CHAPTER II
LITERATURE REVIEW

2.1 General Reservoir Fluids Physical Properties

Physical properties of reservoir fluids can either describe physical behavior of the fluids in the reservoir or relate the surface volumes to reservoir volumes. These properties can be determined through laboratory testing on fluid samples, and field production data. In addition, they can be calculated from empirical correlations with respect to available information. Some general reservoir fluids properties are described as follow.

2.1.1 Bubble Point Pressure (P_b, psia)

P_b is the pressure at which the first bubble of gas evolves from the reservoir fluid as the crude oil pressure decreases. It can also be called as saturation pressure. P_b is various depending on temperature and type of the reservoir fluid system (Beggs, 1987).

2.1.2 Oil Formation Volume Factor (B_o, bbl/STB)

B_o is the ratio between the volume of oil plus solution gas occupied at the reservoir conditions and the volume of the oil at stock-tank conditions (atmospheric pressure of 14.696 psi, temperature of 60 °F). B_o is a measure of the oil-shrinkage as from the reservoir conditions to the surface conditions (Cosentino, 2001).

2.1.3 Viscosity (μ, cp)

μ is known as internal resistance of fluid to flow. For reservoir fluids, viscosity parameter is needed to describe how the fluids flow in reservoir. It is expressed in centipoises (cp). There are two categories of oil viscosity, which are dead oil viscosity (μ_{od}) and live oil viscosity (μ_{ol}). The μ_{ol} can be classified into two types, which are bubble point oil viscosity (μ_{ob}) and undersaturated oil viscosity (μ_o). Moreover, the viscosity of oil in saturated pressure range is called saturated oil viscosity (μ_{sat}).

2.1.4 Solution Gas Oil Ratio (R_s, scf/stb)

R_s is the production ratio between the volume of gas evolving from the oil and the volume of the oil at the surface with standard conditions.

2.1.5 Oil and Gas Specific Gravity (γ_o, γ_g)

γ_o is ratio of the crude oil density to the water density, while γ_g is the ratio of the gas density to the air density. They both are measured at the same condition.

2.1.6 API Gravity (American Petroleum Institute, °API)

API is an inverse measure of the γ_o. It is used to compare the crude oil as the different type of crude oil can have different value of API. It is defined as Equation 2.1.

$$°API = \frac{141.5}{\gamma_o} - 131.5 \qquad (2.1)$$

2.2 Fluid Samplings

For determining crude oil properties through laboratory testing, it is essential to have a good sample. In order to obtain the crude oil sample representing its reservoir conditions, care must be taken. Firstly, the reservoir must be checked for sampling in the early life of the reservoir before the occurrence of significant loss in pressure and the well should have steady productivity with stabilized gas oil ratio and no water cut (Moses, 1986). Secondly, these wells should be checked if they are properly conditioned before sampling for original fluid flow (Reudelhuber, 1957). There are two main methods to collect reservoir fluid samples as follow.

2.2.1 Subsurface or Bottom Hole Sample

This type of sample is collected by wire line equipment in the bore hole at the actual depth of reservoir where it can represent the bottom hole condition. For the accurate result, the sample must retain the reservoir conditions at all time using special chambers within the sampling tube to compensate the pressure of the sample while the sample is being returned to the surface. In case of undersaturated reservoir (reservoir with pressure is staying above bubble point pressure) with the flowing pressure at the wellbore above bubble point pressure, it would be considered as simple sampling case since the sample can be collected when the well is still flowing. Nevertheless, this is not the case for the saturated reservoir or slightly undersaturated reservoir, which is the major problem in subsurface sampling, the well may have to be shut-in to allow the accumulated excess gas caused by pressure drop in the bore hole (Katz, 1938) to be reabsorbed before the sample can be collected (Consentino, 2001).

2.2.2 Surface or Recombined Fluid Sample

This type of sample can be obtained faster and more convenient than subsurface sample. Oil and gas, which are collected separately at the wellhead condition, separator condition, or stock tank condition, will be recombined in the laboratory. However, it is difficult to obtain the result that corresponding to the reservoir conditions. Some of the gas could be liberated from the oil at the surface pressure and large sample of separator gas is required to be collected. Therefore, to ensure that recombined fluid samples can represent its original characteristic, the accuracy of the separator flow measurements and the stability of separator conditions are required.

2.3 Laboratory Testing

2.3.1 PVT Analysis

PVT analysis will be performed at laboratory where the reservoir fluid sample will be tested for its phase behavior through different expansion stages as it undergoes from the reservoir to the stock tank conditions. There are three main PVT laboratory tests for reservoir fluid samples, which are:

2.3.1.1 Flash Expansion

Flash expansion or flash vaporization test is used to define P_b or PVT properties at different pressure step in isothermal expansion process. First, the sample is filled into the PV cell and the pressure and temperature will be increased to the initial reservoir conditions as the gas will be dissolved into the oil. Then the pressure in the cell is decreased through isothermal expansion process while the data is being recorded. When P_b is reached, the first bubble of gas will be liberated from the oil. The pressure will be decreased continuously and phase behavior of the sample will be observed until there is no additional gas liberated from the oil. Thus, for the flash expansion test, no fluid will be withdrawn from the cell for the entire experiment (Cosentino, 2001).

2.3.1.2 Differential Vaporization

Differential vaporization test is used to determine B_o, R_s, γo, and γ_g. The procedure of this test is similar to the flash expansion test from the beginning until the bubble point pressure is reached. The gas liberated from the oil at each pressure step is removed from the cell as the changing in composition of the sample causes the oil to shrink (Cosentino, 2001).

