding *i*-butane, an average error of 29% in pressure prediction can be expected. The error in predicting pressure for *i*-butane was 120% [24]. For multi-component systems, an average error of 25% may be expected also in pressure prediction. The approach is not purely theoretical and many fitted parameters are incorporated in it. In addition, the number and type of phases should be known a priori before the statistical thermodynamic approach can furnish predictions.

From the above discussion on the available models for predicting hydrate formation conditions, it is clear that there is a research need for developing a new model. This model should (1) require the least amount of input information, (2) give high accuracy, (3) be robust and less sensitive to noisy input data, and (4) can be continuously retrained (adapted) to a new input–output information. Neural network models offer all of the above desirable characteristics. There has been recently a resurgence of interest in the use of artificial neural networks, mainly because of their ability to handle complex and nonlinear problems. Artificial neural networks can be taught to learn correlative patterns between variables and can subsequently be used to predict outputs from new inputs.

One of the main objectives of the present study was to develop a neural network model that can predict the hydrate formation conditions for various pure gases, gas mixtures and various inhibitors. The use of neural networks was shown recently to be successful in the accurate prediction of phase equilibria [25] and for the prediction of PVT properties [26]. Various neural network models for the prediction of hydrate formation conditions will be presented. The first model is a gravity model that uses as its input variables the gas gravity and temperature and predicts the hydrate formation pressure. Various other neural network models that are based on the gas composition are also considered. These models consider pure hydrocarbon hydrate formers only, hydrocarbon and non-hydrocarbon hydrate formers only, or all hydrocarbons, non-hydrocarbon, and hydrate inhibitors. The latter model is the most general model. These models were trained using from 1012 to 2387 data points that are

published in the research literature. The predictions of the neural network models are compared to the available correlations and also to experimental data. The remainder of this paper will first give a brief overview of artificial neural networks and then presents the design, training, testing and validation of the various neural network models discussed previously. The testing and validation of the models indicate that the use of neural networks is promising for the accurate prediction of hydrate formation conditions for generalized gas systems. This apparent success leads us to discuss the integration of such models with inhibitor dosing devices for accurate hydrate prevention and to minimize the cost of inhibitors used.

2. Artificial neural networks

Neural networks are so named because they mimic the behavior of biological neurons and learn by trial and error. Neural networks are first subjected to a set of training data consisting of input data together with corresponding outputs. After a sufficient number of training iterations, the neural network learns the patterns in the data fed to it and creates an internal model, which it uses to make predictions for new inputs.

The foundation of an artificial neural network is the neuron or processing element (PE) (Fig. 1). Each processing element i receives an input vector $\boldsymbol{I}_i$ ($\boldsymbol{I}_i = [I_{i1}, I_{i2}, \ldots, I_{in}]$) which is weighted with a weight vector $\boldsymbol{W}_i$ ($\boldsymbol{W}_i = [W_{i1}, W_{i2}, \ldots, W_{in}]$). The processing element then calculates an output O_i, usually by using a sigmoidal function, which is a monatomic, continuously differentiable function, i.e.,

$$O_i = f(\boldsymbol{I}_i) = \frac{1}{1 + \exp(-\boldsymbol{I}_i)} \tag{6}$$

The PEs are arranged in a special topology or architecture. There are several possible architectures that can be used, however the most commonly used one is the multi-layer back-propagation

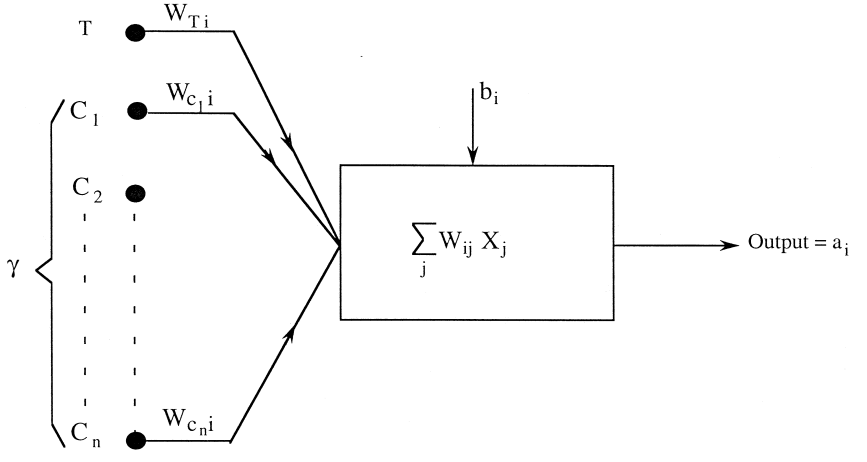

Fig. 1. A typical neuron i in the hidden layer.

architecture (Fig. 2). This type of network contains a layer of input neurons, a layer of output neurons, and one or several hidden layers of intermediate neurons. All the neurons are connected with different weights. In addition, there is a bias neuron that is connected to each neuron in the hidden and output layers. The input layer receives input from the outside world and processes them and transmits them to the hidden layer. The output from a neuron in the hidden and the output layers is also given by Eq. (6). In this case, the input $\boldsymbol{I}_i$ to a neuron in these layers is calculated from:

$$\boldsymbol{I}_i = \left(\sum_j W_{ij} O_j \right) + W_{i\mathrm{B}} O_{\mathrm{B}} \tag{7}$$

where the sum over j represents all neurons in the previous layer and O_{B} is an invariant output from the bias neuron. The weight W_{ij} represents the connection weight between neurons i and j and the weight $W_{i\mathrm{B}}$ is the connection weight between neuron i and the bias neuron.

The back-propagation neural network is applicable to a wide variety of problems and is the predominant supervised training algorithm. The term ‘back-propagation’ indicates the method by which the network is trained, and the term ‘supervised learning’ implies that the network is trained through a set of available input–output patterns. The network is repeatedly exposed to these input–output relationships and the weights among the neurons are continuously adjusted until the network ‘learns’ the correct input–output behavior. Initially, all the weights are set randomly and the difference or error between the desired and calculated outputs is calculated. This error is propagated back through the network and is used to update the weights among neurons iteratively. This update

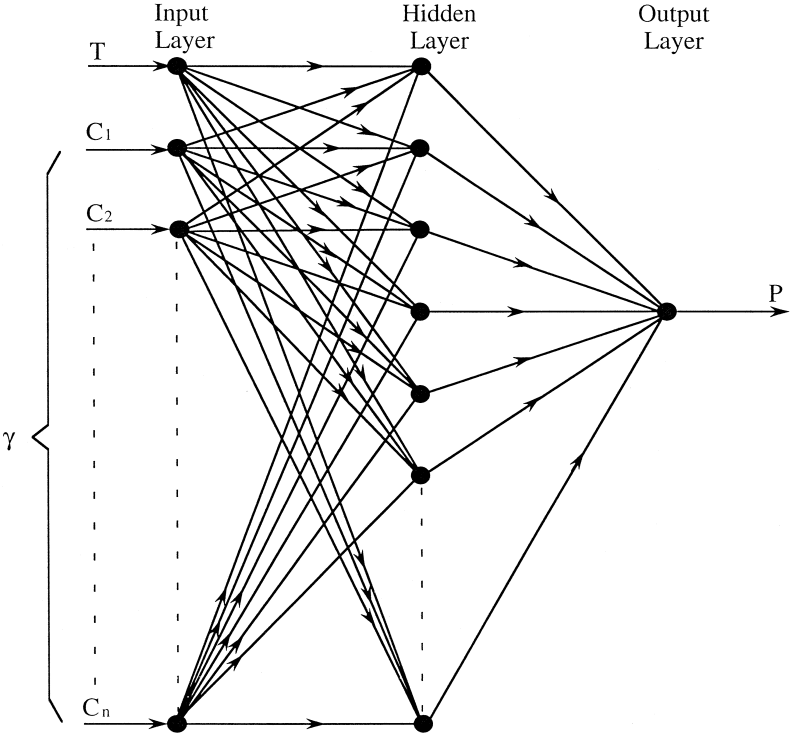

Fig. 2. A one hidden layer feed forward neural network.

process is repeated for all patterns many times until the network becomes capable of reproducing the input–output relationships to within an acceptable small tolerance.

