

**NEW MODELS FOR ESTIMATING THE DEW-POINT PRESSURE OF GAS  
CONDENSATE RESERVOIRS**

BY

**MOHAMMED ABDULLAH AL-DHAMEN**

A Thesis Presented to the  
DEANSHIP OF GRADUATE STUDIES

**KING FAHD UNIVERSITY OF PETROLEUM & MINERALS**

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the  
Requirements for the Degree of

**MASTER OF SCIENCE**

In

**PETROLEUM ENGINEERING**

June 2010

**KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS  
DHAHRAN 31261, SAUDI ARABIA**

**DEANSHIP OF GRADUATE STUDIES**

This thesis, written by **MOHAMMED ABDULLAH AL-DHAMEN** under the direction of his thesis advisor and approval by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN PETROLEUM ENGINEERING**

Thesis Committee

*M. A. Marhoun*

Prof. Muhammad A. Al-Marhoun (Thesis Advisor)

*H. Y. Al-Yousef*

Dr. Hasan Y. Al-Yousef (Member)

*A. A. Al-Majed*

Dr. Abdulaziz A. Al-Majed (Member)

*A. A. Al-Majed*

Dr. Abdulaziz A. Al-Majed  
(Department Chairman)

*[Signature]*

Dr. Salam Zummo  
(Dean of Graduate Studies)



*13/12/10*

Date

## **ACKNOWLEDGEMENT**

All praise is due to Allah the most beneficent, the most compassionate, and peace is upon his prophet Mohammed. I wish to express my sincere thanks to my main advisor professor Mohammed Al-Marhoun for providing me full support during the entire research period and giving up some of his private time to meet with me during the weekends. My deep thanks go to Dr. Hasan Al-Yousef and Dr. Abdulaziz Al-Majed for serving as committee members.

## List of Contents

<b>Acknowledgment</b> .....	III
<b>List of Content</b> .....	IV
<b>List of figures</b> .....	VII
<b>List of tables</b> .....	VIII
<b>Abstract In Arabic</b> .....	IX
<b>Abstract In English</b> .....	X
<b>CHAPTER 1: Introduction</b> .....	1
<b>CHAPTER 2: Literature Review</b> .....	4
2.1 Empirical Correlations.....	4
2.2 Artificial Neural network Models.....	6
<b>CHAPTER 3: Statement of the Problem and Objectives</b> .....	8
3.1 Statement of the Problem .....	8
3.2 Objectives.....	9
<b>CHAPTER 4: Correlations and Regression Theory</b> .....	10
4.1 Regression Theory.....	10
4.2 Nonlinear Multiple Regression.....	10
4.3 Linear Multiple Regression.....	12
4.4 Nonparametric Regression Model (ACE Technique).....	13
<b>CHAPTER 5: Neural Networks</b> .....	14
5.1 The Use of Artificial Neural Networks in Petroleum Industry.....	14
5.2 Artificial Intelligence .....	16
5.3 Artificial Neural Network .....	17
5.3.1 Historical Background.....	17

5.3.2 Definition .....	18
5.3.3 Brain system .....	20
5.4 Fundamentals .....	20
5.4.1 Network Learning.....	23
5.4.2 Network Architecture .....	23
5.4.2.1 Feed forward networks.....	25
5.4.2.2 Recurrent networks.....	25
5.4.3 General Network Optimization .....	28
5.4.4 Activation Functions.....	32
5.5 Back-Propagation Training Algorithm.....	34
5.5.1 Generalized Delta Rule.....	37
5.5.1.1 Update of Output-Layer Weights.....	38
5.5.1.2 Output Function.....	40
5.5.1.3 Update of Hidden-Layer Weights.....	41
5.5.2 Stopping Criteria.....	42
<b>CHAPTER 6: Error Analysis.....</b>	<b>43</b>
6.1 Statistical Error Analysis.....	43
6.2 Graphical Error Analysis.....	45
6.3 Trend Analysis.....	46
<b>CHAPTER 7: Development of New Models.....</b>	<b>47</b>
7.1 Traditional Correlation Model.....	47
7.2 Nonparametric Model (ACE).....	48
7.3 Artificial Neural network Model.....	51

7.3.1 Artificial Neural network Model in Matrix Form.....	52
<b>CHAPTER 8: Results and Discussions</b> .....	<b>55</b>
8.1 Published Correlations Evaluation .....	55
8.2 New Models Evaluation .....	56
8.2.1 Traditional Correlation Model .....	57
8.2.2 Nonparametric Approach (ACE).....	57
8.2.3 Artificial Neural network Model .....	57
<b>CHAPTRE 9: Conclusions</b> .....	<b>79</b>
References.....	81
Appendix A.....	83
Appendix B.....	84
Vitae.....	87

## List of Figures

Figure No.	Figure Name	Page No.
1.1	Phase Diagram of Typical Gas Condensate .....	3
5.1	Major Structure of Biologic Nerve Cell (after Freeman).....	21
5.2	Artificial Neuron (after Freeman).....	22
5.3	Supervised Learning Model.....	24
5.4	Fully Connected Network with Two Hidden Layers and One Output Layer.....	27
5.5	Jordan Recurrent Network.....	30
5.6	Elman Recurrent Network.....	31
5.7	Activation Functions.....	36
7.1	Optimal Transform of Reservoir Temperature.....	49
7.2	Optimal Transform of Gas-Oil Ratio.....	49
7.3	Optimal Transform of Gas Specific Gravity.....	50
7.4	Optimal Transform of Condensate Specific Gravity.....	50
8.1	Cross Plot (Nemeth and Kennedy).....	61
8.2	Cross Plot (Elsharkawy).....	62
8.3	Cross Plot (Homud and Al-Marhoun).....	63
8.4	Cross Plot (Marruffo, Maita, Him and Rojas).....	64
8.5	Cross Plot (New Correlation).....	65
8.6	Cross Plot (ACE model).....	66
8.7	Cross Plot (New Artificial Neural Network Model).....	67
8.8	Error Distrbution (Nemeth and Kennedy).....	68
8.9	Error Distrbution (Elsharkawy).....	69
8.10	Error Distrbution (Homud and Al-Marhoun).....	70
8.11	Error Distrbution (Marruffo, Maita, Him and Rojas).....	71
8.12	Error Distrbution (New Correlation).....	72
8.13	Error Distrbution (ACE model).....	73
8.14	Error Distrbution (New Artificial Neural Network model).....	74
8.15	Accuracy of Correlations for Ranges of Dew-Point Pressures.....	75
8.16	Sensitivity of New Models to Reservoir Temperture.....	76
8.17	Sensitivity of ACE Model to Reservoir Temperture.....	77
8.18	Sensitivity of New Models to Gas-Oil Ratio.....	78

## List of Tables

<b>Table No.</b>	<b>Table Name</b>	<b>Page No.</b>
8.1	Error Statistics with New Coefficients.....	59
8.2	Error Statistics with Original Coefficients.....	60



## خلاصة الرسالة

اسم الطالب كامل: محمد عبدالله الضامن

عنوان الرسالة: علاقات رياضية جديدة لحساب ضغط التكثف لمكامن الغاز الطبيعي

التخصص: هندسة بترول

تاريخ التخرج : يونيو 2010 م

علاقات رياضية جديدة و بأساليب مختلفة تم تطويرها لحساب ضغط التكثف لمكامن الغاز. العلاقات الرياضية التقليدية و العلاقات ذات العناصر الغير محددة، بالإضافة الى الشبكات العصبية تم استخدامها في هذه الدراسة . هذه العلاقات الجديدة تعتمد على مجموعة معلومات يمكن الحصول عليها بسهولة ( درجة حرارة المكامن ، الكثافة النوعية للغاز، الكثافة النوعية للزيت المتكثف و نسبة الغاز للزيت ). لقد تم استخدام في هذه الدراسة ما مجموعه 113 مجموعة معلومات من حقول غاز مختلفة في الشرق الأوسط. و لقد استخدم التحليل الاحصائي لتقييم العلاقات المنشورة و ايضاً التي تم تطويرها في هذه الرسالة. و لقد برهنت النتائج الإحصائية ان العلاقة الرياضية التي تم اشتقاقها من الشبكات العصبية تفوق في الدقة و الأداء العلاقات الرياضية الأخرى.

درجة ماجستير العلوم

جامعة الملك فهد للبترول و المعادن

الظهران-المملكة العربية السعودية

يونيو 2010 م

## **Thesis Abstract**

**Full Name of Student: Mohammed Abdullah Al-Dhamen**

**Title of Study: NEW MODELS FOR ESTIMATING THE DEW-POINT PRESSURE OF GAS CONDENSATE RESERVOIRS**

**Major Field: Petroleum Engineering**

**Date of Degree: June 2010**

New Models with three different techniques have been developed to predict the dew-point pressure for gas condensate reservoirs. Traditional correlations, non-parametric approaches and artificial neural networks have been utilized in this study. The new models are functions of easily obtained parameters (reservoir temperature, gas specific gravity, condensate specific gravity and gas-oil ratio). A total number of 113 data sets obtained from Constant Mass Expansion (CME) experiment were collected from the Middle East fields have been used in developing the models. The data used for developing the models covers a reservoir temperature from 100 to 309 °F, gas oil ratios from 103,536 to 3,321 SCF/STB, gas specific gravity from 0.64 to 0.82 and condensate specific gravity from 0.73 to 0.81. Graphical and statistical tools have been utilized for the sake of comparing the performance of the new models and other empirical models. The results showed that artificial neural network developed in this study has the best results among all other models.

**MASTER OF SCIENCE DEGREE  
KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS  
Dhahran-Saudi Arabia  
Date: June 2010**

## CHAPTER 1

### INTRODUCTION

In reservoir engineering a variety of data is needed to accurately estimate reserves and forecast production. Field characterization consists of reservoir rock analysis and fluid analysis. The determination of gas condensate dew-point pressure is essential for fluid characterization, gas reservoir performance calculations, and for the design of production systems.

The phase diagram of a condensate gas is somewhat smaller than that for oils, and critical point is further down the left side of the envelope. These changes are a result of condensate gases containing fewer of the heavy hydrocarbons than do the oils. The phase diagram of a gas condensate has a critical temperature less than the reservoir temperature and a cricondenthem greater than the reservoir temperature (Figure 1.1). Initially, the gas condensate is totally gas in the reservoir, point 1. As reservoir pressure decreases, the gas condensate exhibits a dew-point, point 2. The dew-point of a gas condensate fluid occurs when a gas mixture containing heavy hydrocarbons is depressurized until liquid is formed, that is, a substantial amount of gas phase exists in equilibrium with an infinitesimal amount of liquid phase. As pressure is reduced, liquid condenses from the

gas to form a free liquid in the reservoir. Normally, there is no effective permeability to the liquid phase and it is not produced.

Traditionally, the dew-point pressure of gas condensate is experimentally determined in a laboratory in a process called constant mass expansion (CME) test using a visual window-type PVT cell. Another study is constant volume depletion test (CVD) which verifies the thermodynamic equilibrium at each pressure depletion level, and describes the change of composition of the reservoir gas with every decreasing pressure step.

The present study focuses on prediction of the dew-point pressure for gas condensate reservoir. Three different approaches will be used to predict the dew-point pressure; traditional correlations, non-parametric approach and artificial neural networks.

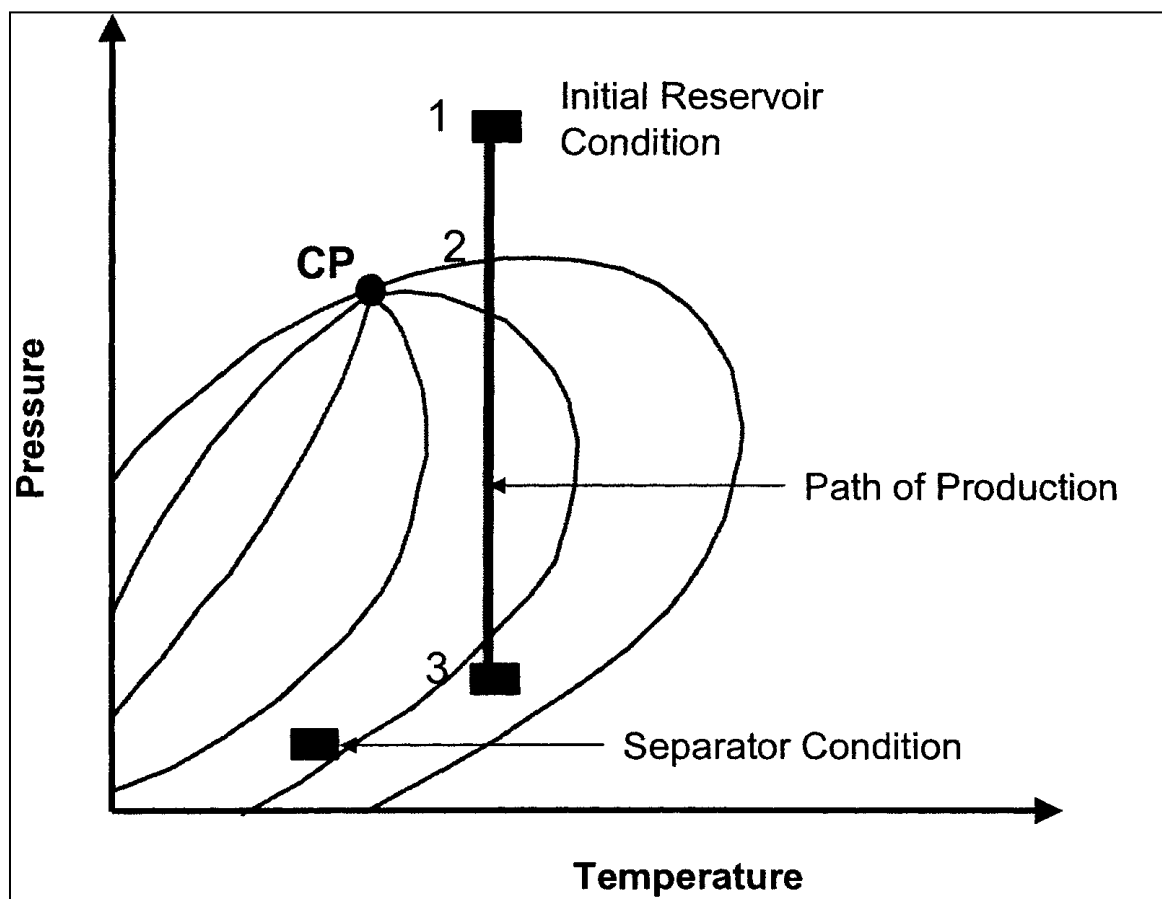


Figure 1.1: Phase Diagram of Typical Gas Condensate

## CHAPTER 2

### LITERATURE REVIEW

This chapter provides a review of the most commonly used correlations and Artificial Neural Network models that are being used to estimate the dewpoint pressure. The first section presents the most commonly correlations while the second section presents the artificial neural network models.

#### *2.1 Empirical Correlations*

In 1947, Sage and Olds studied experimentally the behavior of five paired samples of oil and gas obtained from wells in San Joaquin fields in California. Their investigations resulted in developing a rough correlation relating the retrograde dew-point pressure to the gas-oil ratio, temperature and stocktank API oil gravity. The results of this correlation were presented in tabulated and graphical forms. This correlation is applicable only for gas-oil ratio of 15,000-40,000 SCF/STB, for temperature of 100-220°F, and for API oil gravity of 52°-64°.

In 1952, Organick and Golding presented a correlation to predict saturation pressures, which could be a dew-point or a bubble point pressure, for gas-condensate and

volatile oil reservoir fluids. Saturation pressure is related directly to the chemical composition of the mixtures with the aid of two generalized composition characteristics: (1) the molal average boiling point (  $B$  ) in  $^{\circ}R$ , and (2) the modified average equivalent molecular weight ( $W_m$ ). These parameters can be calculated from the composition of the gas mixture. The correlation was given in the form of 14 working charts, and on each chart the saturation pressure is plotted against temperature. Each chart is for a specific value of  $W_m$  and gives a set of curves representing different values of  $B$ .

In 1967, Nemeth and Kennedy developed a correlation in the form of an equation, which relates the dewpoint pressure of a gas-condensate fluid to its chemical composition, temperature and characteristics of  $C7+$ . The final form of the equation contains eleven constants; See the Appendix. The dewpoint pressure and temperature ranges varied from 1,270- 10,790 psia, and 40-320 $^{\circ}F$  respectively.

In 1996, Potsch and Braeuer presented a graphical method for determining the dewpoint pressure as a backup for the laboratory visual reading of dewpoint pressure during a CME test. The key idea of this method is to plot the number of moles, calculated as a function of single-phase compressibility factor ( $Z$ -factor), versus pressure. Above dewpoint pressure, the plot yields a straight line, and below dewpoint pressure the plot shows a curve. The point of intersection marks the dewpoint pressure.

In 2001 Humoud and Al-Marhoun developed new model using 74 experimental data points relates the dewpoint pressure to the reservoir temperature, primary separator pressure and temperature, gas specific gravity, heptanes plus specific gravity, Gas-Oil ratio and pseudoreduced pressure and temperature

In 2001, Elsharkawy presented a new empirical model to estimate dewpoint pressure for gas condensate reservoirs using experimental data from 340 gas condensate samples covering a wide range of gas, properties and reservoir temperature. Elsharkawy's empirical model contains 19 terms. It correlates dewpoint pressure with reservoir temperature, gas condensate composition as mole fraction and molecular weight and specific gravity of C7+.

In 2002, Marruffo, Maita, Him and Rojas developed a model to estimate the dewpoint pressure. The model correlates dewpoint pressure to Gas-Condensate ratio, C7+ content as mole fraction and reservoir temperature. Also, other models were developed to estimate C7+ content from Gas-Condensate ratio and specific separator gas gravity.

## ***2.2 Artificial Neural network***

In 2003, Barrufet, Gonzalez and Startzman developed an artificial neural network model to estimate the dewpoint pressure. The hydrocarbon and non-hydrocarbon gas condensate composition (C1 - C7+, N, CO<sub>2</sub>, H<sub>2</sub>S), reservoir temperature, molecular weight and specific gravity of C7+ are used as an input to feed the neural network. The neural network architecture consists of three layers; one input layer with 13 neurons, one hidden layer with 6 neurons and one output layer with one neuron. The backpropagation technique and the conjugate gradient decent training algorithm are used to minimize the mean-square error.

In 2007, Akbari, Farahani and Yasser Abdy developed an artificial neural network model to estimate the dew-point pressure. The hydrocarbon and non-hydrocarbon gas



condensate composition (C1 - C7+, N, CO<sub>2</sub>, H<sub>2</sub>S), reservoir temperature, molecular weight of C7+ are used as an input to feed the neural network. The neural network architecture consists of three layers; one input layer with 14 neurons, one hidden layer with 8 neurons and one output layer with one neuron. The backpropagation technique and the Levenberg-Marquardt training algorithm are used to minimize the mean-square error.

## CHAPTER 3

### STATEMENT OF THE PROBLEM AND OBJECTIVES

This chapter describes the problem of predicting dew-point pressure for gas condensate reservoir. The need for developing a model that can overcome the previous difficulties faced in utilizing empirical correlations is addressed through stating the objectives of this work.

#### *3.1 Statement of the Problem*

The need of accurate prediction of the dew-point pressure is very essential for fluid characterization, gas reservoir performance calculations, and for the design of production systems. Also, it is important in avoiding unnecessary stimulation jobs. When a well starts flowing below the dew-point pressure, condensate dropout accumulates around the wellbore. This phenomenon is known as condensate banking and it causes a severe decline in gas production. It is very important to know the causes of the production decline of a gas well; whether it is due to formation damage or condensate banking to make the right course of action.

