Chapter 3
Methodology

3.1. Common Empirical Correlations

Since the 1940's engineers have realized the importance of developing empirical correlation for bubble point pressure. Studies carried out in this field resulted in the development of new correlations. Several studies of this kind were published by Katz, Standing, Lasater and Cronquist. For several years, these correlations were the only source available for estimating bubble point pressure when experimental data were unavailable. In the last thirty years there has been an increasing interest in developing new correlations for crude oils obtained from the various regions in the world. Vazquez & Beggs, Glaso, Al-Marhoun, McCain, Al Shammasi and Dindoruk carried out some of the recent studies.

The common empirical correlations that had been used in this research are:

3.1.1. Standing Correlation, (1947)

Standing used two stage flash liberation tests to obtain experimental data values of gas-oil ratio (at the bubble point),gas gravity ,oil gravity and oil formation volume factor (at P_b) were used for correlation purposes. The gases evolved during the flash separation experiments were essentially free of nitrogen (N_2) and hydrogen sulphide (H_2S) although a few samples did contain carbon dioxide (CO_2) in quantities of less than 1 mole percent. in short, standing correlations should be considered valid only for black oil systems with trace compositions of any non-hydrocarbon components.

$$P_b = 18.2[(R_s/\gamma_g)^{0.83}(10)^a - 1.4] \qquad (3-1)$$

$$a = 0.00091(T - 460) - 0.0125(API)$$

Where:

P_b = bubble-point pressure, psia

T = reservoir temperature, °R

API=oil gravity in API degree

γ_g = average specific gravity of the total surface gases

R_s = gas solubility, SCF/STB

3.1.2. Glaso Correlation, (1980)

The Glaso correlation contains equation for estimating bubble point pressure, solution gas oil ratio, oil formation volume factor for North Sea oils. The author claims that the correlations should be valid for all types of oil and gas mixture after correcting for non-hydrocarbons in the surface gases and the paraffinicity of the oil. According to the author, the correlation more accurately predicts the oil properties of North Sea oils than the standing correlations.

$$\log(p_b) = 1.7669 + 1.7447\log(p_b{}^*) - 0.30218[\log(P_b{}^*)]^2 \qquad (3-2)$$

With:

$$P_b{}^* = (R_s/\gamma_g)^a (t)^b (API)^c$$

Where:

R_s = gas solubility, SCF/STB

t = reservoir temperature, °F

γ_g = average specific gravity of the total surface gases

API=oil gravity in API degree

a, b, c = coefficients of the above equation having the following values:

a=0.816

b = 0.172

c = -0.989

3.1.3. Marhoun's Correlation, (1988)

Al-Marhoun published his correlation for determining bubble point pressure based on 160 data points from Middle East oil samples. The following general relation of bubble point pressure was proposed:

$$P_b = f(R_s, API, \gamma_g, T)$$

In this model, the bubble point pressure is predicted as a direct function of solution gas-oil ratio, specific oil and gas gravity and temperature. Best results were obtained by multiple regression analysis from the following empirical relation:

$$p_b = 5.338088 * 10^{-3}\ R_s^{0.715082} \gamma_g^{-1.87784} \gamma_o^{3.1437}\ T^{1.32657} \qquad (3-3)$$

Where:

T= temperature, $^\circ$ R

γ_o= stock-tank oil specific gravity

γ_g = gas specific gravity

R_s = gas solubility, SCF/STB

p_b = bubble point pressure, psia

3.1.4. Petrosky and Farshad Correlation, (1993)

The Petrosky and Farshad correlation equation for estimating bubble point pressure for Gulf of Mexico, the correlation was developed using fluid sample taken from offshore region in Texas and Louisiana (Galveston island eastern through main pass), the authors claims that the correlation provide improved result over other correlation for the Gulf of Mexico, including those published by Standing, Vasgues and Beggs, Glaso, Al-Marhoun.

$$p_b = \left[\frac{112.727 R_s^{0.577421}}{\gamma_g^{0.8439}.(10)^x}\right] - 1391.051 \qquad (3-4)$$

$$x = 7.916\,(10^{-4})\,(API)^{1.5410} - 4.561(10^{-5})\,(T - 460)^{1.3911}$$

Where:

p_b = bubble point pressure, psia

T = temperature, $^\circ$ R

γ_g = gas specific gravity

R_s = gas solubility, SCF/STB

API=oil gravity in API degree

3.1.5. Hanafy et.al, (1997)

Consist of equations for estimating bubble point pressure, solution gas-oil ratio, and oil formation volume factor and oil compressibility, oil density and oil viscosity for Egyptian crude oil.