2.3.1.3 Flash Separator Tests

Flash separator tests are used to determine the phase behavior of the reservoir fluid at the surface as it travels through the separator(s) to the stock tank as it is not isothermal expansion process. First, the PVT cell is connected to separator(s) system. Then, the sample is flashed through the system to stock tank conditions. At the end of the experiment, the volume of gas and oil will be measured (Cosentino, 2001).

2.3.2 Viscosity Test

Viscosity of crude oil can be measured within laboratory using viscometer or in the bore hole using wire line equipments. In case of laboratory measurement, for oil viscosity, rolling ball viscometer or capillary viscometer is usually used. Since the viscosity is a strong function of pressure and temperature, crude oil should be tested at constant conditions repeatedly for accurate result. Therefore, it is very tedious to obtain gas viscosity since it is difficult to obtain the accurate result. Thus, gas viscosity is usually estimated using correlations as function of specific gravities data obtained from differential vaporization (Mccain, 1990).

2.3.3 Compositional Analysis

Compositional data as well as physical data in separator or stock tank conditions are needed for determining physical properties of reservoir fluid using equation of state or correlation involving composition data as input parameters. Gas or liquid chromatography techniques are commonly used in laboratory in order to measure fluid composition, specific gravity, and molecular weight of oil and gas.

2.4 Empirical Models and Correlation Evaluations

Since the last century, many empirical correlations have been developed and used for determining crude oil properties. Generally, there are two main types of correlations for predicting crude oil properties. The first type is those correlations that use available oil field data as the input parameters, such as reservoir temperature (T_{res}), P_b, API, γ_g, and R_s. The second type is the correlations that use some parameters apart from the first type, such as reservoir fluid composition, pour point temperature, molecular weight, normal boiling point, critical temperature, and acentric factor of components (Coats and Smart, 1986; Houpeurt and Thelliez, 1976; Lasater, 1958; Little and Kennedy, 1968; Lohrenz et al., 1964; Ng and Egbogah, 1983); Wu and Rosenegger (1999, 2000); (Wu et al., 2003). Nevertheless, the availability of the composition data is usually insufficient initially, and even the data are available, they could not fully identify all of the components in the crude oil. The input parameters for the correlations from the first type are typically used to approximate the effect of the composition on these PVT properties (Almehaideb, 1997). There are three main approaches for developing these correlations, such as universal correlation based on worldwide data sets, regional empirical correlation based on data set in each region, and correlations based on specific type of crude oil or range of crude oil conditions. Since characteristic of the crude oil and the range of data in each region are different as the global correlations are very difficult to obtain (Dokla and Osman, 1992; Dokla and Osman, 1991), large volume of correlation developments and evaluation studies have been presented to outperform each other.

D.L.Katz (1942) introduced graphical correlations based from PVT measurements on 117 crude oils for predicting B_o as functions of P, T, R_s, γ_o and γ_g.

However, the correlations were hard to use since the correlations were combination of calculations and graphs.

Beal (1946) presented graphical correlations for determining μ_{od}, μ_{ob}, and R_s based on data mostly from the United States.

Standing (1947) developed correlations and charts for determining P_b, B_{ob} and total formation volume factor (B_t) using 105 data samples from 22 crude oils obtained from California, United States. These correlations are the most widely used in the petroleum industry since they are the first correlations to use T_{res}, γ_o, γ_g, and R_s as input parameters. Later, Lasater (1958) used Standing's basic assumption with additional of derived gas mole fraction term based on standard physical chemical equations of solution to develop P_b correlation (presented in lookup chart form) for black oil systems based on 158 data points collected from 137 reservoirs from Canada, United States, and South America.

Chew and Connally Jr. (1959) presented a μ_{ol} correlation as a function of μ_{od} and R_s based on crude oil data from the United States, Canada, and South America. The relation between μ_{ol} and corresponding μ_{od} at fixed R_s was reported to be a straight line on logarithmic coordinate.

Braden (1966) used data set collected from 8 crude oils and 7 processed oils to develop μ_{od} correlation as a function of API measured at 60 °F and viscosity at one temperature.

Cronquist (1973) presented a dimensionless graphical correlations based on 80 data points from 30 Gulf Coast reservoirs for analysis of depletion-drive reservoirs.

Beggs and Robinson (1975) presented μ_{od} and μ_{ol} correlations following the methods from Beal (1946) and Chew and Connally Jr. (1959) respectively. The developed μ_{od} correlation was a function of API and T, while the developed μ_{ol} correlation was a function of R_s and μ_{od}.

Glaso (1980) developed P_b, B_{ob}, and B_t correlations using data from 45 flash separator oil samples mostly from the North Sea. These correlations were functions of T, Rs, γ_g, and API. Correction factors from some non-hydrocarbon gases and a correction factor for paraffinicity of oil were used with these equations in order to be valid for all types of reservoir fluids worldwide.

Vazquez and Beggs (1980) categorized crude oils into two ranges (oil with API above and below 30 °API) and developed correlations for predicting R_s, μ_o, and B_o. The γ_g was stated to be a strong correlating factor for PVT properties but it was very sensitive to the conditions where it was determined. More than 6,000 data points from 600 laboratory test were used. All the γ_g measurements were converted into the reference separator conditions of 100 psig as the method to develop these correlations.

Ostermann and Owolabi (1983) employed crude oil data from Alaska to evaluate various correlations for prediction of P_b, B_{ob}, μ_{od}, and μ_{ol}.