The laboratory measurements of gas condensate properties provide the most accurate and reliable determination of reservoir fluid properties. However, due to

economical and technical reasons, quite often this information cannot be obtained from laboratory measurements. The experimental determination of these properties requires a representative sample of the reservoir gas with a sufficient volume to complete the analysis, which sometimes is difficult to obtain. The measurements are relatively time consuming, expensive and sometimes subjected to errors. Thus, there is a need for simple accurate method of predicting the dew-point pressure for gas condensate reservoir.

Numerous attempts have been tried to predict the dew-point pressure using correlation and artificial neural network. However, most of these models are utilizing the gas composition and C7+ properties.

In this study, new models have been developed for predicting the dew-point pressure. Some of the best models are reviewed carefully using graphical and statistical analysis. These models are compared against the generated artificial neural network model.

### ***3.2 Objectives***

One of the objectives of this work is to evaluate the most commonly used models to estimate the dew-point pressure of condensate gas. Another objective is to develop new models utilizing the three approaches; traditional correlation, non-parametric approach and artificial neural networks to predict the dew-point pressure as a function of easily obtained parameters such as gas-oil ratio, reservoir temperature, gas specific gravity and heptanes plus gravity. Two types of analysis will be carried out to achieve the objectives: Error analysis and Graphical analyses.

## CHAPTER 4

### CORRELATIONS AND REGRESSION THEORY

Correlation refers to the degree of association between one variable and another or several others. Regression deals with the nature of the relation between these variables. In evaluating the degree of regression, all the error or imprecision is assumed to be in the measurement of one variable called the “dependent”, while the other variables are assumed to be precisely known. These precise variables are called the “independent” variables.

#### *4.1 Regression Theory*

The basic concept of regression analysis is to produce a linear or nonlinear combination of independent variables that will correlate as closely as possible with dependent variable.

#### *4.2 Linear Multiple Regression*

Consider a set of observation of size  $n_d$  on which the properties  $y, x_1, x_2, x_3, x_4, \dots, x_n$  are measured. The  $x$ 's and  $y$  are the independent and dependent variables, respectively. The linear regression equation will then be written as follows:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n \dots\dots\dots (4.1)$$

which represents a hyperplane in ( n + 1) dimensional space. Equation (4.1) can be written for any observation point *i* as:

$$y = a_0 + a_1x_{i1} + a_2x_{i2} + \dots + a_nx_{in} ; i = 1, n_d \dots\dots\dots (4.2)$$

The *n<sub>d</sub>* equations for the *n<sub>d</sub>* experimental measurements can be expressed in matrix form as:

$$\begin{array}{c|c|c|c}
 \mathbf{1} & \mathbf{x}_{11} & \mathbf{x}_{12} & \dots & \mathbf{x}_{1n} & \mathbf{a}_0 & \mathbf{y}_1 \\
 \mathbf{1} & \mathbf{x}_{21} & \mathbf{x}_{22} & \dots & \mathbf{x}_{2n} & \mathbf{a}_1 & \mathbf{y}_2 \\
 \mathbf{1} & \mathbf{x}_{31} & \mathbf{x}_{32} & \dots & \mathbf{x}_{3n} & \mathbf{a}_2 & \mathbf{y}_3 \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \mathbf{1} & \mathbf{x}_{nd1} & \mathbf{x}_{nd2} & \dots & \mathbf{x}_{ndn} & \mathbf{a}_n & \mathbf{y}_{nd}
 \end{array}
 \dots\dots\dots (4.3)$$

or in simpler form

$$X\vec{a} = \vec{y} \dots\dots\dots (4.4)$$

Where

*X* = *n<sub>d</sub>* × (*n* + 1) matrix

*a* = (*n* + 1) vector

*y* = *n<sub>d</sub>* vector

$n$  = total number of independent variables

Therefore, the objective is to solve for the vector  $\vec{a}$  for which  $X\vec{a}$  is as close as possible to vector  $y$  since the exact solution cannot be found. Such a vector is that least-squares solution. The unique least-square solution to this system presented in equation (4.4) is:

$$\hat{a} = (X^T X)^{-1} X^T \vec{y} \dots\dots\dots (4.5)$$

where  $\hat{a}$  is the least-square solution to the system  $X\vec{a} = \vec{y}$  and  $X^T$  is the transpose of the matrix  $X$ .

**4.3 Nonlinear Multiple Regression**

Although most of the dewpoint pressure correlations are nonlinear equations; however, they can be modified slightly to give a form of multiple linear equation. The equation is as follows:

$$S = a'_0 q^{a_1} p^{a_2} \gamma^{a_3} \dots\dots\dots (4.6)$$

and therefore,

$$\log(S) = \log(a'_0) + a_1 \log(q) + a_2 \log(p) + a_3 \log(\gamma) \dots\dots\dots (4.7)$$

and therefore,

$$y = a_0 + a_1 x_1 + a_2 x_2 + \dots\dots\dots + a_n x_n \dots\dots\dots (4.8)$$

where:

$$y = \log(S)$$

$$a_0 = \log(a'_0)$$

$$x_1 = \log(q)$$

$$x_2 = \log(p)$$

$$x_3 = \log(\gamma)$$

Equation (4.8) can be solved by the method of multiple-linear regression, as outlined earlier.

#### ***4.4 Nonparametric Regression Model (ACE Technique)***

The ACE (Alternating Conditional Expectations) algorithm, originally proposed by Breiman and Freiman et al. (1985), provides a methods for estimating optimal transformation for multiple regression that results in a maximum correlation between a dependent variable  $y$  and multiple independent variables  $x_1, x_2, \dots, x_m$ .

A model predicting the value of  $y$  from the values of  $x_1, x_2, \dots, x_m$  is written in genetic form

$$y = f^{-1}(z) \text{ where } z = \sum_{n=1}^m z_n \text{ and } z_n = f_n(x_n) \dots\dots\dots(4.9)$$

The functions  $f_1(\cdot), f_2(\cdot), \dots, f_m(\cdot)$  are called variable transformations yielding the transformed independent variables  $z_1, z_2, \dots, z_m$ . The function  $f_1(\cdot)$  is the transformation for the dependent variable. In fact the main interest is its inverse:  $f^{-1}(\cdot)$ , yielding the dependent variable  $y$  from the transformed dependent variable  $z$ . Given  $N$  observation points can help to find the transformation functions  $f_1(\cdot), f_2(\cdot), \dots, f_m(\cdot)$ . The method of ACE constructs and modifies the individual transformations to achieve maximum correlation in the transformed space.

## CHAPTER 5

### ARTIFICIAL NEURAL NETWORK

This chapter deals with addressing the concept of artificial neural networks. First, the applications of ANN in petroleum industry will be presented. After that, historical background will be introduced, then, the fundamentals of ANN along with a deep insight to the mathematical representation of the developed model and the network optimization and configuration will be also discussed in details. The relationship between the mathematical and biological neuron is also explained. Finally, the chapter concludes with presenting the robust learning algorithm used in the training process.

#### *5.1 The Use of Artificial Neural Networks in Petroleum Industry*

Within recent years there has been a steady increase in the application of neural network modeling in engineering. ANNs have been used to address some of the fundamental problems, as well as specific ones that conventional computing has been unable to solve, in particular when engineering data for design, interpretations, and calculations have been less than adequate. Also, with the recent strong advances in pattern recognition, classification of noisy data, nonlinear feature detection, market forecasting, and process modeling, neural network technology is very well suited for



solving problems in the petroleum industry. Within the last decade, works have been published covering the successful and potential application of ANNs in many different areas of the geosciences. For example, Ali (1994) highlighted the key factors in the design or selection of neural networks and the limitations of the frequent used ANN models. Kumoluyi and Daltaban (1994) presented a general overview of pattern recognition and a special case of conventional feed-forward or back-propagation networks. Romeo et al. (1995) used a simple multilayer perception with 23 neurons to identify seismic data. Miller et al. (1995) outlined the use of ANNs in classification of remote sensing data. Fletcher et al. (1995) presented models that can predict oil-well cement properties using an artificial neural network approach. Trained with diffuse reflectance Fourier Transform spectra of different cements, the proposed ANN models successfully correlated particle size distributions and cement-thickening time with reasonable accuracy. Vukelic and Miranda (1996) presented a case study of the development of a neural network that would decide if a reservoir would produce gas, liquid or nothing. Another implementation of ANNs, presented by Mohaghegh et al. (1994), was the characterization of reservoir heterogeneity. The ANN was able to predict rock permeability, porosity, oil, water and gas saturations with accuracies comparable to actual laboratory core measurements. Similarly, Wong et al. (1995) and Zhou et al. (1993) combined separate back-propagation neural networks trained with wireline logs and lithofacies information to give improved predictions porosity and permeability in petroleum reservoirs. Aside from back-propagation ANNs, radialbasis-function (RBF) ANNs were also used to estimate porosity distribution (Wang et al. 1999). In their study,

Wang et al. combined RBF ANNs with kriging techniques to estimate different yet equally probable porosity distributions. Other applications of ANNs in the petroleum industry include papers that employ ANN to pick the proper reservoir model for well testing purposes (AlKaabi and Lee, 1993; Juniardi and Ershaghi, 1993), analyze and classify beam pumping unit dynamometer diagrams (Rgers et al. 1990), identification of flow regime in pipes with band spectra ( van der Spek and Thomas et al. 1998; A.Garrouch et al. 1998; M.Nikravesch et al. 1998; T.Ertekin et al. 2001 and R.A.Startzman et al. 2001) used neural network models for the prediction of the constant volume depletion behavior of gas condensate reservoirs, estimating tight gas permeability, rock properties estimation based on well log, two phase relative permeability estimation and prediction of U.S. natural gas production, respectively. (Al-Marhoun and Osman et al. 2002), presented a neural network model to predict the bubble point pressure and the formation volume factor at the bubble point pressure.

## ***5.2 Artificial Intelligence***

The science of artificial intelligence or what is synonymously known as soft computing shows better performance over the conventional solutions. Sage et al. 1949 defined the aim of artificial intelligence as the development of paradigms or algorithms that require machines to perform tasks that apparently require cognition when performed by humans. This definition is widely broadened to include preceptrons, language, and problems solving as well as conscious, unconscious processes. Many techniques are classified under the name of artificial intelligence such as genetic algorithms, expert

systems, and fuzzy logic because of their ability, one at least, to make certain reasoning, representation, problem solving, and generalization. Artificial neural network is also considered one of the important components of artificial intelligence system.

### ***5.3 Artificial Neural Network***

#### ***5.3.1 Historical Background***

The research carried on neural network can be dated back to early 1940s. Specifically, McCulloch and Pitts et al. 1943 have tried to model the low-level structure of biological brain system. Hebb et al. 1949 published the book entitled “*the organization of behavior*” in which he focused mainly towards an explicit statement of a physiological learning rule for synaptic modification. Also, he reposed that the connectivity of the brain is continually changing as an organism learns differing functional tasks and the neural assemblies are created by such changes. The book was a source of inspiration for the development of computational models of learning and adaptive systems. However, Ashby et al. 1952 published another book entitled “*design for a brain; the origin of adaptive behavior*”. The book focused on the basic notion that the adaptive behavior is not inborn but rather learned. The book emphasized the dynamic aspects of living organism as a machine and the related concepts of stability. While Gabor et al. 1954 proposed the idea of nonlinear adaptive filters. He mentioned that learning was accomplished in these filters through feeding samples of stochastic process into the machine, together with the target function that the machine was expected to produce. After 15 years of McCulloch and

Pitt's paper, a new approach to the pattern recognition problem was introduced by Rosenblatt et al. 1958 through what's called later, *perceptrons*. The latter, at the time when discovered, considered as an ideal achievement and the associative theorem "*perceptron convergence theorem*" was approved by several authors. The perceptron is the simplest form of a neural network that has been used for classifying pattern. This achievement followed by the introduction of *LMS* "least mean square algorithm" and *Adaline* "adaptive linear element" that followed by *Madaline* "multiple-Adaline" in 1962. Minsky and Papert et al. 1969 showed that there are several problems that cannot be solved by the theorem approved by Rosenblatt and therefore countless effort to make such type of improvement will result in nothing. A decade of dormancy in neural network research was witnessed because of the Minsky's paper results. In 1970s, a competition learning algorithm was invented along with incorporation of self organizing maps. Since that time, several networks and learning algorithms were developed. A discovery of backpropagation learning algorithm was one of these fruitful revolutions that developed by Rumelhart *et al.* 1986.

### **5.3.2 Definition**

Generally, ANN is a machine that is designed to model the way in which the brain performs a particular task or function of interest. The system of ANN has received different definitions. A widely accepted term is that adopted by Alexander and Morton et al. 1958: "*A neural network is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use*".

ANN resembles the brain in two aspects; knowledge is acquired by the network through a learning process, and the interneuron connection strengths known as synaptic weights are used to store the knowledge. In other way, neural networks are simply a way of mapping a set of input variables to a set of output variables through a typical learning process. So, it has certain features in common with biological nervous system. The relationship between the two systems and the brain system mechanism is further explained in the next subsection.

### ***5.3.3 Brain system***

Human brain is a highly complex, nonlinear, and parallel information-processing system. It has the capability of organizing biological neurons in a fashion to perform certain tasks. In terms of speed, neurons are five to six orders of magnitude slower than silicon logic gates. However, human brain compensates for this shortcoming by having a massive interconnection between neurons. It is estimated that human brain consists of 10 billion neurons and 60 trillion synapses. These neurons and synapses are expected to grow and increase in both number and connection over the time through learning. Figure 5.1 is a schematic representation of biological nerve cell. The biological neuron is mainly composed of three parts; dendrite, the soma, and the axon. A typical neuron collects signals from others through a host of fine structure (dendrite). The soma integrates its received input (over time and space) and thereafter activates an output depending on the total input. The neuron sends out spikes of electrical activity through a long, thin strand known as an axon, which splits into thousands of branches (tree structure). At the end of

each branch, a synapse converts the activity from the axon into electrical effects that inhibit or excite activity in the connected neurons. Learning occurs by changing the effectiveness of synapses so that the influence of one neuron on another changes. Hence, artificial neuron network, more or less, is an information processing system that can be considered as a rough approximation of the above mentioned biological nerve system. Figure 5.2 shows a typical neuron in an artificial neuron network. This mathematical neuron is a much simpler than the biological one; the integrated information received through input neurons take place only over space. Output from other neurons is multiplied by the corresponding weight of the connection and enters the neuron as an input; therefore, an artificial neuron has many inputs and only one output. All signals in a neural network are typically normalized to operate within certain limit. A neuron can have a threshold level that must be exceeded before any signal is passed. The net input of the activation function may be increased by employing a bias term rather than a threshold; the bias is the negative of threshold. The inputs are summed and therefore applied to the activation function and finally the output is produced.

#### ***5.4 Fundamentals***

In this section, artificial neural network basics will be presented, along with the close relationship between the technology and the biological nervous system. A full mathematical notation of the developed model and the network topology are also provided.

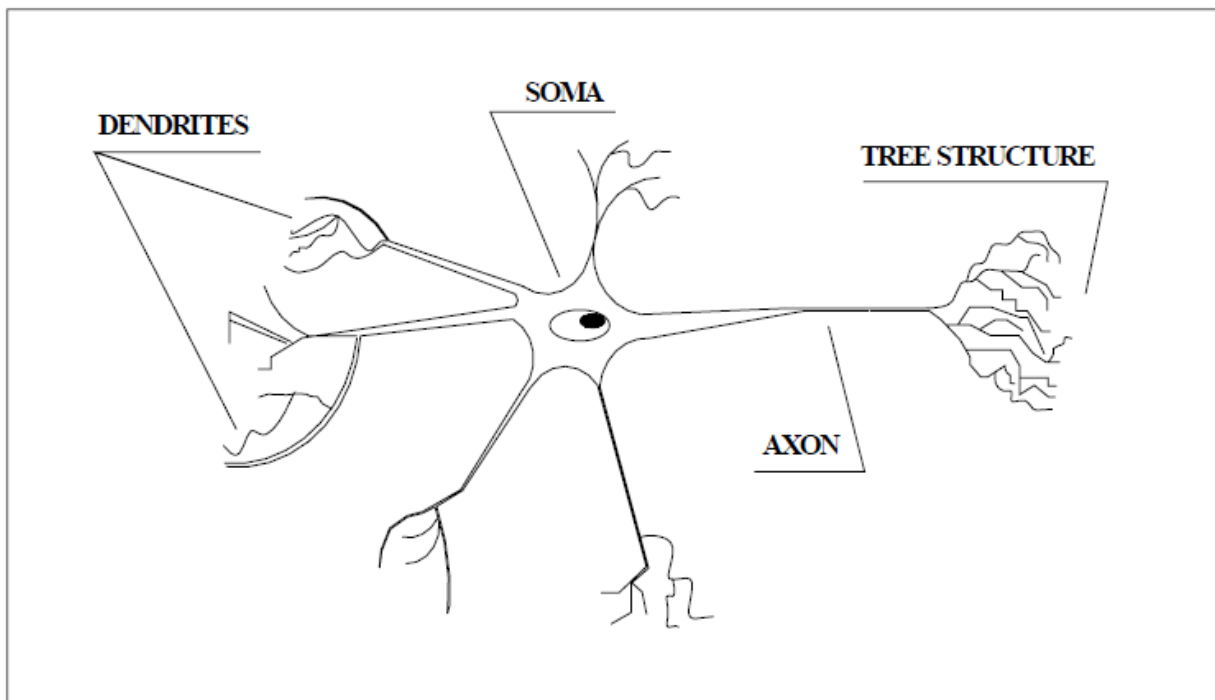


Figure 5.1: Major Structure of Biologic Nerve Cell (after Freeman).

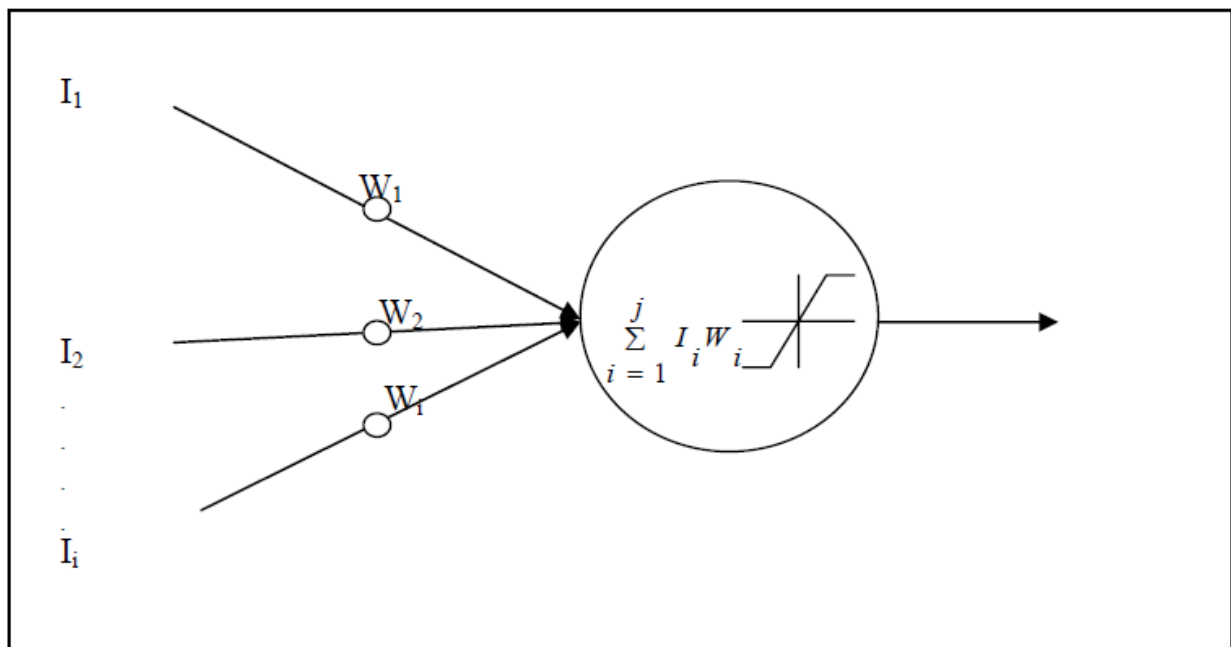


Figure 5.2: Artificial Neuron (after Freeman).