$$p_b = 3.205 * R_s + 157.27 \qquad (3-5)$$

Where:

P_b = bubble-point pressure, psia

R_S =gas solubility, SCF/STB

3.1.6. The Vasquez-Beggs Correlation

The correlation was developed from 600 laboratory PVT analyses from fields all over the world .the data used in development of the correlation covers a wide range of pressure, temperature, and oil properties. The correlation divides the data into two groups: one for oil gravity over 30 API and one at below 30 API

$$p_b = \left(\frac{R_S}{C_1 . \gamma_g . EXP\left(C_3\left(\frac{\gamma_o}{T + 459.67}\right)\right)}\right)^{\left(\frac{1}{C_2}\right)} \qquad (3-6)$$

Where:

Table 3-1: Vasquez-Beggs coefficients,(Tarek Ahmed, 2001)

Coefficient	$\gamma_o \leq 30°API$	$\gamma_o > 30°$ **API**
C_1	0.0362	0.0178
C_2	1.0937	1.1870
C_3	25.7240	23.9310

R_S =gas solubility, SCF/STB

T = temperature, ° F

γ_o= oil specific gravity

In this research, those common empirical correlations will applied for all datasets (212 datasets). the measured and calculated values will be evaluated by using 45 degree fitting method and statistical analysis (mentioned in part 3.3).

3.2. Development of new correlation using polynomial neural network (PNN)

In this research VariReg software (see appendix A) was used, it is a software tool for general-purpose multidimensional regression modeling with the main emphasis on methods used in surrogate modeling, VariReg is primarily intended for use on small and moderately sized numerical data sets, Gints Jekabsons at the Riga Technical University was developed it.

The polynomial neural network (PNN) is a part of tools was included by VariReg.

3.2.1 Polynomial Neural Networks (PNN)

The PNN method is mainly referred as a method for self-organizing polynomial neural networks. The most widely known of its variations work exclusively with polynomials and therefore also the result of the method can be written in polynomial form.

The building blocks of group method of data handling (GMDH) usually are second or third degree polynomials of two or three input variables (See Figure 3-1). Such building blocks, also called partial descriptions (PDs), like neurons in neural networks, are arranged in layers. The coefficients of each PD are calculated using the Ordinary Least-Squares (OLS) by trying to approximate the original dependent variable **y** of training data. The exact number of layers and connections of the network as well as the structure of PDs is not set a priori but is the object of search layer by layer. The maximal number of PDs selected in each layer is usually set equal to the number of original input variables (Gints Jekabsons, 2010)

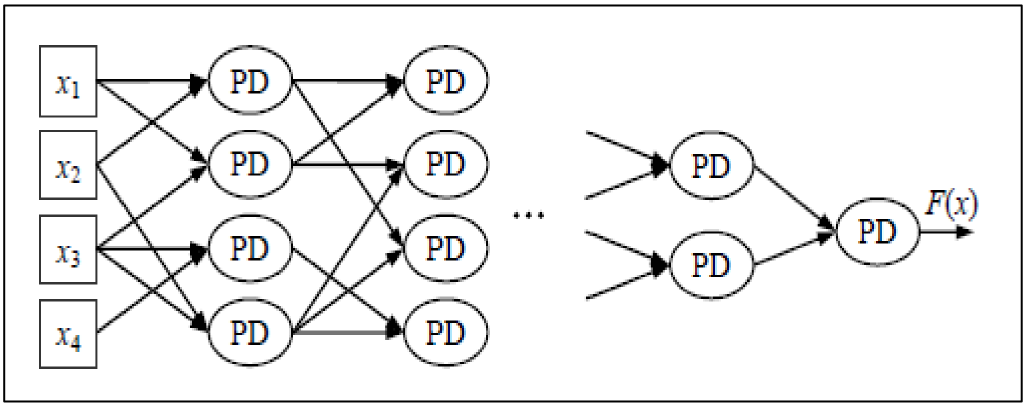

Figure 3-1: An example of GMDH polynomial neural network structure(Gints Jekabsons, 2010)