The network depicted in Fig. 2 shows that in order to train the neural network, several weights have to be adjusted. One might therefore suggest that given several parameters in the usual regression, an accurate model could be prepared to predict the output from the given inputs. What is therefore the advantage of using neural networks? ANNs have several advantages over statistical methods. First, every node in the network encodes only a micro-feature of the network. That is, each node is not totally responsible for mapping an output to given inputs. Only when the effect of all nodes is combined, the network becomes able to predict an output corresponding to given inputs. Therefore, if the value of one-input is slightly off, the performance of the network will still be acceptable. However, in the usual empirical modeling, if the value of a variable is off, the model will usually predict inaccurate results. Secondly, an ANN can be designed to continuously retrain itself and adjust the connection weights among the neurons whenever it is subjected to a new input–output situation. This flexibility resulting in a self-correcting model does not exist in the usual empirical modeling. A third advantage of ANN modeling over statistical modeling is that an ANN can handle data consisting of multiple inputs–multiple outputs, unlike empirical modeling where the model is restricted to predicting only one output at a time. A last advantage of using ANNs, and probably the most important one, is that the training does not require the specification of the forms of the correlating functions. In the usual statistical regression, however, the user must specify the forms of the functions governing the correlation among the data. This process is usually time-consuming and requires a lot of physical insight. To prepare an ANN the only physical expertise needed is to decide what are the important inputs and outputs of the system.

The inputs to the neural network depicted in Fig. 2 are the temperature and composition of the hydrocarbons, non-hydrocarbons and inhibitors. It is not necessary to consider all of the above inputs at once. For instance, the gravity can be used instead of gas composition. With this motivation in mind, four ANN models (models-A, -B, -C, and -D) are considered depending on the input variables. These models are discussed in detail below.

In model-A, the functional relationship is essentially based on the assumption that pressure P depends on temperature T and specific gas gravity γ, or

$$P = f_1(T, \gamma) \tag{8}$$

where γ is determined from dividing the molecular weight for pure gas or apparent molecular weight (AMW) for gas mixture by the AMW of air, i.e., $\gamma = \text{AMW}/28.9625$ or $\gamma = \sum(y_i M_i)/28.9625$. This model has been tested for the pure hydrocarbon hydrate-formers; methane, propane, *i*-butane, and *n*-butane and their mixtures in order to compare with Katz gravity charts.

In model-B, the correlation is expressed in terms of the composition of pure hydrocarbon components and their mixtures, i.e.,

$$P = f_2\left(T, y_{C_1}, y_{C_2}, y_{C_3}, y_{i\text{-}C_4}, y_{n\text{-}C_4}\right) \tag{9}$$

where y is the mol% of each component.

In model-C, the non-hydrocarbon hydrate-formers: carbon dioxide, nitrogen, and hydrogen sulfide are also taken into account so that

$$P = f_3\left(T, y_{C_1}, y_{C_2}, y_{C_3}, y_{i\text{-}C_4}, y_{n\text{-}C_4}, y_{CO_2}, y_N, y_{H_2S}\right) \tag{10}$$

Finally in model-D, the hydrocarbon-non-hydrate-formers such as *i*-pentane, *n*-pentane (lumped as pentane), and hexane plus; and the inhibitors including methanol, sodium chloride, ethylene glycol, calcium chloride, and ethanol are collectively included to the previous model, i.e.,

$$P = f_4\left(T, y_{C_1}, y_{C_2}, y_{C_3}, y_{i\text{-}C_4}, y_{n\text{-}C_4}, y_{CO_2}, y_N, y_{H_2S}, y_{C_5}, y_{C_6^+}, y_{MeOH}, y_{EG}, y_{EtOH}, y_{NaCl}, y_{CaCl_2}\right) \quad (11)$$

In order to achieve appropriate computation stability during the learning phase of ANN, the domain of input variation should be within the range of the used data. This was performed by the following scaling rule:

$$T_{new} = \frac{T_{old} - T_{old,min}}{T_{old,max} - T_{old,min}} \quad (12)$$

where $T_{old,max}$ and $T_{old,min}$ are the maximum and minimum input values of previous (or initial) variable, whereas T_{old} and T_{new} are the scaled values of old and new variables, respectively. The output variable was transformed by using the logarithmic rule [27], i.e.,

$$P_{new} = \ln(P) \quad (13)$$

2.1. Data used

The hydrate phase–equilibrium data used in this study were mainly obtained from the comprehensive book by Sloan [22]. The author provided a magnificent overview of the experimental data available in the literature from 1934 to 1990. Additional data were also collected from other sources [18–21,23,28–30]. A total of 2389 data entries have been examined. The data were divided into single, binary, ternary, and multi-component systems (Table 1). The experimental data on pure hydrocarbon components include methane, ethane, propane, *i*-butane, and pentane (lumped *i*-pentane and *n*-pentane). Heavier components including crude oil constituents are lumped in hexane plus. Because of the lack of success in forming hydrates of pure *n*-butane, no reliable data were reported in the literature for this component. Even the role of *n*-butane in multi-component hydrocarbon systems was not obvious [31]. However, the data available for gas mixtures containing *n*-butane were considered. The pure non-hydrocarbon components involve carbon dioxide, nitrogen, and hydrogen sulfide. The consistency of the data was examined by Sloan [22] and assessed in the present study for the new data.

The phase equilibrium data for binary, ternary, and mixtures of these components were investigated. Depending on the pressure and temperature of the system, the data represented one or more of

Table 1
Types of the gases and materials used in the correlations

Type	Component
Hydrocarbon hydrate-formers	C_1, C_2, C_3, *i*-C_4, *n*-C_4
Non-hydrocarbon hydrate-formers	CO_2, N, H_2S
Hydrocarbon non-hydrate-formers	*i*-C_5, *n*-C_5, C_6^+
Inhibitors	MeOH, EG, EtOH, NaCl, $CaCl_2$

Table 2
Range of the measured data

Parameter	Minimum	Maximum
Pressure (kPa)	42.0	397,000
Temperature (K)	148.8	320.1
Gravity	0.554	2.01
Condensates (mol%)		
Pentane *	0.15	1.01
Hexane plus * *	0.05	78.74
Inhibitors (wt.%)		
Methanol	5.0	73.7
Ethylene glycol	10.0	50.0
Ethanol	15.0	16.5
Sodium chloride	1.09	26.4
Calcium chloride	10.0	36.0

* Lumped *i*-pentane and *n*-pentane.
* * Including *n*-octane and *n*-decane.

the following phase combinations: I–H–V, L_w–H–L_{hc}, L_w–H–L_{hc}–V, and L_w–H–L_{hc} on the general equilibrium pressure–temperature line. Experimental data on the retrograde behavior for some gas systems such as the binary hydrates of methane and propane were considered in the correlation. A similar behavior was indicated by Sloan [22] for the methane–carbon dioxide system. The data on the phase equilibria of various gas systems with inhibitors were also included. Data were collected for inhibitors including methanol, ethylene glycol, ethanol, and common electrolytes such as the aqueous solutions of sodium chloride and calcium chloride of various concentrations. The prohibitive effect of crude oil existing in natural gas on hydrate formation has been considered in hexane plus. The ranges of data utilized in this study are given in Table 2.

