APPLICATION OF TIMES SERIES ANALYSIS IN FORECASTING FUTURE PERFORMANCE OF A RESERVOIR UNDER WATER INJECTION

By

Alfian Rachman

15791

Dissertation submitted in partial fulfillment of

The requirements for the

Bachelor of Engineering (Hons)

(Petroleum)

January 2015

Universiti Teknologi PETRONAS

32610 Bandar Seri Iskandar

Perak Darul Ridzuan

Malaysia

CERTIFICATION OF APPROVAL

APPLICATION OF TIMES SERIES ANALYSIS IN FORECASTING FUTURE PERFORMANCE OF A RESERVOIR UNDER WATER INJECTION

By

Alfian Rachman
15791

A project dissertation submitted to the
Petroleum Engineering Programme
Universiti Teknologi PETRONAS
in partial fulfilment of the requirement for the
BACHELOR OF ENGINEERING (Hons)
(Petroleum)

Approved by,

____

(Berihun Mamo Negash)

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK
January, 2015

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

ALFIAN RACHMAN

ACKNOWLEDGEMENT

I would like to take this opportunity to express my deepest gratitude to Universiti Teknologi PETRONAS for giving me a chance to complete my undergraduate study.

I would like to thank my supervisor, Mr Berihun Mamo Negash, who constantly takes good care of me under his careful supervision. Deepest regards to Dr Shiferaw for additional information in Decline Curve Analysis, and my Junior Shodiq Khoirur Rofieq for constant lesson on History Matching.

The blessing, help and guidance, given by them from time to time shall carry me a long way in the journey of life on which I am about to embark. Lastly, I would like to thank Almighty ALLAH, my family and friends for their constant encouragement in which without it, this journey will be very difficult.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

ABBREVIATION AND NOMENCLATURE

DCA = Decline Curve Analysis

HM = History Matching

TSA = Time Series Analysis

OE = Output-Error

BJ = Box-Jenkins

NMRSE = Normalised Squared Root Error

RMSE = Root mean Squared Error

Q(t) = Production rate at time t

Qi = initial production.

B, Di = constant.

y(k) = output function

u(k) = input function

e(k) = error function

μ = a constant or intercept

B, C, D, F = Matrix Polynomial

ABSTRACT

Decline Curve Analysis (DCA) and History Matching (HM) are classical methods used in predicting reservoir performance. While both methods are widely used, they have certain limitations and strengths. DCA is only applicable for reservoir with primary drive and assumes that all mechanical conditions of a well remain constant. HM, on the other hand, is very complex, takes longer time, and require experience. Hence, a new simpler and faster technique is required. In this work, a technique called Time Series Analysis (TSA) is proposed for predicting the reservoir performance. Time series analysis is widely used in predicting future patterns in economics and weather forecasting, where factors influencing output are too many to consider. Other examples of the application of time series analysis are prediction of equipment prognostic and process of quality control. Two types of TSA were tested: Output-Error (OE) and Box-Jenkins (BJ). Eight models are developed by varying the order of each models. Two of the best models were chosen based on the resulting normalized root mean square error (NRMSE) and are compared with the conventional reservoir forecasting methods. The NRMSE from the selected models, OE (1-2-1) and BJ (1-2-1-2-1), showed a comparable result with DCA and HM. The result of this study shows that, TSA has a very good potential for use in reservoir performance prediction under water injection and hence it can be utilized as alternative reservoir forecasting tool.

CHAPTER 1
INTRODUCTION

1.1 Project Background

Forecasting is a process where an observation of the future is conducted in which the actual output has not been yet observed from one point of time to another specific point of time. Although the definition is linked with predication, the output for forecasting cover more possible outcomes and also more specific.

Forecasting methods are wide; started from qualitative vs. quantitative methods, time series analysis method, naïve method, judgmental methods and many others. The objectives of forecasting is to predict the conflict that may occur during future development. Forecasting relates with planning since it can interpret on what the future will look like, while planning shows on how the future should look like.

In reservoir management, reservoir forecasting is one of significant areas that play a very important rule for predicting the future trend, activities and the behaviour that might be overcome in the future. There are several methods that have been used by petroleum engineers to forecast the reservoir performance, such as decline curve analysis (DCA) and history matching (HM). Decline curve analysis is a long established method that has been widely used to predict the reservoir performance by approaching the production decline-curve rate mathematically. Also, history matching is a method where the numerical set of data is constructed to fit with production history.

Decline Curve Analysis firstly introduced by Arps at 1944. It is the oldest method known for reservoir forecasting, however many modifications of decline curve analysis has been done for a better result. Decline curve analysis applies means value based on past production history, a graph of data will be constructed from a groups of wells to detect a trend to aid in predicting the future performance. The analysis is conducted on semi-log paper or log-paper before the availability of computer.

Nowadays, decline curve analysis software on PC computers is utilized to plot production decline curves for petroleum economics analysis.

Second most used method for reservoir forecasting is history matching. History Matching is a process of adjusting the geological model thus it may reproduce the measured pressure and production data (Lind, 2013). This act of fitting the reservoir behaviour between the model and previous reservoir performance expects the accurate future prediction. The accuracy of this method itself is depending heavily on the data of the pressure and production, therefore it is expected that the fitting of those two parameters is matched as accurate as possible.

Lastly, is material balance, which is an expression of conversion of mass in reservoir. The equation mathematically defines the different producing mechanisms and effectively relates the reservoir fluid and rock expansion to the consecutive fluid withdrawal. The application of material balance may include estimation the performance of reservoir as well.

Although these three methods have been around, each of the method has their own weakness and limitations such as overestimating the value for DCA (Horner& Li, 2005), non-uniqueness factor and uncertainty assessment for History Matching (Cancelliere, 2011), and no spatial information is being used for material balance.

To cover up the limitations of those three methods, it is now a good time to introducing a new solution for reservoir forecasting. One of many methods for forecasting is time series analysis (TSA). Time series analysis is a set of observations that have been thoroughly observed in several time frame (Chatfield, 2013). Any science and engineering that involves with time measurement frequently use time series analysis application. The examples are weather forecasting, astronomy, mathematical finance, statistic, signal processing, earthquake prediction, communication engineering, and econometrics and pattern recognition.

Despite the fact that the application of time series analysis in petroleum industry is quite new, some applications have been used for understanding crude oil price (Torkowe, 2012), and estimation of productivity index (Marcary, 2003). The most recent work for time series analysis that has been done by Olominu (2014), showed that the forecasting by time series gave a better result compare to decline curve analysis.

Time series analysis may ignore the physics condition that happens at reservoir, however this method can be used for reservoir forecasting, in which give a simpler time procedure compare to the simulation. In addition, the time series analysis is lied on the reliability as well. The wider the period of the time, the more reliable the time series analysis is. The last but not least, the time series analysis is able to recognize trend. For example, the trend tendencies may help the managers of franchise store to measure the upward trend due to some fluctuation sales of some particular good. Hence the same approach can be made as an approximation for similarly situated store.

1.2 Problem Statements

Reservoir forecasting is a very crucial elements in petroleum industry. However, there are some limitations in the conventional forecasting method such as:

a. Decline Curve Analysis

   The difficulties to foresee which equation to follow and there is no clearance method on choosing the type of the curve. In addition, can only be used in under assumption that mechanical conditions and reservoir drainage remain constant while it produces at a capacity.

b. History Matching

   Time consuming process, difficulty in parameterization, and difficulty in identifying the uncertainties.

1.3. Objective and Scope of Study

A. Objective

The objective of this study is to conduct a study by applying the time series analysis in predicting a reservoir performance under water injection.

B. Scope of Study

Basic of linear system identification is required to understand more about TSA. Reservoir modelling is also required to accompany the project, although some limitations occur in this particular area: finding a reservoir model that can show the real phenomenon of reservoir. A few software are involved in this research such as: CMG for reservoir modelling and HM, Excel for Decline Curve Analysis, and MATLAB tools for Time Series Analysis.

CHAPTER 2
LITERATURE REVIEW

Petroleum reservoir is one complex system that consisting of large and geologically intricate, and often contain many wells. All of this complexity of the nature reservoir is relied on reservoir modelling and now it plays a very important role for every petroleum engineering to understand the complexity of it. Many of international oil company rely on the hydrocarbon production forecasting for their business planning (Obidike, 2014). Forecasting in petroleum industry hold a very important role; forecast will decide the investment decision, design of facilities, system pipelines, processing and refinery, and the export system. By having a good forecast, it will set a very good bridge communication among authorities, partner, and operators to achieve their production and financial target.

The hardest issues of reservoir forecasting is how to make all of the dynamic part to move in synchronize. Although some of the input, such as facility and storage capacity, can be a static input; however there are certain parameters that can change from time to time (management steers, optimization, etc.). In addition, there are also uncertainty and grisk. Uncertainty can be defined as ‘not known beyond doubt’, and grisk is good risk. Grisk is possibility that may exceeds the expectations that is presented in financial outcome and probabilities. Some of risks that occur are low exchange rate, low plateau, high operating cost and etc. A project with larger grisk than risk is much more preferred. There are many uncertainties that can make the process of reservoir forecast to be quite unsmooth, including size and shape reservoir, factor that affect the depletion and etc. Environmental changes is one of the major factors that may affect the numbers of uncertainty. If the future production is known, imagine how boring the life of E&P is?

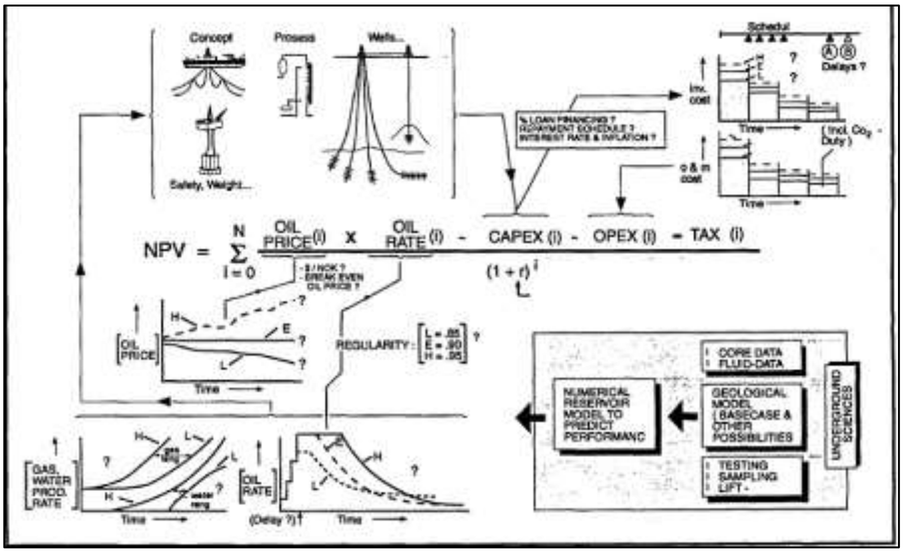

Figure 2.1: Production forecast in an economic context

The decline curve analysis is the oldest method of reservoir forecasting that was created by Arps (1945). This method has been used since very early age even when the viability of computer is not existed yet. Arps stated that the extrapolation of production rate against the certain time frame can assist in predicting the reservoir performance instead of volumetric calculations. An assumption was made that the rate at any preceding date is a constant fraction of production rate which implies that at given constant interval the production will drop. There are four types of decline rate that related to the simple arithmetic relationship, which are Constant decrement, Exponential decline, Harmonic decline and Fractional power decline. Furthermore, it is now simplified into three, which are exponential, harmonic and hyperbolic decline.