Khan (1987); Khan et al. (1987) developed correlations for determining μ_{ob}, μ_{sat}, and μ_{ob} of Saudi Arabian crude oil based on flash separation data obtained from 75 bottom-hole samples. Based on the data used in the work, their correlations (Khan (1987); Khan et al. (1987)) were reported to be the most accurate. However, the μ_{ob} correlation from Beggs and Robinson (1975) was found to be acceptable, and the μ_o correlation from Beal (1964) was reported to have good estimation performance.

Saleh et al. (1987) evaluated the accuracy of the existed PVT correlations in that time to determine P_b, B_o, R_s, B_t, and μ_{od} employing data from Gulf of Suez and Western Desert field, Egypt.

Al-Marhoun (1988) proposed regional correlations and nomographs developed from 160 data points obtained from 69 Middle Eastern reservoirs for estimation of P_b, B_{ob}, and B_t as the functions of T, γ_g, R_s, and γ_o.

Abdul-Majeed et al. (1988) presented B_{ob} correlation as a function of R_s, γ_g, γ_o, and T, using 420 experimental data points resulted from 119 samples obtained from different unpublished sources in Iraq.

Sutton and Farshad (1990) compared the accuracy of various PVT correlations for application in Gulf of Mexico. Correlations from Glaso (1980) including P_b correlation for crude oil with P_b below 7,000 psi, R_s correlation for crude oil with R_s below 1,400 scf/stb, B_o correlation, and μ_{od} correlation were recommended. Correlations from Vazquez and Beggs (1980) including μ_{ob} and μ_o correlations were recommended for crude oil with P_b above 7,000 psi and R_s above 1,400 scf/stb. Moreover, μ_{od} correlation from Beggs and Robinson (1975) were also recommended.

Abdul-Majeed et al. (1990) presented μ_{od} correlation as a function of P, P_b, μ_{ob}, bubble point R_s, and API using data from North Africa and Middle East.

Labedi (1990b) used crude data from Libya, Nigeria and Angola to generate γ_g, P_b, R_s, and composition correlations. The developed correlations are functions of measurable production parameters. Later, Labedi also presented μ_{od}, μ_{ob}, μ_{sat}, and μ_o correlations for light oil (Labedi, 1992).

Dokla and Osman (1991) used original form of Al-Marhoun (1988)'s correlations to develop P_b and B_o correlations for UAE crude based on 51 data points. However, the modified P_b correlation was later found to contradict physical trends (Al-Shammasi, 2001; Dokla and Osman, 1992).

Al-Marhoun (1992) presented correlations for predicting B_{ob}, B_o, and B_t as functions of T, P, R_s, γ_g, and γ_o using experimentally obtained field data from all over the world.

Miadonye et al. (1992) presented one-parameter equation for estimating viscosity of bitumen, heavy crude oil, and high viscosity oil using viscosity measurement at reference temperature of 30°C generated from on data from Alberta, Canada.

Omar and Todd (1993) presented black oil correlations based on Standing's approaches (Standing, 1947) for predicting P_b and B_{ob} using data from Malaysian crudes.

Petrosky Jr. and Farshad (1993) employed 81 laboratory PVT analyses on crude samples obtained from offshore Texas and Louisiana to develop correlations for estimating P_b, R_s, B_{ob}, and c_o for crude oil from Gulf of Mexico.

Kartoatmodjo and Schmidt (1994); Kartoatmodjo and Schmidt (1991) used global data to develop PVT correlations for determining B_o, Rs, P_b, μ_{od}, μ_{sat}, μ_o, and isothermal compressibility of undersaturated oil (c_o, psi^{-1}) as functions of field parameters. In addition, a correction factor of γ_g and a conversion factor from flash data to differential data were also developed.

De Ghetto and Villa (1994) developed modified correlations for determining of P_b, R_s, B_o, c_o, μ_{od}, μ_{ob}, and μ_o for four different API classes using data from Mediterranean Basin, Africa, Persian Gulf, and North Sea. Since the extra-heavy oil correlations except viscosity were reported to be unavailable, the new equations for extra-heavy oil were proposed.

Petrosky and Farshad (1995) proposed μ_{od}, μ_{ob}, and μ_o correlations based on 126 bubble point data from differential liberation and two-stage separation tests. These correlations were developed especially for predicting Gulf of Mexico crude oils.

De Ghetto et al. (1995) presented heavy and extra heavy oil correlations for predicting P_b, R_s, c_o, μ_{od}, μ_{ob}, and μ_o using data from Mediterranean Basin, Africa and Persian Gulf.

Frashad et al. (1996) presented correlations for estimating P_b, R_s, B_o, and c_o of Columbian crude oil using different sets of data obtained from Columbian crude analyses.

Mahmood and Al-Marhoun (1996) evaluated empirical correlations for predicting P_b, B_{ob}, c_o, and μ_o using a total of 185 data points collected from 22 bottom hole samples of Pakistan.

Velarde et al. (1997) introduced P_b, R_s, B_{ob}, and B_o correlations developed from differential vaporization and separator data.

Almehaideb (1997) used crude oil data of U.A.E. to evaluate PVT correlations and developed correlations for estimation of B_{ob}, c_o, P_b, μ_{ob}, and μ_o.

Hanafy et al. (1997a) evaluated various correlations for predicting properties of crude oil using 324 data points collected from 75 fields from Egypt including the oil fields from Gulf of Suez, Western Dessert, and Sinai. These areas were reported the necessity for regional correlations to be developed. Therefore, Hanafy et al. (1997b) used the same data to develop correlations for determining P_b, R_s, B_o, c_o, ρ_o, and μ_{ob} of Egyptian crude oil. Moreover, the developed correlations were later improved by Hanafy et al. (2005).

