A Comparative Study on the Prediction of Compresibility Factor for Generalized Equation of State using Artificial Neural Network

M. Ravi Kumar¹, Ermias Girma Aklilu²
Department of Chemical Engineering,
College of Engineering and Technology,
Samara University, Samara, Afar, Ethiopia.

Abstract - The functional relationship among reduced pressure, compressibility, acentric factor, and reduced temperature of a gas was known as generalized equation of state. Though several generalized equations of state were available in literature each equation has its inherent disadvantages. There is no single generalized equation of state, which was mathematically friendly and can satisfy the experimental compressibility data with engineering accuracy. Hence a two parameter \( z(T, P) \) of generalized EOS model was developed using Artificial Neural Network in the present work. A feed forward network with back propagation of errors algorithm was used in training the network. In the supervised learning, sum squared error was used as lapnov function. The model was found to be in good agreement with the experimental results. In this work Comparison of various neurons were used to find the compressibility factor.

Keywords: Artificial Neural Network, Compressibility Factor, Sum Squared Error, Acentric Factor.

INTRODUCTION

The development of Artificial Neural Network began approximately 58 years ago (Mc Culloch and Pitts, 1943), inspired by a desire to understand the human brain and emulate its functioning. Within the last decade, it has experienced a huge resurgence due to the development of more sophisticated algorithms and emergence of more powerful computational tools. Extensive research has been devoted to investigate the potential of ANN as computational tools that acquire, represent and compute a mapping from one multivariate input space to another. Mathematically an ANN is often viewed as a Universal Approximator. The ability to identify a relationship from given patterns make it possible for ANNs to solve large scale complex problems such as pattern recognition, nonlinear modeling, classification, association and control. Although the idea of ANNs was proposed over fifty years ago, a tremendous growth in the interest of this computational mechanism has occurred since Rumelhart et al (1986) rediscovered a mathematically rigorous theoretical framework for neural networks, i.e., back-propagation algorithm. Consequently ANNs have found applications in diverse areas such as

1. Neuro-Psychology
2. Physics
3. Various Engineering Fields
4. Computer Science
5. Acoustics
6. Robotics
7. Image Processing
8. Financing and Others.

A neural network is characterized by its architecture that represents the pattern of connection between the nodes, its method of determining connection weights and the activation function. A typical ANN consists of a number of nodes that are organized according to a particular arrangement. One way of classifying a neural network is by the number of layers.

1. Single layer – Hop field et al
2. Bilayer – Grosberg adaptive resonance network

Input layer: The input layer receives the input variables for the problem at hand. This consists of all the quantities that can influence the output. The input layer is thus transparent and is a means of providing information to the network.

Hidden layer: The layers processing information are in between the input and output layers, which has any number of nodes.

Output layer: This layer gives the output of the network to external receptor.

Each layer consists of certain number of nodes, which are connected by links. No two nodes in the same layer will be having any links. A synaptic weight is assigned to each link to represent the relative connection strength of two nodes at both ends in predicting the input-output relationship. In this present study comparison of neurons used to find the compressibility factor.

Compressibility factor (Z) Theory:
Virial Equation: -

A useful auxiliary thermodynamic property is defined by the equation.

\[
Z = \frac{PV}{RT}
\]
This dimensionless ratio is called compressibility factor.

\[ Z = 1 + B'P + C'P^2 + D'P^3 + \ldots \]  

Equation (2)

An alternative expression for \( Z \) is also in common use.

\[ Z = 1 + \frac{B}{V} + \frac{C}{V^2} + \frac{D}{V^3} + \ldots \]  

Equation (3)

Both of these equations are known as virial expansions, and the parameter \( B', C', D', \) etc., and \( B, C, D, \) etc., are called virial coefficients. Parameters \( B' \) and \( B \) are second virial coefficients; \( C' \) and \( C \) are third virial coefficients; etc. For a given gas the virial coefficients are functions of temperature only.