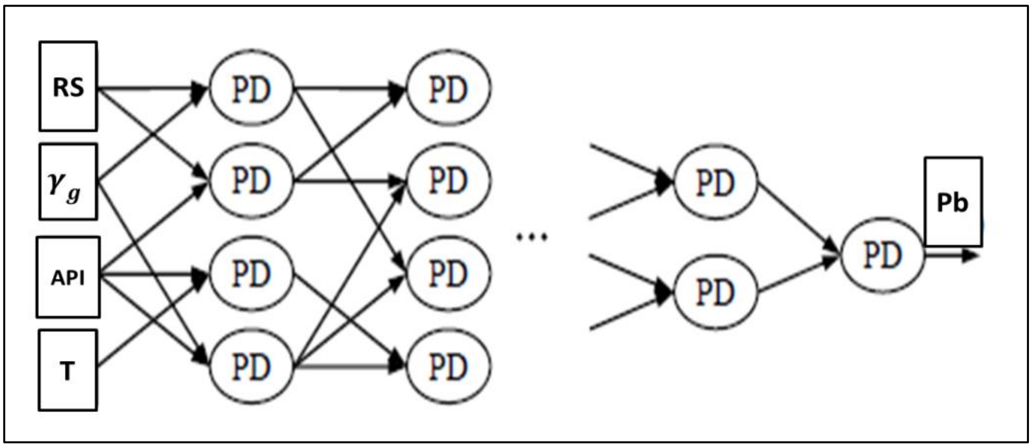

Figure 3-2: Polynomial neural network structure with two inputs for PDs scenario from this case study

In this study, after the loading the data(see appendix A), to achieve best model we need to adjust the Maximum degree for polynomials in PDs, Maximum number of inputs for PDs (2 or 3) (See Figure 3-2 for using datasets as inputs data) , whether the PDs are full polynomials or their structure is generated by a subset selection Algorithm, Criterion for subset selection and stopping of network building (note that in VariReg the Complexity of the network is estimated by calculating the sum of the number of parameters in all the PDs which are directly or indirectly connected to the best PD in the last layer), The maximum number of PDs in each layer, and whether in each successive layer PDs take inputs from PDs of immediately preceding layer or also from the original input variables (see appendix A).

The new model will be generated by using 70% of all datasets. To test the new model, 30% of datasets will be used.

Statistical analysis for bubble point pressure prediction will be done for the best empirical correlations and the new model.

3.3. Statistical Analysis (Error Analysis)

Sum of Squared Error(SSE),Mean Squared Error (MSE),Root Mean Squared Error(RMSE) Standard Deviation (STD),Relative Root Mean, Squared Error (RRMSE) and correlation coefficient (R^2) are used to compare and evaluate the prediction ability of correlations, which are defined as below:

3.3.1. Sum of Squared Error

$$\mathrm{SSE} = \sum_{i=1}^{n} ((\mathrm{y}_{(i)} - \mathrm{F}(\mathrm{X}_{(i)}))^2 \qquad (3-7)$$

Where:

$y_{(i)} \equiv$ bubble point pressure P_b obtained experimentally

$F(X_{(i)}) \equiv$ Predicted bubble P_b point pressure

n = number of data points

3.3.2. Mean Squared Error

$$\mathrm{MSE} = \frac{\mathrm{SSE}}{\mathrm{n}} = \frac{1}{\mathrm{n}}\sum_{i=1}^{n} ((\mathrm{y}_{(i)} - \mathrm{F}(\mathrm{X}_{(i)}))^2 \qquad (3-8)$$

Where:

$y_{(i)} \equiv$ bubble point pressure P_b obtained experimentally

$F(X_{(i)}) \equiv$ Predicted bubble P_b point pressure

n = number of data points

3.3.3. Relative Mean of Squared Error

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}} = \sqrt{\frac{1}{\mathrm{n}}\sum_{i=1}^{n} ((\mathrm{y}_{(i)} - \mathrm{F}(\mathrm{X}_{(i)}))^2} \qquad (3-9)$$

Where:

$y_{(i)} \equiv$ bubble point pressure P_b obtained experimentally

$F(X_{(i)}) \equiv$ Predicted bubble P_b point pressure

n = number of data points

3.3.4. Standard Divination

$$\mathrm{STD} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{(i)} - \bar{y})^2} \qquad (3-10)$$

Where:

$y_{(i)} \equiv$ bubble point pressure P_b obtained experimentally

$F(X_{(i)}) \equiv$ Predicted bubble P_b point pressure

n = number of data points

3.3.5. Relative Root of Mean Squared Error

$$\mathrm{PRMSE} = \frac{\mathrm{RMSE}}{\mathrm{STD}} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}((y_{(i)} - F(X_{(i)}))^2}}{\frac{1}{n}\sum_{i=1}^{n}(y_{(i)} - \bar{y})^2} \qquad (3-11)$$

Where:

$y_{(i)} \equiv$ bubble point pressure P_b obtained experimentally

$F(X_{(i)}) \equiv$ Predicted bubble P_b point pressure

n = number of data points

3.3.6. Variance

$$\mathrm{VAR} = \frac{1}{n}\sum_{i=1}^{n}(y_{(i)} - \bar{y})^2 \qquad (3-12)$$

Where:

$y_{(i)} \equiv$ bubble point pressure P_b obtained experimentally

$F(X_{(i)}) \equiv$ Predicted bubble P_b point pressure

n = number of data points

3.3.7. Correlation Factor

$$R^2 = 1 - \frac{MSE}{VAR} = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n}((y_{(i)} - F(X_{(i)}))^2}{\frac{1}{n}\sum_{i=1}^{n}(y_{(i)} - \bar{y})^2} \qquad (3-13)$$

Where:

$y_{(i)} \equiv$ bubble point pressure P_b obtained experimentally

$F(X_{(i)}) \equiv$ Predicted bubble P_b point pressure

n = number of data points

3.4. Creating Guide User Interface (GUI) using MATLAB Software

MATLAB (matrix laboratory) is a multi-paradigm numerical computing environment and fourth-generation programming language. A proprietary programming language developed by MathWorks, MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of *guide user interfaces* (GUI), and interfacing with programs written in other languages, including C, C++, C#, Java, Fortran and Python.

In 2004, MATLAB had around one million users across industry and academia. MATLAB users come from various backgrounds of engineering, science, and economics. (MathWorks, 2013)

GUIs (also known as graphical user interfaces or GUIs) provide point-and-click control of software applications, eliminating the need to learn a language or type commands in order to run the application.

MATLAB Software is self-contained MATLAB programs with GUI front ends (See Figure3-3) that automate a task or calculation. The GUI typically contains controls such as menus, toolbars, buttons, and sliders. Many MATLAB products, such as Curve Fitting Toolbox, Signal Processing Toolbox, and Control System Toolbox include apps with custom user interfaces. You can also create your own custom apps, including their corresponding UIs, for others to use. (MathWorks, 2014)

In this Study, the common empirical correlations as well as the new developed model will be programmed in MATLAB software and creating GUI by helping of professional programmer for bubble point pressure evaluation and converting it into window standalone application (exe file extension).