### ***5.4.1 Network Learning***

The network is trained using supervised learning “providing the network with inputs and desired outputs”. The difference between the real outputs and the desired outputs is used by the algorithm to adapt the weights in the network. Figure 5.3 illustrates the supervised learning diagram. The net output is calculated and compared with the actual one, if the error between the desired and actual output is within the desired proximity, there will be no weights' changes; otherwise, the error will be back-propagated to adjust the weights between connections (feed backward cycle). After the weights are fixed the feed forward cycle will be utilized for the test set. The other learning scheme is the unsupervised one where there is no feedback from the environment to indicate if the outputs of the network are correct. The network must discover features, rules, correlations, or classes in the input data by itself. As a matter of fact, for most kinds of unsupervised learning, the targets are the same as inputs. In other words, unsupervised learning usually performs the same task as an auto-associative network, compressing the information from the inputs.

### ***5.4.2 Network Architecture***

Network topology (architecture) is an important feature in designing a successful network. Typically, neurons are arranged in layers, each layer is responsible for performing a certain task. Based on how interconnections between neurons and layers are; neural network can be divided into two main categories (feed forward and recurrent).

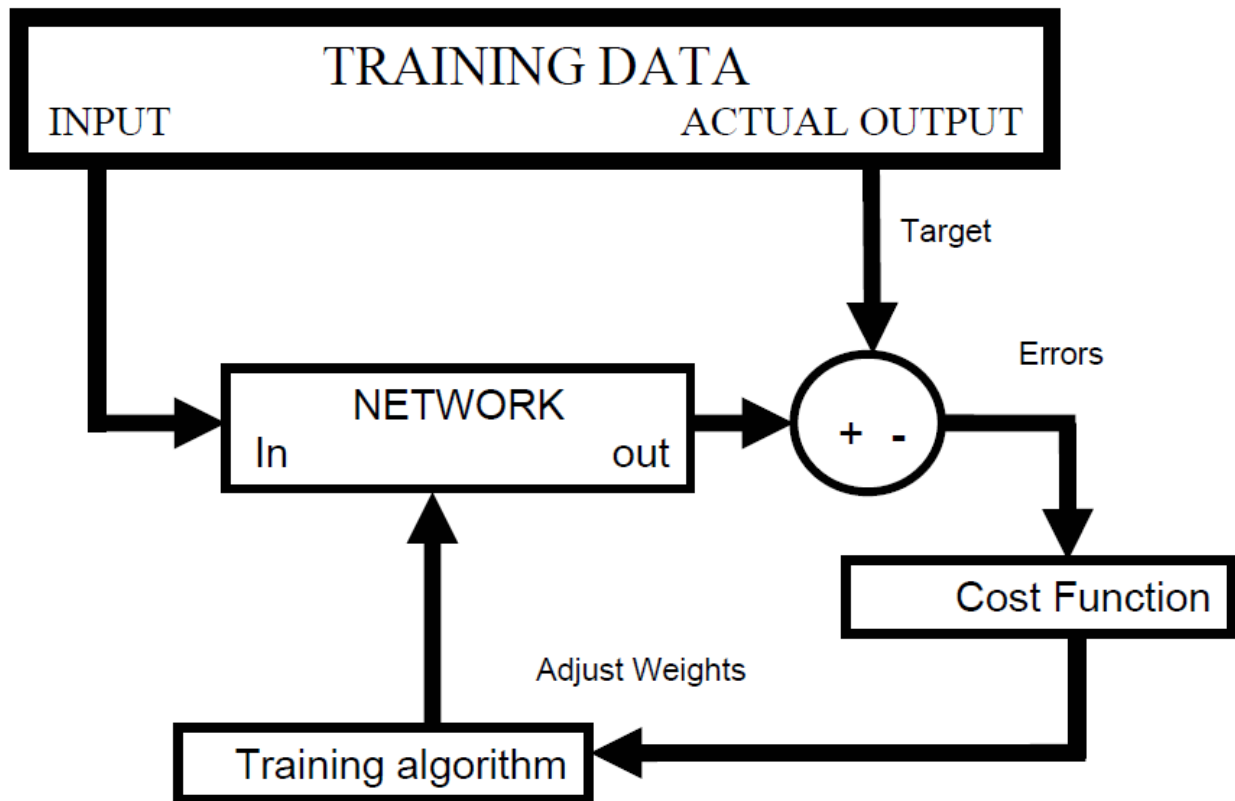


Figure 5.3: Supervised Learning Model.

#### ***5.4.2.1 Feed forward networks***

In these networks the input data sweep directly through hidden layers and finally to the output layer. Hence, it does not allow an internal feedback of information. The essence of connectivity is primarily related to the fact that every node (neuron) in each layer of the network is connected to every other node in the adjacent forward layer. The number of neurons in the input layer should be equivalent to the number of input parameters being presented to the network as input. The same thing is correct for output layer, while the function of hidden layer is to intervene between the external input and the network output. Figure 5.4 is a schematic diagram of a fully connected network with two hidden layer and output layer. The overall response of the network is achieved through the final layer.

#### ***5.4.2.2 Recurrent networks***

Feed-forward networks can be only used for dynamic relationship between input and output variable by including lagged values of input and output variables in the input layer. However, Recurrent Neural Network (RNN) allows for an internal feedback in the system. Internal feedback is a more successful way to account for dynamics in the model. It contains the entire history of inputs as well as outputs. Two types of recurrent neural networks are presented here as examples; Jordan Recurrent Neural Network (JRNN) and Elman Recurrent Neural Network (ERNN). In JRNN, the output feeds back into the hidden layer with a time delay. The output of the previous periods becomes input in the

current period as illustrated in Figure 5.5 Thus, the current period output carries the history of past outputs, which in turn contain past values of inputs.

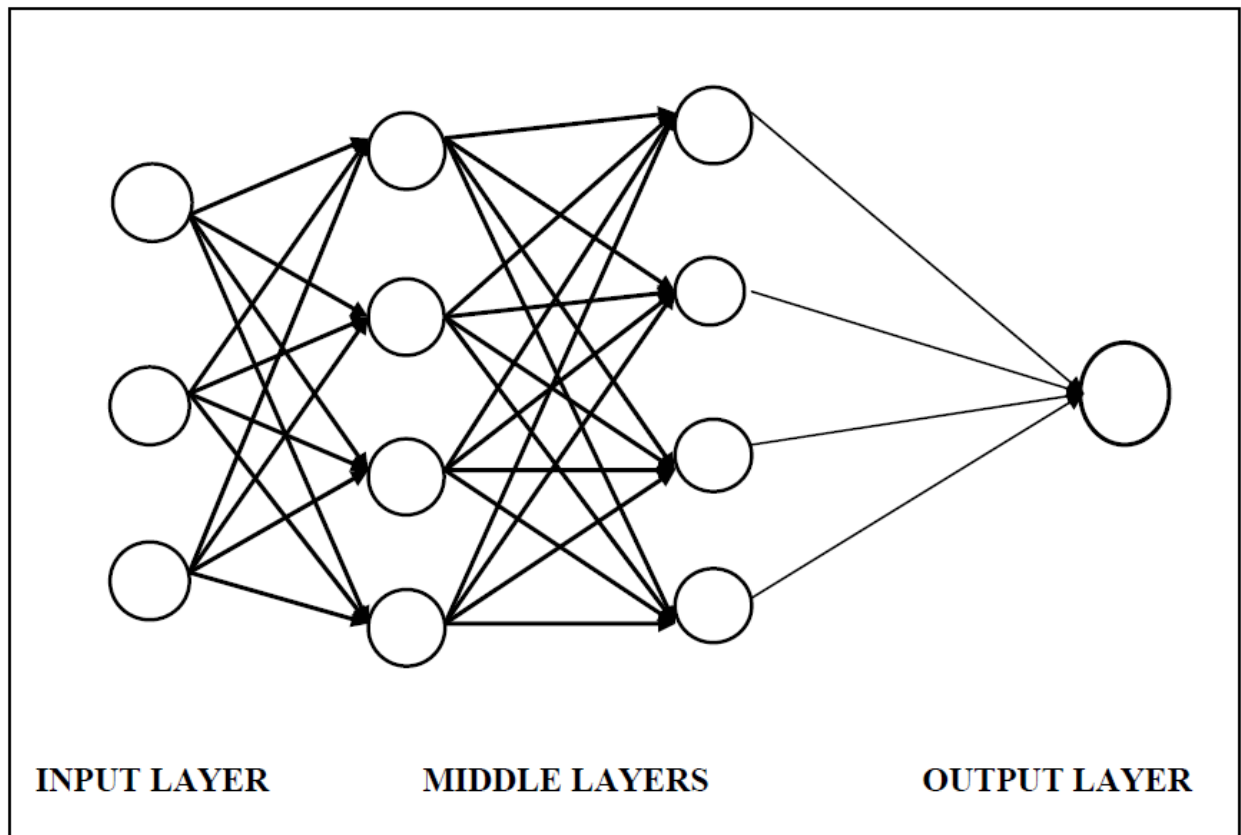


Figure 5.4: Fully Connected Network with Two Hidden Layers and One Output Layer

While a two-layer Elman Recurrent Neural Network (ERNN) is depicted in Figure 5.6. The ERNN accounts for internal feedback in such a way that the hidden layer output feeds back in itself with a time delay before sending signals to the output layer. RNN, however, requires complex computational processes that can only be performed by more powerful software. The back-propagation algorithm is used during the training process in the computation of estimates of parameters.

#### ***5.4.3 General Network Optimization***

Any network should be well optimized in different senses in order to simulate the true physical behavior of the property under study. Certain parameters can be well optimized and rigorously manipulated such as selection of training algorithm, stages, and weight estimation. An unsatisfactory performance of the network can be directly related to an inadequacy of the selected network configuration or when the training algorithm traps in a local minimum or an unsuitable learning set. In designing network configuration, the main concern is the number of hidden layers and neurons in each layer. Unfortunately, there is no sharp rule defining this feature and how it can be estimated. Trial and error procedure remains the available way to do so, while starting with small number of neurons and hidden layers “and monitoring the performance” may help to resolve this problem efficiently. Regarding the training algorithms, many algorithms are subjected to trapping in local minima where they stuck on it unless certain design criteria are modified. The existence of local minima is due to the fact that the error function is the superposition of nonlinear activation functions that may have minima at different points,

which sometimes results in a nonconvex error function. Using randomly initialized weight and inversion of the algorithm may become a solution for this problem. The two most frequent problems that often encountered in network designing are the bad or unrepresentative learning set and overtraining. Therefore, selecting global ratios of data division may resolve it by using 2:1:1 or 3:1:1 or even 4:1:1 as suggested by Haykin. Overtraining refers to the phenomenon when the network starts to model the noise associated with the training data. This phenomenon affects the generalization of network (network is able to accurately generalize when new cases that have not been seen during training are submitted to it). For this reason, cross-validation data are kept aside during training to provide an independent check on the progress of training algorithm. Besides, more confidence is gained where cross-validation data can minimize the error function as training progresses.

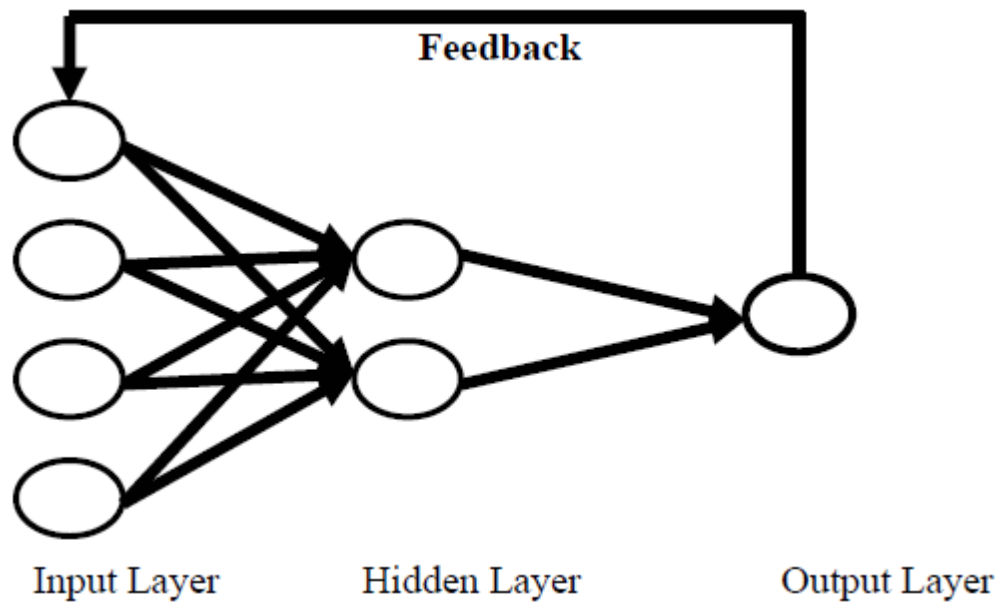


Figure 5.5: Jordan Recurrent Network.



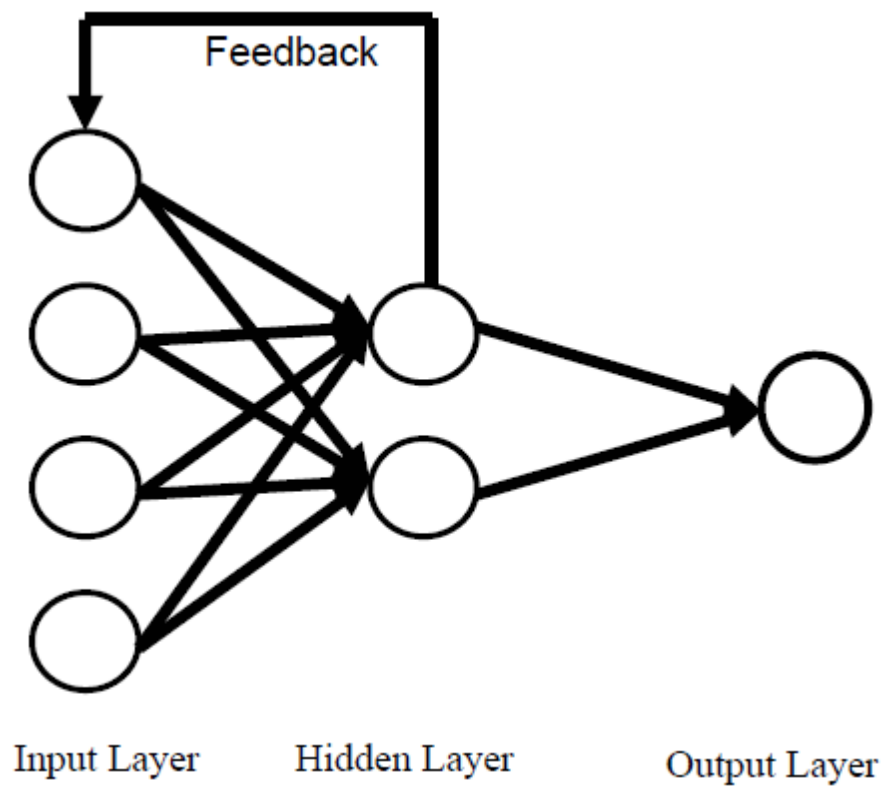


Figure 5.6: Elman Recurrent Network.

#### 5.4.4 Activation Functions

As described earlier, the four basic elements of the neural network model are; synapses (that may receive a signal), adder (for summing up the input signals, weighted by respective synapses), an activation function, and an externally applied threshold. An activation function that limits (the amplitude of) the output of a neuron within a normalized value in a closed interval, say, between [0, 1] or [-1, 1], (see Figure 5.5). The activation function *squashes* the output signal in a 'permissible' (amplitude) range. When a neuron updates it passes the sum of the incoming signals through an activation function, or transfer function (linear or nonlinear). A particular transfer function is chosen to satisfy some specification of the problem that the neuron is attempting to solve. In mathematical terms, a neuron  $j$  has two equations that can be written as follows:

$$NET_{pj} = \sum_{i=1}^N w_{ji}x_{pj} \dots\dots\dots(5.1)$$

and

$$y_{pj} = \varphi(NET - \varphi_{pj}) \dots\dots\dots(5.2)$$

Where;  $x_{p1}, x_{p2}, \dots, x_{pN}$  are the input signals;  $w_{j1}, w_{j2}, \dots, w_{jk}$  are the synaptic weights of neuron  $j$ ;  $NET_{pj}$  is the linear combiner output,  $\varphi_{pj}$  is the threshold,  $\varphi$  is the activation function; and  $y_{pj}$  is the output signal of the neuron.

Four types of activation functions are identified based on their internal features. A simple threshold function has a form of:

$$y_{pj} = k(NET)_{pj} \dots\dots\dots(5.3)$$

Where  $k$  is a constant threshold function, i.e.:

$$y_{pj} = 1 \text{ if } (NET)_{pj} > T$$

$y_{pj} = 0$  otherwise.

T is a constant threshold value, or a function that more accurately simulates the nonlinear transfer characteristics of the biological neuron and permits more general network functions as proposed by McCulloch-Pitts model. However, this function is not widely used because it is not differentiable. The second type of these transfer functions is the Gaussian function, which can be represented as:

$$y_{pj} = ce^{\left[ \frac{-NET_{pj}^2}{\sigma^2} \right]} \dots \dots \dots (5.4)$$

Where:

$\sigma$  is the standard deviation of the function.

The third type is the Sigmoid Function, which is being tried in the present study for its performance. It applies a certain form of squashing or compressing the range of  $(NET)_{pj}$  to a limit that is never exceeded by  $y_{pj}$  this function can be represented mathematically by:

$$y_{pj} = \frac{1}{(1+e^{-a \times NET_{pj}})} \dots \dots \dots (5.5)$$

Where;

$a$  is the slope parameter of the sigmoid function.

By varying the slope parameter, different sigmoid function slopes are obtained. Another commonly used activation function is the hyperbolic function, which has the mathematical form of:

$$y_{pj} = \tan(x) = \left( \frac{1-e^{-NET_{pj}}}{1+e^{-NET_{pj}}} \right) \dots \dots \dots (5.6)$$

This function is symmetrically shaped about the origin and looks like the sigmoid function in shape. However, this function produced good performance when compared to sigmoid function. Hence, it is used as an activation function for the present model. Other functions are presented in Figure 5.7.