3. Results and discussion

In comparing the overall performance of the developed ANN models and the previous methods, these methods will be presented in two groups. The first group comprises the correlations, which are based on gas gravity of the hydrocarbon hydrate-formers. In this group, the ANN gravity model-A will be compared with the correlations suggested by Katz [3], Holder et al. [5], Makogon [6], and Kobayashi et al. [7]. The second group includes the methods based on gas composition such as the *K*-value approach and statistical thermodynamic model. Because the new charts of the *K*-value method were based on the statistical thermodynamic model [2], the compositional models developed in the present study were compared only with the statistical thermodynamic model (STM).

3.1. ANN gravity model-A

In the first stage of this study an attempt was made to develop an ANN model for predicting hydrate formation pressure in terms of gas gravity and temperature. This simplified correlation is useful when the gas composition is not known. It is developed for more data than those used with the

Table 3
Properties and statistical parameters of the neural network models

Model	No. of input variables	No. of data entries	No. of neurons	No. of hidden layers	AAPE	SSE
ANN-A	2	1012	25	1	32.657	0.9770
ANN-B	9	1012	50	1	40.265	0.9521
ANN-C	11	1791	40	1	20.860	0.9864
ANN-D	16	2389	40	1	19.416	0.9794

Katz [3] correlation presented in the year 1945. In order to compare with Katz [3] gravity charts, the model was trained by data for the pure hydrate-formers: C_1, C_2, C_3, *i*-C_4, and *n*-C_4, and their mixtures. Using 1012 input data entries, a correlation for ANN model-A was obtained with one hidden layer and 25 neurons. As illustrated in Table 3, the model correlated the experimental data with an average error of 32.657% and sum-squares error (SSE) of 0.977.

ANN gravity model-A was used to prepare an analogous chart to that of Katz [3]. Fig. 3 shows the hydrate formation conditions for various gas gravities: 0.554, 0.6, 0.7, 0.8, 0.9, and 1.0. While the first three curves are monotonically decreasing with gravity over a wide range of temperature variation, the latter three curves ($\gamma \geq 0.8$) show a limited range of temperature due to the lack of data. For gas gravities greater than 1.0, the corresponding curves (not shown on the graph for simplicity), become closer and ambiguous. This ambiguity can be attributed in part to the lack of data over that wide range of temperature shown in Fig. 3 and to the fact that using gravity only is not enough to define the corresponding thermodynamic system.

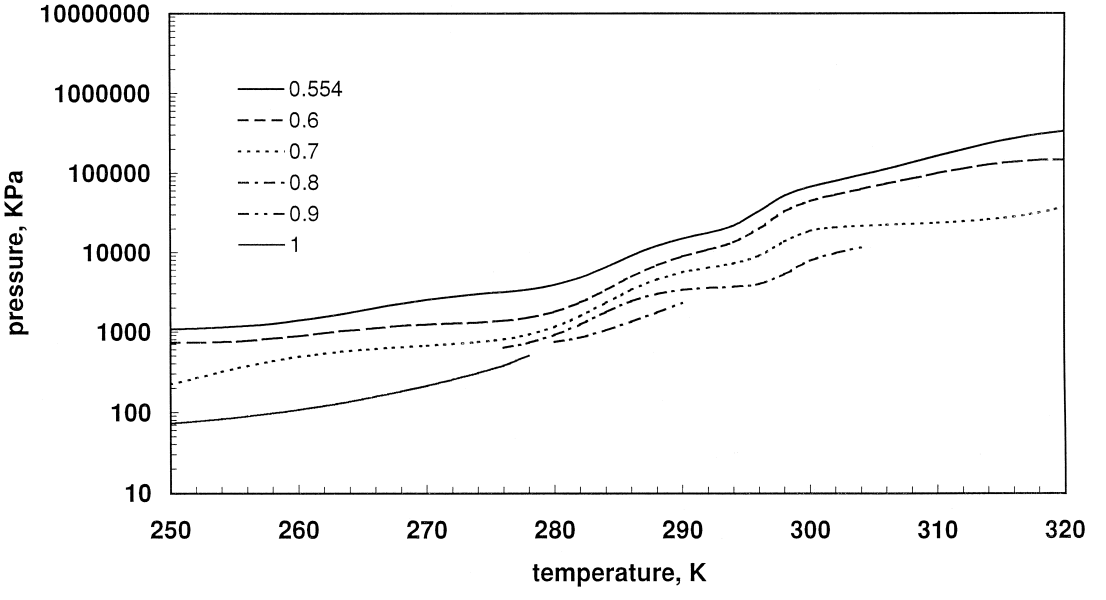

Fig. 3. Gas gravity chart by ANN model-A.

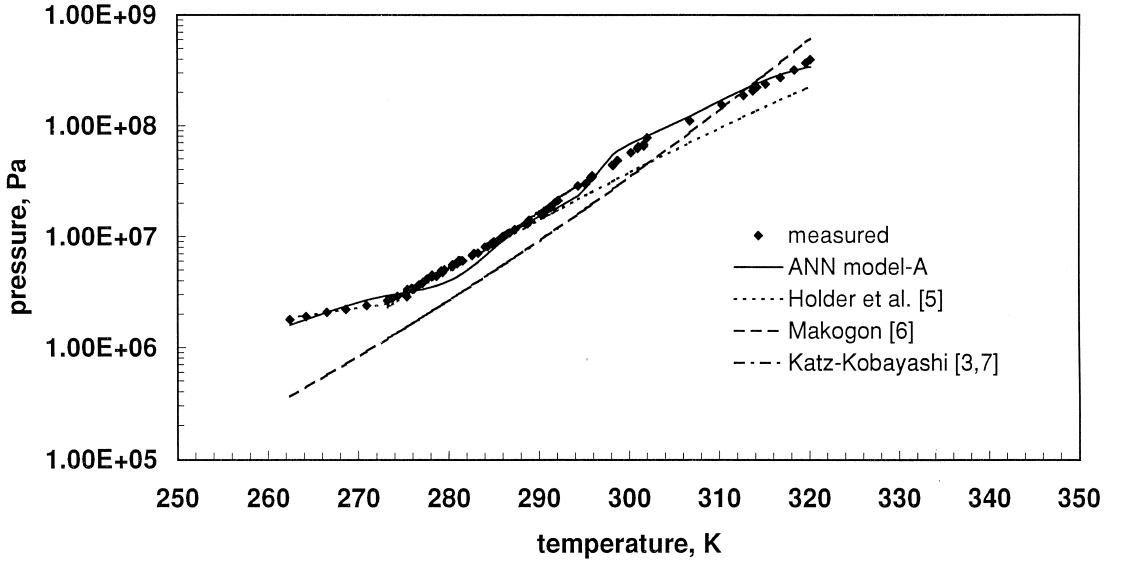

Fig. 4. Predictions by ANN gravity model-A and other correlations for 0.554 gas gravity.