Many corrections for Arps equation is conducted to obtain a better result. Fetkovich (1954) introduced the pressure term for account of relative permeability effect for solution gas drive system by deriving the hyperbolic expression with an Arps exponent.

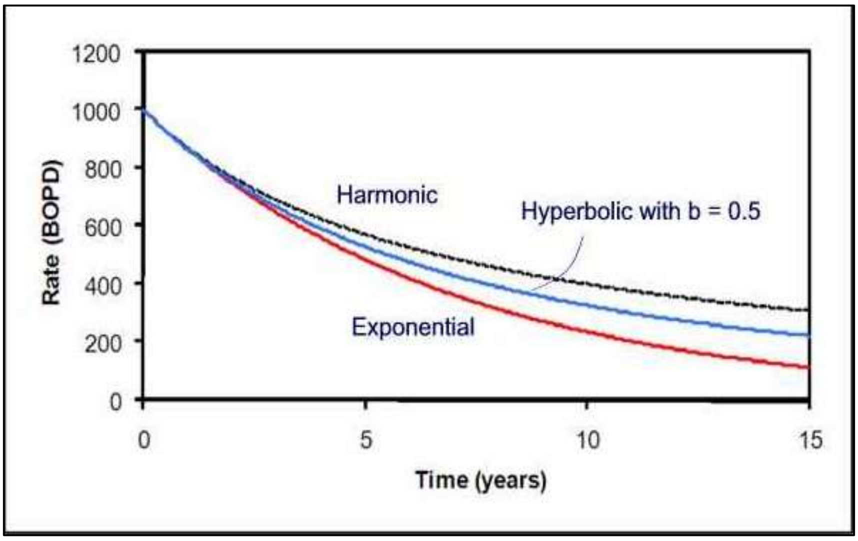

Figure 2.2: Decline Rate Curve

Blasingame also combined both of the previous work by using material balance principle that allows the constant depletion appears as it is constant at constant flow rate.

Decline Curve analysis is legitimately used due to the simplicity; it is straight forward, acceptable for quantitative reservoir determination and available on many software (Ling, 2013). However, there are certain downfalls for using decline curve analysis. Firstly, Ling stated that there will be many uncertainties rises during the analysis of waterflooding for radial injection. This happened because of the production-injection pattern is adjusted, therefore the result of the calculation might be wrong. In addition, in water flooding system, the oil relative permeability cannot be approximated by a function of pressure due to the value fluctuation (Li & Horner, 2005). Secondly, for gas reservoir system, the decline curve analysis ignore the pore compaction that may overestimates the GIIP value. Lastly, it may ignore the effect of infill wells for saving time that may cause of underestimating the production value.

Li and Horner (2005) stated that the decline curve analysis model is heuristic. However, each of the curve has its own disadvantages, which are the exponential decline curve analysis that has a tendency to underestimate the reserves and production rates, and the harmonic exponential curve with a tendency to over predict the reservoir performance. In some cases, the graph is neither following both of the curves but instead cross-over entire curves. There is no clearance on how the decline curve is chosen hence it was difficult for the engineers to foresee which equation should be used.

Second most common method for reservoir forecasting is history matching. The objectives of history matching are to improve and validate the reservoir simulation model for a better understanding process. Integrating the model input including the geology, geophysics, and engineering data is required to obtain accurate prediction in this method.

Lind (2013) had done an analysis for computed history matching on several model. One of the most common overlook issue on history matching is the error in allocated history data. Hence, the objective of Lind's experiment is to compare the conventional history matching and the assisted history matching result. The model that was used for this experiment is a coarse grid model that was matched with global behaviour. The computer assisted matching then is used for individual tune well matches and the result was spectacular. While there are significant improvement from assisted history matching, there are also certain aspects that need to be reviewed, such as the geological model and etc.

A recent study by Katterbauer *et al., (*2014), mentioned a significant improvement for history matching method in heated heavy oil reservoir. A simulation at a heavy oil reservoir is conducted whereas some obstacles occur along the way: high viscosity that limits the viability and recovery factor. By the assistance of electromagnetic radiation, the heat-loss issue from the thermal heating can be solved but the fluid displacement and production history is hardly understood. Katterbauer *et al.,* introduced the cross-well seismic imaging to help the conventional history matching overcome this problem. Combine with the ensemble of Kalman Filter, it decreases the uncertainty and significantly enhanced the accuracy of forecasting.

History matching also has some limitations. Tomomi (2000) tested three different case with a various scenario where it was concluded that every reservoir must be treated separately and has crucial information for whatever data. A different scenario is necessary even it is elaborated. In addition study by Tavasolli (2004) also stated that history matching stated that data required is very enormous, and very time consuming.

An extensive study by Cancelliere (2011) also had done a conclusive research on some limitations of history matching. The first issues is the non-uniqueness of the models. When the conventional history matching started with matching the global parameters with the adjusted well or near wellbore data, this procedure can be flexible since the matching can be carried out by the engineers with good judgment and experience, however this process of trial will take a very long time. Assisted history matching however may shorten the process by multiplying the various calibration at the same time, the solution that is given is usually one corresponding to the minimum given parameters that may give us good production data but not necessarily good estimation.

Tavasolli (2004) stated that the best production-matched model does not essentially have a good fit for the parameters of the reservoir. On the other hand, a model with its parameters close to those of the base case might not have a good match to the production data. In summary, all the results seem to suggest that in using the conventional history-matching methods, one cannot practically promise to improve the true model, which embodies the real geological structure of the reservoir.

There is no limit for each algorithm applicability's (Cancelliere *et al.*, 2011). Although some algorithm model is proven to be highly efficient, however when it faces complex reservoir the majority of this model are failed. Evolutionary algorithm only applicable for small number parameter since if it faces a large one, there will be severe loss of efficiency. Despite the adaptability into any simulator, the ensemble of Kalman Filter however requires a special parameterization as well.

Lastly, the material balance equation. Although it is one of the effective forecasting method, however material balance has some restriction in terms of information usage. Besides the information that is used is no spatial, there is not spatial distribution and saturations data as well. The properties that is used is an average and no time addressing issues arise.

With all of those limitations, a new approach for reservoir forecasting is required. A new approach called time series analysis is proposed to overcome this issue. A time series analysis combine the applicability of the conventional statistic with time correlations where the observations are identically distributed and also independent. The time series analysis is well known in economy forecasting, and some area in industry as well. Although this concept for time series analysis forecasting is still quite new in petroleum industry, the application of time series analysis in petroleum industry is not limited only in forecasting.

Research on time series correlation had been done by Macary (2003), which to discover the productivity index auto-correlation function, where it is important to describe the variable of mutual dependence in different variable at different time period. The result has been compared with common productivity calculation and generate a very accurate result.

A time series model ARIMA (autoregressive integrated moving average) was also used by Tokowei (2012) to understand crude oil price variation. Babinec & Pospichal (2010) conducted a study for dynamic reservoir forecasting by using feedforward neural network for time series analysis forecasting. By choosing laser fluctuations and Mackey-Glass time series as a testing data, the original Echo State neural networks has no possibility to stop the training algorithm, which to avoid the over-fitting problem and combining both tools increased the quality of forecasting.

Olominu *et al.,* (2014) proposed the application of time series analysis prediction versus decline curve analysis. In his case study, Olominu *et al.,* used reservoir output data and model it into mathematical form, called autoregressive integrated moving average or known as ARIMA. Four models were created and one of the best match that gives closer value to the cumulative oil production was chosen and compared with the result of decline curve analysis. The result showed that the chosen model can be used to forecast the reservoir performance for short-long period. The accuracy of TSA has better result compare to the decline curve analysis.

CHAPTER 3

METHODOLOGY

3.1 Research Methodology

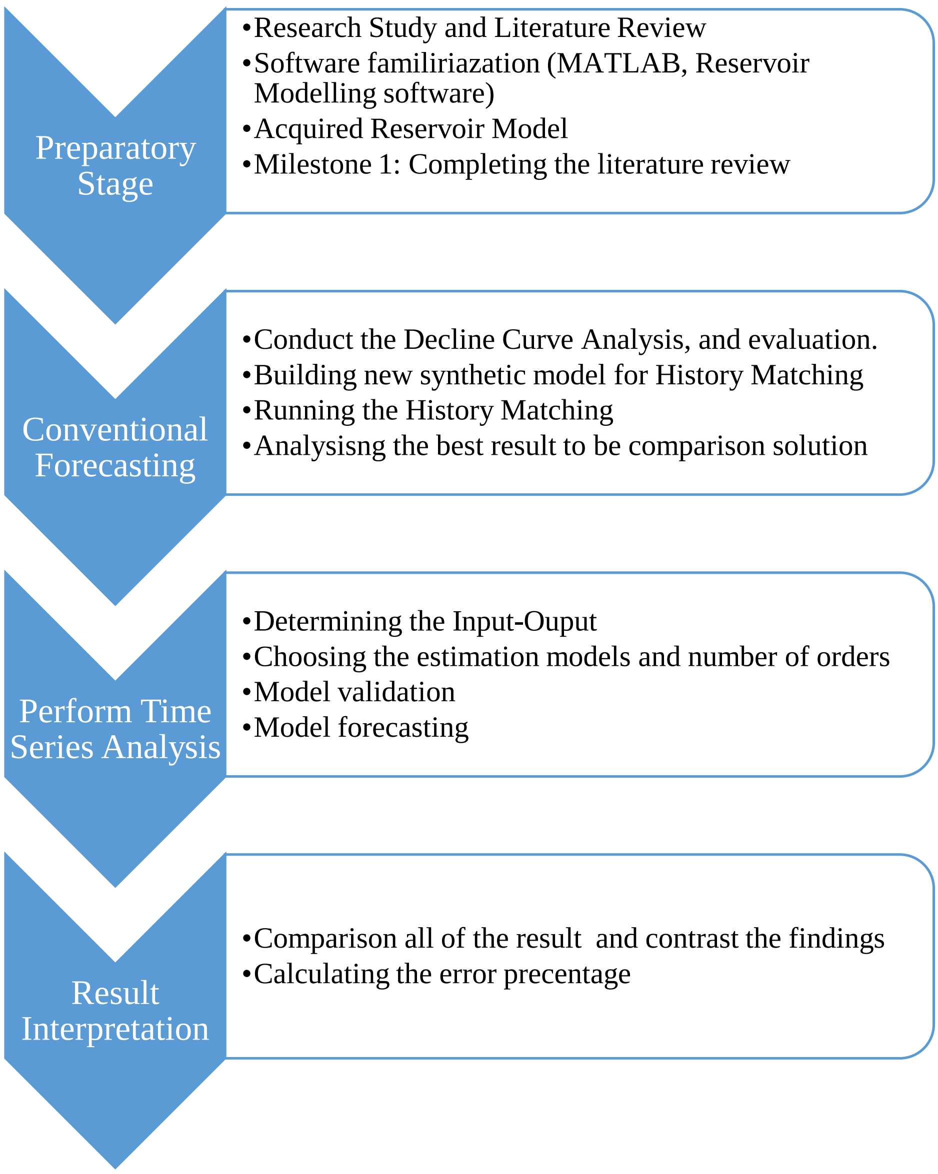

Figure 3.1: Research Methodology

Time Series Approach

:

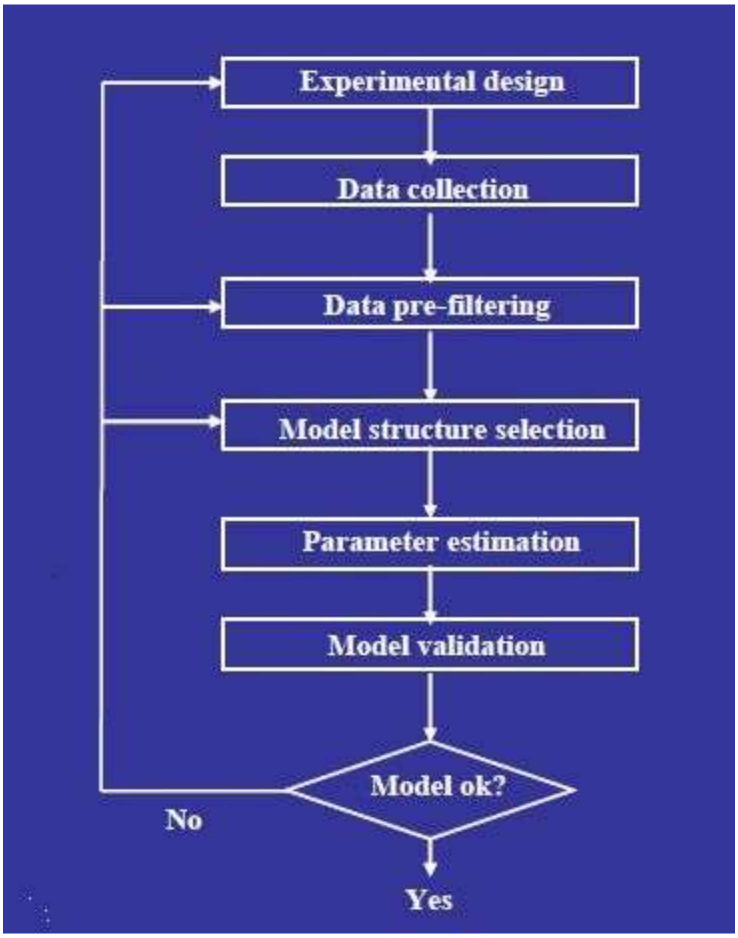

Figure 3.2: Time Series Activities

System identification includes the following steps:

- ***Experiment design:*** to obtain good experimental data, and it includes the choice of the measured variables and of the character of the input Signals.
- ***Selection of model structure:*** A suitable model structure is chosen using prior knowledge and trial and error.
- ***Choice of the criterion to fit:*** A suitable cost function is chosen, which reflects how well the model fits the experimental data.
- ***Parameter estimation:*** An optimisation problem is to obtain the numerical values of the model parameters.
- ***Model validation:*** The model is tested in order to reveal any inadequacies.

3.2 Decline Curve Analysis

Decline Curve Analysis (DCA) is the oldest method of reservoir forecasting that was created by Arps (1954). This method has been used since very early age even when the viability of computer is not existed yet. The empirical Arps equation can be shown as:

$$Q(t) = \frac{Q_i}{1 + bD_i t} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

Where Q(t) is the production rate at time t, and Qi is initial production. b and D_i are constant. The first equation can be simplified into three parts depending on the b value: b=0 is exponential, b=1 is harmonic, and b>0 but b<1 is hyperbolic.

$$Q(t) = Q_i e^{D_i t} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

$$Q(t) = \frac{Q_i}{1 + D_i t} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

In order to determine the Q(t) for hyperbolic case, Fetkovich (1973) plotted the $D_i t$ versus Q/Q_i with b variance started from 0.2 to 0.8. Those curves will be matched against the actual Q versus t curve that have the same trend with any b curves. The calculation for D_i *and* Q for the actual data can be done once the match point is acquired.

Although it is believed as the simplest and long established method for reservoir forecasting, DCA has certain limitations. DCA is only applicable for reservoir with primary drive and assumes that all mechanical conditions of a well remain constant. Secondly, there are tendencies of overestimating and underestimating the reservoir performance (Li& Horner, 2005). Lastly, no justification on which type of curve is selected. In some cases, the exponential will be chosen due to its simplicity, while the hyperbolic usually chosen due to the higher accuracy compare to the rest of the methods.

3.3 History Matching

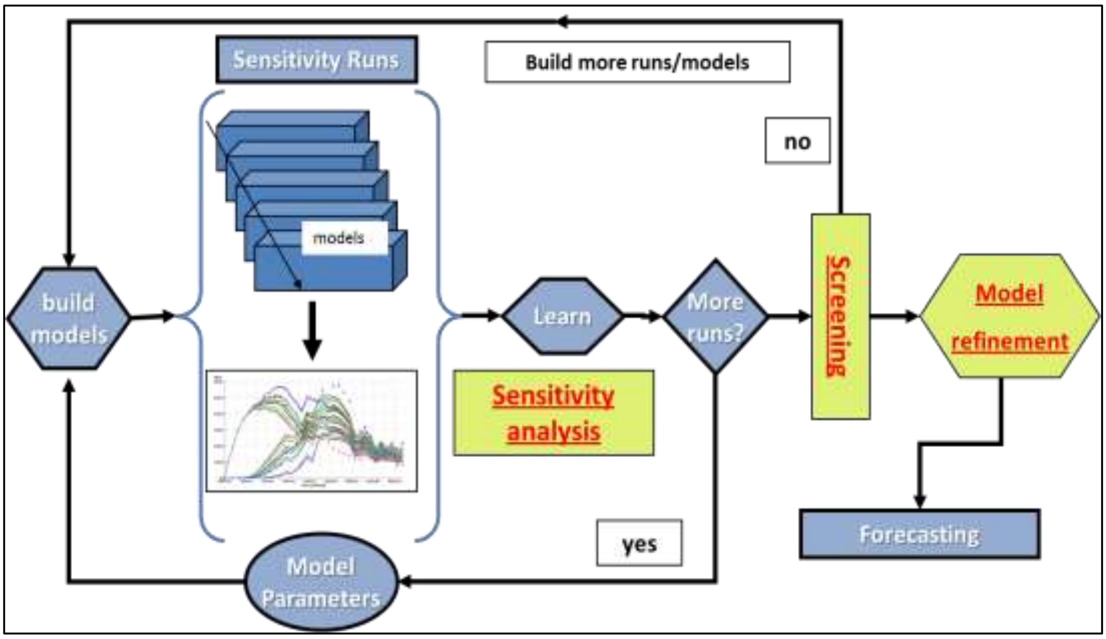

Figure 3.3: History Matching Steps

The objective of the history matching is to adjust the model and its parameters (e.g., permeability and porosities) that the simulation model is capable to generate the well-flow-rate and pressure histories reasonably.

The result of the history-matching process is a new simulation model that can be very dissimilar from the original geological model. One issue with History Matching is no unique solution is provided. Different scenarios of reservoir parameters can provide many simulation models tuned to the available past data. Although, each matched reservoir model is capable of reproducing the observed data, these various geological models can generate different production forecasts. The greatest task is acquiring multiple efficient history-matched models for accurate uncertainty approximation.

Conventional history matching is a trial-and-error process. The mismatch between observed and predicted values is minimized by adapting reservoir parameters over consecutive simulation runs. The process is very time consuming even for professional reservoir engineer. On the other hand, modern history-matching techniques apply numerical optimization and produce multiple geologically constant adjusted scenarios.

Besides matching the previous data, reservoir engineer also need to forecast the reservoir performance with including all of the reservoir development scenarios and various geological models. Parameters that allows the equation fluid flow to produce the output are:

- The porosity;
- The absolute permeability,
- The relative permeability
- The productivity indices of the wells, etc.

At each well, the observed data are mainly the pressure, the fluid rate of the different phases and the composition of the fluids. Any combination of these variables such as water-cut, gas-oil and water-oil ratios can also be used.

In the inversion, it is well-known that the answer may not be unique. Hence, it is important to incorporate geological knowledge in the history match procedure to reduce the space of possible solutions. Taking into account geological data, the production forecasts should be more predictive.

In the conventional HM procedure, the main steps are:

- Building of an Initial model; usually, the model is deterministic.

  On building of new model, sensitivity analysis can be used as assisting tools. It is a study in how uncertainty in the output can be apportioned to different sources of uncertainty in its inputs (Salteli, A., et al 2008). Here is where the issue lies: the multiplier. There is no exact rules on what is the limit of multiplier. Ma and Pointe (2011), using ±0.03 for porosity and ±0.05 for permeability multiplier. Another finding by using tunneling method for computational optimization stated that the most common permeability multiplier is ranging from $0.1 < x < 5.0$ and $0.5 < x < 3.0$ for permeability. This process is necessary to determine which parameters inside the model that sensitivity with changes of value at certain range.

- Matching Observed Data; the observed data are well pressures, flow rates, etc.

  To obtain the match, the conventional procedure is done on a try and error procedure, i.e. given a set of parameter, the simulator is run and the results are com-pared with the observations. When this stage is done manually, the reservoir engineer modifies the values of the parameters with respect of reservoir knowledge and of his understanding of the behavior of the reservoir.

By using the automatic procedure, it is possible to speed up the process. this process allow us to alter iteratively the parameter's value to acquire a better agreement between the observed data and predicted data.

- Simulating Production Forecast with Matched Model
  When there is a suitable match of the available data (history match), the same simulator is used to forecast the behaviour of the reservoir. Sometimes, sensitivity studies are done around the parameters obtained after the match, but this does not directly give a quantification of the uncertainty on the forecasts.

3.4 Time Series Analysis

On design control, a certain mathematical model of dynamic system is required for the process. The model of a system is a description of its properties, suitable for a certain objective. Often the dynamic modelling is difficult to acquire because of the complexity of the process. To select the model in order to serve the certain objective, system identification tool is utilized to solve this issue. System Identification is a tool to build mathematical model of the dynamic system from measured data. .