Elsharkawy and Alikhan (1999) presented regional correlations for predicting μ_{od}, μ_{ob} and μ_o using 254 data points collected from Middle East.

Boukadi et al. (2002) presented correlations for predicting P_b, R_s, B_o, and μ_{ob} based on crude oil data from Oman. However, for P_b correlation, Standing (1947) correlation gave better result.

Valkó and Mccain Jr (2003) stated that it was not necessary to develop geographical correlations since carefully prepared universal correlations could give adequate results. Therefore, Valkó and Mccain Jr (2003) used large set of fluid property data obtained from a service company and presented universal correlations for predicting P_b (developed from a total of 1,745 data points), R_s (developed from a total of 881 data points), and γ_g (developed from a total of 618 data points).

Hashim and Hassaballah (2003) proposed viscosity-temperature correlation based on published viscosity measurements data. The correlation shown absolute average errors of 2.8% after it was applied to 64 viscosity measurements data points from Ahrabi et al. (1989).

Al-Marhoun (2004) evaluated correlations for Middle East crude oil using data obtained from flash vaporization, separator tests, viscosity measurements, and gas analysis on 186 bottom hole samples. Al-Marhoun (1988) correlations for P_b and R_s prediction were the recommended approaches. The B_o and B_t correlations from Al-Marhoun (1992) were recommended. For c_o correlation, Al-Marhoun (2003) correlation was recommended. For oil viscosities, Al-Marhoun (2004) also proposed μ_o correlation, and recommended μ_{ob} correlation from Beggs and Robinson (1975) and μ_{od} correlation from Glaso (1980). However, since high errors of μ_{ob} and μ_{od} correlations for Middle East crude oil, these properties were recommended for more research.

Dindoruk and Christman (2004) correlations for predicting P_b, R_s, B_o, c_o, μ_{od}, μ_{ob}, and μ_o for Gulf of Mexico crude oils. Although these correlations were similar to the correlations from Standing (1947), Petrosky Jr. and Farshad (1993), and Petrosky and Farshad (1995), these correlations gave better results over wider applicable range.

Naseri et al. (2005) used Iranian reservoirs data to develop μ_{od}, μ_{ob}, and μ_o correlations for Iranian crude oils.

Hossain et al. (2005) developed correlations for predicting μ_{od}, μ_{ob}, and μ_o for heavy oil (crude oils with API ranging from 10-22.3 °API) using three data sets from Chevron, Kartoatmodjo and Schmidt (1994), and De Ghetto et al. (1995). They also investigated the influences of wax, asphaltene, and resin contents on oil viscosity. However, Chevron data did not confirm the effects of these components as found in the experiment from Argillier et al. (2002).

Isehunwa et al. (2006) developed μ_{od}, μ_o, and μ_o correlations for light crude oil using data collected from over 400 oil reservoirs of the Niger Delta.

Ikiensikimama et al. (2006) evaluated numerous correlations for predicting oil properties of the Niger Delta crude. Al-Shammasi (2001)'s B_{ob} correlation, Lasater (1958)'s P_b correlation, and Dindoruk and Christman (2004)'s c_o correlation gave the best results. Petrosky Jr. and Farshad (1993)'s c_o correlation and Kartoatmodjo and Schmidt (1991)'s B_o correlation were reported to give the best results. Glaso (1980)'s B_t correlation was reported to give the best results, however, still lack of accuracy. Dindoruk and Christman (2004)'s μ_{od} correlations, Petrosky Jr. and Farshad (1993)'s μ_{ob} correlation, Beal (1946)'s μ_o correlation, and Khan (1987)'s μ_{sat} correlation were reported to give the best results.

Sutton and Bergman (2006) evaluated μ_o correlations based on 10,248 data points from 1,399 oil samples and found that the correlations from De Ghetto and Villa (1994) and Kartoatmodjo and Schmidt (1991) could give negative value of μ_o under high μ_{ob} value and high differential pressure. Moreover, μ_{od} correlation was proposed to use with the wider range of μ_{ob} and differential pressure.

Sutton (2006); Sutton (2008) presented correlations for bubble point oil density, undersaturated oil densitiy, B_{ob}, and c_o based on 11,960 data points obtained from 1,099 worldwide oil PVT reports. However, B_{ob} correlations from Velarde et al. (1997) and Frashad et al. (1996) were found to be more accurate.

Ugbe et al. (2006) developed generalized correlations with one adjustable parameter for predicting μ_{ob} and μ_o. The correlations were employed with Niger Delta crudes.

Hemmati and Kharrat (2007) presented P_b, B_{ob}, and R_s correlations for Iranian crude oil based on 287 data points collected from more than 30 oil fields.

Bergman and Sutton (2007a); Bergman and Sutton (2009) constructed a large database of 9,837 data points from 3,047 samples of crude oil, petroleum fractions, and pure component properties collected from both public and private sources to evaluate 23 μ_{od} correlations. Bergman's μ_{od} correlation (Whitson et al., 2000) was recommended. New μ_{od} method was also developed from the constructed database for results over a wide range of conditions. Another large worldwide database of 12,474 data points from 1,849 oils samples was also used by Bergman and Sutton (2007b) to evaluate 21 μ_{ob} methods. Bergman's μ_{ob} correlations (Whitson et al., 2000) were also recommended. However, another μ_{ob} method was also developed.

Zhang et al. (2007) proposed a simple correlation for predicting μ_{od} at different pressures as a function of T and μ_{od} measured at 50°C using heavy crude oil data from Liaohe Basin, NE China.