Fig. 4 illustrates the hydrate formation pressures for an example gas gravity of 0.554. Results are presented as measured and calculated by ANN model-A and other methods including Holder et al. [5], Makogon [6], Katz [3], and Kobayashi et al. [7]. The latter three methods were chosen because they are mainly based on gravity graphical plots. As shown in Fig. 4, while the methods of Holder et al.

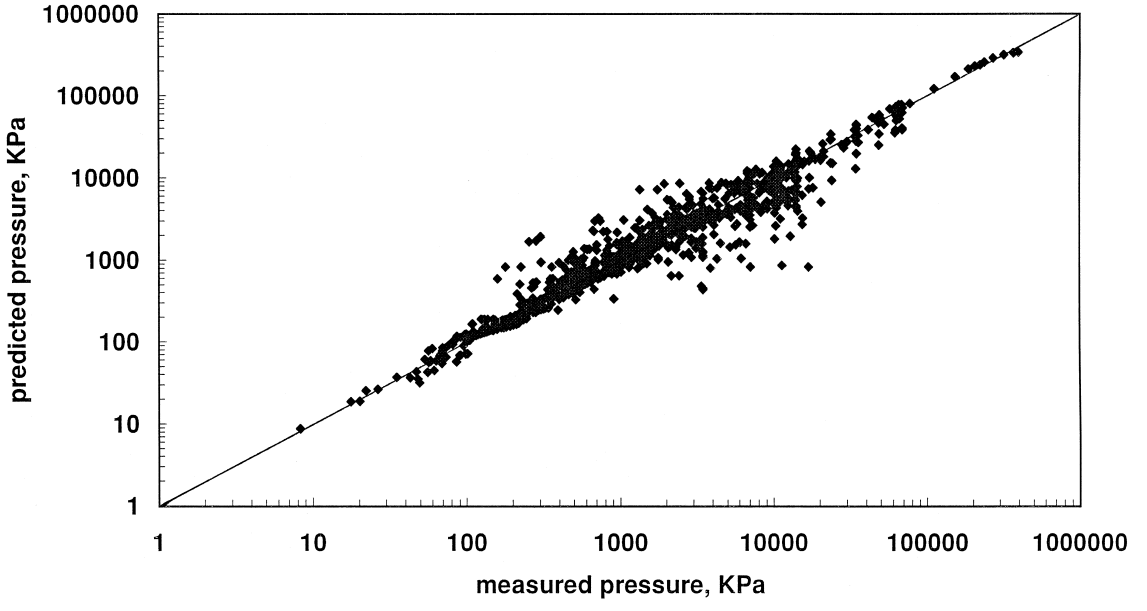

Fig. 5. Cross plot of ANN model-A.

[5], Katz [3] and Kobayashi et al. [7] are accurate over a limited data range, the Makogon [6] correlation gave the least accuracy for the examined data. The ANN model-A provided reasonably accurate results over a wide range of hydrate formation conditions.

The cross plot shown in Fig. 5, compares the measured hydrate formation pressure with that predicted by ANN model-A. The data are presented for various phase equilibria of pure hydrocarbon components and their mixtures. As can be seen, relatively little scatter is obtained around the 45°-line. This indicates a good agreement between the measured and calculated results.

3.2. ANN compositional models

In the next phase of this study, three ANN models were developed to consider the relationships indicated in Eq. (9) through Eq. (11). Several neural network architectures were attempted to find out the best accuracy. A one-hidden layer network was found to be suitable. The number of neurons in the hidden layer was varied until a minimum sum squared error was obtained. Table 3 demonstrates the final neural network properties and statistical parameters of these models. The number of neurons tabulated for the different models delivered acceptable results. They were justified by the corresponding cross plot verifications (Figs. 6–8).

The ANN model-B was developed to include the composition of the hydrocarbon hydrate-formers as indicated in Eq. (11). The model showed lower average error and sum-squares error than those of the ANN model-A. This may be due to the lack of data for pure *n*-butane, though data on its mixture with other components were available [31]. The model is capable to predict the hydrate formation pressure, which fitted well the experimental pressure for the methane–ethane system at different proportions as shown in Fig. 9.

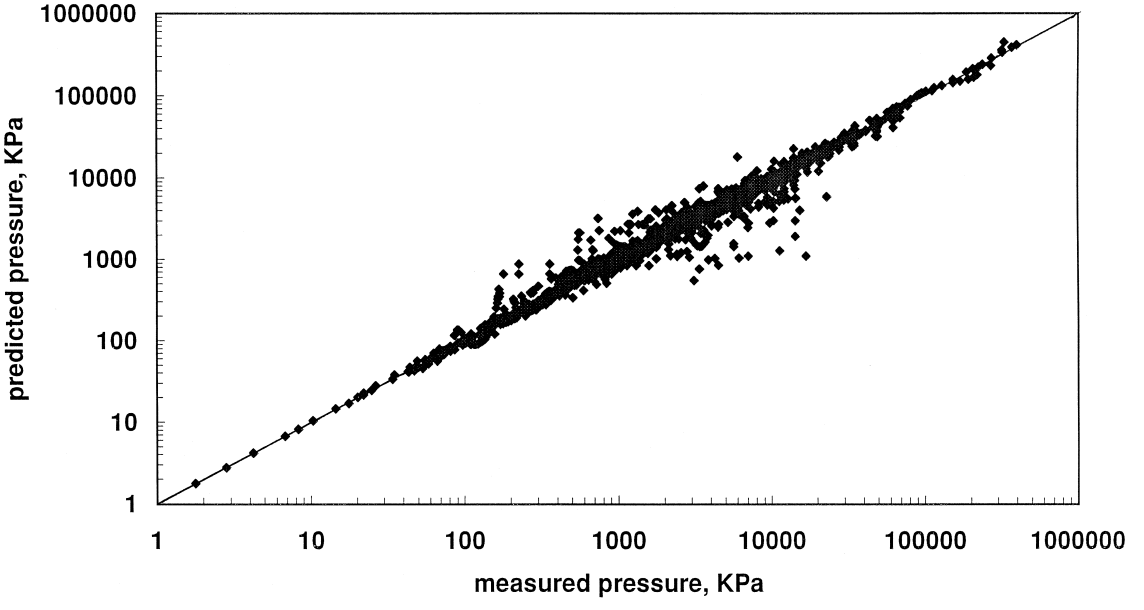

Fig. 6. Prediction of hydrate formation conditions by ANN model-B for methane–ethane mixture.

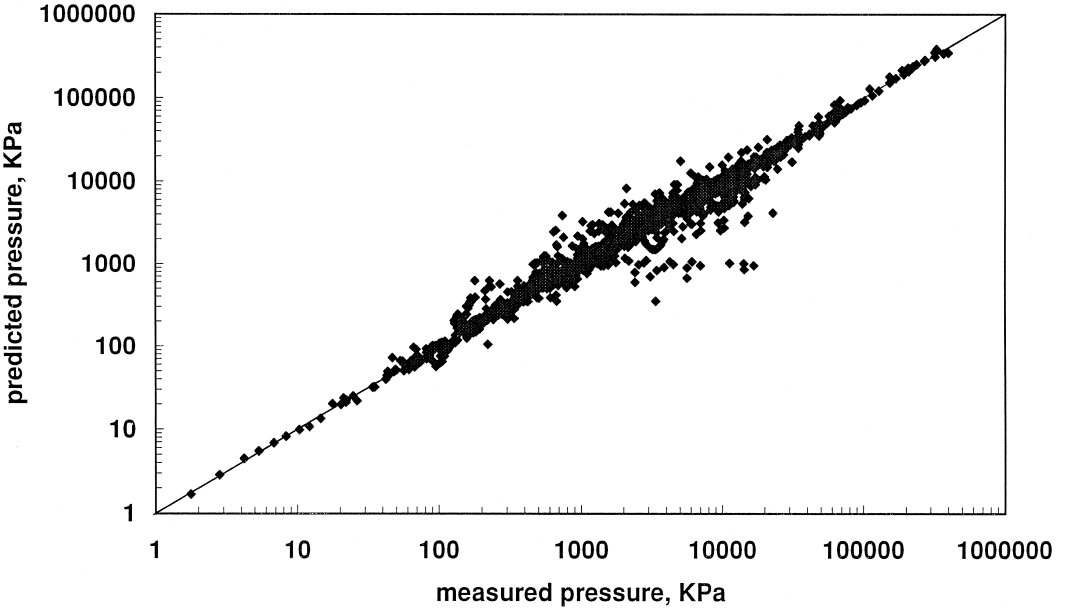

Fig. 7. Cross plot of ANN model-B.