### ***5.5 Back-Propagation Training Algorithm***

Is probably the best known, and most widely used learning algorithm for neural networks. It is a gradient based optimization procedure. In this scheme, the network learns a predefined set of input-output sample pairs by using a two-phase propagate-adapt cycle. After the input data are provided as stimulus to the first layer of network unit, it is propagated through each upper layer until an output is generated. The latter, is then compared to the desired output, and an error signal is computed for each output unit. Furthermore, the error signals are transmitted backward from the output layer to each node in the hidden layer that mainly contributes directly to the output.

However, each unit in the hidden layer receives only a portion of the total error signal, based roughly on the relative contribution the unit made to the original output. This process repeats layer by layer, until each node in the network has received an error signal that describes its relative contribution to the total error. Based on the error signal received, connection weights are then updated by each unit to cause the network to converge toward a state that allows all the training set to be prearranged. After training, different nodes learn how to recognize different features within the input space. The way

of updating the weights connections is done through the generalized delta rule "GDR". A full mathematical notion is presented in the next subsection.

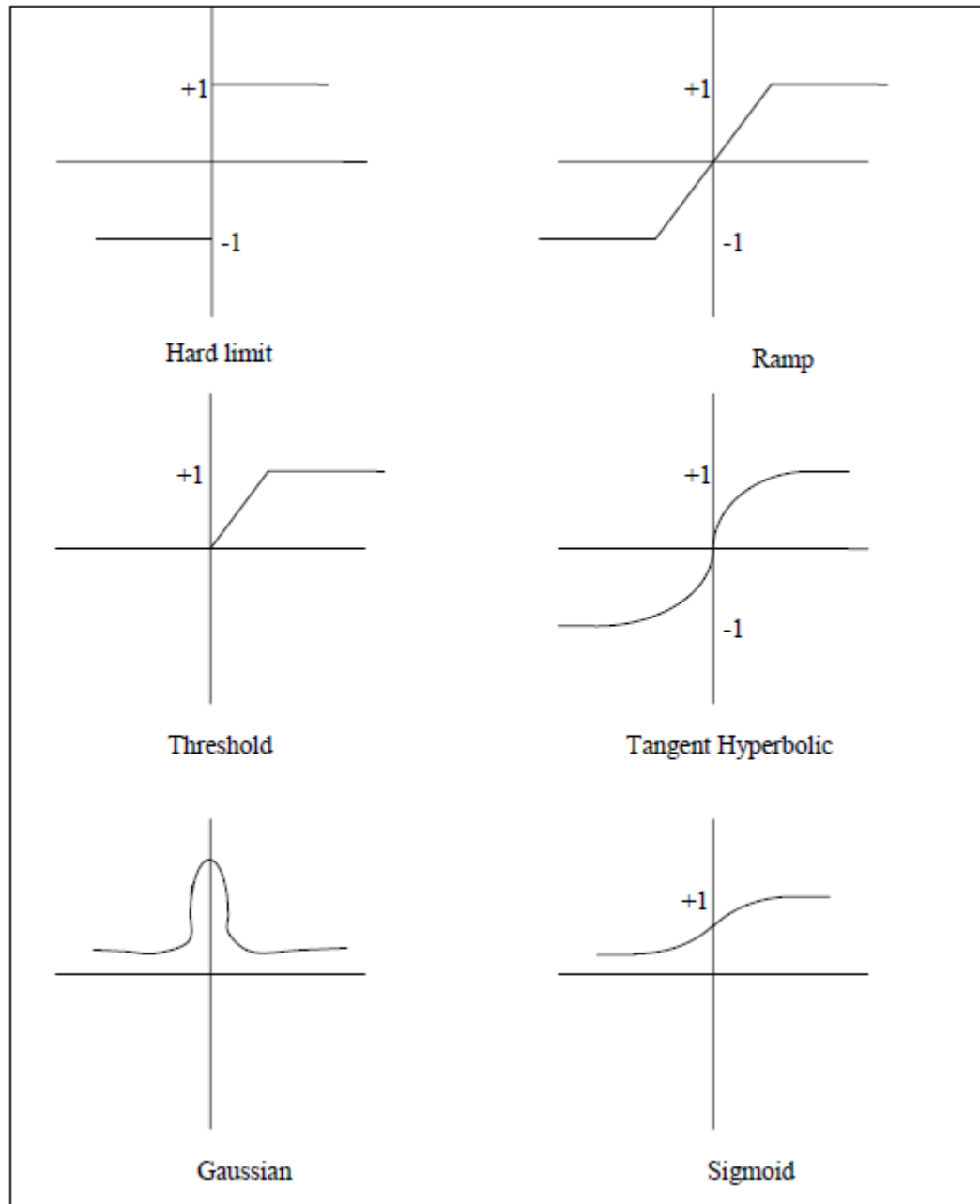


Figure 5.7: Activation Functions

### 5.5.1 Generalized Delta Rule

This section deals with the formal mathematical expression of Back-Propagation Network operation. The learning algorithm, or generalized delta rule, and its derivation will be discussed in details. This derivation is valid for any number of hidden layers. Suppose the network has an input layer that contains an input vector;

$$x_p = (x_{p1}, x_{p2}, x_{p3}, \dots, x_{pN})^t \dots \dots \dots (5.7)$$

The input units distribute the values to the hidden layer units. The net output to the  $j_{th}$  hidden unit is:

$$NET_{pj}^h = \sum_{i=1}^N w_{ji}^h x_{pi} + \theta_j^h \dots \dots \dots (5.8)$$

Where;

$w_{ji}^h$  is the weight of the connection from the  $i_{th}$  input unit, and

$\theta_j^h$  is the bias term

$h$  is a subscript refer to the quantities on the hidden layer.

Assuming that the activation of this node is equal to the net input; then the output of this node is

$$i_{pj} = f_j^h(NET_{pj}^h) \dots \dots \dots (5.8)$$

The equations for the output nodes are:

$$NET_{pk}^o = \sum_{j=1}^L w_{kj}^o i_{pj} + \theta_k^o \dots \dots \dots (5.10)$$

$$o_{pk} = f_k^o(NET_{pk}^o) \dots \dots \dots (5.11)$$

Where:

$o$  superscript refers to quantities of the output layer unit.

The basic procedure for training the network is embodied in the following description:

1. Apply an input vector to the network and calculate the corresponding output values.
2. Compare the actual outputs with the correct outputs and determine a measure of the error.
3. Determine in which direction (+ or -) to change each weight in order to reduce the error.
4. Determine the amount by which to change each weight.
5. Apply the correction to the weights.
6. Repeat steps 1 to 5 with all the training vectors until the error for all vectors in the training set is reduced to an acceptable tolerance.

#### ***5.5.1.1 Update of Output-Layer Weights***

The general error for the  $k^{\text{th}}$  input vector can be defined as;

$$\varepsilon_k = ( d_k - y_k ) \dots \dots \dots (5.12)$$

Where:

$d_k$  = desired output

$y_k$  = actual output

Because the network consists of multiple units in a layer; the error at a single output unit will be defined as

$$\delta_{pk} = ( y_{pk} - o_{pk} ) \dots \dots \dots (5.13)$$

Where;



$p$  subscript refers to the  $p^{\text{th}}$  training vector

$k$  subscript refers to the  $k^{\text{th}}$  output unit

So,

$y_{pk}$  = desired output value from the  $k$ th unit.

$o_{pk}$  = actual output value from the  $k$ th unit.

The error that is minimized by the GDR is the sum of the squares of the errors for all output units;

$$E_p = \frac{1}{2} \sum_{k=1}^M \delta_{pk}^2 \dots\dots\dots(5.14)$$

To determine the direction in which to change the weights, the negative of the gradient of  $E_p$  and  $\nabla E_p$ , with respect to the weights,  $w_{kj}$  should be calculated.

The next step is to adjust the values of weights in such a way that the total error is reduced.

From equation (4.14) and the definition of  $\delta_{pk}$ , each component of  $\nabla E_p$  can be considered separately as follows;

$$E_p = \frac{1}{2} \sum_k (y_{pk} - o_{pk})^2 \dots\dots\dots(5.15)$$

and

$$\frac{\partial E_p}{\partial w_{kj}^o} = -(y_{pk} - o_{pk}) \frac{\partial f_k^o}{\partial (NET_{pk}^o)} \frac{\partial (NET_{pk}^o)}{\partial w_{kj}^o} \dots\dots\dots(5.16)$$

The chain rule is applied in equation (4.16)

The derivative of  $f_k^o$  will be denoted as  $f_k^{ol}$

$$\frac{\partial (NET_{pk}^o)}{\partial w_{kj}^o} = \frac{\partial}{\partial w_{kj}^o} \sum_{j=1}^L w_{kj}^o i_{pj} + \theta_k^o = i_{pj} \dots\dots\dots(5.17)$$

Combining equations (4.16) and (4.17) yields the negative gradient as follows

$$\frac{-\partial E_p}{\partial w_{kj}^o} = (y_{pk} - o_{pk}) f_k^{ol}(NET_{pk}^o) i_{pj} \dots\dots\dots(5.18)$$

As far as the magnitude of the weight change is concerned, it is proportional to the negative gradient. Thus, the weights on the output layer are updated according to the following equation;

$$w_{kj}^o(t + 1) = w_{kj}^o(t) + \Delta_p w_{kj}^o(t) \dots\dots\dots(5.19)$$

Where;

$$\Delta_p w_{kj}^o(t) = \eta (y_{pk} - o_{pk}) f_k^{ol}(NET_{pk}^o) i_{pj} \dots\dots\dots(5.20)$$

The factor  $\eta$  is called the learning-rate parameter, ( $0 < \eta < 1$ ).

### 5.5.1.2 Output Function

The output function  $f_k^o(NET)_{jk}^o$  should be differentiable as suggested in *section 5.4.4*. This requirement eliminates the possibility of using linear threshold unit since the output function for such a unit is not differentiable at the threshold value. Output function is usually selected as linear function as illustrated below

$$f_k^o(NET)_{jk}^o = (NET)_{jk}^o \dots\dots\dots(5.21)$$

This defines the linear output unit.

In the first case:

$$f_k^{ol} = 1$$

$$w_{kj}^o(t + 1) = w_{kj}^o(t) + \eta (y_{pk} - o_{pk}) i_{pj} \dots\dots\dots(5.22)$$

The last equation can be used for the linear output regardless of the functional form of the output function  $f_k^o$ .

### 5.5.1.3 Update of Hidden-Layer Weights

The same procedure will be followed to derive the update of the hidden-layer weights. The problem arises when a measure of the error of the outputs of the hidden-layer units is needed. The total error,  $E_p$ , must be somehow related to the output values on the hidden layer. To do this, back to equation (4.15):

$$E_p = \frac{1}{2} \sum_k (y_{pk} - o_{pk})^2 \dots\dots\dots(5.15)$$

$$E_p = \frac{1}{2} \sum_k (y_{pk} - f_k^o(NE_{pk}^o))^2 \dots\dots\dots(5.23)$$

$$E_p = \frac{1}{2} \sum_k (y_{pk} - f_k^o(\sum w_{kj}^o i_{pj} + \theta_k^o))^2 \dots\dots\dots(5.24)$$

Taking into consideration,  $i_{pj}$  depends on the weights of the hidden layer through equations (4.10) and (4.11). This fact can be exploited to calculate the gradient of  $E_p$  with respect to the hidden-layer weights

$$\begin{aligned} \frac{\partial E_p}{\partial w_{ji}^h} &= \frac{1}{2} \sum_k \frac{\partial}{\partial w_{ji}^h} (y_{pk} - o_{pk})^2 = \\ &= - \sum_k (y_{pk} - o_{pk}) \frac{\partial o_{pk}}{\partial (NE_{pk}^o)} \frac{\partial (NE_{pk}^o)}{\partial i_{pj}} \frac{\partial i_{pj}}{\partial (NE_{pj}^h)} \frac{\partial (NE_{pj}^h)}{\partial w_{ji}^h} \dots\dots\dots(5.25) \end{aligned}$$

Each of the factors in equation (4.25) can be calculated explicitly from the previous equations. The result is;

$$\frac{\partial E_p}{\partial w_{ji}^h} = - \sum_k (y_{pk} - o_{pk}) f_k^{ol}(NE_{pk}^o) w_{kj}^o f_j^{hl}(NE_{pj}^h) x_{pj} \dots\dots\dots(5.26)$$

### ***5.5.2 Stopping Criteria***

Since back-propagation algorithm is a first-order approximation of the steepest-descent technique in the sense that it depends on the gradient of the instantaneous error surface in weight space. Weight adjustments can be terminated under certain circumstances. Kramer and Sangiovanni-Vincentelli et al. 1989 formulated sensible convergence criterion for back-propagation learning; the back-propagation algorithm is considered to have converged when:

- 1.** The Euclidean norm of the gradient vector reaches a sufficiently small gradient threshold.
- 2.** The absolute rate of change in the average squared error per epoch is sufficiently small.
- 3.** The generalization performance is adequate, or when it is apparent that the generalization performance has peaked.

## CHAPTER 6

### ERROR ANALYSIS

The statistical parameters used in the present work are: average percent relative error, average absolute percent relative error, minimum and maximum absolute percent error, root mean square error, standard deviation of error, and the correlation coefficient.

Graphical tools aid in visualizing the performance and accuracy of a correlation or a model. Three graphical analysis techniques are employed; those are crossplots, error distribution, and residual analysis. Also, the error trend will be studied.

#### *6.1 Statistical Error Analysis*

The statistical parameters used in the present work are:

##### **1. Average Relative Error**

It is the measure of relative deviation from the experimental data, defined by:

$$E_r = \frac{1}{n} \left[ \frac{P_{d_{Measured}} - P_{d_{Estimated}}}{P_{d_{Measured}}} \right] \dots\dots\dots(6.1)$$

##### **2. Average Absolute Percent Relative Error**

It measures the relative absolute deviation from the experimental values, defined by:

$$E_a = \frac{1}{n} \sum_{i=1}^n |E_i| \dots\dots\dots(6.2)$$

Where;  $E_i$  is the relative deviation of an estimated value from an experimental value

$$E_i = \left[ \frac{Pd_{Measured} - Pd_{Estimated}}{Pd_{Measured}} \right] \times 100 \dots\dots\dots(6.3)$$

### 3. Minimum and Maximum Absolute Percent Error

$$E_{min} = \min_{i=1}^n |E_i| \dots\dots\dots(6.4)$$

$$E_{max} = \max_{i=1}^n |E_i| \dots\dots\dots(6.5)$$

### 4. Root Mean Square Error

Measures the data dispersion around zero deviation, defined by:

$$RMSE = \left[ \frac{1}{n} \sum_{i=1}^n E_i^2 \right]^{0.5} \dots\dots\dots(6.6)$$

### 5. Standard Deviation of Error

It is a measure of dispersion and is expressed as:

$$STD = \sqrt{\left[ \left( \frac{1}{m-n-1} \right) \sum_{i=1}^m \left[ \left( \frac{Pd_{Measured} - Pd_{Estimated}}{Pd_{Measured}} \right) \times 100 \right]^2 \right]} \dots\dots\dots(6.7)$$

Where; (m-n-1) represents the degree of freedom in multiple- regression. A lower value of standard deviation indicates a smaller degree of scatter.

## 6. The Correlation Coefficient

It represents the degree of success in reducing the standard deviation by regression analysis, defined by:

$$R^2 = 1 - \frac{\sum_{i=1}^n [P_{dMeasured} - P_{dEstimated}]^2}{\sum_{i=1}^n [P_{dMeasured} - \bar{P}_d]^2} \dots\dots\dots(6.8)$$

‘R<sup>2</sup>’ values range between 0 and 1. The closer value to 1 represents perfect correlation whereas 0 indicates no correlation at all among the independent variables.

### 6.2 Graphical Error Analysis

Graphical analysis techniques employed are:

#### 1. Crossplot

In this graphical based technique, all estimated values are plotted against the measured values and thus a crossplot is formed. A 45° straight line between the estimated versus actual data points is drawn on the crossplot, which denotes a perfect correlation line. The tighter the cluster about the unity slope line, the better the agreement between the experimental and the predicted results.

#### 2. Error Distribution

The errors are said to be normally distributed with a mean around the 0%. Hence, most investigated models show either slight negatively skewed error distributed or positively ones.

#### 3. Residual Analysis

Analysis of residual is an effective tool to check model deficiencies.

### ***6.3 Trend Analysis***

A trend analysis was carried out to check whether the developed model is physically correct or not. For this purpose, synthetic sets were prepared where in each set only one input parameter was changed while other parameters were kept constant. To test the developed model, the effects of reservoir temperature and gas-oil were determined and plotted.



## CHAPTER 7

### DEVELOPMENT OF THE NEW MODELS

A total of 113 data sets were used in development the new models to estimate the dew-point pressure as a function of gas-oil ratio, reservoir temperature, gas specific gravity and heptanes plus specific gravity. The first model was developed using the traditional correlation techniques. The ACE algorithm was applied to develop the second model. Finally, an artificial neural network model was constructed to estimate the dew-point pressure.