Bello et al. (2008) evaluated correlations for predicting P_b (Al-Marhoun, 1988; Dokla and Osman, 1992; Glaso, 1980; Lasater, 1958; Standing, 1947) and B_{ob} (Al-Marhoun, 1992; Dokla and Osman, 1992; Glaso, 1980; Labedi, 1990a; Standing, 1947) using 23 black oil PVT data points obtained from different oil fields in the Niger Delta. Since the errors resulted from this evaluation work, development of more accurate P_b and B_{ob} correlations were recommended for Niger Delta oils (Bello et al., 2008).

Ikiensikimama and Ogboja (2009) used functional form of the Lasater (1958) correlation with general relationship from Standing (1947) to develop correlation for predicting P_b for the Niger Delta crude based on 250 differential liberation reports from oil fields in the Niger Delta.

Okoduwa and Ikiensikimama (2010) divided Niger Delta crudes data into 5 different API ranges and developed correlations for determination of P_b at each API ranges.

Elmabrouk et al. (2010) proposed correlations in the absence of PVT analysis based on 476 data points collected from separator(s) tests with 118 reservoir fluid studies from Libyan oil fields in the Sirte Basin for predicting P_b and B_{ob} as functions of separator R_s (R_{SP}), separator pressure (P_{SP}), γ_o, and T_{res}.

Moradi et al. (2010) used a total of 1,801 global data sets (1,170 data sets from 9 published literature, 634 unpublished data sets from Iranian reservoirs) to develop P_b correlation.

Abedini et al. (2010) employed PVT data from 5 oil samples obtained from Iranian oil reservoirs and developed μ_o correlation as a function of P_b and μ_{od}.

Nikpoor and Khanamiri (2011) presented correlations for predicting P_b, R_s, and B_{ob} of Iranian crudes based on 421 differential liberation data points of southwestern Iranian oilfields. Nikpoor and Khanamiri (2012) also developed correlation for predicting μ_o using 572 data points from Iranian oilfields.

Alomair et al. (2011) presented μ_{od} correlation for Kuwaiti oils based on 360 data points from 33 heavy oil samples.

Singh and Hosein (2012) evaluated correlations for predicting P_b, B_{ob}, and R_s for Trinidad oils offshore the Southwest Coast. Velarde et al. (1997) correlations were recommended as they stated that there was no need to develop new correlations for Trinidad oils offshore the Southwest Coast.

Zahaby et al. (2012) used 35 Egyptian crude data to develop guidelines for choosing the best correlation for predicting black oil properties at bubble point, saturated, and undersaturated conditions.

Godefroy et al. (2012) compared and evaluated over 30 published correlations for predicting P_b, R_s, and B_{ob} to show how different parameters could affect the result with different correlations.

2.5 Artificial Neural Network

This section introduces artificial neural network, one of the most widely used artificial intelligent approaches for crude oil properties prediction. It has interconnected group of artificial neurons organized in a series of layers in order to transfer connection weights, which are the memory of the system between its processing elements (nodes) working in parallel. ANN is useful and robust method to solve numerous problems dealing with uncertain, inexact and obscure data where the relations between each parameter are unknown as formal analysis by human or conventional computer is difficult or impossible. ANN can be a tool for functional approximation, pattern recognition, nonlinear system identification, and control.

2.5.1 Inspiration of ANN

ANN is inspired by practical working aspect of the biological structure of nervous system as shown in Figure 2.1. A biological neuron comprises three components including cell body, dendrite, and axon. The dendrites are branches of nerve fibers that transfer and integrate electrical signals (synaptic inputs) received through other upstream neural cells at synapses areas (point of contacts) into the cell body (soma) hence the neuron is activated. The cell body has a role to sum the received synaptic inputs. Finally, the axon delivers the sum of signals from the cell body to dendrites of other neurons connected at their synapses (Hagan et al., 1996).

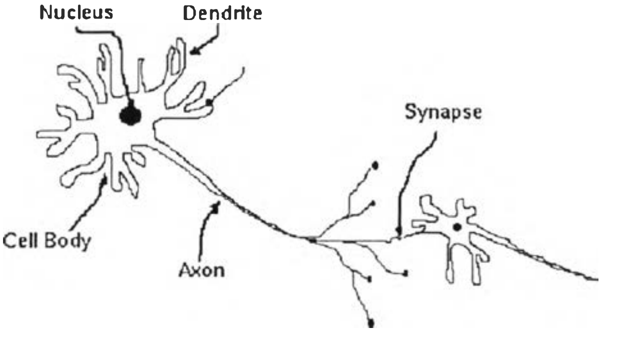

Figure 2.1 Biological structure of nervous system (Engelbrecht, 2007)

2.5.2 ANN Architecture

Generally, ANN is supervised system that can be adjusted or trained based on a given set of data to find patterns between its input and output as a result to desired target output. Like the human brain, in learning phase, ANN can accumulate knowledge and learn from the past experience from flows of information through the learning algorithm (Dutta and Gupta, 2010). Typically, large data set with input and output pairs is needed to train a network efficiently. ANN structure and a single neuron model are depicted in Figure 2.