A linear system identification can be distinguished into two major parts: parametric models and non-parametric models. On this paper, the parametric models will be the main object of the research. The model can describe the true process behaviors with exactly finite parameters that often related with physical quantities. Typical examples of this model are difference or differential equation, transfer functions and state-space functions.

$$y(k) = q^{-k} G(q^{-1}, \theta) + H(q^{-1}, \theta) e(k) \quad \text{……………………..} \quad (4)$$

Where *u(n)* and *y(n)* are the input and output of the system, *e(n)* is zero-white noise or disturbance of the system, $G(q^{-1}, \theta)$ is the transfer function of the deterministic part of the system and *H* is the transfers function of stochastic part of the system.

The function of *G* and *H* can be divided into their numerator and denominator polynomials. The notations of general linear identification can be written as:

$$y(k) = \frac{B(q)}{F(q)A(q)} u(k) + \frac{D(q)}{C(q)A(q)} e(k) \quad \text{……………………...} \quad (5)$$

By making assumptions of polynomials *A, B, C, D* and *F*, many type of linear models can be generated from this formula.

3.5 Mathematical Model

3.5.1 Autoregressive (AR)

Autoregressive or AR model is one of common time series model that only have denominator polynomial.

$$y(k) = \frac{1}{D(q)} e(k) \quad \text{………………………………………} \quad (6)$$

It is the most common tool used in linear prediction. AR model only depending on previous output to produce the new output.

3.5.2 Moving Average (MA)

Moving Average is time series model that just only have numerator polynomial.

$$y(k) = C(q) e(k) \quad \text{………………………………………} \quad (7)$$

Fitting the MA model is slightly more difficult compare the AR model due to the MA model is non-linear that cause the error terms is not observable. The MA model result also has less obvious interpretation compare to the AR model.

3.5.3 Autoregressive Moving Average (ARMA)

As we have remarked, dependence is very common in time series observations. To model this time series dependence, we start with univariate ARMA models. ARMA model combine the AR and MA or we can call it as time series model that has numerator and denominator polynomial.

$$y(k) = \frac{C(q)}{D(q)} e(k) \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(8)}$$

3.5.4 Autoregressive Integrated Moving Average (ARIMA)

This time series model is generalisation model of ARMA model in which the model are more fitted to gain a better accuracy on forecasting.

$$y(k) = \mu + \frac{C(q)}{D(q)} e(k) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(9)}$$

The model is generally meant to as an ARIMA (*p*,*d*,*q*) model in which parameters *p*, *d*, and *q* are non-negative integers that indicate to the order of the autoregressive, integrated, and moving average parts of the model in succession. ARIMA models form a crucial part of the Box-Jenkins approach to time-series modelling.

Four of the models above are the models without any input. In this case, the input-output models are going to be tested as comparison study with conventional reservoir forecasting methods. The models that can be used are: Output-Error (OE) and Box-Jenkins (BJ). Typically, the input-output models have a higher accuracy in term of prediction.

3.5.5 Output-Error (OE)

This model is characterised by white noise that does not contain process dynamic (*H* transfer function is not available). It is one of the most widely used model besides ARX and ARMA. There is a bit confusion to distinguish the output-error: whether it is the model or the class. To clarify the confusion, abbreviation OE is always referred to the special model above. The output-error can be enhanced by adding ARMA filter which eventually become the Box-Jenkins model. OE model can be described as:

$$y(k) = \frac{B(q)}{F(q)}u(k) + e(k) \quad \text{......................................} (10)$$

3.5.6 Box-Jenkins (BJ)

Box-Jenkins model is combination between ARMA and ARIMA model to get the best fitting value for time series.

$$y(k) = \frac{B(q)}{F(q)}u(k) + \frac{C(q)}{D(q)}e(k) \quad \text{.................................} (11)$$

From all of the linear system, BJ model is the most general and the most flexible model. It allows one to approximate separate transfer function with arbitrary denominators and numerators from the disturbance to the output and also from input to the output. However, the flexibility of this model needs to approximate large number of parameter, consequently makes BJ model rarely used in practice.

3.6 POLYNOMIAL MODELLING

Polynomial modelling is an objective function that is generally used when a simple empirical model is required. Characterizing data by global fit can be used by polynomial modelling. Polynomials are usually used for a single empirical mode. The main advantages by using the polynomials modelling are it includes reasonable flexibility for data that is not too complicated.

$$y(t) = \sum_{i=1}^{n+1} P_i x^{n+1-i} \quad \text{..} (12)$$

where n is the *degree* of the polynomial ,$n + 1$ is the *order* of the polynomial, and $1 \le n \le 9$.

For OE model, the polynomial order can be written as:

$$y(t) = \sum_{i=0}^{nu} \frac{B_i(q)}{F_i(q)} u_i(t - nk_i) + e(t) \quad (13)$$

The Box-Jenkins Model can be modeled as:

$$y(t) = \sum_{i=0}^{nu} \frac{B_i(q)}{F_i(q)} u_i(t - nk_i) + \frac{C_i(q)}{D_i(q)} e(t) \quad (14)$$

Where,

$$nb : B(q) = b_1 + b_2 q^{-1} + ... + b_{nb} q^{-nb+1} \quad (15)$$

$$nc : C(q) = 1 + c_2 q^{-1} + ... c_{nb} q^{-nc} \quad (16)$$

$$nd : D(q) = 1 + d_2 q^{-1} + ... d_{nd} q^{-nd} \quad (17)$$

$$nf : F(q) = 1 + f_2 q^{-1} + ... f_{nf} q^{-nf} \quad (18)$$

To choose the polynomial order, it is advisable to try out various available choices and use the one that seems to work the best. However, the higher degree of the polynomial can be unstable, hence the higher degree of polynomial need to be taken care carefully.

3.7 PARAMETER ESTIMATION

Prediction-error methods (PEMs) are a broad family of parameter estimation methods that can be useful to quite uninformed model parameterizations.

The PEM has a number of advantages:

- It can be applied to an extensive spectrum of model parameterizations
- It gives models with exceptional asymptotic properties, due to its similarity with maximum likelihood.
- It can handle systems that operate in closed loop (the input is partly determined as output feedback, when the data are collected) without any special tricks and techniques.

The general predictor model is given:

$$y(k) = G(q,\theta) + H(q,\theta)e(k) \quad \text{.......................................} (19)$$

where it can be constructed as:

$$G(q,\theta) = \frac{B(q)}{F(q)} = \frac{b_1 + b_2 q^{-1} + \ldots + b_{nb} q^{-nb+1}}{1 + f_2 q^{-1} + \ldots f_{nf} q^{-nf}} \quad \text{...........................} (20)$$

$$H(q,\theta) = \frac{C(q)}{D(q)} = \frac{1 + c_2 q^{-1} + \ldots c_{nc} q^{-nc}}{1 + d_2 q^{-1} + \ldots d_{nd} q^{-nd}} \quad \text{...................................} (21)$$

The parameter thus comprises the coefficients *bi, ci, di* and *fi* of transfer functions. The model that is described above is the prediction error for Box-Jenkins method. In special case, such as Output-Error, the disturbance $H(q,\theta)$ is not modelled hence the order of *nc* and *nd* are equal to one.

3.8 NORMALIZED ROOT MEAN SQUARE ERROR (NRMSE)

It is frequently used measure of the difference between the predicted model and the actual model. This tool is normalized version of the sum of absolute difference. It is often expressed as a percentage, where lower values indicate less residual variance.

A measure of the incremental quality of the simulated production and pressure is to calculate the root-mean-square (RMS_a) of the modeled production ($q_{modeled}$, $q_{observed}$) at the field and well levels,

$$RMS_a = \sqrt{\frac{\sum_{t=i}^{n} (q_{modeled} - q_{observed})^2}{n}} \quad \text{..........................} (22)$$

where n is the number of time steps. This error represents the absolute average deviation of the simulated results from the actual results. To show it term of percentage, the RMS_a can be normalized and expressed as

$$NRMSE = \frac{RMS_a}{q_{observed}} \quad \text{....................................} (23)$$

where lower values indicate less residual variance.