3.3. Comparison of developed compositional ANN models

ANN model-D is the most comprehensive of the compositional models considered. The accuracy of this model is checked against the other more specialized models (Models-B and -C). These latter

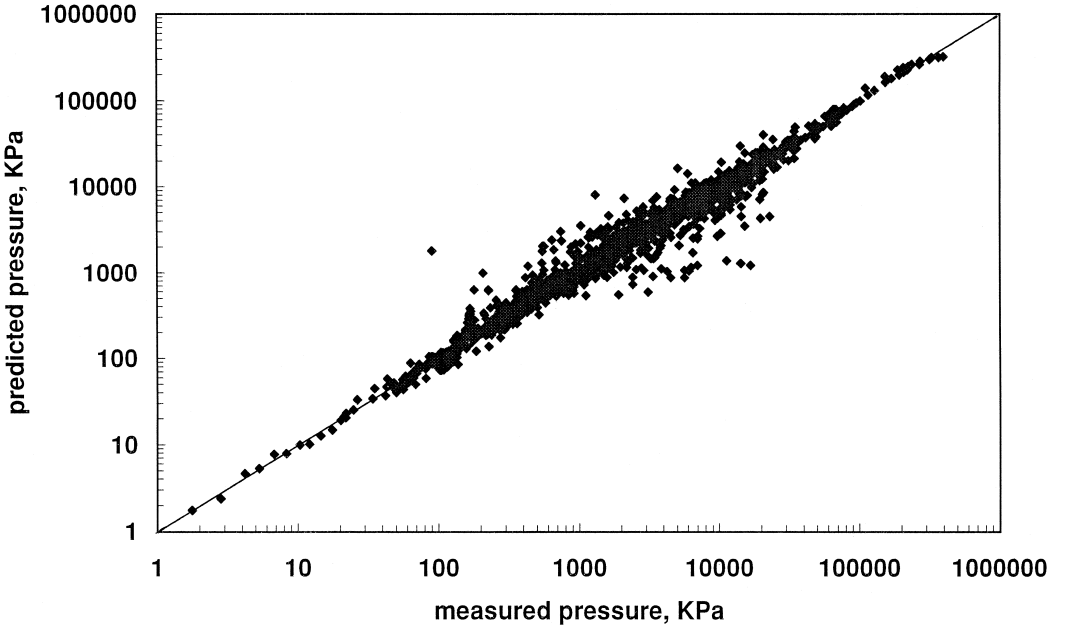

Fig. 8. Cross plot of ANN model-C.

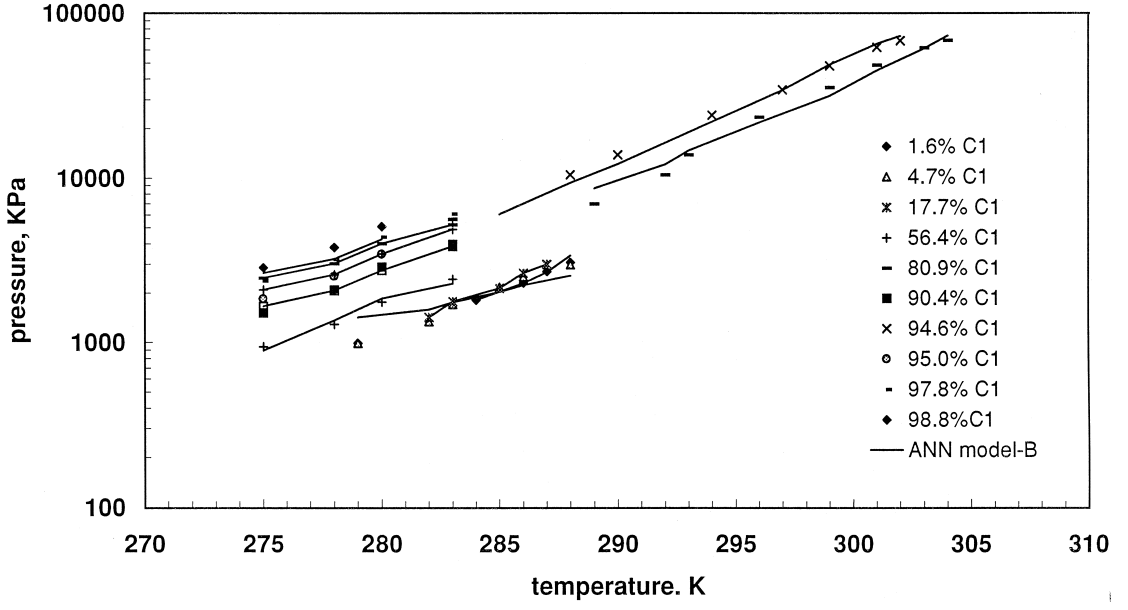

Fig. 9. Cross plot of ANN model-D.

models were considered to facilitate comparisons with models available in the literature and also to check whether or not more specialized models might lead to better accuracies than a generalized model. Table 3 compares the statistical parameters for these models. The results show that the comprehensive ANN model-D deliver a competitive correlation coefficient. For a total number of data entries of 2387, this model gives an average absolute error of 19.4% and sum-squares error of 0.9794. The cross plot of this model (Fig. 9) confirms this fact.

To further check the accuracy of the generalized model-D compared to models-C and -D, a small subset of the data was considered. In Table 4 the experimental hydrate-forming pressure is tabulated against the pressure predicted by different compositional models for this data subset. The results reveal that the ANN model-D fit is good and compares well with the other models. Therefore, it can be concluded that the comprehensive ANN model-D can be generally used to predict the hydrate formation pressure for all system compositions. The predictions of the statistical thermodynamic model (STM) are also included in Table 4 for comparison purposes.

The new model can predict the hydrate formation conditions for gas systems existing in one or more of the following phase regions: I–H–V, L_w–H–L_{hc}, L_w–H–L_{hc}–V, and L_w–H–L_{hc} on the general equilibrium pressure–temperature plot. Contrary to the statistical thermodynamic model, the developed model does not necessarily require a priori knowledge of the number and type of phases before prediction.