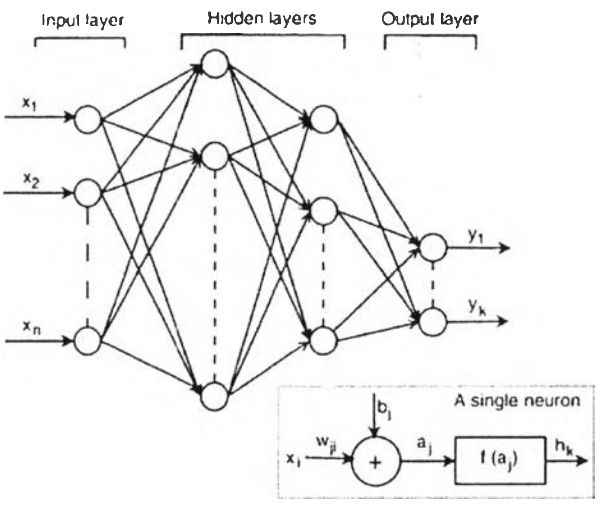

Figure 2.2 FFNN structure and a single neuron model (Dutta and Gupta, 2010)

Figure 2 presents feed-forward neural network (FFNN), one of the most popular ANN structures for function estimation. From the single neuron illustrated in Figure 2, input of the neuron (x_i) is multiplied by neuron weight (w_{ji}) and added with bias (b_j) to form net input (a_j). Then a_j is passed through transfer function ($f(a_j)$), which gives the neuron output (h_k). h_k can either be the input to other neurons in the next layer or the output of the ANN for the h_k from the output layer depending on the position of the neuron. During training phase, network error will be computed using selected training algorithm. Therefore, back propagation algorithm with momentum is the most commonly used as training algorithm for this type of FFNN for function estimation. Therefore, w_{ji} in each neuron will be adjusted depending on magnitude of error. Training will stop when neuron weights are adjusted until error cannot be further decreased; hence good agreement between input and output is achieved. Different types of transfer function, number of hidden-layer, number of neuron in each hidden-layer, and training algorithm, can be chosen at the beginning by the user for different objective (Hagan et al., 1996).

The ANN architecture has to be optimized with the size of the training data to avoid any overfitting ANN, which could give good results for only the data used in training phase. However, the overfiting ANN (or overtrained ANN) could not give a generalized result from the data from other sources (Al-Marhoun and Osman, 2002).

2.5.3 Transfer Functions

Different transfer function ($f(a_j)$) can be selected based on the problem of the particular ANN. Some of the transfer functions including linear transfer function (Purelin), Log-Sigmoid transfer function (Logsig), and Hyperbolic Tangent Sigmoid transfer function (Tansig), are the most widely used for function estimation.

2.6 Artificial Intelligence Techniques

As described in the Section 2.4, large volume of empirical correlations was developed to estimate crude oil properties. Most of them were developed based on specific data source and previous works, which attempted to change coefficients and equations to be applied with each situations (Asadisaghandi and Tahmasebi, 2011). However, artificial intelligence approaches can be used as powerful tools with large degrees of freedom and ability to solve complex, non-linear problems where the relation between each parameter is unknown (Al-Marhoun et al., 2012; Gharbi et al., 1999).

Gharbi and Elsharkawy (1999) presented two back-propagation neural networks (BPNN) with two hidden layers for prediction of P_b and B_{ob} of Middle East crude oil. The models used R_s, γ_g, γ_o, and T_{res} as input parameters. 498 data points were used for training and additional of 22 data points were used for cross-validation both ANN models. The P_b ANN had 8 neurons in the first hidden layer and 4 neurons in the second hidden layer. For the B_{ob} ANN, 6 neurons in both hidden layers were used. The developed ANN models outperformed the conventional correlation methods in term of average error. These developed P_b and B_{ob} ANNs by Gharbi and Elsharkawy (1999) were later validated using data by Al-Shammasi (2001) and Al-Shammasi (1999), and were reported to give physical trend. Again, Gharbi and Elsharkawy (2003); Gharbi and Elsharkawy (1997); Gharbi et al. (1999) presented another universal BPNN model with three hidden layers for predicting P_b and B_{ob} using data sets from 350 different crude oils from all over the world. 5,200 data points were used for training process, and 234 data points for verifying the model. The model used back propagation with momentum as learning algorithm. Comparing to the conventional correlation methods, the universal P_b ANN model was reported to be lower in term of average error, and the universal B_{ob} ANN model was reported to be better in term of correlation coefficient (R-value). However, both universal ANNs had less improvement over the Middle East ANNs (Gharbi and Elsharkawy, 1999).

Al-Shammasi (2001) and Al-Shammasi (1999) presented universal P_b and B_{ob} correlations as well as ANNs. P_b correlation from Standing (1947) and B_o correlation from Petrosky Jr. and Farshad (1993) were the recommended methods. However, the developed ANNs were reported to give small improvements similar to the developed correlations. Therefore, the limitation of the data used in developing ANNs was reported culprit (Al-Shammasi, 1999; Al-Shammasi, 2001).

Elsharkawy (1998, 2003) presented a radial basis function (RBF) neural network model for predicting B_{ob}, Rs, μ_o, μ_{ob}, c_o, and evolved γ_g (more detail of RBF concept can be found in Chen et al. (1991)). The input parameters are P_{res}, T, API, and separator γ_g. The presented model used 90 differential PVT data points for training, and 10 data points for testing the model.

Varotsis et al. (1999) presented a novel ANN approach, two hidden-layers BPNN for predicting the complete PVT behavior of reservoir oil and gas condensate using reservoir fluid composition and field measurements as the input parameters. A total of 650 PVT data points including 400 oil data and 250 gas condensate data were used. The data were randomly classified into training data set (80%), testing data set (10%), and validation data set (10%).

Elsharkwy and Gharbi (2000) presented a comparison between classical regression techniques (CRT) and ANN techniques for prediction of μ as a function of P_{res}, T_{res}, API. and γ_g. The using 805 viscosity measurements of crude oil samples (700 data points for training and 105 data points for cross-validation). General regression neural network (GRNN) was selected to simulate behavior of crude oil viscosity better than other CRT, BP, Ward and Levenberg-Marquardt (LM) back-propagation techniques.