3.9 Milestones and Gantt Chart Project

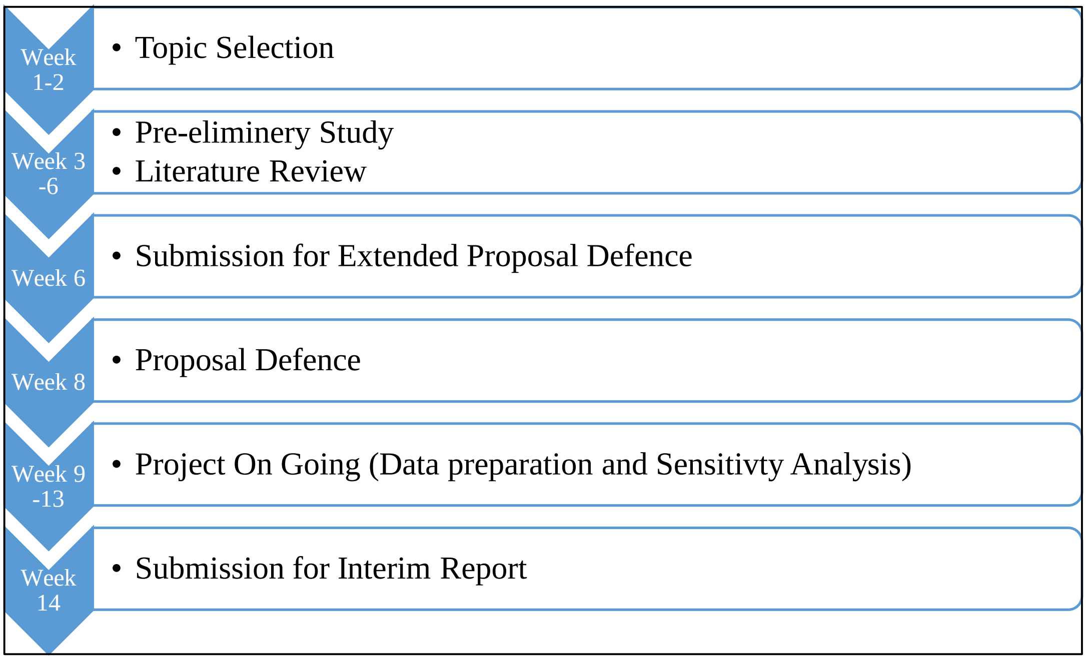

Figure 3.4: Milestone for FYP 1

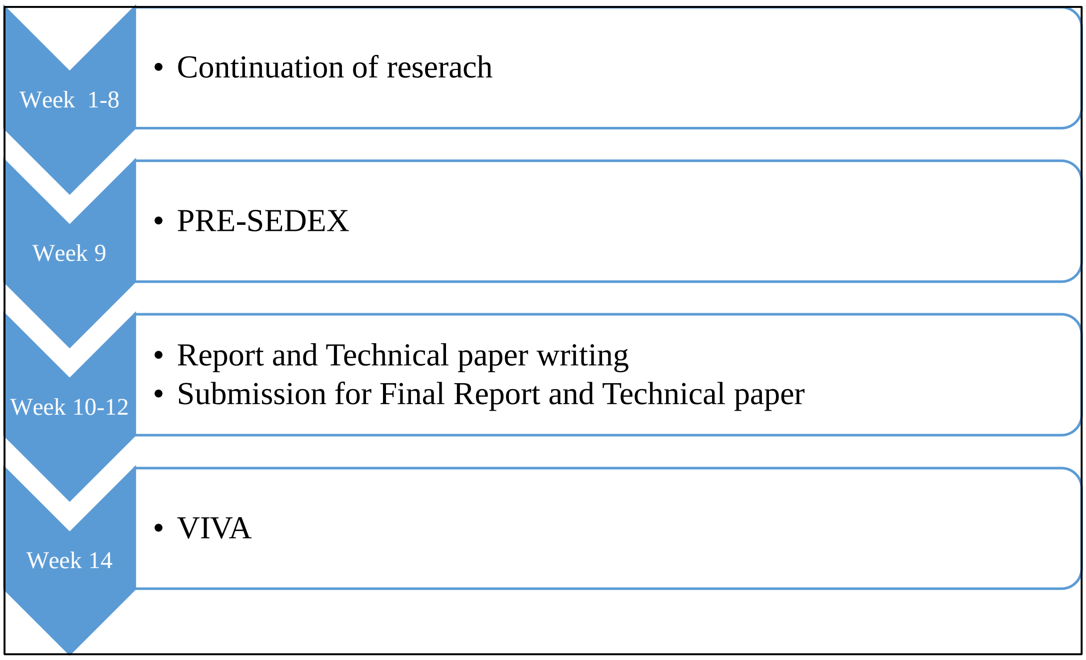

Figure 3.5: Milestone for FYP 2

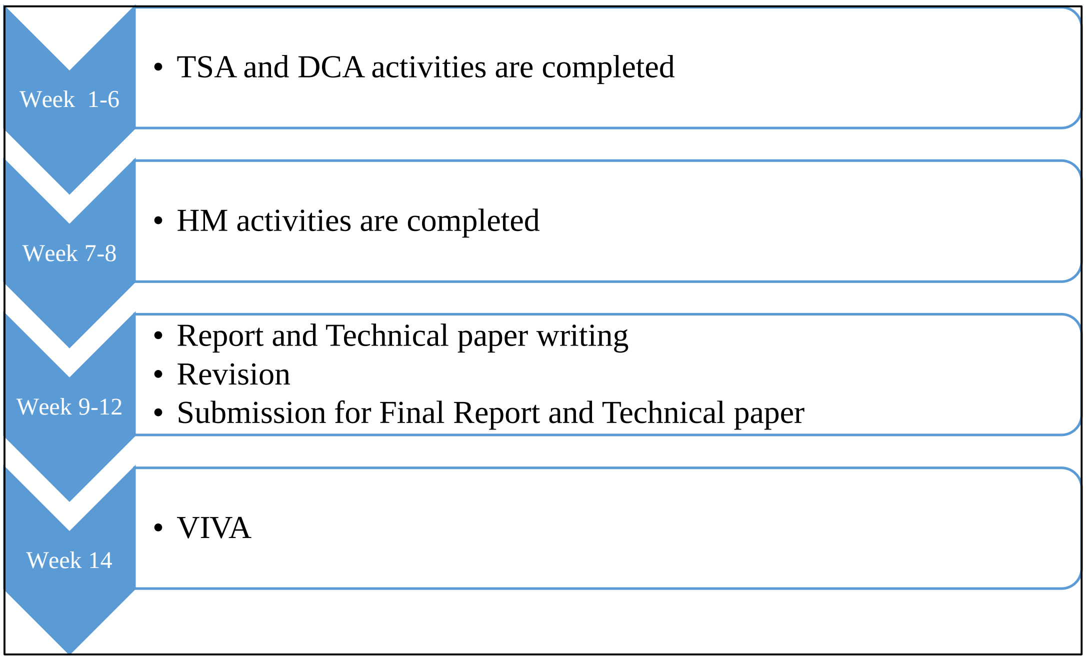

Figure 3.6: Milestone for Project

Table 1: Gantt Chart of FYP 2

Detail	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Project Work Continues	■	■	■	■	■	■	■							
Progress Report Submission							●							
Pre-SEDEX									●					
Submission of Draft Final Report											●			
Revision of Final Report											■	■		
Submission of Dissertation and Technical Paper												●		
Viva														●

Table 2: Project Gantt chart

Details	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Decline Curve Analysis Activities														
Time Series Analysis Activities														
Progress Report Submission														
History Matching Activities														
Writing of Technical Paper														
Submission of Draft Final Report														
Revision of Final Report														
Submission of Dissertation and Technical Paper														

CHAPTER 4
RESULTS AND DISCUSSION

4.1 Model Description

SPE comparative solution number one is used for this study. Comparative solution is a product from SPE that contain a synthetic reservoir model to match up-gridding and upscaling approaches and the ability to forecast performance through a million cell geological model. This is 10 x 10 x 3 Cartesian grid model with injection well located on block (1, 1, 3), while the producer is located on (10, 10, 1). A slight altercation is needed by changing the type of injector into water to serve the objectives of this study. There is no skin factor in this model with initial pressure is 4800psia and temperature 220° F. The reservoir has a constant porosity even though there is large contrast in each layer of permeability. Each of the I, J, and K permeability are constant within each layer of the reservoir, but vary from layer to layer. The rock compressibility for this model is $3x10^{-6}$

Table 3: Reservoir Model Properties

Parameter	Base
Porosity	0.3
PermI_1	200
PermI_2	50
PermI_3	500
PermJ_1	200
PermJ_2	50
PermJ_3	500
PermK_1	20
PermK_2	40
PermK_3	60

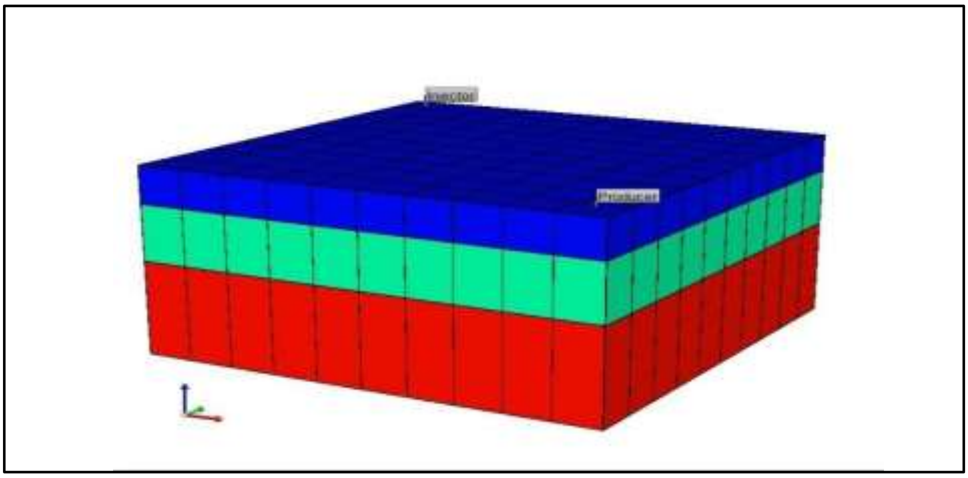

Figure 4.1: 3D View of Model

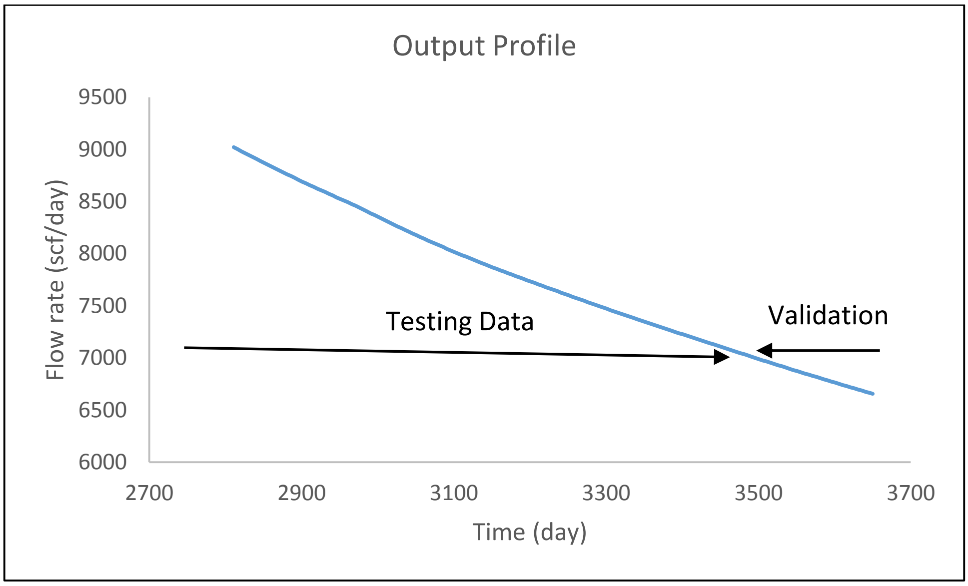

Figure 4.2: Output profile of model

Simulation was conducted over 10 years with a 5-day intervals. Initially, 741 timesteps were produced, where only 424 were taken due to the constants output at earlier period. Data validation was done to avoid overfitting, which normally occurs when a model is excessively complex, such as having too many parameters relative to the number of observations. A model that has been overfitted will normally have poor predictive performance, as it can overstate minor fluctuations in the data. There is no correct percentage for training/test split where the common ratios are 80/20, 70/30, and 60/40. For this study, 60/40 ratio data will be used, and the same action is also done for the input data.