3.4. Validation of the developed models

The validation of the new models was examined by using data that have not been considered in their developments. Random samples of extra measured data points, which represent different phases

Table 4
Calculations of hydrate formation pressure for methane–ethane system by various compositional models

Temp. (K)	Composition (mol%)		Pressure (kPa)								
	Methane	Ethane	Measured	STM		ANN model-B		ANN model-C		ANN model-D	
				Predicted	APE	Predicted	APE	Predicted	APE	Predicted	APE
284	1.60	98.4	1810	1716	5.19	1810	0.131	1720	4.89	1720	5.13
286	1.60	98.4	2310	2249	2.64	2300	0.227	2360	2.36	2190	5.21
288	1.60	98.4	3080	3026	1.75	3400	10.4	3790	23.1	3240	5.12
279	4.70	95.3	990	924	6.67	1420	43.6	855	13.6	1180	19.0
283	4.70	95.3	1710	1508	11.81	1750	2.40	1490	13.2	1660	3.07
288	4.70	95.3	2990	2951	1.30	2560	14.3	3340	11.8	2520	15.6
282	17.7	82.3	1420	1385	2.46	1340	5.48	1350	4.68	1780	25.6
285	17.7	82.3	2140	1983	7.34	2200	2.73	2190	21.8	1840	14.2
287	17.7	82.3	3000	2545	15.17	8902	3.69	2990	0.339	3180	6.00
275	56.4	43.6	945	893	5.50	893	5.48	1290	36.4	946	0.0958
278	56.4	43.6	1290	1233	4.42	1360	5.69	1410	910	1290	0.0237
283	56.4	43.6	2430	2134	12.18	2280	6.34	1870	23.2	2230	8.46
304	80.9	19.1	68,600	34,451	49.78	73,300	6.93	10,100	43.9	8080	15.4
296	80.9	19.1	23,500	20,994	10.66	21,800	6.98	34,500	3.16	33,600	5.62
289	80.9	19.1	7000	6870	1.86	8700	24.3	64,900	5.33	79,600	16.0
275	90.4	9.60	1520	1913	25.86	1660	8.81	2710	77.9	1980	29.6
280	90.4	9.60	2890	3214	11.21	2760	4.43	3650	26.2	3020	4.40
283	90.4	9.60	3970	4447	12.02	3880	2.02	4650	17.3	3950	0.445
275	95.0	5.00	1840	2345	27.45	2100	14.1	1820	1.20	2140	16.1
278	95.0	5.00	2530	3174	25.45	2600	2.79	2290	9.50	2640	4.28
283	95.0	5.00	4770	5365	12.47	4900	2.74	4320	9.40	4580	3.98
275	97.1	2.90	2160	2623	21.44	2380	10.3	2230	3.45	2460	13.8
278	97.1	2.90	2960	3535	19.43	2920	1.24	2830	4.17	3040	2.89
280	97.1	2.90	4030	4329	7.42	3870	4.17	3810	5.63	3940	2.29
275	97.8	2.20	2370	2732	15.27	2480	4.96	2440	3.18	2600	9.73
280	97.8	2.20	4410	4495	1.93	4010	9.02	4120	6.74	4170	5.51
283	97.8	2.20	6090	6138	0.79	5630	7.46	5800	4.68	5650	7.12
275	98.8	1.20	2860	2907	1.64	2650	7.49	2740	4.19	2820	1.60
278	98.8	1.20	3810	3894	2.20	3240	15.0	3420	10.2	3490	8.32
280	98.8	1.20	5090	4753	6.62	4270	16.1	4500	11.6	4520	11.2
Average absolute percent error =					10.998		4.53		7.14		4.83

of hydrate equilibria, were used to test the accuracy of models. Results were compared to those of the available correlations.

Table 5 illustrates the statistical accuracy parameters obtained by models-A and -B compared with those for the statistical thermodynamic model (STM), modified by Paranjpe et al. [15] and Makogon [6] empirical correlation. Data were taken from Paranjpe et al. [15] and chosen within the range of applicability of the latter models, i.e., between 233 and 300 K, and pressures up to 27 MPa. Katz correlation was not included in the comparison because the gas gravity used here was out of its application range. The results show that the ANN models fit, well the cross validation experimental data and are superior to the other models.

Table 5
Comparison of hydrate formation pressure predicted by various correlations

Temp. (K)	Mol%		Mixture gravity	Pressure (kPa)								
	Propane	*i*-butane		Exp.	STM	APE	ANN-A	APE	ANN-B	APE	Makogon	APE
277	47.5	52.5	1.7768	355	302	14.929	4.16	17.2	370	4.23	2680	655.7
277	48.8	51.2	1.7705	365	311	14.79	3.72	1.91	386	5.75	2600	612.2
277	65.3	34.7	1.6906	426	375	11.97	3.48	18.4	505	18.5	1790	319.5
278	79.4	20.6	1.6223	490	469	4.29	3.93	19.7	463	5.51	1580	221.6
Average absolute percent error						11.495		14.30		8.51		452.2

The validity of the comprehensive ANN model-D was tested on two example systems. The first was the inhibition of hydrate formation of the methane–propane system by methanol. Data were collected from Ng and Robinson [12]. As shown in Fig. 10, the model correlated the experimental data with an average absolute error of 15.8%. The second example was on the sodium chloride inhibition of methane hydrates. Experimental data were taken from Sloan [22]. The comprehensive model fitted well the measured data, as indicated in Fig. 11. It gave an average error of 6.03%.

3.5. Relative importance of temperature and gas components in the formation of hydrates

In order to check the effect of the various gas components and also the temperature on the hydrate formation pressure, the partition method proposed by Garson [32] was employed. The general compositional model-D was used for this purpose because it takes into account the interactions among

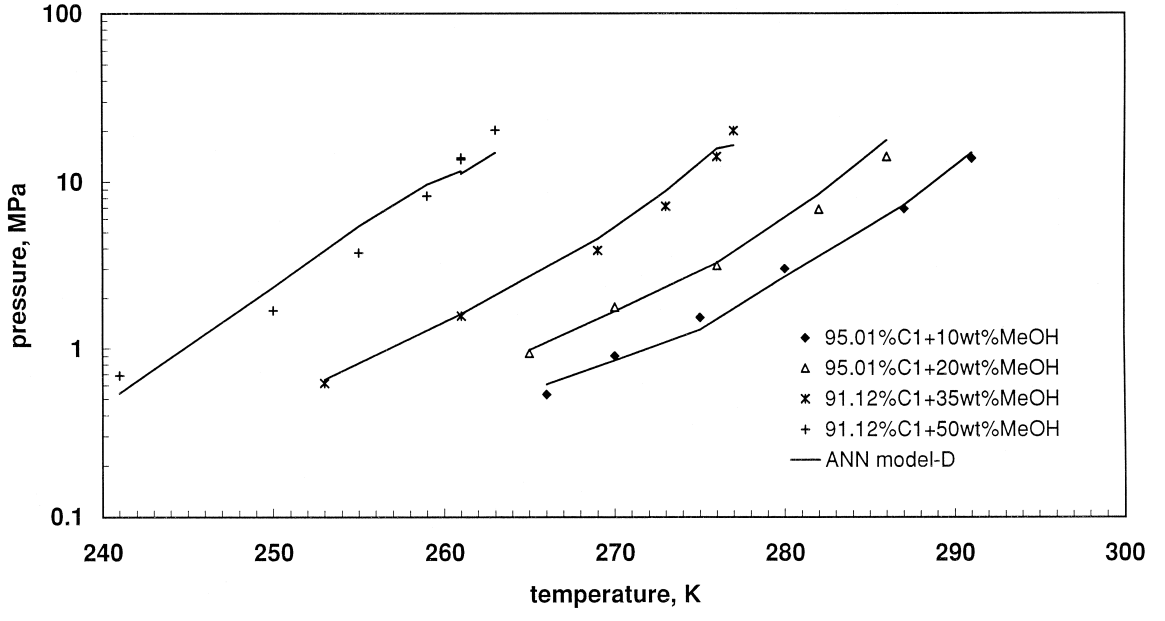

Fig. 10. Methanol inhibition of methane–propane hydrates.