#### *7.1 Traditional Correlation Model*

Non-linear multiple least square regression analysis was used to develop this correlation. Several models were tested to reach to the final form of the correlation:

$$\ln (P_d) = a_1 + a_2 \ln \left[ \frac{GOR \times \gamma_g}{\gamma_{7+}} \right] + a_3 \ln (T_R) + a_4 \gamma_g + \frac{a_5}{\gamma_{cond.}} + a_6 e^{a_7 + a_8 \ln (GOR)}$$

.....(7.1)

Where:  $a_1 = 18.6012$        $a_2 = -0.1520$        $a_3 = -0.1674$

$a_4 = 0.0685$        $a_5 = -5.8982$        $a_6 = -0.0559$

$a_7 = 8.4960$        $a_8 = -0.7466$

$P_d$  : Dew-point pressure, psia

$GOR$ : Gas-oil ratio, SCF/STB

$T_R$  : Reservoir temperature, °F

$\gamma_g$  : Gas specific gravity

$\gamma_{cond.}$  : Condensate specific gravity

## 7.2 Nonparametric Model (ACE)

The transforms were developed using this technique. The plots (Figures 7.1 to 7.4) present the transforms of each independent variable. Finally the following model was developed:

$$P_d = e^{c_1 T(P_d)^3 + c_2 T(P_d)^2 + c_3 T(P_d) + c_4} \dots\dots\dots(7.2)$$

Where

$$T(P_d) = \ln [T(T_R) + T(GOR) + T(\gamma_g) + T(\gamma_{cond.}) + 10] \dots\dots\dots(7.3)$$

And the transforms of the independent variables are:-

$$T(T_R) = p_1 T_R^3 + p_2 T_R^2 + p_3 T_R + p_4 \dots\dots\dots(7.4)$$

$$T(GOR) = r_1 \ln (GOR) + r_2 \dots\dots\dots(7.5)$$

$$T(\gamma_g) = q_1 \gamma_g^2 + q_2 \gamma_g + q_3 \dots\dots\dots(7.6)$$

$$T(\gamma_{cond.}) = s_1 \gamma_{cond.}^3 + s_2 \gamma_{cond.}^2 + s_3 \gamma_{cond.} + s_4 \dots\dots\dots(7.7)$$

$$C_1 = 49.1377, \quad C_2 = -336.5699, \quad C_3 = 770.0995, \quad C_4 = -580.0322$$

$$p_1 = -0.35014 \times 10^{-6}, \quad p_2 = 0.18048 \times 10^{-3}, \quad p_3 = -0.32315 \times 10^{-1}, \quad p_4 = 1.2058$$

$$r_1 = -0.3990, \quad r_2 = 5.1377, \quad q_1 = -23.8741, \quad q_2 = 36.9448, \quad q_3 = -12.0398$$

$$s_1 = -30120.78, \quad s_2 = 69559, \quad s_3 = -53484.21, \quad s_4 = 13689.39$$

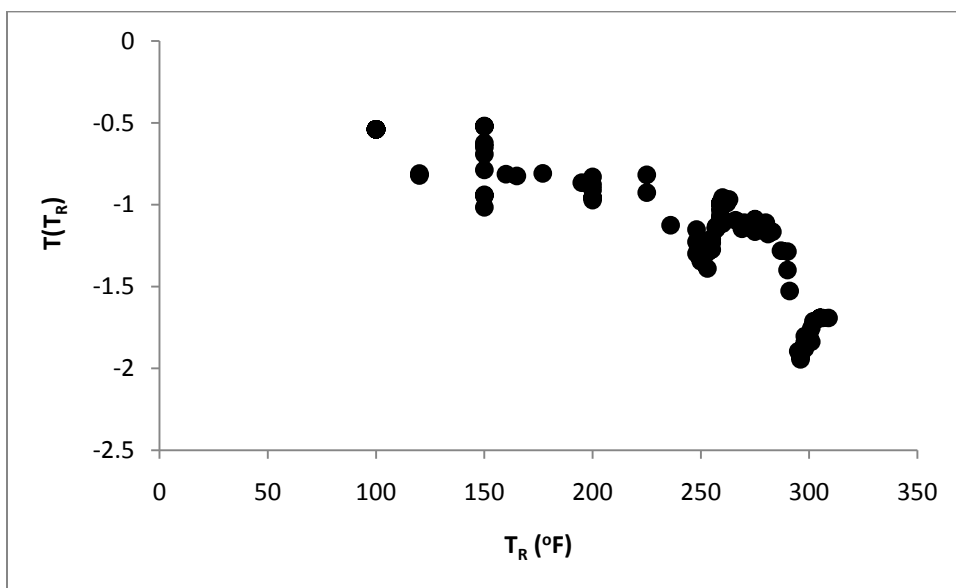


Figure 7.1: Optimal Transform of Reservoir Temperature

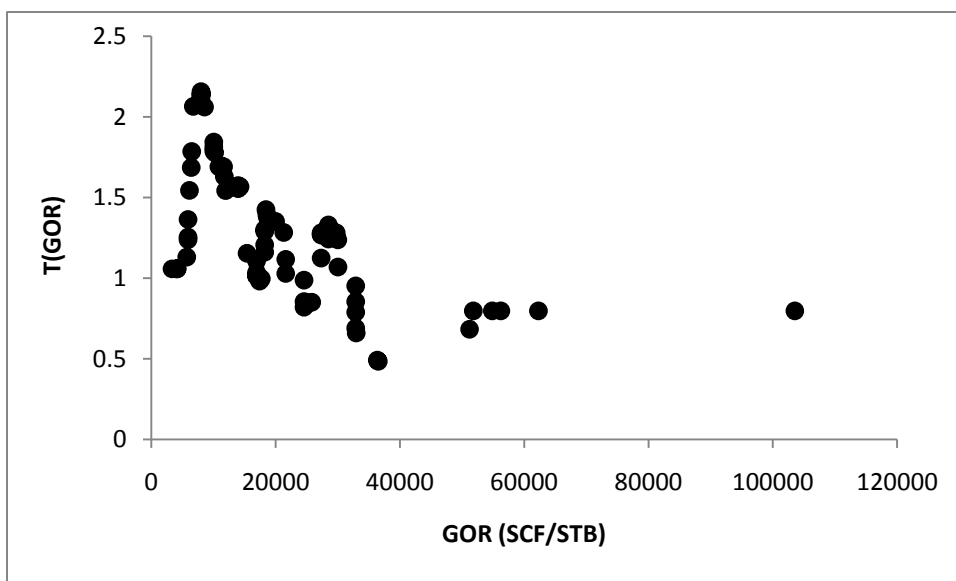


Figure 7.2: Optimal Transform of Gas-Oil Ratio

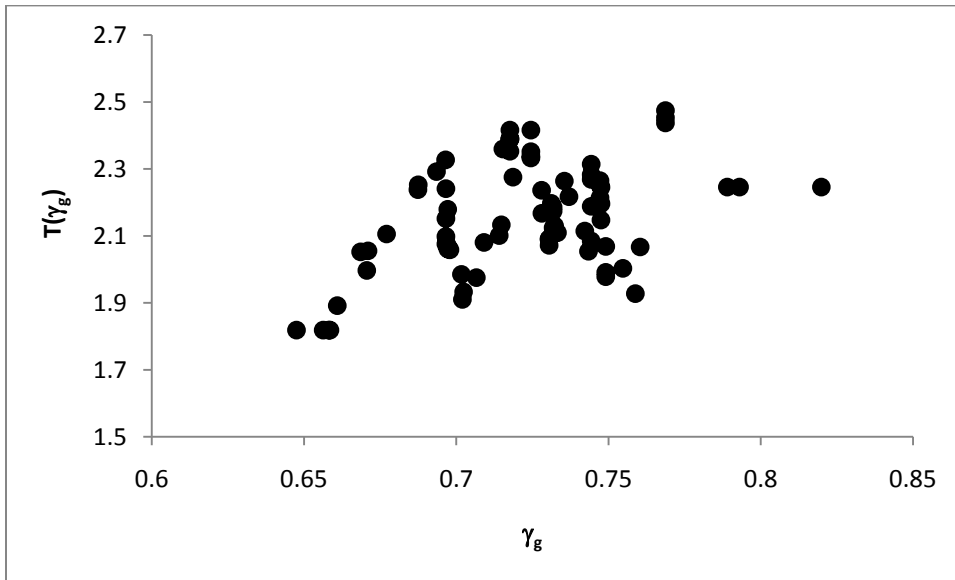


Figure 7.3: Optimal Transform of Gas Specific Gravity

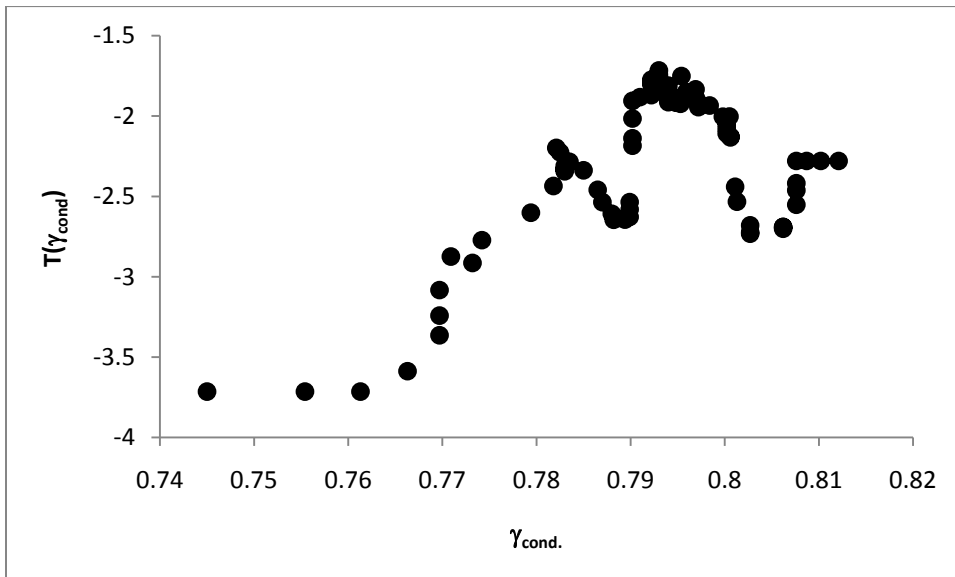


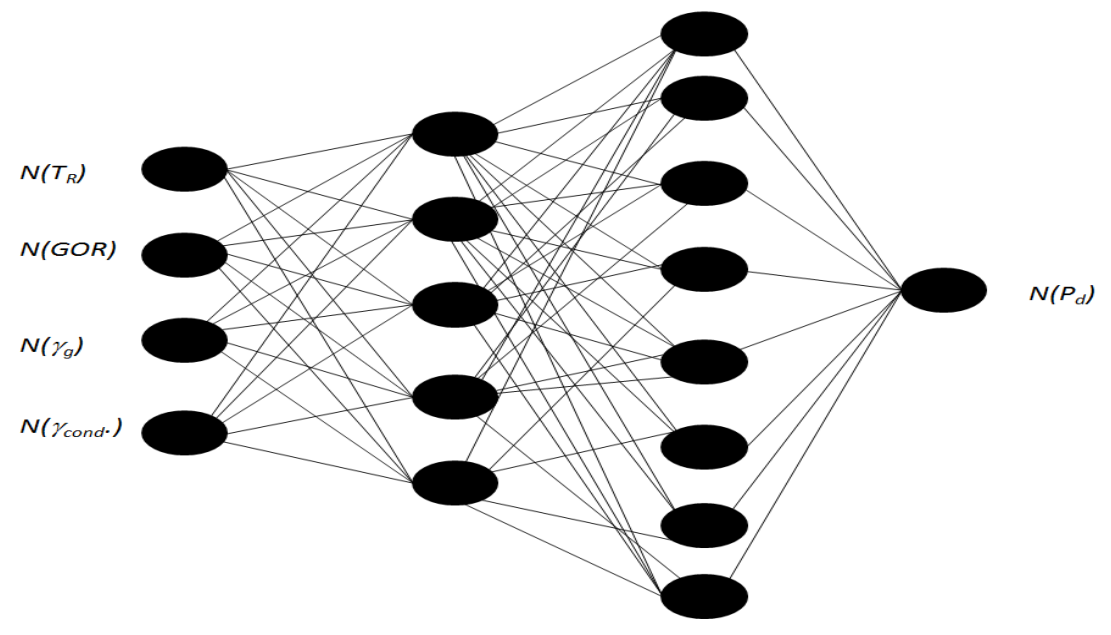
Figure 7.4: Optimal Transform of Condensate Specific Gravity

### 7.3 Artificial Neural Network Model

New artificial neural network model was developed to estimate the dew-point pressure. Gas-oil ratio, reservoir temperature, gas specific gravity and heptanes plus specific gravity are used as inputs to feed the neural network. The neural network architecture consists of three layers; one input layer with 4 neurons, two hidden layers with 5 and 8 neurons respectively; and one output layer with one neuron. The backpropagation technique and the Levenberg-Marquardt training algorithm are used to minimize the mean-square error.

The data were normalized between (0.2 and 0.8) to avoid ill-conditioning and to alleviate saturation problem by an equation such as:

$$N(X) = \left( \frac{X - \text{Min}(X)}{\text{Max}(X) - \text{Min}(X)} \right) \times (0.8 - 0.2) + 0.2 \dots \dots \dots (7.8)$$



### 7.3.1 Artificial Neural network Model in Matrix Form

The artificial neural network method has been converted into Matrix form. Putting the model in this form will help in programming the model without using sophisticated software. The following steps summarize how to estimate the dew-point pressure using matrices.

Step#1: Normalize the input data

$$N(T_R) = \left( \frac{T_R - 100}{309 - 100} \right) \times (0.8 - 0.2) + 0.2 \dots \dots \dots (7.9)$$

$$N(GOR) = \left( \frac{GOR - 3321.176}{103539 - 3321.176} \right) \times (0.8 - 0.2) + 0.2 \dots \dots \dots (7.10)$$

$$N(\gamma_g) = \left( \frac{\gamma_g - 0.6475}{0.8199 - 0.6475} \right) \times (0.8 - 0.2) + 0.2 \dots \dots \dots (7.11)$$

$$N(\gamma_{cond.}) = \left( \frac{\gamma_{7+} - 0.7303}{0.8121 - 0.7303} \right) \times (0.8 - 0.2) + 0.2 \dots \dots \dots (7.12)$$

Step#2: Calculate the first hidden layer ( $L1$ ) in ( $5 \times 1$ ) matrix

$$L1 = \begin{vmatrix} 6.6129 & 1.1447 & 1.304 & -4.8224 \\ -5.871 & -1.2771 & 4.3998 & 2.2958 \\ 6.8838 & -0.3654 & -1.383 & 3.6311 \\ -2.2535 & 0.0535 & 4.0037 & 2.7576 \\ -5.0607 & 2.2438 & 2.2576 & 5.4482 \end{vmatrix} \times \begin{vmatrix} N(T_R) \\ N(GOR) \\ N(\gamma_g) \\ N(\gamma_{7+}) \end{vmatrix} + \begin{vmatrix} -2.7753 \\ 1.554 \\ -4.9073 \\ -4.0399 \\ -4.089 \end{vmatrix} \dots \dots \dots (7.13)$$

Step#3: Calculate  $Tansig(L1)$  as follow

$$Tansig(L1) = \begin{bmatrix} Tansig(l_{11}) \\ Tansig(l_{21}) \\ Tansig(l_{31}) \\ Tansig(l_{41}) \\ Tansig(l_{51}) \end{bmatrix} \dots\dots\dots(7.14)$$

Where

$$Tansig(x) = \frac{2}{1+e^{-2x}} - 1 \dots\dots\dots(7.15)$$

Step#4: Calculate the second hidden layer ( $L2$ ) in ( $8 \times 1$ ) matrix

$$L2 = \begin{bmatrix} -0.5687 & 0.3812 & 0.2953 & 1.7564 & 1.0291 \\ -0.9498 & 0.2475 & -0.5088 & -1.71 & -0.56 \\ -1.3044 & 0.2327 & 0.2349 & -1.7294 & 0.1741 \\ 2.2099 & -0.3945 & 0.2802 & 0.9415 & 1.3975 \\ 2.1921 & 1.3 & 2.7866 & -0.3257 & 0.7323 \\ 0.9597 & 0.3262 & 2.3822 & -0.4447 & -0.42 \\ 1.0194 & -3.4971 & -1.037 & -3.0858 & -1.5253 \\ -0.4748 & -0.6921 & 0.4353 & 1.7615 & 0.4089 \end{bmatrix} \times Tansig(L1) + \begin{bmatrix} 0.7179 \\ 1.4592 \\ 1.3767 \\ -1.2187 \\ 0.7673 \\ -1.2508 \\ 1.4816 \\ 3.0874 \end{bmatrix} \dots\dots\dots(7.16)$$

Step#5: Calculate  $Tansig(L2)$  as follow

$$Tansig(L2) = \begin{bmatrix} Tansig(N_{11}) \\ Tansig(N_{21}) \\ Tansig(N_{31}) \\ Tansig(N_{41}) \\ Tansig(N_{51}) \\ Tansig(N_{61}) \\ Tansig(N_{71}) \\ Tansig(N_{81}) \end{bmatrix} \dots\dots\dots(7.17)$$

Step#6: Calculate  $N(P_d)$

$$N(P_d) = \left| \begin{array}{cccccccc} 0.6402 & -0.5186 & 0.3587 & -0.1736 & 0.1586 & -0.1472 & 0.4849 & 0.447 \end{array} \right| \times \text{Tansig}(L2) + \left| \begin{array}{c} 0.221 \end{array} \right|$$

.....(7.18)

Step#7: Calculate the dew-point pressure ( $P_d$ )

$$P_d = N^{-1}(P_d) = \left( \frac{N(P_d) - 0.2}{0.8 - 0.2} \right) \times (7450 - 2562) + 2562 \dots \dots \dots (7.19)$$



## CHAPTER 8

### RESULTS AND DISCUSSIONS

The dewpoint pressure correlations were evaluated in two stages. In the first stage, they were evaluated using their original coefficients which were published in the original papers. In the second stage, the coefficients of these correlations were recalculated in order to have a better performance in fitting the used data.

For the neural network models the data were divided into three groups; training 70%, validation 10% and testing 20%. Also, the data were normalized between 0.2 and 0.8 in order to avoid the ill condition.

#### *8.1 Published Correlations Evaluation*

Nemeth and Kennedy correlation has reasonable results. The average absolute error with the original coefficients is 13.3 % while it is 6.7 % for the new coefficients. The error distribution is shifted to the left with skewness of -0.5 (figure.8.8). Figure.8.1 presents the cross-plot of Nemeth and Kennedy correlation.

On the other hand, Elsharkawy correlation shows that the average absolute errors for the original and the new coefficients are 16 % and 10.7 % respectively. The error

distribution exhibits clear shift to the left with skewness of -2 (figure.8.9). Figure 8.2 shows that this correlation is not as accurate as Nemeth and Kennedy correlation.

Homud and Al-Marhoun correlation relates the dew-point pressure to gas properties. One of these properties is pseudoreduced pressure. Although the correlation statistically works (figure.8.3), the physics of this relationship is questionable. Pseudoreduced pressure is a function of the reservoir pressure; therefore, the correlation is relating indirectly the dew-point pressure to the reservoir pressure. It is well known that dew-point pressure is a function of gas composition. Therefore, any gas with the same composition would have the same dew-point pressure regardless of the original reservoir pressure. The average absolute error with the original coefficients is 30% while it is 9.7 % for the new ones. The error distribution is shifted to the left with skewness of -1.1 (figure.8.10).

Marruffo, Maita, Him and Rojas model has an average absolute error with the original coefficients of 23% while it is 9.8 % for the new ones. The error distribution is shifted to the left with skewness of -2.5 (figure.8.11). Figure.8.4 presents the cross-plot.

In general, all of the pervious correlations have better accuracy with the modified coefficients as per tables 8.1 and 8.2.

## ***8.2 New Models Evaluation***

Three new models were developed to estimate the dew-point pressure as a function of reservoir temperature, gas-oil ratio, gas specific gravity and C7+ specific gravity. The first model was developed as traditional correlation while the second model

was developed using nonparametric regression method and finally the third model is artificial neural network model. The first model shows reasonable results; however, the nonparametric model was not successful due to difficulty in fitting the transformation parameter. The neural network model is the best among the developed models.

### ***8.2.1 Traditional Correlation Model***

This new correlation has an average absolute error of 9.6%. The error distribution is shifted to the left with skewness of -1.9 (figure.8.12). Figure.8.5 presents the cross-plot error.

### ***8.2.2 Nonparametric Approach (ACE)***

This model was developed using ACE model. This algorithm creates new transformation functions from the dependent and independent variables. In general, ACE model has better results than the conventional models (models that depend on the fluid properties). The average absolute error is 9.5% with small skewness of -0.5. Figures 8.6 and 8.13 show the graphical errors of this model.

### ***8.2.3 Artificial Neural network Model***

The structure of this artificial neural network model consists of one input layer with 4 nodes, two hidden layers with 5 and 8 nodes respectively and one output layer with one node. This model shows excellent results and it is the best among all previous

models (figure 8.7). The average absolute error is 6.5% and the error distribution is shifted a little bit to the left with skewness of -0.5. (Figure 8.14).

Figure 8.16 and Figure 8.17 show the dependency of dew-point pressure on the reservoir temperature when the other variables were fixed at average values. The new correlation shows that the dew-point pressure is decreasing as function of the reservoir temperature. ACE model shows that the dew-point pressure is decreasing as function of the reservoir temperature till at high temperature there will be no dew-point pressure (dry gas). The dew-point pressure with ANN exhibits similar behavior to what Akbari and Farahani found out in their ANN model. They drew this conclusion: The dew-point pressure is increasing function with respect to the reservoir temperature until the cricondenbar and then pressure decreases with temperature until cricondentherm point.

Figure 8.18 shows the dependency of dew-point pressure on the gas-oil ratio. The new correlation and the non-parametric model (ACE) present exponential relationship between dew-point pressure and gas-oil ratio. While, the artificial neural network model show that the relationship between dew-point pressure and gas-oil ratio is similar to sigmoid function.