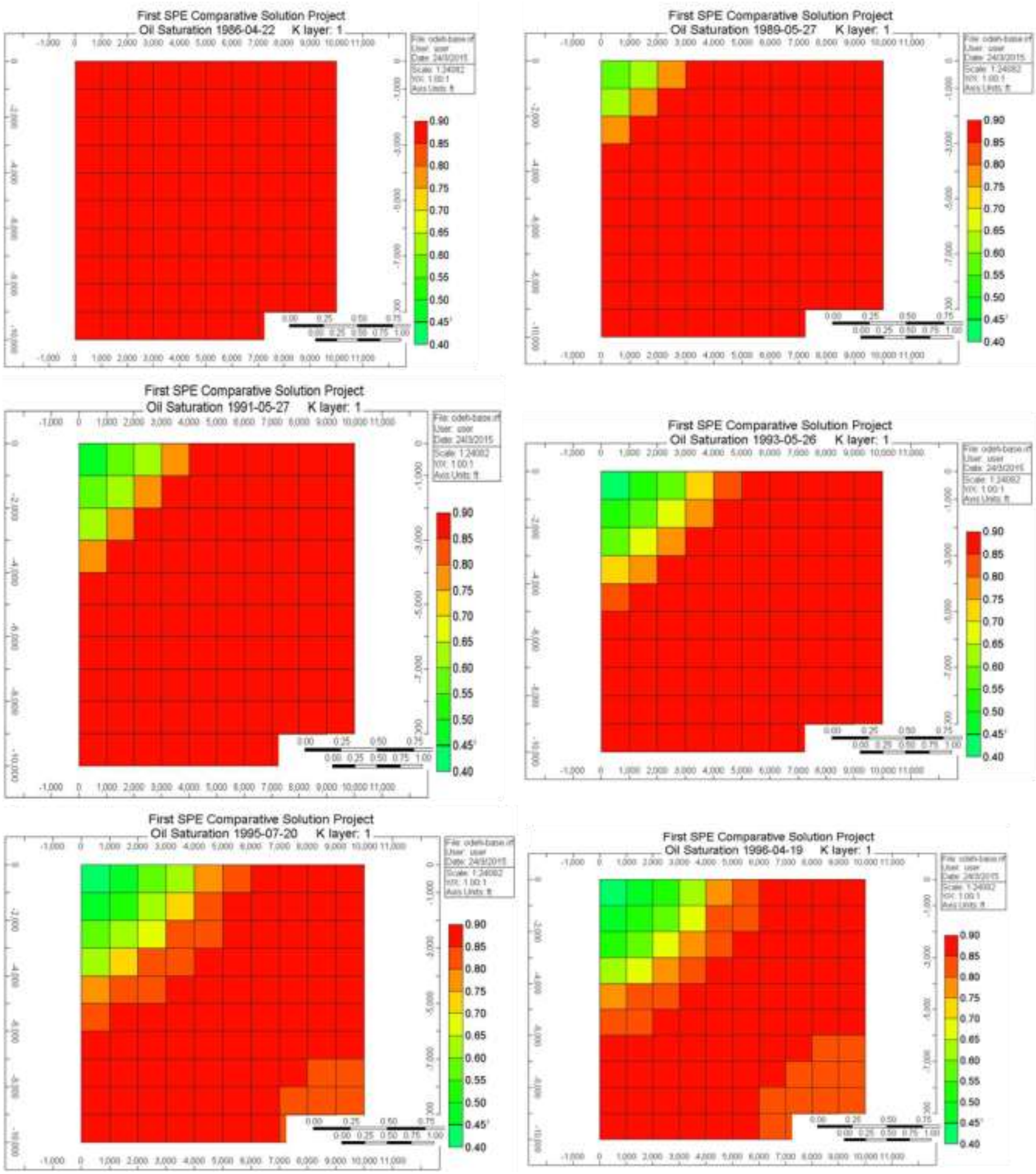

Figure 4.3: Saturation profile over times

4.2 Conventional Reservoir Forecasting Result

4.2.1 Decline Curve Analysis

For this study, DCA is not applicable due to the existence of secondary drive of water-flood. However for the sake's of comparison, this method still going to be constructed.

Alternatively, the selection of the decline curve type can be conducted by studying the trend of curve. The exponential curve or constant percentage decline is characterized by a straight line, while the harmonic curve is indicated by concave upward curve. It is witnessed that the curve generated is straight line, hence exponential decline will be used.

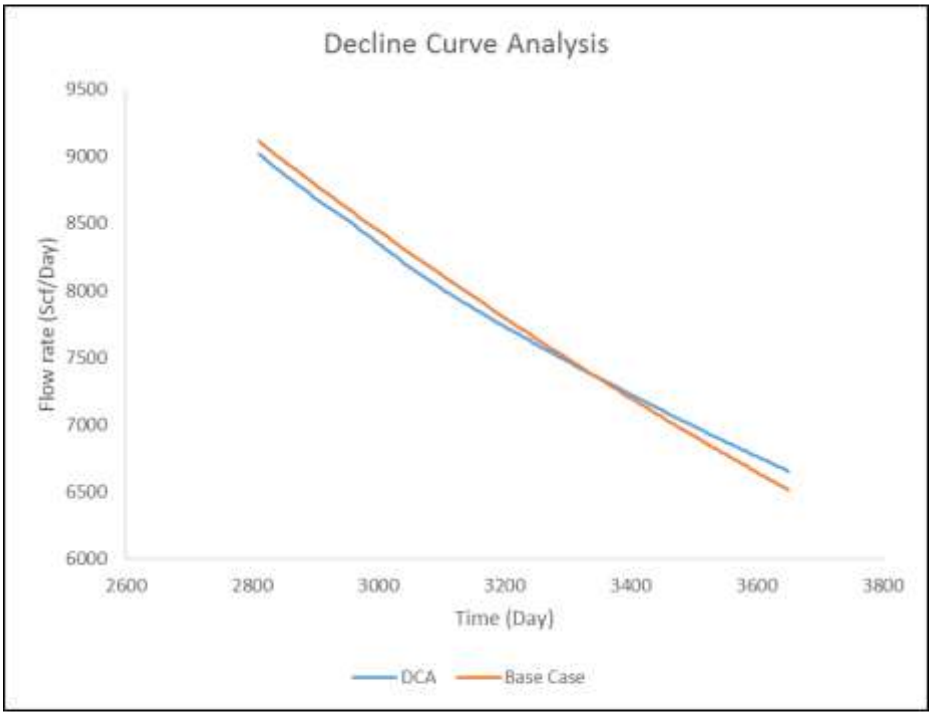

Figure 4.4: DCA Result

Exponential decline rate can be constructed as

$$Q(t) = Q_i e^{-\alpha t} \quad \text{.......................................} \quad (24)$$

where,

$$\alpha = -\ln(1-D) \quad \text{...(25)}$$

$$D = \frac{Q_{t=0} - Q_{t=1}}{Q_{t=0}} \quad \text{.. (26)}$$

D is a constant and α is decline rate. The decline rate formula however can be directly obtain from Excel by simply showing the equation of the curve. Despite it is not applicable in reservoir with secondary drive, figure 4.4 showed that the base and predicted curve are not too much different. By using equation (23), it is found that the error percentage between the base and predicted case is **3.85%.**

4.2.2 History Matching

Table 4: Parameters Comparison

Parameter	Base	New Model
Porosity	0.3	0.25
PermI_1	200	300
PermI_2	50	45
PermI_3	500	400
PermJ_1	200	150
PermJ_2	50	25
PermJ_3	500	350
PermK_1	20	35
PermK_2	40	15
PermK_3	60	30

SPE comparative solution number one is synthetic model without previous production history, hence a new model is needed to be created as comparison with base model. On creating the new model, sensitivity analysis will be conducted to measure which parameters are sensitive.

Response surface method will be used as a method in sensitivity analysis, which the basic concept of defining the most sensitive parameter in the response surface is by observing at the value of its coefficient in the equation. It is probable that one parameter has the biggest coefficient value in a specific time step but not in the other time steps at a particular response variable. It is also probable to find one parameter which has the biggest coefficient in one response variable but not in the other response variables at a specific time step

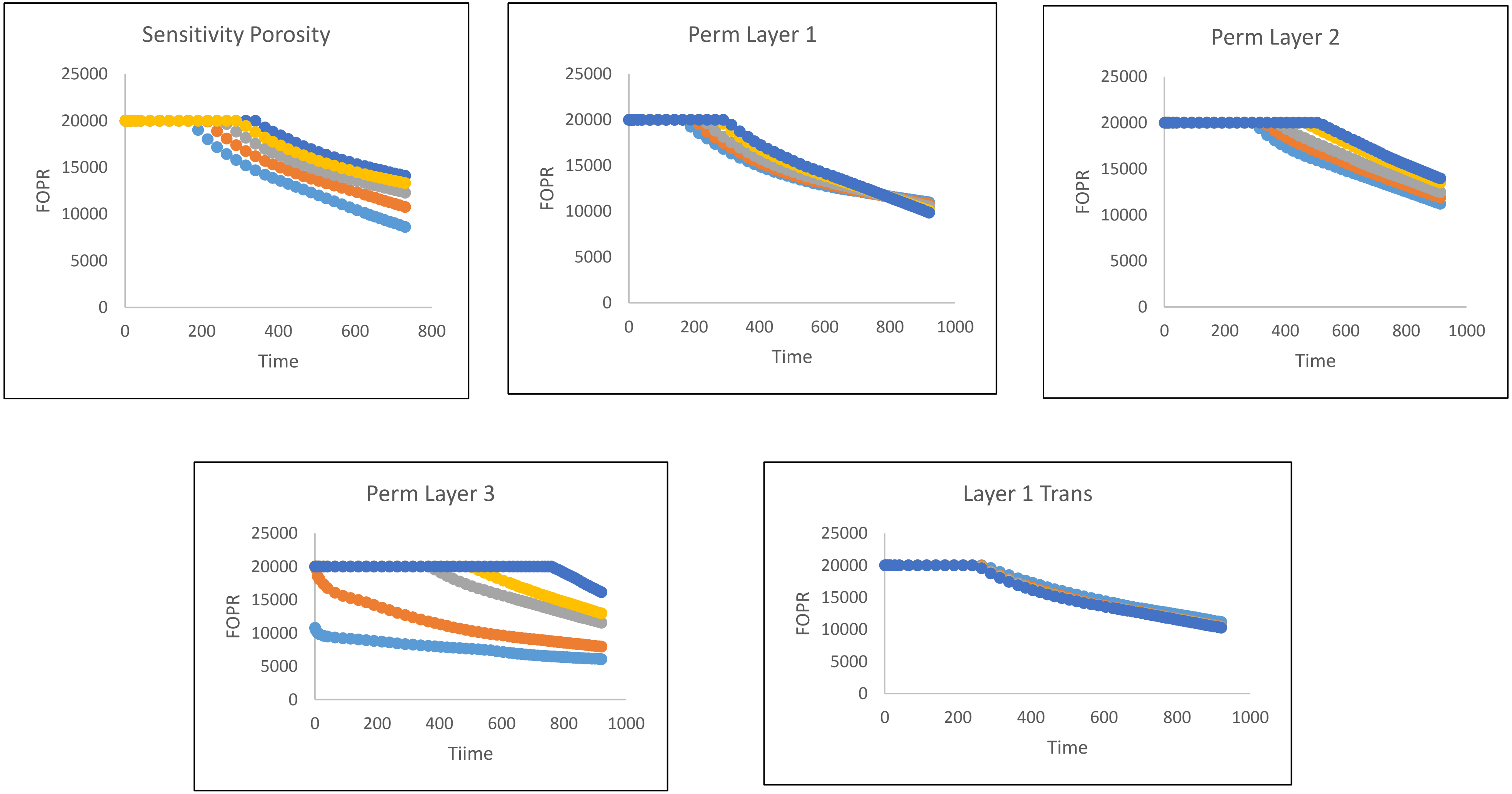

Figure 4.5: Sensitivity Analysis results

. It is found that the sensitive parameters are porosity and permeabilities in every layer. The properties of the new model can be seen at table 4.2. Another important aspect in HM is selection of optimization algorithm, which genetic algorithm is used. It is population based algorithm, which at each iteration, more than one solution are created. 500 cases were conducted within the parameter range listed below and a statistical method, called latin hypercube, will assist in this process by generating samples of a series of value from multidimensional distributions.