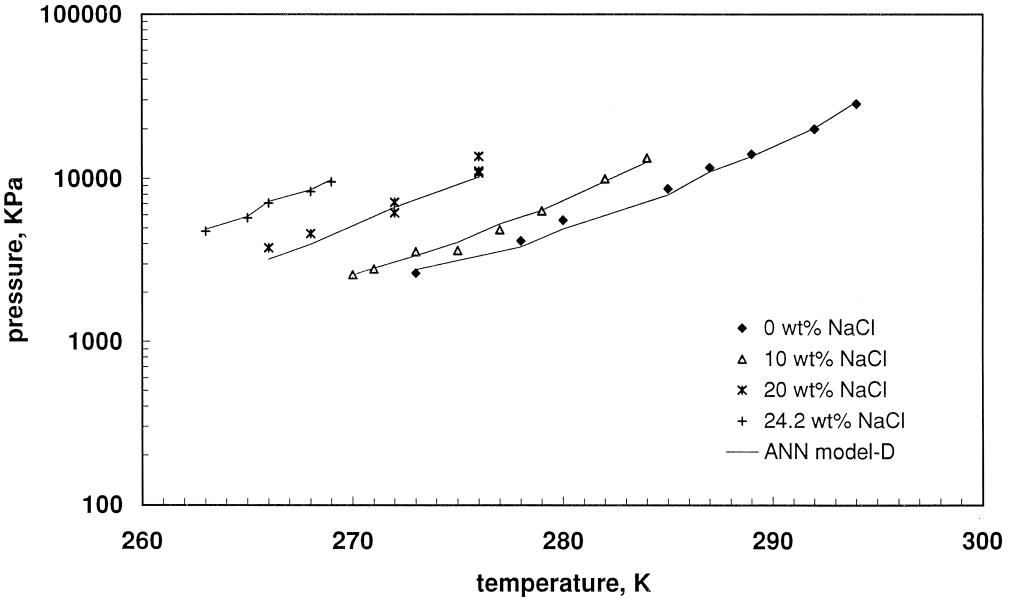

Fig. 11. NaCl inhibition of methane hydrates.

all the considered components in the hydrate formation system. The connection weights in the ANN architecture were used to assess the desired relative importance of the variables. The importance of each variable as proposed by Garson [32] is given by the following formula:

$$\mathrm{IM}(x_P) = \frac{\sum_{j=1}^{n_h}\left[\left(|I|_{P_j}/\sum_{k=1}^{n_P}|I|_{P_{j,k}}\right)|O|_j\right]}{\sum_{i=1}^{n_P}\left(\sum_{j=1}^{n_P}\left[\left(|I|_{P_{i,j}}/\sum_{k=1}^{n_P}|I|_{P_{i,j,k}}\right)|O|_j\right]\right)} \tag{14}$$

where $\mathrm{IM}(x_P)$ is the importance measure for the Pth input variable x_P. n_P is the number of input variables, and n_h is the number of neurons in the hidden layer. The term $|I|_{P_j}$ is the absolute value of the weight in the neural network corresponding to the Pth input variable and the j hidden layer, and the term $|O|_j$ is the absolute value of the output layer weight corresponding to the jth hidden layer. The formula essentially calculates the relative share of the output prediction associated with each input variable. The results of this calculation are given in the last row of Table 6. These results indicate that the composition of the different components play a great role in the hydrate formation. Methane has the biggest share, followed by propane, butane and ethane. It is interesting to note that the non-hydrocarbon components CO_2 and N_2 have also a significant importance in hydrate formation. Of the different inhibitors that can be used, methanol and NaCl seem to have the biggest inhibition role. The table indicates also that *n*-butane is comparable in importance with *i*-butane. Earlier studies did not indicate the significance of this component in hydrate formation.

The effect of temperature on hydrate formation conditions seems to be of less importance than composition. To check this further, the partition procedure was applied on the ANN gravity model-A to assess the importance of temperature against gravity, which reflects the gross gas composition. The

Table 6
Relative importance of the various inputs in the general compositional model

Input	Relative importance (%)
T	14.1
C_1	11.2
C_2	7.85
C_3	8.96
i-C_4	4.96
n-C_4	4.19
C_5	5.51
C_6^+	3.92
CO_2	6.30
N_2	6.36
H_2S	4.90
MeOH	8.48
Nacl	6.21
EG	3.71
$CaCl_2$	2.20
EtOH	1.45

results showed that temperature has a relative importance of 42% while gravity has a bigger share of 58%.

3.6. Integration with a dosing control system

Hydrate formation control should be considered to be an on-line service activity. The formation of hydrates can be avoided by introducing computerized systems along with monitoring instruments and on-line data collection. The objective is to formulate prompt and reliable remedial actions by injecting the appropriate amount of inhibitor for the measured temperature and pressure of the system. This type of 'modern strategy' for inhibition will minimize the cost of chemicals used for inhibition of hydrate formation, and maximize systems effectiveness.

In order to determine the required amount of an inhibitor to avoid the formation of hydrates within a system, the comprehensive compositional neural network model (Model-D) presented in this paper can be used. This can be done by invoking a valuable property of a neural network: the ability to solve the inverse problem of determining the amount of an inhibitor (for instance) that will lead to no hydrate formation at the current temperature, pressure and composition of other components. Thus, one can envision the proposed neural network model-D to be integrated within a generalized inhibitor control system consisting of monitoring instrumentation and on-line data collections. The study of such integration will be the subject of further research.

4. Concluding remarks

Utilization of the neural network technique for predicting the hydrate formation conditions has been investigated in the present study. The technique has been applied on a total of 2387 experimental

data sets obtained for phase equilibrium of pure, binary, and multi-component systems. These systems were composed of hydrate-forming and non-hydrate-forming constituents with and without inhibitors. Various types of inhibitors commonly used in petroleum industry were tested. Based on the results obtained, the following conclusions can be drawn.

(1) Several neural network architectures have been investigated to achieve the highest accuracy. One hidden layer network, with 25 neurons in the hidden layer showed reasonably accurate results for the gravity model. A similar network, with 40 (or 50) neurons in the hidden layer, was found to yield the least error for the compositional models.

(2) A new neural network model based on gas gravity has been developed. The model utilizes the updated experimental data on hydrate formation conditions. It achieved better accuracy when compared with the existing gravity-graphic method and empirical correlations. However, for gas gravities between 0.8 and 1.0, the model was valid for a limited range of available data. For gravities ≥ 1.0, the gravity neural network model showed ambiguous results. This ambiguity may be due to lack of experimental data, and that the use of gravity only in correlation is not sufficient.

(3) The developed compositional neural network models are simple empirical relationships that can predict hydrate formation pressure as a function of temperature and composition. These models delivered satisfactorily accurate results in comparison with the available correlations and models. The models require much less input information when compared with the previous models. The STM for instance, requires the knowledge of the number and type of phases before it can furnish predictions. These input data are not needed for the new models.

(4) Finally, because of the lack of sufficient experimental data especially for some binary and multi-component hydrate systems, the developed models have to be updated by being retrained by using extra collected data.