The bubble point pressure evaluation workflow for the GUI shown in Figure 3-4

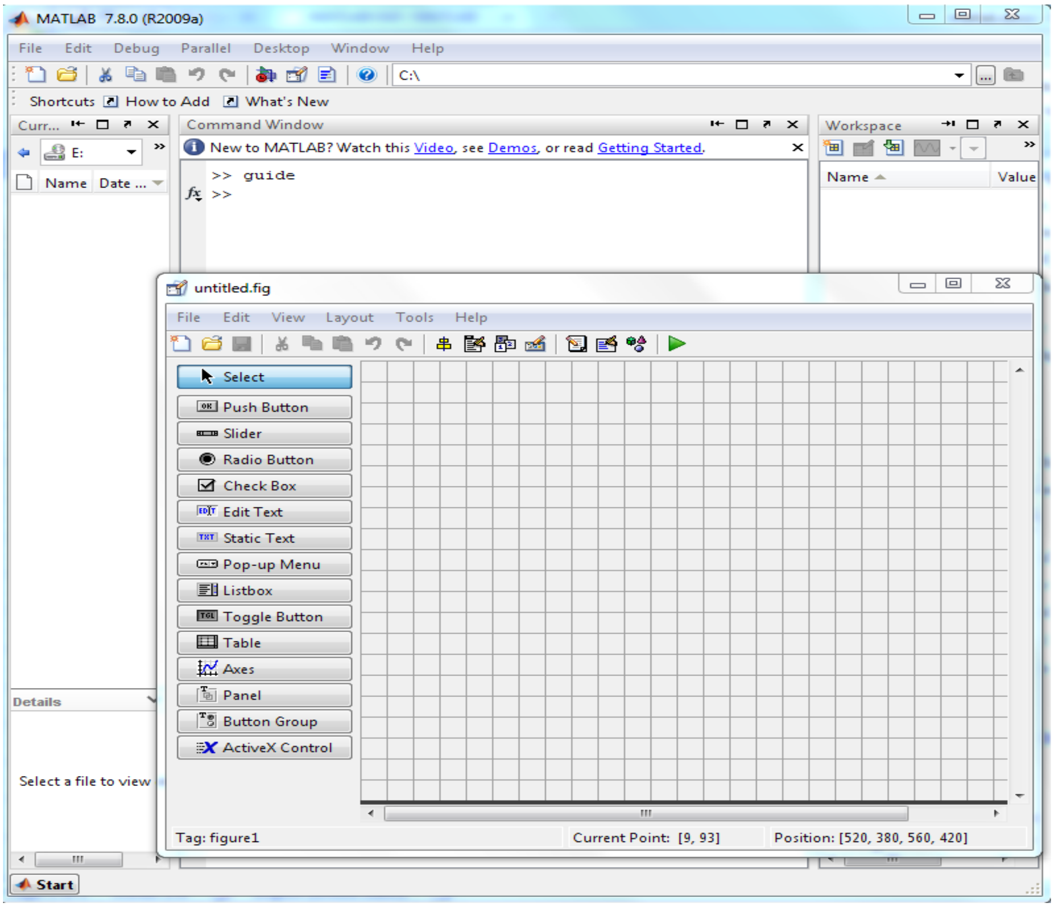

Figure (3-3): New Guide user interface (GUI) in Matlab R2009a software.

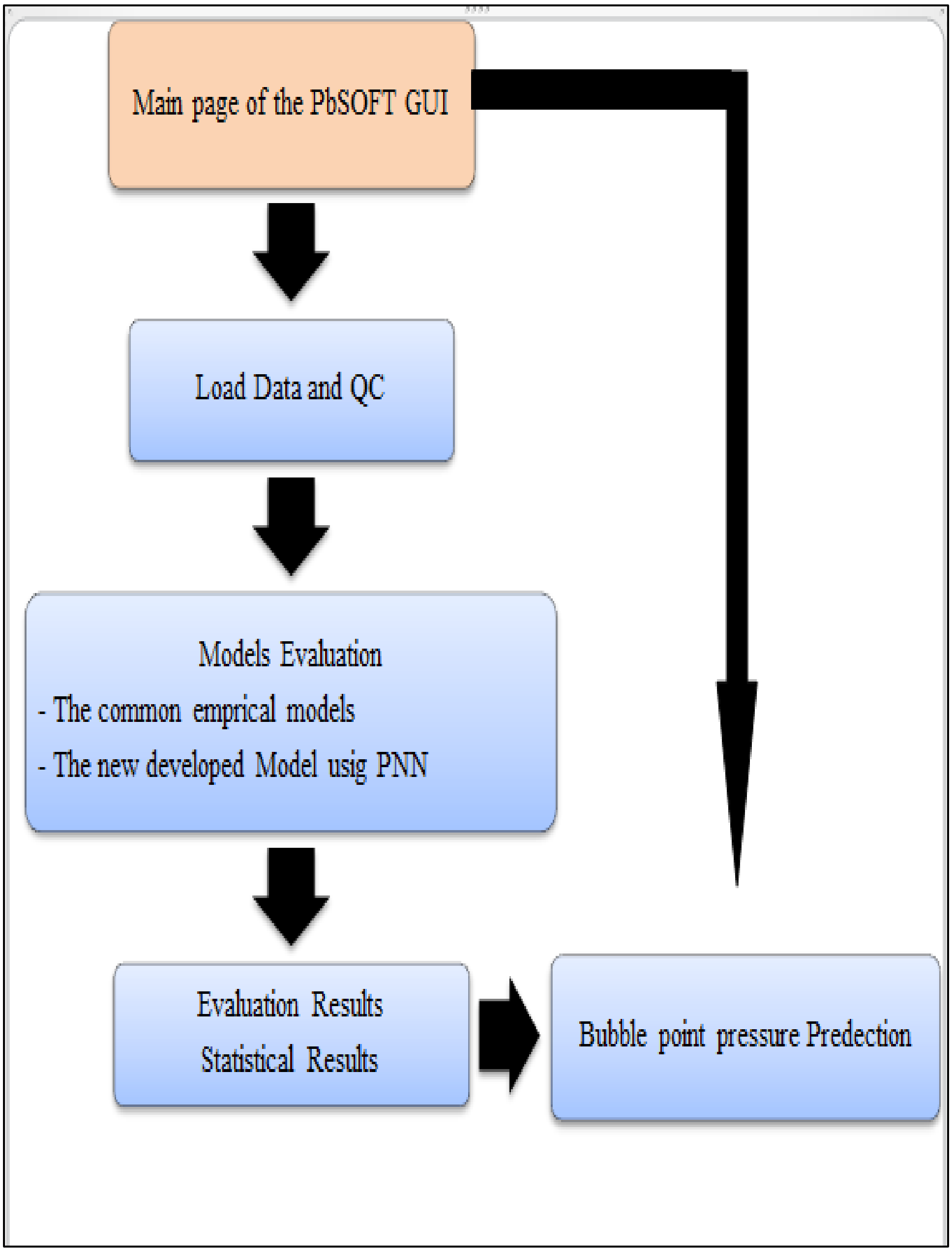

Figure 3-4 show the workflow of the created GUI.