Alcocer and Rodrigues (2001) used data visualization and ANN modeling for predicting μ_o and API gravity based on nuclear magnetic resonance signal data from 24 Venezuelan oil samples. Moreover, specific models were also created using 15 oil samples from the same origin for verifying the presented model.

Osman et al. (2001) used 803 published data sets from the Middle East, Malaysia, Colombia, and the Gulf of Mexico to develop an ANN model for predicting B_{ob} with 402 data used for training, 201 data for cross validation and remaining 200 data for testing the model. The model was one hidden-layer feed-forward neural network (FFNN) with 4 nodes in input layer, 5 nodes in a hidden layer, and one node in the output layer (4-5-1).

Osman and Abdel-Aal (2002) presented abductive network model based on self-organizing group method of data handling (GMDH) approach as an alternative model for predicting P_b and B_{ob}. A total of 283 data sets were split into a set of 198 data points for training, and another set of 85 data points for testing the model. These 283 data sets were reported to be the data collected from Saudi crude oils by Al-Marhoun and Osman (2002). The developed network was different from other neural networks since the GMDH-based abductive network produced a set of high-degree polynomial correlations instead of estimated figures. The same data sets were also used to develop two BPNN models for predicting these properties separately (4-7-1 for P_b, 4-8-1 for bubble oil FVF) using 142 data points for training, 71 data points for cross-validation, and remaining 70 data points for testing both models.

Goda et al. (2003) developed four-layer BPNNs for predicting P_b and B_{ob}. The P_b model used T_{res}, API, γ_g, and γ_o as input parameters, two hidden-layers with 10 nodes in each hidden-layer, and an output layer of P_b (4-10-10-1). The network was trained with 160 datasets and tested with other 20 data sets collected from Middle East crude oils. However, the B_{ob} model was different from other works since the estimated P_b from the first model was used with T_{res}, R_s, API, and γ_g as the input parameters in 4-8-8-1 ANN architecture.

Osman and Al-Marhoun (2005) presented two BPNN with RBF models for predicting PVT properties of oil field brines using 1,040 published data points. The data set were divided for training, cross-validation, and testing for both models at the ratio of 2:1:1. The first model had 3-38-3 ANN architecture for predicting brine compressibility factor, brine formation volume factor, and brine density using T_{res}, P, and salinity as input parameters, while the second model had 2-2-1 architecture for predicting brine viscosity using T_{res} and salinity as input parameters.

El-Sebakhy (2009b); El-Sebakhy et al. (2007) used 782 data sets after dropping the redundant data from the total of 803 used in Osman et al. (2001) and Goda et al. (2003) to develop support vector regression (SVR) modeling scheme (see Cortes and Vapnik (1995) for more detail in SVR technique) for predicting P_b and B_{ob} based on Rs, T_{res}, γ_o and γ_g as input parameters. They claimed that the developed SVR model was better than both abductive networks and FFNNs in term of stability and accuracy. Moreover, the same data sets were used in El-Sebakhy (2009a) for predicting these properties using type-1 fuzzy inference system with the same input parameters as El-Sebakhy (2009b); El-Sebakhy et al. (2007). Details for type-1 fuzzy inference system concept can be found in Zadeh (1965). The developed type-1 fuzzy model gave high accuracy in predicting B_{ob} with stable performance.

Hajizadeh (2007b) presented generic algorithms for predicting viscosity of crude oils using 89 data points collected from 2 PVT reports and 3 fluid characterization reports from Iranian oil fields. The input parameters were Rs, pressure, temperature, and oil density. The model took about 18 hours to run. Hajizadeh (2007a) used the same data sets to develop type-1 fuzzy inference system and two-layer FFNNs for predicting crude oil viscosity. These two approaches were reported to be successful to predict and model crude oil viscosity with capability to recognize possible relations between input and output parameters in large volume of data, in the case where the system architecture was not known.

Obanijesu and Araromi (2008) presented FFNN with LM and BP algorithms for predicting P_b and B_{ob} of Niger Delta crudes using 542 published data sets. The model used R_s, T_{res}, γ_o and γ_g as input parameters. The model could predict P_b and B_{ob} accurately within the range of data.

Dutta and Gupta (2009) present SVR models for predicting Rs, B_{ob}, B_o, μ_{ob}, and μ_o of Indian crudes based on 372 data points for P_b model and R_s model, 530 data points for B_{ob} model, 263 data points for B_o model, 435 data points for μ_{ob} model, and 252 data points for μ_o models. All the developed models outperformed most other conventional correlations. The same data sets were used by Dutta and Gupta (2010) to develop ANN models with BP and BR algorithms for predicting these PVT properties. Each model used 80% of each data for training, and remaining 20% of each data for testing. The developed ANNs had different architecture including 4-6-4-1 for P_b model, 4-6-5-1 for R_s model, 4-9-1 for B_{ob} model, 7-6-1 for B_o model, 4-6-4-1 for μ_{ob} model, and 4-9-1 for μ_o model. All the ANN models gave better performance compared to most other conventional correlations as well.

Ahmadloo et al. (2009) presented response surface methodology (RSM) based approach for predicting μ_{od}, μ_l, and μ_o of extra heavy to heavy crude oils using 45 medium-heavy crude oil PVT data sets selected from a total of 170 data sets from Alberta and Saskatchewan, and 63 heavy-extra heavy data sets selected from 110 data points reported by De Ghetto et al. (1995). The data selected were analyzed, normalized, and divided into training, testing and validation sets in the ratio of 4:1:1. The RMS-based approach could successfully predict μ_{od} and μ_o. However, the developed correlation for μ_o was outperformed by the correlation from Beal (1946) as it gave higher average absolute error (24.4% to 19.4%).