5. List of symbols

$A_1, A_1, \ldots, A_{15}$	empirical coefficients in Eq. (5)
AAPE	average absolute percent error
APE	absolute percent error
a, b	empirical coefficients in Eq. (3)
$\boldsymbol{I}_i$	input received by the processing element i
IM(x_P)	importance measure of the input variable x_P
$\lvert I\rvert_{P_j}$	absolute weight corresponding to Pth input variable and jth layer
$K_{\mathrm{vs},i}$	vapor–solid equilibrium ratio of component i, dimensionless
k	empirical coefficient in Eq. (4)
n_{h}	number of neurons in the hidden layer
n_P	number of input variables
O_i	output calculated by the processing element i
O_j	output calculated from neuron in the previous layer
O_{B}	invariant output from the bias neuron
$\lvert O\rvert_j$	absolute output weight corresponding to the jth hidden layer
P	pressure, Pa
SSE	sum-squares error

T	temperature, K
$\boldsymbol{W}_i$	weight vector of the input received by the processing element i
W_{ij}	connection weight between neurons i and j
$W_{i\mathrm{B}}$	connection weight between neuron i and the bias neuron B
x_i	mole fraction of component i in the hydrate phase, dimensionless
y_i	mole fraction of component i in the vapor phase, dimensionless
β	empirical coefficient in Eq. (4)
γ	gas gravity, dimensionless

References

[1] D.B. Carson, D.L. Katz, Natural gas hydrates, Trans., AIME 146 (1942) 150.
[2] S.L. Mann, L.M. McClure, F.H. Poettmann, E.D. Sloan, Vapor–Solid Equilibrium Ratios for Structure I and II Natural Gas Hydrates, Proc. 68th Ann. Gas Proc. Assoc. Conv., San Antonio, TX, March 13–14, 1989, pp. 60–74.
[3] D.L. Katz, Predictions of conditions for hydrate formation in natural gases, Trans., AIME 160 (1945) 140–144.
[4] E.D. Sloan, Phase Equilibrium of Natural Gas Hydrates, Proceedings of 63rd GPA Convention, 1984, 163–169.
[5] G.D. Holder, S.P. Zetts, N. Pradhan, Phase behavior in systems containing clathrate hydrates, Reviews in Chemical Engineering 5 (1) (1988) .
[6] Y.F. Makogon, Hydrates of Natural Gas, PennWell, 1981, pp. 12–13.
[7] R. Kobayashi, Y.S. Kyoo, E.D. Sloan, in: H.B. Bradley et al. (Eds.), Phase Behavior of Water/Hydrocarbon Systems, Petroleum Engineering Handbook, SPE, Dallas, Richardson, TX, 1987, pp. 25–13.
[8] D.L. Katz, A Look Ahead in Gas Storage Technology, Proceedings AGA Trans. Conference, 1981, T283–T289.
[9] J.H. van der Waals, J.C. Platteeuw, Clathrate solutions, Advances in Chemical Physics 11 (1959) 1–57.
[10] W.R. Parrish, J.M. Prausnitz, Dissociation pressures of gas hydrates formed by gas mixtures, Ind. and Eng. Chem. Process Design and Development 11 (1972) 26–34.
[11] I. Nagata, R. Kobayashi, Predictions of dissociation pressures of mixed gas hydrates from data for hydrates of pure gases with water, Ind. and Eng. Chem. Fundamentals 5 (6) (1966) 466–469.
[12] H.J. Ng, D.B. Robinson, The measurement and prediction of hydrate formation in liquid hydrocarbon–water system, Ind. Eng. Chem. Fundamentals 15 (4) (1976) 293–298.
[13] V.T. John, K.D. Papadopoulos, G.D. Holder, A generalized model for predicting equilibrium conditions for gas hydrates, AIChE J. 31 (2) (1985) 252–259.
[14] P.B. Dharmawardhana, W.R. Parrish, E.D. Sloan, Experimental thermodynamic parameters for the prediction of natural gas hydrate dissociation conditions, Ind. Eng. Chem. Fundamentals 19 (4) (1980) 410–414.
[15] S.G. Paranjpe, S.L. Patil, V.A. Kamath, S.P. Godbole, Hydrate Equilibria for Binary and Ternary Mixtures of Methane, Propane, Isobutane, and *n*-Butane: Effect of Salinity, SPERE, Nov. 1989, 446–454.
[16] A. P Mehta, E.D. Sloan, Structure H hydrate phase equilibria of methane + liquid hydrocarbon mixtures, J. Chem. Eng. Data 38 (4) (1993) 580–582.
[17] A.P. Mehta, E.D. Sloan, A thermodynamic model for structure-H hydrates, AIChE J. 40 (2) (1994) 312–320.
[18] B. Tohidi, Phase Equilibria in the Presence of Saline Water Systems and Its Application to the Hydrate Inhibition Effect of Produced Water, paper SPE 28884 presented at the European Petroleum Conference, London, Oct. 25–27, 1994.
[19] B. Tohidi, A. Danesh, R.W. Burgass, A.C. Todd, Hydrates Formed in Unprocessed Wellstreams, paper SPE 28478 presented at the SPE Annual Technical Conference and Exhibition, New Orleans, Sept. 25–28, 1994.
[20] B. Tohidi, A. Danesh, A.C. Todd, Modeling single and mixed electrolyte solutions and its applications to gas hydrates, Chem. Eng. Res. and Des. 73A (1995) 464–472.
[21] B. Tohidi, A. Danesh, A.C. Todd, R.W. Burgass, Measurement and Prediction of Hydrate-Phase Equilibria for Reservoir Fluids, SPE Production and Facilities, May 1996, 69–76.
[22] E.D. Sloan Jr., Clathrate Hydrates of Natural Gas, Marcel Dekker, New York, 1990, pp. 285–386.

[23] P. Englezos, Computation of the incipient equilibrium carbon dioxide hydrate formation conditions in aqueous electrolyte solutions, Ind. Eng. Chem. Res. 31 (1992) 2232–2237.
[24] API Technical Data Book, 1982, pp. 9–163.
[25] M.J. Lee, J.T. Chen, Fluid property prediction with the aid of neural networks, Ind. Eng. Chem. Res. 32 (1993) 995–997.
[26] W.A. Habiballah, R.A. Startzman, M.A. Barrufet, Use of Neural Networks for Prediction of Vapor/Liquid Equilibrium *K* Values for Light-Hydrocarbon Mixtures, SPERE, May 1996, 121–126.
[27] R. Stein, Preprocessing Data for Neural Networks, AI Expert, March, 1991, 33–37.
[28] K.Y. Song, R. Kobayashi, Final hydrate stability conditions of methane and propane mixture in the presence of pure water and aqueous solutions of methanol and ethylene glycol, Fluid Phase Equilibria 47 (1989) 295–308.
[29] L.F.S. Rossi, C.A. Gasparetto, Prediction of Hydrate Formation in Natural Gas Systems, SPE 28478 presented at the 64th Annual Technical Conf. And Exhib., Dallas, TX., October 6–9, 1991.
[30] P. Englezos, Z. Huang, P.R. Bishnoi, Prediction of natural gas hydrate formation conditions in the presence of methanol using the Trebble–Bishnoi equation of state, J. Canadian Petroleum Technology 30 (2) (1991) 148–155.
[31] H.J. Ng, D.B. Robinson, The role of *n*-butane in hydrate formation, AIChE J. 22 (4) (1976) 656–661.
[32] G.D. Garson, Interpreting Neural-Network Connection Weight(s), AI Expert, April, 1991, 47.