Table 8.1: Error Statistics with New Coefficients

	$E_r$	$E_{max}$	$E_{min}$	$E_a$	RMSE	R	STD	skewness
Nemth&Kennedy	-4.3	22.6	0.2	6.7	9.3	0.80	6.6	-0.5
Elsharkawy	-0.6	53.7	0.4	10.7	14.9	0.42	10.7	-2.0
Humoud	-0.3	33.5	1.2	9.7	12.1	0.69	8.0	-1.1
Marruffo&Rojas	-4.1	63.5	0.2	9.8	16.4	0.62	13.4	-2.3
New Correlation	-3.5	50.2	0.2	9.6	13.9	0.69	11.0	-1.9
ACE Model	-0.8	39.2	0.3	9.5	12.7	0.58	8.4	-0.5
New ANN (testing)	-1.7	23.8	0.9	6.5	8.6	0.82	5.8	-0.5

Table 8.2: Error Statistics with Original Coefficients

	$E_r$	$E_{max}$	$E_{min}$	$E_a$	RMSE	R	STD
Nemth&Kennedy	11.6	40.21	0.06	13.3	16.4	-0.12	9.1
Elsharkawy	12.2	43.1	0.46	16	19.3	-0.51	11
Humoud	-28.7	160.7	0.63	30.7	42.4	-3.5	29.4
Marruffo&Rojas	16	54.8	1.04	23.3	24.8	8.7	-0.56

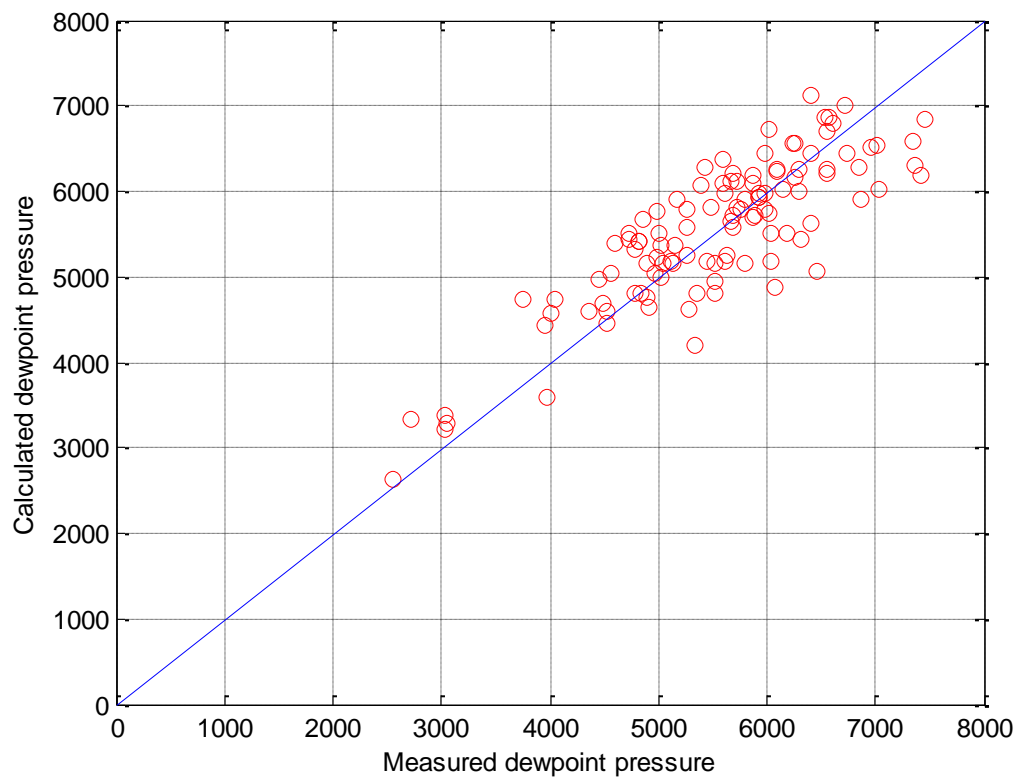
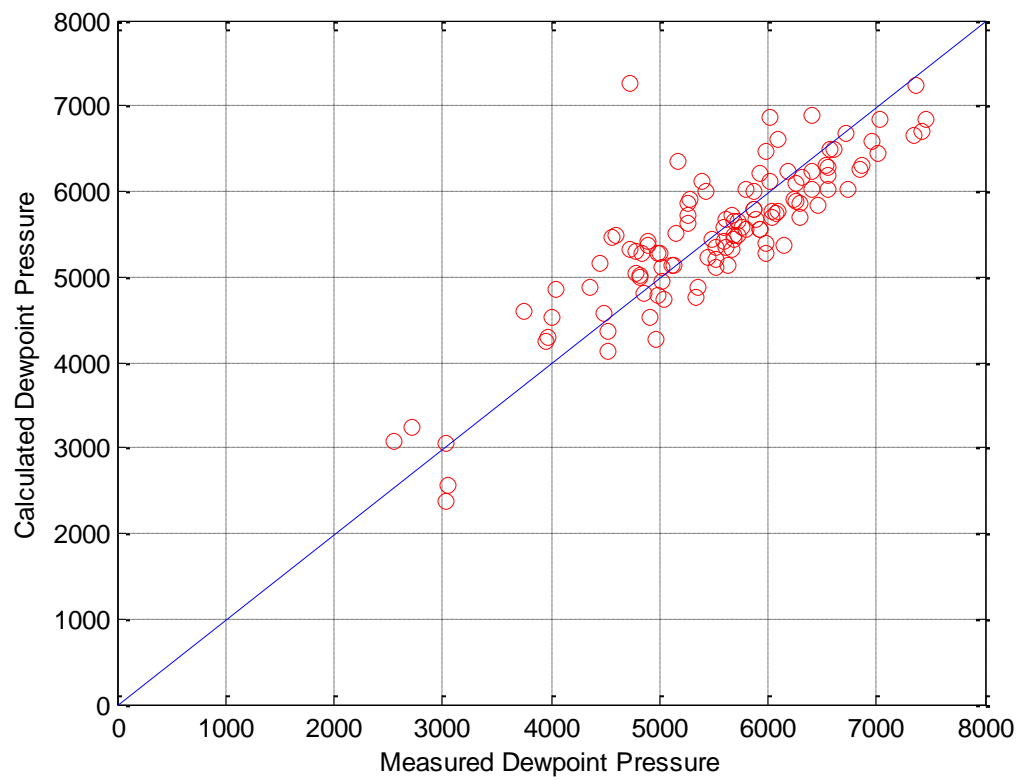


Figure 8.1: Cross Plot (Nemeth and Kennedy)





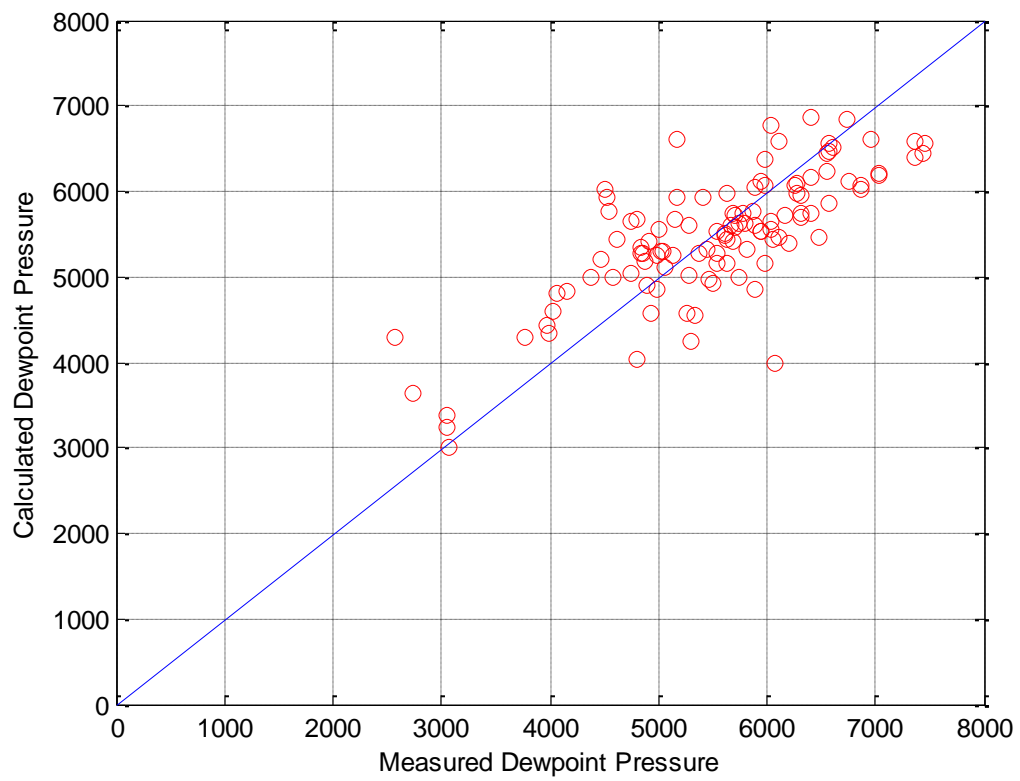


Figure 8.3: Cross Plot (Homud and Al-Marhoun)

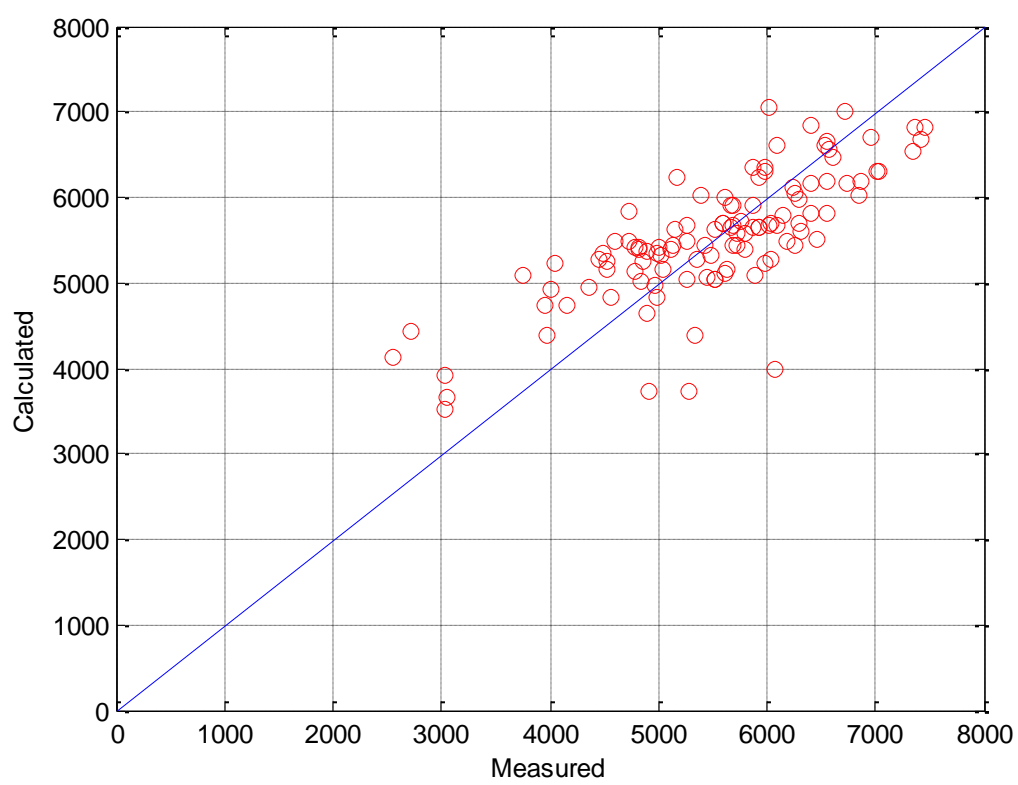


Figure 8.4: Cross Plot (Marruffo, Maita, Him and Rojas)

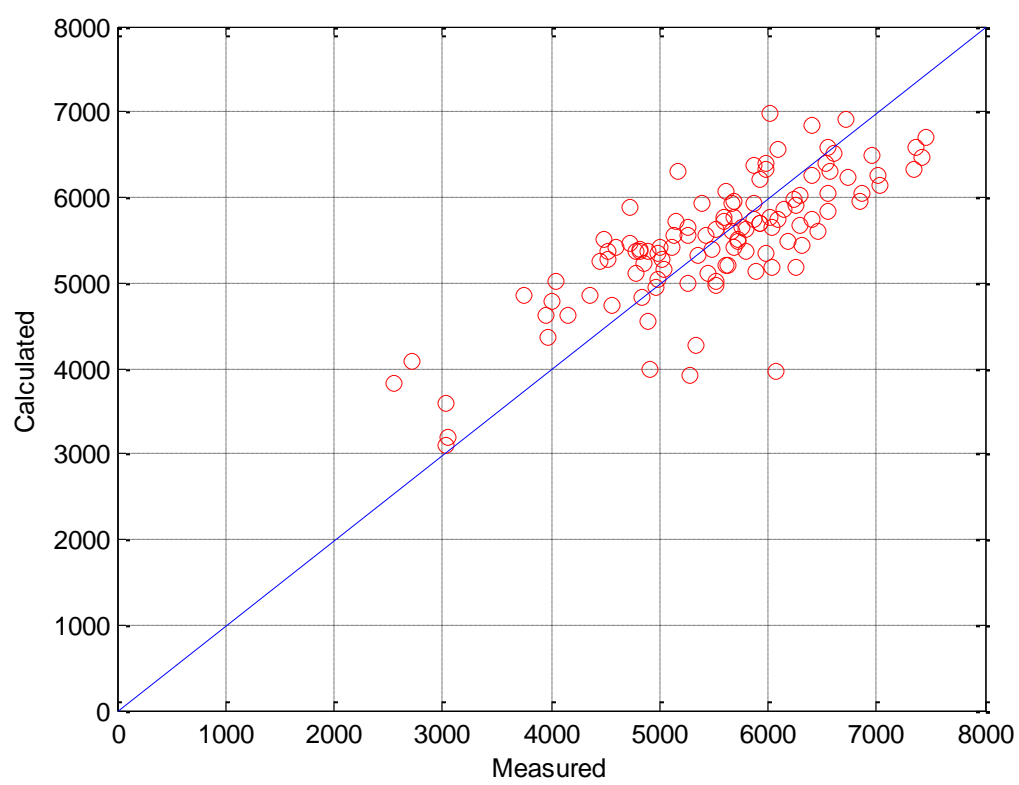


Figure 8.5: Cross Plot (New Correlation)

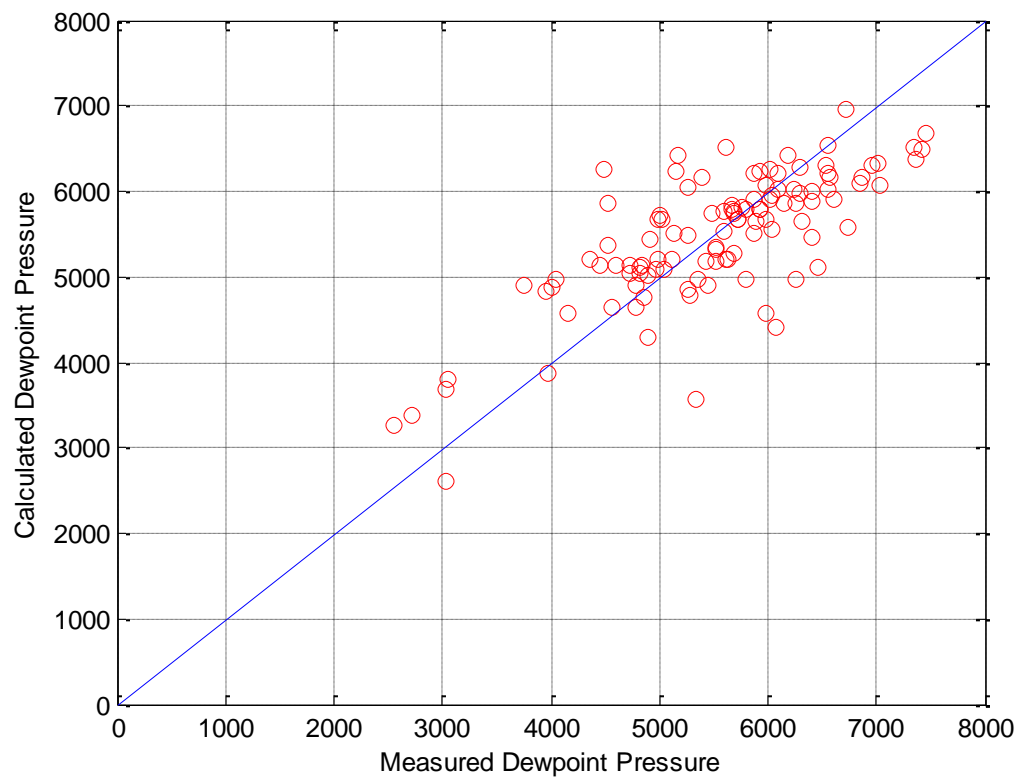


Figure 8.6: Cross Plot (ACE model)

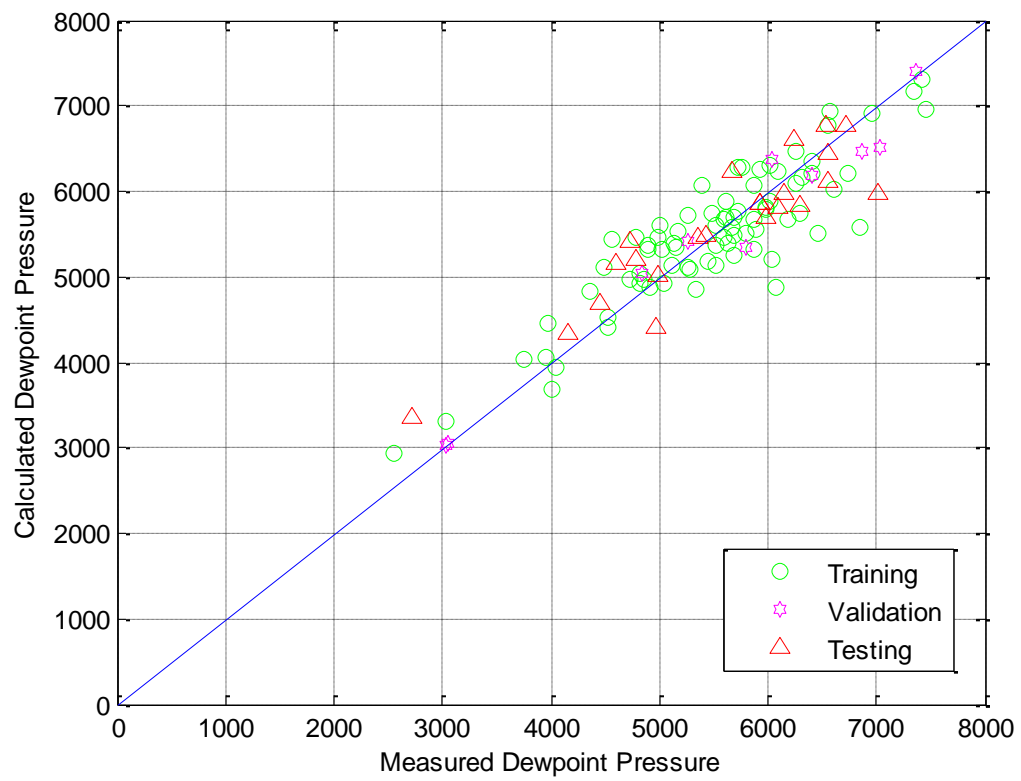


Figure 8.7: Cross Plot (New Artificial Neural Network Model)