The two sets of coefficients is 2 and 3 are related as follows:

\[ B' = \frac{B}{RT}, \quad C' = \frac{C-B^2}{(RT)^2}, \quad D' = \frac{D-3BC+12B^2}{(RT)^3} \quad \text{etc.} \]

Ideal gas:

Since, the terms \( B/V, C/V^2, \) etc., of the virial expansion [Eq. (3)] arise on account of molecular interactions, the virial coefficient \( B, C, \) etc., would be zero if no such interactions existed. The virial expansion would then reduce to:

\[ Z = 1 \quad \text{or} \quad PV = RT \]

For a real gas, molecular interactions do exist, and exert an influence on the observed behavior of the gas. As the pressure of a real gas is reduced at constant temperature, \( V \) increases and the contributions of the terms \( B/V, C/V^2, \) etc., decrease. For a pressure approaching zero, \( Z \) approaches unity, not because of any change in the Virial coefficient, but because \( V \) becomes infinite. Thus in the limit as the pressure approaches zero, the equation of state assumes the same simple form as for the hypothetical case of \( B = C = \ldots = 0; \) i.e.,

\[ Z = 1 \quad \text{or} \quad PV = RT \]

Differentiation of Eq.(2) for a given temperature gives:

\[ \left( \frac{\partial Z}{\partial P} \right)_T = B' + 2C'P + 3D'P^2 + \ldots \]  

Equation (4)

From which,

\[ \left( \frac{\partial Z}{\partial P} \right)_{T,P=0} = B' \]

Thus the equation of tangent line is:

\[ Z = 1 + BP' \]

a result also given by truncating the Eq. 2 to two terms. A more common form of this equation results from the substitution, \( B' = B/RT; \)

\[ Z = \frac{PV}{RT} = 1 + \frac{BP}{RT} \]  

Equation (5)

Equation (3) may also be truncated to two terms for application at low pressure:

\[ Z = \frac{PV}{RT} = 1 + \frac{B}{V} \]  

Equation (6)

and virial equation is truncated to three terms, the appropriate form is:

\[ Z = \frac{PV}{RT} = 1 + \frac{B}{V} + \frac{C}{V^2} \]  

Equation (7)

Pitzer Correlations for the Compressibility Factor:

The correlation for \( Z \) takes the form:

\[ Z = Z^0 + \omega Z^1 \]  

Equation (8)

Where \( Z^0 \) and \( Z^1 \) are functions of both \( T \) and \( P \). When \( \omega = 0 \), as is the case for the simple fluids, the second term disappears, and \( Z^0 \) becomes identical with \( Z \). Thus a generalized correlation for \( Z \) as a function of \( T \) and \( P \) based on data for just argon, krypton, and xenon provides the relationship \( Z^0 = F^0(T, P) \). By itself, this represents a two-parameter corresponding states correlation for \( Z \). Since the second term of Eq. (8) is a relatively small correlation to this correlation, its omission does not introduce large errors, and a correlation for \( Z^1 \) may be used alone for quick but less accurate estimates of \( Z \) than are obtained from a three parameters correlation.

Equation (8) is simple linear relation between \( Z \) and \( \omega \) for given values of \( T \), and \( P \), do indeed yield approximately straight line, and their slopes provide values for \( Z^1 \) from which the generalized function \( Z^1 = F^1(T, P) \) can be constructed.

The simplest form of the Virial equation has validity only to low to moderate pressures where \( Z \) is linear in pressure. The generalized Virial-coefficient correlation is therefore useful only where \( Z^0 \) and \( Z^1 \) are at least approximately linear functions of reduced pressure.

**Levenberg Marquardt Method:**

This ANN design model uses Levenberg Marquardt method as multi variable optimization technique to predict the sum of squared errors for the designed target value.

**Neural Network Architecture:**

A neural network is characterized by its architecture that represents the pattern of connection between the nodes, its method of determining connection weights and the activation function. A typical ANN consists of a number of nodes that are organized according to a particular arrangement. One way of classifying a neural network is by the number of layers.

1. Single layer – Hop field et al.

Input layer: The input layer receives the input variables for the problem at hand. This consists of all the quantities that can influence the output. The input layer is thus transparent and is a means of providing information to the network.
Hidden layer: The layers processing information are in between the input and output layers, which has any number of nodes.

Output layer: This layer gives the output of the network to external receptor.
Each layer consists of certain number of nodes, which are connected by links. No two nodes in the same layer will be having any links. A synaptic weight is assigned to each link to represent the relative connection strength of two nodes at both ends in predicting the input-output relationship.
Fundamentals of Neural Computations: -

A node/neuron/processing element is an abstraction for a computational procedure, consists of three steps.
1. Calculating the weighted sum of input (net input) signals.
2. Applying a threshold (the value which excites or inhibits the nodes firing) on the weighted sum of input signals.
3. Calculating the output according to particular output function called transfer function or activation function.