Table 5: Parameterization range

Parameter	Min	Max
Porosity	0.1	0.5
PermI_1	100	500
PermI_2	10	700
PermI_3	50	600
PermJ_1	50	500
PermJ_2	10	700
PermJ_3	50	600
PermK_1	5	400
PermK_2	5	300
PermK_3	20	300

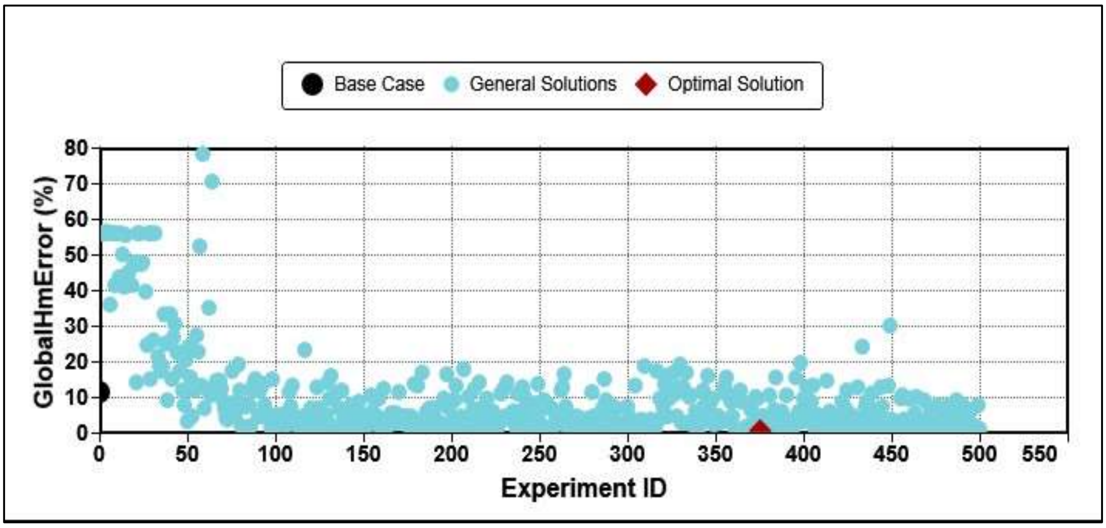

Figure 4.6: HM Result

From 500 cases, the best fitting solution will be generated after the estimation process that took around two hours until completion. It is not recommended to take one result only as a comparison due to possibility of a case that has best solution with different set of properties compared to base case model. The problem occurs as every reservoir must be treated separately and different scenario is required. In addition, the whole process required a long time compare to DCA.

Table 6: Properties of selected cases

Parameter	Case 306	Case 314	Case 375	Case 429	Case 492
Porosity	0.278	0.274	0.274	0.278	0.278
PermI_1	324	348	312	312	300
PermI_2	65.2	65.2	37.6	65.2	79
PermI_3	583.5	578	567	572.5	567
PermJ_1	86	81.5	104	99.5	99.5
PermJ_2	289.45	261.85	248.05	206.65	206.65
PermJ_3	399.25	377.25	388.25	322.25	322.25
PermK_1	186.7	210.4	178.8	178.8	186.7
PermK_2	94.975	112.67	89.075	130.37	130.37
PermK_3	85.8	91.4	85.8	69	69

In order to select the best model from HM, error quality, such as NRSME (*eq. 23)*, is utilized as the selection criteria to compare the predicted and base case that will be plotted on Excel. By using the same equation in DCA, Case 314 is selected to be comparison model with the lowest value of NRMSE.

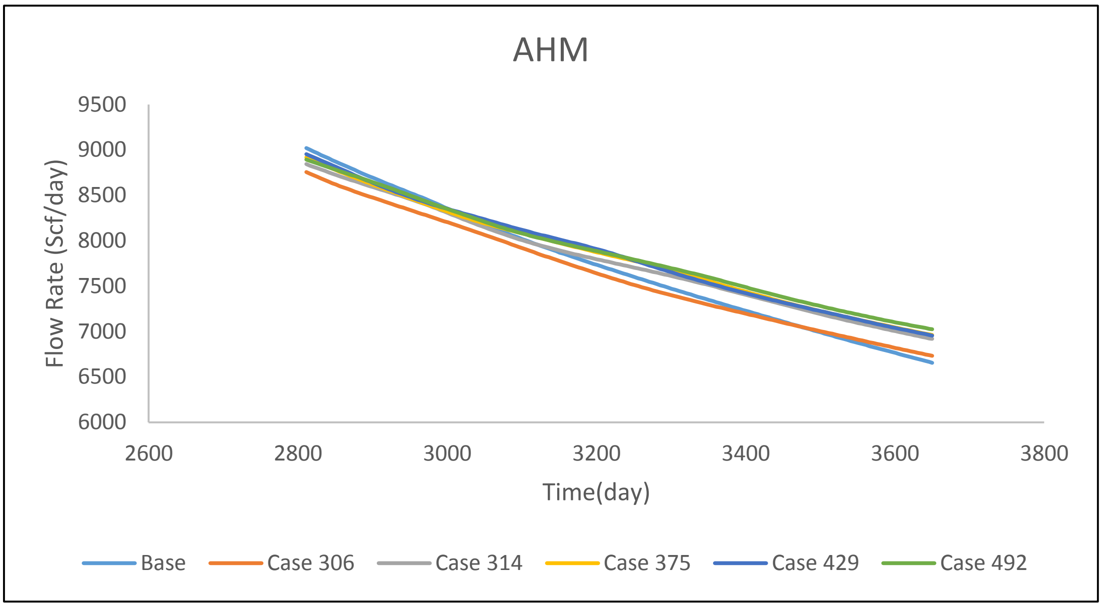

Figure 4.7: History Matching Cases Comparison

Table 7: HM NMRSE

Cases	NRMSE
306	14.23%
314	**11.83%**
375	19.14%
429	21.2%
492	24.2%

4.3 Time Series Analysis Result

Polynomial modelling is used as the objective function when a simple empirical model is required for characterizing data by global fit. The main advantage is it includes reasonable flexibility for data that is not too complicated.

.

In general, the polynomial model can be constructed as

$$A(t)y(t) = \sum_{i=0}^{nu} \frac{B_i(q)}{F_i(q)} u_i(t - nk_i) + \frac{C_i(q)}{D_i(q)} \frac{1}{1-q^{-1}} e(t) \quad \ldots\ldots\ldots\ldots \quad (27)$$

The variables *A*, *B*, *C*, *D*, and *F* are polynomials expressed in the time-shift operator q^{-1}, known as lag operator, as an element of time series to produce previous element. The polynomials of OE and BJ model are shown in equation (13 and (14).

Furthermore, fot estimation process, defining order of polynomials is essential. The *model order* must be specified as a set of integers that represent the number of coefficients for each polynomial include in the selected structure— *na* for *A*, *nb* for *B*, *nc* for *C*, *nd* for *D*, and *nf* for *F*. The details for every order are shown in equation (15) to (18).

To choose the order of polynomial, it is advisable to try out various available choices. However, the higher degree of the polynomial can be unstable, thus the higher degree of polynomial need to be taken care carefully.

Four models for OE and BJ are created. For OE, the models are: OE(1-2-1), OE(2-2-1), OE(3-3-1), and OE(1-3-1). For BJ models: BJ(1-2-1-2-1), BJ(2-2-1-1-1), BJ(2-2-2-2-1), and BJ(1-1-2-2-1). On estimating every polynomial's parameter prediction-error identification will be used. It is a broad family of parameter estimation methods that can be useful to quite uninformed model parameterizations. Estimation parameters of both models are complicated non-linear solution. However, by using MATLAB the parameters can be directly refined and the error predictor for both model can be expressed as below:

Table 8: OE PEM

Model	Error Predictor
OE 121	• $B(z) = -0.08304z^{-1} + 0.08408z^{-2}$ • $F(z) = 1 - 1.946z^{-1} + 0.9468\ z^{-2}$
OE 131	• $B(z) = 0.000208\ z^{-1}$ • $F(z) = 1 - 2.534z^{-1} + 2.087z^{-2} - 0.5528z^{-3}$
OE 221	• $B(z) = -0.08304z^{-1} + 0.08408z^{-2}$ • $F(z) = 1 - 1.946z^{-1} + 0.9468z^{-2}$
OE 331	• $B(z) = 22.85\ z^{-1} - 46.34\ z^{-2} + 23.5z^{-3}$ • $F(z) = 1 - 0.7162z^{-1} - 0.9952z^{-2} + 0.721\ z^{-3}$

Table 9: BJ PEM

Model	Error Predictor
bj11221	• $B(z) = 0.01088\ z^{-1}$ • $C(z) = 1 + 0.1678z^{-1}$ • $D(z) = 1 - 1.825z^{-1} + 0.8315z^{-2}$ • $F(z) = 1 - 1.168\ z^{-1} + 0.1763z^{-2}$
bj22111	• $B(z) = -1.154z^{-1} + 1.154z^{-2}$ • $C(z) = 1 + 1.224z^{-1} + 0.7575z^{-2}$ • $D(z) = 1 - 0.9911z^{-1}$ • $F(z) = 1 - z^{-1}$
bj12121	• $B(z) = 0.002104z^{-1}$ • $C(z) = 1 + 1.062z^{-1} + 0.5948z^{-2}$ • $D(z) = 1 - 0.9652z^{-1}$ • $F(z) = 1 - 1.803z^{-1} + 0.8049z^{-2}$
bj22221	• $B(z) = -0.3805z^{-1} + 0.3847z^{-2}$ • $C(z) = 1 + 0.0175z^{-1} + 0.006893z^{-2}$ • $D(z) = 1 - 1.871z^{-1} + 0.8738z^{-2}$ • $F(z) = 1 - 1.831z^{-1} + 0.8344z^{-2}$