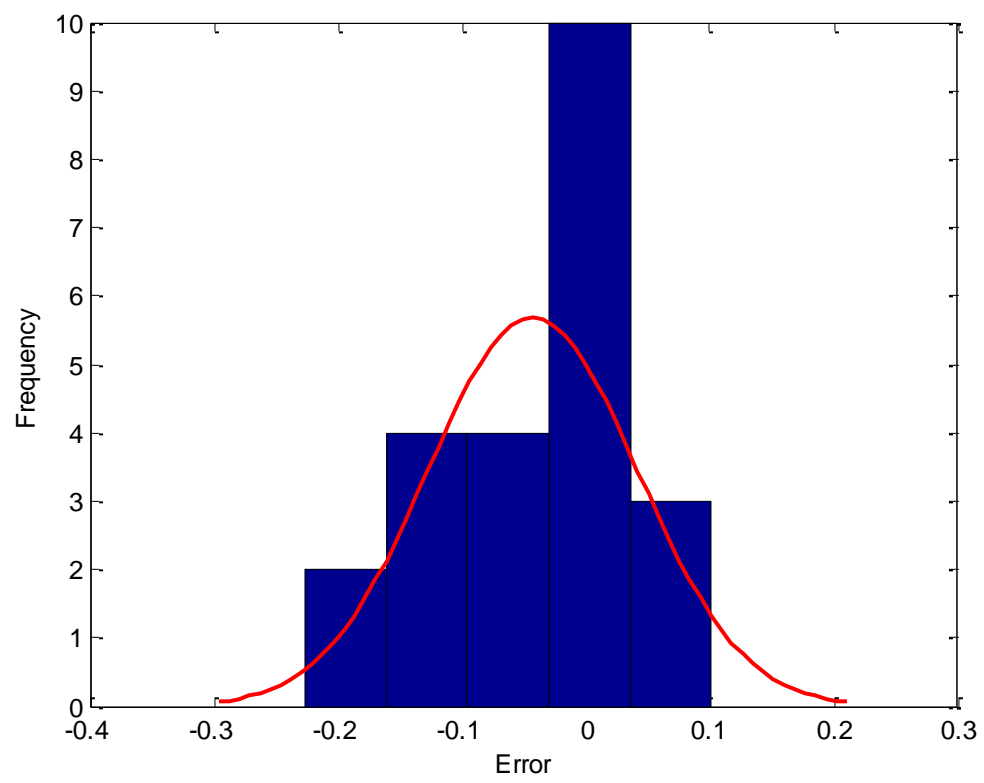


Figure 8.8: Error Distribution (Nemeth and Kennedy)

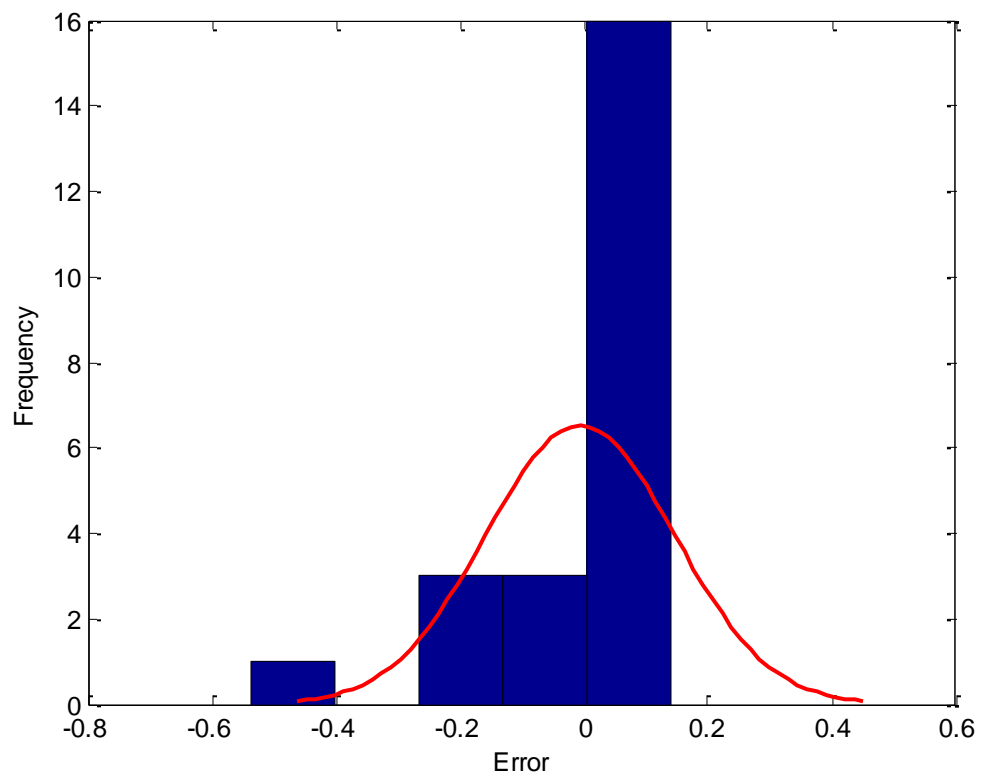


Figure 8.9: Error Distribution (Elsharkawy)

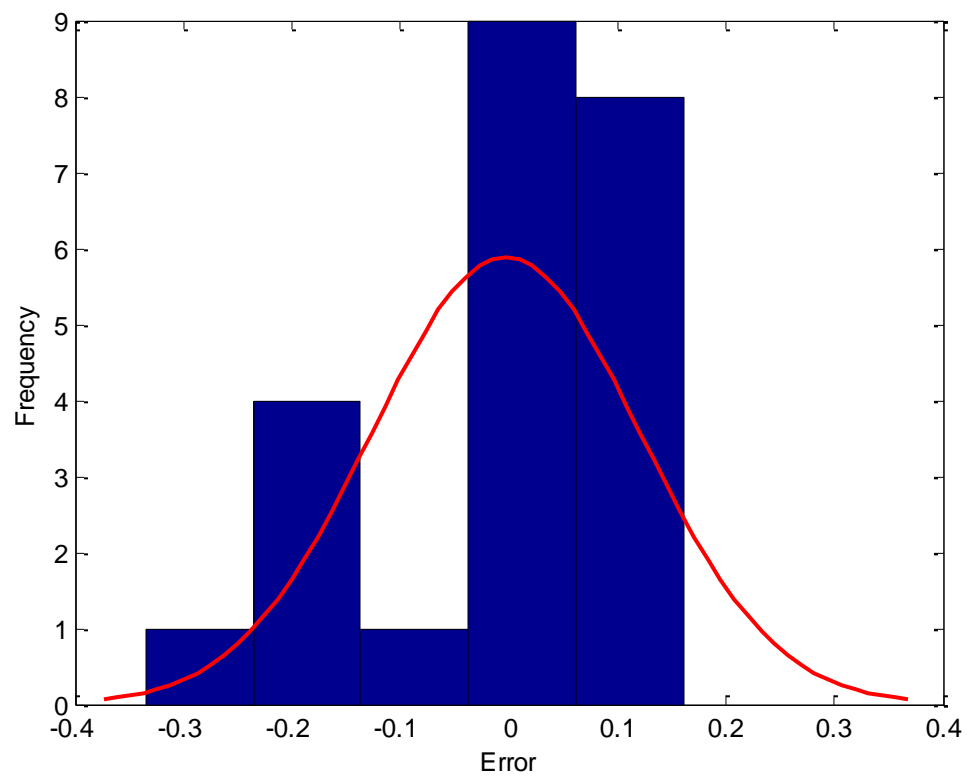


Figure 8.10: Error Distribution (Homud and Al-Marhoun)



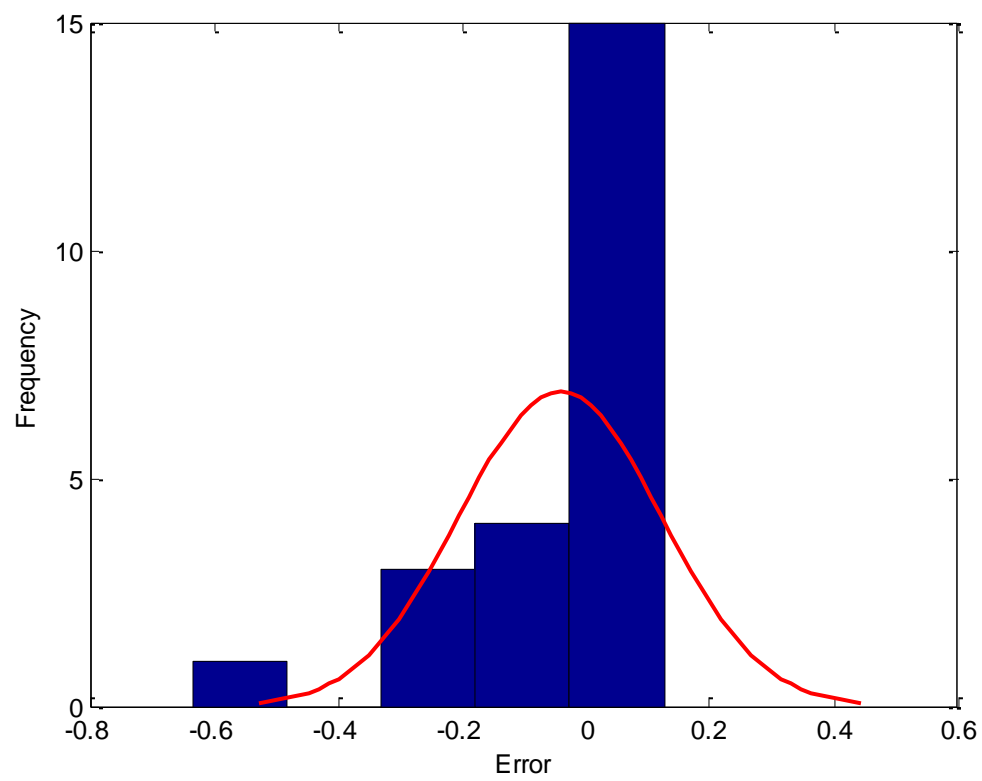


Figure 8.11: Error Distribution (Marruffo, Maita, Him and Rojas)

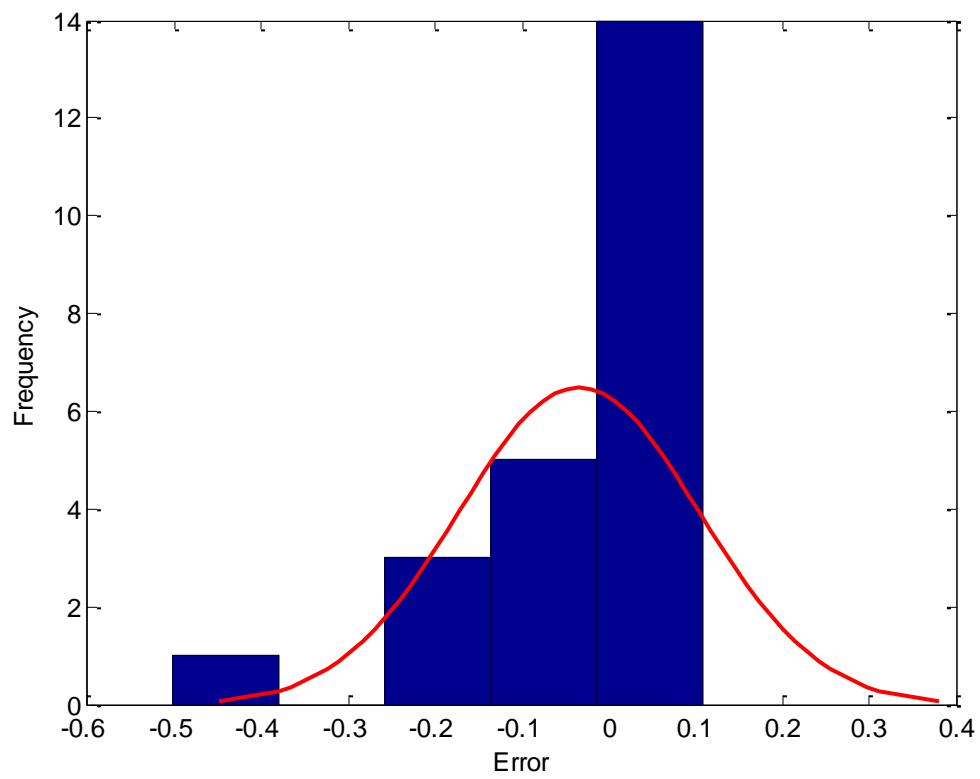


Figure 8.12: Error Distribution (New Correlation)

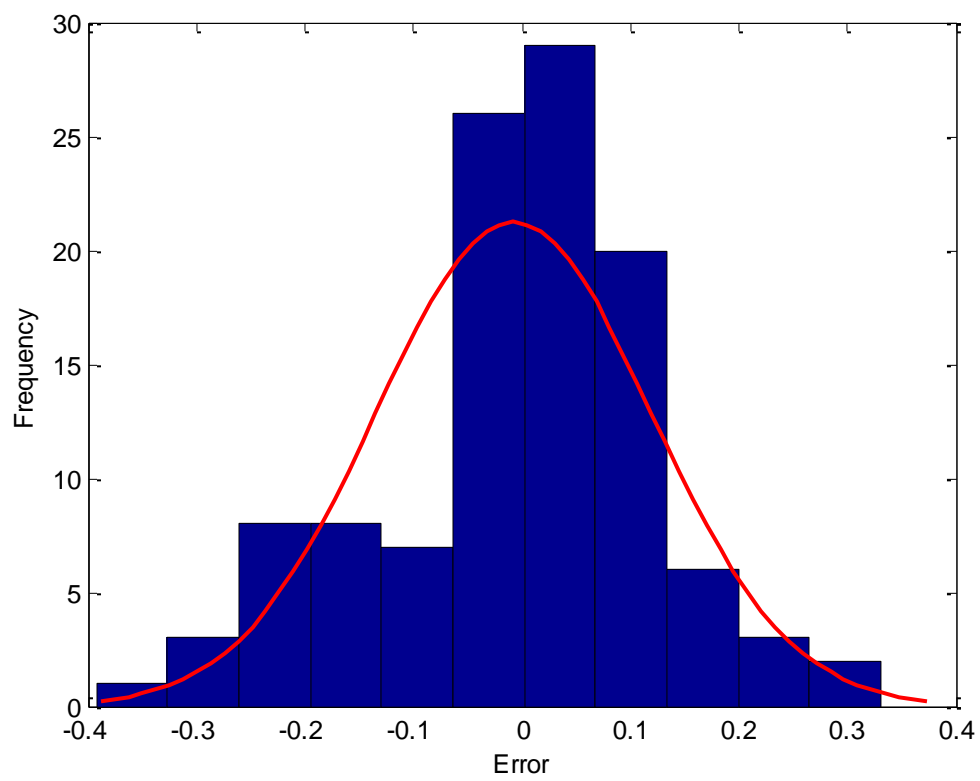


Figure 8.13: Error Distribution (ACE model)

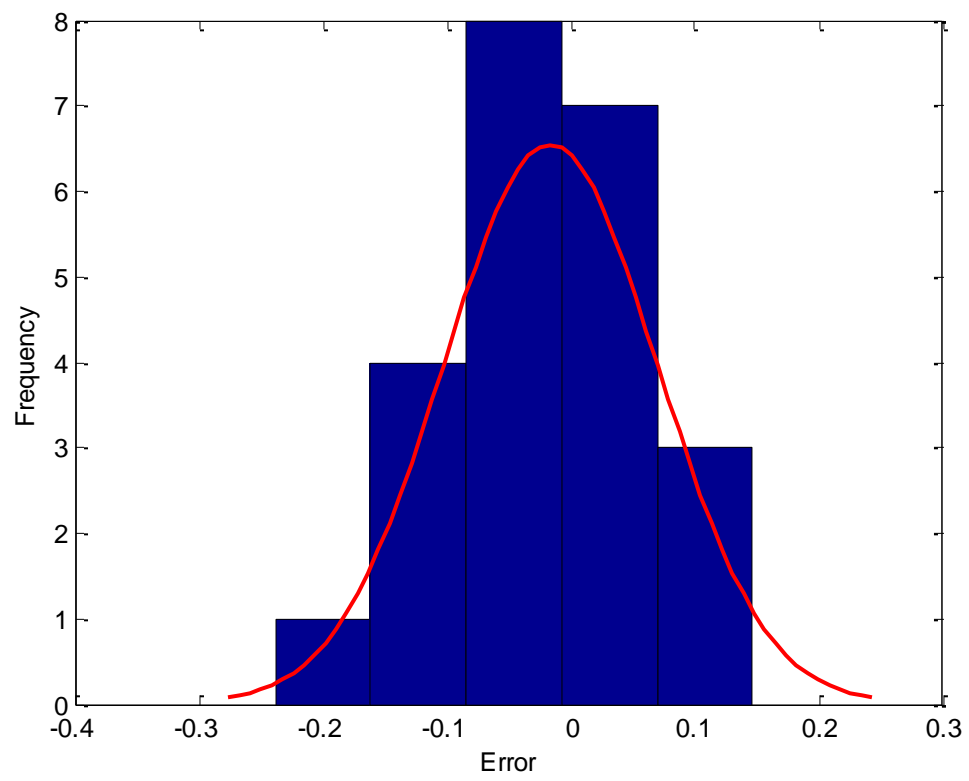


Figure 8.14: Error Distribution (New Artificial Neural Network model)