Weight factor is nothing but the strength of an interconnection between two neurons expressed numerically.

Transfer function is nothing but the mapping relation between the output of node and its threshold applied weighted sum of input signals.

Examples of threshold functions are
- Radial basis transfer function: \( f(x) = e^{-\frac{x^2}{a}} \)
- Logarithmic sigmoid transfer function: \( f(x) = \frac{1}{1+e^{-x}} \)

Normalizing the input-output data: The normalization of input and output data is very critical for successful training of network. If input and output variables do not have same order of magnitude some variables may appear to have more significance than they actually do. The training algorithm has to compensate for this and this is not very effective in many training algorithms. And also transfer functions like sigmoid and hyperbolic tangent are very effective in many training algorithms. And also variables may appear filled transfer function or calculated gradient descent correction term. The equation for adjusting the weight factors is:

\[
\Delta w_{ij} = \alpha \Delta w_{ij}^{old} + \Delta w_{ij}^{new} 
\]

If \( T_j \) is the bias for \( j^{th} \) neuron in output layer, \( \Delta T_j = \beta_0 \epsilon_j \)

When momentum correction is applied, \( \Delta w_{ij} = \beta_0 \epsilon_j \)

When learning rate is very slow, the training requires more number of iterations and the effect is very pronounced for complex networks and the training may be trapped at local minima (of obj. fn.) on the contrary if learning rate is very high the minima is never reached. Hence initially reasonably high and safe learning rate should be taken which goes on decreasing with iterations. General suggestion for learning rate and momentum coefficient is 0.3 and 0.4 respectively.

Selecting proper transfer function: Sigmoid and hyperbolic tangent functions are recommended for prediction/modeling and radial basis for fault diagnosis and feature categorization applications.

Selecting training data (Representative data): A representative data set prepared for training should contain information of all peaks, troughs and middle points also. Total number of data points available is divided into 4:1 ratio of training to testing data sets. Testing data sets should also contain some data points that belongs peaks, troughs and middle points.

Back propagation of error algorithm: -

The back propagation or errors learning algorithm is nothing but changing the weight factors based between layer of interest and its previous layers based or error back propagated from the layer preceded by the layer of interest.

The weight factors in the network between any two layers are calculated as follows for minimizing the sum-squared error using steepest descent method.

Let \( w_{ij} \) be the weight factor between \( i^{th} \) hidden neurons and \( j^{th} \) output neuron. Then the change in \( w_{ij} \) i.e., \( \Delta w_{ij} \) is given by \( \Delta w_{ij} = \beta_0 \epsilon_j \)

When momentum correction is applied, \( \Delta w_{ij} = \beta_0 \epsilon_j \)

where \( \beta_0 \) is learning rate for output layer \( y_j \) is output of \( j^{th} \) neuron in the hidden layer previous to output layer. \( \epsilon_j \) is error of \( j^{th} \) neuron in output layer given by the difference between its actual and network outputs.

\[
W_{ij(new)} = W_{ij(old)} + \Delta w_{ij} 
\]

If \( T_j \) is the bias for \( j^{th} \) neuron in output layer, \( \Delta T_j = \beta_0 \epsilon_j \)

If \( T_j \) is the bias for \( j^{th} \) neuron in hidden layer, \( \Delta T_h = \beta_0 \epsilon_h \)

When momentum correction is applied, \( \Delta w_{ij} = \beta_0 \epsilon_j \)

where \( t \) indicates iteration number and \( \alpha \) is the momentum coefficient.

\[
E_2 = \epsilon_2 \Delta w_{12} = \beta_0 x_1 \epsilon_2 \epsilon_{12} = y_2 (1-y_2) [w_{21} \epsilon_1 + w_{22} \epsilon_2] \
\Delta w_{21} = \beta_0 x_2 \epsilon_2 \Delta w_{12} = \beta_0 x_1 \epsilon_2 
\]
The back propagation learning algorithm:

Step 0: Initialize weights.

Step 1: While stopping condition is false.

For each pair of set:

- a. Calculate outputs of all neurons in the network
- b. Calculate new weight factors and biases

Step 2: While stopping condition is false (go to step 1) else break

RESULTS AND DISCUSSIONS

The Program code is run for a target goal of 0.002 and simulated for different number of Neurons as follows:

- Target Error = 0.002
- No of Neurons = 1

A Three dimensional plot of $T_r$, $P_t$ and $Z$ values for the first run is drawn as shown below.