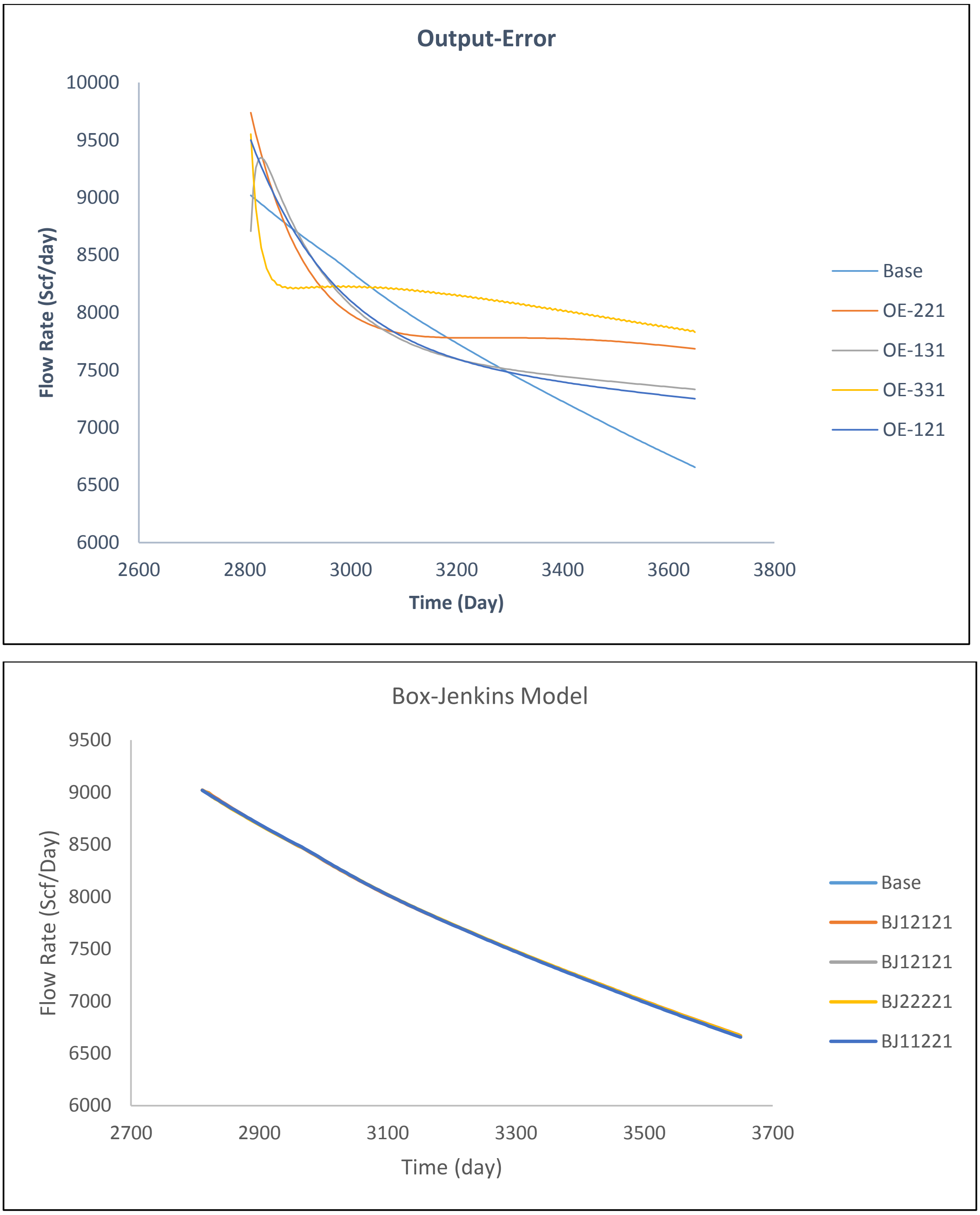

Figure 4.8: OE and BJ Model Comparison

Compare to HM, the whole process of TSA relatively short. It takes 5 seconds for every model to estimate the output.

The next step after the estimation process is validating the model to select the best model and to observe how well the simulated or predicted output of the model matches the measured output.

$$Error = \frac{y - \breve{y}}{y - \bar{y}} \times 100 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (28)$$

Where y is the measured output, $\breve{y}$ is predicted output, and $\bar{y}$ is the mean of y. Small value of error is desirable.

Table 10: Best fit error of OE model

Model	Error
OE 131	34.7%
OE 221	44.94%
OE 331	97.62%
OE 121	37.9%

Table 11: Best fit error of BJ model

Model	Error
BJ 12121	1.05%
BJ 22111	4%
BJ 22221	1.5%
BJ11221	1.14%

Besides using *best fir error* method, NRMSE for all the models are conducted as well. The result are listed below.

Table 12: OE and BJ NMRSE

Model	Error
OE 131	41.35%
OE 221	14.21%
OE 331	64.6%
OE 121	**11.06%**

Model	Error
BJ 12121	0.64%
BJ 22111	0.70%
BJ 22221	0.38%
BJ11221	**0.32%**

It is showed that the accuracy of OE and BJ is very different. This happen because BJ model is modified version from OE; it is much more complex where both of stochastic and deterministic function are involved hence give a better prediction.

4.4 RESULT AND COMPARISON

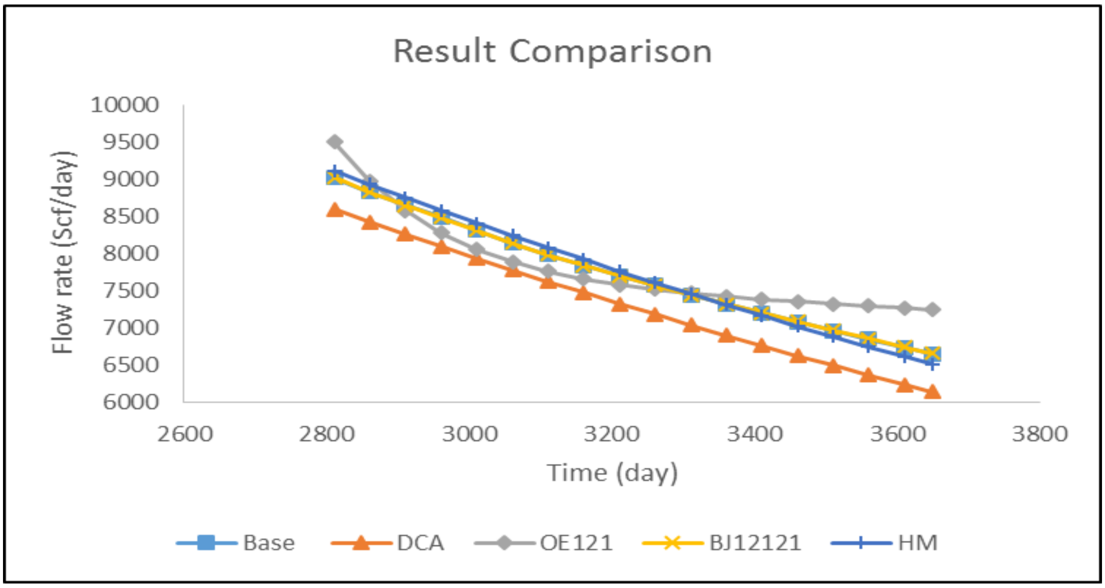

Figure 4.9: Graph Comparison

A comparison among best case from each respective methods are compared by using error quality; where in this study NRMSE will be used. From table below, BJ model showed a very outstanding result compare to other methods with 0.32% error. Despite of the restriction of the secondary drive, DCA come as a second best prediction with 3.85% error. HM showed the least accurate result compare to the others; however this result occurs due to the fact that the change of porosity might alter the value of predicted STOIIP.

Table 13: NMRSE for all methods

Model	NMRSE
DCA	3.85%
TSA BJ 12121	11.06%
TSA OE 121	0.32%
HM Case 314	11.83%

CHAPTER 5
CONCLUSION

On this study, a comparison among forecasting tools for reservoir performance has been conducted to discover which tools have the best accuracy.

Below are the major conclusions from this research:

- Time Series Analysis can be used as a reservoir performance prediction tools. From the NRMSE value, especially BJ model where it showed a very accurate prediction compare to the conventional reservoir prediction tools.
- In term of easiness, DCA is the easiest method to conduct the reservoir forecasting. However, the TSA process is also easier compare to the HM process. Besides, of the easiness both DCA and TSA require a smaller amount of time compare to the HM.
- The BJ model has the smallest error compare to the others method. It indicates that this model is very accurate tools to predict the reservoir performance. Even though the OE model has larger error compare to the BJ model, the model still can be used as well as a prediction tool in reservoir forecasting

REFERENCES

Aitokhuehi, I. (2007, January 1). Maximizing the Potential of Decline Curve Analysis. Society of Petroleum Engineers. doi:10.2118/111916-MS

Arps, J. J. (1945, December 1). Analysis of Decline Curves. Society of Petroleum Engineers. doi:10.2118/945228-G

Babinec, Š. and J. Pospíchal (2009). Combination of Dynamic Reservoir and Feedforward Neural Network for Time Series Forecasting. Advances in Neuro-Information Processing. M. Köppen, N. Kasabov and G. Coghill, Springer Berlin Heidelberg. 5507: 284-291.

Biu, V. T., & Ikiensikimama, S. S. (2012, January 1). Understanding Crude Oil Price Variation Using Time Series Analysis. Society of Petroleum Engineers. doi:10.2118/162964-MS

Cancelliere, M., Verga, F., & Viberti, D. (2011, January 1). Benefits and Limitations of Assisted History Matching. Society of Petroleum Engineers. doi:10.2118/146278- MS

Ebrahimi, M. (2010, January 1). Enhanced Estimation of Reservoir Parameters Using Decline Curve Analysis. Society of Petroleum Engineers. doi:10.2118/133432-MS

Fetkovich, M. J., Vienot, M. E., Bradley, M. D., & Kiesow, U. G. (1987, December 1). Decline Curve Analysis Using Type Curves: Case Histories. Society of Petroleum Engineers. doi:10.2118/13169-PA

Li, K., & Horne, R. N. (2005, January 1). Verification of Decline Curve Analysis Models for Production Prediction. Society of Petroleum Engineers. doi:10.2118/93878-MS

Lind, R. H., Gaaim, A. M., Almuallim, H., & Allottai, O. (2013, September 16). Computer Assisted History Matching - A Field Example. Society of Petroleum Engineers. doi:10.2118/165979-MS

Ljung, L. (2002). "Prediction error estimation methods." Circuits, Systems and Signal Processing **21**(1): 11-21.

Ma, Z.Y., & Pointe, P.L. (2011). Uncertainity Analysis and Reservoir Modelling. AAPG Memoir 96

Obidike, P., Emelle, C., & Vlaardingerbroek, R. (2014, August 5). Integrated Production Forecasting for Business Planning. Society of Petroleum Engineers. doi:10.2118/172479-MS

Olominu, O., & Sulaimon, A. A. (2014, August 5). Application of Time Series Analysis to Predict Reservoir Production Performance. Society of Petroleum Engineers. doi:10.2118/172395-MS

Saltelli, A., & Annoni, P. (2010). How to avoid a perfunctory sensitivity analysis. Environmental Modelling & Software, 25(12), 1508-1517.

Tavassoli, Z., Carter, J. N., & King, P. R. (2004, September 1). Errors in History Matching. Society of Petroleum Engineers. doi:10.2118/86883-PA

Tomomi, Y. (2000, January 1). Non-Uniqueness of History Matching. Society of Petroleum Engineers. doi:10.2118/59434-MS

Wu, X., He, J., Zhang, H., & Ling, K. (2013, March 23). Tactics and Pitfalls in Production Decline Curve Analysis. Society of Petroleum Engineers. doi:10.2118/164503-MS