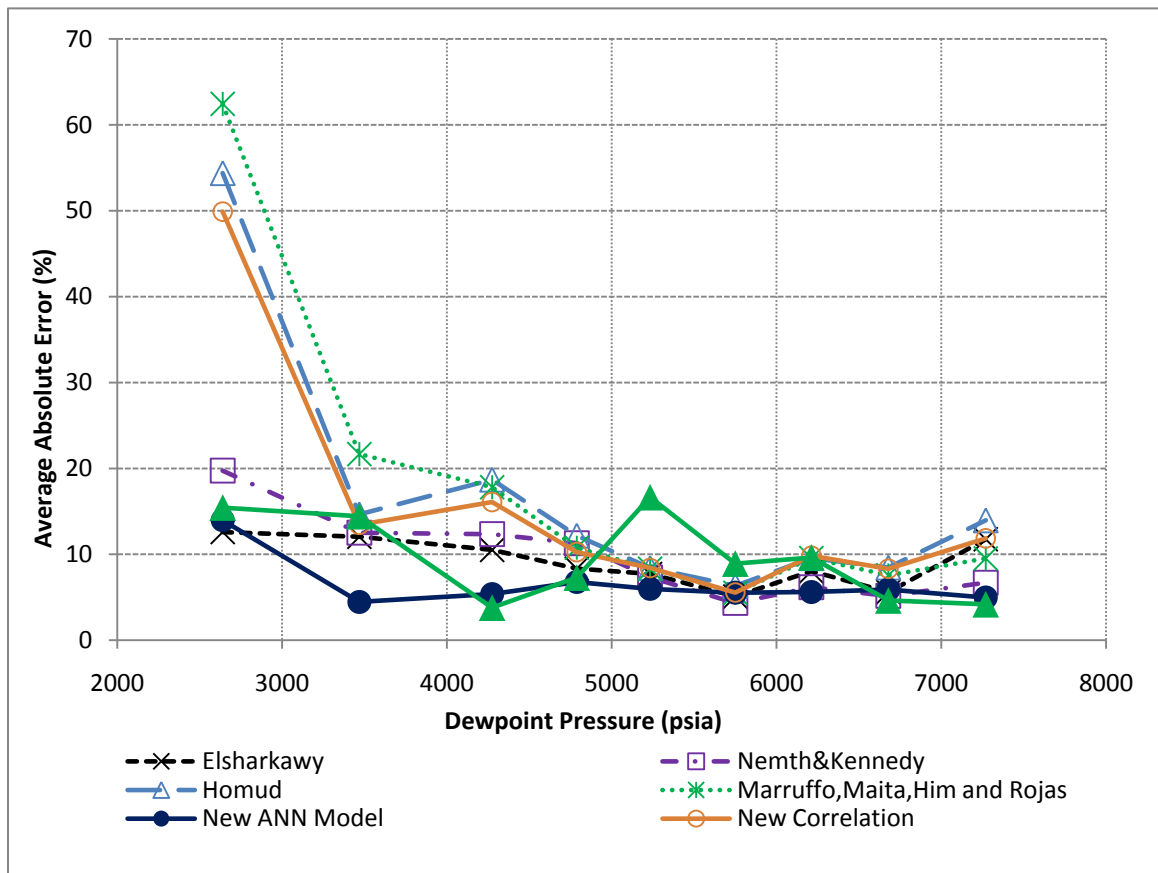


Figure 8.15: Accuracy of Correlations for Ranges of Dew-Point Pressures

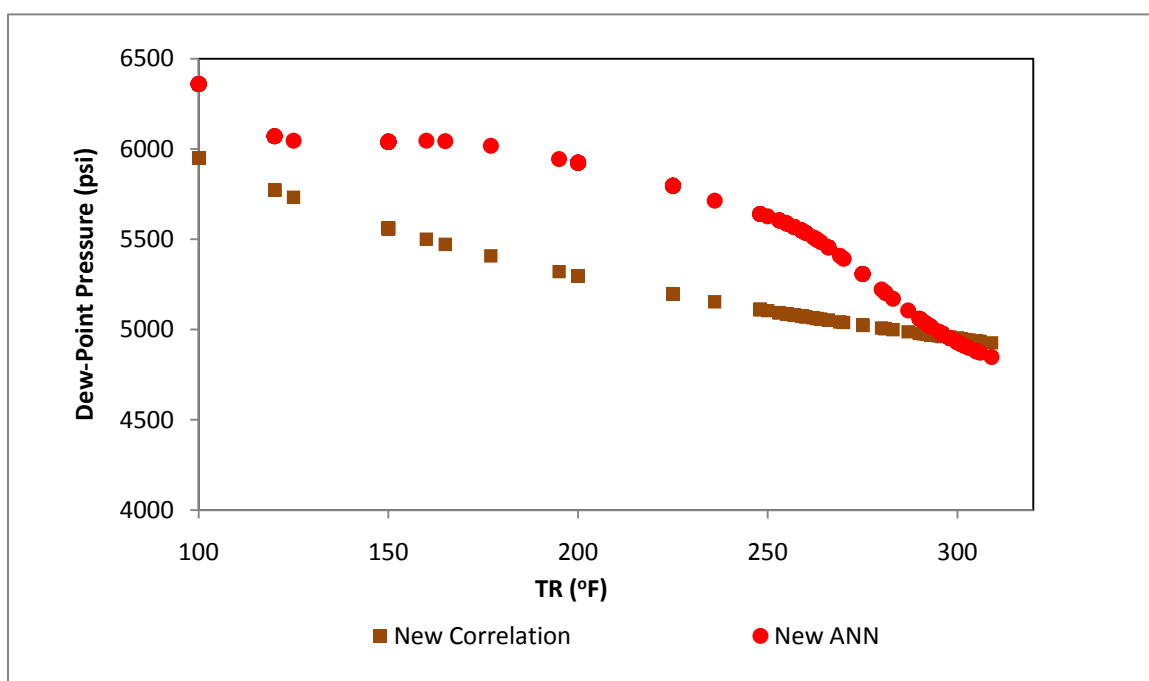


Figure 8.16: Sensitivity of New Models to Reservoir Temperature

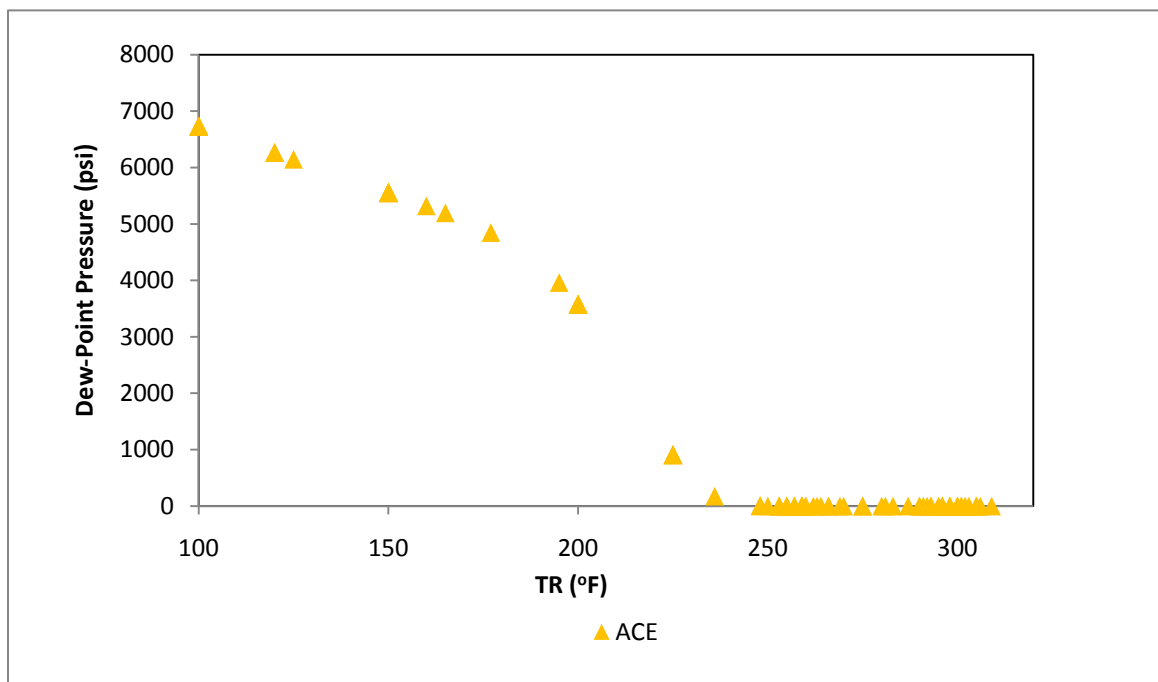


Figure 8.17: Sensitivity of ACE Model to Reservoir Temperature

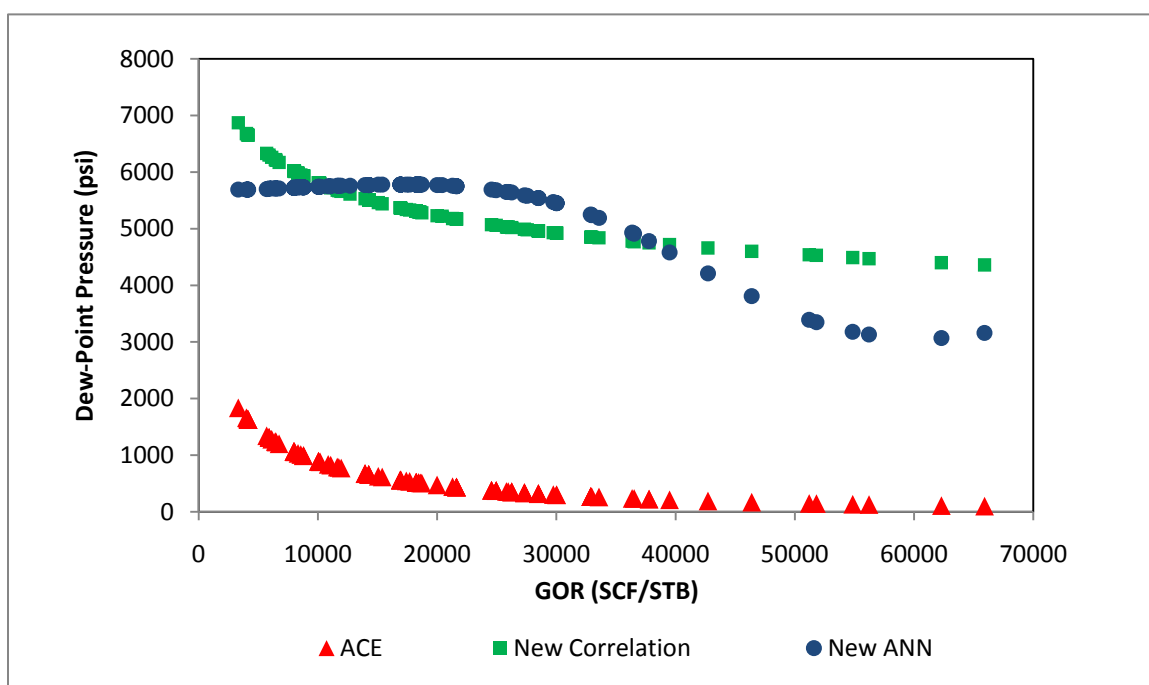


Figure 8.18: Sensitivity of New Models to Gas-Oil Ratio



## CHAPTER 9

### CONCLUSIONS

Three new models were developed in this study to predict the dew-point pressure for gas condensate reservoir: traditional correlation, nonparametric model using ACE algorithm and artificial neural network model. Based upon the literature review and work performed in this thesis, the following conclusions were drawn:

1. The artificial neural network has the best results among all other models.
2. In general, the correlations that depend on the gas composition perform better than the correlation that depends on fluid properties only.
3. All conventional correlations that depend on the fluid properties have failed in predicting the dew-point pressure below 4000 psia.
4. The new correlation and the non-parametric model (ACE) present exponential relationship between dew-point pressure and gas-oil ratio. While, the artificial neural network model shows that the relationship between the dew-point pressure and gas-oil ratio is similar to sigmoid function.
5. The dew-point pressure with ANN exhibits similar behavior to what Akbari and Farahani found out in their ANN model. The dew-point pressure is increasing

function with respect to the reservoir temperature until the cricondenbar and then pressure decreases with temperature until cricondentherm point.

**REFERENCES:**

1. Akbari, M. and Farahani, F. "Dewpoint Pressure Estimation of Gas Condensate Reservoir Using Artificial Neural Network". Paper SPE 107032, presented at the 2007 SPE Europe/EAGE Annual Conference and Exhibition held in London, 11–14 June 2007.
2. Breiman, L. and Friedman, J.H.: "Estimating Optimal Transformations for Multiple Regression and Correlation," Journal of American Statistical Association, September 1985, 580-619.
3. Elsharkawy, A.M., "Characterization of the Plus Fraction and Prediction of the Dewpoint Pressure for Gas Condensate Reservoirs" Paper SPE 68776, presented at the 2001 SPE Western Regional Meeting held in Bakersfield, California, USA, 26-30, March 2001.
4. Gonzalez A., Barrufet M. and Startzaman R, "Improved Neural Network Models Predicts Dewpoint Pressure of Retrograde Gas". Journal of Petroleum Science and Engineering 37 (2003),183-194
5. Humod, A.A. and Al-Marhoun, M.A., "A New Correlation for Gas-condensate Dewpoint Pressure Prediction" Paper SPE 68230, presented at the 2001 SPE Middle East Oil Show held in Bahrain, 17–20 March 2001.
6. ILSis Marruffo, PDVSA, Jose Maita, Jesus Him, Gonzal Rojas "Correlation To Determine Retrograde Dew Pressure and C7+ Percentage of Gas Condensate reservoirs on Basis of Production test Data of Eastern Venezuelan Fields" Paper SPE 75686, presented at the SPE Gas Technology Symposium held in Calgary, Alberta, Canada, 30 April 2002.
7. Kramer, A. H. and Sangiovanni, V. A.: "Efficient Parallel Learning Algorithms for Neural Networks," in Advances in Neural Information Processing Systems 1 (D.S. Touretzky, Edition), San Mateo, CA, Morgan Kaufmann. (1989) 40-78.
8. Nemeth, L.K. and Kennedy, H.T.: "A Correlation of Dewpoint Pressure with Fluid Composition and Temperature," paper SPE 1477 presented at SPE 41<sup>st</sup> Annual Fall Meeting held in Dallas, Texas, 1966.
9. "MATLAB," Mathwork, Neural Network Toolbox Tutorial, ([http://www.mathtools.net/MATLAB/Neural\\_Networks/index.html](http://www.mathtools.net/MATLAB/Neural_Networks/index.html)).

10. Potsch, K.T. and Braeuer, L., "A Novel Graphical Method for Determining Dewpoint Pressures of Gas Condensates," Paper SPE 36919, presented at the 1996 SPE European Conference held in Italy, October 22-24,1996.

## APPENDIX A

### RANGE OF VARIABLES USED IN THIS STUDY

Mole %	Min	Max
Methane	57.71	84.95
Ethane	4.19	10.71
Propane	1.31	5.99
Butane	0.59	3.45
Pentane	0.22	1.85
Hexane	0.15	2.03
Heptane plus	0.53	13.03
Carbon dioxide	0.4	18.29
Nitrogen	0.12	5.71
Hydrogen Sulfide	0	9.32
Reservoir Temperature (°F)	100	309
Gas-Oil ratio (SCF/STB)	3321	103536
Gas Specific Gravity	0.6475	0.8199
Heptane plus Specific Gravity	0.7303	0.8121
Dew-point Pressure (psia)	2726	8800

## APPENDIX B

### PUBLISHED CORRELATIONS IN LITERATURE

1. Nemeth and Kennedy (1967)

$$\begin{aligned} \ln(P_d) = & A1[xC_2 + xCO_2 + xH_2S + xC_6 + 2(xC_3 + xC_4) + xC_5 + 0.4xC_1 \\ & + 0.2xN_2] + A2\gamma_{c7+} + A3[xC_1/(xC_{7+} + 0.002)] + A4T_f \\ & + A5(xC_{7+}M_{c7+}) + A6(xC_{7+}M_{c7+})^2 + A7(xC_{7+}M_{c7+})^3 \\ & + A8[M_{c7+}/(\gamma_{c7+} + 0.0001)] + A9[M_{c7+}/(\gamma_{c7+} + 0.0001)]^2 \\ & + A10[M_{c7+}/(\gamma_{c7+} + 0.0001)]^3 + A11 \end{aligned}$$

Where:

$$\begin{aligned} A1 &= -2.0623054 & A2 &= 6.6259728 \\ A3 &= -4.4670559 \times 10^{-3} & A4 &= -1.0448346 \times 10^{-4} \\ A5 &= 3.2673714 \times 10^{-2} & A6 &= -3.6453277 \times 10^{-3} \\ A7 &= -7.4299951 \times 10^{-5} & A8 &= -1.1381195 \times 10^{-1} \\ A9 &= -6.2476497 \times 10^{-4} & A10 &= -1.1381195 \times 10^{-1} \\ A11 &= -1.0746622 \times 10 \end{aligned}$$

2. Elsharkawy (2001)

$$\begin{aligned} P_d = & A0 + A1T_f + A2xH_2S + A3xCO_2 + A4xN_2 + A5xC_1 + A6xC_2 + A7xC_3 \\ & + A8xC_4 + A9xC_5 + A10xC_6 + A11xC_{7+} + A12M_{c7+} + A13\gamma_{c7+} \\ & + A14(xC_{7+}M_{c7+}) + A15\left(\frac{M_{c7+}}{\gamma_{c7+}}\right) + A16\left(\frac{xC_{7+}M_{c7+}}{\gamma_{c7+}}\right) \\ & + A17\left[\frac{xC_{7+}}{(xC_1 + xC_2)}\right] + A18\left[\frac{xC_{7+}}{(xC_2 + xC_3 + xC_4 + xC_5 + xC_6)}\right] \end{aligned}$$

Where:

A0 = 4268.85	A1 = 0.094056
A2 = -7157.87	A3 = -4540.58
A4 = -4663.55	A5 = -1357.56
A6 = -7776.10	A7 = -9967.99
A8 = -4257.10	A9 = -1417.10
A10 = 691.5298	A11 = 40660.36
A12 = 205.26	A13 = -7260.32
A14 = -352.413	A15 = -114.519
A16 = 8.13300	A17 = 94.916
A18 = 238.252	

### 3. Humoud and Al-Marhoun (2001)

$$\ln(P_d) = \beta_0 + \beta_1 \ln(T) + \beta_2 \ln(R_m) + \beta_3 \ln(P_{SP} \cdot T_{SP}) + \frac{\beta_4}{T_{pr}} + \frac{\beta_5}{P_{pr}} + \frac{\beta_6}{\gamma_{C7+}}$$

Where,

$$\beta_0 = 43.777183, \beta_1 = -3.594131, \beta_2 = -0.247436, \beta_3 = -0.053527$$

$$\beta_4 = -4.291404, \beta_5 = -3.698703, \beta_6 = -4.590091$$

$$R_m = \frac{R_{SP} \cdot \gamma_{gSP}}{\gamma_{C7+}}$$

## 4. Marruffo, Maita, Him and Rojas (2002)

$$P_d = K1 \left[ \frac{GCR^{K2}}{\%C_{7+}^{K3}} \times K8 \times API^{(K4 \times T_r^{K5} - K6 \times C_{7+}^{K7})} \right]$$

Where:

$$K1 = 346.7764689$$

$$K2 = 0.0974139$$

$$K3 = -0.294782419$$

$$K4 = -0.047833243$$

$$K5 = 0.281255219$$

$$K6 = 0.00068358$$

$$K7 = 1.906328237$$

$$K8 = 8.4176216$$

$$\%C_{7+} = \left( \frac{GCR}{70680} \right)^{-0.8207}$$

### Nomenclature

$\gamma_{c7+}$ : heptane-plus specific gravity

$\gamma_{gsp}$ : primary separator gas specific gravity

API: API Condensate gravity

GCR: Gas Condensate ratio (SCF/STB)

$M_{c7+}$ : heptane-plus molecular weight

$P_d$ : Dew-point pressure (psi)

$P_{pr}$ : pseudoreduced pressure

$P_{sp}$ : primary separator pressure (psi)

$R_{sp}$ : Primary separator gas-oil ratio (scf/Sp bbl)

$T_{pr}$ : pseudoreduced temperature

$T_{sp}$ : primary separator temperature ( $^{\circ}R$ )

$T_R$ : Reservoir temperature ( $^{\circ}R$ )

$T_f$ : Reservoir temperature ( $^{\circ}F$ )



## Vitae

Name: Mohammad Abdullah Al-Dhamen  
Nationality: Saudi  
Field: Petroleum Engineering

Address: P.O.Box 12636  
Dhahran 31311  
Saudi Arabia  
E-mail: [mohammad.dhamen@aramco.com](mailto:mohammad.dhamen@aramco.com)  
Mobile: (+966) 506845571

Education: 2007: B.S in Petroleum Engineering, King Fahd University of  
Petroleum & Minerals  
2010: M.S in Petroleum Engineering, King Fahd University of  
Petroleum & Minerals

Occupation: Petroleum Engineer in Aramco

Experience: 2007 – 2009 working in gas reservoir management  
2009-2010 working in gas production engineering  
2010-2011 working in gas drilling and work over engineering