Sum-squared error goal (SSE) = 2.92067.

TRAINLM: Network error did not reach the error goal. Further training may be necessary, or try different initial weights and biases and/or more hidden neurons.

Fig. 1. 3-D Plot between ($T_r$, $P_t$, $Z$)

Fig. 2 ANN Prediction Vs Actual data points in testing for one number of neurons with an error of 0.002
Fig. 3  Experimental Vs Model Predicted for testing Data points for one number of neurons with an error of 0.002

<table>
<thead>
<tr>
<th>Expt. values</th>
<th>Model values</th>
<th>Expt. values</th>
<th>Model values</th>
<th>Expt. values</th>
<th>Model values</th>
<th>Expt. values</th>
<th>Model values</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0029</td>
<td>0.1159</td>
<td>0.8206</td>
<td>0.8991</td>
<td>0.1158</td>
<td>0.1139</td>
<td>0.9946</td>
<td>0.9476</td>
</tr>
<tr>
<td>0.0130</td>
<td>0.1138</td>
<td>0.6635</td>
<td>0.5256</td>
<td>0.1564</td>
<td>0.1138</td>
<td>0.9768</td>
<td>0.9525</td>
</tr>
<tr>
<td>0.0239</td>
<td>0.1187</td>
<td>0.1321</td>
<td>0.1703</td>
<td>0.1904</td>
<td>0.1138</td>
<td>0.9573</td>
<td>0.9531</td>
</tr>
<tr>
<td>0.0442</td>
<td>0.1188</td>
<td>0.1614</td>
<td>0.1192</td>
<td>0.2200</td>
<td>0.1138</td>
<td>0.9174</td>
<td>0.9489</td>
</tr>
<tr>
<td>0.9804</td>
<td>0.2214</td>
<td>0.8561</td>
<td>0.9389</td>
<td>0.0778</td>
<td>0.1189</td>
<td>0.9965</td>
<td>0.9585</td>
</tr>
<tr>
<td>0.0093</td>
<td>0.2903</td>
<td>0.7574</td>
<td>0.8111</td>
<td>0.1109</td>
<td>0.1162</td>
<td>0.9826</td>
<td>0.9586</td>
</tr>
<tr>
<td>0.0178</td>
<td>0.3757</td>
<td>0.6138</td>
<td>0.3573</td>
<td>0.1415</td>
<td>0.1149</td>
<td>0.9659</td>
<td>0.9577</td>
</tr>
<tr>
<td>0.0344</td>
<td>0.3786</td>
<td>0.1844</td>
<td>0.1557</td>
<td>0.1703</td>
<td>0.1143</td>
<td>0.9322</td>
<td>0.9558</td>
</tr>
<tr>
<td>0.9935</td>
<td>0.9313</td>
<td>0.0021</td>
<td>0.1612</td>
<td>0.0611</td>
<td>0.3844</td>
<td>0.0882</td>
<td>0.1149</td>
</tr>
<tr>
<td>0.9725</td>
<td>0.9429</td>
<td>0.0110</td>
<td>0.1294</td>
<td>0.0983</td>
<td>0.2393</td>
<td>0.1429</td>
<td>0.1139</td>
</tr>
<tr>
<td>0.9528</td>
<td>0.9487</td>
<td>0.0239</td>
<td>0.1187</td>
<td>0.1321</td>
<td>0.1703</td>
<td>0.2084</td>
<td>0.1138</td>
</tr>
<tr>
<td>0.9115</td>
<td>0.9445</td>
<td>0.0522</td>
<td>0.1149</td>
<td>0.1664</td>
<td>0.1327</td>
<td>0.2892</td>
<td>0.1138</td>
</tr>
<tr>
<td>0.9963</td>
<td>0.9582</td>
<td>0.9922</td>
<td>0.8897</td>
<td>0.8338</td>
<td>0.9176</td>
<td>0.0687</td>
<td>0.1677</td>
</tr>
<tr>
<td>0.9821</td>
<td>0.9579</td>
<td>0.9505</td>
<td>0.7136</td>
<td>0.7360</td>
<td>0.7481</td>
<td>0.1063</td>
<td>0.1190</td>
</tr>
<tr>
<td>0.9648</td>
<td>0.9573</td>
<td>0.0178</td>
<td>0.3757</td>
<td>0.6138</td>
<td>0.3573</td>
<td>0.1476</td>
<td>0.1143</td>
</tr>
<tr>
<td>0.9300</td>
<td>0.9550</td>
<td>0.0371</td>
<td>0.1664</td>
<td>0.2901</td>
<td>0.1717</td>
<td>0.1939</td>
<td>0.1138</td>
</tr>
<tr>
<td>0.0825</td>
<td>0.1161</td>
<td>0.9961</td>
<td>0.9575</td>
<td>0.0026</td>
<td>0.1183</td>
<td>0.8059</td>
<td>0.8720</td>
</tr>
<tr>
<td>0.1322</td>
<td>0.1140</td>
<td>0.9790</td>
<td>0.9553</td>
<td>0.0119</td>
<td>0.1210</td>
<td>0.1006</td>
<td>0.3903</td>
</tr>
<tr>
<td>0.1904</td>
<td>0.1138</td>
<td>0.9528</td>
<td>0.9487</td>
<td>0.0221</td>
<td>0.1244</td>
<td>0.1301</td>
<td>0.1391</td>
</tr>
<tr>
<td>0.2604</td>
<td>0.1138</td>
<td>0.8810</td>
<td>0.8975</td>
<td>0.0413</td>
<td>0.1246</td>
<td>0.1626</td>
<td>0.1163</td>
</tr>
<tr>
<td>0.0670</td>
<td>0.2363</td>
<td>0.9968</td>
<td>0.9591</td>
<td>0.9849</td>
<td>0.3543</td>
<td>0.8509</td>
<td>0.9349</td>
</tr>
<tr>
<td>0.1027</td>
<td>0.1251</td>
<td>0.9832</td>
<td>0.9585</td>
<td>0.9377</td>
<td>0.4825</td>
<td>0.7471</td>
<td>0.7819</td>
</tr>
<tr>
<td>0.1415</td>
<td>0.1149</td>
<td>0.9648</td>
<td>0.9573</td>
<td>0.8958</td>
<td>0.6043</td>
<td>0.5887</td>
<td>0.3224</td>
</tr>
<tr>
<td>0.1842</td>
<td>0.1139</td>
<td>0.9523</td>
<td>0.9532</td>
<td>0.0336</td>
<td>0.6078</td>
<td>0.1779</td>
<td>0.1494</td>
</tr>
</tbody>
</table>