Oloso et al. (2009) developed ANN models, SVR models, and functional network (FN) models for predicting fitting coefficients for crude oil viscosity and R_s curves using crude oil compositions and reservoir properties. The 12-12-5-1 ANN structures and SVR models were used for predicting fitting coefficients. 12-12-6-1 ANN structures were used to predict oil viscosity and R_s. FN models were used for predicting fitting coefficients and oil viscosity. 99 composition data points, 1706 viscosity-pressure data points and 841 R_s -pressure data points were used. SVR and FN models gave the better performances than ANN models while SVR gave the best result for predicting R_s curves.

Khoukhi et al. (2011a); Khoukhi et al. (2011b) used 99 composition data sets to develop SVR models, FN models, and 12-12-5-1 FFNN model for predicting fitting coefficients of viscosity and R_s curves. In addition, two 12-13-6-1 FFNN models were used for predicting μ_{ob} and R_s. Similarly, both SVR and FN models were reported to outperform the developed FFNNs.

Abedini et al. (2011) employed data from five Iranian oil samples to develop type-1 fuzzy model for determination of μ_o. The model used P, P_b, and μ_{ob} as input parameters. After using 86 experimental data sets for testing, the developed type-1 fuzzy as reported to give better results compared to the published correlations from Beal (1946), Kartoatmodjo and Schmidt (1994), Khan (1987), and Vazquez and Beggs (1980).

Olatunji et al. (2011a) introduced sensitivity based linear learning method (SBLLM) for predicting P_b and B_{ob}. The model used T, Rs, γ_g, and γ_o as input parameters. The data used for developing this model were the same data used in the published works from Al-Marhoun (1988), Al-Marhoun and Osman (2002), Goda et al. (2003), El-Sebakhy et al. (2007), and El-Sebakhy (2009a, 2009b). In addition, Olatunji et al. (2011b) also employed the similar data with to develop adaptive type-2 fuzzy logic inference system for prediction of P_b and B_{ob}. The developed SBLLM and type-2 fuzzy logic inference system from Olatunji et al. (2011b) and Olatunji et al. (2011a) were reported to give competitive results comparing to ANNs developed in their work with better performance compared with those correlations from Standing (1947), Glaso (1980) and Al-Marhoun (1992).

Asadisaghandi and Tahmasebi (2011) presented two ANN models for predicting P_b and μ_{ob} of Iranian crudes using 130 data sets from 23 different oilfields in Iran. Reservoir temperature, R_s, γ_o, and γ_g were normalized to be input parameters in the models. Both models used back propagation (BP) and Bayesian regularization (BR) as training algorithms. The model architecture for P_b was 4-10-1, while for the B_{ob} model was 4-8-1. The results shown that both ANN models gave the best result as well as Glaso (1980)'s correlation for P_b prediction and Al-Marhoun (1988)'s correlation for B_{ob} prediction.

Torabi et al. (2011) developed three ANN models with LM training algorithm for predicting μ_{od}, μ_{ob}, and μ_o using data sets from five Iranian oil reservoir. These models worked in parallel with each other. Firstly, the μ_{od} ANN used API and T as input parameters. Secondly, R_s, P_b, and the predicted μ_{od}, (which was the output from the μ_{od} ANN model) were used as input parameters for μ_{ob} ANN model. Last, the μ_o model then used the μ_{od} and μ_{ob} obtained from the μ_{od} ANN and the μ_{ob} ANN, and P as input parameters. These ANN models were reported to outperform the regional correlations from Beal (1946), Beggs and Robinson (1975), Glaso (1980), Labedi (1992), Kartoatmodjo and Schmidt (1994), Elsharkawy and Alikhan (1999), Chew and Connally Jr. (1959), Vazquez and Beggs (1980).

Al-Marhoun et al. (2012) presented artificial intelligent techniques for predicting oil viscosity curve using data from 42 PVT reports obtained from

Canadian oil fields. Functional network forward selection (FNFS) technique was reported to give the best results followed by SVR.

Khoukhi (2012) proposed adaptive neuro-fuzzy inference system (ANFIS), generic adaptive neuro-fuzzy inference system (GANFIS), genetically optimized neural networks (GONN) as hybrid soft computing systems for predicting P_b and B_{ob} based on the data used in Emad A. El-Sebakhy (2009b). According to comparative results between hybrid systems, neural network as well as correlations by Al-Marhoun (1988), Osman and Al-Marhoun (2005), and Al-Marhoun and Osman (2002), GANFIS and GONN gave the best results with very competitive performance.

Asoodeh and Bagheripour (2012) utilized ANN, fuzzy logic system and ANFIS with power law committee with intelligent systems (PLCIS) for predicting P_b using worldwide 361 PVT data points taken from Ostermann and Owolabi (1983), Dokla and Osman (1991), Omar and Todd (1993), and De Ghetto and Villa (1994). PLCIS could integrate and enhance the precision results from multiple prediction systems.

Finally, Ikiensikimama and Azubuike (2012) presented BPNN with LM algorithm with 4-5-1 network architecture for predicting B_{ob} of Niger Delta crudes using 802 data sets collected from Niger Delta region. 482 data sets were used for training, 160 data sets were used for cross-validation, and remaining 160 data sets were used for testing the model. The input parameters were reservoir temperature, oil API gravity, γ_g, and R_s. The results from statistical and trend analyses shown that the developed neural model had good agreement with physical trend with better performance compared to other empirical correlations.