Table 1  Experimental Vs Model Values for one number of Neurons with an Error of 0.002
Target Error = 0.002   No of Neurons = 10
SSE = 0.0182752.

TRAINLM: Network error did not reach the error goal. Further training may be necessary, or try different initial weights and biases and/or more hidden neurons.
Fig 6 Experimental Vs. Model Predicted for testing Data points for 10 number of neurons with an error of 0.002

Fig 7 Experimental Data points Vs Model Predictions (All data sets) for 10 numbers of neurons with an error of 0.002
Table 2 Experimental vs. Model Values for 10 number of Neurons with an Error of 0.002

<table>
<thead>
<tr>
<th>Number of Neurons</th>
<th>Sum-squared error (SSE)</th>
<th>Is it a desired SSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.92067</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>0.0182752</td>
<td>No</td>
</tr>
<tr>
<td>20</td>
<td>0.0178607</td>
<td>No</td>
</tr>
<tr>
<td>30</td>
<td>0.0218734</td>
<td>No</td>
</tr>
<tr>
<td>40</td>
<td>0.0019945</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 3 SSE values for different number of neurons with target goal of 0.002

As shown in table 2 the predicted SSE is matches with desired target value for forty (40) number of neurons which is the optimum number of neurons for the designed ANN for equation of state model.

CONCLUSION:
In the present work, as the functional relationship is extremely complex an ANN can be suggested as best method to solve with regard to its accuracy. We suggest an ANN with 37 hidden neurons, 2 input neurons and one output neuron to find compressibility factor from given $T_r$ and $P_r$ using Levenberg Marquardt second order optimization technique. The network proposed with above parameters is found to be very accurate and better than R-K or Vander Waal two parameter model.

As shown in table 2 the predicted SSE is matches with desired target value for forty (40) number of neurons which is the optimum number of neurons for the designed ANN for equation of state model.
As shown in table 3 the predicted SSE is maches with desired target value for forty five (45) numbers of neurons which is the optimum number of neurons for the designed ANN for equation of state model.

REFERENCES


